Automation of qPCR Analysis

Matthew Dodson
University of Michigan, Ann Arbor

John Slater
Lawrence Technological University

Paper 2015-3375
March 2015

Abstract

Quantitative real-time polymerase chain reaction, qPCR, is a technique used in gene research. Technological advances have allowed wide access to automated machinery, allowing a multitude of experiments to be conducted in a short time frame. Unfortunately, spreadsheets are often used for manipulation and analysis of the data. This task is tedious, and may result in calculation errors[1]. A software program has been created to analyze the output of a common qPCR system, the Viia™-7 Real-Time PCR System[2].

Introduction

The software is written with the purpose of future modification by other researchers. Thus, a simple and inexpensive programming language is desired. Two popular and free rapid application development environments are Xojo[3] and Qt[4] (both tools charge to compile and deploy executable files). Xojo utilizes an object oriented version of BASIC, while Qt employs C++. Xojo is used because most researchers do not have an extensive computer science background, and the BASIC language is easier to learn than C++. Another reason for choosing Xojo is the ease of changing operating systems. Research labs often have a mix of operating systems, and Xojo can run on Windows©, Macintosh, or Linux operating systems. SQLite[5] is an open source database that is used in conjunction with Xojo to group target genes and quickly compute averages and sample standard deviations using structured query language (SQL).

Terminology

The following terminology is used for the data analysis and computer program.

- *k*: the number of genes in the study
- *n*: the number of treatments being studied
- *m*: the number of samples in each treatment
- *Ct*: the expression level
- GAPDH: control gene
- *f*: fold

Data Analysis

The Viia™-7 Real-Time PCR System outputs data in the following format: target gene, treatment, expression level, average expression level. The following steps are used for the analysis.

1. Compute the expression level difference between the control gene and each sample of each treatment (Δ_{ijl} ,where *i*, *j* and *l* represent the gene, treatment and sample numbers.)
2. Compute the average of this difference for the control treatment for each of the *k* genes ($\overline{\Delta_i}$).
3. Compute the fold for each sample of each treatment. The fold is computed using the following expression.

$$f_i = 2^{-(\Delta_i - Ct_{ijk})} \qquad (1)$$

4. Compute the average fold for each combination of gene and treatment.
5. Compute the fold sample standard deviation for each combination of gene and treatment.
6. Construct a bar graph including error bars with a group of treatments for each gene.

Table 1 displays an example of the analysis computations.

Table 1. Example computations.

Gene	Treatment	GAPDH *Ct*	*Ct*	GAPDH - *Ct*	Average Control(Scr) GAPDH - *Ct*	Fold	Average Fold	Fold STD
PPAT	Scr	16.79	18.33	1.5470	1.3100	0.8485	1.0066	0.1370
PPAT	Scr	16.79	17.97	1.1840		1.0913		
PPAT	Scr	16.79	17.99	1.1990		1.0800		
PPAT	mir320	18.04	19.26	1.2179		1.0660	1.0868	0.0276
PPAT	mir320	18.04	19.25	1.2039		1.0763		
PPAT	mir320	18.04	19.19	1.1489		1.1182		
PPAT	mir335	18.14	18.54	0.3995		1.8797	1.8732	0.0147
PPAT	mir335	18.14	18.56	0.4175		1.8564		
PPAT	mir335	18.14	18.54	0.3965		1.8836		
PPAT	mir340	17.55	18.40	0.8507		1.3749	1.4036	0.0269
PPAT	mir340	17.55	18.36	0.8167		1.4077		
PPAT	mir340	17.55	18.34	0.7957		1.4283		
PPAT	mir374	18.31	20.39	2.0850		0.5844	0.5874	0.0427
PPAT	mir374	18.31	20.49	2.1820		0.5464		
PPAT	mir374	18.31	20.28	1.9730		0.6316		
GARS	Unt	18.01	25.07	7.0519		0.0904		
GARS	Unt	18.01	25.03	7.0179		0.0926		
GARS	Unt	18.01	24.97	6.9509		0.0970		
GARS	Scr	16.79	20.39	3.6060	3.5853	0.9855	1.0002	0.0345
GARS	Scr	16.79	20.32	3.5290		1.0396		
GARS	Scr	16.79	20.41	3.6210		0.9754		
GARS	mir320	18.04	21.95	3.9109		0.7978	0.8126	0.0231
GARS	mir320	18.04	21.88	3.8379		0.8392		
GARS	mir320	18.04	21.95	3.9059		0.8006		
GARS	mir335	18.14	21.37	3.2295		1.2794	1.2248	0.0474
GARS	mir335	18.14	21.47	3.3285		1.1946		
GARS	mir335	18.14	21.46	3.3215		1.2004		
GARS	mir340	17.55	24.02	6.4757		0.1348	0.1345	0.0027
GARS	mir340	17.55	24.00	6.4527		0.1370		
GARS	mir340	17.55	24.06	6.5097		0.1317		
GARS	mir374	18.31	24.67	6.3640		0.1457	0.1516	0.0053
GARS	mir374	18.31	24.57	6.2640		0.1561		
GARS	mir374	18.31	24.60	6.2950		0.1528		

Software Demonstration

Figure 1 shows the opening screen of the software. The data from the output file of the Viia^{TM}-7 Real-Time PCR System can be pasted here.

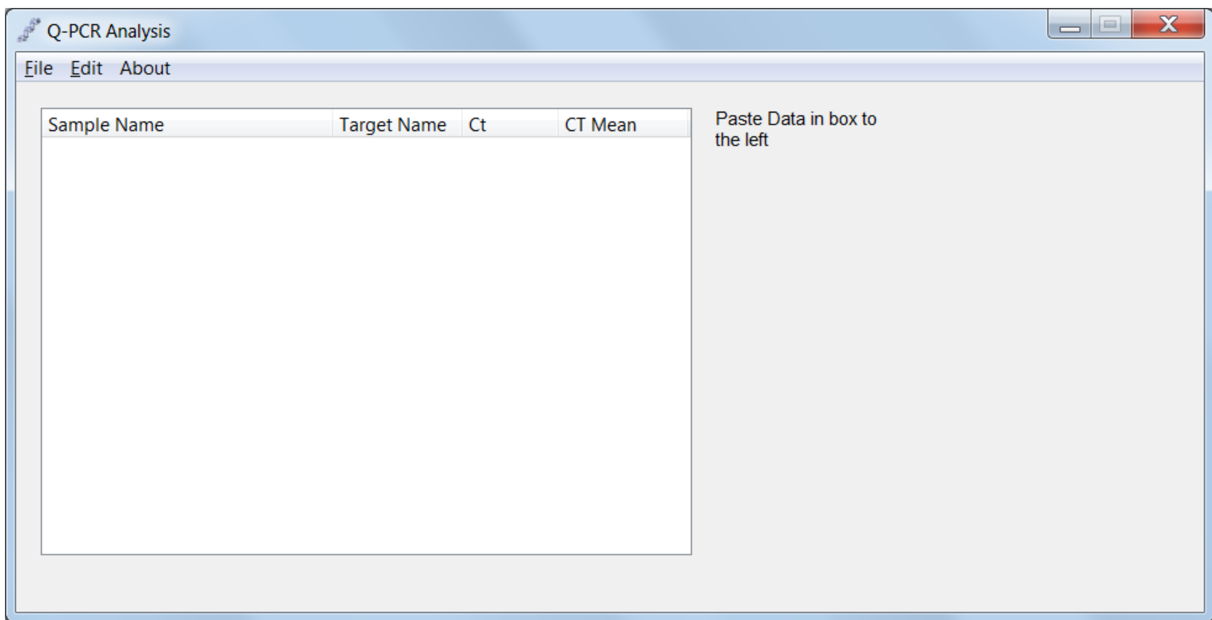

Figure 1. Opening screen.

When the output data is pasted into the opening screen, an SQL command is executed, and a summary of treatments is given in a second list box as shown in Figure 2.

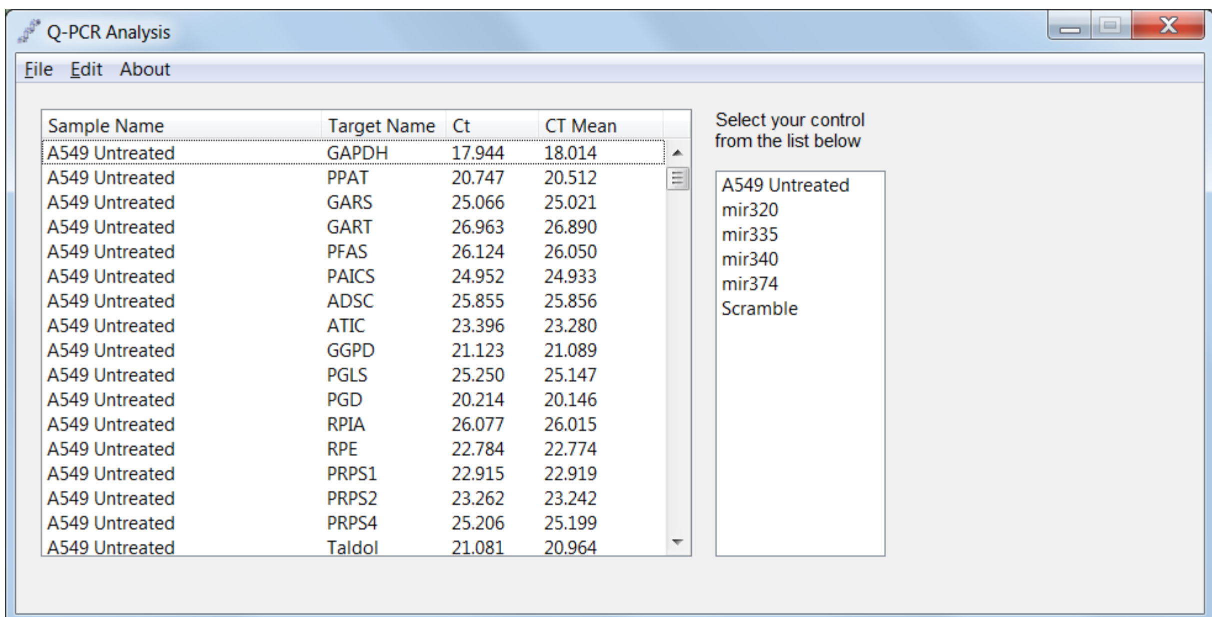

Figure 2. Treatment summary.

Selecting the control treatment from this list box executes another SQL command populating a list box with a summary of target genes. Two additional list boxes are provided for genes to be included in an analysis. Usually only a subset of genes is included because the amount of data contained in a single graph must be limited to maintain legibility. To

include genes in the analysis, simply drag the genes from the third list box to the fourth list box. The fourth list box provided a pop-up menu to remove selected or all genes so results for additional target genes can be graphed. Figure 3 displays an analysis with six target genes.

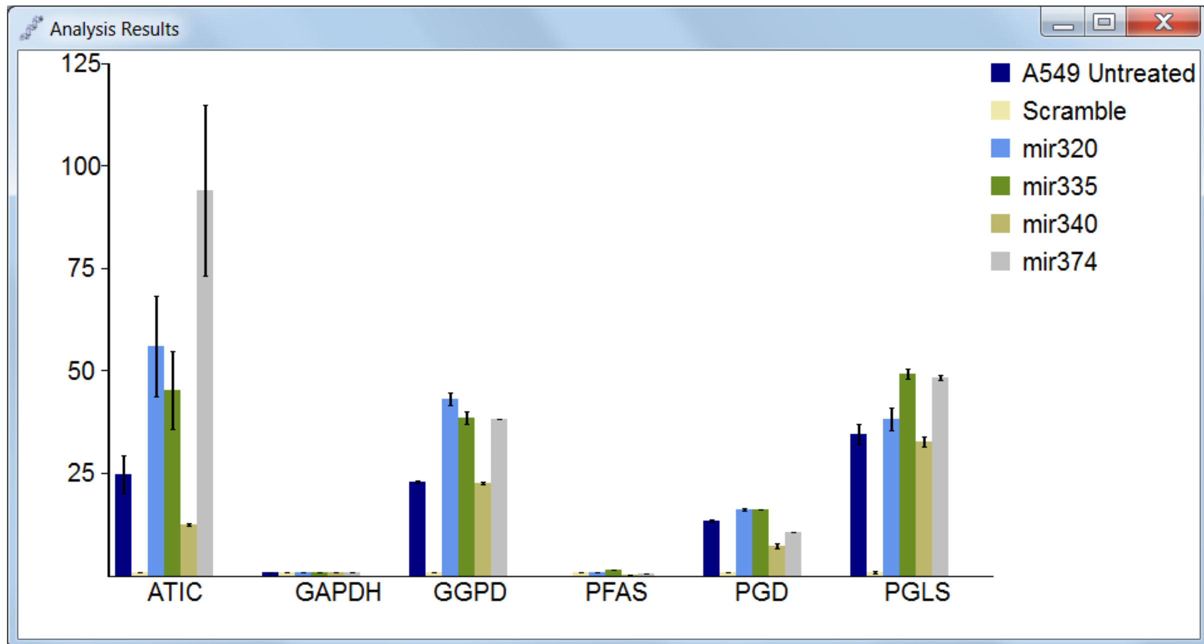

Figure 3. Sample graph.

This graph is enabled with a pop-up menu that allows it to be copied.

Software Availability

The source code for the software may be downloaded at the link below.

http://iqf.org/journal/paper2015-3375/sourceCode.zip

A Windows© executable can be downloaded at the link below. This is a stand alone version that requires no installation. Simply unzip the file, and double-click the executable file (QPCR Analysis.exe).

http://iqf.org/journal/paper2015-3375/qPCRanalysis.zip

A sample data set is provided in spreadsheet format at the link below.

http://iqf.org/journal/paper2015-3375/sampleData.xlsx

Suggested Improvements

It is hoped that the research will improve and customize this software. Suggestions include the ability to process an entire file with no manual involvement. Parameters could be set for the number of target genes per graph, and the software could process all target genes and save all graphs to a desired location. This could be further improved by allowing a target

folder to be identified with multiple output files that could all be processed automatically. Another improvement is to expand the graphics capabilities, allowing users to customize the graph titles, colors, fonts, etc. Printing of the graphics could also be provided.

References

1. Becker, C., Hammerle-Fickinger, A., Riedmaier, I., & Pfaffl, M. W. (2010). mRNA and microRNA quality control for RT-qPCR analysis. Methods. doi:10.1016/j.ymeth.2010.01.010
2. ViiA™ 7 Real-Time PCR System | Life Technologies. (n.d.). Retrieved from http://www.lifetechnologies.com/us/en/home/life-science/pcr/real-time-pcr/real-time-pcr-instruments/viia-7-real-time-pcr-system.html
3. App Creation Software & Mobile App Development - Xojo. (n.d.). Retrieved from http://xojo.com/
4. Qt Project. (n.d.). Retrieved from http://qt-project.org/
5. SQLite Home Page. (n.d.). Retrieved from http://sqlite.org/

Matthew Dodson is a senior majoring in cellular and molecular biology at the University of Michigan, Ann Arbor.

John Slater is a senior majoring in electrical engineering at Lawrence Technological University.