Associations of Serum Tumor Biomarkers with Integrated Genomic and Clinical Characteristics of Hepatocellular Carcinoma

Keun Soo Ahn a Daniel R. O’Brien b Yong Hoon Kim a Tae-Seok Kim a Hiroyuki Yamada c Joong-Won Park d Sang-Jae Park e Seoung Hoon Kim e Cheng Zhang f Hu Li f Koo Jeong Kang a Lewis R. Roberts g

aDepartment of Surgery, Keimyung University Dongsan Hospital, Keimyung University School of Medicine, Daegu, Republic of Korea; bDivision of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA; cGlobal Clinical Research Management, FUJIFILM Wako Pure Chemical Corporation, Tokyo, Japan; dCenter for Liver and Pancreatobiliary Cancer, National Cancer Center, Goyang, Republic of Korea; eDepartment of Surgery, National Cancer Center, Goyang, Republic of Korea; fDepartment of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN, USA; gDivision of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN, USA

Keywords
α-Fetoprotein · Des-γ-carboxyprothrombin · Hepatocellular carcinoma · TP53 · CTNNB1

Abstract

Introduction: Serum α-fetoprotein (AFP), Lens culinaris agglutinin-reactive AFP (AFP-L3), and des-γ-carboxyprothrombin (DCP) are useful biomarkers of hepatocellular carcinoma (HCC). However, associations among molecular characteristics and serum biomarkers are unclear. We analyzed RNA expression and DNA variant data from The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) to examine their associations with serum biomarker levels and clinical data. Methods: From 371 TCGA-LIHC patients, we selected 91 seen at 3 institutions in Korea and the USA and measured AFP, AFP-L3, and DCP from preoperatively obtained serum. We conducted an integrative clinical and molecular analysis, focusing on biomarkers, and validated the findings with the remaining 280 patients in the TCGA-LIHC cohort. Results: Patients were categorized into 4 subgroups: elevated AFP or AFP-L3 alone (1AFP&L3), elevated DCP alone (1DCP), elevation of all 3 biomarkers (elevated levels of all 3 biomarkers [↑All]), and reference range values for all biomarkers (RR). CTNNB1 variants were frequently observed in ↑DCP patients (53.8%) and RR patients (38.5%), but ↑DCP patients with a CTNNB1 variant had worse survival than RR patients. TP53 sequence variants were associated with ↑AFP (30.8%) and ↑DCP (30.8%). The Wnt-β-catenin signaling pathway was activated in the ↑AFP&L3, whereas liver-related Wnt signaling was activated in the RR. TGF-β and VEGF signaling were activated in 1AFP&L3, whereas dysregulated bile acid and fatty acid metabolism were dominant in 1DCP. We validated these findings by showing similar results between the test cohort and the remainder of the TCGA-LIHC cohort. Conclusions: Serum AFP, AFP-L3, and DCP levels can help predict variants in the genetic profile of HCC, especially for TP53 and CTNNB1. These findings may facilitate development of an evidence-based approach to treatment.

© 2021 The Author(s).
Published by S. Karger AG, Basel

This is an Open Access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC) (http://www.karger.com/Services/OpenAccessLicense), applicable to the online version of the article only. Usage and distribution for commercial purposes requires written permission.

Correspondence to:
Koo Jeong Kang, kjkang@dsmc.or.kr
Lewis R. Roberts, roberts.lewis@mayo.edu
Introduction

Hepatocellular carcinoma (HCC) is the predominant form of liver cancer and the second most common cause of cancer death worldwide. Useful serum biomarkers for HCC include α-fetoprotein (AFP), Lens culinaris agglutinin-reactive α-fetoprotein (AFP-L3), and des-γ-carboxyprothrombin (DCP; also termed prothrombin induced by vitamin K absence-II [PIVKA-II]). Single-marker assessment has low sensitivity for patients with early-stage or even advanced HCC [1], but a combination of 2 or 3 tumor markers may improve sensitivity, and provide better specificity for the prediction of prognosis of HCC [1–4].

Because HCC is a highly heterogeneous tumor, the number of elevated biomarkers varies among tumors. About 20% of patients with HCC have elevated levels of more than 2 biomarkers, but 20–30% of patients have all tumor marker levels within the reference range [1]. Therefore, biomarker-specific clinicopathologic characteristics may exist. A few studies of clinicopathologic features have shown that high levels of DCP are associated with alcoholic liver disease [5] and more aggressive pathology [6, 7] and that high levels of AFP are associated with positive hepatitis B surface antigen. However, these studies did not comprehensively investigate the relationships among various parameters, and the associations were not validated in other studies.

Recent molecular studies of HCC genomic alterations have identified frequently mutated genes, including the telomerase reverse transcriptase (TERT) promoter, tumor protein 53 (TP53), and catenin β 1 (CTNNB1) [8]. Several multiplatform analyses of HCC have identified molecular subtypes and suggested potential therapeutic targets on the basis of molecular characteristics [9–12]. However, analysis of molecular characteristics in clinical settings is still not feasible because of the high cost. The association between molecular characteristics and serum biomarkers of HCC has not been evaluated to date. Serum biomarkers are easy to measure, so identification of any specific molecular features that can be predicted from biomarker levels may be useful for directing the approach to treatment. For this purpose, we obtained data from The Cancer Genome Atlas Hepatocellular Carcinoma (TCGA-LIHC) project, measured levels of AFP, DCP, and AFP-L3 from patients with serum samples available for analysis, and conducted an integrative analysis of clinical data (including the 3 biomarkers), RNA expression, and DNA sequence variants.

Materials and Methods

The study was approved by the institutional review board at each institution (Keimyung University Dongsan Hospital [2013-6-035], Korean National Cancer Center [NCCNCS13701], and Mayo Clinic [707-03]). Samples were collected for the respective biospecimen repositories after written patient consent was obtained.

Study Population and Genomic Data Collection

All data were obtained from TCGA-LIHC through their data portal (https://tcga-data.nci.nih.gov) and explored using the Firehose Browser (http://firebrowse.org/). Serum DCP and AFP-L3 data were not available in the TCGA-LIHC data set, and AFP data were missing for many individuals. Therefore, from the total sample of 377 patients, we selected 91 with preoperatively obtained, frozen serum samples available in 3 institutions for analysis. Benign, adjacent liver tissue (i.e., matched samples) was available for 8 patients (3 histologically normal, 1 with fibrosis, and 4 with cirrhosis). Patients were first seen at Keimyung University Dongsan Hospital (Daegu, Korea; n = 51), Mayo Clinic (Rochester, MN; n = 26), and Korean National Cancer Center (Goyang, Korea; n = 14) from January 2002 through December 2013.

Measurement of Serum AFP, DCP, and AFP-L3

Serum samples were stored at –80°C until they were tested for AFP, AFP-L3, and DCP. The assays were performed simultaneously by using a microchip capillary electrophoresis and liquid-phase binding assay on a μTAS Wako i30 automated analyzer (FUJIIFILM Wako Pure Chemical Corporation). With this instrument, if the total AFP level exceeded 0.6 ng/mL, the amount of AFP-L3 was reported as a percentage of the total AFP. The reportable range for AFP was 0.3–1,000 ng/mL; for AFP-L3, 0.5–99.5% of total AFP; and for DCP, 0.1–950 ng/mL. If biomarker levels were above the upper limit of detection, serum samples were serially diluted until the actual level of the biomarker could be determined. The interassay coefficient of variation was 0.7–1.5% for AFP; 0.3–5.6% for AFP-L3; and 1.3–7.9% for DCP.

Differentially Expressed Gene Analysis

Differentially expressed gene (DEG) analysis was performed with R statistical software (https://cran.r-project.org/) and the DESeq2 package. Before analysis, genes with a median of less than 30 raw reads in any group were excluded. After normalization of read counts, we identified DEGs among groups with an absolute log-fold change greater than 2.0 or less than −2.0 between 2 groups and a false discovery rate less than 0.05.

Statistical Analysis

Statistical analyses were performed with R statistical software. Groups were compared by using an independent t test for continuous variables and the χ² test for categorical variables. Correlation analysis for assessment of associations among the 3 serum biomarkers was done using the Pearson correlation coefficient. Overall and disease-free survival rates were calculated with the Kaplan-Meier method and plotted with Prism software (GraphPad Software, Inc).
Results

Patient Characteristics and Serum Biomarkers

The patient cohort consisted of 64 men and 27 women, who had a median age of 59 years (range, 23–85 years). Sixty-six patients (72.5%) were Asian, 23 (25.3%) were white or Hispanic, and 2 (2.2%) were black. The most common underlying liver diseases were hepatitis B (n = 55 [60.4%]) and alcoholic liver disease (n = 33 [36.2%]). During the median follow-up period of 47.1 months (range, 1.0–141.1 months), the 5-year disease-free survival rate was 47.2% and the 5-year overall survival rate was 77.8%. Clinical features are shown in online suppl. Table 1 (see www.karger.com/doi/10.1159/000516957 for all online suppl. material).

When we assessed associations among the 3 serum biomarkers, AFP was highly correlated with AFP-L3 (r = 0.70) (shown in Fig. 1a). However, correlations were lower between AFP and DCP (r = 0.37) and between AFP-L3 and DCP (r = 0.26) (shown in Fig. 1b, c).

We categorized patients as having high or low levels of each biomarker. Cutoff values for each category were selected after assessing the hazard ratios for disease-free survival and overall survival. For AFP, the cutoff value was 20 ng/mL; for DCP, 7.5 ng/mL (corresponding to 40 mAU/mL of PIVKA-II); for AFP-L3, 10%. Patients with high biomarker levels had significantly worse survival than patients with values below the cutoff thresholds (shown in online suppl. Fig. 1a–c). At least 1 biomarker was elevated for 65 patients (71.4%), and all 3 biomarkers were elevated for 21 patients (23.1%). However, all biomarkers were within the reference range for 26 patients (28.6%) (shown in Fig. 1d). The number of elevated biomarkers was associated with survival (shown in online suppl. Fig. 1d). When we excluded the 21 patients with elevated levels of all 3 biomarkers, we noted a high over-
lap of patients with high AFP and AFP-L3 (52.0% of AFP and 68.4% of AFP-L3), but a low overlap of patients with high DCP and AFP, or high DCP and AFP-L3 (less than 15% in each case) (shown in online suppl. Fig. 2).

**DEG Analysis, Stratified by Tumor Biomarkers**

We categorized patients into 3 subgroups by biomarker levels: “elevated AFP alone” (1AFP, n = 9), “elevated AFP-L3 alone” (1AFP-L3, n = 4), and “elevated DCP alone” (1DCP, n = 13). To identify tumor marker-specific genes, we searched for DEGs by comparing those 3 groups with the 8 matched adjacent benign liver samples. We identified 1,162 DEGs for the 3 tumor marker groups. Of the 438 genes that were upregulated or downregulated in 1AFP-L3 HCCs, 286 (65.3%) were altered in the same direction in 1AFP HCCs (shown in Fig. 1e).

Since many of the DEGs of 1AFP-L3 and 1AFP HCCs overlapped, pathway analysis was performed with gene set enrichment analysis (GSEA) [13] to highlight functional differences between those 2 groups. Tumors with 1AFP alone were associated with dysregulation of cell cycle and Notch signaling, while those with 1AFP-L3 alone were associated with angiogenesis and epithelial mesenchymal transition. However, 65% of the dysregulated pathways of the 2 subgroups overlapped with each other (shown in online suppl. Table 2). Therefore, the genetic features of patients with high AFP and high AFP-L3 are similar.

**Clinical Characteristics of Patients in Serum Biomarker Subgroups**

For the next analyses, patients with 1AFP alone, 1AFP-L3 alone, or increases in AFP and AFP-L3 but not DCP were combined into a single subgroup (1AFP&L3). We believed that this categorization was reasonable because AFP-L3 is an isoform of AFP and the serum levels of AFP and AFP-3 were correlated, resulting in a high overlap of patients with 1AFP and 1AFP-L3 (Fig. 1d; online suppl. Fig. 2); further, the genetic features of these patient groups were similar (Fig. 1e). We re-categorized patients into 4 biomarker subgroups; elevated AFP and/or AFP-L3 alone (1AFP&L3; n = 26), elevated DCP alone (1DCP; n = 13), elevation of all 3 biomarkers (1All; n = 21), and reference range values for all biomarkers (RR; n = 26). We excluded 3 patients with concurrent 1AFP and 1DCP and 2 patients with concurrent 1DCP and 1AFP-L3 from these subgroup analyses.

Patient characteristics are reported in Table 1. Mean tumor size was significantly larger in the 1All and 1DCP groups than the 1AFP&L3 and RR groups (p < 0.001). Patients in the RR group had the best prognosis, whereas patients in the 1All group had the worst. The 1AFP&L3 and 1DCP groups showed similar overall survival rates (shown in Fig. 1f).

**DNA Variant Profiles**

The genomic landscape of the full cohort (91 patients) is shown in Figure 2a. Nonsilent sequence variants were commonly observed in tumor suppressor genes (TP53 [28.6%], AXIN1 [11.0%], RB1 [6.6%]), Wnt pathway genes (CTNNB1 [28.6%]), and albumin-related genes (ALB [11.0%, APOB [9.9%]). Most sequence variants in TP53 and CTNNB1 were mutually exclusive, except in just 4 cases (4.4%).

TP53 variants were most frequently identified in the 1AFP&L3 group (n = 8 [30.8%]) and the 1All group (n = 8 [38.1%]). In the 1DCP group, CTNNB1 (n = 7 [53.8%]) and TP53 (n = 4 [30.8%]) variants were most common. In the RR group, the most common variants were in CTNNB1 (n = 10 [38.5%]) and ALB (n = 5 [19.2%]). Among the 23 samples for which TERT promoter sequence data were available, 10 (43.5%) showed sequence variants. The frequency of these TERT promoter variants was similar in all 4 biomarker subgroups.

In summary, TP53 variants were common in the 1AFP&L3 and 1DCP groups, and CTNNB1 variants were common in the 1DCP and RR groups. Interestingly, the overall survival rates were similar for patients with TP53 variants in the 1AFP&L3 and 1DCP groups (Fig. 2b, c), but patients with CTNNB1 variants in the 1DCP group showed significantly worse survival than those in the RR group (shown in Fig. 2d, e).

When we assessed variants associated with specific pathways, we noted that most pathway alterations were evenly distributed among the 4 biomarker subgroups. However, the prevalence of sequence variants differed

**Fig. 2.** Mutational profiles of 4 biomarker subgroups. **a** The plot was generated by the maftools package in R. (TERT promoter variant data were available for 23 cases in the current cohort.) **b** Domain hot spots for TP53 variants in the 1AFP&L3 and 1DCP groups. **c** Survival analysis of TP53 variants in the 1AFP&L3 and 1DCP groups. **d** Domain hot spots for CTNNB1 variants in the 1DCP and RR groups. **e** Survival analysis of CTNNB1 variants in the 1DCP and RR groups. Abbreviations are defined in the legend to Figure 1.

(For figure see next page.)
Molecular Features of HCC according to Serum Biomarkers

Liver Cancer 2021;10:593–605
DOI: 10.1159/000516957

TP53 variant, \( p = 0.39 \)

CTNNB1 variant, \( p = 0.02 \)

Overall survival, %

Time after surgery, mo

TP53

CTNNB1

0 100 200 300 400
Amino acid position

0 5 10 15 20 25
Missense mutation
Multi hit
Nonsense mutation
Splice site
TERT promoter mutation
Frame shift del
In frame del

0 200 400 600 800
Amino acid position

0 100 200 300 400
Amino acid position

0 20 40 60 80 100 120 140
Overall survival, %

0 20 40 60 80 100 120 140
Overall survival, %

0 20 40 60 80 100 120 140
Overall survival, %
for cell cycle progression genes (TP53, RB1) and the Wnt pathway gene CTNNB1 (shown in online suppl. Fig. 3).

**RNA Expression Profiling and Pathway Analysis, Stratified by Biomarker Subgroups**

The 358 upregulated and downregulated DEGs among the 4 biomarker subgroups are stratified by groups and summarized in online suppl. Table 3. EPCAM (a marker of stem cells) and IGF2 (a suppressor of interferon target genes) and the classical Wnt target genes MMP7 and RUNX2 were upregulated in the ↑AFP&L3 group. In the RR group, liver-related Wnt target genes GLUL, REG3A, TBX3, SLClA2, EPHB2, and SPARCL1 [14, 15] were highly expressed.

To classify the DEGs using GSEA [13] and the Database for Annotation, Visualization and Integrated Discovery (DAVID) [16], we focused on several liver-specific oncogenic pathways [17, 18] and a pan-cancer analysis [19] to identify genes that were differentially enriched in each biomarker subgroup (false discovery rate <0.05). In the ↑AFP&L3 group, PI3K-AKT-mTOR, NOTCH, VEGF, and TGF-β signaling pathways were activated, whereas oxidative phosphorylation pathways were suppressed (shown in Fig. 3a).

The Wnt-β-catenin signaling pathway was activated in the RR group compared with the ↑AFP&L3 and ↑All groups. d p53 pathways were deactivated in the ↑DCP group. b–d Generated by gene set enrichment analysis. Abbreviations are defined in the legend to Figure 1.AFP, α-fetoprotein.
Molecular Features of HCC according to Serum Biomarkers

functions were different between the groups (shown in Fig. 3d). Cell cycle dysregulation was predominant in the ↑AFP&L3 group (shown in online suppl. Fig. 4a), whereas dysregulation of RNA editing and DNA repair was evident in the ↑DCP group (shown in online suppl. Fig. 4b).

In the ↑DCP and RR groups, bile acid and fatty acid metabolism was dysregulated. The ↑All group showed mixed features of the other 3 groups (shown in Fig. 3a; online suppl. Table 4).

Validation with the Full TCGA-LIHC Cohort

The full TCGA-LIHC cohort included 371 cases, of which 91 were used for the clinical and molecular analysis of biomarkers described above. We next sought to validate our findings by applying the 358 DEGs associated with the 4 subgroups to data from the remaining 280 patients in the TCGA-LIHC cohort. Unsupervised cluster analysis identified 4 distinct subgroups (shown in Fig. 4a). Subgroup A was defined as the RR-like subgroup; B, the ↑DCP-like subgroup; C, the ↑All-like subgroup; and D, the ↑AFP&L3-like subgroup.

Patient characteristics and molecular features of the new patient groups resembled those of our previously identified categories. Hepatitis B was frequent in subgroups C and D (↑All-like and ↑AFP&L3-like), and non-alcoholic fatty liver disease was observed only in subgroup B (↑DCP-like). The level of AFP was significantly higher in subgroups C and D (↑All-like and ↑AFP&L3-like) than subgroups A and B (RR-like and ↑DCP-like) (p < 0.001) (shown in Fig. 4b). However, we could not determine differences in DCP and AFP-L3 levels among the 4 subgroups because those data were not available in TCGA-LIHC. Although the prevalence of TP53 variants was significantly lower in the A (RR-like) subgroup than the others, the prevalence of CTNNB1 variants was higher in the A (RR-like) and B (↑DCP-like) subgroups (shown
in Fig. 4b). Although the survival outcome was not significantly different among the 4 groups, the subgroup analysis showed that subgroup A (RR-like) had significantly better overall survival than subgroup C (TAll-like) (p = 0.03) (shown in Fig. 4c). Similar to the 91 HCC test cohort, survival of patients with a CTNNB1 variant was significantly better in subgroup A (RR-like) than in subgroup B (1DCP-like) (p = 0.03) (shown in Fig. 4d).

These validated data using the remainder of the TCGA-LIHC cohort showed similar results to the test cohort; CTNNB1 variants were associated with low AFP, and HCCs with CTNNB1 variants and high serum DCP were associated with poorer survival than those with CTNNB1 variants and RR biomarkers. In addition, patients with RR biomarkers had a low frequency of TP53 variants.

**Comparison with Previously Identified Molecular Subclasses**

In a TCGA-LIHC study of 183 patients, HCC data were analyzed with multiplatform integrative molecular subtyping, and 3 subclasses (iClust subclasses 1, 2, and 3) were identified [17]. We compared the clinical and molecular characteristics of the iClust subclasses and the subgroups identified in the present study (shown in Fig. 5a). iClust 1 is characterized by high serum AFP, overexpression of proliferation genes (MYBL2, PLK1, and MKI67), and a low frequency of CTNNB1 variants. Those clinical and molecular characteristics are similar to those of the 1AFP&L3 group of the present study, and therefore, iClust 1 may correspond to the 1AFP&L3 group. In iClust 2 and 3, the frequency of TP53 and CTNNB1 variants is high. However, serum AFP is lower in iClust 2 than iClust 3. We propose that iClust 2 includes features of the 1DCP (low AFP, frequent TP53 and CTNNB1 variants) and RR (low AFP, frequent CTNNB1 variants). As we expected, in the present cohort (23 patients with available iClust data) and the validating TCGA-LIHC cohort (160 with available iClust data), iClust 1 consisted predominantly of patients in the 1AFP&L3 group, and iClust 2 included patients in the 1DCP and RR groups predominantly.

**Fig. 5.** Comparison with other molecular subclasses. a Comparison with iClust in TCGA-LIHC. b Comparison with the Hoshida liver cancer classification by gene set enrichment analysis (GSEA). c, d According to GSEA, 1AFP&L3 and TAll groups had molecular similarity with S2 (e), while the 1DCP and RR groups showed molecular characteristics similar to those of S3 (d). e Unsupervised clustering using Hoshida genets in current cohort identified 3 subgroups that have similar gene expression with Hoshida subclasses. f Distribution of 4 subgroups of current cohort and level of serum biomarkers in S1-, S2-, and S3-like subclasses. Abbreviations are defined in the legend to Figure 1.
iClust 3 had heterogeneous characteristics (shown in Fig. 5a).

We also compared our biomarker subgroups with subclasses identified using transcriptome meta-analysis by Hoshida and collaborators [10, 11, 14, 20]. Patients in the ↑AFP&L3 group had specific features associated with their S1 group (enriched TGF-β and Wnt pathways, high expression of BIRC5, moderate elevation of serum AFP) and S2 group (enriched MYK and AKT pathway, high expression of EPCAM, IGF2, AFP, and GPC3, high serum AFP). Meanwhile, the RR group resembled the S3 group, with similar clinical features (low serum AFP, good survival) and molecular features (high expression of liver-related Wnt target genes, CTNNB1 variants). As our expectation, according to GSEA, gene expression patterns of the ↑AFP&L3 group showed high molecular similarity with S1 and S2 (shown in Fig. 5b, e), and RR group was similar in molecular features with S3 (shown in Fig. 5b, e). The ↑All group was also similar in molecular features with S1 and S2, but it was associated with a higher risk of late recurrence and poorer survival than the ↑AFP&L3 group (shown in Fig. 5b). As with the RR group, the ↑DCP group also showed similar gene expression patterns with the S3 group, but it was associated with a higher risk of late recurrence and better survival than the RR group (shown in Fig. 5b, e). For validation, we applied the Hoshida genesets from the Molecular Signatures Database v. 7.2 of GSEA to data from the current cohort. Unsupervised cluster analysis identified 3 distinct subgroups and compared the serum biomarker levels (shown in Fig. 5e, f). The ↑AFP&L3 group was most closely associated with the S1-like subclass, followed by the S2-like subclass. Most of the RR and ↑DCP groups were associated with the S3-like subclass. Interestingly, the ↑All group was most highly associated with the S2-like subclass (shown in Fig. 5f). In the S2-like subclass, the values of the AFP, AFP-L3, and DCP biomarkers were all more highly elevated than in the other subclasses. The values of the AFP and AFP-L3 in the S1-like subclass were higher than those in S3-like subclass but lower than for the S2-like subclass. In the present study, the serum AFP and AFP-L3 levels were higher in the ↑AFP&L3 group than in the ↑DCP and RR subgroups, but lower than in the ↑All group (shown in Table 1). Considering these molecular characteristics and patterns of serum biomarkers, overall, it appears that Hoshida subclass S2 has overlapping molecular biological characteristics with the ↑All group as well as with the ↑AFP&L3 group, while the S1 subclass shares characteristics with the ↑AFP&L3 group. The S3 subclass shares characteristics predominantly with the RR and ↑DCP groups, but these 2 subgroups have different risks of recurrence and survival (shown in Fig. 5, 6).

### Discussion/Conclusion

High levels of serum biomarkers are usually associated with advanced and poorly differentiated HCC [21], but exceptions have been observed of aggressive HCC with normal biomarker levels [22] and early-stage HCC with highly elevated biomarkers [23]. Therefore, serum bio-

| Characteristics | ↑AFP and/or AFP-L3 | ↑DCP | ↑AFP and/or AFP-L3 + DCP | RR |
|-----------------|--------------------|------|--------------------------|----|
| Variant         | TPS3               | TPS3 | TPS3, CTNNB1 (poor prognosis) | Heterogenous, CTNNB1 (good prognosis) |
| Molecular pathways | WNT, TGF-beta, Notch, VEGF, PI3K-AKT-mTOR, MYC, KRAS | Metabolism, P53 | MYC, PI3K-AKT-mTOR, VEGF, Metabolism | Liver-related WNT, Metabolism |
| Clinical phenotypes | Hepatitis B, Intermediate survival | Fatty liver, Alcohol, Intermediate survival | Mixed risk factors, Poorly differentiated, Large tumor, Poor survival | Fatty liver, Alcohol, Well differentiated, Good survival |
| Published subclass (TCGA-LIHC, hoshida) | iClust 1 S1, S2 (low risk) | iClust 2 S3 (high risk) | S2 (high risk) | iClust 2 S3 (low risk) |

Fig. 6. Summary of clinical and molecular characteristics of biomarker subgroups. Abbreviations are defined in the legend to Figure 1.
markers are thought to be associated with additional factors other than solely clinical features. The present study is novel because of the newly identified associations of serum biomarkers with genetic variants and gene expression profiles in HCC. Characteristics of the 3 biomarkers were delineated by the mutually exclusive TP53 and CTNNB1 variants, which are common in HCC, followed by TERT promoter variants [8, 17, 24, 25]. Our data strongly suggest that tumors bearing TP53 and CTNNB1 variants represent clinically distinct subtypes of HCC, and serum biomarkers might be helpful for predicting the variant profile of HCC (Fig. 6).

The tumor marker AFP is a plasma glycoprotein encoded by the AFP gene. AFP expression is regulated by a complex network of transcriptional regulators, including oncoproteins such as MYC family genes and TP53 [26]. A TP53 loss-of-function sequence variant activates AFP expression [27]. In addition, the hepatitis B virus X protein induces AFP receptor expression and may activate PI3K-AKT-mTOR signaling [28, 29]. This mechanism would explain our finding that HCCs with ↑AFP&L3 were associated with hepatitis B etiology, TP53 variants, and activation of oncogenic signals such as MYC and PI3K-AKT-mTOR. DCP production results from an acquired defect in the posttranslational carboxylation at glutamic acid (Glu, E) residues within the N-terminal domain of the prothrombin precursor in HCC cells. Meanwhile, intracellular fatty acids are necessary for appropriate post-

Table 1. Clinical and genetic characteristics, stratified by biomarker subgroup (n = 86)

| Characteristic                              | RR (n = 26) | ↑AFP&L3 (n = 26) | ↑DCP (n = 13) | ↑All (n = 21) | p value |
|---------------------------------------------|-------------|-----------------|--------------|--------------|---------|
| Sex, n (%)                                  |             |                 |              |              |         |
| Female                                      | 6 (23.1)    | 7 (26.9)        | 4 (30.8)     | 9 (42.9)     | 0.50    |
| Male                                        | 20 (76.9)   | 19 (73.1)       | 9 (69.2)     | 12 (57.1)    |         |
| Age, mean (SD), yr                          | 59.2 (11.5) | 53.2 (11.5)     | 62.7 (10.1)  | 57.8 (14.9)  | 0.12    |
| Body mass index, mean (SD), kg/m²           | 26.3 (5.6)  | 24.8 (2.9)      | 25.7 (5.9)   | 28.8 (4.2)   | 0.74    |
| AFP, ng/mL, log, mean (SD)                  | 1.5 (0.3)   | 3.4 (1.2)       | 1.5 (0.3)    | 4.2 (1.4)    | <0.001  |
| DCP, ng/mL, log, mean (SD)                  | 0.9 (0.6)   | 0.9 (0.5)       | 2.5 (0.5)    | 3.0 (0.7)    | <0.001  |
| AFP-L3 (%), log, mean (SD)                  | 1.0 (0.5)   | 2.1 (0.5)       | 1.0 (0.5)    | 2.6 (0.3)    | <0.001  |
| BCLC stage, n (%)                           |             |                 |              |              |         |
| A                                           | 20 (76.9)   | 24 (92.3)       | 11 (84.6)    | 16 (76.2)    | 0.10    |
| B or C                                      | 6 (23.1)    | 2 (7.7)         | 2 (15.4)     | 5 (23.8)     |         |
| Cirrhosis, n (%)                            | 18 (69.2)   | 16 (61.5)       | 8 (61.5)     | 12 (57.1)    | 0.86    |
| Risk factors, n (%)                         |             |                 |              |              |         |
| Hepatitis B                                 | 14 (53.8)   | 20 (76.9)       | 6 (46.2)     | 14 (66.7)    | 0.30    |
| Hepatitis C                                 | 2 (7.7)     | 3 (11.5)        | 3 (23.1)     | 1 (4.8)      | 0.37    |
| Alcohol                                     | 9 (34.6)    | 6 (23.1)        | 5 (38.5)     | 10 (47.6)    | 0.37    |
| Fatty liver (NAFLD)                         | 3 (11.5)    | 0 (0)           | 1 (7.7)      | 0 (0)        | 0.15    |
| Tumors identified, n, patients, n (%)       |             |                 |              |              |         |
| 1                                           | 24 (92.3)   | 25 (96.2)       | 10 (76.9)    | 18 (85.7)    | 0.27    |
| >1                                          | 2 (7.7)     | 1 (3.8)         | 3 (23.1)     | 3 (14.3)     |         |
| Tumor size, mean (SD), cm                   | 4.9 (2.8)   | 4.0 (2.3)       | 7.0 (5.1)    | 8.7 (3.9)    | <0.001  |
| Vascular invasion, n (%)                    | 5 (19.2)    | 7 (26.9)        | 3 (23.1)     | 7 (33.3)     | 0.25    |
| Histologic grade, n (%)                     |             |                 |              |              |         |
| G1 or G2                                    | 18 (69.2)   | 10 (38.5)       | 7 (53.8)     | 4 (19.0)     | 0.01    |
| G3 or G4                                    | 8 (30.8)    | 16 (61.5)       | 6 (46.2)     | 17 (81.0)    |         |
| AJCC stage, n (%)                           |             |                 |              |              |         |
| I                                           | 19 (73.1)   | 19 (73.1)       | 8 (61.5)     | 15 (71.4)    |         |
| II                                          | 6 (23.1)    | 5 (19.2)        | 2 (15.4)     | 2 (9.5)      | 0.43    |
| III & IV                                    | 1 (3.8)     | 2 (7.7)         | 3 (23.1)     | 3 (14.3)     |         |
| TP53 variant, n (%)                         | 5 (19.2)    | 8 (30.8)        | 4 (30.8)     | 8 (38.1)     | 0.51    |
| CTNNB1 variant, n (%)                       | 10 (38.5)   | 2 (7.7)         | 7 (53.8)     | 5 (23.8)     | 0.01    |

↑, elevated biomarker level; AFP, α-fetoprotein; AFP-L3, Lens culinaris agglutinin-reactive α-fetoprotein; AFP&L3, combined group of AFP and AFP-L3; AJCC, American Joint Committee on Cancer; All, all 3 biomarkers; BCLC, Barcelona Clinic Liver Cancer; DCP, des-γ-carboxyprothrombin; NAFLD, nonalcoholic fatty liver disease; RR, reference range values for all biomarkers.
Molecular Features of HCC according to Serum Biomarkers

translational protein modifications, but dysregulated fatty acid metabolism results in abnormally increased serum levels of the prothrombin protein DCP [30]. In cancer, lipid metabolism is regulated by p53, and loss of p53 activity is associated with increased cholesterol and fatty acid synthesis, which accelerates growth of cancer cells [31]. Therefore, DCP-related HCC development differs from that associated with elevated AFP or AFP-L3, and the role of p53 may differ in AFP- versus DCP-related HCC. In the 1DCP and RR groups, several metabolic pathways were activated and CTNNB1 variants were common. CTNNB1 variants are associated with dysfunctional glutamine and cholesterol metabolism [32], which explains why CTNNB1 variants and nonalcoholic fatty liver disease are associated with 1DCP and RR HCC. Early-stage HCC without activated oncogenic signaling appeared to be associated with the RR group, whereas more advanced disease showed associations with the defective carboxylation and dysregulated oncogenic pathways of the 1DCP group. The 1All group showed advanced clinical features and heterogeneous molecular features. The poor survival outcome of the 1All subgroup may be associated with increased intratumoral heterogeneity of tumors producing AFP, AFP-L3 and DCP, with associated more aggressive phenotypes. Characteristics of the 1All group may vary with disease progression and the dominant serum biomarker.

In the present study, CTNNB1 variants were frequently observed in patients in the RR and 1DCP groups. Clinical implications of CTNNB1 missense variants, in terms of tumor aggressiveness and patient survival, are not yet clear. Although CTNNB1 variants have been observed in patients with well-differentiated tumors and a good prognosis [10], other studies have reported that CTNNB1 variants are associated with high rates of vascular invasion and a poor prognosis [33, 34]. Notably, our study showed that survival of patients in the 1DCP group was worse than that of patients in the RR group. CTNNB1-variant HCCs may have 2 different subtypes with different clinical outcomes, and the biomarker status may indicate the contradictory clinical phenotypes. In patients with CTNNB1 mutation alone with normal biomarkers, dysregulation of fatty acid was also noted. We hypothesize that patients with CTNNB1 mutation with DCP elevation may reflect more advanced HCC with more severe defects in carboxylation and dysregulation of fatty acid pathways than HCC patients with CTNNB1 mutation alone without DCP elevation. However, a direct molecular mechanism explaining the association between CTNNB1 variants and the serum DCP in the 1DCP and RR groups is not yet elucidated, and it is not definitively established whether CTNNB1 variants are critical factors regulating DCP elevation. Nevertheless, the serum DCP may be a useful clinical parameter for assessing the prognosis of a patient with a CTNNB1 variant.

The present study showed that elevated AFP (or AFP-L3) and DCP are associated with TP53 variants. Pathway analysis showed different phenotypes of TP53 variants, including dysregulation of cell cycle function in the 1AFP&L3 group and diminished DNA repair function in the 1DCP group (shown in online suppl. Fig. 4). Although the survival rates are similar for TP53-variant HCCs between the 2 groups and the clinical implications of these sequence variants are unclear, the differences in phenotype are likely to be linked to serum biomarker levels.

Previous research showed that 2 different Wnt pathways are involved in HCC [14]. In 1 pathway, TGF-β-related Wnt signaling in the absence of a CTNNB1 variant has clinical features of aggressive disease and poor survival. In the other pathway, CTNNB1 variant-related Wnt signaling is associated with activation of liver-related Wnt signaling target genes such as GLUL and LGR5 and a good prognosis [11, 14]. In the present study, TGF-β-related Wnt signaling [35] without a CTNNB1 variant was the main pathway in the 1AFP&L3 group, whereas CTNNB1 variant-related Wnt signaling was activated in the RR group. These findings were consistent with those of previous studies showing that CTNNB1 variants do not always regulate canonical Wnt target genes [10, 14]. Our results suggest that these different Wnt signaling pathways are associated with serum biomarkers: 1AFP&L3 for TGF-β-related Wnt signaling and RR biomarkers for CTNNB1 variant-related Wnt signaling.

Although the present study is novel because it showed the relevance of serum biomarkers for different genetic variants of HCC, we acknowledge a few drawbacks. First, the sample size of 91 patients might have been insufficient for identifying specific genetic features associated with various combinations of biomarkers. Although we validated our findings by using the whole TCGA-LIHC cohort, we could not identify other public genetic data sets for validation because there are no data sets with available DCP and AFP-L3 values. As another limitation, although we revealed that elevated AFP was associated with TP53 variants, while elevated DCP was associated with TP53 and CTNNB1 variants, not all patients with these mutations had the corresponding associated biomarker elevations. Since not all patients with elevated serum biomarkers have the corresponding associated mutations, other molecular factors may contribute to elevations of the se-
Serum biomarkers. Nevertheless, although our findings cannot fully explain the molecular mechanisms of the serum biomarker elevations, our study demonstrated that not only genetic mutation, but also particular dysregulated oncogenic pathways were associated with specific serum biomarkers. Therefore, different mechanisms of dysregulation of oncogenic pathways other than mutations may contribute to specific biomarker elevations. For example, in the TCGA HCC marker paper, it was shown that a number of HCCs without TP53 mutations nevertheless had gene expression patterns consistent with inactive p53, suggesting the existence of nonmutational p53 inactivating mechanisms. MDM4, a p53 inhibitory protein, was significantly increased in copy number and expression in low signature wild-type TP53 HCC patients relative to other HCC patients, providing a possible mechanism for low p53 signatures in non-TP53 mutated HCC patients [17]. We believe that our findings partially unmask the molecular relevance of the serum biomarker elevations, and will facilitate further studies examining the molecular mechanisms associated with serum biomarker elevations in the future. For example, high expression of the AFP gene is related to de-methylation of the AFP promoter, and in the present study, promoter methylation of AFP was negatively correlated with AFP gene expression (data was not shown) [36]. Although it was not proven in the present study (data not shown), deregulated intracellular miRNA 122 is associated with activation of AFP protein in tissue [37]. The present study has the limitation that it does not reveal all genetic factors that contribute to elevation of the serum biomarkers. Therefore, further clinical and molecular studies are needed to fully elucidate the mechanistic basis for the biomarker elevations. Nevertheless, use of serum biomarkers as a diagnostic modality is cheap, feasible, and safe; therefore, when considered in conjunction with integrative clinical and pathologic features, serum biomarkers may have a role in predicting genetic characteristics and guiding the approach to targeted therapy.

In conclusion, serum AFP, AFP-L3, and DCP may be helpful for predicting the genetic profile of HCC. CTNNB1 variants were associated with low AFP and AFP-L3 levels, and the prognosis was good for those with RR biomarkers, whereas it was poorer for patients with 1DCP. TP53 variants were associated with 1AFP or 1DCP and showed different activating pathways, depending on the biomarker. These clinically oriented findings may facilitate development of an evidence-based approach to treatment.

Acknowledgment

The biospecimens and data used in this study were provided by the Biobanks of Keimyung University Dongsan Medical Center, member of the Korea Biobank Network, the Korean National Cancer Center, and the Mayo Clinic SPORE in Hepatobiliary Cancer.

Statement of Ethics

The study was approved by the Institutional Review Board at each institution (Keimyung University Dongsan Hospital [2013-6-035], Korean National Cancer Center [NCCNCS13701], and Mayo Clinic [707-03]). Samples were collected for the respective biospecimen repositories after written patient informed consent was obtained. This research complied with the guidelines for human studies and was conducted ethically in accordance with the World Medical Association Declaration of Helsinki.

Conflicts of Interest Statement

L.R.R. is an editorial board member of Liver Cancer. L.R.R. has received grant funding from Bayer, BTG International, Exact Sciences, Gilead Sciences, GlycoTest, Redhill, TARGET PharmaSolutions, and FUJIFILM Medical Systems and has served in consulting roles for AstraZeneca, Bayer, Exact Sciences, Gilead Sciences, GRAIL, QED Therapeutics, and TAVEC. J.-W. Park has served in a consulting or advisory role for AstraZeneca, Bayer, BMS, Roche, Ipsen and Eisai and has served on a speakers’ bureau for Bayer. The other authors have no conflict of interest to declare.

Funding Sources

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MSIP) (No. 2018R1C1B3004435) to K.S. Ahn, and by FUJIFILM Wako Pure Chemical Corporation, the Mayo Clinic Center for Cell Signaling in Gastroenterology (NIDDK P30 DK084567), the Mayo Clinic Center for Clinical and Translational Science (UL1TR000135), and the Mayo Clinic Hepatobiliary SPORE (P50 CA210964) to L.R.R. The sponsors had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Author Contributions

Conceptualization: K.S. Ahn, K.J. Kang, L.R. Roberts. Data curation: D.R. O’Brien and H. Yamada. Formal analysis: K.S. Ahn, C. Zhang, and H. Li. Investigation: K.S. Ahn and Y.H. Kim. Methodology: T.-S. Kim and D.R. O’Brien. Software: D.R. O’Brien, C. Zhang, and H. Li. Validation: J.W. Park, S.J. Park, and S.H. Kim. Supervision: L.R. Roberts and K.J. Kang. Visualization: K.S. Ahn. Writing – review and editing: K.S. Ahn and L.R. Roberts.
Molecular Features of HCC according to Serum Biomarkers

References

1 Toyoda H, Kumada T, Tada T, Niinomi T, Ito T, Kaneoka Y, et al. Prognostic significance of a combination of pre- and post-treatment tumor markers for hepatocellular carcinoma curatively treated with hepatectomy. J Hepatol. 2012;57(6):1251–7.

2 Chaiteerakij R, Zhang X, Addissie BD, Mohamed EA, Harmesen WS, Theobald PJ, et al. Combinations of biomarkers and Milan criteria for predicting hepatocellular carcinoma recurrence after liver transplantation. Liver Transpl. 2015;21(5):599–606.

3 Toyoda H, Kumada T, Osaki Y, Oka H, Urano F, Kudo M, et al. Staging hepatocellular carcinoma by a novel scoring system (BALAD score) based on serum markers. Clin Gastroenterol Hepatol. 2006;4(12):1528–36.

4 Choi JY, Jung SW, Kim HY, Kim M, Kim Y, Kim DG, et al. Diagnostic value of AFP-L3 and PIYK-A-II in hepatocellular carcinoma according to total-AFP. World J Gastroenterol. 2013;19(39):339–46.

5 Kang KH, Kim JH, Kang SH, Lee BJ, Seo YS, Yim HJ, et al. The influence of alcoholic liver disease on serum PIYK-A-II levels in patients without hepatocellular carcinoma. Gut Liver. 2015;9(2):224–30.

6 Okuda H, Nakashim T, Takatsu K, Saito A, Hayashi N, Yamamoto M, et al. Comparison of clinicopathological features of patients with hepatocellular carcinoma seropositive for alpha-fetoprotein alone and those seropositive for des-gamma-carboxyprothrombin alone. J Gastroenterol Hepatol. 2001;16(11):1290–6.

7 Kurokawa T, Yamazaki S, Mitsuka Y, Moriguchi M, Sugitani M, Takayama T. Prediction of vascular invasion in hepatocellular carcinoma by next-generation des-r-carboxy prothrombin. Br J Cancer. 2016;114(11):53–9.

8 Lee JS. The mutational landscape of hepatocellular carcinoma. Clin Mol Hepatol. 2015;21(3):220–9.

9 Lee JS, Chiu 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(3):667–76.

10 Hoshida Y, Nijman SM, Kobayashi M, Chan JA, Brunet JP, Chiang DY, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. Cancer Res. 2009;69(18):7385–92.

11 Goossens N, Sun X, Hoshida Y. Molecular classification of hepatocellular carcinoma: potential therapeutic implications. Hepat Oncol. 2015;2(4):371–9.

12 Calderaro J, Couchy G, Imbeaud S, Amaddeo G, Letouze E, Blanc JF, et al. Histological subtypes of hepatocellular carcinoma are related to gene mutations and molecular tumour classification. J Hepatol. 2017;67(4):727–38.

13 Lachenmayer A, Alsinet C, Savic R, Cabellos L, Toffanin S, Hoshida Y, et al. Wnt-pathway activation in two molecular classes of hepatocellular carcinoma and experimental modulation by sorafenib. Clin Cancer Res. 2012;18(18):4997–5007.

14 Zucman-Rossi J, Benhamiche S, Godard C, Boyault S, Grimer B, Balabaud C, et al. Differential effects of inactivated Axin1 and activated beta-catenin mutations in human hepatocellular carcinomas. Oncogene. 2007;26(5):774–80.

15 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50.

16 Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, et al. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. 2007;35(Web Server issue):W169–75.

17 Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. Cell. 2017;169(7):1327–41.e23.

18 Schulze K, Imbeaud S, Letouze E, Alexandrov LB, Calderaro J, Rebourssou S, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat Genet. 2015;47(5):505–11.

19 Sanchez-Vega F, Mina M, Armenia J, Chatila W, Shimizu K, et al. Oncogenic signaling pathways in the cancer genome atlas. Cell. 2017;2018;173(2):321–37.e10.

20 Tan PS, Nakagawa S, Goossens N, Venkatesh LB, Calderaro J, Rebouissou S, et al. Exome sequencing of hepatocellular carcinomas with distinct molecular subclasses reveals recurrent FRRK-activating mutations and the mechanisms of malignant transformation. Cancer Cell. 2014;25(4):428–41.

21 Sanchez-Vega F, Mina M, Armenia J, Chatila W, Shimizu K, et al. Oncogenic signaling pathways in the cancer genome atlas. Cell. 2014;2014;25(4):428–41.

22 Colombo M. Screening for cancer in viral hepatitis. Clin Liver Dis. 2001;5(1):109–22.

23 Dai JS, Zhang C, Chen P, Jin SJ, Jiang GQ. The prognostic correlation of AFP level at diagnosis with pathological grade, progression, and survival of patients with hepatocellular carcinoma. Sci Rep. 2017;7(1):12870.

24 Pilati C, Letouze E, Nault JC, Imbeaud S, Boulaïi A, Calderaro J, et al. Genomic profiling of hepatocellular adenomas reveals recurrent FRRK-activating mutations and the mechanisms of malignant transformation. Cancer Cell. 2014;25(4):428–41.