



# Cross Analysis of Genomic-Pathologic Features on Multiple Primary Hepatocellular Carcinoma

Fei Ren<sup>1</sup>, Depin Wang<sup>1</sup>, Xueyuan Zhang<sup>1,2</sup>, Na Zhao<sup>2</sup>, Xiaowen Wang<sup>2</sup>, Yu Zhang<sup>3\*</sup> and Li Li<sup>4\*</sup>

<sup>1</sup>High Performance Computer Research Center, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, <sup>2</sup>Zhijian Life Co. Ltd., Beijing, China, <sup>3</sup>Department of Medical Oncology and Radiation Sickness, Peking University Third Hospital, Beijing, China, <sup>4</sup>Department of Oncology, Peking University International Hospital, Beijing, China

## OPEN ACCESS

### Edited by:

Sun Liang,  
Shandong First Medical University,  
China

### Reviewed by:

Kevin Waters,  
Cedars-Sinai Medical Center,  
United States  
Dongbo Chen,  
Peking University, China

### \*Correspondence:

Yu Zhang  
13683371277@163.com  
Li Li  
lili1@pkuhl.edu.cn

### Specialty section:

This article was submitted to  
Computational Genomics,  
a section of the journal  
Frontiers in Genetics

Received: 31 December 2021

Accepted: 25 May 2022

Published: 20 June 2022

### Citation:

Ren F, Wang D, Zhang X, Zhao N,  
Wang X, Zhang Y and Li L (2022) Cross  
Analysis of Genomic-Pathologic  
Features on Multiple Primary  
Hepatocellular Carcinoma.  
Front. Genet. 13:846517.  
doi: 10.3389/fgene.2022.846517

Hepatocellular carcinoma (HCC) is a prevalent malignancy cancer worldwide with a poor prognosis. Hepatic resection is indicated as a potentially curative option for HCC patients in the early stage. However, due to multiple nodules, it leads to clinical challenges for surgical management. Approximately 41%–75% of HCC cases are multifocal at initial diagnosis, which may arise from multicentric occurrence (MO-HCC) or intrahepatic metastasis (IM-HCC) pattern with significantly different clinical outcomes. Effectively differentiating the two mechanisms is crucial to prioritize the allocation of surgery for multifocal HCC. In this study, we collected a multifocal hepatocellular carcinoma cohort of 17 patients with a total of 34 samples. We performed whole-exome sequencing and staining of pathological HE sections for each lesion. Reconstruction of the clonal evolutionary pattern using genome mutations showed that the intrahepatic metastogenesis pattern had a poorer survival performance than independent origins, with variants in the TP53, ARID1A, and higher CNV variants occurring more significantly in the metastatic pattern. Cross-modality analysis with pathology showed that molecular classification results were consistent with pathology results in 70.6% of patients, and we found that pathology results could further complement the classification for undefined patterns of occurrence. Based on these results, we propose a model to differentiate the pattern of multifocal hepatocellular carcinoma based on the pathological results and genome mutations information, which can provide guidelines for diagnosing and treating multifocal hepatocellular carcinoma.

**Keywords:** hepatocellular carcinoma, whole exome sequencing, whole slide images, molecular profiling and subtyping, hepatitis B virus

## INTRODUCTION

Liver cancer is ranked as the sixth most common malignancy cancer, and its incidence is rising (Sung et al., 2021). Hepatocellular carcinoma (HCC) is the most common form of liver cancer, accounting for approximately 90% of liver cancer cases (Llovet et al., 2021). Roughly 41%–75% of patients with HCC present with multiple intrahepatic tumors (Miao et al., 2014; Vogel et al., 2018). Despite there existing standardized guidelines for multifocal HCC and indications for surgical resection, surgical suggestions for individual patients remain complicated owing to the difficulty of accurately predicting future tumor progression. These uncertainties for the recurrence of primary lesions or metastatic possibility provide challenges to the prognosis after surgery for individual patients (Viganò et al., 2019).

Multifocal HCC may arise synchronously or metachronously as a separate primary tumor (multicentric occurrence) or develop due to intrahepatic metastases from the same primary cancer (Baffy, 2015). Since the prognosis of hepatocellular carcinoma patients under these two types varies greatly, it is crucial to construct the correct diagnostic approach for these patients. Several assessment methods, including pathological examination, integration of hepatitis B virus (HBV) DNA by PCR and DNA blot analysis, and heterozygosity analysis of DNA microsatellite loci, have been recently developed to distinguish between these two types of multifocal HCC (Study, 2012). However, the combination of molecular and pathological profiling with analytical methods systematically used to distinguish between these two patterns are still lacking.

The advantage and rapid progress of next-generation sequencing, such as whole-exome sequencing (WES), has made it possible to comprehensively characterize the disease mechanisms and altered genes in multiple cancers (Ally et al., 2017; Yan et al., 2018; Nanki et al., 2020). This approach allows the identification of novel molecular markers and the definition of underlying biological mechanisms, thus facilitating the stratification and characterization of cancers (Cortés-Ciriano et al., 2022). In this study, we selected representative patients of HBV-associated multifocal HCC who underwent tumor resection and exhibited a variable postoperative course.

These HCC samples were conducted whole-exome sequencing (WES) to obtain a complete genetic alteration profiling for each patient. We then performed a systematic analysis of integrated genomics and further correlated these with clinic pathological data. We sought to comprehensively unravel the molecular differences between the two multifocal HCC models as well as differences in pathological features and identify molecular markers for diagnostic, prognostic, and potential therapeutic targets to guide the clinical diagnosis and treatment of multifocal hepatocellular carcinoma.

## METHODS

### Mutation Analysis

First, we aligned the exome sequencing clean reads against the human reference genome hg19 download from UCSC (<http://www.genome.ucsc.edu/>) using BWA (Li and Durbin, 2009) with the default parameters. To reduce systematic (non-random) technical error, we applied base quality recalibration with the Genome Analysis Toolkit (GATK) 4.0 (McKenna et al., 2010). The duplicated reads was removed from the alignment files using the Picard tools. Somatic variants, including single nucleotide variants (SNVs) and small insertions and deletions changes (Indels), were detected by Mutect2 of GATK 4.0 on the paired tumor and normal samples. High confidence variants were screened using the criteria of TLOD >10, and then they were annotated by the vcf2maf tool (<https://github.com/mskcc/vcf2maf>) to obtain nine types of mutations, including “Missense Mutation”, “Nonsense Mutation”, “Nonstop Mutation”, “Splice Site”, “Splice Region”, “In Frame Ins”, “In Frame Del”, “Frame Shift Ins” and “Frame Shift Del” mutations.

### Copy Number Analysis

The bioinformatics tool facet-suite (R package) (Shen and Seshan, 2016) was utilized to detect CNVs on paired sequencing reads of tumor and normal samples from the same patient. We first assessed the copy number of different segments and then filtered those segments with a total copy number greater than twice the DNA ploidy level as the amplification (AMP), and segments with a total copy number equal to zero as deletions (DEL). These AMP or DEL segments were the annotated with genes located in the genome context to obtain gene-level copy number alteration. To summarize total copy number variation at the level of the whole exome, we calculated a CNV score, which is similar to the TMB, simply by multiplying the length of CNV segments by their relative average altered weight.

### Tumor Mutational Burden Analysis

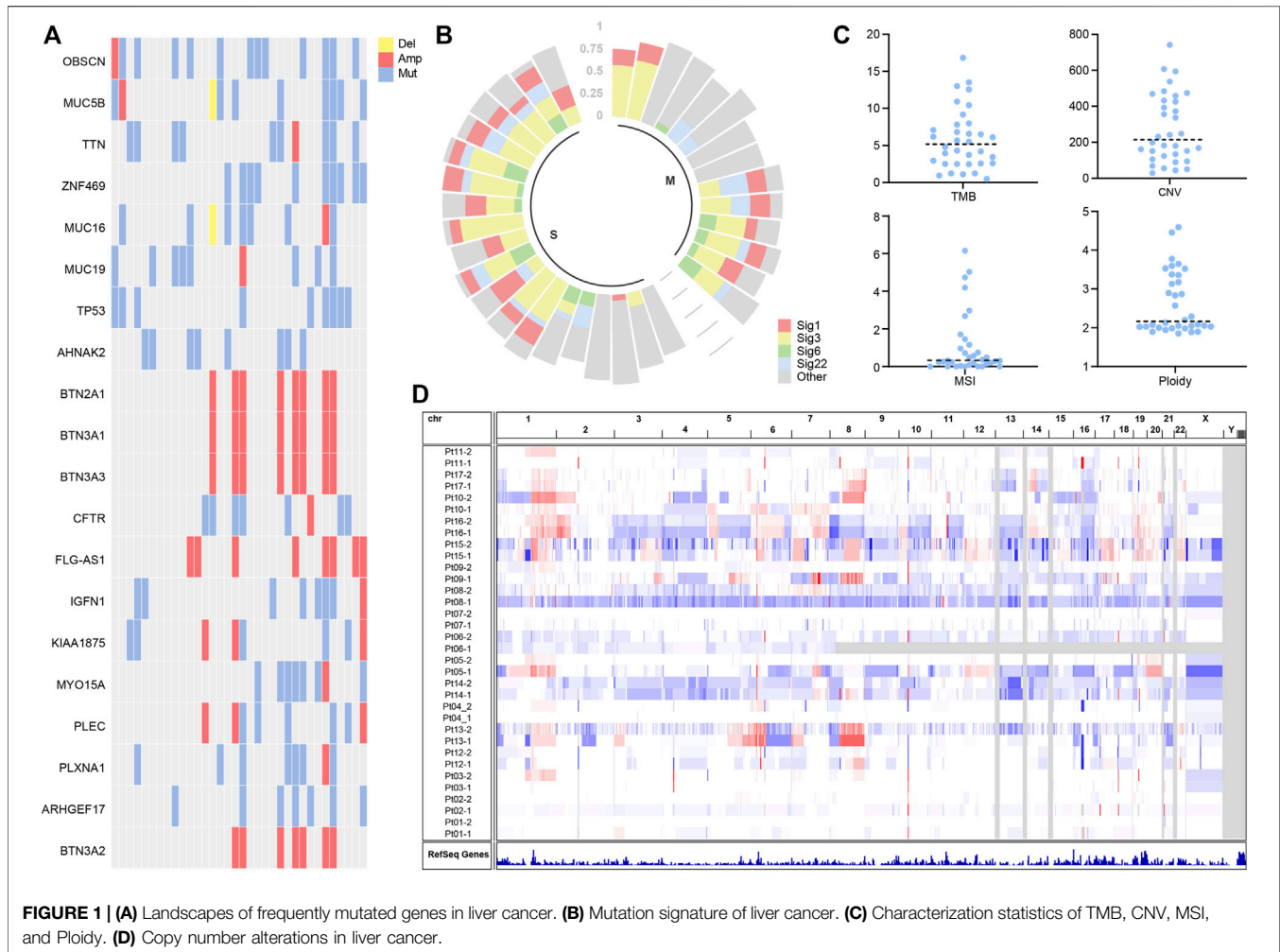
As the predictive biomarker in solid tumors (Wu et al., 2019), the tumor mutational burden (TMB), was calculated for all tumor samples by counting the non-synonymous mutation rate per megabases. We screened nine types of non-silent mutations from the analysis of the vcf2maf annotation tool. Those nine types of variants include “Splice Site”, “Splice Region”, “Missense Mutation”, “Nonstop Mutation”, “Nonsense Mutation”, “Frame Shift Ins” and “Frame Shift Del”, “In Frame Ins”, “In Frame Del”. Then these variants were all counted for TMB calculation, and the values were normalized by the total length of the CDS regions (36 Megabases) covered by the Agilent V6 whole exome (Wang et al., 2020).

### Microsatellite Instability and Mutational Signature Analysis

We evaluated the MSI status of the tumor samples with the bioinformatics tool Msisensor (Niu et al., 2014), and screened MSI-H samples with the criteria of an Msisensor score greater than 20 (Shimozaki et al., 2021). We determined the frequency of 96 mutated triplets per tumor sample based on the distribution of the six substitution patterns (C > A, C > G, C > T, T > A, T > C, T > G) and the neighbor 5' base and 3' base (Alexandrov et al., 2013). Together with their frequency, these triplets were summarized to construct a 96 × N mutation type frequency matrix, where N is the number of variants. We took the matrix as the input to determine the 1–30 mutational signatures (Alexandrov et al., 2020) from the Cosmic database (Tate et al., 2019) and to assess the proportion of specific mutational signatures in the samples using the bioinformatics tool DeConstructSig (Rosenthal et al., 2016). 30 mutational signatures were then reduced and classified to mutational signature 1, mutational signature 3, mutational signature 6, mutational signature 10 and others according to their different frequencies in the HCC samples (others represent the less frequently mutated mutational signatures in the HCC samples).

### Phylogenetic Analysis

Phylogenetic analyses were performed to elucidate genes essential for promoting tumor recurrence. We compared mutant variants in samples from different cancer samples, counted unique and



shared mutations, and used the common and unique mutations in two cancer samples to construct a phylogenetic tree. The phylograms were inferred using the R Bioconductor package phangorn (Schliep, 2011). Through phylogenetic tree analysis, we were able to identify early driver mutations and *de novo* mutations at different stages, thus providing a comprehensive interpretation of the relationships between different tumors.

## Pathology Image Analysis

Feature extraction of cell nucleus from pathology images mainly includes cancer region labeling, patch segmentation, color normalization, nucleus segmentation, nucleus-level and image-level feature extraction (Cheng et al., 2020): 1) The whole slide images (WSIs) were labeled the cancer region manually. 2) Non-overlapping image tiles with a size of 2048\*2048 pixels with a resolution of 0.5  $\mu\text{m}$  per pixel were extracted from Whole Slide Images (WSIs). To remove the bias of different staining procedures, all tiles were normalized based on one reference image using the Macenko normalization method. 3) Use a hierarchical multilevel thresholding approach to segment the nucleus for each tile. 4) Calculate 10 features of each nucleus in each image patch. 5) For the nuclei

of all patches in one WSI, each type of nucleus-level features was dissected into 15 image-level features by combining a 10-bin histogram and 5 distribution statistics (mean, std, skewness, kurtosis, and entropy). In total, we calculated 100 image-level features for each whole-slide image.

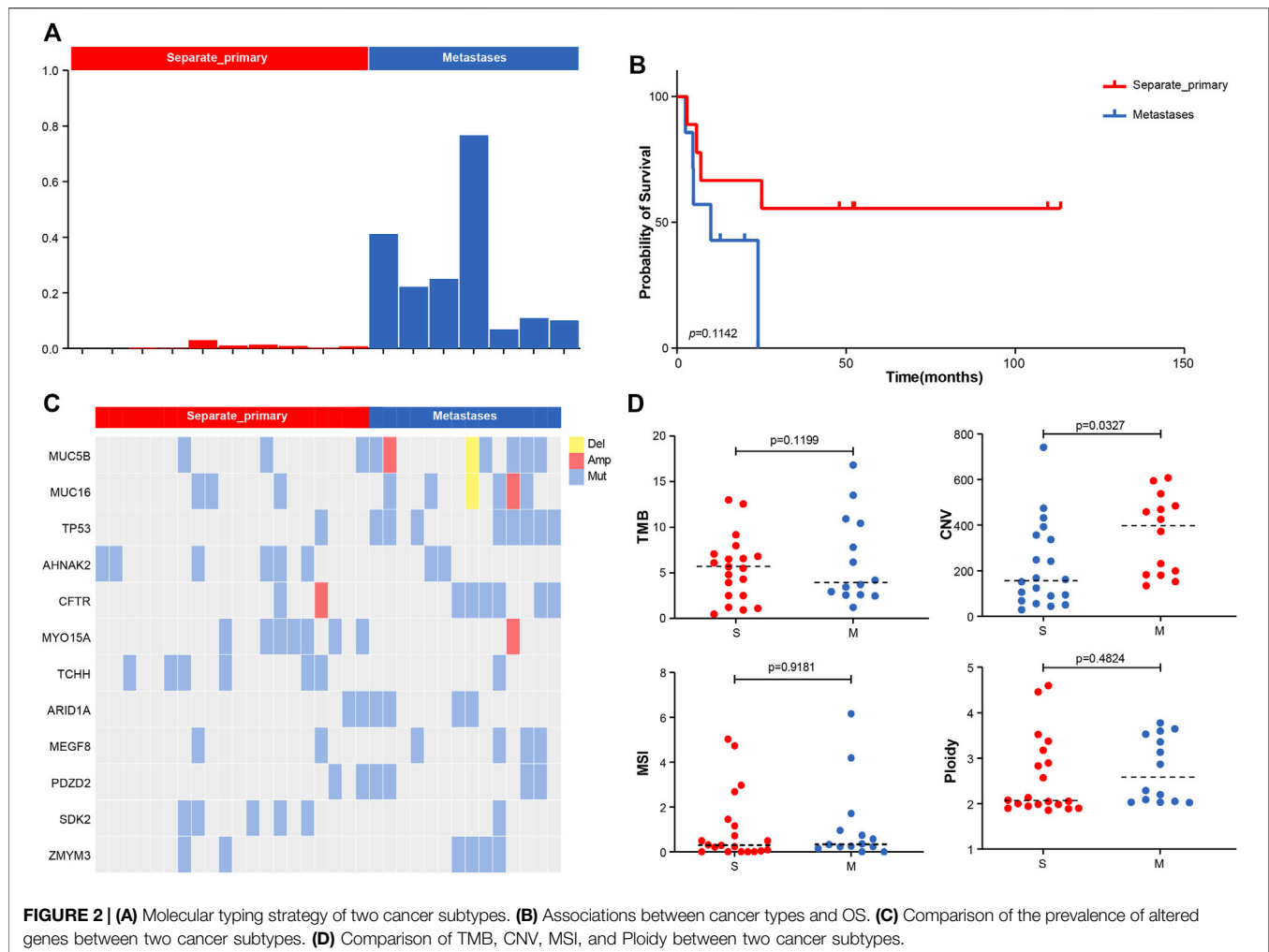
## Statistical Analysis

We use the student *t* test to compare the difference between two continuous variables. Kaplan-Meier survival analysis was used to obtain survival curves reflecting the differences in prognosis among tumor subtypes. Log-rank test was conducted to assess the correlation. Mann-Whitney *U* test was utilized to analyze the relationship between the two classification variables.

## RESULTS

### Collection of Multifocal HCC Samples

The clinical outcome of patients with HCC undergoing radical surgery are closely related to the number of intrahepatic tumors. The main purpose of this study was to explore genomic and pathological characteristics among the different intrahepatic



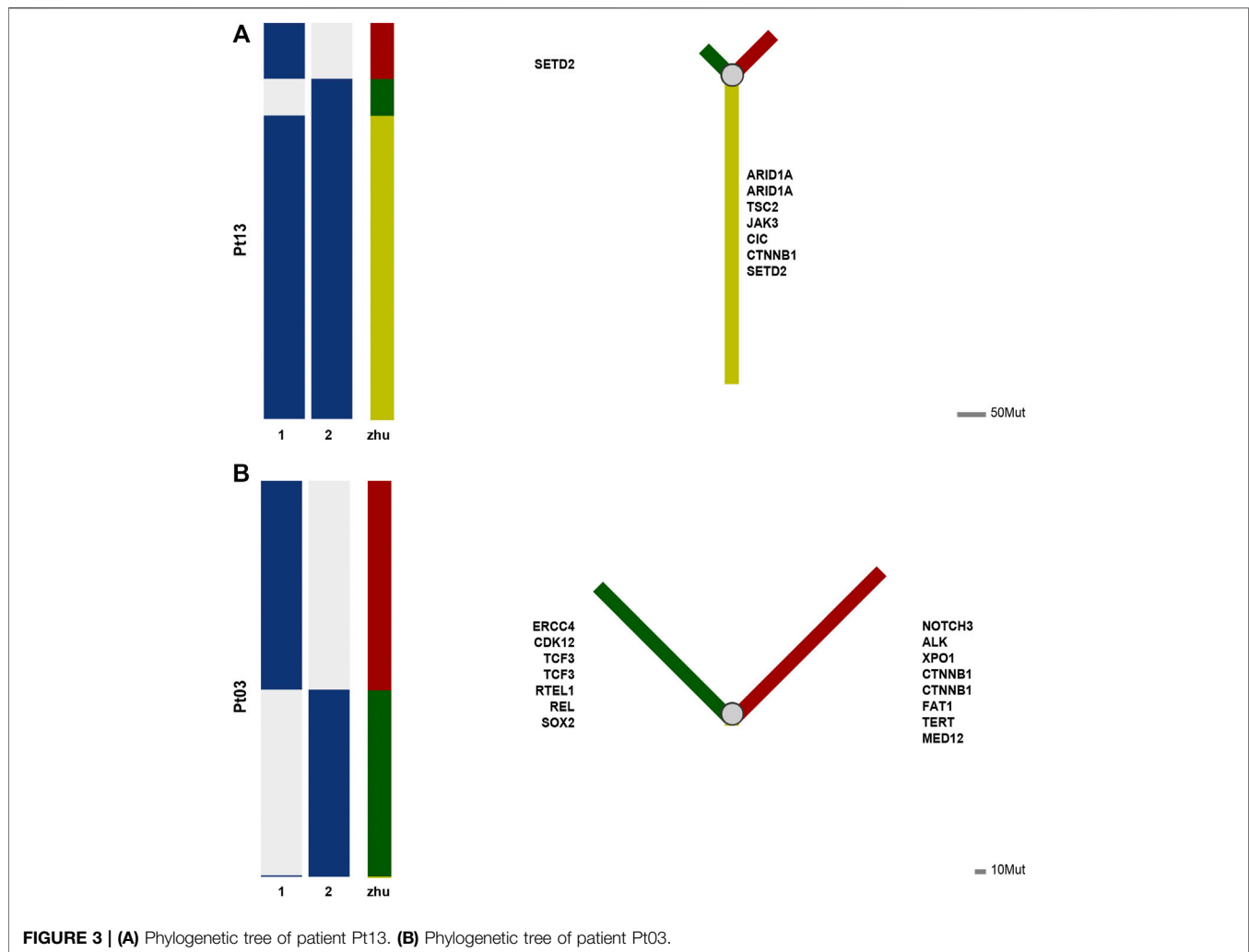
tumors and discover multiple modality indicators, thus we specified the intrahepatic tumor numbers to be 2. Multifocal HCC samples are collected from Peking University International Hospital and Peking University Third Hospital, and the inclusion criteria is as follows. The tumor satisfies the criteria for surgical indications defined by the Chinese CSCO Guidelines for primary HCC. Postoperative pathology confirmed that the tumor was hepatocellular carcinoma. Tumors were radically resected (R0), and the number of tumors was two. A total of 17 cases of patients met the criteria, and 34 tumor samples were performed whole-exome sequencing.

## Molecular Profiling of Liver Cancer

A total of 17 patients with multifocal liver cancer were recruited for our study, with two cancer foci collected per patient. Sixteen of them were male, one was female, and the cohort's median age was 45 years (distribution 43–67 years). The median follow-up time was 42 months. WES was performed on 34 tumor samples and paired samples of FFPE specimens, with a average 200× coverage depth for both the tumor and normal samples. The detailed clinical and pathological information of all patients used in this research is given in **Supplementary Table S1**.

To disentangle somatic mutations and molecular characteristics of multifocal hepatocellular carcinoma, mutation analysis of 34 tumor samples identified 7,752 individual mutations, including 6,378 single nucleotide variants (SNVs) and 1,374 small insertions and deletions changes (Indels) (**Supplementary Table S2**). The mean number of non-synonymous mutations per sample was 77 (range: 10–176), corresponding to 3.5 non-synonymous mutations per Megabyte (Mb), comparable to the TMB in the TCGA cohort. To explore potential driver mutations in patients, we summarized multiple genes with the highest mutation frequency (**Figure 1A**). The most commonly mutated genes in these patients were OBSCN, MUC5B, TTN, ZNF469, MUC16, TP53, with VAF greater than 25%. The frequency of TP53 variants is comparable to that observed in the TCGA cohort. Deletions were not widespread in genes with high mutation rates, while *BTN2A1*, *BTN3A1*, *BTN3A3*, *BTN3A2*, and *FLG-AS1* were amplified in several samples.

Mutation signature analysis showed mutation signature 1, 3, 6, and 22 to be more prevalent in patients (**Figure 1B**). According to published reports (Koh et al., 2021), Signature 6 is associated with DNA mismatch repair defects and MSI tumors. Signature 1 is associated with age at cancer diagnosis and has been detected in



most types of cancer samples. Signature 3 is associated with homologous repair and correlates with BRCA gene function. Signature 22 has been found in urothelial (renal pelvis) carcinoma and liver cancers.

We characterize molecular features of TMB, CNV, MSI, and Ploidy. It is shown that the median TMB was around 5.15 (Figure 1C), and the median CNV was assessed at 214.9 (Figure 1C). The MSI analysis showed that most liver cancer samples had low MSISensor scores, all less than 10 (Figure 1C). Most of the samples had a ploidy around 2 (Figure 1C). The copy number of chromosomes has more amplification events on chromosomes 1 and 8 (Figure 1D).

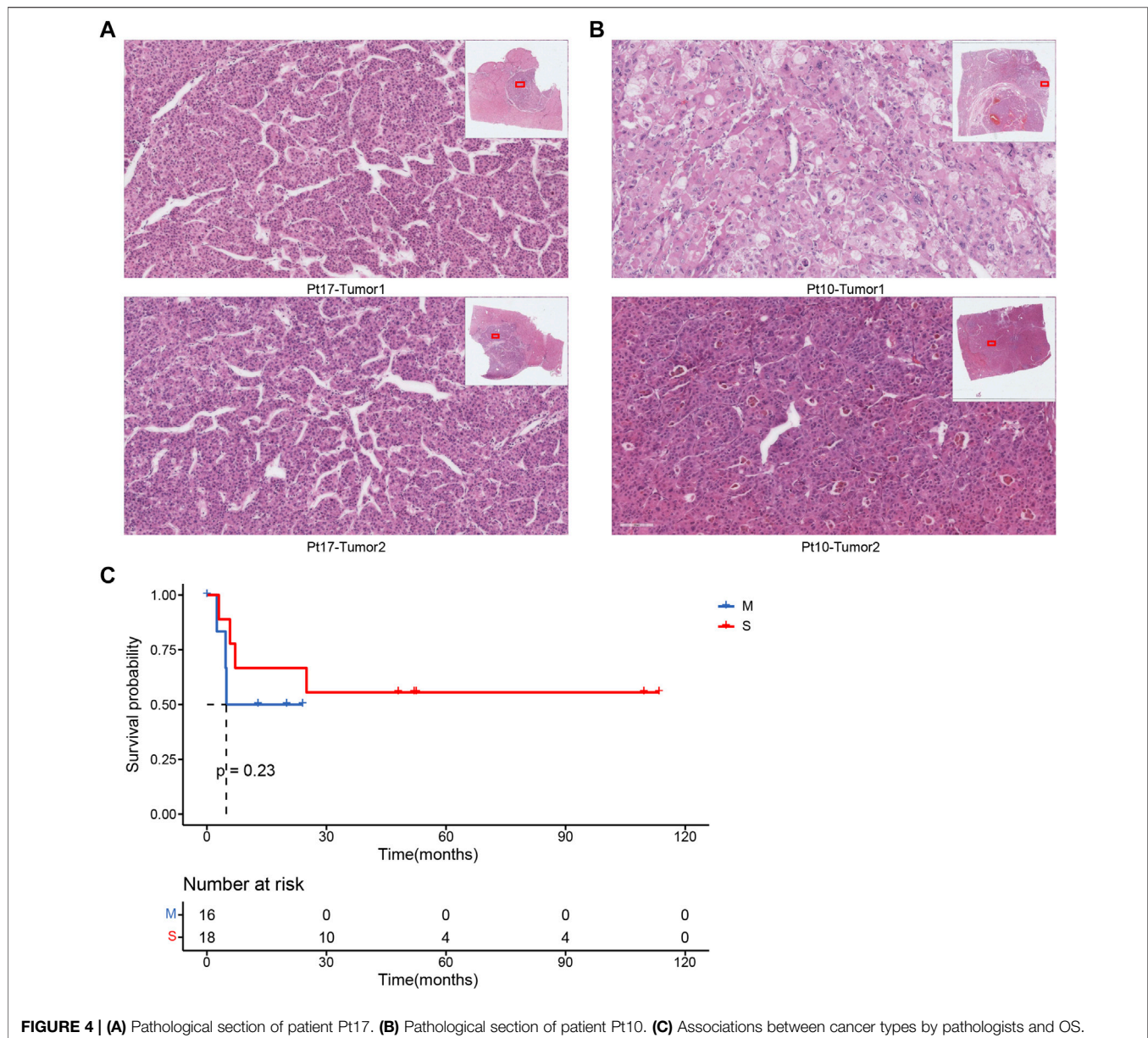
## Identification of Hepatocarcinogenesis Pattern by Genomic Signature

We calculated the Jaccard similarity coefficient (Jaccard Index) (Bu et al., 2021b) of two tumors in the sample patient based on the analysis of shared mutations. An index of 0.01 was taken as the screening threshold, and 17 patients were divided into two groups in total. Among them, we defined those with index <0.01 as

separate primary hepatocellular carcinoma, 10 cases in total, and those with index >0.01 as metastatic, 7 cases in total (Figure 2A). The index of the metastatic group ranged from 0.08 to 0.7. The analysis of PFS showed that patients with metastatic pattern showed worse survival ( $p$ -value = 0.1142) (Figure 2B). Analysis of differences in mutations between the two subgroups showed that TP53 was more inclined to be present in the subgroup with the metastatic pattern, with a  $p$ -value of 0.0212 (8/14 vs. 1/20). ARID1A had a slight elevation in metastasis, with  $p$ -value = 0.2022 (4/14 vs. 2/20) (Figure 2C). The analysis of the difference among TMB, CNV, MSI, and Ploidy showed a slight increase in TMB ( $p$ -value = 0.1199, average = 6.35 vs. 5.44) and a significant increase in CNV ( $p$ -value = 0.0327, average = 358.13 vs. 225.68) in the metastatic group. At the same time, there was no significant difference between MSI and Ploidy (Figure 2D).

## Phylogenetic Analysis of Hepatocarcinogenesis

According to Jaccard's similarity coefficient, seventeen individuals were divided into two groups, of which seven were



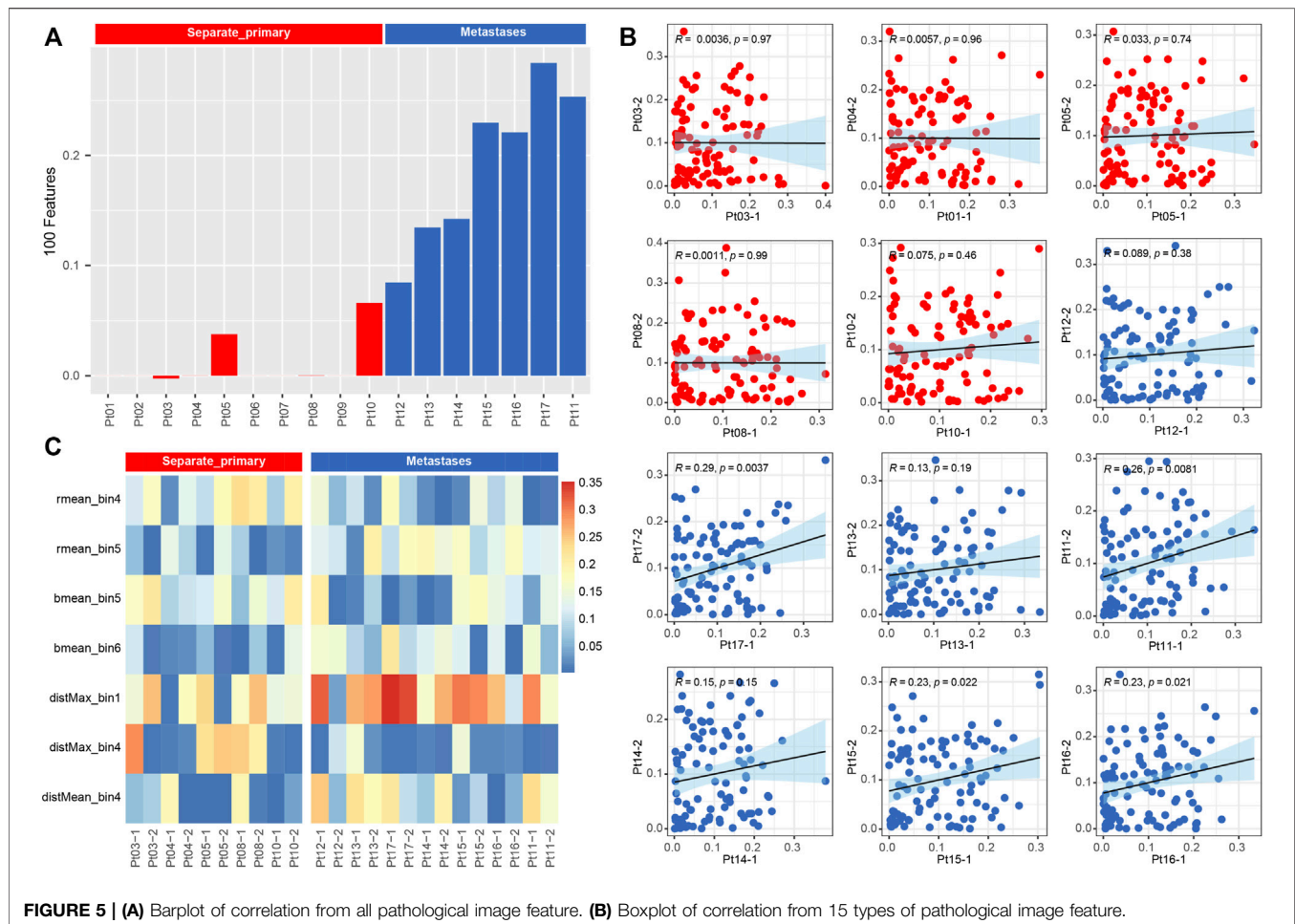
branching evolutionary (metastatic) patterns, and ten were independent occurrence patterns (**Supplementary Figure S1**). We further used phylogenetic tree analysis to show the evolutionary patterns of different cancer lesions and discover essential driver genes. The results shown by the phylogenetic tree were consistent with the Jaccard similarity coefficient.

For example, in the case of Pt13, the two tumors are highly similar according to the Jaccard similarity coefficient. Moreover, according to the results of the phylogenetic tree, a total of 107 mutations occurred in the two lesions, of which 82 mutations were shared in both samples (76.6%), i.e., located in the branching part of the shared phylogenetic tree (**Figure 3A**). Among them, ARID1A, TSC2, JAK3, CIC, CINNB1, and SETD2 were mutated at the early stage of carcinogenesis, which played an essential role in advancing early cancer development and progression.

In contrast, case of Pt03 had low level of similarity between the two tumors. As shown by the phylogenetic tree, 345 mutations occurred in either of tumor, while only TNIP2 was a shared early mutation (**Figure 3B**). And TNIP2 is less reported in cancer and is more like passenger mutation, so the mutation sharing here may be due to technical bias of sequencing or some accumulated alterations due to HBV infection. The phylogenetic trees of remaining cases are available at **Supplementary Figures S1–S5**.

### Pathological Cross-Analysis

The results of molecular testing can provide precise results for accurate diagnostic typing. However, more accessible in the clinic, pathology testing require simpler processing and short time consuming than molecular testing. Therefore, we attempted to



compare pathology results from different individuals to determine what percentage of molecular typing could be consistently distinguished by pathology typing. In simple words, assuming molecular typing results as the standard, we wanted to see how much of a typing indication the pathology could achieve. From there, we can determine the scenario in which pathology and molecules are used in combination with each other.

We had a mid-level pathologist interpret the pathological images of these 17 patients and then compared the results of pathological typing with those of molecular testing. The analysis of the results of the 17 cases showed that 70.6% (12 cases) of the molecular typing results could be distinguished by pathological indicators, using the typical indicators of the nuclei of the pathological sections as important measures. Case Pt17 was classified as metastatic by pathological typing because the cell morphology of the two tumors was very similar (Figure 4A). This result is consistent with the results of molecular typing. In contrast, case Pt10 was classified as the separate primary HCC because the cell morphology of the two tumors was quite different, such as the cellular atypia and sinusoids (Figure 4B). In addition, the survival analysis results of pathological typing showed that pathological interpretation could slightly distinguish between the two cancer subtypes (Figure 4C).

Meanwhile, we used a machine learning approach (Cheng et al., 2020) to extract 100 features of the pathological images, represented by matrix vectors, to discriminate between two subtypes by comparing the pathological features of two foci slides. First, all features were combined to calculate the correlation between the two foci of the same patient (Figure 5C), and the results showed that the subtypes could be distinguished by correlation ( $p < 0.001$ ) (Figure 5A). Second, all 100 pathological features were compared between two groups, and we found that the features of *rmean\_bin4*, *rmean\_bin5*, *bmean\_bin5*, *bmean\_bin6*, *disMax\_bin1*, *disMax\_bin4* and *distMean\_bin4* are significantly different between the two groups (Figure 5B).

## DISCUSSION

Hepatocellular carcinoma is a cancer with a high degree of malignancy (Chidambaranathan-Reghupaty et al., 2021). In this study, we collected a cohort of 17 patients with multifocal hepatocellular carcinoma. Then we utilized the bioinformatics approach to analyze the whole-exome molecular data and image data of H&E stained histology slides. By calculating the Jaccard

Index between two tumor and reconstructing the tumor clonal evolution, we revealed that intrahepatic metastasis and separate primary patterns reflected from the unique gene mutations and copy number alterations. We also utilized a machine learning approach to extract 100 features of the pathological images, to discriminate between two subtypes by comparing the pathological features of two focal H&E slides.

As two standard approaches for accurate diagnosis in the clinic, we explored the consistency between molecular testing and pathological testing. We confirmed that the pathology results could have 70.5% agreement with those of molecular testing. Based on these results, we propose a multi-modality way to differentiate the pattern of multifocal hepatocellular carcinoma using molecular or pathology testing in different clinical scenarios to provide guidelines for diagnosing and treating multifocal hepatocellular carcinoma.

Due to the scarcity of samples for multifocal hepatocellular carcinoma, only 34 samples were collected in this study, which may limit our construction of a more effective mathematical model for molecular subtyping. We may not achieve a significant outcome if the sample size is not large enough. Therefore, in this study, we mainly took a differential comparison to discover possible molecular biomarkers, and analyzed molecular and clinical features to explore how well the molecules testing is consistent with the pathology testing. Following this work, we are conducting a clinical study of multifocal HCC, yielding a more extensive data collection in the future. We will use advanced computational techniques such as artificial intelligence to optimize further the mathematical model of molecular typing (Tanaka et al., 2021), and some biological intelligent interpreters (Bu et al., 2021a) to generate multiple biomedical knowledge. Moreover, decision tools with multimodal combinations (Patel et al., 2021) could also be developed to optimize the diagnosis of multifocal HCC and thus guide the clinical treatment of liver cancer.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

## REFERENCES

- Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A., Behjati, S., Biankin, A. V., et al. (2013). Signatures of Mutational Processes in Human Cancer. *Nature* 500 (7463), 415–421. doi:10.1038/nature12477
- Alexandrov, L. B., Kim, J., Kim, J., Haradhvala, N. J., Huang, M. N., Tian Ng, A. W., et al. (2020). The Repertoire of Mutational Signatures in Human Cancer. *Nature* 578 (7793), 94–101. doi:10.1038/s41586-020-1943-3
- Ally, A., Balasundaram, M., and Carlsen, R. (2017). Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* 169 (7), 1327–1341.e23. doi:10.1016/j.cell.2017.05.046
- Baffy, G. (2015). Decoding Multifocal Hepatocellular Carcinoma: an Opportune Pursuit. *Hepatobiliary Surg. Nutr.* 4 (3), 206–210. doi:10.3978/j.issn.2304-3881.2014.12.05
- Bu, D., Luo, H., Huo, P., Wang, Z., Zhang, S., He, Z., et al. (2021a). KOBAS-i: Intelligent Prioritization and Exploratory Visualization of Biological Functions for Gene Enrichment Analysis. *Nucleic Acids Res.* 49 (W1), W317–w325. doi:10.1093/nar/gkab447

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of Peking University International Hospital (Beijing, China; approval no. 2020-059). The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

FR and DW are responsible for paper writing and data analysis. XZ assists in data analysis. LL and YZ are responsible for the paper's overall design, review, and revision.

## FUNDING

This research was supported by grants from the National Key Research and Development Program of China (2021YFF1201005), Beijing Xisike Clinical Oncology Research Foundation (grant no. Y-HS2019-33 and Y-BMS2019-015), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA16021400), and the Informatization Plan of Chinese Academy of Sciences (CAS-WX2021SF-0101).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.846517/full#supplementary-material>

**Supplementary Figures S1–S5** | The phylogenetic tree of the remaining patients except for patient Pt13 and Pt03.

**Supplementary Table S1** | The detailed clinical pathological information of all patients.

**Supplementary Table S2** | The detailed somatic mutation information of all samples.

- Bu, D., Xia, Y., Zhang, J., Cao, W., Huo, P., Wang, Z., et al. (2021b). FangNet: Mining Herb Hidden Knowledge from TCM Clinical Effective Formulas Using Structure Network Algorithm. *Comput. Struct. Biotechnol. J.* 19, 62–71. doi:10.1016/j.csbj.2020.11.036
- Cheng, J., Han, Z., Mehra, R., Shao, W., Cheng, M., Feng, Q., et al. (2020). Computational Analysis of Pathological Images Enables a Better Diagnosis of TFE3 Xp11.2 Translocation Renal Cell Carcinoma. *Nat. Commun.* 11 (1), 1778. doi:10.1038/s41467-020-15671-5
- Chidambaranathan-Reghupaty, S., Fisher, P. B., and Sarkar, D. (2021). Hepatocellular Carcinoma (HCC): Epidemiology, Etiology and Molecular Classification. *Adv. Cancer Res.* 149, 1–61. doi:10.1016/bs.acr.2020.10.001
- Cortés-Ciriano, I., Gulhan, D. C., Lee, J. J.-K., Melloni, G. E. M., and Park, P. J. (2022). Computational Analysis of Cancer Genome Sequencing Data. *Nat. Rev. Genet.* 23 (5), 298–314. doi:10.1038/s41576-021-00431-y
- Koh, G., Degasperis, A., Zou, X., Momen, S., and Nik-Zainal, S. (2021). Mutational Signatures: Emerging Concepts, Caveats and Clinical Applications. *Nat. Rev. Cancer* 21 (10), 619–637. doi:10.1038/s41568-021-00377-7



- Li, H., and Durbin, R. (2009). Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* 25 (14), 1754–1760. doi:10.1093/bioinformatics/btp324
- Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., et al. (2021). Hepatocellular Carcinoma. *Nat. Rev. Dis. Prim.* 7 (1), 6. doi:10.1038/s41572-020-00240-3
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data. *Genome Res.* 20 (9), 1297–1303. doi:10.1101/gr.107524.110
- Miao, R., Luo, H., Zhou, H., Li, G., Bu, D., Yang, X., et al. (2014). Identification of Prognostic Biomarkers in Hepatitis B Virus-Related Hepatocellular Carcinoma and Stratification by Integrative Multi-Omics Analysis. *J. Hepatol.* 61 (4), 840–849. doi:10.1016/j.jhep.2014.05.025
- Nanki, K., Fujii, M., Shimokawa, M., Matano, M., Nishikori, S., Date, S., et al. (2020). Somatic Inflammatory Gene Mutations in Human Ulcerative Colitis Epithelium. *Nature* 577 (7789), 254–259. doi:10.1038/s41586-019-1844-5
- Niu, B., Ye, K., Zhang, Q., Lu, C., Xie, M., McLellan, M. D., et al. (2014). MSIsensor: Microsatellite Instability Detection Using Paired Tumor-Normal Sequence Data. *Bioinformatics* 30 (7), 1015–1016. doi:10.1093/bioinformatics/btt755
- Patel, R. K., Fontana, M., and Ruberg, F. L. (2021). Cardiac Amyloidosis: Multimodal Imaging of Disease Activity and Response to Treatment. *Circ. Cardiovasc. Imaging* 14 (6), e009025. doi:10.1161/circimaging.121.009025
- Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S., and Swanton, C. (2016). DeconstructSigs: Delineating Mutational Processes in Single Tumors Distinguishes DNA Repair Deficiencies and Patterns of Carcinoma Evolution. *Genome Biol.* 17, 31. doi:10.1186/s13059-016-0893-4
- Schliep, K. P. (2011). Phangorn: Phylogenetic Analysis in R. *Bioinformatics* 27 (4), 592–593. doi:10.1093/bioinformatics/btq706
- Shen, R., and Seshan, V. E. (2016). FACETS: Allele-specific Copy Number and Clonal Heterogeneity Analysis Tool for High-Throughput DNA Sequencing. *Nucleic Acids Res.* 44 (16), e131. doi:10.1093/nar/gkw520
- Shimozaki, K., Hayashi, H., Tanishima, S., Horie, S., Chida, A., Tsgaru, K., et al. (2021). Concordance Analysis of Microsatellite Instability Status between Polymerase Chain Reaction Based Testing and Next Generation Sequencing for Solid Tumors. *Sci. Rep.* 11 (1), 20003. doi:10.1038/s41598-021-99364-z
- Study, E. A. (2012). EASL-EORTC Clinical Practice Guidelines: Management of Hepatocellular Carcinoma. *J. Hepatol.* 56 (4), 908–943. doi:10.1016/j.jhep.2011.12.001
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660
- Tanaka, I., Furukawa, T., and Morise, M. (2021). The Current Issues and Future Perspective of Artificial Intelligence for Developing New Treatment Strategy in Non-small Cell Lung Cancer: Harmonization of Molecular Cancer Biology and Artificial Intelligence. *Cancer Cell Int.* 21 (1), 454. doi:10.1186/s12935-021-02165-7
- Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., et al. (2019). COSMIC: the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* 47 (D1), D941–D947. doi:10.1093/nar/gky1015
- Viganò, L., Costa, G., and Di Tommaso, L. (2019). Liver Resection for Multifocal Hepatocellular Carcinoma: Is it an Option? *Hepatobiliary Surg. Nutr.* 8 (5), 530–533. doi:10.21037/hbsn.2019.05.12
- Vogel, A., Cervantes, A., Chau, I., Daniele, B., Llovet, J. M., Meyer, T., et al. (2018). Hepatocellular Carcinoma: ESMO Clinical Practice Guidelines for Diagnosis, Treatment and Follow-Up. *Ann. Oncol.* 29 (Suppl. 4), iv238–iv255. doi:10.1093/annonc/mdy308
- Wang, L., Ge, J., Lan, Y., Shi, Y., Luo, Y., Tan, Y., et al. (2020). Tumor Mutational Burden Is Associated with Poor Outcomes in Diffuse Glioma. *BMC Cancer* 20 (1), 213. doi:10.1186/s12885-020-6658-1
- Wu, H.-X., Wang, Z.-X., Zhao, Q., Chen, D.-L., He, M.-M., Yang, L.-P., et al. (2019). Tumor Mutational and Indel Burden: a Systematic Pan-Cancer Evaluation as Prognostic Biomarkers. *Ann. Transl. Med.* 7 (22), 640. doi:10.21037/atm.2019.10.116
- Yan, H. H. N., Siu, H. C., Law, S., Ho, S. L., Yue, S. S. K., Tsui, W. Y., et al. (2018). A Comprehensive Human Gastric Cancer Organoid Biobank Captures Tumor Subtype Heterogeneity and Enables Therapeutic Screening. *Cell Stem Cell* 23 (6), 882–897.e11. doi:10.1016/j.stem.2018.09.016

**Conflict of Interest:** XZ, NZ, and XW were employed by Zhijian Life Co. Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer DC declared a shared parent affiliation with the authors LL, YZ to the handling editor at the time of review.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Ren, Wang, Zhang, Zhao, Wang, Zhang and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.