Briefings in Functional Genomics, 20(4), 2021, 196–197

https://doi.org/10.1093/bfgp/elab033
Editorial

EDITORIAL

Briefings in functional genomics special section editorial: analysis of integrated multiple omics data

With the rapid development of high-throughput sequencing technology, omics data that are available as genomics, transcriptomics, proteomics and metabolomics, have been widely used to study the underlying mechanisms of important biological processes. In recent times, multi-omics data integration has attracted extensive attention in the field of medicine and biology, such as conceptual introduction of multi-omics [1–4], multi-omics data processing techniques [5–8], and disease treatment and prevention [5, 9]. Compared with a single omics, multi-omics data can provide more comprehensive and extensive information, thus integrated analysis of multi-omics data can verify and help to complement each other, which can help us identify better biomarkers for high-precision diagnosis, prognosis and predictive treatment.

In the era of big biological data and precision medicine, we organized this Special Issue of Briefings in Functional Genomics on the computational analysis of multi-omics data aiming at bringing researchers together to disseminate their novel techniques for analysis of multiple omics data while exploring the significance from big biomedical data analysis for complex pathological conditions.

Five articles focus on the analysis methods and techniques of multi-omics data and integrative analysis for complex diseases, such as human brain disorders, cancer and virus infection. The article by Augustyn *et al.* [10] presents the application of Cloud computing in integrative analysis of multi-omics data, which can be as solutions for scaling and building independent analysis pipelines for omics data. A verified concept model that exhibits the potentials for performing integrative analysis of multiple omics data sources is introduced as a universal solution. The article by Wang [11] summarizes the general principles and approaches to dissect the regulatory mechanisms of post-transcriptional processes by integrating multi-omics data. The studies of post-transcriptional processing of RNAs based on integrative analyses of multiple omics data, such as RNA binding proteins, epigenomic data and ribosome profiling (ribo-seq) data, are introduced. In the article by Kaur *et al.* [12], various of computational resources and tools for identifying cancer biomarkers based on multiple omics data are summarized. The computational resources and tools are divided into several categories, including cancer-associated multi-omics data repositories, visualization or analysis tools and algorithms for omics data, machine learning-based diagnostic, prognostic, and predictive biomarker tools. Various algorithms for the identification of cancer biomarkers are also listed, such as algorithm based on machine learning, deep learning, and survival analysis algorithms. The article by Dong *et al.* [13] reviews the recent research of human brain disorders by using multi-omics data resources, including schizophrenia, autism, bipolar disorder, Alzheimer's disease, Parkinson's disease, progressive supranuclear palsy, etc. Single-cell omics data in recent brain research are also mentioned, such as single-nucleus RNA-seq, single-cell ATAC-seq and spatial transcriptomics. The limitations of the multi-omics study about human brain disorders, such as high-dimension reduction challenge, and harmonization and heterogeneity of the omics data source, are also discussed. The article by Wang *et al.* [14] propose the strategy of 'precision omics', i.e. the combinatorial strategy of omics technologies and precision medicine, and illustrate how large-scale molecular omics data can be performed with the support of interoperable ontologies for disease treatment and prevention. In the article, interoperable ontology-supported precision COVID-19 omics studies are illustrated to explain the proposed precision omics ontology hypothesis, assuming that the effectiveness of precision omics is positively correlated with the interoperability of ontologies.

Two articles focus on the prediction of gene function or biomolecule-disease associations by using network-based computational methods or machine learning methods. In the article by Chen *et al.* [15], the network technologies used for gene functional prediction are investigated, and the multi-data source fusion is proved to be an effective manner to improve gene function prediction. The network-based gene function prediction methods are classified into four types, namely neighborhood-based method, kernel function-based method, random walk technology-based method and matrix decomposition technology-based method. Advantages of network-based method for gene function prediction are elaborated from several aspects. The article by Ding *et al.* [16] presents machine learning approaches for predicting associations between biomolecules (such as microRNAs, circRNAs, lncRNAs and genes) and diseases. It mainly introduces three parts: data sources for prediction models, feature representation methods for machine learning-based prediction models, and machine learning algorithms for

biomolecule-disease association prediction. Eight feature representation methods are summarized, including one hot encoding, autoencoder representation, node2vec representation, graph representation, manifold learning representation, etc. Three categories of machine learning-based prediction methods are summarized, namely, basic machine learning prediction methods, matrix completion-based methods, and deep learning-based methods. Moreover, advantages and disadvantages of these categories of prediction methods are discussed. Suggestions such as the multiple data source integration strategies are also provided.

The article by Hasan *et al.* [17] evaluates the web applications for DNA N6-methyladenosine (6 mA) site prediction. Eleven 6 mA prediction web available tools are assessed on seven different species-specific datasets. Furthermore, advantages and disadvantages of the web-based 6 mA site prediction tools are discussed, and it shows that no universal best web tools are available for all genomes, especially for three species, namely, *Diospyros lotus*, *Caenorhabditis elegans* and *Escherichia coli*.

The development of data integration in omics needs to be fully advanced in experimental science, analytical science, biomedicine, mathematics, and computer science. In particular, more powerful analytical calculation methods are needed. The effective integration of multi-layered omics data by calculating sails can help us to comprehensively interpret biological systems. We hope that our Special Section can generate interest from researchers and trigger more valuable research results. Finally, we wish to thank all of the authors for their contributions, the scientific communities for peer reviewing, and the staff at the Briefings in Functional Genomics editorial office for their work on this Special Issue.

Conflict of interest

All authors declare no conflict of interest.

Feifei Cui, *Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 610054, China; Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou 324000, China*

Liang Cheng, *NHC and CAMS Key Laboratory of Molecular Probe and Targeted Theranostics, Harbin Medical University, Harbin, Heilongjiang, 150028, China; College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, Heilongjiang, 150081, China*

Quan Zou, *Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 610054, China; Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou 324000, China*

Corresponding authors: Quan Zou, Email: zouquan@nclab.net and Liang Cheng, Email: liangcheng@hrbmu.edu.cn

References

1. Subramanian I, Verma S, Kumar S, *et al*. Multi-omics data integration, interpretation, and its application. *Bioinfor Biol Insights* 2020;**14**:1177932219899051.
2. Huang S, Chaudhary K, Garmire LX. More is better: recent progress in multi-omics data integration methods. *Front Genet* 2017;**8**:84.
3. Conesa A, Beck S. Making multi-omics data accessible to researchers. *Scientific Data* 2019;**6**:1–4.
4. Perez-Riverol Y, Bai M, da Veiga Leprevost F, *et al*. Discovering and linking public omics data sets using the Omics Discovery Index. *Nat Biotechnol* 2017;**35**:406–9.
5. Yan J, Risacher SL, Shen L, *et al*. Network approaches to systems biology analysis of complex disease: integrative methods for multi-omics data. *Brief Bioinform* 2018;**19**:1370–81.
6. Bersanelli M, Mosca E, Remondini D, *et al*. Methods for the integration of multi-omics data: mathematical aspects. *BMC Bioinformatics* 2016;**17**:167–77.
7. Meng C, Zeleznik OA, Thallinger GG, *et al*. Dimension reduction techniques for the integrative analysis of multi-omics data. *Brief Bioinform* 2016;**17**:628–41.
8. Kaur P, Singh A, Chana I. Computational techniques and tools for omics data analysis: state-of-the-art, challenges, and future directions. *Arch Computat Method Eng* 2021;1–37. doi: 10.1007/s11831-021-09547-0.
9. Vasaikar SV, Straub P, Wang J, *et al*. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res* 2018;**46**:D956–63.
10. Augustyn DR, Wyciślik Ł, Mrozek D. Perspectives of using cloud computing in integrative analysis of multi-omics data. *Brief Funct Genomics* 2021. doi: 10.1093/bfgp/elab007.
11. Wang J. Integrative analyses of transcriptome data reveal the mechanisms of post-transcriptional regulation. *Brief Funct Genomics* 2021. doi: 10.1093/bfgp/elab004.
12. Kaur H, Kumar R, Lathwal A, *et al*. Computational resources for identification of cancer biomarkers from omics data. *Brief Funct Genomics* 2021. doi: 10.1093/bfgp/elab021.
13. Dong X, Liu C, Dozmorov M. Review of multi-omics data resources and integrative analysis for human brain disorders. *Brief Funct Genomics* 2021. doi: 10.1093/bfgp/elab024.
14. Wang Z, He Y. Precision omics data integration and analysis with interoperable ontologies and their application for COVID-19 research. *Brief Funct Genomics* 2021. doi: 10.1093/bfgp/elab029.
15. Chen Q, Li Y, Tan K, *et al*. Network-based methods for gene function prediction. *Brief Funct Genomics* 2021. doi: 10.1093/bfgp/elab006.
16. Ding Y, Lei X, Liao B, *et al*. Machine learning approaches for predicting biomolecule–disease associations. *Brief Funct Genomics* 2021. doi: 10.1093/bfgp/elab002.
17. Hasan MM, Shoombuatong W, Kurata H, *et al*. Critical evaluation of web-based DNA N6-methyladenine site prediction tools. *Brief Funct Genomics* 2021. doi: 10.1093/bfgp/elaa028.