

Chapter 13

Molecular Typing Techniques: State of the Art

Richard V. Goering

Introduction

The treatment of infectious disease centers around the goals of both curing the patient and preventing or at least restricting the spread of disease. In a perfect world, health care professionals would know that these goals have been achieved when the patient's health is restored and there are no new occurrences of infected patients. However, the real world of infectious disease is far from perfect. The individual patient may present with evidence of recurring or additional infection by a pathogen (e.g., at a different body site). Different members of a patient population may yield cultures of the same organism. In both instances, the question commonly asked is whether multiple isolates of a given pathogen represent the same strain. In the individual patient, this question commonly relates to issues of therapeutic efficacy while in a patient population the concern is infection control. However, in both settings, the resolution of these questions is aided by specific epidemiological assessment. In the past, a variety of methods based on phenotypic characteristics have been used for this purpose including biotype, serotype, susceptibility to antimicrobial agents, or bacteriophages, etc. [1–4]. However, in the 1970s, techniques developed for the recombinant DNA laboratory began to find application in the molecular characterization of clinical isolates. These included comparing protein molecular weight distributions by polyacrylamide gel electrophoresis, relative mobility of specific enzymes by starch-gel electrophoresis (multi-locus enzyme electrophoresis), specific antibody reactions with immobilized cellular proteins (immunoblotting), and cellular plasmid content (i.e., plasmid fingerprinting) [2, 5, 6]. However, by the 1980s it was clear that comparisons at the genomic level would provide the most fundamental measure of epidemiological relatedness. Thus, molecular typing was born.

R.V. Goering (✉)
Department of Medical Microbiology and Immunology,
Creighton University School of Medicine, Omaha, NE 68178, USA
e-mail: rgoeri@creighton.edu

While a wide range of etiological agents are of clinical concern, this review focuses on molecular approaches to the epidemiological analysis of bacterial pathogens.

What Does “State of the Art” Mean?

In any area of scientific investigation, state of the art methodology may be viewed from two different perspectives. There are cutting-edge techniques requiring specialized equipment and expertise that perform remarkably well but are of limited availability to many investigators. Alternatively, there are functional state of the art approaches, meaning that one is using the best method available within the prevailing (financial, geographic, technical expertise, etc.) environment. In this context, it is important to recognize that while one may not have access to the most recently published sophisticated methods, from an epidemiological standpoint, it is better to do something rather than nothing. Thus, this review begins with examples of established molecular typing techniques which, depending on the (financial, geographic, scientific) environment, may still be viewed as state of the art while also considering more recently described cutting-edge approaches.

The Ultimate Foundation for Epidemiological Comparison: The Bacterial Genome

Advances in DNA sequencing have shown that what was once thought of as the bacterial chromosome is actually a core genome plus a variety of inserted mobile genetic elements [7–9]. Nevertheless, the totality of these sequences makes the cell a specific strain of *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Escherichia coli*, etc. Thus, the bacterial genome represents the most fundamental molecule of identity in the cell and the common goal of molecular typing approaches is to provide a measure of isolate genomic relatedness [10]. While the methodological aspects of these techniques differ, they can generally be grouped into two categories of data output, either electrophoretic “bands” or DNA sequences.

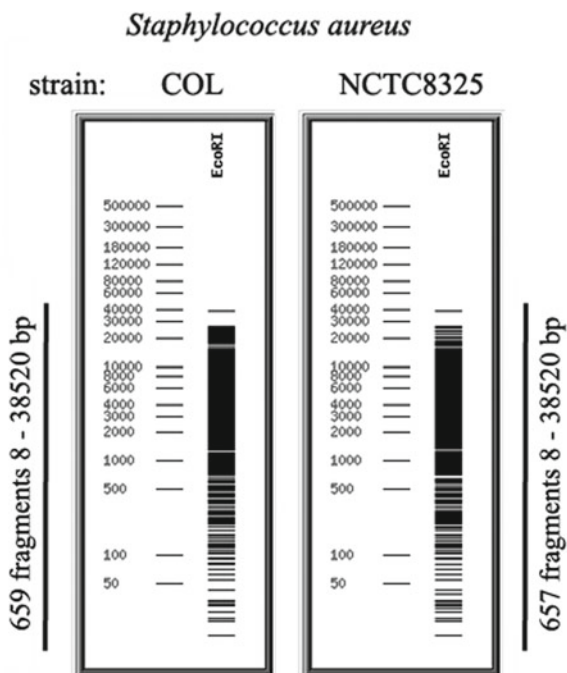
Methods with Electrophoretic Output

Restriction Enzyme-Based

Chromosomal Restriction Enzyme Analysis

The ubiquitous presence of chromosomal DNA in all bacterial pathogens made restriction enzyme analysis (REA) an attractive early approach to molecular strain typing. While all bacterial cells can theoretically be analyzed by such a process, the

Fig. 13.1 Diagrammatic representation of REA with chromosomal DNA from *S. aureus* strains COL and NCTC8325 digested with the restriction enzyme *EcoRI*. Data were generated using the Comprehensive Microbial Resource of the J. Craig Venter Institute Web site: <http://cmr.jcvi.org/cgi-bin/CMR/shared/Menu.cgi?menu=genome>



DNA sequences recognized by common restriction enzymes such as *EcoRI*, *HindIII*, etc., are abundantly dispersed (e.g., on average >600 copies) throughout a typical 2–3 Mb bacterial chromosome. This is illustrated in Fig. 13.1 with a comparison of *EcoRI* REA for *S. aureus* strains COL and NCTC8325. Thus, the resulting challenge is to accurately compare electrophoretic patterns that comprise hundreds of restriction fragments, often co-migrating in clusters of similar size, and potentially including resident plasmid DNA [11]. Consequently, at the present time this method continues to be recommended only for use with *Clostridium difficile* [12].

Since the mid 1970 Southern hybridization [13] has been a staple of molecular biology, and its power to probe for specific DNA sequences soon began to find clinical application. For diagnostic purposes, tests to detect the presence or absence of clinically relevant sequences (e.g., related to organism identification, antibiotic resistance) began to be developed. For epidemiological analysis, probes specific for sequences found at multiple chromosomal locations can be hybridized against chromosomal restriction enzyme fragments which have been electrophoretically separated. The resulting hybridization patterns (restriction fragment length polymorphisms (RFLPs)) provide an indication of chromosomal relatedness between different bacterial isolates. However, at present this approach is not widely used for epidemiological analysis with the exception of probes for the insertion sequence IS6110 in the RFLP analysis of *Mycobacterium tuberculosis* [14, 15].

Pulsed-Field Gel Electrophoresis

In contrast to conventional REA, rare-cutting restriction enzymes cleave the bacterial chromosome into a relatively small number of fragments (e.g., 10–30) due to the length and/or DNA base composition of their recognition sequences,. However, electrophoretic analysis of the megabase-size restriction fragments generated is complicated by their size-independent migration during conventional agarose-gel electrophoresis [16, 17]. In 1980, alternative electrophoretic approaches were developed based on the principal of periodic reorientation of the electric field (and DNA migration) relative to the direction of the gel. The pulsed electric field separates DNA fragments over a wide range of sizes from kilobytes to megabytes (Fig. 13.2)

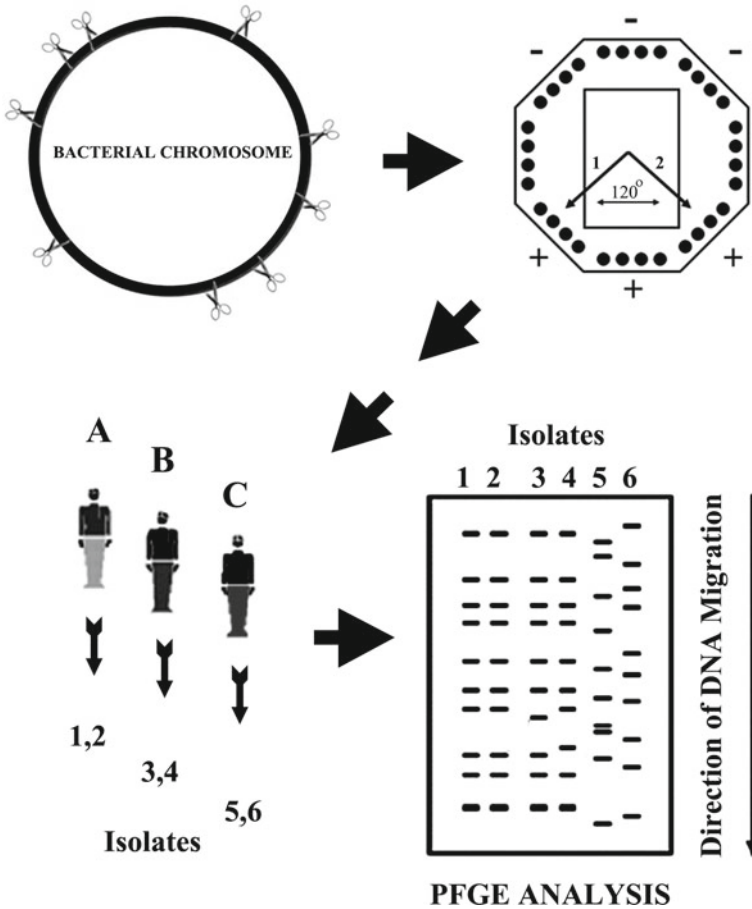
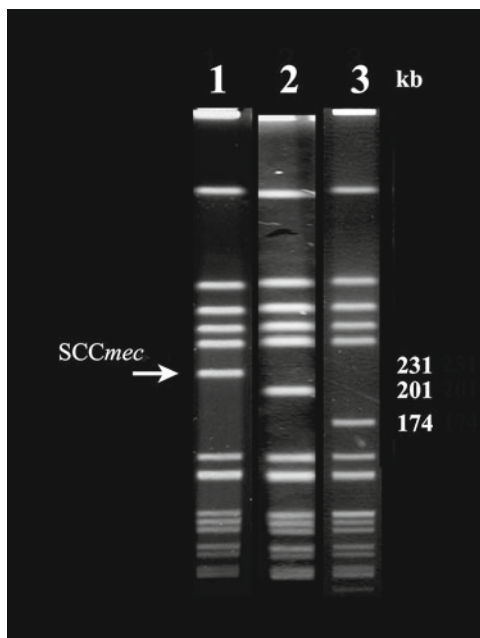


Fig. 13.2 Illustration of PFGE workflow moving from chromosomal digestion with rare-cutting restriction enzymes to macro-restriction fragment separation by PFGE to the final analysis of fragment patterns from different (patient) sources

Fig. 13.3 *Sma*I-digested chromosomal DNA from USA300 *S. aureus* isolates which are (lane 1) methicillin resistant and PCR positive for the adjacent ACME *arcA* gene (lane 2), methicillin susceptible due to loss of *SCCmec* but *arcA* positive, or (lane 3) negative for both *SCCmec* and *arcA*. (Modified from Goering et al. [21])



thus allowing a more manageable comparison of isolate patterns. The usefulness of pulsed-field gel electrophoresis (PFGE) for molecular typing has been recently reviewed [18, 19]. However, it is important to note that while the method is far from new, PFGE has exhibited enormous staying power as a valuable method of genomic analysis and comparison. This is especially true for molecular typing where for the majority of bacterial pathogens it remains the acknowledged “gold standard” for assessing isolate interrelationships. The reason for this longevity is multifold. Overall, the method for chromosomal DNA isolation (i.e., the *in situ* lysis of bacterial cells encased in agarose plugs) requires only minor variation with different bacterial species. A wide range of bacterial pathogens can be analyzed using a small number of different restriction enzymes (commonly *Sma*I and *Xba*I for gram-positive and -negative isolates, respectively). Despite the fact that PFGE obviously does not detect every genetic change and macro-restriction fragment, for most organisms analyzed the sum of the visible fragment sizes represents greater than 90% of the chromosome. This visual sense of global chromosomal monitoring can be highly informative not only for isolate comparisons, but also in associating characteristic PFGE patterns with specific (e.g., internationally recognized) bacterial strains [20]. In addition, the chromosomal overview provided by PFGE allows visualization of genomic rearrangements as in the case of *S. aureus* strain USA300 where changes in PFGE patterns can be specifically associated with loss of the staphylococcal chromosomal cassette encoding methicillin resistance (*SCCmec*) or the adjacent arginine catabolic mobile element (ACME) [21] (Fig. 13.3).

Optical Mapping

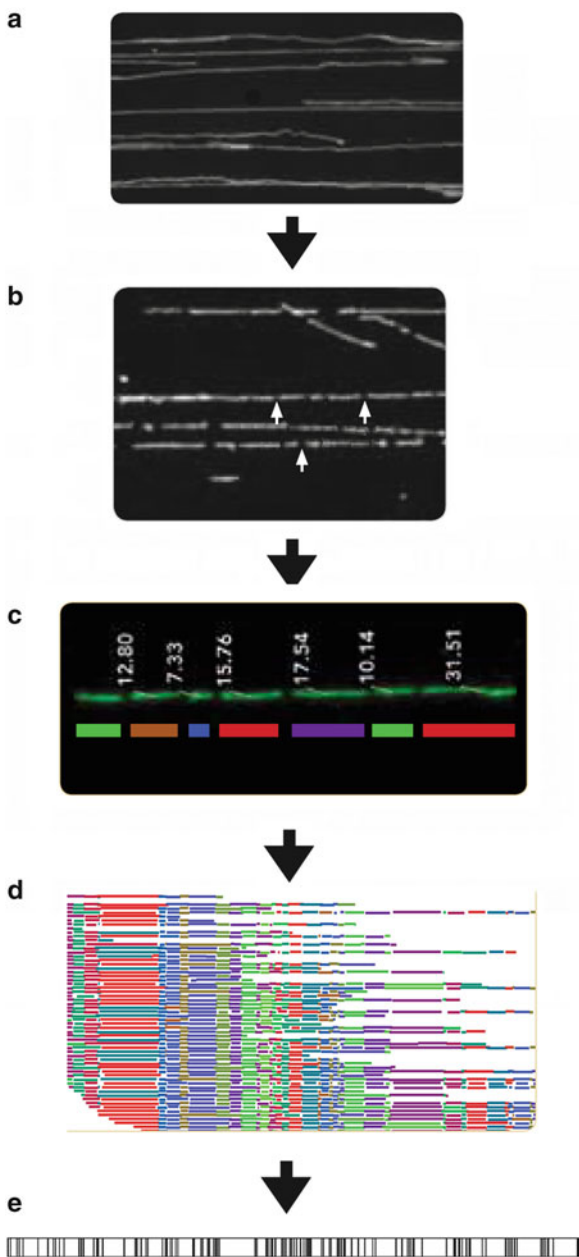
Optical mapping (OM) is an interesting variation of the REA-PFGE method with potential epidemiological application. Similar to PFGE, high molecular weight genomic DNA is obtained by agarose-encased lysis of cells. As illustrated in Fig. 13.4, single DNA molecules are then electrostatically fixed to the surface of material amenable to scanning by fluorescent microscopy. The DNA molecules are exposed to restriction endonuclease digestion but the order of the resulting fragments is maintained since each molecule is immobilized. After staining with a fluorescent dye, fluorescent microscopy coupled with appropriate software converts the optical image to a digital format producing restriction maps of the individual molecules. The overlapping maps are then assembled to produce an ordered restriction map of the entire chromosome. Overall hands-on time of a few hours with final genomic data output in less than 2 days from start to finish makes OM an interesting technology which has shed recent light on a variety of microbial strain interrelationships [22, 23]. However, a per-isolate cost of several thousand dollars (total instrumentation list price >\$200,000) currently makes OM impractical for infection control surveillance or routine multi-isolate epidemiological analysis.

PCR-Based

Amplified Fragment-Length Polymorphism

Amplified fragment-length polymorphism (AFLP) remains in current use as an interesting approach that combines the use of restriction enzymes and PCR to potentially analyze a wide range of bacterial pathogens [24]. The process involves creation of typing patterns based on PCR amplification of a subset of chromosomal restriction fragments (Fig. 13.5). This is accomplished by digesting isolated DNA with two different restriction endonucleases, usually chosen so that one cuts more frequently than the other (e.g., *EcoRI* and *MseI*). While a large group of restriction fragments are initially created, only specific subsets are utilized for isolate comparison. Adapters specific for the cleaved restriction-sites are ligated to the fragment ends thus extending the length of the known end sequences and serving as primer binding sites for PCR. The adapter design includes extra nucleotides beyond the restriction-site sequence allowing only a subset of fragments to be amplified. Using labeled primers the specificity of the process may be further controlled, ultimately leading to an electrophoretic pattern of amplified products that becomes the basis for assessing isolate interrelationships. Recent AFLP improvements have included multiple enzyme–adapter combinations and either fluorescent or radioactively labeled primers, allowing high-throughput analysis to be achieved using an automated DNA sequencer, phosphoimager, etc. [24, 25]. However, issues regarding data analysis and inter-laboratory sharing, and the specialized equipment required for electrophoresis have limited the use of this method in the clinical setting.

Fig. 13.4 Protocol for optical mapping. (a) DNA is electrostatically immobilized, (b) digested, DNA and fluorescently imaged. (c) Restriction fragments are sized and (d) assembled into a (e) consensus optical map. (Modified from the OptiGen®, LLC Web site <http://www.optigen.com>)



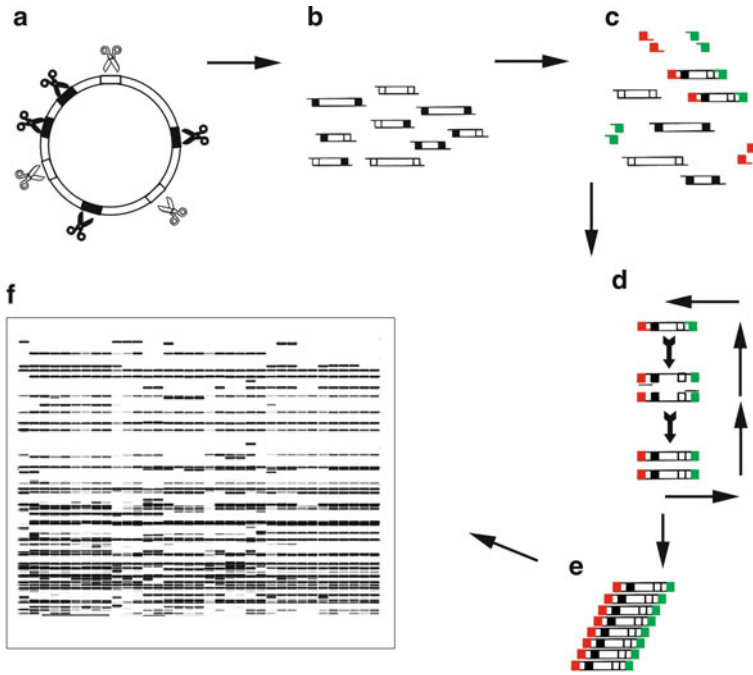


Fig. 13.5 AFLP protocol. (a) Genomic DNA is restricted using two different enzymes to yield fragments (b) with a mixture of restriction sequence ends. (c) Restriction-site specific adapters are ligated to the fragment ends. (d) PCR primers complementary to the adapters with additional bases at their 3' ends restrict amplification to a subset of fragments (e) the sizes of which are then analyzed by electrophoresis (f). (Modified from Rademaker and Savelkoul [84])

Repetitive Sequence-Based PCR

Well before our current level of technology and understanding regarding bacterial genomics, specific DNA sequences were known to be repeated at multiple chromosomal sites in a variety of clinically important pathogens. Enterobacteria were found to contain several hundred copies of repetitive extragenic palindromic (REP) elements and enterobacterial repetitive intergenic consensus (ERIC) sequences [26]. Repeated BOX element sequences were observed in the chromosome of *Streptococcus pneumoniae* [27]. Multiple copies of IS256 were found in staphylococcal genomes [28]. These and other repeat elements represent genomic landmarks of known sequence to which PCR primers may be specifically anchored in an outwardly oriented direction. The resulting amplicons represent inter-repeat distances that do not exceed the capability of the *Taq* polymerase (Fig. 13.6). Thus, strain typing by repetitive sequence-based PCR (rep-PCR) is accomplished by comparing the chromosomal distribution of such repeated sequences as reflected by the resulting pattern of amplicon sizes. Performed under relatively stringent conditions, rep-PCR is much more reproducible than other more generic PCR approaches such as

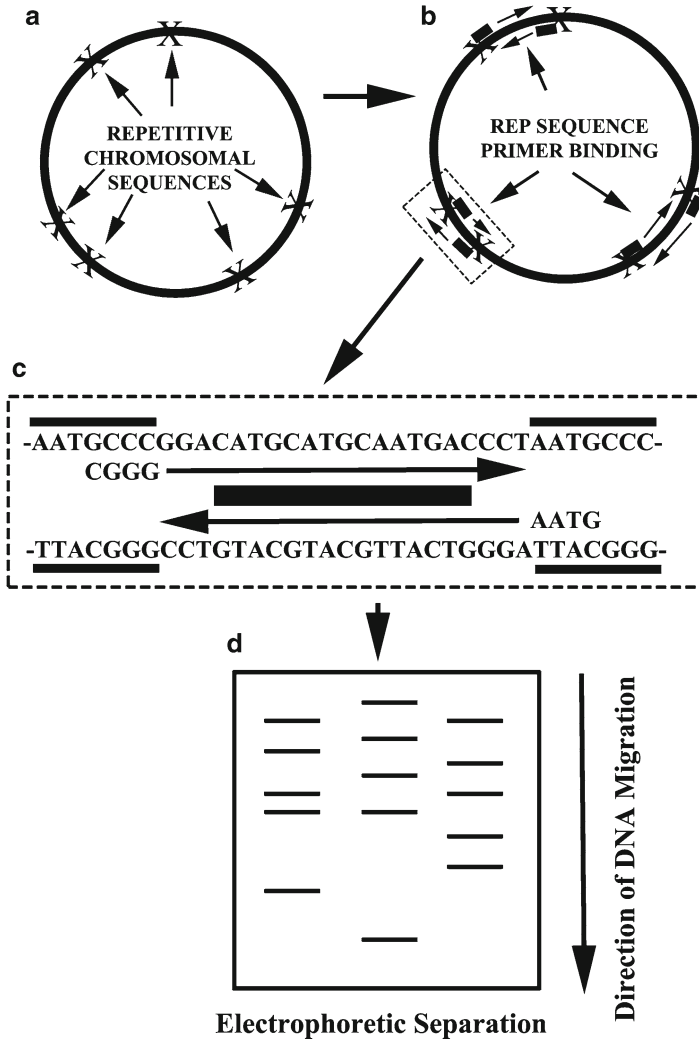


Fig. 13.6 Illustration of rep-PCR. (a) Repetitive sequences in the bacterial chromosome are recognized by outwardly directed primers (b) allowing PCR amplification of the inter-repetitive regions (c) when are then analyzed by electrophoresis (d)

randomly amplified polymorphic DNA (RAPD) and arbitrarily primed-PCR (AP-PCR) which are not considered here [1, 5]. Initial “home brew” efforts at rep-PCR encountered issues such as appropriate primer combinations, PCR conditions, and optimum visualization of amplicon fragment patterns by agarose gel electrophoresis [29]. However, the process has become highly reproducible via commercial automation. The DiversiLab System (bioMérieux) employs optimized protocols, separation of PCR products in a charged microfluidic field (i.e., on a chip) rather

than by conventional agarose gel electrophoresis, and software for data analysis. While in some instances less discriminatory than PFGE [30, 31], rep-PCR remains an interesting typing method although issues regarding database libraries and inter-laboratory data sharing [4] as well as costs associated with the commercial approach are factors to be considered.

PCR Ribotyping

Bacteria typically contain multiple chromosomal copies of rRNA genes. Conventional ribotyping exploits the fact that strain-to-strain differences in the chromosomal regions flanking rRNA genes affect restriction enzyme recognition sites producing different RFLP hybridization patterns with rRNA probes [32]. However, this approach is no longer considered state of the art. PCR-ribotyping, based on primers amplifying polymorphisms in the 16–23S intergenic spacer region, continues to be used as an important tool in the epidemiological monitoring of *C. difficile* [33]. However, it is important to note that the amplicons generated typically include a variety of similar sizes which are a challenge to separate by agarose gel electrophoresis. Nevertheless, the patterns obtained are amenable to databasing and inter-laboratory comparison especially with regard to highly toxigenic strains such as *C. difficile* ribotype 027 [34–36].

Staphylococcal Cassette Chromosome *mec* Typing

Staphylococci resistant to the antibiotic methicillin, especially *S. aureus* (MRSA), represent an infectious disease problem of global concern. Central to this issue is the mobile genetic element staphylococcal cassette chromosome *mec* (SCC*mec*) encoding the altered penicillin-binding protein (known as PBP2a or PBP2') responsible for resistance [37]. Increased understanding of staphylococcal genomics has revealed SCC*mec* variations (termed SCC*mec* types) which differ with regard to their internal organization and total size (<30–>60 kb) [38]. A variety of multiplex PCR approaches have been developed with primers positioned to detect type-specific differences reflected by amplicon banding patterns in agarose gels [39–41]. However, SCC*mec* represents one of the most highly recombinogenic regions in the staphylococcal genome. This is reflected in the increasing complexity associated with newly described SCC*mec* types and subtypes and the multiplex PCR protocols required for their detection (http://www.sccmec.org/Pages/SCC_TypesEN.html) [38]. Nevertheless, SCC*mec* typing represents an important means of studying the element's organization, persistence, and movement in staphylococcal populations. In this context, SCC*mec* type has become a landmark trait in the definition of specific staphylococcal epidemic strains (especially MRSA). However, the method is not discriminating enough to stand alone as an approach to epidemiological monitoring and SCC*mec* differences do not significantly impact anti-staphylococcal chemotherapy [42].

Multiple-Locus VNTR Analysis

Similar to the repetitive sequences discussed earlier (i.e., rep-PCR), advances in bacterial genomics have revealed the presence of chromosomal regions consisting of tandemly repeated sequence “units” varying both in the number and sequence of the individual repeats (Fig. 13.7). These occur by slipped strand mispairing during chromosomal replication resulting in the insertion or deletion of repeat units [43–45]. Bacterial genomes may contain different variable number tandem repeats (VNTR) at multiple chromosomal sites. Properly designed multiplexed PCR primers thus produce multiple-locus VNTR analysis (MLVA) banding patterns by electrophoresis with potential application for strain typing [46]. Finding and validating the epidemiological usefulness of specific MLVA approaches is a deliberative process which varies depending on a number of factors including the degree of VNTR polymorphisms, the organism being analyzed, etc. [3, 46]. Nevertheless, MLVA strain typing has been described for a variety of clinically important bacterial pathogens including *Bacillus anthracis*, *Brucella* spp., *E. coli*, *Legionella pneumophila*, *Leptospira interrogans*, *Mycobacterium tuberculosis*, *P. aeruginosa*, *Yersinia pestis*, *Shigella* spp., *S. aureus*, and *S. pneumonia* (see [46] for a review). This trend has been facilitated by a number of advances including digitized MLVA pattern nomenclature based on VNTR repeat numbers, improved accuracy with pattern visualization by capillary, rather than agarose-gel, electrophoresis, and proper molecular size standards.

Overall, with some exceptions such as PFGE, electrophoretic-based typing methods tend to be relatively simple to perform and also benefit from the potential for decreased cost when agarose-gels are used for analysis. However, it is important to emphasize that strain typing based on electrophoretic banding patterns is primarily a comparison of chromosomal fragment sizes rather than specific genomic content. With the exception of PCR ribotyping and MLVA this is true for both restriction enzyme and PCR-based methods but is especially the case with the former where equivalent-sized fragments in different isolate patterns may or may not represent the same chromosomal sequence. Electrophoresis-based typing approaches are also challenged

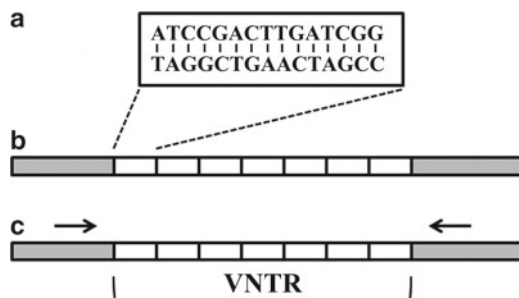


Fig. 13.7 Diagram of a chromosomal VNTR where (a) a sequence unit of “X” base pairs is (b) tandemly repeated “Y” number of times during chromosomal replication. PCR primers anchored to chromosomal regions adjacent to the VNTR (c) allow amplification with subsequent electrophoretic analysis to determine the VNTR “Y” repeat number

regarding issues of typing pattern nomenclature, databasing, and interlaboratory sharing. Nevertheless, as noted earlier, in the context of locally available economic and scientific resources these methods continue to remain of value as options for the epidemiological evaluation of problem bacterial pathogens.

DNA Sequence-Based Methods

Since the bacterial chromosome is the most fundamental molecule of identity in the cell, strain typing based on DNA sequence analysis is the most direct approach to assessing isolate relatedness. Sequence-based approaches have a number of additional advantages over electrophoresis-based typing methods including:

1. Simplicity and reproducibility.

Older molecular methods for epidemiological analysis involve numerous experimental variables including types of equipment, reagents, experimental protocols, etc., all of which affect inter- and intra-laboratory reproducibility. With enough time and effort, any epidemiological method can be standardized as evidenced by classical bacteriophage typing of staphylococci [47] or the success of the nationwide PFGE Pulse-Net System for the investigation of foodborne outbreaks designed by the United States Centers for Disease Control [48]. However, DNA sequence analysis is a more straightforward process that can be performed in a more controlled, uniform, and reproducible manner with specific known chromosomal loci.

2. Data sharing and storage.

Electronic storage and sharing of data from electrophoresis-based typing methods is accomplished using bitmapped (e.g., .tiff, .jpeg) computer images. However, the larger the number of isolates the more unwieldy the process can become. In addition, some form of nomenclature must be used to identify and interrelate isolate banding patterns. With large data sets, the use of appropriate computer software is essential to accomplish this process. However, the framework for data sharing, storage, and retrieval is necessarily based on visual images and the limits that format imposes. Conversely, nucleotide sequences represent simple, highly portable, quaternary data that can much more easily be shared, stored, and retrieved.

3. Data interpretation and detection of significant differences.

As will be discussed more fully later, the most crucial aspect of any typing method is its ability to detect significant (epidemiologically-relevant) differences between isolates. While the goal of molecular typing is a comparison of chromosomal similarity, electrophoretic banding patterns only indirectly address this issue. Despite computer programs which can assist the process, there is always an element of end user judgment that can affect the final evaluation. In contrast, nucleotide sequence data allows direct and unambiguous genomic comparison.

Advances in DNA sequencing and the rapidly expanding database of sequenced microbial genomes have served as the foundation for a variety of typing approaches which can generally be categorized as single-locus, multiple-locus, or whole-genome sequence typing.

Single-Locus Sequence Typing

Since the genome of bacterial pathogens is mega-base in size, it is remarkable to think that a single locus of ca. 1,000 bases could contain sufficient information to be epidemiologically relevant. Nevertheless, three instances where this is the case are detailed below.

***S. aureus* Protein A Typing**

The production of protein A is a hallmark characteristic of *S. aureus*. Thus, the gene for protein A (*S. aureus* protein A, *spa*) is found in all *S. aureus* strains. The 3' end of the *spa* locus (i.e., the polymorphix “X” region) contains a 24-bp VNTR which can be amplified with appropriate primers (e.g., see Fig. 13.7) and sequenced to determine the specific *spa* type. Software packages such as StaphType (Ridom GmbH, Münster, Germany) and BioNumerics (Applied Maths NV, Sint-Martens-Latem, Belgium) are available to assist with the sequence analysis process. Numerous studies have shown that comparisons of *S. aureus spa* types, facilitated by an Internet-based *spa* server (<http://spaserver.ridom.de>), provide epidemiologically-relevant information that correlates well with other typing methods such as PFGE [42, 49–52]. In Europe this has led to the formally organized use of *spa* typing in the epidemiological monitoring of specific *S. aureus* strains (i.e., SeqNet; <http://www.seqnet.org>) involving 60 laboratories from 39 countries.

***Streptococcus pyogenes* M Protein (*emm*) Typing**

The cell surface M protein is an important virulence factor in *S. pyogenes* [53]. Genomic analysis has revealed that the M protein locus (*emm*) is variable and can encode at least 100 different M protein types which were initially detected and cataloged serologically. However, PCR primers flanking the hypervariable region of the *emm* gene allow direct sequencing to determine specific isolate *emm* types. As a result, sequence-based *emm* typing is currently the most widely used approach to group A streptococcal epidemiology [53–56]. As with *S. aureus spa* typing, *emm* typing is facilitated by an Internet-based server (hosted by the US Centers for Disease Control) which houses the *S. pyogenes emm* sequence database (<http://www.cdc.gov/ncidod/biotech/strep/strepblast.htm>). This resource has allowed the CDC to follow specific *S. pyogenes* epidemiological trends such as the proportion

of *emm* types contributing to specific disease in different global regions (e.g., Africa, Asia, Latin America, Middle East, Australia/Pacific Island) (http://www.cdc.gov/ncidod/biotech/strep/emmtypes_proportions.htm).

***mec*-Associated Direct Repeat Unit Typing**

In 1991, Ryffell et al. [57] identified a cluster of repeated imperfect 40-bp sequences (i.e., direct repeat units, *dru*) adjacent to *IS431* within the *SCCmec* element of *S. aureus* isolates. While the *dru* VNTR is absent in a minority of MRSA isolates [58], its constant location in different *SCCmec* types of both coagulase-positive and -negative staphylococcal species represents a valuable and stable internal *SCCmec* characteristic [59]. As with staphylococcal *spa* typing, properly positioned PCR primers allow amplification and sequencing of the *dru* region. Software such as BioNumerics (Applied Maths NV, Sint-Martens-Latem, Belgium) and DruID (<http://www.mystrains.com/druid>) is available to assist with assignment of *dru* types the central repository for which is an Internet-based server (<http://www.dru-typing.org>). As with *SCCmec* typing, *dru* typing has become an increasingly important means of characterizing the persistence and movement of *SCCmec* in staphylococcal populations. While not discriminating enough to serve as a standalone approach to epidemiological monitoring, *dru* typing has proven helpful in assessing movement of *SCCmec* in staphylococcal populations and in subtyping highly clonal (i.e., difficult to differentiate) staphylococcal strains [58, 60, 61]. In addition, a combination of *dru* typing and analysis of *SCCmec* (*ccr*) recombinase genes, has proved highly informative with regard to the phylogeny of specific *S. aureus* MRSA strains [62, 63].

Multi-Locus Sequence Typing

Since its initial description in 1998 [64] multi-locus sequence typing (MLST) has become one of the most popular approaches to microbial strain typing with demonstrated utility for a wide range of clinically relevant pathogens (<http://www.mlst.net/databases/default.asp>). The method is based on PCR amplification and subsequent sequencing of the internal regions (450–500 bp) of multiple essential (i.e., housekeeping) genes. Seven genes are typically employed, the sequences of which are assigned numeric allelic designations (Fig. 13.8a). Individual strains are thus characterized by a seven digit MLST sequence type (ST). For a given organism, individual STs are interrelated based on an algorithm that identifies a parent or “founding” ST as that which has the greatest number of single-locus variants (SLV). Using online graphic tools (eBURST; <http://saureus.mlst.net/eburst/>) the STs can be further grouped into clonal complexes (CC) where members of the group share a minimum of five or six of the seven allelic designations (Fig. 13.8b). The highly portable nature of such data and availability of online databases has facilitated the use of MLST for global epidemiological analysis [65, 66] and long-term

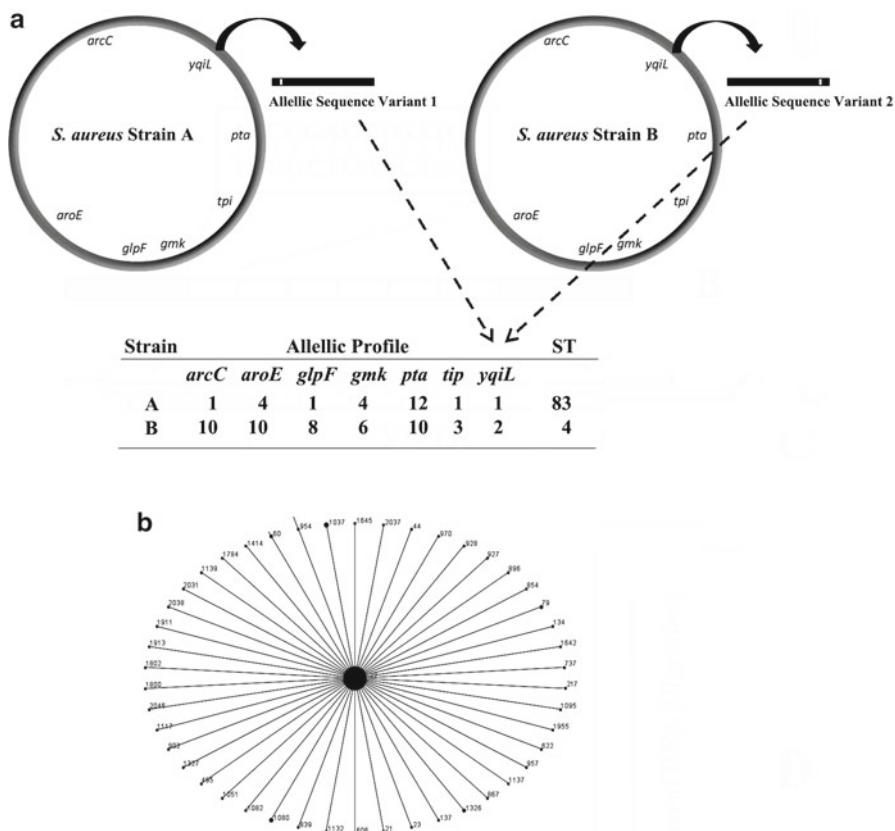


Fig. 13.8 Illustration of MLST with hypothetical *S. aureus* strains A and B depicting the seven chromosomal housekeeping genes with an example of allelic differences (e.g., in *yqiL*) constituting different STs (a). An eBURST example of a clonal complex with central founding ST and associated SLVs is also shown (b)

(i.e., phylogenetic) investigation of bacterial lineages [67–69]. However, the method has not found routine clinical application since MLST housekeeping gene sequences are too conserved to reliably differentiate the closely related isolates typically encountered during short-term outbreaks (e.g., MRSA and MSSA could both have the same ST). The time and cost associated with multiple-gene sequencing (a total of ca. 3–4 kb for 7 loci) has also been a disincentive to routine use.

Other Multi-Locus Approaches: Hybridization-Based Typing

As noted earlier, only a small number of loci may be simultaneously queried using DNA hybridization with restriction fragment-based typing. However, this is not the case with array-based methods where thousands of specific oligonucleotide probes

(e.g., representing species-specific, antimicrobial resistance, and virulence-associated genes) can be anchored to solid surfaces such as glass, plastic, or silicone chips. The hybridization pattern of labeled genomic DNA from isolates to be analyzed thus has the potential to provide a wealth of information regarding genomic content (e.g., the presence or absence of specific genes). Depending on the length of the anchored array sequences even minor sequence variations including insertions, deletions, or changes in a single base of a sequence (single nucleotide polymorphism, SNP) can be detected. The power of this approach has been applied to the characterization of a wide variety of clinically relevant organisms [3, 70–72]. However, while microarrays have the potential for high-throughput genomic analysis their use is not cost-effective for routine clinical use. In addition, a high level of technical expertise is required especially for data analysis which can be complicated by “background” noise due to partial hybridization, etc. An interesting variation on microarray analysis, developed by Luminex (Luminex, Austin), involves the use of flow cytometry to detect hybridization of test DNA to fluorochrome-labeled beads conjugated with specific sequence probes [73]. However, the utility of this suspension-based approach for strain typing remains to be thoroughly evaluated.

Whole Genome Sequence Typing

As noted earlier, the goal of molecular strain typing is epidemiological assessment based on the most fundamental molecule of identity in the cell—the bacterial chromosome. Thus, the ability to compare whole genome sequences (WGS) represents the ultimate molecular typing approach. While this was impossible with older dideoxy/chain termination sequencing technology [74], newer (i.e., next-generation sequencing, NGS) methods have made this goal a reality. The technology behind NGS is discussed in Chap. 37 of this book and is not considered further here. However, from a strain typing standpoint it is important to note that revolutionary developments in NGS have made (WGS) possible with benchtop instrumentation such as the Ion Torrent PGM (Life Technologies, Guilford) and the MiSeq (Illumina, San Diego). Such instrumentation now allows WGS to be completed in only a few hours with extensive multifold coverage allowing isolates to be compared down to the level of SNPs. However, for NGS, as for previous sequencing iterations, the critical issues are throughput, quality, read length and cost. All of these are currently in a state of flux as commercial technology improves and positions itself in the scientific marketplace. However, an example of these concerns is seen in the application of WGS to the analysis of a recent *E. coli* outbreak in Europe which claimed more than 50 lives. One report, based on sequencing with the Ion Torrent PGM, concluded that the outbreak strain and an older 2001 isolate arose from a common ancestor with the current outbreak resulting from gene loss [75]. However, another study, using single-molecule real-time (third-generation) DNA sequencing (Pacific Biosciences, Menlo Park), proposed that the outbreak strain evolved by acquiring the gene for Shiga toxin [76]. These conflicting reports underscore what will clearly become the greatest need as

WGS-based strain typing rapidly develops—bioinformatic data interpretation and analysis. Nevertheless, these are exciting “problems” to have and the scientific stage is clearly set for additional remarkable developments in this most fundamental approach to determine isolate epidemiological interrelationships.

Non-Sequence-Based Whole Cell Typing

While strain typing is firmly directed toward sequence-based analysis, two whole-cell methods deserve mention: Raman spectroscopy and MALDI-TOF mass spectrometry. Both technologies are not new but are finding renewed emphasis in applications for application of strain typing.

The SpectraCell RA Bacterial Strain Analyzer (River Diagnostics, Rotterdam) employs Raman spectroscopy for isolate characterization. Sir C.V. Raman received the Nobel Prize in Physics in 1930 for his discovery of this light-scattering technology. Since every molecule in the cell contributes to the generated spectrum of scattered laser light, in principle different bacterial strains would be expected to generate different Raman spectra while highly related isolates would not. Thus, the SpectraCell system seeks to accomplish strain typing based on the quantitation of these spectral measurements. Early reports suggest that the method has promise in the typing of problem pathogens such as *P. aeruginosa* and *S. aureus* [77, 78]. However, uniformity of pre-analysis bacterial growth conditions as well as method reproducibility and discriminatory power are key issues for the future of this approach to typing.

MALDI-TOF mass spectrometry is considered in detail in Chap. 10 of this book. The method has generated intense interest as a means of rapid microbial identification via the detection of unique cellular protein biomarkers. As a related issue, MALDI-TOF is also being investigated as an approach to bacterial strain typing [79–82]. However, as with Raman spectroscopy, experimental parameters (e.g., loading of the target plate, matrix composition) must be carefully controlled with optimized post-processing and analysis of the mass spectra. Nevertheless, as an adjunct to microbial identification, strain typing is a logical goal for MALDI-TOF technology which will most certainly see additional refinement and application in the future.

Strain Typing in the Context of the Epidemiological Window

In the final analysis, regardless of the quantity or quality of strain typing data, the issue ultimately comes down to data interpretation. In this context, it is important to note that while the term “molecular” epidemiology implies a precise process, this is not always the case regardless of the method employed since such investigations have an unavoidable context-driven component. A variety of environmental factors as well as interaction between the host and infectious agent may all influence the

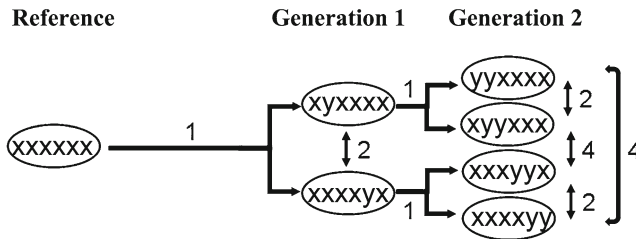


Fig. 13.9 Depiction of the interrelationships between a reference genome containing six characteristics (each designated “x”) with two subsequent generations each experiencing sequential single genetic events ($x \rightarrow y$). Within and between each generation, the number of resulting genetic differences is indicated

course of disease transmission. In addition, infectious disease issues benefiting from epidemiological evaluation are not typically given advance warning. Thus, identification of the index patient in a particular outbreak is a common problem in epidemiological analysis. Nevertheless, the analysis must be conducted in the context of the available isolate data (i.e., the epidemiological window [5, 83]) which, unfortunately, does not always include the outbreak source. Thus, the analytical process is commonly one of attempting to work backward in time which, depending on the available data, may necessitate conclusions based on probabilities rather than hard data. For this reason, regardless of the sophistication of the typing approach, epidemiological analysis commonly contains an element of educated guess. Nevertheless, for most clinical scenarios (i.e., outbreak investigation) the key issue is whether or not a series of bacterial isolates are the result of person-to-person transfer. At the heart of this question is the concept of significant difference which for chromosomal comparison relates to epidemiologically relevant genomic clock speed. In the absence of an index case or isolate, all strain typing methods are challenged as the opportunity for chromosomal change over time increases the potential for genetic distance between epidemiologically related isolates. As illustrated in Fig. 13.9, if one considers a simple reference genome of six characteristics (“x”) evolving through two generations with sequential genetic events of unknown complexity ($x \rightarrow y$), the resulting second-generation genomes would vary from each other by four differences. The potential complexity of the scenario obviously increases further over time. These issues underscore the potential difficulties one may encounter in attempting to discern lineages of infectious agent transmission, regardless of the typing method employed. Thus, for an optimum outcome (e.g., in an outbreak setting) analysis of strain typing data and its epidemiological relevance requires knowledge of (1) the limitations of the typing method, (2) the etiological agent (e.g., genomic clock speed), and (3) the clinical setting within which the issue is being studied.

For the future, it is exciting to consider the advances in strain typing that will continue to be made. The persistence and spread of problem pathogens in patient populations will obviously continue to occur. Thus, perhaps the most important point of all is to emphasize that, more than ever before, strain typing and epidemiological

analysis benefit from communication. It is when all interested parties participate (e.g., physician, nurse, infection control specialist, laboratory) that the epidemiological educated guess is most likely to be correct, and that most certainly is a key goal for the treatment of infectious disease both now and in the future.

References

1. Tenover FC, Arbeit RD, Goering RV (1997) How to select and interpret molecular strain typing methods for epidemiological studies of bacterial infections: a review for healthcare epidemiologists. *Infect Control Hosp Epidemiol* 18(6):426–439
2. Riley LW (2004) *Molecular epidemiology of infectious disease: principles and practices*. ASM, Washington, DC
3. Chen Y, Brown E, Knabel SJ (2011) Molecular Epidemiology of Foodborne Pathogens. In: Wiedmann M, Zhang W (eds) *Genomics of foodborne bacterial pathogens*. Springer, New York, pp 403–453
4. Van Belkum A, Tassios PT, Dijkshoorn L et al (2007) Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin Microbiol Infect* 13(Suppl 3):1–46
5. Goering RV (2000) The molecular epidemiology of nosocomial infection: past, present, and future. *Rev Med Microbiol* 11:145–152
6. Goering RV (2000) Molecular strain typing for the clinical laboratory: current application and future direction. *Clin Microbiol News* 22:169–173
7. Halpern D, Chiapello H, Schbath S et al (2007) Identification of DNA motifs implicated in maintenance of bacterial core genomes by predictive modeling. *PLoS Genet* 3(9):1614–1621
8. Lindsay JA, Moore CE, Day NP et al (2006) Microarrays reveal that each of the ten dominant lineages of *Staphylococcus aureus* has a unique combination of surface-associated and regulatory genes. *J Bacteriol* 188(2):669–676
9. Lindsay JA (2010) Genomic variation and evolution of *Staphylococcus aureus*. *Int J Med Microbiol* 300(2–3):98–103
10. Goering RV (2002) The influence of genomics on the molecular epidemiology of nosocomial pathogens. In: Shaw KJ (ed) *Pathogen genomics: impact on human health*. Humana, Totowa, pp 113–131
11. Bialkowska-Hobrzanska H, Jaskot D, Hammerberg O (1990) Evaluation of restriction endonuclease fingerprinting of chromosomal DNA and plasmid profile analysis for characterization of multiresistant coagulase-negative staphylococci in bacteremic neonates. *J Clin Microbiol* 28:269–275
12. Tenover FC, Akerlund T, Gerding DN et al (2011) Comparison of strain typing results for *Clostridium difficile* isolates from North America. *J Clin Microbiol* 49(5):1831–1837
13. Schwartz EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98(3):503–517
14. Van Belkum A (2003) High-throughput epidemiologic typing in clinical microbiology. *Clin Microbiol Infect* 9:86–100
15. Thorne N, Borrell S, Evans J et al (2011) IS6110-based global phylogeny of *Mycobacterium tuberculosis*. *Infect Genet Evol* 11(1):132–138
16. Schwartz DC, Saffran W, Welsh J, Haas R, Goldenberg M, Cantor CR (1983) New techniques for purifying large DNA's and studying their properties and packaging. *Cold Spring Harbor Symp Quant Biol* 47:189–195
17. Schwartz DC, Koval M (1989) Conformational dynamics of individual DNA molecules during gel electrophoresis. *Nature* 338:520–522

18. Goering RV, Ribot EM, Gerner-Smidt P (2011) Pulsed-field gel electrophoresis: laboratory and epidemiologic considerations for interpretation of data. In: Persing DH, Tenover FC, Tang YW, Nolte FS, Hayden RT et al (eds) *Molecular microbiology*, 2nd edn. ASM, Washington, DC, pp 167–177
19. Goering RV (2010) Pulsed field gel electrophoresis: a review of application and interpretation in the molecular epidemiology of infectious disease. *Infect Genet Evol* 10(7):866–875
20. McDougal LK, Steward CD, Killgore GE, Chaitram JM, McAllister SK, Tenover FC (2003) Pulsed-field gel electrophoresis typing of oxacillin-resistant *Staphylococcus aureus* isolates from the United States: establishing a national database. *J Clin Microbiol* 41(11):5113–5120
21. Goering RV, McDougal LK, Fosheim GE, Bonnstedter KK, Wolter DJ, Tenover FC (2007) Epidemiologic distribution of the arginine catabolic mobile element among selected methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* isolates. *J Clin Microbiol* 45(6):1981–1984
22. Schwan WR, Briska A, Stahl B et al (2010) Use of optical mapping to sort uropathogenic *Escherichia coli* strains into distinct subgroups. *Microbiology* 156(Pt 7):2124–2135
23. Petersen RF, Litrup E, Larsson JT et al (2011) Molecular characterization of *Salmonella typhimurium* highly successful outbreak strains. *Foodborne Pathog Dis* 8(6):655–661
24. Goering RV, Stemper ME, Shukla SK, Foley SL (2011) Restriction analysis techniques. In: Foley SL, Chen AY, Simjee S, Zervos MJ (eds) *Molecular techniques for the study of hospital acquired infection*. Wiley-Blackwell, Hoboken, N.J., pp 135–144
25. Melles DC, Schouls L, Francois P et al (2009) High-throughput typing of *Staphylococcus aureus* by amplified fragment length polymorphism (AFLP) or multi-locus variable number of tandem repeat analysis (MLVA) reveals consistent strain relatedness. *Eur J Clin Microbiol Infect Dis* 28(1):39–45
26. Versalovic J, Koeuth T, Lupski JR (1991) Distribution of repetitive DNA sequences in Eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* 19: 6823–6831
27. Van Belkum A, Sluijter M, De Groot R, Verbrugh H, Hermans PWM (1996) Novel BOX repeat PCR assay for high-resolution typing of *Streptococcus pneumoniae* strains. *J Clin Microbiol* 34:1176–1179
28. Deplano A, Vaneechoutte M, Verschraegen G, Struelens MJ (1997) Typing of *Staphylococcus aureus* and *Staphylococcus epidermidis* strains by PCR analysis of Inter-IS256 spacer length polymorphisms. *J Clin Microbiol* 35:2580–2587
29. Frye SR, Healy M (2011) Repetitive sequence-based PCR typing of bacteria and fungi. In: Persing DH, Tenover FC, Tang YW, Nolte FS, Hayden RT, Van Belkum A (eds) *Molecular microbiology: diagnostic principles and practice*. ASM, Washington, DC, pp 199–212
30. Ross TL, Merz WG, Farkosh M, Carroll KC (2005) Comparison of an automated repetitive sequence-based PCR microbial typing system to pulsed-field gel electrophoresis for analysis of outbreaks of methicillin-resistant *Staphylococcus aureus*. *J Clin Microbiol* 43(11): 5642–5647
31. Roussel S, Felix B, Colaneri C et al (2010) Semi-automated repetitive-sequence-based polymerase chain reaction compared to pulsed-field gel electrophoresis for *Listeria monocytogenes* subtyping. *Foodborne Pathog Dis* 7(9):1005–1012
32. Bouchet V, Huot H, Goldstein R (2008) Molecular genetic basis of ribotyping. *Clin Microbiol Rev* 21(2):262–273
33. Tenover FC, Novak-Weekley S, Woods CW et al (2010) Impact of strain types on detection of toxigenic *Clostridium difficile*: comparison of molecular diagnostic and enzyme immunoassay approaches. *J Clin Microbiol* 48(10):3719–3724
34. Valiente E, Dawson LF, Cairns MD, Stabler RA, Wren BW (2011) Emergence of new PCR-ribotypes from the hypervirulent *Clostridium difficile* 027 lineage. *J Med Microbiol* 61 (Pt 1):49–56
35. Solomon K, Fanning S, McDermott S et al (2011) PCR ribotype prevalence and molecular basis of macrolide-lincosamide-streptogramin B (MLSB) and fluorquinolone resistance in Irish clinical *Clostridium difficile* isolates. *J Antimicrob Chemother* 66(9):1976–1982

36. Gerding DN (2010) Global epidemiology of *Clostridium difficile* infection in 2010. *Infect Control Hosp Epidemiol* 31(Suppl 1):S32–S34
37. Katayama Y, Ito T, Hiramatsu K (2000) A new class of genetic element, staphylococcus cassette chromosome *mec*, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 44:1549–1555
38. IWG-SCC (2009) Classification of staphylococcal cassette chromosome *mec* (SCC*mec*): guidelines for reporting novel SCC*mec* elements. *Antimicrob Agents Chemother* 53(12):4961–4967
39. Kondo Y, Ito T, Ma XX et al (2007) Combination of multiplex PCRs for staphylococcal cassette chromosome *mec* type assignment: rapid identification system for *mec*, *ccr*, and major differences in junkyard regions. *Antimicrob Agents Chemother* 51(1):264–274
40. Milheirico C, Oliveira DC, De Lencastre H (2007) Multiplex PCR strategy for subtyping the staphylococcal cassette chromosome *mec* type IV in methicillin-resistant *Staphylococcus aureus*: ‘SCC*mec* IV multiplex’. *J Antimicrob Chemother* 60(1):42–48
41. Oliveira DC, De Lencastre H (2002) Multiplex PCR strategy for rapid identification of structural types and variants of the *mec* element in methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 46(7):2155–2161
42. Deurenberg RH, Stobberingh EE (2008) The evolution of *Staphylococcus aureus*. *Infect Genet Evol* 8:747–763
43. Van Belkum A, Scherer S, van Alphen L, Verbrugh H (1998) Short-sequence DNA repeats in prokaryotic genomes. *Microbiol Mol Biol Rev* 62:275–293
44. Lindstedt BA (2005) Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. *Electrophoresis* 26(13):2567–2582
45. Hammerschmidt S, Muller A, Sillmann H et al (1996) Capsule phase variation in *Neisseria meningitidis* serogroup B by slipped-strand mispairing in the polysialyltransferase gene (*siaD*): correlation with bacterial invasion and the outbreak of meningococcal disease. *Mol Microbiol* 20(6):1211–1220
46. Pourcel C, Vergnaud G (2011) Strain typing using multiple “variable number of tandem repeat” analysis and genetic element CRISPR. In: Persing DH, Tenover FC, Tang YW, Nolte FS, Hayden RT, Van Belkum A (eds) *Molecular microbiology: diagnostic principles and practice*, 2nd edn. ASM, Washington, DC, pp 179–197
47. Bannerman TL, Hancock GA, Tenover FC, Miller JM (1995) Pulsed-field gel electrophoresis as a replacement for bacteriophage typing of *Staphylococcus aureus*. *J Clin Microbiol* 33:551–555
48. Swaminathan B, Barrett TJ, Hunter SB, Tauxe RV (2001) PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerg Infect Dis* 7:382–389
49. Harmsen D, Claus H, Witte W et al (2003) Typing of methicillin-resistant *Staphylococcus aureus* in a university hospital setting by using novel software for *spa* repeat determination and database management. *J Clin Microbiol* 41:5442–5448
50. Church DL, Chow BL, Lloyd T, Gregson DB (2011) Comparison of automated repetitive-sequence-based polymerase chain reaction and *spa* typing versus pulsed-field gel electrophoresis for molecular typing of methicillin-resistant *Staphylococcus aureus*. *Diagn Microbiol Infect Dis* 69:30–37
51. Frenay HME, Bunschoten AE, Schouls LM et al (1996) Molecular typing of methicillin-resistant *Staphylococcus aureus* on the basis of protein A gene polymorphism. *Eur J Clin Microbiol Infect Dis* 15:60–64
52. Shopsin B, Gomez M, Montgomery SO et al (1999) Evaluation of protein A gene polymorphic region DNA sequencing for typing of *Staphylococcus aureus* strains. *J Clin Microbiol* 37:3556–3563
53. Bessen DE (2009) Population biology of the human restricted pathogen, *Streptococcus pyogenes*. *Infect Genet Evol* 9:581–593
54. Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR (2009) Global *emm* type distribution of group A streptococci: systematic review and implications for vaccine development. *Lancet Infect Dis* 9:611–616

55. Wajima T, Murayama SY, Sunaoshi K, Nakayama E, Sunakawa K, Ubukata K (2008) Distribution of *emm* type and antibiotic susceptibility of group A streptococci causing invasive and noninvasive disease. *J Med Microbiol* 57:1383–1388
56. Tanaka D, Gyobu Y, Kodama H et al (2002) *emm* typing of group A streptococcus clinical isolates: identification of dominant types for throat and skin isolates. *Microbiol Immunol* 46(7):419–423
57. Ryffel C, Bucher R, Kayser FH, Berger-Bächi B (1991) The *Staphylococcus aureus mec* determinant comprises an unusual cluster of direct repeats and codes for a gene product similar to the *Escherichia coli sn*-glycerophosphoryl diester phosphodiesterase. *J Bacteriol* 173:7416–7422
58. Ionescu R, Mediavilla JR, Chen L et al (2010) Molecular characterization and antibiotic susceptibility of *Staphylococcus aureus* from a multidisciplinary hospital in Romania. *Microb Drug Resist* 16:263–272
59. Goering RV, Morrison D, Al-Doori Z, Edwards GF, Gemmell CG (2008) Usefulness of *mec*-associated direct repeat unit (*dru*) typing in the epidemiological analysis of highly clonal methicillin-resistant *Staphylococcus aureus* in Scotland. *Clin Microbiol Infect* 14:964–969
60. Fessler A, Scott C, Kadlec K, Ehrlich R, Monecke S, Schwarz S (2010) Characterization of methicillin-resistant *Staphylococcus aureus* ST398 from cases of bovine mastitis. *J Antimicrob Chemother* 65:619–625
61. Shore AC, Rossney AS, Kinnevey PM et al (2010) Enhanced discrimination of highly clonal ST22-methicillin-resistant *Staphylococcus aureus* IV isolates achieved by combining *spa*, *dru*, and pulsed-field gel electrophoresis typing data. *J Clin Microbiol* 48(5):1839–1852
62. Smyth DS, Wong A, Robinson DA (2011) Cross-species spread of SCC*mec* IV subtypes in staphylococci. *Infect Genet Evol* 11:446–453
63. Smyth DS, McDougal LK, Gran FW et al (2010) Population structure of a hybrid clonal group of methicillin-resistant *Staphylococcus aureus*, ST239-MRSA-III. *PLoS One* 5:e8582
64. Maiden MCJ, Bygraves JA, Feil E et al (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* 95:3140–3145
65. Feil EJ, Spratt BG (2001) Recombination and the population structures of bacterial pathogens. *Annu Rev Microbiol* 55:561–590
66. Aanensen DM, Spratt BG (2005) The multilocus sequence typing network: mlst.net. *Nucleic Acids Res* 33(Web Server issue):W728–W733
67. Yan Y, Cui Y, Han H et al (2011) Extended MLST-based population genetics and phylogeny of *Vibrio parahaemolyticus* with high levels of recombination. *Int J Food Microbiol* 145:106–112
68. Ch'ng SL, Octavia S, Xia Q et al (2011) Population structure and evolution of pathogenicity of *Yersinia pseudotuberculosis*. *Appl Environ Microbiol* 77:768–775
69. Litrup E, Torpdahl M, Malorny B, Huehn S, Christensen H, Nielsen EM (2010) Association between phylogeny, virulence potential and serovars of *Salmonella enterica*. *Infect Genet Evol* 10:1132–1139
70. Miller MB, Tang YW (2009) Basic concepts of microarrays and potential applications in clinical microbiology. *Clin Microbiol Rev* 22:611–633
71. Musser JM, Shelburne SA III (2009) A decade of molecular pathogenomic analysis of group A Streptococcus. *J Clin Invest* 119:2455–2463
72. Li W, Raoult D, Fournier PE (2009) Bacterial strain typing in the genomic era. *FEMS Microbiol Rev* 33:892–916
73. Dunbar SA (2006) Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. *Clin Chim Acta* 363:71–82
74. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74:5463–5467
75. Mellmann A, Harmsen D, Cummings CA et al (2011) Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. *PLoS One* 6:e22751

76. Rasko DA, Webster DR, Sahl JW et al (2011) Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. *N Engl J Med* 365:709–717
77. Willemse-Erix DF, Jachtenberg JW, Schut TB et al (2010) Towards Raman-based epidemiological typing of *Pseudomonas aeruginosa*. *J Biophotonics* 3:506–511
78. Wulf MW, Willemse-Erix D, Verduin CM, Puppels G, van Belkum A, Maquelin K (2012) The use of Raman spectroscopy in the epidemiology of methicillin-resistant *Staphylococcus aureus* of human- and animal-related clonal lineages. *Clin Microbiol Infect* 18(2):147–152
79. Moura H, Woolfitt AR, Carvalho MG et al (2008) MALDI-TOF mass spectrometry as a tool for differentiation of invasive and noninvasive *Streptococcus pyogenes* isolates. *FEMS Immunol Med Microbiol* 53:333–342
80. Wolters M, Rohde H, Maier T et al (2011) MALDI-TOF MS fingerprinting allows for discrimination of major methicillin-resistant *Staphylococcus aureus* lineages. *Int J Med Microbiol* 301:64–68
81. Williamson YM, Moura H, Woolfitt AR et al (2008) Differentiation of *Streptococcus pneumoniae* conjunctivitis outbreak isolates by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Appl Environ Microbiol* 74:5891–5897
82. Dieckmann R, Malorny B (2011) Rapid screening of epidemiologically important *Salmonella enterica* subsp. *enterica* serovars by whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Appl Environ Microbiol* 77(12):4136–4146
83. Goering RV (1998) The molecular epidemiology of nosocomial infection: an overview of principles, application, and interpretation. In: Specter S, Bendinelli M, Friedman H (eds) *Rapid detection of infectious agents*, 1st edn. Plenum, New York, pp 131–157
84. Rademaker JL, Savelkoul P (2004) PCR amplification-based microbial typing. In: Persing DH, Tenover FC, Versalovic J, Tang YW, Unger ER, Relman DA et al (eds) *Molecular microbiology: diagnostic principles and practice*. ASM, Washington, DC, pp 197–221