

Chapter 5

Multi-Omics Data Analysis for Cancer Research: Colorectal Cancer, Liver Cancer and Lung Cancer



Hantao Zhang, Xun Gong, and Min Tang

5.1 Introduction

Cancer is a disease mainly caused by the accumulation of mutations in two gene classes, which are proto-oncogenes and tumor suppressor genes (Weinberg 1996). With its incidence growing rapidly, cancer is regarded as an important obstacle to human life extension (Torre et al. 2016b). In terms of cancer deaths worldwide for both men and women, lung cancer, colorectal cancer, and liver cancer are top three cancer types (Sung et al. 2021).

Since the twentieth century, lung cancer started to become the most common cause of cancer death as well as the second most commonly occurring cancer in both men and women internationally (Alberg and Samet 2003). It also ranks the most frequently diagnosed cancer and the leading cause of cancer mortality in men (Sung et al. 2021). In the United States, every year the number of patients who die from lung cancer is higher than the combined death toll from colon, breast, and prostate cancer (Spiro and Silvestri 2005). Tobacco smoking is regarded as the leading cause of lung cancer (Salgia and Skarin 1998). Compared with non-smokers, smokers have a 20- to 30-fold increase in lung cancer risk (Minna et al. 2002). Hence, the

Publisher's Disclaimer: The work described has not been published before; that it is not under consideration for publication anywhere else; that its publication has been approved by all co-authors, if any, as well as by the responsible authorities—tacitly or explicitly—at the institute where the work has been carried out. The publisher will not be held legally responsible should there be any claims for compensation.

H. Zhang · M. Tang (✉)
School of Life Sciences, Jiangsu University, Zhenjiang, Jiangsu, China
e-mail: mt3138@ujs.edu.cn

X. Gong
Department of Rheumatology & Immunology, the Affiliated Hospital of Jiangsu University, Zhenjiang, China

industrialized countries, where the smoking prevalence first took place, have the highest lung cancer incidence rates (Alberg et al. 2005). Over the past several decades, because of tobacco control policies and smoking cessation, the smoking prevalence keeps decreasing in those countries (De Groot et al. 2018), thus the burden of lung cancer shifts to developing countries (Torre et al. 2016a). People also gain more knowledge in lung cancer biology. The majority of lung cancers have been divided into four histological types, which are small-cell lung cancer (SCLC) and three non-small-cell lung cancer (NSCLC) types including squamous cell carcinoma, adenocarcinoma, and large cell carcinoma (Wistuba and Gazdar 2006). However, the mortality rates of lung cancer still remain high (Barta et al. 2019), which might be explained by nonspecific symptoms of this disease at early stages (Van Meerbeeck et al. 2011). When seek medical treatment, most patients present with advanced disease which is nearly incurable (Patz et al. 2000).

Nowadays, colorectal cancer (CRC) is the second most common cause of cancer-related death worldwide and the third most common malignant disease (Center et al. 2009). Generally, colorectal cancer has been thought as a disease of the elderly, with rare people being diagnosed before 50, but it also strikes younger people (O'connell et al. 2004). In addition, colorectal cancer is the only type that strikes both men and women with approximately equal frequency (Potter 1999), since it is the second most common cancer in females and the third most common cancer in males (Siegel et al. 2014). What's more, the incidence rates of colorectal cancer vary greatly around the world (Stintzing 2014). It is well-known that most cases of colorectal cancer are detected in western countries (Mármol et al. 2017), because people in longstanding developed countries often exhibit same factors playing important roles in the development of colorectal cancer, which might include obesity, unhealthy diet, smoking, alcohol consumption, and physical inactivity (Fearon 1995; Weinberg and Schoen 2014). However, in recent years, high incidence rates of CRC have been observed in newly developed countries where the risk of suffering from colorectal cancer was once quite low (Mármol et al. 2017).

Liver cancer is regarded as an aggressive and heterogeneous tumor which ranks the third most common cause of cancer-related death as well as the second leading cause of cancer-related death in man around the world (Yamashita and Wang 2013; Gao et al. 2019). In addition, liver cancer can be divided into primary liver cancer and secondary liver cancer in nature (Mckillop and Schrum 2005). As for primary liver cancer (PLC), based on different histological features, it can be categorized into six subtypes, which are hepatocellular carcinoma (HCC), intrahepatic cholangiocarcinoma (iCCA), mixed hepatocellular-cholangiocarcinoma (HCC-CCA), fibrolamellar HCC, and the pediatric neoplasm hepatoblastoma (Mcglynn et al. 2001; Srivatanakul et al. 2004). Among these histological types, HCC is the commonest primary liver cancer worldwide which accounts for nearly 90% of all cases of primary liver malignancies (Ariff et al. 2009). The second most frequent type of primary liver cancer is iCCA, the incidence rates increase steadily (Sia et al. 2017). What's more, the incidence rates of liver cancer in different countries vary significantly (Bosch et al. 1999). Blaming for hepatitis B virus (HBV) infection, the East and Southeast Asia as well as the Middle and Western

Africa have the highest liver cancer rates (Bosch et al. 2004). Thanks to HBV vaccine, liver cancer incidence rates is decreasing in several highest-risk areas (Chen and Zhang 2011). However, in some low-risk western countries, the rates continue to increase. Risk factors such as obesity, cigarette smoking, hepatitis C virus (HCV) infection and chronic alcohol abuse are believed to be related to liver cancer in these areas (Bishayee 2014). Gender is another risk factor for liver cancer development, males are more susceptible than females, as the incidence rates of liver cancer among men is over twice that among women (Liu et al. 2015).

Through molecular and genetic studies of cancer, multiple biomarkers of colorectal cancer, liver cancer, and lung cancer have been identified (Zochbauer-Muller and Minna 2000; Bishayee 2014; Dienstmann et al. 2017). However, it is quite difficult to find diagnostic, prognostic, and therapeutic targets from these outcomes, and the morality is still high for patients all over the world (Chakraborty et al. 2018). With the advancement of high-throughput omics technologies, researchers are now able to study genomics, transcriptomics, proteomics, and phosphoproteomic data at the same time (Ahmed 2020). Although through analyzing single omics data set, one can observe the alternation and association of biological entities at that level, the interaction between multiple molecular layers cannot be fully assessed (Biswas and Chakrabarti 2020). Hence, in lung cancer, liver cancer, and colon cancer research, many multi-omics analyses have been conducted in order to gain a holistic view of the molecular dynamics underlying cancer progression and to make a progress in early detection and prognosis (Sun and Hu 2016). Also, because of the heterogeneous nature of cancer, different patient may have different clinical responses to the same treatment (Du and Elemento 2015). For this problem, multi-omics studies at an individual level have been conducted to develop precision cancer medicine (Ghosh et al. 2018; Mantini et al. 2021).

In this review, we introduced different types of omics data used in the research of colorectal cancer, liver cancer, and lung cancer. In addition, we summarized currently used technologies for high-throughput multi-omics data analysis. We also reviewed integrative analyses using genomic, epigenomic, transcriptomic, proteomic, and metabolomics data that helped reveal the molecular pathology of colorectal cancer, liver cancer, and lung cancer. Finally, we discussed challenges and envisioned the future of precision cancer medicine.

5.2 Various Multi-Omics Data Types and Selected Repositories

With the advent of sequencing technologies, biomolecules in a given biological samples can be identified and quantified at multiple omics levels (Das et al. 2020). Next-generation sequencing (NGS) is now frequently used for whole-genome or whole-exome sequencing (Behjati and Tarpey 2013). ChIP-seq (chromatin immunoprecipitation) and DNase I-seq (DNase I hypersensitive sites-sequencing) are used

for detection of DNA-protein interactions. RNA-seq can be used to identify and quantify RNA molecules (Kim and Dekker 2018; Lu et al. 2019). As for proteomic and metabolomic study, mass-spectrometry based techniques are widely used (Domon and Aebersold 2006). Omics data generated by these techniques, including but not limited to genomic, epigenomic, transcriptomic, proteomic, and metabolomic, is together called as multi-omics data (Liu et al. 2019). There are several publicly accessible databases listed in references (Huang et al. 2017; Subramanian et al. 2020), which accommodate multiple omics data sets and serve as rich resources for understanding the etiology of human cancer.

5.2.1 DriverDB v3

The DriverDB database (<http://ngs.ym.edu.tw/driverdb/>) contains numerous exome-seq data that was extracted from The Cancer Genome Atlas (TCGA), The International Cancer Genome Consortium (ICGC), Prostate Cancer Genetics Project (PCGP), The Therapeutically Applicable Research to Generate Effective Treatments (TARGET), and published papers (Cheng et al. 2014). More exome-seq data as well as additional RNA-seq data from TCGA, ICGC, and published papers were added to updated DriverDB v2 (Chung et al. 2016). DriverDB v3, the latest version, incorporated not only new exome-seq and RNA-seq datasets but also copy number variation (CNV), methylation, and smRNA-seq datasets. By applying various bioinformatic tools it contains, users can identify abnormalities at multi-omics levels and discover driver genes and mutations (Liu et al. 2020a).

5.2.2 TCGA Portal

The Cancer Genome Atlas (TCGA) was launched by The National Institute of Health (NIH) in 2006 aiming to reveal genomic and epigenomic alternations associated with 32 types of human cancers (Wang et al. 2016). For each type of human cancer, various kinds of data including gene expression, exon expression, miRNA expression, protein expression, single nucleotide polymorphism (SNP), copy number variation (CNV), loss of heterozygosity (LOH), and DNA methylation has been generated and processed (Tomczak et al. 2015). The aforementioned data are stored in a free-access database, namely the TCGA Data Portal (<https://tcga-data.nci.nih.gov/tcga/>). Without a doubt, the wealth of TCGA data has led to the discovery of diagnostic biomarkers and development of new cancer therapies (Colaprico et al. 2016).

5.2.3 ICGC

The International Cancer Genome Consortium (ICGC; <https://icgc.org/>) mainly contains mutational genomic data in nearly 50 cancer types. The International Cancer Genome Consortium Data Portal (<https://dcc.icgc.org>) is a user-friendly platform which helps users visualize, analyze, and interpret cancer-related genetic, molecular, and clinical data it contains. This may lead to deeper understanding of tumor biology as well as development of better diagnostic methods and drugs (Zhang et al. 2019).

5.2.4 CCLE

In order to promote the translation of genetic and pharmacological data generated by cancer cell line studies into understanding of cancer progression and development of novel therapies, Cancer Cell Line Encyclopedia (<https://portals.broadinstitute.org/ccle>) was built by the collaboration between the broad Institute and the Novartis Institute (Barretina et al. 2012). The original release of CCLE contains a large-scale genomic data set from 947 human cancer cell lines and pharmacological profiling of 24 anticancer drugs across 479 of those cell lines. Later, whole genome sequencing, RNA-seq, miRNA profiling, and histone profiling were added to it (Nusinow et al. 2020).

5.2.5 LinkedOmics

The LinkedOmics database (<http://www.linkedomics.org>) contains mass spectrometry (MS)-based global proteomics data which was downloaded from the Clinical Proteomic Tumor Analysis Consortium (CPTAC). Multi-omics data including genomic, epigenomic, and transcriptomic data as well as clinical data for 32 TCGA cancer types which were downloaded from The Cancer Genome Atlas (TCGA) project were also added to this database. Aiming to allow users to analyze these data in detail, LinkedOmics provided three analysis modules, namely LinkFinder, LinkCompare, and LinkInterpreter. For each cancer cohort, the LinkFinder module allows user to find associations between an attribute of interest and all other attributes. These associations can be compared with query attributes through the LinkCompare module and interpreted through the LinkInterpreter module. The results are presented in the form of plot or heatmap, which may effectively help users gain biological understanding (Vasaikar et al. 2018).

5.2.6 *RHPCG*

Consisting of a group of kinases, hippo signaling pathway is a highly conserved pathway which plays important roles in controlling cell proliferation, apoptosis, and migration. Dysregulation of Hippo signaling pathway is involved in the initiation and progression of cancer, such as breast cancer, lung cancer and so on. The Regulation of the Hippo Pathway in Cancer Genome database (<http://www.medsysbio.org/RHPCG>) can serve as an open resource for visualizing alternations of Hippo pathway genes as well as understanding the roles of Hippo pathway in cancer, because RHPCG was designed to allow users easily search, view, and download alternations of core Hippo-protein-encoding genes in 33 cancer types at levels of genomics, epigenomics, and transcriptomics (Wang et al. 2019).

5.2.7 *MOBCdb*

The Multi-Omics Breast Cancer Database (<http://bigd.big.ac.cn/MOBCdb/>) was constructed in order to facilitate identification of breast cancer subtypes and discovery of novel biomarkers. MOBCdb contains SNV, gene expression, microRNA expression, DNA methylation, clinical, and drug response data that were downloaded from the TCGA data portal, GENECODE, miRBase, PharmGKB, and NCBI. With more than 10,000 files stored in the database, MOBCdb provides several methods to help users effectively gain information. In addition, by using the genome-wide browser in MOBCdb, users can visualize different omics data easily. The survival module was designed to help users find new biomarkers (Xie et al. 2018).

5.2.8 *Target*

The Therapeutically Applicable Research to Generate Effective Treatments database (<https://ocg.cancer.gov/programs/target>) was built by the cooperation of extramural and NCI investigators. TARGET originated with two pilot projects, now it contains the clinical information, gene expression, miRNA expression, copy number, and sequencing data of 24 molecular types of cancer. The effort of TARGET researchers has undoubtedly accelerated discoveries of genomic alterations in cancer and facilitated rapid translation of those findings into the clinic (Wu et al. 2021a).

There is much information that can be obtained from the data sets stored in the aforementioned databases. For instance, genomic studies can reveal the associations between tumorigenesis and genetic mutations (Ghosh et al. 2018). Also, Epigenomic data can lead to knowledge regarding how chemical modifications of DNA and protein drive tumorigenesis (Rhee 2018). Similarly, transcriptomic profiling can be

used to detect the association between cancer and dysregulated genes (Canzler et al. 2020). Proteomic data can help researchers better understand its function in human cancer (Matthiesen and Jensen 2008). Because each omics data type only provides a partial view of the complexity of cancer, biological mechanisms can be fully captured only through integrating different omics data types (Hao et al. 2019).

5.3 Selected Integrative Tools for Multi-Omics Analysis

Cancer is a consequence of malfunction and alteration in multiple molecular layers (Hausman 2019). With decreasing time and cost to generate multiple omics datasets from biological samples, an increased need for large-scale omics analysis tools emerged to explore relationships between different biological readouts (Altenbuchinger et al. 2020). Usually, steps to conduct an integrative analysis of these readouts include data normalization, variable selection, cluster analysis, and dimensional reduction (Meng et al. 2016; Chauvel et al. 2020; Nicora et al. 2020). In this section, we review eight computational integrative tools that are capable of multi-omics data analysis. The first five tools were designed to reveal the biological mechanisms connecting identified key drivers and pathways to diseases. The remaining three tools can be used to discover new therapeutic interventions or support clinical decision making.

Integrative Omics Data Analysis (iODA) is a software for omics data analysis, which is written in Java and able to run on Windows or Linux operating systems. iODA can integrate and refine data generated by RNA-seq, miRNA-seq, and ChIP-seq, which leads to the revelation of complex pathogenesis of human cancer. There are six statistical methods included, namely Least Sum of Ordered Subset Squared, Cancer Outlier Profile Analysis, Maximum Ordered Subset T-statistics, Outlier Robust T-statistics, Outlier Sum, and t-test, which can be selected by users to process their input data. Then, differentially expressed genes and miRNAs as well as transcription factor binding sites are extracted for the following pathway enrichment analysis and consistency analysis. The dysfunctional molecules are mapped on the KEGG pathway, and the consistent molecular signatures are identified as key pathogenic factors in cancer. The source code as well as executable file of iODA can be downloaded at <http://www.sysbio.org.cn/iODA> for free (Yu et al. 2020).

The interactive tool for statistical analysis of omics and clinical data (IOAT in short) is a R and Python-based Windows application for analyzing and visualizing multi-omics and clinical data. IOAT is a user-friendly tool designed for non-programmers. It can perform feature screening, risk assessment, clustering, and survival analysis after reading a comma-separated value text file imported by users and preprocessing the multi-omics and clinical data contained in the file. All results are displayed in a report, which enables users to view the outcomes of each step and thus gain a better understanding of their data. Additionally, IOAT considers data breaches. After downloading an executable file from <https://github.com/WISun>

shine/IOAT-software, users can use this desktop tool without the need for network connectivity, ensuring the security of their personal data (Wu et al. 2021b).

MEXPRESS is a simple and user-friendly web tool for visualizing and interpreting multiple omics data that does not require clinical researchers to be programmers. Users can view gene expression, DNA methylation, and clinical data extracted from TCGA by entering a gene name and selecting a cancer type. MEXPRESS can also be used to conduct statistical analyses on these datasets and determine their correlation, which is extremely useful for biomarker discovery (Koch et al. 2015). While the core functions of MEXPRESS remain unchanged in the new version released in 2019, new data types, statistical methods, and options are included. All code is available for free download at <https://github.com/akoch8/mexpress> (Koch et al. 2019).

PROMO is a powerful and integrative Windows software written in Matlab that is designed to analyze large genomic and clinical datasets contained in multi-omics databases effectively. It includes several features such as data preprocessing, exploration and visualization, clustering, enrichment analysis, biomarker discovery, and classification of cancer subtypes. After importing a multi-omics dataset into PROMO, users can discover correlations between features at various multi-omics levels as well as the genes involved in biological differences, resulting in a better understanding of biological mechanisms and the discovery of new biomarkers. PROMO is freely accessible to the public at <http://acgt.cs.tau.ac.il/promo/> (Netanelly et al. 2019).

Chromatin structures, such as topologically associating domains (TAD) and TAD boundaries, are critical for gene expression regulation. Changes in the structure of chromatin may contribute to the progression of human cancer (Valencia and Kadoch 2019). PredTAD is a machine learning tool that uses the Gradient Boosting Machine (GBM) algorithm to predict 3D chromatin structures. It makes use of genomic and epigenomic data to predict and detect TAD boundary variants in normal and cancer cell genomes. Correlations between TAD boundary alternations and the expression of nearby genes can be identified using RNA-seq data analysis. Because genes located near altered boundaries may be involved in a cascade of oncogenic signaling pathways, PredTAD is an effective tool for transforming genomic and ChIP data into an understanding of the roles of chromatin structures in cancer progression. The source code for PredTAD is available at <https://github.com/jchyr-sbmi/PredTAD/> (Chyr et al. 2021).

IOBR is a computational tool for interpreting multi-omics data; its application in immuno-oncology biological research has the potential to shed new light on tumor-immune interactions and accelerate the development of immunotherapies. It is composed of four functional modules: signature and tumor microenvironment (TME) estimation, phenotype estimation, mutation estimation, and module construction. IOBR is capable of identifying signature genes and phenotype-relevant signatures, analyzing signature-associated mutations, and building models using previously identified signatures. These models can be used to forecast therapy response, prognosis for cancer, and tumor resistance. The IOBR R package can be downloaded from <https://github.com/IOBR/IOBR> (Zeng et al. 2021).

DrugComboExplorer, a computational systems biology tool, predicts drug combinations for specific cancer types by integrating DNA-seq, RNA-seq, methylation, and gene copy number data. It processes multi-omics data from cancer patients, identifies driver signaling networks, and quantifies the efficacy of combinatorial drugs on these networks using multiple algorithms. Combinations of optimal drugs that target driver signaling networks may be a way to copy resistance progression. The source code for DrugComboExplorer is available at <https://github.com/Roosevelt-PKU/drugcombinationprediction> (Huang et al. 2019).

OncoPDSS is a system that interprets multi-omics variants detected in cancer samples as supporting evidence for clinical pharmacotherapy decision-making. It contains the OncoPDSS knowledgebase (OncoPDSSkb), which was created to store data on drug-drug interactions, clinical trials for cancer, and drug indications. OncoPDSS imports user-uploaded variants. It uses a classification strategy to determine whether pharmacotherapies are potentially effective or not based on OncoPDSSkb mutation records, cancer records, and drug records that serve as oncology pharmacotherapy evidence. As a result, this tool will significantly aid clinicians and physicians in making clinical decisions, while also providing cancer researchers with novel treatment strategies. OncoPDSS is accessible via the following link: <https://oncopdss.capitalbiobigdata.com> (Xu et al. 2020a).

Recent cancer projects as well as multi-omics databases provide the research community with a wealth of omics data and clinical information on cancer patients (Cieslik and Chinnaiyan 2020). Integrative analysis of these data is challenging and requires bioinformatics, statistical, and programming skills (Chakraborty et al. 2018; Park et al. 2020). Numerous tools have been built to solve this problem. However, some limitations still exist. For instance, iODA only supports the analysis of mRNA, miRNA, and CHIP-seq data (Yu et al. 2020). Efforts should be devoted to develop new tools that can be applied for all omics data types. In addition, several tools utilize the R language, which is not friendly for researchers with limited biostatistical or programming knowledge (Eicher et al. 2020; Graw et al. 2021). Web-based interfaces should be created to allow fundamental researchers to leverage the merits of multi-omics tools.

5.4 Overview of Cancer Multi-Omics Research

5.4.1 Lung Cancer

Lung cancer is a highly complex and heterogeneous disease (De Sousa and Carvalho 2018). In recent decades, cancer researches focusing on the discovery of prognostic indicators and therapeutic targets have already been made (Jones and Baldwin 2018). Li proposed a novel method for mining cancer-related gene modules based on multi-omics data. First, genome-wide regulatory networks were constructed using key regulatory factors identified by feature selection method. Second, dysregulated gene sets were identified by comparing regulatory networks in variant

and non-variant samples, which were then used to generate cancer-related gene modules. This new mining method has been proved to be applicable to lung cancer research (Li et al. 2019). By analyzing genomic, transcriptomics, and proteomic data, Kong et al. identified abnormal expressed membrane proteins in highly metastatic lung cancer cells. The high expression level of *CDH2*, *EGFT*, *ITGA3*, *ITGB1*, *ITGA5* and low expression level of *CALR* were found to be associated with cancer metastasis (Kong et al. 2020).

Small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) are two major types of lung cancer (Wu et al. 2020). Patients diagnosed as NSCLC accounts for nearly 85% of all lung cancer patients, which makes NSCLC the most common histological type of lung cancer (Wang et al. 2018). Chen et al. performed gene expression, prognosis, DNA methylation, and gene mutation analysis of *NUF2* gene. It was shown that that the more *NUF2* expressed, the poorer prognosis patients had. Thus, *NUF2* might be considered as a prognostic biomarker of NSCLC and can be used for cancer treatment (Chen et al. 2014). Luan et al. integrated DNA methylation, RNA, miRNA and DNA copy number data to construct a survival risk model. Based on this, the chromosome regions 17q24.3 and 11p15.5 were identified as the copy number variation regions that were associated with NSCLC patient survival (Luan et al. 2020).

NSCLC can be further divided into three main subtypes, lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), and large cell carcinoma (LCC) (Herbst et al. 2018).

Numerous potential biomarkers have been identified as a result of advancements in the molecule biology of LUAD. Paula et al. used proteomic data, ChIP-seq and RNA-seq assays to demonstrate that *MGA* gene, which is mutated and copy number deleted in LUAD, acts as a tumor suppressor by repressing genes activated by the MYC pathway. This discovery may open new therapeutic avenues (Llabata et al. 2020). Zhang et al. estimated different tumor microenvironment infiltration patterns and the correlation between these patterns and the genetic or epigenetic alterations by analyzing expression, RNA-seq, WES, and DNA methylation profiles. A prognosis model was constructed using the detected genetic and epigenetic alternations, which may aid in the development of a more accurate prognostic predictor for human LUAD (Zhang et al. 2020b). Ken built a SVM to subclass patients based on their survival based on clinical data from LUAD. By combining RNA expression and miRNA expression data of these subtypes, six genes were efficiently identified to be associated with LUAD patient survival: *ERO1B*, *DPY19L1*, *NCAM1*, *RET*, *MARCH1*, and *SLC7A8* (Asada et al. 2020). Lee et al. applied mRNA, miRNA, DNA methylation and CNV data to develop a deep learning autoencoding approach for survival risk stratification. They successfully identify significant prognostic difference between two groups of LUAD patients using this model (Lee et al. 2020).

LUSC has a worse prognosis than LUAD (Zhang et al. 2020a). Numerous studies have already been conducted to ascertain the molecular characteristics of this subtype. According to Zhang, an integrative analysis of methylation and gene expression data revealed that 113 methylation features and 23 gene expression features are strongly associated with lung cancer. *SFTA3* and *LPP* may serve as

molecular markers for subtyping NSCLC (Zhang et al. 2020a). Additionally, Xu et al. investigated the gene expression changes associated with DNA copy number or DNA methylation in LUSC patients by integrating genomic, transcriptomic, and epigenetic data. Seven genes expressed at a high level, which could be due to CNV or methylation and result in a poor prognosis (Xu et al. 2020c). Additionally, Hu et al. examined multi-omics differences between LUSC patients with high and low levels of programmed death 1 expression (PD1). It was discovered that 178 genes involved in immunity were significantly upregulated in the high expression group, which may contribute to a better understanding of the relationship between PD1 and immunotherapy effect (Hu et al. 2020).

Pulmonary sarcomatoid carcinomas (PSC) is a rare tumor in the family of NSCLC (Antoine et al. 2016). Yang et al. conducted multi-omics analysis of PSC samples and found out that PSC may be converted from the epithelial components and can be divided into five subtypes based on different histological morphologies (Yang et al. 2020b). Also, it was delineated that a large portion of patients had mutations in the p53, RTK/RAS, and PI3K pathways, suggesting that targeted therapy could be an option for patients with PSC (Yang et al. 2020b). Totally, their study shed light on the biological nature and brought entry points for the treatment of this rare malignancy (Yang et al. 2020b).

5.4.2 Colorectal Cancer

Colorectal cancer (CRC) is a heterogeneous disease (Berg et al. 2017; Almusawi et al. 2021). Various studies performed in recent years have provided insights into the molecular characteristics of CRC. Xu et al. explored genes related to CRC prognosis and incidence (Xu et al. 2020b). Genes annotated with single nucleotide mutation sites, copy number variation sites, and methylation sites along with differentially expressed genes were identified as candidate genes (Xu et al. 2020b). Moreover, a weighted gene co-expression network analysis was performed to search for hub genes (Xu et al. 2020b). Finally, *LRRC26* and *REP15* were identified as CRC-specific driving genes (Xu et al. 2020b). Yuan et al. attempted to link genetic variants, genes, and risk of CRC (Yuan et al. 2021). They conducted expression quantitative trait loci (eQTL) analysis, meta-analysis, and methylation quantitative trait loci (mQTL) analysis of 131 lead SNPs to explore potential target genes (Yuan et al. 2021). In addition, a colocalization analysis of genes identified in the previous step was performed, which revealed 66 putative susceptibility genes in CRC (Yuan et al. 2021). Ayiomamitis et al. investigated the roles of cyclooxygenase 2 (COX-2), an enzyme that promotes prostaglandin E2 (PGE2) production, and human telomerase reverse transcriptase (hTERT), a component of telomerase, in the onset of CRC (Ayiomamitis et al. 2019). By analyzing the expression levels of COX-2, PGE2, and hTERT along with telomerase activity, they demonstrated that COX-2 plays a key role in the initial stages of CRC development (Ayiomamitis et al. 2019). Also, high COX-2 expression was found to be associated with low hTERT expression and a

better survival among CRC patients (Ayiomamitis et al. 2019). To gain a better understanding of the clinical relevance between obesity and CRC, Holowatyj et al. performed transcriptomic analysis on visceral adipose and tumor tissues and metabolomics analysis on blood samples of CRC patients (Holowatyj et al. 2020). Combining results generated by each omics measurement, they elucidated that glycolytic metabolism, GPVI signaling, and fibrosis participated in the adipose-tumor crosstalk and could promote CRC development (Holowatyj et al. 2020). Ghaffari et al. investigated the underlying mechanisms that drive metastatic progression (Ghaffari et al. 2021). They performed RNA-seq, ChIP-seq, and ATAC-seq on a CRC cell line (Ghaffari et al. 2021). Then, a statistical model was used to comprehensively analyze these multi-omics profiles along with TF-DNA binding information (Ghaffari et al. 2021). It was elucidated that JunD, a TF, plays a crucial role in CRC migration and invasion (Ghaffari et al. 2021).

It is widely accepted that most colorectal cancers arise as a result of transformation from adenoma to adenocarcinoma (Lam et al. 2016), which is triggered by the stepwise accumulation of genetic and epigenetic mutations (Aarons et al. 2014). Using the deep learning framework, Lv et al. constructed a prognostic model for patients with colon adenocarcinoma (COAD) using the TCGA and GEO databases (Lv et al. 2020). After applying this model to the TCGA dataset, it was discovered that two subgroups with significantly different survival rates existed. Further analysis of these two subgroups revealed 1217 differentially expressed genes and ten differentially expressed miRNAs, which may aid in deciphering the mechanisms underlying COAD development (Lv et al. 2020). Yin et al. proposed an approach to detect potential prognosis risk biomarkers (PRBs) (Yin et al. 2020). First, based on gene expression, exon expression, DNA methylation, and somatic mutation profiles along with clinical information of COAD patients, the multi-omics-based prognostic analysis (MPA) model was used to select features closely related to the prognosis of COAD patients (Yin et al. 2020). Second, they applied the protein-protein interaction (PPI) network to annotate the functions of these features (Yin et al. 2020). Finally, 13 features were identified as PRBs through the further validation, which may serve as drug targets in COAD treatment (Yin et al. 2020).

CRC is also known as bowel and colon cancer, which makes colon cancer (CC) a subset of it (Jahanafrooz et al. 2020). Tong et al. successfully constructed a prognostic prediction model of CC patients by integrating clinical features, gene expression, miRNA expression, and DNA methylation data extracted from TCGA (Tong et al. 2020). Compared with models based on clinical and gene expression data, this integrative prognostic model was more effective, suggesting that the more types of omics data integrated, the better the cancer prognostic model would perform (Tong et al. 2020). Yang et al. also established a prognostic model for CC (Yang et al. 2020a). They first conducted an identification of differentially methylated genes, differentially expressed genes and miRNAs between tumor samples and normal samples (Yang et al. 2020a). Then, using omics features correlated with prognosis, the prognostic model was built, which might be helpful for CC research (Yang et al. 2020a). Yi et al. explored the underlying mechanisms of Wnt/ β -catenin signaling regulating EMT program (Yi et al. 2020). It was validated that the RUNX2

expression activated by Wnt signaling pathway would lead to an increase in the expression of EMT-associated genes (Yi et al. 2020). Because EMT has been proved to be highly correlated with metastasis formation and tumorigenesis (Pastushenko and Blanpain 2019), RUNX2 might serve as a prognostic biomarker for CC. Arora et al. detected the dysregulated expression pattern of seven classical non-homologous end joining (c-NHEJ) pathway genes in CC (Arora et al. 2020). Compared to normal tissues, *XRCC5*, *XRCC6*, *PRKDC*, and *PAXX* were observed to be overexpressed in tumor tissues, whereas the expression level of *LIG4* and *NHEJ1* were downregulated (Arora et al. 2020). In addition, *PAXX* was identified as a prognostic biomarker (Arora et al. 2020). Thus, their study may help reveal the clinical significance of c-NHEJ pathway genes in CC. Using a novel upstream analysis strategy, Kel et al. deciphered the molecular mechanisms of the resistance to methotrexate (MTX) in CC (Kel et al. 2016). This strategy mainly contains two steps, i.e., the identification of transcription factors (TFs) and master regulators that activate these TFs (Kel et al. 2016). After applying this approach to transcriptomics, proteomics, and ChIP-seq data, PKC-alpha, TGF-alpha, TGF-beta, and alpha9-integrin were identified as anti-resistance targets (Kel et al. 2016). Their findings would provide new insight into oncology drug resistance research.

Left-sided colon cancer (LCC), which originates from the hindgut, and right-sided colon cancer (RCC), which originates from the midgut, are two subtypes of CC (Song et al. 2020). In addition to the different tumor locations, there are many differences between them (Shen et al. 2015). To gain a better understanding of these differences, Huang et al. analyzed transcriptomics, clinical, and somatic mutation data of patients with CC (Huang et al. 2021). A total of 360 differentially expressed genes were observed (Huang et al. 2021). Among them, it was indicated that *BRAF* and *KRAS* mutations were frequently presented in RCC, whereas *APC* mutation was frequently presented in LCC (Huang et al. 2021). In addition, the 4-mRNA and 6-mRNA were identified as prognostic signatures for LCC and RCC, respectively (Huang et al. 2021). Similarly, Hu et al. conducted a study on the differences in molecular features between LCC and RCC (Hu et al. 2018). It was revealed that *PARC* was hypermethylated in RCC, whereas *CDX2* was hypermethylated in LCC (Hu et al. 2018). Also, the expression levels of miR31, miR155, and miR625 were observed to be upregulated in RCC, whereas the expression levels of miR-296 and miR592 were downregulated in LCC (Hu et al. 2018). In addition, compared with LCC, the mutation rate of *KRAS* and *BRAF* was higher in RCC, which was believed to be associated with a worse prognosis (Hu et al. 2018). Yi et al. performed a systematic analysis on the regulatory mechanisms between gene mutations and tumor immune microenvironment (TIME) in LCC and RCC cells (Yi et al. 2021). It was revealed that the mutations of top mutated genes were strongly correlated with TIME, DNA methylation levels of some immune checkpoints, and immune-related genes and miRNAs in RCC. However, these associations were less significant in LCC (Yi et al. 2021).

5.4.3 Liver Cancer

Liver cancer, one of the extraordinarily heterogeneous diseases, is caused by the interplay of various internal and environmental factors (Li and Wang 2016; Marengo et al. 2016). The development of omics strategies has helped us gain a holistic view of tumor biology. Shen et al. distinguished two molecular subtypes by analyzing genomic, epigenomic, and transcriptomic data from patients with liver cancer (Shen et al. 2021b). In addition, two prognostic molecular targets, ANXA2 and CHAF1B, were highly expressed in tumor tissues and identified to be strongly related to the prognosis of liver cancer patients (Shen et al. 2021b). Their research findings could provide new insight into the exploration of key biomarkers and mechanisms of liver cancer (Shen et al. 2021b).

Primary liver cancer is a serious public health issue, with HCC as the most common pathological subtype (Lin et al. 2016). Significant effort has been made to reveal the biological nature of HCC. Based on multi-omics datasets of HCC samples downloaded from TCGA and GEO databases, Liu et al. conducted an investigation on the methyltransferase-like 3 (METTL3) as well as methyltransferase-like 14 (METTL14), which were both core molecules of a multicomponent methyltransferase complex (MTC) that catalyzed the formation of N6-methyladenosine (m6A) (Liu et al. 2020b). It was clarified that METTL3 and METTL14 influence distinct signaling pathways and biological processes, thus may play opposite regulatory roles in HCC (Liu et al. 2020b). Using several databases, Jin et al. investigated the impact of the expression levels of *CDK1*, *CCNB1*, and *CCNB2* in the survival of HCC patients (Zou et al. 2020). The upregulation of *CDK1*, *CCNB1*, and *CCNB2*, which might be caused by low levels of methylation or genomic alternations, was found to be highly correlated with poor prognosis in HCC patients (Zou et al. 2020). Using multi-omics analysis of metabolomics and absolute quantification proteomics, Dan et al. conducted an investigation on the effects of canagliflozin (CANA) on the proliferation of HCC cell lines (Nakano et al. 2020). It was shown that CANA, the sodium glucose co-transporter 2 (SGLT2) inhibitor, mainly altered oxidative phosphorylation metabolism, fatty acid metabolism, and DNA synthesis, which may suppress cell proliferation of Hep3B and Huh7 cells (Nakano et al. 2020). Shen et al. performed a multi-omics analysis to explore the metabolic impact of estrogen and its receptors in HCC cells (Shen et al. 2021a). It was suggested that estrogen acts on its receptors to suppress HepG2 cell growth via altering glucose and lipid metabolism, which might be part of the reason why women have a lower risk of HCC development as compared to men worldwide (Shen et al. 2021a). Woo et al. integrated CNV, DNA methylation, and mRNA expression data of a cohort of HCC patients to identify DNA copy-number-correlated (CNVcor) and methylation-correlated (METcor) genes (Woo et al. 2017). The frequencies of CNVcor gene aberration were indicated to be significantly correlated with frequencies of METcor gene aberration, demonstrating that the concomitant regulation of transcriptomes by alternations in DNA copy numbers and methylation should be taken into consideration in liver cancer research (Woo et al. 2017).

In developing countries prevalent for hepatitis B virus (HBV) infection, HBV still remains the most common etiologic agent of HCC (Chang 2014). Much work also has been done to uncover the direct and indirect mechanisms that are involved in HCC oncogenesis by HBV (Xie 2017). Through the integration of proteomics and metabolomics assays, Xie et al. conducted an exploration on the mechanisms of HBV-induced HCC (Xie et al. 2017). They demonstrated that HBV core protein might contribute to the progression of HCC by modifying the metabolism of glycolysis and amino acid (Xie et al. 2017). Consequently, HBV core protein could represent a promising target for antiviral therapy (Xie et al. 2017). Aiming to identify novel biomarkers in HCC, Miao et al. performed multi-omics analyses integrating genomic, transcriptomics, and clinicopathological data of patients with HBV-related multifocal HCC (Miao et al. 2014). Six genes with abnormal expression levels were identified (Miao et al. 2014). Among them, *TTK* might be an overall prognostic indicator for HCC, because the expression level of *TTK* was shown to be highly correlated with metastatic potential, postsurgical recurrence, and survival of HCC patients (Miao et al. 2014). Gao et al. conducted a comprehensive proteogenomic characterization of tumor and adjacent liver samples from 159 HCC patients with HBV infection (Gao et al. 2019). Two metabolic enzymes, *PYCR2* and *ADH1A*, were identified to participate in HCC metabolic reprogramming (Gao et al. 2019). Because the upregulation of *PYCR2* or downregulation of *ADH1A* may result in HCC progression, they were also validated as potential prognostic biomarkers (Gao et al. 2019).

Since accurate stratification is essential for clinical decision making (Preisser et al. 2020), different stratification methods applied to cohorts of HCC patients have been developed. Kumardeep et al. proposed a deep learning-based model derived from RNA-seq, miRNA-seq, CpG methylation and clinical data of HCC samples to identify two subgroups with significantly different survival (Chaudhary et al. 2018). It was illuminated that the more aggressive subgroup is associated with *TP53* inactivation mutations and Wnt pathway activation (Chaudhary et al. 2018). Therefore, this risk stratification model may be useful at HCC prognosis prediction as well as therapeutic intervention (Chaudhary et al. 2018). Xiao et al. formed an integration method and used this method for an analysis of mRNA expression data, DNA-methylation data, somatic mutation data, and clinical information of HCC samples (Ouyang et al. 2020). 34 differentially expressed genes (DEGs) were identified, some of them were verified as diagnostic biomarkers for HCC (Ouyang et al. 2020). According to the gene expression data of the aforementioned DEGs, tumor samples were divided into three subtypes that displayed different biological processes (Ouyang et al. 2020). Hence, what they found out might help improve precision medicine regarding HCC (Ouyang et al. 2020).

The advanced molecular biological techniques as well as improving understanding of complex mechanisms of liver cancer has driven the development of precision medicine (Yoo et al. 2018). Yildiz analyzed datasets generated by high-throughput drug screening and genomic and transcriptomic studies on HCC cell lines (Yildiz 2018). He divided HCC cells into two subtypes that responded differently to the same drug treatments (Yildiz 2018). 6 molecular targets were revealed to be

associated with drug sensitivity, which could aid the development of effective molecular therapies (Yildiz 2018). Also, the EGFR/PI3K/AKT/mTOR signaling pathway was believed to play a central role in the regulation of sensitivity and resistance to drug treatments in HCC (Yildiz 2018). Christos et al. utilized a computational approach to explore the novel drug targets in mTOR-driven HCC (Dimitrakopoulos et al. 2021). 74 mediators under the impact of upstream genetic aberrations and changes in miRNA expression were identified, among which YAP1, GRB2, HDAC4, SIRT1, and LIS1 were validated to be dysregulated in human HCC (Dimitrakopoulos et al. 2021). Thus, inhibitors of these mediators may be potentially useful in HCC treatment (Dimitrakopoulos et al. 2021).

5.5 Conclusion

Multiomics clearly has advantages when it comes to translating the biological characteristics of cancer into understandable and clinically interpretable data. The advancement of multiomics research in the context of a specific cancer reveals numerous “invisible” but critical correlations. Multiple biomarkers have a higher specificity than previous single-gene markers, laying the groundwork for future research in this field. The identification of specific markers enables the diagnosis of cancer and subsequent treatment, as well as better stratifying patients and developing more effective and personalized treatment methods.

As mentioned previously, multi-omics methods have been successfully applied to colorectal cancer, liver cancer, and lung cancer, yielding a wealth of biological data. As methods and resources for multi-omics analysis mature, multi-omics research will play an increasingly important role in understanding the pathogenesis of cancer and developing effective treatment measures.

However, there is a growing gap between the ability to integrate, process, and interpret data and the ability to generate large amounts of omics data. The majority of data standardization efforts and development of a central public database of omics data have been abandoned. Simultaneously, the majority of tools for multi-omics integration are insufficiently robust, prone to errors, and are only suitable for advanced users with programming expertise. There is still a long way to go before multi-omics analysis is widely applied and its value is maximized in cancer research.

Acknowledgments We thank all the individuals who have helped us in this study. We acknowledge the valuable work of the many investigators whose published articles we were unable to cite owing to space limitations.

Conflict of Interest

The authors declare that they have no conflict of interest.

Fundings This work was financially supported by the grants from the National Natural Science Foundation of China (31861143051, 31872425, 32002235, 31602008) and Senior Talent Foundation of Jiangsu University (19JDG039).

References

- Aarons CB, Shanmugan S, Bleier JI. Management of malignant colon polyps: current status and controversies. *World J Gastroenterol.* 2014;20:16178–83.
- Ahmed Z. Practicing precision medicine with intelligently integrative clinical and multi-omics data analysis. *Hum Genomics.* 2020;14:35.
- Alberg AJ, Brock MV, Samet JM. Epidemiology of lung cancer: looking to the future. *J Clin Oncol.* 2005;23:3175–85.
- Alberg AJ, Samet JM. Epidemiology of lung cancer. *Chest.* 2003;123:21S–49S.
- Almusawi S, Ahmed M, Nateri AS. Understanding cell-cell communication and signaling in the colorectal cancer microenvironment. *Clin Transl Med.* 2021;11:e308.
- Altenbuchinger M, Weihs A, Quackenbush J, Grabe HJ, Zacharias HU. Gaussian and mixed graphical models as (multi-)omics data analysis tools. *Biochim Biophys Acta Gene Regul Mech.* 2020;1863:194418.
- Antoine M, Vieira T, Fallet V, Hamard C, Duruisseaux M, Cadranel J, Wislez M. Pulmonary sarcomatoid carcinoma. *Ann Pathol.* 2016;36:44–54.
- Ariff B, Lloyd CR, Khan S, Shariff M, Thillainayagam AV, Bansal DS, Khan SA, Taylor-Robinson SD, Lim AKP. Imaging of liver cancer. *World J Gastroenterol.* 2009;15:1289–300.
- Arora M, Kumari S, Singh J, Chopra A, Chauhan SS. PAXX, not NHEJ1 is an independent prognosticator in colon cancer. *Front Mol Biosci.* 2020;7:584053.
- Asada K, Kobayashi K, Joutard S, Tubaki M, Takahashi S, Takasawa K, Komatsu M, Kaneko S, Sese J, Hamamoto R. Uncovering prognosis-related genes and pathways by multi-omics analysis in lung cancer. *Biomol Ther.* 2020;10
- Ayiomamitis GD, Notas G, Vasilakaki T, Tsavari A, Vederaki S, Theodosopoulos T, Kouroumalis E, Zaravinos A. Understanding the interplay between COX-2 and hTERT in colorectal cancer using a multi-omics analysis. *Cancers (Basel).* 2019;11:1536.
- Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehar J, Kryukov GV, Sonkin D, Reddy A, Liu M, Murray L, Berger MF, Monahan JE, Morais P, Meltzer J, Korejwa A, Jane-Valbuena J, Mapa FA, Thibault J, Bric-Furlong E, Raman P, Shipway A, Engels IH, Cheng J, Yu GK, Yu J, Aspesi P Jr, De Silva M, Jagtap K, Jones MD, Wang L, Hatton C, Palesscandolo E, Gupta S, Mahan S, Sougnez C, Onofrio RC, Liefeld T, Macconail L, Winckler W, Reich M, Li N, Mesirov JP, Gabriel SB, Getz G, Ardlie K, Chan V, Myer VE, Weber BL, Porter J, Warmuth M, Finan P, Harris JL, Meyerson M, Golub TR, Morrissey MP, Sellers WR, Schlegel R, Garraway LA. The cancer cell line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature.* 2012;483:603–7.
- Barta JA, Powell CA, Wisnivesky JP. Global epidemiology of lung cancer. *Ann Glob Health.* 2019;85:8.
- Behjati S, Tarpey PS. What is next generation sequencing? *Arch Dis Child Educ Pract Ed.* 2013;98:236–8.
- Berg KCG, Eide PW, Eilertsen IA, Johannessen B, Bruun J, Danielsen SA, Bjørnslett M, Meza-Zepeda LA, Eknaes M, Lind GE, Myklebost O, Skotheim RI, Sveen A, Lothe RA. Multi-omics of 34 colorectal cancer cell lines - a resource for biomedical studies. *Mol Cancer.* 2017;16:116.
- Bishayee A "The Role of Inflammation in Liver Cancer," in *Inflammation and Cancer*, eds. B.B. Aggarwal, B. Sung & S.C. Gupta., 2014 401–435.
- Biswas N, Chakrabarti S. Artificial intelligence (AI)-based systems biology approaches in multi-omics data analysis of cancer. *Front Oncol.* 2020;10:588221.

- Bosch FX, Ribes J, Borrás J. Epidemiology of primary liver cancer. *Semin Liver Dis.* 1999;19:271–85.
- Bosch FX, Ribes J, Diaz M, Cleries R. Primary liver cancer: worldwide incidence and trends. *Gastroenterology.* 2004;127:S5–S16.
- Canzler S, Schor J, Busch W, Schubert K, Rolle-Kampczyk UE, Seitz H, Kamp H, Von Bergen M, Buesen R, Hacker Müller J. Prospects and challenges of multi-omics data integration in toxicology. *Arch Toxicol.* 2020;94:371–88.
- Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal. *Cancer.* 2009;59:366–78.
- Chakraborty S, Hosen MI, Ahmed M, Shekhar HU. Onco-multi-OMICS approach: a new frontier in cancer research. *Biomed Res Int.* 2018;2018:9836256.
- Chang MH. Prevention of hepatitis B virus infection and liver cancer. *Recent Results Cancer Res.* 2014;193:75–95.
- Chaudhary K, Poirion OB, Lu L, Garmire LX. Deep learning-based multi-omics integration robustly predicts survival in liver cancer. *Clin Cancer Res.* 2018;24:1248–59.
- Chauvel C, Novoloaca A, Veyre P, Reynier F, Becker J. Evaluation of integrative clustering methods for the analysis of multi-omics data. *Brief Bioinform.* 2020;21:541–52.
- Chen JG, Zhang SW. Liver cancer epidemic in China: past, present and future. *Semin Cancer Biol.* 2011;21:59–69.
- Chen Z, Fillmore CM, Hammerman PS, Kim CF, Wong KK. Non-small-cell lung cancers: a heterogeneous set of diseases. *Nat Rev Cancer.* 2014;14:535–46.
- Cheng WC, Chung IF, Chen CY, Sun HJ, Fen JJ, Tang WC, Chang TY, Wong TT, Wang HW. DriverDB: an exome sequencing database for cancer driver gene identification. *Nucleic Acids Res.* 2014;42:D1048-1054.
- Chung IF, Chen CY, Su SC, Li CY, Wu KJ, Wang HW, Cheng WC. DriverDBv2: a database for human cancer driver gene research. *Nucleic Acids Res.* 2016;44:D975-979.
- Chyr J, Zhang Z, Chen X, Zhou X. PredTAD: a machine learning framework that models 3D chromatin organization alterations leading to oncogene dysregulation in breast cancer cell lines. *Comput Struct Biotechnol J.* 2021;19:2870–80.
- Cieslik M, Chinnaiyan AM. Global genomics project unravels cancer's complexity at unprecedented scale. *Nature.* 2020;578:39–40.
- Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, Ceccarelli M, Bontempi G, Noushmehr H. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 2016;44:e71.
- Das T, Andrieux G, Ahmed M, Chakraborty S. Integration of online omics-data resources for cancer research. *Front Genet.* 2020;11:578345.
- De Groot PM, Wu CC, Carter BW, Munden RF. The epidemiology of lung cancer. *Transl Lung Cancer Res.* 2018;7:220–33.
- De Sousa VML, Carvalho L. Heterogeneity in lung cancer. *Pathobiology.* 2018;85:96–107.
- Dienstmann R, Vermeulen L, Guinney J, Kopetz S, Tejpar S, Taberero J. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat Rev Cancer.* 2017;17:79–92.
- Dimitrakopoulos C, Hindupur SK, Colombi M, Liko D, Ng CKY, Piscuoglio S, Behr J, Moore AL, Singer J, Ruscheweyh HJ, Matter MS, Mossmann D, Terracciano LM, Hall MN, Beerenwinkel N. Multi-omics data integration reveals novel drug targets in hepatocellular carcinoma. *BMC Genomics.* 2021;22:592.
- Domon B, Aebersold R. Mass spectrometry and protein analysis. *Science.* 2006;312:212–7.
- Du W, Elemento O. Cancer systems biology: embracing complexity to develop better anticancer therapeutic strategies. *Oncogene.* 2015;34:3215–25.
- Eicher T, Kinnebrew G, Patt A, Spencer K, Ying K, Ma Q, Machiraju R, Mathe AEA. Metabolomics and multi-omics integration: a survey of computational methods and resources. *Meta.* 2020;10:202.

- Fearon ER. "Molecular genetics of colorectal cancer," in *Cancer prevention: from the laboratory to the clinic: implications of genetic, molecular, and preventive research*, eds. H.L. Bradlow, M.P. Osborne & U. Veronesi.), 1995 101–110.
- Gao Q, Zhu HW, Dong LQ, Shi WW, Chen R, Song ZJ, Huang C, Li JQ, Dong XW, Zhou YT, Liu Q, Ma LJ, Wang XY, Zhou J, Liu YS, Boja E, Robles AI, Ma WP, Wang P, Li YZ, Ding L, Wen B, Zhang B, Rodriguez H, Gao DM, Zhou H, Fan J. Integrated proteogenomic characterization of HBV-related hepatocellular carcinoma. *Cell*. 2019;179:561.
- Ghaffari S, Hanson C, Schmidt RE, Bouchonville KJ, Offer SM, Sinha S. An integrated multi-omics approach to identify regulatory mechanisms in cancer metastatic processes. *Genome Biol*. 2021;22:19.
- Ghosh D, Bernstein JA, Khurana Hershey GK, Rothenberg ME, Mersha TB. Leveraging multilayered "omics" data for atopic dermatitis: a road map to precision medicine. *Front Immunol*. 2018;9:2727.
- Graw S, Chappell K, Washam CL, Gies A, Bird J, Robeson MS 2nd, Byrum SD. Multi-omics data integration considerations and study design for biological systems and disease. *Mol Omics*. 2021;17:170–85.
- Hao Y, Li D, Xu Y, Ouyang J, Wang Y, Zhang Y, Li B, Xie L, Qin G. Investigation of lipid metabolism dysregulation and the effects on immune microenvironments in pan-cancer using multiple omics data. *BMC Bioinformatics*. 2019;20:195.
- Hausman DM. What is cancer? *Perspect Biol Med*. 2019;62:778–84.
- Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature*. 2018;553:446–54.
- Holowatyj AN, Haffa M, Lin T, Scherer D, Gigic B, Ose J, Warby CA, Himbert C, Abbenhardt-Martin C, Achaintre D, Boehm J, Boucher KM, Gicquiau A, Gsur A, Habermann N, Herpel E, Kauczor HU, Keski-Rahkonen P, Kloor M, Von Knebel-Doeberitz M, Kok DE, Nattenmuller J, Schirmacher P, Schneider M, Schrotz-King P, Simon T, Ueland PM, Viskochil R, Weijenberg MP, Scalbert A, Ulrich A, Bowers LW, Hursting SD, Ulrich CM. Multi-omics analysis reveals adipose-tumor crosstalk in patients with colorectal cancer. *Cancer Prev Res (Phila)*. 2020;13:817–28.
- Hu W, Yang Y, Li X, Huang M, Xu F, Ge W, Zhang S, Zheng S. Multi-omics approach reveals distinct differences in left- and right-sided colon cancer. *Mol Cancer Res*. 2018;16:476–85.
- Hu Z, Bi G, Sui Q, Bian Y, Du Y, Liang J, Li M, Zhan C, Lin Z, Wang Q. Analyses of multi-omics differences between patients with high and low PD1/PDL1 expression in lung squamous cell carcinoma. *Int Immunopharmacol*. 2020;88:106910.
- Huang L, Brunell D, Stephan C, Mancuso J, Yu X, He B, Thompson TC, Zinner R, Kim J, Davies P, Wong STC. Driver network as a biomarker: systematic integration and network modeling of multi-omics data to derive driver signaling pathways for drug combination prediction. *Bioinformatics*. 2019;35:3709–17.
- Huang S, Chaudhary K, Garmire LX. More is better: recent progress in multi-omics data integration methods. *Front Genet*. 2017;8:84.
- Huang Y, Duanmu J, Liu Y, Yan M, Li T, Jiang Q. Analysis of multi-omics differences in left-side and right-side colon cancer. *PeerJ*. 2021;9:e11433.
- Jahanafrooz Z, Mosafer J, Akbari M, Hashemzaei M, Mokhtarzadeh A, Baradaran B. Colon cancer therapy by focusing on colon cancer stem cells and their tumor microenvironment. *J Cell Physiol*. 2020;235:4153–66.
- Jones GS, Baldwin DR. Recent advances in the management of lung cancer. *Clin Med (Lond)*. 2018;18:s41–6.
- Kel AE, Stegmaier P, Valeev T, Koschmann J, Poroikov V, Kel-Margoulis OV, Wingender E. Multi-omics "upstream analysis" of regulatory genomic regions helps identifying targets against methotrexate resistance of colon cancer. *EuPA Open Proteom*. 2016;13:1–13.
- Kim TH, Dekker J. CHIP-seq. *Cold Spring Harb Protoc*; 2018.
- Koch A, De Meyer T, Jeschke J, Van Criekinge W. MEXPRESS: visualizing expression, DNA methylation and clinical TCGA data. *BMC Genomics*. 2015;16:636.

- Koch A, Jeschke J, Van Criekinge W, Van Engeland M, De Meyer T. MEXPRESS update 2019. *Nucleic Acids Res.* 2019;47:W561–5.
- Kong Y, Qiao Z, Ren Y, Genchev GZ, Ge M, Xiao H, Zhao H, Lu H. Integrative analysis of membrane proteome and MicroRNA reveals novel lung cancer metastasis biomarkers. *Front Genet.* 2020;11:1023.
- Lam K, Pan K, Linnekamp JF, Medema JP, Kandimalla R. DNA methylation based biomarkers in colorectal cancer: a systematic review. *Biochim Biophys Acta.* 2016;1866:106–20.
- Lee TY, Huang KY, Chuang CH, Lee CY, Chang TH. Incorporating deep learning and multi-omics autoencoding for analysis of lung adenocarcinoma prognostication. *Comput Biol Chem.* 2020;87:107277.
- Li L, Wang H. Heterogeneity of liver cancer and personalized therapy. *Cancer Lett.* 2016;379:191–7.
- Li P, Guo M, Sun B. Integration of multi-omics data to mine cancer-related gene modules. *J Bioinforma Comput Biol.* 2019;17:1950038.
- Lin S, Yin YA, Jiang X, Sahni N, Yi S. Multi-OMICs and genome editing perspectives on liver cancer signaling networks. *Biomed Res Int.* 2016;2016:6186281.
- Liu CY, Chen KF, Chen PJ. Treatment of liver cancer. *Cold Spring Harb Perspect Med.* 2015;5:a021535.
- Liu SH, Shen PC, Chen CY, Hsu AN, Cho YC, Lai YL, Chen FH, Li CY, Wang SC, Chen M, Chung IF, Cheng WC. DriverDBv3: a multi-omics database for cancer driver gene research. *Nucleic Acids Res.* 2020a;48:D863–70.
- Liu X, Qin J, Gao T, Li C, Chen X, Zeng K, Xu M, He B, Pan B, Xu X, Pan Y, Sun H, Xu T, Wang S. Analysis of METTL3 and METTL14 in hepatocellular carcinoma. *Aging (Albany NY).* 2020b;12:21638–59.
- Liu XN, Cui DN, Li YF, Liu YH, Liu G, Liu L. Multiple "omics" data-based biomarker screening for hepatocellular carcinoma diagnosis. *World J Gastroenterol.* 2019;25:4199–212.
- Llabata P, Mitsuishi Y, Choi PS, Cai D, Francis JM, Torres-Diz M, Udeshi ND, Golomb L, Wu Z, Zhou J, Svinkina T, Aguilera-Jimenez E, Liu Y, Carr SA, Sanchez-Cespedes M, Meyerson M, Zhang X. Multi-omics analysis identifies MGA as a negative regulator of the MYC pathway in lung adenocarcinoma. *Mol Cancer Res.* 2020;18:574–84.
- Lu Y, Li Q, Zheng K, Fu C, Jiang C, Zhou D, Xia C, Ma S. Development of a high efficient promoter finding method based on transient transfection. *Gene X.* 2019;2:100008.
- Luan M, Song F, Qu S, Meng X, Ji J, Duan Y, Sun C, Si H, Zhai H. Multi-omics integrative analysis and survival risk model construction of non-small cell lung cancer based on The Cancer Genome Atlas datasets. *Oncol Lett.* 2020;20:58.
- Lv J, Wang J, Shang X, Liu F, Guo S. Survival prediction in patients with colon adenocarcinoma via multi-omics data integration using a deep learning algorithm. *Biosci Rep.* 2020;40:BSR20201482.
- Mantini G, Pham TV, Piersma SR, Jimenez CR. Computational analysis of Phosphoproteomics data in multi-omics cancer studies. *Proteomics.* 2021;21:e1900312.
- Marengo A, Rosso C, Bugianesi E. Liver cancer: connections with obesity, fatty liver, and cirrhosis. *Annu Rev Med.* 2016;67:103–17.
- Mármol I, Sánchez-De-Diego C, Pradilla Dieste A, Cerrada E, Rodríguez Yoldi MJ. Colorectal carcinoma: a general overview and future perspectives in colorectal cancer. *Int J Mol Sci.* 2017;18:197.
- Matthiesen R, Jensen ON. Analysis of mass spectrometry data in proteomics. *Methods Mol Biol.* 2008;453:105–22.
- Mcglynn KA, Tsao L, Hsing AW, Devesa SS, Fraumeni JF. International trends and patterns of primary liver cancer. *Int J Cancer.* 2001;94:290–6.
- Mckillop IH, Schrum LW. Alcohol and liver cancer. *Alcohol.* 2005;35:195–203.
- Meng C, Zeleznik OA, Thallinger GG, Kuster B, Gholami AM, Culhane AC. Dimension reduction techniques for the integrative analysis of multi-omics data. *Brief Bioinform.* 2016;17:628–41.

- Miao R, Luo H, Zhou H, Li G, Bu D, Yang X, Zhao X, Zhang H, Liu S, Zhong Y, Zou Z, Zhao Y, Yu K, He L, Sang X, Zhong S, Huang J, Wu Y, Mihsad RA, Robson SC, Jiang C, Zhao Y, Zhao H. Identification of prognostic biomarkers in hepatitis B virus-related hepatocellular carcinoma and stratification by integrative multi-omics analysis. *J Hepatol.* 2014;61:840–9.
- Minna JD, Roth JA, Gazdar AF. Focus on lung cancer. *Cancer Cell.* 2002;1:49–52.
- Nakano D, Kawaguchi T, Iwamoto H, Hayakawa M, Koga H, Torimura T. Effects of canagliflozin on growth and metabolic reprogramming in hepatocellular carcinoma cells: multi-omics analysis of metabolomics and absolute quantification proteomics (iMPAQT). *PLoS One.* 2020;15:e0232283.
- Netanel D, Stern N, Laufer I, Shamir R. PROMO: an interactive tool for analyzing clinically-labeled multi-omic cancer datasets. *BMC Bioinformatics.* 2019;20:732.
- Nicora G, Vitali F, Dagliati A, Geifman N, Bellazzi R. Integrated multi-omics analyses in oncology: a review of machine learning methods and tools. *Front Oncol.* 2020;10:1030.
- Nusinow DP, Szpyt J, Ghandi M, Rose CM, McDonald ER 3rd, Kalocsay M, Jane-Valbuena J, Gelfand E, Schweppe DK, Jedrychowski M, Golji J, Porter DA, Rejtar T, Wang YK, Kryukov GV, Stegmeier F, Erickson BK, Garraway LA, Sellers WR, Gygi SP. Quantitative proteomics of the cancer cell line Encyclopedia. *Cell.* 2020;180(387–402):e316.
- O'Connell JB, Maggard MA, Livingston EH, Yo CK. Colorectal cancer in the young. *Am J Surg.* 2004;187:343–8.
- Ouyang X, Fan Q, Ling G, Shi Y, Hu F. Identification of diagnostic biomarkers and subtypes of liver hepatocellular carcinoma by multi-omics data analysis. *Genes (Basel).* 2020;11:1051.
- Park M, Kim D, Moon K, Park T. Integrative analysis of multi-omics data based on blockwise sparse principal components. *Int J Mol Sci.* 2020;21:8202.
- Pastushenko I, Blanpain C. EMT transition states during tumor progression and metastasis. *Trends Cell Biol.* 2019;29:212–26.
- Patz EF, Goodman PC, Bepler G. Current concepts—screening for lung cancer. *N Engl J Med.* 2020;343:1627–33.
- Potter JD. Colorectal cancer: molecules and populations. *J Natl Cancer Inst.* 1999;91:916–32.
- Preisser F, Cooperberg MR, Crook J, Feng F, Graefen M, Karakiewicz PI, Klotz L, Montironi R, Nguyen PL, D'Amico AV. Intermediate-risk prostate cancer: stratification and management. *Eur Urol Oncol.* 2020;3:270–80.
- Rhee EP. How omics data can be used in nephrology. *Am J Kidney Dis.* 2018;72:129–35.
- Salgia R, Skarin AT. Molecular abnormalities in lung cancer. *J Clin Oncol.* 1998;16:1207–17.
- Shen H, Yang J, Huang Q, Jiang MJ, Tan YN, Fu JF, Zhu LZ, Fang XF, Yuan Y. Different treatment strategies and molecular features between right-sided and left-sided colon cancers. *World J Gastroenterol.* 2015;21:6470–8.
- Shen M, Xu M, Zhong F, Crist MC, Prior AB, Yang K, Allaire DM, Choueiry F, Zhu J, Shi H. A multi-omics study revealing the metabolic effects of Estrogen in liver cancer cells HepG2. *Cell.* 2021a;10:455.
- Shen Y, Xiong W, Gu Q, Zhang Q, Yue J, Liu C, Wang D. Multi-omics integrative analysis uncovers molecular subtypes and mRNAs as therapeutic targets for liver cancer. *Front Med (Lausanne).* 2021b;8:654635.
- Sia D, Villanueva A, Friedman SL, Llovet JM. Liver cancer cell of origin, molecular class, and effects on patient prognosis. *Gastroenterology.* 2017;152:745–61.
- Siegel R, Desantis C, Jemal A. Colorectal cancer statistics, 2014. *CA-a Cancer J Clin.* 2014;64:104–17.
- Song J, Yang J, Lin R, Cai X, Zheng L, Chen Y. Molecular heterogeneity of guanine nucleotide binding-protein gamma subunit 4 in left- and right-sided colon cancer. *Oncol Lett.* 2020;20:334.
- Spiro SG, Silvestri GA. One hundred years of lung cancer. *Am J Respir Crit Care Med.* 2005;172:523–9.
- Srivatanakul P, Sriplung H, Deerasamee S. Epidemiology of liver cancer: an overview. *Asian Pac J Cancer Prev.* 2004;5:118–25.
- Stintzing S. Management of colorectal cancer. *F1000Prime Rep.* 2014;6:108.

- Subramanian I, Verma S, Kumar S, Jere A, Anamika K. Multi-omics data integration, interpretation, and its application. *Bioinform Biol Insights*. 2020;14:1–24.
- Sun YV, Hu YJ. Integrative analysis of multi-omics data for discovery and functional studies of complex human diseases. *Adv Genet*. 2016;93:147–90.
- Sung H, Ferlay J, Siegel RL, Laversanne M, and Bray, F.J.C.a.C.J.F.C.. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. 2021, 71.
- Tomczak K, Czerwinska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)*. 2015;19:A68-77.
- Tong D, Tian Y, Zhou T, Ye Q, Li J, Ding K, Li J. Improving prediction performance of colon cancer prognosis based on the integration of clinical and multi-omics data. *BMC Med Inform Decis Mak*. 2020;20:22.
- Torre LA, Siegel RL, Jemal A. Lung cancer statistics. *Adv Exp Med Biol*. 2016a;893:1–19.
- Torre LA, Siegel RL, Ward EM, Jemal A. Global cancer incidence and mortality rates and trends—an update. *Cancer Epidemiol Biomark Prev*. 2016b;25:16–27.
- Valencia AM, Kadoch C. Chromatin regulatory mechanisms and therapeutic opportunities in cancer. *Nat Cell Biol*. 2019;21:152–61.
- Van Meerbeeck JP, Fennell DA, De Ruyscher DKM. Small-cell lung cancer. *Lancet*. 2011;378:1741–55.
- Vasaikar SV, Straub P, Wang J, Zhang B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res*. 2018;46:D956–63.
- Wang C, Yang F, Chen T, Dong Q, Zhao Z, Liu Y, Chen B, Liang H, Yang H, Gu Y. RHPCG: a database of the regulation of the hippo pathway in cancer genome. *Database (Oxford)*. 2019;2019:baz135.
- Wang Z, Jensen MA, Zenklusen JC. A practical guide to the cancer genome atlas (TCGA). *Methods Mol Biol*. 2016;1418:111–41.
- Wang Z, Wei Y, Zhang R, Su L, Gogarten SM, Liu G, Brennan P, Field JK, McKay JD, Lissowska J, Swiatkowska B, Janout V, Bolca C, Kontic M, Scelo G, Zaridze D, Laurie CC, Doheny KF, Pugh EK, Marosy BA, Hetrick KN, Xiao X, Pikielny C, Hung RJ, Amos CI, Lin X, Christiani DC. Multi-omics analysis reveals a HIF network and hub gene EPAS1 associated with lung adenocarcinoma. *EBioMedicine*. 2018;32:93–101.
- Weinberg DS, Schoen RE. Screening for colorectal cancer. *Ann Intern Med*. 2014;160
- Weinberg RA. How cancer arises. *Sci Am*. 1996;275:62–70.
- Wistuba I, Gazdar AF. Lung cancer preneoplasia. *Annu Rev Pathol*. 2006;1:331–48.
- Woo HG, Choi JH, Yoon S, Jee BA, Cho EJ, Lee JH, Yu SJ, Yoon JH, Yi NJ, Lee KW, Suh KS, Kim YJ. Integrative analysis of genomic and epigenomic regulation of the transcriptome in liver cancer. *Nat Commun*. 2017;8:839.
- Wu B, Wang Z, Lin N, Yan XB, Lv ZC, Ying ZM, Ye ZM. A panel of eight mRNA signatures improves prognosis prediction of osteosarcoma patients. *Medicine*. 2021a;100:e24118.
- Wu L, Liu F, Cai H. IOAT: an interactive tool for statistical analysis of omics data and clinical data. *BMC Bioinformatics*. 2021b;22:326.
- Wu Y, Yang Y, Gu H, Tao B, Zhang E, Wei J, Wang Z, Liu A, Sun R, Chen M, Fan Y, Mao R. Multi-omics analysis reveals the functional transcription and potential translation of enhancers. *Int J Cancer*. 2020;147:2210–24.
- Xie B, Yuan Z, Yang Y, Sun Z, Zhou S, Fang X. MOBCdb: a comprehensive database integrating multi-omics data on breast cancer for precision medicine. *Breast Cancer Res Treat*. 2018;169:625–32.
- Xie Q, Fan F, Wei W, Liu Y, Xu Z, Zhai L, Qi Y, Ye B, Zhang Y, Basu S, Zhao Z, Wu J, Xu P. Multi-omics analyses reveal metabolic alterations regulated by hepatitis B virus core protein in hepatocellular carcinoma cells. *Sci Rep*. 2017;7:41089.
- Xie Y. Hepatitis B virus-associated hepatocellular carcinoma. *Adv Exp Med Biol*. 2017;1018:11–21.

- Xu Q, Zhai JC, Huo CQ, Li Y, Dong XJ, Li DF, Huang RD, Shen C, Chang YJ, Zeng XL, Meng FL, Yang F, Zhang WL, Zhang SN, Zhou YM, Zhang Z. OncoPDSS: an evidence-based clinical decision support system for oncology pharmacotherapy at the individual level. *BMC Cancer*. 2020a;20:740.
- Xu X, Gong C, Wang Y, Hu Y, Liu H, Fang Z. Multi-omics analysis to identify driving factors in colorectal cancer. *Epigenomics*. 2020b;12:1633–50.
- Xu Y, She Y, Li Y, Li H, Jia Z, Jiang G, Liang L, Duan L. Multi-omics analysis at epigenomics and transcriptomics levels reveals prognostic subtypes of lung squamous cell carcinoma. *Biomed Pharmacother*. 2020c;125:109859.
- Yamashita T, Wang XW. Cancer stem cells in the development of liver cancer. *J Clin Investig*. 2013;123:1911–8.
- Yang H, Jin W, Liu H, Wang X, Wu J, Gan D, Cui C, Han Y, Han C, Wang Z. A novel prognostic model based on multi-omics features predicts the prognosis of colon cancer patients. *Mol Genet Genomic Med*. 2020a;8:e1255.
- Yang Z, Xu J, Li L, Li R, Wang Y, Tian Y, Guo W, Wang Z, Tan F, Ying J, Jiao Y, Gao S, Wang J, Gao Y, He J. Integrated molecular characterization reveals potential therapeutic strategies for pulmonary sarcomatoid carcinoma. *Nat Commun*. 2020b;11:4878.
- Yi H, Li G, Long Y, Liang W, Cui H, Zhang B, Tan Y, Li Y, Shen L, Deng D, Tang Y, Mao C, Tian S, Cai Y, Zhu Q, Hu Y, Chen W, Fang L. Integrative multi-omics analysis of a colon cancer cell line with heterogeneous Wnt activity revealed RUNX2 as an epigenetic regulator of EMT. *Oncogene*. 2020;39:5152–64.
- Yi T, Zhang Y, Ng DM, Xi Y, Ye M, Cen L, Li J, Fan X, Li Y, Hu S, Rong H, Xie Y, Zhao G, Chen L, Chen C, Ni S, Mi J, Dai X, Liao Q. Regulatory network analysis of mutated genes based on multi-omics data reveals the exclusive features in tumor immune microenvironment between left-sided and right-sided colon cancer. *Front Oncol*. 2021;11:685515.
- Yildiz G. Integrated multi-omics data analysis identifying novel drug sensitivity-associated molecular targets of hepatocellular carcinoma cells. *Oncol Lett*. 2018;16:113–22.
- Yin Z, Yan X, Wang Q, Deng Z, Tang K, Cao Z, Qiu T. Detecting prognosis risk biomarkers for colon cancer through multi-omics-based prognostic analysis and target regulation simulation Modeling. *Front Genet*. 2020;11:524.
- Yoo BC, Kim KH, Woo SM, Myung JK. Clinical multi-omics strategies for the effective cancer management. *J Proteome*. 2018;188:97–106.
- Yu C, Qi X, Lin Y, Li Y, Shen B. iODA: an integrated tool for analysis of cancer pathway consistency from heterogeneous multi-omics data. *J Biomed Inform*. 2020;112:103605.
- Yuan Y, Bao J, Chen Z, Villanueva AD, Wen W, Wang F, Zhao D, Fu X, Cai Q, Long J, Shu XO, Zheng D, Moreno V, Zheng W, Lin W, Guo X. Multi-omics analysis to identify susceptibility genes for colorectal cancer. *Hum Mol Genet*. 2021;30:321–30.
- Zeng D, Ye Z, Shen R, Yu G, Wu J, Xiong Y, Zhou R, Qiu W, Huang N, Sun L, Li X, Bin J, Liao Y, Shi M, Liao W. IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. *Front Immunol*. 2021;12:687975.
- Zhang H, Jin Z, Cheng L, Zhang B. Integrative analysis of methylation and gene expression in lung adenocarcinoma and squamous cell lung carcinoma. *Front Bioeng Biotechnol*. 2020a;8:3.
- Zhang J, Bajari R, Andric D, Gerthoffert F, Lepsa A, Nahal-Bose H, Stein LD, Ferretti V. The international cancer genome consortium data portal. *Nat Biotechnol*. 2019;37:367–9.
- Zhang Y, Yang M, Ng DM, Haleem M, Yi T, Hu S, Zhu H, Zhao G, Liao Q. Multi-omics data analyses construct TME and identify the immune-related prognosis signatures in human LUAD. *Mol Ther Nucleic Acids*. 2020b;21:860–73.
- Zochbauer-Muller S, Minna JD. The biology of lung cancer including potential clinical applications. *Chest Surg Clin N Am*. 2000;10:691–708.
- Zou Y, Ruan S, Jin L, Chen Z, Han H, Zhang Y, Jian Z, Lin Y, Shi N, Jin H. CDK1, CCNB1, and CCNB2 are prognostic biomarkers and correlated with immune infiltration in hepatocellular carcinoma. *Med Sci Monit*. 2020;26:e925289.