



SOFTWARE TOOL ARTICLE

REVISED PSFC: a Pathway Signal Flow Calculator App for Cytoscape [version 2; referees: 2 approved]

Lilit Nersisyan¹, Graham Johnson^{2,3}, Megan Riel-Mehan^{2,3}, Alexander R Pico⁴, Arsen Arakelyan ^{1,5}

- ¹Institute of Molecular Biology NAS RA, Yerevan, Armenia
- ²California Institute for Quantitative Biosciences (QB3), San Francisco, CA, 94158, USA
- ³University of California, San Francisco, CA, 94158, USA
- ⁴Gladstone Institutes, San Francisco, CA, 94158, USA
- ⁵American University of Armenia, Yerevan, Armenia

v2 First published: 05 Aug 2015, 4:480 (doi: [10.12688/f1000research.6706.1](https://doi.org/10.12688/f1000research.6706.1))
 Latest published: 03 Apr 2017, 4:480 (doi: [10.12688/f1000research.6706.2](https://doi.org/10.12688/f1000research.6706.2))

Abstract

Cell signaling pathways are sequences of biochemical reactions that propagate an input signal, such as a hormone binding to a cell-surface receptor, into the cell to trigger a reactive process. Assessment of pathway activities is crucial for determining which pathways play roles in disease versus normal conditions. To date various pathway flow/perturbation assessment tools are available, however they are constrained to specific algorithms and specific data types. There are no accepted standards for evaluation of pathway activities or simulation of flow propagation events in pathways, and the results of different software are difficult to compare. Here we present Pathway Signal Flow Calculator (PSFC), a Cytoscape app for calculation of a pathway signal flow based on the pathway topology and node input data. The app provides a rich framework for customization of different signal flow algorithms to allow users to apply various approaches within a single computational framework.



This article is included in the [Cytoscape Apps gateway](#).

Open Peer Review

Referee Status:

	Invited Referees	
	1	2
REVISED		
version 2 published 03 Apr 2017		
version 1 published 05 Aug 2015	 report	 report

- 1 **David Ruckerbauer**, University of Natural Resources and Applied Life Sciences, Austria
- 2 **Maarten Altelaar**, Utrecht University, Netherlands
Linsey Raaijmakers, Utrecht University, Netherlands

Discuss this article

Comments (0)

Corresponding author: Arsen Arakelyan (arakelyan@sci.am)

Competing interests: The authors disclose no competing interests.

How to cite this article: Nersisyan L, Johnson G, Riel-Mehan M *et al.* **PSFC: a Pathway Signal Flow Calculator App for Cytoscape [version 2; referees: 2 approved]** *F1000Research* 2017, 4:480 (doi: [10.12688/f1000research.6706.2](https://doi.org/10.12688/f1000research.6706.2))

Copyright: © 2017 Nersisyan L *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 work was funded by Google Summer of Code 2014 (Student: LN, Mentor: GJ), and the Armenian National Science and Educational Foundation research grant molbio-3818 (PI: LN).

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

First published: 05 Aug 2015, 4:480 (doi: [10.12688/f1000research.6706.1](https://doi.org/10.12688/f1000research.6706.1))

REVISED Amendments from Version 1

The new version of this manuscript contains small amendments to the text addressing the respective changes in Figures and the Supplementary material.

The [Figure 2](#) and [Figure 5](#) have been updated. [Figure 2](#) has been corrected for mistakes in the figure and in the legend ('source-target' rule replaced by 'target-source', and 'update' replaced by 'addition'). The [Figure 5](#) has been updated to reflect the two-color scheme of PSFC 1.0.2 and 1.1.2.

We have added the [Supplementary material](#) 'figure2_inputs_and_instructions.rar', which contains an xgml formatted example network for [Figure 2](#) and instructions for replicating each of the subfigures. An additional [Supplementary material](#), 'loop_example.xml' is a toy network to replicate the loop handling options, as presented in [Figure 3](#). The PSFC User manual has been updated to the version 1.1.2 of the app.

See referee reports

Introduction

Cell signaling pathways are sets of directed interactions between biological molecules, that are initiated by a particular signal (e.g. a ligand binding to a receptor) and result in realization of certain target processes (e.g. transcription of genes). Pathways can be represented as graphs, with nodes as biological entities (proteins, other biomolecules, chemical compounds, other pathways), and edges as physical or regulatory interactions between them. In contrast to protein-protein interaction networks, biomolecular pathways have directionality, input nodes, intermediate nodes and branches, and output or sink nodes.

Pathway Signal Flow (PSF), or perturbation, is the flux generated by propagation of the signal starting from input nodes, flowing through intermediate nodes in branches and accumulating at sink nodes. Thus, PSF can be an indicator of pathway activity state. Assessment of changes in pathway activity is of major interest for identification of processes involved in the formation of certain phenotypes (healthy and diseased states), and assessment of cell response to drugs and other stimuli. First attempts to globally evaluate the pathway activity changes were performed in parallel with the appearance of high-throughput gene expression measurement experiments. Pathway involvement is typically analyzed by over-representation analysis (ORA)¹ or gene set enrichment analysis (GSEA)². The major drawback of these widely used approaches is that they operate on gene sets involved in the pathway, but do not account for the pathway topology and ignore the interactions between the nodes.

A number of techniques and tools have recently emerged, aimed at determining pathway activities based on topological information of pathways and gene expression/protein activity levels. One of the pioneering papers in this direction was the Pathway Impact Analysis algorithm, which combines GSEA with gene position in the network³. Other approaches apply specific rules to model flow or signal propagation through the pathway and evaluate the amount of the signal reaching the sink nodes⁴⁻⁶.

The above mentioned algorithms and tools are implemented using various programming and scripting languages, making their use and

result comparison difficult in the common context. Moreover, they often work with programming environment specific objects, and are not flexible for using biological pathways that appear in various formats. Cytoscape, on the other hand, is a powerful and flexible platform that, together with its diverse collection of available apps, provides a rich environment for parsing, visualization and analysis of networks⁷.

Herein we present Pathway Signal Flow Calculator (PSFC), a Cytoscape app for computation of pathway signal flow based on input data and pathway topology. PSFC provides a variety of options for signal propagation, both used in already published signal flow algorithms³⁻⁶, as well as in new ones. Thus, it allows experimenting with the results obtained by various (existing and customizable) approaches within a single framework, and evaluating their ability to simulate real life situations.

Methods

Implementation

PSFC packages and data structures. PSFC is implemented in Java and is available as an app for Cytoscape 3. The main module consists of two main packages, *logic* and *gui*. The package *logic* is designed to handle PSFC-inherent structures and algorithms, while the *gui* package is responsible for user communication via the PSFC tab in the Cytoscape GUI west panel, and for mapping Cytoscape inherent data structures to PSFC data structures (*Graph*, *Node* and *Edge*) contained in the *logic* package ([Figure 1](#)).

PSFC algorithms

Graph sorting. Graph sorting is the first step before proceeding to signal flow calculation. The aim of sorting is to assign levels to the nodes, to propagate the signal from lower to higher level nodes.

We have modified the topological sort algorithm implemented in Java JGraphT library [<http://jgraph.org/>], to handle multiple input node containing graphs.

Recall that biological networks often contain feedback loops, which create cycles in graphs. PSFC firstly performs depth first search traversal and removes backward edges from the graph, and performs the topological sorting on the resulting acyclic graph, after which the backward edges are restored. Finally, node levels of the sorted graph are mapped to the Cytoscape node attributes table.

Pathway signal flow calculation. In biological signaling networks, the signal is propagated via interactions between source-target node pairs. The outcome of signal propagation events is the signal (PSF value) accumulated at each network node. [Figure 2](#) provides an example of how the signal propagates through a sample network, with various signal propagation options applied.

Rules for simple source-target interactions. Functional interaction types can be broadly defined as activation or inhibition, while the range of physical and regulatory interactions is much wider (phosphorylation, binding, dimerization, ubiquitination, etc.). An edge in a graph carries a signal transfer function, which depends on the interaction type. PSFC allows the user to define the interaction type of each edge in the network, as well as define the edge-type specific mathematical functions of wide complexity.

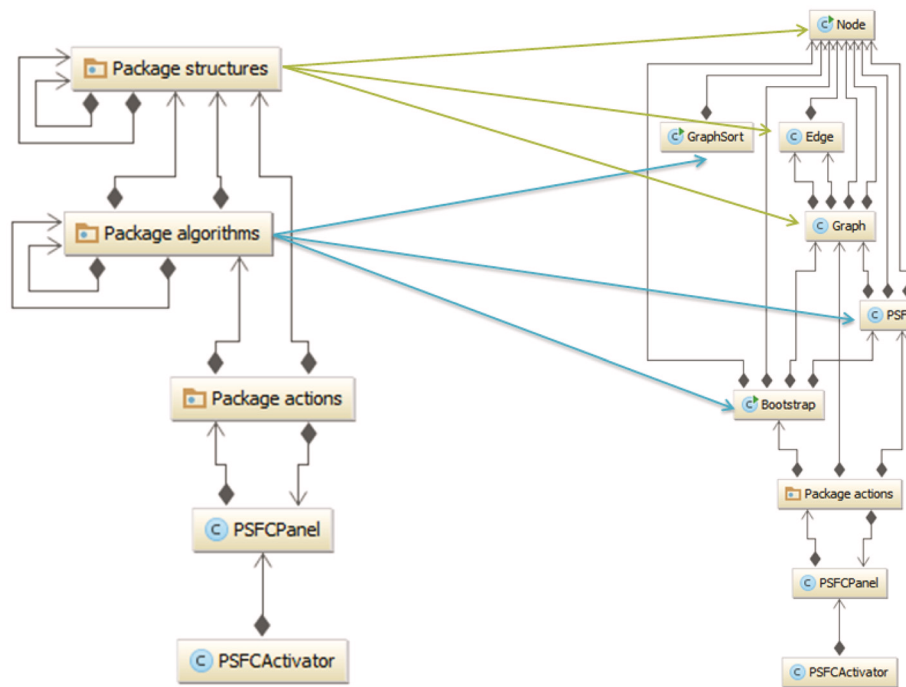


Figure 1. The PSFC class diagram. Only the main packages are displayed. In the right-hand diagram, the packages *logic.algorithms* and *logic.structures* are extended to respective classes. The communication from and to the user is performed via the *PSFCPanel* class and the *gui.actions* package. PSFC inherent structures and algorithms are implemented in the packages *logic.algorithms* and *logic.structures*. The classes extended from the packages *logic.structures* and *logic.algorithms* are shown with green and blue arrows, respectively. Class dependencies are indicated with grey arrows.

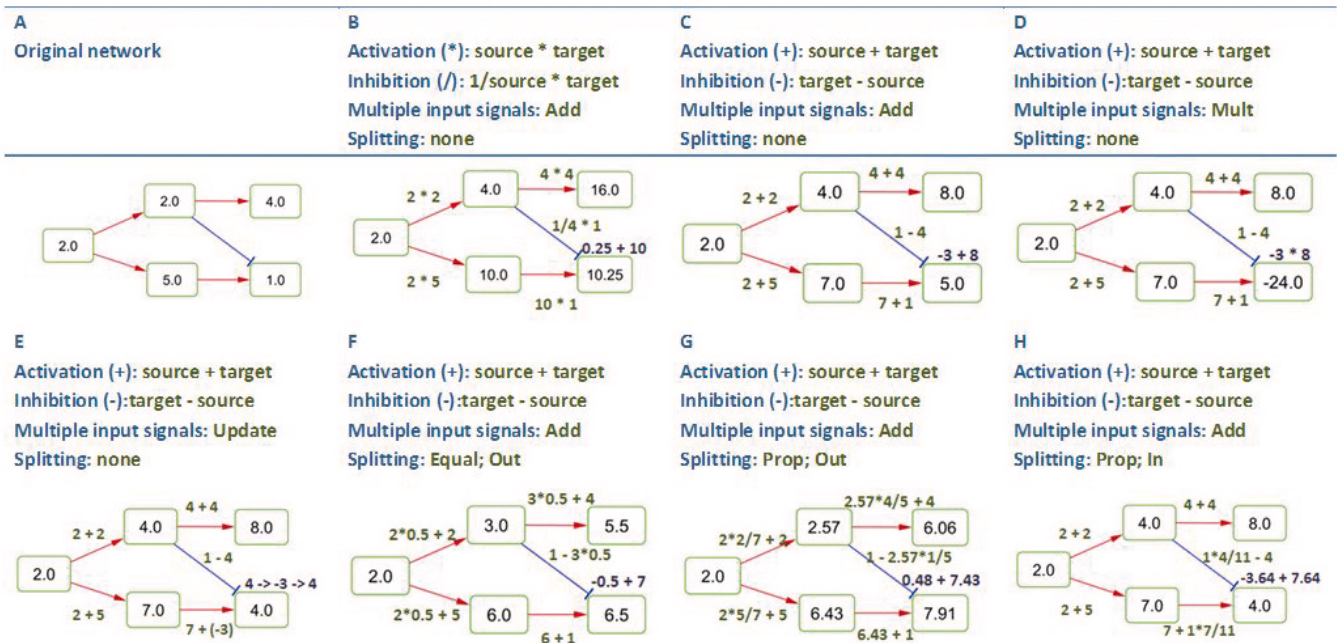


Figure 2. Demonstration of PSF computation on a sample network with different flow propagation rules. Red and blue edges are of types “activation” and “inhibition”, respectively. Multiple signals at a target node are computed by addition (Add), multiplication (Mult), or by updating target node signals (Update). Signal splitting is either set to “none” or “equal” (Equal) or “proportional” (Prop) rules, and is either performed on multiple outgoing edges (Out) or multiple incoming edges (In).

These functions should have the form $f(\text{source}, \text{target})$, where the *source* and *target* variables stand for the source node signal and the target node value. The functions are parsed with Exp4j Java library for symbolic operations [<http://www.objecthunter.net/exp4j/>]. Function assignment for different edge types is shown in Figure 2.

Rules for multiple incoming and outgoing signals. Generally, the intensity of interactions between molecules largely depends on their concentration and activation state. However, if a node has several interacting partners, those may compete with each other, and the interaction capacity of the node may be “split” between those partners. Thus, there is the option to proportionally split the signal among multiple edges starting from a single source or ending on a single target node. The signals on multiple edges ending on a single target node may be processed in one of the following three ways: the signals may be computed separately at each edge and added (i) or multiplied (ii) to each other, or they may be processed in order (iii), by updating the signal at a target node each time a single edge is processed. The order, in which the edges are processed in the last case, may be adjusted by user defined edge ranks (Figure 2). The example network presented in Figure 2, and instructions for replicating the rules and results are provided in the [Supplementary material](#), figure2_inputs_and_instructions.rar.

Handling of feedback loops. The presence of negative and positive feedback loops in biomolecular pathways is of paramount importance for pathway functionality and regulation. However, currently it is a major obstacle for developing optimal algorithms for pathway activity assessment. To our knowledge, there is no single solution for treatment of loops in signal propagation algorithms, thus PSFC provides several options for loop handling:

Ignore feedback loops: In this case cycle-forming backward edges are ignored during PSF calculations (Figure 3B).

Precompute signals at loops: In this mode, the algorithm firstly finds cycle-forming backward edges, computes their signals, and updates their target node values. Afterwards, the algorithm runs on the whole graph in the “ignore feedback loops” mode (Figure 3C).

Iterate until convergence: The algorithm runs for several rounds, until convergence of signal flow values is reached (Figure 3D–F). Convergence is reached if the percentage of signal changes between two iterations is less than the specified convergence threshold at all the nodes. If convergence is not achieved, the algorithm stops after running for a defined number of iterations. The user may check the convergence status of the calculations in the PSFC log file and in the command prompt.

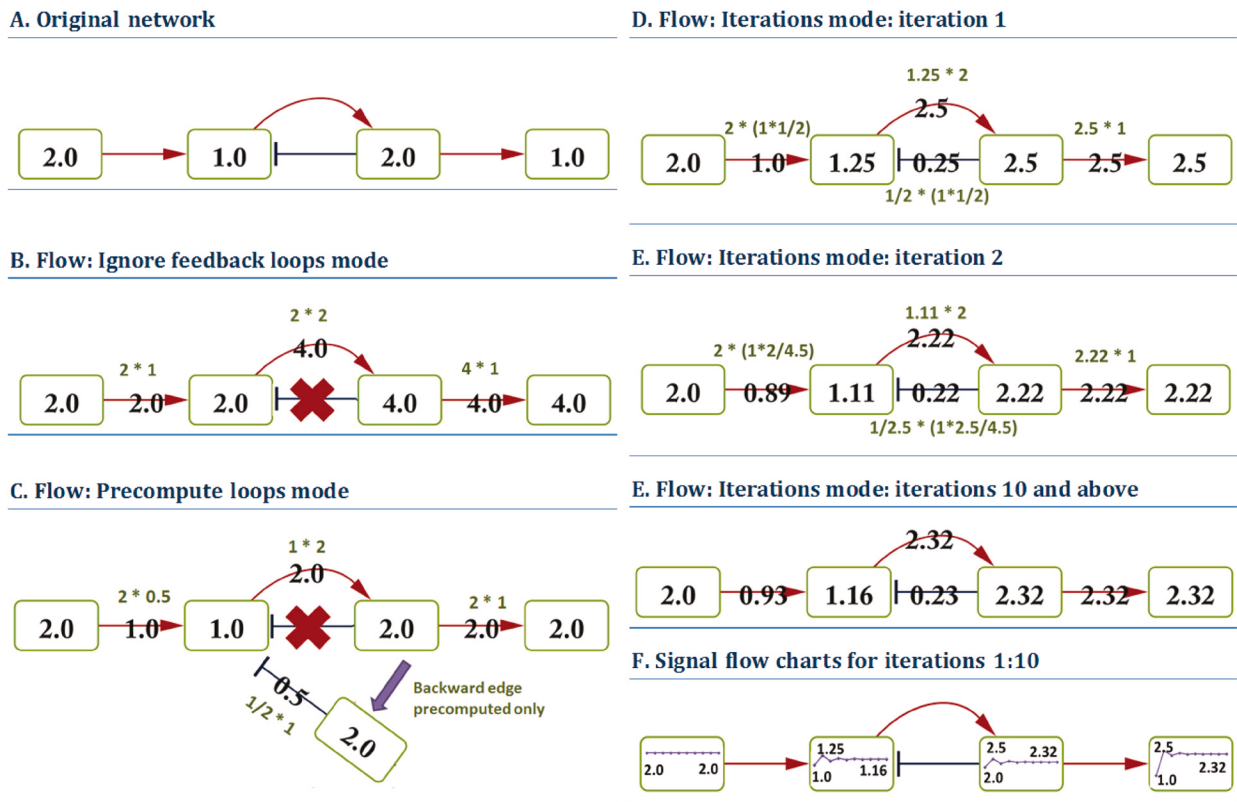


Figure 3. Signal flow calculation on a sample network with different options for loop handling. The PSF flow rules are: for red edges of type activation (*): $\text{source} * \text{target}$; for blue edges of type inhibition (/): $1/\text{source} * \text{target}$; Multiple input signals: Addition; Splitting: Proportional; Split on: Incoming edges. The “Ignore feedback loops” mode (B) does not account for backward edges. In the “Precompute loops” mode (C), the backward edge signal first updates the target node value, and is ignored in the following single iteration. (D–E): signal flow at different iterations with “Iterate until convergence” mode. The network converges at iteration 10. The dynamics of flow changes at each node during 10 iterations are shown as line charts in (F). Edges ignored during the computation are indicated by red X symbols.

The example network for replicating the loop handling options presented in Figure 3 is available in the [Supplementary material](#), `loop_example.xml`.

Significance calculation. The significance value of signal flows at each node is computed using bootstrapping. The user may choose between sample-centric or gene-centric bootstrapping modes. In the sample-centric mode, the values of the nodes in the network will be reshuffled among each other during resampling. In the gene-centric mode, the value of each node is randomly chosen from a supplied distribution of node values, e.g., from measurements of a given gene's expression across multiple samples.

Operation

PSFC is implemented for Cytoscape version 3.2 and higher, with Java 1.7 or higher. PSFC may be installed with either the Cytoscape App Manager or by direct download of the jar file from <http://apps.cytoscape.org/apps/psfc>. The whole functionality of PSFC is accessible to users via a single tab in the Cytoscape GUI west panel.

Use cases

General use case of PSFC. The main use case of the app is presented in Figure 4. PSFC operates on any network loaded into the Cytoscape environment. Node data and edge types should be loaded into Cytoscape attribute tables, while signal propagation

rules should be set in respective PSFC GUI tabs (Figure 4). PSF computation is performed with the “Compute flow” button. The resulting PSF values are stored both in Cytoscape attribute tables, and in PSFC output files (the score backup file and `psfc.log` file in text formats). The signal propagation may be visualized via node color and edge width mapping, where continuous values are mapped to color gradients and width ranges at a chosen level or across all levels in a sequence.

PSF calculation on MAPK signaling pathway: a use case. We evaluated signal flow changes in the MAPK signaling pathway network taken from previously published papers^{8,9}. In their paper, Nelander *et al.* have performed a series of experiments, where they have downregulated one or many of the MAPK pathway proteins, and measured the changes of protein phosphorylation levels, and the states of G1-arrest and apoptosis⁸. Feiglin *et al.*⁹ have compared the results of the experimental data with their predictions, based on a wiring algorithm described in their paper⁹. We have repeated the same experimental simulations, to compare the performance of PSFC with the wiring algorithm and with the experimental data. The node values were presented as gene expression fold change (FC) values that show the relative increase or decrease of gene expression compared to the reference state. In the reference state, the amount of PSF should be 1, corresponding to the normal level of pathway activity necessary to realize the target biological process. Departure

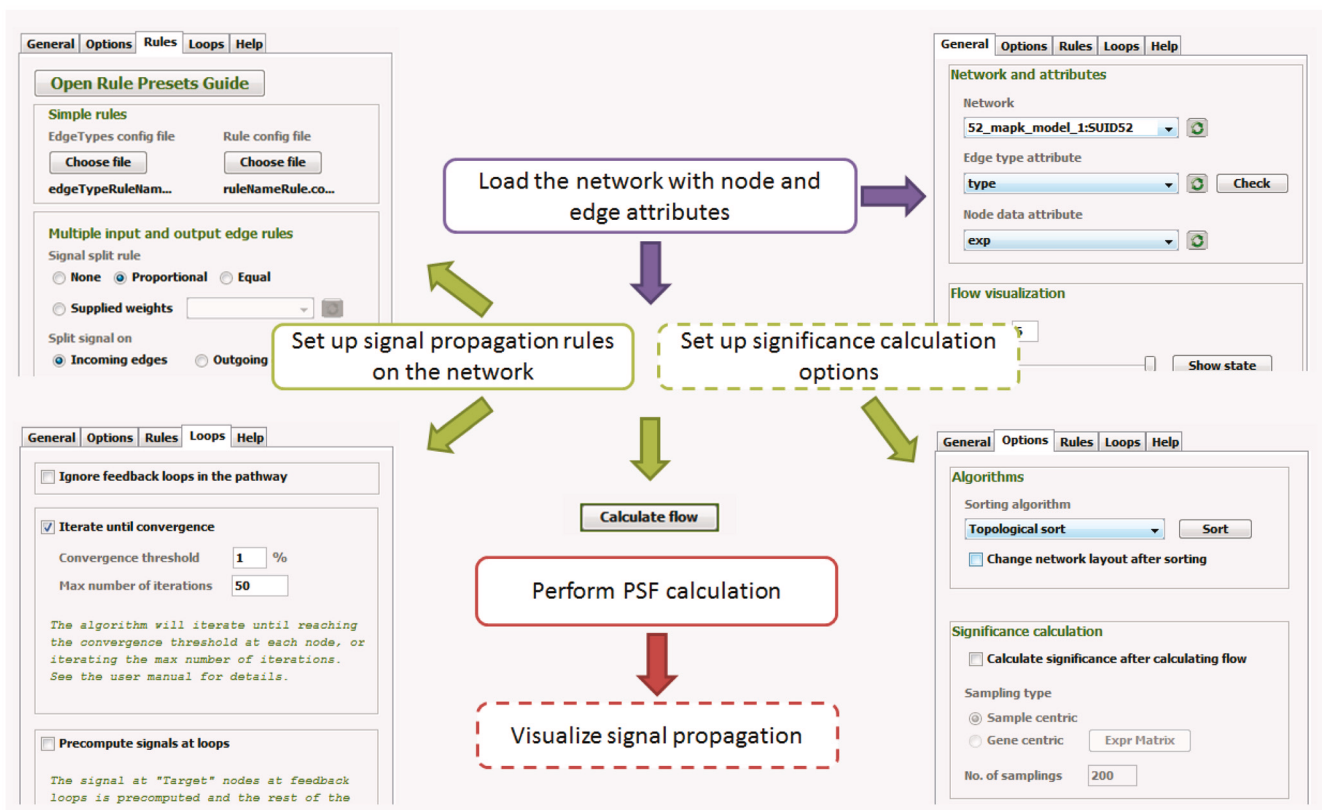


Figure 4. The general use case of PSFC. The user should load the network into Cytoscape, and import node and edge attributes using the Cytoscape environment. Further, the user sets the rules for signal propagation, loop handling and significance calculation. After PSF calculation, the signal propagation may be visualized in Cytoscape. The dashed rectangles are optional.

of PSF values from 1 is indicating an up- or down-regulation of the pathway. To simulate this situation, we have applied the following rules for signal propagation. The single edges were treated with $(source*target)$ and $(1/source*target)$ functions for edges of types activation and inhibition, respectively, ensuring that an FC change on a node propagates proportionally via signal perturbations to downstream nodes. Furthermore, we have applied splitting on incoming edges and addition of multiple incoming signals on a single target node. This is based on the speculation that, in signaling pathways, the capacity of a protein to interact with upstream agents depends on the relative frequency of co-occurrence and the interaction strength with those agents, which is this case, is represented as the PSF signals of the source nodes. Finally, loop handling was in “iterate until convergence” mode, since the absence of positive feedback loops in the MAPK pathway and FC representation of the node values ensure that the algorithm will converge.

We have performed PSF calculations in 6 different experiments. In each of these experiments one of the IGF1R, PI3K, mTOR,

PKCdelta, p-MEK, or EGFR nodes was assigned a value of 0.1 (down-regulated), while the rest of the nodes had “fc” values of 1, which corresponds to the unchanged state compared to the control (Figure 5A).

Under all the experimental settings we have predicted up-regulation of the “G1-arrest” and “apoptosis” nodes, which is in full accordance with the predictions of Feiglin *et al.*⁹. These predictions deviated from the experimental outcomes⁸ in only one case (Figure 5B). The network xml file and all configuration files are available in [Supplementary material](#), MAPK_psf_configurations.rar.

Summary

We have developed PSFC, a Cytoscape app for PSF calculation. The main purpose of the app is to evaluate the signal flow propagation in pathways and assess activity states of pathway components, based on input data and the topology. PSFC may be used for the purpose of assessment of pathway activity deregulations in different conditions, for simulation studies on network dynamics, etc.

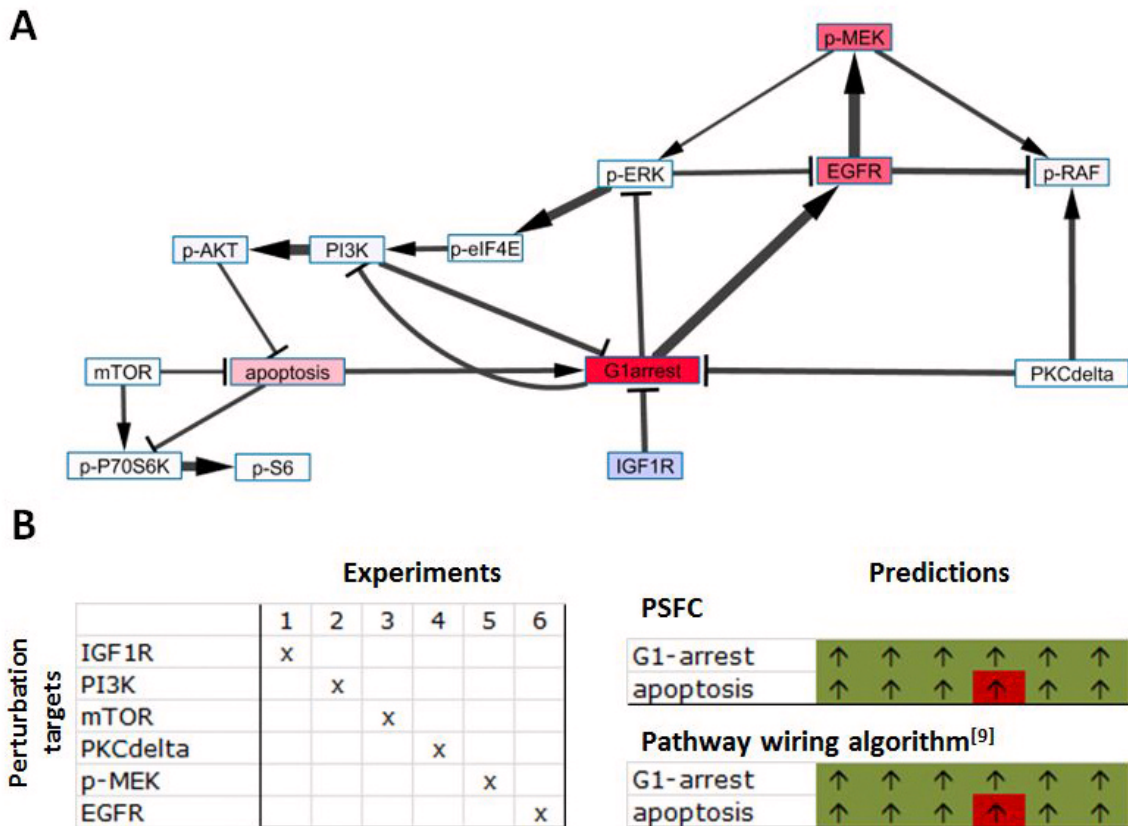


Figure 5. PSF calculation results on MAPK signaling pathway. MAPK signaling pathway is generated by the model used in 8. **Part A** shows the pathway, as visualized in Cytoscape after PSF calculations. Edges of activation and inhibition types have Delta and T shaped target arrows. The darker colors of nodes indicate greater PSF values, and the line width corresponds to the signal intensities at the edges. The prediction results of six perturbation experiments by PSFC and the pathway wiring algorithm⁹ are presented in **part B**, similar to the representation in 9. In each of the six experiments, one node was down-regulated, as indicated by the “x” sign. The predicted perturbations at the “G1-arrest” and “apoptosis” nodes are shown with up and down arrows. Green and red colors indicate consistency and non-consistency, respectively, of the predictions with the experimental results presented in 8.

Compared to other similar software, PSFC stands out with a wide set of rules and options for signal propagation, which makes it possible to use the app for the majority algorithms that could possibly be applied for pathway flow calculations in different biological contexts. It is, thus, not constrained with preset algorithm design, but allows users to apply their own algorithms. Thus, PSFC can be used in routine data analysis by bench biologists using available presets, but also can become a powerful tool for sophisticated pathway analyses in the hands of a bioinformatics skilled person.

Software availability

Software available from: <http://apps.cytoscape.org/apps/psfc>

Latest source code

<https://github.com/lilit-nersisyan/psfc/>

Source code as at the time of publication

<https://github.com/F1000Research/PSFC>

Archived source code as at the time of publication [version 1 of manuscript]

<http://dx.doi.org/10.5281/zenodo.19465>¹⁰

Archived source code as at the time of publication [version 2 of manuscript]

<https://doi.org/10.5281/zenodo.439419>¹¹

Supplementary material

Archive for 'PSF calculation on MAPK signaling pathway: a use case'

The archive contains source and configuration files used to generate the data described in the "PSF calculation on MAPK signaling pathway: a use case" subsection of the manuscript. The MAPK_network.xml file contains the network model, which should be imported into Cytoscape. The EdgeTypeRuleName.config and RuleNameRule.config files are configuration files defining simple rules for flow propagation. The psfc.props contains the properties used by PSFC during calculations. The Flow_propagation_rules.pdf describes the applied rules and options.

[Click here to access the data.](#)

PSFC user manual

The user manual covers installation notes, a generic use case and a specific example, in addition to an overview of the graphical user interface and details about network sorting, flow rule configurations, pathway flow calculation and flow visualization. This supplemental file is a static version of the manual. For the latest online version, click on the Tutorial link on the PSFC app page: <http://apps.cytoscape.org/apps/psfc>.

[Click here to access the data.](#)

Figure2_inputs_and_instructions.rar

Network file and instructions to replicate the examples of Figure 2.

The archive contains the following files:

InstructionsToFigure2.pdf: contains instructions to load the files in the archive and replicate the examples in Figure 2.

figure2.xgmml: the network file that can be imported into Cytoscape.

figure2.rules a property file containing the proper settings to replicate the examples of Figure 2

edgeTypeRuleName1.config, edgeTypeRuleName2.config, ruleNameRule – standard configuration files for PSFC

[Click here to access the data.](#)

License

PSFC is free software; it can be distributed and/or modified under the terms of the GNU General Public License version 3. The license can be found at <http://www.gnu.org/licenses/gpl.html>. The exp4j library is distributed under Apache License 2.0, while the JGraphT library is dual-licensed GNU Lesser General Public License and Eclipse Public License.

Author contributions

LN and AA performed software design and development, testing and analyses, and manuscript preparation. GJ, MR-M, and AP contributed to software design and manuscript preparation. All the authors have read and approved the final manuscript.

Competing interests

The authors disclose no competing interests.

Grant information

This work was funded by Google Summer of Code 2014 (Student: LN, Mentor: GJ), and the Armenian National Science and Educational Foundation research grant *molbio-3818* (PI: LN).

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

Loop example xml

A toy network to replicate loop handling options presented in Figure 3.

The xml file can be imported into Cytoscape. The loop handling options can then be applied as described in the manuscript and the legends of Figure 3.

[Click here to access the data.](#)

References

1. Khatri P, Draghici S: **Ontological analysis of gene expression data: current tools, limitations, and open problems.** *Bioinformatics.* 2005; **21**(18): 3587–3595.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
2. Hung JH: **Gene Set/Pathway enrichment analysis.** *Methods Mol Biol.* 2013; **939**: 201–13.
[PubMed Abstract](#) | [Publisher Full Text](#)
3. Draghici S, Khatri P, Tarca AL, *et al.*: **A systems biology approach for pathway level analysis.** *Genome Res.* 2007; **17**(10): 1537–45.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
4. Isik Z, Ersahin T, Atalay V, *et al.*: **A signal transduction score flow algorithm for cyclic cellular pathway analysis, which combines transcriptome and ChIP-seq data.** *Mol Biosyst.* 2012; **8**(12): 3224–31.
[PubMed Abstract](#) | [Publisher Full Text](#)
5. Arakelyan A, Aslanyan L, Boyajyan A: **High-throughput Gene Expression Analysis Concepts and Applications.** *Sequence and Genome Analysis II – Bacteria, Viruses and Metabolic Pathways.* iConcept Press. 2013; ISBN: 978-1-480254-14-5.
[Reference Source](#)
6. Haynes WA, Higdon R, Stanberry L, *et al.*: **Differential expression analysis for pathways.** *PLoS Comput Biol.* 2013; **9**(3): e1002967.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. 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](#)
8. Nelander S, Wang W, Nilsson B, *et al.*: **Models from experiments: combinatorial drug perturbations of cancer cells.** *Mol Syst Biol.* 2008; **4**: 216.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Feiglin A, Hacohen A, Sarusi A, *et al.*: **Static network structure can be used to model the phenotypic effects of perturbations in regulatory networks.** *Bioinformatics.* 2012; **28**(21): 2811–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
10. Nersisyan L: **PSFC: Pathway Signal Flow Calculator (PSFC), a Cytoscape app for calculation of a pathway signal flow based on the pathway topology and node input data [Data set].** *Zenodo.* 2015.
[Data Source](#)
11. Nersisyan L, Arsen Arakelyan: **PSFC 1.1.2 source code [Data set].** *Zenodo.* 2017.
[Data Source](#)

Open Peer Review

Current Referee Status:  

Version 1

Referee Report 27 November 2015

doi:10.5256/f1000research.7203.r10923

 **Maarten Altelaar**¹, **Linsey Raaijmakers**²

¹ Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, Netherlands

² Biomolecular Mass Spectrometry and Proteomics, Utrecht University, Utrecht, Netherlands

The authors present PSFC, a cytoscape app for the analysis of pathway signal flow to indicate pathway activity states. A user manual is presented with some example datasets.

In the manual the various steps are very well described.

My assessment for this paper is 'Approved', the app helps in visualization pathway flow and generates an easy interpretable network.

Just some minor comments and recommendations:

- The flow visualization part and the Rules tab in the control panel is not completely visible. The bottom part falls outside the panel. And also the 'Min', 'Mid', 'Max' text boxes are very small, making it difficult to modify the numbers inside these boxes.
- When pathways are loaded from ReactomeFIPlugin, the edge annotation column contains often more than 1 entry. This causes problems in the flow calculation steps. An idea is to split this column into the separate entries and check which entry is described in the edgeType config file.
- When the pathway is loaded from a plugin inside Cytoscape, like Reactome, the network file is not saved yet. This caused problems with writing of the results output file. It would be better if the network file could be automatically saved at this step.
- Some interactions could represent a complex, instead of an 'inhibition' or 'activation' signal, in this case there is not really a rule that would describe this situation, besides maybe source == target, but unfortunately I get an error when I use the '=' sign to describe an edge rule.

Competing Interests: No competing interests were disclosed.

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.

Author Response 29 Mar 2017

Lilit Nersisyan, Institute of Molecular Biology NAS RA, Armenia

Response to the reviewers

We would like to thank the reviewer for their comments and recommendations. We have addressed those in the version 2 of the manuscript and in version 1.1.2 of PSFC.

Here, the paragraphs containing responses to the reviewer's comments are indicated with "Response:" in the beginning.

- The flow visualization part and the Rules tab in the control panel is not completely visible. The bottom part falls outside the panel. And also the 'Min', 'Mid', 'Max' text boxes are very small, making it difficult to modify the numbers inside these boxes.

Response: We have changed the layout of PSFC 1.1.2 accordingly. It is still possible that the user experiences distortions depending on the screen resolution. We would like to receive comments from users in such cases to make appropriate changes in further versions of the app.

- When pathways are loaded from ReactomeFIPlugin, the edge annotation column contains often more than 1 entry. This causes problems in the flow calculation steps. An idea is to split this column into the separate entries and check which entry is described in the edgeType config file.

Response: We understand the concern brought by the reviewers. However, while this feature would be useful in the particular case when pathways are loaded from the ReactomeFIPlugin, in other cases this would cause unnecessary complications. Therefore, we have left this split operation to the user: they may create an additional column containing the proper edge annotation entries.

- When the pathway is loaded from a plugin inside Cytoscape, like Reactome, the network file is not saved yet. This caused problems with writing of the results output file. It would be better if the network file could be automatically saved at this step.

Response: In PSFC the user should specify which network to perform calculations on. There may be many networks loaded into the Cytoscape session, while the user would desire to perform PSF on only one network. Therefore, we find that automatically saving the networks should not be a feature of PSFC. In case of issues, we advise the user to send us the details of the errors to perform appropriate fixes.

- Some interactions could represent a complex, instead of an 'inhibition' or 'activation' signal, in this case there is not really a rule that would describe this situation, besides maybe source == target, but unfortunately I get an error when I use the '=' sign to describe an edge rule.

Response: We thank the reviewers for this comment. We understand that the case of "source == target" can be handled by simply putting "source" in the edge rule. This will lead to source signals to be transferred to the targets as is. The cases of complex nodes are handled with the new feature of PSFC 1.1.2 – operator nodes. In addition to the rules assigned to edges, each specific node can also be assigned a function. E.g. if this node is a complex node it can take on the 'min' function, which means that the minimum signal of all the incoming nodes will be assigned to the target. Functions 'min', 'max', 'mean' and 'prod' are supported. The details are available in the user manual.

Competing Interests: No competing interests were disclosed.

Referee Report 20 August 2015

doi:10.5256/f1000research.7203.r9853



David Ruckerbauer

Department of Biotechnology, University of Natural Resources and Applied Life Sciences, Vienna, Austria

The authors introduce an app for Cytoscape called 'Pathway Signal Flow Calculator' (PSFC). PSFC allows to simulate and visualize the propagation of signals through a pathway, considering both the pathway topology and input data for 'node activity', such as protein phosphorylation. The aim is to provide a flexible environment that supports the integration of various custom algorithms and data sets.

The paper describes several rules and algorithms for signal flow and demonstrates their effect on two small toy models as well as a bigger use case (MAPK signaling pathway), all in context of using PSFC.

Furthermore, the authors provide the use case with the corresponding input files as supplementary material, a very detailed step-by-step manual for the different options of PSFC and the source code of the app.

I have tested the app under Linux Mint 17.2 (64-bit), Mac OS X 10.7.5 (64-bit) and Windows Vista (32-bit) (see minor issues below).

While my assessment of this paper is 'Approved', I have come across some issues that are detailed below. Additionally to major and minor issues I have added a section with suggestions that I personally feel might improve PSFC. I would like to stress, however, that PSFC has full usability without considering any of the points in the latter section.

Major Issues

Figure 2: The definition of inhibition as 'source – target' seems counterintuitive, as a strong signal from the source should decrease the result in the target node. Should it be 'target – source'? This definition is also used by Isik *et al.* (Reference 4 in the paper) as well as in the document 'rules_presets.pdf' provided by the authors. Please be aware that both the text and the figures would need to be updated.

Figure 2 (F,G,H): 'Multiple input fields' is set to 'Update', the figures, however, correspond to 'Add'. If it should indeed be 'Update' the values for the bottom right box are wrong in all three figures.

While all necessary input files for the use case are provided in the supplementary material the output is missing (apart from the screenshot of a log-file in the manual, which is not easy to read). This information should definitely be provided for all perturbations in order to be able to compare numerical values; a qualitative comparison is not enough, especially since all results are the same (up-regulation).

Ideally the toy model from Figure 2 would also be provided as an .xgmml-file, as this would enable

readers to test and verify different settings.

Minor Issues

Clicking the link for reference 5 gives an error (page not found).

It is possible to install PSFC in Cytoscape running on Windows Vista 32-bit (with current versions of Cytoscape and Java), but running the calculation gives an error. I am aware that Cytoscape generally recommends using a 64-bit system, but if PSFC is for 64-bit systems only then this information could be added to the manual.

Redoing a calculation shuffles the sorting of the columns in the node / edge attribute table (in terms of the order of `psf_l0`, `psf_l1`, etc.), in Linux Mint.

Setting mTOR to an `fc` of 2 in the MAPK-example will lead to a final state where the box 'apoptosis' is completely white, even though the value is 0.75, and the box for 'p-RAF' is completely black (with a value of 4.59). This may or may not be an issue with the Cytoscape installation under Linux Mint.

I don't understand the purpose of the 'level textbox' in the flow visualization window. Does it only show the current maximum level?

Page numbers should be added to the manual.

Suggestions

Using a two color scheme (or maybe an easily customizable one) would improve the visualization. Right now even negative values are shown in the same color as positive ones and they are indistinguishable from low positive numbers.

An option to visualize the values for nodes and edges might be helpful.

Clicking the 'Play flow' button in the flow visualization window causes a pop-up which covers part of the network visualization. Maybe this pop-up could be forced to a bottom corner of Cytoscape.

The 'play flow' feature is on a very high speed if there is only a small number of levels. Ideally either the speed should be easily adjustable (and an endless mode available) or maybe the 'Show state' button could be removed and the visualization updated directly via the slider. This would allow to control the speed and direction of the visualization very easily. It would also solve the above mentioned issue of the pop-up.

Is the use of two config files (Edge Types config file and Rule config file) a requirement of Cytoscape or the parser? If not it would seem simpler to merge them and remove the redundant 'Function names'.

The file 'rules_presets.pdf' could be added to the manual instead of being a separate file, especially since there already is an extra pdf for the use case.

Additional toy models with the relevant input/output in order to test and verify 'significance calculation', handling feedback loops and edge weights might be helpful.

The paper explicitly mentions 'bench biologists' as potential users of PSFC. I feel that the app is described sufficiently in order to allow a user unfamiliar with modeling and/or Cytoscape to use it successfully. The bottleneck would probably be generating the necessary input model. Maybe a short paragraph in the manual pointing to the relevant sections in the Cytoscape manual as well as describing the creation and changing of node and edge attribute tables would remove this obstacle.

Competing Interests: No competing interests were disclosed.

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.

Author Response 29 Mar 2017

Lilit Nersisyan, Institute of Molecular Biology NAS RA, Armenia

Response to the reviewer

We would like to thank the reviewer for their thorough revision of the manuscript, and for the useful suggestions. We have addressed the issues and suggestions and have incorporated them in the version 2 of the manuscript and in version 1.1.2 of PSFC.

Here, the paragraphs containing responses to the reviewer's comments are indicated with "Response:" in the beginning.

Major Issues

- Figure 2: The definition of inhibition as 'source – target' seems counterintuitive, as a strong signal from the source should decrease the result in the target node. Should it be 'target – source'? This definition is also used by Isik *et al.* (Reference 4 in the paper) as well as in the document 'rules_presets.pdf' provided by the authors. Please be aware that both the text and the figures would need to be updated.

Response: We agree with the reviewer. Although the user may set any rule they desire, "target – source" would be more intuitive for inhibition type of edges. We have updated the example and the figures respectively.

- Figure 2 (F,G,H): 'Multiple input fields' is set to 'Update', the figures, however, correspond to 'Add'.
If it should indeed be 'Update' the values for the bottom right box are wrong in all three figures.

Response: We agree with the reviewer. The rule for multiple inputs is addition, while its mentioned as "update" in the figures 2 F,G,H. We have made respective changes.

- While all necessary input files for the use case are provided in the supplementary material the output is missing (apart from the screenshot of a log-file in the manual, which is not easy to read). This information should definitely be provided for all perturbations in order to be able to compare numerical values; a qualitative comparison is not enough, especially since all results are the same (up-regulation).

Response: We have presented the PSF scores of all the experiments in a separate document in the supplementary material.

- Ideally the toy model from Figure 2 would also be provided as an .xgmml-file, as this would enable readers to test and verify different settings.

Response: The example network in Figure2 is provided as an xgmml file. In addition we have provided instructions to replicate the settings applied in different subfigures.

Minor Issues

Clicking the link for reference 5 gives an error (page not found).

Update the link with

<https://www.icconceptpress.com/books/genomics-ii--bacteria-viruses-and-metabolic-pathways/1100006>

It is possible to install PSFC in Cytoscape running on Windows Vista 32-bit (with current versions of Cytoscape and Java), but running the calculation gives an error. I am aware that Cytoscape generally recommends using a 64-bit system, but if PSFC is for 64-bit systems only then this information could be added to the manual.

Response: This was explicitly mentioned in the Manual.

Redoing a calculation shuffles the sorting of the columns in the node / edge attribute table (in terms of the order of psf_l0, psf_l1, etc.), in Linux Mint.

Response: Unfortunately, currently it is impossible to order the columns in Cytoscape. Currently, this as an ongoing task and will soon be addressed by the Cytoscape development team.

Setting mTOR to an fc of 2 in the MAPK-example will lead to a final state where the box 'apoptosis' is completely white, even though the value is 0.75, and the box for 'p-RAF' is completely black (with a value of 4.59). This may or may not be an issue with the Cytoscape installation under Linux Mint.

Response: The versions 1.0.2 and 1.1.2 of PSFC allows for setting the color scheme based on custom preferences. So the user may get rid of complete white and black nodes.

I don't understand the purpose of the 'level textbox' in the flow visualization window. Does it only show the current maximum level?

Response: The flow visualization is based on sequential mapping of the node and edge signals, from the input nodes (level 0), to the sink nodes (highest level). The level textbox may thus be considered as an indication of time-step. In PSFC 1.0.2 and 1.1.2 we have made it non-editable and named it "time-step", so that is more understandable to the user.

Page numbers should be added to the manual.

Response: According to the reviewer's comment, we have added page numbers to the manual.

Suggestions

Using a two color scheme (or maybe an easily customizable one) would improve the visualization. Right now even negative values are shown in the same color as positive ones and they are indistinguishable from low positive numbers.

Response: Thank you for the suggestion. We have added a two-color scheme in PSFC 1.0.2, with option to specify also the middle color and signal value. Additionally, the user may also choose the range of edge widths. In PSFC 1.1.2 the minimum and maximum values may also be modified by the user.

An option to visualize the values for nodes and edges might be helpful.

Response: The initial values of nodes and edges can be visualized by flow visualization at level of time-step 0. We have additionally described this in the manual.

Clicking the 'Play flow' button in the flow visualization window causes a pop-up which covers part of the network visualization. Maybe this pop-up could be forced to a bottom corner of Cytoscape. The 'play flow' feature is on a very high speed if there is only a small number of levels. Ideally either the speed should be easily adjustable (and an endless mode available) or maybe the 'Show state' button could be removed and the visualization updated directly via the slider. This would allow to control the speed and direction of the visualization very easily. It would also solve the above mentioned issue of the pop-up.

Response: According to the reviewers suggestion, we have removed the "Play flow button" in PSFC 1.0.2 and added control buttons, so that the user is able to visualize the pathway states manually. This also removed the problem with the pop-up.

Is the use of two config files (Edge Types config file and Rule config file) a requirement of Cytoscape or the parser? If not it would seem simpler to merge them and remove the redundant 'Function names'.

Response: The edge types config file describes the mapping between edge types and function names, while the rule config files assigns mathematical functions to each function name. The reason these two files are separate, is that multiple edge types can be treated with the same rule, and writing the same function for that rule multiple types in a file may produce errors. Thus, we think that keeping the mappings separate is more optimal. The instructions for using Figure 2 toy example of the manuscript, also describe the advantage of having two separate config files.

The file 'rules_presets.pdf' could be added to the manual instead of being a separate file, especially since there already is an extra pdf for the use case.

Response: The "rules_preset.pdf" is to assist the user specifically in setting up the options for flow propagation rules. While the user may be familiar with the rest of the options in PSFC, this is intended to be a quick look-up guide for rules. Thus, we made it as a separate file, to free the user from scrolling through the manual every time they have a doubt which rule to use.

Additional toy models with the relevant input/output in order to test and verify 'significance calculation', handling feedback loops and edge weights might be helpful.

Response: The toy network model for loop-testing has also been added in the supplementary material of the second version of the manuscript.

The paper explicitly mentions 'bench biologists' as potential users of PSFC. I feel that the app is described sufficiently in order to allow a user unfamiliar with modeling and/or Cytoscape to use it successfully. The bottleneck would probably be generating the necessary input model. Maybe a short paragraph in the manual pointing to the relevant sections in the Cytoscape manual as well as describing the creation and changing of node and edge attribute tables would remove this obstacle.

Response: We agree with the reviewer. We have referred the user to respective pages of the Cytoscape manual, in the PSFC User Manual, where appropriate.

Competing Interests: No competing interests were disclosed.
