Multi-omics data integration in upper gastrointestinal cancers research: A review of concepts, approaches, and application

Radman Mazloomnejad  
Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Centre, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran.  
https://orcid.org/0000-0001-8888-299X

Amin Ahmadi  
Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Centre, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran.  
https://orcid.org/0000-0003-0562-9708

Moein Piroozkhah  
Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Centre, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran.  
https://orcid.org/0000-0003-3061-388X

Ayda Zhian Omranian  
Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Centre, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Mohammad Reza Zabihi  
Department of Immunology, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran  
https://orcid.org/0000-0002-9357-2942

Ehsan Nazemalhosseini-Mojarad  
Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Kaveh Kavousi (kkavousi@ut.ac.ir)  
Laboratory of Complex Biological Systems and Bio-informatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Iran

Zahra Salehi (zahra.salehi6463@yahoo.com)  
Laboratory of Complex Biological Systems and Bio-informatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Iran  
https://orcid.org/0000-0002-0839-2729

Systematic Review
Abstract

Upper gastrointestinal (GI) cancers, including esophageal, gastric, liver, and pancreatic cancers, are a major medical and economic burden worldwide. Despite significant advances in radiotherapy, chemotherapy, and targeted treatments for upper GI cancers in the past decade, a high recurrence rate and poor prognosis are still challenging in upper GI cancer management. This trouble is rooted in the current diagnosis methods and the lack of adequate and reliable diagnostic/prognostic biomarkers. The diagnosis of almost every disease of the upper GI tract still depends on invasive investigations such as endoscopy of the upper GI tract, manometry of the stomach and esophagus, or radiography. Although cancer was considered a single disease in the organ of origin in the past, today, it is accepted that cancer is a heterogeneous disease assuming the same organ of origin. Therefore, to conduct precision/personalized medicine, it seems necessary to have suitable biomarkers to make an accurate diagnosis, appropriate patient classification, prognosis assessment, and drug response in cancers. Systems biology and multi-omics research are strategies adopted to provide genetic and molecular biomarkers in cancer. Toward studying complex biological processes, multi-omics data analysis provides an opportunity to gain a deeper and more comprehensive understanding of cancer development and progression. Multi-omics approaches are new frameworks that integrate omics datasets, including genome, epigenome, transcriptome, proteome, metabolome, and metagenome, on the same set of samples to understand cancer's molecular and clinical characteristics better. Therefore, in this review, we focus on the integrated multi-omics studies conducted on esophageal, gastric, liver, and pancreatic cancers and discuss the results regarding diagnostic and prognostic biomarkers, as well as biomarkers that determine the response to treatment.

Key Points

- Emerging precision medicine initiatives in cancer have raised many hopes for declining cancer morbidity and mortality.
- Systems biology approaches, as a quintessential part of precision medicine, encourage understanding of biochemical networks, the discovery of biomarkers, and patient stratification based on unique genomic (e.g., transcriptomic, epigenomic, proteomic, metabolic, metagenomic) and non-genomic profiles.
- Complete access to electronic health records, patient genotypes, microbial and molecular data of patients is one of the requirements for bioinformaticians to apply systems biology in cancers and achieve precision medicine.
- Bringing system biology studies to the clinic and conducting as many clinical trials as possible based on the multi-omics, multi-modal imaging, and clinical demographic data of patients are necessary to achieve precision medicine in complex disease.

Introduction
Global burden of upper GI cancers and research efforts

Cancers of the upper gastrointestinal (GI) tract, including the esophagus, stomach, liver, bile duct, gallbladder, and pancreas, are among the main causes of cancer-related deaths worldwide (1). In 2020, more than 604,100 patients died from esophageal cancer, more than 466,003 from pancreatic cancer, and more than 830,180 from hepatocellular carcinoma, which is rising annually (2). Therefore, early diagnosis and proper treatment of malignant tumors in the GI tract are substantial and challenging issues in these types of cancers (3). Remarkably, GI cancers are often clinically silent in their development (4). So often, the emergence of symptoms and discovery of the disease takes place when the disease has already reached an advanced stage (5, 6). In addition, in the pathogenesis of GI cancers, micro-metastasis is formed in the early stages of development, which is another obstacle in the treatment and prognosis of these cancers. So that only about 10–20% of patients with pancreatic and stomach cancer are candidates for surgery at the time of diagnosis (7, 8). Hence, it leads to a high local and systemic recurrence rate within five years after surgical removal of solid tumors, which is as high as 80–90% for pancreatic cancer and 60–70% for gastric cancer (9–11). Therefore, it seems essential to utilize innovative methods in the direction of accurate diagnosis and prognosis estimation and prescribing the best treatment regimen tailored to each patient to reduce the morbidity and mortality of patients.

Currently, our understanding of cancer is improving and advancing dramatically. However, by comprehending more and more about cancer, more complications of this disease are revealed to us. In this regard, it seems necessary to understand the physiological and pathophysiological processes at the cellular and molecular levels to overcome these complications and address the disease (12, 13). These processes, which operate in a dynamic and interconnected manner, are driven by genetic and metabolic events in cancer cells and the microenvironment they interact with (14). As observed in many types of cancer, there has been a significant heterogeneity of genetic and molecular changes in the patient population that affect the classification, prognosis, and treatment responses (15, 16). Consequently, personalized and precision medicine in cancer treatment, where diagnostic, prognostic, and therapeutic features are selected according to specific molecular deregulation of patients, seems necessary and influential for the most promising disease management and treatment efficacy (17). In aiming to achieve precision medicine, there are significant challenges in interpreting and discovering cellular systems at multiple levels (17). Accordingly, Systems biology has been developed and tries to harness some of these cellular complexities by employing mathematical and computational science techniques and plays a prominent role in the implementation of precision and personalized cancer medicine (18, 19).

Systems Biology and its use in research of upper GI cancers

Systems biology, by investigating cellular-molecular interactions in biological systems and evaluating the characteristics that emerge from them, attempts to construct a more comprehensive and precise view of the development and progress of diseases, including cancer (20, 21). So, the use of systems biology approaches in cancer help to understand how the complex deregulations correlated to cancer coordinately shape malignant states and phenotypes (20). In this regard, systems biology studies use
omics data to discover biomarkers and reveal the pathophysiology of diseases. Omics technologies are high-throughput biochemical assays that facilitate the comprehensive and simultaneous measurement of molecules of one type from a biological sample. Genome, epigenome, transcriptome, proteome, and metabolome analysis are among the accomplishments of this technology, which have been obtained by measuring all or almost all samples of the target molecular space. The use of omics approaches in investigations has assisted a lot in comprehending the functional mechanism of complex biological systems and discovering the underlying molecular signatures of complicated cell phenotypes (22). In this regard, single-omics studies focused on measuring one type of data (e.g., genome) were initially used. Since the systems' biological behaviors result from the interactions of different molecular-cellular layers, single omics studies faced many obstacles in their understanding. Hence, recently, with the progress made in this field, multi-omics data integration have been employed as the weapon of choice to describe aberrant cellular functions in multifactorial diseases such as cancer (20, 22).

Multi-omics data integration for upper GI cancers

Multi-omics, known as pan-omics, is the cross-integration of omics dimensions performed systematically across samples and can be employed for more profound systems biology analyses to confine the origins, relationships, and impacts of biological processes (23). It is longitudinal mainly in design, and they have broad applications and potential for use in clinical research. Multi-omics approaches have been applied in various fields, including cardiovascular diseases, autoimmune diseases, pregnancy and obstetrics, and cancers (24–27). Several integration approaches, including conceptual, statistical, and model-based, have been used to integrate omics data and gain biological insight. Applications of multi-omics to date include biomarker discovery, patient classification, identification of therapeutic targets, characterization of microbial dynamics, and host-microbiome interactions (28, 29). In recent years, omics-based data integration (integrated omics, pan-omics, and trans-omics) has been widely used in cancer research. These studies have successfully identified multiple mutations, gene expression differences, epigenetic alterations, protein abundance differences, and metabolite concentrations associated with cancer heterogeneity and staging (28). Hence, they have significantly revolutionized our understanding of various cancers' etiology and mechanisms. Undoubtedly, integrated-omics studies give a chance to accelerate diagnosis, estimate prognosis more accurately, and prescribe the most effective and personalized therapy in patients with upper GI cancers (30).

In this state-of-the-art review, we investigated the studies generating or evaluating multi-omics data in human upper GI cancers, including esophageal, gastric, liver, and pancreatic cancer. The following sections are organized based on the type of upper GI cancer. Then, we shall discuss the studies that used multi-view data for subtyping, prognosis, and diagnosis of the cancers. We hope that the current review could provide a comprehensive overview of available integrated-omics studies as well as the pitfalls in upper GI tumors. We recommend interested readers see Box 1 to better understand the terms and concepts frequently used in the present review.

**Box 1. Terms and concepts frequently used in the present review.**
<table>
<thead>
<tr>
<th>Terms</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision Medicine</td>
<td>Precision medicine, known as &quot;personalized medicine&quot;, is an innovative strategy for tailoring disease prevention and treatment that considers differences in people's genes, environments, and lifestyles. Precision medicine aims to target the right treatments to the right patients at the right time.</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Identifying a disease, condition, or injury from its signs and symptoms. A health history, physical exam, and tests, such as blood tests, imaging tests, and biopsies, may aid in making a diagnosis.</td>
</tr>
<tr>
<td>Prognosis</td>
<td>The possible outcome or course of a disease; the prospect of recovery or recurrence.</td>
</tr>
<tr>
<td>Disease free survival</td>
<td>The length of time after primary treatment for a cancer ends that the patient survives without any signs or symptoms of that cancer.</td>
</tr>
<tr>
<td>Overall Survival</td>
<td>The length of time from either the date of diagnosis or the start of treatment for a disease, such as cancer, that patients diagnosed with the disease are still alive.</td>
</tr>
<tr>
<td>Biomarker</td>
<td>A biological molecule discovered in blood, other body fluids, or tissues is a character of a normal or abnormal operation or a condition or disease. A biomarker may be utilized to see how well the body responds to a treatment for a disease or condition. Also named a molecular marker and signature molecule.</td>
</tr>
<tr>
<td>Systems Biology</td>
<td>Systems biology is an integrated research approach to investigating complex biological systems by exploring interactions and networks at the molecular, cellular, community, and ecosystem levels.</td>
</tr>
<tr>
<td>Multi-Omics Data Integration</td>
<td>Integrated approaches combine individual omics data sequentially or simultaneously to understand the interplay of molecules.</td>
</tr>
<tr>
<td>Machine Learning</td>
<td>Machine learning is a class of artificial intelligence that permits software applications to become more precise at predicting consequences without being explicitly programmed. Machine learning algorithms utilize historical data as input to predict new output values.</td>
</tr>
<tr>
<td>Deep Learning</td>
<td>Deep learning is a subset of machine learning, essentially a neural network with three or more layers. These neural networks endeavor to simulate the behavior of the human brain—albeit far from matching its ability—allowing it to &quot;learn&quot; from large amounts of data. While a neural network with a single layer can still make approximate predictions, additional hidden layers can help to optimize and refine for accuracy.</td>
</tr>
</tbody>
</table>

**Method**

We used the free-access PubMed search engine to gather studies from MEDLINE, using the search terms “omics”, “multi-omics”, and “integrated omics”, “systems biology” and “Esophageal cancer”, “Gastric cancer”, “Pancreatic cancer”, “Hepatocellular carcinoma”, combined with the Boolean operator AND. This search was conducted in June, July and August 2022 and the obtained studies were reviewed. We did not establish limitations on the type of journal (eg, quartile, impact factor) or the year of publication. Furthermore, we excluded review articles (to avoid duplicate articles), but did use them to identify studies not retrieved by the search engine. We focused on studies that two or more omics were examined. We
also filtered out some individual omics datasets for the candidate disorders. The investigations discussed in this review pursue goals such as recognition and discovery of 'disease pathway', 'diagnostic signature', 'prognostic signature', 'patients' stratification', 'drug discovery', and 'therapeutic response' in esophageal, gastric, liver, and pancreas cancers (Fig. 2.). We must emphasize that there may be other important omics projects not listed here due to limitations in conducting this study.

**Esophageal cancer**

Esophageal cancer, the eighth most common cancer worldwide, is one of the most aggressive GI malignancies. It is estimated that 20,640 new esophageal cancer cases will be diagnosed in the United States in 2022 (31). The prevalence of esophageal cancer is higher in men than in women. So that, the risk of developing esophageal cancer in the United States is estimated to be approximately 1 in 125 for men and approximately 1 in 417 for women (31).

The two leading histological subtypes of esophageal cancer are esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC) (32). It is well comprehended that the incidence of both subtypes varies in geographic regions. ESCC is more common in East Asia, East and Southern Africa, and Southern Europe, while EAC is much more prevalent in other parts of Europe and North America (32, 33). EAC originates from columnar epithelial cells in the lower part of the esophagus (34). However, ESCC arises from the squamous cell lineage of the esophagus through the progression of premalignant precursor lesions that develop in the presence of risk factors that cause chronic irritation and inflammation (35, 36). From the anatomical point of view, ESCC occurs mainly in the middle third of the esophagus and, less frequently, in the lower third. Moreover, there are modifiable and non-modifiable risk factors that are influential in developing ESCC. Genetics, age, and gender, which have been mentioned, are placed in the non-modifiable category, and chewing tobacco, smoking, alcohol consumption, and low fruit and vegetable diet are identified as the modifiable risk factors independently associated with ESCC emergence (35, 36).

The poor prognosis and high mortality rate in esophageal cancer have placed it sixth in cancer-related deaths (37). However, the progress and improvement of treatment approaches in recent years have significantly affected the patients' survival rate. So that during the 1960s and 1970s, the 5-year OS (OS) rate was only about 5%, while today, the 5-year OS rate varies from 15 to 25% worldwide. Since most patients with esophageal cancer are diagnosed in the late stages of clinical course and have a high mortality rate due to its aggressiveness, fast growth rate, and anatomical condition, the discovery of accurate, low-cost, and clinically applicable diagnostic and prognostic molecular signatures has a significant impact on a better understanding of pathophysiology as well as treatment outcomes (37). Considering the complexity and dynamics of the pathogenic process in ESCC, including genetic interaction, epigenetics, gene expression, and proteins, integrative multi-omics studies have contributed significantly to comprehending the disease (30, 38, 39).
Table 2
Multi-omics data integration studies performed in ESCC

<table>
<thead>
<tr>
<th>No</th>
<th>Authors, Year</th>
<th>Disease</th>
<th>Sample</th>
<th>Omics</th>
<th>Finding</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cao W et al., (2020)</td>
<td>ESCC</td>
<td>93 ESCC patients, 202 ESCC samples from TCGA</td>
<td>genomic, transcriptomic, epigenomic, proteomic</td>
<td>Prognostic signature</td>
<td>(40)</td>
</tr>
<tr>
<td>2</td>
<td>Xi Y et al., (2022)</td>
<td>ESCC</td>
<td>91 ESCCC patients, TCGA</td>
<td>genomic, transcriptomic, epigenomic</td>
<td>Diagnostic signature, Prognostic signature</td>
<td>(41)</td>
</tr>
<tr>
<td>3</td>
<td>Jin Y et al., (2020)</td>
<td>ESCC</td>
<td>198 ESCC samples from TCGA</td>
<td>transcriptomic, epigenomic</td>
<td>Prognostic signature, Patients classification</td>
<td>(42)</td>
</tr>
<tr>
<td>4</td>
<td>YU J et al., (2020)</td>
<td>ESCC</td>
<td>202 ESCC samples from TCGA</td>
<td>transcriptomic, epigenomic</td>
<td>Prognostic signature, Patients classification</td>
<td>(43)</td>
</tr>
<tr>
<td>5</td>
<td>Liu G et al., (2019)</td>
<td>ESCC</td>
<td>40 ESCCC patients, TCGA</td>
<td>transcriptomic, epigenomic</td>
<td>Prognostic signature</td>
<td>(44)</td>
</tr>
<tr>
<td>6</td>
<td>Chen Y et al., (2019)</td>
<td>ESCC</td>
<td>15 ESCC patients, 172 ESCC samples from TCGA, 125 ESCC samples from GEO</td>
<td>transcriptomic</td>
<td>Prognostic signature</td>
<td>(45)</td>
</tr>
<tr>
<td>7</td>
<td>Zhou P et al., (2021)</td>
<td>ESCC</td>
<td>159 ESCC samples from TCGA, 179 ESCC samples from GEO</td>
<td>genomic, transcriptomic, epigenomic</td>
<td>Prognostic signature, Patients classification</td>
<td>(46)</td>
</tr>
<tr>
<td>8</td>
<td>Li L et al., (2021)</td>
<td>ESCC</td>
<td>10 ESCC patients, 81 ESCC samples from TCGA</td>
<td>genomic, proteomic</td>
<td>Prognostic signature</td>
<td>(47)</td>
</tr>
<tr>
<td>9</td>
<td>Jin X et al., (2021)</td>
<td>ESCC</td>
<td>(24 + 41 + 100) ESCC patients, 78 ESCC samples from TCGA</td>
<td>transcriptomic, proteomic, metabolomic</td>
<td>Prognostic signature</td>
<td>(48)</td>
</tr>
<tr>
<td>10</td>
<td>Yang H et al., (2022)</td>
<td>ESCC</td>
<td>60 ESCC samples from TCGA</td>
<td>transcriptomic, metagenomic</td>
<td>Prognostic signature, Patients classification</td>
<td>(49)</td>
</tr>
</tbody>
</table>
Multi-omics Data Integration For Evaluation Of ESCC Diagnosis And Prognosis

The unclarity of how the epigenome affects the development of ESCC encourages scientists to investigate epigenome alterations to discover molecular signatures (40). Differentially methylated CpG sites (DMCs) are considered better than other genetic features for cancer diagnosis due to the characteristic of source tissue and type of cancer, early emergence during carcinogenesis, and relative stability in fixed samples over time. The distribution of these DMCs varies across chromosomes. It has been revealed that they are predominantly on chromosome 8 but often absent on chromosome 22. Highly methylated sites are often on chromosomes 18 and 19, while hypo-methylated sites are on chromosome 8 (41). In this regard, Xi Y et al. aimed to determine diagnostic and prognostic ESCC biomarkers by integrating genomic and gene expression data with DMCs of 91 Chinese ESCC patients. Ultimately, by examining 35,577 DMCs and using random forest and LASSO, a panel of 16 DMCs was obtained; 12 diagnostic and 4 prognostic. Of 12 diagnostic DMCs, 8 of them were in AFF3, PDE4D, SYNE3, SLC8A3, CPS1, HOXC10, LDB2 and PACRG genes loci. Except AFF3, high methylation level in them will lead to higher expression level of these protein coding genes. Regarding the remaining four diagnostic DMCs, despite the previous genes, the high methylation of each marker (cg10085326, cg24276395, cg05446471, cg21553182), which are located in the promoters of MMP13, YEATS2, HDAC11, and ZNF578 genes, respectively, will cause lower expression of them. Moreover, four prognostic DMCs including cg23378365, cg06090867 and cg03244277 were in the promoters of CYFIP2, UBXN10, AREG, respectively, and cg02370667 in NECAB2 were significantly associated with patients OS. Based on these prognostic DMCs and their methylation levels, a prognostic model was presented according to which patients were classified into two groups with high or low prognostic risk, in which high-risk patients had a shorter average life expectancy than the other (41).
One of the most fundamental challenges in ESCC is estimating the prognosis and unraveling disease progress. Accordingly, most studies have attempted to discover prognostic biomarkers by employing ESCC multi-omics data integration. To this end, Jin et al. examined a total of 198 samples through DNA methylation and mRNA expression and clinical pathology data. Of 6,261 differentially expressed genes (DEGs), by combining DNA methylation and mRNA expression data 71 methylation-related DEGs were obtained. Using univariate and multivariate COX analysis, a prognostic epigenetic signature, including FAM24B and FAM200A, was determined in ESCC. A risk score was defined based on the epigenetic signature and independent of clinical pathology variables with favorable prognostic ability, based on which patients were classified into two groups of high-risk scores and low-risk scores. Subsequently, a nomogram based on risk score and 3 pathological clinical factors was presented to predict prognosis in patients with ESCC. Examination of this nomogram shows its higher sensitivity and specificity compared to the American Joint Committee on Cancer (AJCC) staging system (42).

Moreover, Yu et al. also used Autoencoder to identify prognostic risk factors and new effective diagnostic methods for early ESCC. To this end, they used early fusion Autoencoder and joint multimodal representation strategies to build an automated encoder and processed a total of 96 ESCCs, including paired RNA and DNA methylation data, from the TCGA database. 16772 genes from RNAseq and 20112 genes from DNA methylation data were achieved. After filtering the properties with CV < 0.1 and univariate Cox pH analysis, the acquired properties were significantly related to the patients' prognosis. Afterwards, ESCCs were divided into two G1 and G2 risk subgroups with different survival based on the properties. A total of 1107 DEGs, including PIWIL2, ZFP57, GPR77, MUC5B, DCC, MUC6, ADAMTS18, FIBCD1, ANXA10, ABCC2, and 199 differentially methylated genes (DMGs) such as ELSPBP1_promoter, REG3G_promoter, PWRN1_promoter, REG1P_promoter, MIR1468_promoter, OR10W1_promoter, OR9I1_promoter, OR2L2_promoter, OR2M4_promoter, and OR2L8_promoter were identified in G2 risk subgroup. Further, KEGG enrichment pathways of the important DEGs were analyzed. Forty-six enrichment pathways, including cytokine-cytokine receptor interaction, cAMP signaling pathway, cell adhesion molecules (CAMs), and PPAR signaling pathway, were identified for the upregulated DEGs, and 24 enrichment pathways, including pathways in cancer, basal cell carcinoma, Ras signaling pathway, transcriptional misregulation in cancer, PI3K-Akt signaling pathway, ECM-receptor interaction, and mTOR signaling pathway were identified for the downregulated DEGs. Most of these KEGG pathways are associated with metastasis and proliferation of various types of cancer and thus may be significant predictors. G2 risk subgroup as an important prognostic factor and independent of any pathological condition, stage, gender, and performed chemotherapy/radiotherapy was significantly associated with patients' prognosis with hazard ratio. Afterward, an SVM model was constructed that showed the robustness and stability of the classification of this study (43).

Another multi-omics study has also designed and conducted a multi-omics study analyzing whole-genome shotgun bisulfite sequencing (WGBS), RNA-seq, and proteome in patients with EC. 295 DMRs were found among ESCC and normal cells that were predominantly hypomethylated. The findings showed a significant relationship between epigenetic alterations and clinical outcomes. The higher variance of DNA methylation change in ESCC cells is associated with poor clinical outcomes and
indicates worse patient conditions. However, a lower variance is related to a much more favorable survival. A higher variance of DNA methylation change in ESCC cells is associated with poor clinical outcomes and indicates a worse prognosis, while a lower variance is related to more favorable survival. Furthermore, 98% of CpGs were shown to be hypomethylated throughout the ESCC genome. Hypomethylated regions are enriched in regions with heterochromatin binding markers (H3K9me3, H3K27me3), whereas hypermethylated regions are enriched in polycomb repressive complex recognition regions (EZH2/SUZ12). Altered methylation in promoters, enhancers, and gene bodies, as well as in polycomb repressive complex occupancy and CTCF binding sites, is associated with cancer-specific gene dysregulation. One of the most hypermethylated promoter regions in the ESCC is WNT2, which unexpectedly results in high expression of WNT2. High expression of this gene activates the WNT2 / β-catenin / MMP pathway and increases the levels of β-catenin, MMP3, and MMP9, causing cancer cells to grow, migrate and invade the extracellular matrix which results in metastasis. Multi-omics data showed that higher expression of WNT2 in patients is a biomarker of tumor progression, and silencing WNT2 suppressed cancer cell growth by reducing the expression of its target genes and reduced cell invasion and migration. Another example of genetic expression disorder due to epigenome involvement in ESCC is the overexpression of ESCCAL-1, which seeks hypomethylation in its promoter and consequently increases the transcription factor binding. Overexpression of ESCCAL-1 / CASC9 has been observed in other types of cancer in addition to ESCC, which causes cancer cells to grow, invade, and metastasize. High expression of ESCCAL-1 in ESCC cells, like high WNT2 level, is associated with disease progression and can be utilized as a biomarker associated with worse clinical outcomes in ESCC (40).

In further investigation, the RNASEq data of 40 patients with esophageal cancer in four stages of cancer and healthy control volunteers from the TCGA database were analyzed using the R-language cluster profiling package. A total of 7,457 genes with different expressions and 13 common DEGs genes were discovered. In the examined samples, CPLX2, DPEP1, EPHA5, SCGB1A1, and ST18 genes were continuously downregulated, and FGF14, KCNH6, LOC100506136, RGS7, SH3GL2, THBS4 genes were continuously upregulated. Differential expression of these genes results in positive regulation of protein transport, gastric acid secretion, insulin-like growth factor receptor binding, and p53 signaling pathway, epidermal growth factor signaling pathway. Further, transcription factors and non-coding RNAs including hypoxia-induced factor 1A and hsa-miR-330-3p were also identified, which significantly regulated gene expression. hsa-miR-330-3p is a regulating factor of the SH3GL2 gene, which has been recognized as an important gene in esophageal cancer pathogenesis. Methylation analysis also showed that the SH3GL2 gene was up-regulated. SH3GL2 gene, one of the shared DEGs of the four-stage time series, may play an essential role in regulating esophageal cancer by methylation alterations. Consequently, one of these prognostic molecules found in esophageal cancer is the G protein subunit gamma transducin 2 (GNGT2). Liu et al. In their multi-omics study of esophageal cancer patients at all four stages showed that GNGT2 was significantly upregulated in esophageal cancer patients and cell lines, enhancing the proliferation of esophageal cancer cell lines and being closely related to esophageal cancer patient's survival (44).

Another research has been done aimed at discovering the prognosis of biomarkers and analyzing potentially misplaced genes in the ESCC samples by examining DNA methylation and mRNA expression
from the RNA-Seq data and WGBS. Overall, 860 DEGs and 16 key DMRs were determined. Whereas VSIG10L, ST6GALNAC1, SCNN1B, PRSS27, PPP1R3C, KRT4, KLK13, KLK11, IL1RN, GPX3, EHD3, and CRABP2 were hypermethylated, hypomethylation was occurred in the SIX4, MFAP2, and COL5A2. It was further found that 11 of these 16 genes (EHD3, IL1RN, KKL13, PRSS27, COL5A2, CRABP2, GPX3, KRT4, MFAP2, SCNN1B, and SIX4) were associated with OS and disease-free survival (DFS) in patients with ESCC. Accordingly, Chen et al. developed and tested a prognostic model (Signature-1) based on these 11 genes. Signature-1 is an independent factor of clinical features, age, gender as well as TNM prognostic factor. They found that Signature-1’s ability to predict the course of the disease and the prognosis of ESCC is significantly better than pTNM (45).

Ma et al. proposed a prognostic model using DNA methylation and copy number variations (CNVs). Since DNA CNVs and epigenetic abnormalities are examples of genomic instability and will predispose patients to malignancy, the esophageal cancer CNVs genes and methylation genes were identified and analyzed from 159 TCGA samples. Finally, patients were divided into three prognostic subgroups, including ic1, ic2, and ic3 based on gene expression, DNA methylation, and CNVs. Patients in subgroup ic1 showed the worst prognosis and patients in subgroup ic3 displayed the best. Studies have shown that although this prognostic classification is independent of TNM, stage, sex, and age due to the association of ic1 with poor prognosis in ESCC, most patients were in the latest stage of the disease. In addition, in the ic1 subgroup that all 6 studied immune cell lines were significantly reduced compared to the others, and the patients of this subgroup are somewhat at the state of immunosuppression. These differences in the tumor immune microenvironment can identify candidates and potential targets for immunotherapy in ESCC patients. Also, univariate survival analysis showed that CLDN3, FAM221A, and GDF15, and YBX2 genes were significantly associated with disease prognosis. Downregulation of these genes within the tumoral samples indicates a better prognosis and disease course. The study identified the lower level of these four prognostic biomarkers within the ic3 subgroup. Also, 61 genes with remarkable higher mutation frequencies in ic1 than in ic3 samples have been found by exploring and analyzing single nucleotide mutations. It suggests that an increase in the frequency of single nucleotide mutations may be associated with a worse course of esophageal cancer in patients (46).

Another multi-omics study investigating proteogenomic properties obtained a carcinogenic pathway with eight cancer-driving waves and genes involved in the development of ESCC with a focus on issues such as DNA damage (e.g. PPP2R5A, IVL, and SERPINB3), cell cycle (e.g. ORC3, RAD21, and FGFR1OP), cell differentiation (e.g. BRCC3, MSH6, and MTOR), cell proliferation (e.g. ERBB2, NOTCH3, and MOV10), metabolism (e.g. ANXA1, CTSB, and HMGCS1), EC (e.g. PGK1, GSK3B, and CTNNB1), lesion invasion (e.g. INSR, IRAK1, and IGFALS), and tumor metastasis (e.g. TK1, VIM, and PIK3R4) (47).

Regarding the high mismatch between mRNAs and proteins, a comprehensive description of transcription, proteomics, phosphoproteomics, and metabolomics on ESCC tissue of adjacent natural tissues paired to identify new molecular vulnerabilities for ESCC, delineate multilayer molecular changes, and perform potential therapeutic goals. Using these data, ESCC-related metabolic pathways and signaling networks were discovered. It was also shown that in the pathogenesis of cancer, specific
pathways are stimulated in RNA transcription, processing, and metabolism in the ESCC. In this regard, a comparison of ESCC and normal adjacent tissues illustrated differences in gene levels, gene isoforms, proteins, phosphoproteins, and metabolite levels. The proteomic data analysis revealed 2890 differentially expressed proteins (DEPs). Of these, only 66 proteins were associated with the risk of disease recurrence or death in ESCC patients. Their alterations in ESCC tissue were analyzed at the corresponding protein and mRNA levels. The expression of proteins is controlled and modulated at different levels including genes, transcription and translation. Consistency of protein-related mRNA expression indicates modulation at the transcriptional level, and inconsistency in protein and mRNA expression levels indicates regulation at the post-transcriptional level. Biological process analysis in ESCC tumor tissues reveals RNA processing and RNA cleavage activities and post-transcriptional regulatory hyperactivity in ESCC tissues. So that, 23 of these 66 proteins showed a coherent expression direction with their respective mRNAs, which indicates that these proteins are modulated at the transcriptional level. While the remaining prognostic proteins showed conflicting expression direction with their respective mRNAs, which implies that these proteins are regulated at the post-transcriptional level. In this regard, the ESCC proteome study has also shown that most proteins are upregulated not only when the corresponding mRNAs are increased, but also when the corresponding mRNAs are unchanged or decreased. Therefore, it is reasonable to hypothesize that Active post-transcriptional and post-translational regulation is a potential oncogenic driver of ESCC.

Further, there were three proteins including XPNPEP3, BPTF and FBL, which were highly and significantly expressed in ESCC tissues and were introduced as prognostic biomarkers in ESCC patients. Among them, FBL protein and its impressive mechanism in ESCC pathogenesis were documented for the first time in this investigation. FBL has been shown to act as a nuclear methyltransferase, stimulate PI3K/AKT signaling, and promote cell cycle progression and growth of ESCC cancer cells. Therefore, FBL was mentioned as a prognostic biomarker in ESCC and can be a potential therapeutic target against ESCC cells. Moreover, SF3A2, CSTF2, RIF1, LSM6, MRPL21 and UBE2A, which act as tumor enhancers and are also up-regulated in ESCC tissues, can be considered and employed as therapeutic markers (48).

By looking at the metagenomics layer, it has been reported that oral and esophageal microbiota may assist in ESCC occurrence (49). Yang et al. performed a multi-omics study and classified ESCC patients into two clusters (A and B) based on ESCC tissue-resident microenvironments and used this method to construct a predictive classification model. This study analyzed data from 60 patients from TCGA and the cancer microbiome atlas (TCMA). Patients were classified based on five types of the richest resident microbiota of esophageal tissue, including Bacteroidetes, Proteobacteria, Firmicutes, Fusobacteria, and Actinobacteria. Thirty-one of these patients, based on a high proportion of Proteobacteria, Firmicutes, and Actinobacteria, and a small proportion of Bacteroidetes and Fusobacteria with relatively better survival, were classified as group I, while another 29 with a low proportion of Proteobacteria, Firmicutes. Moreover, in cluster II, Actinobacteria and a high proportion of Bacteroidetes and Fusobacteria with relatively poor survival were included. The study showed apparent differences in genetic degree and gene expression in the two clusters. The genes with the highest mutation in cluster I were CSMD3, TP53, LAMA1, DNAH5, PCLO, and TENM4, and the genes with the highest mutation in cluster II were TP53, CSMD3, FLG,
DNAAH5, and PIK3CA. A total of 133 DEGs were identified in cluster II (63 upregulated genes and 70 downregulated genes). The study showed that these differences in gene expression were related to tumorigenesis, intraepithelial neoplasia, and differences in microbiota. They even designed a formula based on 10 DEGs, including SNX3, AKIRIN2, TMEM87B, STEAP3, PPME1, LGALS7B, ARFRP1, STX11, RP11-295P9.3, and RP11-434D12.1 to classify ESCC patients into one of the clusters which were established based on the esophageal microbiota. This approach paved the way to identify the potential patients' resident microbiota based on the DEGs, and use it as a predictive formula (49).

**Multi-omics Data Integration For Evaluation Of ESCC Disease Progression, Drug Discovery, And Therapeutic Response**

Several treatments, including surgery, chemotherapy, immunotherapy, and radiotherapy, have been used for ESCC patients. However, the data obtained from the treatment results, as well as the survival rate of the patients, indicate their inadequacy. Therefore, many investigations attempt to explore and identify new therapeutic targets and also discover biomarkers and factors affecting the therapeutic response. In this regard, another multi-omics study aimed to understand and reconstruct transcriptional regulation networks and protein-protein interaction in the ESCC to identify biomarkers and novel therapeutic targets. By examining healthy and ESCC samples, a total of 1582 and 1078 genes were identified as low and high regulation, respectively. 30 down-regulated and 21 up-regulated genes, the so-called DEG core, were differentially expressed in all ESCC patients. These can be considered promising candidate biomarkers and potential drug targets. These genes play a role in different categories. Down-regulated ACPP, C2orf54, DYNLT3, KANK1, ENDOU, FM02 and up-regulated HOMER3, RFC4, COL10A1, FNDC3B and MARCKSL1 were reported for the first time in ESCC. In addition, transcription factors, proteins and metabolites were also discussed in this study. The study of transcription factors representing 11 and 12 transcription factors (TFs) were set up and down, respectively. The most important family of these transcription factors were E2Fs, which were upregulated. It has been also observed that ESCC-related receptors such as nuclear receptors, interleukin receptors, and olfactory receptors were downregulated. However, lysophosphatidic acid receptors, lipoprotein receptors, immunoglobulin receptors, and tumor necrosis factor (TNF) receptors were upregulated. In the next step, the analysis of metabolites in ESCC samples showed the presence of three metabolites, proton (H +) and 10-Hydroxyeicosatetraenoate (10-HETE) in all five datasets. Also, vitamins that play a role as a coenzyme in many biological processes are impaired in the samples, so it was proposed that the reduction of vitamins E and C is one of the influential factors in the course of ESCC (50).

Furthermore, several studies have suggested the role of intertumoral hypoxia as one of the prognostic factors in the course of cancers. Intratumoral hypoxia has been introduced as one of the factors inducing resistance to conventional chemotherapy and radiation therapy in cancer patients. It has been revealed that the presence of hypoxic conditions in the treatment of patients significantly contributed to the emergence of chemical resistance in many cancers, including ESCC, and is ultimately associated with a poor prognosis in cancer patients. Therefore, since hypoxia-stemness coexists in the tumor
microenvironment, it seems efficient to have a prognostic biomarker based on transcriptional expression patterns to evaluate the prognostic status in ESCC. Hence, Tang et al. employed the t-SNE algorithm, LASSO method, and Cox regression model to screen prognostic genes from TCGA ESCC RNA-seq data from 161 patients and 11 controls. Among the 242 DEGs obtained, eight genes related to prognosis (FBLN2, IL17RB, CYP2W1, AMTN, FABP1, FOXA2, GAS1, and CTSF) were chosen to build a risk-scoring system. A risk score formula was developed for each ESCC patient, and the patients were split into low-risk and high-risk groups. Most of the mutation profiles of patients in high-risk and low-risk groups were investigated. The outcomes revealed that the ratio of TP53 mutations in the high-risk group was remarkably higher than in the low-risk group. Studies have demonstrated that tumors with TP53 mutations have a higher hypoxia score in different tumor types, including breast and lung adenocarcinoma, thus supporting the hypothesis that TP53 mutations are a genomic consequence of tumor hypoxia. Further, the evaluation of this risk score in drug response and its relationship with hypoxia was also convincing, so a higher risk score was significantly associated with a lower sensitivity of ESCC patients to chemotherapy drugs, including ceritinib, bexarotene, dasatinib, and imatinib (51).

**Gastric Carcinoma (GC)**

With a mortality rate of over 784000 annually, GC is known as the third-highest cause of malignancy mortality worldwide (52). GC is often presented as an asymptomatic disease. This results in significant delays in diagnosis of the disease, which in turn contributes to more advanced stages and higher mortality (53). Management of late-stage disease is usually performed as palliative treatment using chemotherapy agents and poor outcomes are common (54). The incidence of GC is known to differ between sexes. A male-to-female age-standardized incidence rate of 2.2 in GC translates to a higher risk of GC in male individuals compared to females. Since anti-estrogen drugs increase cancer risk in females, and increased fertility and delayed menopause decrease this risk, estrogen has been suggested as a protective factor for GC (55, 56).

GC is defined as any malignancy that arises from the gastroesophageal junction to the pylorus. 95 percent of GCs originate from epithelial cells. Histologically, this disease can be categorized into intestinal type, which is the more prevalent type and emerges from intestinal metaplasia and has a high correlation with lifestyle, and diffuse type, which is the less prevalent type known to be endemic to specific areas and correlates more with genetic cues (57). Traditionally, H.Pylori has been thought of as the cause of GC. While it still remains one of the major causes of this disease, several other factors including other pathogens as well as genetic and environmental cues have been found to play important roles in GC pathogenesis (53). Interactions between these factors often account for variable clinical outcomes and complex situations (58). TCGA has proposed a molecular subtyping system for GC which divides it into the following categories: microsatellite-instable (MSI), Epstein- Barr-virus-positive (EBV), chromosomal-instable (CIN), genomically-stable (GS), and hypermutated-single-nucleotide-variants (HM-SNV). However, these subtypes share similar characteristics and more precise categorizations are required (59–61). In this section, the recent advances in molecular subtyping of GC based on multi-omics
integration methods are reviewed. Utilization of these methods has provided important insights in several areas of cancer research including subtyping, diagnosis, prognosis, and identifying disease mechanism and therapeutic targets.
<table>
<thead>
<tr>
<th>No</th>
<th>Authors, Year</th>
<th>Disease</th>
<th>Sample</th>
<th>Omics</th>
<th>Finding</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Erawijantari PP et al., (2020)</td>
<td>GC</td>
<td>50 GC patients</td>
<td>metabolomic, metagenomic</td>
<td>Prognostic signature</td>
<td>(62)</td>
</tr>
<tr>
<td>2</td>
<td>Wu X et al., (2022)</td>
<td>GC</td>
<td>GEO datasets, TCGA</td>
<td>genomic, transcriptomic</td>
<td>Prognostic signature</td>
<td>(63)</td>
</tr>
<tr>
<td>3</td>
<td>Yoon SJ et al., (2019)</td>
<td>GC</td>
<td>3951 GC samples from GEO, AGS cell line</td>
<td>transcriptomic, metabolomic, lipidomic</td>
<td>Disease pathway, Prognostic signature</td>
<td>(64)</td>
</tr>
<tr>
<td>4</td>
<td>Liu N et al., (2021)</td>
<td>GC</td>
<td>293 GC sample from TCGA</td>
<td>genomic, transcriptomic</td>
<td>Prognostic signature</td>
<td>(28)</td>
</tr>
<tr>
<td>5</td>
<td>Huang Y-J et al., (2021)</td>
<td>GC</td>
<td>GEO datasets, TCGA</td>
<td>genomic, transcriptomic, epigenomic</td>
<td>Prognostic signature</td>
<td>(65)</td>
</tr>
<tr>
<td>6</td>
<td>Hu X et al., (2021)</td>
<td>GC</td>
<td>1357 GC samples from GEO, and TCGA</td>
<td>genomic, transcriptomic, epigenomic</td>
<td>Prognostic signature</td>
<td>(66)</td>
</tr>
<tr>
<td>7</td>
<td>Chen Y et al., (2021)</td>
<td>GC</td>
<td>407 GC samples from TCGA</td>
<td>genomic, transcriptomic, epigenomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(67)</td>
</tr>
<tr>
<td>8</td>
<td>Song P et al., (2020)</td>
<td>GC</td>
<td>GEO datasets, TCGA</td>
<td>transcriptomic, epigenomic</td>
<td>Prognostic signature</td>
<td>(68)</td>
</tr>
<tr>
<td>9</td>
<td>Saha SK et al., (2019)</td>
<td>GC</td>
<td>236 GC samples from TCGA, GEO datasets</td>
<td>genomic, transcriptomic, epigenomic</td>
<td>Prognostic signature</td>
<td>(69)</td>
</tr>
<tr>
<td>10</td>
<td>Guo C et al., (2022)</td>
<td>GC</td>
<td>839 GC samples from TCGA, GEO datasets</td>
<td>genomic, transcriptomic</td>
<td>Prognostic signature</td>
<td>(70)</td>
</tr>
<tr>
<td>No</td>
<td>Authors, Year</td>
<td>Disease</td>
<td>Sample</td>
<td>Omics</td>
<td>Finding</td>
<td>Ref</td>
</tr>
<tr>
<td>----</td>
<td>---------------</td>
<td>---------</td>
<td>--------</td>
<td>-------</td>
<td>---------</td>
<td>-----</td>
</tr>
<tr>
<td>11</td>
<td>Wang H et al., (2021)</td>
<td>GC</td>
<td>70 GC patients</td>
<td>multidimensional genomic characteristics (mutation, CNV, neoantigen, clonality, and essential genomic alterations)</td>
<td>Prognostic signature</td>
<td>(71)</td>
</tr>
<tr>
<td>12</td>
<td>Chen T et al., Preprint (2021)</td>
<td>GC</td>
<td>361 GC samples from TCGA</td>
<td>transcriptomic, epigenomic</td>
<td>Prognostic signature</td>
<td>(72)</td>
</tr>
<tr>
<td>13</td>
<td>Zhang H et al., (2021)</td>
<td>GC</td>
<td>77 GC patients, 733 GC samples from TCGA, GEO dataset</td>
<td>transcriptomic, epigenomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(73)</td>
</tr>
<tr>
<td>14</td>
<td>Song Y et al., (2017)</td>
<td>GC</td>
<td>4 GC patients</td>
<td>transcriptomic, proteomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(74)</td>
</tr>
<tr>
<td>15</td>
<td>Tanaka Y et al., (2021)</td>
<td>GC</td>
<td>98 GC patients, 440 GC samples from TCGA</td>
<td>genomic, transcriptomic, epigenomic</td>
<td>Prognostic signature</td>
<td>(75)</td>
</tr>
<tr>
<td>16</td>
<td>He Y et al., (2020)</td>
<td>GC</td>
<td>TCGA</td>
<td>genomic, transcriptomic, proteomic</td>
<td>Prognostic signature, Therapeutic response</td>
<td>(76)</td>
</tr>
<tr>
<td>17</td>
<td>Zhu Y et al., (2022)</td>
<td>GC</td>
<td>996 GC samples from TCGA, GEO datasets</td>
<td>genomic, transcriptomic</td>
<td>Prognostic signature, Therapeutic response</td>
<td>(77)</td>
</tr>
<tr>
<td>18</td>
<td>Zeng D et al., (2019)</td>
<td>GC</td>
<td>1524 GC samples from TCGA, GEO datasets</td>
<td>genomic, transcriptomic</td>
<td>Therapeutic response, Patients stratification</td>
<td>(78)</td>
</tr>
<tr>
<td>19</td>
<td>Li Z et al., (2020)</td>
<td>GC</td>
<td>35 GC patients</td>
<td>genomic, transcriptomic</td>
<td>Therapeutic response, Patients stratification</td>
<td>(79)</td>
</tr>
</tbody>
</table>
Multi-omics Data Integration For Evaluation Of GC Prognosis

Due to asymptomatic presentation, GC is generally known as a cancer with a relatively poor prognosis (53). Many multi-omics studies have been done with the aim to discover novel prognostic biomarkers in GC. GC is usually treated by resection of tumors in a procedure called gastrectomy. Gastrectomy can result in changes in the gut microbiome. It has been found that the gut microbiome can affect outcomes of gastrectomy. Patients with GC are at higher risk of metachronous cancers after gastrectomy. This might be due to the changes in the gut microbiome. Erawijantari et al. investigated the effects of gastrectomy in GC patients on the gut microbiome and metabolome. By assessing fecal samples from 50 cancer patients and 56 controls, they found that gastrectomy changes microbial diversity and function. In regards to metabolome, the surgery group had higher levels of secondary bile acids in the fecal samples. On the other hand, the control group had higher levels of primary bile acids. In addition, several amino acids were more enriched in the surgery group. To assess the relationship between metabolome and microbiome, they utilized Bi-clustering of correlations. Briefly, different clusters were found in regard to amino acids and bile acids and the associated microbiome. The altered microbiome in the gastrectomy group is similar to that of metachronous colorectal cancer (CRC) patients. Although the mechanism for metachronous cancer following gastrectomy is not clear, the same pathogens seem to be present in the case of CRC. These results can provide a prognostic tool for GC patients undergone gastrectomy (62). Nonetheless, the results need to be further examined by long-term follow-up of patients and evaluation of outcomes. Helicobacter Pylori (HP) infection by approximately 50% infecting of the world population’s stomachs is renowned worldwide as a leading cause of GC. Wu et al used transcriptomics data from 6
cohorts to develop a prognostic model for GC. 73 HP-related genes were evaluated and two distinct prognostic groups were found. Moreover, an assessment of the tumor microenvironment, which consists of various immune cell types, endothelial cells, cancer cells, fibroblasts, and smooth muscle cells, revealed that the two recognized prognostic groups differed in regard to immune cell types. A higher concentration of B cells was observed in the high-risk group. This could point out the potential for anti-B-cell immunotherapy in this group. The results from this study enable clinicians to have a better understanding of GC (63). Besides HP, EBV is also widely recognized to give rise to GC. EBV-associated GC (EBVaGC) is typically known as a subtype of GC with a good prognosis. However, the molecular mechanism through which EBVaGC progresses and the reasons for the better prognosis of this disease is poorly understood. Yoon et al. examined transcriptomics, metabolomics, and lipidomics of EBVaGC using the data from Gene Expression Omnibus (GEO) and in-vitro analysis of an EBVaGC cell line. This study shows that a total of 58 genes that are associated with metabolic pathways were downregulated in EBVaGC. These pathways include lipid metabolism, amino acid metabolism, and nucleotide synthesis pathways. Moreover, EBVaGC differs from EBV not-associated GC (EBVnGC) in 88 metabolites. Integration of these data revealed that the downregulated genes and metabolites were highly correlated. Several of the genes that were found to differ in expression in EBVaGC in this study which were previously known to be related to better outcomes in GC. Collectively, the altered transcriptomics and metabolomics found by this study may be the mechanism through which EBVaGC has different outcomes from EBVnGC (64). Changes in the genomic characteristics (e.g. mutations) can contribute to an altered prognosis of GC. Liu et al. used data from 293 GC patients extracted from TCGA in order to develop a prognostic model based on genomic variations. By integrating somatic mutation data and copy number alterations (CNAs), they aimed to identify key driver genes as genomic variation molecular markers. They recognized 31 molecular markers of genomic variation which contributed to different prognoses. It was found that these key driver genes promoted programmed cell death. All in all, the results of this study suggest a panel of prognostic key driver genes that cause different outcomes in GC patients (28). Different mutational profiles can contribute to altered prognosis of cancer. Mutation of MUC16, which contains cancer antigen 125 (CA-125), is a type I transmembrane mucin, has previously been found to contribute to a better prognosis in GC. However, the mechanism for this has been unclear. Huang et al. investigated the mechanisms through which MUC16 mutation can result in longer survival of cancer patients using multi-omics approaches. To this end, they used single nucleotide polymorphism (SNP), DNA methylation, and mRNA data from TCGA. They utilized Cox Regression and Random Survival Forest algorithm to analyze the data. They found that MUC16 mutation correlates to lower stromal scores and a higher survival rate. This score can be translated to higher cancer stem cell levels in the mutation group. In addition, an assessment of immune cell infiltration revealed that, unlike the wild-type group, the mutation group had high levels of CD56^{dim}NK cells and activated CD4 + T cells and low levels of plasmacytoid dendritic cells and memory CD4 T cells. Evaluation of gene expression level was also performed and results indicate that NPY1R expression was higher in the poor prognosis group. The ratio of immune cells and NPY1R expression matched in terms of prognosis. Also, more p53 and DNA repair pathways were active in the low expression group. NPY1R was examined as a therapeutic target by using the Genomics of Drug Sensitivity in Cancer (https://www.cancerrxgene.org/). Three drugs were found to
be related to low \textit{NPY1R} expression. These drugs were WZ-184, Roscovitine, and WH-4-023. All in all, it was found that \textit{MUC16} correlated with better prognosis, and p53 activation might be the mechanism for this (65). These results were confirmed by a study by Hu et al. Hu et al. attempted to discover novel subtypes of GC by analyzing five different omics data from TCGA. DNA methylation, gene mutation, mRNA, lncRNA, and miRNA data were utilized and ten clustering algorithms, including iClusterBayes, PINSPlus, CIMLR, IntNMF, ConsensusClustering, 146 COCA, NEMO, moCluster, LRA and SNF were used. They successfully characterized two GC subtypes. The subtypes were distinguishable by survival and the activated carcinogenesis pathways. It was observed that in the first subtype, which was associated with lower survival time, epithelial to mesenchymal transition and extracellular matrix pathways were more prominent. On the other hand, patients with the second subtype had a longer OS time. Moreover, this subtype had a higher number of genetic mutations compared to the first one. Particularly, \textit{TTN, MUC16}, and \textit{ARID1A} mutations were among the ones which were increased in the subtype with the better prognosis. In regard to treatment, an increased number of immune cells was observed in the second subtype which can be translated to an enhanced response to immunotherapy. This subtype was also found to respond to chemotherapy with cisplatin, 5-fluorouracil, and paclitaxel more favorably (66). As mentioned above, p53 activation correlates with a more favorable prognosis. Although, whether or not p53 mutation contributes to altered prognosis was not mentioned. Chen et al. studied the long non-coding RNA (lncRNA) profile of patients with GC. In this study, they gathered 375 tumor samples and 32 samples from sites adjacent to the tumor and identified 1547 novel lncRNAs in the patients. Moreover, based on 1235 lncRNAs, they developed three subtypes with distinct prognostic outcomes. To assess the genomic and epigenomic factors in each subtype, they integrated mutational and methylation profiles into the data. The subtype with the worst prognosis was strongly correlated with \textit{TP53} mutations while the subtype with the best prognosis exhibited \textit{ARID1A, KMT2B, PIK3CA, KRAS, and FBXW7} mutations more frequently. In regard to DNA methylation, higher levels of \textit{TDG, TET1, and TET3}, which are involved in DNA demethylation, were observed in the subtype with the worse prognosis. Collectively, genome-wide hypomethylation and TP53 mutation was found to correlate with higher expression of lncRNAs specific to the subtype with the worst prognosis. These RNAs were usually involved in transcriptional and post-transcriptional regulation of gene expression (67). Song et al. used methylation and lncRNA expression data from TCGA and GEO in order to find novel biomarkers for GC by applying multi-omics approaches. A total of 548 differentially methylated CpG islands and 2399 differentially expressed lncRNAs were identified. Of them, two lncRNAs and three CpG sites were significantly associated with patient survival. This study found several lncRNAs which were modulated by abnormal DNA methylation, to be potential prognostic biomarkers for GC development (68).

Another factor affecting cancer prognosis is altered ion channel activity and expression level. The tweety family of genes (\textit{TYYH}s), which are a group of chloride channel responsive genes, play roles in several cellular functions including proliferation and tumorigenesis. Saha et al. studied the gene and protein expression of \textit{TYYH3} and the relationship between mRNA levels and clinical outcomes in GC. It was found that gene and protein expression of \textit{TYYH3} was elevated in patients in several publicly available datasets. Co-analysis of DNA methylation and \textit{TYYH3} gene expression exhibited lower promoter
methylation and higher TTYH3 expression in cancerous tissue. A total of 8 mutations were found in the patients. Copy number alterations positively correlated with TTYH3 expression in the TCGA database. Finally, they revealed that higher expression levels significantly correlated with a worse prognosis. These results indicate that the altered TTYH3 expression might be regulated through methylation and copy number alteration and higher expression levels result in a worse prognosis. This suggests that TTYH3 can be a potential prognostic biomarker for GC (69). Further, Guo et al. studied the effects of IncRNAs on GC prognosis using data from TCGA and GEO cohorts and integrated them within a robust multiarray averaging (RMA) algorithm. The results from the Cox regression algorithm exhibited that 15 pyroptosis-derived IncRNAs contributed to cancer prognosis. Interestingly, it was found that the low-risk group had a higher number of gene mutation frequencies compared to the high-risk group. Also, the high-risk group showed lower levels of immune infiltration compared to the low-risk group. No significant difference was observed regarding CNVs in the groups (70).

Focusing on molecular classification can also be a proper method to evaluate the prognosis of GC. Wang et al. studied the molecular classification of GC based on multi-dimensional genomic profiles of cancer patients. They obtained biopsies from 70 GC patients and used WES to collect mutational signature, neoantigen, CNV, clonality, and essential genomic alterations data. An unsupervised clustering algorithm was utilized to cluster the patients based on an integrative analysis of the mentioned data. The patients were divided into four groups based on phenotype and prognosis. The patients differed in regard to tumor phenotype (intestinal-type vs. diffuse type) and probability and location of metastasis. Collectively, these results propose a novel classification of GC based on the integration of multi-dimensional genomic characteristics (71). Chen et al. developed a deep learning-based model to evaluate GC survival rate and discover molecular classifications by integrating mRNA and miRNA expression as well as DNA methylation. To do this, they utilized data from 361 GC samples from the TCGA database. The data were used to train an autoencoding network and K-means clustering to generate a survival risk. Moreover, an SVM model was created to assess the differences in subtypes using the results of the previous stage of the study. It was found that eight features possessed the highest weights in terms of decision-making. Interestingly, none of these features belonged to the miRNA data; four belonged to mRNA expression and four belonged to DNA methylation. Therefore, this study provide a valuable deep understanding of GC subtypes and survival rates based on molecular classification (72).

In an attempt to develop a method to distinguish metastatic tumors from non-metastatic ones, Zhang et al. created a prognostic model using 19 gene pairs with reversal relative expression ordering (REO). 77 stage II and III GC patients were selected for this study and the results were confirmed in two other independent cohorts. Based on the mentioned REOs, patients were classified into two high and low-risk groups. Evaluation of the epigenomics characteristics of the group revealed that a total of 2634 genes were hypermethylated in the high-risk group. Also, the assessment of mutational profiles of the 170 samples classified into the two mentioned groups indicated that mutation levels of 4 genes (ADCY3, KCNU1, PPP2R2B, and GNAS) were higher in the low-risk group. Also, the high-metastasis group exhibited hyperactivation of metastasis-related pathways such as focal adhesion, Rap1, and PI3K-Akt. This study provides a method to categorize GC in metastatic and non-metastatic groups, and using multi-omics
methods provides a potential underlying mechanism for that (73). One of the most prominent predictors of GC prognosis is the tendency of the tumor for lymph node metastasis (LNM). The LNM and the number of involved lymph nodes impact prognosis both in the early and late stages of the disease. Although the mechanism of LNM in GC is not thoroughly understood, LNM in many other cancers is affected by protein functions. Protein post-translational modifications (PTMs) greatly influence protein functions. Lysine succinylation is one of the PTMs of proteins which is known to play important roles in cellular metabolic pathways. Recently, lncRNAs have been found to control several PTMs. Moreover, lncRNAs are known to be potential biomarkers for GC. Song et al. explored the proteomics and transcriptomics of GC by assessing IncRNA profiling and lysine succinylome of this disease. They collected one biopsy sample of GC without LNM and 3 samples of patients with LNM. It was found that as cancer progresses, higher expression of proteins with nuclear functions including cell cycle proteins and DNA replication components are present. Also, lysine succinylation was evaluated as a potential biomarker for GC. It was observed that lysine succinylation plays distinct roles in LNM cancer vs non-LNM through alterations in the TCA cycle and pentose phosphate pathway. Moreover, assessment of the relationship between lncRNAs and succinylome identified two lncRNAs with crosstalks with one succinylation site. The results of this study provide important information in regard to GC prognosis and lymph node metastasis (74). GC also has a tendency to metastasize to the peritoneum. Peritoneal metastases are one of the hallmarks of advanced metastatic GC. This complication is especially important in diffuse-type GC and causes high mortality and limited treatment options. In order to study the molecular characteristics of this phenomena, Tanaka et al. collected ascitic fluids from 98 patients and purified the cancer cells. This was followed by comprehensive multi-omics evaluations using RNA-seq, WGS, chromatin immunoprecipitation followed by sequencing (ChIP-seq) and methylation analysis integrated within various computational models to search for novel molecular targets. They found that compared to the primary form of the disease, peritoneal metastatic cells exhibit higher levels of mitogen-activated tyrosine kinase and receptor tyrosine kinase levels. Moreover, they were able to find two distinct subtypes of this disease based on active super-enhancers (SE). One subtype had active SEs at EHF, ELF3, and KLF5 whereas the other had active SE at SMAD3. Transcriptional enhancer factor TEF-1 (TEAD-1) had high expression in the second subtype. It was suggested that through inhibition of TEAD-1, resistance to treatment could be bypassed. Therefore, TEAD-1 inhibition was introduced as a potential molecular target for the management of this disease (75).

## Multi-omics Data Integration For Evaluation Of GC Therapeutic Response

Immunotherapy is becoming an increasingly promising method in cancer treatment. Currently, nivolumab and pembrolizumab are being used to treat microsatellite-instable variants of GI cancers. However, a large proportion of cancer patients fail to benefit from immunotherapy due to genetic and genomic variations. He et al. studied multi-omics characteristics of GC in relation to tumor immunity with the goal of identifying molecular markers enhancing or suppressing tumor immunity. They used mutational profiles, expression and proteomic data and analyzed cancer-associated signaling pathways from the
TCGA database, which could potentially contribute to altered tumor immunity. Several genes were discovered whose mutation (\textit{MTOR}, \textit{RNF213}, \textit{PIK3CA}, \textit{TP53}, \textit{PTEN}, \textit{ATM}, \textit{ARID1A}, and \textit{CDH1}) or expression (\textit{CXCL9}, \textit{CXCR6}, \textit{CXCL13}, \textit{GUCY2C}, \textit{CCL5}, \textit{MAP3K9}, \textit{PAK6}, \textit{STK35}, \textit{NEK3}, and \textit{WNK2}) promoted tumor immune properties. While expression of several proteins exhibited a similar relationship to that of the mentioned genes (caspase-7, \textit{PREX1}, Lck, \textit{PI3K-p85}, transglutaminase, and \textit{Bcl-2}) expression of acetyl-CoA carboxylase showed an opposite pattern (76). Furthermore, a number of miRNAs, IncRNAs, and signaling pathways were recognized which correlated with tumor immunity. These results provide a molecular indication of whether or not a patient might respond to immunotherapy and yield a better strategy for immunotherapy patient stratification (66). Tumor immune microenvironment (TIME), which is usually characterized as the rate of infiltration of cytotoxic immune cells at the tumor site, affects the outcomes of immunotherapy. Zhu et al. used SNPs and RNA-seq data from 352 TCGA GC cases to study TIME phenotypes in this disease by employing k-means clustering and linear discriminant analysis algorithm and assessed the clinical relevance of the subtypes. They were able to classify the patients into three subtypes which varied in survival. The subtype with the more favorable prognosis was found to exhibit higher immune scores, immune cell infiltration, and immune signaling pathways. Additionally, this subtype responded to immunotherapy in a more favorable manner compared to the two others. Using weighted gene co-expression network analysis (WGCNA) resulted in the identification of 14 genes (\textit{DYSF}, \textit{MAN1C1}, \textit{HTRA3}, \textit{EMCN}, \textit{RFLNB}, \textit{KANK3}, \textit{MAGEH1}, \textit{CD93}, \textit{PCAT19}, \textit{FUT11}, \textit{BMP1}, \textit{FOSB}, \textit{DCHS 1}, and \textit{TCF3}) that contributed to the aforementioned subtypes. Five of these genes (\textit{HTRA3}, \textit{EMCN}, \textit{BMP1}, \textit{TCF3}, and \textit{FOSB}) are among the genes which are previously known to play roles in GC genesis (77). Evaluation of transcriptomic data has also been utilized to assess TIME index in GC. Zeng et al. developed a TIME score using transcriptomic data from 1524 patients in 6 cohorts and used unsupervised hierarchical clustering to categorize the patients based on the TME score in 3 subtypes. It was found that TIME scores correlated with immune-therapeutic response in the patients. Additionally, they investigated the relationship between mutational burden and TME score. It was found that the TIME score had a positive correlation with the mutational burden. Using a random forest algorithm, they were able to recognize 33 highly variant mutate genes with relationships with the TIME score. These results give a better insight into the mechanism of altered TME in different cancer patients (78).

Immunotherapy is not the only cancer treatment that the multi-omics approach has benefited. Neoadjuvant chemotherapy is considered a valuable method of treatment for GC. However, patients have exhibited varied results. Understanding the factors which affect neoadjuvant chemotherapy outcomes in GC can aid patient selection for this treatment. Li et al. studied the outcomes of neoadjuvant chemotherapy in 35 GC patients by taking tumor biopsies. They used a multi-omics data integration approach with RNASeq, WES, and WGS methods. They found that microsatellite instability greatly affects the treatment of GC. In particular, higher grades of MSI correlate with weaker responses. Moreover, \textit{IRS1} and \textit{C10orf71} mutations were found to relate to tumor treatment resistance. \textit{MYC} and \textit{MDM2} amplification were found to be biomarkers associated with tumor response with the former contributing to treatment sensitivity and the latter with resistance. C10orf71 mutation was found to be related to cisplatin resistance (79). Another therapeutic marker which has been under investigation using the multi-
omics approach is the ASC amino acid transporter 2 (ASCT2). ASCT2 is a Na+-dependent glutamine/neutral amino acid transporter, which is upregulated in several cancers including gastric and liver cancer. ASCT2 not only can contribute to cancer growth by transporting L-glutamine inside cancer cells, but also has been found to play roles in cancer development through crosstalk with several intracellular signaling pathways like mTOR. Therefore, it is considered a potential therapeutic target for cancer. Kasai et al. investigated the effects of KM8094, an ASCT2 blocker, in GC and explored possible biomarkers for efficacy evaluation by using a patient-derived xenograft (PDX) mouse model. It was observed that different cell lines of PDX models exhibited different results. In order to explore biomarkers for patient stratification, a multi-omics study was done using DNA methylation and gene expression data. A total of 11 genes were found to be upregulated more than 5-fold in the responder group compared to the non-responder one. Notably, TFF2 showed the strongest difference in this regard. However, none of the mentioned genes exhibited a differential methylation profile compared to the non-responder group. Moreover, the ASCT2 gene was found to be downregulated more than 5-fold in the responder group. Also, metabolomics analysis revealed a clear distinction in redox and energy status between the groups. The results from this study provide insights for future clinical studies on ASCT2 blockers in GC treatment and lay the basics for patient stratification based on genomic and transcriptomic biomarkers (80).

**Multi-omics Data Integration For Evaluation Of GC Diagnosis And Progression**

Utilization of multi-omics integration methods not only has been applied in clinical settings, but has also been employed to bring to light the complex mechanisms of GC progression. It is clear that the interactions between the tumor and the tumor microenvironment account for cancer development. In the case of GC, native myofibroblasts of the cancer tissue, often affect tumor progression through their secretome. One major player that synergizes with tumor-promoting effects of native tissue cells is hypoxia. Najgebauer et al. studied the effects of hypoxia on GC myofibroblasts (GCMs) and assessed the transcriptome and proteome of the mentioned cells. They acquired samples from 3 cancer patients and three controls and cultured them under hypoxic conditions. The results from their study indicate that GCMs, cancer adjacent myofibroblasts, and normal tissue myofibroblasts have distinct transcriptome and proteome profiles. Hypoxia seemed to influence GCMs gene expression to a weaker extent compared to other cell lines. Furthermore, analysis of transcriptome and proteome in cell lines revealed that transcriptome is a good indicator for the prediction of the proteome in myofibroblasts. However, this correlation was weaker in GCMs. The study provides novel insights into molecular mechanisms for the progression of GC (81). It seems that another factor to affect GC progression is vagal innervation. In a study performed by Rabben et al., the metabolomic characteristics of a mouse GC model and the effects of vagotomy on the metabolic reprogramming of cancer cells were studied. Vagotomy was found to significantly reverse metabolic reprogramming of cancer cells in the mentioned model. In order to explore the mechanism for this phenomenon, they performed an integrative multi-omics study using transcriptomic and metabolomic data derived from this model. They found that vagotomy altered glutaminolysis, WNT/b-catenin signaling, neuronal signaling, and mTOR signaling. This suggests that
blocking or activating these pathways might change the progression course of GC. Therefore, these pathways are considered potential therapeutic targets for GC (82). Another factor contributing to GC progression is the gut microbial properties, mainly HP infection. In recent years, other intragastric bacteria have also been found to associate with GC. Moreover, host gene expression in precancerous stages is found to be different than that of cancer. Park et al. studied the relationship between host gene expression and gastric microbiome using multi-omics approaches. The study comprised two different cohorts. In the first cohort, biopsy samples were taken from 3 cancer patients and 27 patients without cancer. Then, the microbiome and transcriptome were compared between the cancer group, HP-related gastritis group, HP-eradication group, and healthy group using a canonical correlation assay. The second cohort consisted of cancer samples as well as non-cancerous tumor-adjacent samples in 40 GC patients. 583 genes and 7 bacterial taxa were identified which differed between HP-related gastritis and healthy stomachs and Helicobacteraceae; STK17B and FAM3D were the most significant ones. Moreover, Neisseriaceae and Pasteurellaceae along with ITM2A and SOCS3 were associated with GC. In the surgery cohort, Lachnospiraceae alongside UBD, and GABRP were the most abundant in the cancer group compared with adjacent severe gastritis samples. Moreover, evaluation of immune cell infiltration levels revealed that in the cancer group, lower levels of immune cell infiltration were observed compared with healthy and gastritis samples. The findings of this study lay the ground for further research with the goal of developing novel diagnostic tools as well as therapeutic targets for GC (83).

Pancreatic cancer

Pancreatic cancer (PC) has the lowest 5-year survival rate of all common solid tumors, at only 11 percent (84). It is the 12th most common malignancy and the 7th formest cause of cancer mortality worldwide and is anticipated to be the second leading cause of cancer-related death in the US by 2030 due to its complicated biology, lack of efficient treatments, and high propensity to metastasize (85). About 90% of all cases of PC are diagnosed as pancreatic ductal adenocarcinomas (PDAC) (86). Hereafter, the term PC will be used throughout this text to refer to PDACs. Modifiable risk factors related to PC development include obesity, type 2 diabetes, and tobacco use. On the other hand, 5–10% of all PCs are estimated to be attributable to non-modifiable risk factors especially inherited predisposition genes such as BRCA1 or BRCA2, PALB2, ATM, STK11, CDKN2A (87).

Even after treatment, the recurrence incidence for PC is substantial; 83.7% of PC patients undergoing surgery (7.8 months of median DFS) (88) and 87% of patients who had adjuvant chemotherapy following surgical resection (13.4 months of median DFS) experience cancer recurrence again (89). Unpredictability in prognosis for patients with PC is a critical barrier in treating them because of the disease's heterogeneity. A number of studies have focused on factors that have been linked to PC prognosis, including clinicopathological characteristics (90), serum CA19-9 levels (91), and gene expression levels (e.g., TP53, bFGF, CD34, and VEGF) (92, 93). However, in order to establish a personalized treatment plan, the information provided by current medical and molecular tests is insufficient.
<table>
<thead>
<tr>
<th>No</th>
<th>Authors, Year</th>
<th>Disease</th>
<th>Sample Description</th>
<th>Omics</th>
<th>Finding</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rajamani, et al. (2016)</td>
<td>PC</td>
<td>GEO and TCGA</td>
<td>Transcriptomic, Epigenomic</td>
<td>Prognostic signature</td>
<td>(94)</td>
</tr>
<tr>
<td>2</td>
<td>Kong et al. (2020)</td>
<td>PC</td>
<td>161 PC patients</td>
<td>Genomic, Transcriptomic, Epigenomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(95)</td>
</tr>
<tr>
<td>3</td>
<td>Baek et al. (2020)</td>
<td>PC</td>
<td>134 PC patients</td>
<td>Transcriptomic, Epigenomic</td>
<td>Prognostic signature</td>
<td>(96)</td>
</tr>
<tr>
<td>4</td>
<td>Jia et al. (2020)</td>
<td>PC</td>
<td>150 PC samples from TCGA</td>
<td>Genomic, Transcriptomic, Epigenomic</td>
<td>Prognostic signature</td>
<td>(97)</td>
</tr>
<tr>
<td>5</td>
<td>Zhu et al. (2021)</td>
<td>PC</td>
<td>105 PC samples from TCGA</td>
<td>Genomic, Transcriptomic</td>
<td>Prognostic signature, Drug discovery</td>
<td>(98)</td>
</tr>
<tr>
<td>6</td>
<td>Xu et al. (2021)</td>
<td>PC</td>
<td>171 PC samples from TCGA</td>
<td>Genomic, Transcriptomic</td>
<td>Prognostic signature</td>
<td>(99)</td>
</tr>
<tr>
<td>7</td>
<td>Sinkala et al. (2020)</td>
<td>PC</td>
<td>45 PC patients</td>
<td>Transcriptomic, Epigenomic, Proteomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(100)</td>
</tr>
<tr>
<td>8</td>
<td>Ju et al. (2021)</td>
<td>PC</td>
<td>146 PC samples from TCGA</td>
<td>Transcriptomic, Epigenomic</td>
<td>Prognostic signature</td>
<td>(101)</td>
</tr>
<tr>
<td>9</td>
<td>Mishra et al. (2019)</td>
<td>PC</td>
<td>153 PC samples from TCGA</td>
<td>Transcriptomic, Epigenomic</td>
<td>Prognostic signature</td>
<td>(102)</td>
</tr>
<tr>
<td>10</td>
<td>Yang et al. (2021)</td>
<td>PC</td>
<td>66 pancreatic cancer PDX mouse models</td>
<td>Genomic, Transcriptomic, Epigenomic</td>
<td>Therapeutic response</td>
<td>(103)</td>
</tr>
<tr>
<td>11</td>
<td>Jang et al. (2017)</td>
<td>PC</td>
<td>BxPC-3 and BxPC-3ER cell lines</td>
<td>Transcriptomic, Metabolomic</td>
<td>Therapeutic response</td>
<td>(104)</td>
</tr>
<tr>
<td>12</td>
<td>Zhang et al. (2021)</td>
<td>PC</td>
<td>AsPC-1 and PANC-1 cell line</td>
<td>Transcriptomic, Proteomic</td>
<td>Therapeutic response, Patients stratification</td>
<td>(105)</td>
</tr>
<tr>
<td>13</td>
<td>Song et al. (2021)</td>
<td>PC</td>
<td>36 Pancreatic cancer cell lines</td>
<td>Transcriptomic, Proteomic</td>
<td>Therapeutic response, Patients stratification</td>
<td>(106)</td>
</tr>
<tr>
<td>No</td>
<td>Authors, Year</td>
<td>Disease</td>
<td>Sample</td>
<td>Omics</td>
<td>Finding</td>
<td>Ref</td>
</tr>
<tr>
<td>----</td>
<td>---------------</td>
<td>---------</td>
<td>--------</td>
<td>-------</td>
<td>---------</td>
<td>-----</td>
</tr>
<tr>
<td>14</td>
<td>Chen et al. (2022)</td>
<td>PC</td>
<td>185 PC samples from TCGA</td>
<td>Genomic, Transcriptomic, Epigenomic</td>
<td>Therapeutic response, Patients stratification</td>
<td>(107)</td>
</tr>
<tr>
<td>15</td>
<td>Long et al. (2014)</td>
<td>PC</td>
<td>6 PC patients, and 6 paired controls</td>
<td>Genomic, Transcriptomic</td>
<td>Diagnostic signature</td>
<td>(108)</td>
</tr>
<tr>
<td>16</td>
<td>Starzynska et al. (2020)</td>
<td>PC</td>
<td>21 PC patients (9 with PNETs, 8 with PDACs, and 4 with INETs)</td>
<td>Genomic, Transcriptomic</td>
<td>Diagnostic signature</td>
<td>(109)</td>
</tr>
<tr>
<td>17</td>
<td>Vandenbrouck et al. (2019)</td>
<td>PC</td>
<td>PC tissue, TCGA</td>
<td>Transcriptomic, Proteomic</td>
<td>Diagnostic signature</td>
<td>(110)</td>
</tr>
</tbody>
</table>

Multi-omics Data Integration For Evaluation Of PC Prognosis

A new era in cancer management began when single-omics data (111–115) which used to predict PC survival and recurrence. However, despite the abundance of research on single-omics data, there is room for advancement in the accuracy of these methods. Subsequently, a number of researchers sought to forecast PC patients’ prognoses by utilizing multi-omics data (116). For instance, Rajamani, et al. (94) attempted, in a multistep fashion, to create a system-level network of PDAC based on the omics data collected from the rank-based meta-analysis (mRNA, miRNA, DNA methylation), and they detected five candidate genes related to survival. Lower levels of *HMGA2*, *ACTN1*, and *SKI* expression were related to better survival probability, while higher levels of *IRS1* and *DLL1* expression were associated with longer survival times. (Log-rank test $p$-value $\leq 0.075$).

In a multi-omics analysis on PC published in 2020, researchers used information from the TCGA to identify four molecular subgroups of PC (iC1, iC2, iC3, and iC4) with differential Clinicopathological characteristics, immunological scores, and survival rates. The iC1 subgroup fared better than the others in terms of prognosis, immune cell infiltration, and genomic stability. Also, three novel genes (*GRAP2*, *ICAM3*, and *A2ML1*) were found to link with prognosis in PC according to this multi-omics analysis (95).

In another investigation, two biological characteristics were identified based on multi-omics datasets to predict survival and recurrence in pancreatic adenocarcinoma (PAAD) patients (96). First, they used TCGA WES data from 134 patients with PAAD and discovered five potential genes (*CDKN2A*, *TP53*, *TTN*, *KCNJ18*, and *KRAS*) altered at an early stage of tumorigenesis with high cellular prevalence. Patients who carried mutations in these potential genes had considerably better survival and recurrence rates.
compared to those whose cancer was caused by other mutations. Second, they used an autoencoder to combine TCGA information on RNA-seq, microRNA-seq, and DNA-MET. K-means clustering was then utilized to divide the patients into two groups (G1 and G2) with similar characteristics. When comparing DFS and OS, G1 had a worse prognosis. Using these two biological variables and nine clinical data, they developed Machine learning prognostic prediction models for predicting cancer recurrence and survival within five years (96).

Further, researchers used a multi-omics approach to identify biomarkers that have a significant impact on patients' prognoses. To this end, they examined TCGA's database of mRNA-seq, miRNA-seq, MET, and SNP data from people with PC. Finally, a PC prognostic model based on 12-biomarkers (hsa-mir-1179, hsa-mir-1224, hsa-mir-1251, hsa-mir-129-1, hsa-mir-129-2, MAPK8IP2, CPE, DPP6, MSI1, IL20RB, S100A2, and FMN2) was created (AUC = 0.683) (97).

Enhanced aerobic glycolysis (the Warburg effect) strongly correlates with tumor aggressiveness and is a poor prognostic indicator of patient survival. Zhu et al. (98) used TCGA data on 109 clinically annotated PDAC to examine the molecular events involved in aerobic glycolysis by matching DNA mutations, CNVs, and expression profiles of mRNA and IncRNA. Numerous IncRNAs implicated in PDAC aerobic glycolysis were found by combining transcriptional profiles with CNV and DNA mutations. They came to the conclusion that LINC01559 and UNC5B-AS1 control PDAC glycolysis and that by knocking down these IncRNAs, the glycolytic capacity is significantly inhibited. Additionally, high levels of LINC01559 and UNC5B-AS1 expression were also associated with a poor prognosis.

Through the integration and screening of genomic and transcriptomic data, Xu et al. (99) identified a signature of 9 characteristic genes, including UNC13B, TSPYL4, MICAL1, KLHDC7B, KLHL32, AIM1, ARHGAP18, DCBLD1, and CACNA2D4 that are strongly correlated with PC prognosis (AUC for 5-year survival was 0.93). Only the expression of these nine genes needs to be detected in clinical settings. Once the patients' prognosis scores have been calculated using this RNA-based risk score formula, the subsequent decisions can be taken based on whether the patient is at high risk or low risk.

In addition, a study has been conducted to determine molecular subtypes of PC using mRNA, microRNA, protein, and DNA-MET profiles by employing Similarity Networks Fusion (SNF), which constructs a similarity matrix of data inside and across different types of omics. On the basis of data from 45 highly purified cases of PC, two subgroups, Subtype-1 and Subtype-2, were established; each subtype with its own unique signaling pathways and clinical outcomes (OS = 16.05 months and 23.06 months, respectively. Nevertheless, their analysis of OS and DFS periods using the Kaplan-Meier methods indicated no statistically significant difference between the subtypes of pancreatic cancer (P-value = 0.18); this could be related to the limited sample size (100). Using a deep learning-based framework called MODEL-P (Multi-Omics Deep Learning Prognosis-correlated subtyping), Jie Ju et al. (101) analyzed mRNA, microRNA, and DNA methylation omics data from 146 PDAC surgically resected primary infiltrating (non-metastatic) patients from the TCGA PAAD cohort and identified two PDAC subtypes with discrete survival outcomes (median survival of 10.1 months and 22.7 months, respectively, log-rank p = 1
* 10^{-6}). Compared to subtypes found in earlier research (111–113, 117), MODEL-P subtypes were demonstrated to be more predictive of patient prognosis. Clinicians can use MODEL-P’s information in therapeutic decision-making.

**Multi-omics Data Integration For Evaluation Of PC Therapeutic Response**

To remedy PC, surgical excision is the sole option. However, only around 20%-30% of PC patients are candidates for surgical intervention at the time of diagnosis, and 70%-80% of patients were declared unresectable, including those with locally progressed or metastatic disease (118). The standard of care for patients with nonresectable PC is still chemotherapy. Chemotherapy regimens vary from patient to patient, but common examples include Gemcitabine, FOLFIRINOX, and albumin-bound paclitaxel, as well as combination therapies like Erlotinib and Gemcitabine. However, the poor prognosis is usually the result of the rapid and widespread emergence of chemoresistance (119, 120). It will be necessary to have a more in-depth understanding of the molecular processes behind the complex biological pathways leading to PC resistance in order to overcome drug resistance. Using WGS, RNA-seq, miRNA-seq, and DNA-MET analysis, Yang and colleagues analyzed the pre- and post-Gemcitabine treated 15 drug-sensitive and 13 drug-resistant PDXs (103). They found that DHTKD1 and CD55 were involved in Gemcitabine resistance. Furthermore, miR-135a-5p was shown to be substantially linked with the prognosis of patients with PC and was identified as a possible molecular marker to predict Gemcitabine response.

Jang et al. (104) compared erlotinib-sensitive (BxPC-3) and erlotinib-resistant (BxPC-3ER) PC cell lines using a multi-omics study of transcriptomics and metabolomics to determine which of these modifications might promote resistance to Erlotinib in PDAC. The research uncovered major changes in metabolic pathways controlling polyamines, amino acids, and fatty acids. Ornithine decarboxylase (ODC) and its major metabolite putrescine were discovered to have a role in the development of Erlotinib resistance in BxPC-3ER cells by further transcriptome research. Restoring sensitivity to Erlotinib was achieved by pharmacological or genetic blocking of ODC.

Numerous anti-angiogenic medicines have been authorized for use in oncology, and these drugs have shown promise in the treatment of multiple forms of cancer (121, 122). To the best of our knowledge, no clinical trials have demonstrated that angiogenesis inhibitors significantly enhance OS in PC (123). To elucidate the mechanism of action of Anlotinib, a new angiogenesis inhibitor, against PC cells, Zhang et al. (105) performed multi-omics studies (transcriptomics, proteomics, and phosphoproteomics). Patients with PC were divided into high-risk and low-risk groups with the help of a gene signature related to Anlotinib (consisting of five essential genes). Anlotinib-regulated signaling pathways were highly enriched in the high-risk category, which also had significantly lower survival times.

In 2021, Song et al. used a multi-omics approach using transcriptome and proteomic data. They classified 36 pancreatic cancer cell lines (PCCLs) into C1 and C2 subgroups, intending to develop anti-cancer medications for the precision treatment of PC patients. Temozolomide and NVP-TAE684 had a
more significant effect on the C1. The SIRT1, SRT-1720, on the other hand, was highly sensitive in C2. Finally, they looked into the links between changes in the genome and responses to certain medications, discovering that PCCLs with CDKN2A, TP53, or SMAD4 mutations were strikingly more susceptible to doxorubicin, sirolimus, and valdecoxib, among others (106).

Resistance to cancer therapies, such as chemotherapy (124), targeted therapy (125), or immunotherapy (126), has been the main focus of autophagy research in oncology. Despite evidence from in vitro and mouse models showing that blocking autophagy improves chemosensitivity to Gemcitabine in PC, the bulk of clinical trials has shown disappointing outcomes (127, 128). Chen et al. (107) combined mRNA, IncRNA, miRNA, methylation, and mutation datasets by employing 10 state-of-the-art multi-omics clustering algorithms. A robust PDAC classification algorithm was built to uncover the genes involved in autophagy, and the PADC patients were split into two groups (CS1 and CS2). Using a more comprehensive regulating network, they identified the top 20 autophagy-related hub genes (GAPDH, MAPK3, RHEB, SQSTM1, EIF2S1, RAB5A, CTSD, MAP1LC3B, RAB7A, RAB11A, FADD, CFKN2A, HSP90AB1, VEGFA, RELA, DDIT3, HSPA5, BCL2L1, BAG3, and ERBB2), 6 microRNAs, 5 transcription factors, and 5 immune infiltrating cells as biomarkers. Finally, each defined subgroup's projected potential drug-targeting signaling pathways were made to improve the advancement of anticancer treatment methods.

Multi-omics Data Integration For Evaluation Of PC Diagnosis

Diagnosing PC at an early stage is a complicated and ongoing medical challenge. Currently, the only treatment option for patients with PC who are not surgical candidates owing to metastases is palliative care; consequently, precise biomarkers for diagnosis at an earlier stage will significantly improve the PC patients’ prognosis. Traditional diagnostic methods are obviously lacking in the case of early detection. Indeed, the non-biopsy tests that are available (serum CA19-9) do not provide the proper sensitivity (50–75%) and specificity (83%) (129, 130). The multi-omics approach holds significant promise for the early diagnosis of PC because it gives a more holistic understanding of the cancerous process and its underlying mechanisms and pathways (108). In light of this, Kwon et al. (131) assessed the utility of multi-omics data analysis to discover biomarkers for early diagnosis based on miRNA and mRNA profiles from 104 cases of PC, employing support vector machine (SVM) modeling and leave-one-out cross-validation, and thereby identifying 705 multi-markers as promising potential biomarkers for PC diagnosis. Using an integrative method along with omics-based data, and supervised machine-learning techniques, Long et al. (132) were able to identify and validate potential biomarkers, which led to the development of a panel consisting of ADAM9, ANXA2, APLP2, and LAMC2 that could accurately detect PDAC in its early stages.

Hepatocellular carcinoma

Primary liver cancer was the third most prevalent cause of cancer death globally in 2020, which is the sixth most frequent cancer to be diagnosed, with 830,000 fatalities and about 906,000 new cases (133). The most frequent form of liver cancer is hepatocellular carcinoma (HCC), which accounts for ~ 90% of
patients with a 5-year OS rate lower than 20% (134). Unraveling the underlying mechanisms regarding the pathogenicity will be necessary to comprehend the HCC variety and develop targeted treatments. Chronic intake of alcohol, obesity-related NASH or diabetes, Hepatitis-B virus (HBV) infection, or Hepatitis-C virus (HCV) can all lead to liver damage and progressive destruction that results in the inflammatory process, fibrosis, and at last, carcinogenesis (135). Around 60% of HCC cases in Africa and Asia and 20% of patients in the West are affected by HBV infection (134). Viral replication relies on the metabolism of infected cells since they are intracellular parasites. Viruses control metabolic pathways to prevent infected cells' metabolic exhaustion and provide enough resources to maintain effective virogenesis, indicating that metabolic reprogramming is a feature of viral oncogenesis (136). The HBV genome encodes the HBV core protein (HBc), which may contribute significantly to the HBV life cycle. Nevertheless, the relevance of HBc's role in the initiation and progression of liver cancer remains to be determined. In an investigation, a multi-omics approach (composed of proteomics and metabolomics) was evolved to determine the metabolic alterations in HepG2 HCC cells transfected with HBc. In other words, HBc boosts the secretion and expression of essential metabolites in HCC cells, mainly by triggering the glycolysis and amino acid metabolism pathways (137). Recently a group of researchers gave a thorough perspective of the cellular endocytosis-associated proteome, transcriptome, and ubiquitylome over HBV infection using multi-omics integration methods to shed insight on the replication and pathogenesis of HPV. Through integrating multi-omics data from HCC cell lines, prominent proteins such as EGFR, HGS, STX4, ICAM1, VAMP8, and SCAMPs family members were discovered. SCAMPs are well known to take part in endocytosis. Consequently, they focused on SCAMPs and summarized their changes across omics data sets of the transcriptome, proteome, and ubiquitylome. Results implied that upregulation of SCAMP1 effectively impeded HBV RNAs and proteins secretion, while SCAMP1 knockdown remarkably grew viral production. Altogether, these two multi-omics studies' findings supply helpful knowledge to develop novel therapeutics and an exhaustive understanding of The role of HBV in HCC pathogenesis (137).

Globally, men are 2–5 times more likely than women to develop HCC (138). Estrogen may play a protective role against HCC carcinogenesis. Liver cancer specimens contain estrogen receptors (ER) which are sensitive to the hormone estrogen. Furthermore, Estrogen-specific ER agonists repress HCC cell proliferation and boost apoptosis (139, 140). To determine the metabolic effects of estrogen in HCC Shen et al. (141) used multi-omics techniques to assess the influence of estradiol and ER agonists on the transcriptome and the metabolome of HepG2 cells. They discovered Common altered genes and metabolites and determined the gene-metabolite interaction map in which SSTR1 and C5AR1 were in the middle of the map. It has been declared that normal hepatocytes do not express SSTR. Further, an elevated expression of SSTR1 in tumor cells was seen in patients diagnosed with advanced-stage HCC (142). Additionally, elevated C5AR1 expression is associated with tumor stage and tumor cell invasion of the liver capsule (143). In this study, SSTR1 expression was decreased by all estradiol and ER agonists, and estradiol also downregulated C5AR1 expression. Overall, this multi-omics study presents that estrogen inhibits the proliferation of liver cancer cells by altering metabolism and also suggests potential therapeutic targets for the treatment of HCC.
Nowadays, The Barcelona Clinic Liver Cancer (BCLC) staging system is the most common method for predicting HCC patients' survival and deciding which treatment options to pursue. Percutaneous ablation, surgical resection, and transplantation are all effective therapies for early HCC based on this method (144). Nonetheless, their efficacy diminishes in advanced HCC (145). Noticeably, the BCLC staging method is not sensitive or specific enough to characterize precisely the molecular and biological features that affect prognosis and responsiveness to therapy, even among tumors of the same stage (146). The high heterogeneity of HCC and the complicated etiologic factors make prognosis prediction challenging and necessitate a further critical need for developing methods and approaches to predict survival (147).
Table 5
Multi-omics data integration studies performed in HCC

<table>
<thead>
<tr>
<th>No</th>
<th>Authors, Year</th>
<th>Disease</th>
<th>Sample</th>
<th>Omics</th>
<th>Finding</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Liu et al. (2016)</td>
<td>HCC</td>
<td>256 HCC samples from TCGA</td>
<td>Genomic, Transcriptomic, Epigenomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(148)</td>
</tr>
<tr>
<td>2</td>
<td>TCGA (2017)</td>
<td>HCC</td>
<td>196 HCC patients</td>
<td>Genomic, Transcriptomic, Epigenomic, Proteomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(149)</td>
</tr>
<tr>
<td>3</td>
<td>Chaisaingmongkol et al. (2017)</td>
<td>HCC</td>
<td>130 ICC and 69 HCC patients</td>
<td>Genomic, Transcriptomic, Metabolomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(150)</td>
</tr>
<tr>
<td>4</td>
<td>Woo et al. (2017)</td>
<td>HCC</td>
<td>64 HCC patients</td>
<td>Genomic, Transcriptomic, Epigenomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(151)</td>
</tr>
<tr>
<td>5</td>
<td>Chaudhary et al. (2018)</td>
<td>HCC</td>
<td>360 HCC samples from TCGA</td>
<td>Transcriptomic, Epigenomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(152)</td>
</tr>
<tr>
<td>6</td>
<td>Benfeitás et al. (2019)</td>
<td>HCC</td>
<td>360 HCC patients</td>
<td>Transcriptomic, Proteomic, Metabolomic</td>
<td>Prognostic signature, Drug discovery</td>
<td>(153)</td>
</tr>
<tr>
<td>7</td>
<td>Gao et al. (2019)</td>
<td>HCC</td>
<td>159 HCC patients</td>
<td>Genomic, Transcriptomic, Proteomic</td>
<td>Prognostic signature, Drug discovery</td>
<td>(154)</td>
</tr>
<tr>
<td>8</td>
<td>Huang et al. (2020)</td>
<td>HCC</td>
<td>369 HCC patients from TCGA</td>
<td>Transcriptomic, Epigenomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(155)</td>
</tr>
<tr>
<td>9</td>
<td>Wang et al. (2020)</td>
<td>HCC</td>
<td>374 HCC patients</td>
<td>Transcriptomic, Proteomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(156)</td>
</tr>
<tr>
<td>10</td>
<td>Suter et al. (2022)</td>
<td>HCC</td>
<td>48 HCC patients</td>
<td>Genomic, Transcriptomic, Proteomic</td>
<td>Prognostic signature, Patients stratification</td>
<td>(157)</td>
</tr>
<tr>
<td>11</td>
<td>Wang et al. (2022)</td>
<td>HCC</td>
<td>287 HCC patients</td>
<td>Transcriptomic, Epigenomic</td>
<td>Prognosis signature</td>
<td>(158)</td>
</tr>
<tr>
<td>No</td>
<td>Authors, Year</td>
<td>Disease</td>
<td>Sample</td>
<td>Omics</td>
<td>Finding</td>
<td>Ref</td>
</tr>
<tr>
<td>----</td>
<td>------------------------</td>
<td>---------</td>
<td>------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>--------------------</td>
<td>------</td>
</tr>
<tr>
<td>12</td>
<td>Cheng et al. (2018)</td>
<td>HCC</td>
<td>375 HCC patients</td>
<td>Transcriptomic, Epigenomic</td>
<td>Diagnosis signature</td>
<td>(159)</td>
</tr>
<tr>
<td>13</td>
<td>Gao et al. (2021)</td>
<td>HCC</td>
<td>371 HCC samples and 49 NATs from TCGA</td>
<td>Genomic, Transcriptomic, Epigenomic</td>
<td>Diagnosis signature</td>
<td>(160)</td>
</tr>
<tr>
<td>14</td>
<td>Caruso et al. (2019)</td>
<td>HCC</td>
<td>34 liver cancer cell lines</td>
<td>Genomic, Transcriptomic, Proteomic</td>
<td>Drug discovery</td>
<td>(161)</td>
</tr>
<tr>
<td>15</td>
<td>Dimitrakopoulos et al. (2021)</td>
<td>HCC</td>
<td>12 mTOR-driven HCC mouse model</td>
<td>Genomic, Transcriptomic, Proteomic</td>
<td>Drug discovery</td>
<td>(162)</td>
</tr>
<tr>
<td>16</td>
<td>Yin et al. (2020)</td>
<td>HCC</td>
<td>374 HCC samples from GDC portal</td>
<td>Genomic, Transcriptomic</td>
<td>Drug discovery</td>
<td>(163)</td>
</tr>
</tbody>
</table>

**Multi-omics Data Integration For Evaluation Of HCC Prognosis**

Multiple biological levels of data integration enable cancer classification, particularly for highly heterogeneous cancer like HCC. To this end, Liu et al. combined CNV, MET, mRNA, and miRNA expression in 256 TCGA HCC cases. These samples were subjected to a "cluster of cluster" analysis, which yielded five different subgroups (S1-S5). The subclasses were linked to both molecular characteristics and overall survival (P-value = 0.0086). S1 tumors had a worse prognosis and were characterized by TP53 mutation, lower lipid metabolism, increased oncoprotein expression, and reduced tumor suppressor proteins. S2 and S3 were distinguished by low TERT and DNMT1/3B expression and telomere hypomethylation. S3 was linked to lower rates of copy number variation and better prognostic biomarkers compared to S2, such as CRP and CYP2E1 (148). By assembling the genome, transcriptome, epigenome, and proteome data of 196 HCC cases, TCGA succeeded in determining three subgroups (iClust 1–3). With a significantly worse prognosis, iCluster1 showed a high tumor grade and evidence of macrovascular invasion. The molecular features of this subgroup contained low TERT and CTNNB1 promoter mutations and upregulation of proliferation marker genes. Clinical characteristics of iCluster2 were low-grade tumors and reduced microvascular invasion. Chromosome instability, the prevalence of TP53 mutations, and hypomethylation at several CpG sites were all present in iCluster3. In addition, The immune checkpoint proteins CTLA-4, PD-1, and PD-L1 as well as WNT signaling are potential therapeutic targets for which drugs are available (149). Genomics, transcriptomics, and metabolomics profiles of 199 Asian liver cancer (130 Intrahepaticcarcinoma and 69 HCC) were integrated and four subtypes of intrahepatic
cholangiocarcinomas (ICC) including ICC C1-4 and three subtypes of HCC (HCC C1-3) were identified. Survival rates were the worst for HCC-C1 tumors and the best for HCC-C2 tumors, with a similar trend in ICC-C1 and ICC-C2 tumors. Expression and prognosis were shown to be quite similar between common C1 (HCC-C1 and ICC-C1), and common C2 (HCC-C2 and ICC-C2). Compared to the common C2 subtype, which was strongly enriched in cell immunity-related pathways, the common C1 subtype was enriched in mitotic checkpoint signaling pathways. Besides suggesting discrepancies in survival between different groups, these findings also demonstrate that despite ICC and HCC being clinically treated as separate entities, they share common molecular characteristics in prognosis prediction and actionable drivers to improve precision therapy (150). In a study by Hu et al. Three prognostic subtypes (C1, C2, and C3) of HCC are identified using multi-omics integration of the CNVcor and METcor genes of 64 HCC patients. RFS and OS were the worst for C1 ($p$-value $= 0.015$ and $p$-value $= 0.001$, respectively). Additionally, they evaluated clinicopathological characteristics between these subgroups and discovered that C1 cancers had vascular invasion more frequently than the others ($P = 0.002$). DNA copy number and DNA methylation abnormalities were most prevalent in group C1, whereas they were least prevalent in group C3 (151).

Chaudhary et al. used DNA MET, RNA-seq, and miRNA-seq data to develop a deep learning (DL)-based survival-sensitive model on 360 HCC patients. An unsupervised deep learning technology, anointed the autoencoder method, was applied in this study. Notably, the autoencoder framework was significantly more effective than principal component analysis (PCA) and individual Cox proportional hazards-based models in discovering survival factors. The DL-based model identified two optimal subgroups with significant differences in survival ($P = 7.13e-6$). S1, the most aggressive subtype, is related to frequent TP53 inactivation mutations, tumor marker BIRC5, high expression of stemness markers (EPCAM and KRT19), and activated Akt and Wnt signaling pathways. The gene expression profiles of KIRT19, EPCAM, and BIRC5, as well as Wnt signaling pathways, were strongly correlated with poor survival after functional analysis of these two subtypes (152).

Redox metabolism regulates signaling pathways, the progression of the cell cycle, and proliferation. Likewise, redox metabolism dysregulation might harm cells in different ways, such as disrupting pathways and promoting mutagenesis and apoptosis (164). The development of redox-based cancer prognoses and treatments has recently been pursued due to its crucial cellular roles (165). Combining Transcriptomics, Proteomics, and Metabolomics data from 360 primary HCC and 50 normal tissue specimens, Benfeitas and colleagues (153) sought to shed light on the critical features related to redox metabolism in HCC. Their integrative methods revealed that HCC tumors have significant metabolic differences. It was found that a 4-gene signature (PKM, MTHFS, G6PD, and HIF1A) led to an accurate prediction of patient survival (log-rank $P < 0.01$). Furthermore, this new survival 4-gene signature might be targeted for HCC treatment.

Proteogenomic analysis was carried out by Gao et al. (154) on 159 HBV-related HCC patients. Three subgroups (metabolic, proliferation, and microenvironmental dysregulation subgroups) were found, each with its own distinctive characteristics and metabolic pathways. Metabolism-related proteins, such as
ACAT1, ADH1A, G6PC, and PGM1 were higher in the metabolic subgroup. The proliferation subgroup had the shortest overall survival (P-value = 0.008) as well as the up-regulation of proteins related to proliferation, such as PARP1, TOP2A, PCNA, and MKI-67. The microenvironmental dysregulation subgroup is characterized by an intermediate expression of proliferative and metabolic proteins, mainly downregulated inflammatory, immune, and stromal proteins, such as CD4, CD8A, SPARC, S100A12, and ITGB3. PYCR2 and ADH1A, two prognostic biomarkers implicated in HCC metabolic reprogramming, were also discovered.

To comprehend the heterogeneity of HCC, Huang et al. defined a classification of HCC by combining data on the transcriptome and methylation of methylation status (MDGs). Consensus clustering determined four subclasses (HS1-4) notably related to prognosis (P value = 0.0057). Each of the four subclasses had distinct clinical characteristics as well. HS1 had a good prognosis and was well-differentiated. The increased serum fetoprotein level in HS2 was linked to its unfavorable outcome. HS3 was well differentiated and had a low serum level of beta-fetoprotein. It was also rich in metabolic signatures. Nonetheless, it was narrowly implicated in immune signatures. HS4 tumors were poor-differentiated, with the worst prognosis among all subtypes, and scarcely involved in metabolism signatures. This study may improve the ability to assess HCC patients’ prognoses (155).

An investigation combined transcriptomics and proteomics data to determine differentially expressed RNA binding proteins (RBPs) between 374 HCC tumor tissues and 50 normal samples. They selected 11 RBPs (BRIX1, DYNC1H1, GTPBP4, PRKDC, RAN, RBM19, SF3B4, SMG5, SPATS2, TAF9, and THOC5) to construct a risk score model. All 11 RBPs were expressed significantly higher in advanced-stage patients. The patients were also divided into low-risk and high-risk groups based on the risk score model, and the high-risk group patients had worse OS than the low-risk group (p < 0.001) (156).

Further, the genomics, transcriptomics, proteomics, and phosphoproteomics data of 48 HCC patients were used to apply bnClustOmics (Unsupervised Bayesian network-based clustering of multi-omics data). The authors distinguished three subtypes of HCC (Cluster 1–3) associated with prognosis and molecular characteristics. The Cluster 2 tumors had the poorest prognosis (p = 0.039) and were mostly composed of samples with mutant TP53. Co-occurrence of CTNNB1 and TP is most common in clusters 1 and 3, and furthermore, prognoses for these two clusters are better than Cluster 2 (157).

Wang et al. used rMKL-LPP to integrate multi-omics data (DNA methylation, miRNA, and mRNA expression) from 287 HCC patients. Two HCC subtypes (high-risk and low-risk) were shown to have a substantial correlation with overall survival (P-value < 0.0001). The 3-year mortality rate for the high-risk and low-risk groups was 51.0% and 23.5%, respectively. Potential prognostic biomarkers were identified by focusing on the two groups. High expression of four hub genes (CDK1, CDCA8, TACC3, and NCAPG) significantly correlated with poor prognosis (P-value < 0.05). These findings may also supply an essential reference for the precision treatment of HCC patients (158).
Multi-omics Data Integration For Evaluation Of HCC Therapeutic Response

Curative treatments, such as surgery, are advantageous in the early stages of HCC. However, most patients have advanced disease at the time of diagnosis and are unresectable, and not greatly benefitted from chemotherapy in terms of survival (166). In clinical practice, Sorafenib is one of the available agents for treating unresectable HCC patients. Nevertheless, only a small proportion of these patients respond satisfactorily to Sorafenib. In two clinical trials, the median survival was only 2–3 months longer in advanced HCC patients treated with Sorafenib than those who received placebo (167, 168). Therefore, to enhance HCC treatment outcomes, more accurate patient classification and the identification of novel pharmacological targets are desired. Employing Network-based Integration of multi-omiCS data (NetICS), Dimitrakopoulos et al. (162) identified genetic alterations in transcriptomics, proteomics, and phosphoproteomic data in twelve tumors from an mTOR-driven HCC murine model. This investigation indicates that targeting pathways, namely YAP1 or GRB2, and pathways regulating histone acetylation like SIRT1 or HDAC4 may be beneficial in the treatment of HCC with hyperactive mTOR signaling.

Over the last few years, immunotherapy has emerged as an accepted treatment for advanced stages of cancers. Despite its considerable success, only a small number of patients benefit from immunotherapy (163). Accordingly, the need for biomarkers for predicting the effectiveness of immunotherapy is urgent. Tumor mutation burden (TMB) is a novel biomarker that may help predict immunotherapy's outcome. TMB defines as the total number of mutations per megabase in the genome. In general, the higher the TMB, the more disparity within the tumor tissues; consequently, patients benefit more from immunotherapy (169). Yin et al. (163) assessed TCGA Transcriptome and gene mutation data, consisting of 374 HCC cases and 50 normal adjacent tissues, to learn more about the TMB and immune infiltration in HCC, as well as to look for potential new biomarkers for immunotherapy. Initially, samples were divided into two subclasses (High and Low-TMB groups), and the Low-TMB group had better OS (AUC = 0.69 and P-value = 0.002). Further, they found that patients lived considerably longer if they had low expression of GABRA3, CECR7, and TRIM16 and high expression of IL7R. The authors also noted that TMB is linked to mutations in GABRA3, CECR7, TRIM16, and IL7R, all of which promote anti-tumor immunity. Moreover, Xu et al. utilized multi-omics analysis comprised of transcriptomics and CNVs to construct the prognostic TMB-based Risk score, which significantly affected OS time, infiltration of immune cells, and immune checkpoint blockade (ICB)-related hub targets. Two subgroups (High and Low risk, in which low-risk patients had significantly higher OS (P-value = 2.407e−02)) were identified based on three hub genes, including HTRA3, OLFM1, and PLN. The correlation of prognostic signature with the expression value of six ICB hub genes was investigated to demonstrate the potential role of the risk score in response to immunotherapy. Results indicated that risk score might act as a non-negligible player in regulating the immune response to further immunotherapeutic efficacy (170).

Conclusion And Future Directions
The era in which cancer was diagnosed and treated only based on the organ of origin is ceasing. Therefore, the need to comprehend the more complex interactions between genomic and molecular abnormalities in tumor cells, tissue-specific gene expression patterns and the tumor microenvironment becomes more apparent (171, 172). Systems biology utilizing mathematical methods and artificial intelligence in cancer has prepared the beginning of an era in which the impacts of molecular abnormalities and interactions within networks are integrated with molecular and pharmacogenomics knowledge (173, 174). Recent years' remarkable growth in genetic sequencing, quantitative measurement of biochemical, imaging, spatial localization of network activities, and the combination of spatial-temporal information obtained in patients with omics data and computational analysis has paved the way for systems biology and, subsequently, precision medicine (175). Systems biology approaches in cancer can enhance our understanding of how tumor genetic heterogeneity, immune response, neovascularization, prognosis, and alterations in the tumor and its microenvironment change over time in response to therapeutic regimens (176, 177). Therefore, the most critical achievement of employing systems biology in cancer soon can be the extraction of diagnostic and prognostic biomarkers, the identification of new therapeutic targets and resistance mechanisms, and identifying combined approaches to prevent drug resistance.

This review demonstrates that applying multi-omics integration and ML algorithms in upper gastrointestinal cancers is an active domain that has developed enormously in recent years. The principal utilization of this promising technology is in the diagnosis, stratification of disease subtypes, prediction of prognosis, and evaluation of patient's treatment response. Investigations have also focused on optimizing strategies or searching for novel indicators to improve clinical evaluation regarding cost, time, and patient well-being.

In this study, we reviewed four main applications of multi-omics integration in upper GI cancers in order to achieve the goals of personalized medicine: 1) early diagnosis of patients with upper GI cancers; 2) Classification of upper GI cancers subtypes; 3) prediction of prognosis and response to treatment of patients with upper GI cancers based on personalized data, and 4) achieving novel and practical treatment goals (which are briefly shown in Fig. 2A). The outcomes have demonstrated that multi-omics integration and computational biology may have supported health specialists in clinical environments and research on upper GI cancers.

Currently, the use of systems biology approaches and multi-omics studies in cancers, especially upper GI cancers, is limited to research and has yet to be adopted in the clinical setting. Therefore, in order to overcome the existing clinical limitations and incorporate systems biology into treatment guidelines as quickly as possible, it seems essential to take additional measures, such as

1. prospective validation of systems biology methods and multi-omics studies in cancers of the upper gastrointestinal tract in independent cohorts,
2. Standardization of systems biology approaches and comparative studies, which evaluate the influence of heterogeneity using different types of data sets on outcomes of interest,
3. Conduct randomized controlled trials to determine whether employing systems biology in cancer results in more significant clinical effectiveness and profitability than standard care in upper GI cancers.

Reviewing the present investigations shows that the focus of multi-omics integration studies in upper GI cancers has been chiefly on genomic, epigenomic, and transcriptomic data (Fig. 2B and 2C). Although few other studies have been conducted in proteomics, metabolomics, and metagenomics, supposedly, more future studies in these fields will provide a more comprehensive and accurate view of the pathogenesis, course of the disease, and response to the treatment of these cancers. Furthermore, it has been clarified that the cellular-molecular conditions of each cell concerning its neighbors and non-cellular structures can provide valuable information to determine the cellular phenotype, cellular status, and cell and tissue function. Hence, applying single omics studies, which often go to the subcellular level, and provide an unbiased map of gene expression, protein, metabolite, etc., across tissue sections will be worthwhile (178, 179). Spatial transcriptomics is a recent approach that, considering the limitations of conventional transcriptomics techniques, tries to depict the situational context of transcriptional activity in intact tissue, either for regions or single cells (180). Therefore, it is hoped that future studies will take a step towards developing systems biology in cancer by using these data and achieving the next generation of diagnostic and prognostic biomarkers and new cancer treatments, which will play an essential role in the development of personalized medicine.

### Abbreviations

- **BCLC** Barcelona clinic liver cancer
- **CIN** Chromosomal-instable
- **CNAs** Copy number alterations
- **CNVs** Copy number variations
- **CRC** Colorectal cancer
- **DALYs** Disability-adjusted life years
- **DEGs** Differentially expressed genes
- **DEPs** Differentially expressed proteins
- **DMCs** Differentially methylated CpG sites
- **DMGs** Differentially methylated genes
- **DMRs** Differentially methylated regions
Declarations

Competing interests
The authors declare that they have no competing interests.

Availability of data and material
The data presented in this study are available on request from the corresponding author.

References


lumber motor impairments in patients with post-stroke hemiparesis: A systematic review and meta-
analyses. Adv Clin Exp Med. 2022 Sep 1;31(12):0–0.


116. Ding MQ, Chen L, Cooper GF, Young JD, Lu X. Precision oncology beyond targeted therapy: Combining omics data with machine learning matches the majority of cancer cells to effective therapeutics. Mol Cancer Res. 2018 Feb;16(2):269–78.


**Figures**

**A.**Deaths, rate per 100k

- **Esophageal cancer**
- **Stomach cancer**
- **Liver cancer**
- **Pancreatic cancer**


**B.**Global, Both sexes, 2019

- **Pancreatic cancer**
- **Liver cancer**
- **Stomach cancer**
- **Esophageal cancer**

**Figure 1**

**(A)** The global death rate of various upper GI cancers (esophageal cancer, gastric cancer, liver cancer, and pancreatic cancer) between 1990 and 2019, and **(B)** Disability-adjusted life years (DALYs) by age.
Figure 2

Multi-omics data integration in upper GI cancers pursue goals such as recognition and discovery of 'disease pathway', 'diagnostic signature', 'prognostic signature', 'patients’ stratification', 'drug discovery', and 'therapeutic response' in esophageal, gastric, liver, and pancreas cancers.
Figure 3

(A) Sankey diagram demonstrating the sequential steps employed to investigate the integration of omics data (including genomics, epigenomics, transcriptomics, proteomics, metabolomics, and metagenomics) in upper GI cancers to stratify patients or obtain diagnostic, prognostic, and therapeutic markers. (B) A cord plot displays the relationship between various omics studies in the multi-omics data integration.
studies in upper GI cancers. (C) R plot indicating the contribution of multi-omics studies considering their specific integrated omics in each upper GI cancers investigation.