SOFTWARE TOOL ARTICLE

miRcomp-Shiny: Interactive assessment of qPCR-based microRNA quantification and quality control algorithms

[version 1; peer review: 3 approved with reservations]

Lauren Kemperman¹, Matthew N. McCall ²¹Goergen Institute for Data Science, University of Rochester, Rochester, NY, 14627, USA²Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, NY, 14642, USA**V1** First published: 23 Nov 2017, 6:2046
<https://doi.org/10.12688/f1000research.13098.1>Latest published: 23 Nov 2017, 6:2046
<https://doi.org/10.12688/f1000research.13098.1>

Abstract

The miRcomp-Shiny web application allows interactive performance assessments and comparisons of qPCR-based microRNA expression and quality estimation methods using a benchmark data set. This work is motivated by two distinct use cases: (1) selection of methodology and quality thresholds for use analyzing one's own data, and (2) comparison of novel expression estimation algorithms with currently-available methodology. The miRcomp-Shiny application is implemented in the R/Shiny language and can be installed on any operating system on which R can be installed. It is made freely available as part of the miRcomp package (version 1.3.3 and later) available through the Bioconductor project at: <http://bioconductor.org/packages/miRcomp>. The web application is hosted at <https://laurenkemperman.shinyapps.io/mircomp/>. A detailed description of how to use the web application is available at: http://lkemperm.github.io/miRcomp_shiny_app

Keywords

microRNA, miRcomp, qPCR, benchmark data

Open Peer Review

Approval Status ? ? ?

	1	2	3
version 1	?	?	?
23 Nov 2017	view	view	view

1. **Levi Waldron** , City University of New York, New York, USA
2. **Simina M. Boca** , Georgetown University Medical Center, Washington, USA
3. **Elana J. Fertig**, Johns Hopkins University, Baltimore, USA

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the **RPackage** gateway.

This article is included in the **Bioconductor** gateway.

Corresponding author: Matthew N. McCall (mccallm@gmail.com)

Author roles: **Kemperman L:** Data Curation, Investigation, Methodology, Software, Visualization, Writing – Original Draft Preparation; **McCall MN:** Conceptualization, Data Curation, Funding Acquisition, Methodology, Project Administration, Resources, Software, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the National Institutes of Health grant to M.N.M (R00-HG006853).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2017 Kemperman L and McCall MN. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Kemperman L and McCall MN. *miRcomp-Shiny: Interactive assessment of qPCR-based microRNA quantification and quality control algorithms [version 1; peer review: 3 approved with reservations]* F1000Research 2017, 6:2046 <https://doi.org/10.12688/f1000research.13098.1>

First published: 23 Nov 2017, 6:2046 <https://doi.org/10.12688/f1000research.13098.1>

Introduction

Quantitative real-time PCR (qPCR) is one of the most widely used methods to measure the expression of a target transcript. A variety of algorithms have been developed to estimate expression from qPCR fluorescence measurements. The vast majority of these algorithms were developed and tested using gene expression data¹⁻³; however, they are now routinely applied to qPCR-based microRNA expression measurements. To evaluate the performance of these methods on microRNA data, we developed a benchmark data set and a collection of statistical assessments⁴. These and other recent assessments⁵ highlight the need to develop qPCR quantification and quality assessment methodology specifically tailored to microRNA expression platforms.

Current methods to estimate expression from raw qPCR amplification data have been developed in a variety of programming languages (e.g. R, Python, SAS) and may be restricted to a particular operating system (e.g. Windows, Mac OS, Unix/Linux)⁶. Furthermore, these algorithms often return different data structures, complicating comparisons between methods. The miRcomp-Shiny web application provides a unified assessment environment that is platform independent and takes simple expression and quality matrices as input. This approach removes barriers to usage and facilitates the comparison of methods.

Methods

Implementation

We have developed a Shiny (<http://shiny.rstudio.com/>) interface to the miRcomp R package (version $\geq 1.3.3$). Currently, six of the most widely-used algorithms to estimate miRNA expression and sample quality are included in the miRcomp-Shiny app. Each method provides both expression estimates and quality metrics. Assessments can be performed on individual algorithms, or two available algorithms can be compared. Researchers can also upload the results from their own method to be assessed. As new methods are developed and tested, we will continue to add these methods to miRcomp-Shiny. The development of a repository of qPCR-based miRNA expression estimation algorithms will be a valuable resource for researchers seeking to develop new methodology or comparison existing algorithms across a wide variety of assessment criteria.

Advantages of an interactive interface to the miRcomp package

The web application framework in R (Shiny) has enabled us to make several aspects of the miRcomp package more interactive than they were previously and facilitate comparisons that would have been difficult to make in R. Below we describe two common use cases that motivated the development of miRcomp-Shiny.

Methodology and quality threshold selection

When selecting methodology to analyze a data set, miRcomp-Shiny can be used to evaluate the performance of existing methods based on the benchmark data. The results of these evaluations can be used to guide the selection of an expression estimation algorithm

and quality threshold based on the assessments most relevant to the user's experiment. Additionally, one can examine the effect of changing quality thresholds on the performance of each method. The result of changes in the quality threshold are then displayed immediately for each assessment. This is particularly useful when selecting a quality threshold for one's own data.

Comparison of novel algorithms with current methods

Another use case is comparison of a new method to an existing method. By providing current methods for comparison, researchers do not have to implement these algorithms themselves, which is often a substantial bottleneck in the development and assessment of novel algorithms. Additionally, we will continue to add new methods to miRcomp-Shiny. This will produce a richer set of available methods in a single location to guide comparisons. The success of this approach has been demonstrated by the affycomp webtool^{7,8}.

Operation

Installation. To access miRcomp-Shiny locally, the miRcomp R package and all required dependencies can be installed from Bioconductor with the following commands:

```
source("http://bioconductor.org/biocLite.R")
biocLite("miRcomp")
```

To access miRcomp-Shiny remotely, simply go to:

<https://laurenkemperman.shinyapps.io/mircomp/>

Input. The miRcomp-Shiny app takes one or two quantification methods as input. These methods can be selected from the drop-down menus on the left panel (Figure 1). Alternatively, the user can upload the results of their own method by selecting the `custom` option from the menu. If the `custom` option is selected, the user is prompted to upload a matrix of quality values (`qc`) and a matrix of expression estimates (`ct`). Once the method or methods have been selected, each assessment plot contains additional assessment-specific options below the plotting window (Figure 1).

Output. The miRcomp-Shiny app produces five plots to assess the performance of the method or methods selected: Limit of Detection, Accuracy, Precision, Quality Assessment, and Titration Response.

Use cases

The assessments performed by miRcomp-Shiny are based on a benchmark data set available at: <http://bioconductor.org/packages/miRcompData/>

Users wishing to assess quantification and quality control metrics beyond those currently implemented, can run any algorithm on those data and upload the resulting matrices of quality values (`qc`) and expression estimates (`ct`). Examples of these matrices are included as R data objects in the miRcomp package for each of the currently implemented methods.

Figure 1. The miRcomp-Shiny interface is organized into two panels: on the left the user selects one or more methods to compare and on the right the user selects an assessment tab. The results of several methods are available or the user can upload the results of their own method (as shown above). After selecting one or more methods, the user can examine the 5 assessment tabs: limit of detection, accuracy, precision, quality assessment, and titration response. Each tab has its own options shown at the bottom of the pane.

Summary

The success of the miRcomp Shiny web application and software will depend on new methods being developed and the application being used to test them. We have already begun encouraging people to use the package, and hope that readers will do the same. Widespread use of these tools will lead to improvements across all benchmarks in microRNA expression estimation, and the accessibility of the web application makes that possible.

Software availability

Software available from: <https://laurenkemperman.shinyapps.io/mircomp/>

Source code available from: <http://bioconductor.org/packages/miRcomp>

Archived source code as at time of publication: <http://doi.org/10.5281/zenodo.1049074>⁹

Software license: GPL-3

Author contributions

L.K. developed the software under the supervision of M.N.M. All authors wrote and approved the final manuscript.

Competing interests

No competing interests were disclosed.

Grant information

This work was supported by the National Institutes of Health grant to M.N.M (R00-HG006853).

The authors confirm that the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We thank Dr. Marc Halushka for his helpful comments on this manuscript. The OpenArray chips were run by the staff at the Genetic Resources Core Facility, Johns Hopkins Institute of Genetic Medicine, Baltimore, MD.

References

- Ritz C, Spiess AN: **qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis**. *Bioinformatics*. 2008; **24**(13): 1549–1551.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Lievens A, Van Aelst S, Van den Bulcke M, Goetghebeur E: **Enhanced analysis of real-time PCR data by using a variable efficiency model: FPK-PCR**. *Nucleic Acids Res*. 2012; **40**(2): e10.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Rödiger S, Burdukiewicz M, Schierack P: **chipPCR: an R package to pre-process raw data of amplification curves**. *Bioinformatics*. 2015; **31**(17): 2900–2.
[PubMed Abstract](#) | [Publisher Full Text](#)
- McCall MN, Baras AS, Crits-Christoph A: **A benchmark for microRNA quantification algorithms using the openarray platform**. *BMC bioinformatics*. 2016; **17**(1): 138.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Mestdagh P, Hartmann N, Baeriswyl L, *et al.*: **Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study**. *Nat Methods*. 2014; **11**(8): 809–815.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Pabinger S, Rödiger S, Kriegner A, *et al.*: **A survey of tools for the analysis of quantitative PCR (qPCR) data**. *Biomol Detect Quantif*. 2014; **1**(1): 23–33.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cope LM, Irizarry RA, Jaffee HA, *et al.*: **A benchmark for affymetrix genechip expression measures**. *Bioinformatics*. 2004; **20**(3): 323–331.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Irizarry RA, Wu Z, Jaffee HA: **Comparison of affymetrix genechip expression measures**. *Bioinformatics*. 2006; **22**(7): 789–794.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kemperman L, McCall MN: **mccallm/miRcomp-shiny: F1000Research Version (Version v1.9.0-F1000)**. *Zenodo*. 2017.
[Data Source](#)

Open Peer Review

Current Peer Review Status: ? ? ?

Version 1

Reviewer Report 18 December 2017

<https://doi.org/10.5256/f1000research.14205.r28321>

© 2017 Fertig E. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Elana J. Fertig

Division of Biostatistics & Bioinformatics, Johns Hopkins University, Baltimore, MD, USA

This paper presents a new Shiny application for analyses of qPCR data to facilitate comparison of methods. This software is important for the community and well suited to publication in *F1000Research*.

Major comments

1. Some details of the software are hard to follow from the manuscript below (e.g., which methods are implemented, what benchmark datasets, etc). These are detailed in minor comments below, and must be fixed for readability of the manuscript. The authors may wish to add a table listing each of the methods and the quality thresholds that they yield.
2. Use cases should report results from applying the method, not merely the datasets used for analysis.
3. The manuscript should describe the range of possible analyses that can be performed with miRcomp-Shiny, expanding upon the "output section".
4. The annotations and help on the software available from <https://laurenkemperman.shinyapps.io/mircomp/> require improvement for greater usability. Some examples are listed below:
 - Dataset descriptions describes the methods employed, but does not indicate which datasets are used.
 - The platform seems limited to analysis of the miRcomp data. The software does not appear to enable input of new datasets which would be critical to its utility.
 - The format of files for qc and ct elements for custom analyses are not specified.
 - There is no ability to export processed datasets and/or assess the quality of specific miRNAs from the preprocessing implemented in this application.

Minor comments

1. The Implementation subsection of the Methods should expand the sentence "Currently, six

of the most widely-used algorithms to estimate miRNA expression and sample quality are included in the miRcomp-Shiny app” to clarify precisely which six algorithms are implemented and include citations to those methods. It should also clarify whether these 6 algorithms are representative of all in the miRcomp R package or a subset of the methods implemented in that package.

2. “The benchmark data” referenced in the Methodology and quality selection threshold should be defined. Which datasets are included as benchmarks? How are they selected?
3. It is unclear what variables the “quality thresholds” in the Methodology and quality section threshold section reference.
4. The subsection “Comparison of novel algorithms...” should edit the sentence “Another use case is comparison of a new method to an existing method.” to read “Another use case is comparison of a data table with results from a new method to the existing methods implemented in the miRcomp-Shiny app.”
5. The sentence “We have already begun encouraging people to use the package, and hope that readers will do the same.” Should be cut.
6. The summary should place this tool in context of others in the literature and discuss its limitations / future work.

Is the rationale for developing the new software tool clearly explained?

Partly

Is the description of the software tool technically sound?

Partly

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

No

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

Partly

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

Partly

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 18 December 2017

<https://doi.org/10.5256/f1000research.14205.r28323>

© 2017 Boca S. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Simina M. Boca

Innovation Center for Biomedical Informatics, Department of Oncology, Georgetown University Medical Center, Washington, DC, USA

I applaud Kemperman and McCall on the work they put forward to make methods comparisons more accessible to researchers interested in microRNA quantification. I generally agree with Dr. Waldron's comments. Specific points where I have reservations include:

1) I am still not sure how users can add their own data to compare various methods on it. It seems like if one selects "custom" for one of the two methods, then one must load a processed version of the same dataset in order to compare a new algorithm to existing algorithms.

2) In general, I think Dr. Waldron's comment on whether users still need to use the Bioconductor package directly is very valuable. It partly depends on what users the authors have in mind for the shiny app in terms of their level of R/bioinformatics expertise.

Two specific examples here:

a) If an unfamiliar user were to use this app, it seems like they would still need to go to the Bioconductor package to get the dataset. Perhaps the authors could include this dataset in the app and make it easier to download as both a CSV file and an .RData file? It would also be worthwhile to include a citation to the dataset, as opposed to needing to go through Bioconductor.

b) More explanations need to accompany the plots and explain exactly what is being plotted and why - each plot should at least include the same type of level of detail used when writing a figure caption in a scientific journal. A more detailed introduction should also be provided, along with links to the Bioconductor package(s) and to this paper, but written so that it is at least somewhat self-contained.

For example, the first plot, "Limit of detection" appears to not be a comparison plot at all, but rather to show just the limit of detection for the first method. What should the user expect to see here for a "good" method? What does "Proportion Poor Quality" even mean (is there a threshold that can be changed to indicate this, why is there never a boxplot for the proportion = 1?) For "Accuracy" it is also not clear what "percentage of data to exclude" means (why is it being excluded? quality issues?) For "Accuracy" and "Precision," the Low/Medium/High values on the x-axis should be described, along with stating that within each category, the methods are being compared (maybe some dashed vertical lines between categories would also help here). In general, it should be indicated what one should look for in terms of one method having better performance compared to another.

I strongly encourage the authors to make these changes/additions in order to allow a larger number of individuals to use and benefit from their tool.

Is the rationale for developing the new software tool clearly explained?

Partly

Is the description of the software tool technically sound?

Partly

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

No

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

Partly

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 04 December 2017

<https://doi.org/10.5256/f1000research.14205.r28320>

© 2017 Waldron L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Levi Waldron

School of Public Health, City University of New York, New York, NY, USA

Kemperman and McCall present a Shiny App that is a straightforward extension of their Bioconductor mirComp package, creating the plots provided by the package with sliders for the options of the plotting functions, using a suite of microRNA qPCR test data.

Major

There are some shortcomings in the app that initially made me think it would still be necessary to use the Bioconductor version directly:

* I couldn't find documentation or an option in the app on how to use one's own data with the app (use case 1 from abstract). **update** I found the "custom" method option, but only by trial and

error – this option should be emphasized within the app

* I also couldn't see documentation on how to add a novel algorithm for comparison (use case 2 from abstract). **update** I think I understand now that a novel algorithm would be applied prior to uploading the data to the app, but please clarify this in the manuscript and app

* The figures don't have axis labels or captions, so one has to look up the miRcomp vignette and reference manual to understand them

It would be helpful to state in the Introduction and on the web app page who the intended users of the app are, and who it's not intended for. It would also be helpful to state in the Introduction how the "simple expression and quality matrices as input" would normally be generated, with specific instructions both for generating (e.g. pointing to documentation for the methods already in the tool and how data from added methods should be formatted), and for uploading to the tool. This should also be coupled with explicitly stating that the tool does not perform normalization of miR expression data (or, better yet, incorporating normalization into the tool), and that the tool only assesses already normalized miR expression data. The point of this comment is to make it clearer up front to a reader whether or not the tool is for them.

As the app takes simple expression and quality matrices as input, it would be helpful to point to instructions on how to prepare these matrices. It would seem that this requires using the R/Bioc command line, so the app may facilitate the comparison of methods but not remove barriers to usage. Again, it should just be clear up front what requirements to the user are for the intended use cases.

This may be outside the scope of the paper, but the tool would be of greater use to wet lab biologists if they could upload raw data, do the comparisons provided by the app, then download normalized data. This would probably significantly expand the number of potential users.

I understand from the introduction that the tool intends to expand on methods for testing qPCR normalization used for miR expression data. But it would be helpful to state how the app is actually specific to miR expression data – is it just that it provides miR datasets for benchmarks? Or are some of the normalization methods miR-specific?

When trying out the app at <https://laurenkemperman.shinyapps.io/mircomp/>, I constantly got the message "Disconnected from the server. Reload" I had to run the app locally to test it usefully. The authors may need an upgraded shinyapps.io account to support public usage. When reloading, all changes made to the settings are reset.

Minor

With qpcRb4 as the first method, I get an error "need finite 'ylim' values". I haven't checked through all the plotting combinations.

Is the rationale for developing the new software tool clearly explained?

Partly

Is the description of the software tool technically sound?

Partly

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

Partly

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: biostatistics, metagenomics, microbiome, software for multi-omics data analysis

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research