Drug Designing: Open Access
ISSN: 2169-0138

Sardari, Drug Design 2012, 1:1
DOI: 10.4172/2169-0138.1000e102

Editorial Open Access

Data Mining Methods for 'OMICs' Applications in Anticancer Drug Design and Discovery

Soroush Sardari*

Drug Design and Bioinformatics Unit, Medical Biotechnology Department, Biotechnology Research Center, Pasteur Institute, Iran

Cancer forces a serious burden on the public health system and has created a challenge to the medical science researchers. Though the century-long drift of cancer mortality in the world was reversed in the middle of 19th century and now cancer remains the second leading cause of death. Annually, more than 10 million new case of cancer are diagnosed based on the World Health Organization (WHO) report [1]. By 2020, the world population is expected to have risen to 7.5 billion; of this number, around 15 million new cancer cases will be diagnosed, and 12 million cancer patients will die. The pointed undesirable data demonstrate that cancer is described as a serious challenge in human healthcare and survival. Although we have witnessed the development of many drugs against cancer, the death rate for the most prevalent cancers has not decreased [2]. The design of novel drugs to treat cancer is an extended and hard process with a very high level of abrasion. Many steps in this lengthy procedure use data generated from various molecular studies and chemical species. One key challenge is to successfully translate the basic findings of target validation further into safety studies in clinical trial stage.

Advanced computational evolutionary analysis methods combined with the increasing accessibility of sequence data enable the application of systematic evolutionary approaches to targets and pathways of attention to drug discovery. Data mining methodology as one of cheminformatics tools is applicable in drug discovery process to analyze related data from many different sources, classifying and summarizing the relationships identified. The high throughput data collectively referred to as 'OMICs' data are ubiquitous during the drug design and discovery process from target identification and validation to the development and testing of novel anticancer drug candidates to solving cancer treatment challenges [3].

Presently large quantities of genomics, proteomics, metabolomics and pharmacogenomics data are being generated both in academia and industry. Gene expression altered in response to a drug or a toxin is usually measures by microarrays. Changes in global gene expression patterns in animals or cells at multiple dose levels and time-points produce related "signature" genes that can be used as predictors or biomarkers in human [4]. There is direct correlation between the evaluation of the compound effect on protein activity/its concentration level and the overall mechanism of its toxicity and distantly with gene expression data. Proteomics involves total quantitative and qualitative rate of protein concentration/expression in whole-tissue samples. This is significant because presence of a mature mRNA transcript is insufficient for having a corresponding active protein due to post-translational modifications, proteolysis and other dynamic processes causing functional changes. A comprehensive analysis of biological systems needs the integration of all biological data that is generated to discover molecular biomarkers [4]. The relation of experimental data and large amounts of literature data on transporters, enzymes, channels and receptors that bind small molecules may need to be interpreted as a network of interactions enabled by expansions in databases and data annotation to eventually reflect the response of the whole system, as well as provide insight into the functional organization of the cell [4]. In combination with whole-genome sequences of human and numerous model organisms, the availability of technologies to change and measure cellular responses at the level of personal transcripts and proteins presents opportunities to accelerate the process of drug discovery across the entire channel, from disease understanding and target identification through clinical trials, postmarketing surveillance and diagnostics [5]. To progress analysis of genetic sequences for quick evaluation of RNA/protein amounts in related tissues for validating novel drug targets, multiple new techniques have been improved in recent times. Such new techniques have to be integrated in order to reach the maximum usefulness in pharmaceutical research and can reach the discovery and development pipeline.

These important recent technical advances are not without limitations and challenges. The generated data are technically and statistically complex; therefore, computational approaches as bio/cheminformatics methods have been developed and adapted to facilitate the processing and analysis of large amounts of generated data. Data mining methodology as one of the most applicable cheminformatics tools have appeared to define associations in many types of databases [5]. These methods have been employed to classes, clusters, associates and patterns of raw achieved data resulted to design and discover effective target with high chance of success in clinical trial stages. Data mining methods applied in drug discovery process generally include artificial neural networks, Bayesian probability approaches, genetic algorithms, decision trees, nearest neighbor methods, rule induction, new data visualization and virtual screening techniques. Database mining clearly has increased the number of putative targets [6,7]. Two applied data mining methods involve database compilation and data visualization employed to class, cluster, associate and model raw data generated from genomics and postgenomics studies for achieving novel anticancer drug like compounds [8]. In fact, data mining is the procedure of analyzing data from different perspectives and summarizing it into practical knowledge.

Recently, the field of research in drug discovery is focused on applying data mining approaches to design and discover effective molecules to affect cancer targets with the aim at success in advanced stages to have better chance in clinical trial stages [5]. In this regard, as the recent investigations, a data mining novel approach in combining cheminformatics, intensive data handling were used together with correlation of biologic data to search for the desired biologic activity in the domain of

***Corresponding author:** Soroush Sardari, Drug Design and Bioinformatics Unit, Medical Biotechnology Department, Biotechnology Research Center, Pasteur Institute, #69 Pasteur Ave., Tehran 13164, Iran, Tel: +98-21-66405535; Fax: +98-21-66465132; E-mail: ssardari@hotmail.com

Received March 28, 2012; **Accepted** March 29, 2012; **Published** March 30, 2012

Citation: Sardari S (2012) Data Mining Methods for 'OMICs' Applications in Anticancer Drug Design and Discovery. Drug Design 1:e102. doi:10.4172/2169-0138.1000e102

natural products that were not explored before [9]. It would be obvious that the *in silico* drug discovery based on data mining approaches play great roles in cancer treatment and are gaining increased applications in drug discovery and development with saving several folds of cost and time parallel to increasing success rate to present novel efficient anticancer drugs through minimum side effects in clinical trial.

As the final indictable aspect, there would be no doubt that one of the essential parts for constricting and designing a fruitful research project based on *in silico* approaches as well as other fields could be defined as complete and accurate literature search that open access scientific resource play critical role in this regard. The open access movement has begun in the first half of the 1990s [10]. Open access to scientific knowledge is vital for several reasons, such as: I. Research is an interdependent procedure that is notified by the earlier works of others. Easy and open access to scholarly and research output let making of extra truthful and progressive research outcomes, providing for scholarly and intellectual progression; II. Broader access to information assists quicker scientific progress, which benefits the community at great, particularly in the area of medicine and drug discovery; III. The global sharing of knowledge and learning supports social unity and cultural development; and IV. The general belief that it is vital for all human beings around the world should have free and equal access to information [10].

The traditional techniques of sharing information, primarily through conventional print publishing, while still applicable, are no longer correctly adapted to the wider distribution of knowledge. Electronic publishing and the internet present the opportunity to exchange information worldwide, instantly and efficiently. OMICS Publishing Group as one of the mentioned Open Access publication models disseminates research articles globally in the field of 'OMICs' with a wide range of critical applications utilized for several scientific procedures of drug discovery and design process.

References

1. Vicker N, Sharland CM, Heaton WB, Gonzalez AM, Bailey HV, et al. (2009) The design of novel 17beta-hydroxysteroid dehydrogenase type 3 inhibitors. Mol Cell Endocrinol 301: 259-265.
2. Zhang X, Li L, Wei D, Yap Y, Chen F (2007) Moving cancer diagnostics from bench to bedside. Trends Biotechnol 25: 166-173.
3. Nikolsky Y, Nikolskaya T, Bugrim A (2005) Biological networks and analysis of experimental data in drug discovery. Drug Discov Today 10: 653-662.
4. Clarke PA, te Poele R, Workman P (2004) Gene expression microarray technologies in the development of new therapeutic agents. Eur J Cancer 40: 2560-2591.
5. Ekins S, Shimada J, Chang C (2006) Application of data mining approaches to drug delivery. Adv Drug Deliv Rev 58: 1409-1430.
6. Cockett M, Dracopoli N, Sigal E (2000) Applied genomics: integration of the technology within pharmaceutical research and development. Curr Opin Biotechnol 11: 602-609.
7. Lutz WM, Warren VP, Gill WR, Searls BD (2005) Managing genomic and proteomic Knowledge. Drug Discov Today 2: 197-204.
8. Cragg GM, Newman DJ, Weiss RB (1997) Coral reefs, forests, and thermal vents: the worldwide exploration of nature for novel antitumor agents. Semin Oncol 24: 156-163.
9. Sardari S, Shokrgozar MA, Ghavami G (2009) Cheminformatics based selection and cytotoxic effects of herbal extracts. Toxicol In Vitro 23: 1412-1421.
10. Pappalardo KM, Fitzgerald AM, Fitzgerald BF, Kiel-Chisholm SD, O'Brien D et al. (2007) A Guide to Developing Open Access Through Your Digital Repository. DEST funded Open Access to Knowledge Law Project OAK Law Project.