Experimental and Computational Approaches in Leveraging Natural Compounds for Network based Anti-cancer Medicine

Bisen PS[*]

School of Studies in Biotechnology, Jiwaji University, Gwalior, India

[*]**Corresponding author**: Prakash S Bisen, School of Studies in Biotechnology, Jiwaji University, Gwalior, India, Tel: +917512462500; E-mail: psbisen@gmail.com

Editorial

Cancer is caused by the complex interplay between various non-genetic (carcinogens, tobacco, chemicals, radiations and infectious organisms), and genetic factors (inherited mutations, hormones, immune conditions and mutations that occur from metabolism). These causal factors may act together, or in sequence to initiate or promote carcinogenesis. At a molecular level, it starts with the activation of oncogenes in the cells leading to subsequent inactivation of tumor suppressor genes [1]. The genes involved in the development of cancer can be grouped under two categories, viz. (i) Oncogenes and (ii) Tumor suppressor genes (TSGs). Oncogenes are responsible for transforming normal cells into their cancerous counterparts. Oncogenes are mutated form of otherwise normal genes known as proto-oncogene. These proto-oncogenes carry out critical functions like cell cycle regulation, and differentiation [2]. The development of cancer is a multistep process enabled by the occurrence of key hallmark events like sustaining proliferative signaling, evading growth suppressors, resisting apoptotic cell death, enabling replicative immortality, inducing angiogenesis, activating invasion, metastasis and inflammation [3]. Targeting a gene involved in multiple hallmark events could be an effective strategy to control cancer. Mutations in proto-oncogenes compromise with critical functions, and thereby lead to uninterrupted growth promoting activities necessary for tumor formation. Conversely, tumor suppressor genes are responsible in maintaining the normal cellular homeostasis by regulating cell division, repair errors in DNA, and inducing apoptosis in the case of accumulation of abnormal signals [4-6]. The development of cancer is facilitated by promotion of oncogenes and silencing of tumor suppressor genes by various mechanisms like mutation, epigenetic silencing etc. The exact sequence of this genetic insult is still a mystery, even after decades of cancer research [2,7].

Nature is an attractive source of therapeutic candidate compounds because of its tremendous chemical diversity, which is found in millions of species of plants, animals, marine organisms, and microorganisms [8-12]. Despite major scientific and technological progress in combinatorial chemistry, drugs derived from natural products still make an enormous contribution to drug discovery today. Drugs derived from natural sources are preferred over synthetic chemotherapeutic compounds because of their efficacy and minimal side effects. The treatment with existing chemotherapeutic agents is often associated with myriad kind's toxicities, because of the indiscriminate nature of compounds to kill cancer as well as normal cells [4,13]. The advent of targeted therapies for the treatment of various cancers can be regarded as an attractive alternative to existing treatment options. Therapeutic action of a drug is mediated through modulation of molecular target(s), involved in the process of genesis and progression of disease. Mechanistic understanding of drug compound helps in medication in various ways; it helps in explaining drug efficacy, adverse drug reaction (ADR), and synthesis of better drug. Natural compounds can be screened and ranked based on its target profile to explain mechanism of drug action of natural compounds predicted/reported to be effective for cancer therapy and propose theoretical framework for screening compounds with desired activity [9,14].

The cancer biology has become exciting field of research, especially after the discovery of DNA, and also, other advances made in the field of high throughput genomics and proteomics technology [1,2,6]. These technological advances have enabled us to simultaneously study expression pattern of entire genome/proteome. Such technologies generate huge amount of data, which is often noisy, therefore, needs bioinformatics and computational methods to manage data, and extract biologically relevant information [9,14]. The potential therapeutic compounds for cancer can be identified through a large-scale mining of bioactive compounds from publicly available databases like NCBI-PubChem, ChEMBL. Initial set of compounds with anti-cancer activity are normally identified by custom build support vector machine (SVM) classifier leading to build and predict task derived from functional groups associated with the compound. Active protein bioassays are normally used to associate molecular targets with compound. Compounds with known indications are collected from DrugBank and published papers; and these compounds are used to train partial least squares regression model [9,14]. The weights associated with the targets are extracted from partial least square regression modelling, and these weights are used to compute a custom score for the input compound. The data produced by such high throughput technologies can be analyzed with the help of novel computational/bioinformatics methods, to generate list of candidate molecular targets to exploit for the development of therapeutic and/or diagnostic applications. The potential therapeutic targets are then taken forward for designing and/or selection of effective therapies for various cancers. The potential compounds are identified on the basis of cancer specific custom score, and the other annotations related to its physicochemical properties and behavior in cancer specific bioassays studies [9,14].

The study of disease at the molecular level equips us with host of targets which are significantly altered in diseased state when compared with normal condition, such targets can be used for identification (or diagnosis) and/or treatment of the disease. The knowledge about therapeutic targets for a disease of interest can be found to be spread across the huge corpus of publications. There have been informatics approaches like creation of highly specific databases for easy access of targets associated with the diseases; the therapeutic target database (TTD) is one such dedicated database which comprehensively stores target information of several diseases including various cancer types [15]. Literature databases like NCBI-PubMed is still the gold standard for finding information about targets involved in less researched

cancers, however, extracting information out of enormous database like PubMed is like searching a needle in a haystack. The high-throughput genomics techniques like microarrays, Next-Gen Sequencing (NGS) are efficient profiling technologies, which can be used to generate a list of molecular targets associated with a disease process [7]. The NCBI-Gene Expression Omnibus (GEO) is a publicly available repository in which gene expression studies related with various pathologies and conditions are available. The study dataset can be of various experimental designs like the paired and unpaired - sample design. The unpaired case/control sample design is the simplest, to which most of the statistical methods can be easily applied, since the methods do not have to account of dependency between case/control samples, as in the case of paired sample design. In order to generate biologically meaningful hypothesis, the statistical power of underlying study should be good. The results obtained from microarray studies does not usually translate into robust applications in clinical settings, which is more than often due to shortcomings in a dataset, which has more variables (or genes) when compared to samples used for profiling transcriptome. Gene expression profiling through microarrays is regarded as promising tool to understand transcriptional changes linked with cancer formation [7].

The inclusion of a large number of samples representing different pathologies or conditions of interest, is one of the solution to improve the statistical power required to generate meaningful hypothesis on which clinical application can be reliably built, however, because of financial and/or logistics constraints it is not always possible to procure a large number of samples required for gene expression studies. The meta-analysis is one such power technology which can be used for improving statistical power by combining pathologically similar samples from different gene expression studies. The meta-analysis combines pathologically similar samples from different studies, and applies statistical tests attempting to extract the gene expression profiles which are consistent across different studies and arrange the genes based on cumulative statistics, which is p value or rank.

The meta-analysis techniques are the perfect solution of biological problems, like finding differential expression patterns observed among different experiments with similar experimental design. However, it cannot be applied in scenarios where you want to use expression values for the downstream analysis. The analytical methods for direct integration of gene expression dataset from two or more studies can be used, when you want to improve the statistical power by integration of dataset, and yet, want to have gene expression values for downstream statistical analysis. The direct integration of datasets from different studies is challenging due to the existence of myriad sources of non-biological variations, often referred as 'batch-effects'. Such probe-level integration of dataset from two different studies is possible by removing batch-effects with the help of cross-platform normalization methods [16]. The batch corrected and merged dataset can be used for the identification of genes which are differentially expressed compared to controls. Differential expression is definitely an important criterion for therapeutic target; however, certain other conditions must be met by a gene to qualify as a potential therapeutic target. Dependency network can be generated based on gene expression correlation among differentially expressed genes. The genes which undergo marked change between diseased and control conditions are good candidate for therapeutic intervention. Two genes can be also connected based on causal relationship (activate/inhibit) between them. Causal reasoning analysis attempts to identify the biologically meaningful hypothesis.

In my view natural compounds have promising target profile with targets dispersed across different hallmark events, thus they can be regarded as ideal candidate of drug development. The physicochemical properties of most of the natural compounds are optimal for clinical use; however, there are certain natural compounds have low bioavailability due to non-optimal physicochemical properties. The poor bioavailability can be regarded as one of the main roadblocks towards realization of natural compounds based therapy. The problem of poor bioavailability can be addressed with the help of various structural optimization methods, and/or with the help of novel targeted drug delivery methods with path-breaking technological advancements to overcome this challenge The availability of therapeutic targets is most important pre-condition for rational drug discovery and equally the effective treatment options are highly desirable. Mesmerized with great biodiversity in nature, I conclude this write-up with a sense of optimism that our scientific endeavors would make world better place for generations to come to control life style diseases such as cancer, obesity, CVD and diabetes etc.

References

1. Bisen PS, Khan Z, Bundela S (2013) Biology of Oral Cancer and Key Apoptotic Regulators, CRC Pres.
2. Khan Z, Bisen PS (2013) Oncoapoptotic signaling and deregulated target genes in cancers: special reference to oral cancer. Biochim Biophy Acta Reviews on Cancer 1836: 123-145.
3. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144: 646-674.
4. Khan Z, Khan N, Tiwari RP, Patro IK, Prasad GBKS, et al. (2010) Down-regulation of survivin by oxaliplatin diminishes radioresistance of head and neck squamous carcinoma cells. Radiotherapy and Oncology 96: 267-273.
5. Khan Z, Khan N, Varma AK, Tiwari RP, Mouhamad S, et al. (2010) Oxaliplatin-mediated inhibition of survivin increase sensitivity of head and neck squamous cell carcinoma cell lines to paclitaxel. Current Cancer Drug Target 10: 660-669.
6. Khan Z, Tiwari RP, Khan N, Prasad GBKS, Bisen PS (2012) Induction of Apoptosis and Sensitization of Head and Neck Squamous Carcinoma Cells to Cisplatin by Targeting Survivin Gene Expression. Current Gene Therapy 12: 444-453.
7. Mousumi D, Prasad GBKS, Bisen PS (2010) Molecular Diagnostics: Promises and Possibilities. Springer Verlag.
8. Bisen PS, Bundela S, Sharma A (2012) Ellagic Acid- Chemoprotective role in oral cancer. Journal of Cancer Science and Therapy 4: 023-030.
9. Bundela S, Sharma A, Bisen PS (2015) Potential Compounds for Oral Cancer Treatment: resveratrol, nimbolide, lovastatin, bortezomib, vorinostat, berberine, pterostilbene, deguelin, andrographolide, and colchicines. PLoS One 10: e0141719.
10. http://www.omicsonline.org/marine-microbes-unexplored-therapeutic-bio-mine-1948-5956.1000e112.php?aid=7690
11. Bisen PS (2013) Mushroom: A panacea for cancer. J Cancer Sci Ther 5: e125.
12. Bisen PS (2016) Therapeutic and Preventive Potential of Functional Food and Cancer. Cancer Med Anticancer Drug 1: e101.
13. Khan Z, Khan AA, Prasad GBKS, Khan Noor, Tiwari RP, et al. (2016) Growth inhibition and and chemo-radiosensitization of head and neck squamous cell carcinoma (HNSCC) by Survivin-siRNA lentivirus treatment. Radiotherapy and Oncology 118: 359-368.
14. Chen X, Ji ZL, Chen YZ (2002) TTD: Therapeutic Target Database. Nucleic Acids Res 30: 412–415.
15. Rudy J, Valafar F (2011) Empirical comparison of cross-platform normalization methods for gene expression data. BMC Bioinformatics 12: 467.

16. Yildirim MA, Goh KI, Cusick ME, Barabási AL, Vidal M (2007) Drug-target network. Nat Biotechnol 25: 1119–1126.