Diagnostic Pathology

RESEARCH

Open Access

Integrating bioinformatics and machine learning methods to analyze diagnostic biomarkers for HBV-induced hepatocellular carcinoma

Anyin Yang[1†], Jianping Liu[1†], Mengru Li[3†], Hong Zhang[1], Xulei Zhang[2*] and Lianping Wu[1*]

Abstract

Hepatocellular carcinoma (HCC) is a malignant tumor. It is estimated that approximately 50–80% of HCC cases worldwide are caused by hepatitis b virus (HBV) infection, and other pathogenic factors have been shown to promote the development of HCC when coexisting with HBV. Understanding the molecular mechanisms of HBV-induced hepatocellular carcinoma (HBV-HCC) is crucial for the prevention, diagnosis, and treatment of the disease. In this study, we analyzed the molecular mechanisms of HBV-induced HCC by combining bioinformatics and deep learning methods. Firstly, we collected a gene set related to HBV-HCC from the GEO database, performed differential analysis and WGCNA analysis to identify genes with abnormal expression in tumors and high relevance to tumors. We used three deep learning methods, Lasso, random forest, and SVM, to identify key genes RACGAP1, ECT2, and NDC80. By establishing a diagnostic model, we determined the accuracy of key genes in diagnosing HBV-HCC. In the training set, RACGAP1(AUC:0.976), ECT2(AUC:0.969), and NDC80 (AUC: 0.976) showed high accuracy. They also exhibited good accuracy in the validation set: RACGAP1(AUC:0.878), ECT2(AUC:0.731), and NDC80(AUC:0.915). The key genes were found to be highly expressed in liver cancer tissues compared to normal liver tissues, and survival analysis indicated that high expression of key genes was associated with poor prognosis in liver cancer patients. This suggests a close relationship between key genes RACGAP1, ECT2, and NDC80 and the occurrence and progression of HBV-HCC. Molecular docking results showed that the key genes could spontaneously bind to the anti-hepatocellular carcinoma drugs Lenvatinib, Regorafenib, and Sorafenib with strong binding activity. Therefore, ECT2, NDC80, and RACGAP1 may serve as potential biomarkers for the diagnosis of HBV-HCC and as targets for the development of targeted therapeutic drugs.

†Anyin Yang, Jianping Liu and Mengru Li contributed equally to this work.

*Correspondence:
Xulei Zhang
nancy85813883@163.com
Lianping Wu
wlp202403@163.com

[1]Department of Pharmacy, Gaochun People's Hospital, Gaochun Hospital Affiliated to Jiangsu University, Nanjing 211300, China
[2]Department of Liver Disease, Gaochun People's Hospital, Gaochun Hospital Affiliated to Jiangsu University, Nanjing 211300, China
[3]Department of Hospital Infection Management Section, Gaochun People's Hospital, Gaochun Hospital Affiliated to Jiangsu University, Nanjing 211300, China

Introduction

Hepatocellular carcinoma, as a malignant tumor, is the sixth most commonly diagnosed cancer and third leading cause of cancer-related deaths globally [1] and Rumgay et al. predicted that the incidence of liver cancer would increase by 55.0% and the number of deaths would increase by 56.4% between 2020 and 2040 [2]. Significant progress has been made in the epidemiology, risk factors, and molecular characteristics of HCC in many countries around the world over the past few decades. The main risk factors for hepatocellular carcinoma include chronic infection with hepatitis B virus (HBV) or hepatitis C virus (HCV), aflatoxin-contaminated food, heavy alcohol consumption, and type 2 diabetes. Chronic infection with hepatitis B virus (HBV) is considered a major risk factor for the occurrence and progression of HCC, accounting for more than half of global HCC cases [3]. In patients with hepatitis B, the incidence of hepatocellular carcinoma increases with viral load, duration of infection, and severity of liver disease [4]. Numerous studies have shown that the presence of other pathogenic factors in conjunction with HBV can increase the incidence of hepatocellular carcinoma [5].

HBV can increase the genomic instability of host cells, leading to epigenetic reshaping of host DNA, chromosomal reshaping, and abnormal expression of oncogenes and tumor suppressor genes through integration or induction of host gene mutations. It can also induce malignant transformation of liver cells by activating various cancer-related signaling pathways, regulating cell metabolism, and other mechanisms. The liver microenvironment undergoes changes induced by chronic inflammation and interactions between the virus and innate immune cells, adaptive immune cells, helping the virus evade immune surveillance and promoting the progression of the disease from inflammation to tumor formation [5]. As a high-risk factor for inducing HCC, HBV influences the occurrence and progression of tumors. Therefore, further research on the molecular mechanisms of HBV infection-induced HCC can help improve the prevention, diagnosis, and treatment of HCC.

With the continuous development of computer technology, artificial intelligence (AI) is also becoming increasingly mature. Machine learning, as a branch of AI, focuses on using mathematical algorithms to identify patterns in data for prediction. Deep learning, as a subfield of machine learning, specifically utilizes multi-layer neural network algorithms inspired by the structure of the brain for prediction [6]. Due to the increasing availability and integration of various types of data such as genomics, transcriptomics, and pathology, cancer treatment is shifting towards precision medicine. Deep learning models have the potential to identify relevant granular features from multiple data types. Deep learning is being applied in the diagnosis, prognosis, and treatment of tumors, providing meaningful insights [7].

In this study, we aim to analyze the pathogenic mechanisms of HBV-induced HCC using a combination of bioinformatics and deep learning methods. We seek to identify valuable diagnostic biomarkers and hope to provide new insights for the diagnosis and treatment of HBV-induced HCC.

Methods

Differentially expressed gene screening

Retrieve gene chip data with the keyword "HBV-HCC" from the National Center for Biotechnology Information (NCBI) public Gene Expression Omnibus (GEO) platform (https://www.ncbi.nlm.nih.gov/geo/), download data with a sample size greater than 100, filter differentially expressed genes using the R software by reading the downloaded matrix file, analyze the tumor group and control group using the "limma" package, obtain differentially expressed genes (DEGs), and screen DEGs with criteria: $|\log FC|>2$, $P<0.05$.

Functional enrichment analysis

GO function enrichment analysis and KEGG pathway enrichment analysis were performed using the ' clusterProfiler ' package of R language to discover the biological functions and pathways that DEGS may be involved in. Both of them used $P<0.05$ as the screening index.

WGCNA co-expression analysis

The 'WGCNA 'R package in R language was used to locate the co-expressed genes in the HBV-HCC dataset.The sample clustering tree algorithm is used to eliminate outlier samples, and the pick Soft Threshold function is used to select the best soft threshold β to ensure the construction of the scale-free network. The blockwiseModules function in the WGCNA package is used to construct the co-expression matrix. The merging threshold of similar modules is set to 0.25 (mergeCutHeight=0.25), the topological overlap matrix (TOM) is deepSplit=1, and the minimum number of genes in each module is set to 30 (minModuleSize=30). Other parameters are set according to the default setting. The samples in the data set were divided into control group and HBV-HCC group, and the modules with high correlation with tumors were screened out.

Screening and verification of diagnostic markers

The intersection genes of DEGS and WGCNA modules with high tumor correlation were screened out. Lasso regression, random forest and SVM-RF were used to screen the variables of the intersection genes, and the intersection of the variables screened by the three algorithms was used as a preliminary diagnostic biomarker.

Table 1 GEO array data information

Group	Data series	Platforms	Normal VS tumor
Training Data	GSE121248 GSE55092	GPL570	128 VS 119
Validation Data	GSE47197	GPL16699	63 VS 61

The relationship between biomarkers and the survival and prognosis of patients with liver cancer was analyzed in the GEPIA2 database. The 'RMS' package was used to construct a nomogram model for the diagnosis of HBV-HCC based on diagnostic biomarkers, and the clinical decision curve (DCA) was drawn. The ROC curve of diagnostic biomarkers was calculated to analyze the accuracy of prediction.

Molecular docking

The 3D structure files of Lenvatinib, Regorafenib, and Sorafenib were obtained from the PubChem database(https://pubchem.ncbi.nlm.nih.gov/) Energy minimization was conducted using the Chem3D 2019 software tool. The protein structure files were acquired from the PDB database(https://www.rcsb.org/)and preprocessing steps such as removing water molecules and small molecule ligands were carried out using PyMol software. Subsequently, Autodock Vina was employed for molecular docking, and the outcomes were visualized using PyMol software.

Results

Differential gene screening results

Download gene array datasets GSE121248 and GSE55092 from the National Center for Biotechnology Information (NCBI) public Gene Expression Omnibus (GEO) platform (https://www.ncbi.nlm.nih.gov/geo/). After merging the datasets, a total of 119 samples of HBV-HCC and 128 control samples were obtained. The merged dataset was then used as a training set for gene differential analysis, resulting in 133 downregulated genes and 64 upregulated genes. Subsequently, the dataset containing HBV-HCC-related data, GSE47197, was used as a validation set (Table 1, 2).

Co-expression gene identification results

WGCNA was used to locate co-expressed genes in the HBV-HCC dataset. The WGCNA co-expression network was constructed after calculating the optimal soft threshold ($\beta=8$) using the WGCNA package (Fig. 1B). By analyzing the correlation between genes and phenotypes, it was found that 518 genes in the black module and 102 genes in the cyan module were highly correlated with HBV-HCC. The correlation between the black module and HBV-HCC was 0.79 ($P<0.001$), and the correlation between the cyan module and HBV-HCC was 0.77 ($P<0.001$) (Fig. 1E, F).

Screening results of diagnostic biomarkers

The differential genes were intersected with the genes highly related to HBV-HCC screened by WGCNA to obtain 14 genes. Random forest, LASSO and SVM _ RF are used for variable screening, and the important variables of each machine learning screening are sorted. After the importance of all variables is sorted by random forest, the importance of RACGAP1, ECT2 and NDC80 is the first three (Fig. 2A, D). The variables retained in the LASSO regression after screening were ECT2, NDC80, CTNNA2, CCNB1, RACGAP1, VNN1, TMEM45E and ASPA (Fig. 2B, E). The most important variables selected by SVM _ RF were RACGAP1, ECT2, NDC80, CDC20 and CDK1 (Fig. 2C, F).

Core genes verification

The overlapping genes selected by the three machine learning methods were RACGAP1, NDC80, and ECT2 (Fig. 3A). In the validation set, the mRNA expression levels of RACGAP1, ECT2, and NDC80 were higher in liver cancer tissues compared to normal liver tissues. Analysis based on the TCGA database in the GEPIA2 database showed that the high expression of RACGAP1, ECT2, and NDC80 was associated with poor prognosis

Table 2 List of differential genes

Status	Gene symbol
Up	CAP2, RACGAP1, HMMR, TOP2A, NDC80, MELK, ASPM, ECT2, PRC1, ROBO1, FAM72A, BUB1B, CDK1, CCNB1, BC017398, GPR158, FAM83D, KIF20A, RRM2, DNAJC6, PBKDTL, NCAPG, GINS1, NEK2, RBM24, E2F7, TTK, CDC20, DUXAP10, LOC344887, ZIC2, NUF2, COL15A1, TRIM16, CR936796, CRNDE, SULT1C2, GPC3, CD109, FAM133A, AK093362, SMPX, NRCAM, FGF13, SSX1, LOC101930288, MAGEA1, CTNNA2, SPINK1, AKR1B10, LCN2, GABBR1, REG3A, COX7B2, MAGEA12, DKK1, MAGEA3, LOC100506403, MAGEA6, ALDH3A1, LINC01419, GAGE1, GAGE12B
Down	CLEC1B, FCN2, OIT3, CLEC4M, GPR128, CLEC4G, CXCL14, CYP26A1, CRHBP, LINC01093, RSPO3, FCN3 PLAC8, CDHR2, CCBE1, SLC25A47, CXCL12, FAM65C, LCAT MARCO, KCNN2, HAMP, CETP, GPM6A, CNDP1, TTC36, NPY1R, CYP39A1, RND3, CYP1A2, FOS, C8orf4, OLFML3, CD5L, HGF, GADD45B, IGFBP3, ESR1, IDO2, ZG16, FBP1, KMO, ASPA, IGHM, CA2, GCH1, SRPX, FOSB, NAT2, TBX15, ID1, HAO2, MT1F, CYP2B6, C7, LPA, SRD5A2, FREM2, BCO2, SPP2, DCN, IGJ, GRAMD1C, APOF, MT1H, MT1E, MT1G, IGF1, IL-13RA2, CYP4A11, GHR, PGLYRP2, SLC22A1, AKR1D1, TMEM27, PLGLB1, IGH, IGF2, TDO2, EGR1, ADH4, HGFAC, ANXA10, CYP2C18, LOC101928916, COLEC11, MT1X, CLRN3, ID4, VNN1, FOLH1B, MT1M, TMEM45A, IGHG1, ALDOB, PRG4, CYP2B6, IGLV1-44, IGLC1, HSD17B2, C9, ATF5, FAM110C, ANK3, SLCO1B3, GBA3, GNMT, HAL, CYP2B7P, BBOX1, RDH16, CYP2A6, PCK1, FGFR2, C6, CNTN3, ACOT12, AFM, GYS2, CYP2C8, SLC51A, C3P1, SLC10A1, MBL2, ADH1A, CYP3A7, LECT2, H19, FABP1, LUM, EPCAM, HPGD, CYP2E1

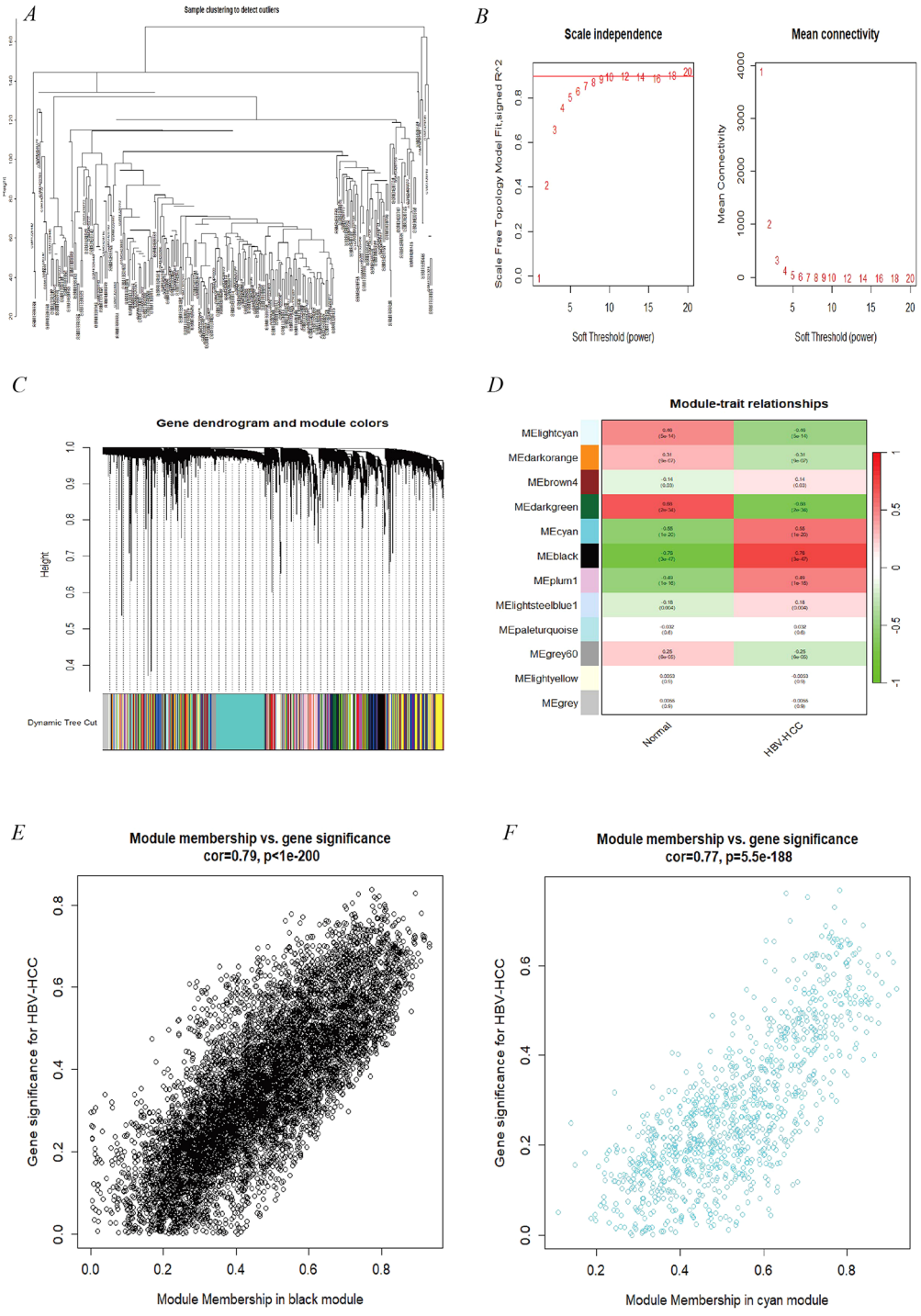


Fig. 1 WGCNA analysis results

in HCC patients (Fig. 4). A clinical diagnostic model for diagnosing HBV-HCC was constructed based on RACGAP1, ECT2, and NDC80, showing good model calibration curves (Fig. 3B). The decision curve analysis (DCA) demonstrated that patients could benefit from the clinical diagnostic model based on RACGAP1, ECT2, and NDC80 (Fig. 3C). A nomogram model based on RACGAP1, ECT2, and NDC80 was constructed and displayed as a calibration plot (Fig. 3D). In the training set, the AUC values for RACGAP1, ECT2, and NDC80 were 0.979, 0.969, and 0.976, respectively (Fig. 3E); in the validation set, the expression levels of RACGAP1, ECT2, and NDC80 were higher in liver cancer tissues compared to normal liver tissues, with AUC values of 0.878, 0.731, and

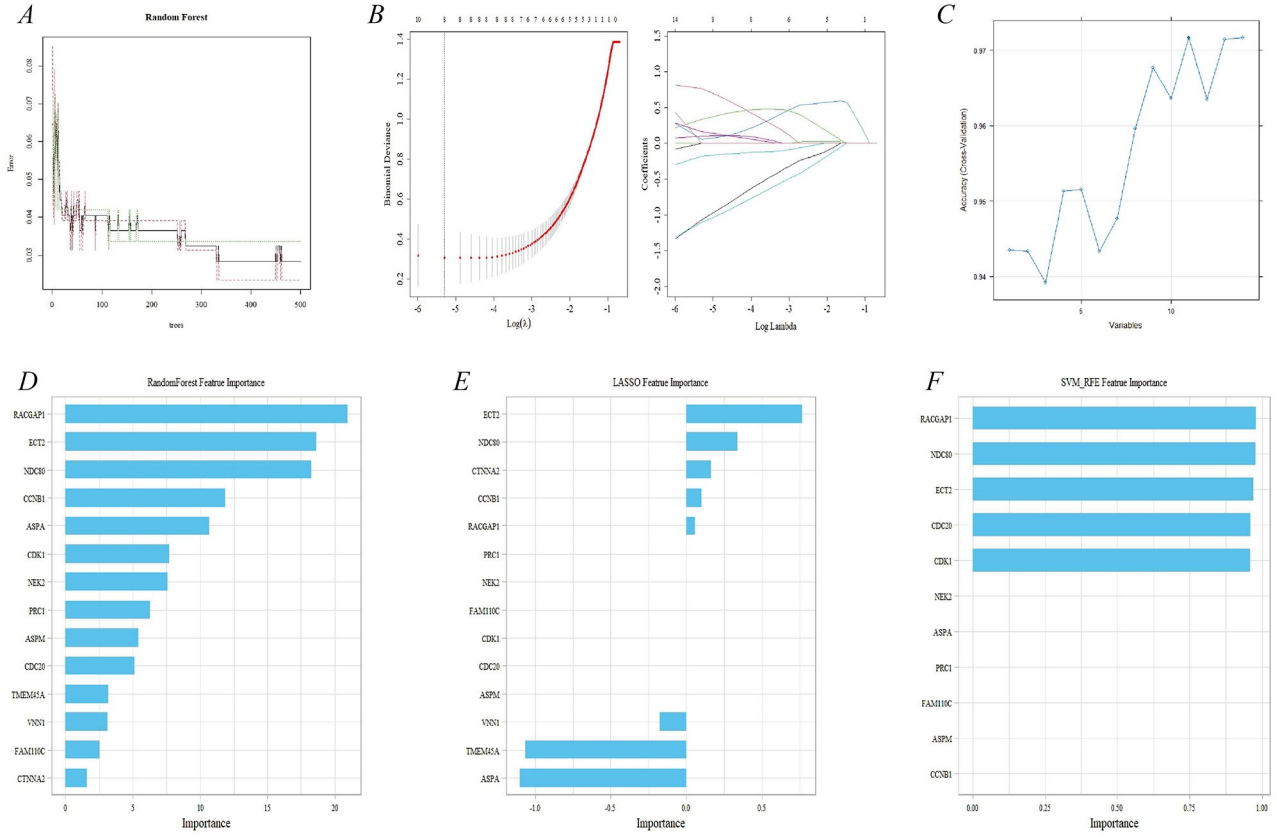


Fig. 2 Variable screening and feature importance ranking

0.915, respectively (Fig. 3J), indicating that these diagnostic biomarkers have high predictive accuracy and diagnostic value.

GO and KEGG enrichment analysis

The results of KEGG and GO enrichment analysis showed that the main enriched BPs were: nuclear chromosome segregation, chromosome segregation and sister chromatid segregation (Fig. 5A).S indle, chromosome, centromeric region and condensed chromosome were the main enriched CCs (Fig. 5B). The main enriched pathways of MF were protein serine kinase activity, microtubule binding and protein tyrosine kinase activity (Fig. 5C). The main enriched pathways were cell cycle, p53 signaling pathway, FoxO signaling pathway and Viral carcinogenesis pathway, and most of the genes in most of the pathways were up-regulated (Fig. 5D, E).

Molecular docking

By molecular docking of the core targets related to HBV-HCC with the commonly used drugs for anti-hepatocellular carcinoma, it was found that the docking energy of the docking binding configuration was less than-5 kcal / mol, which proved that the binding configuration had good activity (Table 3). Except that the binding energy of NDC80 and Lenvatinib is greater than-7 kcal / mol, the binding energy of other target proteins and drugs is less than-7 kcal / mol, which proves that the binding configuration has strong activity (Fig. 6).

Discussion

Hepatocellular carcinoma (HCC) is the most common primary liver cancer and an important medical problem. The mortality rate has increased in recent years. [8]. As a major risk factor for the occurrence and progression of HCC, HBV infection poses a threat to human life and health. HBV infection can directly or indirectly promote hepatocellular carcinogenesis. At the genetic level, HBV can increase the instability of the host cell genome, cause epigenetic remodeling of the host DNA, and lead to chromosomal remodeling and abnormal expression of oncogenes and tumor suppressor genes by integrating or inducing host gene mutations. It can also activate various cancer-related signaling pathways, regulate cell metabolism and other mechanisms to cause malignant transformation of liver cells. It is of great significance to study the specific mechanism of the occurrence and progression of HBV-HCC for its prevention and treatment. In the liver microenvironment, chronic inflammation induced by HBV infection, changes in the interaction between the

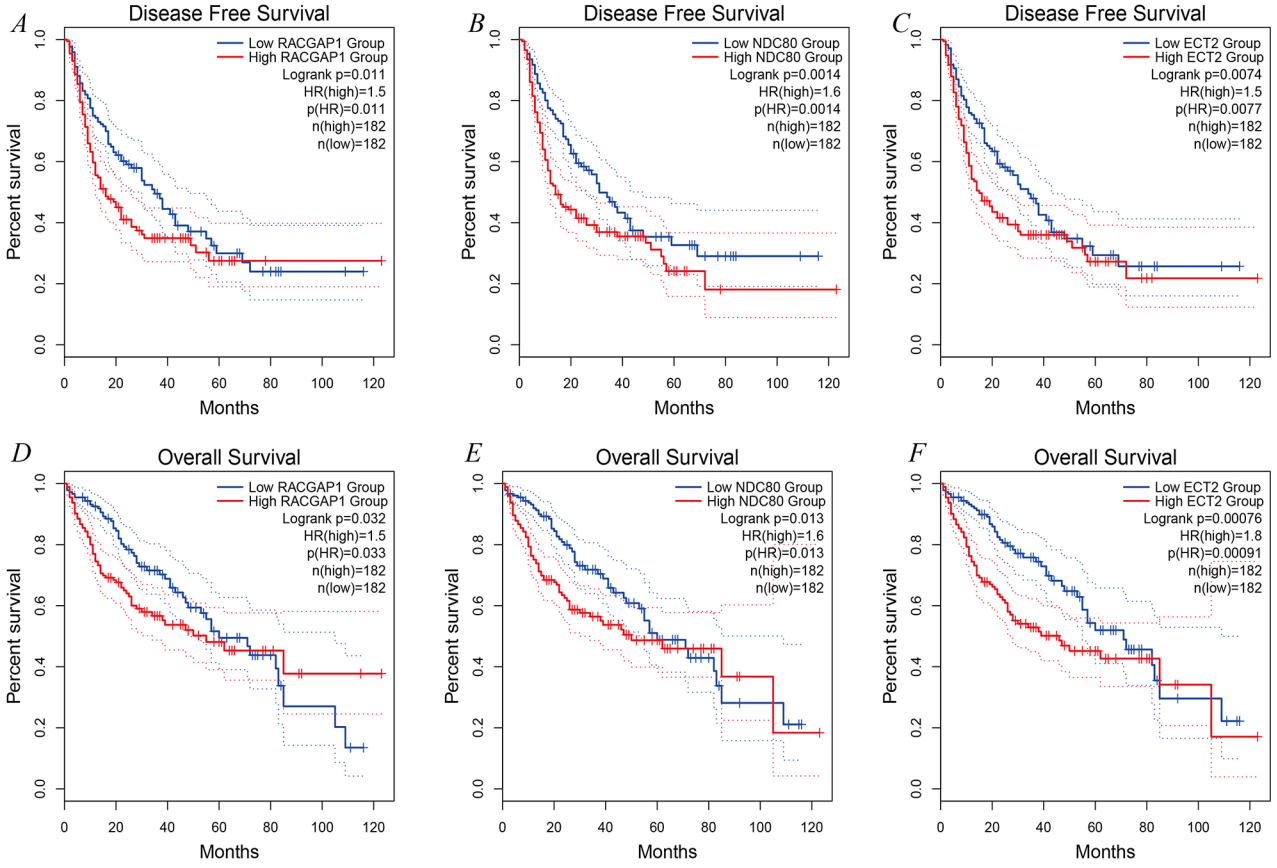
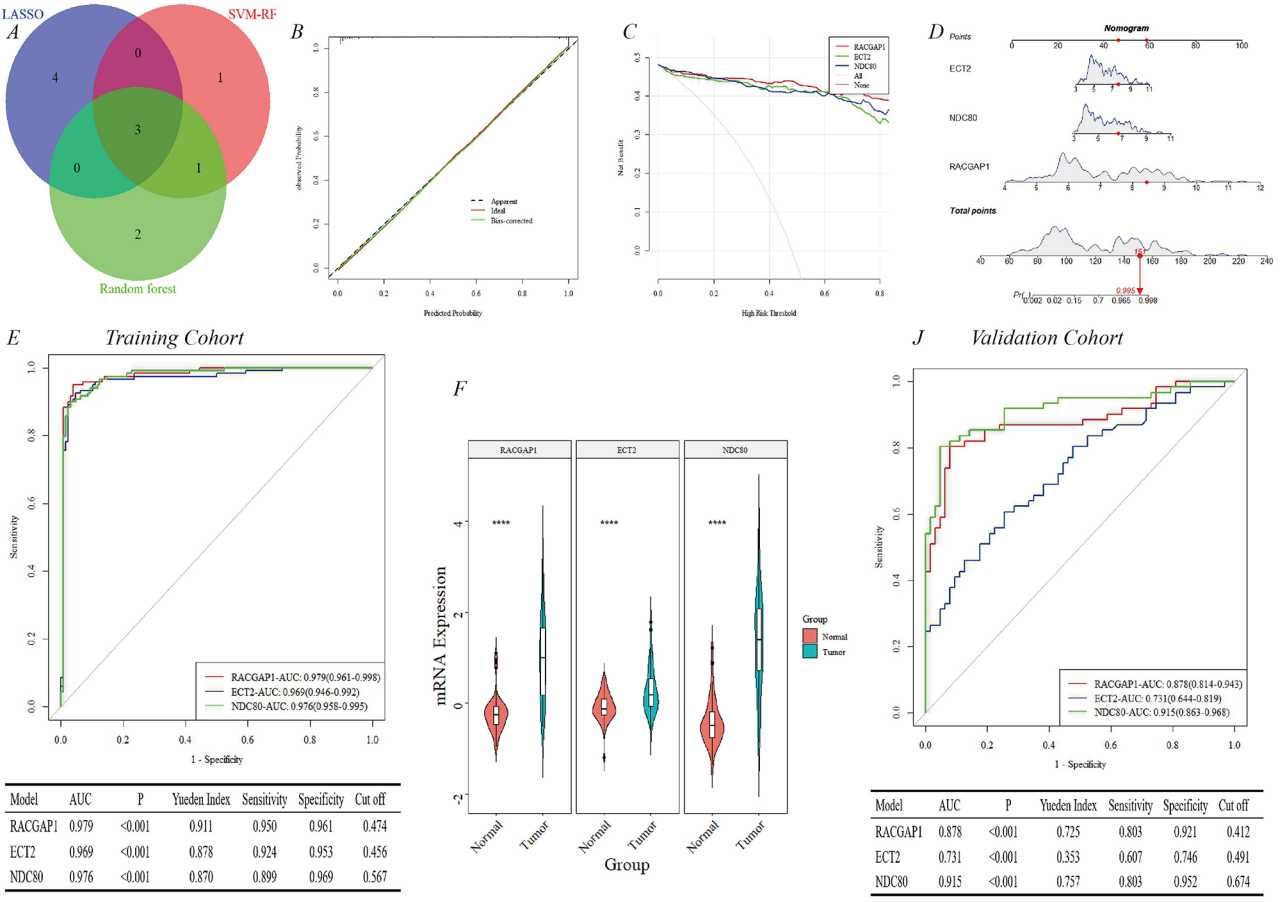


Fig. 3 Core target protein expression and survival analysis

virus and innate immune cells and adaptive immune cells help the virus evade immune surveillance and promote the evolution of the disease from inflammation to tumor formation [9]. Further study of the mechanism of HBV infection-induced HCC can provide reliable new ideas and methods for the prevention, diagnosis and treatment of HBV-HCC.

With the development of gene sequencing technology and various deep learning algorithms, it provides a method for identifying new biomarkers in diseases. In this study, three deep learning methods, random forest, Lasso regression and SVM-RF, were used to identify the key genes RACGAP1, ECT2 and NDC80 in HBV-HCC-related gene sequencing data. Through difference analysis, it was found that they were highly expressed in tumor tissues compared with normal tissues. Survival analysis showed that its high expression was associated with poor prognosis in patients with liver cancer. The selected three key genes were used to construct a clinical diagnostic model. In the training set, the key genes showed high accuracy in the diagnosis of HBV-HCC, and also had good accuracy in the validation set. It can be concluded from the DCA curve that patients can get better benefits from the model. Lenvatinib, Sorafenib and Regorafenib are clinically used drugs for the systematic treatment of hepatocellular carcinoma, which can improve the survival and prognosis of patients with hepatocellular carcinoma [10]. Through molecular docking, it was found that the docking configurations of Lenvatinib, Sorafenib and Regorafenib with RACGAP1, ECT2 and NDC80 had strong activity, indirectly support that the 3 genes are likely key players in the oncogenesis/progression of HBV-HCC, also indicating that the target may be a potential therapeutic target for hepatocellular carcinoma.

RACGAP1 is an important cellular protein. It is a GTPase-activating protein that acts on the Rho GTPase family. It belongs to the GTPase-activating protein family and participates in many cellular processes, including cell division, transformation, and invasive migration [11], Studies have found that RACGAP1 is highly expressed in a variety of cancers, such as the poor prognosis and adverse clinicopathological features of gastrointestinal stromal tumors with high expression of RACGAP1 [12]. RACGAP1 can drive breast cancer metastasis by regulating ECT2-dependent mitochondrial quality control [13]. RACGAP1 is used as a biomarker for lymphatic metastasis and poor prognosis of colon cancer [14]. In hepatocellular carcinoma, high expression of RACGAP1 promotes



Model	AUC	P	Youden Index	Sensitivity	Specificity	Cut off
RACGAP1	0.979	<0.001	0.911	0.950	0.961	0.474
ECT2	0.969	<0.001	0.878	0.924	0.953	0.456
NDC80	0.976	<0.001	0.870	0.899	0.969	0.567

Model	AUC	P	Youden Index	Sensitivity	Specificity	Cut off
RACGAP1	0.878	<0.001	0.725	0.803	0.921	0.412
ECT2	0.731	<0.001	0.353	0.607	0.746	0.491
NDC80	0.915	<0.001	0.757	0.803	0.952	0.674

Fig. 4 Core target screening and verification

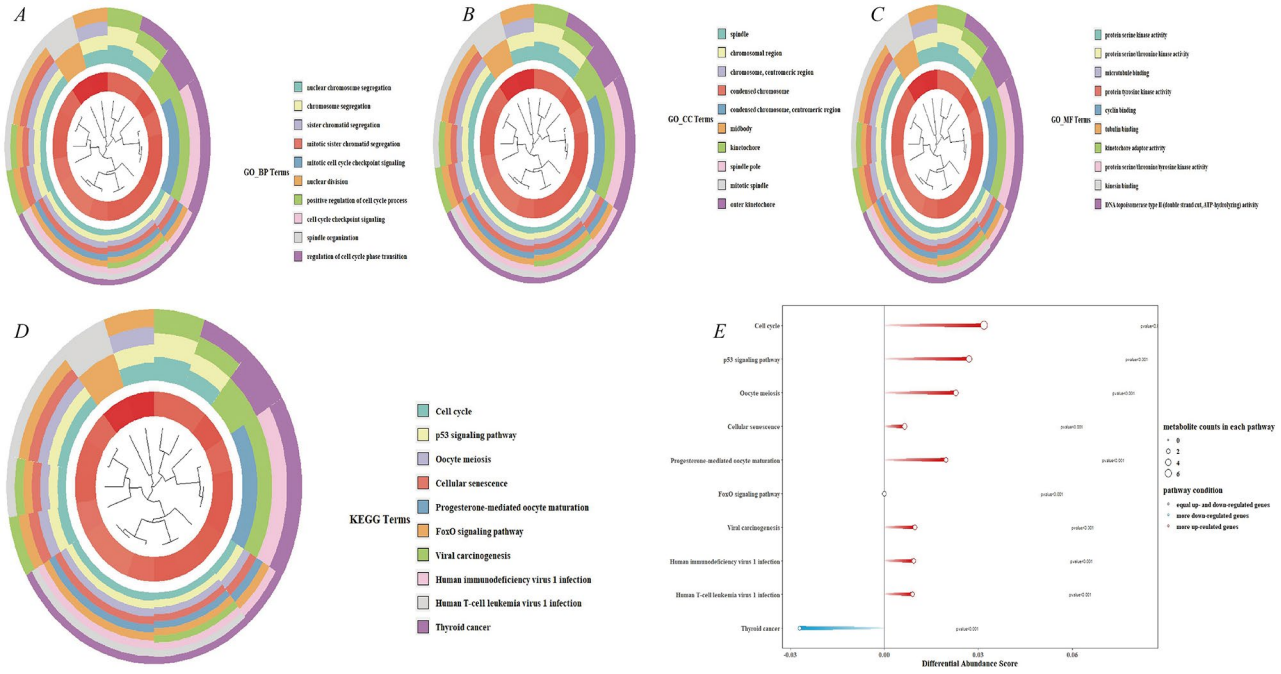


Fig. 5 KEGG and GO enrichment analysis

Table 3 Parameters of molecular docking box and docking binding energy

Target	Ligand	Grid center			NPTs			Binding energie
ECT2	Lenvatinib	59	72	72	-31.42	-20.35	-48.11	-7.3 (kcal/mol)
ECT2	Regorafenib	59	72	72	-31.42	-20.35	-48.11	-8.0 (kcal/mol)
ECT2	Sorafenib	59	72	72	-31.42	-20.35	-48.11	-8.5 (kcal/mol)
NDC80	Lenvatinib	36	47	31	3.74	27.09	44.14	-6.7 (kcal/mol)
NDC80	Regorafenib	36	47	31	3.74	27.09	44.14	-7.9 (kcal/mol)
NDC80	Sorafenib	36	47	31	3.74	27.09	44.14	-7.6 (kcal/mol)
RACGAP1	Lenvatinib	35	39	36	13.48	1.87	12.16	-7.6 (kcal/mol)
RACGAP1	Regorafenib	35	39	36	13.48	1.87	12.16	-9.3 (kcal/mol)
RACGAP1	Sorafenib	35	39	36	13.48	1.87	12.16	-9.1 (kcal/mol)

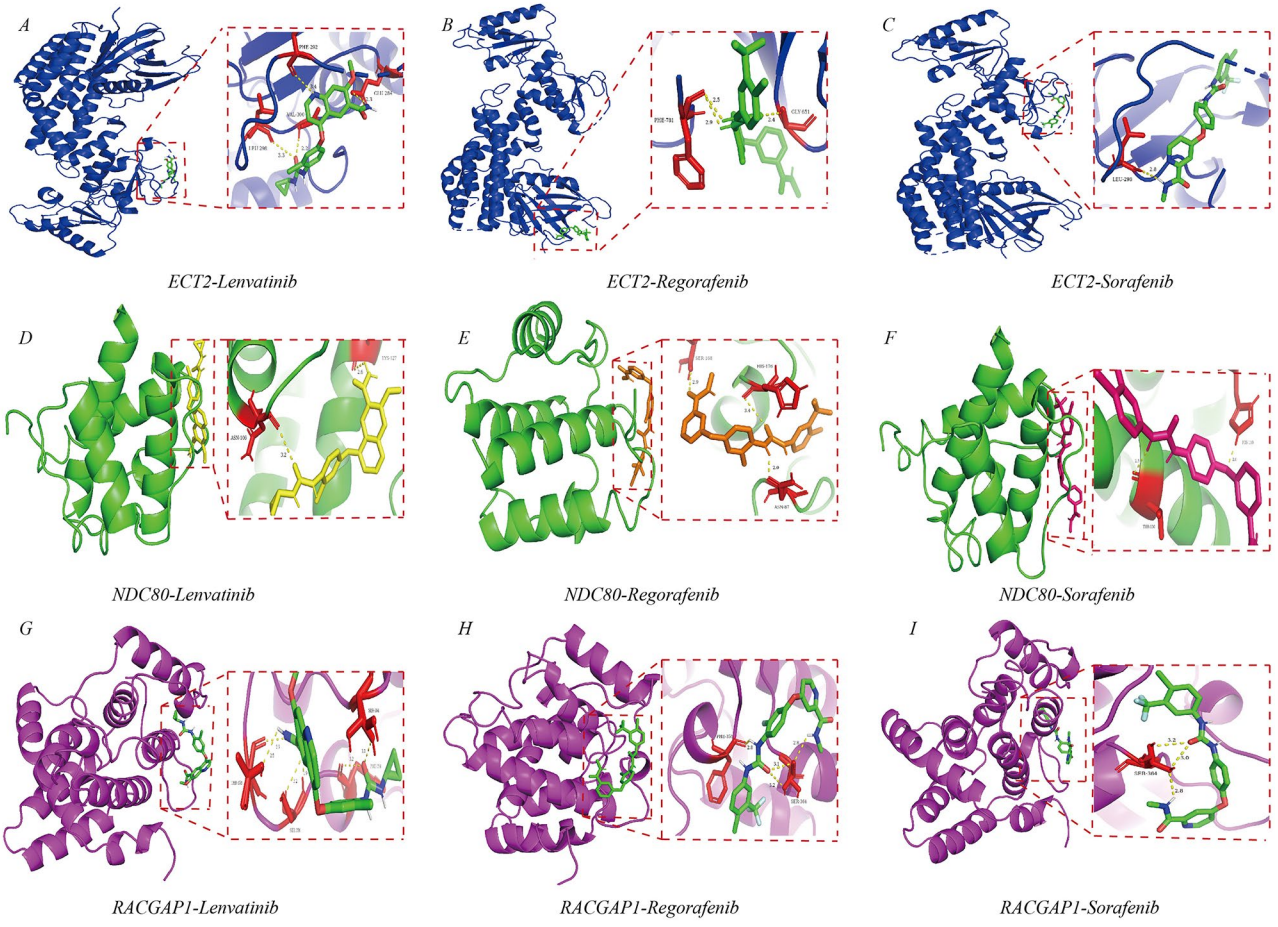


Fig. 6 Molecular docking results

tumor progression, and studies have shown that up-regulation of RACGAP1 is significantly associated with early recurrence of hepatocellular carcinoma [15, 16]. ECT2 is a guanine nucleotide dissociation exchange factor. It is a high incidence area of chromosomal abnormalities in malignant tumors. It is widely present in cells and tissues, and has the effects of regulating cell proliferation, apoptosis and DNA damage repair [17]. ECT2 has been reported to be overexpressed in a variety of human tumors, such as hepatocellular carcinoma [18], prostatic cancer [19], ovary carcinoma [20], oral cancer [21] And gastric cancer [22]. Promoting the expression of ECT2 will enhance the proliferation of HCC cells and enhance the metastasis of cancer cells. [23, 24]. NDC80 is a core component of the outer kinetochore and mitogen regulators and is involved in the migration, proliferation, invasion and apoptosis of various types of tumor cells [25, 26]. High expression of NDC80 enhances cisplatin resistance in triple-negative breast cancer [27]. Overexpression of NDC80 can lead to decreased apoptosis of HCC cells and overcome cell cycle arrest to promote the development of HCC and is associated with poor prognosis

in HCC patients [26, 28]. According to the literature, the expression of RACGAP1 [29], ECT2 [22] and NDC80 [30] can be detected in serum. This suggests that RACGAP1, ECT2, and NDC80 have the potential to serve as serum biomarkers for diagnosing and monitoring the recurrence of HBV-induced HCC in future studies.

The p53 signaling pathway plays an important role in cell cycle regulation, metabolism, aging development, reproduction and inhibition of tumor expression [31–33]. It has been found that p53, as a tumor suppressor, mutates or loses in nearly half of cancers. In the other half of the tumor, although the p53 protein is normal, the upstream regulatory factors and downstream mediators are disordered, resulting in the destruction of the entire p53 pathway [34]. Cell cycle is a highly regulated process that makes cell growth, genetic material replication and cell division possible. In the normal cell cycle, the expression of various cell cycle proteins is strictly regulated. However, in tumor cells, the mechanism of cell cycle regulation is disordered, resulting in abnormal activation of cyclins, which plays a pathogenic role in the development of most tumor types [35].

The results of GO enrichment analysis showed that the genes interacting with key genes RACGAP1, NDC80 and ECT2 were mainly enriched in chromosome-related pathways. The p53 and cell cycle pathways mainly enriched by KEGG played an important role in the occurrence and progression of cancer. Previous studies have also shown that overexpression of RACGAP1, ECT2 and NDC80 is associated with malignant progression and poor prognosis of HCC. Previous studies have shown that HBV leads to chromosomal remodeling and abnormal expression of oncogenes and tumor suppressor genes by integrating or inducing host gene mutations. It can also activate various cancer-related signaling pathways to promote the occurrence and progression of cancer [3, 5]. The clinical prediction model for the diagnosis of HBV-HCC based on RACGAP1, ECT2 and NDC80 also showed good accuracy.

In summary, according to the results of machine learning and molecular docking, we speculate that HBV may induce gene mutations in RACGAP1, ECT2 and NDC80, affect the normal function of chromosomes, affect the normal regulation of p53 and cell cycle signaling pathways, and then lead to the occurrence and progression of HCC. Moreover, NDC80, RACGAP1 and ECT2 may be valuable diagnostic biomarkers for HBV-HCC and potential therapeutic targets.

Acknowledgements
Not applicable.

Author contributions
Lianping Wu and Xulei Zhang conceived and designed the study and critically revised the manuscript. Anyin Yang, Jianping Liu, Mengru Li performed the experiments, analyzed the data, and drafted the manuscript, Yang Anyin, Liu Jianping, Li Mengru have a common contribution to the article. Hong Zhong participated in study design, study implementation and manuscript revision.

Funding
This work was no funding.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Competing interests
The authors declare no competing interests.

Received: 21 June 2024 / Accepted: 21 July 2024
Published online: 02 August 2024

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–63. https://doi.org/10.3322/caac.21834.
2. Rumgay H, Arnold M, Ferlay J, Lesi O, Cabasag CJ, Vignat J, et al. Global burden of primary liver cancer in 2020 and predictions to 2040. J Hepatol. 2022;77(6):1598–606. https://doi.org/10.1016/j.jhep.2022.08.021.
3. Jia L, Gao Y, He Y, Hooper JD, Yang P. HBV induced hepatocellular carcinoma and related potential immunotherapy. Pharmacol Res. 2020;159:104992. https://doi.org/10.1016/j.phrs.2020.104992.
4. Chen CJ, Yang HI, Su J, Jen CL, You SL, Lu SN, et al. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. JAMA. 2006;295(1):65–73. https://doi.org/10.1001/jama.295.1.65.
5. Jiang Y, Han Q, Zhao H, Zhang J. The mechanisms of HBV-Induced Hepatocellular Carcinoma. J Hepatocell Carcinoma. 2021;8:435–50. https://doi.org/10.2147/jhc.S307962.
6. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–44. https://doi.org/10.1038/nature14539.
7. Kleppe A, Skrede OJ, De Raedt S, Liestøl K, Kerr DJ, Danielsen HE. Designing deep learning studies in cancer diagnostics. Nat Rev Cancer. 2021;21(3):199–211. https://doi.org/10.1038/s41568-020-00327-9.
8. Vogel A, Meyer T, Sapisochin G, Salem R, Saborowski A. Hepatocellular carcinoma. Lancet. 2022;400(10360):1345–62. https://doi.org/10.1016/s0140-6736(22)01200-4.
9. Levrero M, Zucman-Rossi J. Mechanisms of HBV-induced hepatocellular carcinoma. J Hepatol. 2016;64(1 Suppl):S84–101. https://doi.org/10.1016/j.jhep.2016.02.021.
10. Villanueva A. Hepatocellular Carcinoma. N Engl J Med. 2019;380(15):1450–62. https://doi.org/10.1056/NEJMra1713263.
11. Saigusa S, Tanaka K, Mohri Y, Ohi M, Shimura T, Kitajima T, et al. Clinical significance of RacGAP1 expression at the invasive front of gastric cancer. Gastric Cancer. 2015;18(1):84–92. https://doi.org/10.1007/s10120-014-0355-1.
12. Imaoka H, Toiyama Y, Saigusa S, Kawamura M, Kawamoto A, Okugawa Y, et al. RacGAP1 expression, increasing tumor malignant potential, as a predictive biomarker for lymph node metastasis and poor prognosis in colorectal cancer. Carcinogenesis. 2015;36(3):346–54. https://doi.org/10.1093/carcin/bgu327.
13. Yin C, Toiyama Y, Okugawa Y, Shigemori T, Yamamoto A, Ide S, et al. Rac GTPase-Activating protein 1 (RACGAP1) as an oncogenic enhancer in esophageal carcinoma. Oncology. 2019;97(3):155–63. https://doi.org/10.1159/000500592.
14. Ren K, Zhou D, Wang M, Li E, Hou C, Su Y, et al. RACGAP1 modulates ECT2-Dependent mitochondrial quality control to drive breast cancer metastasis. Exp Cell Res. 2021;400(1):112493. https://doi.org/10.1016/j.yexcr.2021.112493.
15. Gu Y, Chen B, Guo D, Pan L, Luo X, Tang J, et al. Up-Regulation of RACGAP1 promotes progressions of Hepatocellular Carcinoma regulated by GABPA via PI3K/AKT pathway. Oxid Med Cell Longev. 2022;2022(3034150). https://doi.org/10.1155/2022/3034150.
16. Zabady S, Mahran N, Soltan MA, Alaa Eldeen M, Eid RA, Albogami S, et al. Cyanidin-3-Glucoside modulates hsa_circ_0001345/miRNA106b/ATG16L1 Axis expression as a potential protective mechanism against Hepatocellular

Carcinoma. Curr Issues Mol Biol. 2022;44(4):1677–87. https://doi.org/10.3390/cimb44040115.
17. Xu D, Wang Y, Wu J, Zhang Z, Chen J, Xie M, et al. ECT2 overexpression promotes the polarization of tumor-associated macrophages in hepatocellular carcinoma via the ECT2/PLK1/PTEN pathway. Cell Death Dis. 2021;12(2):162. https://doi.org/10.1038/s41419-021-03450-z.
18. Chen J, Xia H, Zhang X, Karthik S, Pratap SV, Ooi LL, et al. ECT2 regulates the Rho/ERK signalling axis to promote early recurrence in human hepatocellular carcinoma. J Hepatol. 2015;62(6):1287–95. https://doi.org/10.1016/j.jhep.2015.01.014.
19. Guo Z, Chen X, Du T, Zhu D, Lai Y, Dong W, et al. Elevated levels of epithelial cell transforming sequence 2 predicts poor prognosis for prostate cancer. Med Oncol. 2017;34(1):13. https://doi.org/10.1007/s12032-016-0872-3.
20. Huff LP, Decristo MJ, Trembath D, Kuan PF, Yim M, Liu J, et al. The role of Ect2 Nuclear RhoGEF activity in Ovarian Cancer Cell Transformation. Genes Cancer. 2013;4(11–12). https://doi.org/10.1177/1947601913514851. 460 – 75.
21. Iyoda M, Kasamatsu A, Ishigami T, Nakashima D, Endo-Sakamoto Y, Ogawara K, et al. Epithelial cell transforming sequence 2 in human oral cancer. PLoS ONE. 2010;5(11):e14082. https://doi.org/10.1371/journal.pone.0014082.
22. Wang HB, Yan HC, Liu Y. Clinical significance of ECT2 expression in tissue and serum of gastric cancer patients. Clin Transl Oncol. 2016;18(7):735–42. https://doi.org/10.1007/s12094-015-1428-2.
23. He CK, Li ZB, Yi D, Zhu XY, Liu RR, Zhang DX, et al. LncRNA FGD5-AS1 enhances the proliferation and stemness of hepatocellular carcinoma cells through targeting miR-223 and regulating the expression of ECT2 and FAT1. Hepatol Res. 2022;52(7):614–29. https://doi.org/10.1111/hepr.13767.
24. Lv E, Sheng J, Yu C, Rao D, Huang W. Long noncoding RNA MAPKAPK5-AS1 promotes metastasis through regulation miR-376b-5p/ECT2 axis in hepatocellular carcinoma. Dig Liver Dis. 2023;55(7):945–54. https://doi.org/10.1016/j.dld.2022.11.024.
25. Chen J, Ünal E. Meiotic regulation of the Ndc80 complex composition and function. Curr Genet. 2021;67(4):511–8. https://doi.org/10.1007/s00294-021-01174-3.
26. Ju LL, Chen L, Li JH, Wang YF, Lu RJ, Bian ZL, et al. Effect of NDC80 in human hepatocellular carcinoma. World J Gastroenterol. 2017;23(20):3675–83. https://doi.org/10.3748/wjg.v23.i20.3675.
27. Li J, Xu X, Peng X. NDC80 enhances cisplatin-resistance in Triple-negative breast Cancer. Arch Med Res. 2022;53(4):378–87. https://doi.org/10.1016/j.arcmed.2022.03.003.
28. Xie W, Wang B, Wang X, Hou D, Su H, Huang H. Nine hub genes related to the prognosis of HBV-positive hepatocellular carcinoma identified by protein interaction analysis. Ann Transl Med. 2020;8(7):478. https://doi.org/10.21037/atm.2020.03.94.
29. Wontakal SN, Britto M, Zhang H, Han Y, Gao C, Tannenbaum S, et al. RACGAP1 variants in a sporadic case of CDA III implicate the dysfunction of centralspindlin as the basis of the disease. Blood. 2022;139(9):1413–8. https://doi.org/10.1182/blood.2021012334.
30. Klatt MG, Dao T, Yang Z, Liu J, Mun SS, Dacek MM, et al. A TCR mimic CAR T cell specific for NDC80 is broadly reactive with solid tumors and hematologic malignancies. Blood. 2022;140(8):861–74. https://doi.org/10.1182/blood.2021012882.
31. Hu W, Feng Z, Teresky AK, Levine AJ. p53 regulates maternal reproduction through LIF. Nature. 2007;450(7170):721–4. https://doi.org/10.1038/nature05993.
32. Kruiswijk F, Labuschagne CF, Vousden KH. p53 in survival, death and metabolic health: a lifeguard with a licence to kill. Nat Rev Mol Cell Biol. 2015;16(7):393–405. https://doi.org/10.1038/nrm4007.
33. Tanikawa C, Zhang YZ, Yamamoto R, Tsuda Y, Tanaka M, Funauchi Y, et al. The Transcriptional Landscape of p53 signalling pathway. EBioMedicine. 2017;20:109–19. https://doi.org/10.1016/j.ebiom.2017.05.017.
34. Huang J. Current developments of targeting the p53 signaling pathway for cancer treatment. Pharmacol Ther. 2021;220:107720. https://doi.org/10.1016/j.pharmthera.2020.107720.
35. Suski JM, Braun M, Strmiska V, Sicinski P. Targeting cell-cycle machinery in cancer. Cancer Cell. 2021;39(6):759–78. https://doi.org/10.1016/j.ccell.2021.03.010.

Publisher's Note