Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

Identification of prognostic biomarkers in hepatitis B virus-related hepatocellular carcinoma and stratification by integrative multi-omics analysis

Ruoyu Miao[1,2,3,†], Haitao Luo[2,†], Huandi Zhou[4,†], Guangbing Li[1], Dechao Bu[2], Xiaobo Yang[1], Xue Zhao[1], Haohai Zhang[1], Song Liu[4], Ying Zhong[4], Zhen Zou[4], Yan Zhao[4], Kuntao Yu[2], Lian He[1], Xinting Sang[1], Shouxian Zhong[1], Jiefu Huang[1], Yan Wu[3], Rebecca A. Miksad[5], Simon C. Robson[3], Chengyu Jiang[4,6,*,‡], Yi Zhao[2,6,*,‡], Haitao Zhao[1,6,*,‡]

[1]*Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China;* [2]*Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China;* [3]*Liver Center and The Transplant Institute, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA;* [4]*State Key Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China;* [5]*Division of Hematology/Oncology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA;* [6]*Center of Translational Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China*

Background & Aims: The differentiation of distinct multifocal hepatocellular carcinoma (HCC): multicentric disease *vs.* intrahepatic metastases, in which the management and prognosis varies substantively, remains problematic. We aim to stratify multifocal HCC and identify novel diagnostic and prognostic biomarkers by performing whole genome and transcriptome sequencing, as part of a multi-omics strategy.

Methods: A complete collection of tumour and somatic specimens (intrahepatic HCC lesions, matched non-cancerous liver tissue and blood) were obtained from representative patients with multifocal HCC exhibiting two distinct postsurgical courses. Whole-genome and transcriptome sequencing with genotyping were performed for each tissue specimen to contrast genomic alterations, including hepatitis B virus integrations, somatic mutations, copy number variations, and structural variations. We then constructed a phylogenetic tree to visualise individual tumour evolution and performed functional enrichment analyses on select differentially expressed genes to elucidate biological processes involved in multifocal HCC development. Multi-omics data were integrated with detailed clinicopathological information to identify HCC biomarkers, which were further validated using a large cohort of HCC patients (n = 174).

Results: The multi-omics profiling and tumour biomarkers could successfully distinguish the two multifocal HCC types, while accurately predicting clonality and aggressiveness. The dual-specificity protein kinase *TTK*, which is a key mitotic checkpoint regulator with links to p53 signaling, was further shown to be a promising overall prognostic marker for HCC in the large patient cohort.

Conclusions: Comprehensive multi-omics characterisation of multifocal tumour evolution may improve clinical decision-making, facilitate personalised medicine, and expedite identification of novel biomarkers and therapeutic targets in HCC.

© 2014 European Association for the Study of the Liver. Published by Elsevier B.V. All rights reserved.

Cancer

Keywords: Hepatocellular carcinoma; Whole-genome sequencing; Transcriptome sequencing.

Received 13 October 2013; received in revised form 8 May 2014; accepted 13 May 2014; available online 22 May 2014

* Corresponding authors. Addresses: State Key Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College (CAMS & PUMC), 5 Dong Dan San Tiao, Beijing 100005, China. Tel.: +86 10 69156908; fax: +86 10 69156551 (C. Jiang). Bioinformatics Research Group, Advanced Computing Research Laboratory, Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, No. 6 Kexueyuan South Road Zhongguancun, Haidian District, Beijing 100190, China. Tel.: +86 10 62601010 (Y. Zhao). Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College (CAMS & PUMC), 1 Shuaifuyuan, Wangfujing, Beijing 100730, China. Tel.: +86 10 69156042; fax: +86 10 69156043 (H. Zhao).

E-mail addresses: chengyujiang@gmail.com (C. Jiang), biozy@ict.ac.cn (Y. Zhao), ZhaoHT@pumch.cn (H. Zhao).

† These authors contributed equally to this work.

‡ These authors share senior co-authorship.

Abbreviations: HCC, hepatocellular carcinoma; SNP, single nucleotide polymorphism; HBV, hepatitis B virus; RNA-Seq, RNA sequencing; CNV, copy number variation; SV, structural variation; RFS, recurrence-free survival; OS, overall survival; TERT, telomerase reverse transcriptase; Indel, small insertion and deletion; CDS, coding sequence; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; HBsAg, hepatitis B surface antigen; HBeAg, hepatitis B e antigen; ALT, alanine aminotransferase; AFP, alpha-fetoprotein; HAL, histidine ammonia-lyase; RASGRF1, Ras protein-specific guanine nucleotide-releasing factor 1; SFN, stratifin; KIF15, kinesin family member 15; TTK, TTK dual-specificity protein kinase; BUB1, budding uninhibited by benzimidazoles 1 homolog (yeast); MCM4, minichromosome maintenance complex component 4.

Downloaded for Anonymous User (n/a) at The Second Hospital of Hebei Medical University from ClinicalKey.com by Elsevier on November 10, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.

Introduction

Hepatocellular carcinoma (HCC) is ranked as the sixth most common malignant cancer and is the third leading cause of cancer-related death worldwide [1,2]. HCC patients often present with multiple intrahepatic tumours. Although standardised guidelines for multifocal HCC and indications for surgical removal are available [3–6], surgical decision-making is still complicated by the difficulty in accurately predicting future tumour development, recurrence of the primary lesion and/or metastatic spread. These uncertainties also make prognostication after surgery to be very difficult for individual patients.

Speculatively, multifocal HCC may arise either synchronously, and metachronously, as primary tumours (multicentric occurrence) or develop as a consequence of intrahepatic metastases of the same primary cancer (Supplementary Fig. 1A) [7], yet molecular mechanistic underpinnings and tools to define and discriminate these two types are still lacking. As the management and prognosis in these different scenarios varies substantially, it is important to obtain the correct diagnosis. However, accurate stratification of multifocal HCC has not been achieved [8,9], despite the recent development of several evaluation approaches (i.e., pathological examination, profiling of integrated hepatitis B virus (HBV) DNA by PCR and Southern blotting, loss-of-heterozygosity analysis of specific DNA microsatellite loci) to differentiate these two development patterns. Clinical decision-making and theoretical considerations are therefore used to dictate treatment strategies for multifocal HCC.

The emergence and rapid progress of next-generation sequencing (NGS) has enabled comprehensive characterisation of cancers, including HCC [10–12]. This approach has allowed the identification of new molecular markers, as well as defining underlying biologic mechanisms, which facilitate the stratification and characterisation of tumours [13,14].

In this study, we selected representative patients with HBV-related multifocal HCC who underwent tumour resection and exhibited distinct postsurgical courses. These samples were subjected to NGS to obtain complete data sets for each patient. We then performed multi-omics analyses integrating genomics and transcriptomics, and further correlated these with the clinicopathological data. We sought to comprehensively decipher molecular differences between the two multifocal HCC models and identify molecular markers for diagnosis, prognosis and, potentially, therapeutic targets.

Materials and methods

Patients and clinical samples

Two representative HBV-HCC patients who had tumour resection were selected for NGS studies. Each patient underwent the same pathologic evaluation on all tumours.

In validation studies, 174 pairs of frozen HCC and non-cancerous livers were randomly selected from HBV-HCC patients with single or multifocal HCC who had undergone hepatectomy at PUMCH between September 1, 2003 and February 28, 2012.

Detailed clinicopathological information is listed in Supplementary Tables 1 and 2, respectively. Sample collection is described in Supplementary data.

The study was approved by the Ethics in Research Committee of PUMCH. Written informed consent was obtained from all study participants.

Whole-genome sequencing, Poly(A) RNA sequencing (RNA-Seq), and SNP Genotyping

Whole-genome sequencing and RNA-Seq were performed on the Illumina HiSeq™ 2000 platform (Illumina, San Diego, CA), and the sequencing data were deposited at the European Genome-phenome Archive (EGA, http://www.ebi.ac.uk/ega/), which is hosted by the EBI, under the accession numbers EGAS00001000325 and EGAS00001000372 respectively. DNA libraries were also genotyped using the Genome-Wide Human SNP 6.0 array (Affymetrix, Santa Clara, CA). Microarray data were submitted to the ArrayExpress Archive (http://www.ebi.ac.uk/arrayexpress/) under the accession number E-MEXP-3705. Further details are outlined in Supplementary data.

PCR, Sanger sequencing, and Quantitative real-time PCR (qRT-PCR)

Validation analyses were carried out as described previously [15] and detailed in Supplementary materials and methods.

Data analyses

NGS data (948.49 Gb in total, Supplementary Table 3) were comprehensively analysed using appropriate analytical tools. Multi-omics data analyses are detailed in Supplementary data and pipelined in Supplementary Fig. 2.

Statistical analyses

All statistical analyses were performed using SPSS 17.0 software (SPSS Inc., Chicago, IL). For statistical comparisons, a one-way analysis of variance, χ^2 test, Fisher's exact test, Student's *t* test, or Mann-Whitney U test was performed when appropriate. Recurrence-free survival (RFS) was calculated from the date of tumour resection until detection of the first HCC recurrence, death, or the last follow-up. Overall survival (OS) was defined as the time between surgery and death or the last follow-up examination. Patients who were lost to follow-up or died from causes unrelated to HCC were considered as censored events. Survival curves were analysed by the Kaplan-Meier method and compared with the log-rank test. Independent factors for RFS or OS were evaluated by multivariate Cox proportional hazards regression analysis. Significance was defined as $p < 0.05$.

Results

HBV integrations and clonality of multifocal HCC

Based on clinicopathological evaluations and postsurgical outcomes (Supplementary Fig. 1, Supplementary Table 1), we identified two patients with distinct multifocal HCC. Patient I (PI) who had cirrhosis and multifocal, poorly differentiated HCC died of recurrent disease at three months after resection. In contrast, Patient II (PII) was non-cirrhotic and presented with well-differentiated multifocal HCC, with no recurrence at two years after surgery. Therefore, we hypothesised that PI had intrahepatic metastases while PII had synchronous primary tumour development, with no spread or metachronous lesions. Tissues from multiple lesions for each patient (PI-P, PI-M1, PI-V, PII-L, PII-R), adjacent non-cancerous liver controls (PI-N, PII-N), and peripheral blood (PI-B, PII-B) were used for NGS followed by PCR validation. Two additional satellite lesions from PI (PI-M2, PI-M3) were also used for validation.

Initially, HBV integration patterns into the respective host genome were evaluated using whole-genome sequencing data. Profiling of HBV integration in these two patients was found to be distinctly different. A single integration site in the HCC carried by PI was noted to be associated with a 3209 bp event confirmed by PCR to be at the intergenic region of 3q26.1 in all tumours (Fig. 1A); clearly suggesting a monoclonal origin of PI metastases.

Downloaded for Anonymous User (n/a) at The Second Hospital of Hebei Medical University from ClinicalKey.com by Elsevier on November 10, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.

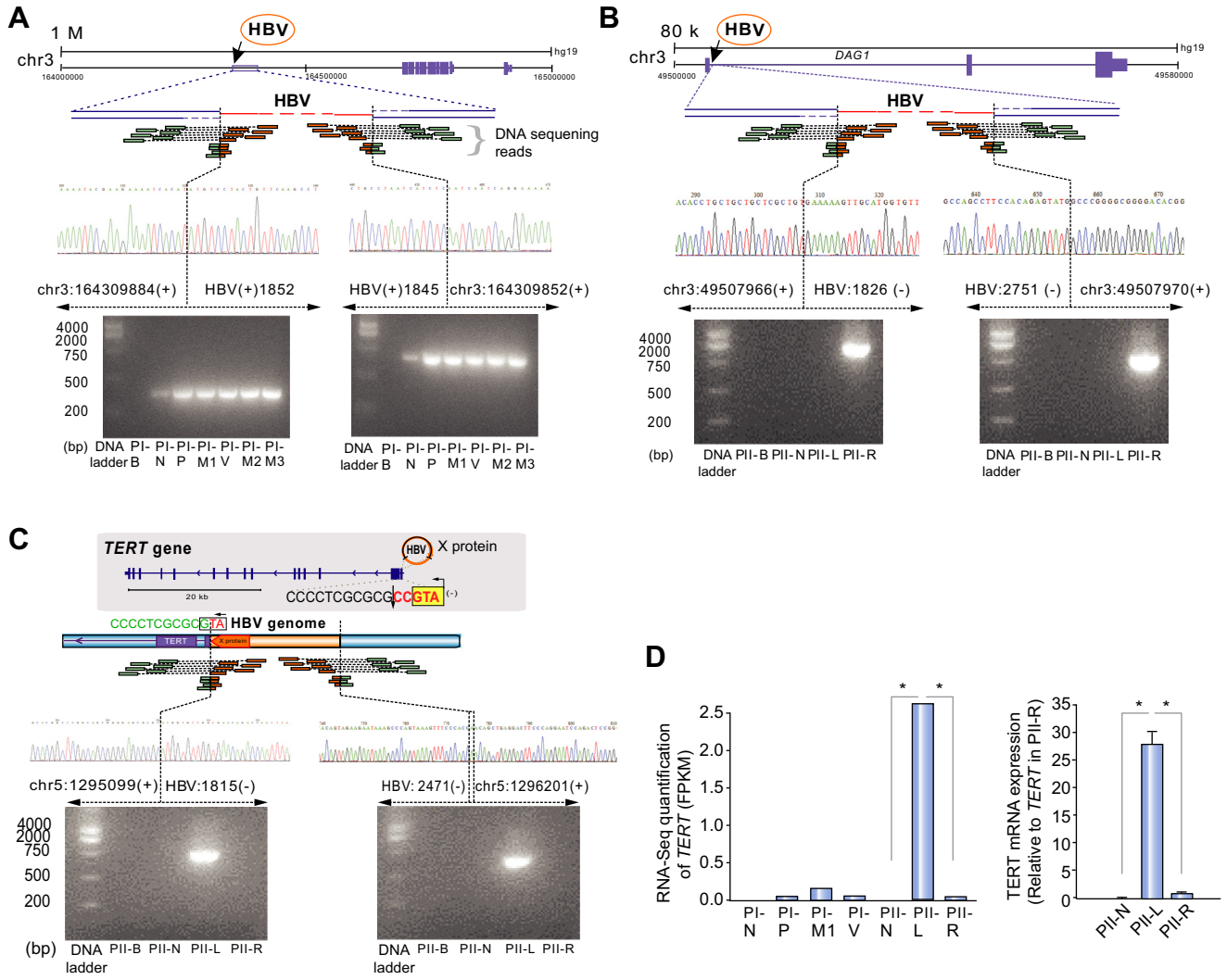


Fig. 1. Differential HBV integrations in different HCCs. (A–C) As predicted, the only 3209 bp HBV integration in 3q26.1 of PI (A) and two select integrations in PII: the 2291 bp integration at the first intron of *DAG1* in 3p21.31 of PII-R (B) and the 2560 bp integration in 5p15.33 of PII-L (C). Putative integrations were validated by PCR (bottom in each panel: images of DNA agarose gels). (D) *TERT* mRNA levels in PII samples quantified by RNA-Seq and qRT-PCR. FPKM, fragments per kilobase of exon per million fragments mapped. Data are given as mean ± SEM (n = 3). **p* <0.0001.

In contrast, among four totally different integration sites noted in PII tumours, two were validated exclusively in PII-L and the other two were only seen in PII-R. These data were indicative of HBV integration in two distinct tumour-initiating clones (Fig. 1B and C, Supplementary Fig. 3). Notably, the 2560 bp integration in 5p15.33 of PII-L, which combines HBVgp3 and *TERT* (telomerase reverse transcriptase), commenced at the sixth nucleotide of the *TERT* open-reading frame. This is the only event that would potentially result in a novel in-frame *TERT* initiation codon leading to differential expression of *TERT* (Fig. 1C). This prediction was further validated by RNA-Seq and qRT-PCR analyses, showing a 27-fold increase in *TERT* mRNA levels in PII-L when contrasted to PII-R (Fig. 1D).

Taken together, HBV integration data above suggest different patterns of tumour clonality to explain the different manifestations of the multiple tumours in these two patients.

Genomic alterations and aggressiveness of multifocal HCC

HBV integrations alone might not fully predict the clonality of multifocal HCC, as subsequent genomic alterations might have evolved independently from one another after metastases. Thus next, genomic alterations by mutational analyses were further used to assess individual tumour development and progression pattern.

First, somatic mutations inclusive of substitutions (Substitution) and small insertions/or deletions (Indel) were examined. Overall, ratios of non-synonymous substitution rate (Ka) to synonymous substitution rate (Ks) are PI-M1 (1.49) >PI-V (1.17) >PI-P (0.96) >PII-R (0.83) >PII-L (0.81) (Supplementary Table 4), suggestive of a substantively higher positive Darwinian selection conveying selective advantages on cancer subclones [16] in PI tumours. Additionally, further mutations (albeit not within the

Downloaded for Anonymous User (n/a) at The Second Hospital of Hebei Medical University from ClinicalKey.com by Elsevier on November 10, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.

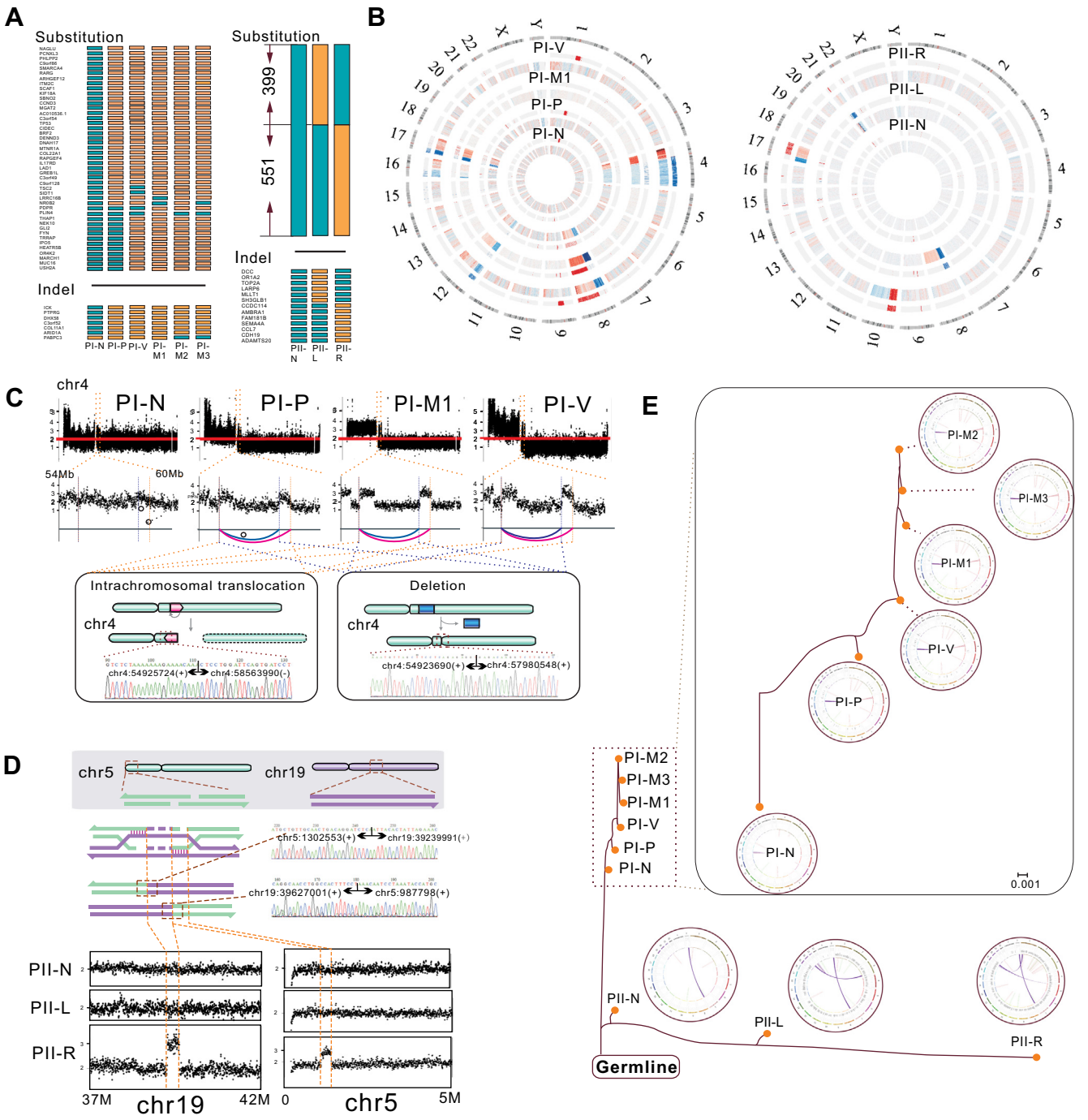


Fig. 2. Genomic alterations in multifocal HCCs. (A) Overlapping of genomic mutations (substitutions and indels) in the coding sequence (CDS) regions in PI (left) and PII (right). Orange and blue indicates the presence of a mutation and a normal allele, respectively. (B) Circos plots of copy number variations (CNVs) in PI (left) and PII (right) genomes. The outer and inner rings represent the CNVs obtained from SNP and whole-genome sequencing, respectively. Amplifications are shown in red, and deletions in blue. Color intensity correlates with the magnitude of a mutation. (C, D) Examples of somatically acquired structural variations. An intra-chromosomal translocation and a 3 Mb deletion in 4q12 of PI-P, PI-M1, and PI-V are shown together with the relevant CNVs in chromosome 4 (C). A balanced inter-chromosomal translocation due to homologous recombination between chromosomes 5 and 19 in PII-R is shown with the relevant 315 kb amplification in 5p15.33 and 387-kb amplification in 19q13.2 (D). (E) Phylogenetic tree and Circos plots of CDS mutations and structural variations. The length of each line is proportional to the genetic distance between the two connected nodes. The further away from the phylogenetic origin, the more severe the mutations.

coding sequence (CDS) regions overlapping with tumours' mutations) were noted in PI-N, when compared to PII-N where minimal mutations were noted (Supplementary Fig. 4, Supplementary Table 4). Moreover in PI, similar mutation patterns in all regions (including coding and non-coding genic regions, and intergenic regions) were observed among all tumour

Downloaded for Anonymous User (n/a) at The Second Hospital of Hebei Medical University from ClinicalKey.com by Elsevier on November 10, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.

tissues (Fig. 2A, Supplementary Fig. 4, Supplementary Table 4). In striking contrast, the two HCC tumours in PII had distinct mutation profiles (Fig. 2A, Supplementary Fig. 4). Further KEGG (Kyoto Encyclopedia of Genes and Genomes) [17] pathway enrichment of CDS mutations revealed significant enrichments of p53 signaling in all PI tumours, whereas no critical cancer-related pathways were enriched in the PII tumours (Supplementary Tables 5–8).

Second, copy number variations (CNVs) were assessed by analysing whole-genome sequencing and SNP genotyping data. As shown by Circos [18] plots in Fig. 2B, comparable results were obtained from both approaches. In PI, four regions with large deletions (4q12-4q35.2, 12p13.33-12p11.1, 16p13.3-16p11.2, and 17p13.3-17p11.2) and one amplified region (8p11.1-8q24.23) were detected in all tumours (PI-P, PI-M1, and PI-V) but not in PI-N. Moreover, the amplification in chromosome 8 and deletions in chromosomes 12 and 17 were more substantial in PI-M1, whereas deletions in chromosomes 4 and 16 were more profound in PI-V, suggesting that tumour subclones further selectively mutated during metastasis (Fig. 2B, Supplementary Table 9). Such examples of whole chromosome instability are characterised by aneuploidy, which is a major form of genomic instability observed in HCC and may promote tumourigenesis [19].

In PII, two large deletions in 8p23.3-8p12 and 21q11.2-21q22.3 were only noted in PII-L and two amplifications in 10p15.3-10p11.1 and 17q11.1-17q25.3 were exclusive for PII-R, implicative of two independent tumourigenic processes (Fig. 2B, Supplementary Table 10), as inferred by the different HBV integration sites.

In parallel, genomic structural variations (SVs) were analysed and further validated by PCR and Sanger sequencing. A total of 13 and 26 SVs were identified in PI and PII, respectively (Supplementary Tables 11 and 12) and examples are shown in Fig. 2C and D and Supplementary Fig. 5. SV patterns in all tissues recapitulate, to a great extent, their CNV patterns.

Lastly, a phylogenetic tree was constructed to predict the temporal development of each tissue, regardless of their germline differences (Fig. 2E). Three of the metastatic tumours of PI (PI-M1-3), in close proximity, were phylogenetically most distant from the putative germline, when compared to other tissues (PI-P, PI-V, or PI-N). In comparison, patterns of the two tumours in PII were distant from germline and also from each other, implicative of synchronous development of distinct clones. More intriguingly, Fig. 2E also demonstrated how PI tumours appear to have sequentially developed from the primary tumour (PI-P) first to the portal vein tumour thrombus (PI-V), and then to the satellite metastatic lesions (PI-M1-3). These data clearly explain the genomic similarities of all PI tumours and imply that these intrahepatic lesions spread via the portal venous blood, which is the most common type of HCC metastasis [20].

Transcriptomic analyses of multifocal HCC

To expand on the genomic studies, we next performed RNA-Seq followed by transcriptomic analysis to profile protein-coding gene expression in HCC. Overall, mRNA expression levels of genesfalling within the large CNV segments strongly correlated with the respective copy numbers (Supplementary Fig. 6A, Supplementary Tables 9 and 10). We also observed that associations of the differentially expressed genes between the PI tumours were more pronounced than those of PII-L and PII-R (Supplementary Fig. 6B). These data are indicative of genetic similarities between metastases and the primary HCC in PI.

Next, gene sets from Gene Ontology (GO) were enriched based on differentially expressed genes and were visualised by Cytoscape [21] plugin Enrichment Map [22]. As noted, several function modules, such as immune and inflammatory responses, coagulation, and normal liver functions (inclusive of amino acid and lipid metabolism, carbohydrate biosynthesis, and drug metabolism), were deregulated in all tumours (Fig. 3, Supplementary Table 13).

In general, the vast majority of functional changes in all PI tumours were comparable. We noted major changes in coenzyme metabolism and energy generation via mitochondrial oxidative phosphorylation. PI tumours further exhibited perturbations of carbohydrate catabolism and aerobic glycolysis, associated with increases in nucleic acid metabolism, protein translation and transport, macromolecular complex assembly, cell cycle and cell proliferation, and cell migration. Moreover, upregulation of genes participated in cytoskeletal remodelling and extracellular matrix organisation, considered essential for metastasis, were exclusively observed in PI-M1.

In contrast, distinct patterns of transcriptomic dysregulation were noted in each lesion in PII tumours albeit with some existing overlap (Fig. 3).

Functional changes in both PII tumours were relatively trivial, with the majority leading to decreased gene activities. Moreover, neither tumour in this patient displayed any suggestion of the molecular signatures of metastasis, observed in PI. These data are suggestive of non-invasive phenotypes of PII tumours emerging from distinct premalignant clones in the non-cirrhotic liver.

Collectively, these findings suggest that transcriptomic analysis corroborated genetic alterations identified at the genomic level and could suggest clonality, aggressiveness, and metastatic potential of multifocal HCC.

Validation studies of biomarkers in large cohort of patients with HCC

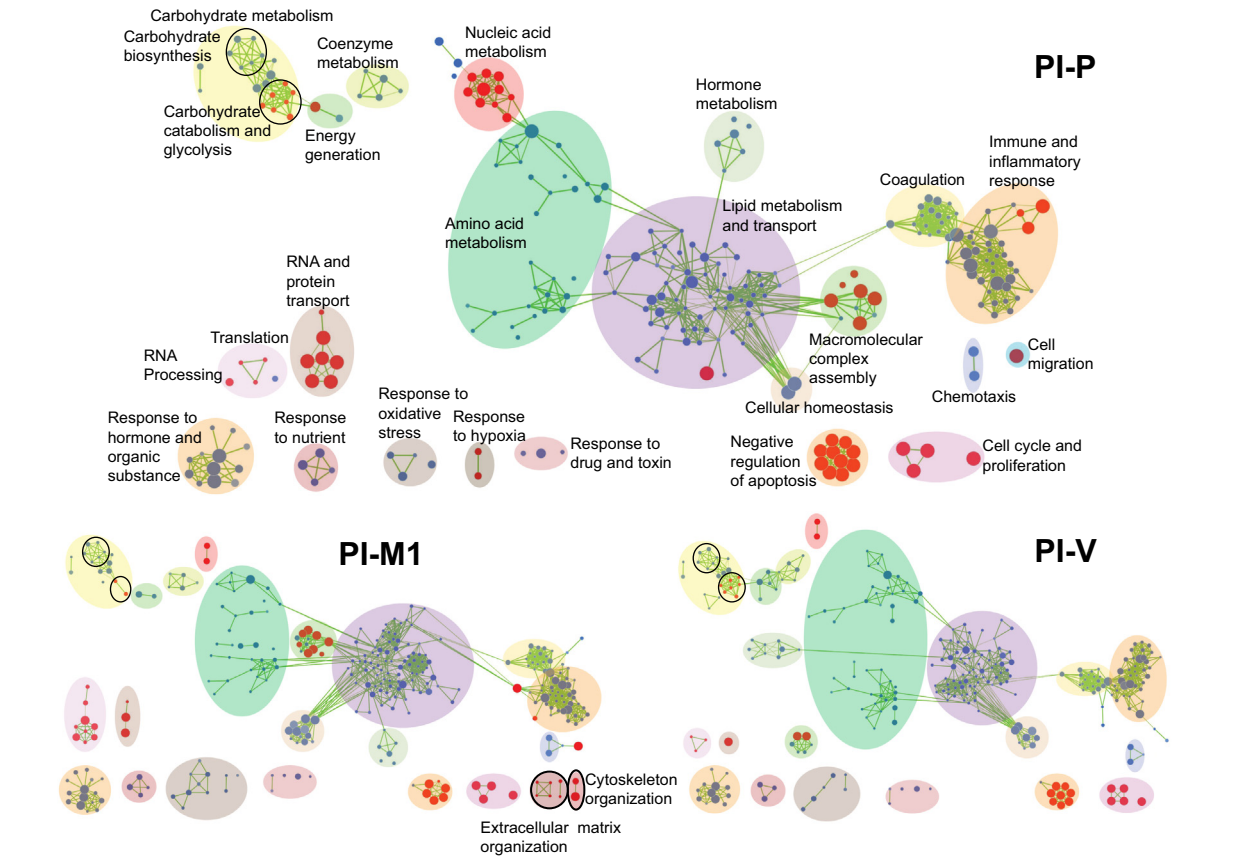
Lastly, to test the potential utility of our novel personalised multi-omics model in HCC biomarker identification, regardless of the types of HCC development, we then applied these multi-omics results to search for novel biomarkers for HCC diagnosis and prognosis by studying a larger cohort of HBV-HCC patients.

Genes with over four-fold differential expression in the PI and PII tumours were evaluated for pathway enrichment. Five KEGG or BioCarta [23] pathways comprising a total of 21 genes, including cell cycle (hsa04110), p53 signaling (hsa04115), histidine metabolism (hsa00340), G2/M checkpoint (g2Pathway), and Ran-mediated mitotic spindle regulation (ranMSpathway), were found to be exclusively enriched in PI tumours (Fig. 4A).

Seven of the highest level expressing genes (>10 fold) were selected for further validation by qRT-PCR using paired tumour/liver tissues from 174 HBV-HCC patients. Alterations in gene expression at mRNA levels were confirmed for six genes inclusive of *HAL*, *SFN*, *KIF15*, *TTK*, *BUB1*, and *MCM4* (Fig. 4B).

In parallel, various clinicopathological characteristics including age, HBsAg, ALT, albumin, Edmondson tumour grade, and satellite lesion were found to be highly associated with at least one gene, i.e., mRNA levels of *SFN* ($p = 0.0004$), *TTK* ($p = 0.0051$), *BUB1* ($p = 0.0084$), and *MCM4* ($p = 0.0126$) were significantly correlated with tumour grade (Supplementary Table 14).

Downloaded for Anonymous User (n/a) at The Second Hospital of Hebei Medical University from ClinicalKey.com by Elsevier on November 10, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.



Downloaded for Anonymous User (n/a) at The Second Hospital of Hebei Medical University from ClinicalKey.com by Elsevier on November 10, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.

Furthermore, we evaluated correlations of gene expression levels with postsurgical prognosis. Remarkably, *TTK* mRNA levels were inversely correlated with RFS and OS of patients (Fig. 4C and D, Supplementary Fig. 7, Supplementary Tables 15 and 16). Empirically, types of multifocal HCC development post hepatectomy may be distinguished according to the interval. Early recurrences ($\leqslant$1 year) are thought to arise mainly from intrahepatic metastases, whereas late recurrences (>1 year) are more likely to be multicentric in origin [24]. We noted that among patients with recurrent HCC, the median RFS was 3.53 months in *TTK*-high group *vs.* 12.48 months in *TTK*-low group ($p = 0.0122$). These data suggest that *TTK* expression might be an independent prognostic indicator for metastatic potential, postsurgical recurrence, and survival of HCC patients.

Discussion

This translational study with the application of comprehensive and personal multi-omics analyses in HCC has two major innovative findings. First, we were able to distinguish between metastatic disease *vs.* synchronous primary tumour development at the levels of genomic and transcriptomic studies. Our findings, for the first time, not only confirmed but also defined the two putative multifocal HCC development models via analysis of tumour clonality. Secondly, using multi-omics data from these initial studies, integrated with detailed clinicopathological information, we were able to efficiently identify HCC biomarkers, which were further validated using a large cohort of HCC patients.

Sequential biopsies from diseased livers and from distinct multifocal HCC could provide valuable tools to uncover stepwise changes during hepatocarcinogenesis [25]. By contrasting multi-omics profiling between tumour nodules as well as with control liver and non-mutated genome of the same patient, we have identified key mutational similarities and differences. Similarities in catastrophic alterations of genomes in multifocal satellite lesions and the presumed primary tumour in PI are suggestive of a single clonal origin for the HCC, albeit with subsequent differential sub-clonal mutations. This type of tumour evolution is similar to metastases of other epithelial cancers such as prostate [26], pancreatic [27], and breast cancer [28].

Moreover, varying degrees of genomic alterations in the control liver tissue of PI are indicative of molecular changes accumulated during the destruction of normal liver structures and nodule regeneration in a severely diseased cirrhotic liver with a likely field effect [29]. Under such disease-predisposing liver conditions [30], additional genetic mutations in critical genes, e.g., p53 pathway components and cell-cycle regulators, may facilitate hepatocyte malignant transformation and HCC progression. Indeed, these multi-omics observations and assumptions are in keeping with clinicopathological features and poor prognosis of this patient who had cirrhosis and died of HCC recurrence.

In stark contrast, trivial genetic and functional changes in the control liver of PII, with milder liver fibrosis and better prognosis, are consistent with distinct mutation patterns in two tumours that lack gene and function enrichments in cancer-related pathways. These findings suggest that two clones independently expanded from different premalignant hepatocytes, likely due to persistent HBV activity in the host genome, associated with low carcinogenicity.

These approaches described here, once standardised could provide valuable molecular implications for personalised treatment and stratification of HCC. Molecular approaches to differentiate multicentric occurrence from intrahepatic metastasis and to evaluate the aggressiveness of existing lesions are not yet available for clinical use. The future availability of such robust and simple tests would be critically important for surgical decision-making. Currently, however, the limited feasibility of these multi-omics approaches in terms of high-throughput and cost largely limits their clinical utility.

Moreover, assessment of individual tumour pathogenesis and aggressiveness of resected nodules by multi-omics profiling may provide useful prognostic information to facilitate personalised postsurgical management (e.g., adjuvant therapy and surveillance), and direct treatment strategies for recurrences. For example, patients with aggressive intrahepatic metastases such as PI could be treated with targeted systemic therapy based on the molecular profile of the original tumour (e.g., restoring p53 function in p53 mutant tumours) [31]. On the other hand, due to the nature of multicentric and synchronous HCC, targeted therapy choices in these patients would need to be based on analysis of molecular changes in the *de novo* tumours instead of the previously analysed lesions. Patients with non-invasive multiple primary HCC such as PII may be followed with regular surveillance imaging, may benefit from treatment of the underlying liver disease, and may be a candidate for resection of any subsequent recurrences.

Additionally, we demonstrate that comparative multi-omics profiling of a complete collection of representative patient specimens is very effective for HCC biomarker identification. Expression levels of four out of seven HCC-related genes identified/examined here, namely *SFN*, *TTK*, *BUB1*, and *MCM4*, are strongly associated with different tumour differentiation patterns. *TTK* is further shown to be a promising prognostic marker for HCC. Intriguingly, high-level expression of *TTK* has been shown to significantly correlate with aggressive clinical courses and low survival rates in HCC. As a dual-specific protein kinase participating in p53 pathway, *TTK* has been reported to contribute to tumourigenesis in many cancers by modulating the mitotic checkpoint [32]. We also note that *TTK* has been previously demonstrated in a microarray study to predict survival in a subgroup of HCC patients [33]. All these observations suggest that *TTK* might be a *bona fide* HCC biomarker with prognostic significance. However, the role of *TTK* in hepatocarcinogenesis (with possible links to p53 signaling) [32] remains elusive.

Prior well-conducted gene-expression profiling studies have suggested several multi-gene score systems that can be used to classify HCC or predict overall survival and/or recurrence [29,34,35]. Our new prognostic maker, a single TTK gene model,

Fig. 3. Functional enrichment maps of the differentially expressed protein-coding genes. The enriched gene sets from Gene Ontology based on two-fold differentially expressed genes were visualised by Cytoscape plugin Enrichment Map. For each sample, each node represents a gene set; size of the node is indicative of the number of genes and the colour intensity reflects the level of significance. Up-regulated gene sets are shown in red, and down-regulated ones in blue. The thickness of each line is proportional to number of genes shared by connected gene sets. Map for each sample is differently magnified for easier visualisation.

Downloaded for Anonymous User (n/a) at The Second Hospital of Hebei Medical University from ClinicalKey.com by Elsevier on November 10, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.

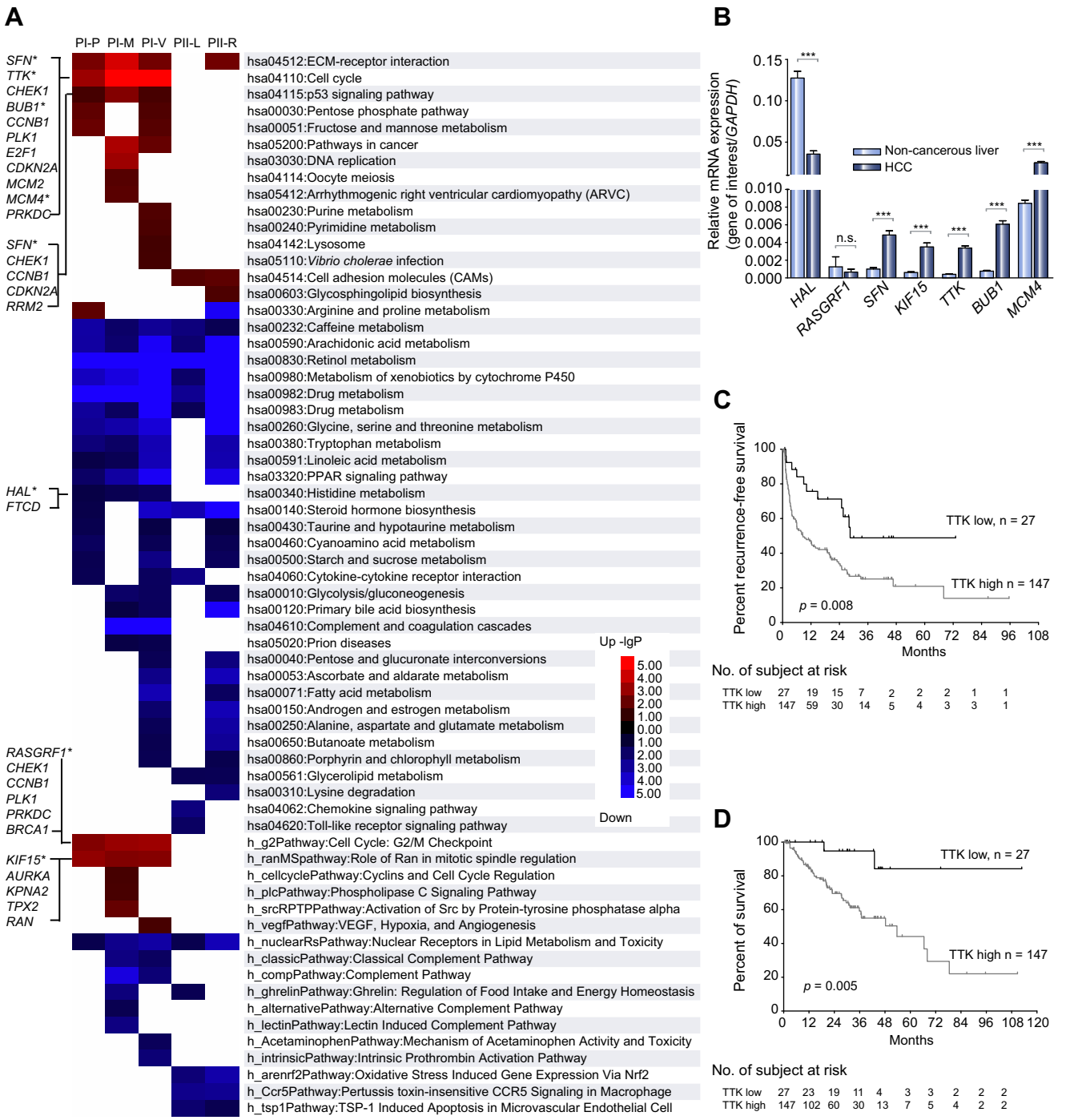


Fig. 4. Validation of HCC biomarker identification data. (A) Pathway enrichment analyses of genes with over four-fold differential expression in the KEGG (upper) and BioCarta (lower) pathways. Significantly enriched pathways with p <0.05 are shown with up-regulated ones in red and down-regulated ones in blue. On the left, genes (contained to five pathways) only enriched in all PI tumours were listed from highest to lowest according to their average fold changes within each pathway. *Genes selected for qRT-PCR validation. (B) qRT-PCR analyses of mRNA expression of seven select genes using paired HCC and non-cancerous tissues from 174 HBV-HCC patients. Data are given as mean ± SEM. ***p <0.0001; n.s., not significant. (C, D) Correlations of tissue *TTK* mRNA levels with recurrence-free survival (C) or overall survival (D) of patients.

can not only accurately predict the recurrence rate, but also the time of recurrence, which is very critical for clinical decision-making. Early interventional therapy is suggested for TTK-high patients (3.53 months of RFS), whereas TTK-low patients (12.48 months of RFS) only require standard surveillance post hepatectomy. In contrast, other current clinical biomarkers such as AFP can only be used as postoperative surveillance factors. Intriguingly, it remains to be determined whether combining

Downloaded for Anonymous User (n/a) at The Second Hospital of Hebei Medical University from ClinicalKey.com by Elsevier on November 10, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.

TTK with other previously reported prognostic signatures [29,34,35] provides an improved predictive potency.

Of course, this novel integrative multi-omics approach needs further validation using large and ideally prospective HCC (inclusive of non-HBV-related) patient cohorts. Further advances in NGS and computational methods, with increasing capacity and decreasing cost, will expedite comprehensive characterisation of the genetic and functional alterations occurring in individual tumours. These advances may accelerate clinical translation of "omics" science into clinical applications that will direct personalised molecular medicine in HCC, and potentially all types of human cancer.

Financial support

This work was supported by the International Science and Technology Cooperation Projects of China (2010DFB33720 and 2010DFA31840), Training Program of the Major Research Plan of the National Natural Science Foundation of China (91229120), Program for New Century Excellent Talents in University of China (NCET-11-0288), and National Institutes of Health of United States (P01HL107152-01).

Conflict of interest

The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

Acknowledgments

We thank Dr Yu Xiao at the Department of Pathology, Peking Union Medical College Hospital (PUMCH) for assisting in pathologic evaluations. The whole-genome and transcriptome sequencing were performed at Beijing Genomics Institute (BGI), and we thank Mr. Mingrong Zhang for coordinating these processes.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhep.2014.05.025.

References

[1] Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin 2011;61:69–90.
[2] Maluccio M, Covey A. Recent progress in understanding, diagnosing, and treating hepatocellular carcinoma. CA Cancer J Clin 2012;62:394–399.
[3] Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. Hepatology 2011;53:1020–1022.
[4] EASL-EORTC clinical practice guidelines. Management of hepatocellular carcinoma. J Hepatol 2012;56:908–943.
[5] Llovet JM, Fuster J, Bruix J. The Barcelona approach: diagnosis, staging, and treatment of hepatocellular carcinoma. Liver Transpl 2004;10:S115–S120.
[6] Omata M, Lesmana LA, Tateishi R, Chen PJ, Lin SM, Yoshida H, et al. Asian Pacific Association for the Study of the Liver consensus recommendations on hepatocellular carcinoma. Hepatol Int 2010;4:439–474.
[7] Budhu A, Forgues M, Ye QH, Jia HL, He P, Zanetti KA, et al. Prediction of venous metastases, recurrence, and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. Cancer Cell 2006;10:99–111.
[8] Yamamoto T, Kajino K, Kudo M, Sasaki Y, Arakawa Y, Hino O. Determination of the clonal origin of multiple human hepatocellular carcinomas by cloning and polymerase chain reaction of the integrated hepatitis B virus DNA. Hepatology 1999;29:1446–1452.
[9] Morimoto O, Nagano H, Sakon M, Fujiwara Y, Yamada T, Nakagawa H, et al. Diagnosis of intrahepatic metastasis and multicentric carcinogenesis by microsatellite loss of heterozygosity in patients with multiple and recurrent hepatocellular carcinomas. J Hepatol 2003;39:215–221.
[10] Tao Y, Ruan J, Yeh SH, Lu X, Wang Y, Zhai W, et al. Rapid growth of a hepatocellular carcinoma and the driving mutations revealed by cell-population genetic analysis of whole-genome data. Proc Natl Acad Sci U S A 2011;108:12042–12047.
[11] Sung WK, Zheng H, Li S, Chen R, Liu X, Li Y, et al. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. Nat Genet 2012;44:765–769.
[12] Fujimoto A, Totoki Y, Abe T, Boroevich KA, Hosoda F, Nguyen HH, et al. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. Nat Genet 2012;44:760–764.
[13] Puente XS, Pinyol M, Quesada V, Conde L, Ordonez GR, Villamor N, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature 2011;475:101–105.
[14] Berger MF, Hodis E, Heffernan TP, Deribe YL, Lawrence MS, Protopopov A, et al. Melanoma genome sequencing reveals frequent PREX2 mutations. Nature 2012;485:502–506.
[15] Sun Y, Li C, Shu Y, Ju X, Zou Z, Wang H, et al. Inhibition of autophagy ameliorates acute lung injury caused by avian influenza A H5N1 infection. Sci Signal 2012;5:ra16.
[16] Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, Jones D, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. Nature 2009;463:184–190.
[17] Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res 2012;40:D109–D114.
[18] Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. Genome Res 2009;19:1639–1645.
[19] Thomas RM, Berman JJ, Yetter RA, Moore GW, Hutchins GM. Liver cell dysplasia: a DNA aneuploid lesion with distinct morphologic features. Hum Pathol 1992;23:496–503.
[20] Blumgart LH, Fong Y. Surgery of the liver and biliary tract. W.B. Saunders; 2000.
[21] Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 2011;27:431–432.
[22] Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PLoS One 2010;5:e13984.
[23] Nishimura D. BioCarta. Biotech software & internet. Report 2001;2:117–120.
[24] Poon RT, Fan ST, Ng IO, Lo CM, Liu CL, Wong J. Different risk factors and prognosis for early and late intrahepatic recurrence after resection of hepatocellular carcinoma. Cancer 2000;89:500–507.
[25] Thorgeirsson SS, Grisham JW. Molecular pathogenesis of human hepatocellular carcinoma. Nat Genet 2002;31:339–346.
[26] Liu W, Laitinen S, Khan S, Vihinen M, Kowalski J, Yu G, et al. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. Nat Med 2009;15:559–565.
[27] Campbell PJ, Yachida S, Mudie LJ, Stephens PJ, Pleasance ED, Stebbings LA, et al. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 2010;467:1109–1113.
[28] Ding L, Ellis MJ, Li S, Larson DE, Chen K, Wallis JW, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature 2010;464:999–1005.
[29] Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N Engl J Med 2008;359:1995–2004.
[30] Aravalli RN, Steer CJ, Cressman EN. Molecular mechanisms of hepatocellular carcinoma. Hepatology 2008;48:2047–2063.
[31] Brown CJ, Cheok CF, Verma CS, Lane DP. Reactivation of p53: from peptides to small molecules. Trends Pharmacol Sci 2011;32:53–62.

Downloaded for Anonymous User (n/a) at The Second Hospital of Hebei Medical University from ClinicalKey.com by Elsevier on November 10, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.

[32] Janssen A, van der Burg M, Szuhai K, Kops GJ, Medema RH. Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. Science 2011;333:1895–1898.
[33] Lee JS, Chu IS, Heo J, Calvisi DF, Sun Z, Roskams T, et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology 2004;40:667–676.
[34] Boyault S, Rickman DS, de Reynies A, Balabaud C, Rebouissou S, Jeannot E, et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. Hepatology 2007;45:42–52.
[35] Nault JC, De Reynies A, Villanueva A, Calderaro J, Rebouissou S, Couchy G, et al. A hepatocellular carcinoma 5-gene score associated with survival of patients after liver resection. Gastroenterology 2013;145:176–187.

Downloaded for Anonymous User (n/a) at The Second Hospital of Hebei Medical University from ClinicalKey.com by Elsevier on November 10, 2021. For personal use only. No other uses without permission. Copyright ©2021. Elsevier Inc. All rights reserved.