dpcR: web server and R package for analysis of dPCR experiments

Michał Burdukiewicz[1]*, Piotr Sobczyk[2], Paweł Mackiewicz[1], Peter Schierack[3], Stefan Rödiger[3]
*michalburdukiewicz@gmail.com

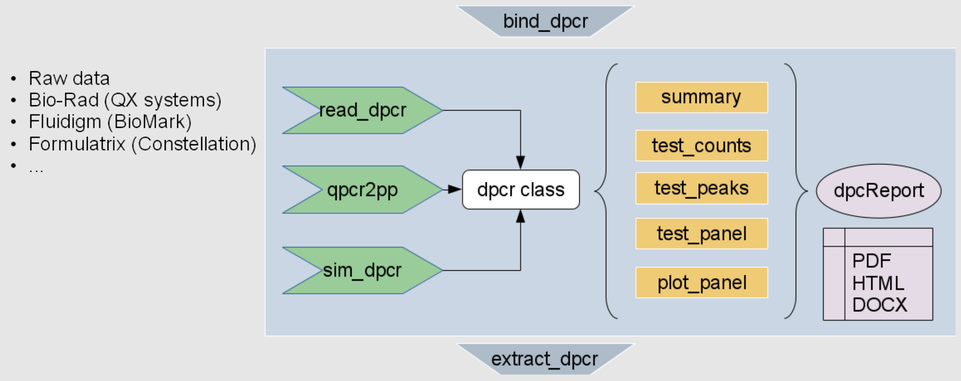
[1]University of Wrocław, Department of Genomics
[2]Wrocław University of Technology, Faculty of Pure and Applied Mathematics
[3]Brandenburg University of Technology Cottbus-Senftenberg, Institute of Biotechnology

Introduction

dPCR reaction consists of multiple amplifications occurring in numerous small partitions. The result of dPCR is a binary vector describing states of partitions (positive in case of detected amplification, negative otherwise). This data is further used to estimate the main parameter, $\boldsymbol{\lambda}$, which may be interpreted as the mean number of template molecules per partition.

We created *dpcR*, an open source **R** package for reproducible analysis of dPCR data, fully compatible with dMIQE requirements.

dpcR workflow



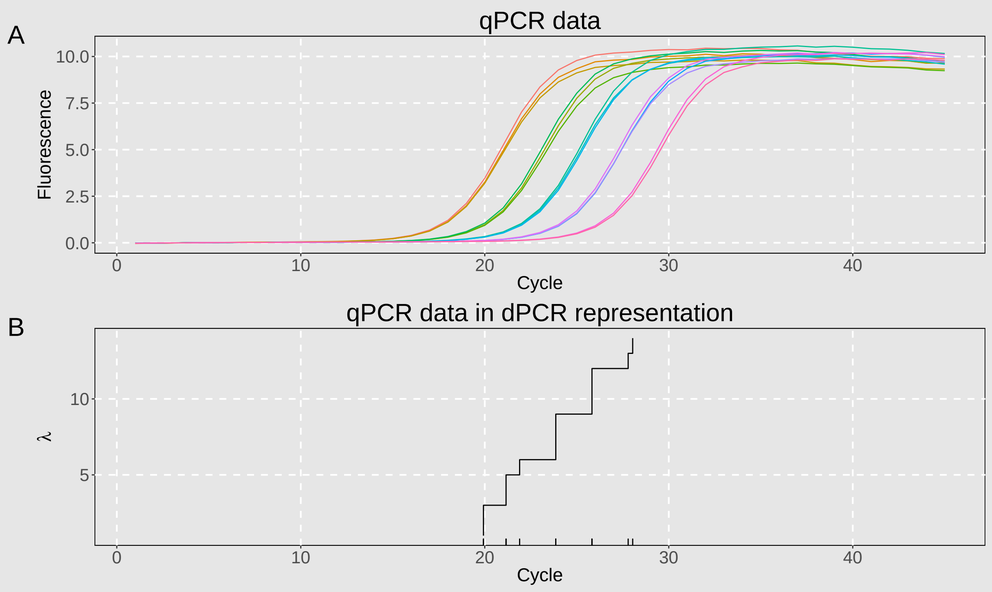
The workflow diagram shows main functions available at the each step of a dPCR data analysis.

Data import

Import functions limit availability of the package by determining which datasets can be easily processed using the provided framework. Since the RDML format for dPCR is not yet established, we wrote function *read_dpcr* streamlining data import from several systems produced by Bio-Rad, Fluidigm and Formulatrix.

To cover experimental or not yet included systems, we created a "raw data" format (see Supplementary Files for description). The user can manually arrange his data in this format and import it to the *dpcR* package. Such input files can be created in a spreadsheet program or a text editor.

Integration of qPCR data



The dPCR methodology may be used to analyze qPCR data (Mojtahedi et al., 2014). Quantification points (Cq) are computed using the real-time measurements of several amplification curves (A). Next, the Cq values are binarized and treated as the status of partitions effectively converting multiple qPCR experiments into a dPCR (B). This functionality is supported by the *qpcr2pp* function.

Availability

dpcReport web server:
`www.smorfland.uni.wroc.pl/dpcReport`
dpcR download:
`http://cran.r-project.org/package=dpcR`
This research was partially funded by KNOW Consortium.

Bibliography

Bhat, S., Herrmann, J., Armishaw, P., Corbisier, P., and Emslie, K. R. (2009). Single molecule detection in nanofluidic digital array enables accurate measurement of DNA copy number. *Analytical and Bioanalytical Chemistry*, 394(2):457–467.

Burdukiewicz, M., Rödiger, S., Sobczyk, P., Schierack, P., and Mackiewicz, P. (2016). Methods of comparing digital PCR experiments (accepted). *Biomolecular Detection and Quantification*, 28(NN):NN–NN.

Dorazio, R. M. and Hunter, M. E. (2015). Statistical Models for the Analysis and Design of Digital Polymerase Chain Reaction (dPCR) Experiments. *Analytical Chemistry*, 87(21):10886–10893.

Dube, S., Qin, J., and Ramakrishnan, R. (2008). Mathematical analysis of copy number variation in a DNA sample using digital PCR on a nanofluidic device. *PloS one*, 3(8):e2876.

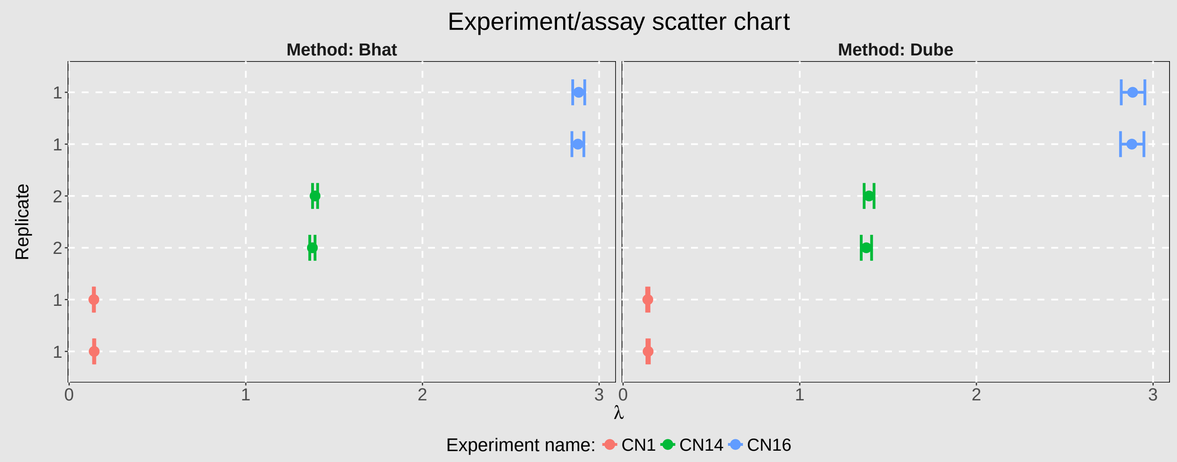
Mojtahedi, M., Fouquierd'Hérouël, A., and Huang, S. (2014). Direct elicitation of template concentration from quantification cycle (Cq) distributions in digital PCR. *Nucleic Acids Research*, 42(16):e126–e126.

Case study: analysis of dPCR data

We present the functionalities of *dpcR* using the previously published data set consisting of three experiments repeated twice (Dorazio and Hunter, 2015).

Experiment	Assay	Positive partitions	Total partitions
CN1	MYC	1687	12765
CN1	MYC	1787	13681
CN14	MYC	9913	13259
CN14	MYC	11117	14794
CN16	MYC	13919	14747
...	...	...	...

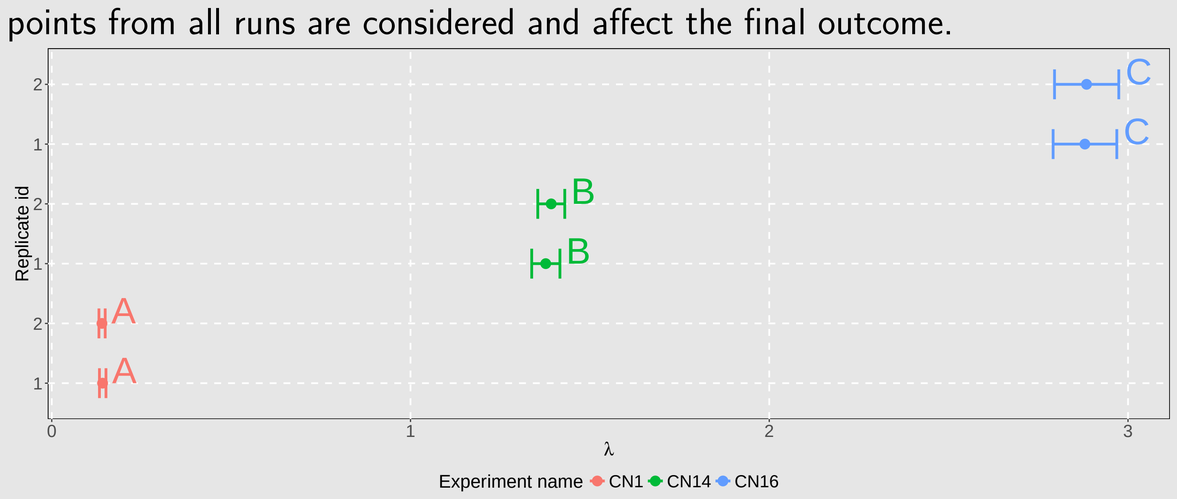
Calculation of the uncertainty



To determine the uncertainty of the estimated $\boldsymbol{\lambda}$ *dpcR* employs two previously published peer-reviewed methods Dube et al. (2008); Bhat et al. (2009).

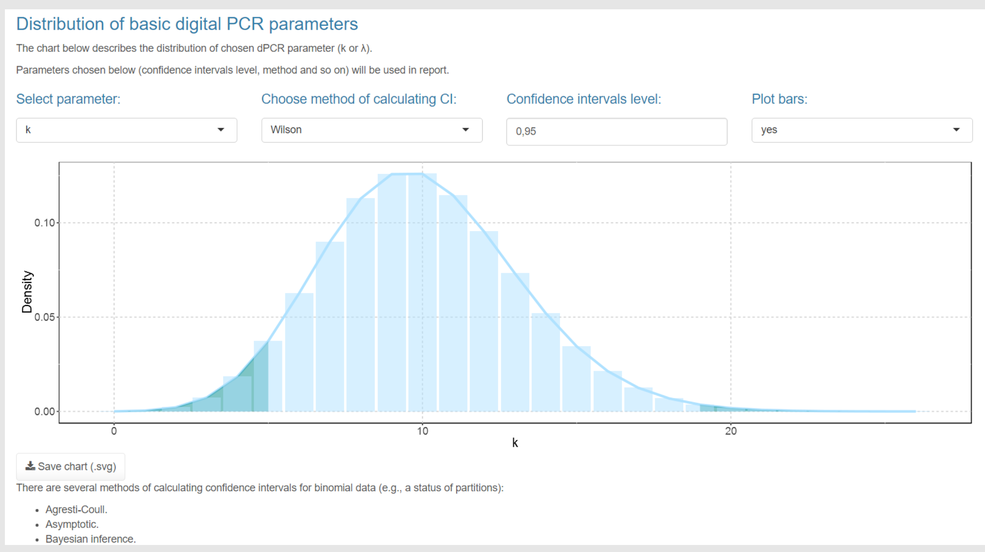
Comparison of individual runs

The *dpcR* package covers peer-reviewed methods of comparing results of dPCR experiments. Here, by the comparison we understand a procedure, where all data points from all runs are considered and affect the final outcome.



Two methods, GLM and MRT, conduct such analysis on the run level by comparing individual runs against each other (Burdukiewicz et al., 2016). Additionally, we also implemented a method for individual dPCR experiments (not runs) (Dorazio and Hunter, 2015). All methods automatically assigns experiments to groups based on their $\boldsymbol{\lambda}$ values (A, B and C on the figure above).

dpcReport



The majority of functions described above is also accessible using the web server *dpcReport* and does not require any experience with **R**. To preserve the reproducible research principle *dpcReport* generates highly customizable reports, which may even include **R** code necessary to fully recreate the analysis.

Funding

This research was partially funded by KNOW Consortium, InnoProfile-Transfer 03IPT611X (BMBF) and KMU-innovativ-16 031B0098B (BMBF) projects.