EDITORIAL

Biomarker Identification Based on Relationships of Protein-Protein Interactions or Solely on Feature Selection Procedures: A Matter of Selecting Descriptive Genes or Only High Performance Models

ELHAM AMJAD[1], SOLMAZ ASNAASHARI[2], BABAK SOKOUTI[3]
[1]Biotechnology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
[2]The first two authors (EA and SA) contributed equally to this research and should be regarded as joint first author.
[3]Correspondence to Dr. BabakSokouti, Email: b.sokouti@gmail.com; sokoutib@tbzmed.ac.ir, Tel: +98 41 33364038 Fax: +98 41 33370420

From many years ago, the biology and the mathematics terms and their contents were known to be completely separate from each other due to the communication ways established between biological and mathematical scientists.To fill the existing gaps between the two abovementioned terms one may think about possible keys in connecting them in real world to achieve correct justifications. To illustrate how complex the position is, it is worth to epitomize the famous systems biology example[1]. In this example, a number of people are sent to a dark room in which they are supposed to recognize an object (i.e., elephant) by touching only one part. After, they have been asked for theirrecognized object; however, none of them were correct so long as they accumulated their knowledge about the object. By simply demonstrating the elephant example, it can be deduced that these people can be regarded as biologists and mathematicians who work separately. And, once they work as a team a life-related problem can be solved both biologically and mathematically.

In this editorial, the purpose is focused on the early diagnosis, prognosis and hence treatment of a particular disease which is the aim of almost all clinical types of studies. So, clinicians and pathologists, biologists and geneticists, as well as the biomedical engineers and bioinformaticians try to investigate these topics in their particular point of views. In this regard, several genome-wide association studies (GWAS) for biomarker identification have been frequently scrutinized based on experimental wet and dry labs in order to achieve the personalized medicine treatments[2-5]. Most of the biomarker identification approaches available in the literature were carried out based on laboratory studies coupled with statistical analysesas well as the machine learning techniques developed for raw genome datasets. However, none of these tactics might provide a comprehensive insight on the identified gene biomarkers without understanding their effective interactions. Recently, BioDiscML tool implemented in java was proposed to automatically select the gene biomarkers using machine learning methods available in WEKA[6]. Although, the tool is a feature selection methodology customized for high-dimensional OMICs data and has outperformed other preceded performers, however, providing interpretable gene signatures might still be a challenge for biologists even by retrieval of the highest correlated genes with the current identified ones.

To ascertain the fact that the identified biomarkers could propose at least the required descriptive gene features and their importance in development and progression of a disease, the systems biology and systems biomedicine would be beneficial in understanding how living systems function and affect each other.

REFERENCE

1. Voit E. A First Course in Systems Biology: Garland Science; 2018. 960 p.
2. Hsu H-M, Chu C-M, Chang Y-J, Yu J-C, Chen C-T, Jian C-E, et al. Six novel immunoglobulin genes as biomarkers for better prognosis in triple-negative breast cancer by gene co-expression network analysis. Scientific reports. 2019;9(1):4484.
3. Maind A, Raut S. Identifying condition specific key genes from basal-like breast cancer gene expression data. Computational biology and chemistry. 2019;78:367-74.
4. Rahman M, Islam T, Gov E, Turanli B, Gulfidan G, Shahjaman M, et al. Identification of prognostic biomarker signatures and candidate drugs in colorectal cancer: Insights from systems biology analysis. Medicina. 2019;55(1):20.
5. Westbrook JA, Wood SL, Cairns DA, McMahon K, Gahlaut R, Thygesen H, et al. Identification and validation of DOCK4 as a potential biomarker for risk of bone metastasis development in patients with early breast cancer. The Journal of pathology. 2019;247(3):381-91.
6. Leclercq M, Vittrant B, Martin-Magniette M-L, Scott-Boyer M-P, Perin O, Bergeron A, et al. Large-scale automatic feature selection for biomarker discovery in high-dimensional OMICs data. Frontiers in genetics. 2019;10:452