SOFTWARE TOOL ARTICLE

Automation of ReactomeFIViz via CyREST API [version 1; peer review: 1 approved, 1 approved with reservations]

Fred Loney¹, Guanming Wu²

¹Knight Cancer Institute, Oregon Health & Science University, Portland, OR, USA

²Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, OR, USA

v1 First published: 02 May 2018, 7:531
<https://doi.org/10.12688/f1000research.14776.1>
 Latest published: 23 May 2018, 7:531
<https://doi.org/10.12688/f1000research.14776.2>

Abstract

Pathway- and network-based approaches project seemingly unrelated genes onto the context of pathways and networks, enhancing the analysis power that cannot be achieved via gene-based approaches. Pathway and network approaches are routinely applied in large-scale data analysis for cancer and other complicated diseases. ReactomeFIViz is a Cytoscape app, providing features for researchers to perform pathway- and network-based data analysis and visualization by leveraging manually curated Reactome pathways and highly reliable Reactome functional interaction network. To facilitate adoption of this app in bioinformatics software pipeline and workflow development, we develop a CyREST API for ReactomeFIViz by exposing some major features in the app. We describe a use case to demonstrate the use of this API in a Python-based notebook, and believe the new API will provide the community a convenient and powerful tool to perform pathway- and network-based data analysis and visualization using our app in an automatic way.

Keywords

ReactomeFIViz, CyREST, Cytoscape, Reactome, Pathway and network

Open Peer Review

Approval Status

	1	2
version 2		
(revision)		
23 May 2018		
version 1		
02 May 2018		

1. **Marina Piccirillo** , National Research Council, Naples, Italy

2. **John H. Morris** , University of California, San Francisco, San Francisco, 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 **Cytoscape** gateway.

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

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

Author roles: **Loney F:** Software, Visualization, Writing – Review & Editing; **Wu G:** Conceptualization, Funding Acquisition, Methodology, Project Administration, Software, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project is supported by a NIH grant (5U41HG003751).

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

Copyright: © 2018 Loney F and Wu G. 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: Loney F and Wu G. **Automation of ReactomeFIViz via CyREST API [version 1; peer review: 1 approved, 1 approved with reservations]** F1000Research 2018, 7:531 <https://doi.org/10.12688/f1000research.14776.1>

First published: 02 May 2018, 7:531 <https://doi.org/10.12688/f1000research.14776.1>

Introduction

Pathway- and network-based computational approaches are now routinely used in large-scale data analysis to uncover hidden patterns that are otherwise impossible to discover. These approaches project significant genes, proteins, metabolites, and other kinds of biological entities collected from pre-analysis onto pathways and networks, knowledge produced by many years' experimental studies. **Cytoscape**¹ is the most popular biological network visualization and analysis platform, widely used in the research community to perform pathway and network analysis and visualization. The release of the CyREST app² enables Cytoscape as an integrative and indispensable tool to build automatic software pipeline and workflow in programming languages widely used by the bioinformatics and computational biology community, including **Python** and **R**, via a RESTful API. The standalone Java-based Cytoscape application thereby functions as a microservice servlet exposing the major features of Cytoscape.

Reactome³ is the most comprehensive open source biological pathway knowledgebase, widely used in the research community, with its web site accessed by roughly 60,000 unique IP addresses per month. To perform genome-scale network-based data analysis and visualization, we have also constructed a highly reliable Reactome functional interaction (FI) network by extracting FIs from manually curated pathways from Reactome and other popular large-scale pathway databases and predicting FIs based on a machine learning approach⁴. Based on this FI network and the high quality Reactome pathways, we have developed a Cytoscape app, called "ReactomeFIViz"⁵, which is one of most popular Cytoscape apps, downloaded over 30,000 times since it was released in September, 2013 into Cytoscape app store.

ReactomeFIViz provides a suite of features to help users to perform pathway- and network-based data analysis and visualization for cancer and other complicated diseases. Users can construct a FI subnetwork based on a set of genes, perform network clustering, annotate found network modules using Reactome pathways and Gene Ontology terms, and perform survival analysis if clinical data is available to search for gene signatures related to patient overall survival. Users can also explore Reactome pathways directly inside Cytoscape, perform pathway enrichment analysis for a set of genes, and conduct pathway modeling using multiple types of omics data based on factor graphs converted automatically from Reactome pathways. Recently we added new features for users to visualize FDA approved cancer drugs and their targets interactions in the contexts of Reactome pathways and the FI network, and perform fuzzy logic based modeling to study the effects of drug application on the pathway activities (see **ReactomeFIViz** wiki page).

To facilitate third-party software tool developers to integrate the powerful network and pathway analysis features provided in ReactomeFIViz, we implemented a new CyREST API. The current version of this API is focused on FI network construction and Reactome pathway enrichment analysis for a set of genes.

Methods

To develop the CyREST API for ReactomeFIViz, we followed the recommended procedures described in **Cytoscape** wiki on **adding Automation to existing apps**. To handle the complex data models used in ReactomeFIViz, we chose the Functions over Commands approach by adding JAX-RS resource onto existing ReactomeFIViz code base. In brief, we added two new Java packages, `org.reactome.cytoscape.rest` and `org.reactome.cytoscape.rest.tasks`, and refactored original tasks into **ObservableTask**. All refactored ObservableTasks are grouped into package `org.reactome.cytoscape.rest.tasks`, and their execution is managed by a `SynchronousTaskManager` object and monitored by their respective `TaskObserver` objects.

The ReactomeFIViz CyREST API is specified in a Java interface, `ReactomeFIVizResource`, and documented using **Swagger UI** as Java annotations for methods defined in the interface. The implementation of `ReactomeFIVizResource` is provided in class `ReactomeFIVizResourceImp`. Both the interface and the implementation are placed in package `org.reactome.cytoscape.rest`.

The ReactomeFIViz CyREST API is powered up by the CyREST app using its embedded light-weight Grizzly HTTP server. CyREST delegates all RESTful API calls to ReactomeFIViz, which then calls the ReactomeFIViz RESTful server via its RESTful API. The ReactomeFIViz server fetches the Reactome content from databases hosted in a MySQL database engine via a **Hibernate** API and the in-house built **Reactome Java API** (Figure 1).

Operation

Currently, the ReactomeFIViz CyREST API provides 8 methods (Table 1). These methods allow third-party workflow and pipeline developers to construct a FI sub-network based on a set of genes listed in a variety of file formats, annotate displayed network using collected pathways, GO biological process, molecular function, or cellular component terms, perform network cluster and then annotate network modules. These methods also allow them to perform Reactome pathway enrichment analysis for a set of genes and then export pathway diagrams. The CyREST API document for ReactomeFIViz, which is accessed via menu **Help/Automation/CyREST** API, provides detailed description about all these resources.

Figure 1. Software architecture for ReactomeFIViz CyREST API.

Table 1. List of ReactomeFIViz CyREST API.

URL	HTTP Method	Function	Note
fiVersions	GET	List supported versions of the FI network	Usually three versions should be listed
buildFISubNetwork	POST	Build a FI subnetwork for a set of genes	Multiple file formats are supported
enrichment/{type}	GET	Perform network enrichment analysis	Four types are supported, which is specified by type in URL: Pathway, GO BP ¹ , MF ² , and CC ³
cluster	GET	Perform network clustering analysis	Clustering is performed for the displayed network
moduleEnrichment/{type}	GET	Perform network module enrichment analysis	Four types are supported, which is specified by type in URL: Pathway, GO BP, MF, and CC
pathwayTree	GET	Show the Reactome pathway tree	All pathways are returned from this resource
ReactomePathwayEnrichment	POST	Perform pathway enrichment analysis for a set of genes	pathwayTree must be called first
exportPathwayDiagram	POST	Export a pathway diagram as a PDF file	Only pathways with hasDiagram values equal to true can export diagrams.

Notes: 1. GO BP: Gene Ontology biological process; 2. MF: Molecular function; 3. CC: Cellular component.

Results

ReactomeFIViz CyREST API provides a set of URL-based language-neutral functions, which accept parameters and return results in the JSON format. As with any other CyREST API, it can be easily integrated into Python, R, or any other programming language as long as it supports HTTP-based function calling. Here we describe a use case based on [The Jupyter Notebook](#) to showcase the usage of this API in a workflow development.

Workflow

Previous study⁶ has shown the genomic similarity between high-grade serous ovarian tumors and basal-like breast tumors based on multiple omics data types, including copy number variants (CNVs), somatic mutations, and mRNA gene expression. To demonstrate use of the ReactomeFIViz CyREST API, we perform a network- and pathway-based comparison analysis between genes having somatic mutations in TCGA ovarian cancer and breast cancer. Our analysis is focused on showing the utility of our API. Therefore, we use all TCGA breast cancer samples without subtyping to simplify our workflow.

[Figure 2](#) shows the workflow of this use case. The TCGA BRCA (breast invasive carcinoma) and OV (ovarian serous cystadenocarcinoma) mutation data was downloaded from the [Broad firehose web site](#) in the mutation annotation file (MAF) format using its RESTful API, and stored in two local files, one for each cancer type. The MAF was then loaded into Cytoscape via the CyREST API to construct a FI sub-network after choosing a

sample cutoff to select genes forming a network composed of about 500 genes. The FI-network was then subject to network clustering analysis. Two sets of network modules from network clustering were compared to find modules shared and not shared between these two cancers. Pathway enrichment analysis was performed to collect pathways not shared between them. These results suggest common and cancer-specific network and pathway patterns, facilitating researchers to understand shared and specific oncogenesis mechanisms in these two cancers.

Analysis results

We performed a network comparison study for the two FI sub-networks constructed separately for TCGA BRCA and OV mutated genes. Network clustering yielded 23 and 38 modules with the TCGA BRCA and OV FI subnetwork, respectively. Overlapping analysis for modules having no less than 3 genes showed 16 out of 18 BRCA modules significantly overlapped with 15 out of 22 OV modules (FDR < 0.01. Results not shown here. See the output in the notebook for details), implying that almost all of BRCA modules can be found in the OV subnetwork and some of BRCA modules are shared with more than one OV module. For example, BRCA Module 1 is significantly overlapped with OV Modules 2 and 3.

For the pathway comparison study, we focused on pathways enriched for network modules not shared between BRCA and OV, searching for possible different oncogenic mechanisms for these two cancers. The module overlapping analysis showed that

Figure 2. The workflow of the comparison analysis use case using TCGA OV and BRCA mutation data.

BRCA modules 13 and 17, which contains 6 and 4 genes, respectively, are not significantly shared with any OV module. Pathway annotation for these two modules revealed genes in these two modules are significantly enriched for some signaling pathways, including **DAG and IP3 signaling**, **Signaling by GPCR**, and **Glucagon signaling in metabolic regulation**. There are 7 modules in the OV subnetwork that are not significantly shared with the BRCA subnetwork. Pathway enrichment analysis showed that genes in these modules are significantly involved in **Ion channel transport**, **O-linked glycosylation**, **C-type lectin receptors**, and several others. For detailed results, see the output from the notebook.

To visualize mutated genes in the context of Reactome pathways, the notebook also generated two pathway diagrams, one for each cancer, and saved into the working directory as PDF files. Entities in pathway diagrams composed of mutated genes are highlighted in purple.

Discussion

Reactome provides a large set of high-quality manually curated pathways. The Reactome FI network provides a genome-scale highly reliable functional interaction network covering over 60% of total human genes. The ReactomeFIViz CyREST API delivers language neutral REST-based functions for third-party software developers to leverage high-valued resources provided by the Reactome project in their own software tools.

The current set of functions implemented in this version of ReactomeFIViz CyREST API focuses on some major features implemented in ReactomeFIViz, related to FI network construction, clustering, and Reactome pathway enrichment analysis. As shown in the above use case, it is very easy to integrate with other CyREST APIs and integrated into a Python or R programming language environment to perform Reactome-related pathway and network analysis and visualization.

We will expose other ReactomeFIViz features in the CyREST API, including gene expression data analysis, network module-based survival analysis, pathway modeling based on Boolean network and probabilistic graph model, and cancer drug visualization

and simulation. We will also develop a Python package for easy third-party tool integration.

Software availability

Home page for user guide: <https://reactome.org/tools/reactome-fiviz>

Cytoscape app store: <http://apps.cytoscape.org/apps/reactomefi-plugin>

Latest source code: <https://github.com/reactome-fi/CytoscapePlugin/tree/path-x>

Use case Jupyter notebook: <https://github.com/reactome-fi/workflows>

Source code at the time of publication: https://github.com/reactome-fi/CytoscapePlugin/releases/tag/f1000_auto_paper for ReactomeFIViz, https://github.com/reactome-fi/workflows/releases/tag/f1000_auto_paper for the Python use case notebook

Archived source code at the time of publication: <http://doi.org/10.5281/zenodo.1226433>⁷ for ReactomeFIViz, <http://doi.org/10.5281/zenodo.1226430>⁸ for the Python use case notebook

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

Data availability

Output from the network comparison are available in the Python use case notebook <http://doi.org/10.5281/zenodo.1226430>⁸

Grant information

This project is supported by a NIH grant (5U41HG003751).

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

Acknowledgments

We would like to thank the whole Reactome team for providing the foundation for ReactomeFIViz.

References

1. Shannon P, Markiel A, Ozier O, *et al.*: **Cytoscape: a software environment for integrated models of biomolecular interaction networks**. *Genome Res.* 2003; **13**(11): 2498–504.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
2. Ono K, Muetze T, Kollishovski G, *et al.*: **CyREST: Turbocharging Cytoscape Access for External Tools via a RESTful API [version 1; referees: 2 approved]**. *F1000Res.* 2015; **4**: 478.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
3. Fabregat A, Jupe S, Matthews L, *et al.*: **The Reactome Pathway Knowledgebase**. *Nucleic Acids Res.* 2018; **46**(D1): D649–D655.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
4. 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](#)
5. Wu G, Dawson E, Duong A, *et al.*: **ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis [version 2; referees: 2 approved]**. *F1000Res.* 2014; **3**: 146.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. Cancer Genome Atlas Network: **Comprehensive molecular portraits of human breast tumours**. *Nature.* 2012; **490**(7418): 61–70.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. Wu G: **reactome-fi/CytoscapePlugin: F1000Research ReactomeFIViz Automation (Version f1000_auto_paper)**. *Zenodo.* 2018.
[Data Source](#)
8. Loney F, Wu G: **reactome-fi/workflows: F1000Research ReactomeFIViz Automation (Version f1000_auto_paper)**. *Zenodo.* 2018.
[Data Source](#)

Open Peer Review

Current Peer Review Status:

Version 1

Reviewer Report 11 May 2018

<https://doi.org/10.5256/f1000research.16077.r33666>

© 2018 Morris J. 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.

John H. Morris

Resource for Biocomputing, Visualization and Informatics, Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA, USA

In this article, the authors have described their extensions to the popular ReactomeFIViz Cytoscape App to support automation via CyREST. They do an excellent job providing the motivation for their extension. The manuscript was clear and well-written, and I believe that the new REST API's will be quite useful.

I do have a problem with the manuscript, however, in that in the Workflow section, they state:

"The MAF was then loaded into Cytoscape via the CyREST API to construct a FI sub-network after choosing a sample cutoff to select genes forming a network composed of about 500 genes. The FI-network was then subject to network clustering analysis. Two sets of network modules from network clustering were compared to find modules shared and not shared between these two cancers. Pathway enrichment analysis was performed to collect pathways not shared between them."

My problem with this is that they don't provide us with enough detail on the actual CyREST calls and parameters they used to allow us to understand how to reproduce it. Which part of their analysis was done in Python vs. which was done in Cytoscape through the GUI. It is true that all of my answers exist in the Jupyter notebook they provided, but that's a lot of work to figure out that they used the "buildFISubNetwork" method with the "MAF" format the fiVersion "2016". My suggestion is to add at least enough information either in the figure or the text to help others understand the REST flow for their App without referring to the notebook.

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

Yes

Is the description of the software tool technically sound?

Yes

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

Partly

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

Yes

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.

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.

Author Response 20 May 2018

Guanming Wu

Dear Dr. Morris,

Thanks a lot for reviewing our paper and your comment. In version 2 of our paper, we have added a new table listing detailed information about actions performed in the use case workflow (Table 2), and added new words in the main text (the second paragraph in the Workflow section) to help readers understand how to reproduce the workflow without going to the Jupyter notebook. We hope our revisions have addressed your concern.

Best regards,

Guanming Wu, Ph.D.

Competing Interests: No competing interests were disclosed.

Reviewer Report 10 May 2018

<https://doi.org/10.5256/f1000research.16077.r33745>

© 2018 Piccirillo M. 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.

Marina Piccirillo

Lab-GTP, High Performance Computing and Networking Institute, National Research Council, Naples, Italy

The authors describe the development of CyREST API for ReactomeFIViz, a Cytoscape app, which exposes some Reactome functions as REST APIs for external software components, to process pathway and network data in automatic and reproducible workflows built using almost any programming languages. The source code provided for ReactomeFIViz and for the Python use case notebook, make the use of CyREST API easier and simpler to understand.

The manuscript is well-organized and the described app will be highly valuable for users working with big data related to enrichment analysis and visualization of networks.

I have only some minor comments:

1. In the Introduction section (pag.2): the sentence “These approaches project significant genes, proteins, metabolites, and other kinds of biological entities collected from pre-analysis onto pathways and networks, knowledge produced by many years’ experimental studies.”, is not quite understandable, I think it should be rephrased.
2. The Methods section (pag.2) is clearly explained and sufficient details are provided to allow replication of the method development and its use by others.
3. The section Operation (pag.2) I think is unnecessary, the authors can add it as the subsection of Methods because here they are explaining the methods available in the CyREST API developed.
4. In the Analysis results subsection (pag.3): I think it's possible to add a figure (such as a Venn Diagram) to better explain the results obtained.
5. In the same subsection (Analysis results: pag.3) more details should be provided about the results. (For example:
 - in BRCA modules 13 and 17 are there overlapping genes? If yes, which are their names?
 - What is the way, in which the enriched signaling pathways are involved in BRCA and/or OV carcinoma?)
6. More references to supporting literature must be provided in the manuscript in order to further substantiate the claims.

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

Yes

Is the description of the software tool technically sound?

Yes

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?

Yes

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: Bioinformatics and computational biology

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.

Author Response 20 May 2018

Guanming Wu

Dear Dr. Piccirillo,

Thanks a lot for reviewing our paper and your comments. We have made changes to address some of your comments. Please see details below:

1). In the Introduction section (pag.2): the sentence "These approaches project significant genes, proteins, metabolites, and other kinds of biological entities collected from pre-analysis onto pathways and networks, knowledge produced by many years' experimental studies.", is not quite understandable, I think it should be rephrased.

We have rephrased this sentence and hope the meaning is clear now.

2). The Methods section (pag.2) is clearly explained and sufficient details are provided to allow replication of the method development and its use by others.

Thanks!

3). The section Operation (pag.2) I think is unnecessary, the authors can add it as the subsection of Methods because here they are explaining the methods available in the CyREST API developed.

This section is suggested by the article template provided by the journal editor.

4). In the Analysis results subsection (pag.3): I think it's possible to add a figure (such as a Venn Diagram) to better explain the results obtained.

Thanks for the suggestion. However, we found that it is difficult to show this type of overlapping result in a Venn diagram. We hope that our text description is clear.

5). In the same subsection (Analysis results: pag.3) more details should be provided about the results. (For example:

- in BRCA modules 13 and 17 are there overlapping genes? If yes, which are their names?

We have added another sentence to address your question in the second paragraph of the Analysis Result section.

- What is the way, in which the enriched signaling pathways are involved in BRCA and/or OV carcinoma?)

This is a great question. However, this short software article is written to showcase how to use the new ReactomeFIViz CyREST API. We have refrained ourselves to discuss the biological implications of our analysis results. However, the exported pathway diagrams may answer questions like this, which will be addressed in our future projects.

6). More references to supporting literature must be provided in the manuscript in order to further substantiate the claims.

Thanks for the comment. Along with listed references, we have also linked to external resources using hyper-links in the main text. We feel that we have cited enough supporting references for this on-line short software article.

Best regards,

Guanming Wu, Ph.D.

Competing Interests: No competing interests were disclosed.

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