



SOFTWARE TOOL

REVISED **ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis [v2; ref status: indexed, <http://f1000r.es/4cf>]**

Previously titled: ReactomeFIViz: the Reactome FI Cytoscape app for pathway and network-based data analysis

Guanming Wu^{1,2}, Eric Dawson³, Adrian Duong¹, Robin Haw¹, Lincoln Stein^{1,4}

¹Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada

²DMICE, Oregon Health & Science University, Portland, Oregon 97239, USA

³Section of Integrative Biology, Institute for Cellular and Molecular Biology, and Center for Computational Biology and Bioinformatics, The University of Texas at Austin, Austin, TX 78712, USA

⁴Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada

v2 **First published:** 01 Jul 2014, 3:146 (doi: [10.12688/f1000research.4431.1](https://doi.org/10.12688/f1000research.4431.1))
Latest published: 12 Sep 2014, 3:146 (doi: [10.12688/f1000research.4431.2](https://doi.org/10.12688/f1000research.4431.2))

Abstract

High-throughput experiments are routinely performed in modern biological studies. However, extracting meaningful results from massive experimental data sets is a challenging task for biologists. Projecting data onto pathway and network contexts is a powerful way to unravel patterns embedded in seemingly scattered large data sets and assist knowledge discovery related to cancer and other complex diseases. We have developed a Cytoscape app called “ReactomeFIViz”, which utilizes a highly reliable gene functional interaction network combined with human curated pathways derived from Reactome and other pathway databases. This app provides a suite of features to assist biologists in performing pathway- and network-based data analysis in a biologically intuitive and user-friendly way. Biologists can use this app to uncover network and pathway patterns related to their studies, search for gene signatures from gene expression data sets, reveal pathways significantly enriched by genes in a list, and integrate multiple genomic data types into a pathway context using probabilistic graphical models. We believe our app will give researchers substantial power to analyze intrinsically noisy high-throughput experimental data to find biologically relevant information.



This article is included in the [Cytoscape App Collection](#)

Open Peer Review

Invited Referee Responses

	1	2
version 1 published 01 Jul 2014	 report 1	 report 1
version 2 published 12 Sep 2014 REVISED		 report

- 1 **Hamid Bolouri**, Fred Hutchinson Cancer Research Center USA
- 2 **Nikolaus Schultz**, Memorial Sloan-Kettering Cancer Center USA, **B. Arman Aksoy**, Memorial Sloan Kettering Cancer Center USA

Latest Comments

No Comments Yet

Corresponding author: Guanming Wu (guanmingwu@gmail.com)

How to cite this article: Wu G, Dawson E, Duong A *et al.* **ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis [v2; ref status: indexed, <http://f1000r.es/4cf>]** *F1000Research* 2014, 3:146 (doi: [10.12688/f1000research.4431.2](https://doi.org/10.12688/f1000research.4431.2))

Copyright: © 2014 Wu G *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

Grant information: This project is supported by a NIH grant (2U41HG003751-05) to LS and a Genome Canada grant (OG1 5458) to LS. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

Competing interests: No competing interests were disclosed.

First published: 01 Jul 2014, 3:146 (doi: [10.12688/f1000research.4431.1](https://doi.org/10.12688/f1000research.4431.1))

First indexed: 22 Sep 2014, 3:146 (doi: [10.12688/f1000research.4431.2](https://doi.org/10.12688/f1000research.4431.2))

REVISED Amendments from Version 1

The major changes we have made to this version according to suggestions from reviewers are below:

- 1). Added a new paragraph in the "Results" section.
- 2). Added a new figure as Figure 3 for the new paragraph.
- 3). Added a new reference as Reference 8.
- 4). Changed the title to avoid confusion.

See referee reports

Introduction

High-throughput experiments, which generate large and complex data sets, are routinely performed in modern biological and clinical studies to unravel mechanisms underlying complex diseases, such as cancer. However, extracting reliable and meaningful results from these experiments is usually difficult and requires sophisticated computational tools and algorithms, which are challenging for experimental biologists to comprehend. A user-friendly software tool is extremely important for both bench and computational biologists to perform high-throughput data analysis related to cancer and other complex diseases.

Many studies have shown that alterations in pathways or networks are better correlated with complex disease phenotypes than any particular gene or gene product^{1,2}. Pathway- and network-based data analysis approaches project information about seemingly unrelated genes and proteins onto pathway and network contexts, and create an integrated view for researchers to understand mechanisms related to phenotypes of interest.

In this paper, we describe a software tool called ReactomeFIViz (also called the Reactome FI Cytoscape app or ReactomeFIPlugIn), which can be used to perform pathway- and network-based data analysis for data generated from high-throughput experiments. This tool uses the highly reliable Reactome functional interaction (FI) network³ for doing network-based data analysis. The FI network was constructed by merging interactions extracted from human curated pathways with interactions predicted using a machine learning approach. This tool can also be used to perform pathway-based data analysis by using high quality human-curated pathways in the Reactome database⁴, the most comprehensive open source pathway database.

Implementation

Software architecture

We used conventional three-tier software architecture to implement ReactomeFIViz (Figure 1). The back-end contains several databases

hosted in the open-source MySQL database engine (<http://www.mysql.com>). The middle server-side application uses hibernate (<http://hibernate.org>) to access the databases storing FIs and cancer gene index data (see below). The server-side application also uses the in-house developed Reactome API for Object/Relational mapping to access pathway-related contents stored in a database using the Reactome database schema. On the server-side, a lightweight servlet container, Spring Framework (<http://projects.spring.io/spring-framework/>), and a Java RESTful framework, Jersey (<https://jersey.java.net>), are used to power a RESTful API for the Cytoscape front-end. The front-end Cytoscape app uses this RESTful API to communicate with the server-side application. Almost all analysis features in the app are provided by this RESTful API, which should also facilitate their use by other front-end applications, such as a web browser or tablet app.

For cancer data analysis, we imported the cancer gene index (CGI, <https://wiki.nci.nih.gov/display/cageneindex>) data into a MySQL database and then developed a hibernate API for the server-side application. The CGI data contains annotations for cancer-related genes. These annotations were extracted by using text-mining technologies and then validated by human curators (<https://wiki.nci.nih.gov/display/cageneindex/Creation+of+the+Cancer+Gene+Index>).

The Reactome FI network is updated annually. We recommend using the latest version of the FI network. Different versions of the FI network may yield different results due to updates to gene interactions, so we have also deployed two older versions of the FI network to use for comparison of legacy data sets and to reproduce published results.

R (<http://www.r-project.org>) is used in the server-side for executing network module-based survival analysis and other statistical computations. ReactomeFIViz uses Java based methods in the server-side to call functions in R. Users of our app don't need to install R in their machines in order to perform the statistical analyses implemented in the app.

ReactomeFIViz is designed and implemented for Cytoscape 3, and includes all features in Reactome FI Cytoscape plug-in for Cytoscape 2. Users are recommended to use the latest version of our app for Cytoscape 3.

Network analysis features

ReactomeFIViz implements multiple features for users to perform network-based data analysis, including FI sub-network construction³, network module discovery³, functional annotation³, HotNet mutation analysis^{5,6}, and network module-based gene signature discovery from microarray data sets⁷. The HotNet algorithm^{5,6} was



Figure 1. The three-tier software architecture used to implement ReactomeFIViz.

implemented by porting python and MatLab code of HotNet_v1.0.0 (downloaded from <http://compbio.cs.brown.edu/projects/hotnet/>) to Java and R. For details about other algorithms and their implementations, please refer to our previous work^{3,7}.

The majority of interactions in the Reactome FI network are extracted from reactions and complexes. In order to display semantic meanings (e.g. catalysis, activation and inhibition) of these interactions, we created a Reactome FI network specific visual style. This visual style is registered as a service using the OSGi API supported by Cytoscape 3, and applied to newly constructed FI sub-network automatically for network analysis.

Pathway analysis features

Since version 4.0.0.beta, released in January 2014, ReactomeFIViz allows users to explore a list of high quality, human curated Reactome pathways, visualize Reactome pathways directly in Cytoscape, and perform pathway enrichment analysis on a list of genes based on a binomial test⁸. In April 2014, we added a new experimental feature for performing integrated pathway analysis for multiple genomic data types by adapting a factor graph based approach called "PARADIGM"⁹ into ReactomeFIViz.

The Reactome database contains several hundred manually laid-out pathway diagrams⁴. Pathway diagrams in Reactome are drawn based on biochemical reactions. A reaction usually contains multiple inputs and outputs, in addition to catalysts, inhibitors and activators. The network model in Cytoscape is designed to support simple graphs containing edges between two nodes only. In order to display Reactome pathway diagrams, we adapted the pathway diagram view in the Reactome curator tool⁴ into the Cytoscape environment, and wrapped it in a JInternalFrame so that a pathway view can be displayed along with a network view in the Cytoscape desktop (Figure 2).

Results

ReactomeFIViz provides a suite of features to assist users to perform pathway- and network-based data analyses (Figure 3). Based on a list of genes loaded from a file, the user can construct a sub-network, perform network clustering to search for network modules related to patient clinical or other phenotypic information, annotate network modules, perform pathway enrichment analysis, and even model pathway activities based on probabilistic graphical models⁹. By performing pathway- and network-based analyses using ReactomeFIViz, researchers will be able to uncover pathway and network

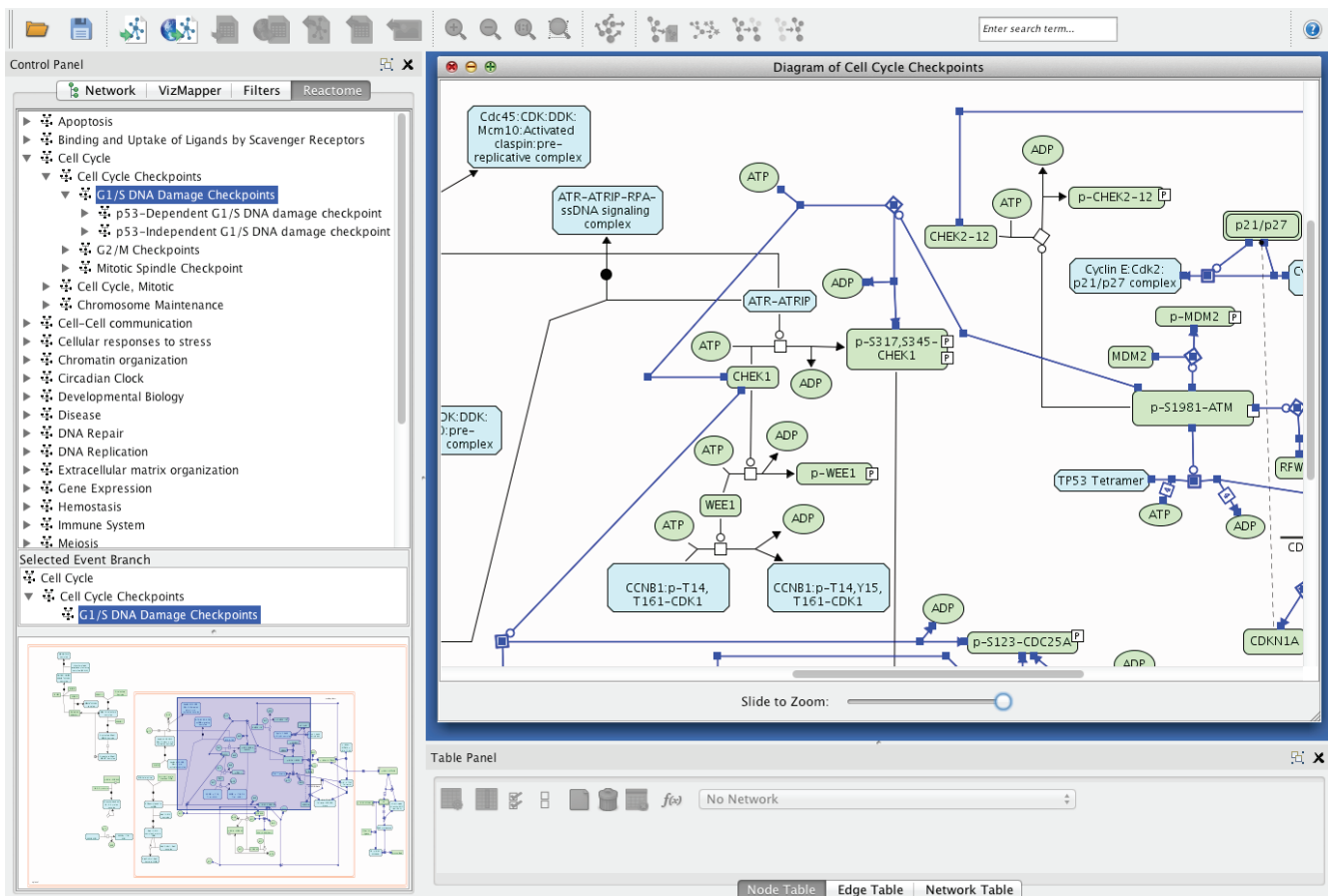


Figure 2. A Reactome pathway diagram displayed in a Reactome diagram view. The diagram view is wrapped in a JInternalFrame and hosted in the Cytoscape desktop.

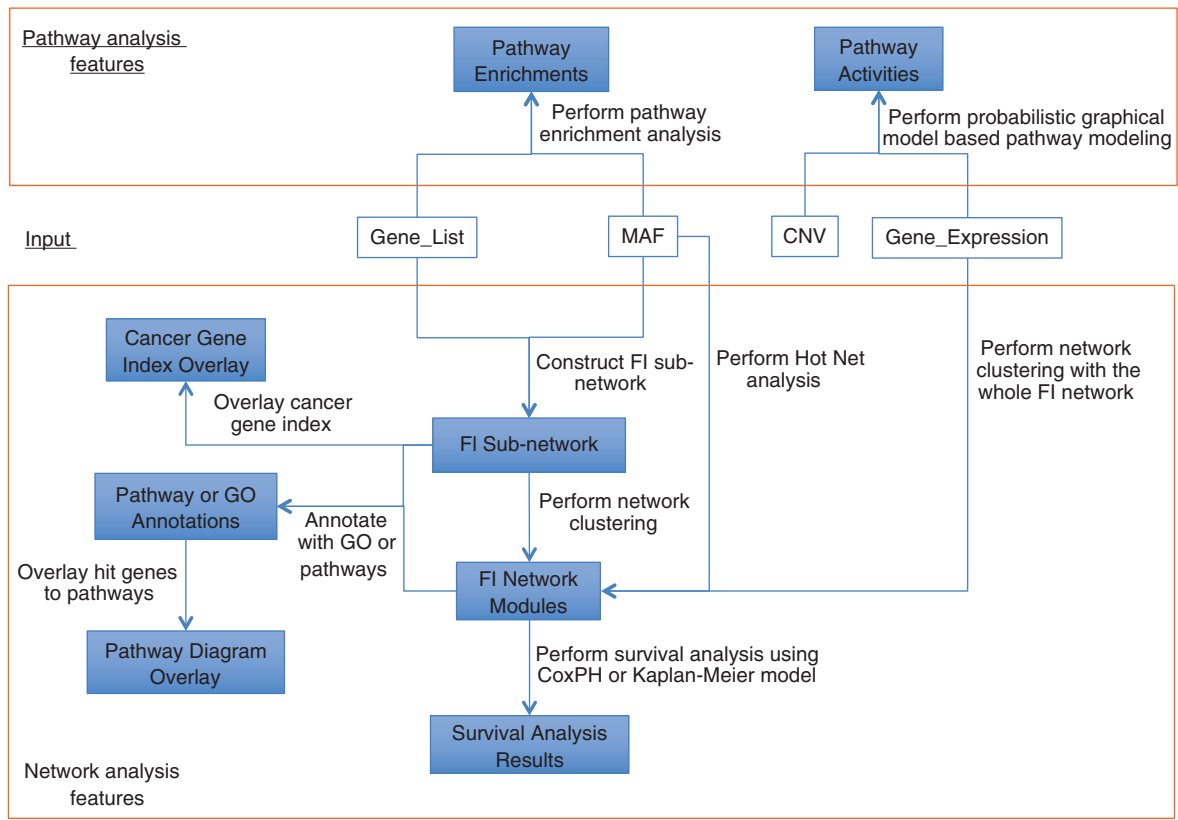


Figure 3. Major features implemented in ReactomeFIViz. Features implemented in ReactomeFIViz are roughly categorized into two types: pathway analysis features (top) and network analysis features (bottom). Input for these features can be a simple gene list contained in a text file, a mutation annotation file (MAF), a text file for gene expression data, or copy number variants (CNVs).

patterns related to their studies and then link found patterns to clinical phenotypes^{3,7}.

As an example, we present results generated from network module based analysis for the TCGA ovarian cancer mutation data¹⁰ using ReactomeFIViz. The TCGA mutation data file and clinical information file were downloaded from the Broad Institute Firehose web site <https://confluence.broadinstitute.org/display/GDAC>, released in July 2012. The clinical information has been pre-processed. For this data set, we chose the 2009 version of the FI network, and picked genes mutated in three or more samples to construct a FI sub-network. We performed a network clustering, followed by survival analysis for each network module by splitting samples into two groups: samples having genes mutated in the module (Group 1) and samples not having genes mutated in the module (Group 0). Our results indicate that group 1 samples (Figure 4, green line in the Kaplan-Meier plot¹¹) have significantly longer overall survival times compared to group 0 samples (Figure 4, red line in the Kaplan-Meier plot) ($p\text{-value} = 3.4 \times 10^{-5}$ based on the CoxPH analysis¹²) based on module 3. Pathway enrichment analysis results imply that module 3 is enriched with genes in calcium signaling pathway (<http://www.genome.jp/kegg/pathway/hsa/hsa04020.html>) and mitotic G2/M transition (http://www.reactome.org/cgi-bin/control_panel_st_id?ST_ID=REACT_2203.2). These results suggest that mutations impacting calcium signaling and the cell cycle may increase the survival of

ovarian cancer patients. However, we may need more samples and independent data sets to validate our conclusion.

Using the same version of ReactomeFIViz but different versions of the FI network may yield different results because of updates of protein interactions in the FI network. We performed the same analysis with the latest version of the FI network (the 2013 version), and found that genes in module 3 from the 2009 version of the FI network have been split among several modules discovered using the newer version of the FI network. The module having the largest overlap with module 3 from the 2009 version of the FI network has the most significant p-value from the survival analysis ($p\text{-value} = 1.1 \times 10^{-3}$ from CoxPH), which implies that our method is fairly robust against updates of the FI network. For details, see the supplementary results.

Discussion

Our Cytoscape app provides a suite of features for users to perform network- and pathway-based analysis for data generated from multiple experiments related to cancer and other complex diseases. Users can use our tool to search for disease-related network and pathway patterns. Our tool is built upon the Reactome database, arguably the most comprehensive human curated open source pathway database, and leverages the highly reliable functional interaction network extracted from human curated pathways. Many

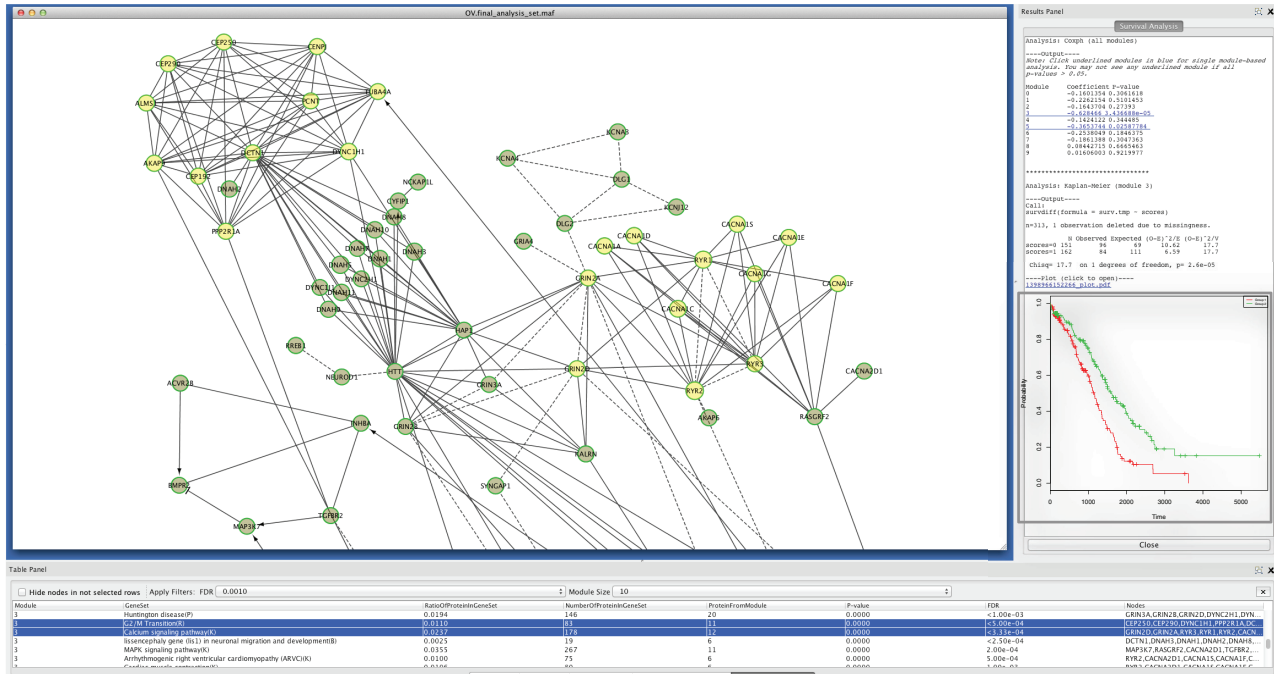


Figure 4. Module 3 generated from the TCGA ovarian cancer mutation data file is significantly related to patient overall survival. The central main panel shows the network view of module 3. The bottom table displays pathway annotations for genes in module 3 with two pathways, Calcium Signaling Pathway and G2/M Transition, highlighted. The right panel shows survival analysis results using both the Cox proportional hazards (CoxPH) model and the Kaplan-Meier model. The Kaplan-Meier plot was added to the figure later.

studies based on the FI network and this app have shown its many applications to cancer and other disease studies¹³⁻¹⁶.

For future development, we will focus on using probabilistic graphical models, such as factor graphs, for performing pathway modeling and linking results to patient clinical information in order to uncover cellular mechanisms related to cancer drug sensitivity, search for cancer biomarkers, and assist new drug development.

Data availability

Data files used in the example: http://reactomews.oicr.on.ca:8080/cabigR3WebApp/TCGA_OV_Firehose_MAF_CLIN_2012.zip.

Use the detailed procedures described in our user guide to reproduce the results described in the example: http://wiki.reactome.org/index.php/Reactome_FI_Cytoscape_Plugin.

Software availability

Homepage: http://wiki.reactome.org/index.php/Reactome_FI_Cytoscape_Plugin

Cytoscape app: <http://apps.cytoscape.org/apps/reactomefiplugin>

Latest source code: <https://github.com/reactome-fi/CytoscapePlugIn>

Source code as at the time of publication: <https://github.com/F1000Research/CytoscapePlugIn>

Archived source code as at the time of publication: <http://www.dx.doi.org/10.5281/zenodo.1038517>

License: the Creative Commons Attribution 3.0 Unported License (http://www.reactome.org/?page_id=362).

Author contributions

GW and LS initiated and guided the project. GW designed the software. GW, ED, and AD implemented the software. GW, ED, RH and LS wrote the paper.

Competing interests

No competing interests were disclosed.

Grant information

This project is supported by a NIH grant (2U41HG003751-05) to LS and a Genome Canada grant (OGI 5458) to LS.

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

Acknowledgements

We would like to thank Irina Kalatskaya, Christina Yung and other members in Dr. Lincoln Stein’s group for software testing and many feedbacks. We also thank Peter D’Eustachio and Irina Kalatskaya for reading and editing the manuscript. We are indebted to Alexander Pico and William Longabaugh from the Cytoscape core development team for reviewing the manuscript and providing many suggestions.

Supplementary results

In this supplementary document, we describe analysis results for the example data set using the 2013 version of the FI network. **Figure S1** shows two modules, module 4 and module 11, having smallest p-values from survival analyses based on the CoxPH model. **Figure S2** shows the Kaplan-Meier plots for Modules 4 and 11. Overlapping analysis (**Table S1**) indicated that 58 genes in the module 3 from the 2009 version of the FI network have been spread into module 4 (18 genes, $p\text{-value} = 5.5 \times 10^{-13}$ based on hypergeometric

test), module 3 (21 genes, $p\text{-value} = 5.8 \times 10^{-8}$), module 26 (2 genes, $p\text{-value} = 0.01$), module 11 (5 genes, $p\text{-value} = 0.02$), module 1 (4 genes, $p\text{-value} = 1.0$), and module 0 (1 gene, $p\text{-value} = 1.0$). It is interesting to see that module 4 has the most significant overlap with module 3 from the 2009 version of the FI network, and also has the most significant p-value from the survival analyses ($p\text{-value} = 1.1 \times 10^{-3}$ from CoxPH), which implies that our method is fairly robust against updates of the FI network.

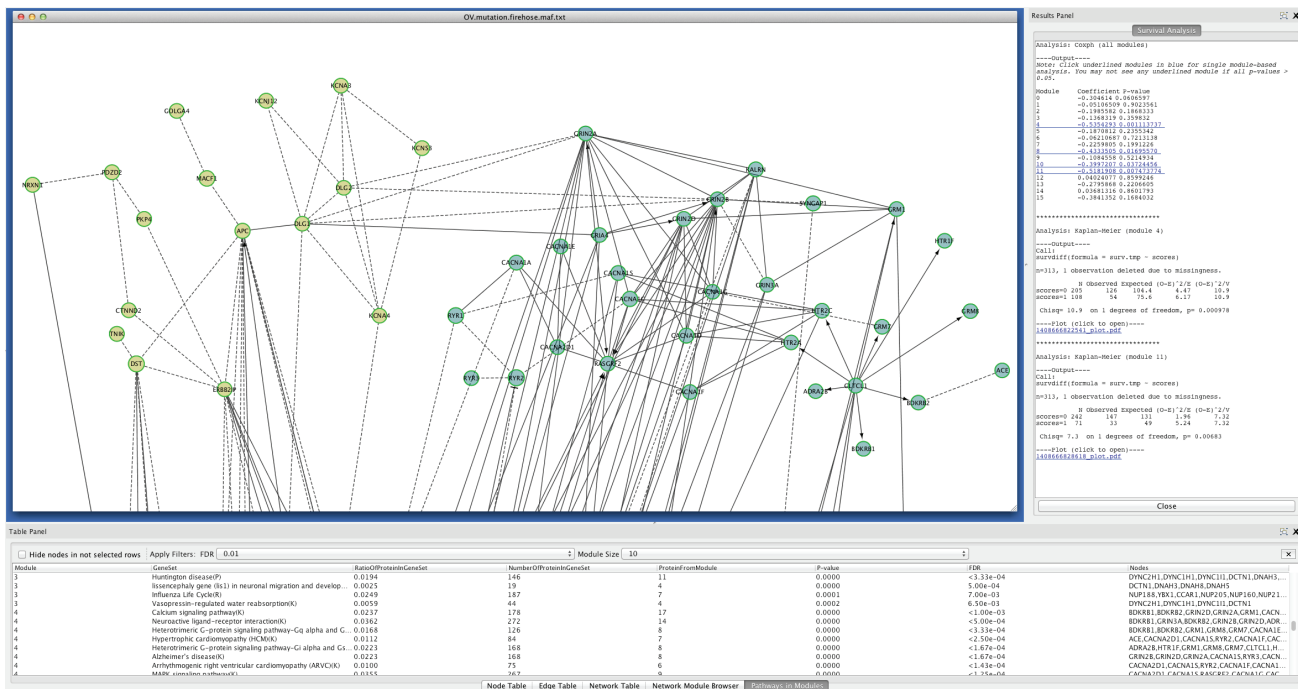


Figure S1. Modules 4 and 11 from the TCGA ovarian cancer mutation data file generated by using the 2013 version of the FI network. The central panel shows the network view of module 4 (right in lightseegreen) and module 11 (left in darkkhaki). The bottom table displays pathway annotations for module 4. The right panel shows survival analysis results using the CoxPH model for modules containing genes no less than 10 and the Kaplan-Meier model for modules 4 and 11.

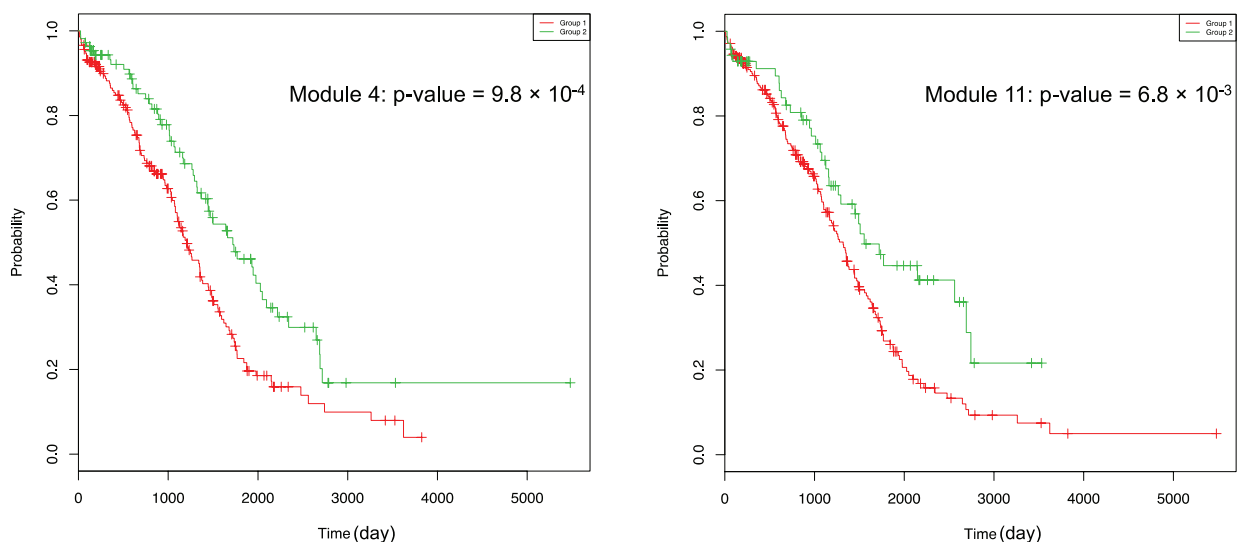


Figure S2. Kaplan-Meier survival plots for modules 4 and 11. P-values are calculated based on the Kaplan-Meier survival model.

Table S1. Distribution of module 3 genes from the 2009 version of the FI network into modules from the 2013 version of the FI network. Column Module is for module indices, Size for numbers of genes contained by modules from the 2013 version of the FI network, Shared for numbers of genes shared between 2009 module 3 and 2013 module, and P-value for significance of sharing based on the hypergeometric test.

Module	Size	Shared	P-value
4	30	19	5.5E-13
3	60	21	5.8E-08
26	2	2	1.2E-02
11	16	5	2.2E-02
1	97	4	1.0E+00
0	110	1	1.0E+00

References

- Barabási AL, Gulbahce N, Loscalzo J: **Network medicine: a network-based approach to human disease.** *Nat Rev Genet.* 2011; **12**(1): 56–68.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Vidal M, Cusick ME, Barabási AL: **Interactome networks and human disease.** *Cell.* 2011; **144**(6): 986–998.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Wu G, Feng X, Stein L: **A human functional protein interaction network and its application to cancer data analysis.** *Genome Biol.* 2010; **11**(5): R53.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Croft D, Mundo AF, Haw R, *et al.*: **The Reactome pathway knowledgebase.** *Nucleic Acids Res.* 2014; **42**(Database issue): D472–7.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Vandin F, Upfal E, Raphael BJ: **Algorithms for detecting significantly mutated pathways in cancer.** *J Comput Biol.* 2011; **18**(3): 507–22.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Vandin F, Clay P, Upfal E, *et al.*: **Discovery of mutated subnetworks associated with clinical data in cancer.** *Pac Symp Biocomput.* 2012; 55–66.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wu G, Stein L: **A network module-based method for identifying cancer prognostic signatures.** *Genome Biol.* 2012; **13**(12): R112.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Huang da W, Sherman BT, Lempicki RA: **Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists.** *Nucl Acids Res.* 2009; **37**(1): 1–13.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Vaske CJ, Benz SC, Sanborn JZ, *et al.*: **Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM.** *Bioinformatics.* 2010; **26**(12): i237–45.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cancer Genome Atlas Research Network: **Integrated genomic analyses of ovarian carcinoma.** *Nature.* 2011; **474**(7363): 609–15.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kleinbaum DG, Klein M: **Survival Analysis: A Self Learning Guide.** New York: Springer. 2005.
[Reference Source](#)
- Cox D: **Regression Models and Life Tables (with Discussion).** *J R Stat Soc B.* 1972; **34**(2): 187–220.
[Reference Source](#)
- Shah SP, Roth A, Goya R, *et al.*: **The clonal and mutational evolution spectrum of primary triple-negative breast cancers.** *Nature.* 2012; **486**(7403): 395–9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Sawey ET, Chanrion M, Cai C, *et al.*: **Identification of a therapeutic strategy targeting amplified FGF19 in liver cancer by Oncogenomic screening.** *Cancer Cell.* 2011; **19**(3): 347–58.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Tamborero D, Gonzalez-Perez A, Perez-Llamas C, *et al.*: **Comprehensive identification of mutational cancer driver genes across 12 tumor types.** *Sci Rep.* 2013; **3**: 2650.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kalatskaya I, Wu G: **Pathway and network based analysis using Functional Interaction network.** *Genomics II Bacteria, Viruses and Metabolic Pathways* iConcept. 2013.
[Reference Source](#)
- Wu G, Dawson E, Duong A, *et al.*: **Reactome FI Cytoscape plugin.** *Zenodo.* 2014.
[Data Source](#)

Open Peer Review

Current Referee Status:



Referee Responses for Version 2



Nikolaus Schultz¹, B. Arman Aksoy²

¹ Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, USA

² Memorial Sloan Kettering Cancer Center, New York, USA

Approved: 22 September 2014

Referee Report: 22 September 2014

doi:[10.5256/f1000research.5631.r6104](https://doi.org/10.5256/f1000research.5631.r6104)

The authors have addressed all of our comments and concerns.

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Referee Responses for Version 1



Nikolaus Schultz¹, B. Arman Aksoy²

¹ Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, USA

² Memorial Sloan Kettering Cancer Center, New York, USA

Approved with reservations: 11 August 2014

Referee Report: 11 August 2014

doi:[10.5256/f1000research.4742.r5319](https://doi.org/10.5256/f1000research.4742.r5319)

In this manuscript, the authors describe details of a Cytoscape 3.0 App: ReactomeFIViz. This app is intended to help users easily investigate genomic alteration data in the context of biological networks. The app allows users to obtain a network from the curated Reactome Functional Interaction database; map mutation, copy-number alteration or gene expression data onto the network; conduct a gene set enrichment analysis or module discovery on the simplified Reactome network; and finally, see the detailed pathway view provided by the Reactome Pathway Browser.

The previous version of the App, which was compatible with Cytoscape 2, was introduced as a supplemental software tool in an earlier study (Wu, Feng and Stein, 2010). Since then, both Cytoscape extensions, the plug-in and the app, have actively been used by users. This has also been evident by the positive reviews the App has received on the Cytoscape App Store. Considering the rich functionality and the ease of cancer genomics analysis that the App provides to the users, we believe this app is of interest to many researchers working in the fields of Computational Biology, Cancer Genomics and Systems

Biology.

There is already extensive documentation about the App on the Reactome web site, however the manuscript fails to provide a general overview for non-experienced users. We have the following suggestions for the authors and if addressed, we believe, these will considerably improve the manuscript:

- In the abstract, the authors say "... *pathways from Reactome and other pathway databases...*". The paper, however, creates the impression that the App is highly dependent on the Reactome infrastructure and does not allow communication with other databases. We suggest that the authors remove "*and other pathway databases*" from the abstract or better clarify this point in the abstract.
- The details about the type of analyses and the statistical tests the App enables are provided in an earlier paper by the same group (Wu, Feng and Stein, 2010), but not in this paper. For readers that are interested in learning these details, we suggest the authors to add a sentence to the manuscript and refer to their earlier work for details. If the functionality and the implementation of these tests have changed since then, we suggest that the authors clearly list these new improved items in the paper for more clarification.
- As an example use case, the authors provide the details of a re-analyzed data set and mention that the results of these two analyses differ due to changes in Reactome FI networks. Curated databases are, of course, subject to changes over time; but it is not clear from the text whether it was the changed network that was causing the problem or the new version of the extension. We suggest that the authors provide the results of these two runs as a supplement to their paper for users to compare and contrast.
- The last sentence of the *Implementation* section says that the analyses were conducted in R, but can the authors clarify the requirements for this App? Do users, for example, need to install R to use this App? Related to this, the authors do not talk about the Cytoscape version that the App targets. Is the Cytoscape 2 Plugin deprecated? Do authors suggest that users install the newer version in Cytoscape 3 as an App?
- Figure 1 provides details about the implementation and the architecture of the ReactomeFIViz App, however we think the manuscript needs a simple user flow diagram that shows where different types of data are obtained, the functionalities of the App and the output the users get. The mutation-based module discovery and differential survival analysis examples mentioned in the paper are good use cases, however it is not clear from the text what the App supports other than these examples.
- Finally, we suggest the authors to provide a supplemental step-by-step guide to replicate the results that they describe in the paper. For a new user, this may provide a good base to start using the software and for many researchers it might be more convenient to have such an article provided as a supplementary file to this paper.

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

Competing Interests: No competing interests were disclosed.

1 Comment

Author Response

Guanming Wu,

Posted: 06 Sep 2014

Thanks a lot for your review of our article. We have made changes according to your thoughtful suggestions. Please see details below:

- *In the abstract, the authors say "... pathways from Reactome and other pathway databases...". The paper, however, creates the impression that the App is highly dependent on the Reactome infrastructure and does not allow communication with other databases. We suggest that the authors remove "and other pathway databases" from the abstract or better clarify this point in the abstract.*

Although the app uses a software infrastructure running on Reactome, the curated pathways it uses for the network and pathway enrichment analysis come in equal parts from Reactome and other pathway databases. We have clarified this in the abstract.

- *The details about the type of analyses and the statistical tests the App enables are provided in an earlier paper by the same group (Wu, Feng and Stein, 2010), but not in this paper. For readers that are interested in learning these details, we suggest the authors to add a sentence to the manuscript and refer to their earlier work for details. If the functionality and the implementation of these tests have changed since then, we suggest that the authors clearly list these new improved items in the paper for more clarification.*

We have added references to our original FI network paper in the paragraph introducing "Network analysis features". Also, we added the date that version 4.0.0.beta was released in order to indicate that pathway analysis features are new features that have not been covered by our previous work.

- *As an example use case, the authors provide the details of a re-analyzed data set and mention that the results of these two analyses differ due to changes in Reactome FI networks. Curated databases are, of course, subject to changes over time; but it is not clear from the text whether it was the changed network that was causing the problem or the new version of the extension. We suggest that the authors provide the results of these two runs as a supplement to their paper for users to compare and contrast.*

We modified the sentence in the last paragraph describing different results using different versions of the FI network to make it clear that the same software, but different versions of the FI network, were used. We also deleted this sentence in the implementation section, "Each version is handled by its own web application on the server-side for easy software maintenance.", to avoid confusing the readers with technical details.

As suggested, added a supplementary document to describe results from the 2013 version of the FI network.

- *The last sentence of the Implementation section says that the analyses were conducted in R, but can the authors clarify the requirements for this App? Do users, for example, need to*

install R to use this App? Related to this, the authors do not talk about the Cytoscape version that the App targets. Is the Cytoscape 2 PlugIn deprecated? Do authors suggest that users install the newer version in Cytoscape 3 as an App?

We have clarified the section to make the requirements clearer, and added a new paragraph to address the relationship of the application to Cytoscape 2 and 3.

- *Figure 1 provides details about the implementation and the architecture of the ReactomeFIViz App, however we think the manuscript needs a simple user flow diagram that shows where different types of data are obtained, the functionalities of the App and the output the users get. The mutation-based module discovery and differential survival analysis examples mentioned in the paper are good use cases, however it is not clear from the text what the App supports other than these examples.*

We have added a new paragraph at the top of the Results section to highlight some major features in the App, and created a new diagram to show these major features as suggested.

- *Finally, we suggest the authors to provide a supplemental step-by-step guide to replicate the results that they describe in the paper. For a new user, this may provide a good base to start using the software and for many researchers it might be more convenient to have such an article provided as a supplementary file to this paper.*

The application's web-based tutorial already has provided very detailed, step-by-step instructions that replicate the results. We have added a new sentence in the Data availability section to point this out, and the README file contained in the zip file for downloading now points readers interested in replicating the analysis to the online tutorial.

Competing Interests: No competing interests were disclosed.



Hamid Bolouri

Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Approved: 23 July 2014

Referee Report: 23 July 2014

doi: [10.5256/f1000research.4742.r5320](https://doi.org/10.5256/f1000research.4742.r5320)

This article describes version 4.1.0 of the Reactome Functional Interaction "plugin"/"app" for the Cytoscape network exploration/analysis Java platform.

The Reactome Functional Interaction network is an established, well-received and widely used resource. The software described here provides useful resources both for exploring Reactome networks and for using them in network enrichment analyses.

My only reservation about this article is that its brevity makes it sketchy. For example, how are pathway enrichment statistics calculated? Does the implementation of HotNet offered here differ in any way from

the original?

I am especially intrigued that the example use-case given is reported to work well with the 2009 version of the ReactomeFI network, but apparently the discovered modules cannot be found with the newer versions of the FI because the underlying interactions have been "*spread into several modules in the newer version of the FI network*". Why does this happen and how can users guard against it?

I was also not able to explore/reproduce the example presented in the paper, because (contrary to its ".txt" extension) the MAF file provided did not appear to be plain text and the needed survival data file was not included in the zip file provided. A 'README' file with instructions may be useful here.

P.S. The multiple names given to this resource - "ReactomeFIViz (also called the Reactome FI Cytoscape app or ReactomeFIPlugIn)" - are a little confusing and seem unnecessary.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

1 Comment

Author Response

Guanming Wu,

Posted: 06 Sep 2014

Thanks a lot for your comments. We have made changes according to your comments. Please see details below:

- *My only reservation about this article is that its brevity makes it sketchy. For example, how are pathway enrichment statistics calculated? Does the implementation of HotNet offered here differ in any way from the original?*

We have now stated that a binomial test is used for pathway enrichment analysis in the "Pathway analysis features" section and cited a reference (reference 8) for it. We also added a sentence to point out that our implementation of HotNet was done by porting the original Python and MatLab code to Java and R, and so the original algorithm is unchanged.

- *I am especially intrigued that the example use-case given is reported to work well with the 2009 version of the ReactomeFI network, but apparently the discovered modules cannot be found with the newer versions of the FI because the underlying interactions have been "spread into several modules in the newer version of the FI network". Why does this happen and how can users guard against it?*

Because new interactions have been added into the latest version of the FI network, these new interactions may change the network clustering results. However, we still find some network modules, which overlap significantly with the original module and have significant p-values, though higher than the original one, from survival analysis. As suggested by another reviewer, we have added a supplementary document to describe results from the 2013 version of the FI network. As stated in the manuscript, we recommend that the user should use the latest version of the FI network, and check with previous versions of the FI network to investigate the stability of found network modules.

- *I was also not able to explore/reproduce the example presented in the paper, because (contrary to its ".txt" extension) the MAF file provided did not appear to be plain text and the needed survival data file was not included in the zip file provided. A 'README' file with instructions may be useful here.*

We apologize for the problem. It turned out that the zip file was collapsed somehow. We have fixed the problem, and added a simple README file as suggested.

- *The multiple names given to this resource - "ReactomeFIViz (also called the Reactome FI Cytoscape app or ReactomeFIPlugin)" - are a little confusing and seem unnecessary.*

Because of the development history of this software, the same app has been called different names. For example, prior to Cytoscape 3, all Cytoscape extensions were called "plug-ins", but the nomenclature has now been changed to "app". We have tried to minimize the confusion by using ReactomeFIViz consistently and referring to the names of earlier versions of the software just once.

Competing Interests: No competing interests were disclosed.