Clinical Microbiology Newsletter

Vol. 31, No. 24 www.cmnewsletter.com December 15, 2009

Validation of New Molecular Tests for Microbiological Testing of Clinical Specimens

Kimberlee A. Musser, Ph.D., Director, Bacteriology Laboratory, and Christina Egan, Ph.D., Director, Biodefense Laboratory, Wadsworth Center, New York State Department of Health, Albany, New York

Abstract

Access to the most rapid, sensitive, and specific testing is of critical importance to the microbiology laboratory. However, the tests of interest are constantly changing, as new pathogens emerge, as new antimicrobial markers and virulence genes are discovered, and as new technologies are utilized in the detection and characterization of infectious diseases. The tests available from commercial sources may not fit the exact needs of a laboratory, or the tests may not be FDA approved for patient testing. These factors have led many laboratories to develop and validate new molecular assays for their current patient testing needs. The task of developing and validating these tests is considerable, and the laboratory must learn to navigate the field of molecular testing. This article describes the process of assay development and validation of real-time PCR and nucleic acid sequencing assays; included are descriptions of important considerations, such as the PCR workflow, inhibition testing, proper use of controls, and multiplexing.

Introduction

The introduction of molecular methods has had a positive impact in many areas of diagnostic microbiology (1-4). Microbiology as a field is embracing molecular-diagnostic testing, because these tests have been proven to be more sensitive and specific than classical testing in some situations (5), and they are particularly useful for specimens that may contain fastidious bacteria and viruses (6,7) or when patient treatment has already been initiated (8). Results from genotypic tests, in some situations, are generally more objective than the interpretation of conventional phenotypic characteristics. Each year has seen advances in instrumentation; these advances give rise to less expensive equipment that is easier to use, as well as new methods for extracting nucleic acid and ever-improving PCR reagents. Although a limited number of FDA-approved or cleared methods are available, many laboratories are compelled to develop their own tests to fit their patient testing needs. The latter are referred to as "home-brew" or "laboratory-developed" assays and can be validated for use in clinical and public health microbiology laboratories.

If microbiology laboratories are to take advantage of these technological advances, a new set of rules must be applied to determine whether the tests will have acceptable performance in any given laboratory. The classical methods of culture and biochemical testing must, like all assays, have appropriate quality control (QC) and quality assurance (QA) practices associated with each assay, but the individual tests do not require additional validation. For the newer nucleic acid-based methods, including real-time PCR and nucleic acid sequence analysis, standards and guidelines have been developed and must be adhered to when the performance of the tests is assessed. Such standards include the Clinical Laboratory Improvement Act (CLIA); state standards (when they are at least as stringent as CLIA), such as the NYS Clinical Laboratory Evaluation Program (CLEP); and consensus guidelines and documents such as those available from the Clinical Laboratory Standards Institute. The present article describes the processes of assay development and validation, as well as discussing other considerations, such as types of specimens tested, PCR inhibition, nucleic acid extraction, and PCR controls.

Mailing address: Kimberlee A. Musser, Ph.D., Director, Bacteriology Laboratory, Wadsworth Center, New York State Department of Health, Albany, NY 12208. Tel.: 518-473-2780, Fax: 518-486-7971. E-mail: musser@wadsworth.org

  0196-4399/00 (see frontmatter)

Diagnostic Assay Development

Establishment of a highly robust molecular assay to be implemented in the clinical microbiology laboratory includes two distinct phases. The first phase is the assay development phase, and the second is the validation phase. The two phases are both important and must be carefully carried out if the end product is to be useful for the laboratory and physician.

One begins the development phase of a home-brew molecular assay with a review of the current literature. This provides an understanding of the choices of target genes used in previous studies, potential specificity problems, and additional information that may be useful. Once an appropriate target is selected, primers and probes can be designed. There are many public sources for primer/probe identification, as well as numerous commercially available programs. The proper choice of an analysis program is a critical component of a well-designed assay. Also essential is evaluation of how the assay will be used in the laboratory. If the assay will be multiplexed with other assays already developed in the laboratory, utilizing a program that can assess multiplexing is beneficial. Some software packages are specifically designed for the development of assays for greater than two targets per assay. Other computer programs are used for hybridization probe technology and melting curve analysis. It is important to determine the type of probe chemistry of interest and then to purchase or use a program that is suited for the application and assay under development.

In our laboratory, as part of the assay development phase, a number of important parameters must be met before starting the validation phase. One critical component of a well-designed assay is the determination of the analytical sensitivity of the assay. After an appropriate target is established, and after detailed in silico analysis has been performed to verify that the target is specific to the pathogen or gene of interest, the analytical sensitivity of the assay must be determined. This evaluation establishes a limit of detection (LOD) for the assay, through testing at least 8 to 10 dilutions of the whole organism (if possible). The evaluation should include dilutions of the bacterial suspension that fall in the linear portion of the curve produced during PCR amplification. This dilution series should also include at least one dilution that yields a negative test result for determination of when the reaction has gone to extinction. This testing provides a basis for the assessment of how the assay will perform when testing is extended to clinical specimen matrices. The theoretical LOD for any PCR-based assay is one intact copy of the target sequence. However, due to stochastic and partitioning issues, a single copy cannot, in practice, be reproducibly detected. When we develop assays, we aim for detection of 5 to 25 gene copies or colony-forming units (CFUs). This target level can usually be attained with minimal deleterious effects on the overall assay performance.

In retrospective analysis of a broad sample of individual real-time PCR assay results, we have observed that a clear and consistent break in cycle threshold (C_T) values occurs between detect and non-detect results at around cycle 40. If we assume that this break corresponds to target nucleic acid presence (i.e., ≥ 1 amplifiable copy) or absence (i.e., <1 amplifiable copy), then we can back-calculate the equipment sensitivity to amplicon accumulation-dependent fluorescence change. In our experience with C_T values of 38 to 40, data are not reproducible. This is to be expected based on the theoretical LOD of one target at a C_T value of 40, assuming 100% PCR efficiency. Therefore, running real-time PCR assays past 40 cycles is not considered effective, and the results can be inconsistent and confusing for the end user. We have found data to be highly reproducible below C_T values of 38. Specimens that produce C_T values in the "gray zone" between 38 and 40 should be tested again. It is extremely important for a laboratory to determine its own precise cutoff levels depending on the factors that comprise the assay. Testing and reporting algorithms should be developed and rigorously followed. These algorithms should take into account whether the testing is performed on cultured material or specimens that may contain low levels of the pathogen.

Optimization

In our experience, in-house-developed real-time PCR assays have shown good linear performance, yet the LOD was >100 CFUs. To achieve an acceptable LOD, the next step in the development phase is assay optimization, which includes the testing of a range of primer and probe concentrations and the evaluation of concentrations of $MgCl_2$. Additionally, the optimization of concentrations of other components of the mastermix, including *Taq* polymerase, can also be evaluated. Simply by determining optimal concentrations for these reaction components, we often can reduce the LOD value to within a range acceptable for use. Ever-improving mastermixes have greatly decreased the work necessary in comparison to 10 years ago.

When an assay designed to detect a bacterial pathogen is being developed, a suspension prepared at a concentration

at or just above the LOD should be used, so that subtle improvements in the assay sensitivity can be tracked. For example, when the $MgCl_2$ concentration is optimized (if the commercially available mastermix system allows this), the bacterial suspension concentration is held constant for each reaction, so that the only variables adjusted are the amounts of $MgCl_2$ and water (with the total volume kept consistent). For $MgCl_2$ titrations, concentrations between 1 mM and 7 mM can be tested in 0.5-mM increments. Such testing should be performed in duplicate and repeated over the course of several days. The C_T values can be averaged and charted, to determine whether one concentration or several provide optimal conditions for detection. It should be noted that many real-time PCR buffers contain some $MgCl_2$; this factor makes optimization extremely difficult unless the manufacturer is willing to prepare a lot without $MgCl_2$. Similarly, primer and probe concentrations are often tested, not only to ascertain the optimal conditions, but also to determine whether the concentrations can be reduced without affecting assay sensitivity. This can achieve a cost savings per test and can render the assay potentially better for future multiplexing. Beyond assessment of the concentrations of primers, some laboratories evaluate multiple ratios of the forward and reverse primers (i.e., 1:1, 1:3, 1:5, 3:1, and 5:1). While these adjustments can indeed improve sensitivity, the added complexity may be unjustified if the assay is to eventually be multiplexed.

Linearity and Amplification Efficiency

A LOD evaluation can provide additional valuable information. Important parameters, such as linearity and amplification efficiency, can be obtained from sensitivity analyses if such analyses are performed correctly. Real-time PCR assays have broad inherent linearity, typically 6 to 8 log units, depending on the optical sensitivity of the detection platform used. Linearity data can be obtained if a dilution series encompassing 8 to10 log units of the bacterial suspension are tested in triplicate. For real-time PCR assays, the amplification efficiency can be determined directly from the linearity data. Data obtained from each dilution can be plotted against the concentration or dilution and converted into an amplification efficiency, usually expressed as a percentage. The efficiency (e) is determined as follows:

$$e = (10^{(-1/m)} - 1) \times 100.$$

For example, if the slope (m) of a 10-log dilution standard curve is –3.38, the efficiency of that assay is 97.6%. In our laboratory, an assay is deemed to satisfy acceptance criteria if the amplification efficiency is >90%.

Specificity

After the LOD for the assay has been established, analytical specificity is the next performance criterion examined. Our laboratory does in silico analyses to determine whether any cross-reactivity could arise with organisms that will potentially be present in the type of specimen to be tested, near neighbors of the target of interest, and organisms that can cause clinical symptoms similar those induced by the target pathogen. The potential for cross-reactivity is further evaluated by challenging the assay with specificity panels that include genetic near neighbors, human nucleic acid, common flora found in the particular specimen type, and other organisms that can produce similar clinical symptoms. We have occasionally found that primers and probes exhibit cross-reactivity with an organism that was not identified in the in silico-based analysis. It is critical that every new assay, including multiplex reactions, be evaluated with a detailed specificity panel. This evaluation needs to occur early in the assay development process; valuable staff time and resources are wasted if specificity is not assessed until the validation phase. Specificity testing should be performed using an excess of the genetic material for organisms being evaluated in the panel, in order to detect the weak cross-reactivity that can occur in actual patient specimens. Our laboratory generally uses a 5-µl aliquot of a suspension of bacteria at a concentration of 1 McFarland (equivalent to 10^6 organisms). Because specificity testing is fundamental to the developmental process, it should be performed at least in duplicate.

Clinical Specimens

Another aspect to consider in the developmental process is the types of specimens or samples that will be assayed. Some research supports (9) testing of unprocessed samples or specimens that have not undergone a process for extraction of genetic material from the specimen. However, it must be proven that the specimen type to be tested can be assayed without potential loss of sensitivity and that the specimen type is not associated with substantial inhibition in the assay. Elimination of the processing steps should only be used for assays that have been thoroughly evaluated with statistically significant numbers of a single sample type and that have been found to show no loss of sensitivity.

A goal of the assay development phase may be to create a test that can be used on a wide range of clinical specimens, ranging from cerebrospinal fluid (CSF) to biopsy material, and possibly other sample types, such as water, food, or environmental swabs. Because of the broad spectrum of potential specimen types, upstream processing methods, such as nucleic acid extraction, must be very robust, and appropriate scientific controls must be integrated into the procedure to monitor inhibition or sample matrix effects. The most time- and cost-effective approach for method development, at least in our hands, is to compare several commercially available nucleic acid isolation kits and to determine which yields the highest quality and most consistent nucleic acid products for subsequent real-time PCR analysis. This testing can include previously determined negative specimens spiked with bacterial suspensions, if the testing is for a bacterial pathogen. It is important for a laboratory to utilize a nucleic acid isolation method that is appropriate for as many different applications as possible, as multiplexing is often a consideration for future testing.

Controls

Critical control issues must be addressed during the development of any new PCR assay. A negative control and a no-template control (NTC) are important components of an assay. A negative control usually consists of a known-negative specimen or a bacterial suspension that does not contain target nucleic acid. An NTC typically consists of water or buffer added to the real-time PCR mixture. The NTC is used as a measure of potential amplification

contamination of the reagents. A positive control containing a known low concentration of the target organism should also be included in each analytical run. The positive control has many purposes: besides acting as a qualitative assessment of assay performance, it can serve as a monitor for inhibition (when seeded into a specimen) (10), and it can be used to reduce assay-to-assay variability when recorded over time. Because one potential problem with inclusion of positive controls in the assay is the risk of sample cross-contamination, only low levels of positive control nucleic acid should be used. Positive and negative extraction controls are also extremely beneficial when performing PCR testing. In our laboratory, these controls are incorporated each time that a PCR assay with an extraction step is utilized. The use of an extraction control ensures that the extraction reagents have adequately lysed the cells and that high-quality nucleic acid was obtained; the negative extraction control can also serve to assess potential cross-contamination or reagent contamination.

In certain instances, it is necessary to use plasmid controls that have target sequences inserted in them for determination of analytical sensitivity. The means to obtain appropriate positive control material for assay development and design is a key aspect of implementation of a molecular assay in the laboratory. For certain organisms that are non-cultivable, fastidious, or slow-growing, it is acceptable to create plasmid controls. These plasmid controls are used in the optimization of assay parameters before the validation phase of assay implementation. Although plasmids containing the target sequence of interest can be purchased commercially, depending on the target size and number of targets per assay, these plasmids can be costly. Creation of a plasmid control can be easily performed (although special precautions must be taken when propagating and working with the control to avoid contamination). However, the procedure requires molecular biological expertise, and a laboratorian may need 1 to 2 weeks to create a proper positive control for use in assay development. It is possible to include several target sequences in a single plasmid. In this way, the positive control material generates positive signals for a number of targets, not all of which would normally be present in a single sample. This type of positive control can then be used for multiplex PCR that are developed later for the simultaneous detection of pathogens in a single tube.

Another fundamental feature of a molecular assay used in the clinical microbiology setting is the inclusion of a control to assess potential PCR inhibition. It is well known that PCR inhibitors are present in many clinical specimens; if the molecular test under development does not eliminate potential interfering substances and assess if such inhibitors are present, false-negative results can be a serious issue. Nucleic acid extraction is an important initial step in reducing potential interfering substances before the specimen is tested. Fortunately, many commercially available reaction components (e.g., mastermixes) are also available for real-time PCR, which include additives that reduce the effects of PCR-specific inhibitors that are commonly found in clinical sample matrices (11,12).

Removal of potential PCR-inhibitory substances requires the use of proper nucleic acid extraction methodology to isolate the amplifiable nucleic acids from the inhibitory compounds (13-17). The many extraction kits available for nucleic acid processing provide varying levels of PCR inhibitor mitigation (17). In addition, certain pathogens, like spore-forming bacteria and mycobacteria, require the use of harsher or additional processing steps for the release of their genetic material for PCR-based detection (18). For all of these reasons, the processing and PCR analysis of clinical specimens require significant preliminary testing and evaluation, in addition to inclusion of proper controls. Clinical specimen types considered to be challenging include blood, stool, and tissue; however, any specimen type that will potentially be collected for use with the molecular test must be considered in the assay development phase.

A number of approaches can be used to assess inhibition; among these are the use of target nucleic acid present in specimens (19); the use of another independent test to which the specimen of interest is added, along with a low level of specific target nucleic acid (20); and methods that utilize plasmids or oligonucleotides (12,21,22). The last method allows the possibility of differentiation between sample cross-contamination and positive control contamination (23). In any method, a critical aspect is to use a level of nucleic acid that is appropriate in the inhibition assessment at only a low level (at or near the LOD of the assay). The appropriate level of nucleic acid permits the visualization of inhibition in situations when the target is present at low levels, as can be anticipated for many clinical specimens.

PCR Workflow

In diagnostic laboratories that use PCR-based methods, the workflow needs to be well defined and procedures to prevent contamination must be strictly followed. Such procedures will minimize the potential for sample cross-contamination, especially from positive control material or post-PCR-amplication products. In order to prevent contamination, separate areas should be maintained for each of the individual steps of the PCR workflow. A three- or four-step unidirectional PCR layout can be integral to the protocol. The first step or area is the PCR clean area, or PCR reagent setup area; it is used for PCR mastermix reaction setup only. This area is only entered once and cannot be re-entered by the same laboratorian after a PCR assay is completed. The second step or area is a space in which the specimen nucleic acid is added to the PCR mastermix. It is advisable to have another area (if space allows) where positive control material can be added; this space is physically separated from the space where patient specimens are added to avoid cross-contamination of unknown samples with-known positive material. If spatial separation is not possible, unknown patient specimens should always be set up and the tubes containing them capped before tubes containing positive control material are opened. Each of these areas should be cleaned after each setup. The third or fourth step or area (depending on space constraints) is the area in which the PCR instruments are located. This should be physically separated from the previously described areas. Once this area has been utilized by an analyst, he or she cannot re-enter the other PCR areas. In addition, the laboratory should also have a

general area used for nucleic acid extraction and specimen processing that is separate from the areas outlined for PCR workflow. All of these areas should contain dedicated equipment (e.g., pipettes, racks, pens, calculators, and personal protective equipment, such as lab coats and gloves). Worksheets utilized by the analyst should only be brought into the PCR clean area if they are newly printed or photocopied. Adherence to these procedures helps to minimize the potential for contamination.

The use of real-time PCR mitigates much of the risk resulting from post-amplification analysis, because the PCR reaction tubes containing target amplicons are not opened. The other methods, including PCR and nucleic acid sequence analysis, are at risk for potential contamination. Surveillance of the separate areas within the PCR workflow should be considered. This process may include a schedule of environmental sampling followed by testing the environmental samples with the PCR assays used in the laboratory to assess any contamination that may exist.

Assay Validation

The process for the validation of new molecular assays varies across institutions. In New York state, the CLEP has outlined the necessary steps that need to be accomplished for any laboratory seeking approval to offer a molecular test that has not been cleared by the FDA. After successful completion of the assay development phase, as detailed above, an assay must undergo a thorough validation evaluation prior to any implementation. For the validation phase, an SOP (standard operating procedure), should be defined and accepted. The SOP must include predefined interpretation and reporting algorithms. Validation studies typically include analyses that confirm the assay's accuracy, precision, sensitivity, and specificity. Validation data summaries are compiled during the process so that if an assay passes and becomes validated, all phases of the development and validation are documented in a single report.

Our approach to the validation of a molecular assay for use on clinical specimens begins with a sensitivity study in the clinical specimen matrix or matrices. Similar to the analytical sensitivity evaluation, the evaluation of clinical or diagnostic sensitivity should use a dilution series consisting of 8 to10 log concentrations of the whole target organism (when possible) assayed in triplicate. Determination of sensitivity in this manner will also produce the necessary accuracy, reportable range, and linearity data for assessment of the performance of the assay in a matrix. These studies should be performed for each potential matrix for which the developing laboratory will use the assay.

The next part of the validation process should assess inter- and intra-assay reproducibility. These studies are used to determine the accuracy and precision of the assay. We typically expect assay accuracy to exceed 95%, a level that limits the false-negative rate to <5%. For intra-assay reproducibility studies, the assay being validated is performed in triplicate on three specimens. The coefficient of variance (CV) must be <5% in order for the assay to be acceptable for use with clinical specimens in our laboratory. We also test these specimens on three separate days to satisfy inter-assay requirements. Alternatively, the positive control used in the assay can also be examined over a longer period of time, e.g., 15 days, for determination inter-assay reproducibility. Again, the CV must be <5%. Besides serving as a standard, for mastermix quality and PCR cycling parameters, the use of a positive control allows the baseline and threshold settings to be adjusted so that the analysis is effectively normalized from run to run, thereby minimizing inter-operator and/or inter-machine variability.

The final step in the assay validation process is to assess the robustness of the assay in a specimen matrix that will eventually be used for diagnosis (e.g., blood or serum for bacteremia diagnosis). This validation process should include a study performed in a blinded, randomized fashion. The blinded validation study needs to include at least 30 positive or seeded specimens and at least 10 negative specimens. It is important that the specimens include those having a target organism concentration that is at or near the LOD of the assay. The sensitivity assays and initial characterization of the assay that were performed in the development phase can be used as guidelines during the ensuing validation phase. Ideally, the blinded validation study should utilize actual specimens known to contain the target organism. For target pathogens, such as *Escherichia coli* or *Salmonella,* stool specimens containing the organisms are readily available in the clinical laboratory. However, other pathogens that may be detected in specimen types, such as CSF or biopsy specimens, may not be readily available in the validating laboratory. If that is the case, the relevant specimen matrices should be seeded with the target organism. The assay to be validated must be tested for each specimen matrix to which it will be applied, because differences in physical properties, inhibitors present, and pathogen loads exist among specimen types. An assay designed to detect *Legionella* spp. in urine can have very different parameters from the same assay designed for sputum specimens. While a comparison of the validated assay against another FDA-approved or validated molecular assay can be performed, the validation study should ideally be compared to a gold standard assay, such as culture-based testing. In addition, since most real-time PCR assays are performed on extracted nucleic acid, the validation study must take into account the potential loss of nucleic acid during the extraction process; thus, the validation study is normally performed using an initial concentration that is slightly higher than the assay's LOD.

Quality Assurance and Quality Control

Of paramount importance to the validation of any laboratory assay is the existence of a solid QA/QC program. Adherence to such a program is crucial for any assay and indeed for the overall functioning of a laboratory, and it is particularly critical for PCR assays. For a PCR assay, the slightest deviation from an SOP can produce amplification of an unintended target that did not originate from the sample or specimen itself. In the present context, lack of a high-quality QA program can lead to serious problems in efforts to validate a new assay.

The laboratory QA program should include regularly scheduled cleaning of areas in which PCR is performed, to ensure there is no amplicon contamination. Some laboratories use monitoring of PCR areas to avoid any spread of contamination. This practice should

0196-4399/00 (see frontmatter)

be put into place whenever a problem has been uncovered through the use of controls run in the assays. A good policy is to perform monitoring on a routine schedule. Such monitoring can include swiping of the areas and performance of PCR assays directly on these swipe materials to ensure that no contamination is present. Also, laboratory areas in which target nucleic acid is used, notably the extraction area and the area where the positive control is added, should always be cleaned before and after every use. There are commercially available products that remove residual nucleic acid from surfaces; the use of bleach to destroy nucleic acid is highly recommended as an additional cleaning or decontamination method.

Not only are the reagents and supplies part of the QA/QC program, but the equipment that is used for PCR assays, whether it is a conventional PCR machine or a real-time instrument, should have routine maintenance performed, to ensure the integrity of each run. The machine should be included in a monthly surveillance program that monitors any amplicon contamination. Also, routine background checks and dye calibrations should be performed as recommended by the real-time PCR equipment manufacturer. Documentation of such maintenance should be kept on record in the laboratory. Often, the maintenance log serves as a valuable troubleshooting reference when machine malfunctions occur.

Another critical component of the QA program is the competency assessment of the analysts who perform the assays. During the validation process for each new assay, multiple analysts should verify the robustness and precision of the assay. Proficiency testing programs should be set up and documented in order to maintain the staff level of competence. For some PCR assays, external PCR proficiency testing programs are available; if not, an internal proficiency testing program should be initiated and documented. Maintenance of competency among the analysts who perform the assays will help to ensure reliable, high-confidence testing.

Nucleic Acid Sequence Analysis

The development and validation of diagnostic assays utilizing nucleic acid sequence analysis follow the criteria described above for PCR assays, but some additional factors are associated with them. These sequencing assays all entail an initial PCR amplification reaction, followed by a second nucleic acid sequencing assay. The two main means of generating nucleic acid sequence data are Sanger sequencing analysis and pyrosequencing analysis. These assays can be further divided into broad-range sequence analysis and target-specific sequence analysis; they each necessitate some different development and validation criteria for evaluation and documentation.

Broad-range nucleic acid sequence analysis assays, including targeting ribosomal RNA and housekeeping genes, have been in use for decades (24), first in studies aiming to elucidate the evolution of bacteria, and later as a diagnostic tool for bacterial identification (25). The CLEP requirements for the approval of a test of this type include the development of a detailed SOP. The SOP needs to incorporate a complete description of controls that are used in the assay, including information pertaining to the type and source of *Taq* polymerase used in amplification reactions. The SOP should include information pertaining to the expected size of the amplicon and the acceptable percentage of the sequence length that will be utilized for sequence analysis. A description of the process used to analyze sequence data, including review of electropherograms, the process for editing the sequence, the expected read length of the sequence, and the protocol used for interpreting ambiguous base calls, must also be provided.

Additionally, a detailed description of the reporting algorithm, including the range of percent identity matches that will be reported (i.e., reporting to the genus level or to the species level and the percent identity that will be used in the assay to denote inability to identify the target organism) is required. Specifics, such as how results will be reported when the sequence analysis identifies multiple organisms with 100% identity, must be addressed. Also, a statement should be made regarding taxonomic nomenclature as to whether common names or taxonomically correct names will be used. Finally, information detailing how the local/proprietary database updates will be performed and how records are maintained should be included. The use of a single public database is not recommended for broad-range sequence-based identification assays; identification should be based on the use of two or more databases. If a commercial sequence database that contains gold-standard validated contents is utilized, use of this single source is acceptable. However, such a database may only contain a subset of the microorganisms that have the potential to be present in clinical specimens. If an in-house-developed database is used in the absence of a commercial database, organism identifications must be verified using strains that have been typed using conventional gold-standard microbiological methods. Alternatively, comparisons between the in-house database and a commercial database can be performed, to validate the in-house database.

The requirements for validation of a sequence-based analysis assay are similar to those for the PCR assays described above. However, data demonstrating specificity and the LOD for use on isolates are not required. If sequenced-based assays will be performed directly on nucleic acid extracted from patient specimens, results from such samples should be submitted. Validation data for at least 30 representative organisms that the laboratory routinely identifies should be submitted. Data need to include a comparison between sequence-based identification and identification by a reference method, such as phenotypic identification. Additionally, a validation data table using the format described below, should be submitted. For each sample analyzed, data should include the top match(es) and the next closest match. If the input sequence and the output sequence length are not the same, an explanation of the difference should be provided. For approval of this type of assay through CLEP, participation in a specialized assessment administered by the Wadsworth Center is required.

Target-specific sequence analysis assays often incorporate PCR of a target gene that was selected because the gene provides some information that characterizes a pathogen beyond its presence or absence. Often these sequence data allow further typing of the pathogen (26), assessment of antibiotic resistance markers (27), or the detection of other

markers of virulence (22,27). Such assays are often paired with other real-time PCR assays that detect the pathogen. Development and validation of the assays are similar to the process described above for PCR. Additionally, all known subtypes or other known sequence variations (mutations, insertions, and deletions) of the pathogen should be included in the specificity testing if feasible. Also, gold-standard methods should be used in the validation when possible, including serotyping and classical antibiotic susceptibility testing.

Acknowledgements

We thank Nick Cirino and Ron Limberger for their work at the Wadsworth Center that is reflected in this manuscript. We acknowledge Adriana Verschoor for critical reading of the manuscript. We thank the committee established to prepare standards and guidelines for the nucleic acid amplification testing for infectious agents at the Wadsworth Center.

References

1. Whelen, A.C. and D.H. Persing. 1996. The role of nucleic acid amplification and detection in the clinical microbiology laboratory. Annu. Rev. Microbiol. 50:349-373.
2. Procop, G.W. 2007. Molecular diagnostics for the detection and characterization of microbial pathogens. Clin. Infect. Dis. 45(Suppl.2):S99-S111.
3. Mothershed, E.A. and A.M. Whitney. 2006. Nucleic acid-based methods for the detection of bacterial pathogens: present and future considerations for the clinical laboratory. Clin. Chim. Acta 363:206-220.
4. Espy, M.J. et al. 2006. Real-time PCR in clinical microbiology: applications for routine laboratory testing. Clin. Microbiol. Rev. 19:165-256.
5. Tondella, M.L. et al. 2002. Development and evaluation of real-time PCR-based fluorescence assays for detection of *Chlamydia pneumoniae*. J. Clin. Microbiol. 40:575-583.
6. Breitkopf, C. et al. 2005. Impact of a molecular approach to improve the microbiological diagnosis of infective heart valve endocarditis. Circulation 111:1415-1421.
7. Riffelmann, M. et al. 2005. Nucleic Acid amplification tests for diagnosis of *Bordetella* infections. J. Clin. Microbiol. 43:4925-4929.
8. Pandit, L. et al. 2005. Diagnosis of partially treated culture-negative bacterial meningitis using 16S rRNA universal primers and restriction endonuclease digestion. J. Med. Microbiol. 54:539-542.
9. Peruski, L.F., Jr. and A.H. Peruski. 2003. Rapid diagnostic assays in the genomic biology era: detection and identification of infectious disease and biological weapon agents. Biotechniques 35:840-846.
10. Fout, G.S. et al. 2003. A multiplex reverse transcription-PCR method for detection of human enteric viruses in groundwater. Appl. Environ. Microbiol. 69:3158-3164.
11. Halse, T.A., K.A. Musser, and R.J. Limberger. 2006. A multiplexed real-time PCR assay for rapid detection of *Chlamydia trachomatis* and identification of serovar L-2, the major cause of lymphogranuloma venereum in New York. Mol. Cell Probes 20:290-297.
12. Nazarian, E.J. et al. 2008. Design and implementation of a protocol for the detection of *Legionella* in clinical and environmental samples. Diagn. Microbiol. Infect. Dis. 62:125-132.
13. de Franchis, R. et al. 1988. A potent inhibitor of Taq polymerase copurifies with human genomic DNA. Nucleic Acids Res. 16:10355.
14. Tsai, Y.L. and B.H. Olson. 1992. Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction. Appl. Environ. Microbiol. 58:2292-2295.
15. Kuske, C.R. et al. 1998. Small-scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. Appl. Environ. Microbiol. 64:2463-2472.
16. Sutlovic, D. et al. 2005. Taq polymerase reverses inhibition of quantitative real time polymerase chain reaction by humic acid. Croat. Med. J. 46:556-562.
17. Filion, M., M. St-Arnaud, and S.H. Jabaji-Hare. 2003. Direct quantification of fungal DNA from soil substrate using real-time PCR. J. Microbiol. Methods 53:67-76.
18. Aldous, W.K. et al. 2005. Comparison of six methods of extracting *Mycobacterium tuberculosis* DNA from processed sputum for testing by quantitative real-time PCR. J. Clin. Microbiol. 43:2471-2473.
19. Brinkman, J.A. et al. 2004. Optimization of PCR based detection of human papillomavirus DNA from urine specimens. J. Clin. Virol. 29:230-240.
20. Oikarinen, S. et al. 2009. PCR inhibition in stool samples in relation to age of infants. J. Clin. Virol. 44:211-214.
21. Hadjinicolaou, A.V. et al. 2009. Molecular beacon-based real-time PCR detection of primary isolates of *Salmonella typhimurium* and *Salmonella enteritidis* in environmental and clinical samples. BMC Microbiol. 9:97.
22. Wroblewski, D. et al. 2009. Rapid molecular characterization of *Clostridium difficile* and assessment of the population of *C. difficile* in stool specimens. J. Clin. Microbiol. 47: 2142-2148.
23. Sachadyn, P. and J. Kur. 1998. The construction and use of a PCR internal control. Mol. Cell Probes 12:259-262.
24. Clarridge, J.E. 2004. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. Clin. Microbiol. Rev. 17:840-862.
25. Woo, P.C. et al. 2003. Usefulness of the MicroSeq 500 16S ribosomal DNA-based bacterial identification system for identification of clinically significant bacterial isolates with ambiguous biochemical profiles. J. Clin. Microbiol. 41:1996-2001.
26. Pai, R., J. Limor, and B. Beall. 2005. Use of pyrosequencing to differentiate *Streptococcus pneumoniae* serotypes 6A and 6B. J. Clin. Microbiol. 43:4820-4822.
27. Somoskovi, A. et al. 2003.Use of molecular methods to identify the *Mycobacterium tuberculosis* complex (MTBC) and other mycobacterial species and to detect rifampin resistance in MTBC isolates following growth detection with the BACTEC MGIT 960 system. J. Clin. Microbiol. 41:2822-2826.

  0196-4399/00 (see frontmatter)