# The Microbe: The Basics of Structure, Morphology, and Physiology as They Relate to Microbial Characterization and Attribution

**2**

Joany Jackman

**Abstract**

This chapter is meant to (1) review classical methods used to characterize and classify microbes and (2) introduce new molecular methods used in microbial characterization. The fundamental composition of microbes is discussed as well as their importance in classification of microbes into genus and species. Classical microbiological methods in general seek to define the common features of specific bacterial groups as a means of classification and identification of microbes. Thus, the focus was to describe the common features which discriminated closely related groups of organisms. In contrast, the newer molecular methods often seek to expand the classification of microbes not only as a means to organize microbial phylogeny but also to differentiate signatures between microbes identified within a species in greater detail. Molecular biology tools are used both as an adjunct to established methods and as replacement for classical methods for detection, discrimination, or identification of bacterial and viral species.

## 2.1 Introduction

As discussed in Chap. 1, the application of forensics to microbiology is important to public security and public health. A large component of threat reduction for biological attacks involves the use of appropriate antibiotics and/or prophylactic vaccination. Application of these mitigating technologies requires some foreknowledge of the threat; hence, one component of microbial forensics is rapid identification of the threat. A second component of the microbial forensics is careful examination with the goal of attribution since encoded in the concept is a law enforcement application. In addition to the chapters in this book, there are several outstanding reviews which address the requirements and gaps in current bioforensics for human and plant biothreats [1–3]. One major difference in the application of forensic assays for public safety and for law enforcement is the requirement for speed and specificity. Assays which impact public safety must be rapid and be suitable for presumptive diagnostic purposes. Presumptive diagnostics are those which are used as an adjunct to clinical

J. Jackman, Ph.D. (✉)
The Johns Hopkins University Applied
Physics Laboratory, Laurel, MD, USA
e-mail: Joany.Jackman@jhuapl.edu

tests and to assist the physician in patient evaluation. Tests which can be used at point of care or at the scene are preferred so that rapid treatment or triage of susceptible populations can be emplaced where they will do the most good. Tests which describe attribution characteristics related to a crime can be slower but must be highly specific. This chapter describes the features of microbes which are used by microbiologists for identification and categorization of biological organisms and intrinsic features of bacteria which can be used in forensic applications for attribution.

The concept of attribution, a critical component of microbial forensics, differs in approach from diagnostics in that the focus of analysis is not necessarily limited to disease-producing organisms. Attribution includes not only disease-producing organisms but organisms which may hitchhike with the sample of interest such as phages and other environmental contaminants (pollen, dust, plant/animal DNA). Likewise, metabolic products and growth media residue are important to attribution and may be loosely associated with collected samples. Classical microbial methods do not address these components, but many of the newer tools for molecular biology and physical chemistry can be used to include characterization of these materials.

## 2.2 How Microbiologists Look at Microbes

Classical microbiological tools are founded in observational techniques for bacteria, some of which originated with the invention of the microscope by Leeuwenhoek in 1674. Notably, less than 5% and perhaps as little as 1% of all bacteria can be cultured in the laboratory. If one includes all viruses infecting all species on earth, the number becomes lower by several orders of magnitude. Hence, one caveat of classical microbial methods is that the tools used to describe microbial development and provide systematic organization are based on a statistically minor portion of all bacteria [4]. This is acceptable for medically important species since classical microbiological techniques, especially those adapted to the clinical laboratory, are biased for analysis

of human pathogens. When microbial analysis techniques require description and typing of environmental organisms rather than microbial pathogens, many laboratory methods especially clinical techniques may be inappropriate or give false information. This is especially true for automated laboratory tools which require matching to known signatures in the database. Environmental microbes, since they do not produce disease and therefore are not isolated from patient samples, are not frequently represented in clinical databases.

### 2.2.1 Classical Methods Applied to Forensic Characterization of Unknown Samples

Classical microbiology uses both gross and microscopic morphology to identify microbes. Gross morphology includes colony shape, size, and surface features (Fig. 2.1). For example, *Bacillus atropheu*s strain *globigii* produces an orange-pigmented colony on tryptic soy agar but produces small white colonies on other media. The structures assigned to bacteria – cocci (round), bacilli (rods), or spirochetes (corkscrew) – can be readily seen via light microscopy with no sample preparation. Generally, bacterial isolates are further differentiated at the time of microscopic examination by staining. The gram stain is one of the most useful and commonly used tools to differentiate bacteria beyond the genus level. This staining procedure named for its inventor, Hans Christian Gram, supplies both biochemical information about the composition of bacteria (see Chap. 3 for more detail concerning the biochemistry of bacterial cell walls) and special information regarding the distribution of chemicals within the cell.

Gram-negative bacteria are unique in that they contain lipopolysaccharide (LPS), a polymeric structure found between the cell wall and the cellular membrane (found in all bacteria). LPS lies internal to the cell wall and helps regulate the permeability of the cell among other functions. The first stain is crystal violet which stains all bacteria blue-purple followed by an iodine fixative. The critical step is then the decolorizer

A. Colony Morphology:
  1. Colony Shape: Circular, Filamentous (fibrous), Rhizoid (thick fibers), Irregular
  2. Colony Height: Raised, Flat, Convex (rounded), Umbonate (peaked), Crateriform (indented)
  3. Colony Margin (Edge): Entire (smooth), Undulating (wavy), Lobular (finger-like projections), Filiform (fibrous projections), Curled (swirled)
  4. Surface refraction: Smooth, Dull, Glistening (mucoid), Rough (ground glass), Rugose (wrinkled)
  5. Opacity and color: Transparent, Translucent, Opaque, Iridescent, Buff, Red, Blue or other color pigmented (very dependent on culture media), hemolytic (clear regions surrounding colony on plates containing blood cells).
B. Cellular Morphology
  1. Shape: Bacilli, Cocci, Spiral
  2. Gram stain: Gram negative, Gram positive
  3. Inclusions and location: Endospores (central, lateral and terminal)
  4. Surface features: Capsule, Flagella

**Fig. 2.1** Morphological considerations of bacteria

which is methanol. Methanol fixes the cells causing the pores in the walls of gram-positive cells to collapse and become sealed thus retaining the blue dye. The methanol also dissolves the lipid portion of the cell wall of gram-negative cells causing them to become more porous and allowing the dye to leak out. At this point, gram-negative cells can be distinguished from gram-positive cells because they are clear. A red counter stain of fuchsin or safranin is used, giving the gram-negative cells a pink appearance to make their shape and size easier to see microscopically. The patterning of staining can be important in categorizing bacteria to the species level. For example, the staining pattern of the biothreat agent, *Yersinia pestis* is described as "safety pin" since the gram-negative rod-shaped bacteria stains pink with the exception of blue staining at the ends of the rods indicating the specific exclusion of LPS from apical ends.

Viruses, in contrast, have many more shapes by which they are categorized but can only be viewed using electron microscopy. Categorization by shape is still useful in this molecular age and is still the gold standard by which new viruses are identified. The newly emergent coronavirus, the causative agent of severe acute respiratory syndrome (SARS), was identified first by electron microscopic examination of patient samples followed by the use of molecular tools for genetic and immunological identification [5]. Therefore, a toolbox for forensic analysis should include a combination of methods based on both

classical methods and new molecular tools. High confidence can be assumed when these tests agree especially since the databases and signature libraries in general were built from and validated by organisms identified by classical methods. Thus, there is an internal dependence of the new tools on the results predicted by the older classical methods.

A great deal of forensic information can be obtained from analysis of the composition of microbes beyond genetic information. Since microbes and microbial communities are dependent upon other community members in biofilms or microbial mats or on their hosts, significant amounts of information about microbial composition may not be directly determined from genome analysis. Growth conditions and growth media can play a significant role in the process of attribution. Some key areas are highlighted below.

## 2.2.2 Culture Conditions as a Consideration in Forensic Microbiology

Unlike mammals and higher eukaryotes, growth medium may significantly affect the molecular and structural composition of bacteria as well as the isotopic profile of its components. For example, the placement of freshly isolated samples of bacteria on laboratory media alters the fatty acid composition of the bacteria [6–8]. As discussed in Chap. 3, fatty acid composition can be an

important tool for microbial identification and strain distinction. It is unclear from these studies whether laboratory media are in themselves selective for new variants in the isolated sample with altered fatty acid profiles or whether the isolate itself changes to meet its changing nutrient availability [6].

Important to address is the issue of clonality in cultured materials. When isolating bacteria on plated media, most procedures are initiated from a single isolated colony on a plate. The incorrect assumption is that each colony arises from a single bacterial cell; however, in reality it can only be stated that a colony arose from *at least one bacterial cell*. Many organisms do not grow as isolated individual strains. The bacilli- and cocci-shaped microbes discussed above do not always grow as individual cells but may grow as chains or clusters, respectively. The average length of the chains or numbers in each cluster is a growth characteristic which is useful in distinguishing bacteria to the genus or species level. Both chain length and cluster size are influenced by culture media effects [9, 10]. Microscopic examination of cultures of *Staphylococcus aureus* show "grape-like" clusters of cocci-shaped bacteria, the smallest of which is generally made of diplococci (two cells) and the largest is 12 or more cells. Without robust efforts to disassociate these clusters prior to plating, it is unlikely that a single colony of *Staphylococcus* has descended from a single cell. Hence, enumeration of bacteria by colony count (serial dilution followed by plating on growth media) is really only an estimate and frequently an underestimate of the total number of viable bacteria present. Depending on the viability of the sample, it may or may not reflect the number of target bacterial genomes present since polymerase chain reaction (PCR)-based technologies (discussed below) do not differentiate between viable and dead cells. Better agreement is expected using direct counting via a Petroff Hauser chamber. The Petroff Hauser chamber is a low-tech solution in that it merely provides a calibrated area in which individual objects are counted microscopically. It does not discriminate between live or dead. Use of the Petroff Hauser chamber and colony counts can provide a good estimate of the health, i.e., relative viability of the sample.

Even simple direct microscopic examination of raw samples combined with microbial culture can provide the investigator with a better estimate of the initial concentration of microbial targets being analyzed in the original sample. Such examination can provide estimates of sampling variability based on the distribution of material in the sample. Unlike chemicals, bacteria and viruses are discrete units which sometimes adhere to each other or other materials present within a sample creating an uneven distribution and distortion of estimates of microbial targets when using indirect methods to evaluate concentration such as culture. It is important to note that culture, long considered the gold standard of microbiology, is being replaced by direct methods of microbial counting and detection.

With the development of culture-independent methods for microbial evaluation of samples, it is possible to evaluate unculturable organisms in a sample. The ability to identify unculturable microbes further enhances the forensic analysis of a sample. Unculturable organisms may not be the target of concern for biocrimes since one of the underlying requirements for nearly all types of infectious agents is that they are easily cultured; however, contaminating unculturable microbes can provide information regarding comparative temporal and geolocation typing of microbial communities. At the species level and, in particular, at the strain level, the overall makeup of microbial communities can be extremely specific. Some microbes will persist in environments with few alterations for years, while other organisms will persist seasonally or be otherwise chronologically limited. In this case, both the dynamics and stability of the population contain useful information for comparative purposes. It is important to stress comparison when considering characterization of unculturable microbes since the application space afforded by characterizing microbial communities and its variations can be infinite. Culture-independent tools include the use of PCR, microarrays, and mass spectrometry (described below) to identify the presence of bacterial species.

### 2.2.3   Viable but Not Culturable as a Microbial Life Stage

Not to give the impression that all unculturable microbes are excluded as targets, culture conditions can produce growth states known as viable but nonculturable (VBNC) in microbes which normally can be cultured on standard growth media (as reviewed by [11]). In some cases, the microbes can be revived to their original culture state only by passage through a live host [12]. First described for *Vibrio cholerae*, VBNC states have been observed for *Escherichia coli*, *Vibrio cholerae, Mycobacterium tuberculosis, Campylobacter jejuni, Helicobacter pylori,* and bioterrorism agents such as *Fransicella tularensis* and *Burkeholderia* species (as reviewed in [13]). VBNC complicates environmental sample analysis when looking for source attribution in waste water, soil, or other environmental samples since standard culture methods do not apply. While culture-independent methods can detect the presence of a pathogen, its virulence is dependent on its viability which must be determined via culture in one of the pathogen's host species.

### 2.2.4   Impact of Culture on Microbial Identification and Attribution Elements

Culture conditions may affect the surface properties of the bacteria itself. Changes in fatty acid composition can change the adhesion properties of the cell surface [8]. Bacteria designed or selected for greater adherence exhibit greater virulence. Therefore, the fatty acid composition of the native sample may be of importance for attribution or intent. If the perpetrator of a biocrime selects a medium known to alter molecular composition in order to increase microbial virulence, this information is resident in the recovered samples with or without the recovery of the growth medium. It can be discerned by comparison of the recovered sample with isolates grown on more standard laboratory media. Of course, such an approach and determination is limited to samples which are recovered in bulk (i.e., stabilized powders or concentrates).

Concurrent with these tests are a number of media-specific growth tests in which observation of the appearance or presentation of bacterial growth is useful. These media are used to exclude the growth of nontarget organisms, accentuate visual features of the colony, or for gross biochemical assays which describe the enzymatic profile of the bacteria. For example, one classic test for *Bacillus anthracis* is the growth of the organism on sheep blood agar (SBA) plates. Lysis of the red blood cells in the agar is known as beta hemolysis and is carried out through the activation of the hemolysin enzyme. Hemolysis appears as a clear ring surrounding the colony on a dark red plate. *Bacillus anthracis,* unlike its near neighbor *Bacillus cereus,* is nonhemolytic on standard SBA plates. Notably, if red blood cells are washed prior to addition to the plate media, *Bacillus anthracis* colonies exhibit a clear zone around the colonies indicating hemolysin enzyme activity. Hemolysin genes are known to be promiscuous among bacterial species [14], so detection without functional assessment is not sufficient for microbial characterization. Point mutation analysis of the three-part hemolysin enzyme complex can be applied using genetic techniques, but in this case, a simpler and less costly biochemical screen for exclusion may suffice.

Early methods in bacterial identification made use of culture biology for identification. Carbon source utilization, synthesis of lipids, and oxygen requirements were standard methods for screening bacteria. So successful and well validated were these approaches in classical microbial analysis that several commercial enterprises simplified these tests reducing them from racks of test tubes to single 96-well-plate or credit-card-sized formats (Table 2.2). Two general biochemical types of assays are highlighted. The first compares the growth of organisms on a variety of simple substrates. Low-oxygen or oxygen-free conditions are replicated by the addition of mineral oil layers over the growth solutions. The combination of individual tests [15–45] is fed into any automated algorithm which compares the test outputs to a database of outputs for known microbes. The resulting test pattern is matched based on a best fit for microbial identification. These systems are generally used for speciation

of bacteria but are typically unable to distinguish bacterial strains within the same species [46–49]. Since its development, 16S RNA analysis (see discussion below) has replaced many of these tests due to its accuracy and speed [15, 16]. An alternative biochemical assay uses Fatty Acid Methyl Ester (FAME) analysis to profile bacterial lipids. MIDI is an automated system which uses gas chromatography to determine the pattern of lipids which can be derived in a bacterial cell. Lipids found in the lipid bilayer of the cell membrane and attached to proteins and sugars are important regulators of specific bacterial cell functions and can be used to discriminate bacteria to the strain level. The advantage of FAME analysis is that it can be used for viruses, whereas the substrate utilization assays cannot [17–19]. More detailed information on fatty acid techniques is available in Chap. 3.

## 2.3 Terminology Important to Distinguishing Microbes

Before getting too far ahead in the characterization of microbes, it is important to make known or to refresh the definition of some common terminology use in discussion of microbes. Microbes are defined as bacteria, viruses, and fungi. Each category has its unique distinctions and subclassifications; however, for simplicity's sake, genus and species are generally sufficient information to order most bacteria, viruses, and fungi. The taxonomic name, i.e., *Bacillus anthracis* or *Yersinia pestis*, refers to the organism or causative agent of a disease, while the more general description, i.e., anthrax or plague, refers to the disease itself. Therefore, to state that someone is growing anthrax or has cultured plague is incorrect. The correct terminology is that they are culturing *Bacillus anthracis* or *Yersinia pestis*. Note that taxonomic names are always italicized.

### 2.3.1 Microbial Speciation Is a Continuum

The definition of species is much fuzzier for microbes than for higher eukaryotes such as plants and mammals. In higher eukaryotes, speciation is defined by the ability of two individuals to successfully interbreed. Failure to produce offspring or failure to produce offspring capable of reproduction is indicative of two different species. Speciation in bacteria is much more nebulous. Originally, the morphological and biochemical differences between organisms were determined to categorize isolated bacteria into groups. Type strains are the strains which exhibit many typical characteristics of a species. Usually, type strains are among the first representatives to define the species and may not represent all of the most common feature of the species. Shape, size, and culture appearance all go into the morphological distinction of bacterial types. Most bacteria exist as actively growing cells called vegetative cells. Only a few subsets of bacteria are able to produce spores, viable but dormant structures which encapsulate the genome of bacteria. Spores enable bacteria to survive in unfavorable conditions until the correct growth requirements are detected, at which time the spore undergoes a process called germination where the spore breaks down and releases a replicating vegetative cell. Using classical microbiological methods, bacteria are further distinguished based on biochemical differences. In bacteria, sugar utilization and other nutrient requirements are noted. Many of the selective media that can be used to enrich mixed samples for a particular bacteria species are based on the restriction of carbon sources to a single type of sugar or preferred sugar which allows the target species to grow fastest, i.e., Yersinia agar which is selective for *Yersinia spp.* or PLET medium which is selective for *Bacillus anthracis* [21]. Appearance on specific agars (e.g., blood agar plates in which sheep or other mammalian red blood cells are added to a basic agar media such as tryptic soy broth) provides additional information.

Hemolysis is notable and distinctive for bacterial species differentiation. Recently, a rapid 2-h culture test for low-cost screening to rule out *Bacillus anthracis* exposure has been developed [20]. The test is based on the fact that *Bacillus* species exhibit differential beta hemolysis or zonal lysis and clearing of red blood cells surrounding the margins of a colony. Microscopic

examination of cultures for gram-positive rods and spores coupled with the presence of active beta hemolysis is enough information to rule out the presence of *Bacillus anthracis*. Since tests such as these are indicators, it would seem simple to develop rapid DNA dipstick screens for the presence of the hemolysin gene to rule out *Bacillus anthracis*; however, as it turns out, *Bacillus anthracis* has a complete hemolysin gene. It is only when the red blood cells are specially washed prior to the preparation of the agar that beta hemolysis can be noted in *Bacillus anthracis* [22]. Hence, partial genetic characterization such as the presence or absence of a gene is not always definitive in microbial differentiation. In the case of exclusionary decisions regarding a sample, the absence of a gene may be sufficient to eliminate a sample from further analysis. For example, in a case of death due to a respiratory infection, the absence of any of the three toxin genes coding for the *Bacillus anthracis* toxins could rule out a sample as the cause of death even if other phylogenic markers indicate *Bacillus anthracis* is present.

Additional names and numbers may follow each designator of genus and species such as *Bacillus anthracis* Sterne or *Yersinia pestis* Kim 10, and additional sets of information may follow these names, usually numbers. These refer to strain and isolate designations. Strain and isolate are often used interchangeably, but there really is a small distinction. Originally, during isolation of an unknown and identification of a new variant of a species, the isolated microbe was given a unique identifier such as the name of the place found, host from which it was isolated, or person who first isolated it.

Strains have all or most of the qualifying characteristics to be part of the species to which they are assigned but may also have some differences between the newest isolate and other members. These differences are relatively minor and are usually not enough to classify the isolated microbe as its own species; however, since they are enough to be notable, the microbe is given a new strain designation. The term isolate refers to a pure colony grown from a culture or environmental sample. The description or identification of the isolate is linked to a particular sample, place, or even culture method. Isolates from different sources are often the basic unit of comparison in many forensically based studies. It is important to remember though that an isolate is not equivalent to an individual.

The smallest unit for a microbe is a bacterium (bacterial cell), or spore, or in the case of viruses, a virion or in the case of fungi, a cell or a spore. The designation of a cell in the kingdom of fungi can be tricky since it is difficult to define the unit of a cell, hence making an entire colony of a non-septated, multinucleated fungi appear as one continuous unit. Currently, the title for the largest organism on earth goes to the discovery of a specimen of *Armillaria ostoyae*. This single organism covers 2.5 $mi^2$ in southern Oregon [23]. Therefore, a colony which appears as a single unit on a plate (or for viruses a single plaque) may arise from a single bacterium or virion. Once it becomes visible, a population of thousands of bacteria or virions is present. This may seem trivial; however, it becomes important when discussing genetic mutations and the application of human forensic tools to microbial forensics. If the average colony or viral plaque contains $10^5$–$10^9$ individuals and the average spontaneous mutation rate is $10^{-9}$ to $10^{-10}$ per base per generation of bacteria (viruses are somewhat higher, $10^{-4}$ to $10^{-8}$ [24, 25]), each colony or plaque will contain at least 3,000 individuals that are genetically different from the starting organism (Table 2.1).

Humans and other higher eukaryotic organisms that have developed better repair systems have mutations 100–10,000 times lower than bacteria and viruses and achieve fewer generations than the average microbe, so genetic drift due to spontaneous mutations is less of a problem in humans than bacteria. Therefore, studies which match genetic elements must use as few culture steps as possible, must limit the analysis to the most stable or least mutable genes, or must take into account genetic drift as an influencing factor in genome comparison.

**Table 2.1** Comparison of mutation rates among different organisms (Adapted from Drake [24])

| Organism class | Example(s) | Mutation rate[a] |
| --- | --- | --- |
| DNA virus | Bacteriophage (T-type) | $2.4 \times 10^{-8}$ |
| RNA virus (negative strand) | Orthomyxoviruses (influenza) | $8.3 \times 10^{-5}$ |
| RNA virus (positive strand) | Picornaviruses (polio) | $1 \times 10^{-4}$ |
| Bacteria | *Escherichia coli* | $5.4 \times 10^{-10}$ |
| Fungi | *Neurospora crassa* | $7.2 \times 10^{-11}$ |
| Higher eukaryote insect | *Drosphilia melanogaster* | $3.4 \times 10^{-10b}$ |
| Higher eukaryote mammal | *Homo sapiens* | $5.0 \times 10^{-11b}$ |

[a]Spontaneous mutations/base pair/replication event
[b]Average for multiple loci

**Table 2.2** Common Commercial Microbial Identification Systems

| System Name (Company) | Detection Method | Interpretation |
| --- | --- | --- |
| Vitek (bioMérieux) | Colorimetric | Automated growth monitoring of bacterial isolates on variety of carbohydrate sources and biochemicals |
| API strips (bioMérieux) | Colorimetric | Measure growth of isolated bacterial on variety of carbohydrate sources and biochemicals |
| GEN III Microbial ID (Biolog) | Colorimetric | Uses chromophoric redox indicator as measure of respiration and growth on various carbohydrates |
| Rainbow Agar (Biolog) | Selective media and chromogenic growth media | Selective growth media which suppresses non target species and uses chromophoretic compounds as redox indicator |
| MALDI-TOF BioTyper (Bruker Daltonics) | Analysis of proteins based on mass | Identification of bacterial protein signature is compared to library of known bacterial signatures for identification |
| Sherlock Microbial Identification System (MIDI) | Gas chromatography | Fatty acid analysis to discriminate bacterial species |
| MicroSEQ® Rapid Microbial Identification System (ABI) | PCR and nucleic acid sequencing | Uses well established method of speciating bacterial based on 16S ribosomal RNA |

## 2.3.2 Adopted Information as a Consideration for Attribution

The hierarchical process by which key features are ordered is often a subject of debate among microbiologists. New molecular methodologies have thrown cladistics and bacterial phylogeny into some disarray. Partly this is due to the promiscuity of gene transfer in microbial species and highlights another feature of microbial genetics which differs significantly from genetics in higher eukaryotes. In eukaryotic species, the only manner of introducing new genes or alleles is through sexual reproduction. This is referred to as a vertical gene transfer in that the new genetic material is derived only from an ancestor.

Bacteria have no such restrictions in that they acquire genetic material from another bacterium by horizontal gene transfer. By this mechanism, there is no ancestral link between the bacterial species which share these genes, and in fact, horizontal or lateral gene transfer is not limited to transfer within members of a species [26]. Speciation in other organisms is often represented by punctuated genetic differences, i.e., appearance of new genes, while in bacteria speciation, the distinction between species is more of a genetic continuum of dubious linearity. As noted by Carl Woese, "one necessarily questions the universal phylogenetic tree. That compelling tree image resides deep in our representation of biology. But the tree is no more than a graphical device; it is not some a priori form that nature imposes upon the evolutionary process. It is not a matter of whether your data are consistent with a tree, but whether tree topology is a useful

way to represent your data" [27]. In other words, evolutionary trees and categorization of microbes into defined units is a tool and not necessarily the rule for defining bacterial identity. The presence of a new gene in a bacterium of known classification may indicate that the target has been in close proximity with other species and may provide forensic information about the life and/or residence of the microbe prior to its isolation and sampling. For example, the presence of partial genes from a soil bacterium found only in specific geological regions in a fairly ubiquitous pathogen may suggest the location where a bacterium was grown or isolated. Unlike growth media and nutrient components, these features may not be lost when the organism is cultured.

Likewise, recovered viruses are good sources for information about the methods and sources of materials used to produce them. Viruses are not considered living, and therefore, their classification is more basic. Viruses are classified based on their structure (visible at the necessary level of detail by electron microscopy), their genetic makeup (one or two strands of genetic material, RNA or DNA and orientation of RNA strand if single stranded), and the presence of a lipid envelope. So although the genetic sequence can give information on viral identification and similarity, the surface composition of the virus can provide information on the source of the material.

Viruses require living hosts to manufacture their components; therefore, the signature from the living host is present in the molecular components surrounding the virion. For example, glycosylation patterns are indicative of the host species in which a virus was grown. Glycosylation is the method by which small subunit sugars are linked to proteins and lipids in long and sometimes complex polymers. The ability of viruses and bacteria to recognize and invade a host is based on recognition of sugar polymers on target proteins [28–30]. The biochemical steps required for addition of specific sugars to proteins is evolutionarily restricted [31, 32]. Simple glycosylation is carried out in bacteria, more complex enzymology is encoded in eukaryotes, and enzymes making highly specific glycosylation linkages diverge in higher eukaryotes like *Drosophila*, mammals, and primates. Therefore, insects and mice make different branched sugar polymers. The appearance of unique sugar linkages can be used to identify which culture system was used to grow a viral preparation. Probing differences in glycosylation can be as simple as monitoring the differences in surface protein confirmation using epitope mapping via immunoassay, or a more detailed analysis can be conducted using ion trap mass spectrometry [33, 34].

Additional surface components found in microbes are not encoded in the microbial genome. Enveloped viruses which include influenza, herpes, measles, HIV, and rabies acquire the lipid membrane which surrounds and protects the virion capsule during the process of viral budding from host cells. Lipids and other components of the host cell membrane hitchhike along with the budding virus. Rough handling, mild warming, and other environmental factors can quickly disassociate the envelope, so care must be used in collection of forensic materials if the membrane components of the virion are to be analyzed. Unlike genetic elements, the envelope is lost during reculture of collected viruses, so minute quantities are not amendable to many analysis methods other than mass spectrometry.

Direct testing of viral samples for mitochondrial DNA can also be informative. Contaminating mitochondria and mitochondrial DNA can be amplified in viral preparations, providing, at the very least, species information regarding the host cell or even tying the production to a specific cell line [35]. This methodology was used to analyze archival polio vaccine in order to confirm the presence of macaque mitochondrial sequences and failed to detect chimpanzee mitochondrial sequences. While glycosylation linkages, lipid profiles, and mitochondrial sequences are not intrinsic to the virion, they are frequently associated with released virus and may warrant further forensic analysis. Hence, sample collection methods must be designed to take advantage of these potential markers of exclusion and inclusion. These methods and the associated quality systems are described in more detail in Chap. 9 (Wilson, this volume).

## 2.4 Nucleic Analysis for Microbial Characterization: The New Gold Standard

### 2.4.1 Polymerase Chain Reaction (PCR): An Enabling Technology for Microbial Genomics

With the revolution in molecular biology, the characterization and typing of microbes has become a matter of some debate. The source of much discussion among systematic biologists is the ordering of bacteria into specific classes. Much of this reassignment is due to the use of genetic methods to identify bacteria based on their DNA profiles rather than their morphology. Like the science of forensics, which is a continuum of methods from exclusion to attribution [2], genetic profiling of microbes is a continuum from which it is sometimes difficult to make clear distinctions. Instead of cutoffs or limits, it is better to compare based on a degree of "sameness" when comparing two strains of bacteria.

PCR is a process by which small amounts of nucleic acid is synthesized in vitro to make large amounts of an exact or nearly exact copy of nucleic acid. As a result, PCR has enabled countless new applications in human forensics, medicine, and agriculture by providing enough material for more robust analyses without the need for a living organism. PCR has been the subject of many excellent recent forensic reviews [1, 36, 48] and so will be discussed here relative to microbial applications. Briefly, PCR works by amplifying DNA or, in the case of reverse transcriptase PCR (RT-PCR), amplifying RNA. This is possible due to the discovery of thermostable polymerases which retain their ability to extend short oligomers of nucleotides known as primers [37]. As for all DNA polymerases, a template is required in order for the polymerase to extend the new complementary DNA strand. Therefore, extension of the primer only occurs when the primer is bound to its DNA target. The number of DNA targets is limiting in the reaction initially. So to increase the number of DNA templates after the first extension is complete, the two DNA strands are dissociated

by raising the temperature to the point where the DNA "melts." When the temperature is lowered, the next primer which is in high concentration in the reaction binds to the DNA template, and the process begins again. Using one primer, the process is rather slow and the increase in the DNA is linear. To make the process of amplification logarithmic, a second primer is added which binds downstream of the first primer on the complementary DNA template and any amplified DNAs which have the DNA template sequence. Therefore, after the first few cycles of primer binding (annealing), primer extension by the polymerase, and DNA melting, amplification of the original DNA template is rare in comparison with amplifications of primers binding and reamplifying newly synthesized DNA. The product of this reaction is a short segment of tens to thousands of base pairs of double-stranded DNA and, in acknowledgement of its synthetic origin, is known as an amplicon.

In their most primitive form, amplicons are visualized by gel electrophoresis and staining to determine if the reaction succeeded in producing product and if the product's size is consistent with predicted results. A whole variety of tools have been developed to make the detection of the amplified products more efficient such as (1) the addition of fluorophores to the end of the primer for direct fluorescent detection of amplicons without staining, (2) the use of internal probes whose exact binding to sequences within the amplicon produces a light-emitting reaction known as Fluorescent Resonance Energy Transfer (FRET) (FRET emission can only occur if the two internal primers are within one base pair of each other and if the amplicon sequence matches perfectly, thus providing sequence information about the amplified product as well as its size), and (3) the addition of intercalating DNA dyes into the reaction. The intercalating dye fluorescence is enhanced only when the dye is able to bind double-stranded DNA. Generally, the template DNA is present in too small a quantity to contribute to the reaction.

Amplicons can then be used in other assays, sequenced directly, or inserted into other genes or

organisms. By carefully designing the complementary primer sets use for amplification, it is possible to amplify any sequence of DNA uniquely. Because of this, PCR-based methods have caused concern in the law enforcement and scientific community, as it has been demonstrated that whole microorganisms can be regenerated by this means [38–40]. Nevertheless, PCR has provided investigators with a molecular tool needed to examine small amounts of materials for low-abundance microbial targets.

PCR has been used to detect pathogens directly by a number of methods. Most common is to target a gene which is important to the function of a pathogen. Panels of primer sets have been published for nearly every pathogen including bioterrorism agents. Most of the gene targets for these panels use virulence factors as a condition of pathogenicity. For example, initial screening of forensic samples to find the source of a suspected *Bacillus anthracis* attack would likely include amplification and detection of genes for protective antigen (pag), lethal factor (lef) or edema factor (ef), and capsule (cap). All of these genes must be present for virulence [41]. Rapid testing of these targets and others such as BA 813, a chromosomal marker specific for the *Bacillus cereus* subspecies, can be accomplished on suspected spore preparations, i.e., white powders, without extensive sample cleanup. As shown in Figs. 2.2 and 2.3, amplification of BA 813 is evident when whole spores or vegetative cells are added directly to the PCR reaction without extraction or extensive methods to destroy the cell structure. As shown in Fig. 2.2, autoclaved spores which have been thermally disrupted and intact spores added to a PCR reaction amplify similarly. This method has been used for the toxin genes named above with similar results and sensitivities (Jackman, personal communication). This method works well because DNA is found on the external surface of spores. The exosporal DNA is intrinsically associated with the spore surface. Copious washing does not remove the DNA even when other DNAs added to the spore preparation are easily removed, suggesting that this DNA is tightly retained by the endospore during the sporulation by an as yet unknown mechanism. Our estimates
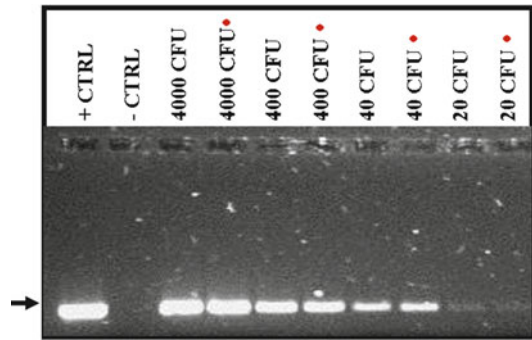


**Fig. 2.2** Rapid exosporal detection of *B. anthracis* sequences: PCR products amplified using a dilution series of *B. anthracis* Sterne strain spores that were plated to determine their concentration (colony-forming units). The samples were either autoclaved which damages the spore coat and exosporia (*dot*) or heated at 95°C which does not damage these structures; then, PCR was performed using the Ba813 chromosomal DNA-targeted assay. This assay produces a 152-bp product marked by the *arrow*
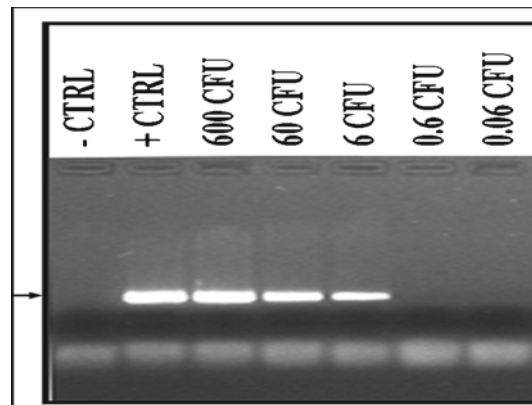


**Fig. 2.3** Rapid detection of chromosomal markers in whole vegetative bacteria: PCR products formed using a dilution series of *B. anthracis* Sterne strain vegetative cells that were plated to determine their concentration. These samples were either autoclaved (+) or heated to 95°C, and then PCR was performed with the Ba813 chromosomal DNA-targeted assay. This assay produces a 152-bp product marked by the *arrow*

based on enumeration of spore preparations and quantification of DNA target copy number indicate that about 1–10% of total genomic DNA is retained on the spore surface.

While the sensitivity of the test based on gene copy number is lower overall by about tenfold as compared to clean extracted DNA, the assumption

for rapid screening for hazard assessment and sample collection involving white powders is that the number of spores is not limiting. Pure dry-powdered *Bacillus anthracis* spore preparations contain between $10^{12}$ and $10^{14}$ spores per gram depending on the sample preparation methods and amount of contaminants remaining. When adding vegetative bacteria directly to the PCR reaction, amplification is likely enhanced by the release of internal DNA by the process of heating and cooling in a typical PCR cycle. Heat shock is used to kill off by lysis vegetative cells in microbial preparations where only the number of spores (typically heat resistant). This property of vegetative cells can be easily exploited by the PCR process to liberate intracellular DNA targets for amplification. Appropriate controls should be used always to verify that contaminants present in the sample do not inhibit the PCR reaction.

### 2.4.2 Ribosomal RNA for Microbial Identification of Species and Strain

One major event which contributed significantly to the identification and discrimination of bacteria is the use of ribosomal RNA sequencing. As mentioned above, lateral gene transfer, whether by natural methods or by the use of genetic engineering, has made the assignment of individual genes (both plasmid borne and chromosomal) to specific organisms less rigid. Since lateral gene transfer involves the transfer of whole genes, duplicated nonfunctional genes, and partial gene segments, the presence of a particular sequence may or may not be significant. For example, plasmid-borne Cry genes which encode the insect toxins are not only found in their natural host *Bacillus thuringiensis* but are found as sometimes functional hitchhikers in the closely related *Bacillus cereus* [42].

The use of one gene to type bacteria has not always held up as a mechanism of scrutiny in the larger microbial community. The most useful genes, as it turns out, may not be the genes which are uniquely associated with species or genus but the more common essential genes found in all organisms. The best example of this appears to be ribosomal RNA (rRNA). Woese first recognized the uniqueness of rRNA as a tool for the classification of all life [43]. rRNA serves a structural and catalytic role in the production of proteins. The ribosome encodes the machinery for the translation of messenger RNA transcripts into their respective protein products. The presence of rRNA in cells is required for viability, and thus, the presence or absence of rRNA is sometimes used as a discriminator of live versus dead cells, respectively. There are three rRNA subunits located in a single gene transcriptional unit. In bacteria, the subunits identified originally by their sedimentation rate expressed in Svedberg unit(s) are 16S, 23S, and 5S. In eukaryotes, the subunits are 18S, 28S, 5.8S, and 5S.

The advantages of rRNA for genotyping are the following: (1) At the DNA level, the gene is represented in many copies per bacterial and fungal genome. (2) At the RNA level, there are hundreds to tens of thousands of copies per organism, making it the most highly amplified nucleic acid in any bacterial cell. (3) The gene coding for rRNA is found on the bacterial chromosome and cannot be deleted due to its overall effect on all translation in a cell. (4) It has been shown to mutate or evolve slowly, demonstrating its stability in terms of millions of years. Woese first suggested that rRNA serves as an evolutionary clock by which all life can be mapped. Over 10,000 publications have described the use and consistency of rRNA as a method for phylogenic mapping of organisms, substantiating the success of this approach. While the approach was solid, the methodology for analysis was not straightforward when rRNA phylogenetic profiling was in its infancy.

Adaptation of techniques related to analysis of restriction fragment length polymorphisms (RFLP) quickened the analysis for rRNA by combining a general technique with a specific gene. RFLP analysis is a rather gross tool for comparing whole genomes. It is dependent on the idea that at some DNA restriction enzyme cleavage sites, single nucleotide polymorphisms (SNPs) will occur, thereby preventing the cleavage of the DNA and creating a different pattern when the PCR products are analyzed. Ribotyping uses a

similar approach in that amplification products are generated from the genomic copies encoding the regions of the rRNA genes within greatest variability in the intergenic spacer regions.

As well conserved as the coding regions of the rRNA genes are, the intergenic regions are highly variable and show significant differences between strains in sequence and length. The amplified fragments are separated by gel electrophoresis to visualize the differences in intergenic regions to discriminate bacteria at the species and strain level. Ribotyping can be a better discrimination tool at the strain level than PCR-based methods to amplify specific pathogen-associated virulence factors [44].

Fortunately, the biotechnology community has worked simultaneously to develop high-throughput techniques along with data-reduction protocols to overcome the technical challenges needed to apply bench level application rRNA analysis to real samples. Two technologies, microarray technology and mass spectrometry, have dramatically changed the manner of microbial analysis using rRNA. Microarrays are a logical presentation of biological molecules laid out on a solid planar platform, generally a glass slide [45]. The resulting presentation is referred to as an array or, as it is sometimes called, a "chip." In this format, individual molecules are separated from one another, allowing the user to interrogate each molecule independently but simultaneously. As a result, it is possible to obtain relational information about hundreds and thousands of different events simultaneously. No other technique offers this amount of data in a single test. In this way, microarrays have enabled similar pattern matching and genetic fingerprinting methods for typing microbes in similar fashion to the classical fingerprint methods.

The types of molecules which may be presented or "arrayed" include DNA oligonucleotides, cDNAs, peptides, or sugar molecules. A full history of microarrays can be found in reviews by Venkatasubbarao [50] and Southern [51]. When used to define rRNA phylogeny, DNA probes called capture probes are bound to the chip and are interrogated with labeled but unknown DNA or RNA sequences derived from a microbe. The sequences selected as capture probes were derived from direct sequencing ribosomal genes of cultured microbes. There are several open databases which have been created to provide this information [52–54] and several microarray designers which have based their products on these databases. Probes are designed to match areas of the 16S gene which provide discrimination between bacteria usually on the bases of single nucleotide polymorphisms (SNPs).

The GeneChip (Affymetrix Corporation) contains over 31,000 oligonucleotides designed complementary to 16S sequences from a diverse number of bacteria [55, 56]. Following PCR using labeled amplification primers directed toward the ribosomal genes, the amplicons are allowed to hybridize to the microarray and the resulting complex pattern of target sequence. The label can be a fluorophore, or isotope, to which imaging mass spectrometry is sensitive. The pattern of hybridization is used to determine the best match to the unknown bacteria. The MAGIChip (Argonne National Laboratory) works in a similar fashion except that PCR is not required. Nucleic acid derived from unknown bacteria is directly labeled and hybridized to the microarray, eliminating potential artifacts resulting from post-isolation amplification [57].

What are the advantages of this approach? First, since rRNA is common to all living organisms, one chip can be used to discriminate all organisms rather than developing unique probes to each. Second, complementary information between the probes is used to derive the correct identification. In other words, the probes are hierarchically ordered in a manner consistent with existing cladistic ordering, giving increased confidence to the microarray assay since all probes must be internally consistent. An example of this is given in Fig. 2.4 which shows the interdependence of the capture probes sequence used on the MAGIChip [58].

Using this strategy, the detection of probe sets for *Bacillus cereus*, *Bacillus anthracis*, and other members of the *Bacillus* group can only occur if probe sets which indicate the organism is a bacteria (*Eubacteria*), is gram-positive, contains high
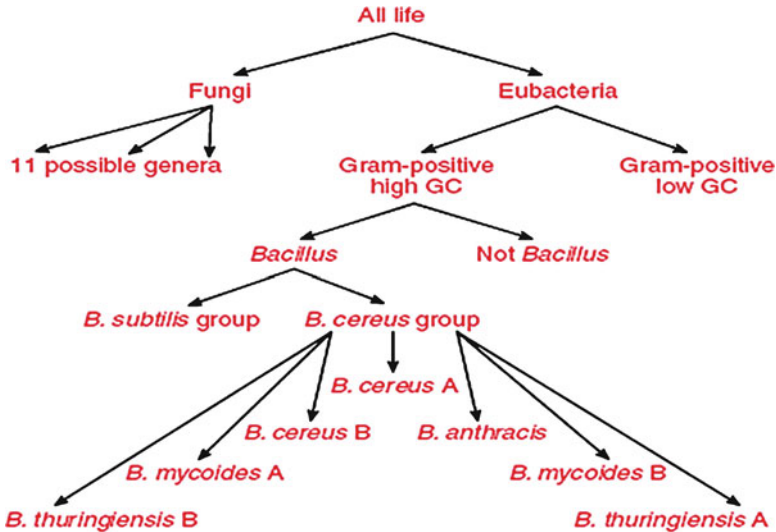
**Fig. 2.4** Phylogenetic tree: a schematic of hierarchical recognition on MAGIChip. Hierarchical-based identification strategy provides an internal consistency check.

Probes are phylogenetically ordered by domain, kingdom, phylum, class, order, family, genus, and species (GC = guanine/cytosine base pair)

guanine/cytosine (GC) content, and is in the *Bacillus* group are also hybridized. This internal consistency check prevents the false identification of unknown bacteria which does not meet all these simultaneous tests. For example, any unknown bacteria which might hybridize to the final probe set for *Bacillus anthracis* would never be detected as a *Bacillus anthracis* unless it also hybridized the probe sets for gram-positive *Bacillus* and failed to hybridize probe sets for gram-negative bacteria and *Proteobacteria*. In over 1,800 tests, the MAGIChip and its associated automated algorithm for interpretation of the hybridized pattern has shown no false alarms for the identification of *Bacillus anthracis* [58].

In addition, a modified MAGIChip format has been shown to identify microbes in environmental samples which cannot be cultured [59]. To do this, the microarray is placed on the microarray reader stage and gently heated. The loss of fluorescent signal is recorded over time for each of several hundred probes. As a result, hundreds of melting curves are generated, and the melting point (Tm) or point at which 50% of the fluorescence is lost indicates the exact sequence which is bound to each capture probe. Melting curves which show multiple humps indicate the presence

of multiple Tm and hence multiple sequences bound to the same capture probe. By combining microarray hybridization with melting curve analysis of the bound targets, it is possible to discriminate mixtures of bacteria [60]. One disadvantage presently is the inability to discriminate between two bacteria from the same species but from different strains. In general, rRNA discriminates to the species level, although the addition of other genetic loci and their associated SNPs is easily accomplished using microarray technology. Thus, the microarrays could, with modification, be used to discriminate differences between strains.

The success of this multiplex approach is not only evident in both these microarray formats but also in a related technology which uses mass spectrometry. In the T5000 system (IBIS Inc), amplicons derived from samples using multiple primer sets are analyzed for both their presence and their mass which can be used to derive the sequence composition [61]. Unlike PCR approaches described above targeting specific gene sequences, the probes are designed not to be unique but to amplify sequences found in all microbes; hence, primers amplify not only the rRNA genes but also housekeeping genes such as

tRNA synthetase, topoisomerase, polymerases, and other genes common among organisms. Following amplification, the products are fed into an electrospray ionization mass spectrometer, and the molecular masses of each strand of each amplified product are determined within 1 Da of its mass to obtain a precise base composition. The resulting profile of all the gene products and accurate masses are matched to a database for triangulation of the most likely organism. Since the smallest difference in molecular weight possible for a single base substitution as occurs within a SNP is greater than a Dalton, the T5000 system can detect single base differences and can be used to distinguish strains. This system also uses an automated hierarchical strategy to order the information derived from the various amplicons and their masses to identify microbes with high confidence [62]. One of the great advantages of this system is that it can successfully identify mixtures of pathogens.

### 2.4.3 Repetitive Element Pattern Analysis for Bacterial Identification

There are three main thrusts of microbial classification and distinction based on pattern analysis of genetic elements other than rRNA. These are Restriction Fragment Length Polymorphisms (RFLP)/Amplified Fragment Length Polymorphisms (AFLP), Variable Nucleotide Tandem Repeat (VNTR), and Multiloci VNTR Analysis (MVLA). An additional application of SNP as described for rRNA genes above has been applied to other genes to create patterns of mutations unique by strain and isolate, although this approach is in its infancy relative to the former approaches. Pattern analysis of whole genomes started and still uses RFLP and AFLP analyses as screening tools for evaluation of mixed strains in a population [63–65]. RFLP as described above uses random pattern analysis created by restriction enzyme digestion of the genome and separation of fragments. Fragments must be visualized by probing with complementary labeled oligonucleotide sequences. AFLP uses the distribution

of repeated elements in the genome to create specific primers for amplification.

Following enzyme digestion, the ends of the digested DNA are annealed with double-stranded sequences which fill in the restriction site and provide the sequence specificity for primer amplification. The pattern is created from alternatively amplified DNA fragments which are separated and visualized on gels or by capillary electrophoresis. Both pattern analysis tools are discriminatory of bacteria at the strain level due to the same molecular mechanisms of gene insertion/deletion (INDELS) or SNPs to create or destroy restriction sites at random positions in the DNA. One of the drawbacks of these approaches is that the patterns are a global representation of the genomic patterns. As a result, it is difficult to predict the stability of the observed pattern or to compare across genomes due to the large differences in frequency of restriction sites for any one restriction enzyme in all microbes and the relatively low characterization of mutation rates in restriction enzyme sites for many organisms.

In contrast, VTNRs are the holy grail of forensic DNA analysis due to the high degree of information available about these specific short sequences. An entire book could be devoted only to the discussion of applications and variations of VTNRs for microbial discrimination at the isolate level, and instead of repeating this information, the reader is directed to reviews by Fan, Budowle, VanBelkum, Linstead, or Grissa [1, 66–69]. The best use and most advanced development of VNTRs for forensic and legal use is based on markers for human DNA. In humans, the most useful VNTRs are repeats of 2–7 base pairs found in microsatellite DNA known as short tandem repeats (STRs) [66]. Despite the large amount of research on STRs, the functions of STRs are as yet unknown; however, the original supposition that STRs are only "junk DNA" is being gradually replaced with functional assignments for specific STRs. There are over 10,000 STRs in the human genome, but only 13 human STRs have been down-selected for use in forensics due to their high degree of variation and large amount of available data. STRs are heritable and can be used to determine relatedness. Using a

combination of sequencing probes, automated sequencing, and well-validated algorithms, information derived from human sequences has been placed in a global DNA database known as the Combined DNA Index System (CODIS). Due to the number of variants for each of the 13 loci, the chance that two people share all 13 loci variants is $5 \times 10^{19}$, therefore making this matching tool one of the most robust personal identification systems in the world.

Fortunately, bacteria and fungi have a similar system of repeat loci so that the genetic tools for comparison can be readily adapted to the process. Multiloci VNTR analysis (MLVA) is the pattern analysis of multiple STR loci found in bacteria. Consider though the following: Humans which represent a single species have over 10,000 STRs alone. Though their genomes are smaller, by comparison to humans, the number of microbial species is staggering, and therefore, the potential number of multi-VTNR for typing all species of bacteria is even larger. As a result, not all bacteria can be readily typed by this method. MLVA analysis is heavily dependent upon a rich database of sequences for a particular bacterial species or subspecies [70]. Additionally, bacterial repeats do not have the unit fidelity seen in other repeat analyses. Thus, analysis of repeat units is complicated since bacterial repeats are not always an integral number [71].

Nevertheless, MLVA analysis has been demonstrated as a robust method to discriminate bacteria of high importance on the basis of species, strain, and even isolate differences [69]. A tutorial of the requirements to develop and carry out MLVA can be found on the web [72]. Briefly, to carry out a successful MLVA analysis, known PCR primers are developed whose product sizes can be predicted based on libraries of sequenced target organisms [72]. Large sets of primers (25 are used for *Bacillus anthracis* and *Yersinia pestis* typing singly and 19 for *Mycobacterium tuberculosis*) are used to amplify repeated sequences [73–75]. PCR products are isolated by gel electrophoresis or capillary electrophoresis, and the observed patterns are compared using commercial software or freeware. To perform this MVLA for diverse species of bacteria requires a great number of PCR primers at this time.

However, great efforts are being made to simplify the approach using a minimal number of primers.

Keim first showed that *Bacillus anthracis* subtyping by the MLVA approach was possible using as few as eight primers sets, but recent studies indicated that these sets may not be sufficient to distinguish subtypes during epidemic outbreaks [76, 77]. As scientists understand more about the distribution of bacterial STRs and more bacterial strains are characterized by this method, it may be possible using a reductionist approach to reduce the number of STR loci needed per species or genus to fewer panels or a smaller number of STRs per panel.

### 2.4.4 Whole Genome Analysis: Canonical Sequences and SNP Analysis

While DNA sequencing is not the newest technology to influence the field of forensics, new developments in sequencing technology have made it possible to sequence and assemble whole bacterial genomes in weeks rather than years. Application of the data from these efforts in the area of comparative genomes has led to unique discoveries regarding the speciation and stability of genomes. For forensic application, whole genome analysis and comparative genomics will be the Rosetta stone for microbial identification for isolate level discrimination as well as for specific attribution. However, as with any new technical advancement, there are challenges to be overcome. Many of these concerns and caveats have been reviewed by Budowle [78].

Of greatest interest to the microbial forensic community is the mapping and distinction of SNPs. Based from the human SNP efforts, for a mutation to be designated as a SNP, the polymorphism (rather than a random or common mutation) must occur in greater than 1% of the population (SNP consortium, [79]). Thus, translation of this description to microbial populations is somewhat more difficult. Is the population of bacteria defined as a mixture of strains, colonies on a plate, or individuals in a culture tube? Classifying true polymorphisms from mutations requires the simultaneous development of statistical tools

to track the stability of polymorphisms in a collection of strains [80]. One of the best examples to illustrate this development occurs in a series of papers published over 5 years, describing the search for SNPs in the *Bacillus anthracis* genome [81]. Prior to these studies, *Bacillus anthracis* was considered to have a relatively monogenetic type. Its low rate of spontaneous mutation supported that conclusion. However, once the genome was assembled, it was clear that *Bacillus anthracis* could be differentiated at the strain and isolate level based on the analysis of approximately 60 SNPs. Further work with complete and redundant sequences led to the discovery of over 3,500 SNPs in the plasmid and chromosomal *Bacillus anthracis* genome [82]. In an effort to make microbial typing easier using comparative genomics based on SNPs, sequencing primers specific to 32 of the chromosomal SNPs which showed the best discrimination among isolates were down-selected and used in subsequent restricted sequencing approaches to highlight these polymorphisms by strain [83].

Still it is important to note that SNP analysis of whole genomes is of greatest utility when the overall mutation rate is low. Both laboratory and environmental effects have been documented which induce the evolution of the mutator phenotype in some bacteria in a population [84]. Culture stress, exposure to antibiotics in the host, and coexistence with other viral pathogens have all been demonstrated to induce the mutator phenotype [85, 86]. Mutator bacteria have significantly higher rates of spontaneous mutation than other members of their community. Deletions or alterations in the genes associated with polymerase function have been demonstrated to induce mutator types in *Bacillus anthracis* [87]. The phenotype is induced presumably by enhanced adaptation under environmentally stressful conditions and by providing the most diverse genetic background for survival selection by speeding adaptation under environmentally stressful conditions and providing a more diverse genetic background. All DNA-containing microbes including viruses have the same rates of spontaneous mutation [24]. As a precautionary step or to aid the interpretation of SNP data sets, general screening methods such as monitoring for rifampicin resistance might be beneficial prior to whole genome analysis especially when extensive culture on highly restrictive media, suboptimal storage, or environmental stressors may act as the foundation for induction of this phenotype [86]. This is particularly true for host-derived materials relative to long-established laboratory adapted strains. Observed elevations in mutations in culture material should be taken as a warning sign that cultures contain mutator phenotypes and relative SNP between strains should be surveyed for artifacts. A caveat in the use of SNPs for forensic analysis is that the mutation rate should be relatively low or the stability of an individual SNP be known in order for it to be useful for establishing strain or isolate identities.

## 2.5 Proteome and Intent as Determined by Mass Spectrometry and Other Multitarget Approaches

It is post genomic ERA, notably based on the completion of the human DNA sequence, the interest has shifted slightly to the development and application of tools for protein sequencing (discussed in Chap. 5), carbohydrate profiling (discussed in Chap. 4), fatty acid analysis (discussed in Chap. 3), and identification of contaminating substances (discussed in Chap. 8). While a variety of methods are possible, mass spectrometry is a technology which is central to all these applications. The most commonly used systems are Matrix Assisted Laser Desorption Ionization Time of Flight mass spectrometry (MALDI-TOF), Electrospray Ionization Time of Flight mass spectrometry (ESI-TOF), and Gas Chromatography mass spectrometry (GC-MS) [88–91]. Two of these technologies, MALDI-TOF and ESI-TOF, are commonly used to detect and identify proteins using sequence-related information, which can sometimes be crucial to identification of microbes or understanding differences observed in microbes from the same lineage.

Why is the sequencing of proteins and peptides necessary if the genome, sequenced and assembled, can be used to predict the amino acid sequence of translated genomic products? The rule

of one gene–one transcript–one protein is frequently ignored in the microbial world [92]. Posttranscriptional modification of RNA called RNA editing has been well documented [93]. Viruses and bacteria can very specifically alter a nucleotide in transcribed RNA resulting in a change in the translation if a coding sequence resulting in the incorporation of a different amino acid from that represented in the coding sequence. Other mechanisms used include genomic stuttering and insertion of sequences at the level of the RNA. As a result, direct protein profiling is essential for determination and identification of proteins and variants thereof. A more thorough discussion of the changes in microbial features as a result of culture effects are discussed in greater detail in Chaps. 3, 4, 5, and 8.

Also, protein profiling can provide information regarding the expression of genetic elements such as virulence factors. As previously noted, microbes are promiscuous in the transfer of genetic sequences but are not always precise. Transferred genes may or may not be associated with the correct genetic elements for expression. Correct and functional protein folding may not occur when transferred into a new microbial host.

Finally, an important concept in the analysis of microbes for forensic analysis is related to the concept of intent. Since natural infections of high-consequence microbial pathogens occur worldwide, it is important to be able to distinguish the origination of materials. This information can be used to determine attribution. Particularly important to the concept of attribution is the addition of specific materials to the microbial preparation intended to enhance microbial virulence or to promote microbial survival. These materials include flow factors added to aerosolized preparation to improve distribution of the target microbe and encapsulants which can be used to stabilize microbes [94, 95]. The purpose of these types of stabilizers is to promote the survival of the microbe in harsh environments or to act as a trigger for microorganism growth until the correct environment (pH, temperature, pressure, etc.) is reached. A thorough discussion of this concept is beyond the scope of this chapter but is addressed by Wahl (Chap. 8, this volume).

## 2.6 Conclusion: A Toolbox for Characterization of Microbes for Forensics

Microbiologists have an extensive combination of tools available to them which include the old standards of culture, microscopy, and biochemistry upon which the field of microbiology was built as well as new standards which are heavily attentive to differences in DNA sequence. While the genetic profiling methods will have a remarkable impact on the ability to rapidly identify an unknown pathogen, speciate microbes, and discriminate isolates, other techniques should be applied in order to capture information for attribution which is not located in the genome. In the postgenomic age, tools such as mass spectrometry hold promise for general application to the area of attribution by providing information regarding contaminants, coassociating viruses, media, or host residues and other hitchhiking factors. The highest confidence approaches will employ a variety of orthogonal and parallel processes in order to gain the greatest argument in support of an attribution source.

## References

1. Budowle B, Johnson MD, Fraser CM, Leighton TJ, Murch RS, Chakraborty R (2005) Genetic analysis and attribution of microbial forensics evidence. Crit Rev Microbiol 31(4):233–254
2. Murch RS (2003) Microbial forensics: building a national capacity to investigate bioterrorism. Biosec Bioterrorism 1(2):117–122
3. Fletcher J, Bender C, Budowle B, Cobb WT, Gold SE, Ishimaru CA, Luster D, Melcher U, Murch R, Scherm H, Seem RC, Sherwood JL, Sobral BW, Tolin SA (2006) Plant pathogen forensics: capabilities, needs, and recommendations. Microbiol Mol Biol Rev 70(2):450–471
4. DeLong EF, Pace NR (2001) Environmental diversity of bacteria and archaea. Syst Biol 50(4):470–478
5. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, Tong S, Urbani C, Comer JA, Lim W, Rollin PE, Dowell SF, Ling AE, Humphrey CD, Shieh WJ, Guarner J, Paddock CD, Rota P, Fields B, DeRisi J, Yang JY, Cox N, Hughes JM, LeDuc JW, Bellini WJ, Anderson LJ, SARS Working Group (2003) A novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med 348(20):1953–1966, Epub 2003 Apr 10

6. Scherer C, Müller KD, Rath PM, Rainer AM (2003) Ansorg influence of culture conditions on the fatty acid profiles of laboratory-adapted and freshly isolated strains of *Helicobacter pylori*. J Clin Microbiol 41(3):1114–1117

7. Casano FJ, Wells JM, van der Zwet T (1987) Effect of growth medium and physiological age on the fatty acid analysis of *Erwinia amylovora*. Acta Hort (ISHS) 217:41–42

8. Kankaanpää P, Yang B, Kallio H, Isolauri E, Salminen S (2004) Effects of polyunsaturated fatty acids in growth medium on lipid composition and on physicochemical surface properties of Lactobacilli. Appl Environ Microbiol 70(1):129–136

9. Rhee SK, Pack MY (1980) Effect of environmental pH on chain length of *Lactobacillus bulgaricus*. J Bacteriol 144(3):865–868

10. Murdoch DR, Greenlees RL (2004) BacT/ALERT blood culture bottles by direct Gram rapid identification of *Staphylococcus aureus* from stain characteristics. J Clin Pathol 57:199–201

11. Colwell R et al (2000) In non-culturable microorganisms in the environment. ASM, Washington DC, pp 325–342

12. Xu HS, Roberts N, Singleton FL, Attwell RW, Grimes DJ, Colwell RR (1982) Survival and viability of nonculturable *Escherichia coli* and *Vibrio cholerae* in the estuarine and marine environment. Microb Ecol 8: 313–323

13. Oliver JD (2005) The viable but nonculturable state in bacteria. J Microbiol 43 Spec No:93–100

14. Prüss BM, Dietrich R, Nibler B, Märtlbauer E, Scherer S (1999) The hemolytic enterotoxin HBL is broadly distributed among species of the *Bacillus cereus* group. Appl Environ Microbiol 65(12):5436–5442

15. Rascoe J, Berg M, Melcher U, Mitchell F, Bruton BD, Pair SD, Fletcher J (2003) Identification, phylogenetic analysis and biological characterization of *Serratia marcescens* strains causing cucurbit yellow vine disease. Phytopathology 93:1233–1239

16. Foschino R, Gallina S, Andrighetto C, Rossetti L, Galli A (2004) Comparison of cultural methods for the identification and molecular investigation of yeasts from sourdoughs for Italian sweet baked products. FEMS Yeast Res 4(6):609–618

17. Blough HA (I971) Fatty acid composition of individual phospholipids of influenza virus. J Gen Virol 12(3):7–320

18. Blough HA, Lawson DEM (1968) The lipids of paramyxoviruses: a comparative study of Sendai and Newcastle disease viruses. Virology 36:286–292

19. Barnes JA, Pehowich DJ, Allen TM (1987) Characterization of the phospholipid and fatty acid composition of Sendai virus. J Lipid Res 28:130–138

20. Papaparaskevas J, Houhoula DP, Papadimitriou M, Saroglou G, Legakis NJ, Zerva L (2004) Ruling out Bacillus anthracis. Emerg Infect Dis. Available from http://wwwnc.cdc.gov/eid/article/10/4/03-0544.htm

21. Knisely RF (1966) Selective medium for *Bacillus anthracis*. J Bacteriol 92:784–786

22. Guttmann DM, Ellar DJ (2000) Phenotypic and genotypic comparisons of 23 strains from the *Bacillus cereus* complex for a selection of known and putative *B. thuringiensis* virulence factors. FEMS Microbiol Lett 188:7–13

23. Casselman A (2007) Scientific American. 4 Oct 2007. (http://www.sciam.com/article.cfm?id=strange-but-true-largest-organism-is-fungus)

24. Drake JW, Charlesworth B, Charlesworth D, Crow JF (1998) Rates of spontaneous mutation. Genetics 148:1667–1686

25. Tago Y, Imai M, Ihara M, Atofuji H, Nagata Y, Yamamoto K (2005) *Escherichia coli* mutator (Delta) polA is defective in base mismatch correction: the nature of in vivo DNA replication errors. J Mol Biol 351(2):299–308

26. Syvanen M (1985) Cross-species gene transfer; implications for a new theory of evolution. J Theor Biol 112:333. doi:10.1016/S0022-5193(85)80291-5, Retrieved on 2007-09-05

27. Woese C (2004) A new biology for a new century. Microbiol Mol Biol Rev 68(2):173–186

28. Bu W, Mamedova A, Tan M, Xia M, Jiang X, Hegde RS (2008) Structural basis for the receptor binding specificity of Norwalk virus. J Virol 82(11):5340–5347, Epub 2008 Apr 2

29. Neu U, Woellner K, Gauglitz G, Stehle T (2008) Structural basis of GM1 ganglioside recognition by simian virus 40. Proc Natl Acad Sci USA 105(13):5219–5224, Epub 2008 Mar 19

30. Chandrasekaran A, Srinivasan A, Raman R, Viswanathan K, Raguram S, Tumpey TM, Sasisekharan V, Sasisekharan R (2008) Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. Nat Biotechnol 26(1):107–113, Epub 2008 Jan 6

31. Thibodeaux CJ, Melançon CE, Liu HW (2007) Unusual sugar biosynthesis and natural product glycodiversification. Nature 446(7139):1008–1016

32. Drickamer K, Taylor ME (1998) Evolving views of protein glycosylation. Trends Biochem Sci 23(9): 321–324

33. Granados-Gonzalez V, Claret J, Berlier W, Vincent N, Urcuqui-Inchima S, Lucht F, Defontaine C, Pinter A, Genin C, Riffard S (2008) Opposite immune reactivity of serum IgG and secretory IgA to conformational recombinant proteins mimicking V1/V2 domains of three different HIV type 1 subtypes depending on glycosylation. AIDS Res Hum Retroviruses 24(2): 289–299

34. Demelbauer UM, Plematl A, Kremser L, Allmaier G, Josic D, Rizzi A (2004) Characterization of glyco isoforms in plasma-derived human antithrombin by on-line capillary zone electrophoresis-electrospray ionization-quadrupole ion trap-mass spectrometry of the intact glycoproteins. Electrophoresis 25(13): 2026–2032

35. Berry N, Jenkins A, Martin J, Davis C, Wood D, Schild G, Bottiger M, Holmes H, Minor P, Almond N (2005) Mitochondrial DNA and retroviral RNA analyses of archival oral polio vaccine (OPV CHAT)

materials: evidence of macaque nuclear sequences confirms substrate identity. Vaccine 23(14): 1639–1648

36. Gill P (2002) Role of short tandem repeat DNA in forensic casework in the UK–past, present, and future perspectives. Biotechniques 32(2):366–372

37. Saiki RK et al (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

38. Couzin J (2002) Bioterrorism. A call for restraint on biological data. Science 297(5582):749–751

39. Cello J, Paul AV, Wimmer E (2002) Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. Science 297(5583):1016–1018, Epub 2002 Jul 11

40. Atlas R, Campbell P, Cozzarelli NR, Curfman G, Enquist L, Fink G, Flanagin A, Fletcher J, George E, Hammes G, Heyman D, Inglesby T, Kaplan S, Kennedy D, Krug J, Levinson R, Marcus E, Metzger H, Morse SS, O'Brien A, Onderdonk A, Poste G, Renault B, Rich R, Rosengard A, Salzberg S, Scanlan M, Shenk T, Tabor H, Varmus H, Wimmer E, Yamamoto K, Journal Editors and Authors Group (2003) Statement on the consideration of biodefence and biosecurity. Nature 421(6925):771, Epub 2003 Feb 15

41. Koehler TM (2002) *Bacillus anthracis* genetics and virulence gene regulation. Curr Top Microbiol Immunol 271:143–164

42. González JM Jr, Brown BJ, Carlton BC (1982) Transfer of *Bacillus thuringiensis* plasmids coding for delta-endotoxin among strains of *B. thuringiensis* and *B. cereus*. Proc Natl Acad Sci USA 79(22): 6951–6955

43. Woese C, Fox G (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA 74(11):5088–5090

44. Kim YR, Batt CA (2008) Riboprint and virulence gene patterns for *Bacillus cereus* and related species. J Microbiol Biotechnol 18(6):1146–1155

45. Maskos U, Southern E (1993) A novel method for the parallel analysis of multiple mutations in multiple samples. Nucleic Acids Res 21(9):2269–2270

46. Sanders CC, Peyret M, Moland ES, Cavalieri SJ, Shubert C, Thomson KS, Boeufgras JM, Sanders WE Jr (2001) Potential impact of the VITEK 2 system and the advanced expert system on the clinical laboratory of a university-based hospital. J Clin Microbiol 39(7): 2379–2385

47. Munson EL, Doern GV (2007) Comparison of three commercial test systems for biotyping *Haemophilus influenzae* and *Haemophilus parainfluenzae*. J Clin Microbiol 45(12):4051–4053, Epub 2007

48. Butler JM (2007) Short tandem repeat typing technologies used in human identity testing. Biotechniques 43(4):ii–v

49. Truu J, Talpsep E, Heinaru E, Stottmeister U, Wand H, Heinaru A (1999) Comparison of API 20NE and Biolog GN identification systems assessed by techniques of multivariate analyses. J Microbiol Methods 36(3):193–201

50. Venkatasubbarao S (2004) Microarrays – status and prospects. Trends Biotechnol 22(12):630–637

51. Southern E (2001) DNA microarrays: history and overview. In: Rampal (Humana) JB (ed) DNA arrays, vol 170., pp 1–16

52. Cole JR, Chai B, Farris RJ, Wang Q, Kulam SA, McGarrell DM, Garrity GM, Tiedje JM (2005) The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Res 33(Database issue):D294–D296

53. Maidak BL, Olsen GJ, Larsen N, Overbeek R, McCaughey MJ, Woese CR (1997) The RDP (Ribosomal Database Project). Nucleic Acids Res 25(1):109–111

54. Wuyts J, Van de Peer Y, Winkelmans T, De Wachter R (2002) The European database on small subunit ribosomal RNA 2002. Nucleic Acids Res 30(1): 183–185

55. Maidak BL, Cole JR, Parker CT Jr, Garrity GM, Larsen N, Li B, Lilburn TG, McCaughey MJ, Olsen GJ, Overbeek R, Pramanik S, Schmidt TM, Tiedje JM, Woese CR (1999) A new version of the RDP (Ribosomal Database Project). Nucleic Acids Res 27(1):171–173

56. Maidak BL, Cole JR, Lilburn TG, Parker CT Jr, Saxman PR, Farris RJ, Garrity GM, Olsen GJ, Schmidt TM, Tiedje JM (2001) The RDP-II (Ribosomal Database Project). Nucleic Acids Res 29(1):173–174

57. Bavykin SG, Mikhailovich VM, Zakharyev VM, Lysov YP, Kelly JJ, Alferov OS, Gavin IM, Kukhtin AV, Jackman J, Stahl DA, Chandler D, Mirzabekov AD (2008) Discrimination of *Bacillus anthracis* and closely related microorganisms by analysis of 16S and 23S rRNA with oligonucleotide microarray. Chem Biol Interact 171(2):212–235, Epub 2007 Sep 12

58. Theodore M, Jackman J, Bethea W (2004) Microarray for counterproliferation. John Hopkin Tech Digest

59. Stahl DA (2004) High-throughput techniques for analyzing complex bacterial communities. Adv Exp Med Biol 547:5–17

60. Liu WT, Mirzabekov AD, Stahl DA (2001) Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. Environ Microbiol 3(10): 619–629

61. Hofstadler SA, Sampath R, Blyn LB, Eshoo MW, Hall TA, Jiang Y, Drader JJ, Hannis JC, Sannes-Lowery KA, Cummins LL, Libby B, Walcott DJ, Schink A, Massire C, Ranken R, Gutierrez J, Manalili S, Ivy C, Melton R, Levene H, Barrett-Wilt G, Li F, Zapp V, White N, Samant V, McNeil JA, Knize D, Robbins D, Rudnick K, Desai A, Moradi E, Ecker DJ (2005) TIGER: the universal biosensor. Int J Mass Spectrom 242(1):23–41

62. Hofstadler SA, Hari KL, Sampath R, Blyn LB, Eshoo MW, Ecker DJ (2006) Detection of microbial agents using broad-range PCR with detection by mass spectrometry: the TIGER concept. Chapter in Encyclopedia of rapid microbiological methods Encyclopedia of Rapid Microbiol Methods 3:287–306

63. Donat V, Biosca EG, Peñalver J, López MM (2007) Exploring diversity among Spanish strains of *Erwinia amylovora* and possible infection sources. J Appl Microbiol 103(5):1639–1649

64. Krishnan MY, Radhakrishnan I, Joseph BV, Madhavi Latha GK, Ajay Kumar R, Mundayoor S (2007) Combined use of amplified fragment length polymorphism and IS6110-RFLP in fingerprinting clinical isolates of *Mycobacterium tuberculosis* from Kerala, South India. BMC Infect Dis 7:86

65. Meudt HM, Clarke AC (2007) Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends Plant Sci 12(3):106–117, Epub 2007 Feb 14. Review

66. Fan H, Chu JY (2007) A brief review of short tandem repeat mutation. Genomics Proteomics Bioinformatics 5(1):7–14

67. Grissa I, Bouchon P, Pourcel C, Vergnaud G (2008) On-line resources for bacterial micro-evolution studies using MLVA or CRISPR typing. Biochimie 90(4):660–668, Epub 2007

68. Lindstedt BA (2005) Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. Electrophoresis 26(13):2567–2582

69. van Belkum A (2007) Tracing isolates of bacterial species by multilocus variable number of tandem repeat analysis (MLVA). FEMS Immunol Med Microbiol 49(1):22–27

70. Le Flèche P, Hauck Y, Onteniente L, Prieur A, Denoeud F, Ramisse V, Sylvestre P, Benson G, Ramisse F, Vergnaud G (2001) A tandem repeats database for bacterial genomes: application to the genotyping of *Yersinia pestis* and *Bacillus anthracis*. BMC Microbiol 1:2, Epub 2001 Mar 30

71. Denoeud F, Vergnaud G (2004) Identification of polymorphic tandem repeats by direct comparison of genome sequence from different bacterial strains: a Web-based resource. BMC Bioinformatics 5:4

72. Al Dahouk S, Le Flèche P, Nöckler K, Jacques I, Grayon M, Scholz HC, Tomaso H, Vergnaud G, Neubauer H (2007) Evaluation of Brucella MLVA typing for human brucellosis. J Microbiol Methods 69(1):137–145

73. Fabre M, Koeck JL, Le Flèche P, Simon F, Hervé V, Vergnaud G, Pourcel C (2004) High genetic diversity revealed by variable-number tandem repeat genotyping and analysis of hsp65 gene polymorphism in a large collection of "*Mycobacterium canetti*" strains indicates that the *M. tuberculosis* complex is a recently emerged clone of "*M. canetti*". J Clin Microbiol 42(7):3248–3255

74. Pourcel C, André-Mazeaud F, Neubauer H, Ramisse F, Vergnaud G (2004) Tandem repeats analysis for the high resolution phylogenetic analysis of *Yersinia pestis*. BMC Microbiol 4:22

75. Lista F, Faggioni G, Valjevac S, Ciammaruconi A, Vaissaire J, le Doujet C, Gorgé O, De Santis R, Carattoli A, Ciervo A, Fasanella A, Orsini F, D'Amelio R, Pourcel C, Cassone A, Vergnaud G (2006) Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis. BMC Microbiol 6:33

76. Keim P, Price LB, Klevytska AM, Smith KL, Schupp JM, Okinaka R, Jackson PJ, Hugh-Jones ME (2000) Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. J Bacteriol 182(10):2928–2936

77. Kenefic LJ, Beaudry J, Trim C, Daly R, Parmar R, Zanecki S, Huynh L, Van Ert MN, Wagner DM, Graham T, Keim P (2008) High resolution genotyping of *Bacillus anthracis* outbreak strains using four highly mutable single nucleotide repeat markers. Lett Appl Microbiol 46(5):600–603, Epub 2008 Mar 18

78. Budowle B, van Daal A (2008) Forensically relevant SNP classes. Biotechniques 44(5):603–608, 610

79. Thorisson GA, Stein LD (2003) The SNP Consortium website: past, present and future. Nucleic Acids Res 31(1):124–127

80. Binnewies TT, Motro Y, Hallin PF, Lund O, Dunn D, La T, Hampson DJ, Bellgard M, Wassenaar TM, Ussery DW (2006) Ten years of bacterial genome sequencing: comparative-genomics-based discoveries. Funct Integr Genomics 6(3):165–185, Epub 2006 May 12

81. Read TD, Salzberg SL, Pop M, Shumway M, Umayam L, Jiang L, Holtzapple E, Busch JD, Smith KL, Schupp JM, Solomon D, Keim P, Fraser CM (2002) Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*. Science 296:2028–2033

82. Pearson T, Busch JD, Ravel J, Read TD, Rhoton SD, U'Ren JM, Simonson TS, Kachur SM, Leadem RR, Cardon ML, Van Ert MN, Huynh LY, Fraser CM, Keim P (2004) Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing. Proc Natl Acad Sci USA 101:13536–13541

83. Van Ert MN, Easterday WR, Simonson TS, U'Ren JM, Pearson T, Kenefic LJ, Busch JD, Huynh LY, Dukerich M, Trim CB, Beaudry J, Welty-Bernard A, Read T, Fraser CM, Ravel J, Keim P (2007) Strain-specific single-nucleotide polymorphism assays for the *Bacillus anthracis* Ames strain. J Clin Microbiol 45(1):47–53, Epub 2006 Nov 8

84. Tanaka MM, Bergstrom CT, Levin BR (2003) The evolution of mutator genes in bacterial populations: the roles of environmental change and timing. Genetics 164(3):843–854

85. Shaver AC, Dombrowski PG, Sweeney JY, Treis T, Zappala RM, Sniegowski PD (2002) Fitness evolution and the rise of mutator alleles in experimental *Escherichia coli* populations. Genetics 162(2):557–566

86. Giraud A, Matic I, Tenaillon O et al (2001) Cost and benefits of high mutation rates. Science 291:2606–2608

87. Yang H, Miller JH (2008) Deletion of dnaN1 generates a mutator phenotype in *Bacillus anthracis*. DNA Repair (Amst) 7(3):507–514, Epub 2008 Jan 31

88. Oberacher H, Parson W (2007) Forensic DNA fingerprinting by liquid chromatography-electrospray ionization mass spectrometry. Biotechniques 43(4):vii–xiii

89. Snow NH (2007) Biological, clinical, and forensic analysis using comprehensive two-dimensional gas chromatography. Adv Chromatogr 45:215–243

90. Ammann AA (2007) Inductively coupled plasma mass spectrometry (ICP MS): a versatile tool. J Mass Spectrom 42(4):419–427

91. Sobrino B, Carracedo A (2005) SNP typing in forensic genetics: a review. Methods Mol Biol 297: 107–126

92. Kiss T (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. EMBO J 20:3617–3622

93. Brennicke A, Marchfelder A, Binder S (1999) RNA editing. FEMS Microbiol Rev 23(3):297–316

94. Gharapetian H, Maleki M, O'Shea GM, Carpenter RC, Sun AM (1987) Polyacrylate microcapsules for cell encapsulation: effects of copolymer structure on membrane properties. Biotechnol Bioeng 30(6): 775–779

95. Shum HC, Kim JW, Weitz DA (2008) Microfluidic fabrication of monodisperse biocompatible and biodegradable polymersomes with controlled permeability. J Am Chem Soc 130:9543–9549