

Red-R: A visual programming and data analysis framework

Anup Parikh^{1,*} and Kyle R Covington¹

1. Baylor College of Medicine, One Baylor Plaza, Houston Texas, USA
- * Contact author: anup.parikh@gmail.com

Keywords: graphical user interface, visual programming, bioinformatics

Cross-discipline collaboration is major driving force for new innovation. As an example, the collaboration between bioinformaticists (computational users) and biologists (non-computational users) to generate and analyze high-throughput datasets has revolutionized the life sciences. However, as the volume of data and technical complexity of the analysis increases, non-computational users can not participate in the data exploration and analysis. Furthermore due to the complexity, data analysis pipelines are difficult to understand and share, even with other computational users. This lack of powerful yet user friendly tools greatly limits analysis interpretability and the ability of non-computational users to participate in the data analysis process.

To address these challenges we have developed Red-R, a user friendly visual programming and data analysis framework for *R*. Red-R makes the advanced functionality of *R* available to the non-computational users by hiding the computational complexity behind a visual programming interface. Analyses are performed by visually linking a series of widgets together that read, manipulate, and interactively display data. These pipelines, representing both the data and analysis, can be easily shared with others. Creating a visual representation of the analysis greatly increases interpretability and gives non-computational users the ability to explore the data and analysis without understanding *R* code. Along with the visual representation, all the *R* code is readily available for review. This functionality in Red-R can create more effective collaborations by allowing users to perform complex analyses and share them so that others can easily understand the analysis and further explore the data.

Red-R is an extension of Orange (<http://www.ailab.si/orange>), a data mining framework written in *Python* and *Qt*. Red-R accesses all the functionality and data in *R*, using the *Python* interface for *R* provided by RPy (<http://rpy.sourceforge.net>). This framework is highly flexible and can be extended to include virtually all the functionally *R* currently offers. To facilitate this, we have created a widget that parses the *R* code of a given function and generate the basic *Python/Qt* code to quickly create a new widget in Red-R. In addition to creating a visual interface to *R* functions, Red-R can harness the power of *Qt* graphics to create interactive visualizations.

The current version of Red-R provides functionality for data manipulation, basic statistics and advanced bioinformatics analysis from the **Bioconductor** package. Red-R also provides widgets to import existing *R* sessions and execute any *R* code. We hope to gain community support in extending the repository of widgets and providing analysis pipeline for common tasks. To that aim, our website (<http://www.red-r.org>) hosts documentation, example schemas, and shared resources for Red-R.

Red-R Interface: Differential expression analysis of Affymetrix microarrays.

