# tidyqpcr: Quantitative PCR analysis in the tidyverse.

**Edward W. J. Wallace** [1] **and Samuel J. Haynes** [1]

**1** Institute for Cell Biology and SynthSys, School of Biological Sciences, The University of Edinburgh,

## Summary

Quantitative polymerase chain reaction (qPCR) is a fundamental technique in molecular biology to detect and quantify DNA and RNA. Here we present the tidyqpcr software package for user-friendly qPCR analysis using the tidyverse suite of R packages. tidyqpcr offers a consistent user interface and structure for qPCR analysis, within the tidyverse paradigm of spreadsheet-like rectangular data frames and generic functions that build up complex analyses in a series of simple steps. tidyqpcr focuses on experimental design in microwell plates, and relative quantification using changes in quantification cycle ($\Delta Cq$). Overall, tidyqpcr empowers scientists to conduct reproducible, flexible, and best-practice compliant quantitative PCR analysis.

## Statement of need

Quantitative PCR is among the most common techniques in biological and biomedical research, used for the quantification of DNA and RNA. There is a critical need for rigorous analysis and reporting of qPCR experiments, codified in the minimum information for publication of quantitative real-time PCR experiments (MIQE) guidelines (Bustin et al., 2009). Yet it is common for qPCR to be analysed either by closed-source software supplied by the manufacturers of PCR machines, or by highly variable, in-house analysis scripts that have not been peer-reviewed.

Our package, tidyqpcr, addresses the need for a qPCR analysis package that fully integrates with the user-friendly tidyverse, encourages the use of MIQE best-practice compliant experimental design, and provides detailed example analysis pipelines as R vignettes. Following the tidy data paradigm integrates tidyqpcr into the wider collection of data analysis packages provided by the tidyverse, while being accessible to novice users of R. Other open-source libraries for qPCR analysis are available with distinct aims. HTqPCR (Dvinge & Bertone, 2009), ReadqPCR/NormqPCR (Perkins et al., 2012), and qpcR (Spiess, 2018) have similar and in some respects greater functionality than tidyqpcr, but follow object oriented approaches with specialised data objects. By contrast, pcr (Ahmed & Kim, 2018) aligns most closely with tidyqpcr but its function inputs are not tidy data frames. Although alternatives are available, tidyqpcr's aims and approach are distinct: to improve the quality of qPCR experiments from plate design to analysis, by exposing all data in a consistent tidy format that integrates with the tidyverse.

tidyqpcr aims to be:

- Empowering: tidyqpcr combines a free, open-source qPCR analysis R package with online teaching materials.
- Reproducible: tidyqpcr scripts produce paper-ready figures straight from raw data with identical results across computers.
- Flexible: tidyqpcr follows the 'tidy' data paradigm to ensure scalability and adaptability.

- Best-practice compliant: tidyqpcr encourages standardised, reliable experimental design by prioritising MIQE-compliant best practices.

tidyqpcr can be used to analyse qPCR data from any nucleic acid source - DNA for qPCR or ChIP-qPCR, RNA for RT-qPCR. Currently tidyqpcr has functions that explicitly support relative quantification by the $\Delta Cq$ method, but not yet absolute quantification.

tidyqpcr's current features allow users to:

- use a single data type for analysis as every object is a tibble / data frame.
- lay out and display 96/384-well plates for easy experimental setup (`label_plate_rowcol`, `create_blank_plate`, …).
- consistently describe samples and target amplicons with reserved variable names (`sample_id`, `target_id`).
- flexibly assign metadata to samples for visualisation with ggplot2 (see vignettes).
- read in quantification cycle (Cq) and raw data from Roche LightCycler machines with single-channel fluorescence (`read_lightcycler_1colour_cq`, `read_lightcycler_1colour_raw`).
- calibrate primer sets including estimating efficiencies and visualization of curves (`calculate_efficiency`).
- visualize amplification and melt curves (`calculate_drdt_plate`)
- perform normalisation and relative quantification to one or more reference targets by the $\Delta Cq$ method (`calculate_normcq`, `calculate_deltacq_bysampleid`).
- estimate differential expression across multiple samples by the $\Delta \Delta Cq$ method (`calculate_deltadeltacq_bytargetid`).
- accelerate further downstream analysis and visualization by writing tidy data frames that are fully compatible with the tidyverse suite.

We have conducted a series of user interviews to improve tidyqpcr's capabilities and documentation. The ease-of-use and documentation of tidyqpcr will enable efficient best-practice analysis of qPCR data by both novice and experienced programmers.

## Acknowledgements

## References

Ahmed, M., & Kim, D. R. (2018). pcr: An R package for quality assessment, analysis and testing of qPCR data. *PeerJ*, *6*, e4473. https://doi.org/10.7717/peerj.4473

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. *Clinical Chemistry*, *55*(4), 611–622. https://doi.org/10.1373/clinchem.2008.112797

Dvinge, H., & Bertone, P. (2009). HTqPCR: High-throughput analysis and visualization of quantitative real-time PCR data in R. *Bioinformatics*, *25*(24), 3325–3326. https://doi.org/10.1093/bioinformatics/btp578

Perkins, J. R., Dawes, J. M., McMahon, S. B., Bennett, D. L. H., Orengo, C., & Kohl, M. (2012). ReadqPCR and NormqPCR: R packages for the reading, quality checking and normalisation of RT-qPCR quantification cycle (Cq) data. *BMC Genomics*, *13*, 1–8. https://doi.org/10.1186/1471-2164-13-296

Spiess, A.-N. (2018). *qpcR: Modelling and analysis of real-time PCR data.* https://CRAN.R-project.org/package=qpcR