WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal Homepage: www.ijrpr.com ISSN 2582-7421

Integrative Bioinformatics and Molecular Docking Approaches for Identifying Biomarkers and Therapeutic Targets in Hepatocellular Carcinoma.

Baswa Akhila, Dr. NVL Suvarchala Reddy.

Department of Pharmacology, Gokaraju Rangaraju College of Pharmacy, Osmania University, Hyderabad, Telangana- 500090.

ABSTRACT

Hepatocellular carcinoma (HCC) is one of the most common and aggressive forms of primary liver cancer and remains a major global health concern due to its high mortality, late diagnosis, tumour heterogeneity and limited treatment response. The complex molecular mechanisms involved in HCC development and progression highlight the need for reliable biomarkers and novel therapeutic targets. In recent years, bioinformatics and molecular docking have emerged as important computational approaches for understanding cancer-associated molecular alterations and supporting early-stage drug discovery. This review focuses on the integration of bioinformatics approaches and molecular docking strategies for identifying potential biomarkers and therapeutic targets in HCC. Publicly available biological databases, including the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA), provide valuable resources for analysing gene expression profiles and identifying differentially expressed genes (DEGs). Functional enrichment analysis, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, and protein–protein interaction (PPI) networks help identify important molecular pathways and hub genes associated with HCC progression. Furthermore, molecular docking techniques provide insights into potential interactions between selected protein targets and therapeutic compounds. Natural bioactive compounds such as curcumin, quercetin, resveratrol, epigallocatechin-3-gallate (EGCG) and luteolin have gained considerable attention due to their reported effects on cancer-associated signalling pathways. The integration of transcriptomic analysis, network biology, molecular docking and molecular dynamics may improve the prioritization of potential drug candidates and support drug repurposing strategies. Artificial intelligence and machine learning approaches may further enhance biomarker discovery and computational drug screening. However, limitations related to data quality, tumour heterogeneity, protein flexibility and docking prediction accuracy remain important challenges. Therefore, experimental and clinical validation is essential to establish the biological and therapeutic relevance of computational findings. Overall, the integration of bioinformatics and molecular docking represents a promising strategy for exploring potential biomarkers, therapeutic targets and future precision-based treatment approaches in hepatocellular carcinoma.

Keywords: Hepatocellular carcinoma; Bioinformatics; Differentially Expressed Genes; Molecular Docking; Biomarkers; Therapeutic Targets; Protein–Protein Interaction; Natural Compounds; Drug Discovery; Artificial Intelligence.

INTRODUCTION

Liver cancer has emerged as the fourth-leading cause of cancer-related deaths worldwide and the sixth-most frequently diagnosed malignancy. The incidence of liver cancer in men is about two times higher compared to women. Several risk factors have been implicated in the pathogenesis of this disease. The most prevalent risk factors for liver cancer are chronic hepatitis B virus (HBV), hepatitis C virus (HCV), aflatoxin-contaminated food, type 2 diabetes, heavy alcohol consumption, smoking, and obesity. There are three commonly recognized types of liver cancer, such as hepatocellular carcinoma (HCC), characterized by the malignant transformation of hepatocytes, accounting for approximately 80% of the global morbidity rate; secondly, intrahepatic cholangiocarcinoma (ICC), arising from biliary duct epithelium tissue; and finally, a rare type of hepatocellular carcinoma called combined HCC-ICC (cHCC-ICC), contributing to fewer cases worldwide. Hepatocellular carcinoma (HCC) constitutes almost 80% of all primary liver cancer cases, posing a significant global health burden. Therefore, to ensure early diagnosis and the prediction of prognosis for patients with HCC, the identification of liver cancer biomarkers can be crucial for improving outcomes. (1)

Due to the absence of early clinical manifestations, approximately three-fourths of primary liver cancer instances are detected and diagnosed in the advanced stage, failing to accommodate radical surgery and other treatment options, leading to poor prognosis and outcome with high recurrence and mortality rate. Despite the identification of various risk factors as contributing factors to HCC, the molecular mechanism of pathogenesis has been poorly understood. Several studies have found alterations in a variety of vulnerable genes, including TP53, CTNNB1, AXIN1, cyclin D1 (CCND1), epidermal growth factor receptor (EGFR), c-myc, and Ras. These genes exhibit aberrant expression and mutations and are, therefore, implicated in the genesis and advancement of HCC. Down-regulated ASPP1 and ASPP2 genes have been found to play roles in the development of HCC. However, despite various

studies aimed at identifying the underlying mechanisms, the current understanding is still insufficient. Therefore, further investigation is of profound importance to unveil the underlying mechanisms of HCC initiation and progression. (2)

To date, for early-stage HCC, different interventions such as surgical resection, liver transplantation, radiofrequency ablation, and various drugs such as cisplatin and doxorubicin by transarterial chemoembolization (TACE) have been used, but the efficacy of these treatments is still insufficient due to recurrence, complex risk factors, and drug resistance. Several targeted therapies, including sorafenib, lenvatinib, regorafenib, cabozantinib, and pembrolizumab, have demonstrated effectiveness in diverse clinical trials either alone or in combination. However, the efficacy of these drugs is compromised due to the adverse effects they entail. Hence, drug repositioning or repurposing through bioinformatic tools can be a cost-effective and time-efficient approach to drug development. This approach involves exploring approved or investigational drugs that can reverse cancer-related abnormal gene expression and can be reused for HCC, thereby expanding the array of treatment alternatives (3).

This study aimed to explore the core DEGs, prospective biomarkers, therapeutic targets, potential drug agents, and enriched pathways in HCC by using an integrated bioinformatics approach. Herein, we focused on screening hub genes by establishing a functional protein–protein interaction (PPI) network alongside regulatory biomarkers (i.e., TFs and miRNA) and potential therapeutic agents via drug repositioning techniques. By elucidating the enriched pathways associated with the core DEGs, the study intended to enhance our understanding of the molecular mechanisms driving HCC development and progression (4)

Molecular Pathogenesis of Hepatocellular Carcinoma

HCC mostly develops in ~80% of cases in the setting of a severely damaged cirrhotic liver that already gathers molecular alterations. In addition, several etiological (HCV and HBV infection, alcohol use, non-alcoholic steatohepatitis) and environmental factors (aflatoxin, aristolochic acid and tobacco) have been identified with distinct specific paths to cancer development. Specific molecular and immune classes have been defined, which integrate the current molecular knowledge of this cancer. In this regard, immune and epigenetic mechanisms might have major consequences in understanding the onset, evolution and treatment of this malignancy. Overall, the main molecular alterations and pathogenic processes involved in HCC development have been extensively reviewed elsewhere Most HCCs develop in patients with cirrhosis. These neoplasms progress through a sequence of well-defined histopathological phases, starting with emergence of dysplastic nodules which can ultimately transform into HCC. Genetic and epigenetic oncogenic alterations likely occur within hepatocytes, a cell type that despite its differentiated features is a facultative stem cell. Although mature hepatocytes are the main cells of origin for HCC, liver stem cells and transit amplifying cell populations have been also implicated in liver oncogenesis. In preclinical models, hepatic oncogenesis is favored by cell death with compensatory regeneration of hepatocytes, and blocking apoptosis reduces HCC formation. Replicative stress within regenerating hepatocytes induces genetic lesions favoring transformation and cancer progression, especially in the context of inflammation and fibrosis. In human NASH, auto-aggressive $CD8^+$ $PD1^+$ T cells induce hepatocyte cell death, promote NASH pathogenesis and impair immune surveillance, thereby favoring HCC occurrence and progression (5).

Bioinformatics Approaches in Hepatocellular Carcinoma Research

Gene Expression Data Source

Four datasets were selected based on the cause and development of HCC-associated conditions. They were DNA methylation, HBV and HCV-associated HCC (viral effect on HCC, Age-related HCC, Alcohol-associated HCC, and Cirrhosis and tumor in association with HCC. All gene expression datasets were obtained from GEO (Gene Expression Omnibus) (https://www.ncbi.nlm.nih.gov/gds). The details of the datasets considered for the current study (6) are described in Table **1**.

Table 1: GEO datasets related to HCC.

S. No.	**Datasets**	**Source**	**Define Groups**
Dataset-1 (DS-1)	GSE19665	DNA methylation in HBV and HCV [GPL570]	Normal *Vs* Methylation
Dataset-2 (DS-2)	GSE45267	Human hepatocellular gene expression profile [GPL570]	Age Specific 35-60 *Vs* Normal
Dataset-3 (DS-3)	GSE25097	Gene expression in HCC [GPL60876]	a. Cirrhosis *Vs* Healthy b. Tumor *Vs* Non-Tumor
Dataset-4 (DS-4)	GSE50579	Expression profiling of HCC [14550]	Alcohol *Vs* Normal

Processing the Data for DEGs Screening

The GEO2R web application (https://www.ncbi.nlm.nih.gov/geo/geo2r/) was used to identify genes that exhibit differential expression in two or more GEO series samples. For each selected dataset, the groups defined as tumor and non-tumor samples, cirrhotic *vs*. normal, age-specific HCC *vs*. normal, and alcohol *vs*. normal were reanalyzed through 'GEO2R Analyze' to identify DEGs. Various parameters like the *p*-value (adjusted p-value <0.05), force normalization, limma precision weight, and the Benjamini Hochberg were verified. The DEGs (upregulating and downregulating genes) were screened based on logFC value ≤2. The DEGs obtained from GEO2R were further analyzed in the EVenn platform (http://www.ehbio.com/test/venn/). It is a versatile tool used to determine genes that are persistently differently expressed across datasets, resulting in the identification of strong, significant biological genes within the intersection of the datasets (7)

Construction and Analysis of Protein-Protein Interaction Networks

Protein-Protein Interaction Network (PPIN) construction and analysis were done using three tools/databases *viz*., Fun Rich Tool, STRING, and Cytoscape. These studies help us to identify the genes that have the highest and strongest interactions mediating HCC (8).

Identification of Nodes using Fun Rich Tool : Functional enrichment tool (Fun Rich tool) (http://www.funrich.org/) is a software package designed for functional enrichment analysis and visualization of genes and proteins. It is particularly useful for analysing large-scale omics datasets, such as those generated from gene expression or proteomics experiments. This study was used to thoroughly visualize the nodes that have strong interactions and no interaction as well. The genes that were obtained from Evenn were subjected to the Fun Rich tool to identify the non-interacting nodes. This process helps to filter the genes that have no interaction and enables us to study the significant nodes further using the STRING database (9).

Construction of Protein-Protein Interaction Networks Using STRING Database: The obtained data of interacting nodes were subjected to STRING (Search Tool for Retrieval of Interacting Genes or Proteins) (https://string-db.org/) database, a web tool that forecasts protein-protein interactions. It is used to predict protein function and its relationship with other proteins and to understand the biological processes underlying the molecular level. It assigns confidence scores to each interaction, indicating the reliability of the supporting evidence. The PPIN of all interacting nodes/DEGs of HCC were constructed and ascertained using the STRING database. Usually, the interaction network is represented by seven types of evidence data: text mining, experiments, databases, coexpression, neighborhood, gene fusion, and cooccurrence. We comparatively analyzed the experimental data with all types of data. The upregulating and downregulating gene networks were separately constructed and downloaded to further identify hub genes. The highest number of nodes, edges, and the average node degree were regarded as hub genes (10).

Visualization and Analysis of Hub Genes Using Cytoscape: Cytoscape is a powerful software platform for visualizing and analysing biological networks (https://cytoscape.org/). The ranking of the hub nodes in the regulation network was located using CytoHubba. The top overlapping genes and their node degrees were determined by MCC (Maximal Clique Centrality). Using the MCODE plug-in, densely connected regions in the network were identified with degree cut-off 2, node score cut-off 2, K-core 2, and max depth = 100 (11).

Biomarker Discovery and Validation in HCC

Omics approach to identify novel biomarkers

The general approach of using omics technologies to diagnose HCC involves analyzing an array of biological molecules, such as proteins, metabolites, and modified RNA sites, to identify specific biomarkers and molecular profiles associated with the disease. This comprehensive and integrative method allows for early detection and improved understanding of HCC pathogenesis, ultimately leading to better patient outcomes and targeted treatment options. By overcoming challenges in data analysis, identification, and accessibility, omics technologies can revolutionize the field of HCC diagnosis and management. Mass spectrometry (MS)-based omics technologies have the potential to diagnose HCC by analyzing biological molecules related to disease development, such as proteomics techniques that identify specific serum protein panels. However, there are challenges to overcome, such as the rapid identification of metabolites, proteins, and modified RNA sites and the development of feasible in-house databases and powerful bioinformatics tools. Addressing these challenges and making multi-omic strategies accessible to non-specialist groups could lead to a better understanding of HCC pathogenesis and improved human health outcomes. Additionally, employing bioinformatics analysis on the current dataset is a valuable method for the secondary discovery of potential targets. Establishing prediction models based on DNA, methylation profiling, and somatic copy number aberrations data is a viable strategy. In recent years, the integration of machine learning algorithms has substantially advanced the field of biomarker discovery. Besides the bioinformatics approach to the studies of epigenetics and ncRNAs, spatial proteomics, or transcriptomics can also contribute to the disease prognosis and management. However, it is crucial to recognize that bioinformatics should be utilized as a supplementary tool in biomarker research rather than a primary source due to the absence of experimental evidence. Consequently, further experimental and clinical validations must be conducted to ensure the efficacy of the identified biomarkers.Taken together, many promising targets can be considered as new liquid biopsy biomarkers for detecting HCC. However, there is still a long way to look for a new universal serological HCC detection biomarker using liquid biopsy to replace AFP completely (12).

Biomarkers for Treatment options and Therapeutic Responses

Some recent review articles have discussed the representative findings of molecular biomarkers to predict HCC prognosis and the current staging system to predict HCC prognosis. The following part highlights the current understanding of biomarkers that can provide insights into treatment options and therapeutic responses. Biomarkers for therapeutic options. To facilitate the decision on the best treatment for HCC patients, biomarkers predicting therapeutic response to different treatments, including surgical resection, TACE, HAIC, liver transplantation, targeted therapies, and immunotherapies, are being investigated. Some of the significant aspects of the studies are highlighted in the illustration, and the predicted response and correlated clinical outcomes are drawn on the diagram (13).

Biomarkers related to liver transplantation

Liver Transplantation is considered when patients have unresectable HCC tumors by conventional surgery, but the tumors are still not widespread, usually at T1. Patients have to have good health conditions and normal liver function. As liver transplantation involves the donation of liver tissues from another individual and is a complicated procedure, it is precious. It requires special attention before the clinician and surgeon decide to operate on the patient. Therefore, reliable biomarkers are needed to help determine if a patient is suitable for transplantation and the predicted prognosis after the surgery. AFP levels are usually considered only in current clinical guidelines for therapeutic options. For example, patients with AFP levels above 1000 ng/mL are not considered for liver transplantation due to poor predicted prognosis after operation. Although AFP is one of the most widely used biomarkers for HCC diagnosis, its stability as an indicator of disease progression and the reference range has been controversial. Studies report a positive correlation between AFP mRNA level and tumor size, circulating malignant hepatocytes, and metastases. However, controversial studies have also been reported, limiting its significance as a prognostic marker for HCC. Parameters, including tumor sizes, numbers, serum AFP levels, and tumor uptake on FDG PET/CT, are recognized post-transplantation predictors of tumor recurrence in HCC patients. A combined set of criteria, including tumor sizes, numbers, serum AFP levels, and maximum tumor-to-background ratio, is the most effective way to predict post-transplant recurrence and select candidates for living donor liver transplantation compared to conventional criteria sets. A study of 103 patients with HCC found that poorly differentiated tumors and high AFP levels were linked with poorer survival ($p < 0.05$) and higher recurrence ($p < 0.05$) after liver transplantation. The research suggests that these tumor biological characteristics should be taken into account before performing the procedure for patients with HBV-associated HCC. This information can help assess HCC patient survival time after liver transplantation. A 30-year comprehensive study calculated the recurrence rate of 865 HCC patients receiving liver transplantation. Multivariate factors predicting recurrence were developed, including tumor grade, macrovascular and microvascular invasion, tumors outside Milan criteria, radiographic maximum diameter, and specific pre-transplant neutrophil-to-lymphocyte ratios. The situation is highly individual and requires more clinical information to help make decisions (14).

Biomarkers related to targeted therapies

For patients who have unresectable HCC tumors and who are also not suitable for liver transplantation, systemic therapies are the first-line treatment. Sorafenib is the most commonly used first-line drug, but the development of sorafenib resistance is usually observed. The study of sorafenib resistance is a hot topic, and many molecular mechanisms leading to sorafenib resistance have been discovered. Therefore, it is crucial to predict the drug response to systemic therapies in HCC patients. In a study using data from the SORAMIC study, researchers utilized EV-based proteomics to predict treatment response. Low GPX3 and high ARHGAP1 in plasma predicted better treatment response to sorafenib with AUC 0.97. From another phase III SHARP trial, the plasma Tregs and $CD56^{Dim}$ NK cells are correlated with shorter OS for patients receiving sorafenib ($p < 0.05$). When sorafenib is combined with mFOLFOX, plasma sVEGFR21 is associated with shorter time-to-progression in HCC patients. Consistent results were observed that VEGF and HIF in HCC have poor predicted treatment outcomes ($p < 0.001$) and OS ($p < 0.001$). Apart from typical serological markers, ncRNAs are also at the hot spot. For instance, the lncRNA Linc01056 is found to be downregulated in HCC tumors, where its expression level is related to sorafenib sensitivity. Besides, from a genomic study using patient-derived tissues, a sub-type of β-Catenin with high Male predominance has low miRNA expression and is sensitive to sorafenib. Similarly, the circulating level of another miRNA, miR-518d-5p, is reported to be a potential sorafenib resistance biomarker in BCLC-C HCC patients. On the other hand, the immune microenvironment also affects systemic drugs' therapeutic effect. A study showed that high $CD4^+$ T effector/Treg ratios and IFN-γ production by $CD8^+Ki67^+$ T cells are potential biomarkers for improved survival and response to sorafenib therapy. Similarly, $pERK^+$/pAkt- CTCs were discovered to be correlated with better prognosis in HCC patients receiving sorafenib treatment. Lenvatinib is sometimes regarded as the second-line therapy after sorafenib resistance has arisen. High expression of FAT1 and LRP1B genes in ctDNA is associated with poor PFS and a higher recurrence rate. Also, serum ST6GAL1, a tumor-secreted protein, is positively correlated with better survival of Lenvatinib. Likewise, a higher level of circMED27 in the serum of HCC patients also correlates with worse prognosis and clinical outcomes when treated with Lenvatinib. Some other biomarkers related to treating HCC with targeted therapies are reviewed in some recent articles (15).

Identification of Therapeutic Targets in HCC

The strategy for liver cancer therapy depends on many factors, including the histology, size of the lesion(s), their multiplicity, and stage of the disease. Comprehensive reviews about the main therapies adopted report surgical resection, liver transplantation, and loco-regional treatments as the main surgical and physical approaches. Systemic chemotherapy is reserved for advanced HCC and includes doxorubicin, gemcitabine, and oxaliplatin. The resulting overall survival is, however, still low, and alternative systemic therapies are under evaluation. Sorafenib (brand name: Nexavar®) and regorafenib (Stivarga®) are kinase inhibitors used to treat unresectable liver carcinoma (https://go.drugbank.com, accessed on 25 January 2024). Sorafenib is a bi-

aryl urea and an oral multikinase inhibitor. It targets cell surface tyrosine kinase receptors and downstream intracellular kinases that are implicated in tumor cell proliferation and tumor angiogenesis. This mechanism of action is shared by regorafenib: in in vitro biochemical or cellular assays, regorafenib or its major human active metabolites, M-2 and M-5, inhibited the activities of RET, VEGFR1, VEGFR2, VEGFR3, KIT, PDGFRα, PDGFRβ, FGFR1, FGFR2, TIE2, DDR2, TrkA, Eph2A, RAF1, BRAF, BRAFV600E, SAPK2, PTK5, and Abl at concentrations of regorafenib that have been achieved clinically (https://go.drugbank.com, accessed on 25 January 2024). In in vivo models, regorafenib demonstrated anti-angiogenic activity in a rat tumor model and inhibition of tumor growth, as well as anti-metastatic activity in several mouse xenograft models, including some for human colorectal carcinoma. Improvement in survival, though statistically significant, is, however, very limited with both drugs, which still imposes the need for better therapies (16).

Molecular Docking in Drug Discovery for HCC

Protein-ligand interactions are central to the in-depth understanding of protein functions in biology because proteins accomplish molecular recognition through binding with various molecules. Drugs often act as inhibitors when interacting with proteins, because many diseases including cancer are attributed to the abnormally active protein-ligand interactions. In this circumstance, inhibitors can be used to prevent the abnormal interactions for specific therapy. For instance, in the treatment of chronic myelogenous leukemia (CML), STI571 (Gleevec™, imatinib mesylate) is used to inhibit Bcr-Abl target protein activation, which also exemplifies the successful development of the rational drug design. Therefore, insights into protein-ligand interactions are vital for the development of drugs. Molecular docking algorithms predict protein-ligand interactions in a complex that is formed non-covalently. In biological systems, hydrogen bonds, ionic bonds, Van der Waals interactions and hydrophobic bonds are the four main types of non-covalent interactions. Non-covalent bonds, ranging from 1–5 kcal/mol, are weak in comparison with covalent bonds. Therefore, most of the time the term “non-covalent interactions” is used, rather than non-covalent bonds. However, the cumulative effect of these non-covalent interactions can be significant, which means multiple non-covalent bonds often act together at the binding interface of a complex to produce highly stable and specific associations (17). Fisher’s **lock-and-key model** theorizes that the binding interface should be matched complementarily. The protein and the ligand are rigid in this model, which means that their conformations are identical before and after binding. Du *et al.* suggested that the lock-and-key model is an entropy-dominated binding process (18). In Koshland’s **induced-fit model**, conformational change occurs to the protein during binding in order to accommodate the ligand with the best amino acid configuration. It should be noted that in most cases only minor conformational changes arise in proteins after ligand binding. Compared with the lock-and-key model, the induced-fit hypothesis can be linked to a “hand in glove” model, adding flexibility upon Fisher’s idea (19). In the **conformational selection model**, ligands bind selectively to the most suitable conformational state among an ensemble substates. In the original conformational selection model, no further conformational rearrangement occurs. However, in an extended conformational selection recognition mechanism, ligands first bind to a protein in the initial favorable protein conformational state, which induces a protein conformational change subsequently. In other words, the ”induced-fit” and ”conformational selection” recognition mechanisms may co-exist in ligand binding (20).

Docking Programs

In the past two decades, more than 60 docking programs have become available, such as SLIDE, X-CSCORE and ConsDock. Most of the existing docking algorithms treat ligands as flexible molecules and proteins as rigid molecules, because a protein usually consists of thousands of atoms. In some programs such as GalaxyDock, protein side-chains can be optimized. Moreover, there are indirect ways to consider the protein induced fit effect in ligand binding. A common approach is to utilize ensemble protein conformations. These ensemble protein structures can be obtained either from computational analysis such as molecular dynamics (MD) simulations or from experimental measurements such as NMR or X-ray diffraction. The MDock software incorporates a method referred to as “ensemble docking” to implicitly account for protein flexibility. In this method, the protein conformational number is introduced as an additional variable during docking optimization. This ensemble docking algorithm is computationally efficient; its computational speed is comparable to that of single-protein docking. Mizutani *et al.* used an enlarged binding pocket in their docking program ADAM to allow for certain protein flexibility. In this approach, the van der Waals energy curve for each atom pair is shiftied to widen the protein cavity uniformly, followed by structure optimization to remove atomic clashes and to improve interaction energy. Similarly, Bottegoni *et al.* adopted the SCARE (SCan Alanines and REfine) protocol, which enlarges the binding pocket by alanine mutation within the binding pocket. To enhance the success of docking procedures, certain docking programs focus on specific protein or ligand categories. Rather than creating a “one-size-fits-all scoring function”, improving prediction accuracy can be achieved by training a system-specific scoring function or incorporating additional factors into energy calculations. For instance, the ITScore in MDock has been effectively retrained using 2897 protein-ligand complexes sourced from the Protein Data Bank, proving valuable in screening for inhibitors and assessing their selectivities across multiple target proteins. Additionally, Vina-carb integrated the Carbohydrate Intrinsic (CHI) energy terms into the AutoDock Vina scoring function (ADV), improving the docking accuracy to 74%, compared with the 55% when relying soley on the ADV score (21).

Structure-based virtual screening and inverse virtual screening

The applications of molecular docking in therapy interventions can be bidirectional depending on the purpose of docking. Conventional docking searches for a compound from a DrugBank or from a compound library, a process referred to as virtual screening (VS). On the other hand, for a specific ligand of interest, inverse docking is useful for identification of the targets from thousands of protein candidates, a process called inverse virtual screening (IVS). Both prediction tasks are challenging and require immense calculations. Virtual screening methods are classified into structure-based virtual screening and ligand-based virtual screening. Structure-based virtual screening is explored by docking methods on a large scale. In contrast, ligand-based virtual screening methods utilize all the structural and chemical information (e.g., by searching the common features or similarity) rather than using a docking

strategy. These two methods are not competitive, and can be used individually or in combination. For instance, a ligand-based prescreening step can be done for preselection in order to reduce the size of the compound database in many structure-based virtual screening applications. Tan *et al.* integrated the two methods into parallel selection and postulated that structure-based and ligand-based screening could be more effective in a complementarity manner than rank fusion. IVS reverse docking has been widely applied to drug development in recent years. Screening from a drug target database can help rapidly identify potential target candidates for a specific molecule. The molecule can be a non-drug ligand or a known drug. The major safety issues of drug candidates are possible side effects (referred to as off-target effects) and toxicity. IVS can help investigate the interactions of the drug candidate with multiple targets in early-stage drug development. The purpose of IVS studies on known drugs is to search for alternative uses of FDA approved drugs, which reduce costs and risks in comparison with new chemical compounds. This approach is referred to as drug-repositioning (also known as drug repurposing or reprofiling). Drug development has been trending from "one gene, one drug, one disease" to polypharmacology, which means that drugs are designed to interact with multiple targets. The main challenge of IVS is the costly computational time. Recently, Zou group integrated their fast docking application, MDock, into a webserver platform in order to undertake the IVS tasks. The server is built-in with 3268 protein-ligand complexes (containing 3349 distinct binding sites) associated with 537 protein targets. A typical IVS run for a given ligand takes 1 to 4 hours. For the example, the IVS study on the progesterone receptor takes approxiamately one hour on a workstation with 24 Intel Xeon cores [Intel(R) Xeon(R) CPU E5–2650 v3 @ 2.30GHz]. In addition, the screening results from our web server provide searchable target diseases information (22).

Challenges of Molecular docking

Molecular docking methodology is rapidly evolving and improving with the exponential increase of computing power. Recently, machine learning algorithms (including neural network and deep learning) and hybrid approaches from high-content screening and virtual screening have been proposed for docking predictions and drug discovery. A critical question in docking is how reliable the docking predictions are. Consequently, benchmarking excises for blind docking prediction have been held, such as Community Structure-Activity Resource (CSAR), Drug Design Data Resource (D3R) and Continuous Evaluation of Ligand Pose Prediction (CELPP) group participated in the weekly CELPP competition using a fully automated systematic strategy. These large-scale protein-ligand complex structure predictions are statistically significant. Through our analysis, the current sampling (80.4%) is much better than the scoring (46.3%). The inaccuracy of current scoring functions likely arises from various sources, such as solvent effects, entropic effects, protonation or tautomer states, and so on. Development of new scoring schemes to address these issues is at large. The biggest hurdle in molecular docking stems from protein flexibility. The computational cost is immense for sampling all possible protein conformations. In conclusion, despite decades of docking algorithm development, prediction accuracy remains a major issue (23).

Integration of Bioinformatics and Molecular Docking

The integration of bioinformatics and molecular docking has become an important approach in modern cancer research, particularly in the identification of potential biomarkers and therapeutic targets in hepatocellular carcinoma (HCC). Bioinformatics tools help researchers analyze large-scale gene expression datasets, identify differentially expressed genes (DEGs), and understand the molecular mechanisms involved in cancer development and progression. Further analysis using functional enrichment and protein–protein interaction (PPI) networks helps identify important genes and proteins associated with HCC. Molecular docking provides an additional approach to investigate the possible interactions between selected protein targets and potential therapeutic compounds. By combining transcriptomic analysis, network-based approaches, and molecular docking, researchers can prioritize promising drug targets and screen compounds based on their predicted binding interactions. This integrated strategy may contribute to the early stages of drug discovery and provide useful information for future experimental investigations (24).

Prioritization of Candidate Biomarkers and Therapeutic Targets

The identification of potential biomarkers and therapeutic targets is an important step in understanding the molecular characteristics of HCC. Publicly available databases such as Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) provide valuable gene expression data for identifying molecular changes between tumor and normal liver tissues. Differentially expressed genes are commonly evaluated using functional enrichment analysis and protein–protein interaction networks. Hub genes identified through network analysis may play important roles in tumor development, cell proliferation, invasion, and disease progression. Genes such as TP53, MYC, CTNNB1, VEGFA, and GPC3 have been widely investigated in HCC research. However, their suitability as biomarkers or therapeutic targets depends on further molecular and clinical validation (25).

Structure-Based Screening of Potential Therapeutic Compounds

Structure-based screening is widely used in computational drug discovery to identify compounds that may interact with specific protein targets. In this approach, the three-dimensional structure of a selected protein is obtained from databases such as the Protein Data Bank (PDB), while potential ligand structures can be retrieved from PubChem. Molecular docking tools, including AutoDock Vina and PyRx, are used to predict the possible binding orientation and interactions of selected compounds within the protein binding site. Natural compounds such as curcumin, quercetin, luteolin, kaempferol, and resveratrol have been investigated in several computational studies for their potential interactions with cancer-associated proteins. These approaches help prioritize compounds for further investigation, although predicted binding scores alone cannot establish their therapeutic effectiveness (26).

Integration of Gene Expression, Network Analysis and Docking

Combining gene expression analysis with network analysis and molecular docking provides a broader understanding of the molecular mechanisms involved in HCC. Transcriptomic analysis helps identify genes showing altered expression in tumor tissues, while functional enrichment analysis provides information about their associated biological processes and signaling pathways. PPI network analysis using resources such as STRING and Cytoscape helps identify highly connected genes and potential regulatory proteins. Selected hub genes can subsequently be evaluated as possible therapeutic targets through molecular docking studies. This approach connects gene-level changes with protein-level interactions and provides a systematic framework for exploring potential therapeutic compounds in HCC (27).

Identification of Potential Drug–Target Interactions

The identification of potential drug–target interactions is an important application of integrated computational approaches. Once suitable therapeutic targets are selected, molecular docking can be performed to investigate the possible binding of candidate compounds to their corresponding proteins. Docking analysis provides information about predicted binding affinity, binding orientation, and interactions such as hydrogen bonds, hydrophobic interactions, and van der Waals forces. Visualization software such as BIOVIA Discovery Studio can be used to examine these interactions in two-dimensional and three-dimensional representations. The results may help researchers compare candidate compounds and select promising molecules for further experimental studies (28).

Advantages of an Integrated Computational Approach

The integration of bioinformatics and molecular docking offers several advantages in early-stage drug discovery research. It enables the analysis of large biological datasets, identification of potential biomarkers, prioritization of therapeutic targets, and preliminary screening of candidate compounds. Compared with experimental screening of large compound libraries, computational approaches can help reduce initial research time and costs. Another advantage is the ability to combine gene expression information with protein interaction networks and structural analysis. This provides a more comprehensive understanding of disease-associated molecular mechanisms. Nevertheless, computational predictions have certain limitations, including protein flexibility, scoring function accuracy, and differences between predicted and actual biological interactions. Therefore, experimental validation and further biological studies are necessary to confirm the potential applications of identified targets and compounds (29).

Natural Compounds and Potential Therapeutic Agents in Hepatocellular Carcinoma

Natural compounds have gained considerable attention in cancer research because of their ability to influence multiple biological processes involved in tumour development and progression. Several plant-derived compounds possess antioxidant, anti-inflammatory and anticancer properties that may be useful in the study of hepatocellular carcinoma (HCC). These compounds can interfere with important cellular mechanisms, including cell proliferation, apoptosis, angiogenesis, oxidative stress and tumour-associated inflammation. In recent years, compounds such as curcumin, quercetin, resveratrol, epigallocatechin-3-gallate (EGCG) and luteolin have been investigated for their potential effects against HCC. Their ability to regulate multiple molecular signalling pathways makes them interesting candidates for further drug discovery research (30).

1. Flavonoids and Polyphenols

Flavonoids and polyphenols are naturally occurring compounds widely distributed in fruits, vegetables, tea and medicinal plants. They have been studied for their protective effects against several diseases, including cancer. In HCC research, these compounds have attracted interest because they can influence different stages of tumour development, particularly oxidative stress, chronic inflammation and uncontrolled cell growth. Flavonoids form an important group of polyphenolic compounds and include quercetin, kaempferol, luteolin, apigenin, genistein and naringenin. Experimental studies have shown that several flavonoids can reduce the proliferation of liver cancer cells and promote programmed cell death. They may also influence signalling pathways such as PI3K/AKT, NF-κB, MAPK and Wnt/β-catenin, which are associated with tumour growth and survival. In addition, certain polyphenols have demonstrated the ability to regulate inflammatory responses and reduce the expression of molecules involved in tumour progression. The anticancer potential of these compounds is mainly supported by laboratory and animal studies. Although their biological activities are promising, further research is required to understand their effective concentrations, safety and suitability for clinical applications in HCC (31).

2. Curcumin, Quercetin and Resveratrol

Curcumin, quercetin and resveratrol are among the natural compounds frequently investigated for their potential anticancer properties. Each compound has a different chemical structure, but all three have demonstrated the ability to influence molecular pathways associated with cancer development. Curcumin is a natural bioactive compound obtained from the rhizome of *Curcuma longa* (turmeric). It is known for its anti-inflammatory and antioxidant properties. In experimental HCC models, curcumin has been reported to suppress cancer cell proliferation and promote apoptosis. These effects are associated with the regulation of NF-κB, STAT3, PI3K/AKT and Wnt/β-catenin signalling pathways. Curcumin has also been investigated for its possible role in controlling cell-cycle progression and reducing tumour-associated inflammation. However, its poor water solubility and low oral bioavailability remain major challenges for its therapeutic development. Quercetin is a naturally occurring flavonol present in foods such as onions, apples, berries and several vegetables. It has received attention because of its antioxidant, anti-inflammatory and antiproliferative activities. Experimental studies suggest that quercetin may inhibit the growth of HCC cells by influencing apoptosis-related proteins and signalling pathways such as PI3K/AKT and STAT3. It

has also been investigated for its potential to increase the sensitivity of cancer cells to certain anticancer agents. Despite these findings, its limited bioavailability and the need for further safety studies restrict its current clinical application. Resveratrol is a natural stilbene polyphenol found in grapes, berries and peanuts. It has been studied for its potential effects on cancer cell growth, inflammation and oxidative stress. In experimental HCC models, resveratrol has demonstrated the ability to promote apoptosis and regulate pathways associated with tumour survival, including PI3K/AKT, STAT3 and NF-κB. Some studies have also reported changes in p53 and PTEN-related signalling following resveratrol treatment. However, its rapid metabolism and limited systemic availability are important factors that need to be addressed before considering its therapeutic use (32).

3. EGCG, Luteolin and Other Bioactive Compounds

Apart from curcumin, quercetin and resveratrol, several other natural compounds have been investigated for their possible role in HCC management. One important example is epigallocatechin-3-gallate (EGCG), the major catechin present in green tea. EGCG has demonstrated antioxidant and anti-inflammatory activities in different experimental models. In HCC studies, it has been associated with reduced cancer cell proliferation, increased apoptosis and modulation of signalling pathways such as PI3K/AKT and ERK. Its possible effects on angiogenesis and tumour invasion have also been investigated. Luteolin is a flavone commonly found in celery, parsley, green peppers and other plant sources. Experimental studies suggest that luteolin may inhibit the growth of liver cancer cells by regulating apoptosis, cell-cycle progression and inflammatory signalling. It has also been associated with changes in pathways such as PI3K/AKT, NF-κB and Wnt/β-catenin. These findings indicate that luteolin may have potential as a candidate for further investigation in HCC research. Other bioactive compounds, including kaempferol, apigenin, genistein, naringenin, baicalein and berberine, have also shown promising activities in preclinical cancer studies. Depending on the compound and experimental model, these molecules may influence oxidative stress, inflammatory responses, cell survival, angiogenesis and metastatic behaviour. However, their effects are not identical, and their biological activity depends on factors such as concentration, absorption, metabolism and the type of cancer model used. Overall, these compounds provide a broad range of chemical structures and biological activities that can be explored in the search for potential therapeutic candidates against HCC (33).

4. Computational and Experimental Evidence for Anticancer Activity

Computational approaches and experimental studies have become important tools for investigating the anticancer potential of natural compounds. Molecular docking, in particular, is widely used to predict the possible interaction between bioactive molecules and selected protein targets involved in cancer progression. It helps researchers examine the binding orientation of a ligand within a protein binding site and estimate its binding affinity. In HCC research, natural compounds such as curcumin, quercetin, resveratrol, EGCG and luteolin can be evaluated against proteins associated with tumour growth, survival, angiogenesis and other cancer-related processes. Tools such as AutoDock Vina, PyRx and Discovery Studio are commonly used for docking analysis and visualization of protein–ligand interactions. The predicted interactions, including hydrogen bonds, hydrophobic interactions and other non-covalent contacts, can provide useful information for prioritizing compounds for further investigation. Experimental studies provide additional evidence by examining the effects of these compounds in cancer cell lines and animal models. In-vitro studies can assess cell viability, apoptosis, cell-cycle changes and alterations in gene or protein expression. In-vivo studies help evaluate their effects on tumour growth and biological responses in living systems. For example, preclinical investigations have reported that curcumin, resveratrol and EGCG can influence apoptosis-related mechanisms (34).

Clinical Relevance and Translational Applications

The application of bioinformatics and molecular docking in hepatocellular carcinoma (HCC) has created new opportunities for understanding disease progression and identifying potential treatment strategies. Although conventional approaches such as surgery, locoregional therapies and systemic treatment are used in HCC management, treatment outcomes can vary considerably among patients. This variation is partly associated with differences in tumour characteristics, genetic alterations, molecular signalling pathways and treatment responses. Integrating transcriptomic analysis, biomarker identification and molecular docking may help researchers understand these differences and identify potential targets for future therapeutic development. Translating these research findings into clinical practice could contribute to more accurate diagnosis, improved treatment selection and better patient management (35).

1. Precision Medicine and Biomarker-Guided Therapy

Precision medicine focuses on selecting appropriate treatment strategies based on the molecular and biological characteristics of an individual patient. In HCC, tumours may show different patterns of gene expression, mutations and molecular pathway alterations. Therefore, patients with the same type of cancer may not always respond similarly to a particular treatment. Biomarkers play an important role in this approach by providing information about disease development, prognosis and possible treatment response. Transcriptomic analysis can help identify differentially expressed genes (DEGs) between normal liver tissue and HCC samples. Further analysis using gene ontology, pathway enrichment and protein–protein interaction networks can help identify genes and proteins associated with tumour progression. These findings may support the identification of potential diagnostic, prognostic and predictive biomarkers. For example, genes such as TP53, MYC, GPC3 and VEGFA have been investigated in relation to HCC development and progression. Alterations in these genes and their associated pathways may provide useful information about tumour behaviour. However, the clinical usefulness of individual biomarkers depends on their validation in independent patient populations and their ability to provide reliable information beyond existing clinical assessments. Biomarker-guided treatment may also help identify molecular features associated with specific therapeutic responses. For instance, information about angiogenesis-related pathways, immune-related markers and growth signalling mechanisms can support research into treatment selection. In this way, combining molecular profiling with clinical information may contribute to the development of more individualized approaches to HCC management (36).

2. Drug Repurposing and Novel Drug Discovery

Drug repurposing involves investigating existing drugs for new therapeutic applications beyond their originally approved indications. This approach has attracted interest in cancer research because the safety profiles, pharmacological properties and some clinical characteristics of existing medicines are already known. Bioinformatics tools can support drug repurposing by identifying disease-associated genes, molecular pathways and possible drug–target relationships. In HCC research, gene expression analysis and network-based approaches can be used to identify important molecular targets involved in tumour progression. These targets can then be compared with information available in drug databases to identify existing compounds that may interact with the selected proteins. Molecular docking can provide additional information by predicting the possible binding orientation and interaction strength of a drug molecule with a target protein. Apart from repurposing existing drugs, computational approaches can also support the discovery of novel therapeutic candidates. Natural compounds such as curcumin, quercetin, resveratrol, luteolin and EGCG have been investigated for their possible interactions with cancer-related proteins. Docking studies can help prioritize these compounds based on their predicted interactions with selected targets. Compounds showing promising computational results can then be taken forward for further experimental investigation. However, drug repurposing and computational screening are early stages of drug discovery. A favourable docking score does not establish therapeutic effectiveness. The selected compounds require further evaluation through biochemical assays, cell-based studies, animal models and appropriate clinical investigations. Their toxicity, pharmacokinetic properties, bioavailability and possible drug interactions must also be considered before clinical use (37).

3. Potential Applications in Personalized HCC Treatment

Personalized treatment aims to provide therapeutic strategies that are appropriate for the individual characteristics of each patient. In HCC, this approach may involve considering tumour stage, liver function, molecular alterations, underlying liver disease and the patient's overall clinical condition. Molecular profiling can provide additional information about the biological behaviour of the tumour and may help identify pathways that could be explored for targeted treatment. Integrated bioinformatics analysis may contribute to this process by identifying disease-associated genes, important molecular networks and potential therapeutic targets. These findings can be combined with molecular docking results to investigate whether selected drug molecules or natural compounds may interact with the identified targets. Such an approach may help prioritize candidate compounds for further drug development and experimental validation. In the future, combining transcriptomic data, clinical information, biomarker profiles and computational drug screening may support the development of more individualized treatment strategies for HCC. Machine learning and other artificial intelligence-based methods may further assist in analysing large biological datasets and identifying patterns associated with disease progression and treatment response. Despite these opportunities, several challenges must be addressed before these approaches can be routinely applied in clinical practice. These include tumour heterogeneity, differences between patient populations, limited availability of validated biomarkers and the need for reliable experimental evidence. Therefore, findings obtained through computational analysis should be confirmed using independent clinical datasets and laboratory studies. Overall, the integration of bioinformatics, molecular docking and clinical information has the potential to strengthen the development of precision-based approaches in HCC. Although these methods cannot replace established clinical decision-making, they provide valuable research strategies for identifying biomarkers, prioritizing therapeutic targets and exploring potential treatment candidates (38).

Challenges and Future Perspectives

Although bioinformatics and molecular docking have become valuable approaches in hepatocellular carcinoma (HCC) research, several challenges still need to be addressed before their findings can be translated into clinical applications. These methods help researchers analyse large biological datasets, identify potential biomarkers and predict interactions between drug molecules and target proteins. However, the reliability of the results depends on the quality of the data, the accuracy of computational methods and proper experimental validation. In addition, the complex molecular characteristics of HCC make it difficult to identify a single biomarker or therapeutic target that is suitable for all patients. Addressing these limitations and improving existing computational techniques will be important for future research and drug discovery (39).

Data Quality, Reproducibility and Tumour Heterogeneity

The quality and reliability of biological data play an important role in bioinformatics analysis. Public databases such as Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) provide valuable gene expression and genomic information for HCC research. However, these datasets may differ in sample size, patient characteristics, experimental methods, sequencing platforms and data processing procedures. Such differences can introduce technical variations and affect the identification of differentially expressed genes (DEGs). Another important challenge is the reproducibility of computational results. Different normalization methods, statistical parameters, software versions and analytical pipelines may produce variations in the final results. Therefore, researchers should clearly describe their data sources, analysis methods and selection criteria. The use of independent datasets for validation can also help determine whether the identified biomarkers are consistently associated with HCC. Tumour heterogeneity is another major difficulty in HCC research. Different regions of the same tumour may contain cells with different genetic and molecular characteristics. In addition, molecular profiles can vary between patients due to differences in disease stage, underlying liver conditions and genetic alterations. This variation may influence tumour progression and treatment response. Single-cell RNA sequencing and spatial transcriptomic approaches may help researchers understand these differences at a more detailed level. These techniques could improve the identification of cell-specific biomarkers and molecular targets in HCC (40).

Limitations of Computational Predictions

Computational approaches, including differential gene expression analysis, protein–protein interaction networks and molecular docking, are useful for identifying potential biomarkers and therapeutic candidates. However, these methods have certain limitations that should be considered while interpreting the results. In transcriptomic analysis, the identification of DEGs depends on the quality of the dataset, sample selection, normalization and statistical thresholds. A gene showing increased or decreased expression does not necessarily have a direct role in tumour development. Similarly, hub genes identified through protein–protein interaction networks may represent highly connected proteins, but their importance does not automatically indicate that they are suitable therapeutic targets. Molecular docking also has several limitations. Most docking methods use simplified representations of protein structures and may not fully account for protein flexibility, solvent effects, changes in protein conformation or the biological environment. The predicted binding affinity is generally an estimate and may not accurately represent the actual binding strength of a compound under experimental conditions. In addition, different docking programs and scoring functions may produce different results for the same protein–ligand interaction. Another limitation is the availability and accuracy of protein structures. Some target proteins may have incomplete structural information, and the selected protein structure may not represent its active biological conformation. Therefore, computational predictions should be interpreted as preliminary findings rather than confirmed evidence of biological activity or therapeutic effectiveness (41).

Molecular Dynamics and Experimental Validation

Molecular dynamics (MD) simulations provide additional information about the stability and behaviour of protein–ligand complexes over time. Unlike conventional molecular docking, which generally predicts a static binding pose, MD simulations examine the movement of atoms and molecules under simulated conditions. They can help researchers understand protein flexibility, ligand stability and changes in protein–ligand interactions. Parameters such as root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg) and hydrogen bond formation are commonly examined during molecular dynamics analysis. These measurements provide information about the structural stability and flexibility of the protein–ligand complex. Methods such as molecular mechanics/generalized Born surface area (MM/GBSA) can also be used to estimate the binding free energy of a complex. However, these computational measurements are dependent on the simulation conditions and should not be considered direct proof of therapeutic activity. Experimental validation is essential to confirm the findings obtained through bioinformatics, docking and molecular dynamics studies. In-vitro experiments using HCC cell lines can help evaluate cell viability, apoptosis, cell-cycle progression and changes in gene or protein expression. Further studies using appropriate animal models can provide information about tumour growth, drug distribution, toxicity and treatment response. For example, a compound showing favourable docking interactions with a selected HCC-associated protein should be tested to determine whether it produces the expected biological effect. Biochemical binding assays and target-specific experiments can help confirm the predicted interaction. Combining computational analysis with laboratory and preclinical studies provides stronger evidence for selecting potential drug candidates. Clinical studies are ultimately required to establish safety and therapeutic effectiveness in patients (42).

Artificial Intelligence in HCC Biomarker and Drug Discovery

Artificial intelligence (AI) and machine learning are increasingly being explored in cancer research because of their ability to analyse large and complex biological datasets. These techniques can identify patterns that may be difficult to recognize using conventional statistical methods. In HCC research, AI-based approaches can be applied to gene expression data, genomic profiles, clinical in formation and molecular interaction datasets. Machine learning algorithms can help identify important genes associated with HCC development, classify tumour and normal samples, and develop potential diagnostic or prognostic biomarker signatures. Methods such as random forest, support vector machines and other classification algorithms can be used to select relevant features and evaluate their ability to distinguish different disease groups. Deep learning approaches may also support the analysis of complex multi-omics data and improve the understanding of molecular patterns associated with tum our progression. AI can also contribute to drug discovery by supporting virtual screening, prediction of drug–target interactions and identification of potential therapeutic compounds. When combined with molecular docking, machine learning may help prioritize compounds for further investigation and reduce the number of candidates requiring experimental screening. In addition, AI-based models may assist in predicting pharmacological properties and identifying possible toxicity-related concerns. In the future, integrating AI with transcriptomic analysis, protein interaction networks, molecular docking and molecular dynamics simulations may provide a more comprehensive approach to HCC research. Such integration could support the identification of reliable biomarkers and the selection of promising therapeutic candidates. However, the performance of AI models depends on the quality, diversity and size of the training datasets. Independent validation, transparent reporting and careful assessment of model performance are necessary to avoid biased or unreliable predictions. Overall, improvements in data quality, computational methods, molecular simulations and experimental validation are expected to strengthen future HCC research. The combination of bioinformatics, molecular docking and artificial intelligence may provide new opportunities for understanding tumour biology and exploring potential therapeutic strategies. However, successful clinical translation will require further laboratory investigations and well-designed clinical studies (43).

CONCLUSION

Hepatocellular carcinoma (HCC) is a complex and heterogeneous malignancy associated with multiple genetic, epigenetic, molecular and environmental factors. Despite advances in diagnosis and treatment, HCC continues to present major challenges due to its late detection, high recurrence rate, tumour

heterogeneity and variable therapeutic responses. Therefore, the identification of reliable biomarkers and potential therapeutic targets is important for improving the understanding and management of this disease. Bioinformatics approaches have become valuable tools in HCC research by enabling the analysis of large-scale biological datasets and the identification of differentially expressed genes, molecular pathways and important regulatory networks. Techniques such as Gene Expression Omnibus (GEO) data analysis, functional enrichment, protein–protein interaction networks and hub gene identification provide useful insights into the molecular mechanisms associated with tumour development and progression. These approaches can support the identification of potential diagnostic, prognostic and predictive biomarkers, although their clinical significance requires further validation. Molecular docking provides an additional computational strategy for investigating interactions between selected protein targets and potential therapeutic compounds. By predicting binding orientation, binding affinity and protein–ligand interactions, docking studies can help prioritize promising molecules for further investigation. Natural bioactive compounds such as curcumin, quercetin, resveratrol, epigallocatechin-3-gallate (EGCG) and luteolin have attracted considerable interest because of their reported biological activities against cancer-associated pathways. However, their therapeutic potential requires further investigation, particularly regarding bioavailability, pharmacokinetic properties, toxicity and clinical effectiveness. The integration of transcriptomic analysis, functional enrichment, network biology and molecular docking provides a systematic approach to connecting gene-level alterations with potential protein targets and therapeutic compounds. Furthermore, molecular dynamics simulations and artificial intelligence-based approaches may improve the interpretation of protein–ligand interactions and support the identification of promising drug candidates. These integrated computational strategies may contribute to drug repurposing, biomarker discovery and the development of more personalized therapeutic approaches for HCC. Nevertheless, computational predictions alone cannot establish the biological activity or therapeutic effectiveness of identified biomarkers and compounds. Differences in patient populations, tumour heterogeneity, data quality, protein flexibility and limitations of docking algorithms remain important challenges. Therefore, independent dataset validation, biochemical studies, in-vitro and in-vivo experiments, and well-designed clinical investigations are essential to confirm computational findings. Overall, the integration of bioinformatics and molecular docking represents a promising research strategy for understanding the molecular complexity of hepatocellular carcinoma and exploring potential biomarkers and therapeutic targets. Further improvements in computational methods, molecular validation and experimental research may help translate these findings into clinically relevant applications and support the future development of precision-based treatment strategies for patients with HCC.

REFERENCES

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [Google Scholar] [CrossRef] [Scilit] [PubMed]
2. Bosch, F.; Ribes, J.; Borràs, J. Epidemiology of Primary Liver Cancer. *Semin. Liver Dis.* **1999**, *19*, 271–285. [Google Scholar] [CrossRef] [Scilit] [PubMed]
3. Piñero, F.; Tanno, M.; Aballay Soteras, G.; Tisi Baña, M.; Dirchwolf, M.; Fassio, E.; Ruf, A.; Mengarelli, S.; Borzi, S.; Fernández, N.; et al. Argentinian clinical practice guideline for surveillance, diagnosis, staging and treatment of hepatocellular carcinoma. *Ann. Hepatol.* **2020**, *19*, 546–569. [Google Scholar] [CrossRef] [Scilit] [PubMed]
4. Pham, C.; Sin, M.-K. Use of Electronic Health Records at Federally Qualified Health Centers: A Potent Tool to Increase Viral Hepatitis Screening and Address the Climbing Incidence of Liver Cancer. *J. Cancer Educ.* **2021**, *36*, 1093–1097. [Google Scholar] [CrossRef] [Scilit] [PubMed]
5. Henry, S.H.; Bosch, F.X.; Bowers, J.C. Aflatoxin, Hepatitis and Worldwide Liver Cancer Risks. In *Mycotoxins and Food Safety*; DeVries, J.W., Trucksess, M.W., Jackson, L.S., Eds.; Advances in Experimental Medicine and Biology; Springer: Boston, MA, USA, 2002; Volume 504, pp. 229–233. ISBN 978-1-4613-5166-5. [Google Scholar]
6. Tahmasebi-Birgani, M.; Ansari, H.; Carloni, V. Defective mitosis-linked DNA damage response and chromosomal instability in liver cancer. *Biochim. Biophys. Acta BBA-Rev. Cancer* **2019**, *1872*, 60–65. [Google Scholar] [CrossRef] [Scilit]
7. *Biological Basis of Alcohol-Induced Cancer: This Proceedings Volume Will Contain Chapters Based upon the Presentation of the 2nd International Conference on Alcohol and Cancer in Colorado, 2013*; Vasiliou, V., Seitz, H.K., Eds.; Advances in Experimental Medicine and Biology; Springer: Cham, Switzerland; Heidelberg, Germany; New York, NY, USA; Dordrecht, The Netherlands; London, UK, 2015; ISBN 978-3-319-09613-1. [Google Scholar]
8. Lee, Y.-C.A.; Cohet, C.; Yang, Y.-C.; Stayner, L.; Hashibe, M.; Straif, K. Meta-analysis of epidemiologic studies on cigarette smoking and liver cancer. *Int. J. Epidemiol.* **2009**, *38*, 1497–1511. [Google Scholar] [CrossRef] [Scilit]
9. Saitta, C.; Pollicino, T.; Raimondo, G. Obesity and liver cancer. *Ann. Hepatol.* **2019**, *18*, 810–815. [Google Scholar] [CrossRef] [Scilit]
10. Gera, S.; Ettel, M.; Acosta-Gonzalez, G.; Xu, R. Clinical features, histology, and histogenesis of combined hepatocellular-cholangiocarcinoma. *World J. Hepatol.* **2017**, *9*, 300. [Google Scholar] [CrossRef] [Scilit]
11. Saffroy, R.; Pham, P.; Reffas, M.; Takka, M.; Lemoine, A.; Debuire, B. New perspectives and strategy research biomarkers for hepatocellular carcinoma. *Clin. Chem. Lab. Med.* **2007**, *45*, 9. [Google Scholar] [CrossRef] [Scilit]

12. Zhou, Y.-M.; Yang, J.-M.; Li, B.; Yin, Z.-F.; Xu, F.; Wang, B.; Liu, P.; Li, Z.-M. Clinicopathologic characteristics of intrahepatic cholangiocarcinoma in patients with positive serum a-fetoprotein. *World J. Gastroenterol.* **2008**, *14*, 2251. [Google Scholar] [CrossRef] [Scilit]

13. Yin, X.; Zhang, B.-H.; Qiu, S.-J.; Ren, Z.-G.; Zhou, J.; Chen, X.-H.; Zhou, Y.; Fan, J. Combined Hepatocellular Carcinoma and Cholangiocarcinoma: Clinical Features, Treatment Modalities, and Prognosis. *Ann. Surg. Oncol.* **2012**, *19*, 2869–2876. [Google Scholar] [CrossRef] [Scilit] [PubMed]

14. Bai, Y.; Long, J.; Liu, Z.; Lin, J.; Huang, H.; Wang, D.; Yang, X.; Miao, F.; Mao, Y.; Sang, X.; et al. Comprehensive analysis of a ceRNA network reveals potential prognostic cytoplasmic lncRNAs involved in HCC progression. *J. Cell. Physiol.* **2019**, *234*, 18837–18848. [Google Scholar] [CrossRef] [Scilit] [PubMed]

15. Calderaro, J.; Couchy, G.; Imbeaud, S.; Amaddeo, G.; Letouzé, E.; Blanc, J.-F.; Laurent, C.; Hajji, Y.; Azoulay, D.; Bioulac-Sage, P.; et al. Histological subtypes of hepatocellular carcinoma are related to gene mutations and molecular tumour classification. *J. Hepatol.* **2017**, *67*, 727–738. [Google Scholar] [CrossRef] [Scilit] [PubMed]

16. Romualdo, G.R.; Grassi, T.F.; Goto, R.L.; Tablas, M.B.; Bidinotto, L.T.; Fernandes, A.A.H.; Cogliati, B.; Barbisan, L.F. An integrative analysis of chemically-induced cirrhosis-associated hepatocarcinogenesis: Histological, biochemical and molecular features. *Toxicol. Lett.* **2017**, *281*, 84–94. [Google Scholar] [CrossRef] [Scilit] [PubMed]

17. Li, L.; Lei, Q.; Zhang, S.; Kong, L.; Qin, B. Screening and identification of key biomarkers in hepatocellular carcinoma: Evidence from bioinformatic analysis. *Oncol. Rep.* **2017**, *38*, 2607–2618. [Google Scholar] [CrossRef] [Scilit]

18. Zhao, J.; Wu, G.; Bu, F.; Lu, B.; Liang, A.; Cao, L.; Tong, X.; Lu, X.; Wu, M.; Guo, Y. Epigenetic silence of ankyrin-repeat-containing, SH3-domain-containing, and proline-rich-region- containing protein 1 (ASPP1) and ASPP2 genes promotes tumor growth in hepatitis B virus-positive hepatocellular carcinoma. *Hepatology* **2010**, *51*, 142–153. [Google Scholar] [CrossRef] [Scilit]

19. Wang, C.-H.; Wey, K.-C.; Mo, L.-R.; Chang, K.-K.; Lin, R.-C.; Kuo, J.-J. Current Trends and Recent Advances in Diagnosis, Therapy, and Prevention of Hepatocellular Carcinoma. *Asian Pac. J. Cancer Prev.* **2015**, *16*, 3595–3604. [Google Scholar] [CrossRef] [Scilit]

20. Bruix, J.; Sherman, M. Management of hepatocellular carcinoma: An update. *Hepatology* **2011**, *53*, 1020–1022. [Google Scholar] [CrossRef] [Scilit]

21. Yang, J.D.; Heimbach, J.K. New advances in the diagnosis and management of hepatocellular carcinoma. *BMJ* **2020**, m3544. [Google Scholar] [CrossRef] [Scilit]

22. Tan, G.S.Q.; Sloan, E.K.; Lambert, P.; Kirkpatrick, C.M.J.; Ilomäki, J. Drug repurposing using real-world data. *Drug Discov. Today* **2023**, *28*, 103422. [Google Scholar] [CrossRef] [Scilit]

23. Zhang, Z.-M.; Tan, J.-X.; Wang, F.; Dao, F.-Y.; Zhang, Z.-Y.; Lin, H. Early Diagnosis of Hepatocellular Carcinoma Using Machine Learning Method. *Front. Bioeng. Biotechnol.* **2020**, *8*, 254. [Google Scholar] [CrossRef] [Scilit] [PubMed]

24. SD, A.; Pasumarthi, D.; Pasha, A.; Doneti, R.; Botlagunta, M.; Pawar, S.C. Identification of Differentially Expressed Genes in Cervical Cancer Patients by Comparative Transcriptome Analysis. *BioMed Res. Int.* **2021**, *2021*, 8810074. [Google Scholar] [CrossRef] [Scilit]

25. Wang, S.; Chen, X. Identification of potential biomarkers in cervical cancer with combined public mRNA and miRNA expression microarray data analysis. *Oncol. Lett.* **2018**, *16*, 5200–5208. [Google Scholar] [CrossRef] [Scilit] [PubMed]

26. Yang, X.; Zhu, S.; Li, L.; Zhang, L.; Xian, S.; Wang, Y.; Cheng, Y. Identification of differentially expressed genes and signaling pathways in ovarian cancer by integrated bioinformatics analysis. *OncoTargets Ther.* **2018**, *11*, 1457–1474. [Google Scholar] [CrossRef] [Scilit] [PubMed]

27. Pfister D et al. NASH limits anti-tumour surveillance in immunotherapy-treated HCC. Nature 592, 450–456 (2021). [DOI] [PMC free article] [PubMed] [Google Scholar]

28. Llovet JM, Burroughs A & Bruix J Hepatocellular carcinoma. Lancet 362, 1907–1917 (2003). [DOI] [PubMed] [Google Scholar]

29. Llovet JM et al. Sorafenib in advanced hepatocellular carcinoma. N. Engl. J. Med 359, 378–390 (2008). [DOI] [PubMed] [Google Scholar]

30. Kudo M et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet 391, 1163–1173 (2018). [DOI] [PubMed] [Google Scholar]

31. Bruix J et al. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet 389, 56–66 (2017). [DOI] [PubMed] [Google Scholar]

32. Abou-Alfa GK et al. Cabozantinib in Patients with Advanced and Progressing Hepatocellular Carcinoma. N. Engl. J. Med 379, 54–63 (2018). [DOI] [PMC free article] [PubMed] [Google Scholar]

33. Zhu AX et al. Ramucirumab after sorafenib in patients with advanced hepatocellular carcinoma and increased α-fetoprotein concentrations (REACH-2): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 20, 282–296 (2019). [DOI] [PubMed] [Google Scholar]

34. Finn RS et al. Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. N. Engl. J. Med 382, 1894–1905 (2020). [DOI] [PubMed] [Google Scholar]

35. Zhu AX et al. Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. Lancet Oncol 19, (2018). [DOI] [PubMed] [Google Scholar]

36. [Press Release] FDA grants accelerated approval to nivolumab and ipilimumab combination for hepatocellular carcinoma (2020). [Google Scholar]

37. Yau T et al. Efficacy and Safety of Nivolumab Plus Ipilimumab in Patients With Advanced Hepatocellular Carcinoma Previously Treated With Sorafenib: The CheckMate 040 Randomized Clinical Trial. JAMA Oncol 6, e204564 (2020). [DOI] [PMC free article] [PubMed] [Google Scholar]

38. Llovet JM et al. Locoregional therapies in the era of molecular and immune treatments for hepatocellular carcinoma. Nat. Rev. Gastroenterol. Hepatol 18, 293–313 (2021). [DOI] [PubMed] [Google Scholar]

39. Llovet JM et al. Trial Design and Endpoints in Hepatocellular Carcinoma: AASLD Consensus Conference. Hepatology 73 Suppl 1, 158–191 (2021). [DOI] [PubMed] [Google Scholar]

40. Hernandez-Meza G et al. DNA-Methylation Profiling of Human Hepatocarcinogenesis. Hepatology 74, 183–199 (2021). [DOI] [PMC free article] [PubMed] [Google Scholar]

41. Meunier L et al. DNA Methylation Signatures Reveal the Diversity of Processes Remodeling Hepatocellular Carcinoma Methylomes. Hepatology 74, 816–834 (2021). [DOI] [PMC free article] [PubMed] [Google Scholar]

42. Hlady RA et al. Interferon drives hepatitis C virus scarring of the epigenome and creates targetable vulnerabilities following viral clearance. Hepatology hep.32111 (2021) doi: 10.1002/hep.32111. [DOI] [PMC free article] [PubMed] [Google Scholar]

43. Michalopoulos GK & Bhushan B Liver regeneration: biological and pathological mechanisms and implications. Nat Rev Gastroenterol Hepatol 18, 40–55 (2021). [DOI] [PubMed] [Google Scholar]