dpcReport

A framework for the read-in, analysis, intra/inter assay comparison and report of dPCR experiments

Michał Burdukiewicz[1], Piotr Sobczyk[2], Paweł Mackiewicz[1] and Stefan Rödiger[3]

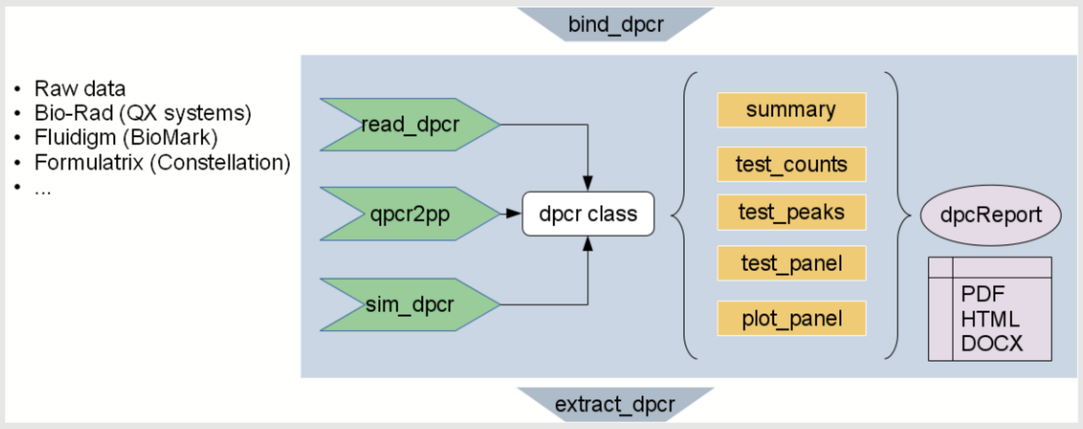
[1]University of Wrocław, Department of Genomics, Poland
[2]Wrocław University of Technology, Department of Mathematics, Poland
[3]Faculty of Natural Sciences, Brandenburg University of Technology Cottbus–Senftenberg, Germany

Introduction

dpcReport is a GUI tool for dPCR data mining and report generation. It is based on dpcR, a versatile open source cross-platform software, which provides functions to process and study dPCR data independent of the hardware. Our software can be used for data analysis and presentation, as a framework for novel technical developments and as reference for statistical methods in dPCR analysis.

We based our framework on the sophisticated statistical computing environment R, so the most fundamental interface is a command-line. To move the hurdle of learning new software from users to developers, we also designed a stand-alone graphical interface, accessible also as the interactive web application.

Workflow



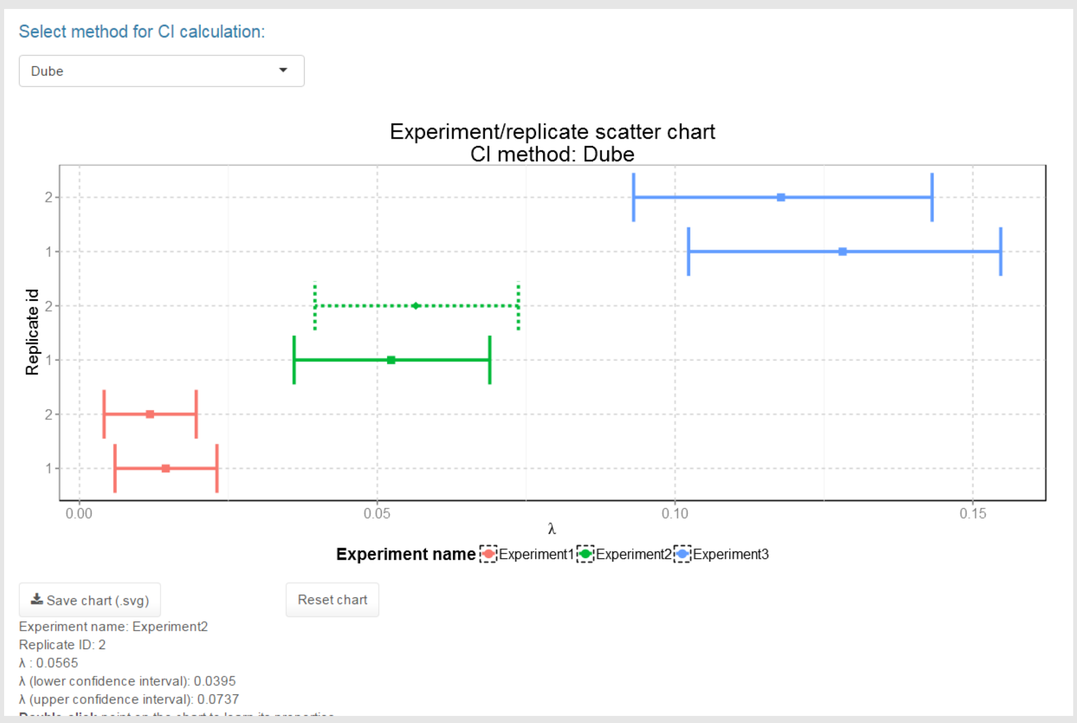
The central concept of the workflow is dpcr object, an abstract representation of dPCR data regardless of its source. It allows integration results obtained from various systems in one framework.

Input data

Vendor	System
Bio-Rad	QX100 & QX200
Fluidigm	BioMark
Formulatrix	Constellation Digital PCR

Results from systems described in the table above may be freely use as the input in the worflow. Moreover, numerous data formats can be processed with the functionality provided by the **R** environment Rödiger et al. (2015).

Data summary



The summaries of input data are available both as interactive tables (allowing filtering and selecting specified runs) and also interactive charts (as presented on the figure above, where replicate 2 of experiment 2 is selected).

Test panel

Array based dPCR experiments provide information about spatial distribution of partitions. Procedures belonging to spatial statistics verify if the status (positive, negative) of partition depends on its location. To address such questions, we implemented a Complete Spatial Randomness test Baddeley and Turner (2005) for dPCR arrays.

Comparison of dPCR experiments

dpcReport uses two methods of comparing dPCR experiments, able to simultaneously compare $\boldsymbol{\lambda}$ values of multiple runs. One of them is based on Generalized Linear Models, where we employ a simplistic model reflecting relationships between variables in dPCR results, given by formula:

$$\log \boldsymbol{Y} = \beta^T \boldsymbol{X}$$

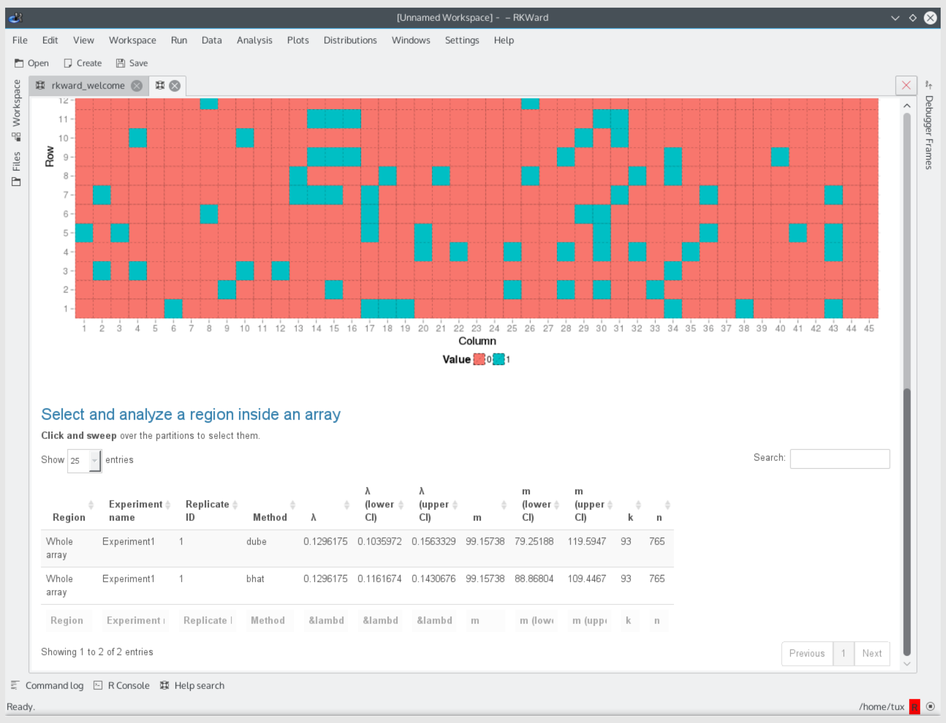
where $\boldsymbol{Y}$ are counts of positive partitions, $\boldsymbol{X}$ are experiments names (categorical data) and $\boldsymbol{\beta}$ are coefficients for every run. Based on the model, estimated means of copies per partitions can be calculated from $\hat{\boldsymbol{\lambda}} = \exp \boldsymbol{\beta}$ and next pairwise compared using multiple t-tests as described elsewhere Bretz et al. (2010).

The second method pairwise compares the $\hat{\boldsymbol{\lambda}}$ from two or more dPCR experiments using the uniformly most powerful (UMP) ratio test (Fay, 2010). We use Benjamini-Hochberg correction for p-values (Benjamini and Hochberg, 1995) to control the family-wise error rate. The UMP ratio test has following null-hypothesis:

$$\boldsymbol{H_0} : \frac{\boldsymbol{\lambda_1}}{\boldsymbol{\lambda_2}} = 1$$

where $\boldsymbol{\lambda_1}$ and $\boldsymbol{\lambda_2}$ are mean numbers of template molecules per partition in two experiments.

Graphical User Interface



The dpcReport may be used as web application, accessible from every modern web browser or an **R** IDE/GUI such as **RKWard** Rödiger et al. (2012) or **RStudio**. Moreover, it may be downloaded and installed as a standalone application.

Summary

The dpcReport is a freeware multi-platform software design to assist researchers in the analysis of dPCR data.

Availability

dpcReport (standalone version for Microsoft Windows ®):

`http://sourceforge.net/projects/dpcreport/`

dpcReport (web server):

`http://www.smorfland.uni.wroc.pl/dpcReport`

dpcR package: `https://github.com/michbur/dpcR`

Bibliography

Baddeley, A. and Turner, R. (2005). spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software*, 12(6):1–42.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300.

Bretz, F., Hothorn, T., and Westfall, P. (2010). *Multiple Comparisons Using R*. Chapman & Hall/CRC Press, Boca Raton, Florida, USA.

Fay, M. (2010). Two-sided exact tests and matching confidence intervals for discrete data. *Proceedings of the National Academy of Sciences of the United States of America*, 2(1):53–58.

Rödiger, S., Burdukiewicz, M., Blagodatskikh, K. A., and Schierack, P. (2015). R as an Environment for the Reproducible Analysis of DNA Amplification Experiments. *The R Journal*, 7(2):127–150.

Rödiger, S., Friedrichsmeier, T., Kapat, P., and Michalke, M. (2012). RKWard: A Comprehensive Graphical User Interface and Integrated Development Environment for Statistical Analysis with R. *Journal of Statistical Software*, 49(9):1–34.

Storer, B. E. and Kim, C. (1990). Exact properties of some exact test statistics for comparing two binomial proportions. *Journal of the American Statistical Association*, 85(409):146–155.