MULTI-OMICS TECHNOLOGY BASED BIOMARKERS

Abstract

Multi-omics techniques, which assimilate data from genomics (DNA), transcriptomics (RNA), proteomics (proteins), and metabolomics (metabolites) are useful in investigation of oncogenesis pathways. High-throughput omics technologies have accelerated the identification of numerous potential biomarkers in recent years. An overview of the current state in this domain is offered with examples of genomics, proteomics, transcriptomics, metabolomics and microbiomics biomarkers in the field of oncology, along with some proposed ways to accelerate their validation. The utilization of multi-omics data has enhanced our understanding of the disease and enabled the identification of valuable biomarkers. These molecular signatures are helpful for elucidating the development and progression of cancer and any disease. In this chapter, efforts are made to emphasize potential applications of multi-omics for finding novel biomarkers and enhancing clinical evaluation.

Keywords: Biomarker, multi-omics, women health, cancer, cancer.

Authors

Pedamallu Niharika
Department of Pharmaceutical Chemistry
SIMS College of Pharmacy
Guntur, Andhra Pradesh, India
niharika.mpharm12@gmail.com

S. Manohar Babu
Department of Pharmacology
SIMS College of Pharmacy
Guntur, Andhra Pradesh, India
manoharsitty@yahoo.co.in

A. Mercy
Department of Doctor of Pharmacy
SIMS College of Pharmacy
Guntur, Andhra Pradesh, India
mercy.avutapalli888@gmail.com

Akula Prabhu Sangeetha
Department of Doctor of Pharmacy
SIMS College of Pharmacy
Guntur, Andhra Pradesh, India
aps.sangeetha222@gmail.com

Satya Narayan Tripathy
Dept. of Pharmaceutical Analysis
School of Pharmacy & Life Sciences
Centurion University of Technology and Management
mrsntripathy@gmail.com

I. INTRODUCTION

Diseases are triggered by shift patterns in the regulation of genes or by the combined effects of multiple genes and the environment on a particular organ or tissue. These diseases make their presence known through significant alterations in human physiology, which provides the foundation for clinical chemistry and enables it to contribute significantly to the diagnostic process and subsequent therapeutic interventions. DNA, RNA, proteins, lipids, and metabolites are all part of the complex network of molecules that are involved in the biological mechanism of disease. The term "biomarker" refers to "a parameter that is objectively measured and evaluated as an indicator of normal biological processes, pathological processes, or pharmacological responses to a therapeutic intervention." In general, an ideal biomarker should be able to detect a fundamental feature of a specific disease; should be precise, validate, confirm and differentiate those disease cases from other similar disease cases or family members of that disease. An ideal biomarker is simple to perform; reliable; non-invasive; and inexpensive, if at all possible. [1].

The omics methodology is a comprehensive examination of the roles that these molecules play in the function of living organisms. Different omic methods accessible include genomics, epigenomics, transcriptomics, proteomics, metabolomics, lipidomics and immunomics. Genomics refers to the study of genetic variants in order to investigate association with diseases, response to treatments, and prognosis. Epigenomics, on the other hand, is the study of the reversible change of DNA and the proteins that are connected with it. [2,3] The field of epigenomics investigates the role of interactions between genes and the environment in the onset, progression, prevention, and treatment of disease. The study of RNA transcripts is known as transcriptomics, and it is used to identify biological pathways, track changes that occur during the progression of disease, and distinguish healthy individuals from diseased ones. Figure 2 gives an overview of the processes that are involved in genomics and transcriptomics. As shown in figure 3, proteomics is both the scientific study and the technological process of isolating and identifying proteins from basic biological samples. The study of proteomics defines the biological and functional role that proteins play, reveals their biological mechanisms and their use as biomarkers and drug targets, and measures the proteins that are involved in both normal and pathological states in order to gain an understanding of the disease process. The study of metabolomics elucidates the low molecular weight proteins that are unique to cells, tissues, and biological fluids, as well as the profile of metabolic pathways that are active in various systems of an organism and to explore biomarkers of drug addiction.[4] The field of immunomics focuses on the comprehensive study of the genes and proteins that make up the immune system. This study contributes to a better understanding of the dynamic link that exists between cells and the web of responses that occur in response to a stimulus or a disease. [5]

A single omic technique on its own cannot completely reveal the intricate details of a disease. As shown in **Figure 1,** an integrative investigation of various omics methodologies, often known as a multi-omics strategy, can give comprehensive information on molecular function, the genesis of disease, diagnostic biomarkers, and potential therapeutic targets for drug discovery. Integrative analysis of multi-omics data paved a way for the discovery and functional studies of complex human disorders. [6,7]

II. BIOMARKER DISCOVERY

The search for one-of-a-kind disease biomarkers has become increasingly important in light of the emergence of diseases that are resistant to detection and treatment. This results in an improvement in the disease's diagnosis, prognosis, as well as the medication discovery process and therapy methods. The search can be carried out in one of two ways: a method based on hypothesis testing or an approach based on discovery, both of which are detailed in table 1. A biomarker can be derived from a mechanistic knowledge of a disease using the hypothesis-based search method, whereas the discovery-based method focuses on the changes that take place in molecular species throughout the illness state. The found biomarkers can be placed into one of three categories: risk, diagnostic, or prognostic biomarkers. When it comes to the early diagnosis of a disease, a panel of molecules, as opposed to a single biomarker, provides greater sensitivity as well as specificity. [8]

Identification of new biomarkers attracts great importance in the areas of women's health and rare disease. Socio-cultural discrimination, poverty, differences in the effect of diseases on women's health and unique health issues necessitate the need for special education, diagnosis, care and research. Annual death rate among women globally is 55.4 million and the 2nd leading cause of death is cancer. Even though the death rate declined, most of the deaths in cancer occur due to improper therapeutic efficacy of drug arise from genetic polymorphism, lack of early diagnosis, lack of specific treatment and invasive therapies. Most common cancers found among women include lung, breast, gynecological and colon cancer. Gynecological cancers are the deadly cancers and multi-omics approach provide platform for identification of specific biomarkers. About 43250, 12810, 4280 and 12550 women in US die annually because of breast, ovarian, cervical and endometrial cancer respectively [9,10,11].

An ailment or illness is considered to be rare when it only affects five people out of every 10000 people in Europe or when it affects less than 200,000 people in the United States. The clinical heterogenicity of the condition, the failure of clinical trials, and the restricted number of therapeutic alternatives are the major problems associated with the development of drugs for rare diseases. There are 6000 to 8000 uncommon diseases known to exist in the globe at this time, yet only 5% of these have therapies that are considered acceptable. Because of this, it is necessary to speed up the process of identifying biomarkers and developing drugs. The application of genomic, epigenomic, and transcriptomic approaches has resulted in an improvement in diagnostic accuracy and contributed to the development of new medicines.

Below are list of the biomarkers identified through integration of omic approaches.

III. BIOMARKERS IDENTIFIED THEOUGH MULTI-OMICS APPROACH

The most common type of cancer that affects women's gynecological systems is ovarian cancer. Ovarian cancer is known as a "silent killer" since its early signs is not easily recognizable, it is often diagnosed in its advanced stages, and it has a high rate of recurrence. cAMP response element binding protein, also known as CREB, is a transcription factor that activates the transcription of genes necessary for embryogenesis and also plays a role in the process of malignant transformation of cells. CREB binds to the DNA sequence that can be found at cAMP response elements. Because CREB is overexpressed in many tumour cells,

including those of non-small cell lung cancer, glioblastoma, leukaemia, breast cancer, and melanoma, CREB1 has emerged as a promising candidate for use as a biomarker for a variety of cancers. Ovarian cancer cells have been shown to have high levels of CREB1 expression, which is linked to a number of negative outcomes associated with the disease, including a poorer prognosis and a lower chance of survival [12].

An extremely rare form of epithelial ovarian cancer, low-grade serous ovarian carcinoma (LGSOC) has a median survival time of just 10 years from the time of diagnosis. Through a multi-omic and data-integration strategy, a correlation between mutations in the gene for mitogen-associated protein kinase (MAPK) and LGSOC was discovered, introducing MEK as a therapeutic target and MEK inhibitor medications for efficient therapy [13].

Cervical cancer is once the deadliest cancer. Second primary malignancies are recurrence of cancer which is a serious complication and results in death of patients. Such second primary malignancies are seen in cervical cancer. Synaptopodin-2 is one such prognostic factor for cervical cancer, hepatocellular carcinoma, breast cancer, melanoma and bladder cancer. Synaptopodin-2 expression regulates PI3K/AKT pathway, epithelial-mesenchymal transition, initiation of estrogen receptors, DNA repair and cell cycle. Low expression of Synaptopodin-2 phosphorylates PI3K/ AKT and activates epithelial-mesenchymal transition pathway which leads to gain in invasive capability and malignant tumor progression[14].

Exosomes are double-membrane vesicles of 30-150nm that are released by immunological, cancer, and mesenchymal stem cells and separated from plasma, serum, breast milk, saliva, urine, and ascites. They include proteins, lipids, viral particles, and RNAs (messenger, micro, long non-coding, circular). Exosomes help cells communicate with one another and share resources and information. They control cellular processes in both normal and abnormal conditions. The unique molecular content, components, and behaviour of exosomes generated by cancer cells as compared to healthy cells position them as promising potential biomarkers for gynecologic cancer risk assessment and detection. Cancer, angiogenesis, metastasis, immunological regulation, and treatment resistance can all flourish in an exosome-friendly environment. Nucleic acids and proteins carried by exosomes of different origins act as prospective biomarkers for gynecologic malignancies. Some of the nucleic acid biomarkers from the exosome include Let-7 family and miR-221-3p from the cervical cancer cells promotes proliferation, invasion and angiogenesis. [15].

Endometrial cancer mainly effects post-menopausal women is usually found early but delayed diagnosis make it difficult to study and treat the cancer [16]. Members of the Minichromosome Maintainance (MCM) family are a collection of genes that regulate early embryogenesis, cell division, as well as genome integrity and participation in homologous recombination repair. Using a multi-omic approach, specifically epigenomics and transcriptomics, researchers were able to determine that MCM plays a role as an oncogene and contributes to the proliferation of tumor cells. Because of this, MCMs serve as a biomarker for cancer. MCM attaches to the enzymes that are responsible for replicating DNA and dividing cells, which speeds up the spread of cancer. The upregulation of MCM2, MCM3, and MCM 7 mRNAs in endometrial cancer is helpful in diagnosis, while MCM 7 is helpful for pathological staging, prognosis, and as a therapeutic target [17].

Uterine corpus endometrial carcinoma is the second most common gynecological cancer and the sixth most common cause of cancer among women around the globe. An epigenetic RNA modification plays an important role in development of cancer. N6-methyl adenosine (m6A) is the most common epigenetic RNA modification. m6A regulators such as METTL3, LEF1, METTL14, YTHDF3, ALCBH5, METTL16, FTO and ALKBH5 were dysregulated in some cancers such as prostate cancer, melanoma, osteosarcoma, colorectal cancer, gastric cancer and plays a key role in disease progression and metastasis. Mutations in m6A regulator genes are also associated with prognosis of uterine corpus endometrial carcinoma and hence can be used as prognostic biomarker [18].

Differential methylation was identified as a possible marker of retinitis pigmentosa, Fuchs endothelial corneal dystrophy, IgA nephropathy, and polycystic kidney disease by multi-omic investigations [19].

IV. DISCUSSION

Biomarker, a molecule revealing the risk, status of a disease plays a vital role in its diagnosis, treatment and prevention. Identification of such molecules is tedious, expensive and complicated process. Overcoming all these obstacles, omics have generated comprehensive data on most of the bio molecules such as genes, mRNA, proteins and metabolites. Genomics, transcriptomics, proteomics and metabolomics is the comprehensive study of respective molecules. The data generated through various omic technologies was integrated to derive complex relationship between the molecule and diseases. Application of these integrative or multi-omics strategies in the areas of women health and rare diseases helped in predicting complexity of pathological process, biomarkers, prognosis and drug targets.

Gynecological cancers are the most deadly diseases among women, whose progression, recovery and death rate can be reduced with identification of a specific drug target and a reliable biomarker. With the advent of multi-omic technologies, predictable biomarkers for ovarian, cervical and endometrial cancers along with rare diseases emerged. Ovarian cancer is a diverse infection among gynecological malignances, has the worst prognosis and the utmost fatality rate. The main streams of people who have OC are typically diagnosed at a later point in the disease's progression because there are no distinct early signs. Cervical cancer one of the most dangerous cancer, death rate is decreased through invention of PAP smear test. Unlike other gynecological cancers endometrial cancer is found in the early stage while in advanced stage treatment is difficult arising the need to study biomarkers for early diagnosis and new therapeutic interventions. Multi-omic technologies made it possible to deal with the complicated role of molecules involved in disease development and treatment. The prognostic biomarker CREB1 and therapeutic agents MEK inhibitors for ovarian cancer, biomarkers synaptopodin-2 and exosomal miR-221-3p for cervical cancer and the MCM2, MCM3, MCM7 and m6A regulators genes were discovered using multi-omics approach to improve diagnosis, therapy and prognosis. This multi-omic technology also correlated DNA modification by methylation to study the complexity of rare disease and its diagnosis.

V. CONCLUSION

Therefore, there is an immediate need for improved biomarkers that may be used in research and clinical practise. Over the course of the past ten years, we have witnessed acceleration in the rate of progress made in sequencing and biotechnology methods. Many women who have been diagnosed with cancers are now receiving excessive treatment because there are no reliable markers to accurately track the course of the disease. There is a need for reliable and predictive biomarkers that may be retrieved from molecular or genetic profiles in order to differentiate between instances that are likely to require medication and those that should be left untreated. Researchers established a structured machine learning approach that enacts multi-omics feature selection and model regularization for the purpose of identifying biomarker combinations that could be used to differentiate low-risk from those with a higher likelihood of progression.

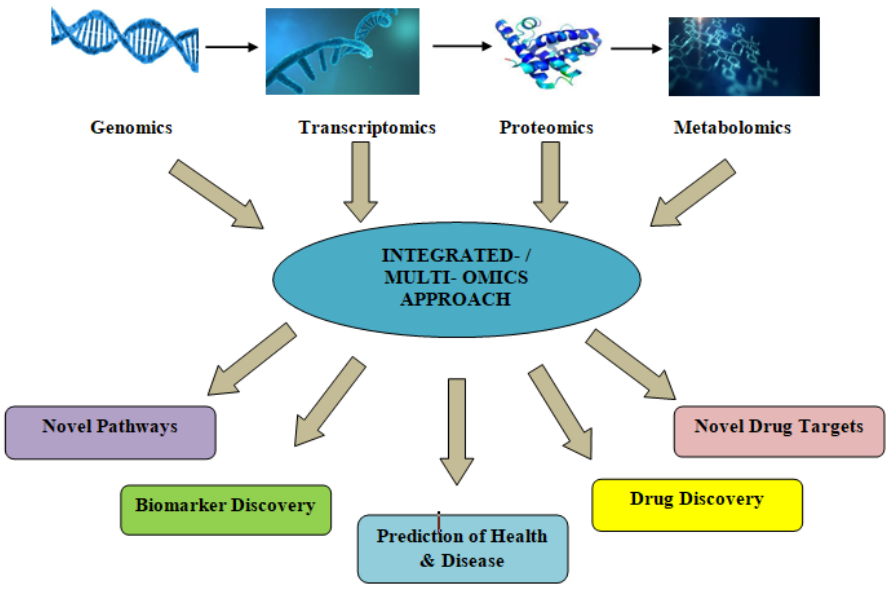

Figure 1: Multi-omics approach and applications

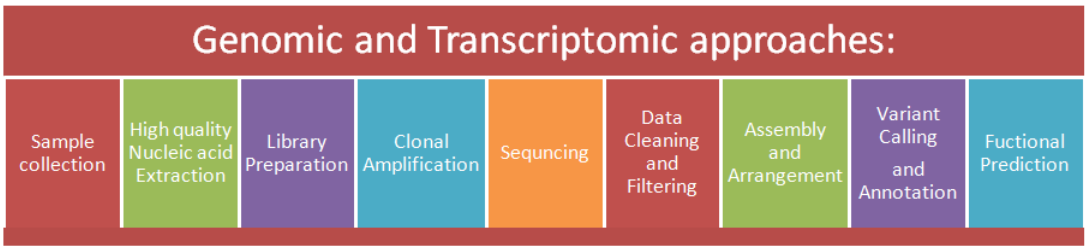

Figure 2: Steps involved in genomic and transcriptomics strategies

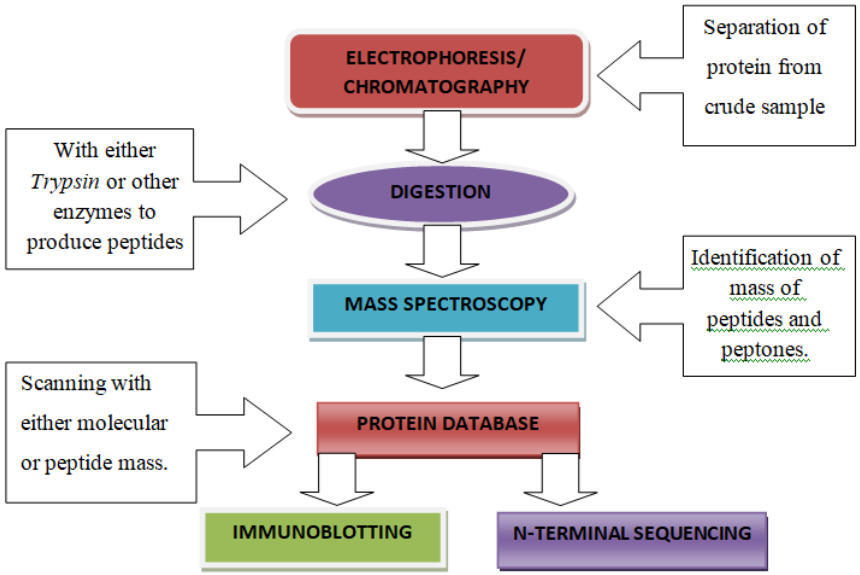

Figure 3: Steps involved in proteomic approach

Table 1: Process for Biomarkers Identifictions and its applications[20]

Sl. No	BIOMARKER IDENTIFICATION PROCEDURE		
	Approaches	**Methods**	**Applications**
1.	**DATA-DRIVEN APPROACH** **Data Reduction** - Complex dataset is made easier to understand - Eliminates noise	**Trend Analysis**	• Determines association between expression levels of particular pair of gene and a variable protein during disease • Reveals statistically relevant patterns in gene expression profiles
		Clustering	• Groups data according to the interactions among the elements • Best suited for initial screening and biomarker verification.
	Classification - Predicts class membership for the existing data	**Regression**	• Most successful method • Selection of biomarkers according to predictive power and estimates the relationship among a set of molecules to create a panel
		Support Vector Machine	• Selects biomarkers from high density datasets
		Decision tree and Random Forest	• Builds a tree of questions to identify markers

Sl. No	BIOMARKER IDENTIFICATION PROCEDURE		
	Approaches	**Methods**	**Applications**
		Artificial Neural Networks	• Machine learning approach which resemble biological neural networks • ANN consists of nodes and links. • Used to acquire panel of biomarkers
		Relative Gene Expression Analysis	• Determines the differences in expression of pair of gene and proteins in disease and non-disease states
1	**Visualisation** - Obtaining results from data through summarising and generating an image	**Principal Component Analysis**	• Reduce the magnitude of complex datasets and reveals the important influencing factors. • Graph visualizes strength of separation of factors in disease and non disease state and helps in successful identification of biomarkers
		Network Analysis	• Network forms are best for representation of large datasets and reveal some biological relationships in expression profiles. • Network form consists of clusters and communities which help in suggesting the new regulatory mechanisms.
2.	**Knowledge-Driven Approach** Integrates heterogenous data, interprets the results and faciltates in undestanding disease and biological process	**Protein-Protein Interactions**	• Helps in identification of underlying mechanisms of disease progression. • Includes experimental identification, characterization and interpretation of protein-protein interactions and application of computational approaches
		Pathway Analysis	• Identifies differentially expressed gene related to function • Done in three steps a) choosing pre-existing gene set, b)asking functionally relevant questions, c) statistical testing • Correlates differently expressed pathways with phenotypes.
		Text mining	• Exploration of knowledge from the existing biomedical literature • Done in three steps a) recognizing terms b) identification of relationship between terms c) discovering new relationship • Powerful biomarker discovery and validation tool.

REFERENCES

[1] Ziad J Sahab, Suzan M Sumaan, Qing-XiangAmy Sang, "Methodology and Applications in disease biomarker identification in human serum," Biomarker insights, vol. 2, pp.21-43,2007
[2] Yehudit Hasin, Marcus Seldin, Aldons Lusis, "Multi-omics approaches to disease," *Genome Biology* vol. 18, pp. 83, 2017
[3] How To Maximize Insights From Precious Samples: The Power Of Multi-Omics ilumina nature research custom media
[4] Ghanbari, Susan Sumner Reza, "Using metabolomics to examine biomarkers of drug addiction," Trends in Molecular Medicine, volume 24, number 197-205, February 2018
[5] Angélique B and J. Philip , "Learning the Human Immunome: The Complexity of Comprehensive Leukocyte Immunophenotyping," (2014).
[6] Yan V. Sun, Yi-Juan Hu, "Integrative analysis of multi-omics data for discovery and functional studies of complex human diseases," Adv Genet., 93, pp. 147–190, 2016
[7] Indhupriya Subramanian, Srikant Verma, Shiva Kumar,Abhay Jere and Krishanpa Anamika, "Multi-omics Data Integration, Interpretation, and Its Application," Bioinformatics and Biology Insights, vol 14, pp. 1–24,
[8] Jason E. McDermott, Jing Wang, Hugh Mitchell, Bobbie-Jo Webb-Robertson, Ryan Hafen, John Ramey, and Karin D. Rodland, " Challenges in Biomarker Discovery: Combining Expert Insights with Statistical Analysis of Complex Omics Data," Expert Opin Med Diagn. , vol. 7, pp. 37–51, January 2013.
[9] https://www.cancer.org/content/dam/CRC/PDF/Public/8773.00.pdf
[10] https://www.cancer.org/content/dam/CRC/PDF/Public/8577.00.pdf
[11] https://www.cancer.org/content/dam/CRC/PDF/Public/8599.00.pdf
[12] Ju-Yueh Li, Chia-Jung Li, Li-Te Lin, Kuan-Hao Tsui, "Multi-omics analysis identifying key biomarkers in ovarian cancer," Cancer Control, vol. 27, pp. 1-10, 2020.
[13] Raunak Shrestha, Marta Llaurado Fernandez, Amy Dawson, Joshua Hoenisch, Stanislav Volik, Yen-Yi Lin, Shawn Anderson, Hannah Kim, Anne M. Haegert, Shane Colborne, Nelson K.Y. Wong, Brian McConeghy, Robert H. Bell, Sonal Brahmbhatt, Cheng-Han Lee, Gabriel E. DiMattia, Stephane Le Bihan, Gregg B. Morin, Colin C. Collins, and Mark S. Carey, "Multiomics Characterization of Low-Grade Serous Ovarian Carcinoma Identifies Potential Biomarkers of MEK Inhibitor Sensitivity and Therapeutic Vulnerability,"Cancer Res, vol. 81, pp.1681–94, 2021.
[14] Yizi Wang, Fang Ren, Zixuan Song, Xiaoying Wang, Xiaoxin Ma, "Multiomics profile and prognostic gene signature of m6A regulators in uterine corpus endometrial carcinoma," *Journal of Cancer*, Vol. 11, pp. 6390-6401, 2020
[15] Miaomiao Ye, Jing Wang, Shuya Pan, Lihong Zheng, Zhi-Wei Wang, Xueqiong Zhu, "Nucleic acids and proteins carried by exosomes of different origins as potential biomarkers for gynecologic cancers," Molecular Therapy: Oncolytics, vol. 24, pp. March 2022
[16] https://www.cancer.org/cancer/endometrial-cancer/about/research.html
[17] Hua Lan, Jing Yuan, Xingyu Chen, Chu Liu, Xiaohui Guo, XinyuWang, Jiarui Song, Ke Cao, Songshu Xia, "Multiomics profiling of the expression and prognosis of MCMs in endometrial carcinoma," *Bioscience Reports, vol.* 41, pp. BSR20211719, 2021
[18] Yizi Wang, Fang Ren, Zixuan Song, Xiaoying Wang, Xiaoxin Ma, "Multiomics profile and prognostic gene signature of m6A regulators in uterine corpus endometrial carcinoma," *Journal of Cancer*, Vol. 11, pp. 6390-6401, 2020
[19] Kerr K, McAneney H,McKnight AJ, "Protocol for a scoping review of multi-omic analysis for rare diseases," BMJ Open, vol. 9, pp. e026278, 2019 doi:10.1136/bmjopen-2018-026278
[20] Zhang-Zhi Hu, Hongzhan Huang, Cathy H. Wu, Mira Jung, Anatoly Dritschilo, Anna T. Riegel, and Anton Wellstein, " Omics-Based Molecular Target and Biomarker Identification Methods," Mol Biol., 719, pp.no. 547–571, 2011.

