



Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb
www.sciencedirect.com



PREFACE

Bioinformatics Commons: The Cornerstone of Life and Health Sciences



Zhang Zhang^{1,*}, Yu Xue^{2,*}, Fangqing Zhao^{3,*}

¹ BIG Data Center and CAS Key Laboratory of Genome Sciences & Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China

² Department of Bioinformatics & Systems Biology, MOE Key Laboratory of Molecular Biophysics, College of Life Science and Technology and the Collaborative Innovation Center for Biomedical Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

³ Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China

Received 21 August 2018; accepted 21 September 2018

Available online 28 September 2018

Bioinformatics, an interdisciplinary field that combines biology, mathematics, computer science, medicine, and health science, to integrate, analyze, and interpret biological data, is now becoming increasingly data-intensive. To dig out the treasure from big data powered by high-throughput sequencing technologies, it is highly dependent on *Bioinformatics Commons* that involves a variety of fundamental resources, including databases, web servers, algorithms, methods, software tools, ontologies, and standards. To be short, *Bioinformatics Commons* has become the cornerstone of life and health sciences, facilitating and accelerating the translation of big data into big discoveries.

Over years, a number of resources in *Bioinformatics Commons* have been developed and publicly released, with the aim to ease users to explore the relationship between genotype and phenotype and improve our understanding of genetic basis for human diseases. Hence, it is pivotal to raise the general awareness of the significance of *Bioinformatics Commons* in knowledge discovery. Toward this end, as a consequence, we

propose to set up the Special Issue on *Bioinformatics Commons* this year. In this issue, *Bioinformatics Commons* (I), we included 9 resources (8 databases and 1 web server; **Table 1**), covering a diversity of topics on DNA methylation, post-translational modifications, circular RNAs, genetic variant annotation, association studies, cancer, and other diseases, as well as plant defense responses. All databases in this issue have been added to Database Commons (<http://bigd.big.ac.cn/databasecommons/>), a catalog of global biological databases, in BIG Data Center [10]. Articles on algorithms, methods, and application notes of software tools will be included in *Bioinformatics Commons* (II) in the next issue.

Clearly, this Special Issue on *Bioinformatics Commons* exhibits several features. First, it is noted that extensive efforts, in coincidence with precision medicine, have been made related to human diseases through literature curation [3], potentially providing valuable clues into precision healthcare and treatment on a variety of diseases, especially autoimmune disease [4], esophageal squamous cell carcinoma [5], hepatocellular carcinoma [6], and collecting tumor-specific neoantigens as potential targets for cancer immunotherapy [7]. Second, several cutting-edge research topics are involved. For instance, CIRCpedia [1] features integration of comprehensive circular RNAs and their annotations, HeteroMeth [2] features single-cell DNA methylomes, and PlaD features transcriptomic characterization of plant defense responses to pathogens [8]. Third, advanced information technology has been applied in

* Corresponding authors.

E-mail: zhfq@biols.ac.cn (Zhao F), xueyu@mail.hust.edu.cn (Xue Y), zhangzhang@big.ac.cn (Zhang Z).

^a ORCID: 0000-0001-6603-5060.

^b ORCID: 0000-0002-9403-6869.

^c ORCID: 0000-0002-6216-1235.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

<https://doi.org/10.1016/j.gpb.2018.09.001>

1672-0229 © 2018 The Authors. Production and hosting by Elsevier B.V. on behalf of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Table 1 Resources in the Special Issue on Bioinformatics Commons

| Name | Short description | Link | Ref. |
|------------|--|---|------|
| CIRCpedia | A database of circular RNAs | http://www.picb.ac.cn/rnomics/circpedia | [1] |
| HeteroMeth | A database of single-cell DNA methylomes | http://qianlab.genetics.ac.cn/HeteroMeth | [2] |
| PTMD | A database of human disease-associated post-translational modifications | http://ptmd.biocuckoo.org | [3] |
| GAAD | An autoimmune disease association database | http://gaad.medgenius.info | [4] |
| CCGD-ESCC | A comprehensive database for genetic variants associated with esophageal squamous cell carcinoma in Chinese population | http://db.cbi.pku.edu.cn/ccgd/ESCCdb | [5] |
| HCCDB | A database of hepatocellular carcinoma expression atlas | http://lifeome.net/database/hccdb | [6] |
| TSNadb | A tumor-specific neoantigen database | http://biopharm.zju.edu.cn/tsnadb | [7] |
| PlaD | A transcriptomics database for plant defense responses to pathogens | http://systbio.cau.edu.cn/plad/index.php | [8] |
| DeepNitro | A web server for predicting protein nitration and nitrosylation sites | http://deepnitro.renlab.org | [9] |

Note: All databases listed here have been added to the Database Commons at <http://bigd.big.ac.cn/databasecommons/>.

bioinformatics, as exemplified by DeepNitro [9] that features adoption of deep learning technologies for prediction of protein nitration and nitrosylation sites. Taken together, *Bioinformatics Commons* provide valuable resources in aid of big data integration, translation, and interpretation, and accordingly are of broad utility in advancing life and health sciences. Thus, the Special Issue on *Bioinformatics Commons* is expected to be issued annually in *Genomics, Proteomics and Bioinformatics* and we look forward to your submissions for consideration of future publication in this Special Issue.

Competing interests

The authors have declared no competing interests.

Acknowledgments

We thank all persons involved in composing this Special Issue, especially the authors, the reviewers, and the handling editors who facilitated the completion of this work.

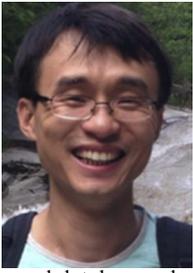
References

- [1] Dong R, Ma XK, Li GW, Yang L. CIRCpedia v2: an updated database for comprehensive circular RNA annotation and expression comparison. *Genomics Proteomics Bioinformatics* 2018;16:226–33.
- [2] Huan Q, Zhang Y, Wu S, Qian W. HeteroMeth: a database of cell-to-cell heterogeneity in DNA methylation. *Genomics Proteomics Bioinformatics* 2018;16:234–43.
- [3] Xu H, Wang Y, Lin S, Deng W, Peng D, Cui Q, et al. PTMD: a database of human disease-associated post-translational modifications. *Genomics Proteomics Bioinformatics* 2018;16:244–51.
- [4] Lu G, Hao X, Chen WH, Mu S. GAAD: a gene and autoimmune disease association database. *Genomics Proteomics Bioinformatics* 2018;16:252–61.
- [5] Peng L, Cheng S, Lin Y, Cui Q, Luo Y, Chu J, et al. CCGD-ESCC: a comprehensive database for genetic variants associated with esophageal squamous cell carcinoma in Chinese population. *Genomics Proteomics Bioinformatics* 2018;16:262–8.
- [6] Lian Q, Wang S, Zhang G, Wang D, Luo G, Tang J, et al. HCCDB: a database of hepatocellular carcinoma expression atlas. *Genomics Proteomics Bioinformatics* 2018;16:269–75.
- [7] Wu J, Zhao W, Zhou B, Su Z, Gu X, Zhou Z, et al. TSNadb: a database for tumor-specific neoantigens from immunogenomics data analysis. *Genomics Proteomics Bioinformatics* 2018;16:276–82.
- [8] Qi H, Jiang Z, Zhang K, Yang S, He F, Zhang Z. PlaD: a transcriptomics database for plant defense responses to pathogens, providing new insights into plant immune system. *Genomics Proteomics Bioinformatics* 2018;16:283–93.
- [9] Xie Y, Luo X, Li Y, Chen L, Ma W, Huang J, et al. DeepNitro: prediction of protein nitration and nitrosylation sites by deep learning. *Genomics Proteomics Bioinformatics* 2018;16:294–306.
- [10] BIG Data Center Members. Database resources of the BIG Data Center in 2018. *Nucleic Acids Res* 2018;46:D14–20.



Zhang Zhang, PhD. Professor and Executive Director of BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences (CAS). Dr. Zhang obtained his PhD degree in Computer Science from Institute of Computing Technology, CAS in 2007. He worked as Postdoctoral Associate at Yale University from 2007–2009 and Research Scientist at King Abdullah University of Science and Technology from 2009–2011. Dr.

Zhang was elected in the CAS 100-Talent Program in 2011 and given as Excellence Award in the final evaluation of the CAS 100-Talent Program in 2017. His research focuses on big data integration and mining, and computational precision health genomics. He acts as the Associate Editor-in-Chief for *Genomics, Proteomics & Bioinformatics* and Asian Associate Editor for *Briefings in Bioinformatics*.



Yu Xue, PhD. Professor of Huazhong University of Science and Technology (HUST), China. He obtained his PhD degree in cell biology from School of Life Sciences, University of Science and Technology of China in 2006. He was elected in the National Program for Support of Top-Notch Young Professionals in 2014, and Young Changjiang Scholars Program of China in 2017. His major interests are focused on the development of novel databases, algorithms, and computational software, such as the GPS (Group-based Prediction System) series algorithms, for understanding the temporally and spatially regulatory roles of post-translational modifications (PTMs) involved in cellular signaling pathways and networks. Dr Xue serves the editorial board of *Genomics, Proteomics & Bioinformatics*.



Fangqing Zhao, PhD. Professor of Beijing Institutes of Life Science (BIOLS), CAS. His research interests include developing novel computational approaches for exploring noncoding RNAs and human microbiomes. He obtained his PhD degree in marine genomics from Institute of Oceanology, CAS in 2006. Dr. Zhao conducted his postdoctoral research in the Department of Biochemistry and Molecular Biology at Pennsylvania State University from 2006 to 2010. Since 2011, he has been working as a Principal Investigator (100-Talent Plan) in the Computational Biology Center, Beijing Institutes of Life Science, CAS. He has published more than 50 research papers in high-profile journals including *Genome Research, Genome Biology, Gut, Nature Communications*, and *Nucleic Acids Research*. Dr Zhao is an editorial board member of *Genomics, Proteomics & Bioinformatics*.