

East Tennessee State University

Digital Commons @ East Tennessee State University

ETSU Faculty Works

Faculty Works

9-1-2017

Resources and Tools for Generating and Mining Big Data

Sushma Naithani

Oregon State University

Dhirendra Kumar

East Tennessee State University, kumard@etsu.edu

Follow this and additional works at: <https://dc.etsu.edu/etsu-works>

Citation Information

Naithani, Sushma; and Kumar, Dhirendra. 2017. Resources and Tools for Generating and Mining Big Data. *Current Plant Biology*. Vol.11-12 1. <https://doi.org/10.1016/j.cpb.2018.01.001>

This Editorial is brought to you for free and open access by the Faculty Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in ETSU Faculty Works by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact digilib@etsu.edu.

Resources and Tools for Generating and Mining Big Data

Copyright Statement

2018 Published by Elsevier B.V.

Creative Commons License



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



Editorial

Resources and Tools for Generating and Mining Big Data



With the advent of the next-generation sequencing technology and high-performance bio-computing, tremendous quantities of genomics data, the so-called ‘Big data’ from crops and model plants have accumulated. Genomic resources, analysis tools, and databases play the pivotal role in hosting, processing, annotating /curating and analyzing the ‘Big data’ and making it accessible, reusable and interoperable across various platforms in order to advance knowledge of plants at the systems level. While some of the processes and protocols for handling ‘Big Data’ are well established, others are emerging, and are being conceptualized as we progress. In this special issue of Current Plant Biology, we have compiled a set of articles focused on generation, processing, curation, and analysis of high-throughput genomic and transcriptome data.

Odell et al. pp 2–11 describe the principles and workflow applied to the curation of genomic data including genome sequences, genes, quantitative trait loci, genetic markers and associated scientific literature by curators of the GrainGenes database (<https://wheat.pw.usda.gov>), a data repository for wheat, barley, rye, oat, and other small grains funded by the USDA Agricultural Research Service. The article also highlights a necessity to establish formal guidelines based on the consensus of the community for indexing, formatting, accessibility, and interoperability of genomic data and a recent initiative in such direction using AgBioData (<https://www.agbiodata.org>) platform.

George et al. pp 12–22 discuss salinity- and drought-induced transcriptome data from a perennial C4 grass *Chrysopogon zizanioides* and how this analysis can bring forth the insights needed for making crop plants more resilient to abiotic stresses.

Goswami et al. pp 23–32 present physiological characterization and genotyping of 27 rice genotypes and one wild rice relative in the context of submergence tolerance traits and *Sub1* loci.

Basantani et al. pp 33–40 review various genomic methods used for sequencing, genome and transcriptome assemblers and various tools available for genomic data analysis.

Geniza and Jaiswal pp 41–45 provide a focused review on the de novo transcriptome assembly and need to deposit generated and analyzed data to public repositories to make it accessible.

He et al. pp 46–51 present a novel approach for transcriptomic data analysis for a set of genes that share homology (gene family members) or genes that participate in the similar biological process.

We hope that the collection of research and review articles compiled in this issue will generate fruitful conversations and insights among plant genomic community.

References

- [1] M.K. Basantani, D. Gupta, R. Mehrotra, S. Mehrotra, V. Vaish, A. Singh, An update on bioinformatics resources for plant genomics research, *Curr. Plant Biol.* 11–12 (2017) 33–40, <http://dx.doi.org/10.1016/j.cpb.2017.12.002>.
- [2] M. Geniza, P. Jaiswal, Tools for building de novo transcriptome assembly, *Curr. Plant Biol.* 11–12 (2017) 41–45, <http://dx.doi.org/10.1016/j.cpb.2017.12.004>.
- [3] S. George, D. Manoharan, J. Li, M. Britton, A. Parida, Drought and salt stress in *Chrysopogon zizanioides* leads to common and specific transcriptomic responses and may affect essential oil composition and benzyloquinoline alkaloids metabolism, *Curr. Plant Biol.* 11–12 (2017) 12–22, <http://dx.doi.org/10.1016/j.cpb.2017.12.001>.
- [4] S. Goswami, R. Kumar Kar, A. Paul, N. Dey, Genetic potentiality of indigenous rice genotypes from Eastern India with reference to submergence tolerance and deepwater traits, *Curr. Plant Biol.* 11–12 (2017) 23–32, <http://dx.doi.org/10.1016/j.cpb.2017.10.002>.
- [5] M. He, P. Liu, C.J. Lawrence-Dill, A hypothesis-driven approach to assessing significance of differences in RNA expression levels among specific groups of genes, *Curr. Plant Biol.* 11–12 (2017) 46–51, <http://dx.doi.org/10.1016/j.cpb.2017.12.003>.
- [6] S.G. Odell, G.R. Lazo, M.R. Woodhouse, D.L. Hane, T.Z. Sen, The art of curation at a biological database: principles and application, *Curr. Plant Biol.* 11–12 (2017) 2–11, <http://dx.doi.org/10.1016/j.cpb.2017.11.001>.

Sushma Naithani

Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR-97331-2902, USA

E-mail address: naitkans@science.oregonstate.edu

Dhirendra Kumar

Department of Biological Sciences, East Tennessee State University, Johnson City, TN 37614-1700, USA

E-mail address: kumard@etsu.edu