Journal of Microbiological Methods 212 (2023) 106795

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Note

A simple method for testing and controlling inhibition in soil and sediment samples for qPCR

Tom Bott [a,b,*], George Shaw [b], Simon Gregory [a,*]

[a] British Geological Survey, Keyworth, Nottingham NG12 5GG, UK
[b] School of Biosciences, Sutton Bonington, University of Nottingham, Loughborough LE12 5RD, UK

ARTICLE INFO

Keywords:
Quantitative-PCR (qPCR)
Inhibition
Method optimisation

ABSTRACT

The presence of polymerase-chain-reaction (PCR) inhibitors in many environmental samples can make reliable and repeatable quantitative-polymerase-chain-reaction (qPCR) analysis difficult without sample dilution. To estimate an optimal sample dilution for qPCR and reduce effects of inhibition, a simple test based on multiple dilution series of samples is presented that avoids the use of internal controls and standards reducing complexity and cost.

Polymerase-chain-reaction (PCR) inhibition occurs when a substance interferes with DNA and polymerase interaction (Wilson, 1997). Inhibitors, for example clays and humic substances, bind strongly to DNA stopping the polymerase from completing amplification, or they can denature the polymerase. As a result, inhibition is a potential problem with DNA extracts from environmental samples such as sediments and soils. With quantitative-polymerase-chain-reaction (qPCR), the presence of inhibitors can impact the accuracy and repeatability of assays through the introduction of stochastic error. Diluting samples offers a potential solution but only where template concentration can be kept above the limit of quantification (LoQ).

Accurate estimates of gene copy number, using qPCR, rely on a doubling of the number of amplicons every cycle. The percentage efficiency (E) of the qPCR assay can be calculated using the slope of the standards' log-linear trendline, plotted using the $\log_{10}$ concentration against the cycle threshold, C_t) value or quantification cycle (C_q), as in Eq. (1). The C_t value is the fractional amplification cycle where the assay's fluorescence, proportional to the concentration of dsDNA, is greater-than the background. A well-designed and optimised assay will have an E between 90 and 110% and the $R^2 \geq 0.95$ (Bustin et al., 2009). If inhibition stops the doubling of the amplicon every cycle, it renders the estimation of sample copy-number from comparison with standards, through a log-linear trendline, unreliable.

qPCR efficiency calculation

$$E = 10^{-\frac{1}{Slope}} - 1 \quad (1)$$

where E = Efficiency and the slope is derived from the log-linear trendline plotted from the standards (see example: Fig. 2).

Many environmental laboratories using qPCR likely face issues with inhibition, but no simple methodology has been published to help overcome the problem. Previous work to control and understand inhibition has included an internal recovery standard. This could take the form of a multiplex reaction, i.e. with two targets and amplicons in one reaction, for example the SPUD assay (Nolan et al., 2006), which adds cost and complexity. The PREXCEL-Q (P-Q) program described by Gallup and Ackermann (2008) sought to improve the reproducibility and uniformity of qPCR assay design, unfortunately, the program has not been maintained and is unavailable for use. Recently, a probabilistic model of interpreting standards and low-concentration samples has been published but this does not seek to control or predict inhibition during qPCR (Schmidt et al., 2023).

Another common approach is to dilute samples to the point where inhibition is no longer affecting the assay. Reducing sample concentration in qPCR can lessen the impact of inhibitors during qPCR (Schneider et al., 2009; Wang et al., 2017). However, diluting samples risks the gene of interest being pushed to an undetectable level. To compound this problem, inhibition can affect genes with low copy numbers to a greater extent (Lindberg et al., 2007).

Reactions with either too little, or too much, target amplicon can exhibit noise or scatter in their reported C_t values. This is generated through error both during setup, for example pipetting, and during the qPCR assay. At high concentrations, inhibition will increase the variation in reported C_t values. In the extreme, high concentrations can

* Corresponding authors at: British Geological Survey, Keyworth, Nottingham NG12 5GG, UK.
E-mail addresses: tom.bott@nottingham.ac.uk (T. Bott), simongr@bgs.ac.uk (S. Gregory).

https://doi.org/10.1016/j.mimet.2023.106795
Received 30 May 2023; Received in revised form 20 July 2023; Accepted 2 August 2023
Available online 3 August 2023

confuse baseline estimation and lead to imperfect amplification efficiency. With low concentrations, limited target also leads to imperfect efficiency, especially at the start of the assay; as the assay progresses samples will reach the quantification threshold earlier if they amplified well in the early cycles, compared to sample replicates that initially amplified poorly. Cumulatively, this results in a sigmoidal curve when C_t is plotted against starting log copy-number as the scatter in reported C_t values leads to a flattening in the expected log-linear trendline at the extremes of concentration. Here, it is proposed that a dilution series is used with difficult starting samples in a preliminary qPCR assay; Ct is plotted against log-dilution and used to predict a log-linear trendline; this is then used to estimate a usable dilution range for that sample set, avoiding both under and over-dilution of samples.

A subset of the samples is diluted with ultra-pure water down to 1:10,000 in 12 increments (Fig. 1). The diluted sample series are used in a qPCR assay of the 16S rRNA gene before the main analysis, which may include other qPCR assays targeting different genes. Samples can be run in triplicate, for each dilution, if a lot of variation is anticipated. Average C_t values for each dilution are generated using the C_t values from all samples and replicates. A complete set of standards for the 16S rRNA gene assay is also run (10^3 to 10^8 copies μl^{-1}), plus a no template control. The samples' C_t values are grouped (Fig. 1) and an E and R^2 value is produced for each grouping. Each grouping covers a 100-fold dilution range. These E and R^2 values are compared to those of the standards. The grouping with an E value closest to that of the standards is considered usable. qPCR assays of the whole sample set can be completed after dilution to highest concentration in that grouping.

An example of this approach is from a gravel rich, clayey-sand soil containing low organic matter. Samples were from a field site adjacent to a river which flooded the site seasonally. Soil DNA extractions were completed using FastDNA™ Spin kit (MP Biomedicals, USA-CA), using the manufacturer's protocol. Dilutions and qPCR were completed as described above. The qPCR assay, for the 16S rRNA gene, used a 341f (5′-CCTACGGGAGGCAGCAG-3′) and 543r (5'-ATTACCGCGGCTGCTGG-3′) primer set (Juck et al., 2000; Nossa et al., 2010), to produce a 202 bp product. Thermocycling conditions consisted of an initial denaturation at 98 °C for 120 s, followed by 35 cycles of denaturation at 98 °C for 10 s, annealing at 64 °C for 30 s and extension at 72 °C for 10 s, before a final extension stage at 72 °C for 120 s. SSoAdvanced™ Universal Inhibitor-Tolerant SYBR® Green Supermix (Bio-Rad, USA-CA) was used.

The average C_t value for each dilution was used for analysis. When plotted, C_t values plateau at the higher concentrations (Fig. 2). The fitted trendline to the group starting as undiluted has a clearly shallower slope than the other two trendlines provided ($E = 191.74\%$, $R^2 = 0.86$). The group starting with a 10-fold dilution ($E = 103.37\%$, $R^2 = 0.99$) was the closest to the standard series ($E = 101.68\%$, $R^2 = 0.99$). Using a 10-fold dilution with the samples will allow the samples to amplify in a similar manner to the standards, increasing the reliability and repeatability of results.

From the analysis of the example sample set, an understanding of the limit of quantification can also be gained. The optimum C_t change between a ten-fold dilution should be 3.3, between x1000 and x10000 dilution the C_t change is 2.51 (C_t values of 20.37 and 22.88 respectively) suggesting a flattening of the trendline and the end of linear amplification. Using the 16S rRNA gene standards, at x10 dilution the samples had a mean 3.10×10^6 copies μl^{-1} (1 SD $\pm$ 1.09×10^6) of the 16S rRNA gene. Therefore, when using qPCR on this sample set, diluted at x10, where an assay estimates an uncorrected copy number below 3×10^3 copies μl^{-1}, i.e. around three orders of magnitude less than the copy number at a x10 dilution, the result should be used with caution as they are likely outside of LoQ suggesting that repeatability and accuracy may be limited, and that overestimation of copy-number may occur due to erroneously low C_t values and increasing scatter in reported values. That said, qPCR will still detect those genes, if present, and may predict which samples have greater copy numbers of the target gene. When working with genes that are believed to have a low copy number, between 10^2 and 10^3 copies, this technique could be repeated with that specific assay, to estimate its quantifiable range with the sample set.

The soil samples discussed were used in further qPCR assays using the sample diluted x10. To estimate relative copy number of the target gene in the bacterial population, qPCR on all samples was completed for both 16S rRNA gene and the genes of interest. Final data was presented as relative abundance, derived from the division of the gene of interest's copy-number by the 16S rRNA gene copy-number, with errors presented as the propagated standard deviation derived from both datasets. Using relative abundance can help overcome issues with differing DNA recovery levels and varying levels of inhibition between samples. This can enable direct comparison between samples by normalising the copy number of the gene of interest against an indicator of recovered and quantifiable DNA, in this case the 16S rRNA gene as a proxy for the total bacterial population.

To reduce time and cost, where multiple field samples from the same site and medium (soil, sediment etc.) are to be analysed, a subset can used. In the example, three of 18 collected samples were used in the inhibition test: these were selected to form a transect across the field with the aim of capturing any soil variability. Selecting samples can also be based upon a knowledge of the sample site, for example selecting samples with suspected inhibition or low DNA recovery.

There are limitations with this approach. Principally, it assumes that inhibition will impact all qPCR assays equally. To improve the likelihood that the inhibition test accurately reflects inhibition in other qPCR assays a 16S rRNA gene target of a similar amplicon length could be used. Inhibition becomes more likely with increasing amplicon length due to the increasing probability that the DNA polymerase will encounter competitively bound material on the target DNA strand. Therefore, using a 16S rRNA gene target with of a similar length to the target

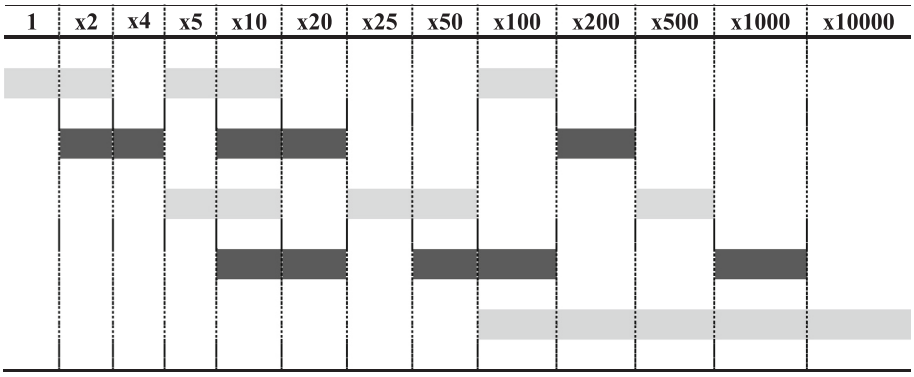

Fig. 1. Dilutions and the groupings used to calculate E values. Dilutions are in the top row. The greyscale horizontal bars indicate five groups of five dilutions which were used for efficiency calculations. These determined the appropriate dilution ranges for field samples.

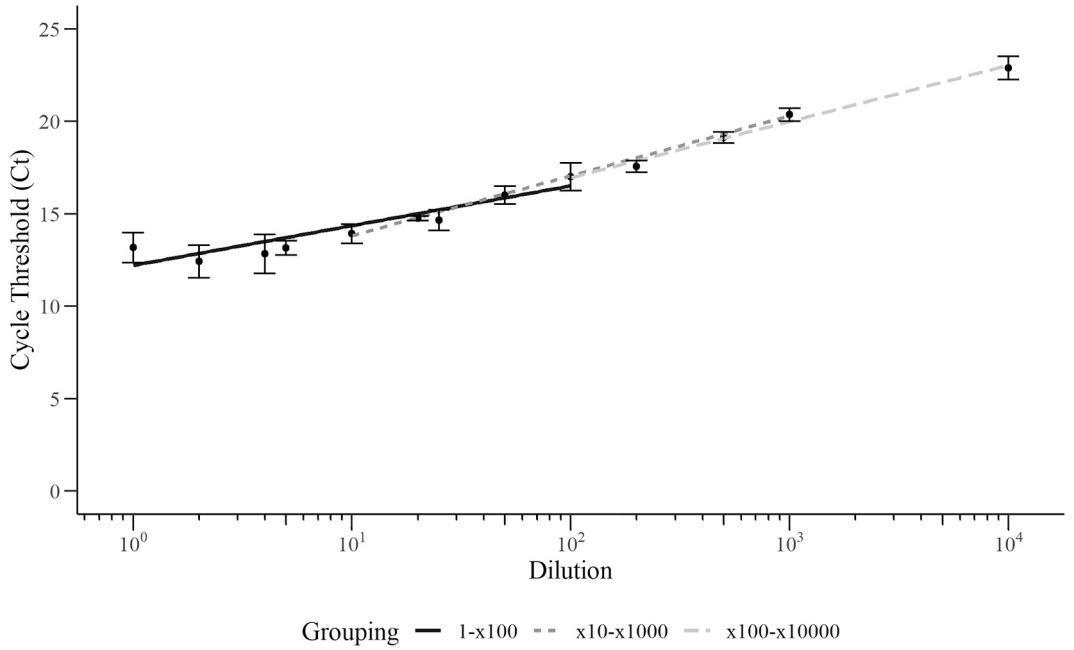

Fig. 2. Example plot of C_t against dilution with three example trendlines. Trendlines are generated using groups of five dilutions; here only three dilution groups are shown for clarity. The legend refers to the range of dilutions used in the series to produce the trendline. This sample is showing inhibition: the undiluted assay does not fit the expected trend with a higher C_t value (13.17) than the x10 dilution (12.42). Mean C_t values are plotted from three soil samples, from the same site. Error bars are ±1 SD.

amplicons in latter qPCR may increase the reliability of dilution estimates. The 16S rRNA gene is principally used in this technique as it should be the most abundant bacterial gene maximising the likelihood of amplification in difficult samples however the availability of different primers for this gene and subsequently different amplicon lengths is also a benefit. Furthermore, the use of 16S rRNA gene and this method can highlight issues with quantifiable DNA recovery which quantification and quality measurements might struggle to achieve by demonstrating the concentration of DNA that is detectable in PCR.

Previous approaches have also used an assay and dilution steps to understand PCR inhibition (Schneider et al., 2009; Wang et al., 2017). These authors used an internal standard to assess how inhibition and subsequently amplification changes with dilution. Wang et al. (2017) spiked DNA extracts with an internal standard and diluted the extracts, then compared the C_t values against the standard when it was amplified in pure water and diluted along the same gradient. Differences in C_t value highlighted the presence of inhibition, in some samples the technique could highlight a linear range of usable dilutions. Their approach enables a clear and statistically robust measure of inhibition, enabling a suitable dilution to be found and could be used to understand DNA recovery losses during sample cleanup. However, the approach is considerably more complex and requires greater time and cost than the method proposed here. Furthermore, the method proposed here can be applied to any qPCR assay.

Cumulatively, this simple approach enables a rapid understanding of inhibition within environmental samples. It provides an estimation for a suitable dilution to prepare samples for qPCR and a potential quantification limit for low copy number targets. This method is particularly useful for soil and sediment samples, where high DNA yields can be expected alongside the presence of inhibitors, but it also useful to identify suitable dilutions where DNA recovery is poor. The small additional cost is paid off in improved reliability and confidence in results and through the avoidance of over- or under- diluting the samples.

Funding sources

This work was supported by BBSRC and NERC through a Soils Training and Research Studentships (STARS) grant (no. NE/M009106/1). STARS is a consortium consisting of Bangor University, the British Geological Survey, the Centre for Ecology and Hydrology, Cranfield University, the James Hutton Institute, Lancaster University, Rothamsted Research and the University of Nottingham.

Ethics statement

This work was completed in line with Elsevier ethics policy and guidelines and required no approval from an ethics committee. No human participants were involved in this research. No animals were used in this research.

CRediT authorship contribution statement

Tom Bott: Conceptualization, Formal analysis, Investigation, Writing – original draft. **George Shaw:** Writing – review & editing, Supervision, Funding acquisition. **Simon Gregory:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Thanks to Jessica Mackie for her useful comments on a draft version of this manuscript.

Simon Gregory published this manuscript with the permission of the Executive Director of the British Geological Survey.

References

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. 55, 611–622.

Gallup, J.M., Ackermann, M.R., 2008. The 'PREXCEL-Q Method' for qPCR. Int. J. Biomed. Sci. 4, 273–293.

Juck, D., Charles, T., Whyte, L.G., Greer, C.W., 2000. Polyphasic microbial community analysis of petroleum hydrocarbon-contaminated soils from two northern Canadian communities. FEMS Microbiol. Ecol. 33, 241–249.

Lindberg, E., Albrechtsen, H.-J., Jacobsen, C.S., 2007. Inhibition of real-time PCR in DNA extracts from aquifer sediment. Geomicrobiol J. 24, 343–352.

Nolan, T., Hands, R.E., Ogunkolade, W., Bustin, S.A., 2006. SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. Anal. Biochem. 351, 308–310.

Nossa, C.W., Oberdorf, W.E., Yang, L., Aas, J.A., Paster, B.J., Desantis, T.Z., Brodie, E.L., Malamud, D., Poles, M.A., Pei, Z., 2010. Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome. World J. Gastroenterol. 16, 4135–4144.

Schmidt, P.J., Acosta, N., Chik, A.H.S., D'Aoust, P.M., Delatolla, R., Dhiyebi, H.A., Glier, M.B., Hubert, C.R.J., Kopetzky, J., Mangat, C.S., Pang, X.-L., Peterson, S.W., Prystajecky, N., Qiu, Y., Servos, M.R., Emelko, M.B., 2023. Realizing the value in "non-standard" parts of the qPCR standard curve by integrating fundamentals of quantitative microbiology. Front. Microbiol. 14.

Schneider, S., Enkerli, J., Widmer, F., 2009. A generally applicable assay for the quantification of inhibitory effects on PCR. J. Microbiol. Methods 78, 351–353.

Wang, H., Qi, J., Xiao, D., Wang, Z., Tian, K., 2017. A re-evaluation of dilution for eliminating PCR inhibition in soil DNA samples. Soil Biol. Biochem. 106, 109–118.

Wilson, I.G., 1997. Inhibition and facilitation of nucleic acid amplification. Appl. Environ. Microbiol. 63, 3741–3751.