Bioinformatics, 35(21), 2019, 4504–4506
doi: 10.1093/bioinformatics/btz419
Advance Access Publication Date: 22 May 2019
Applications Note

OXFORD

Systems biology

PathFXweb: a web application for identifying drug safety and efficacy phenotypes

Jennifer L. Wilson [1,*], Mike Wong[2], Ajinkya Chalke[3], Nicholas Stepanov[3], Dragutin Petkovic[2,3] and Russ B. Altman[1,4]

[1]Department of Bioengineering, Stanford University, Stanford, CA 94035, USA, [2]CoSE Computing for Life Sciences, [3]Department of Computer Science, San Francisco State University, San Francisco, CA 94132, USA and [4]Department of Genetics, Stanford University, Stanford, CA 94035, USA

*To whom correspondence should be addressed.

Associate Editor: Lenore Cowen

Received on February 1, 2019; revised on April 19, 2019; editorial decision on May 10, 2019; accepted on May 15, 2019

Abstract

Summary: : Limited efficacy and intolerable safety limit therapeutic development and identification of potential liabilities earlier in development could significantly improve this process. Computational approaches which aggregate data from multiple sources and consider the drug's pathways effects could add to identification of these liabilities earlier. Such computational methods must be accessible to a variety of users beyond computational scientists, especially regulators and industry scientists, in order to impact the therapeutic development process. We have previously developed and published PathFX, an algorithm for identifying drug networks and phenotypes for understanding drug associations to safety and efficacy. Here we present a streamlined and easy-to-use PathFX web application that allows users to search for drug networks and associated phenotypes. We have also added visualization, and phenotype clustering to improve functionality and interpretability of PathFXweb.

Availability and implementation: https://www.pathfxweb.net/.

Contact: pathfx_support@googlegroups.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

The drug discovery process is long, difficult and expensive, and lack of drug efficacy or intolerable side effects prevent drugs from reaching the market. Multiple studies suggest that drugs are more successful when genetic evidence supports the selection of a drug target (Nelson *et al.*, 2015, 2016). Further, investigations of drug adverse events have demonstrated that the targets themselves and pathways surrounding the targets may be sources of side effects, including more serious, adverse events (Burkhart *et al.*, 2015; Sakellaropoulos *et al.*, 2016). These studies suggest that identifying the disease and safety phenotypes associated with drug targets and target pathways are essential components of understanding drug efficacy and potential safety risks (Campillos, 2016; Guney *et al.*, 2016; Lorberbaum *et al.*, 2014). Computational tools can improve understanding of drug mechanisms by integrating multiple data sources and leveraging networks of information around drug targets. To this effect, we developed PathFX and demonstrated its ability to identify drug-relevant phenotypes (Wilson *et al.*, 2018). PathFX is designed to look at the pathways-level effects (pathway FX) of a drug intervention. Under the hood, PathFX is an interaction-network tool that searches for the most relevant protein–protein interactions around a drug's target(s), and then analyzes for which phenotypes the network is enriched relative to the entire interaction network (summarized in Supplementary Fig. S1). The method is intended to help scientists understand how a drug's target(s) are associated to safety and efficacy phenotypes.

However, downloading source code of the application is not ideal for all stakeholders and computational tools should be accessible to non-computational scientists to achieve greatest impact. We were specifically motivated to inform the regulatory review process and collaborated closely with regulators at the US Food and Drug

Administration to first design PathFX (Wilson *et al.*, 2018) and second, to design an easy-to-use PathFX web application which allows users to easily search drug networks and associated phenotypes.

Here, we present the PathFXweb application and highlight its features. The application has been developed using modern software engineering methods including engaging relevant stakeholders and employing User Centered Design to ensure ease of use. PathFXweb uses Open Source components and is available on Amazon Web Services (https://www.pathfxweb.net/).

PathFXweb complements other network-based online tools including GUILDify (Guney *et al.*, 2014) and NetworkPrioritizer (Kacprowski *et al.*, 2013). These tools prioritize genes associated to a query phenotype using protein–protein interactions between the gene products and each tool uses distinct methods to prioritize these gene sets. In comparison, PathFX and PathFXweb first construct networks associated with drug targets and then search for phenotype associations to these networks. For these reasons, we believe that PathFXweb will be a valuable asset to the research and regulatory communities. There is more information about PathFX and relevant data sources included in the Supplementary Material.

2 Accessing and running PathFXweb

Login (pathfxweb.net)

To create a private record of individual queries, users are prompted to login when first submitting an analysis. The login enables users to revisit past queries and results, as well as check the status of current jobs. Logging in also allows the server to associate an email with the analysis and contact the user when analysis is complete (Supplementary Fig. S2A). Google Chrome is the preferred browser for accessing PathFXweb, though we have also tested on Safari and Firefox.

Job submission

The user navigates to the form submission page and is prompted for three required parameters—an analysis name, analysis description and a drug name—and two optional parameters—a list of drug targets, and a toggle for implementing phenotype clustering (Supplementary Fig. S2B). The description is not used for analysis but is recommended for keeping track of the motivation for a particular query and for data provenance. The description is stored with the analysis and recorded in the README file provided with all results (Supplementary Table S1).

The user specifies a drug name as free text. Because we constructed PathFX using drug–protein binding data from DrugBank (Wishart *et al.*, 2006), the drug name may auto-populate if a name matches a compound name in DrugBank or a DrugBank identifier, and if that drug binds at least one protein in our interaction network. However, a DrugBank name or identifier is not required, and the user can input the name of a novel compound. If a drug name is recognized, PathFX will initialize the drug target list with all targets listed in DrugBank. If the drug name is not recognized, PathFXweb will initialize an empty target list. Until the user specifies a target list, PathFXweb will return an error message prompting the user to add a target list. In the case where targets are misspelled or do not exist in the interactome network, the 'pathfx log' file (found on the 'PathFX Jobs' page) will report targets that are or are not included in the network analysis.

The optional target list is a space for users to specify the gene names of targets of interest. We see this option used in a couple scenarios: (i) the user is querying a novel compound not in DrugBank (indeed, it may not exist at all) and wants to explore a unique set of targets, (ii) the user is querying a combination of products in which one or neither compound may exist in DrugBank, and the user wants to query the union of both drug products. In the case that additional targets are provided in the target list input, PathFXweb queries the union of targets from DrugBank and user inputs; redundant targets are not subject to additional analysis. We have provided example queries on the 'Run PathFX' page (Supplementary Fig. S2B) as radio buttons and a pre-computed example available on the homepage (no login required). Clicking the radio buttons will auto-populate analysis parameters and initiate analysis in real-time. Results can be found on the user's 'PathFX Jobs' page (Supplementary Fig. S2C). We additionally highlight these example queries and parameters in the Supplementary Figures and Tables.

Turning the phenotype clustering feature to 'on' initiates the phenotype clustering feature. This feature takes the top 50 phenotypes from the association table result file (explained in Supplementary Table S1) and clusters the phenotype associations based on semantic similarity using the umls-interface.pl and umls-similiarity.pl (McInnes *et al.*, 2009) tools for the UMLS Metathesaurus. We constrained this analysis to 50 phenotypes because computation time with more phenotypes reached several days with our current computational server configuration. Users interested in changing this parameter can download and install the command-line application (https://github.com/jenwilson521/PathFX).

3 Results files and descriptions

PathFXweb emails users once analysis is completed. The email contains a zipped file of results. The user accesses the zipped file and a network image from their 'PathFX Jobs' page (Supplementary Fig. S2C). For a full description of included results files, we refer the users to Supplementary Table S1 and https://www.pathfxweb.net/about.php. Additionally, the user visualizes the PathFX network through links on their 'PathFX Jobs' page (example for metformin and metformin with atorvastatin drug combo shown in Supplementary Figs S3 and S4, respectively). Users can toggle members of the graph, rearrange network layout and export network images for further analysis.

4 Conclusions

Here we presented PathFXweb, a streamlined interface for PathFX. This web-server version of PathFX enables users to easily query drug targets for associated phenotypes, visualize network models and summarize phenotype groups.

Acknowledgements

The authors would like to thank our FDA collaborators for helpful inputs about the design and implementation of PathFXweb, namely, Michael Pacanowski, Oluseyi Adeniyi, Jielin Sun and Rebecca Racz. This article contents are solely the responsibility of the authors and do not necessarily represent the official views of the HHS or FDA.

Funding

R.B.A. was supported by National Institutes of Health LM005652, GM102365, and GM061374. This work was also made possible by Food and Drug Administration U01FD004979, which supports the UCSF-Stanford Center for Excellence in Regulatory Sciences and Innovation.

Conflict of Interest: none declared.

References

Burkhart,K.K. *et al*. (2015) Data mining FAERS to analyze molecular targets of drugs highly associated with Stevens-Johnson syndrome. *J. Med. Toxicol*., **11**, 265–273.

Campillos,M. (2016) Computational approaches for the study of the role of small molecules in diseases. *Perspect. Sci*., **9**, 49–52.

Guney,E. *et al*. (2014) GUILDify: a web server for phenotypic characterization of genes through biological data integration and network-based prioritization algorithms. *Bioinformatics*, **30**, 1789–1790.

Guney,E. *et al*. (2016) Network-based in silico drug efficacy screening. *Nat. Commun*., **7**, 10331.

Kacprowski,T. *et al*. (2013) NetworkPrioritizer: a versatile tool for network-based prioritization of candidate disease genes or other molecules. *Bioinformatics*, **29**, 1471–1473.

Lorberbaum,T. *et al*. (2014) Systems pharmacology augments drug safety surveillance. *Clin. Pharmacol. Ther*., **97**, 151–158.

McInnes,B.T. *et al*. (2009) UMLS-Interface and UMLS-Similarity: open source software for measuring paths and semantic similarity. *AMIA Annu. Symp. Proc*., **2009**, 431–435.

Nelson,M.R. *et al*. (2016) The genetics of drug efficacy: opportunities and challenges. *Nat. Rev. Genet*., **17**, 197–206.

Nelson,M.R. *et al*. (2015) The support of human genetic evidence for approved drug indications. *Nat. Genet*., **47**, 1–7.

Sakellaropoulos,T. *et al*. (2016) Computed biological relations among five select treatment-related organ/tissue toxicities. *Chem. Res. Toxicol*., **29**, 914–923.

Wilson,J.L. *et al*. (2018) PathFX provides mechanistic insights into drug efficacy and safety for regulatory review and therapeutic development. *PLoS Comput. Biol*., **14**, e1006614.

Wishart,D.S. *et al*. (2006) DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res*., **34**, D668–72.