International Seed Federation
Seed is Life

Technical Guidelines for PCR Assay Development for Seed Health Tests

MARCH 2024, VERSION 1

Developed by
International Seed Health Initiative (ISHI) of the International Seed Federation (ISF)

TECHNICAL GUIDELINES AND THEIR PURPOSE

The following technical guidelines have been developed to facilitate in the design of PCR assays for seed health tests and help the ISHI group to keep track of performance data on PCR assays.

These guidelines are not intended to serve as a legal reference. They are not binding on ISHI nor ISF members.

DISCLAIMER

This document has been developed based on the current technical state of the art. ISF cannot be held liable for any possible claims associated with the application of these guidelines.

CONTENTS

INTRODUCTION

Polymerase chain reaction (PCR) is a biochemical process that uses short synthetic single stranded DNA fragments, called primers, to select a genetic region of interest (i.e. target), called an amplicon. During the process, repeated cycles of heating and cooling result in exponential synthesis of the desired amplicon. In the scope of International Seed Health Initiative (ISHI), the PCR processes designated in the current document only imply conventional PCR, Real-Time PCR (qPCR) and reverse transcriptase (RT-)PCR with the aim of targeting seed borne pathogens.

Primers are oligonucleotides specifically designed to match the target sequence for amplification. Considering that primers determine the specificity of the PCR amplification, primer design is a very important part in PCR assay development. Each PCR assay requires at least one forward and one reverse primer.

There are several steps in the PCR process. In conventional PCR, the double stranded DNA is first heat-denatured and then cooled in the first cycle, allowing the primers to anneal to the 3' of each DNA strand of the amplicon. Subsequently, new complementary strands will be synthetized starting at the 5' end of the molecule, resulting in the doubling of DNA molecules of the target sequence for each PCR cycle. Real Time PCR (qPCR) requires, in addition, a fluorescent reporter alongside the primers. The amplified target sequences can be detected using gel electrophoresis for conventional PCR or by fluorescence detection in the case of qPCR. The cycle quantification (Cq) value indicates the cycle number in qPCR at which the fluorescence signal crosses a given threshold that serves to discriminate the amplification from the background. For amplification of RNA samples, (q)PCR is used following an initial reverse transcription (RT) step. The RT step is necessary to produce complementary DNA (cDNA) from an RNA template followed by DNA amplification.

In the case of qPCR based strategy, different chemistries can be used to detect the PCR product. The most commonly used are DNA binding dyes or a specific oligonucleotide probe with a bound fluorophore. The DNA-binding dyes, such as SYBR® Green, will bind to all double-stranded nucleic acids (including non-specific PCR products, primer dimers, etc.). As a result, they are less suitable for multiplex PCR assays, which use multiple primers in a single PCR reaction. The specific oligonucleotide probe with a bound fluorophore, such as TaqMan® chemistry, will bind only to DNA containing the complementary sequence, increasing the specificity significantly in comparison to the DNA-binding dyes. The probe consists of a quencher that is downstream of a reporter dye, which is at the 5' end. Multiplex assays are possible with this chemistry due to the availability of different reporter dyes, thus allowing for the detection of several target sequences in the same reaction. Complex multiplex mixtures can be less robust due to an exponential increase in oligonucleotide (primer/probe) dimer interactions and potential non-specific amplification of plant and microbial matrix DNA/RNA (Life Technologies, 2012). These oligo dimers or non-specific amplification may have an inhibitory effect on the amplification efficiency of the targets by competition for PCR reagents.

Developing PCR assays, including the design of primers and probes can be challenging. The present guidelines will discuss several elements and critical steps to take into consideration to minimize non-specific amplification. The document will mainly focus on qPCR and RT-qPCR, as these are most used in seed health testing currently. For efficiency, the term PCR will be used in the document as a generic term involving all types of PCR described.

1. GOAL OF THE PCR ASSAY

When developing a PCR assay, the goal(s) of the assay should first be clearly defined. Essential elements of the goal(s) are the target to be detected, and the required performance of the PCR assay. Furthermore, it needs to be defined if the target will be detected with multiplex (typically duplex or triplex) assays or singleplex assays.

1.1. Defining targets and non-targets

The target may be a single species, multiple species, or multiple genetic lineages within a given species (e.g., pathogenic members of a single phylogenetic clade). For a single species target, the listing of the species name would be sufficient. In case of more complex targets (e.g., in case of a generic assay), a more detailed definition of the target would be required. Potentially non-targets should be listed as well (e.g., species which should not be detected with the PCR assay).

1.2. Performance criteria

The required performance of the PCR assay concerning analytical specificity (exclusivity and inclusivity) should be established before developing the assay. The final intended use of the assay within the overall seed health method (e.g., whether it will be a pre-screen PCR, or serve as confirmation PCR) should be taken into consideration when establishing the required performance. For example, in the case of a pre-screen assay, such as a seed extract PCR, a lower level of exclusivity may be acceptable than in the case of a confirmation PCR, which is the final assay of a seed health method.

ISHI has identified three performance criteria, analytical specificity, amplification efficiency and diagnostic performance, for which definition of minimum requirements is advised (Table 1). By establishing the performance of a PCR assay, the fitness of a PCR assay for its intended purpose can be determined.

Table 1. Performance characteristics definition and description for PCR assays.

Performance characteristics	Definition
Analytical specificity (critical)	The ability of a PCR assay to detect the target(s) pathogens (inclusivity) while excluding non-targets (exclusivity).
Analytical specificity is evaluated using nucleic acid derived from pure cultures of representative isolates. When a pathogen cannot be cultured, alternatives such as infected host material or synthetic DNA/RNA representing the target of interest can be used. An assay should be evaluated with a culture collection that represent the known phenotypic and genotypic diversity of the pathogen. In addition, the assay should be evaluated with a diverse, characterized collection of non-target strains associated with the host seed, such as those originating from the same host and location, and related species. The recommended size of a characterized collection is 20-30 strains each of target isolates, as well as a minimum of 20 non-target isolates. The target of the isolates to be tested should be determined to ensure a consistent detection of target isolates and be similar between isolates.	
Amplification efficiency (critical)	The fraction of target molecules that are copied in one PCR cycle.
Amplification efficiency is evaluated using a dilution series of nucleic acid derived from pure cultures of strains of the target of interest. Efficiency of 90 up to 110% and a linear regression of $R2 \geqslant 0.98$	

are accepted as good performance characteristics and sub-optimal values need additional optimisation. Dilution series should have a maximum dilution factor of 10× per step and consist of minimally four steps, aiming at Cq values ranging from ~20 until no amplification. A dilution series with a minimum of two replicates should be tested to build a standard curve.	
Diagnostic performance (optional)	An evaluation of the performance of the new assay in comparison to the reference assay to discriminate between positive and negative samples.
The performance of an assay is evaluated using DNA/RNA samples and compares the results of the reference assay *versus* the new assay when available.	

2. LITERATURE REVIEW AND SEQUENCE DATA ASSESMENT AND USE

Once the goal(s) of the new assay is defined, it is recommended to review the literature, notably on the biology of the pathogen, and sequence data available in public and private domains, before starting the development of a new assay.

A literature screen should contain a search for reports on existing PCR assays in multiple databases such as NCBI/PubMed, Google scholar, University libraries and conference/meeting reports. It is recommended to capture all described detection methods, and to align all published PCR assays to a sequence consensus and investigate the assay performance.

A sequence consensus should be created on all available pathogen data, including different isolates/strains/pathotypes, etc. using curated reference data repositories based on published scientific work, such as NCBI or any other database. Erroneous annotations can also be found in these reference databases and should be considered. The quality of genomes should be prioritized rather than the number or genomes. Genome assemblies consisting of a single contig (closed, complete, or finished genomes) and genomes with only a few contigs are generally high quality genomes. Sequences, which are associated with publications, are preferred due to background information and metadata.

If no reference data is available, generating new sequencing data should be preferred over identifying a sequence in closely related organisms. New sequencing data should be of sufficient quality for publication in a peer-reviewed scientific journal and the data should preferably be made available in a public database. When generating sequencing data, define the metadata extensively (information about the isolates, e.g., origin, year of collection, host) and information on the sequencing technique used as well as details on the assembly and, if performed, annotations. It is important to confirm the exact sequence of the pathogen that will serve as the positive control, for example by Sanger sequencing of the target regions of the resulting amplicon. Any putative mismatches within primers or probe binding sites may affect the PCR performance.

3. DESIGN OF PRIMERS AND PROBES

The PCR assay developed should be compatible with the ISHI Best Practices for PCR assays in Seed Health Tests, particularly regarding the inclusion of ISHI-recommended controls. A SE-PCR pre-screen should be developed according to the ISHI guidelines for pre-screens (Real-time PCR, an 'indirect' test used for pre-screening in seed health methods).

A successful PCR reaction requires highly specific and efficient primers and probes. The use of two primer sets for each target pathogen that amplify different genomic regions is highly recommended in the ISHI Best Practices for PCR Assays in Seed Health Tests to reduce the risk of false negative results.

3.1. Genome location

The genome location for primers and probes design depends on the target(s). If a general identification is required (e.g., for a family of viruses), then the primers should target a sequence present in all the species/variants in the group. Typically, these are on a conserved region containing polymorphisms that allow discrimination of the target from non-target groups, so called out-groups. In case the target is a specific species, primers should amplify a sequence that is present solely in the target pathogen and is not present in closely related species. Alignments of target and non-target sequences are helpful to indicate suitable target sequences.

It is recommended that primers target the genomic loci linked to pathogenicity of a pest in question (e.g., (a)virulence gene), to avoid detection of closely related non-pathogenic species. It is also important to ensure that the targeted sequence is present in all isolate variants of the target pathogen.

3.2. Melting temperature

The melting temperature (Tm) of a primer/probe, the temperature at which one half of the DNA duplex will dissociate to become single-stranded, indicates the duplex stability. It relies directly on the length and composition of the primer/probe. A longer primer strand and a higher GC content result in a higher Tm. The annealing temperature (Ta), used in the annealing step of a PCR, is highly dependent on the Tm of primers and used PCR reagents. Differences in the Tm of the forward and reverse primers should be minimized, also when designing assays which are to be multiplexed in the same reaction. When detecting RNA by RT-qPCR, it is not advised to design the reverse primer as the most discriminative element of the assay as the lower annealing temperature during the reverse transcription step may allow for unspecific annealing and thus may decrease the specificity of the overall assay.

Different Tm calculators might provide different Tm's for the same primer/probe and other factors, such as salt concentration, can play a role. For probes, special extensions can be used to increase the Tm, such as minor groove binding (MGB) and Locked Nucleic Acid (LNA). While being non-specific, an MGB-probe can help to align Tm for multiplex applications (Kutyavin *et al.*, 2000). LNA-probes are typically shorter probes and have LNA oligonucleotides that ensure higher thermal stability and increase the maximum annealing temperature of LNA-probes (Latorra *et al.*, 2003). However, regular probes are preferred over MGB or LNA probes because of costs and flexibility in the choice of fluorophores. Recommended Tm for the primers and probes are presented in Table 2.

3.3. Primers and probes characteristics

Publicly available tools can be used to assist with primer/probe design (e.g., Primer3Plus (Rozen and Skaletsky, 2000; Untergasser *et al.*, 2007), Primer BLAST (NCBI)) and Tm calculation (e.g., Multiple Primer Analyzer (Thermo Fisher)). Many bioinformatic applications also contain tools for assay development (e.g., CLC bio (Qiagen), DNAStar (Lasergene), Geneious (Dotmatics), Netprimer (Premier Biosoft)).

The designed primers must match the target sequence and avoid non-specific amplification. General rules and guidelines to design good primers and probes have been extensively described (Apte and Daniel, 2009; Green and Sambrook, 2019; Kusser *et al*., 2006; "Multiplex Real-Time PCR," 2006; Sachse and Frey, 2002). Recommended characteristics for primers and probes by ISHI are described in Table 2.

Table 2. Recommended characteristics for primers and probes.

Recommended characteristics	Primers	Probes
Length	18-30 nucleotides	13-25 nucleotides
GC content	40-60%	
Melting temperature (Tm)	58-68 °C	At least 5 °C higher than primers
3' end		Optional extension (e.g., MGB)
5' end		No G allowed
Amplicon length		50-150 nucleotides

Nucleotides that facilitate discrimination between targets and non-targets should be positioned close to the 3' end of primers. Probes should not contain a "G" at the 5' end since it has a quenching effect on the fluorophore at that position (Lietard *et al*., 2022). When designing primers and probes, ensure the probe does not overlap with the primer binding regions.

3.4. Degenerate primers

A primer is termed specific if it targets one unique sequence and is termed degenerate if it targets a collection of unique sequences. Degenerate primers may be positioned on several possible bases, which are represented by single characters as a standard representation of DNA bases by the International Union of Pure and Applied Chemistry (IUPAC; Johnson, 2010). Degenerate primers can be used if conserved areas are not present in the target to be detected. When working with degenerate primers, a few points should be considered:

- Primer degeneracy should be kept to a minimum.
- One to two mismatches in the primers and probe can be tolerated if they are not at the 3' end of the primer.
- If single nucleotide polymorphisms (SNPs) occur on multiple positions in the target sequences and are linked together, primers should be separately designed with the observed combinations and mixed when setting up the PCR. This separate design is preferred over incorporation of (mixed) nucleotides during synthesis as it is the case with degenerate primers, since this will lead to the creation of additional primers that have no direct contribution in the assay.
- In exceptional cases, if positions have three or four possible nucleotides, an inosine residue can be inserted at that position. Inosine functions as a 'neutral' nucleotide minimizing the complexity of primer mixes.

3.5. *In silico* study

It is recommended to perform a BLAST comparison for the newly designed assay(s), using the default parameters or change them if necessary (https://www.ncbi.nlm.nih.gov/tools/primer-blast/).

In the BLAST results report the following parameters should be considered:

- Query cover - percentage of the target sequence that overlaps the reference sequence. Only one to two mismatches are allowed. Mismatches should not be present in the 3' end of the primer sequence.
- Expected (E) value - the number of expected hits of similar quality (score) that could be found just by chance. The smaller the E-value, the better the match.
- Percentage identity - reports on the percentage of base pairs that are the same between the target sequence and that of the reference sequence. The higher the number, the more identical the sequences are.

Care needs to be taken that primers and probes do not interact with each other. Folding and secondary structures may impact the PCR sensitivity. Publicly available tools can be used to predict structural features of target sequences (e.g., MFold (Zuker, 2003)). Furthermore, primer/probe dimers, two oligonucleotides that have attached to each other because of strings of complimentary bases, may also impact PCR sensitivity. Primer/probe dimers can be predicted using publicly available tools (e.g., Multiple Primer Analyzer (Thermo Fisher), OligoAnalyzer (IDT DNA)).

4. PCR ASSAY OPTIMISATION AND EVALUATION

It is important to test the designed primers and probes in wet lab tests to determine if the *in silico* design performs as expected. Some PCR designs push the boundaries of being able to amplify specific targets and avoid amplification of non-specific products. Optimising and testing different PCR parameters helps in finetuning these PCR boundaries and in reducing the chance of false positive or false negative outcomes. Screening the PCR amplicons on length and number of fragments using a visualisation method, like gel electrophoresis, is important for assay evaluation. Assessing PCR performance criteria will help to determine the fitness of a PCR assay for its intended purpose.

4.1. Quality of primers and probes

Order primers and probes from a trusted provider and control their quality tested. Ask for specification sheets when ordering. In the specification sheets, the following information of the primers and probes may be indicated:

- Name and order number
- Sequence
- Properties
  - Tm (at certain molarity of NaCl and other salts)
  - Molecular weight
  - Amount and mass (as determined by absorbance at OD260)
- Extinction coefficient
- Amount of oligo (whether dissolved or not in buffer)
- Purification method (Standard desalting, HPLC, or other)
- Method of delivery (pre-dissolved to a standard molarity or not, solvent)
- For probes: type of modification (Fluorophore, quencher, MGB etc.)

Purification: For primers from 18 to 35 nucleotides in length, standard desalted grade is sufficient for qPCR applications. HPLC or PAGE purification is only recommended for very long primers (over 35 nucleotides). For probes, use standard reverse phase HPLC.

When received, always check the integrity of the package and tubes. Check labels, compare the order with the sequence received and the amount of solution in the tubes if ordered pre-dissolved. After dissolving the probes in the buffer recommended by the supplier, protect them from exposure to strong light sources and store primers and probes at -20 °C. Avoid repetitive thawing and freezing of solutions by preparing aliquots.

4.2. Optimisation of PCR

Run the assay(s) using default primer and probe final concentrations of 0.4 µM primer and 0.2 µM of probe using the recommended mix composition and PCR program provided by the supplier of the chemistry and equipment used. Use an annealing temperature (Ta) based on the Tm of the primers that is calculated or provided by the supplier of the primers and run the program for 40 cycles. Perform this test on RNA/DNA with at least a single positive process control (PPC) and a negative process control (NPC) for the assay(s) and a non template control (NTC) (see Best Practices for PCR assay). It is recommended to perform this PCR on DNA/RNA extracts of pure isolates or symptomatic material confirmed by a different assay. When no positive/suspected results are obtained with the PPC, the PCR conditions are suboptimal, potentially due to the primers and probes design, target DNA/RNA material used, extraction method, master mix chemistry and/or Ta of the PCR program.

For Ta optimisation, run a gradient PCR with Ta temperatures ranging from 58 to 68 °C on RNA/DNA of at least a single target and non-target sample for each assay, and a non template control (NTC). It is recommended to perform this test with a target sample which gives a Cq value of ~32. Lower DNA/RNA concentrations give better insight on the effects of Ta on a PCR run.

When additional optimisation is needed for a target assay, repeat the test with the same probe concentration, but different primer concentrations (e.g., 0.8 µM, 0.6 µM, 0.2 µM, 0.1 µM). Using the optimum primer concentration, perform this optimisation step as well with varying concentrations of the probe (e.g., 0.4 µM, 0.3 µM, 0.1 µM and 0.05 µM).

4.3. Performance evaluation

Evaluate and verify performance of the newly developed assay according to the performance criteria defined in Section 1.2, Table 1.

For the evaluation of the analytical specificity of a newly developed PCR assay, it is important to test as many isolates as possible. In addition, different isolates might be present in different countries and a globally collected set of isolates is necessary. Considering that isolates from different countries may vary genetically, it is important to have background information on the biology and epidemiology of the target pathogen. Genetically different isolates collected from different geographical regions, time periods, and hosts should be used in performance evaluation.

Determine the amplification efficiency of singleplex or multiplex assay by testing a 10-fold dilution series of the target sample DNA/RNA in PCR with a minimum of two repetitions. At least four dilutions of the dilution series need to result in a Cq value. It is recommended to create a dilution series until no amplification is reproducibly observed before 40 cycles. If the multiplex assay includes a positive extraction control (PEC), then test this dilution series with and without

spiked PEC. Include the spike by adding a certain amount of DNA/RNA of that PEC target to the dilution series to obtain a consistent Cq value, preferably between 28 to 32. Confirm that the PEC spike has no negative effect on the target amplification caused by competition or by unwanted interaction between oligonucleotides. There is no negative effect when the Cq value of the target, in singleplex reaction, is comparable to the Cq value of the target assay when performed in a multiplex reaction including the PEC assay. It is acceptable when the PEC is negatively affected by high concentration of DNA/RNA of the target, but not by DNA/RNA of a non-target. Use the dilution series as standard curve and calculate the efficiency rate of the PCR for the assay using the equipment software.

When applicable, assess diagnostic performance of the newly developed PCR assay by comparison against the old assay, using the same set of negative and positive samples.

When factors, such as the number of target pathogen strains and availability of infected material in case of non-culturable, new, or emerging pathogens limit the ability to establish performance of a PCR assay as required, this should be justified.

REFERENCES

Apte, A. and Daniel, S. (2009). PCR Primer Design. *Cold Spring Harbor Protocols*, **2009**(3), pdb.ip65. https://doi.org/10.1101/PDB.IP65.

Green, M. R. and Sambrook, J. (2019). Polymerase Chain Reaction. *Cold Spring Harbor Protocols*, **2019**(6), pdb.top095109. https://doi.org/10.1101/PDB.TOP095109.

Johnson, A. D. (2010). An extended IUPAC nomenclature code for polymorphic nucleic acids. *Bioinformatics*, **26**(10), 1386-1389. https://doi.org/10.1093/bioinformatics/btq098.

Kusser, W., Javorschi, S. and Gleeson, M. A. (2006). Real-Time PCR. *Cold Spring Harbor Protocols*, **2006**(1), pdb.prot4112. https://doi.org/10.1101/PDB.PROT4112.

Kutyavin, I. V., Afonina, I. A., Mills, A., Gorn, V. V. Lukhtanov, E. A., Belousov, E. S., Singer, M. J., Walburger, D. K., Lokhov, S. G., Gall, A. A., Dempcy, R., Reed, M. W., Meyer, R. B. and Hedgpeth, J. (2000). 3'-Minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. *Nucleic Acids Research*, **28**(2), 655-661. https://doi.org/10.1093/nar/28.2.655.

Latorra, D., Arar, K. and Hurley, J. M. (2003). Design considerations and effects of LNA in PCR primers. *Molecular and Cellular Probes*, **17**(5), 253-259. https://doi.org/10.1016/S0890-8508(03)00062-8.

Lietard, J., Ameur, D. and Somoza, M. A. (2022). Sequence-dependent quenching of fluorescein fluorescence on single-stranded and double-stranded DNA. *RSC Advances*, **12**, 5629-5637.

Life Technologies (2012). Real-time PCR handbook. Chapter 1: Basics of real-time PCR. 70pp. Available at: https://www.gene-quantification.de/real-time-pcr-handbook-life-technologies-update-flr.pdf. Last accessed 07 July 2023.

Multiplex Real-Time PCR. (2006). *Cold Spring Harbor Protocols*, **2006**(1), pdb.prot4113. https://doi.org/10.1101/PDB.PROT4113.

Rozen, S. and Skaletsky, H.J. (2000). Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J.A.M (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research*, **35**, W71-W74. https://doi:10.1093/nar/gkm306.

Sachse, K. and Frey, J. (2002). PCR Detection of Microbial Pathogens. Humana Totowa, NJ, USA. https://doi.org/10.1385/1592593445.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, **31**(13), 3406-3415. https://doi.org/10.1093/nar/gkg595.

Seed is Life