

RDML: real-time PCR data markup language

Discussion forum Tuesday March 27th, 17:20-18:40

1. Introduction

It is currently very cumbersome to share qPCR data between different laboratories, or exchange data between different software packages or analysis tools. The problem is founded in the data collection software packages that, depending on the instrument provider, export information in various file formats (.csv, .txt, .xls), with different layout and data field terminology. A common language and universal format would enable easy exchange of raw annotated data and analysis settings for qPCR experiments. This would make it possible to include qPCR data and analysis settings in scientific papers, enabling both reviewers and readers to re-analyse the data, much like the MIAME guidelines propose for microarray experiments. In principle, the universal data format should contain sufficient information to understand the experimental setup, re-analyse the data and interpret the results. This data format could be part of yet to be established MIARE guidelines (Minimal Information About Real-time Experiments).

2. Discussion

We have drafted a proposal for a universal data format for the exchange of real-time PCR data, termed RDML (Real-time PCR Data Markup Language - <http://medgen.ugent.be/rdml/>). We propose to use an XML-based file type because this format is independent of computer hardware, operating system or available software package, and can be extended in the future to include additional information if required. The aim of this session is to discuss and finalize the RDML specifications. Points to discuss include:

1. What information must be included?
2. Which fields are required and which are optional?
3. Which fields are free text, and which have predefined values?
4. What terminology should be used?
5. How to obtain broad acceptance of this standard?
6. Is there a need for an RDML working group to continue the development of RDML?

Because a fruitful discussion often takes place with a limited number of people closely participating, we would like you to register prior to the RDML session. The results of the discussion will be made available to the general public soon after this symposium.

RDML registration form

for discussion forum on Tuesday March 27th, 17:20-18:40

Name:

Organization:.....

Notes:

.....

Please return the completed registration form to the reception desk!

3. RDML terminology

- abbreviations: UNKN: unknown, NTC: no template control, NAC: no amplification control, STD: standard, REF: reference target, TOI: target of interest
- target: universal term for nucleic acid sequence to be amplified (including but not limited to genes)
- Cq: quantification cycle (Ct, Cp or TOP depending on software), fractional PCR cycle at which fluorescence crosses threshold value above background (to be left determined by the instrument software), should be in the exponential phase of the PCR
- exclusion: indicates that this well should be excluded for data-analysis (e.g. bad replicate, suspicious melt curve)
- quantity: should be provided for STD samples
- experiment: group of one or multiple runs that require to be processed together and from which conclusions can be drawn
- probe sequences: Probe 1 field is intended for single probe detection chemistries (e.g. hydrolysis probe); Probe 2 is intended for two-probe assays (e.g. hybridisation probes).
- commercial assay: to accommodate the use of commercial primer and probe assays where the exact sequences of the primers and probes is not provided
- raw fluorescence: values without background correction
- amplification efficiencies: values between 0.5 and 1.5 (1=100% amplification efficiency)
- priming method: any of the three methods, or a combination of these can be chosen
- Replicates: by definition wells with identical sample and target names are considered PCR replicates, other types of replicates should have differential sample or target names.
- Quantification fluorescence = (background corrected) fluorescence at quantification cycle

4. runRDML

- Run (ID)
 - Run annotation
 - Name
 - ? Description
 - Instrument
 - PCR format {single-well, 48-well, 96-well, 384-well, 32-rotor, 72-rotor, 100-rotor}
 - Data collection software
 - Program name
 - Version
 - Background determination method
 - Ct/Cp detection method {automated threshold and baseline settings, manual threshold and baseline settings, maximum second derivative}
 - PCR program
 - ? Reverse transcriptase step
 - Temperature
 - Duration
 - ? Pre incubation step
 - Temperature
 - Duration
 - Initial denaturation
 - Temperature
 - Duration
 - Cycles
 - Number of cycles
 - Cycle steps (2, 3 or 4)
 - Temperature
 - Duration
 - Final step
 - Temperature
 - Duration
 - ? Melting analysis
 - Start temperature
 - End temperature
 - Temperature interval OR number of steps
 - DataCollectionTemperature
 - Run data
 - + Well
 - Well name OR Well number
 - Sample type {UNKN, NTC, NAC, STD}
 - Sample name
 - + Dye
 - Dye name
 - Target name
 - Quantification cycle
 - ? Quantity
 - Exclusion (default = false)
 - ? Sample specific amplification efficiency
 - ? Raw amplification fluorescence intensities
 - + Fluorescence value (index = PCR cycle number)
 - ? Raw melting curve fluorescence intensities
 - + Fluorescence (index = Temperature)
 - ? Background fluorescence
 - ? Quantification fluorescence
 - Inter-run calibrator (default = false)

5. *experimentRDML*

- Experiment
 - Experiment annotation
 - Name
 - ? Annotation
 - Target annotation
 - + Target name (as used in Run element)
 - ? Official gene symbol
 - ? Entrez gene ID
 - Target type {REF, TOI}
 - Target specific amplification efficiency
 - ? Sequences
 - Forward primer
 - Reverse primer
 - ? Probe 1
 - ? Probe 2
 - ? Amplicon
 - ? Commercial assay
 - Company
 - Order number
 - ? RTprimerDB ID
 - Sample annotation
 - + Sample name (as used in Run element)
 - ? Description
 - ? Template isolation method
 - ? DNase treatment (Y/N)
 - ? cDNA synthesis method
 - Enzyme
 - Priming method {oligo-dT, random, target specific}
 - ? Template quantity
 - ? Template quality
 - Method
 - Result
 - Runs
 - + Run (ID) The contents of this section is described in runRDML