Check for updates

SOFTWARE TOOL ARTICLE

CyTargetLinker app update: A flexible solution for network extension in Cytoscape [version 1; peer review: 1 approved, 1 not approved]

Martina Kutmon [ID] [1,2], Friederike Ehrhart [ID] [1,3], Egon L. Willighagen [ID] [1], Chris T. Evelo [ID] [1,2], Susan L. Coort [ID] [1]

[1]Department of Bioinformatics - BiGCaT, NUTRIM, Maastricht University, Maastricht, 6229 ER, The Netherlands
[2]Maastricht Centre for Systems Biology (MaCSBio), Maastricht University, Maastricht, 6229 ER, The Netherlands
[3]GKC-Rett Expertise Centre, Maastricht University Medical Center, Maastricht, 6200 MD, The Netherlands

Abstract

Here, we present an update of the open-source CyTargetLinker app for Cytoscape (http://apps.cytoscape.org/apps/cytargetlinker) that introduces new automation features. CyTargetLinker provides a simple interface to extend networks with links to relevant data and/or knowledge extracted from so-called linksets. The linksets are provided on the CyTargetLinker website or can be custom-made for specific use cases. The new automation feature enables users to programmatically execute the app's functionality in Cytoscape (command line tool) and with external tools (e.g. R, Jupyter, Python, etc). This allows users to share their analysis workflows and therefore increase repeatability and reproducibility. Three use cases demonstrate automated workflows, combinations with other Cytoscape apps and core Cytoscape functionality. We first extend a protein-protein interaction network created with the stringApp, with compound-target interactions and disease-gene annotations. In the second use case, we created a workflow to load differentially expressed genes from an experimental dataset and extend it with gene-pathway associations. Lastly, we chose an example outside the biological domain and used CyTargetLinker to create an author-article-journal network for the five authors of this manuscript using a two-step extension mechanism. With 300 downloads per month in the last year and over 12,000 downloads in total, CyTargetLinker shows the adoption and relevance of the app in the field of network biology. In April 2018, the original publication was cited in 57 articles demonstrating the applicability in biomedical research.

Keywords

Cytoscape, CyTargetLinker, network extension, network visualization, regulatory networks, data integration

Open Peer Review

Reviewer Status ✓ ✓

	Invited Reviewers	
	1	**2**
REVISED **version 2** published 13 Aug 2019		✓ report
		↑
version 1 published 14 Jun 2018	✓ report	✗ report

1 **Alex Gutteridge** [ID], GlaxoSmithKline (GSK), Brentford, UK

2 **John H. Morris** [ID], 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 ELIXIR gateway.

This article is included in the Cytoscape Apps gateway.

Corresponding author: Martina Kutmon (martina.kutmon@maastrichtuniversity.nl)

Author roles: Kutmon M: Conceptualization, Data Curation, Methodology, Project Administration, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Ehrhart F**: Data Curation, Formal Analysis, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Willighagen EL**: Data Curation, Writing – Review & Editing; **Evelo CT**: Conceptualization, Supervision; **Coort SL**: Conceptualization, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

How to cite this article: Kutmon M, Ehrhart F, Willighagen EL *et al*. **CyTargetLinker app update: A flexible solution for network extension in Cytoscape [version 1; peer review: 1 approved, 1 not approved]** F1000Research 2018, **7**:743 (https://doi.org/10.12688/f1000research.14613.1)

First published: 14 Jun 2018, **7**:743 (https://doi.org/10.12688/f1000research.14613.1)

Introduction

The Cytargetlinker app provides a flexible and simple way to extend networks in Cytoscape[1] with links to (prior) knowledge from external sources. Since its first release in 2013[2], CyTargetLinker has been downloaded more than 12,000 times and used in numerous studies. These applications in biological studies range from the creation of a microRNA-gene association network for lipid diseases[3] or Alzheimer's disease[4] to the application of algorithms for drug sensitivity prediction[5].

While the app was originally intended to be used for the extension of biological networks with regulatory interactions, researchers have used CyTargetLinker to integrate knowledge about many different types of relationships (e.g. pathway associations and disease annotations). Therefore, we renamed the previously used *Regulatory Interaction Networks* (RegINs) to *linksets* to make the broader applicability more explicit. Moreover, the generation of linksets, either manually or in an automated manner, has become more user-friendly.

In this new version of CyTargetLinker, we introduce an automation feature that allows programmatic access to the app functionality. In the Results section, we present three use cases that highlight the app's purpose, how it can be easily combined with other Cytoscape apps and the advantages of the automation. Whereas the first two use cases have a biological nature, the third use case demonstrates the broader applicability with a non-biological example.

Methods

The newest version of CyTargetLinker (4.0.0+) was developed for Cytoscape (3.6.0+) which introduces a new interface for automation that can make apps callable as services by the Cytoscape Command scripts, Python and R. This promotes open and reproducible data analysis, and simple integration with other apps. CyTargetLinker can be installed through the Cytoscape app store.

LinkSets

On the CyTargetLinker website, we provide a variety of linksets for regulatory interactions, pathway associations and disease annotations (https://projects.bigcat.unimaas.nl/cytargetlinker/linksets/). Additionally, we deliver a simple Java program to convert tab delimited text files into XGMML linksets that can be used with CyTargetLinker (https://github.com/CyTargetLinker/linksetCreator). Using BridgeDb[6], a framework for finding and mapping database identifiers, the script enables the support of multiple identifier systems for biological entities.

Application programming interface

While CyTargetLinker can still be used through the Cytoscape graphical user interface (see online tutorials), we would like to highlight the novel application programming interface (API) that allows the programmatic execution of the app's functionality.

CyTargetLinker provides a set of API methods to automise the extension of networks, see Table 1. The key function is the "extend" function, which parses the provided LinkSets and extracts relevant interactions for the selected network. The user can then choose to use the CyTargetLinker visual style and the force-directed layout. Often, users want to integrate knowledge for the same interaction type from different resources. With the "filterOverlap" function, users can visualise only those interactions that are supported by multiple resources.

Use cases

The broad applicability of CyTargetLinker will be demonstrated in three different use cases. The focus lies on the automation of the analysis and the R scripts for each use case are provided in the tutorial repository on GitHub. We chose to present two biological and one non-biological use cases to demonstrate the flexibility of the app.

Use cases highlighting the new automation functionality

Use case 1: Investigating drug-targets and disease associations for a Rett syndrome protein-protein interaction network. Rett syndrome is a rare disease caused by a mutation in the methyl-CpG-binding protein 2 (MECP2) gene[7]. In this use case, we used the stringApp[8] of Cytoscape to create a protein-protein interaction (PPI) network for Rett syndrome (Disease Query). The PPI network is then extended using

Table 1. **CyTargetLinker API.** List of the API methods of CyTargetLinker, their parameters and a general description.

API Method	Parameters	Task
version		Returns the version of the CyTargetLinker app
extend	direction, idAttribute, linkSetDirectory, linkSetFiles, network	Extend network with interactions from linksets
applyVisualstyle	network	Apply the CyTargetLinker visual style to the extended network
applyLayout	network	Apply a force-directed layout
filterOverlap	network, overlap	Filter extension and show only links supported by *n* linksets
showPanel	show	Show the CyTargetLinker result panel

CyTargetLinker with compound-target interactions from ChEMBL[9,10] and disease-gene associations from a manually curated subset for rare diseases from OMIM[11]. ChEMBL is an open online bioactivity database containing information about compounds, their bioactivity and their possible targets (including proteins). OMIM is a comprehensive collection of human genetic phenotypes and their associated human genes. Both linksets are available on the linkset download page on the CyTargetLinker website.

First, the stringApp was used to create a Rett syndrome PPI (query="Rett syndrome", cutoff=0.4, limit=20). Using CyTargetLinker, the network was extended with 37 compound-target interactions from ChEMBL and 18 gene-disease associations from OMIM (see Figure 1).

Use case 2: Pathway associations for differentially expressed genes in Rett syndrome. For this use case, we selected a list of differentially expressed genes in the Purkinje cells located in the cerebellar cortex of the brain of a Mecp2$^{-/y}$ mouse model[12,13] for Rett syndrome. Next, we investigated in which biological processes these altered genes are involved. Using the pathway annotations from the WikiPathways database[14], CyTargetLinker adds the pathway information and creates a pathway-gene network.

From the dataset, we extracted 65 genes with an absolute log2 fold change larger than 1. Only 16 genes are present in one or more pathways of the curated mouse pathway collection from WikiPathways. Figure 2 shows the resulting gene-pathway network. Genes without pathway annotations have been removed.
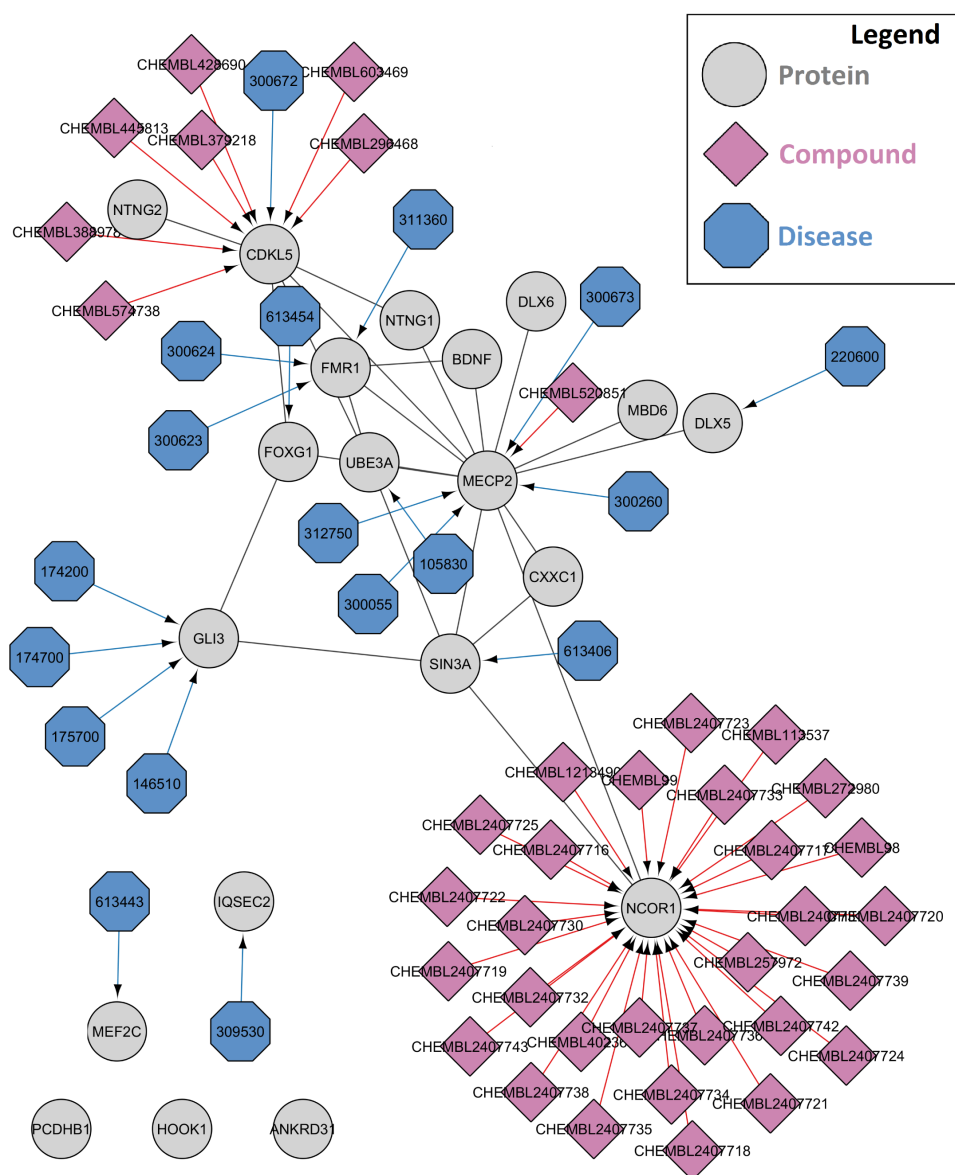

Figure 1. Compound-gene-disease network for Rett Syndrome. The protein-protein interaction network for Rett syndrome was created using the disease query option of the stringApp for Cytoscape. The proteins are represented as gray circles. Then, CyTargetLinker was used to extend the network with compounds from ChEMBL (purple diamonds) and disease annotations from OMIM (blue octagons).
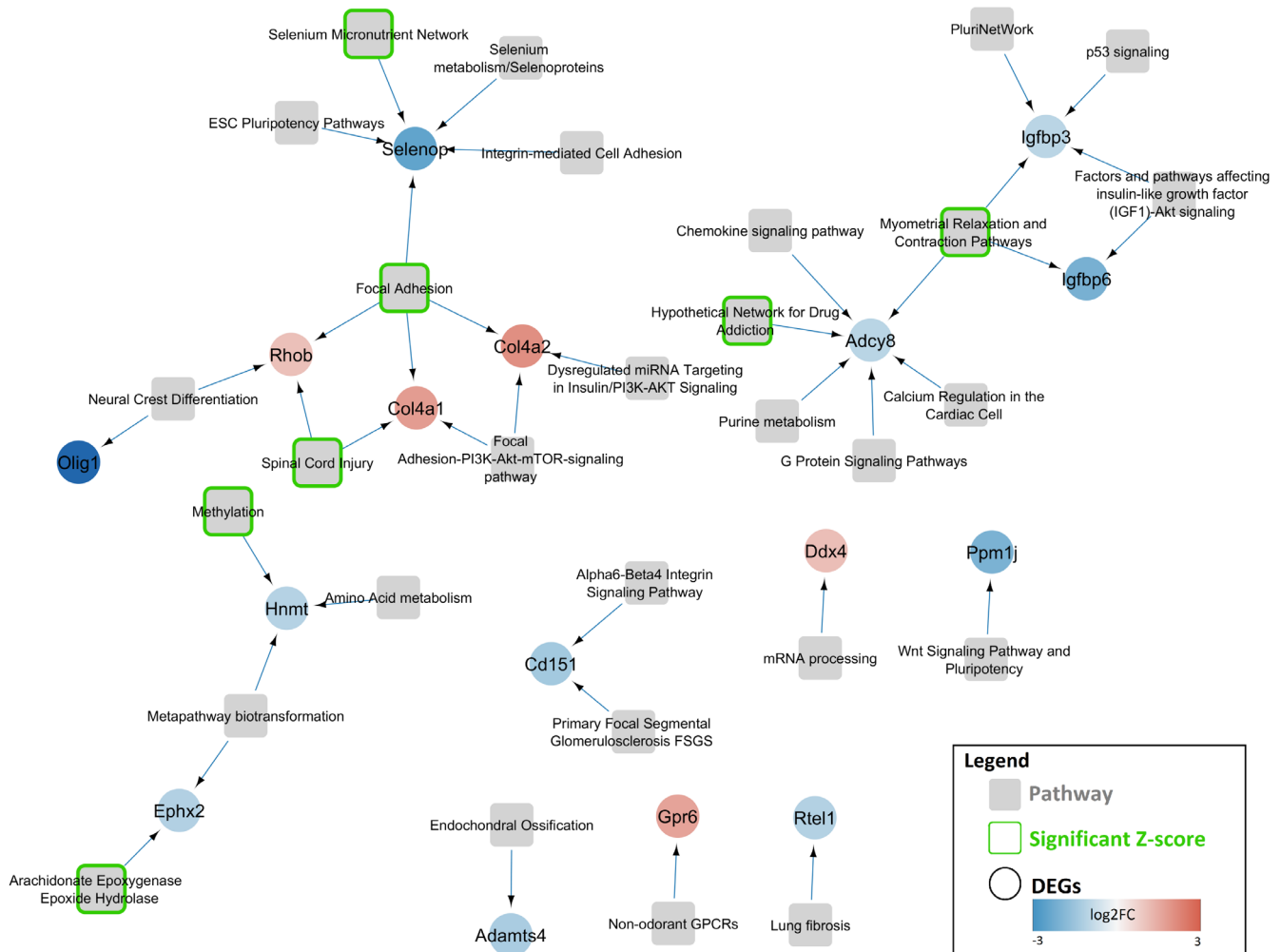
Figure 2. Gene-Pathway network for differentially expressed genes (DEGs) in Purkinje cells of Mecp2$^{-/y}$ vs. wild type. DEGs in a mouse model for Rett syndrome (Mecp2$^{-/y}$ mice) were selected and imported in Cytoscape (circular nodes). The genes are colored based on changes in gene expression (blue=down-regulated, red=up-regulated). Thereafter, gene-pathway associations from WikiPathways were added and the pathways are shown as gray rectangles. Genes without pathway annotations have been removed. The green border color indicates pathways that are significantly altered based on over-representation analysis in PathVisio (Z-Score > 1.96).

Differential gene expression is shown on the gene nodes (blue = down, red = up) and green border color of the pathway nodes indicates that the pathway has been identified as significantly affected through overrepresentation analysis in the pathway analysis tool PathVisio[15].

Use case 3: Author-publication-journal network. This example uses two custom made linksets for author-article and article-journal relationships from Wikidata[16–18]. After loading the initial five author nodes in Cytoscape, we performed a two-step extension with CyTargetLinker. We first added publications from the author-article linkset and then the journals from the article-journal linkset, see Figure 3. Author nodes are colored in gray, articles in yellow and journals in green. The network clearly shows the collaborations and diversity between the authors. Layout and visual style was slightly adapted manually in the graphical user interface to improve the readability of the network.

Discussion

One of the major challenges in science is the reproducibility of results presented in articles. Besides the challenges in reproducibility of experiments, the computational analyses are also often unclear and insufficiently described[19]. Automation of analysis workflows enables researchers to share the details of their computational analyses and enables simple reproducibility of the results.

Here, we introduce the new version of the CyTargetLinker app, which provides full programmatic execution of the functions from within Cytoscape (command line tool) and external tools (e.g. R, Jupyter, Python, etc). The network
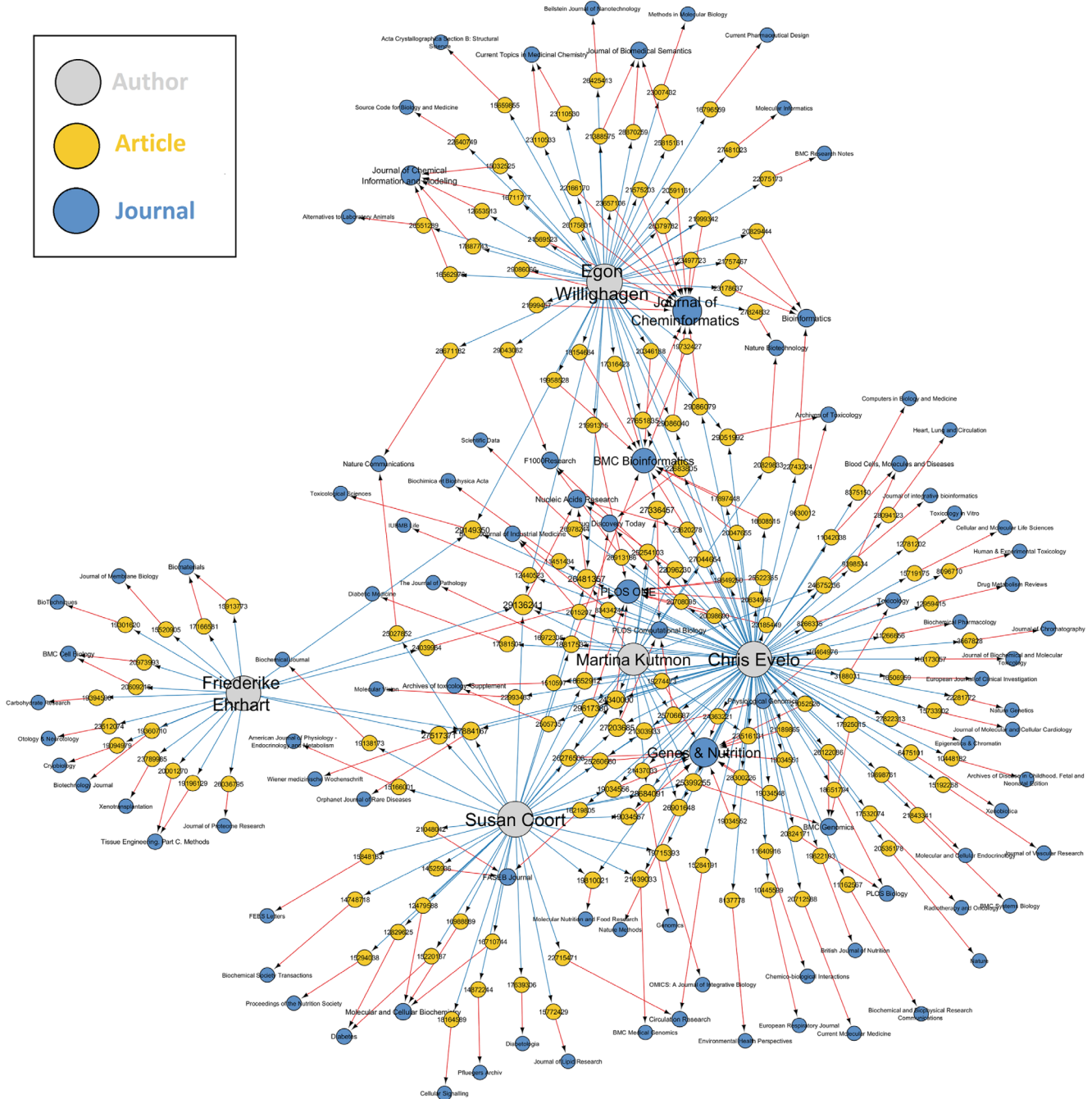
Figure 3. Author-Article-Journal network. For the five authors of this manuscript, two custom linksets were used to first add author-article relationships and in a second step add article-journal relationships.

extension is therefore reproducible and repeatable with other input data. Consequently, users can build scripts that run common analysis workflows and combine CyTargetLinker with other apps, as shown in Use case 1 (stringApp). The integration of CyTargetLinker in Cytoscape gives access to a powerful set of visualization options, as demonstrated in Use case 2.

As part of the Cytoscape tutorials collection for online presentations, we developed a CyTargetLinker tutorial presentation

using Reveal.js. This tutorial presentation can be reused and adapted for specific teaching activities. Together with our tutorials for the Cytoscape desktop application and the automation example scripts, relevant documentation for users is provided to get familiar with the functionality of CyTargetLinker.

The generic nature of CyTargetLinker has been highlighted by renaming *RegINs* to *linksets*, and we now provide a variety of different linksets on our website. The XGMML structure

of the linksets is simple, instructions how to create them from tab-delimited text files are available, and CyTargetLinker could therefore be used for non-biological networks as well (shown in use case 3).

Conclusions

In this paper, we showcase the latest update of the CyTargetLinker app for Cytoscape and its new automation feature. The ability to programmatically execute the app's functionality opens up the possibility to build complex workflows that are repeatable and reproducible. We also explored the broader applicability of the app besides the originally intended use for regulatory network extension. Due to the flexible design of the app and the linksets, we are now also showcasing other use cases, including non-biological networks.

Data availability

Linksets, tutorials and link to source code (app and linkset creator) are available from the CyTargetlinker app website: https://projects.bigcat.unimaas.nl/cytargetlinker/.

Software availability

1. **The app is available from the Cytoscape app store:** http://apps.cytoscape.org/apps/cytargetlinker.

2. **Link to source code:** https://github.com/CyTargetLinker/cytargetlinker.

3. **Archived source code at time of publication:** https://doi.org/10.5281/zenodo.1218721[20].

4. **Tutorials including R code for use cases:** https://github.com/CyTargetLinker/cytargetlinker-tutorials.

5. **Software license:** https://www.apache.org/licenses/LICENSE-2.0.

Competing interests

No competing interests were disclosed.

Grant information

Acknowledgements

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): 2498–2504.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

2. Kutmon M, Kelder T, Mandaviya P, *et al.*: **CyTargetLinker: a cytoscape app to integrate regulatory interactions in network analysis.** *PLoS One.* 2013; **8**(12): e82160.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

3. Kandhro AH, Shoombuatong W, Nantasenamat C, *et al.*: **The MicroRNA Interaction Network of Lipid Diseases.** *Front Genet.* 2017; **8**: 116.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

4. Moradifard S, Hoseinbeyki M, Ganji SM, *et al.*: **Analysis of microRNA and Gene Expression Profiles in Alzheimer's Disease: A Meta-Analysis Approach.** *Sci Rep.* 2018; **8**(1): 4767.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. De Niz C, Rahman R, Zhao X, *et al.*: **Algorithms for drug sensitivity prediction.** *Algorithms.* 2016; **9**(4): 77.
   Publisher Full Text

6. van Iersel MP, Pico AR, Kelder T, *et al.*: **The BridgeDb framework: standardized access to gene, protein and metabolite identifier mapping services.** *BMC Bioinformatics.* 2010; **11**(1): 5.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7. Leonard H, Cobb S, Downs J: **Clinical and biological progress over 50 years in Rett syndrome.** *Nat Rev Neurol.* 2017; **13**(1): 37–51.
   PubMed Abstract | **Publisher Full Text**

8. Szklarczyk D, Morris JH, Cook H, *et al.*: **The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible.** *Nucleic Acids Res.* 2017; **45**(D1): D362–D368.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

9. Bento AP, Gaulton A, Hersey A, *et al.*: **The ChEMBL bioactivity database: an update.** *Nucleic Acids Res.* 2014; **42**(Database issue): D1083–D1090.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

10. Willighagen EL, Waagmeester A, Spjuth O, *et al.*: **The ChEMBL database as linked open data.** *J Cheminform.* 2013; **5**(1): 23.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

11. Hamosh A, Scott AF, Amberger JS, *et al.*: **Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders.** *Nucleic Acids Res.* 2005; **33**(Database issue): D514–D517.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

12. Sugino K, Hempel CM, Okaty BW, *et al.*: **Cell-type-specific repression by methyl-CpG-binding protein 2 is biased toward long genes.** *J Neurosci.* 2014; **34**(38): 12877–12883.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

13. Ehrhart F, Coort SL, Cirillo E, *et al.*: **New insights in rett syndrome using pathway analysis for transcriptomics data.** *Wien Med Wochenschr.* 2016; **166**(11–12): 346–352.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

14. Slenter DN, Kutmon M, Hanspers K, *et al.*: **WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research.** *Nucleic Acids Res.* 2017; **46**(D1): D661–D667.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

15. Kutmon M, van Iersel MP, Bohler A, *et al.*: **PathVisio 3: an extendable pathway analysis toolbox.** *PLoS Comput Biol.* 2015; **11**(2): e1004085.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

16. Vrandečić D, Krötzsch M: **Wikidata: a free collaborative knowledgebase.** *Communications of the ACM.* 2014; **57**(10): 78–85.
    Publisher Full Text

17. Nielsen FÅ, Mietchen D, Willighagen E: **Scholia, scientometrics and wikidata.** In *Lecture Notes in Computer Science.* Springer International Publishing, 2017; 237–259.
    Publisher Full Text

18. Taraborelli D, Mietchen D: **Wikicite 2017 report.** 2017.
    Publisher Full Text

19. **Reality check on reproducibility.** *Nature.* 2016; **533**(7604): 437.
    PubMed Abstract | **Publisher Full Text**

20. Kutmon M: **CyTargetLinker: Release 4.0.1.** 2018.
    Data Source

Open Peer Review

Current Peer Review Status: ✔ ✘

Version 1

Reviewer Report 05 December 2018

✘ **John H. Morris** 🆔

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

The authors present an update to their successful CyTargetLinker app for Cytoscape. This update extends the basic capabilities of CyTargetLinker by generalizing some concepts and adding automation functionality. The article is clear and well-written and succinctly outlines the major point of the article and the features the article is going to focus on. They present three use cases, which outline the automation capabilities nicely.

I did have several issues that I believe should be addressed to support users trying to follow the instructions and to elucidate more clearly the functionality of the new automation features.

First, the instructions point to a GitHub repository for the three use cases, however, I could only find two use cases in the repository. This should absolutely be fixed before approval.

Second, the authors state that both the ChEMBL and OMIM linksets are available on the CyTargetLinker website. Unfortunately, only the ChEMBL linkset is provided. The OMIM linkset **is** provided in the github repository, however, once the file is unzipped. It wasn't clear to me why the authors chose to provide the data as a zip file in UseCase1, but just provide the unzipped directories in UseCase2. I would suggest just to leave the files unzipped in UseCase1 in the same way you do for UseCase2.

Third, to support general understanding of the automation, I would add the CyTargetLinker commands you actually execute to achieve the results. For UseCase1, this would be:

"cytargetlinker extend idAttribute=\"display name\" linkSetFiles=\"/tmp/cytargetlinker-tutorials/R/UseCase1/LinkSets/chembl_23_hsa_20180126.xgmml,/tmp/ network=current". Obviously, the paths could be truncated, so something like: "cytargetlinker extend idAttribute=\"display name\" linkSetFiles=\"*path* /chembl_23_hsa_20180126.xgmml,*path*/rare-disease-gene-associations-hsa-20180411.xgmml\" network=current"
This would make it much easier for app writers to utilize your command features without having to look

through the R source code.

And, as mentioned before, I couldn't find UseCase3.

Finally, as a suggestion, it would be nice to have the functionality to create a linkset from a csv file as part of the app so the user wouldn't need to download and execute it separately.

Overall, I'm enthusiastic about the app and the new capabilities, and would be happy to approve it after the above major changes are made.

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.

Reviewer Expertise: Network Biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 13 Aug 2019

Martina Summer-Kutmon, Maastricht University, Maastricht, The Netherlands

Thank you for your comments. We completely restructured and updated the website. We also made sure the third use case script is available in the automation repository.
We provided the data for use case 1 as a zip file because the ChEMBL linkset is too large to be uploaded to Github. I now only zipped that specific linkset and gave instructions in the R script to first unzip that specific xgmml file. This makes it more consistent with the other use cases.
We added the general commands used to execute the extension for each of the use cases to the manuscript.
Since we are using the linkset creator script to generate all linksets regularly, we opted for keeping it separate. Most users use the linksets provided or contact us to add a new one and only few

generate their own.
We hope that we addressed all of your requests in version 2 and the new updated website.

Competing Interests: No competing interests were disclosed.

Reviewer Report 28 June 2018

https://doi.org/10.5256/f1000research.15902.r35060

✔ **Alex Gutteridge** 🆔

GlaxoSmithKline (GSK), Brentford, UK

The CyTargetLinker Cytoscape app presented here provides a useful mechanism for extending existing biological networks. The rationale for this functionality is well explained by the authors and the general utility is demonstrated by the citation of several previous studies using the tool and the variety of use cases presented.

The API interface, which is the focus of this update, is a very useful addition which allows the scripting and hence reproducibility of this kind of analysis. The API is simple, but covers the fundamental use case of the app. I can only find partial documentation of the API method parameters however. They are named in Table 1, but I could not find more extensive documentation on the app website (https://projects.bigcat.unimaas.nl/cytargetlinker/) which mostly refers to an earlier version of the app from what I can tell. The API methods should be documented somewhere more completely. This is why I have answered 'Partly' to 'Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?'.

The interpretation of the results and conclusions about the tool are straight forward. I have not tested the performance of the app outside the use cases the authors provide, but in those contexts performance is perfectly adequate. Perhaps the authors could comment on whether the app has performance limits on the size of network that can be annotated in this way and how that compares to Cytoscape's underlying limits.

I tested the first use case script with Cytoscape 3.6.1, stringApp 1.1.1 and CyTargetLinker 4.0.1 (the latest versions at time of writing). This revealed a small API incompatibility. The script attempts to link on idAttribute 'display name', but no such attribute comes from this version of stringApp. Changing the idAttribute to 'name' allows the script to work. I can't find which version of stringApp the authors are using, but maybe it needs to be specified or the script updated. This is why I have marked 'Partly' to the question 'Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?' Otherwise the use case reproduced fine.

I have not tested the CyTargetLinker Java program, but it seems reasonably well documented and straight forward to run.

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.

Author Response 13 Aug 2019

Martina Summer-Kutmon, Maastricht University, Maastricht, The Netherlands

Thank you for your comments. We completely restructured and updated the website and all tutorials and example scripts should now contain the relevant information to reproduce the results.

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