



# A Deep Look into the Program of Rapid Tumor Growth of Hepatocellular Carcinoma

Jie Wang<sup>+1,2</sup>, Yi Lou<sup>+3</sup>, Jianmin Lu<sup>4</sup>, Yuxiao Luo<sup>4</sup>, Anqian Lu<sup>1</sup>, Anna Chen<sup>1</sup>, Jiantao Fu<sup>1</sup>, Jing Liu<sup>1,2</sup>, Xiang Zhou<sup>\*1,2</sup> and Jin Yang<sup>\*1,2</sup>

<sup>1</sup>Department of Translational Medicine, Affiliated Hospital of Hangzhou Normal University, Institute of Hepatology and Metabolic Diseases of Hangzhou Normal University, Hangzhou Normal University, Hangzhou, Zhejiang, China; <sup>2</sup>Department of Liver Disease, Affiliated Hospital of Hangzhou Normal University, Hangzhou Normal University, Hangzhou, Zhejiang, China; <sup>3</sup>Department of Occupational Medicine, Hangzhou Red Cross Hospital, Zhejiang Provincial Integrated Chinese and Western Medicine Hospital, Hangzhou, Zhejiang, China; <sup>4</sup>Department of Orthopedics, Affiliated Hospital of Hangzhou Normal University, Hangzhou Normal University, Hangzhou, Zhejiang, China

## Abstract

**Background and Aims:** Great efforts have been made towards increasing our understanding of the pathogenesis involved in hepatocellular carcinoma (HCC), but the rapid growth inherent to such tumor development remains to be explored. **Methods:** We identified distinct gene coexpression modes upon liver tumor growth using weighted gene coexpression network analysis. Modeling of tumor growth as signaling activity was employed to understand the main cascades responsible for the growth. Hub genes in the modules were determined, examined *in vitro*, and further assembled into the growth signature. **Results:** We revealed modules related to the different growth states in HCC, especially the fastest growth module, which is preserved among different HCC cohorts. Moreover, signaling flux in the cell cycle pathway was found to act as a driving force for rapid growth. Twenty hub genes in the module were identified and assembled into the growth signature, and two genes (*NCAPH*, and *RAD54L*) were tested for their growth potential *in vitro*. Genetic alteration of the growth signature affected the global gene expression. The activity of the signature was associated with tumor metabolism and immunity in HCC. Finally, the prognosis effect of the growth signature was reproduced in nine cancers. **Conclusions:** These results collectively demonstrate the molecule organization of rapid tumor growth in HCC, which is a highly synergistic process, with implications for the future management of patients.

**Keywords:** Hepatocellular carcinoma; Tumor growth; Coexpression network; Cell cycle; Metabolism.

**Abbreviations:** CNV, copy number variation; GEO, Gene Expression Omnibus; GS, gene significance; KEGG, Kyoto Encyclopedia of Genes and Genomes; LIHC, liver hepatocellular carcinoma; ME, module eigengene; MM, module membership; NES, normalized enrichment score; OS, overall survival; PCA, principal components analysis; SCNA, somatic copy-number alternations; sh, short hairpin; ssGSEA, single-sample gene-set enrichment analysis; TCGA, The Cancer Genome Atlas; TMB, tumor mutational burden; WGCNA, weight gene coexpression network analysis.

Received: 7 September 2020; Revised: 12 November 2020; Accepted: 1 December 2020

<sup>†</sup>These authors contributed equally to this work.

\*Correspondence to: Jin Yang, Department of Translational Medicine, Affiliated Hospital of Hangzhou Normal University, Institute of Hepatology and Metabolic Diseases of Hangzhou Normal University, Hangzhou Normal University, Hangzhou, Zhejiang 310015, China. Tel: +86-571-88358062, E-mail: hz\_zhiy@163.com; Xiang Zhou, Department of Liver Disease, Affiliated Hospital of Hangzhou Normal University, Hangzhou Normal University, Hangzhou, Zhejiang 310015, China. Tel: +86-571-88303403, E-mail: heijack@163.com

**Citation of this article:** Wang J, Lou Y, Lu J, Luo Y, Lu A, Chen A, et al. A deep look into the program of rapid tumor growth of hepatocellular carcinoma. *J Clin Transl Hepatol* 2021;9(1):22–31. doi: 10.14218/JCTH.2020.00084.

## Introduction

Hepatocellular carcinoma (HCC) represents a leading cause of cancer-related death worldwide.<sup>1</sup> Currently, though several staging systems can stratify the HCC patients into appropriate risk categories, a great deal of divergence remains within each risk group, due to the molecular heterogeneity of tumor cells and microenvironment. One way to describe the progression of HCC is tumor growth, since it might hold for the long-held view that rapidly growing tumors are more likely to metastasize and become lethal than slow-growing tumors;<sup>2</sup> although, the fundamental question regarding the ability of tumor cells to rapidly grow remains to be answered.

In fact, tumor cells adapt to changing environmental conditions and profoundly shape the dependencies of individual cells. For instance, through aerobic glycolysis, cancer cells produce energy by taking up glucose at much higher rates than other cells, while, at the same time, using a smaller fraction of the glucose for energy production. This allows cancer cells to function more like fetal cells, promoting extremely rapid growth.<sup>3</sup> However, the underlying molecular basis of the intertwined interactions among tumor immunology, oncogenic signaling, and tissue/biochemical context, upon tumor growth remains largely unknown.<sup>4</sup>

Currently, a robust gene coexpression network for the mining of hub genes that drive pivotal signaling pathways in terms of large-scale gene expression profiles can be built through weighted gene coexpression network analysis (WGCNA).<sup>3</sup> Previous studies have applied WGCNA to provide functional explanations of systems biology, proposing candidate therapeutic targets or diagnostic biomarkers for cancer in recent years.<sup>5</sup> For example, Zhao *et al.*<sup>6</sup> utilized WGCNA to investigate the relationships underlying the molecular and clinical characteristics of cholangiocarcinoma. However, a robust WGCNA network for cancer growth has not yet been established.

In the present study, the HCC transcriptome and tumor growth-related modules were explored by WGCNA. Focusing on rapid tumor growth, the integrative functional analy-

**Table 1. Hub genes in the HCC rapid growth module**

Top10 GS	GS.Fastest	p.GS.Fastest	Top10 K.in	K.in	Top10 MM	MM.green	p.MM.green
CCNA2	0.5795	7.97E-16	CDCA5	51.7516	CCNB1	0.9639	2.53E-93
NDC80	0.5733	1.89E-15	CCNB1	50.1223	CDCA5	0.9611	8.49E-91
CDCA8	0.5700	2.99E-15	PRC1	50.1064	PRC1	0.9503	1.69E-82
CENPE	0.5580	1.48E-14	CDK1	48.4094	CDK1	0.9492	9.73E-82
KIF11	0.5576	1.55E-14	EXO1	47.6809	EXO1	0.9478	7.54E-81
H2AFX	0.5571	1.66E-14	BIRC5	46.9120	DLGAP5	0.9431	6.61E-78
RAD54L	0.5534	2.69E-14	DLGAP5	46.5064	CENPF	0.9404	2.24E-76
NCAPH	0.5520	3.20E-14	CENPF	44.2346	BIRC5	0.9354	1.11E-73
RFC4	0.5495	4.42E-14	PTTG1	44.0655	PTTG1	0.9352	1.32E-73
HJURP	0.5491	4.61E-14	CDCA8	43.8984	NUF2	0.9350	1.83E-73

sis was expanded to the levels of the growth signature, associated molecular events, and corresponding modulations.

## Methods

### Data preparation

The transcription profile of HCC was downloaded from the Gene Expression Omnibus (commonly referred to as GEO) with accession number GSE54236, which includes 78 primary HCC tumor samples representing the different speeds of tumor growth.<sup>7</sup> Briefly, patients underwent two computed tomography scans 6 weeks apart to determine tumor volumes and HCC doubling time, which ranged from 30 to 621 days and were divided into the following quartiles: ≤53 days ( $n=19$ ), 54–82 days ( $n=20$ ), 83–110 days ( $n=20$ ), and ≥111 days ( $n=19$ ). Based on these quartiles, tumor growth was classified into slow, fast, faster, and fastest states, respectively. Low and non-expressed genes were removed by selecting probes with a mean expression in the top 50% of all probes. Next, genes with expression variance above average level were selected. Different probes targeting the same gene were collapsed. These steps finally resulted in 5511 genes to infer coexpression networks.

In addition, GSE14520, GSE25097, GSE62232, GSE36376 datasets were obtained from the GEO database. RNA-seq expression profiles from nine cancer types, including adrenocortical carcinoma, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma, brain lower grade glioma, liver hepatocellular carcinoma (LIHC), lung adenocarcinoma, mesothelioma, pancreatic adenocarcinoma, and sarcoma, were obtained from The Cancer Genome Atlas (commonly known as TCGA) database. Detailed information about the datasets is shown in Supplementary Table 1.

### Coexpression network construction

We constructed the coexpression network using the WGCNA package.<sup>8</sup> Briefly, the steps included: (1) defining the similarity matrix; (2) choosing the soft threshold,  $\beta$ , and inferring the adjacency matrix; (3) defining the topological overlap matrix; (4) performing hierarchical clustering; (5) performing the dynamic tree cut method to identify the modules; and (6) computing the module eigengene (ME) of each module. The ME can be considered as a representative of the gene expression profiles in a module. The average-linkage hierarchical clustering method was employed

to cluster the MEs of all modules, and the modules with high similarity were merged to obtain the coexpression network.<sup>9</sup> Another tool, the CEMITool package, was used to validate the gene modules, as described previously.<sup>10</sup>

The module preservation statistic Zsummary was used to assess the overlap between network modules,<sup>8</sup> which takes into account the overlap in module membership (MM), the density (mean connectivity) and connectivity (sum of connections) patterns of modules. A module was considered not being preserved if preservation Zsummary < 2, moderately preserved if  $2 \leq Zsummary < 10$ , and highly preserved if Zsummary ≥ 10.

### Identification of hub genes and growth signature

Hub genes (genes that are highly interconnected with the nodes of the module) are of functional importance. MM was defined as the correlation between the ME and gene expression values. The MM measure is highly related to intramodular connectivity (K.in).<sup>8</sup> Highly connected intramodular hub genes tend to have high MM values to the respective module. In short, the larger the MM value of the gene, the higher the correlation between the gene and a given module. In addition, the gene significance (GS) was defined as mediated  $p$ -value of each gene ( $GS = \lg P$ ) in the linear regression between gene expression and the clinical traits.<sup>5</sup>

We used the network screening function based on GS (representing the correlation between the gene and a given clinical trait) and MM or K.in (representing the correlation between the gene and a given module) in the WGCNA package to directly identify the top hub genes in the fastest growth module, and further assembled them into the growth signature. The growth activity was quantified by applying the single-sample gene-set enrichment analysis (ssGSEA).<sup>6</sup> We defined the growth signature as either high or low by using median cut-off.

### Functional annotation

Functional annotations of the gene sets were performed using webgestalt<sup>11</sup> or Enrichr.<sup>12</sup>

### Pathway activity computation

Using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways as templates for signal simulation, a ca-

nonical circuit was defined as any possible route the signal can traverse to be transmitted from a particular input to a specific output node.<sup>13</sup> The computation of the signal intensity across the different circuits of the entire pathways was performed by the Hipathia program.<sup>13</sup> In addition, the growth signature associated cancer pathway activity in pancreatic cancer was performed using GSCALite.<sup>14</sup>

### Modeling genetic-gene expression

Multi-layered profiles for DNA copy numbers, mRNA expression, and mutations in the LIHC data were obtained from the Xena portal (<http://xena.ucsc.edu/>). A linear modeling approach that measured the association of expression levels was used on a gene-by-gene basis with a number of potential predictors, including gene mutations or genomic alterations, as previously described.<sup>15</sup> Somatic mutations and genomic alterations were presented by using Maftools<sup>16</sup> and encoded as being present/absent. Linear expression models were fit with the Limma package.<sup>17</sup> For the expression of gene  $k$  in patient  $i$ ,  $Y_{ik}$  is fitted by the following equation:

$$Y_{ik} = \sum_{j=1} X_{ij} \beta_{jk} + \beta_{0k} + \varepsilon$$

$X_{ij}$  is the mutation matrix for patient  $i$  and mutation  $j$ , with entries  $X_{ij}=1$  denoting patient  $i$  has a mutation  $j$  and 0 otherwise. The coefficients  $\beta_{jk}$  measure the expression change in gene  $k$  induced by the presence of a mutation  $j$ . The entry  $\beta_{0k}$  implies the baseline expression level of gene  $k$ .<sup>15</sup>

### Immune and metabolism signatures

For each tumor sample, ESTIMATE was used to assess tumor purity.<sup>18</sup> The cytolytic activity or the interferon score of the local immune infiltrate was calculated as previously described.<sup>19</sup> Gene signatures of 28 tumor-infiltrating lymphocytes, including CD8 or CD4 T cells, B cells, natural killer cells, as well as markers from multiple types of oncoimmunology-containing genes were referenced in a prior study.<sup>20</sup> The ssGSEA method was utilized to quantify the enrichment levels of metabolism-related signatures based on Reactome (<https://reactome.org>) gene-sets.

### Cellular experiments

The human liver cancer cell line HepG2 was purchased from the Cell Bank of Chinese Academy of Sciences (Shanghai, China). Cells were cultured in DMEM (Gibco, Life Technologies, Waltham, MA, USA) supplemented with 10% fetal bovine serum (Gibco), and used within 20 passages after culture.

Lentivirus production was performed as previously described.<sup>21</sup> The targeting sequences of short hairpin (sh)RNAs were as follows: 5'-CACCGAACCAACCAACTTTAA-3' for sh-NCAPH#1, 5'-ACTGACTCACCTCGCTTATTG-3' for sh-NCAPH#2; and 5'-CTTTGTAATCATCCAGCTCTA-3' for sh-RAD54L#1, 5'-TGGTCTGGGTGTAGCTCTTAG-3' for sh-RAD54L#2. Knockdown efficiency was verified by quantitative PCR. The Cell Counting Kit-8 (Dojindo Molecular Technologies, Kumamoto, Japan) assays were used to measure cell proliferation.

### Statistical analysis

Experimental data are represented as the average  $\pm$  stand-

ard deviation. Unless otherwise indicated, the Student's two-tailed unpaired  $t$ -test was used to determine statistical significance. The significance threshold was set at 0.05. For survival analysis, the LIHC data were analyzed with the GEPIA database.<sup>22</sup> Samples were split into high and low groups based on the median value. Kaplan-Meier survival analysis was calculated using the log rank test, with a  $p$  value for significance of  $<0.05$ .

## Results

### Rapid growth module in HCC

In defining the gene clusters involved in HCC growth, 11 distinct gene modules were explored using WGCNA, as shown in Fig. 1A.

Module stability was verified by repeating network construction and module identification on expression data that consists of resampled sets of the original dataset or alternatively by another tool (the CEMITool) running the same dataset<sup>10</sup> (Supplementary Fig. 1). The results proved the robustness of module assignments.

Next, we evaluated the relationship between each module and the growth status. Notably, the cirrhotic features of modules (red, blue, yellow, brown) and the aggressive proliferative HCC features of modules (magenta and green for the faster and fastest state respectively) were identified (Fig. 1B and C). Network features such as GS, MM and K.in<sup>23</sup> of each module in different growth states were computed (Supplementary Table 2).

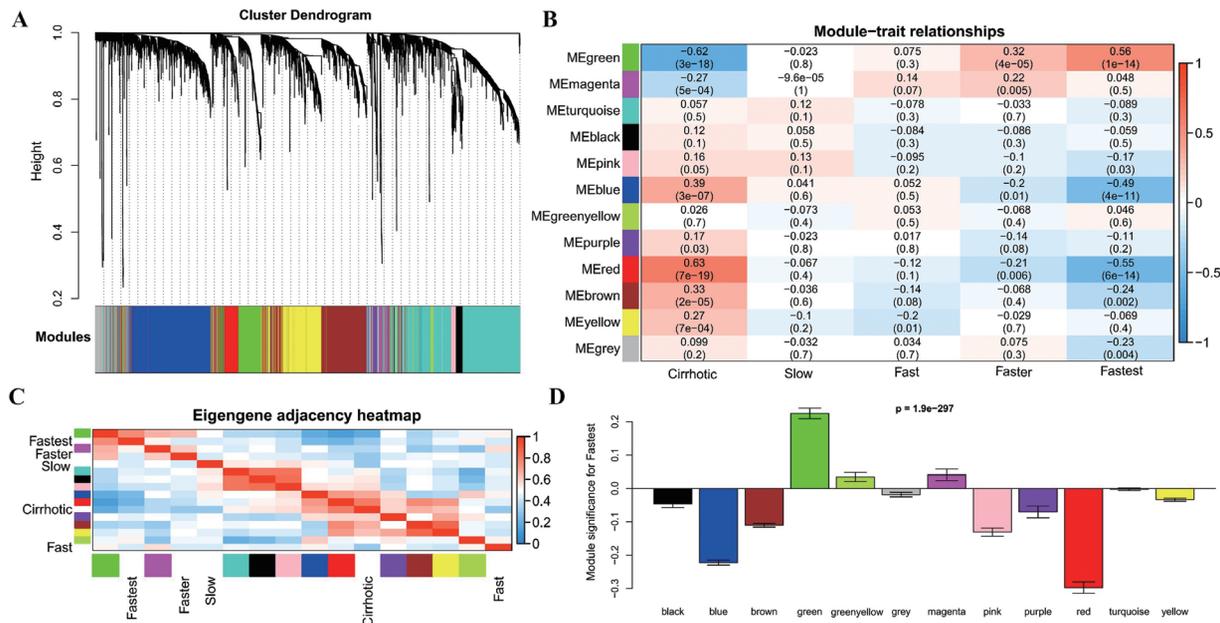
Moreover, we narrowed down our analyses to the fastest HCC growth. Module green (referred to as the growth module hereafter) was the best, as reflected by its strongly positive correlations to the fastest tumor growth (Fig. 1B-D).

### Signal fluxes in the rapid growth module

To ask the question of whether the growth module was highly preserved across independent HCC datasets, external validation was performed. Using five different HCC cohorts, the growth module showed a higher preservation statistics summary than expected by random chance using bootstrapping procedures (Fig. 2A). Thus, the growth module was deemed to hold promise in independent tumor profiles from different patient cohorts. For function enrichment, the cut-off set with the false discover rate of  $<0.01$ , cell cycle, DNA replication, Fanconi anemia, etc. constituted the main KEGG pathways in the growth module (Fig. 2B).

To decompose the KEGG pathway into detail, a canonical circuit was defined as any possible route the signal can traverse to be transmitted from a particular input to a specific output node.<sup>13</sup> Effector nodes at the end of the circuits trigger specific functions in the cell. Using gene expressions as proxies of node activation values, computation of the signal intensity across the different circuits of the pathways was performed by canonical circuit activity analysis to compute the transmission of the signal along the network.<sup>13</sup> Thus, we estimated the level of activity of subpathways (signaling circuits) using the Hipathia program,<sup>13</sup> and detected several pathways with perturbed activity in the growth module (Supplementary Fig. 2).

Focusing on the cell cycle pathway, five effector circuits were deemed ultimately responsible for the functions of DNA replication and cell cycling. These circuits were highlighted in the fastest state, one of them ending in the node including *RB1*, one including *RAD21*, one containing *TFDP1* and *E2F4*, one ending in the node with protein genes for



**Fig. 1. Growth-related modules in HCC.** (A) Clustering dendrograms of all genes, with dissimilarity based on topological overlap, together with assigned module colors. (B) Correlation between module eigengene and tumor growth state of HCC. (C) Heatmap representation of the module-module relationship. (D) Module significance of each module in the fastest growth state of HCC. The higher the mean GS in a module is, the more significantly associated with tumor growth the module will be.

*CDC45, MCM7, MCM6, MCM5, MCM4, MCM3* and *MCM2*, and the last one ending in the node with protein genes for *ORC* and *MCM* (Fig. 2C). Indeed, most nodes in the effector circuits have adverse outcome in the LIHC cohort (Supplementary Fig. 3). These results clearly suggested the signaling in the green module as providing multiple routes and broader activity to promote cell cycle progression, thus accelerating tumor growth.

### Rapid growth program in HCC

Next, we examined the hub genes in the growth module. Hub genes,<sup>5</sup> including the top 10 GS, K.in or MM genes, are shown in Fig. 3A and Table 1. Among them, 20 genes were identified.

There was step-wise activation of all these genes that accompanied increased speed of tumor growth (Fig. 3A and Supplementary Fig. 4). The trend was clearly consistent and coordinated. As expected, all these hub genes were involved in the advanced prognosis of HCC, as evidenced by the results from our survival analysis (Supplementary Fig. 4).

Since the top 20 hub genes were densely interacted by protein-protein interaction analysis,<sup>24</sup> we categorized them among the rapid growth signature (Fig. 3B) and applied the ssGSEA algorithm to infer the growth activity for each sample.

Recent advances in molecular biomarkers of HCC have indicated various oncogenes and tumor suppressor genes.<sup>25</sup> Indeed, growth<sup>high</sup> patients in our study were more likely to show higher expression of many known adverse prognostic biomarkers, such as *AFP, DCP1A, GPC3, MDK, MCM6*. In contrast, growth<sup>low</sup> patients were likely to show higher expression of tumor suppressor genes, such as *GPR155* and *IFIT3* (Fig. 3C).

Furthermore, we found that rapid growth activity has a bad survival prognosis, both in LIHC and GSE14520 cohorts (Fig. 3D). These findings suggested a negative regulation relationship between the HCC growth program and HCC

prognosis.

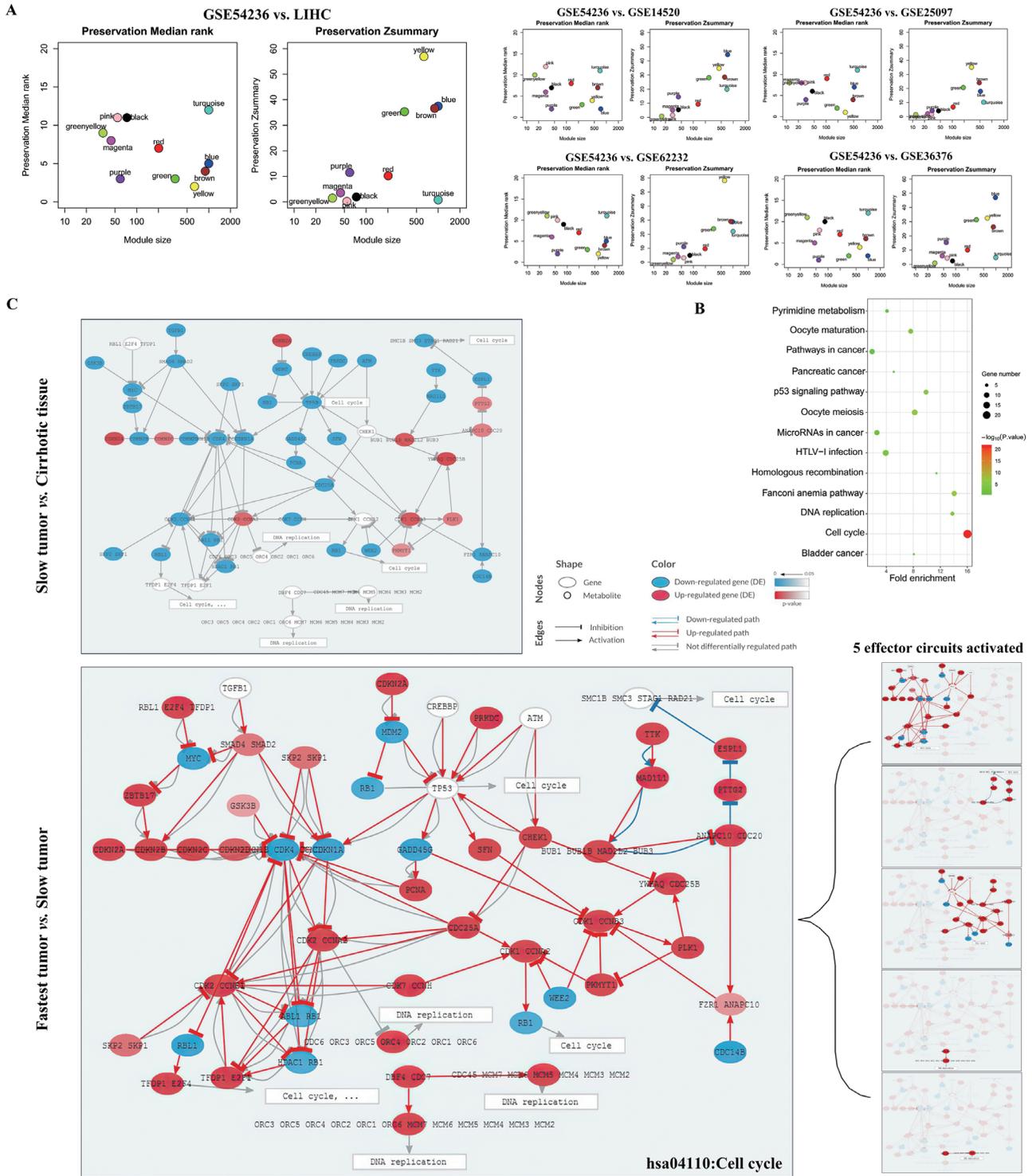
Next, we reasoned that genes in the signature would have a higher degree of association with cell proliferation. To test this hypothesis, we selected genes (*NCAPH* and *RAD54L*) for experimental validation. As expected, knockdown of *NCAPH* or *RAD54L* significantly suppressed the proliferation of HepG2 cells (Fig. 3E).

### Somatic mutations and copy number alterations of the growth signature

We next investigated the growth signature at the genomic level. By focusing on the somatic non-silent mutation or copy number variation (commonly known as CNV) genes (Fig. 4A), the principal components analysis (commonly known as PCA) was computed to maximize the stability of the components. The first two principal components, respectively, account for 13.6% and 11.2% of the total variability in gene expression; the first 20 principal components cumulatively explain 67% of the variance (Fig. 4B). Notably, overlaying the status of the genetic alterations of hub genes (growth<sup>mut</sup>) on to the first two principal components demonstrated that mutations or CNV alterations correlated with general gene expression profiles (Fig. 4C).

The transcriptome was globally perturbed by growth<sup>mut</sup>, with expression levels of 6,565/15,569 (42%) genes significantly associated with at least one genetic change of the hub genes ( $r > 0.3$ , false discovery rate of  $< 0.001$ ). For instance, genetic change of *NUF2* co-occurred with other hub genes' alteration, which led to 1069 genes' differential expression (Fig. 4D-E). The observed variability can be largely explained by the presence of other hub genes' alteration leading to strong up-regulation of *NUF2* mRNA (Fig. 4D). The expression changes of the growth<sup>mut</sup>-related genes are summarized in Supplementary Table 3.

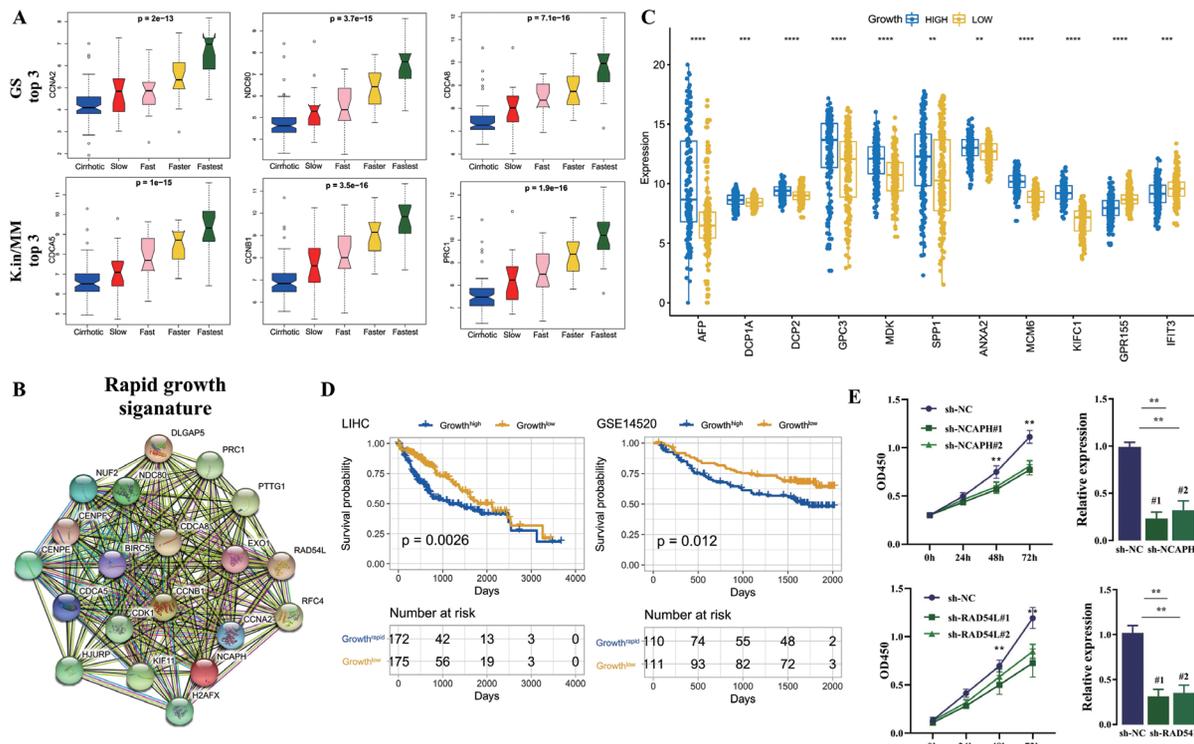
To understand the mechanism underlying growth<sup>mut</sup>, webgestalt<sup>11</sup> analysis showed that growth<sup>mut</sup>-related genes were enriched in the KEGG pathways of cell cycle, oocyte



**Fig. 2. Signaling events in the rapid growth module.** (A) Preservation of growth modules between different datasets. Each module is represented by the color-code and label. The left panel shows the composite statistic of preservation median rank. High median ranks indicated low preservation. The right panel shows preservation of the Zsummary statistic. (B) Enrichment of KEGG pathways in the green module. (C) The cell cycle pathway (hsa04110). The upper panel indicates the comparison between slow tumors and the adjacent cirrhotic tissues. The lower panel compares the fastest tumors with the slow tumors. Genes in red represent genes that were up-regulated, and blue represents genes that were down-regulated. The right panel highlights five effector circuits for the function output.

meiosis and DNA replication, etc. (Fig. 4G). Given that oncogenic signaling accelerates cell cycle progression,<sup>26</sup> these

data indicated that growth<sup>mut</sup> cancers can amplify growth signaling to maintain cell proliferation.



**Fig. 3. Hub genes in the growth module.** (A) Expression level of top three hub genes from cirrhotic, slow to the fastest growth state in HCC. (B) Highly interconnected top 20 hub genes constitute the rapid growth signature. (C) Expression of well-known biomarkers of HCC between the growth<sup>high</sup> and growth<sup>low</sup> group. The *p* values between the two groups were calculated by the Wilcoxon rank sum test. \*\**p*<0.01. (D) Kaplan-Meier survival plot of growth<sup>high</sup> vs. growth<sup>low</sup> in the LIHC cohort or GSE14520 cohort respectively. (E) Knockdown of *NCAPH* or *RAD54L* suppressed HepG2 cells proliferation. The right panel denotes knockdown efficiency. Student's *t* test was used to compare the differences. \**p*<0.05, \*\**p*<0.01.

**Rapid growth is associated with tumor immune-metabolism**

We then explored the correlation between growth activity and tumor-infiltrating lymphocytes (Fig. 5A). We observed a significant negative correlation of growth signature with Th1 cell types and positive correlation with Th2 cell types. Next, significant negative correlations between rapid growth and natural killer cells, plasmacytoid dendritic cells or macrophages were found (Fig. 5A and Supplementary Table 4).

Moreover, dysregulation of diverse immune signatures in HCC, including HLA expression, cytokines or chemokines, and interferon response were identified between growth<sup>high</sup> and growth<sup>low</sup> groups (Fig. 5C). Notably, the growth signature showed strong correlation with neoantigens (*r*=0.37, *p*=1.52E-12 in LIHC; *r*=0.38, *p*=5.22E-09 in GSE14520) (Fig. 5B). In addition, tumor rapid growth showed no significant correlation with immunoinhibitors or immunostimulators, including well-known checkpoint genes (Table S4). Additionally, immune cytolytic activity or tumor burden was not significantly different between high versus low rapid growth tumors (Supplementary Fig. 5). Thus, the growth signature itself might not elicit the active immune response.

Next, we investigated the metabolic configuration according to the growth activity. Using curated metabolic gene sets from Reactome as indicators, we found rapid tumor growth was highly associated with diverse metabolism processes. For example, the rapid tumor growth was significantly positively correlated with synthesis of DNA (*r*=0.73, *p*=1.37E-58 in LIHC; *r*=0.85, *p*=2.52E-64 in GSE14520) and negatively correlated with bile acid metabolism (*r*=-0.43,

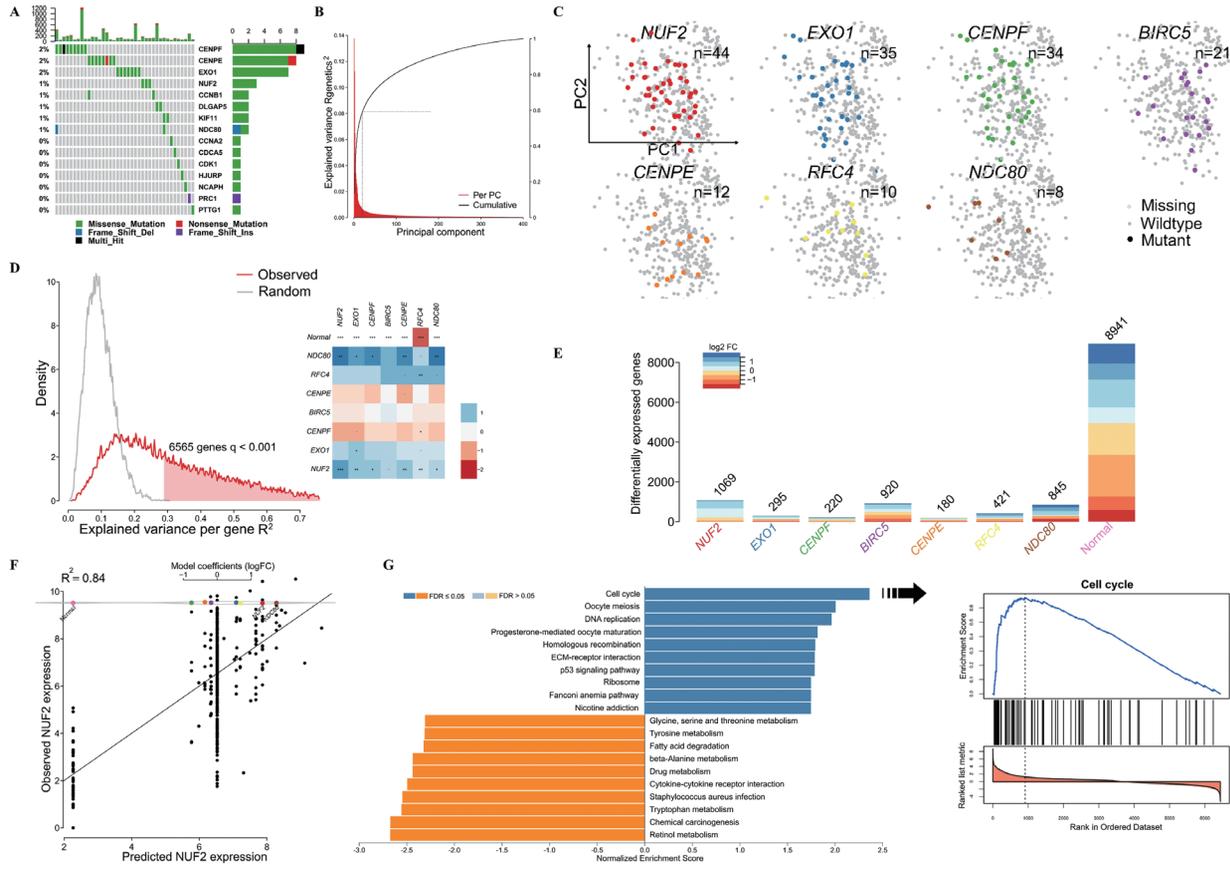
*p*=1.70E-17 in LIHC; *r*=-0.46, *p*=1.92E-13 in GSE14520). Detailed growth-metabolism correlations are provided in Supplementary Table 5.

The representative metabolic activity between growth<sup>high</sup> and growth<sup>low</sup> tumors is provided in Fig. 5D. For instance, dysregulation of cytochrome P450, xenobiotics, and biological oxidation are shown to be associated with poor prognosis. These results suggested that growth<sup>high</sup> tumors tend to present an immunetolerant and metabolism reconfigured microenvironment in HCC.

**Prognostic role of growth signature in pan-cancer**

When extending the growth signature to the pan-cancers, we found a significant hazard ratio between the growth genes and overall survival or recurrence-free survival in multiple cancers, including adrenocortical carcinoma, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma, brain lower grade glioma, LIHC, lung adenocarcinoma, mesothelioma, pancreatic adenocarcinoma, and sarcoma (Fig. 6A). Gene expression was significantly higher in the growth<sup>high</sup> group (Fig. 6D) and the growth<sup>high</sup> indicated advanced prognosis in these cancers (Fig. 6B-C). The pathway relation network also indicated that the signature was mostly involved in the cell cycle in nine cancers (Fig. 6E).

To investigate the clinical implications of the growth signature, we searched for targets of candidate drugs by using the L1000 project.<sup>27</sup> The top 10 associations are presented in Fig. 6F. Palbociclib functions as a *CDK4/6* inhibitor in multiple cancers.<sup>28</sup> It also acts as a novel radiosensitizer,



**Fig. 4. Genetic change of the growth signature correlates with differential expression.** (A) Overview of the somatic non-silent mutation and CNV of the growth signature using Maftools. Samples are displayed in columns, and different colors indicate the different types of somatic mutations. The bar plots show the recalibrated mutation frequencies. (B) Distribution of the variance explained by the genetic alterations across genes. (C) Scatter plot of the first two principal components of the gene expression data overlaid with mutation status. (D) Distribution of the variance explained by genetic alterations across genes. The right panel indicated the heatmap of observed pairwise mutation patterns (odds ratio). Blue color denotes preferential co-mutation/high overlap, while red color indicates mutual exclusivity. (E) Number of differentially expressed genes for each hub gene. (F) Scatter plot of expression predictions for the *NUF2* gene vs. observed expression values. The inset shows the model coefficients, indicating the predicted magnitudes of expression changes when a given alteration is present. (G) Bar plot showing the GSEA enriched KEGG pathways of growth<sup>mut</sup>-affected genes. Enrichment is represented as  $-\log_{10}(p \text{ value})$ .

dovitinib, inducing apoptosis and overcoming sorafenib resistance in HCC through *SHP-1*-mediated inhibition of *STAT3*.<sup>29–32</sup> As the topoisomerase inhibitor, mitoxantrone, was widely used in clinic.<sup>33</sup> Canertinib is an irreversible *EGFR* inhibitor in cancers.<sup>34</sup>

**Discussion**

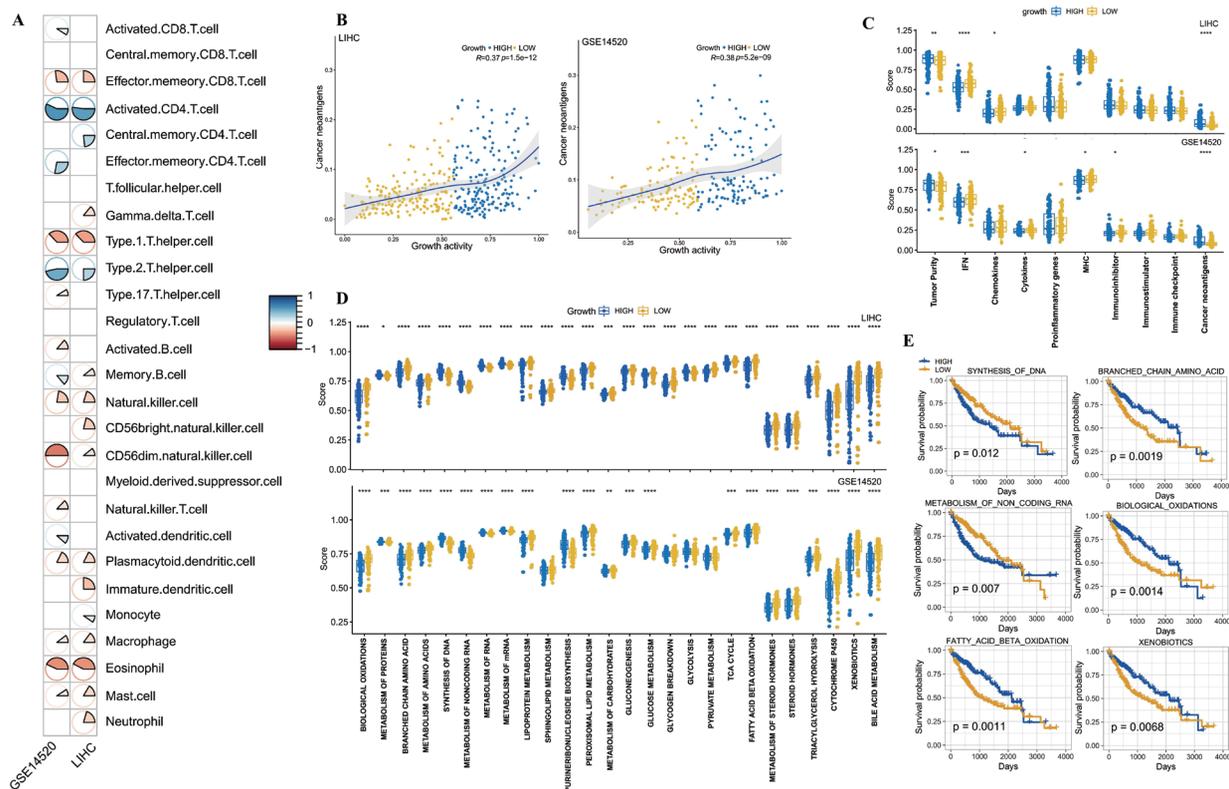
The theory that patients with a rapidly growing cancer have a poor prognostic outlook has remained persistent;<sup>2</sup> thus, detailing the biological structure in the hidden layer of rapid growth is an interesting question.

As illustrated in this study, to understand the growth rate of HCC at the modular level and toward uncovering the critical drivers of the disease in a comprehensive manner, HCC transcriptomes were explored in the context of modular pattern, in which the green module was responsible for rapid tumor growth after internal and external validation. Briefly, the top enriched functional classes of this program are consistent with our existing knowledge of the cell cycle, as well as DNA repair, replication and cell proliferation in cancer.

Due to the presence of the highly interconnected top 20

hub genes, we assembled them into the growth signature. At the gene level, we demonstrated that the expression levels of all these genes were increased in coordination with the state of growth-rate. Accordingly, high expression of these genes predicted the adverse outcome of HCC. Indeed, various lines of evidence showed the involvement of these hub genes in previous studies. For example, *CCNA2* is the leading gene according to the GS rank. A recent study revealed a new poor-prognosis HCC entity and a rearrangement signature related to replication stress, due to *CCNA2* alterations.<sup>35</sup> The top K.in gene, *CDCA5*, transcribed by *E2F1*, promotes oncogenesis by enhancing cell proliferation and inhibiting apoptosis via the AKT pathway in HCC.<sup>36</sup> The top MM gene, *CCNB1* was highly expressed in the samples of recurrent HCC, which was associated with significantly reduced recurrence-free survival.<sup>37</sup> In addition, the vast majority of genes within the signature, such as *RFC4*, *HJURP*, *ECT2*, *KIFC1*, *NUSAP1*, *CDK1*, *PRC1*, *KIF4A*, etc., contribute to the pathogenesis of HCC, as reported previously.<sup>38–43</sup> At the time of preparation of this paper, little is known about *NCAPH* and *RAD54L* in HCC. In this study, however, we found that knockdown of *NCAPH* and *RAD54L* expression is associated with growth inhibition.

Genetic background and in particular genomic alteration



**Fig. 5. Tumor rapid growth is associated with immune-metabolism.** (A) Correlation between growth activity and tumor-infiltrating lymphocytes' infiltration. (B) Tumor growth was highly correlated with neoantigen production both in the LIHC and GSE14520 cohort. (C) Dysregulation of immune signatures between growth<sup>high</sup> and growth<sup>low</sup> groups in LIHC and GSE14520 respectively. (D–E) Metabolic reconfiguration between the growth<sup>high</sup> and growth<sup>low</sup> groups in LIHC and GSE14520 respectively, which have prognosis effect.

might contribute to gene expression in HCC related to tumor growth speed. Taking a genetic-centric view, we specified the set of gene expression changes that correlate with the alteration of the signature. The number of target genes whose expression are differentially affected varies widely across the different genetic change of growth signature. For example, *BIRC5*, playing dual roles in mitosis and cell survival of HCC,<sup>44</sup> was independently correlated with expression levels of 920 genes. These data supported the notion that genetic change of these hub gene results in a rapid tumor growth.

Moreover, the growth signature exhibited a significant correlation with certain genomic features, such as tumor purity and neoantigens. Further, no obvious association between rapid growth and active antitumor immune signatures was found. In addition, a significant negative correlation was observed between growth activity and energy metabolism integration, biological oxidation, vitamins, fatty acid, and glucose metabolism, *etc.* (Supplementary Table 5). The above results supported the idea of a link between tumor growth, metabolic landscape reconfiguration, and clinical progression of cancer. Accordingly, aggressive malignant phenotypes of cancer cells obtained by accumulated mutations change metabolic phenotypes for proliferation represented by nucleotide and amino acid metabolism. Indeed, this hypothesis of highly coordinated growth program warrants further study.

Finally, candidate drugs have been inferred based on the growth signature. Numerous cell cycle inhibitors have been designed over the past decades.<sup>45</sup> The current findings of growth-specific drugs for HCC would have potential implica-

tions in warranting future studies toward developing targeted and combinatorial therapeutics for HCC.

## Conclusions

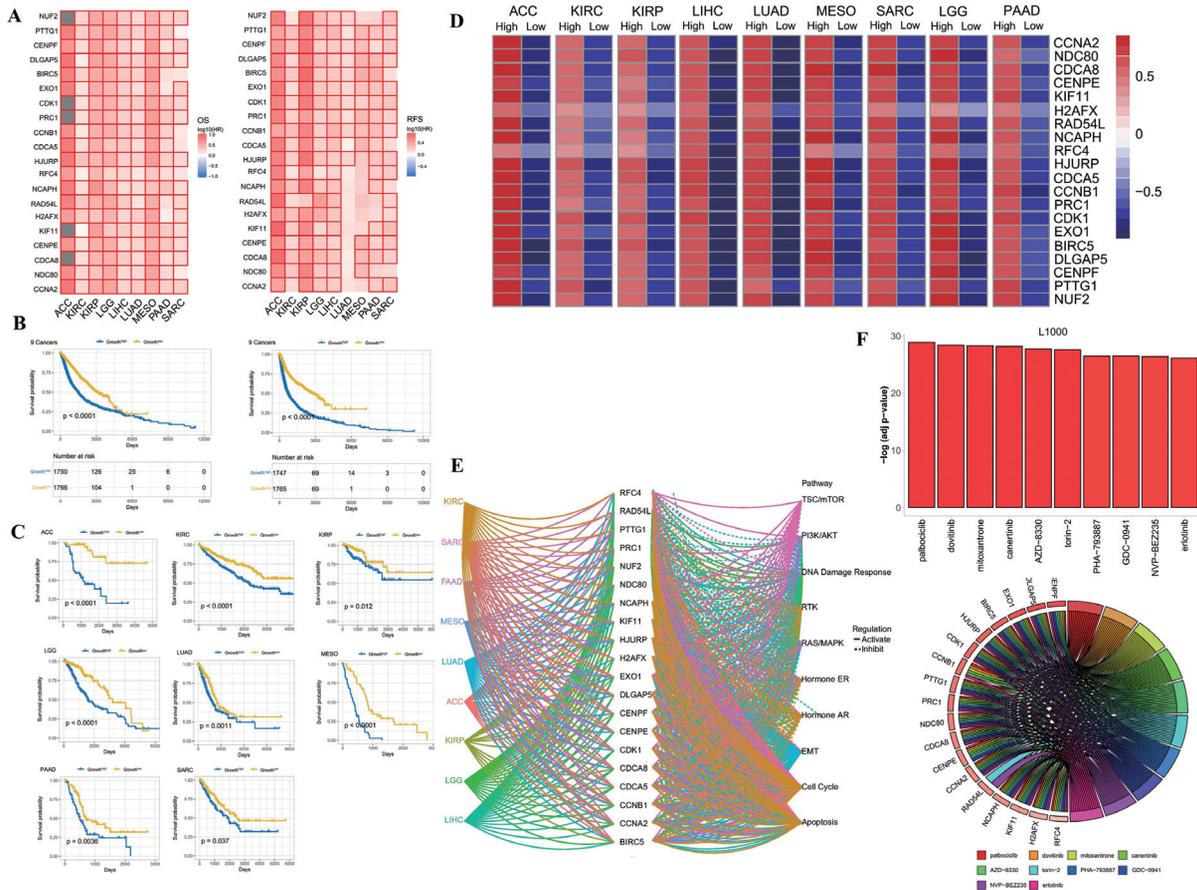
Given the inherently modular profile of tumor growth, the present study revealed a unifying portrait of the growth signature of HCC, and could be extended to pan-cancer. Next, the study offered information to better define how specific organizations of genes are able to orchestrate rapid growth. Further, the growth signature has potential prognostic and therapeutic intervention value for HCC, lightening the way toward tailoring the targeted therapeutics for HCC.

## Funding

This work was supported by the National Natural Science Foundation of China (No. 81772520), Zhejiang Provincial Natural Science Foundation (No. LGF19H030004), Zhejiang Medical and Health Technology Project (No. 2018PY039), and Hangzhou Science and Technology Project (No. 2017A36, 20180533B46).

## Conflict of interest

The authors have no conflict of interests related to this publication.



**Fig. 6. Growth signature in diverse cancers and harbor clinically actionable genes.** (A) Survival map of the hub genes within the signature for overall survival or recurrence-free survival in nine cancer types. (B) Survival plot of growth<sup>high</sup> vs. growth<sup>low</sup> in nine cancers, using median cut-off. (C) Survival plot of rapid growth across nine cancer types. (D) The average mRNA expression in growth<sup>high</sup> and growth<sup>low</sup> samples across nine cancers. Red indicates relatively high expression, while blue indicates relatively low expression. (E) The network showing the relationship between genes and pathways in different cancer types. The solid line indicates activation, and the dashed line indicates inhibition. (F) The top 10 drugs (upper panel) enriched for the growth signature by L1000 and their targeted genes (lower panel).

**Author contributions**

Contributed to study concept and design (XZ and JY), acquisition of the data (JmL and YxL), assay performance and data analysis (JW, YL, AqL, AnC, and JtF), drafting of the manuscript (JW and YL), critical revision of the manuscript (JW and XZ), supervision (JY).

**References**

[1] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin* 2015;65:87–108. doi:10.3322/caac.21262.  
 [2] Adami HO, Csermely P, Veres DV, Emilsson L, Løberg M, Bretthauer M, et al. Are rapidly growing cancers more lethal? *Eur J Cancer* 2017;72:210–214. doi:10.1016/j.ejca.2016.11.011.  
 [3] Morita M, Sato T, Nomura M, Sakamoto Y, Inoue Y, Tanaka R, et al. PKM1 Confers Metabolic Advantages and Promotes Cell-Autonomous Tumor Cell Growth. *Cancer Cell* 2018;33:355–367.e7. doi:10.1016/j.ccell.2018.02.004.  
 [4] Bi J, Wu S, Zhang W, Mischel PS. Targeting cancer’s metabolic co-dependencies: A landscape shaped by genotype and tissue context. *Biochim Biophys Acta Rev Cancer* 2018;1870:76–87. doi:10.1016/j.bbcan.2018.05.002.  
 [5] Lou Y, Tian GY, Song Y, Liu YL, Chen YD, Shi JP, et al. Characterization of transcriptional modules related to fibrosing-NAFLD progression. *Sci Rep* 2017;7:4748. doi:10.1038/s41598-017-05044-2.  
 [6] Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7. doi:10.1186/1471-2105-14-7.  
 [7] Villa E, Critelli R, Lei B, Marzocchi G, Cammà C, Giannelli G, et al. Ne-

oangiogenesis-related genes are hallmarks of fast-growing hepatocellular carcinomas and worst survival. Results from a prospective study. *Gut* 2016;65:861–869. doi:10.1136/gutjnl-2014-308483.  
 [8] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008;9:559. doi:10.1186/1471-2105-9-559.  
 [9] Long J, Huang S, Bai Y, Mao J, Wang A, Lin Y, et al. Transcriptional landscape of cholangiocarcinoma revealed by weighted gene coexpression network analysis. *Brief Bioinform* 2020;bbaa224. doi:10.1093/bib/bbaa224.  
 [10] Russo PST, Ferreira GR, Cardozo LE, Bürger MC, Arias-Carrasco R, Maruyama SR, et al. CEMiTool: a Bioconductor package for performing comprehensive modular co-expression analyses. *BMC Bioinformatics* 2018;19:56. doi:10.1186/s12859-018-2053-1.  
 [11] Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res* 2019;47:W199–W205. doi:10.1093/nar/gkz401.  
 [12] Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016;44:W90–W97. doi:10.1093/nar/gkw377.  
 [13] Hidalgo MR, Cubuk C, Amadoz A, Salavert F, Carbonell-Caballero J, Dopazo J. High throughput estimation of functional cell activities reveals disease mechanisms and predicts relevant clinical outcomes. *Oncotarget* 2017;8:5160–5178. doi:10.18632/oncotarget.14107.  
 [14] Liu CJ, Hu FF, Xia MX, Han L, Zhang Q, Guo AY. GSCALite: a web server for gene set cancer analysis. *Bioinformatics* 2018;34:3771–3772. doi:10.1093/bioinformatics/bty411.  
 [15] Gerstung M, Pellagatti A, Malcovati L, Giagounidis A, Porta MG, Jädersten M, et al. Combining gene mutation with gene expression data improves outcome prediction in myelodysplastic syndromes. *Nat Commun* 2015;6:5901. doi:10.1038/ncomms6901.  
 [16] Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 2018;28:1747–1756. doi:10.1101/gr.239244.118.

- [17] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, *et al*. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47. doi:10.1093/nar/gkv007.
- [18] Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-García W, *et al*. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* 2013;4:2612. doi:10.1038/ncomms3612.
- [19] Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 2015;160:48–61. doi:10.1016/j.cell.2014.12.033.
- [20] Chen X, Xu C, Hong S, Xia X, Cao Y, McDermott J, *et al*. Immune cell types and secreted factors contributing to inflammation-to-cancer transition and immune therapy response. *Cell Rep* 2019;26:1965–1977.e4. doi:10.1016/j.celrep.2019.01.080.
- [21] Li Z, Lou Y, Tian G, Wu J, Lu A, Chen J, *et al*. Discovering master regulators in hepatocellular carcinoma: one novel MR, SEC14L2 inhibits cancer cells. *Aging (Albany NY)* 2019;11:12375–12411. doi:10.18632/aging.102579.
- [22] Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res* 2017;45:W98–W102. doi:10.1093/nar/gkx247.
- [23] Langfelder P, Mischel PS, Horvath S. When is hub gene selection better than standard meta-analysis? *PLoS One* 2013;8:e61505. doi:10.1371/journal.pone.0061505.
- [24] Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, *et al*. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 2015;43:D447–D452. doi:10.1093/nar/gku1003.
- [25] Umeda S, Kanda M, Koderia Y. Recent advances in molecular biomarkers for patients with hepatocellular carcinoma. *Expert Rev Mol Diagn* 2019;19:725–738. doi:10.1080/14737159.2019.1638254.
- [26] Burotto M, Chiou VL, Lee JM, Kohn EC. The MAPK pathway across different malignancies: a new perspective. *Cancer* 2014;120:3446–3456. doi:10.1002/cncr.28864.
- [27] Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, *et al*. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* 2017;171:1437–1452.e17. doi:10.1016/j.cell.2017.10.049.
- [28] Whittle JR, Vaillant F, Surgenor E, Policheni AN, Giner G, Capaldo BD, *et al*. Dual targeting of CDK4/6 and BCL2 pathways augments tumor response in estrogen receptor-positive breast cancer. *Clin Cancer Res* 2020;26:4120–4134. doi:10.1158/1078-0432.CCR-19-1872.
- [29] Cheng AL, Thongprasert S, Lim HY, Sukeepaisarnjaroen W, Yang TS, Wu CC, *et al*. Randomized, open-label phase 2 study comparing frontline dovitinib versus sorafenib in patients with advanced hepatocellular carcinoma. *Hepatology* 2016;64:774–784. doi:10.1002/hep.28600.
- [30] Huang CY, Tai WT, Wu SY, Shih CT, Chen MH, Tsai MH, *et al*. Dovitinib acts as a novel radiosensitizer in hepatocellular carcinoma by targeting SHP-1/STAT3 signaling. *Int J Radiat Oncol Biol Phys* 2016;95:761–771. doi:10.1016/j.ijrobp.2016.01.016.
- [31] Tai WT, Cheng AL, Shiau CW, Liu CY, Ko CH, Lin MW, *et al*. Dovitinib induces apoptosis and overcomes sorafenib resistance in hepatocellular carcinoma through SHP-1-mediated inhibition of STAT3. *Mol Cancer Ther* 2012;11:452–463. doi:10.1158/1535-7163.MCT-11-0412.
- [32] Huynh H, Chow PK, Tai WM, Choo SP, Chung AY, Ong HS, *et al*. Dovitinib demonstrates antitumor and antimetastatic activities in xenograft models of hepatocellular carcinoma. *J Hepatol* 2012;56:595–601. doi:10.1016/j.jhep.2011.09.017.
- [33] Prosperini L, Mancinelli CR, Solaro CM, Nociti V, Haggiag S, Cordioli C, *et al*. Induction versus escalation in multiple sclerosis: A 10-year real world study. *Neurotherapeutics* 2020;17:994–1004. doi:10.1007/s13311-020-00847-0.
- [34] García M, Rodríguez-Hernández CJ, Mateo-Lozano S, Pérez-Jaume S, Gonçalves-Alves E, Lavarino C, *et al*. Parathyroid hormone-like hormone plays a dual role in neuroblastoma depending on PTH1R expression. *Mol Oncol* 2019;13:1959–1975. doi:10.1002/1878-0261.12542.
- [35] Bayard Q, Meunier L, Peneau C, Renault V, Shinde J, Nault JC, *et al*. Cyclin A2/E1 activation defines a hepatocellular carcinoma subclass with a rearrangement signature of replication stress. *Nat Commun* 2018;9:5235. doi:10.1038/s41467-018-07552-9.
- [36] Chen H, Chen J, Zhao L, Song W, Xuan Z, Chen J, *et al*. CDCA5, transcribed by E2F1, promotes oncogenesis by enhancing cell proliferation and inhibiting apoptosis via the AKT pathway in hepatocellular carcinoma. *J Cancer* 2019;10:1846–1854. doi:10.7150/jca.28809.
- [37] Weng L, Du J, Zhou Q, Cheng B, Li J, Zhang D, *et al*. Identification of cyclin B1 and Sec62 as biomarkers for recurrence in patients with HBV-related hepatocellular carcinoma after surgical resection. *Mol Cancer* 2012;11:39. doi:10.1186/1476-4598-11-39.
- [38] Arai M, Kondoh N, Imazeki N, Hada A, Hatsuse K, Matsubara O, *et al*. The knockdown of endogenous replication factor C4 decreases the growth and enhances the chemosensitivity of hepatocellular carcinoma cells. *Liver Int* 2009;29:55–62. doi:10.1111/j.1478-3231.2008.01792.x.
- [39] 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;62:1287–1295. doi:10.1016/j.jhep.2015.01.014.
- [40] Fu X, Zhu Y, Zheng B, Zou Y, Wang C, Wu P, *et al*. KIFC1, a novel potential prognostic factor and therapeutic target in hepatocellular carcinoma. *Int J Oncol* 2018;52:1912–1922. doi:10.3892/ijo.2018.4348.
- [41] Cai J, Li B, Zhu Y, Fang X, Zhu M, Wang M, *et al*. Prognostic biomarker identification through integrating the gene signatures of hepatocellular carcinoma properties. *EBioMedicine* 2017;19:18–30. doi:10.1016/j.ebiom.2017.04.014.
- [42] Chen J, Rajasekaran M, Xia H, Zhang X, Kong SN, Sekar K, *et al*. The microtubule-associated protein PRC1 promotes early recurrence of hepatocellular carcinoma in association with the Wnt/ $\beta$ -catenin signalling pathway. *Gut* 2016;65:1522–1534. doi:10.1136/gutjnl-2015-310625.
- [43] Huang Y, Wang H, Lian Y, Wu X, Zhou L, Wang J, *et al*. Upregulation of kinesin family member 4A enhanced cell proliferation via activation of Akt signaling and predicted a poor prognosis in hepatocellular carcinoma. *Cell Death Dis* 2018;9:141. doi:10.1038/s41419-017-0114-4.
- [44] Li D, Fu J, Du M, Zhang H, Li L, Cen J, *et al*. Hepatocellular carcinoma repression by TNF $\alpha$ -mediated synergistic lethal effect of mitosis defect-induced senescence and cell death sensitization. *Hepatology* 2016;64:1105–1120. doi:10.1002/hep.28637.
- [45] Jäger K, Walter M. Therapeutic targeting of telomerase. *Genes (Basel)* 2016;7:39. doi:10.3390/genes7070039.