Journal of Aquatic Animal Health

Quantitative Polymerase Chain Reaction (PCR) for Detection of Aquatic Animal Pathogens in a Diagnostic Laboratory Setting

Maureen K. Purcell a, Rodman G. Getchell b, Carol A. McClure c e & Kyle A. Garver d

a U.S. Geological Survey, Western Fisheries Research Center, 6505 North East 65th Street, Seattle, Washington, 98115, USA
b Department of Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Upper Tower Road, Ithaca, New York, 14853, USA
c Center for Aquatic Health Sciences, University of Prince Edward Island, 550 University Avenue, Charlottetown, Prince Edward Island C1A 4P3, Canada
d Fisheries and Oceans Canada, Aquatic Animal Health, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, British Columbia V9T 6N7, Canada
e AquaEpi Research, 300 Heartz Road, Charlottetown, Prince Edward Island C1A 7J7, Canada

Available online: 10 Oct 2011

To cite this article: Maureen K. Purcell, Rodman G. Getchell, Carol A. McClure & Kyle A. Garver (2011): Quantitative Polymerase Chain Reaction (PCR) for Detection of Aquatic Animal Pathogens in a Diagnostic Laboratory Setting, Journal of Aquatic Animal Health, 23:3, 148-161

To link to this article: http://dx.doi.org/10.1080/08997659.2011.620217

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
ARTICLE

Quantitative Polymerase Chain Reaction (PCR) for Detection of Aquatic Animal Pathogens in a Diagnostic Laboratory Setting

Maureen K. Purcell*
U.S. Geological Survey, Western Fisheries Research Center, 6505 North East 65th Street, Seattle, Washington 98115, USA

Rodman G. Getchell
Department of Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Upper Tower Road, Ithaca, New York 14853, USA

Carol A. McClure
Center for Aquatic Health Sciences, University of Prince Edward Island, 550 University Avenue, Charlottetown, Prince Edward Island C1A 4P3, Canada

Kyle A. Garver
Fisheries and Oceans Canada, Aquatic Animal Health, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, British Columbia V9T 6N7, Canada

Abstract

Real-time, or quantitative, polymerase chain reaction (qPCR) is quickly supplanting other molecular methods for detecting the nucleic acids of human and other animal pathogens owing to the speed and robustness of the technology. As the aquatic animal health community moves toward implementing national diagnostic testing schemes, it will need to evaluate how qPCR technology should be employed. This review outlines the basic principles of qPCR technology, considerations for assay development, standards and controls, assay performance, diagnostic validation, implementation in the diagnostic laboratory, and quality assurance and control measures. These factors are fundamental for ensuring the validity of qPCR assay results obtained in the diagnostic laboratory setting.

The popularity of real-time, or quantitative, polymerase chain reaction (qPCR) has increased rapidly over the past decade across many different scientific fields (Walker 2002). Quantitative PCR technology is quickly supplanting other molecular methods for detecting the nucleic acids of human and other animal pathogens. The elimination of post-PCR manipulation of amplified products, high throughput, and the robustness of the technology has contributed to the swift adaptation of qPCR in a diagnostic setting (Gunson et al. 2006).

The past decade has seen numerous publications describing qPCR assays that detect a range of bacterial, viral, and parasitic pathogens of finfish (Getchell and Bowser 2011). The majority of these new assays were developed by research laboratories, which have the freedom to implement new methods to address specific research questions. Different constraints exist in a diagnostic laboratory setting. Diagnostic assays must perform consistently and reliably over time, as they are used by different staff members, who often use different tissue types.
or host species and samples of varying quality (Hoorfar et al. 2004).

Fish health diagnostic centers in North America typically rely on established methods specified by the Blue Book of the American Fisheries Society’s Fish Health Section (AFS–FHS 2010); Fisheries and Ocean Canada’s Manual of Compliance (DFO 2004); or the World Organization for Animal Health’s (OIE) Manual of Diagnostic Tests for Aquatic Animals (OIE 2009). These sources prefer that extensive validation occurs before tests are accepted for diagnostic purposes. The AFS Blue Book (AFS–FHS 2010) distinguishes between tests that can be used for inspections (establishing freedom from infection) versus tests that can be used for surveillance or diagnosis. Inspection assays typically rely on a “gold standard” assay for initial detection (presumptive test) followed by a confirmatory test based on a different diagnostic principle. Inspection results affect movement of fish for natural resource management or trade, as well as lead to possible eradication or control actions. Thus, diagnostic tests used for inspections must have sufficient validation to be legally defensible. To date, molecular-based tests for aquatic animal pathogens have not achieved widespread regulatory acceptance for inspections except when used as a confirmatory method. A primary concern when using PCR-based tests in an inspection setting is that government regulators may adopt a highly precautionary approach to interpreting low-level positive results.

The choice of tests that can be used for surveillance or diagnoses of clinical cases is more flexible, but there can also be regulatory consequences when a true or false positive result indicates the presence of a pathogen in a new geographical range, species, or environment. The increasing use of qPCR for aquatic animal pathogen surveillance has raised concerns, particularly when the results cannot be confirmed by a nonmolecular diagnostic test. In contrast, qPCR is used as a primary screening tool for many terrestrial pathogens (Pestana et al. 2010). Two PCR techniques have been approved by the OIE as prescribed tests for international trade; a qPCR test for infectious bovine rhinotracheitis and a conventional PCR test for bluetongue (OIE 2010). The National Animal Health Laboratory Network (NAHLN) and the Canadian Animal Health Surveillance Network conduct surveillance for important terrestrial animal pathogens. A major goal of NAHLN is rapid diagnosis and qPCR technology has met this need, serving as a primary detection and surveillance tool for avian influenza, Newcastle disease, classical swine fever, swine influenza, African swine fever, foot-and-mouth disease, and rinderpest (www.aphis.usda.gov/animal_health/nahln). The NAHLN has outlined specific testing algorithms for each priority pathogen with precise guidelines on how positive results should be confirmed, interpreted, and communicated. The algorithm changes depending on the pathogen and takes into account factors such as the characteristics of the diagnostic tests and the level of risk represented by the pathogen.

Diagnostic PCR tests operate in fundamentally the same manner whether used for the detection of a terrestrial or an aquatic pathogen. However, the policy and regulatory frameworks that guide interpretation of results is better developed for terrestrial animal pathogens. Furthermore, there is a long history of terrestrial animal culture involving a limited number of species worldwide, in contrast to the culture of aquatic animal species. Many of the concerns regarding the validity of qPCR results for aquatic animal pathogens, as well as the actions that should be taken based on those results, will need to be addressed on a pathogen-by-pathogen basis. However, the first step in addressing these questions for any pathogen is a thorough evaluation of all the variables that may affect the validity of test results. The goal of this review is to neither endorse nor reject the use of qPCR technology for aquatic animal pathogen diagnostics. Rather, our intention is to review the basic scientific principles of this technology, the appropriate controls and standards, and the ideal implementation and validation procedures needed to ensure the validity of test results in the aquatic animal health laboratory.

QUANTITATIVE PCR

Basic Principles

The basic concepts of qPCR have been reviewed extensively elsewhere (Bustin 2004; Kubista et al. 2006). The technological advance that makes qPCR possible is the ability to monitor PCR amplification in “real time” by measuring the fluorescence change during the reaction. During real-time PCR amplification, intercalation of a fluorescent dye or release of a reporter dye from a hydrolysis probe causes increased fluorescence in the reaction. A background threshold line is drawn across the fluorescence plots and a cycle number is assigned to each sample, representing the point at which the sample crossed the threshold. Multiple terms are used by different real-time instrument manufacturers to describe the fractional value at which the sample crosses the threshold including cycle threshold (C\text{T}), crossing point (C\text{p}), and cycle quantification (C\text{Q}) (Bustin et al. 2009); here we will employ the most commonly used term, \text{C}_T. The \text{C}_T value is a continuous number with lower \text{C}_T values indicating higher quantities of pathogen template in the sample. The PCR reaction is stopped after a fixed number of cycles, and if no \text{C}_T value is assigned to a sample, then the sample is considered negative because it has a pathogen copy number below the detection threshold of the assay.

Assay Types

The two most commonly used qPCR technologies are the 5’ nuclease assay (e.g., Taqman assay) or the dye intercalation assay (e.g., SYBR-green) (Bustin 2004). The 5’ nuclease assay uses two primers and an internal hydrolysis probe, dual-labeled with a 5’ fluorescent reporter dye molecule and a 3’ quencher molecule. The dual-labeled probe is cleaved during amplification and the fluorescent reporter is released from close proximity with the quencher, resulting in increased fluorescence of the reporter that is measured by the real-time PCR instrument. The
internal hydrolysis probes can be manufactured by a variety of companies, each having proprietary modifications to facilitate specific needs. For instance Applied Biosystems (ABI; Carlsbad, California) incorporates a minor-groove-binding (MGB) linker that raises the melting temperature of the probe, thereby enabling MGB probes to be shorter in length and have increased specificity. In the intercalating dye assays, fluorescent dyes are incorporated into newly synthesized DNA and fluoresce when bound to double-stranded DNA. Reduced costs are commonly mentioned as an advantage of intercalating dye assays. However, the role of intercalating dye assays in diagnostic pathogen testing has been questioned (Beld et al. 2007). In general, this method has reduced specificity because any amplification product incorporates the dye, and melting temperature analysis cannot match the specificity that is provided by the internal hydrolysis probes (Gunson et al. 2006). Nontarget amplicons may preferentially consume PCR reagents, reducing analytical sensitivity (ASe; Table 1). However, intercalating dye assays may be better for pathogens with high sequence diversity, for which an internal probe design is not possible.

Quantitative PCR assays can be used to detect pathogens with DNA genomes (e.g., DNA viruses, bacteria, and parasites) or RNA genomes (RNA viruses). Assays that target RNA first require enzymatic reverse transcription to convert RNA into complementary DNA (cDNA). Detection of mRNA from pathogens with DNA-based genomes can be done with the rationale that detecting gene expression may be a better proxy for presence of live pathogen (Suzuki and Sakai 2007); however, genes encoding virulence factors may be modulated during the infection cycle. A comparison of DNA- or RNA-based qPCR assays for detecting the parasite Kudoa thyrsites observed higher ASe in the DNA-based assay (Funk et al. 2007).

Semantics

A number of different PCR-based assays are now available and a variety of terms have been used in the literature. The term conventional PCR (cPCR) has emerged to distinguish standard PCR from qPCR. When the starting template is RNA, a reverse transcriptase step is included and the abbreviation “RT” is added to the name (e.g., RT-cPCR). The use of the RT abbreviation to denote “real-time” should be avoided in the literature to prevent confusion. Similarly, nested PCR (nPCR) and RT-nPCR are also distinguished from the conventional PCR assays. Nested PCR employs two pairs of PCR primers per target, which increases both analytical sensitivity and specificity. All of these assays are typically detected by gel-based electrophoresis and provide only a positive or negative result. In contrast, qPCR and RT-qPCR provide quantitative results in the form of C_T or pathogen copy number. At times, the quantitative results of these assays are discarded and results are reported as simply positive or negative; the term real-time PCR (rPCR) or reverse-transcriptase real-time PCR (RT-rPCR) will be used to describe this type of application. For ease of discussion, we will use the general term qPCR throughout this review.

Advantages and Disadvantages of qPCR

A major advantage of qPCR is the elimination of second round nPCR amplification and post-PCR manipulation of amplified products, two sources of contamination in the diagnostic laboratory. These assays are amenable to high throughput, and qPCR assays that use an internal probe may have an extra level of analytical specificity (ASp) (Table 1) over assays that use primers alone. In general, qPCR assays have a large dynamic range, low interassay variation, and high reliability. However, the ASe of qPCR-based assays is not necessarily greater or less than cPCR or nPCR assays, since sensitivity is assay-specific and varies depending on reaction conditions and primer sequences (Bastien et al. 2008). The main disadvantage of qPCR has been the initial start-up and operating costs, but competition has reduced these costs. The other disadvantages of qPCR technology are the same as any PCR-based diagnostic test, including (1) there is a possibility of failure to detect pathogen template owing to genetic variation at primer, probe, or both sites, (2) that PCR inhibitors are present in samples, (3) there is a high risk of contamination, (4) there is no indication of pathogen viability, (5) there is a confirmed presence of nucleic acid but no infection, and (6) that only a small proportion of the tissue is examined per reaction (Millar et al. 2002; Rådström et al. 2004).

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical sensitivity (ASe)</td>
<td>The minimum number of copies reliably detected by the assay.</td>
</tr>
<tr>
<td>Diagnostic sensitivity (DSe)</td>
<td>The proportion of animals that test positive given that the animal is infected (true positives).</td>
</tr>
<tr>
<td>Analytical specificity (ASp)</td>
<td>The degree to which the assay does not detect (amplify) other pathogens.</td>
</tr>
<tr>
<td>Diagnostic specificity (DSp)</td>
<td>The proportion of animals that test negative given that the animal is not infected (true negatives).</td>
</tr>
<tr>
<td>Predictive value (PV)</td>
<td>The probability that the test result reflects the true disease status.</td>
</tr>
<tr>
<td>Limit of detection (LOD)</td>
<td>Another term to describe analytical sensitivity.</td>
</tr>
<tr>
<td>Repeatability</td>
<td>Agreement between sample replicates, both within an assay run and between independent assay runs, when tested by the same laboratory.</td>
</tr>
<tr>
<td>Reproducibility</td>
<td>Agreement among test results when the same sample is tested by different laboratories.</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>Reproducibility of an assay using different reagent brands or batches and different equipment.</td>
</tr>
</tbody>
</table>
Most laboratories with existing molecular diagnostic capabilities will find few hurdles in incorporating qPCR technology into their program (Beld et al. 2007). Most instrumentation manufacturers have good technical support for operation and data analysis, and can assist in custom assay development. Laboratories that have molecular biology expertise will be best able to troubleshoot start-up problems.

**INTRODUCTION TO ANALYTICAL AND DIAGNOSTIC VALIDATION**

Validation is a multistaged process that encompasses assay development, assay optimization, analytical performance at the bench-top scale, and diagnostic performance to ultimately establish the fitness of a new diagnostic assay under the conditions in which it will be employed. Well-established guidelines for validating new diagnostic methods are available from the OIE, including specific information on validating PCR-based tests (OIE 2009). An overview of the validation pathway is provided in Figure 1. Common terms used in analytical and diagnostic validation are defined in Table 1. Despite the importance of analytical and diagnostic validation, few diagnostic tests for aquatic animal pathogens have undergone complete validation as outlined by the OIE. The lack of validated tests for aquatic pathogens is probably due to the logistical difficulties of designing and performing validation experiments and the limited funding available for this area of research. Furthermore, obtaining sufficient field samples for diagnostic validation can be difficult if the pathogen is not endemic or at low prevalence. In these cases, provisional recognition of assays can be granted based on thorough analytical (bench-top) validation (OIE 2009).

**ASSAY DEVELOPMENT**

**Determining Fitness for Purpose**

An assay can only be validated for its intended purpose, so determining how the assay will be used should occur before assay development begins (OIE 2009). The OIE defines the most common purposes for a diagnostic test as (1) inspections that demonstrate freedom from infection in a population, (2) inspections that demonstrate freedom from infection in an individual, (3) eradication of the disease agent from a population, (4) diagnoses of clinical disease cases or use as a confirmatory test, and (5) surveillance to estimate prevalence in a population.

**Designing the Assay**

Proper development of a qPCR assay involves considering all the factors that affect its performance at an early stage in the development, such as target sequence, primer and probe design, reaction conditions, sample collection, storage conditions, and extraction procedures. There are an increasing number of software options for designing qPCR primers and probes, including commercially and freely available programs. Most qPCR assays are designed to work with a universal set of cycling conditions and most researchers rely on commercially available assay reagents, typically referred to as master mixes. Optimization of assay conditions is usually limited to adjusting primer and probe concentrations and testing different master mix brands. Adjusting thermocycling conditions is less common (Gunson et al. 2006). Often, re-designing the primer and probe set is the best solution for improving an assay that performs poorly.

**Choice of Target Genes for qPCR Development**

The choice of target genes is the most critical element in establishing a sensitive and specific qPCR assay. A good target gene choice requires that one knows (1) the extent of sequence variation within the target primer and probe sites, (2) whether the target gene is single or multi-copy, and (3) the extent of similarities with nontarget sequences from other pathogens or sources. Ribosomal DNA genes (e.g., small subunit rDNA) are a common choice for qPCR assay development (Corbeil...
Pathogen Diversity and Diagnostic Challenges

Pathogen diversity may affect assay efficiency or even lead to false negative results (Whiley et al. 2008). For example, dye-intercalation qPCR assays based on both the N and G genes of infectious hematopoietic necrosis virus (IHNV) were unable to detect the virus as effectively as standard cell culture (Dhar et al. 2008); the authors suggest that high level of genetic heterogeneity among field isolates was the cause. Sequence diversity can also have subtle effects reducing the accuracy of quantification when either the primer or probes are mismatched, ultimately reducing the sensitivity of the assay (Whiley et al. 2008). For many fish pathogens, the extent of sequence diversity among isolates of a given taxon is not known. Some researchers have used a multiplex PCR assay that targets separate genes or gene regions to compensate for high pathogen diversity (Whiley et al. 2003; Herrmann et al. 2004). The use of live-attenuated or killed vaccines in a given population can also lead to positive results and methodologies to distinguish the vaccine product from pathogen may be required. A single validated assay must be constantly reviewed to ensure that it continues to detect circulating strains and be linked to current taxonomy or regulatory definitions. It is advisable to regularly confirm the target primer and probe sequences to ensure that they remain stable.

Initial Characterization of Assay Dynamics

Optimization of the qPCR and determination of its analytical characteristics is necessary before proceeding to the validation stages (Figure 1). The linear range of an assay determines the degree to which the concentration of the target can vary and still be accurately quantified; the typical linear range of a qPCR assay is 5–7 log_{10} per reaction. Correlation coefficients (r^2 values) derived from the linear regression of serially diluted standards (standard curve) can be used to evaluate whether the assay is optimized. A typical qPCR standard curve would have an r^2 value greater than 0.95.

There are three stages of a PCR reaction: the stochastic, exponential, and plateau phases; early exponential amplification is the best stage for accurate quantification of a sample. An amplification efficiency (E) of 2 (100% efficiency) indicates that the reaction is optimized and each DNA template copy is doubled per cycle. The mean efficiency from several runs is a measure of an assay’s repeatability. An assay that possesses 100% amplification efficiency will have a standard curve slope of −3.32. Efficiency can affect the accuracy and precision of quantification (Bustin 2004). Slopes that vary between −3.1 and −3.6 are typically considered acceptable, but there is little empirical support for these limits (Raymaekers et al. 2009). Poor efficiency (outside this slope range) may be a sign of nonspecific amplification, primer–dimer formation, presence of inhibitors, suboptimal annealing temperature, or poorly designed primers (Bustin and Nolan 2004).

STANDARDS, CONTROLS, AND NORMALIZATION

Standards

Standards are samples of known concentration or copy number that are typically used to construct a standard curve for qPCR applications. Absolute standards are desired for diagnostic purposes. A good standard is stable, mimics the biological target, can be accurately quantified, can be reliably reproduced in new batches, and is not a high contamination risk for the laboratory. Typical standards used for DNA targets...
are plasmid DNA containing the target region, single-stranded oligodeoxynucleotides that represent the entire target, or DNA from a quantified pathogen culture. Assays targeting RNA can use similar standards as assays targeting DNA, as well as RNA transcribed in vitro from plasmids encoding the target gene of interest. Plasmid DNA and oligodeoxynucleotide standards are highly stable, easy to prepare, can be distributed in large quantities, are noninfectious, and can be accurately quantified. However, oligodeoxynucleotides can pose a contamination risk to the laboratory and should be ordered on a separate date from the primer and probes or from a separate company. Dilutions of oligodeoxynucleotide and plasmid standards should be prepared in an area removed from locations where PCR reagents or samples are handled. Some debate exists whether circular plasmid DNA should be cleaved by restriction enzymes because supercoiled circular molecules interact differently than linear molecules (Beld et al. 2007). In our experience, it is generally not necessary to digest plasmids but we have encountered plasmid preparations that do not dilute in a linear fashion; in these latter cases, restriction digestion of the plasmid typically resolved the issue. The use of in vitro transcribed RNA as a standard for RT-qPCR has the advantage that the efficiency of the reverse transcription step is assessed, along with the amplification for each set of reactions. However, in vitro transcribed RNA standards are less stable and do not withstand freeze–thaw cycles well. Quantified pathogen cultures have been used as standards in a number of studies, such as parasite spores or bacteria stained with fluorescent antibodies (Chase et al. 2006; Arsan et al. 2007). Viral RNA extracted from quantified culture may be used as a standard, but there can be a several log difference between infectious units and genome copies (Purcell et al. 2006). Choosing the standard is dependent on the goal and existing capabilities in the laboratory. There can be advantages to using at least two different standards. For instance, researchers may be interested in converting $C_T$ results into a scale relevant to the purified pathogen (e.g., parasite spore), but genome copy number may vary in the pathogen; this copy number variation can be assessed with a synthetic standard (e.g., plasmid DNA) (Arsan et al. 2007).

In some applications, results are simply reported as raw $C_T$ units; in the case of rPCR, results are reported as only positive or negative. The absolute $C_T$ of a sample can vary among qPCR platforms and is dependent on analysis parameters (e.g., threshold levels). Thus, standardization within or across laboratories requires consistent analysis parameters or the development of a reference standard for distribution. If standard curves are not used, a reference sample should be included on each plate to monitor interassay variation (Table 2).

### Controls

Controls ensure the validity of positive or negative results (Table 2). These controls should be used to monitor the entire procedure from sample extraction through qPCR detection. Similar to standards, control materials should behave like the real sample, be readily available, be stable over time, and vary little in concentration from aliquot to aliquot or vial to vial. A positive processing control is important for assessing the consistency of extraction efficiency. These processing controls are known positive or spiked samples and should optimally be near the lower range of assay detection, ideally no greater than 30 $C_T$ (Fryer et al. 2008). A negative processing control is important for detecting false positives due to contamination during the extraction process. During the PCR set-up, a no template control (e.g., water control) is used to detect false positives due to contamination. Additionally, each assay should include a minimum of one reference standard sample as described above. Most commercial qPCR master mix formulations include a dye that is used as a reference signal to normalize variation in master mix concentrations or volumes. These master mix dyes are referred to as passive references in contrast to active references, which are endogenous or exogenous controls that have their own set of primers and probe (as discussed below).

The presence of PCR inhibitors or poor nucleic acid quality may lead to false negative results. Inhibition can be detected by including an internal positive control (IPC) in each PCR reaction. The IPC assay targets an exogenous artificial template that is unrelated to the specific pathogen target and is often referred to as the active reference. Separate IPC primer and probes are spiked into each reaction along with the artificial template. Since the IPC is at a known concentration, inhibition can be detected by decreased amplification. The addition of an IPC can reduce the sensitivity of a specific assay because it creates a multiplex reaction. Thus, it is preferable to use commercial qPCR master mixes suitable for multiplex reactions, and the effect of the IPC on the $A_S$ of the specific assay must be empirically evaluated. Several studies have used IPCs when detecting important fish viruses by RT-qPCR (Matejusova et al. 2008; Snow et al. 2009). Amplification of an endogenous reference gene, often called a housekeeping gene, can also be used as an active reference for detecting false negative results owing to poor RNA quality or inefficient cDNA synthesis (Matejusova et al. 2010). Since endogenous reference genes are highly expressed and the exact copy number is typically unknown, IPCs have greater sensitivity for detecting the presence of PCR inhibitors.

### Normalization

Normalization with exogenous (external) variables or endogenous (internal) reference genes is used to standardize the reporting of pathogen load and to correct for variation in template quantity. Examples of exogenous normalizing variables are tissue weight or volume extracted or concentration of template. One advantage of normalizing to tissue weight is that this unit is typically used for other diagnostic assays (e.g., colony forming units per gram tissue). The disadvantage of tissue weight is that extraction efficiency may vary among samples and it is not a measure of template
TABLE 2. Recommendations for the use of controls, standards, and normalization.

<table>
<thead>
<tr>
<th>Category</th>
<th>Type</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standards</td>
<td>Standard curve</td>
<td>Always recommended when quantitative results are desired.</td>
</tr>
<tr>
<td></td>
<td>Reference sample</td>
<td>Always recommended to include a minimum of one positive reference sample per assay run.</td>
</tr>
<tr>
<td>Controls</td>
<td>Positive processing sample</td>
<td>Always recommended to verify nucleic acid extraction effectiveness.</td>
</tr>
<tr>
<td></td>
<td>Negative processing sample</td>
<td>Always recommended to detect contamination during the extraction process.</td>
</tr>
<tr>
<td></td>
<td>No template control for PCR reaction</td>
<td>Always recommended on every assay run to detect contamination in PCR reagents.</td>
</tr>
<tr>
<td></td>
<td>Internal positive control (IPC)</td>
<td>Good practice for detecting false negative results if IPC does not interfere with analytical sensitivity.</td>
</tr>
<tr>
<td>Normalization</td>
<td>Amplification of endogenous gene</td>
<td>Good practice for ensuring nucleic acid integrity and troubleshooting.</td>
</tr>
<tr>
<td></td>
<td>Exogenous normalization variables</td>
<td>Good practice to track tissue weight and nucleic acid concentration; normalizing copy number to these variables is dependent on goals.</td>
</tr>
<tr>
<td></td>
<td>Normalization to endogenous gene</td>
<td>Not recommended to normalize copy number to endogenous gene expression in field samples.</td>
</tr>
</tbody>
</table>

quantity or quality. When tissue weight is standardized before extraction, there is typically a good correlation between tissue weight and nucleic acid concentration (M. K. Purcell, personal observation). Alternatively, a number of studies report pathogen copies per microgram of DNA or RNA. Spectrophotometric analysis of nucleic acid is a good practice for monitoring extraction efficiency and template quality, but contaminating nucleic acids in samples (e.g., contaminating genomic DNA in RNA preparations) can reduce the accuracy of this method when used for normalization. Standardization of pathogen copy number to exogenous variables or no standardization is most common among published studies. Since pathogens levels typically differ on a logarithmic scale, small errors in normalization often have a negligible effect on the final values.

Amplification of endogenous reference genes is a useful control for verifying template quality and can be performed for either DNA- or RNA-based assays (as discussed above). Normalization to an endogenous reference gene is an accepted practice when measuring gene expression, where it is assumed that the reference gene expression is an indicator of template quantity. However, mRNA expression of endogenous reference genes can vary considerably among individual animals, even in laboratory-based animal experiments, and this variation may be much greater in animals collected in the field. Thus, while amplification of an endogenous host gene improves confidence that nucleic acid extraction was successfully performed (e.g., as a control), the results may not be appropriate for “normalizing” pathogen load (Table 2). Additionally, normalization of pathogen load to a reference gene usually results in fractional units since reference genes tend to be highly expressed, and interpretation of these fractional units is not intuitive.

ANALYTICAL VALIDATION

Analytical Sensitivity (ASe)

It is important to determine the ASe of a new assay, which is defined as the minimum number of copies reliably detected by the assay (Table 1). Every assay has a limit of detection (LOD) that is defined empirically for each new assay. Theoretically, at least one copy of the pathogen must be in a qPCR reaction for amplification to occur, but it is unlikely that this copy number will be reliably detected. Extending the number of qPCR cycles in an attempt to increase the LOD may result in degradation of the probe, leading to fluorescence increases that are interpreted as a positive amplification. Practical ASe of an assay can be limited by the volume or tissue quantity analyzed in each sample. Pathogens that create highly focal infections or have differing tissue tropisms may not be detected if only a small tissue sample is analyzed by PCR. Samples at or below the LOD typically have low repeatability and may require a high number of technical replicates to generate confident results.

Choosing a $C_T$ positive–negative cut-off value may be required for certain diagnostic calculations or if results are to be interpreted as simply positive or negative (rPCR). If absolute standards are available, the cut-off value can be assigned as 1 copy of pathogen molecule (typically at or near 40 $C_T$). Alternatively, the cut-off can be adjusted to the minimum copy number or $C_T$ that is reliable and repeatable across technical replicates (e.g., a $C_T$ at which 50% of the technical replicates can be detected) (OIE 2009). Choosing a cut-off is controversial because any detectable PCR amplification is often considered positive even if the result is not repeatable among technical replicates. Whether to consider these samples positive depends on the pathogen and the implications of the finding. This question
should be addressed when designing the overall testing scheme for a pathogen (screening and confirmatory testing) and the testing scheme should not be applied until all outcome scenarios are considered (Caraguel et al. 2011).

Analytical Specificity (ASp)

Analytical specificity (ASp) is defined as the degree to which the assay detects the target template but not nontarget templates. A number of studies rely solely on in silico sequence analysis, but specificity is ideally tested empirically (Bustin et al. 2009). Obtaining a diverse set of phylogenetically or phenotypically related pathogens for specificity testing can be difficult. Certain pathogens are available from type culture collections, but pathogen propagation may require specialized facilities, reagents, or regulatory permits. Ideally, one would also test isolates of the target pathogen from different geographical regions or isolates representing different defined genotypes or strains. If clinical material is not available, this should be stated in the publication. At a minimum, validation should be done with samples derived from the target host species and geographic region, and validation must be repeated to expand the assays to new species or regions.

Analytical Repeatability

Repeatability is defined as agreement between replicates, both within an assay run and between independent assay runs, and generally refers to results generated within a single laboratory (OIE 2009). In general, qPCR assays provide highly repeatable results but repeatability declines near the LOD. When copy number in the starting reaction is low, the distribution of template copies follows the Poisson distribution, causing some replicates to amplify while others do not. Furthermore, a PCR efficiency of less than 100% leads to small statistical errors that affect the accuracy of quantification. These statistical errors have very little effect when the starting copy number is high, but when the starting copy number is less than $10^2$, these statistical errors affect both the accuracy and precision of quantification (Bustin 2004). The OIE manual provides guidance on the number of replicates that should be performed to assess variations within and between runs; the coefficient of variation (CV = SD/mean) is a useful metric for comparing within- and between-run variations.

DIAGNOSTIC VALIDATION

Diagnostic Sensitivity (DSe) and Diagnostic Specificity (DSP) Overview

Diagnostic performance of an assay is measured by diagnostic sensitivity (DSe; the conditional probability that a test result will detect disease in an infected animal) and diagnostic specificity (DSP; the conditional probability that a normal test result will exclude the disease in an uninfected animal) (Table 1) (McClure et al. 2005). The DSe, DSP, and disease prevalence in the population are used to calculate the predictive values of a test. The test’s ability to accurately predict the disease status is the test’s positive and negative predictive values. Ultimately, inference is made on whether the test result is positive or negative. Therefore, accurate estimates of DSe and DSP are highly important (OIE 2009). Disease in the context of these metrics refers to exposure to or infection with a pathogen, and does not necessarily imply clinical disease.

Evaluation of DSe and DSP requires knowledge of the true infection status determined by a perfect reference test (gold standard). When infection status cannot be indisputably determined owing to the lack of a perfect reference assay, only an estimate of DSe and DSP is obtained for the new assay. Poor DSe and DSP estimates can be mitigated if the reference assay is well-validated and the biases of the tests are quantified. For aquatic pathogens, there are very few tests that are validated to this level and often the current reference assays have lower sensitivity than do the new assay being tested (e.g., a new qPCR test can be more sensitive than traditional culture methods). In these cases, DSe and DSP can be determined by means of other statistical methods or by creating samples of known infection status in the laboratory. Here we will review the basic concepts and assumptions of various methods to estimate DSe and DSP.

Selecting the Reference Population

Estimation of DSe and DSP of any new assay should be made with known reference populations. Ideally, the reference population should have the same or similar features as the eventual target animal population (Greiner and Gardner 2000). Different variables of the target populations should be considered such as species, age, sex, stage of infection, and other relevant history. The DSe should be determined with field samples derived from fish populations with known exposure to the pathogen. Conversely, DSP should be determined with a field population of fish with no exposure to the pathogen. For example, samples of positive and negative abalone populations collected around Australia and Tasmania were used to determine the DSe and DSP of a qPCR test for the detection of abalone herpes-like virus (Corbeil et al. 2010). If there are no field samples from exposed and naive reference populations available, then experimental animals that have been deliberately exposed to the pathogen can be used. A single sampling time point of individual experimental animals should be used for the calculations. The use of experimentally infected fish is not optimal as these samples are not subjected to the numerous natural conditions and variability experienced by fish in the target population, which results in an overestimation of either DSe and DSP (Jacobson 1998; Greiner and Gardner 2000). Additionally, since calculations are based on a single time point, the stage of infection may affect the DSe estimate. Thus, experimental animals may be useful for initial calculations of DSe and DSP, but as field samples and test results are collected over time, these initial estimates of DSe and DSP should be revised. Further discussion and formulas can be found elsewhere (Jacobson 1998; Banoo et al. 2006).
Calculation of Diagnostic Sensitivity and Specificity

The traditional method of calculating DSe and DSp is relatively simple when assay results are interpreted as a binary result (positive–negative; Smith 2006). In this case, a positive–negative cut-off value for the CT or copy number values must be determined. This cut-off value can be determined analytically during the assay development by clearly defining the LOD. The CT cut-off influences DSe and DSp; for instance, if the positive–negative cut-off is lowered and more reference samples are considered negative, then DSe typically decreases while DSp typically increases. Methods for determining cut-off values for qPCR results and the effect of the cut-off on DSe, DSp, and PV have been reviewed extensively by Caraguel et al. (2011). Deciding the acceptable values for DSe and DSp is dependent on the purpose of the test. For example, if the test will be used to screen for freedom for disease, it may be more important to have a high DSp to minimize false positives (Greiner et al. 2000). However, if the test is used for surveillance to monitor the spread of an endemic virus into new geographical locations or new hosts, then a high DSe may be more important. Alternatively, different cut-offs can be chosen to maximize DSe or DSp, and samples that fall between these CT cut-off values can be confirmed with secondary (adjunct) tests. The high ASE of qPCR makes confirming test results with a nonmolecular method difficult; in these cases, there is precedent for using DNA sequencing to confirm qPCR results (Arsan et al. 2007; OIE 2009).

A number of methods are available to estimate DSe and DSp in the absence of a gold standard; a flow chart to aid in choosing the best method based on the reference tests available is presented by Reitsma et al. (2009). Three of the most common approaches are (1) adjustment of DSe and DSp when comparing the new test with a reference test of known DSe and DSp, (2) the composite reference standard method, and (3) latent class modeling. A fourth approach is discrepant resolution analysis, but this method is not recommended as it is subject to bias because the infection status is conditional on the results of the new assay (Hadgu 1999). The first method, adjusting DSe and DSp, requires a reference test that has a defined DSe and DSp, and the reference and new test should ideally be independent (Enoe et al. 2000; Gardner et al. 2000). The dependence of two tests is briefly discussed below and is reviewed by Dohoo et al. (2009). The second approach, composite reference standard, combines multiple imperfect reference tests results to define a new composite reference standard to compare the new assay (Alonzo and Pepe 1999). If two comparison reference tests are being used, they can be combined in parallel (e.g., both tests must be negative for the fish to have a negative disease status). The composite reference standard method was used to determine DSe and DSp of assays for ISA (Nerette et al. 2008).

The final approach is latent class modeling, which can be used when there is no appropriate reference test available (Hui and Walter 1980). Latent class modeling allows for estimations of DSe and DSp as well as the prevalence in a population by employing either maximum likelihood or Bayesian estimation procedures (Nerette et al. 2005; Gustafson et al. 2008; Nerette et al. 2008). There are three assumptions that must be satisfied to use latent class models: (1) at least two populations must be tested and prevalence must vary between these two populations, (2) DSe and DSp should be constant across the populations, and (3) the tests should be conditionally independent given the infection status (although methods to account for conditional dependence of tests in latent class models were reported by Qu et al. 1996). Conditional dependence of test sensitivity or specificity occurs when the second test has a different sensitivity or specificity for infected fish that test positive on the first test compared with those that test negative on the first test (McClure et al. 2005).

REPRODUCIBILITY AND RUGGEDNESS

Diagnostic reproducibility is a measure of variation in test results from identical samples obtained by using the same protocol in different laboratories (Dohoo et al. 2009; OIE 2009; Table 1). Ruggedness is very similar in concept to reproducibility, but describes how well the assay performs when exposed to numerous different operating conditions such as differences in sample shipping conditions, change in reagent batches, or use of different laboratory equipment (Crowther et al. 2006; OIE 2009).

Calculating diagnostic reproducibility is straightforward if results are interpreted as positive or negative and agreement among samples can be directly calculated. Kappa, or Cohen’s kappa statistic, is a better measurement of agreement because it accounts for the agreement due to chance (Cohen 1960). However, the kappa statistic is dependent on prevalence of the pathogen in the population, which can confound results; kappa will be lower if the prevalence is very high or very low (Feinstein and Cicchetti 1990). Agreement among tests can also be evaluated using CT values or pathogen copy number, rather than using dichotomous positive or negative results. A variety of approaches can be used including coefficient of variation, intraclass correlation coefficient, Pearson’s correlation coefficient, concordance correlation coefficient, and the limits of agreement plot (Dohoo et al. 2009). However, these approaches can only assess the agreement among positive samples. A comparison of these approaches is presented by Quinn et al. (2009).

IMPLEMENTATION OF qPCR IN THE DIAGNOSTIC LABORATORY

Sample Collection, Storage, and Extraction

How fish tissues are handled, processed, and stored can have a substantial effect on the quality and quantity of nucleic acids available during the extraction process. Sample acquisition constitutes the first potential source of experimental variability. Diagnostic laboratories should have protocols to ensure that
samples collected are appropriate for the test being performed. Furthermore, there should be clear acceptance–rejection criteria in the event of departure from the standardized collection procedure and the testing should not be performed if the samples are not suitable. Given the high sensitivity of qPCR testing, it is essential that collection procedures avoid nucleic acid carryover between specimens. Work surfaces and nondisposable tools must be decontaminated with a nucleic acid degrading solution such as sodium hypochlorite (bleach) between each sample. Fish tissues may be obtained from remote field locations and protocols must ensure that sample integrity is maintained through the collection period, transport, and receipt of the sample. Samples can be stabilized and stored in a variety of manners, but the chosen methodology should be empirically evaluated for each tissue type. Samples for RNA can be stabilized by freezing in liquid nitrogen or by the addition of stabilizing solutions such as RNA-later (Ambion, Austin, Texas), followed by long-term storage at −80°C. It is important to note that some viruses preserved in RNA stabilizing solutions may still be infectious (Uhlenhaut and Kracht 2005). Tissue samples for DNA can be preserved in 95% ethanol or directly frozen at −20°C or −80°C. Methods that do not require cold storage or ethanol, such as those that use FTA Cards (Whatman, Piscataway, New Jersey), may represent a useful approach for stabilizing DNA in tissue samples collected in the field (Sudhakaran et al. 2009). Samples should be clearly labeled and accompanied with specimen collection data such as the name of collector, the date, time, and location of collection, the number of specimens, and environmental conditions. Sample identity and accompanying information is maintained throughout the life of the sample and housed in a laboratory information management system (LIMS), thereby enabling sample tracking and archiving of relevant data. Manuscripts and reports should indicate how tissues were sampled and stabilized and under what conditions were they stored.

A variety of methods for efficient extraction of RNA and DNA exist, and many commercial kits are available. The best extraction efficiency is obtained when researchers adhere to the manufacturer’s recommendations, particularly to the suggested tissue weight or volume. Extraction efficiency can be determined by spiking known numbers of pathogens into negative control tissue before extraction. Nucleic acids and contaminants co-extracted from tissues can overwhelm or interfere with extraction efficiency. Determining final nucleic acid concentration and quality by spectrophotometric analysis is useful for monitoring extraction efficiency and quality control.

Nucleic acid extraction is a major bottleneck in sample throughput, and there are a number of automated systems designed to partner with robotic workstations. However, these high throughput methodologies should be carefully evaluated to ensure no cross-contamination occurs among samples. Multiple positive and negative extraction controls placed throughout the plate are informative for detecting cross-contamination.

**Laboratory Design**

All PCR-based methods are prone to false positive results due to sample contamination (Schweiger et al. 1997). Thus, it is important for laboratories to have separate dedicated spaces for reagent preparation, sample preparation and handling, and PCR amplification (Kwok 1990; McCreedy and Callaway 1993). Activities are optimally divided into three separate work areas to minimize contamination potential: (1) a clean room, which is used for storage of reagents and preparation of master mixes; (2) a sample preparation room where all samples and controls are processed; and (3) a dirty room where qPCR amplifications take place. If plasmid DNA or oligodeoxynucleotides are used as controls, manipulation of these controls should be done in an area separate from sample preparation. In addition to the physical separation between rooms there should be no exchange of materials and instruments such as laboratory coats, gloves, and pipettes between rooms. Moreover, the workflow between rooms is unidirectional, proceeding from the clean room to dirty room but never the other way.

**Assay Performance Evaluation: Internal Auditing**

Ongoing quality assurance and quality control (QA/QC) programs are essential for establishing, maintaining, and guaranteeing that a laboratory is producing accurate test results. General QA/QC guidelines for fish health laboratories are provided in the AFS Blue Book (AFS–FHS 2010). The introduction of molecular diagnostic tests such as qPCR requires that QA/QC programs further accommodate the increased technical demands and potential contamination hazards.

As discussed earlier, a variety of controls must be carried out to monitor the performance of single steps during qPCR analysis (i.e., nucleic acid extraction, cDNA production, and template amplification). The inclusion of such controls is essential for arriving at a definitive test result, avoiding false positive and false negative results. Additionally, if control material that is stable over time and varies little in concentration from vial to vial is repetitively used in each assay, then the mean (x) and standard deviation (SD) can be determined. These statistics generally calculated from a minimum of 20 measurements (one assay per day for 20 d) define the upper and lower control limits (x ± 2SD) of control material. For each assay, these control limits are used to judge whether the observed control measurement represents typical or atypical performance of the qPCR assay (Westgard et al. 1981). Positive and negative controls need to be included with every run or plate and subsequently monitored for compliance with acceptable values. These data can be plotted graphically with a control chart, and changes in the assay’s precision and accuracy should be investigated (OIE 2009). This monitoring is essential for maintaining the validated status of the assay.

**Assay Performance Evaluation: External Auditing**

Internal quality control schemes monitor assay performance through process controls, but they do not provide a measure of the diagnostic performance of the assay. To assess a laboratory’s
capability to conduct a specific qPCR diagnostic assay, an external quality assessment is required. Numerous studies investigating diagnostic performance characteristics of PCR-based methods revealed that a high degree of variability in performance exists among laboratories carrying out the same protocol (Noordhoek et al. 1996; Dequeker and Cassiman 2000; Lemmer et al. 2004; Apfalter et al. 2005; Raggi et al. 2005; Ramsden et al. 2006). In all cases, incorrect results or interlaboratory discrepancies arose owing to technical reasons, data analysis (misinterpretation of results), or reporting mistakes. Thus, the validity of test results is influenced by both assay performance characteristics and technical competence. External quality assessment (EQA) programs are also referred to as laboratory proficiency testing or ring-testing schemes and evaluate the technical competence of personnel and ultimately laboratory diagnostic capacity. Unknown samples are analyzed by the test laboratory and results are sent to the evaluator where they are compared with intended results to determine the ability of participating laboratories to produce reliable results by using specific tests. Participation in EQA programs facilitates certifying the performance of individual technicians, provides a panel of samples to evaluate new test methods and harmonize existing methods, and provides a means to resolve interlaboratory differences. Participating in EQAs is often voluntary; however, for many laboratory accreditation programs, involvement in regularly scheduled proficiency testing schemes is mandatory.

Accreditation and Certification Bodies

Various accreditation programs have been constructed to provide a formal process for recognition of laboratory quality and capability by independent authorities. The internationally recognized standard for laboratory competency is ISO/IEC 17025 (ISO 2005). The OIE’s *Quality Standard and Guidelines for Veterinary Laboratories* is an interpretation of the ISO/IEC 17025 standard, specifically for the diagnosis of infectious diseases in animals (OIE 2008). For recognition to be awarded and maintained, laboratories must be compliant with all stipulated requirements as laid out in the accreditation program guidelines. In short, these programs provide confidence that laboratories operate to a recognized standard of quality. Moreover, recognition status facilitates acceptance of test results and diagnostic interpretations between laboratories that comply with this standard and other accreditation bodies who have entered into mutual recognition agreements with equivalent bodies in other countries. To date, there are few aquatic animal diagnostic laboratories that have formal accreditation. The OIE recognizes that accreditation is costly, time-consuming, and not feasible for many laboratories, but lack of accreditation to an international standard should not prevent a laboratory from performing diagnostic tests or prevent countries from entering into trade agreements (www.oie.int/en/our-scientific-expertise/reference-laboratories/proficiency-testing/).

CONCLUSIONS

The increased use of real-time or qPCR for surveillance or as a screening test for aquatic animal pathogens has raised concerns about the credibility of results as well as the implications of findings. This is particularly problematic when qPCR results cannot be confirmed by any other diagnostic test and when identification of a particular pathogen has major repercussions for industry, trade, or natural resource management. Despite these concerns, it is very likely that the trend toward increasing use of qPCR technology will continue. The Canadian National Aquatic Animal Health Plan (NAAHP) has recently been established into Canadian law and is comanaged by the Canadian Food Inspection Agency and Fisheries and Oceans Canada. The Canadian National Aquatic Animal Health Laboratory System (NAAHLS) is responsible for the diagnostic testing, targeted research, and scientific advice in support of the NAAHP in Canada. The NAAHLS is moving forward in its goal to develop, standardize, and validate qPCR tests for priority aquatic animal pathogens following OIE guidelines. In the United States, the Joint Subcommittee on Aquaculture convened a task force consisting of representatives from the U.S. Departments of Agriculture, Commerce, and Interior to develop the U.S. NAAHP. As the United States moves forward to implement this plan and the associated National Aquatic Animal Pathogen Testing Network (NAAPTN), it will need to evaluate how qPCR technology should be used. The NAAPTN has recently convened the first aquatic technical working group with the goal of evaluating and validating rapid diagnostic screening methods for VHSV. Comprehensive standardization and rigorous validation are still the key challenges to be addressed for the majority of PCR-based diagnostic tests for aquatic pathogens.

ACKNOWLEDGMENTS

The review was produced by the American Fisheries Society, Fish Health Section ad hoc committee on quantitative PCR. The authors thank Jim Winton (U.S. Geological Survey), Gael Kurath (U.S. Geological Survey), Andy Goodwin (University of Arkansas, Pine Bluff), Geof Grocock (Cornell University), Peter Wright (Fisheries and Ocean Canada), E. Scott Weber (University of California–Davis), and Emir Hadzoc (University of California–Davis) for their critical reviews of the manuscript. The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the U.S. Department of Interior, U.S. Geological Survey, or Fisheries and Oceans Canada of any product or service to the exclusion of others that may be suitable.

REFERENCES

QUANTITATIVE PCR FOR DETECTION OF AQUATIC ANIMAL PATHOGENS

159


