GENERAL PAPER

CrossMark

Validation guidelines for PCR workflows in bioterrorism preparedness, food safety and forensics

Johannes Hedman[1,2] · Moa Lavander[3] · Emelie Näslund Salomonsson[4] ·
Tomas Jinnerot[5] · Lina Boiso[1] · Bertil Magnusson[6] · Peter Rådström[2]

Abstract

The polymerase chain reaction (PCR) is the backbone of contemporary DNA/RNA analysis, ideally enabling detection of one or just a few target molecules. However, when analysing food or forensic samples the analytical procedure is often challenged by low amounts of poor quality template molecules and complex matrices. Applying optimised and validated methods in all steps of the analysis workflow, i.e. sampling, sample treatment, DNA/RNA extraction and PCR (including reverse transcription for RNA analysis), is thus necessary to ensure the reliability of analysis. In this paper, we describe how in-house validation can be performed for the different modules of the diagnostic PCR process, providing practical examples as tools for laboratories in their planning of validation studies. The focus is analysis of heterogeneous samples with interfering matrices, with relevance in food testing, forensic DNA analysis, bioterrorism preparedness and veterinary medicine. Our objective is to enable rational in-house validation for reliable and swift quality assurance when results are urgent, for example in the event of a crisis such as a foodborne outbreak or a crime requiring the analysis of a large number of diverse samples. To that end, we explain the performance characteristics associated with method validation from a PCR and biological sample matrix perspective and suggest which characteristics to investigate depending on the type of method to be validated. Also, we include a modular approach to validation within the PCR workflow, aiming at efficient validation and a flexible use of methods.

✉ Johannes Hedman
johannes.hedman@tmb.lth.se

1   Swedish National Forensic Centre, Linköping, Sweden

2   Applied Microbiology, Department of Chemistry, Lund University, Lund, Sweden

3   Science Division, Biology Department, National Food Agency, Uppsala, Sweden

4   CBRN Defence and Security, Swedish Defence Research Agency, Umeå, Sweden

5   National Veterinary Institute, Uppsala, Sweden

6   Chemistry, Materials and Surfaces, Bioscience and Materials, RISE – Research Institutes of Sweden, Borås, Sweden

Introduction

The polymerase chain reaction (PCR) is widely applied for the analysis of DNA/RNA from humans and microorganisms. Hence, PCR techniques are crucial for several sectors and applications, such as the investigation of crime and to ensure human safety through analysis of foods or suspected bioterrorism samples. In forensics as well as food safety, a false-positive or false-negative result can have dire consequences. Optimised and validated analysis workflows are necessary to minimise the risk for such events. One mutual challenge in the analysis of food, feed, and forensic samples is the wide variety of possible sample types and the heterogeneous nature of the samples. Sample matrices can have a negative impact on the analysis by bringing PCR-inhibitory molecules into sample extracts or by trapping the target cells/DNA [1, 2].

🖄 Springer

Further, the target DNA/RNA is often partially degraded and present at low levels.

Validation guidelines for chemical analysis have been developed, for example by Eurachem [3], to simplify and standardise method validation, helping testing and calibration laboratories to improve their quality assurance and apply for accreditation according to the ISO/IEC 17025 standard [4]. PCR differs from classical chemical analysis as it is based on the capacity of an enzyme, a DNA polymerase, to amplify specific DNA fragments. In PCR, the target nucleic acid sequence is amplified and subsequently analysed through a process consisting of physical as well as biochemical factors. A few validation guidelines directed towards PCR-based analysis have been published, mainly for analysis of Genetically Modified Organism (GMO) content in foods [5–7] but also for forensic DNA analysis [8, 9]. Most guidelines have been focused on the actual PCR assay, but some documents incorporate considerations also for the upstream modules of the analysis chain, i.e. sampling, sample treatment, and DNA/RNA extraction and purification.

For the individual laboratory, the method validation process commonly starts with a new demand, creating a need to analyse a certain sample type in a certain way, and ends with a laboratory decision whether or not the requirements are fulfilled by the applied method [3]. The steps in the validation process are (1) set the requirements, (2) modify an existing method or develop a new method to handle new targets or sample types, (3) prepare a method instruction, (4) evaluate performance characteristics through validation experiments, and finally (5) a decision regarding "fitness for purpose". If the set requirements are not met, the requirements may need to be updated or the method improved. Following implementation in routine analysis, quality control measures are used to ensure the continuous performance of the method. Here, we address the scenario of a DNA/RNA laboratory that has a validated PCR workflow when faced with a new demand, such as a request from the Police to start analysing crime scene DNA from new matrices. The validation process may be handled as a part of the continuous developmental work, or handled urgently if connected with a crisis situation. Rational procedures for in-house validation are important in both cases. In many instances, the result of an analysis is critical, since actions may have to be taken depending on them; e.g. recalls of foods from stores, alerts concerning microbiological risks, or identification of culprits from crime scene samples. Should such an urgent analysis be requested for a new sample type, for which the existing methods have not been validated, there will be little time to perform validation.
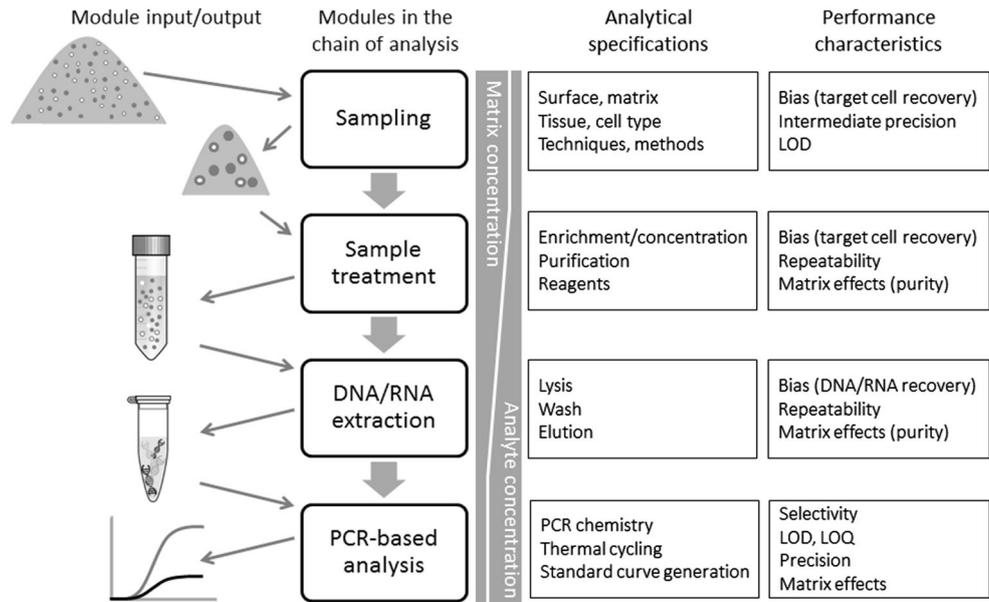
The objective of this paper is to provide validation guidelines for the different modules of the PCR workflow (Fig. 1), focusing on analysis of the challenging samples encountered in for example food testing, forensic DNA analysis, bioterrorism preparedness and veterinary medicine. In these sectors, the sample matrix has a substantial impact on the analytical success. We include a modular approach to method validation within the chain of analysis, aiming at efficient validation and a flexible use of methods. The objective is to enable rational validation of new or improved methods, or for analysis of a new sample type with an existing method. To that end, we explain the performance characteristics associated with method validation from a PCR and biological sample matrix perspective and propose which characteristics to investigate depending on the type of method to be validated. We also suggest experimental setups including which sample types to apply in single-laboratory validation of the different modules. A specific application of the guide is the need for urgent validation in the event of a crisis such as a foodborne outbreak.

The workflow in PCR diagnostics

The PCR analysis chain can be divided into four modules: (1) sampling, (2) sample treatment, (3) DNA/RNA extraction and purification, and (4) PCR-based analysis (including reverse transcription for RNA analysis) (Fig. 1). Sampling must generate a representative sample from a large surface or background material, maximise the uptake of target cells/DNA and ideally minimise the uptake of PCR inhibitors [10]. Sample treatment serves to concentrate target cells, and/or separate them from a background of other cells or matrices prior to cell lysis and nucleic acid extraction. Sample treatment may be performed with different types of methods, e.g. using ultrafiltration for large water samples when testing for pathogenic microbes [11] or using laser capture microdissection to pick up individual human cells in forensic investigations [12]. Cultivation is often needed in food testing to meet the requirement to confidently determine the absence of pathogens in 25 g of background material [13]. However, sample treatment is often time-consuming and costly, why performing extraction/purification directly after sampling is preferable, when possible. Extensive nucleic acid purification should also be avoided as it leads to loss of DNA/RNA [14]. An inhibitor-tolerant DNA polymerase-buffer system may be applied to lower the need for purification [1, 10]. This approach is part of a concept called pre-PCR processing [1], aiming at reaching an optimal limit of detection for challenging samples and at the same time keeping the analytical procedure efficient and simple. For RNA analysis, a reverse transcription (RT) step is needed prior to PCR, either as a stand-alone process or integrated with the PCR. The

Fig. 1 The PCR analysis chain described by four modules: sampling, sample treatment, DNA/RNA extraction and PCR-based analysis. The sample flow is shown to the left, starting with cells/viruses in a matrix and ending with DNA in the PCR tube. While the sample is processed, the matrix concentration ideally decreases and the analyte concentration increases (middle). Analytical specifications and performance characteristics (to the right) are included for each module for investigation in method validation (important but in-exhaustive examples)

Module input/output	Modules in the chain of analysis	Analytical specifications	Performance characteristics
	Sampling	Surface, matrix Tissue, cell type Techniques, methods	Bias (target cell recovery) Intermediate precision LOD
	Sample treatment	Enrichment/concentration Purification Reagents	Bias (target cell recovery) Repeatability Matrix effects (purity)
	DNA/RNA extraction	Lysis Wash Elution	Bias (DNA/RNA recovery) Repeatability Matrix effects (purity)
	PCR-based analysis	PCR chemistry Thermal cycling Standard curve generation	Selectivity LOD, LOQ Precision Matrix effects

success of the RT-qPCR analysis is to a large extent determined by the efficiency of the reverse transcription [15], making it vital to control this step in validation. Reverse transcription yield is for example highly affected by the primer type (e.g. random hexamers or specific primers), the RNA target and the type of RT enzyme applied [15, 16].

Depending on the aim of the analysis, one of these three technological platforms may be applied in PCR diagnostics: (1) Conventional PCR followed by electrophoresis detection (slab gel or capillary) or sequencing of amplicons, (2) Real-time PCR (qPCR), or (3) Digital PCR (dPCR). Nucleic acid analysis may be qualitative or quantitative, depending on the need, the applied platform and the analysis process. qPCR, RT-qPCR and dPCR enable quantitative analysis, but when applied following cultivation of bacteria, for example, they are used qualitatively for detection of the specific target species. Guidelines for reporting of qPCR and dPCR results have been published, with the aim to improve the quality the scientific literature and enable justified conclusions to be drawn from PCR results [17, 18]. These guidelines may be helpful also in method development and in-house validation.

In each module of the PCR workflow, there are numerous factors that affect the analytical performance and the measurement uncertainty (Fig. 2). For example, the cell type and matrix affect sampling and sample treatment, the applied cell lysis reagents and thermal conditions affect DNA extraction, and standard curve generation and DNA quality affect the PCR measurement. The relevant sources of variation should be considered when designing validation experiments.

Module-based method validation

The modular nature of the PCR workflow lends itself well for a modular approach to method validation as proposed by Holst-Jensen and Berdal [19] (Fig. 1). There, a module is defined as a method to be used in a certain step of the analysis chain. If the modules are independent, each module may be validated separately, not as a part of the complete procedure. This increases flexibility, as a validated module may be used in several different workflows without the need for re-validation of the whole workflow. However, the complete independence between modules cannot be assumed in all instances. Only limited work has been directed towards proving the generality of the modular approach, and to the best of our knowledge only in the GMO field [20, 21]. Holst-Jensen and Berdal propose to evaluate the performance of each module by applying non-PCR methods, e.g. optical density (OD) absorbance measurements to estimate DNA concentration and purity for validation of DNA extraction protocols [19]. This tactic is valid if the modules are truly independent. In our case, analysing samples containing heterogeneous matrices that may disturb PCR, it is important to verify the compatibility between the existing PCR workflow and the method to be validated. We, therefore, suggest the application of previously validated methods from the PCR analysis chain when validating a new module. Thus, the method performance can be confirmed in a relevant context, without the need for complete validation of the workflow, keeping the flexibility provided by the modular approach.

Additionally, methods such as OD or fluorometry for measuring DNA concentration and purity may not give relevant results with respect to PCR. PCR inhibition, for

Fig. 2 Sources of variation in PCR diagnostics. Factors that affect the performance, variation and measurement uncertainty of PCR analysis in each of the four modules are shown. Other factors than the ones mentioned may also affect variation, such as the reverse transcription step in RNA analysis
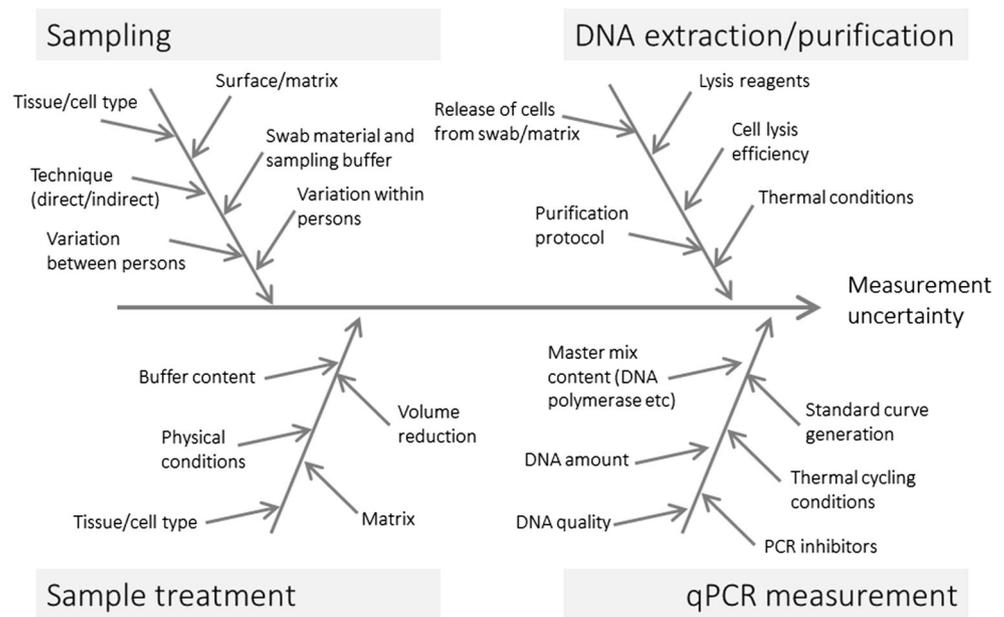
example, is largely dependent on the applied DNA polymerase-buffer system and is not directly reflected by OD measured sample impurities [22, 23]. Also, for mammalian cells, viruses and some bacteria, for which culture-based methods are not applicable, there are no readily available methods for estimating the performance of sampling or sample treatment without applying DNA/RNA extraction and PCR.

The impact on total measurement uncertainty from a certain module may be estimated during or after validation, if necessary for the application. For example, if the variation coming from sample treatment, DNA extraction and PCR is known, the variation from sampling can be deduced from experiments performed as described above. In the PCR community, it is widely accepted that the upstream processes of sampling, sample treatment and extraction/purification, as well as reverse transcription in RNA analysis, add more to the variation than the PCR assay [15, 19].

Performance characteristics

The first step in single-laboratory validation of a new or improved module in the analysis workflow should be to state the requirements on the method. The requirements are generally given as limits for a set of performance characteristics, i.e. selectivity, limit of detection (LOD), limit of quantification (LOQ), working range, analytical sensitivity, trueness, precision, ruggedness, and matrix effects. We have also included contamination risk and carry-over as these two are important parameters in PCR diagnostics. Ideally, the investigated performance characteristics

together span all the requested properties of the method, ensuring that the right target is analysed and that it can be confidently detected from low level samples containing relevant matrices. The performance characteristics are defined in the International Vocabulary of Metrology [24] and interpreted for validation of analytical chemistry methods by Eurachem [3, 25]. We build on the VIM and Eurachem guides and describe the performance characteristics from a PCR perspective, including examples and suggested experimental setups. Our descriptions are intended as support for establishing a validation plan prior to commencing the practical validation work. Different parameters may be important depending on the module to be validated. In Table 1, we suggest which performance characteristics to investigate in validation of the different modules in the analysis chain and for the different PCR technologies.

Selectivity

In analytical chemistry, selectivity is defined as "the extent to which the method can be used to determine particular analytes in mixtures or matrices without interferences from other components of similar behaviour" [26]. In the PCR context, this is related to the ability of the method to detect target DNA/RNA sequences in a background of non-target nucleic acids. In microbial analysis, detecting the variants that should be detected is referred to as inclusivity, and excluding those that should not be detected is referred to as exclusivity [27]. Here, we choose to separate selectivity from matrix effects, i.e. the impact of the matrix substances (here defined as non-nucleic acid content) in the samples.

Table 1 Performance characteristics to be evaluated in the validation of different modules of the PCR analysis workflow. Parameters that are important to investigate for a certain module type are marked with "+", those that may be tested depending on the situation are marked with "+/−", and less important/not applicable parameters are marked with "−"

Performance characteristics	Module to be validated					
	Sampling	Sample treatment	DNA/RNA extraction	PCR-based analysis		
				Conventional PCR	qPCR/RT-qPCR	dPCR
Selectivity	−	−	−	+	+	+
LOD	+/−	+/−	+/−	+	+	+
LOQ	+/−	+/−	+/−	+/−	+	+
Working range	+/−	+/−	+/−	+/−	+	+
Analytical sensitivity	−	−	−	−	+/−	+/−
Trueness						
Bias/recovery	+	+	+	+/−	+	+
Precision						
Repeatability	+	+	+	+	+	+
Intermediate precision	+	+	+	+	+	+
Reproducibility	−	−	−	+/−	+/−	+/−
Ruggedness	+	+	+	+	+	+
Matrix effects	+	+	+	+	+	+
Contamination risk	+	+	+	+	+	+

Thus, selectivity testing as described below is performed with purified DNA to distinguish the outcome from matrix effects. Matrix effects and in particular PCR inhibition are further discussed later.

An initial step in evaluating selectivity is to ensure that the generated signal originates from the requested analyte, i.e. confirmation of identity (Fig. 3). In PCR diagnostics, the analyte is a specific DNA or RNA sequence which is amplified to enable identification of, for example, a particular individual in a forensic investigation or pathogen in

Fig. 3 Confirmation of identity in PCR-based analysis. In this example, the source of the qPCR output (amplification curve) is verified by performing gel electrophoresis to determine the size of the generated DNA fragment. The grey amplification curve and gel bands are the result of correct amplification, confirming that the assay detects the target it is supposed to detect. The black amplification curve, on the other hand, comes from the detection of incorrect (smaller) amplicons (unspecific products or primer-dimers)

the food supply chain or in veterinary medicine. The identity can be confirmed by applying reference strains or purified reference material with known DNA profiles [28], or reference methods. As stated in the Eurachem validation guide [3], an independent method should be used to confirm that the analysis method identifies the analyte it is designed to detect. In qPCR, amplification curves are generated that should reflect the amplification of the target. However, this signal could also be caused by the amplification of non-specific products or artefacts such as primer-dimers, especially when non-specific DNA binding dyes such as SYBR Green I are used for detection. Here, the confirmation of identity can be achieved by determining that the PCR product has the expected size, for example applying gel electrophoresis (Fig. 3) or melt curve analysis. For further confirmation, the product may also be sequenced and identified in a nucleotide sequence database, if deemed necessary.

Confirmation of identity testing should preferably be performed using pure DNA/RNA from a specific target microorganism or a human individual, depending on the application. Confirmation of identity testing can be performed as a limited and simple experiment, often done when the PCR assay is first set up at the laboratory.

For PCR assays, selectivity is determined primarily by the constructed primers and probes, supposed to bind specifically only to the intended sequences of the target region(s) (Fig. 4). However, selectivity is also affected by
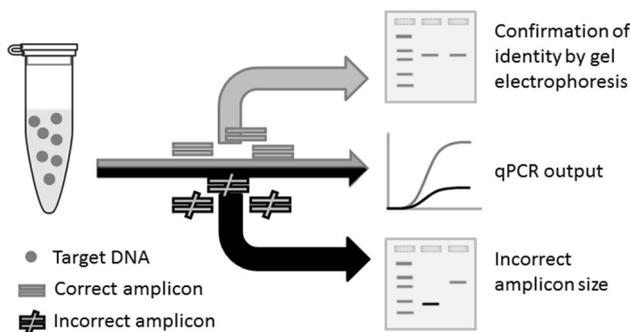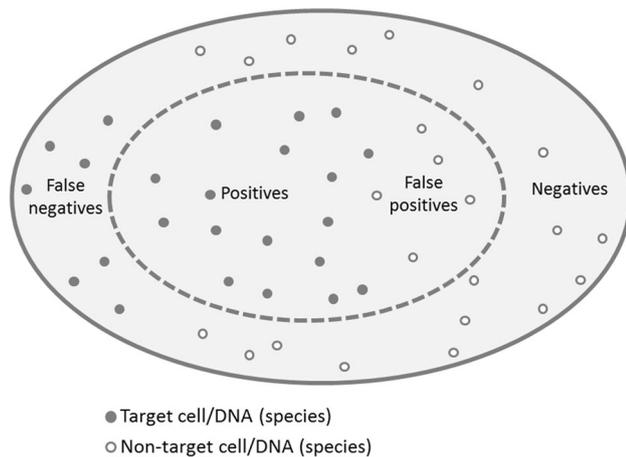
Fig. 4 Selectivity of a PCR assay. The samples/strains detected by the assay are visualised with the dashed line circle, showing true-positive results (filled grey circles inside the dashed line), false-negative results (filled grey circles outside the dashed line), false-positive results (white circles inside the dashed line) and true-negative results (white circles outside the dashed line)

physical and chemical factors such as annealing temperature and the applied concentration of magnesium ions in the assay. A lower annealing temperature or higher magnesium ion concentration generally elevates the risk of generating faulty products through increased stability of primer-DNA binding (i.e. the primer may bind to DNA even if several bases are mismatched). Thus, selectivity must be re-evaluated if changing any of these conditions for a validated method. When an assay is designed, the selectivity is usually tested in silico using an appropriate reference genome sequence database. This gives a prediction of whether or not the designed primers will bind only to the target sequence. However, the true selectivity should be determined empirically, by PCR analysis of DNA extracted from target organisms, not only by in silico analysis [29].

For microbial methods, a panel of nucleic acid samples from relevant strains is usually set up to evaluate inclusivity and exclusivity. To determine the inclusivity in pathogen testing (defined as "the strains or isolates of the target analyte(s) that the method can detect" [27]), the panel should preferably include a diversity of organisms (genus, species, subspecies, serotypes, etc.) that the assay is intended to detect. For exclusivity (defined as "the non-target strains, which are potentially cross-reactive, that are not detected by the method" [27]), the panel should include: (1) closely related strains, (2) strains that are commonly found in relevant samples and (3) non-related agents which may give similar symptoms or may occur in the same environment [27]. In the ISO 22118:2011 standard for PCR detection and quantification of foodborne pathogens [30], it is recommended to use at least 50 strains

for the inclusivity test and at least 30 strains for the exclusivity test. For qPCR assays targeting human DNA, a number of human individuals and samples from other species may be tested. Selectivity is generally only relevant for the PCR modules (Table 1).

For determination of selectivity, an amount of DNA/RNA that does not challenge the limit of detection of the PCR assay should be used. The above-mentioned ISO 22118:2011 standard states that: "a clearly detectable amount of DNA, e.g. representing DNA of 10^6 cells, should be used for the selectivity testing" [30]. For bacteria, 1 ng of DNA per reaction generally meets this criterion, corresponding to approximately $1.5 \times 10^5 - 1.5 \times 10^6$ genome copies.

Limit of detection (LOD), limit of quantification (LOQ) and working range

LOD refers to the smallest concentration of analyte that can be detected by the method with a given probability. Commonly, both for PCR-based methods and in other contexts, LOD_{95} is used, which is defined as the lowest concentration of analyte at which 95 % of the positive samples are detected by the analysis method [17]. Limit of quantification (LOQ) refers to the lowest analyte concentration that can be determined with acceptable uncertainty. Working range refers to the range of analyte concentrations that can be quantified with acceptable accuracy. The lowest point in the working range is the LOQ.

LOD, LOQ and working range for a PCR assay can be determined by means of a dilution series containing known amounts of target DNA/RNA (Fig. 5). The dilution series should include several replicates and concentrations of nucleic acid to give a useful estimate of the LOD and/or LOQ and working range. More replicates may be
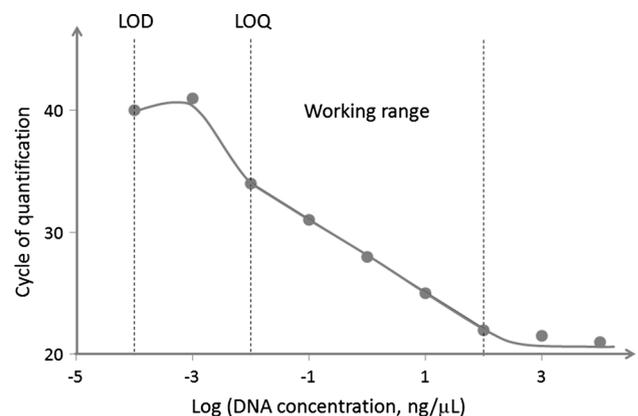

Fig. 5 Determination of LOD, LOQ and working range in qPCR. Quantification cycle (Cq) values from a dilution series of DNA are plotted against log of the DNA concentration to generate a standard curve covering the working range

introduced close to the critical levels in order to improve the LOD/LOQ estimations. The dilution series may consist of pure standard DNA or, preferably, target cells/DNA in a relevant matrix. The latter ensures that amplification efficiency is similar for the prepared samples as for the "real" samples, making the LOD, LOQ and working range estimations relevant for the routine analysis situation. In fact, quantification with qPCR builds on the assumption of identical amplification efficiencies for standards and unknown samples. In dPCR, no standard curve is needed for absolute quantification, making the technology less affected by differing amplification efficiencies, e.g. due to impurities [31]. LOD and LOQ may also be investigated when validating pre-PCR modules, if deemed necessary (Table 1).

LOD can be determined for the PCR assay separately, but in general it is more relevant to determine the LOD for the whole analysis chain. Then, more modules and aspects of the workflow must be considered and, if relevant, included in the tests. A common test design for the evaluation of LOD is to spike (i.e. add) target cells (or nucleic acid) in different levels to relevant matrices. The samples are then processed according to the analytical procedure which can include sample treatment steps such as culture enrichment and concentration, and DNA/RNA extraction steps such as cell lysis, filtration, and elution. Spiking is not as ideal as using real samples, but as real samples are often lacking and also have unknown contents, spiking is often the best choice available.

Analytical sensitivity

Analytical sensitivity refers to the change in instrument response signal as a function of change of analyte concentration. Note that this differs from diagnostic sensitivity, which refers to the ability to diagnose correctly. The word sensitivity should be avoided when referring to LOD, to avoid any confusion. In general, less importance can be given to evaluating analytical sensitivity for PCR-based analysis; it is rarely interesting to determine which of a pair of unknown samples that contains the highest amount of target cells.

Trueness

For quantitative methods, trueness is "an expression of how close the mean of an infinite number of results (produced by the method) is to a reference value" [3]. Thus, it is connected with the systematic variation of a method. Trueness cannot be measured directly, but may be estimated as bias. Bias refers to the proximity between the measurement value and the true value or, alternatively, a reference value. For a qPCR assay, the reference value may be the DNA concentration of a certified reference material, e.g. as provided by NIST for human DNA [32].

In validation of sampling or DNA/RNA extraction methods, bias may be measured in recovery experiments. These spiking tests can be performed by adding a certain amount of target cells to blank matrices before DNA/RNA extraction and measure the recovered proportion. In this case, recovery is a measure of the efficiency of DNA/RNA extraction. Comparisons may also be made against an established reference method, where the reference method result may be set to 100 %. Alternatively, cells can be counted before spiking and the theoretical DNA amount used as a reference value. For example, it is estimated that one human haploid cell contains around 6 pg DNA [33]. Recovery is also referred to as yield.

Precision

Precision refers to the random variation of a method and may be determined as repeatability, intermediate precision or reproducibility, depending on what is most appropriate for the particular module. Distribution measurements such as standard deviation or coefficient of variation may be applied for all the three precision parameters. Precision is an important parameter for modules in all steps of the analysis chain (Table 1).

Repeatability is the variation between analyses conducted in an identical way, for example replicates within a DNA extraction batch or a PCR run. Thus, the analyses for repeatability testing are performed with identical reagents and applying the same instruments, within a short period of time.

Intermediate precision is the variation between analyses performed at one laboratory under somewhat different conditions [3], for example with different persons performing DNA/RNA extraction or applying different reagent lots or PCR instruments. Separation in time between analyses also counts as intermediate precision conditions.

Reproducibility refers to variation between measurements performed at different locations/laboratories [3]. This is a required part of validation of some newly developed analysis methods, e.g. new qPCR assays targeting pathogens. Inter-laboratory studies are, for example, required in the validation of alternative methods to be used in the official control of food and feed, replacing standardised reference methods [34]. Reproducibility may be determined through ring trials, i.e. by analysing replicated samples in different laboratories and comparing the results. For more established methods, it is generally not necessary for the individual laboratory to further investigate reproducibility as part of in-house validation.

Ruggedness

Ruggedness, sometimes referred to as robustness, is the method's insensitivity for small, consciously made changes in the experimental conditions. Ruggedness is evaluated during validation by varying key parameters or reagent concentrations and studying the effects. For PCR methods, the effects of slightly varying the temperatures and incubation times during thermal cycling or applying different primer/probe amounts may be evaluated. For reagents, deviations of around ± 10 % from the optimal concentration are frequently applied. This type of test provides information on how robust the method is in regard to pipetting errors. The outcome of the ruggedness test may be used to determine the limits of the method, for example concerning incubation time ranges in different steps of DNA/RNA extraction.

Matrix effects: PCR inhibition

Matrix effects refer to the possibility of obtaining a true positive result when the analyte is present in a certain matrix, and a true negative result if it is absent. In that way, it resembles selectivity, with the distinction that the focus is on the background material, the matrix, rather than on the design of the PCR assay. The matrix effects may improve detection, such as a matrix that acts as a carrier for the analyte or a matrix that promotes growth of a target bacterium, but it is more common that a matrix disturbs analysis. A negative matrix effect may cause false-negative results, partial results, or incorrect quantification through lowering of amplification efficiency. Determining the limitations and understanding matrix effects is a vital part of the validation of methods in the PCR workflow. Among the possible negative effects to look into are: trapping of cells in DNA extraction/purification (e.g. cells binding tightly to cotton or soil), impaired culture of microorganisms (e.g. from heroin samples), inhibition of PCR amplification (e.g. from humic acid in soil, blood, faeces, feed [10, 35]), and blocked amplicon detection (e.g. from denim fabric, blueberries, soil [36]).

PCR inhibition, i.e. disturbing amplification or amplicon detection, is arguably the most important matrix effect in PCR diagnostics. PCR inhibitory molecules may emanate from the sample, the background material, or be added in the analytical chain (Fig. 6). Examples of the latter are DNA extraction ingredients such as phenol, SDS, EDTA, Chelex, all of which are known PCR inhibitors with different modes of disturbing the reaction [10, 37]. All relevant sources of PCR inhibitors should be investigated in validation, through experiments applying relevant matrices at relevant levels. To limit the amount of experiments, matrices with varying effects are preferably chosen. See for example Ref. [35] for a list of PCR inhibitors and their respective mechanisms. The choice of matrices [13] for testing should also be determined by the nature of the target to be analysed. For a *Francisella tularensis* assay, for example, relevant PCR-inhibitory background materials include soil, mosquito, water, and clinical samples, as these reflect the environments where the bacterium may be found [38]. *Francisella tularensis* could also appear as an agent in bioterrorism [39], with other possible disturbing matrices such as various surfaces (through aerosols) and carcasses.

Contamination risk and carry-over

Contamination risk is the risk of detecting analytes not derived from the original sample, but instead being added along the analysis chain. Contaminating cells/DNA may come from the person performing sampling or DNA extraction, especially when human DNA is targeted, or from consumables and reagents used, such as swabs, plastic tubes and buffers (Fig. 7). From the perspective of food safety, contamination may have a different meaning, i.e. that the tested food stuffs contain the target microorganism
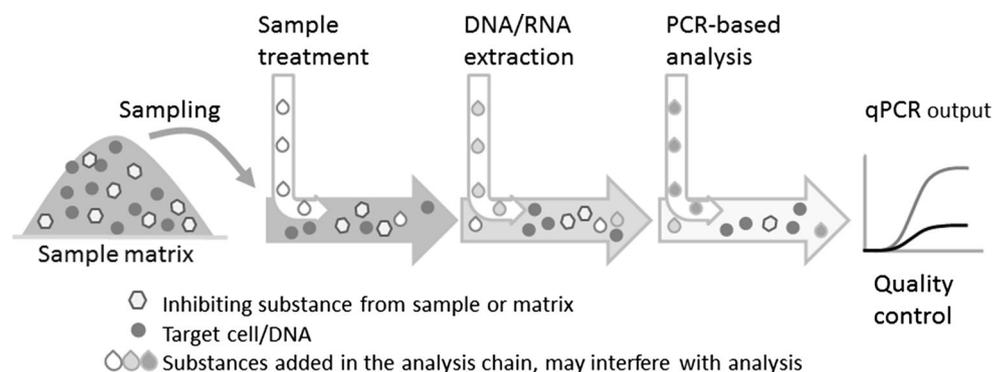

Fig. 6 Sources of matrix effects in PCR. The sample flow in the PCR analysis chain is shown. Substances that disturb PCR (i.e. PCR inhibitors) may be added to the samples in any of the modules. The grey amplification curve signifies ideal amplification, and the black curve signifies amplification affected by inhibitors (lowered amplification efficiency)

Fig. 7 Contamination risks in the analytical procedure. Contaminating cells or molecules may be added to the sample in any of the modules leading up to analysis
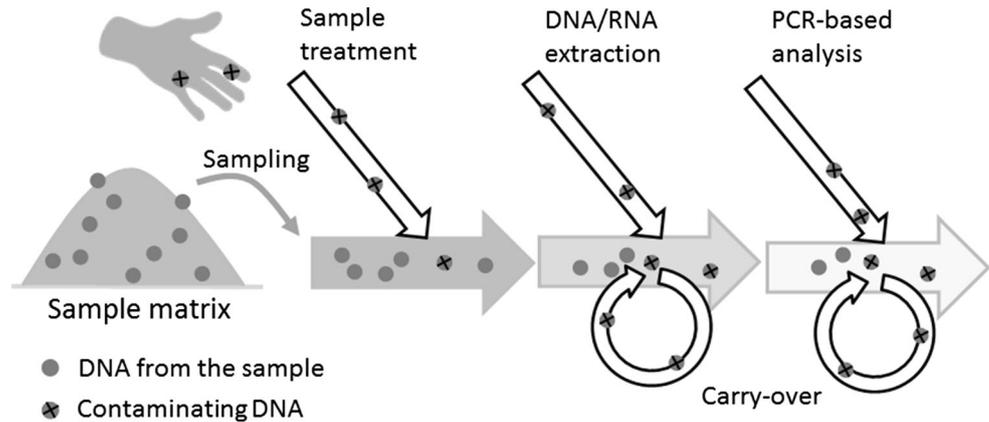
due to poor food hygiene. Here, we use the word contamination in the analytical sense described above.

Carry-over refers to the risk that a sample analysed in an instrument affects the next test. A specific carry-over issue in PCR-based analysis is the enormous multiplication of target molecules, creating a risk that amplicons from one reaction contaminate another prior to amplification. Therefore, pre- and post-PCR areas must be separated, preferably in different rooms with different air pressures [40]. Carry-over within capillary electrophoresis instruments can be evaluated by analysing blank samples following samples with high amounts of amplicons. In general, the contamination risk is investigated by including negative controls in the validation study to monitor the relevant modules. Negative controls are treated the same as the samples, with the only difference that no target cells/DNA are consciously added to them.

Planning the validation study: practical considerations

Validation can be a laborious undertaking, creating a need for rational validation design, relevant for the method at hand. Considering all possible matrices that salmonella or human culprit DNA may appear in, the theoretical scope of a perfect validation study includes an almost infinite number of samples. Hence, key (i.e. very common or particularly challenging) sample types should be chosen to make validation relevant as well as reasonable concerning time and resources. Another challenge is to limit the samples to a manageable number. Hence, the number of sample types, nucleic acid levels and replicates must be determined to get (a) the information needed to assess the performance characteristics, and (b) a feasible experimental setup. For example, analysis of 50 samples in total has been suggested for the internal validation of commercial methods in forensic DNA analysis [41].

The design of the single-laboratory validation study and number of analyses performed relies on how established

the method is, for example, whether it is a commercially available method that has been quality assured by a manufacturer or a new, in-house method. ISO methods have normally been validated through inter-laboratory testing and do not need to be extensively validated by the testing laboratory. However, the performance of ISO methods should be verified at the laboratory. A recurring question for DNA/RNA laboratories, for example in a crisis situation such as a foodborne outbreak, is: "if a method has been validated for analysis of agent X in sample type A, can the same method be applied for analysis of sample type B?" If A and B are distinct, it may be necessary to perform a limited validation study to verify the performance for B. Six different cases for method validation are listed below. In each case, the laboratory must determine how extensive the validation needs to be in their particular case.

1. Standard method (e.g. ISO)
2. Commercial method/kit, validated by the manufacturer
3. Method published in scientific journal
4. In-house developed method
5. Modified method, of type 1-4
6. Validated method to be used with new sample types/matrices

Choosing relevant sample types is an important part in the planning of a validation study. Spiking experiments are highly useful to that end, as spiking reduces the number of unknowns and enables quantitative analysis of the performance characteristics including precisions measures. However, it is difficult to mimic the full complexity of "real" samples with spiking experiments. Therefore, a range of different samples from routine analysis (or prepared samples mimicking routine samples) should preferably be applied to complement the replicated spiking experiments.

To set up a feasible validation study for a pathogen testing method, a few relevant strains of the target organism must be chosen. One such example is the inter-laboratory validation of the ISO 10272-1 method for detection

and enumeration of campylobacter [42]. The method was designed for detection of species of campylobacter in samples from the food supply chain. Seventeen laboratories participated in the validation study. Five different sample types were used (broiler caecal material, frozen spinach, frozen minced pork/beef, raw milk, and chicken skin). Each laboratory received eight samples per sample type containing high level, low level or no *Campylobacter* (i.e. 24 samples in total per sample type). One strain of *Campylobacter jejuni* or of *C. coli* was used per sample type, presumably to keep the total number of samples at a manageable number. In this example, the bacterial species were chosen since they are relevant food contaminants and also good representatives for their species. Both these factors should be considered when designing the study. In general, it should also be considered whether or not the organism is expected to be persistent (biofilm formation, resistance, etc.) in routine testing and preferably this should be reflected in the validation experiments.

Through module-based validation it will only be necessary to validate the actual method that has been added or modified, not the entire analytical procedure. This saves time and cost. However, it is still necessary to verify the performance of the whole analysis chain, to ensure the compatibility with the new method. Each step of the workflow has its own specific challenges concerning validation planning. Below we give some practical advice for each of the modules concerning experimental setup and choice of matrices for testing.

Validation of a sampling method

Sampling may be direct, meaning that a piece of a material/matrix is taken directly for further processing, or indirect, meaning that a sampling device is used to lift the sample from the material. Swabbing is arguably the most common approach for indirect sampling, in forensic DNA analysis as well as in microbial testing. To validate a sampling method, relevant matrices, free from target analyte, may be spiked with a known amount of target cells/viruses. In forensic DNA analysis, a certain amount of saliva or blood may be put on a relevant surface and sampled after drying. The outcome may be compared against a reference method, or against a theoretical value coupled to the number of target cells applied. Spiking with known amounts of target may not be applicable in all instances, for example in some forensic DNA analysis applications. When validating a method for sampling of shed human cells on clothes, reference material may instead be prepared by someone wearing a set of identical garments in a controlled fashion for a specified amount of time [43]. Trueness and precision are arguably the most important parameters to investigate in the validation of a

sampling module (Table 1). Recovery gives an estimate of the efficiency of sampling, and intermediate precision may be applied to study variation between individuals performing sampling. It may also be of importance to look into LOD, e.g. when investigating LOD for the whole analysis chain.

Validation of a sample treatment method

In many instances, DNA/RNA extraction is performed directly after sampling. However, in some cases sample treatment may be needed as a link between sampling and DNA/RNA extraction, e.g. to concentrate the target cells. In validation, it is important to apply relevant samples concerning both sample matrix, cell type and sampling method. For pathogen testing of water, this may include applying clean water as well as water with different amounts of humic substances [44]. Recovery is the most important parameter in relation to sample treatment.

Validation of a DNA/RNA extraction method

The bulk of experiments in validation of DNA/RNA extraction methods may be performed applying relevant matrices spiked with known numbers of target cells (e.g. microorganism or human cells). This approach makes it possible to quantify recovery, precision and matrix effects (Table 1). In forensics, for example, cigarette filter paper from a certain brand may be spiked with a certain volume of saliva to investigate matrix effects. This may be complemented with a set of smoked cigarettes of different brands (i.e. real samples) to pick up any other matrix effects.

Recovery is a key parameter—how much of the available target DNA is successfully recovered by the method? Recovery may be investigated by comparing the amount of target cells or DNA/RNA added to the sample with the amount retrieved after extraction, or be calculated as a ratio against a reference method. Any variation linked to the technical setup and to individuals performing the pipetting should be investigated, making both repeatability and intermediate precision important. Matrix effects may partly be evaluated while performing the DNA/RNA extraction. E.g. is the matrix compatible with the reagents and instruments used for extraction? Some matrices may for example clog pipette tips and hence disable the method. Matrices may also interfere with the downstream analysis, e.g. if the extraction method does not remove PCR inhibitors in a satisfactory way. This may be analysed by spiking generated extracts (free of target) with a certain

amount of pure DNA and investigate if the extracts cause impaired PCR amplification.

Validation of a PCR or RT-PCR method

PCR inhibition is a main limiting factor in PCR diagnostics and should be carefully studied in method validation. By spiking PCRs with relevant, homogenised matrices and adding DNA of high quality, PCR inhibition effects may be determined in a straightforward and reproducible way. This approach enables quantification of matrix effects as well as ensures similar effects over time. One challenge in validation is to choose appropriate reference materials that together give a broad range of relevant inhibitory effects. In forensic DNA analysis, manufacturers often validate the inhibitor tolerance of their DNA profiling systems by applying purified hematin as a model for blood and humic acid as a model for soil [45, 46]. An alternative strategy is to prepare casework-like reference materials containing solutions of different matrices, such as cigarettes, chewing gum, and soil, giving a more complex content [47]. In RNA analysis, the RT step must be included in inhibition testing. The efficiency of the RT, generating complementary DNA (cDNA), is generally not measured directly due to a lack of suitable methods. Instead, PCR/qPCR is applied for the measurements. RT generally adds more to the variation between measurements than the PCR step [15].

Summary

In this paper, we present guidelines for single-laboratory validation of methods applied in the PCR analysis chain. The specific focus is analysis of DNA/RNA in sectors such as food safety, forensics, bioterrorism preparedness and veterinary medicine, where the target is often in low levels and mixed with high amounts of complex matrices. These guidelines serve to help laboratories to ensure the performance of their new or modified methods using relevant sample matrices. The choice of matrices to test during validation is of great importance. Relevant matrices should be chosen based on for example which sample types that are expected for a certain target in routine analysis. The guideline may be applied in a crisis situation, such as a foodborne outbreak, requiring urgent analysis of new sample types. In that case, there is little time to perform and evaluate validation experiments, meaning that a strategy for method validation should be present beforehand. By applying a modular approach to validation the methods can be used with more flexibility and the validation studies can be made less laborious. The compatibility between the existing workflow and the new method is verified by applying previously validated methods in the validation study.

References

1. Rådström P, Knutsson R, Wolffs P, Lövenklev M, Löfström C (2004) Pre-PCR processing: strategies to generate PCR-compatible samples. Mol Biotechnol 26:133–146
2. Alaeddini R (2011) Forensic implications of PCR inhibition—a review. Forensic Sci Int Genet 6:297–305
3. Magnusson B, Örnemark U (eds) (2014) Eurachem Guide: the fitness for purpose of analytical methods—a laboratory guide to method validation and related topics, 2nd edn. ISBN 978-91-87461-59-0. http://www.eurachem.org. Accessed 15 Sept 2017
4. ISO/IEC 17025:2005 General requirements for the competence of testing and calibration laboratories. International Organization for Standardization, Geneva
5. Mazzara M, Savini C, Charles Delobel C, Broll H, Damant A, Paoletti C, Van Den Eede G (2008) Definition of minimum performance requirements for analytical methods of GMO testing European network of GMO laboratories. Publication office of the European Union - OPOCE, Luxembourg
6. Broeders S, Huber I, Grohmann L, Berben G, Taverniers I, Mazzara M, Roosens N, Morisset D (2014) Guidelines for validation of qualitative real-time PCR methods. Trends Food Sci Technol 37:115–126
7. Hubner P, Waiblinger HU, Pietsch K, Brodmann P (2001) Validation of PCR methods for quantitation of genetically modified plants in food. J AOAC Int 84:1855–1864
8. SWGDAM (2016) Scientific Working Group on DNA Analysis Methods: Validation Guidelines for DNA Analysis Methods. Retrieved from https://docs.wixstatic.com/ugd/4344b0_813b241e8944497e99b9c45b163b76bd.pdf. 15 Sept 2017
9. ENFSI (2010) Recommended minimum criteria for the validation of various aspects of the DNA profiling process, European Network of Forensic Science Institutes, DNA working group. Retrieved from http://enfsi.eu/wp-content/uploads/2016/09/minimum_validation_guidelines_in_dna_profiling_-_v2010_0.pdf. 15 Sept 2017
10. Hedman J, Rådström P (2013) Overcoming inhibition in real-time diagnostic PCR. Methods Mol Biol 943:17–48
11. Polaczyk AL, Narayanan J, Cromeans TL, Hahn D, Roberts JM, Amburgey JE, Hill VR (2008) Ultrafiltration-based techniques for rapid and simultaneous concentration of multiple microbe classes from 100-L tap water samples. J Microbiol Methods 73:92–99
12. Elliott K, Hill DS, Lambert C, Burroughes TR, Gill P (2003) Use of laser microdissection greatly improves the recovery of DNA from sperm on microscope slides. Forensic Sci Int 137:28–36
13. EC (2005) Commission Regulation (EC) No 2073/2005 of 15 November 2005 on microbiological criteria for foodstuffs (Text with EEA relevance), pp 1–26. Retrieved from: http://eur-lex.europa.eu/eli/reg/2005/2073/oj. 15 Sept 2017
14. Miller DN, Bryant JE, Madsen EL, Ghiorse WC (1999) Evaluation and optimization of DNA extraction and purification

procedures for soil and sediment samples. Appl Environ Microbiol 65:4715–4724

15. Ståhlberg A, Håkansson J, Xian X, Semb H, Kubista M (2004) Properties of the reverse transcription reaction in mRNA quantification. Clin Chem 50:509–515

16. Ståhlberg A, Kubista M, Pfaffl M (2004) Comparison of reverse transcriptases in gene expression analysis. Clin Chem 50:1678–1680

17. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611–622

18. Huggett JF, Foy CA, Benes V, Emslie K, Garson JA, Haynes R, Hellemans J, Kubista M, Mueller RD, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT, Bustin SA (2013) The digital MIQE guidelines: minimum information for publication of quantitative digital PCR experiments. Clin Chem 59:892–902

19. Holst-Jensen A, Berdal KG (2004) The modular analytical procedure and validation approach and the units of measurement for genetically modified materials in foods and feeds. J AOAC Int 87:927–936

20. Bellocchi G, De Giacomo M, Foti N, Mazzara M, Palmaccio E, Savini C, Di Domenicantonio C, Onori R, Van den Eede G (2010) Testing the interaction between analytical modules: an example with Roundup Ready soybean line GTS 40-3-2. BMC Biotechnol 10:55

21. Lievens A, Jacchia S, Kagkli D, Savini C, Querci M (2016) Measuring digital PCR quality: performance parameters and their optimization. PLoS ONE 11:e0153317

22. Roussel Y, Wilks M, Harris A, Mein C, Tabaqchali S (2005) Evaluation of DNA extraction methods from mouse stomachs for the quantification of H. pylori by real-time PCR. J Microbiol Methods 62:71–81

23. Gryson N, Dewettinck K, Messens K (2007) Influence of cocoa components on the PCR detection of soy lecithin DNA. Eur Food Res Technol 226:247–254

24. International vocabulary of metrology – Basic and general concepts and associated terms (VIM), JCGM 200:2012, www.bipm.org. A previous version is published as ISO/IEC Guide 99:2007, ISO Geneva. Accessed 15 Sept 2017

25. Barwick VJ, Prichard E (eds) (2011) Eurachem guide: terminology in analytical measurement—introduction to VIM 3. ISBN: 978-0-948926-29-7. www.eurachem.org. Accessed 15 Sept 2017

26. Vessman J, Stefan R, van Staden J, Danzer K, Lindner W, Burns D, Fajgelj A, Müller H (2001) Selectivity in analytical chemistry (IUPAC recommendations 2001). Pure Appl Chem 73:1381–1386

27. Brunelle S, LaBudde R, Nelson M, Wehling P (2012) AOAC INTERNATIONAL Methods Committee Guidelines for Validation of Microbiological Methods for Food and Environmental Surfaces. Retrieved from http://www.aoac.org/aoac_prod_imis/AOAC_Docs/StandardsDevelopment/AOAC_Validation_Guidelines_for_Food_Microbiology-Prepub_version.pdf. 15 Sept 2017

28. Kline MC, Butts ELR, Hill CR, Coble MD, Duewer DL, Butler JM (2011) The new Standard Reference Material1 2391c: PCR-based DNA profiling standard. For Sci Int Genet Suppl Ser 3:e355–e356

29. Morales SE, Holben WE (2009) Empirical testing of 16S rRNA gene PCR primer pairs reveals variance in target specificity and efficacy not suggested by in silico analysis. Appl Environ Microbiol 75:2677–2683

30. ISO 22118:2011 Microbiology of food and animal feeding stuffs—Polymerase chain reaction (PCR) for the detection and quantification of foodborne pathogens—Performance characteristics. International Organization for Standardization, Geneva

31. Sidstedt M, Romsos EL, Hedell R, Ansell R, Steffen CR, Vallone PM, Rådström P, Hedman J (2017) Accurate digital polymerase chain reaction quantification of challenging samples applying inhibitor-tolerant DNA polymerases. Anal Chem 89:1642–1649

32. Kline MC, Duewer DL, Travis JC, Smith MV, Redman JW, Vallone PM, Decker AE, Butler JM (2009) Production and certification of NIST Standard Reference Material 2372 Human DNA Quantitation Standard. Anal Bioanal Chem 394:1183–1192

33. Dolezel J, Bartos J, Voglmayr H, Greilhuber J (2003) Nuclear DNA content and genome size of trout and human. Cytometry A 51:127–128; author reply 129

34. ISO 16140:2003 Microbiology of food and animal feeding stuffs—Protocol for the validation of alternative methods. International Organization for Standardization, Geneva

35. Hedman J, Knutsson R, Ansell R, Rådström P, Rasmusson B (2013) Pre-PCR processing in bioterrorism preparedness: improved diagnostic capabilities for laboratory response networks. Biosecur Bioterror 11:S87–S101

36. Sidstedt M, Jansson L, Nilsson E, Noppa L, Forsman M, Rådström P, Hedman J (2015) Humic substances cause fluorescence inhibition in real-time polymerase chain reaction. Anal Biochem 487:30–37

37. Boiso L, Hedman J (2017) Overcoming sodium dodecyl sulfate induced PCR inhibition. Forensic Sci Int Genet 29:e16–e18

38. Keim P, Johansson A, Wagner DM (2007) Molecular epidemiology, evolution, and ecology of Francisella. Ann N Y Acad Sci 1105:30–66

39. Dennis DT, Inglesby TV, Henderson DA, Bartlett JG, Ascher MS, Eitzen E, Fine AD, Friedlander AM, Hauer J, Layton M, Lillibridge SR, McDade JE, Osterholm MT, O'Toole T, Parker G, Perl TM, Russell PK, Tonat K, Working Group on Civilian B (2001) Tularemia as a biological weapon: medical and public health management. JAMA 285:2763–2773

40. Mifflin TE (2007) Setting up a PCR laboratory. CSH Protoc 2007:5–14

41. Butler JM (2009) Debunking some urban legends surrounding validation within the forensic DNA community. Promega Profiles in DNA 9:3–6

42. ISO 10272-1:2017 Microbiology of the food chain—Horizontal method for detection and enumeration of Campylobacter spp.—Part 1: Detection method. International Organization for Standardization, Geneva

43. Forsberg C, Wallmark A, Hedell R, Jansson L, Ansell R, Hedman J (2015) Reference material for comparison of different adhesive tapes for forensic DNA sampling. Forensic Sci Int Genet Suppl Ser 5:454–455

44. Borgmästars E, Jazi MM, Persson S, Jansson L, Rådström P, Simonsson M, Hedman J, Eriksson R (2017) Improved detection of norovirus and hepatitis A virus in surface water by applying pre-PCR processing. Food Environ Virol 9:395–405

45. Green RL, Lagacé RE, Oldroyd NJ, Hennessy LK, Mulero JJ (2013) Developmental validation of the AmpFlSTR NGM SElect PCR Amplification Kit: a next-generation STR multiplex with the SE33 locus. Forensic Sci Int Genet 7:41–51

46. Tucker VC, Hopwood AJ, Sprecher CJ, McLaren RS, Rabbach DR, Ensenberger MG, Thompson JM, Storts DR (2012) Developmental validation of the PowerPlex ESX 16 and PowerPlex ESX 17 Systems. Forensic Sci Int Genet 6:124–131

47. Sanga M, Boiso L, Lindsten H, Rådström P, Ansell R, Hedman J (2015) A panel of PCR-inhibitory reference materials for quality evaluation of multiplex STR analysis kits. Forensic Sci Int Genet Supp Ser 5:e317–e319