The Role of Bioinformatics in Cancer Treatment

Jingnian Shi
Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu, China
909930066@qq.com

ABSTRACT

Cancer is a very aggressive and complicated malignant tumor. It is difficult to treat because many people have it, and its molecular mechanism is very complicated. In the past, in order to find cancer drugs, people often used the method of "blind screening" and tested natural products. But it takes a long time and costs a lot, and we don't always know exactly how these drugs work. With the advent of bioinformatics, the way of doing cancer research has changed a lot. This paper explains how bioinformatics is very important to treat cancer. In particular, he talked about the help of large databases, such as cancer genome map (TCGA), to find different types of tumors and important genes. Next, he discussed how to use multi-omics comprehensive analysis to find new drug targets. This helps to better understand how cancer cells work and find more accurate treatments. Although there are still problems, such as the data are not always clear or it is difficult to use these findings to treat patients, bioinformatics-combining modern technologies such as artificial intelligence and single cell genomics-will help push cancer treatment towards precision medicine and personalized treatment.

Keywords: Bioinformatics; Cancer Treatment; Precision Medicine; Multi-Omics Analysis; Tcga; Drug Target Discovery; Tumor Heterogeneity; Artificial Intelligence.

1. OVERVIEW OF CANCER

Cancer is a disease in which some cells of the body grow uncontrollably and spread to other parts of the body [1]. This is a large group of diseases, in which cells grow too much, invade and may metastasize. Cancer comes exclusively from epithelial tissue. According to the origin of tissue, it can be divided into cancer, sarcoma and hematologic malignancies. Global statistics show that there are more than 20 million new cases every year, and lung cancer, breast cancer and colorectal cancer are the most common [2]. Traditional treatment methods, such as surgery, chemotherapy and radiotherapy, often have the problem of "suitable for everyone", which is based on anatomy rather than molecular characteristics, resulting in a lot of toxicity. Traditional methods can't track the dynamic evolution of tumors because of their great differences in space and time. Bioinformatics has become the main driving force of precision medicine. It combines multiple omics data to understand the heterogeneity of tumors and find the responsible genes, so that patients can carry out personalized treatment, such as targeted therapy or immunotherapy.

2. OVERVIEW OF TRADITIONAL DRUG DEVELOPMENT

2.1 Characteristics and Limitations of Traditional Drug Discovery Stage

Before bioinformatics intervention, the discovery of early cancer drugs followed the empirical model of "blind screening-phenotypic verification-post-mechanical clarification". Researchers have tested a large number of natural products or synthetic chemicals to observe tumor inhibition in vitro, without precise assumptions about molecular targets. For example, the discovery of paclitaxel and adriamycin comes from a large-scale screen without direction; Their unique mechanism was only clarified later, during their development. The lack of systematicness in this path makes this process blind, time-consuming and expensive.

2.2 Limitations of Traditional Target Research Stage

Early target verification depends largely on simple detection techniques, such as radioactive binding assay [3] or northern block [4]. Although they are very important for finding targets such as estrogen receptor (ER) or c-myc oncogene, these methods are slow, unable to process a large number of samples at a time, and unable to display the complex interaction network of the whole cancer genome. This "fragmentation" research paradigm means that many potential hidden targets have been missed, and there is no systematic modeling of bioinformatics.

3. CONCEPT OF BIOINFORMATICS

Bioinformatics is a very important interdisciplinary field. It combines life science with computer science and mathematics. Its value lies in understanding the biological significance of massive data to help disease research and drug development. Specifically, it can help to discover and develop the following aspects of anticancer drugs.

3.1 Methods and Advantages of Using Bioinformatics in Modern Biology

3.1.1 Bioinformatics-Driven Tumor Database Construction: Taking TCGA as an Example

Cancer Genome Map (TCGA) successfully showed the tumors of 33 different types of cancer. For glioblastoma, TCGA found important genes such as ERBB2 and NF1, and showed the relationship between methylation of MGMT promoter and hypermutation phenotype [5]. For colorectal cancer, he found mutations in ARID1A and SOX9, and confirmed the role of MYC regulation in invasion [6]. It also classifies the subtypes of primary prostate cancer, which shows that there are great molecular heterogeneity and potential targeting defects [7]. These results show that large databases are very important for discovering biomarkers and defining molecular subtypes. These molecular classifications help to group patients, discover new biomarkers, and wisely choose targeted or immune-based therapies.

3.1.2 Modern Target Research Stage: Multi-omics Integrative Analysis

Multiomics integration has become the main strategy of modern drug target research. This method goes beyond the limit of single integer data by using algorithms such as iCluster, MOFA or PARADIGM to create models. It calculates the "multi-omics consistency score" and requires the candidate target to show the mutation, expression level and important attributes of clinical results at the same time. For example, the NetICS algorithm has determined that YAP1 and GRB2 are important in liver cancer [8]. The joint modeling of TWAS and PWAS identified 24 genes related to cancer risk and the core role of NRF2 pathway [9]. Algorithms like JDINAC also found important genes related to the survival of breast cancer, which directly helped to create new drugs for clinical trials [10].

4. SUMMARY AND OUTLOOK

Bioinformatics has shifted from an additional tool to the main engine of precision oncology. Research is no longer based on experience, but on data. Multiomics strategy effectively reduced biological noise and false alarm in early research. In the future, we will focus on using single cell sequencing and spatial transcriptomics together. This will give us a better understanding of how cells communicate and how tumors are resistant to drugs. In addition, the generative AI model used to predict the structure of protein will accelerate the development of lead compounds. By combining it with multi-modal AI that analyzes clinical images, omics data and electronic records, the dose and tracking of each patient will be more accurate. This will provide necessary technical support for more effective and safer cancer treatment.

REFERENCES

[1] National Cancer Institute. (n.d.). What is cancer? U.S. Department of Health and Human Services. https://www.cancer.gov/about-cancer/understanding/what-is-cancer

[2] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024 May-Jun;74(3):229-263. DOI: 10.3322/caac.21834. Epub 2024 Apr 4. PMID: 38572751.

[3] Jensen EV, DeSombre ER. "Estrogen-receptor interaction." Science. 1973;182(4108):126-134. DOI:10.1126/science.182.4108.126.

[4] Pachl C, Schubach W, Eisenman R, Linial M. Expression of c-myc RNA in bursal lymphoma cell lines: identification of c-myc-encoded proteins by hybrid-selected translation. Cell. 1983 Jun;33(2):335-44. DOI: 10.1016/0092-8674(83)90415-4. PMID: 6190569.

[5] The Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455, 1061-1068 (2008). https://doi.org/10.1038/nature07385

[6] The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 487, 330-337 (2012). https://doi.org/10.1038/nature11252

[7] The Molecular Taxonomy of Primary Prostate Cancer, Cell. 2015;163(4):1011-1025. DOI: 10.1016/j.cell.2015.10.025.

[8] Dimitrakopoulos, C., Hindupur, S.K., Colombi, M. et al. Multi-omics data integration reveals novel drug targets in hepatocellular carcinoma. BMC Genomics 22, 592 (2021). https://doi.org/10.1186/s12864-021-07876-9

[9] Jin, X., Mei, Y., Yang, P. et al. Prioritization of therapeutic targets for cancers using integrative multi-omics analysis. Hum Genomics 18, 42 (2024). https://doi.org/10.1186/s40246-024-00571-2

[10] Fan Y, Kao C, Yang F, Wang F, Yin G, Wang Y, He Y, Ji J, Liu L. Integrated Multi-Omics Analysis Model to Identify Biomarkers Associated with Prognosis of Breast Cancer. Front Oncol. 2022 Jun 10; 12:899900. DOI: 10.3389/fonc.2022.899900. PMID: 35761863; PMCID: PMC9232398.