DNA Assays for Detection, Identification, and Individualization of Select Agent Microorganisms

Susan W. Jones, Michael E. Dobson, Stephen C. Francesconi, Richard Schoske, Robert Crawford

The Armed Forces Institute of Pathology, Division of Microbiology, Washington, DC, USA

Abstract

The purpose of this article is to review the status of DNA assays used for the detection, identification, and individualization of Bacillus anthracis, Yersinia pestis, Francisella tularensis, Burkholderia mallei, and Brucella abortus. These select agent microorganisms are historically significant as they have either been used or experimented with as a bioweapon or as a terrorist agent and are the subject of intense research in the areas of biodefense and bioforensics. If the presence of a biological agent is suspected, sensitive and specific assays for rapid detection and identification are necessary. However, DNA methods for identification of the sample may also be applied in order to individualize the strain and potentially determine the source of the microorganism. Methods used at the Armed Forces Institute of Pathology (AFIP) for select agent microbial DNA analyses include DNA extraction, DNA quantitation, real-time polymerase chain reaction (real-time PCR) of genetic targets unique to the select agent microorganism, microbial 16S ribosomal RNA gene DNA sequencing, amplified fragment length polymorphism polymerase chain reaction (AFLP-PCR), and more recently, repetitive element polymerase chain reaction (REP-PCR) DNA fingerprinting. The methodologies of 16S ribosomal RNA gene DNA sequencing and DNA fingerprinting of microorganisms are well established within the field of diagnostic microbiology for DNA identification purposes, as well as DNA typing for epidemiological and genetic relatedness studies. 16S ribosomal RNA gene DNA sequencing and AFLP DNA fingerprinting have been validated at the Armed Force Institute of Pathology (AFIP) laboratory for identification purposes and can be used as a possible strain typing tool for Bacillus anthracis, Yersinia pestis, Francisella tularensis, as well as Brucella and Burkholderia species. The continued development and implementation of new DNA based methods with increased sensitivity and defined specificity will be particularly useful for the detection of residual microbial DNA signature in situations where the microorganism has been rendered nonviable by decontamination procedures or not able to be cultured on microbiological media.

Biological warfare, defined as using microorganisms in offensive warfare, has been documented since the Middle Ages. Many of the select agent microorganisms have been weaponized by various countries in more recent times (1). Though the Biological Weapons Convention was signed in 1972, this treaty could not stop the use of microorganisms in acts of bioterrorism, by either rogue nations or terrorist groups, to cause illness or elicit psychological effects such as fear (2). Recently, several terrorist acts using Bacillus anthracis Ames spores have been perpetrated, including the attack on the US postal service in 2001, the Bacillus anthracis attacks in Japan by the Aum Shinrikyo, and others, including hoaxes (3,4). These attacks have renewed fears that these dangerous and highly pathogenic microorganisms may be used for bioterrorism. Several articles reviewing the use of these organisms as a possible weapon in bioterrorist events have been published (5-9).

The need to detect microbial agents that could be used in bioterrorist acts has motivated
molecular biologists to develop increasingly sensitive, rapid methods of detection for DNA identification and more highly discriminating assays to genetically characterize and potentially track these microorganisms. This need becomes more urgent because growth of some these organisms in culture may require several days.

Molecular microbiological testing methods particularly with DNA polymerase chain reaction (PCR) are used for rapid DNA identification of organisms in clinical samples. At the AFIP laboratory, DNA is isolated from cultured organism by detergent lysis (home-made buffers) combined with organic extraction and DNA precipitation (10). The DNA is then quantitated using the Pico-Green® DNA dye binding assay kit (Molecular Probes, Inc. Eugene, OR, USA) with the fluorescence read from DNA samples placed in a 96-well plate format in the BioRad Icycler®. The Pico-Green dye is specific for double stranded DNA, easily standardized and able to detect picogram quantities of DNA. It should be noted that the PicoGreen DNA dye binding assay is not specific for any particular type of DNA such as microbial, human, or plant DNA (11).

Environmental samples (usually swabs) submitted to the AFIP for testing are cultured for the presence of organism. Then both the original collection swab and a microbiological broth culture from the swab (after incubation overnight) are boiled to release the DNA, and real-time PCR screening for select agent microorganisms is performed on the original and the cultured samples. If either the overnight microbiological broth culture or the original collection swab is positive for organism, then more DNA identification testing, as well as additional culture is performed for confirmation. Several PCR based methodologies are utilized for the detection, characterization, and individualization of genetic signature from these select agent microorganisms.

**Real-time Polymerase Chain Reaction**

The most sensitive DNA detection assay is the real-time PCR technique. These assays are performed using PCR primers and an internal probe containing a fluorescent reporter and quencher molecule. The progress of the reaction is followed on a cycle-to-cycle basis by the measurement of increasing fluorescence, using a real-time PCR instrument. Appropriate genomic or plasmid clone DNA controls are used to generate a standard curve in order to determine an approximate number of DNA chromosomal copies of an organism, or the number of plasmid DNA copies of a plasmid-borne gene that are present. Real-time PCR assays are routinely performed at AFIP to detect the presence of DNA from *Bacillus anthracis* in environmental samples that consist of swabs, contaminated documents, or suspicious powders. These environmental samples are tested, using real-time PCR for unique plasmid-borne and chromosomal genes in the *B. anthracis* genome, such as the lethal factor (lef), edema factor (cya), and protective antigen gene A (pagA) located on the pXO1 plasmid, the capsular protein gene A, B, and C (capA, capB, and capC) located on the pXO2 plasmid, and the chromosomal SASP (small acid soluble protein) gene. These assays are performed in the international network of Laboratory Response Network (LRN) laboratories. The LRN was established in order to be prepared for bioterrorism events and leadership for the LRN is provided by the Centers for Disease Control and Prevention in Atlanta, GA (12).

Chromosomal real-time PCR assays such as the DNA gyrase A gene (gyrA) and other chromosomal loci are useful for detecting the presence of *B. anthracis* genome and determining gene copy numbers present in a sample (13). The ability to determine the number of copies of an organism is of particular interest with environmental samples that may contain only a few non-culturable cells or spores that are not able to be cultured on microbiological media, or have been rendered nonviable by irradiation or decontamination procedures. Real-time PCR assays for chromosomal genes, in general, are not as sensitive as the plasmid-borne gene assays due to the presence of more than one copy of the pXO1 and/or pXO2 plasmid per *B. anthracis* chromosome in the bacterial cell. The presence or absence of these plasmids in certain *B. anthracis* strains is an attribute that can be used for DNA identification screening and initial detection of virulent strains. The *B. anthracis* Ames strain contains both plasmids, the *B. anthracis* Sterne strain contains only the pXO1 plasmid, and the *B. anthracis* Pasteur strain contains only the pXO2 plasmid, however it is possible that either one or both of these plasmids could
be lost during extended laboratory maintenance and repetitive subculture (12,14,15).

Yersinia pestis, Francisella tularensis, Brucella spp., and Burkholderia spp. could also potentially be used in a bioterrorism event. Therefore, real-time PCR assays for these organisms have been developed for rapid screening and detection of genetic signature. Real-time PCR assays for Y. pestis, include detection of the presence of the plasminogen activator (pla) gene (16), the murine toxin (mlt) gene, and the fraction 1 antigen (f1a) gene (17). Real-time PCR assays for Y. pestis have been studied at AFIP, for military applications, using a field-deployable thermalcycler (RAPID™, Ruggedized Advanced Pathogen Identification Device from Idaho Technologies, Inc., Salt Lake City, UT, USA) (18). PCR reagent kits for several bioterror organisms are commercially available from Idaho Technologies, Inc. and are utilized in various biodefense-related studies at the AFIP. The AFIP laboratory utilizes the Applied Biosystems, Inc. (ABI) Sequence Detection System 7900 HT, the ABI Sequence Detection System 7700, and the Idaho Technologies, Inc. RAPID platforms for microbial real-time PCR assays for select agent bacteria. Real-time PCR testing is performed on the Y. pestis plasminogen activator gene (pla) and the murine toxin gene (mlt). The Y. pestis murine toxin gene is located on the largest ~100 kb plasmid (pMT1) of three plasmids, pMT1, pCD1, and pPCP1, contained in Y. pestis strains, but not the two closely related species of Yersinia pseudotuberculosis and Yersinia enterocolitica (19). Y. pestis has been divided into three biogroups (biovars), Orientalis, Medievalis, and Antiqua, based on metabolic tests for nitrate reduction and the ability to grow on glycerol, as well as PCR assays (20).

Real-time PCR targets for chromosomal genetic regions in Francisella tularensis, such as the Francisella outer protein A (ftopA) and a F. tularensis 13-kDa protein and 17 kDa major membrane protein precursor (tu14) gene have been utilized to detect as low as 25 genetic copies (≈50 fg DNA) and detect all three Francisella biogroups: F. tularensis type A subspecies tularensis, F. tularensis type B subspecies holarctica and F. tularensis subspecies novicida (21). These loci are utilized for Francisella DNA identification at the AFIP laboratory with the RAPID thermal cycler (22). Francisella species may contain cryptic plasmids, such as pFNL10, however these plasmids have not been associated with virulence, like the plasmids contained in B. anthracis and Y. pestis (23).

Real-time PCR assays for DNA signature from Brucella abortus were performed at the AFIP for sensitive, specific detection of levels as low as 7.5 fg of DNA from this microbe (24). Additional assays have been developed, by other laboratories, for detection of different Brucella species (25-27), including a well characterized genus specific Brucella species real-time PCR assay specific for the perosamine synthetase (per) gene contained within all species of Brucella (28).

Real-time PCR assays for the sensitive detection of Burkholderia pseudomallei and Burkholderia mallei, the causative agents of the diseases melioidosis and glanders, respectively, have been developed by others (29-31). Melioidosis and glanders are diseases endemic to Asia, however the potential use of B. pseudomallei or B. mallei bacteria as bioweapons underscores the need for a sensitive and specific DNA identification assay using real-time PCR (1).

16S Ribosomal RNA gene DNA Sequencing

DNA identification using 16S Ribosomal RNA gene DNA sequencing (16S rDNA sequencing) involves microbial DNA isolation, DNA quantitation, as well as PCR amplification of the 16S rDNA gene, with subsequent sequencing of the PCR product. Following 16S rDNA PCR amplification, the PCR product is cleaned using gel filtration columns. Then, forward and reverse sequencing primers that bind to various locations across the approximately 1.6-kb 16S rDNA genetic locus are utilized in separate forward and reverse sequencing reactions with reagents provided in the Applied Biosystems, Inc. (ABI) BigDye® Terminator v3.1 Cycle Sequencing Kit (32). The sequencing reaction is cleaned by gel filtration to remove unincorporated dideoxynucleotides and the fluorescently labeled fragments are electrophoresed through an 80-cm capillary array filled with POP-4 polymer using the default settings on the ABI Prism® 3100 Genetic Analyzer, a capillary electrophoretic apparatus. The sequencing data sets are exported from the instrument and aligned using the Sequencher™ sequencing alignment program (GeneCodes, Inc., Ann Arbor, MI, USA). It is very important that multiple PCR and DNA se-
sequencing reactions be performed, so that when the sequence data sets are imported into the alignment program, overlapping sequence data and appropriate base-pair coverage is obtained across the genetic locus, and for each DNA base-pair. The presence of mixed bases, particularly if they are present in B. anthracis 16S rDNA sequence data, should be noted. The consensus DNA sequence from the alignment is then copied into a Basic Local Alignment Search Tool (BLAST) query on the Internet (33). A computerized BLAST query is available at various internet locations such as The Institute for Genomic Research (TIGR) website (http://www.tigr.org) and the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/BLAST) at the National Institute of Health. The BLAST query is a computerized search that contains a complex sequence search algorithm designed to find the best sequence matches in a database of various gene sequences. Once the computerized query is performed, the identities of the sequences in the database that match the questioned sequence are listed in the query results in order from the best match to the worst match. The statistics of the search methods used are well defined at the NCBI website under the BLAST tutorial; however, the statistical significance of a sequence match should not be overstated with regard to bacterial strain DNA identification. The 16S rDNA sequencing technique has been useful for DNA identification of bacteria for verification of the identity of strains assayed by biochemical methods, for determination of the identity of an unknown bacterial species, and for confirmation of identity of difficult to culture microorganisms contained in clinical specimens.

16S rDNA sequencing has been used in studies for the rapid DNA identification and differentiation of B. anthracis from other Bacillus species such as B. cereus and B. thuringiensis. The 16S rDNA sequencing assay produced identical 16S rDNA sequences for the 86 strains of B. anthracis tested, therefore the primary use is for the rapid DNA identification but not strain individualization, of suspect isolates of B. anthracis (34). 16S rDNA sequencing studies have also been performed on Y. pestis, F. tularensis, B. mallei, and B. pseudomallei (35-37).

The 16S rDNA sequencing data and the BLAST results must be carefully examined before making a determination of a DNA identification of Y. pestis. This is because some strains of Y. pseudotuberculosis cannot be reliably discriminated from Y. pestis, because they are so closely related (35). In fact, it has been suggested by multiple genetic studies that Y. pestis is a clone that evolved from Y. pseudotuberculosis thousands of years ago, and that American strains are less diverse due to a genetic bottleneck that occurred when Y. pestis was transferred and spread throughout the Americas (38-40).

The 16S rDNA sequencing studies with Francisella species demonstrated that 16S rDNA sequencing is useful for DNA identification of Francisella tularensis, however the 16S rDNA sequencing technique was not discriminating enough for subspecies or strain differentiation. Moreover, the 16S rDNA sequence data was identical for 54 F. tularensis subspecies holarctica strains that were tested at the AFIP laboratory (36).

Another interesting 16S rDNA sequencing study was performed for DNA identification of B. pseudomallei and B. mallei, the causative agent of melioidosis and glanders, respectively (37). In this study, 22 out of 23 strains of B. mallei had identical 16S rDNA sequences; however, the B. pseudomallei could be divided into nine different 16S rDNA sequencing types based on six positions containing base pair differences. This testing allows further differentiation of the two strains of Burkholderia, however further DNA testing would be needed for more definitive genetic discrimination of strains.

16S rDNA sequencing has also been described for rapid confirmatory identification of Brucella isolates (41). This is because biochemical methods of identification of Brucella species can take up to a week. Therefore, a rapid DNA identification test is needed for DNA identification of Brucella species. The 16S assay is a reliable method for rapid DNA identification of Brucella, however more discriminating DNA identification tests would have to be performed for definitive strain differentiation, as many Brucella strains have identical 16S sequences.

 Amplified Fragment Length Polymorphism PCR

Amplified fragment length polymorphism polymerase chain reaction (AFLP-PCR) is a technique that has been validated and used at the
AFIP for strain differentiation for individualization. Individualization is the process by which morphological, biochemical or genetic characteristics are associated with a microorganism, such that a combination of characteristics are unique to that microorganism and exclude other microorganisms. The AFLP-PCR technique is particularly useful when little genomic sequence information on the bacterial strain is available. The technique involves restriction enzyme digestion of the bacterial genomic DNA, DNA ligation of the digested DNA to oligonucleotide linker molecules, and amplification of the DNA fragments using primers specific for the linkers attached to the DNA fragments. The PCR amplification primers are labeled with a fluorescent label on the 5' end of the primer, so that the amplified DNA fragments are labeled in the PCR reaction. These fluorescent labeled primers are available in an AFLP kit that is sold commercially by Applied Biosystems, Inc. (ABI, Foster City, CA, USA). The fluorescent labeled AFLP DNA fragments are then electrophoresed on the ABI Prism® 3100 Genetic Analyzer using the 50cm capillary with POP-4 polymer and the GeneScan® 500 ROX size standard. The AFLP technique has been studied with various *Bacillus anthracis* strains, as well as other closely related *Bacillus* species, such as *Bacillus cereus* and *Bacillus thuringiensis*. These closely related species are often referred to as “nearest neighbors” although they can be readily distinguished from *B. anthracis* using AFLP (42,43). This study demonstrated that relative to the *B. cereus* and *B. thuringiensis* genome, the genome of *B. anthracis* is highly monomorphic. This monomorphism means that all strains of *B. anthracis* have genomes that are very similar (44).

An AFLP study has not been published for evaluation of diversity of *Y. pestis* strains. Previous studies using pulsed field gel electrophoresis and restriction fragment length polymorphism demonstrated some genome diversity in *Y. pestis* strains from the United States (45). AFLP studies were performed on *F. tularensis* subspecies holarctica compared with other *Francisella* species and these studies revealed that *F. tularensis* has some diversity that allows for genetic discrimination using AFLP (36). No AFLP studies have been published evaluating the diversity of *Brucella* species or *Burkholderia mallei* and *Burkholderia pseudomallei* strains. AFLP is a useful DNA technique in the context of a database collection of AFLP types of bacterial strains from the same or closely related species to see if the AFLP typing of an unknown strain from the same genus or species is phylogenetically related and where the strain fits into the phylogenetic analysis. However, without a database collection of AFLP genotypes it is not a stand-alone DNA identification test.

**Other DNA Fingerprinting Assays for Strain Individualization**

Other DNA fingerprinting assays have been utilized by other laboratories for studying strain diversity and the strain discrimination of *B. anthracis, Y. pestis, F. tularensis*, *Brucella* species, and *Burkholderia* species. These assays include Repetitive Element PCR (REP-PCR) and the Multilocus Variable Number of Tandem Repeats Assay (MLVA).

REP-PCR is currently under investigation in our laboratory and has been described for a number of select agent bacteria including *B. anthracis, Franciscella* and *Brucella* species (46-48). REP-PCR uses oligonucleotide primers that are homologous to the repetitive sequences that are scattered throughout the bacterial genome. The genomic regions between these repetitive sequences are amplified in the PCR reaction, and the analysis of the amplified PCR products has been standardized in an automated system that increases the ability to rapidly and reliably reproduce highly sensitive REP-PCR DNA fingerprints by using microcapillary electrophoresis chip (49).

MLVA is a discriminating PCR assay that has been studied in other laboratories and developed for the rapid genetic strain typing of *B. anthracis, Y. pestis*, and *F. tularensis* strains (20, 50-55). The MLVA assay utilizes oligonucleotide DNA primers that flank tandem repeat regions in the genomes of each species of bacteria in a species-specific PCR amplification reaction for differentiating the strains within the species. Various bacterial strains within a species may contain allelic variation at the different genetic loci being tested, and can subsequently be divided into similar groups, called clades, based on the strain allelic DNA typings. These phylogenetic groupings allow one to determine the level of diversity contained within a bacterial species. The MLVA assay involves the measurement of discrete tandem repeat alleles within the bacterial genome. It is, therefore,
easily standardized and highly amenable to automated detection methods and quality assurance. MLVA has also been described for Brucella strains using an octameric repeat present in various genes of the Brucella genome (56). This study allowed for separation of the Brucella species, particularly B. abortus and the various biovars.

Another technique that is very similar to MLVA is multilocus sequence typing that has been described for the study of Burkholderia strains. The multilocus sequence typing assay is based on sequence polymorphisms contained within functional genes that are evaluated by a multiplex PCR assay specific for the various genetic regions (57).

All of the previously mentioned molecular biology techniques provide genetic information that can be used for DNA identification of a particular select agent bacterial genus and, for some assays, DNA identification can be established at the species level. However, none of these techniques alone can provide strain-level DNA identification and differentiation for absolute source attribution with 100% confidence. Whole genomic sequencing of all bacterial select agent strains, genomic microarrays that query all of the most genetically informative polymorphisms, including single nucleotide polymorphisms (SNP’s) among strains, and bioinformatic genetic information comparisons in a very large database, will no doubt contribute to the effort to obtain DNA identification for individualization and strain source attribution. Some of these studies have already been performed with B. anthracis in order to define polymorphisms to discriminate between individual strains of B. anthracis Ames acquired from various locations (58-60).

Perhaps another approach would be to use multiple different types of assays for DNA testing for signature of select agent microorganisms, a method similar to the deductive algorithm testing procedures used in toxicological testing. Deductive reasoning and selection of the best DNA assay testing algorithm may be based on knowledge of what each DNA assay will yield in terms of genetic information, the significance of the information relevant to the other bacterial strains within the same species, and the DNA information obtained from these assays when testing is performed using closely related nearest neighbor bacterial species. A stepwise strategy for real-time PCR screening for unique bacterial targets, 16S rDNA sequencing for rapid species identification, combined with multiple layers of DNA fingerprinting by AFLP, REP-PCR, and MLVA, as well as recently developed assays, and the statistical power of each technique combined may allow for more definitive species-level DNA identification (Fig. 1).

The quest for individualization for strain attribution of select agent bacterial strains, particularly B. anthracis Ames strains, continues. More definitive strain individualization DNA assays, such as detection of single nucleotide polymorphisms and comparative whole genomic sequencing assays that can further define SNP’s, indels, or other genomic changes in B. anthracis Ames strains from different locations, have already been developed (58-60). These assays and revised phylogenetic comparisons with the B. anthracis Ames strains further refine the individual DNA strain characteristics of B. anthracis Ames in order to get closer to “strain attribution.” The ability to distinguish enough differences between these strains to be able to determine with certainty that a particular strain was definitively used in an act of bioterrorism is very important. These discriminating assays have not yet been developed for other select agent bacterial strains, are necessary for biological threat characterization, and will no doubt be the subject of future research endeavors involving these microorganisms.

**Acknowledgement**

The authors acknowledge the support by the American Registry of Pathology. The opinions and assertions expressed herein are solely those of the authors and should not be construed as official or as views of the United States Department of Defense or the United States Department of the Army or Navy. There are no conflicts of interest, commercial affiliations or other arrangements that constitute a conflict of interest.
References


Received: April 28, 2005
Accepted: May 16, 2005

Correspondence to:
Susan W. Jones
Armed Forces Institute of Pathology
Department of Microbiology
6825 16th Street NW, BLDG 54 RM 54315
Washington, DC 20306, USA
jonessw@afip.osd.mil