SOFTWARE

Open Access

Biocaiv: an integrative webserver for motif-based clustering analysis and interactive visualization of biological networks

Dong-Xu Li[1,2,3], Peng Zhou[4], Bo-Wei Zhao[1,2,3], Xiao-Rui Su[1,2,3], Guo-Dong Li[1,2,3], Jun Zhang[1,2,3], Peng-Wei Hu[1,2,3] and Lun Hu[1,2,3*]

*Correspondence:
hulun@ms.xjb.ac.cn

[1] The Xinjiang Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, Ürümqi, China
[2] University of Chinese Academy of Sciences, Beijing, China
[3] Xinjiang Laboratory of Minority Speech and Language Information Processing, Ürümqi, China
[4] School of Computer Science and Artificial Intelligence, Wuhan University of Technology, Wuhan, China

Abstract

Background: As an important task in bioinformatics, clustering analysis plays a critical role in understanding the functional mechanisms of many complex biological systems, which can be modeled as biological networks. The purpose of clustering analysis in biological networks is to identify functional modules of interest, but there is a lack of online clustering tools that visualize biological networks and provide in-depth biological analysis for discovered clusters.

Results: Here we present BioCAIV, a novel webserver dedicated to maximize its accessibility and applicability on the clustering analysis of biological networks. This, together with its user-friendly interface, assists biological researchers to perform an accurate clustering analysis for biological networks and identify functionally significant modules for further assessment.

Conclusions: BioCAIV is an efficient clustering analysis webserver designed for a variety of biological networks. BioCAIV is freely available without registration requirements at http://bioinformatics.tianshanzw.cn:8888/BioCAIV/.

Keywords: Clustering analysis, Biological networks, Bioinformatics

Background

In biology, many complex biological systems can be modeled as networks where individual biological entities, such as different kinds of molecules, are nodes and the connections among them are links [1, 2]. The purpose of clustering analysis in biological networks is to identify functional modules of interest, such as protein complexes and gene co-expression modules. These modules are of great significance to provide valuable insights into the understanding of corresponding biological systems [3].Therefore, many related methods and works on biological networks or clustering have been proposed [4–6].

Network motifs, that is, patterns of mutual connections that appear in complex networks [7]. In the context of networks, higher-order structures typically refer to the connectivity patterns involving more than just pairwise connections between nodes. Motifs are recurring and small non-random subgraphs that represent higher-order structures in a network, and their occurrences are more frequently than expected in a random network. To better illustrate the meaning of motifs, we also provide a graphical explanation in Fig. 1. We have recently presented a more general Higher-order Structural Clustering Framework (HiSCF), which aims to identify functional modules in a given biological network based on a variety of specific network motifs. HiSCF is introduced as a powerful clustering analysis tool designed for biological networks, with a focus on capturing higher-order structures, or motifs. It begins by constructing a tensor that encodes all instances of a given motif within a biological network. This step is crucial for capturing the structural characteristics of the motif. To address the clustering problem, HiSCF formulates it as a higher-order Markov chain model, which is computationally intractable. Hence, it leverages the concept of spacey random walk theory to approximate the higher-order Markov chain by a first-order Markov chain. The Markov Clustering Algorithm is then employed by using the transition matrix of the first-order Markov chain, allowing for the effective identification of clusters specific to the network motif under consideration [8]. However, the implementation of HiSCF integrates several complex programming environments python, Julia, and java, thus limiting its wide applicability. The compatibility issue is also difficult to address for non-programmer biological researchers who are not familiar with Python, Julia or Java. It is for this reason that we would like to develop an online framework that facilitate the implementation and usage of HiSCF in a concise and reproducible manner.

Currently, only few web tools have been made available to the scientific community on the task of clustering analysis for biological networks. They are either in the form of plugins or web-based. For example, GraphWeb [9] is a public web server developed for discovering functional modules through biological network analysis. To improve the clustering efficiency, SPICi [10] is dedicated as a web tool to provide fast clustering analysis for large biological networks. ClusterViz [11] is designed as a plugin package of Cytoscape to conduct the clustering analysis of biological networks. To our knowledge, no new online clustering tools have emerged recently for biological networks. In particular, none exists for visualizing biological networks and providing an in-depth analysis for discovered clusters.

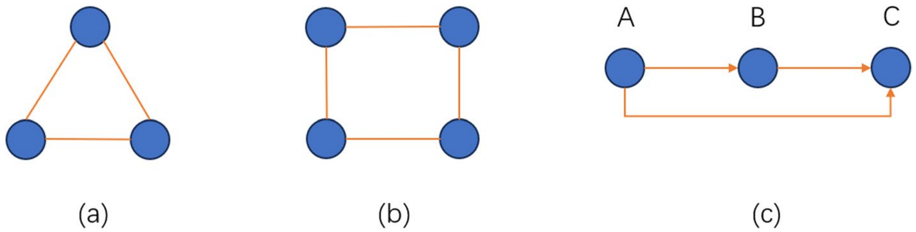


Fig. 1 Three representative motifs in biological networks, **a** triangle motif, **b** rectangle motif and **c** the motif of feed-forward loop

In this context, we present a more comprehensive and user-friendly web server, namely BioCAIV, which integrates HiSCF to offer motif-based clustering analysis for biological networks. BioCAIV is composed of three modules: (1) a backend service to reimplement HiSCF, (2) a web interface to upload network files and present clustering results, (3) a visualization tool able to visualize the input network, and (4) a functional enrichment analysis service to investigate the functional correlation of molecules in the clusters.

Currently, BioCAIV offers two distinct operational modes: the Normal Mode and the Custom Mode, each catering to specific network characteristics and user preferences. In the Normal Mode of BioCAIV, the platform is designed to support undirected networks. This limitation arises from the input format of network data, which does not accommodate directional information. For users who require the flexibility to work with both directed and undirected networks, BioCAIV offers the Custom Mode. In this mode, the platform provides comprehensive support for both network types. Users can process directed and undirected networks within BioCAIV, provided that all motif instances extracted from the biological network are correctly formatted in the input file.

BioCAIV is a promising tool capable of performing clustering analysis on a wide range of biological networks. However, there are specific considerations when it comes to functional enrichment analysis, depending on the nature of the biological network. First of all, BioCAIV is capable of conducting clustering tasks on a variety of biological networks, regardless of their types. For biological networks such as protein–protein interaction networks, gene regulatory networks, and gene co-expression networks, users can leverage BioCAIV's clustering capabilities for these network types without restrictions. Secondly, functional enrichment analysis within BioCAIV is specifically tailored for biological networks composed of genes, or gene-coded proteins. Hence, biological networks containing gene-related nodes, such as protein–protein interaction networks, gene regulatory networks, and gene co-expression networks, are eligible for functional enrichment analysis. Residue interaction networks, characterized by nodes representing individual protein residues, are not suitable for functional enrichment analysis provided by g:Profiler, as their nodes do not involve gene-level entities.

This new web server addresses the drawbacks of HiSCF as it is easy to use with a web interface. Specifically, when compared with HiSCF, BioCAIV is preferred from several aspects. First, it can be easily accessed through internet without compiling sources codes. Second, unlike HiSCF that only supports triangle and rectangle motifs, it allows users to perform clustering analysis for any higher-order structures they are interested in. Last, we also provide the results of functional enrichment analysis, thus enabling users to refine identified modules that are functionally significant and should be experimentally verified for potential applications. In particular, there are several distinct advantages to using BioCAIV in comparison to other programs. These advantages make BioCAIV a highly attractive tool for users in the field of bioinformatics, and they are listed as below.

1 *User-Friendly Accessibility*: BioCAIV eliminates the need for users to possess specialized knowledge on compiling source code written in various programming languages. Unlike standalone programs like HiSCF, BioCAIV offers an accessible and hassle-free experience for researchers, regardless of their programming expertise.

2 *Seamless Computational Resources*: Traditional standalone clustering software, like Cytoscape, often rely on the computational power of users' local computers. In contrast, BioCAIV centralizes all necessary computational resources on its dedicated server. This approach ensures consistent and efficient performance, irrespective of the user's hardware capabilities.
3 *Advanced Clustering Technique and User-Centric Features*: BioCAIV distinguishes itself from existing online clustering tools in two significant ways. Firstly, it employs a state-of-the-art clustering algorithm, resulting in more accurate and insightful analyses. Secondly, it provides a range of user-friendly functions, including email notifications, functional enrichment analysis, and network visualization. These features enhance the overall usability and productivity of the platform.

In light of these compelling advantages, we firmly believe that BioCAIV stands as an valuable tool for the clustering analysis of biological networks.

Implementation

As a modular web server, BioCAIV is more advanced than existing clustering tools by providing three features enabling it to visualize and analyze biological networks, and search functionally significant clusters. The overall framework of BioCAIV is presented in Fig. 2.

Input module

BioCAIV currently operates in two modes. One is *Default* mode where two kinds of motifs, i.e., Triangle and Rectangle, are supported by default. The other is *Custom* mode where clustering analysis can be performed on any particular motif specified by users. To use BioCAIV, users are required to input biological networks or motif instances corresponding to different modes by specifying the following information. The workflow of input module is presented in Fig. 3.

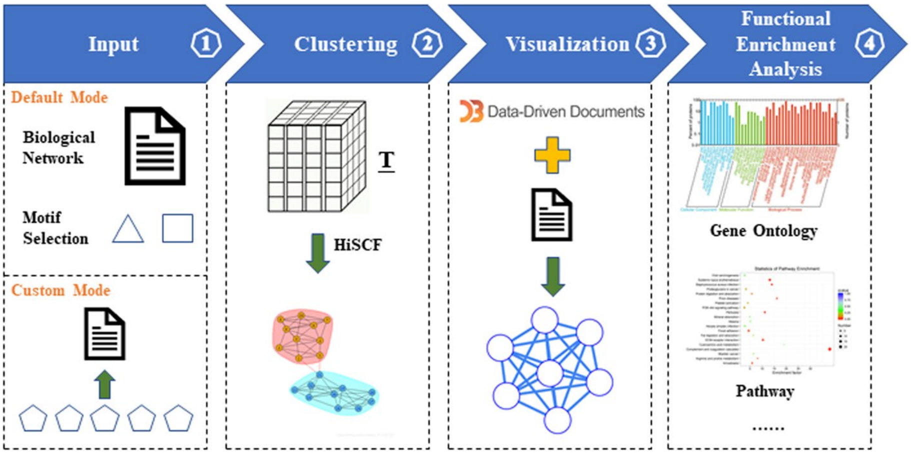
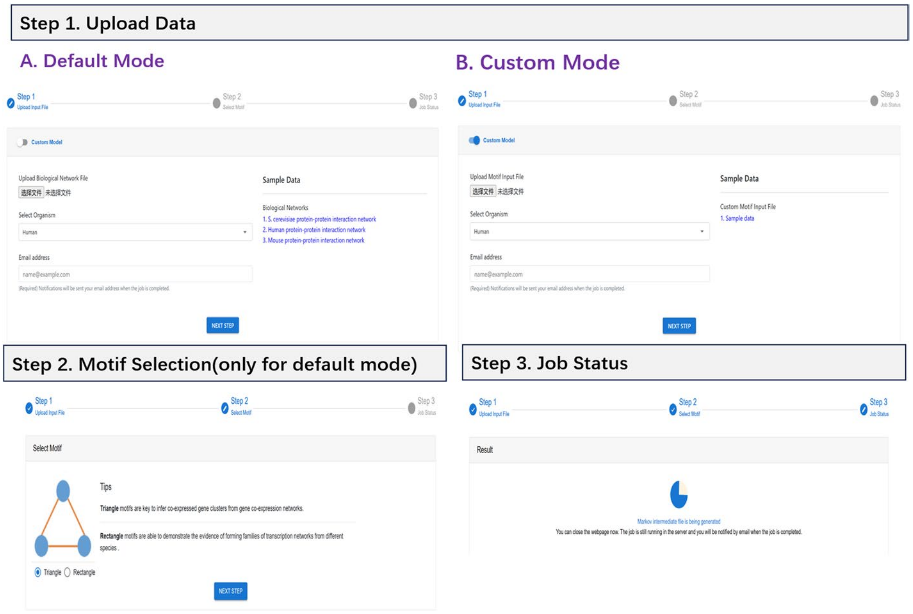


Fig. 2 Overview of the BioCAIV pipeline. Step 1. Input is provided in the formatted file that specifies details on the biological network to be analyzed. Step 2. A tensor-based variable is constructed to apply HiSCF for discovering clusters. Step 3. BioCAIV makes use of D3.js to fastly visualize the input network with interactive functions. Step 4. Clustering results are further analyzed to search functionally significant clusters



Select Organism: Human

Email address: name@example.com

(Required) Notifications will be sent your email address when the job is completed.

Triangle motifs are key to infer co-expressed gene clusters from gene co-expression networks.

Rectangle motifs are able to demonstrate the evidence of forming families of transcription networks from different species .

● Triangle ○ Rectangle

Markov intermediate file is being generated

You can close the webpage now. The job is still running in the server and you will be notified by email when the job is completed.

Fig. 3 The workflow of input module

1. Default Mode

   - Biological Network File: the file can be uploaded via BioCAIV. Three sample network files are provided as an illustration to the data format. They are, S. cerevisiae protein–protein interaction network, Human protein–protein interaction network and Mouse protein-protein interaction network. Incorrect biological network files will lead to the failure of clustering analysis. With respect to the input format, BioCAIV utilizes a straightforward and distinctive representation. Two specific letters, ’v’ and ’e’, are employed to indicate nodes and edges, respectively, within the input file. In particular, rows beginning with ’v’ are used to denote node identifiers. These ’v’-prefaced rows contain information about individual nodes within the network. Rows that start with ’e’ are employed to describe the connections between pairwise nodes. These ’e’-prefaced rows specify the relationships or edges between nodes in the network.
   - Motif Selection: users need to select a motif, either Triangle or Rectangle, used for clustering analysis.

2. Custom Mode
   - Motif Input File: the file is uploaded via BioCAIV. Users have to create this file based on the motif they are interested in. A sample motif file is provided to describe the data format. In particular, each line is composed of nodes separated by tab and it is an instance of that particular motif in the biological network.

One should note that in the either biological network or motif input files, molecules have to be represented by their corresponding Entrez Gene IDs if functional enrichment analysis is applied. In addition, the following information should be provided by users for both modes.

- Organism Selection: currently BioCAIV supports Three organisms, i.e., Human, Saccharomyces Cerevisiae and Mouse for functional enrichment analysis. More organisms will be included upon the requests of users.
- Email: once the job is submitted, a link to the job status page will be sent to the user. Moreover, a notification will also be sent to the user when the job is completed.The algorithm takes a certain amount of time to run. During the running time, there is no need to wait on the page all the time. You can close the page to check the prompt email in the email and click the "View Results" button to jump to the result page.

Upon successful submission of the input data, users are redirected to the job page, where the current status is presented and the results will be displayed upon completion.

Visualization module

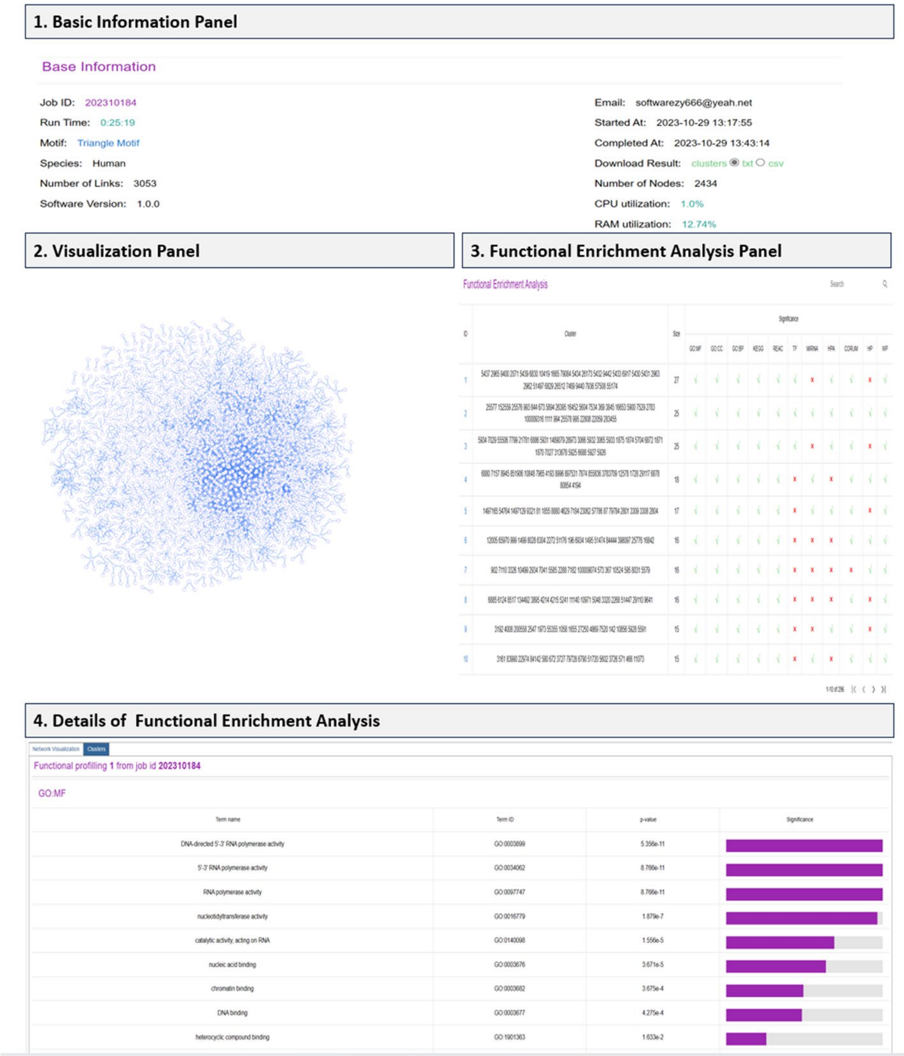
In the post-genomics era, biological networks have been becoming increasingly larger. Their visualization can be troublesome and time-consuming, thus requiring efficient and easy-to-use interfaces to present their topological structures in an autonomous manner. In BioCAIV, we visualize the input biological network by using D3.js [12], which is a JavaScript library having full capabilities of modern browsers without installing any plugins. Due to the minimal overhead taken by D3.js, BioCAIV is able to support the visualization of large biological networks with fast interactive behaviors.

Clustering analysis module

Before performing the clustering analysis, BioCAIV first defines a tensor variable to indicate all instances of the selected motif. Taking the triangle motif as an example, a three-mode tensor is defined as $\underline{\boldsymbol{T}} = \left(\underline{t}(i,j,k)\right)$. Regarding the value of $\underline{t}(i,j,k)$, we have $\underline{t}(i,j,k) = 1$ if nodes v_i, v_j and v_k construct a triangle, and $\underline{t}(i,j,k) = 0$ otherwise. BioCAIV then follows the procedure of HiSCF to identify clusters upon $\underline{\boldsymbol{T}}$.

The clustering analysis job is usually completed less than half an hour for a biological network composed of thousands of nodes. Even users close the webpage, the job is still running in the server side and users will be notified by email when the job is completed. The clustering results are composed of three panels displaying the following information. An overview of the output page is reported in Fig. 4.

- Basic Information. In addition to Job ID and Email address, we also provide the time information for the job. The clustering result is available to download by clicking the "clusters" link. we have included additional metrics in the result page. Alongside the running time, the average usage of CPU and RAM resources is now available. This information offers a more detailed perspective on the computational demands of the analysis.
- Network Visualization. The input biological network is visualized in this panel, which also allows users to interact with the network.
- Clusters. In this panel, we list the discovered clusters as well as their component nodes. Besides, we also present the overall assessment regarding the functional significance of clusters.



1. Basic Information Panel

Base Information

Job ID: 202310184
Run Time: 0:25:19
Motif: Triangle Motif
Species: Human
Number of Links: 3053
Software Version: 1.0.0

Email: softwarezy666@yeah.net
Started At: 2023-10-29 13:17:55
Completed At: 2023-10-29 13:43:14
Download Result: clusters ● txt ○ csv
Number of Nodes: 2434
CPU utilization: 1.0%
RAM utilization: 12.74%

2. Visualization Panel

3. Functional Enrichment Analysis Panel

Functional Enrichment Analysis

ID	Cluster	Size	Significance										
			GO:MF	GO:CC	GO:BP	KEGG	REAC	TF	MIRNA	HPA	CORUM	HP	MP
1	5437 2965 9400 2071 5439 6830 10419 1665 79084 5434 26173 5432 9442 5433 6917 5430 5431 2963 2962 51497 6829 26512 7469 9440 7936 57508 55174	27	√	√	√	√	√	√	x	√	√	x	√
2	25577 152559 25576 993 844 673 5894 26395 16452 5604 7534 369 3845 16653 5900 7529 2783 100009316 1111 994 25578 995 22808 22059 283455	25	√	√	√	√	√	√	√	√	√	√	√
3	5604 7029 55506 7799 21781 6886 5931 1489079 28973 3066 5932 3065 5933 1875 1874 5704 6872 1871 1870 7027 313678 5925 6688 5927 5926	25	√	√	√	√	√	√	x	√	√	x	√
4	6880 7157 8945 851906 10848 7965 4193 8996 867531 7874 855836 3783709 12578 1728 29117 6878 80854 4194	18	√	√	√	√	√	x	√	x	√	√	√
5	1467165 54764 1497129 9321 81 1855 8880 4629 7184 23082 57796 67 79784 2801 3309 3308 2904	17	√	√	√	√	√	x	√	√	√	x	√
6	12005 65970 999 1499 8028 6304 2272 51176 196 6934 1495 51474 84444 398097 25776 16842	16	√	√	√	√	√	x	x	x	√	√	√
7	902 7110 3326 10499 2934 7041 5585 2288 7182 100009874 573 367 10524 595 8031 5579	16	√	√	√	√	√	x	x	x	x	√	√
8	6885 6124 8517 134492 3895 4214 4215 5241 11140 10971 5048 3320 2288 51447 29110 9641	16	√	√	√	√	√	x	x	x	√	x	√
9	3192 4008 290558 2547 1973 55355 1058 1655 27250 4869 7520 142 10856 5928 5591	15	√	√	√	√	√	x	x	√	√	x	√
10	3161 83990 22974 84142 580 672 3727 79728 6790 51720 5902 3726 571 466 11073	15	√	√	√	√	√	x	√	x	√	√	√

1-10 of 296

4. Details of Functional Enrichment Analysis

Network Visualization | Clusters

Functional profilling 1 from job id 202310184

GO:MF

Term name	Term ID	p-value	Significance
DNA-directed 5'-3' RNA polymerase activity	GO:0003899	5.356e-11	
5'-3' RNA polymerase activity	GO:0034062	8.766e-11	
RNA polymerase activity	GO:0097747	8.766e-11	
nucleotidyltransferase activity	GO:0016779	1.879e-7	
catalytic activity, acting on RNA	GO:0140098	1.556e-5	
nucleic acid binding	GO:0003676	3.671e-5	
chromatin binding	GO:0003682	3.675e-4	
DNA binding	GO:0003677	4.275e-4	
heterocyclic compound binding	GO:1901363	1.633e-2	

Fig. 4 BioCAIV provides (1) basic information panel of the task, (2) the visualization of input biological network, (3) the overall picture of functional enrichment analysis of all clusters, and (4) the details of functional enrichment analysis for each cluster

Output file

The output file generated by BioCAIV adopts a simplified structure. Each row in the output file comprises node identifiers separated by spaces. These rows effectively represent functional modules or clusters identified by BioCAIV during the clustering analysis. In addition to the plain text output format, we have introduced an enhanced output format: CSV (Comma-Separated Values). This CSV format is designed for compatibility with various data analysis applications, including popular tools like Cytoscape, ensuring seamless integration for downstream analysis.

Functional enrichment analysis module

In this module, users can retrieve the result of functional enrichment analysis, which is driven by Python script of g:Profiler [13], for each cluster. In particular, BioCAIV additionally provides functional enrichment analysis for identified modules from different resources of biological information, including functional categories from Gene Ontology [14], pathways from WikiPathways [15], regulatory motif matches from TRANSFAC [16] and human disease phenotypes from Human Phenotype Ontology [17]. In the panel of functional enrichment analysis, an overall picture is presented to indicate the significance of all clusters under different biological information. By clicking the cluster ID, significant terms identified by g:profiler in each resource are displayed for the corresponding cluster. When BioCAIV retreives the functional enrichment analysis from g:profiler, the network interruption may occur. To overcome this issue, BioCAIV first popups a message box showing the text "Failed to retrieve the results of functional enrichment analysis from g:profiler. Reload or not?" If users face the data loading problem. Once the OK button is clicked, BioCAIV will reload the current page to retrieve the results from g:profiler again.

Use of case

In particular, we take the sample data of Human protein–protein interaction network as the input biological network. Since feed-forward loops are neither triangle nor rectangle motifs, we opt to select the Custom Model. Then a customized input file should be prepared for providing the feed-forward loops in the network. In this file, each row represents a unique feed-forward loop extracted from the network, and it should consist of node identifiers that constitute that loop and are separated by tabs. Upon uploading the customized input file, users will be prompted to input their email addresses. This is essential to ensure that users receive timely notifications regarding the processing status of their job. Once users click "Next Step," they will be redirected to the result page. This page provides real-time information about the status of the current job, allowing users to monitor the progress of their analysis. In addition to that, we set up the "Functional enrichment analysis on disease analysis" example on the Use of Case page. In this case study, we take the clustering results, obtained from Human protein–protein interaction network by using the Rectangle Motif, as an example to provide disease analysis through functional enrichment analysis.

The organization of BioCAIV

As for BioCAIV, the *Home* page is organized with a top banner displaying several available sections. Dedicated tabs help users easily navigate to browse the content of each section. Moreover, it also describes the general information of BioCAIV and presents the input form to initialize a new clustering analysis task. It is also possible to run a test job by using the sample data provided in *Home*.

Besides *Home*, the *Contact*, *References* and *Help* pages are also provided. Explanations regarding the usage of BioCAIV and the required input formats are reported in *Manual*. The implemented clustering algorithm, i.e., HiSCF, is introduced in *Method*. The *References* page provides the relevant literature related to BioCAIV and the *Help* page provides links to feedback the problems encountered by users as well as their suggestions on this web tool.

The BioCAIV web server is written in Python 3.7 based on the Tornado framework. The server is containerized with Virtualvenv and deployed on a dedicated GNU/Linux server.

Results and discussion

In BioCAIV, the integration of tensor-based data structures and efficient clustering algorithm allows users to study the higher-order organization patterns in biological networks. On the one hand, BioCAIV simplifies the clustering analysis by offering easy-to-use functionalities, and also creates an intuitive interface to visualize the biological networks. BioCAIV renders a variety of motifs into a uniform analysis pipeline, with which clusters can thus be identified upon a specific motif. On the other hand, looking at the topological structures of clusters may admittedly fail to provide valuable insights into biological systems. To overcome this problem, BioCAIV allows researchers to approach a number of functional enrichment analysis tasks from different biological information perspectives. Functionally significant clusters are worth further investigation. The main advantage of BioCAIV is its simple interface and fast speed in visualizing and handling large biological networks for clustering analysis. Hence, we have reason to believe that BioCAIV is a useful web server in discovering and analyzing clusters from biological networks, since it allows users to reduce the analysis complexity. Such a reduction may broaden the feasibility of BioCAIV to more potential applications in the field of bioinformatics.

Regarding the future work, we would like to unfold it from four aspects. First, BioCAIV currently provides two motifs, i.e., triangle and rectangle, for clustering analysis, and more motifs will be implemented in *Default* mode to facilitate the usage of BioCAIV. Second, two organisms are supported by BioCAIV to perform functional enrichment analysis on discovered clusters, and we are interested in converting more organisms in future. Third, more interactive operations will be incorporated into BioCAIV when we visualize the biological networks as well as discovered clusters. Such an improvement may provide access to more topological analysis methods. Last, we also would like to explore more potential applications by using BioCAIV, such as drug repositioning [18] and miRNA-disease association prediction [19].

Conclusions

We have presented here a novel web server, namely BioCAIV, that allows users to perform an accurate clustering analysis for biological networks by using the higher-order structural information provided by users. As an extension of HiSCF, this web server needs only biological network file or custom motif file, not requiring any additional data and it is preferred for clustering analysis over HiSCF due to its user-friendly interface and free availability. Thus, its applicability will be maximized by targeting to a large audience, especially for biological researchers with limited computational resources. Moreover, with the creation of BioCAIV, we wish to be dedicated to addressing the clustering tasks for understanding the functional mechanisms of different biological networks, and providing a tool by emphasizing its ease of use.

Abbreviation

HiSCF Higher-order structural clustering framework

Acknowledgements

The authors would like to thank all anonymous reviewers for their constructive advice.

Author contributions

D.-X.L., P.Z. and L.H. conceived the webserver, D.-X.L. and P.Z. implemented the webserver. B.-W.Z., X.-R.S., and G.-D.L. performed the analyses. D.X.L. and L.H. wrote the main manuscript. J.Z. and P.-W.H. revised the manuscript. L.H. supervised the project. All authors read and approved the final manuscript.

Funding

Natural Science Foundation of Xinjiang Uygur Autonomous Region [2021D01D05]; Tianshan Youth Project–Outstanding Youth Science and Technology Talents of Xinjiang [2020Q005]; CAS Light of the West Multidisciplinary Team Project [xbzg-zdsys-202114]; Pioneer Hundred Talents Program of Chinese Academy of Sciences.

Availability of data and materials

Project name: BioCAIV.
Project home page: http://bioinformatics.tianshanzw.cn:8888/BioCAIV/.
Operating system(s): Platform independent.
Programming language: Python.
Other requirements: Python 3.7 or higher.
License: GNU GPL.
Any restrictions to use by non-academics: No restrictions.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 5 July 2023 Accepted: 20 November 2023
Published online: 29 November 2023

References

1. Girvan M, Newman ME. Community structure in social and biological networks. Proc Natl Acad Sci. 2002;99(12):7821–6.
2. Alon U. Biological networks: the tinkerer as an engineer. Science. 2003;301(5641):1866–7.
3. Hu L, Yuan X, Liu X, Xiong S, Luo X. Efficiently detecting protein complexes from protein interaction networks via alternating direction method of multipliers. IEEE/ACM Trans Comput Biol Bioinf. 2019;16(6):1922–35.
4. Hu L, Yang Y, Tang Z, He Y, Luo X. Fcan-mopso: an improved fuzzy-based graph clustering algorithm for complex networks with multi-objective particle swarm optimization. IEEE Trans Fuzzy Syst. 2023.
5. Wang X, Yang W, Yang Y, He Y, Zhang J, Wang L, Hu L. Ppisb: a novel network-based algorithm of predicting protein–protein interactions with mixed membership stochastic block model. IEEE/ACM Trans Comput Biol Bioinf. 2022;20(2):1606–12.
6. Zhao B-W, Su X-R, Hu P-W, Huang Y-A, You Z-H, Hu L. igrldti: an improved graph representation learning method for predicting drug-target interactions over heterogeneous biological information network. Bioinformatics. 2023;39(8):451.
7. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. Network motifs: simple building blocks of complex networks. Science. 2002;298(5594):824–7.
8. Hu L, Zhang J, Pan X, Yan H, You ZH. Hiscf: leveraging higher-order structures for clustering analysis in biological networks. Bioinformatics. 2021;37(4):542–550.
9. Reimand J, Tooming L, Peterson H, Adler P, Vilo J. Graphweb: mining heterogeneous biological networks for gene modules with functional significance. Nucleic Acids Res. 2008;36(suppl 2):452–9.
10. Jiang P, Singh M. Spici: a fast clustering algorithm for large biological networks. Bioinformatics. 2010;26(8):1105–11.
11. Wang J, Zhong J, Chen G, Li M, Wu F-X, Pan Y. Clusterviz: a cytoscape app for cluster analysis of biological network. IEEE/ACM Trans Comput Biol Bioinf. 2015;12(4):815–22.
12. Zhu NQ. Data visualization with D3. Js cookbook. Birmingham: Packt Publishing Ltd; 2013.
13. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. g:profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 2019;47(W1):191–8.
14. Consortium GO. The gene ontology resource: 20 years and still going strong. Nucleic Acids Res. 2019;47(D1):330–8.
15. Slenter DN, Kutmon M, Hanspers K, Riutta A, Windsor J, Nunes N, Mélius J, Cirillo E, Coort SL, Digles D. Wikipathways: a multifaceted pathway database bridging metabolomics to other omics research. Nucleic Acids Res. 2018;46(D1):661–7.

16. Matys V, Fricke E, Geffers R, Gößling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV. Transfac®: transcriptional regulation, from patterns to profiles. Nucleic Acids Res. 2003;31(1):374–8.
17. Köhler S, Vasilevsky NA, Engelstad M, Foster E, McMurry J, Aymé S, Baynam G, Bello SM, Boerkoel CF, Boycott KM. The human phenotype ontology in 2017. Nucleic Acids Res. 2017;45(D1):865–76.
18. Zhao BW, Wang L, Hu PW, Wong L, Su XR, Wang BQ, You ZH, Hu L. Fusing higher and lower-order biological information for drug repositioning via graph representation learning. IEEE Trans Emerg Top Comput. 2023. https://doi.org/10.1109/TETC.2023.3239949.
19. He Y, Yang Y, Su X, Zhao B, Xiong S, Hu L. Incorporating higher order network structures to improve mirna-disease association prediction based on functional modularity. Brief Bioinform. 2023;24(1):562.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.