SOFTWARE TOOL ARTICLE

REVISED GASOLINE: a Cytoscape app for multiple local alignment of PPI networks [version 2; referees: 2 approved, 1 approved with reservations]

Giovanni Micale[1], Andrea Continella[2], Alfredo Ferro[3], Rosalba Giugno[3], Alfredo Pulvirenti[3]

[1]Department of Computer Science, University of Pisa, Pisa, 56127, Italy
[2]Department of Electronics, Information and Bioengineering, Polytechnic University of Milan, Milan, 20133, Italy
[3]Department of Clinical and Molecular Biomedicine, University of Catania, Catania, 95125, Italy

v2 **First published:** 01 Jul 2014, **3**:140 (doi: 10.12688/f1000research.4537.1)
Latest published: 23 Sep 2014, **3**:140 (doi: 10.12688/f1000research.4537.2)

Abstract

Comparing protein interaction networks can reveal interesting patterns of interactions for a specific function or process in distantly related species. In this paper we present GASOLINE, a Cytoscape app for multiple local alignments of PPI (protein-protein interaction) networks. The app is based on the homonymous greedy and stochastic algorithm. GASOLINE starts with the identification of sets of similar nodes, called seeds of the alignment. Alignments are then extended in a greedy manner and finally refined. Both the identification of seeds and the extension of alignments are performed through an iterative Gibbs sampling strategy. GASOLINE is a Cytoscape app for computing and visualizing local alignments, without requiring any post-processing operations. GO terms can be easily attached to the aligned proteins for further functional analysis of alignments. GASOLINE can perform the alignment task in few minutes, even for a large number of input networks.

This article is included in the Cytoscape apps channel.

Open Peer Review

Referee Status: ✔ ? ✔

	Invited Referees 1	2	3
REVISED version 2 published 23 Sep 2014	✔ report		✔ report
version 1 published 01 Jul 2014	? report	? report	? report

1 **Antonio J. Perez Pulido**, Centro Andaluz de Biología del Desarrollo (CABD) Spain, **Pablo Mier**, Universidad Pablo de Olavide Spain
Pablo Mier, Universidad Pablo de Olavide Spain

2 **Rintaro Saito**, University of California San Diego USA

3 **Byung-Jun Yoon**, Texas A&M University USA

Discuss this article

Comments (0)

Corresponding authors: Giovanni Micale (gmgmicky@gmail.com), Alfredo Pulvirenti (apulvirenti@dmi.unict.it)

How to cite this article: Micale G, Continella A, Ferro A *et al.* **GASOLINE: a Cytoscape app for multiple local alignment of PPI networks [version 2; referees: 2 approved, 1 approved with reservations]** *F1000Research* 2014, **3**:140 (doi: 10.12688/f1000research.4537.2)

Grant information: The author(s) declared that no grants were involved in supporting this work.

Competing interests: No competing interests were disclosed.

First published: 01 Jul 2014, **3**:140 (doi: 10.12688/f1000research.4537.1)

REVISED **Amendments from Version 1**

ABSTRACT

- The section was extended with a slightly more detailed description of GASOLINE.

INTRODUCTION

- Acronym for GASOLINE was defined.
- Missing references concerning the available PPI datasets, the most important alignment methods and the importance of using network alignment were added.
- A new paragraph to describe what a network alignment algorithm is and which is the main goal was included.
- A new paragraph describing the differences between local and global alignment methods was added.
- A small paragraph with a general introduction to Cytoscape and its advantages was included.
- More details about GASOLINE and its input parameters were included.
- A small paragraph describing the main strength and limitations of GASOLINE was added.

IMPLEMENTATION

- Acronyms for COG, NOG and KOG were defined.
- The sentence "GASOLINE is very fast; it computes the results on 25 networks in a few minutes" was removed.
- A more detailed description of ISC score was provided.

RESULTS

- The technical characteristics of the machine used for the test were reported.
- Figure 2 was revised, highlighting and describing each panel within GASOLINE app.

CONCLUSIONS

- A new paragraph describing some interesting applications of GASOLINE, concerning annotation transfer and protein function prediction, was added.

See referee reports

Introduction

In the last few years there has been a rapid growth of biological network data, including protein-protein interaction (PPI) networks, metabolic networks and regulatory networks[1–6]. Among these, PPI networks are important in several biological phenomena such as signaling, transcriptional regulation and formation of multi-enzyme complexes[7].

Comparing PPI networks of evolutionary distant species can help to understand the regulation of a specific function or process, which the sequence comparison alone cannot explain[8]. Network alignment detects functionally orthologous proteins across networks of different species. Alignment methods can be local or global. In local alignment, networks are compared in order to find conserved protein complexes or pathways[8–10]. Global alignment methods[10–13] try to find a complete one-to-one mapping between functional orthologous to infer evolutionary relationships among different species.

Most network alignment algorithms have been provided with their implementations, however none of them is fully integrated within Cytoscape. Cytoscape[14] is an open source platform for visualizing and analyzing molecular interaction networks and integrating data with gene expression profiles and other kind of knowledge. Cytoscape functionality can be easily extended through Java apps.

Here, we describe a Cytoscape app based on GASOLINE (Greedy And Stochastic algorithm for Optimal Local alignment of Interaction NEtworks), a greedy and stochastic algorithm for multiple local alignment of PPI networks, introduced in[15]. We chose to build a Cytoscape app for GASOLINE because Cytoscape is a very popular tool in bioinformatics and is written in Java, therefore multi-platform.

The algorithm starts by identifying a set of similar nodes, one from each network, by using a Gibbs Sampling strategy (*bootstrap phase*). Then, a similar technique is applied in the *iterative phase*, where the alignment is extended until the local density of the aligned subgraphs increases. A single extension step consists of adding one node from each network to the current subgraph alignment. Finally, in the *refinement phase*, the alignment is refined by iteratively removing aligned nodes and adding new mapped ones.

GASOLINE is very accurate and scalable with respect to the number of aligning networks. However, it produces a one-to-one mapping between aligned nodes and its running time is moderately affected by the size and the average density of input networks.

The app enables the 2D visualization of local alignments and aligned proteins can be associated with GO annotations for further functional analysis. To the authors knowledge, it is the first Cytoscape tool for the computation and visualization of local alignments in a user-friendly way, without requiring any post-processing operations.

Implementation

GASOLINE app encloses the original Java implementation of the namesake algorithm[15], with the addition of a visualization module, which exploits the Cytoscape interface.

The app has been written in Java version 7 and designed following a classic Model-View-Controller (MVC) model. The Model part is represented by the classes implementing the algorithm and the auxiliary data structures. The View part is composed by two Java Panels; one for setting all the input and output parameters, and one for listing local alignments and handling their visualization.

The Controller part ensures the communication between the Model and the View and is implemented by different Cytoscape Task classes, one for each process performed by GASOLINE (i.e. checking file format, computing alignments, importing networks, protein description and GO annotations, building the alignment graphs). Each Task class properly notifies the corresponding view class when a task has been completed.

Input networks are imported as text files and then internally represented in two different ways to optimize the performance of our algorithm. We used CyNetwork and CyNetworkView objects for network alignment visualization and custom classes for computing alignments. For all the imported networks, the corresponding Cytoscape view is initially disabled to reduce the memory consumption.

The main component of GASOLINE is represented by a tabbed panel named "GASOLINE" located in the Control Panel of Cytoscape (see Figure 1, panel A). Through the interface the users can provide the following information:

- "Similarity information", to upload orthology similarity scores between proteins of different species;
- "Networks", for selecting two or more networks to align;
- "Parameters setting", to modify the default GASOLINE parameters;
- "Optional parameters setting", for setting other advanced input parameters;
- "Ontologies", to upload GO terms linked to the proteins of the aligned networks;
- "Output", to specify the folder where the final alignments will be saved.

The button labeled "?", when present, explains the meaning of a specific function or parameter of GASOLINE, whenever the mouse arrow hovers.

In the following subsections, we will describe all the required steps to run GASOLINE on a set of PPI networks.

Loading input data

Before running GASOLINE, the user needs to upload input data, including:

a) Two or more networks to be aligned;

b) A file of orthology BLAST bit scores between proteins of different species;

c) A set of GO terms linked to the proteins of each network.

The GO terms file is not mandatory and can be omitted.

Networks are given as a list of weighted edges and can be uploaded from the "Networks" panel. There is no theoretical limitation on the size and the number of uploadable networks.

Orthology data can be uploaded through the "Similarity information" panel. They can be supplied in two different formats: "BLAST Bit scores" or "COG groups".

The "BLAST Bit scores" format is a text file where each row has a couple of proteins of different species followed by their corresponding BLAST bit score.

Files in the "COG groups" associate a list of orthology groups to the proteins of aligning networks. Groups include KOG (euKaryotic

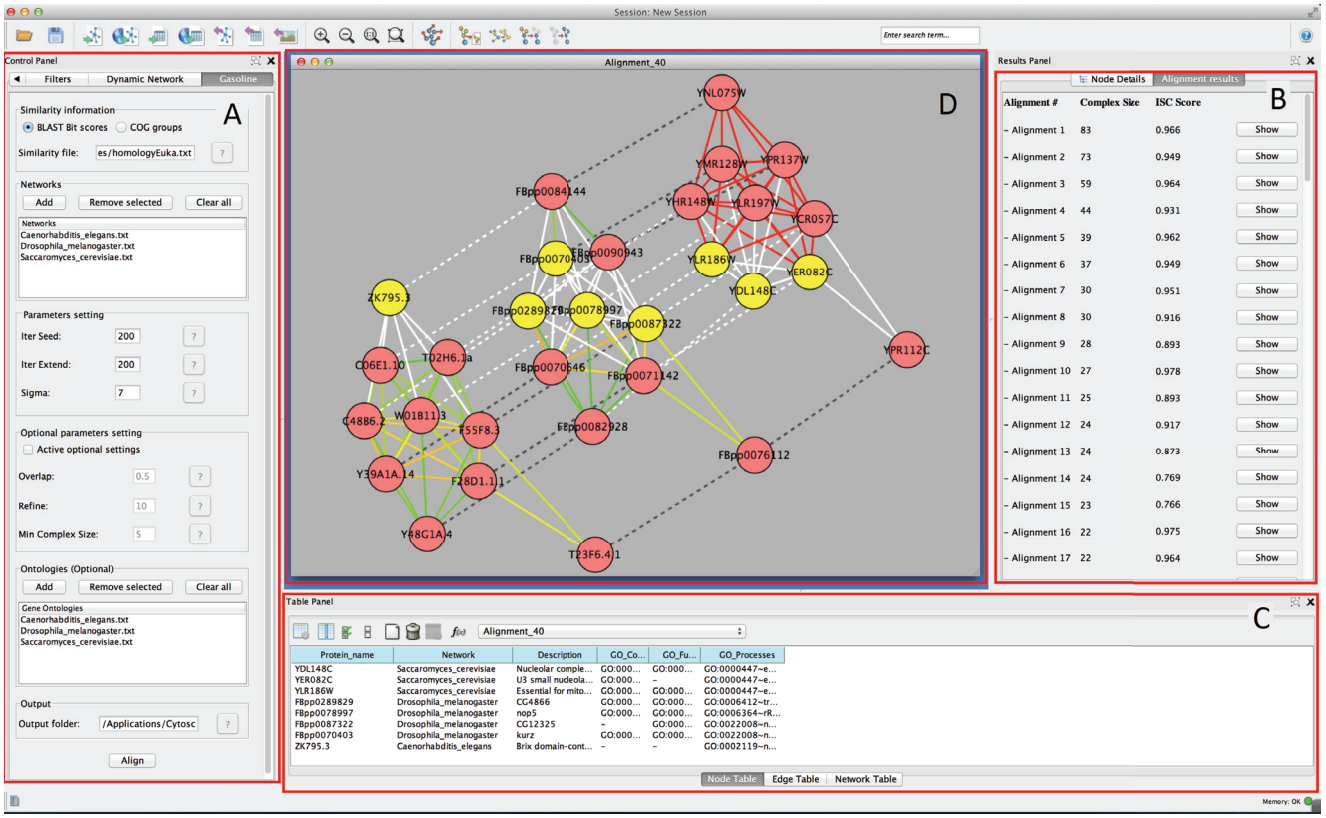

Figure 1. GASOLINE Cytoscape panels: A) GASOLINE parameters; **B**) Alignment results; **C**) Description of selected proteins and associated GO terms; **D**) Alignment visualization: intra-edges are represented with solid lines and coloured according to their weight (green for low values, yellow for medium values and red for high values); inter-edges are drawn with dashed lines.

Orthologous Groups), NOG (Non-supervised Orthologous Group) and COG (Cluster of Orthologous Groups) terms. When aligning many networks, the "COG groups" format can be more suitable, since the number of possible pairwise bit scores grows exponentially with the number of aligning networks.

GO categories can be optionally uploaded from "Ontologies" panel. They are provided as text files, where a list of GO cellular components, processes and functions is associated to each protein. Note that GO terms are not used for computing the alignments and can only facilitate the understanding of the alignments returned by GASOLINE.

Whenever the GO categories are provided as input, the list of GO terms for a specific protein is added as node attribute in Cytoscape, these are accessible from the "Node Browser" tabbed panel once GASOLINE ends the computation and the local alignments are ready to be visualized.

Setting the parameters

The main input GASOLINE parameters are specified in the "Parameters setting" panel. These include:

- "Iter Seed": the number of iterations of Gibbs sampling in the bootstrap phase;
- "Iter Extend": the number of iterations of Gibbs sampling in each extension step of the iterative phase;
- "Sigma": minimum network degree of nodes that can be selected as seeds in the bootstrap phase.

Values for "Iter Seed" and "Iter Extend" depend on the number of aligning networks: the more networks we have, the higher these values should be. However, based on the experiments performed on real PPI networks and reported in[15], we empirically established that 200 iterations of Gibbs sampling in both phases are enough to produce reliable results for up to 25 networks.

The choice of "Sigma" implies a trade-off between speed and accuracy of GASOLINE: the higher the σ, the faster is GASOLINE but the lower is accuracy. If networks are very sparse (like most of the existing PPI networks), low values of sigma (1 or 2) are recommended.

The "Optional parameters setting" panel contains three more input parameters:

- "Overlap": a value between 0 and 1, denoting the maximum allowed fraction of common nodes between two alignments, in order to be considered distinct. If two alignments have many nodes in common, the one with the least number of nodes is discarded from the final set;
- "Refine": the number of iterations of GASOLINE in the refinement phase;
- "Min Complex Size": the minimum size of conserved complexes in the final set of local alignments.

These parameters can be modified by checking the box "Active optional settings", otherwise the default values will be used.

Note that a high value of the "Refine" parameter can be used to increase the accuracy of the local alignments, but the algorithm will be more time consuming. In our tests[15], we experienced that a value of 10 guarantees the best trade-off between speed and accuracy of GASOLINE.

For the "Overlap" and "Min Complex Size" parameters, we suggest 0.5 and 5 as default values, respectively.

Finally, the user can specify an output folder for the final alignments, by clicking on the text field next to the "Output folder" label. Each local alignment will be stored in a separate text file into the specified folder, containing the list of aligned sub-graphs and the one-to-one mapping between aligned nodes.

Running GASOLINE

Once all the required input files are provided and all the parameters are set up, GASOLINE can be executed by clicking on the "Align" button. Then, a progress bar will appear.

Visualizing local alignments

When GASOLINE ends, a table containing all the computed local alignments is shown on the right side of the "Results panel" of Cytoscape (see Figure 1, panel B). The table reports, for each alignment, the size of the aligned complexes and the ISC (Index of Structural Conservation) score. ISC index measures the average degree of structural conservation of an alignment and can be used to evaluate its quality.

Each row of the table contains a "Show" button, for the visualization of the corresponding alignment graph on the left side of the "Results Panel" of Cytoscape (see Figure 1, panel D).

In the alignment graph, each node is labeled with the ID of the corresponding protein. If the GO annotations have been provided, the user can select one or more nodes and view the description of the proteins and their corresponding GO terms (see Figure 1, panel C).

Two kinds of edges are shown (panel D):

- Intra-edge, linking proteins of the same network, which are represented with solid colored lines;
- Inter-edge, linking proteins of different networks in the local alignment, which are drawn with dashed lines.

Colors of intra-edges depend on the probability p of the corresponding protein-protein interaction: for low values of p colors range from green to yellow, for high values of p colors range from yellow to red. Weights are automatically associated to edges as attributes, therefore the user can select an edge and retrieve its weight from the "Edge Attribute Browser" (see Figure 1, panel C).

Layout "Kamada-Kawai" has been used for the visualization of the alignment graph.

Results

Finally, we show an example of the workflow, using three well known biological networks (*C. elegans*, *D. melanogaster*, *S. cerevisiae*) taken from the STRING database[3], considering only experimentally validated interactions. We also annotated proteins by using a set of GO annotations and protein descriptions taken from BioDBnet[16]. This test was performed in a Intel Core i7-2670 2.2Ghz CPU with a RAM of 8 GB.

Following the steps described in the Implementation section, we loaded the three networks and ran GASOLINE, using the default parameters (*IterSeed* = 200, *IterExtend* = 200, *Sigma* = 7, *Overlap* = 0.5, *Refine* = 10, *MinComplexSize* = 5).

GASOLINE took 110 seconds to complete the task and returned many known conserved complexes with a high degree of topological conservation (ISC between 80 and 90%). These include the large and small subunits of ribosomes (64 proteins), a serine/threonine kinase complex (34), the spliceosome (28), a DNA repair complex (24 proteins, Figure 2), the V-ATPase complex (17) and the ARP2/3 complex (16).

Conclusions

In this paper, we presented GASOLINE, an app for Cytoscape 3 for computation and visualization of multiple local alignments of protein-protein interaction networks. To the best of our knowledge, it is the first Cytoscape plugin for computation and visualization of multiple local alignment of biological networks.

GASOLINE offers a user-friendly interface and an easy 2D visualization of local alignments. Moreover, alignments can be further investigated, by attaching GO terms to the proteins of aligning networks.

We believe that GASOLINE can be a valuable tool to better understand the functioning of some biological processes. It can help to transfer functional annotations from one species to another, to

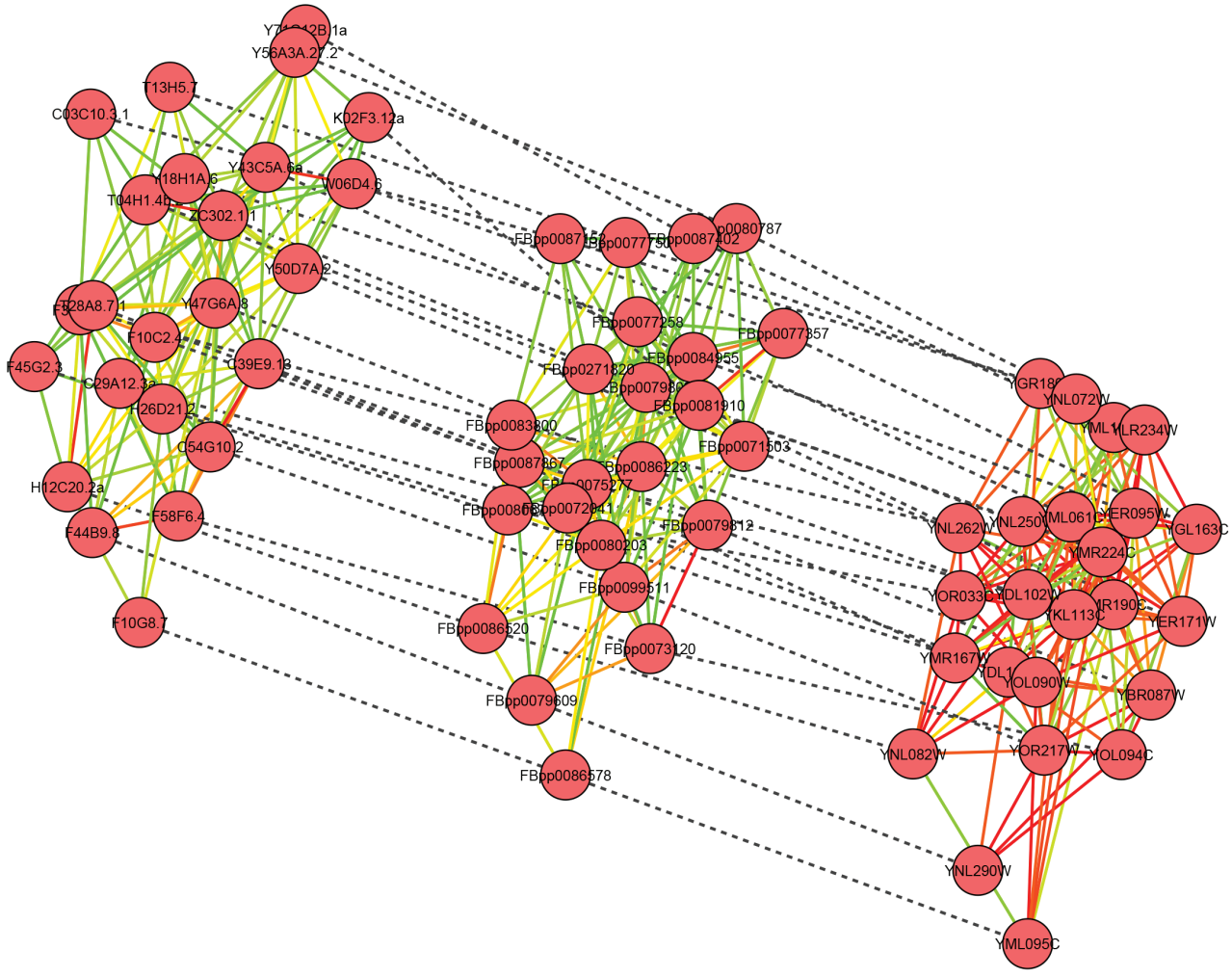

Figure 2. Alignment of DNA repair complex in *C. elegans*, *D. melanogaster* and *S. cerevisiae*.

predict new complexes or to help in establishing the function of unknown human proteins within the cells.

Software availability

The GASOLINE app, as well as datasets of real PPI networks and orthology files that can be directly used to run the algorithm, can be downloaded from the GASOLINE website: http://ferrolab.dmi.unict.it/gasoline/gasoline.html.

The GASOLINE plugin can also be downloaded from the Cytoscape App Store: http://apps.cytoscape.org/apps/gasoline.

On our website, there is also a complete documentation on the GASOLINE plugin, with more details about the format of input and output data, and a JAR file for running our algorithm in local with any platform.

Latest source code: https://github.com/GMicale/GASOLINE

Source code as at the time of publication: https://github.com/F1000Research/GASOLINE/releases/tag/V1.0

Archived source code as at the time of publication: http://www.dx.doi.org/10.5281/zenodo.104629

Author contributions

GM and AP conceived the project. GM and AC designed the architecture of Cytoscape plugin for previous versions of Cytoscape (up to version 2.8). AC wrote a first implementation of the plugin. GM enabled the porting of the code to the new versions of Cytoscape (from version 3.0 onward) and extended the functionality of the application. AP, RG and AF coordinated the project. GM, AP, RG and AF wrote the paper.

Competing interests

No competing interests were disclosed.

Grant information

The author(s) declared that no grants were involved in supporting this work.

References

1. Stark C, Breitkreutz BJ, Reguly T: **BioGRID: a general repository for interaction datasets.** *Nucleic Acids Res.* 2006; **34**(Database issue): D535–D539.
PubMed Abstract | Publisher Full Text | Free Full Text
2. Xenarios I, Salwinski L, Duan XJ, *et al.*: **DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions.** *Nucleic Acids Res.* 2002; **30**(1): 303–305.
PubMed Abstract | Publisher Full Text | Free Full Text
3. Szklarczyk D, Franceschini A, Kuhn M, *et al.*: **The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored.** *Nucleic Acids Res.* 2011; **39**(Database issue): D561–D568.
PubMed Abstract | Publisher Full Text | Free Full Text
4. Ceol A, Chatr Aryamontri A, Licata L, *et al.*: **MINT, the molecular interaction database: 2009 update.** *Nucleic Acids Res.* 2010; **38**(Database issue): D532–D539.
PubMed Abstract | Publisher Full Text | Free Full Text
5. Hodges PE, McKee AH, Davis BP, *et al.*: **The Yeast Proteome Database (YPD): a model for the organization and presentation of genome-wide functional data.** *Nucleic Acids Res.* 1999; **27**(1): 69–73.
PubMed Abstract | Publisher Full Text | Free Full Text
6. Cerami EG, Gross BE, Demir E, *et al.*: **Pathway Commons, a web resource for biological pathway data.** *Nucleic Acids Res.* 2011; **39**(Database issue): D685–D690.
PubMed Abstract | Publisher Full Text | Free Full Text
7. Barabasi AL, Oltvai ZN: **Network biology: understanding the cell's functional organization.** *Nat Rev Genet.* 2004; **5**(2): 101–113.
PubMed Abstract | Publisher Full Text
8. Kelley BP, Yuan B, Lewitter F, *et al.*: **PathBLAST: a tool for alignment of protein interaction networks.** *Nucleic Acids Res.* 2004; **32**(Web Server issue): W83–8.
PubMed Abstract | Publisher Full Text | Free Full Text
9. Kalaev M, Bafna V, Sharan R: **Fast and accurate alignment of multiple protein networks.** *J Comput Biol.* 2009; **16**(8): 989–999.
PubMed Abstract | Publisher Full Text
10. Flannick J, Novak A, Srinivasan BS, *et al.*: **Graemlin: general and robust alignment of multiple large interaction networks.** *Genome Res.* 2006; **16**(9): 1169–81.
PubMed Abstract | Publisher Full Text | Free Full Text
11. Liao CS, Lu K, Baym M, *et al.*: **IsoRankN: spectral methods for global alignment of multiple protein networks.** *Bioinformatics.* 2009; **25**(12): i253–258.
PubMed Abstract | Publisher Full Text | Free Full Text
12. Kuchaiev O, Milenkovic T, Memisevic V, *et al.*: **Topological network alignment uncovers biological function and phylogeny.** *J R Soc Interface.* 2010; **7**(50): 1341–1354.
PubMed Abstract | Publisher Full Text | Free Full Text
13. Sahraeian SM, Yoon BJ: **SMETANA: accurate and scalable algorithm for probabilistic alignment of large-scale biological networks.** *PLoS One.* 2013; **8**(7): e67995.
PubMed Abstract | Publisher Full Text | Free Full Text
14. Shannon P, Markiel A, Ozier O, *et al.*: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res.* 2003; **13**(11): 2948–2504.
PubMed Abstract | Publisher Full Text | Free Full Text
15. Micale G, Pulvirenti A, Giugno R, *et al.*: **GASOLINE: a Greedy And Stochastic algorithm for Optimal Local multiple alignment of Interaction NEtworks.** *PLoS One.* 2014; **9**(6): e98750.
PubMed Abstract | Publisher Full Text | Free Full Text
16. Mudunuri U, Che A, Yi M, *et al.*: **bioDBnet: the biological database network.** *Bioinformatics.* 2009; **25**(4): 555–556.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Referee Status: ✔ ? ✔

Version 2

Referee Report 22 October 2014

doi:10.5256/f1000research.5731.r6213

✔ **Byung-Jun Yoon**
Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA

The authors have addressed most of the comments in my previous review, and I believe the paper is suitable for indexing in its current form.

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

Competing Interests: No competing interests were disclosed.

Referee Report 26 September 2014

doi:10.5256/f1000research.5731.r5293

✔ **Antonio J. Perez Pulido**[1], **Pablo Mier**[2]
[1] UPO Genetics Bioinformatics Group, Centro Andaluz de Biología del Desarrollo (CABD), Sevilla, Spain
[2] Universidad Pablo de Olavide, Sevilla, Spain

We are agreed with the last sent changes, and we think that the article can be indexed in its current state.

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

Competing Interests: No competing interests were disclosed.

Version 1

Referee Report 15 July 2014

doi:10.5256/f1000research.4854.r5295

?

Byung-Jun Yoon
Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA

In this paper, the authors introduce a Cytoscape plug-in that can be used to perform local alignments of protein-protein interaction (PPI) networks using GASOLINE. The local network algorithm algorithm called GASOLINE was developed by the same team and it was presented in a separate journal paper. As the authors claim, a Cytoscape app for the alignment and visualization of PPI networks may serve as a useful tool for researchers working with multiple PPI networks. However, there are also several issues that need to be properly addressed to make the paper more accessible and informative for potential readers.

Abstract:

- Readers may not be familiar with GASOLINE. Please provide a more detailed description of the algorithm.

Introduction:

- "In the last few years there has been a rapid growth of biological network data, including protein-protein interaction (PPI) networks, metabolic networks and regulatory networks."

  Please add references for available PPI databases and resources.

- "Comparing PPI networks of evolutionary distant species can help to understand some mechanisms underlying a specific function or process, which the sequence comparison alone cannot explain."

  Please provide an example or cite a reference that shows when a simple sequence comparison does not suffice.

- "Local network alignment aims to compare networks of different species, in order to find conserved protein complexes or pathways."

  Define network alignment and briefly describe and compare different types of network alignments (e.g. local vs. global).

- "In literature, several network alignment algorithms have been described together with their implementations"

  At least a brief description of the popular existing methods would be helpful.

- "Here, we describe a Cytoscape app implementing the GASOLINE algorithm for multiple local alignment of PPI networks."

  Please provide at least a short introduction to GASOLINE. A brief overview of the algorithm and its pros/cons would be useful.

  Please explain what type of local alignment is obtained by GASOLINE. (e.g., one-to-one vs. many-to-many).

Implementation

- Define acronyms before use (e.g, COG, KOG, NOG, ISC)
- "The “COG groups” format can be more convenient when aligning many networks since the all the possible pairwise bit scores are many."

  Meaning unclear. Please revise.
- "GO categories can be optionally uploaded from “Ontologies” panel."

  What are the added benefits of uploading GO information? Are they used in any way during the alignment? Does it provide additional information regarding the predicted local alignments? If it does, what kind of extra info can be obtained by uploading GO files?
- Please describe the capabilities and limitations of the developed Cytoscape app on a typical personal computer. How many networks can be handled by the GASOLINE Cytoscape plug-in? How about the maximum size of the networks? (e.g., number of nodes, number of interactions). How does it scale with the number (size) of networks?
- Please describe the main difference between the original GASOLINE and the GASOLINE Cytoscape app discussed in this paper. Are there any technical differences? Are there any differences in terms of their capabilities/limitations? Any recommendation on when one should use the desktop version instead of the online version?

Conclusions

- Please discuss potential applications of the Cytoscape app.

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, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Author Response 23 Sep 2014

Giovanni Micale,

We thank the reviewer for his helpful comments and suggestions.

- ABSTRACT

Readers may not be familiar with GASOLINE. Please provide a more detailed description of the algorithm.

We added a paragraph to the Abstract section, providing a more detailed description of GASOLINE.

- INTRODUCTION

"In the last few years there has been a rapid growth of biological network data, including protein-protein interaction (PPI) networks, metabolic networks and regulatory networks."

Please add references for available PPI databases and resources.
"Comparing PPI networks of evolutionary distant species can help to understand some mechanisms underlying a specific function or process, which the sequence comparison alone cannot explain."

Please provide an example or cite a reference that shows when a simple sequence comparison does not suffice.
We added all the references.

"Local network alignment aims to compare networks of different species, in order to find conserved protein complexes or pathways."

Define network alignment and briefly describe and compare different types of network alignments (e.g. local vs. global).
We included a small paragraph describing the network alignment problem and the differences between local and global alignment methods.

"In literature, several network alignment algorithms have been described together with their implementations"

At least a brief description of the popular existing methods would be helpful.

We do not give details of most important network alignment methods, but we added references to them.

"Here, we describe a Cytoscape app implementing the GASOLINE algorithm for multiple local alignment of PPI networks."

Please provide at least a short introduction to GASOLINE. A brief overview of the algorithm and its pros/cons would be useful.

Please explain what type of local alignment is obtained by GASOLINE. (e.g., one-to-one vs. many-to-many).

We provided more details concerning GASOLINE algorithm and all its phases. Moreover, we added a new paragraph, which highlights the virtues and the current limitations of GASOLINE. In this paragraph we also stress that GASOLINE produces a one-to-one mapping.

- IMPLEMENTATION

Define acronyms before use (e.g, COG, KOG, NOG, ISC).

We defined all the acronyms.

"The “COG groups” format can be more convenient when aligning many networks since the all the possible pairwise bit scores are many."

Meaning unclear. Please revise.

We stressed the benefits of using the "COG groups" format in the alignment of many networks.

"GO categories can be optionally uploaded from “Ontologies” panel."

What are the added benefits of uploading GO information? Are they used in any way during the alignment? Does it provide additional information regarding the predicted local alignments? If it does, what kind of extra info can be obtained by uploading GO files?

We clarified that GO information are never used by GASOLINE in the alignment task. They only give, when provided by the user, useful information to better understand the final set of local alignments returned by GASOLINE.

Please describe the capabilities and limitations of the developed Cytoscape app on a typical personal computer. How many networks can be handled by the GASOLINE Cytoscape plug-in? How about the maximum size of the networks? (e.g., number of nodes, number of interactions). How does it scale with the number (size) of networks?

We specified that theoretically there is no limit on the number and size of input networks. At the end of the “Introduction” section we specified that the running time of GASOLINE is moderately affected by the size and the average density of input networks, but it is scalable with the number of aligning networks.

Please describe the main difference between the original GASOLINE and the GASOLINE Cytoscape app discussed in this paper. Are there any technical differences? Are there any differences in terms of their capabilities/limitations? Any recommendation on when one should use the desktop version instead of the online version?

At the beginning of this section we underlined that there is no technical difference between the original GASOLINE and the Cytoscape App implementation. The Cytoscape App of GASOLINE wraps the original algorithm and simply includes a visualization module. So overall, the two versions are substantially equivalent.

- CONCLUSIONS

Please discuss potential applications of the Cytoscape app.

In the last paragraph of this section we sketch some interesting applications of our app, concerning annotation transfer and protein function prediction.

Competing Interests: No competing interests were disclosed.

Referee Report 15 July 2014

doi:10.5256/f1000research.4854.r5342

? **Rintaro Saito**
School of Medicine, University of California San Diego, La Jolla, CA, USA

I was unable to install the App GASOLINE to Cytoscape 3.0.

In particular, I tried to install GASOLINE to Cytoscape 3.0.2 on Mac OS X 10.9.4 via "install from file" and from the App store via App manager. However, I got error messages, and the GASOLINE menu did not appear. I also tried the installation on another machine (Mac OS X snow leopard), but I still got the error.

Other Apps such as MCODE or BiNGO could be installed, thus I do not think that it is the problem of my specific environment.

Therefore, I would like to request authors to solve this problem.

Here is error messages:
The task could not be completed because an error
has occurred.

java.lang.Exception: java.lang.UnsupportedClassVersionError: PluginWrapper : Unsupported major.minor version 51.0
org.cytoscape.work.internal.task.JDialogTaskManager$TaskRunnable.run(JDialogTaskManager.java:307

java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:439)
java.util.concurrent.FutureTask$Sync.innerRun(FutureTask.java:303)
java.util.concurrent.FutureTask.run(FutureTask.java:138)
java.util.concurrent.ThreadPoolExecutor$Worker.runTask(ThreadPoolExecutor.java:895)
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:918)
java.lang.Thread.run(Thread.java:695)
java.lang.UnsupportedClassVersionError: PluginWrapper : Unsupported major.minor version 51.0
java.lang.ClassLoader.defineClass1(Native Method)
java.lang.ClassLoader.defineClassCond(ClassLoader.java:637)
java.lang.ClassLoader.defineClass(ClassLoader.java:621)
java.security.SecureClassLoader.defineClass(SecureClassLoader.java:141)
java.net.URLClassLoader.defineClass(URLClassLoader.java:283)
java.net.URLClassLoader.access$000(URLClassLoader.java:58)
java.net.URLClassLoader$1.run(URLClassLoader.java:197)
java.security.AccessController.doPrivileged(Native Method)
java.net.URLClassLoader.findClass(URLClassLoader.java:190)
java.lang.ClassLoader.loadClass(ClassLoader.java:306)
java.lang.ClassLoader.loadClass(ClassLoader.java:247)
org.cytoscape.app.internal.manager.SimpleApp.createAppInstance(SimpleApp.java:94)
org.cytoscape.app.internal.manager.SimpleApp.install(SimpleApp.java:173)

org.cytoscape.app.internal.manager.AppManager$1.onFileCreate(AppManager.java:339)
org.apache.commons.io.monitor.FileAlterationObserver.doCreate(FileAlterationObserver.java:379)
org.apache.commons.io.monitor.FileAlterationObserver.checkAndNotify(FileAlterationObserver.java:345)

org.apache.commons.io.monitor.FileAlterationObserver.checkAndNotify(FileAlterationObserver.java:304)

org.cytoscape.app.internal.manager.AppManager.checkForFileChanges(AppManager.java:563)
org.cytoscape.app.internal.manager.AppManager.installApp(AppManager.java:631)
org.cytoscape.app.internal.task.InstallAppFromFileTask.run(InstallAppFromFileTask.java:36)
org.cytoscape.work.internal.task.JDialogTaskManager$TaskRunnable.run(JDialogTaskManager.java:279

java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:439)
java.util.concurrent.FutureTask$Sync.innerRun(FutureTask.java:303)
java.util.concurrent.FutureTask.run(FutureTask.java:138)
java.util.concurrent.ThreadPoolExecutor$Worker.runTask(ThreadPoolExecutor.java:895)
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:918)
java.lang.Thread.run(Thread.java:695)

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, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Author Response 15 Jul 2014

Giovanni Micale,

The error you encountered was due to the fact that our app was compiled with Java 7, while the last version of Java released by Apple was version 6. So, we recompiled GASOLINE with Java 6 and we tested this new version on all operative systems and it ran without problems. This new version of GASOLINE app (version 1.2) is now available at http://apps.cytoscape.org/apps/gasoline

Competing Interests: No competing interests were disclosed.

Referee Report 10 July 2014

doi:10.5256/f1000research.4854.r5404

Antonio J. Perez Pulido[1], **Pablo Mier**[2]
[1] UPO Genetics Bioinformatics Group, Centro Andaluz de Biología del Desarrollo (CABD), Sevilla, Spain
[2] Universidad Pablo de Olavide, Sevilla, Spain

The article describes a useful plugin for Cytoscape which can be used to align protein interaction networks, including GO annotations. It is an extension of a previous paper where they presented the algorithm. This fact should be clearly mentioned, and the functionality and use of this kind of analysis should also be showed. The authors should also solve the following points:

Abstract

- The abstract should better explain what kind of alignments it makes. It could be confused with sequence local alignments.

Introduction

- What does “GASOLINE” mean? It appears on the web site but not in the actual manuscript.
- “*Comparing PPI networks of evolutionary distant species can help to understand some mechanisms*”. Indicate what mechanisms.
- “*Local network alignment aims to compare networks of different species, in order to find conserved protein complexes or pathways*”. Include reference.
- The introduction section is too short. Cytoscape should be introduced. What is it, what are the advantages from it and why was it chosen. They should also explain what a network alignment algorithm is. It is a not a usual algorithm. Its utility must be adequately motivated prior to show the new algorithm. In addition, the introduction should show that GASOLINE algorithm was previously published in another article. Is there any difference with the original algorithm? This is a very important issue because the difference with the previous article must be clear.

Implementation

- “*the number of iterations of Gibbs sampling in the bootstrap phase / in each extension step of the iterative phase*”. Explain what that is.
- “*GASOLINE is very fast; it computes the results on 25 networks in a few minutes*”. If it is a result from the previous article, please indicate this or move this to the results.
- What is the ISC score? “*Index of Structural Conservation*”. But it is not shown in this manuscript.

Results

- The test took 350 seconds on our computers. Perhaps the authors should indicate what kind of computer was used.
- Figure 2 should be used to highlight and describe the different panels and information in them.

Conclusions

- It should indicate potential and practical uses of GASOLINE to solve biological problems.

Typos

- Please, review the syntax. For example, "*estabilshed*" and some English sentences.

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

Competing Interests: No competing interests were disclosed.

Author Response 23 Sep 2014

Giovanni Micale,

We thank the reviewer for his comments and suggestions.

The article describes a useful plugin for Cytoscape which can be used to align protein interaction networks, including GO annotations. It is an extension of a previous paper where they presented the algorithm. This fact should be clearly mentioned, and the functionality and use of this kind of analysis should also be showed.

We clearly state that GASOLINE app is an extension of a previous paper and we provided some examples of application of our tool for annotation transfer and protein prediction.

In what follows we answer to each question.

- ABSTRACT

The abstract should better explain what kind of alignments it makes. It could be confused with sequence local alignments.

We clarified that GASOLINE has been implemented for solving local network alignment of PPI networks

- INTRODUCTION

What does "GASOLINE" mean? It appears on the web site but not in the actual manuscript.

We defined the acronym for GASOLINE.

"Comparing PPI networks of evolutionary distant species can help to understand some mechanisms". Indicate what mechanisms.

We defined what kind of mechanisms (regulation of specific processes of functions within the cell) can be better explained through PPI network alignment.

"Local network alignment aims to compare networks of different species, in order to find conserved protein complexes or pathways". Include reference.

We provided references for the most important local alignment methods.

The introduction section is too short. Cytoscape should be introduced. What is it, what are the advantages from it and why was it chosen.

We added a small paragraph, where we briefly introduced Cytoscape, explaining its advantages.

They should also explain what a network alignment algorithm is. It is a not a usual algorithm. Its utility must be adequately motivated prior to show the new algorithm.

The network alignment problem has been described more clearly and the differences between local and global alignment methods have been remarked.

In addition, the introduction should show that GASOLINE algorithm was previously published in another article. Is there any difference with the original algorithm? This is a very important issue because the difference with the previous article must be clear.

In the "Implementation" section we clarified that there is no technical difference between GASOLINE Cytoscape app and the original algorithm (for which we provided the reference).

- IMPLEMENTATION

"the number of iterations of Gibbs sampling in the bootstrap phase / in each extension step of the iterative phase". Explain what that is.

All the GASOLINE parameters have been better described in the "Introduction" section, where we provided more details of GASOLINE. We indicated that GASOLINE is based on an iterative Gibbs sampling algorithm, which is applied both in the bootstrap and in the iterative phase (extension steps).

"GASOLINE is very fast; it computes the results on 25 networks in a few minutes". If it is a result from the previous article, please indicate this or move this to the results.

This result is taken from the original paper of GASOLINE. However, we decided to remove it from this section and from the whole paper. We specified that GASOLINE is scalable with the number of networks in the "Introduction" section, and we only indicated the running time of GASOLINE for the test reported in the "Results" section.

What is the ISC score? "Index of Structural Conservation". But it is not shown in this manuscript.

We defined what the ISC score is and how it is used to evaluate the quality of local alignments.

- RESULTS

The test took 350 seconds on our computers. Perhaps the authors should indicate what kind of computer was used.

We indicated the technical characteristics of the machine used for the tests.

Figure 1 should be used to highlight and describe the different panels and information in them.

We revised Figure 1 highlighting and describing each GASOLINE panel.

- CONCLUSIONS

It should indicate potential and practical uses of GASOLINE to solve biological problems.

We give few more details concerning the potential applications of GASOLINE in annotation transfer and protein function prediction.

- TYPOS

Please, review the syntax. For example, "estabilshed" and some English sentences.

Fixed.

Competing Interests: No competing interests were disclosed.