

Molecular genomic landscapes of hepatobiliary cancer

Tatsuhiko Shibata^{1,2}  | Yasuhito Arai² | Yasushi Totoki²

¹Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan

²Division of Cancer Genomics, National Cancer Center Research Institute, Tokyo, Japan

Correspondence

Tatsuhiko Shibata, Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan.

Email: tshibata@ims.u-tokyo.ac.jp

Funding information

National Cancer Center Research and Development Funds, Grant/Award Number: 29-A-6; Japan Agency for Medical Research and Development

Hepatocellular carcinoma (HCC) and biliary tract cancer are more frequent in East Asia including Japan than in Europe or North America. A compilation of 1340 multi-ethnic HCC genomes, the largest cohort ever reported, identified a comprehensive landscape of HCC driver genes, comprised of three core drivers (*TP53*, *TERT*, and *WNT* signaling) and combinations of infrequent alterations in various cancer pathways. In contrast, five core driver genes (*TP53*, *ARID1A*, *KRAS*, *SMAD4*, and *BAP1*) with characteristic molecular alterations including fusion transcripts involving fibroblast growth factor receptor 2 and the protein kinase A pathway, and *IDH1/2* mutation constituted the biliary tract cancer genomes. Consistent with their heterogeneous epidemiological backgrounds, mutational signatures and combinations of non-core driver genes within these cancer genomes were found to be complex. Integrative analyses of multi-omics data identified molecular classifications of these tumors that are associated with clinical outcome and enrichments of potential therapeutic targets, including immune checkpoint molecules. Translating comprehensive molecular-genomic analysis together with further basic research and international collaborations are highly anticipated for developing precise and better treatments, diagnosis, and prevention of these tumor types.

KEYWORDS

biliary tract cancer, driver gene, hepatocellular carcinoma, molecular classification, mutational signature

1 | GEOGRAPHICAL AND EPIDEMIOLOGICAL DIVERSITIES OF HEPATOBILIARY CANCERS

Hepatocellular and biliary tract cancers are both dismal tumors. The frequencies of these tumors are geographically more common in

East Asia including Japan and BTC is especially more frequent in Southeast Asian countries such as Thailand.

The epidemiological backgrounds of these tumor types are also unique. Hepatitis virus infection, aflatoxin B exposure, alcohol intake, and other metabolic diseases (including obesity, diabetes mellitus, and hemochromatosis) are well-known risk factors for HCC.^{1,2} Hepatitis B virus infection is a worldwide health problem, whereas HCV is more prevalent in Japan. However, the number of patients infected with HCV has been rapidly increasing in developed countries, especially the USA. Biliary tract cancer encompasses three subtypes based on the location: ICC, ECC, and GB. The former two are collectively called CCA. Primary sclerosing cholangitis and bile duct anomaly are the well-known predisposing factor for ICC.³ In South-East Asia, especially in Northeast Thailand, parasite infection,

Abbreviations: 1,2-DCP, 1,2-dichloropropane; AA, aristolochic acid; AAV2, adeno-associated virus type 2; ARID, AT-rich interaction domain; BTC, biliary tract cancer; CCA, cholangiocarcinoma; CDKN2A, cyclin-dependent kinase inhibitor 2A; COSMIC, Catalogue of Somatic Mutations in Cancer; CTNNB1, catenin β 1; ECC, extrahepatic cholangiocarcinoma; GB, gallbladder cancer; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; ICC, intrahepatic cholangiocarcinoma; TCGA, The Cancer Genome Atlas; TERT, telomerase reverse transcriptase; TPM, *TERT* promoter mutation; WES, whole exome sequencing; WGS, whole genome sequencing.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2018 The Authors. *Cancer Science* published by John Wiley & Sons Australia, Ltd on behalf of Japanese Cancer Association.

particularly the endemic liver flukes such as *Opisthorchis viverrini*, is a major epidemiological factor for CCA.⁴

2 | DRIVER GENE LANDSCAPE IN THE LARGEST HCC COHORT

Several landmark studies of WES and WGS, including those from the International Cancer Genome Consortium and TCGA, have already analyzed large-scale cohorts of HCC in different epidemiological and ethnic backgrounds.⁵⁻⁸

At present, genomic data from more than 1000 HCC cases are publicly available. We compiled these data from four previous studies that included 1340 cases from different ethnic groups, including 642 cases from a Japanese cohort, 451 cases from TCGA, and 247 cases from a French cohort (Table S1). Using the largest HCC cohort ever reported, we attempted to identify driver genes that could evaluate the most precise frequencies of mutations in HCC. We defined a driver gene as a statistically significantly mutated gene after adjusting the background mutation rate (Appendix S1). Thirty-three significantly mutated driver genes (q -value < 0.1; Table S2) are listed in Figure 1. The two most frequently mutated genes were *TP53* (29.1% of total cases) and *CTNNB1* (28.6%). WNT signaling regulates cell proliferation, stem cell maintenance, and other oncogenic signals, and *CTNNB1* plays a central role in activating WNT signaling. Previous sequencing analyses of different ethnic groups reported that these two genes were the most frequent, confirming that they are master driver genes of HCC beyond the diversities of ethnic/epidemiological backgrounds.

Telomerase reverse transcriptase restores telomere length during replication, and is required for cell transformation or escaping from senescence. Recently, TPMs were reported in a broad range of human cancers including HCC.⁹ Because TPMs were analyzed in part by studies we compiled (TPMs were analyzed in 1093

cases) and are difficult to be precisely identified by Illumina sequencing due to their presence in the highly GC rich genome area, the true frequencies of TPMs could not be established in the current study. However, four large-scale studies reported highly frequent TPMs (>50%) in Japanese, French, and US (TCGA) cohorts,^{5,7-10} suggesting that *TERT* also belongs to the master HCC driver genes. *TERT* promoter mutations are reported to occur in dysplastic nodules or early HCC, suggesting that TPMs could be founding driver genes in HCC.¹¹ Chiba et al¹² reported that TPMs have a two-step mechanism in cell transformation: promoting immortalization in the initial phase, and induction of genomic instability in the later stage. Collectively *TP53*, *CTNNB1* (WNT signal), and *TERT* (immortalization) are the core driver genes in HCC of all ethnic groups (Figure 2A).

The second most frequently mutated group (>7%) of HCC driver genes included *ALB* (10.2%), *APOB* (9.8%), *ARID1A* (8.8%), *ARID2* (8.2%), and *AXIN1* (7.5%). Both *ARID1A* and *ARID2* encode components of SWI/SNF chromatin remodeling machinery, and their genetic alterations are also reported in other tumor types.¹³ Because SWI/SNF machinery regulates chromatin conformation in an ATP-dependent manner, their alterations are supposed to affect epigenetic changes in cancer cells; however, the biological roles of *ARID1A/ARID2* mutations in HCC remain to be clarified. Recently, Sun et al¹⁴ reported context-dependent roles of *ARID1A* in hepatocarcinogenesis by analyzing animal models. Axin 1 promotes degradation of *CTNNB1* protein and negatively regulates WNT signaling. Therefore, its inactivation is largely complimentary to *CTNNB1* activating mutations for WNT signal activation in HCC.¹⁵ In contrast to these well-characterized tumor suppressor genes, the biological roles of highly frequent mutations of *ALB* and *APOB* in HCC are uncertain. Because both *ALB* and *APOB* mutations were strongly enriched in indels, these indels could be caused by replication slippage errors and may result from conflicts between the replication and transcription machineries.¹⁶

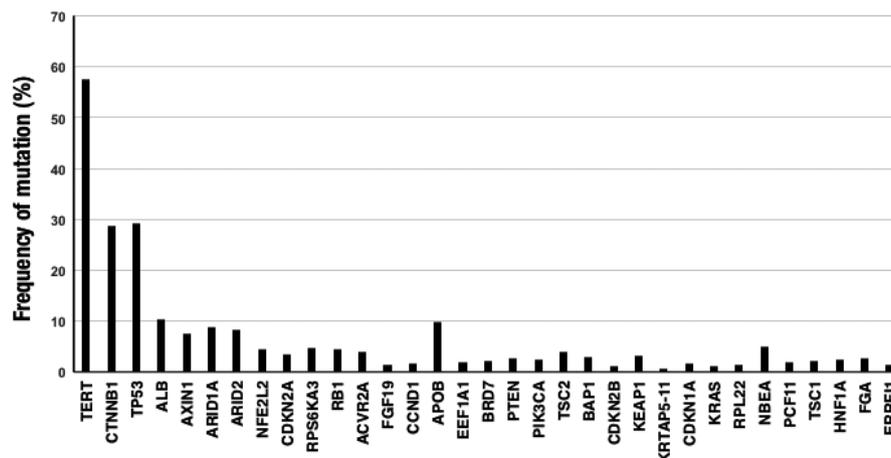


FIGURE 1 Frequencies of significantly mutated driver genes in 1340 hepatocellular carcinoma genomes. Genes were sorted by q -value. Somatic mutations in the coding exons of 1340 hepatocellular carcinoma genomes were accumulated and the statistical significance of mutation frequencies in driver genes were estimated. Briefly, the P -values from each individual significance test were calculated as the geometric mean, as previously reported.⁵ Multiple testing adjustments (q -value) were carried out as described.⁴⁵ Note that the frequency of *TERT* promoter mutations was calculated by combining data from four studies^{5,7-9}

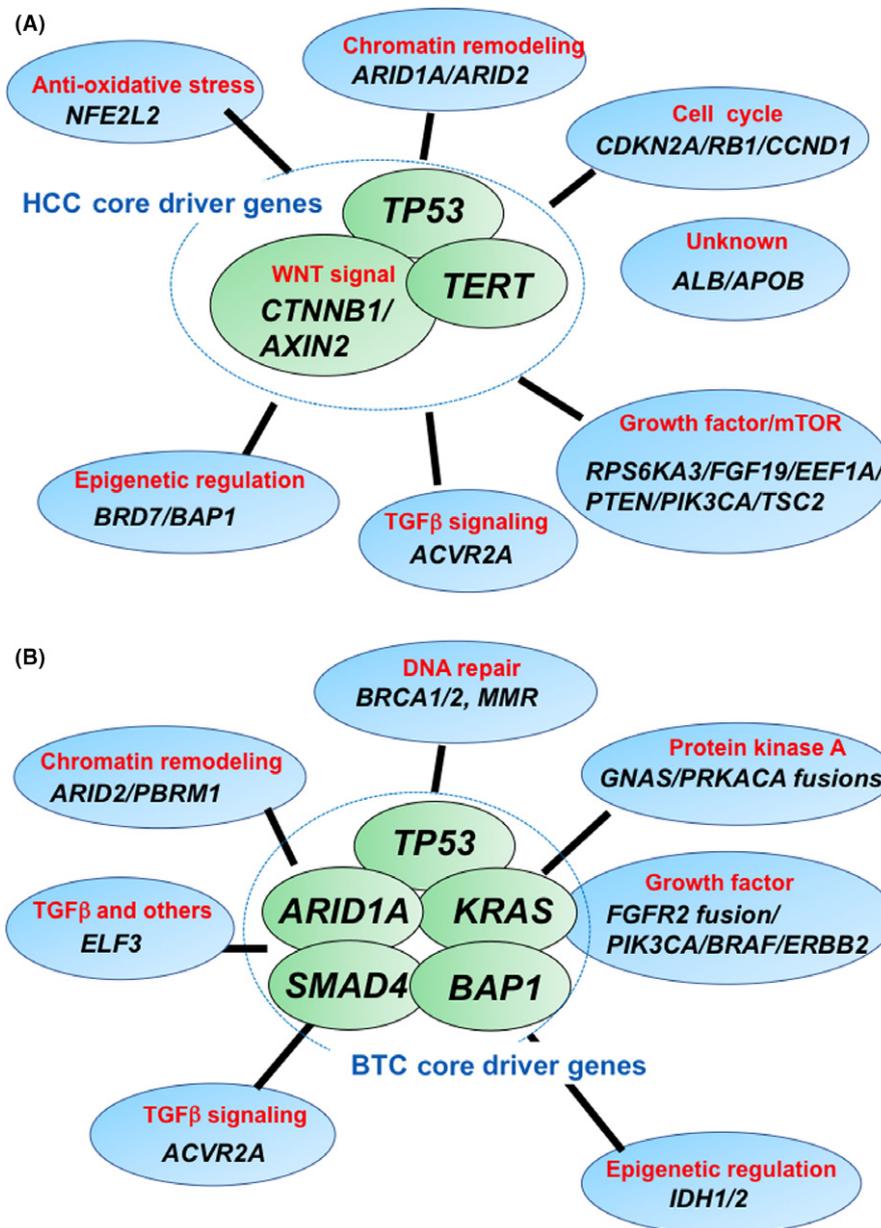


FIGURE 2 Landscapes of major driver genes in hepatocellular carcinoma (HCC) and biliary tract cancer (BTC). A, Combination of three core driver genes (*TP53*, *TERT*, and WNT signal) and other driver pathways in HCC. B, Combination of five core driver genes (*TP53*, *ARID1A*, *KRAS*, *SMAD4*, and *BAP1*) and other driver pathways in BTC. TGFβ, transforming growth factor-β

The remaining top driver genes were mutated in a fraction of cases (<5%) with various biological functions (Figure 2A). *CDKN2A*, *RB1*, and *CCND1* regulated cell cycle checkpoints, whereas *RPS6KA3*, *PTEN*, *PIK3CA*, and *TSC2* are implicated in growth factor signaling and the mTOR pathway. *EEF1A1* encodes the alpha subunit of the elongation factor-1 complex, which is responsible for the enzymatic delivery of aminoacyl tRNAs to the ribosome, which has oncogenic and invasive activities.¹⁷ They also include therapeutic targets such as *PIK3CA*, *PTEN*, and *CCND1*. These infrequently mutated driver genes might contribute to specific cellular phenotypes such as aggressive growth, invasiveness, or metastasis. Alternatively, combined with other types of alterations or alterations of other genes involved in the same functional pathways, they could play more major roles in HCC.

There are little data regarding non-coding driver genes in HCC. Fujimoto et al⁷ reported the largest WGS of 300 Japanese cases of HCC and identified recurrent non-coding mutations. They found that TPM was the most frequent and significant mutation, followed by *NEAT1* and *MALAT1*, both of which are non-coding RNAs.

3 | STRUCTURAL REARRANGEMENTS AND VIRUS INTEGRATION IN HCC

Copy number alterations and structural rearrangements occur frequently in HCC.⁵⁻⁸ Although rare, chromothripsis was also reported in HCC.⁷ A broad range of cancer-related genes including oncogenes

(including *MYC*, *TERT*, *CCND1*, and *MET*) and tumor suppressor genes (including *TP53*, *RB1*, *ARID1A*, *CDKN2A*, and *APC*) were affected by these molecular events.^{5,7,8} In addition to somatic rearrangements, HBV genome integration also affects gene expression in HCC. Recurrent targets of HBV integration include *TERT*, *MLL4*, and *CCNE1*, and their expressions were elevated by virus genome integrations.^{18,19} Copy number alterations were also significantly increased at HBV breakpoint locations, implying that HBV genome integration was associated with chromosomal instability.^{18,19}

Three groups reported integration of the AAV2 genome in HCC.^{7,20,21} Integrations of the AAV2 genome were detected 5.7% of French HCC cases.²⁰ It was reported that AAV2-positive tumors tended to occur in non-cirrhotic liver without known risk factors, suggesting that AAV2 genome integration might play an oncogenic role in some of the cases currently classified as risk-free. Integration of AAV2 was reported to be rarer (~1%) in Japanese (3/300) and Korean (2/289) cohorts.^{7,21}

4 | MUTATIONAL SIGNATURES IN THE HCC GENOME

Endogenous and exogenous mutational processes induced by various factors, such as defects in DNA repair or exposure to smoking-associated carcinogens, leave unique nucleotide changes (substitutions, insertion, or deletions) within characteristic sequence contexts (Figure 3). Examining such patterns of somatic mutations, called “mutational signatures”, can provide an indication of the etiology of the operative mutational processes in each cancer case. Analyzing more

than 7000 cancer genome sequences uncovered more than 30 distinctive mutational signatures operating in human cancers.²²

Totoki et al⁵ extracted mutational signatures in WES data of 503 HCC cases from both Japanese and TCGA cohorts. They isolated three mutational signatures (COSMIC signatures 1, 16, and 22) in that dataset, and reported unique associations between mutational signature and ethnic backgrounds. Although COSMIC signature 1 (NpCpG>NpTpG) was common in all ethnic groups, COSMIC signature 16 (ApTpN>ApCpN) was more frequent in Japanese male patients, and COSMIC signature 22 (AA exposure signature) (T>A substitution) was preferentially detected in Asian patients living in the USA. Surprisingly, virus status (HBV, HCV, and non-viral) did not show significant associations with any mutational signature. Fujimoto et al⁷ extracted six COSMIC mutational signatures (COSMIC signatures 1, 4, 9, 12, 16, and 19) and one new signature (signature W6) WGS data from 300 cases. Each mutational signature was associated with epidemiological, clinical, and driver gene status, which was revealed by multiple linear regression analysis. COSMIC signature 4 (C>A substitution) was associated with smoking status, patient age, clinical history of double cancer, and the presence of *TP53* mutation. COSMIC signature 16 was associated with alcohol consumption and the presence of *TPM* and *CTNNB1* mutation, and also male gender and older age. No association between virus status and each mutation signature was also confirmed in this study.

Schulze et al⁶ analyzed mutational signatures in 233 cases of HCC from Europe and Africa. They identified eight mutational signatures, including two new ones that were later classified as COSMIC signatures 23 and 24. COSMIC signature 24 was found to be an aflatoxin B1-associated signature, and was more frequent in African

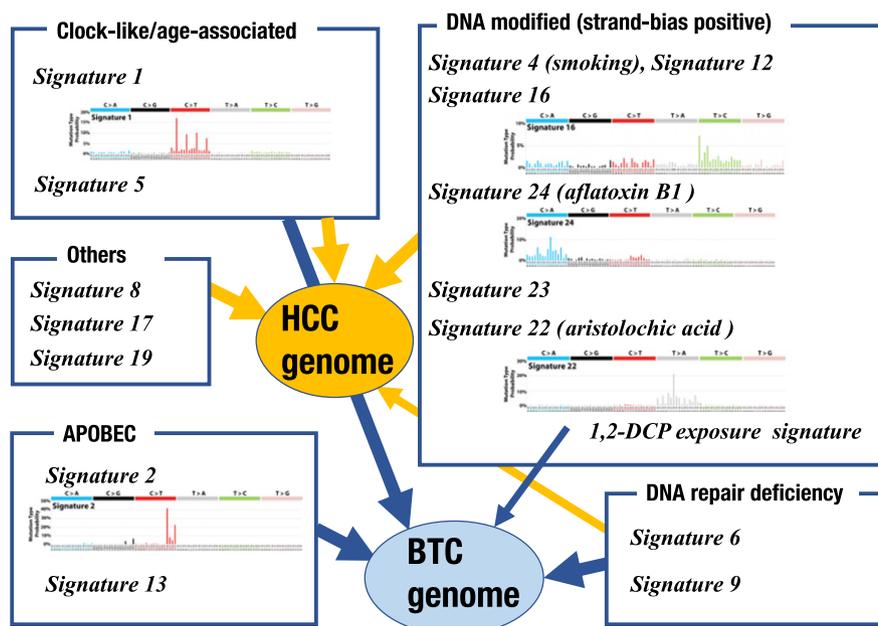


FIGURE 3 Mutational signatures in the hepatocellular carcinoma (HCC) and biliary tract cancer (BTC) genomes. Mutational signatures extracted from previous sequencing analyses of HCC and BTC genomes are classified by their mechanisms. Direct causes of these signatures, if identified, are indicated. Representative mutational spectra with sequence context (from the Catalogue of Somatic Mutations in Cancer database) are shown

cases with significant enrichment of a somatic *TP53* p. Arg249Ser mutation, which is a typical mutation of aflatoxin B1-exposed HCC. Letouzé et al²³ further analyzed WGS of European HCC cases and identified 10 mutational signatures (COSMIC signatures 1, 4, 5, 6, 12, 17, 16, 22, 23, and 24). Of note, they affirmed a significant association between COSMIC signature 16 and epidemiological backgrounds such as male gender, alcohol intake, and tobacco consumption. They also reproduced the significant association between the presence of *CTNNB1* mutation and the frequency of COSMIC signature 16.

Recently, Ng et al²⁴ reported broad prevalence of the AA signature (COSMIC signature 22) in HCC cases in Taiwan. They undertook WES of 98 HCC cases from two hospitals in Taiwan, and found that 78% of cases harbored the AA signature. Further searching in 1400 HCCs from diverse geographic regions found that 47% of Chinese HCCs and 29% of HCCs from Southeast Asia showed the AA signature, which is somewhat consistent with exposure through known herbal medicines. The AA signature was also but infrequently detected in 13% and 2.7% of HCCs from Korea and Japan, respectively, as well as in 4.8% and 1.7% of HCCs from North America and Europe, respectively. The TCGA HCC group reported 4.6% of cases with significant contribution of the AA signature.⁸ Although the frequency of AA-associated HCC is low in Japan and Western countries, exposure to AA or their derivatives would play an important role in HCC globally, especially in East Asia, implying substantial opportunities for prevention and intensive screening of HCC in these endemic areas.

5 | DRIVER GENE LANDSCAPE OF BTC

Nakamura et al²⁵ reported the first large-scale WES and transcriptome sequencing of 260 BTC cases that contained three anatomical

subtypes (145 ICC cases, 86 ECC cases, and 29 GB cases). Three top driver genes were *TP53* (26% of all cases), *KRAS* (18%), and *ARID1A* (11%), followed by *SMAD4*, *BAP1*, *PIK3CA*, *ARID2*, and *GNAS*. Similar to HCC, most driver genes were altered in a fraction (<5%) of cases, suggesting of genetic heterogeneity in BTC. *TERT* promoter mutations were detected in 21% of GB cases, whereas only one ICC case had these mutations. This is quite a contrast with the high frequency of TPMs in HCC as described above.

The list of frequently mutated driver genes in BTC cases looks partly similar to that of HCC (*TP53*, *ARID1A*, and *ARID2*) and pancreatic cancer (*TP53*, *KRAS*, and *SMAD4*) (Figure 4). By contrast, characteristically, alterations in *WNT* signaling and *TERT* were infrequent and *IDH1* mutations (3%) were unique in BTC. Nakamura et al further classified driver genes among anatomical subtypes, and reported subtype-specific and shared driver genes (Figure 4). For example, *IDH1* and *BAP1* mutations were ICC-specific, whereas *ELF3* and *ARID1B* mutations were more common in ECC. *EGFR*, *ERBB2*, *PTEN*, and *ARID2* mutations with TPMs occurred more frequently in GB cases. The molecular mechanisms underlying the selection of specific driver genes along anatomical locations and characteristic differences between HCC and ICC remain completely unknown. It could be possible that cell type-specific expression levels or signaling cascades, or carcinogenic causes, might affect the selection of driver genes. Whole genome sequencing of liver cancer displaying biliary phenotypes revealed a mixture genetic profiles of HCC and ICC, suggesting that this category could be a marginal mixture of two tumor types.²⁶ Activation or inactivation of some driver genes might exert cytotoxicity or a differentiation switch in specific lineages of liver and bile duct cancer, as reported in the case of *IDH1* mutations.²⁷

The Cancer Genome Atlas CCA project analyzed 34 CCA samples that were fluke-negative and HBV/HCV-negative.²⁸ Due to the limited size of the cohort, frequencies of driver genes were variable compared to other large-scale studies; however, notably, this cohort

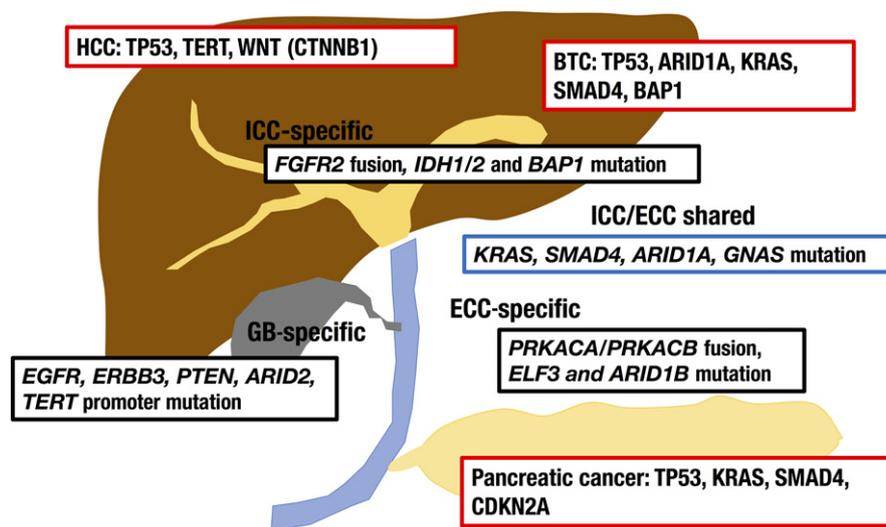


FIGURE 4 Specific driver genes across anatomical subtypes of biliary tract cancer (BTC), and comparison with core driver genes in hepatocellular carcinoma (HCC) and pancreatic cancer. ECC, extrahepatic cholangiocarcinoma; GB, gallbladder cancer; ICC, intrahepatic cholangiocarcinoma

represented a higher frequency of *IDH1* (13%) and *IDH2* (5%) mutations compared to Asian cohorts. Other studies also reported high frequencies of *IDH1/2* mutations in Caucasian CCA cases,²⁹ suggesting that *IDH1/2* could be ethnic-preferential driver genes in BTC. The CCA cells with *IDH1* mutations were reported to be more sensitive to a multikinase inhibitor dasatinib.³⁰

Jusakul et al³¹ compiled molecular data of 489 CCA cases (310 ICCs and 173 ECCs) from 10 countries including Japan, Thailand, China, Korea, and Italy. This is the largest CCA cohort to date and covers liver fluke-positive and -negative cases across various regions including both Asian and Caucasian ethnicities. The top five driver genes were *TP53* (32%), *ARID1A* (17.4%), *KRAS* (16.5%), *SMAD4* (13%), and *BAP1* (8.5%), which are the same as those identified in the Japanese cohort, followed by *APC*, *PBRM1*, and *ELF3*. Infrequent (<5%) but actionable mutations include those of *BRAF*, *IDH1*, and *BRCA2*. Because this cohort contains the largest number of fluke-positive (133 cases) and fluke-negative (356 cases) samples, more accurate fluke-associated driver genes were identified. They found that *KRAS*, *FBXW7*, and *PTEN* mutations and *ERBB2* amplification were significantly more frequent in fluke-positive cholangiocarcinoma, whereas *IDH1*, *PBRM1*, and *ACVR2A* mutations occurred more frequently in fluke-negative cases. Mutations of *TP53*, *ARID1A*, and *SMAD4*, three of the top five driver genes, were common in both types.

There are limited genomic data of GB cancers. Li et al³² reported WES of 57 Chinese GB samples. *TP53* (47.1%), *ERBB3* (11.8%), and *KRAS* (7.8%) were the top three driver genes, and ErbB signaling (including *EGFR*, *ERBB2*, *ERBB3*, and *ERBB4* and their downstream genes) was the most extensively mutated pathway. These observations were consistent with the above report of a Japanese cohort.³²

Whole transcriptome sequencing revealed fusion genes in CCA. Among them, fusion genes containing *FGFR2* tyrosine kinase and cyclic AMP-dependent PKA signaling components (*PRKACA* and *PRKACB*) they are recurrently reported in previous papers and they

are also therapeutic targets.^{25,31} The *FGFR2* fusion genes were generally exclusive to genetic alterations of other receptor kinase signal components such as *KRAS*, *NRAS*, and *BRAF*.²⁵ Protein kinase A component fusions (*ATP1B1-PRKACA/PRKACB*, and *LINC00261-PRKACB*) presented exclusiveness to other genetic alterations of PKA signaling, such as *GNAS*.²⁵ Other kinase fusion genes reported in BTC include *FIG-ROS1*, *FGFR3-TACC3*, *PTPRK-RSPO3*, and *EML4-ALK*.^{25,31,33}

The *FGFR2* fusion gene is one of characteristic molecular signatures of CCA with a frequency of approximately 10% of cases.^{25,28,34-36} Previous studies reported at least 19 different fusion partners, and various types of structural alterations, including intrachromosomal (inversion and interstitial deletion) or interchromosomal (translocation) rearrangements, are involved (Figure 5). Invariably, fusion partners replace the last exon of the *FGFR2* gene, which encodes C-terminal protein and 3'-UTR. This fusion event is supposed to cause two molecular changes; one is to add a protein interaction domain for promoting ligand-independent *FGFR2* dimerization, and the other is to stabilize the *FGFR2* transcript, probably by deleting micro-RNA target sequences in the 3'-UTR. Interestingly, Jusakul et al³¹ reported interchromosomal translocations that only deleted the last exon of *FGFR2* (3'-UTR loss) in two cholangiocarcinoma cases. They showed that cases with 3'-UTR-truncated *FGFR2* exhibited higher expression levels compared with those of wild-type *FGFR2* transcript. The *FGFR2* fusion gene is also important for clinical management of CCA. In vivo and early clinical studies showed that kinase inhibitors targeting *FGFR2* could be effective for these kinase fusion genes,³⁵⁻³⁷ and clinical trials of several *FGFR2* inhibitors are ongoing. The response rate of a recent phase II clinical trial was 18.8% for *FGFR2* fusion-positive CCA,³⁸ and acquired resistant secondary *FGFR2* mutations have been identified.³⁹ Clinical features of *FGFR2* fusion-positive CCA remain unclear, however, fluke-positive CCA was reported to be negative for *FGFR2* fusion genes.³¹

3'-gene	Position	Rearrangement	Cases
WAC	10p12	inversion	1
OPTN	10p13	inversion	1
BICC1	10q21	inversion	9
CCDC6	10q21	interstitial deletion	1
MGEA5	10q24	interstitial deletion	1
C10orf118	10q25	interstitial deletion	1
KIAA1598	10q25	interstitial deletion	3
AHCYL1	1p13	translocation	7
ZMYM4	1p34	translocation	1
TXLNA	1p35	translocation	1
TBC1D1	4p14	translocation	1
TACC3	4p16	translocation	2
FRK	6q22	translocation	1
PARK2	6q26	translocation	1
CREB5	7p15	translocation	1
PPHLN1	12q12	translocation	1
KCTD1	18q11	translocation	1
NOL4	18q12	translocation	1
STK26	Xq26	translocation	1

FIGURE 5 Diverse fusion partners with *FGFR2* reported in biliary tract cancer. Nineteen fusion partner genes with *FGFR2* identified so far are shown with structural rearrangements. References are listed in Table S3

6 | MOLECULAR CARCINOGENESIS OF BTC (FIGURE 3)

Nakamura et al²⁵ extracted two mutational signatures (COSMIC signature 1 and the APOBEC signature: COSMIC signatures 2/13) from the Japanese BTC cohort. Contributions of these signatures were significantly different among anatomical subtypes; COSMIC signature 1 was more frequent in ICC, whereas the APOBEC-associated process was more prevalent in ECC and GB. Jusakul et al³¹ analyzed WGS data of 71 CCA cases and extracted 10 mutational signatures comprising COSMIC signatures 1, 5, 8, 16, and 17, APOBEC signatures (C>T and C>G substitutions), mismatch-repair deficiency-associated signatures (COSMIC signatures 6 and 20), and AA-exposure signature 22. Fluke-positive cholangiocarcinoma were enriched for APOBEC signatures. In contrast, in TCGA cohort, a collection of both fluke and hepatitis-negative CCA cases showed only COSMIC signatures 6 and 1.²⁸ Possibly chronic infection caused by fluke and other pathogen infections may be prone to induce APOBEC signatures, however, further comprehensive study will be required for determining precise associations.

Recently, an outbreak of CCA was reported among workers in an offset color proof-printing company in Japan.⁴⁰ These cases had been exposed to high concentrations of chemical compounds, including 1,2-DCP and/or dichloromethane. The International Agency for Research on Cancer categorized 1,2-DCP as a Group 1 (carcinogenic to humans) carcinogen and dichloromethane as Group 2A (probably carcinogenic to humans).⁴¹ Whole exome sequencing of four CCA cases of the printer workers revealed unique and predominant mutational spectra that are distinct from known COSMIC mutational signatures and partly reproduced by *in vitro* 1,2-DCP exposure.⁴² These unique mutation spectra could be useful for understanding carcinogenesis processes by these chemicals and evaluating chemical exposure and screening.

7 | INTEGRATIVE MOLECULAR CLASSIFICATION OF HEPATOBILIARY CANCER (FIGURE 6)

Integrating genome data (mutation, copy number, fusion gene, and mutational signature) with transcriptome, epigenome, proteome, and metabolome data will contribute to identifying unique molecular subtypes in cancer.

The Cancer Genome Atlas HCC group assembled the largest collection of genome, transcriptome, epigenome, and proteome data in 196 HCC cases.⁸ By integrating these omics data, they identified three subclasses (iClust 1-3). These clusters partly overlapped with the gene expression-based subtypes (S1-3) reported by Hoshida et al.⁴³ Tumors in iClust 1 tended to have lower frequency of *CTNNB1* mutation, TPMs, and *CDKN2A* silencing, whereas iClust3 was characterized by a higher degree of chromosomal instability with high frequency of *TP53* mutation and CpG hypomethylation.

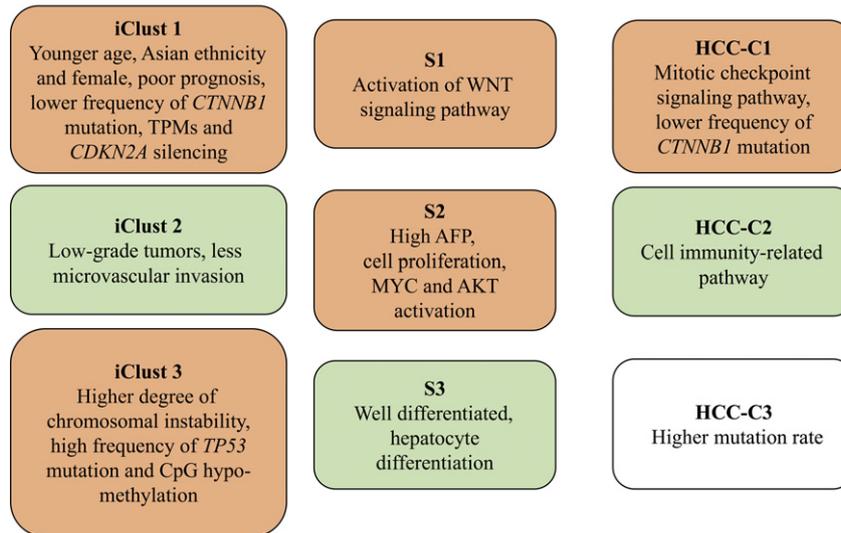
Chaisaingmongkol et al⁴⁴ reported genome, transcriptome, and metabolome analyses of 199 liver cancers (130 ICC and 69 HCC) from Thailand as the TIGER consortium. Based on gene expression data, they identified three HCC subtypes (HCC C1-3), four ICC subtypes (ICC C1-4), and a Japanese-specific ICC subtype (UM subtype). By comparing expression signatures in these subtypes, they found significant similarity between HCC-C1 and ICC-C1 (common C1), and HCC-C2 and ICC-C2 (common C2). Genes highly expressed in common C1 tumors were enriched with mitotic checkpoint signaling pathway and Hoshida's S1-2-related genes, whereas common C2 tumors tended to express more cell immunity-related pathway. They further collected molecular data of 847 liver cancers (561 HCC and 286 ICC, 582 Asian and 265 Caucasian), and found that C1/C2 clusters were reproducible in Asian samples but were less stable in Caucasian cases.

Transcriptome sequencing and WES of 160 Japanese BTC identified four BTC subtypes (Nakamura clusters 1-4), that were associated with prognosis and specific driver gene profiles.²⁵ Nakamura cluster 3 was enriched with *IDH1*, *BAP1*, and *NRAS* mutations and *FGFR2* fusion genes, but rarely harbored *TP53*, *KRAS*, and *SMAD4* mutations. This cluster partly overlaps to the UM subtype identified in the above study,⁴⁴ which was enriched in *IDH1* and *BAP1* mutations. Nakamura cluster 4 showed higher immune checkpoint activity and was enriched with hypermutated and more neo-antigen-presenting cases, suggesting immune evasion processes were more active in this subtype.

Integration of genome, transcriptome, and epigenome data from 95 CCA cases revealed four subtypes (CCA clusters 1-4) that were associated with the presence of liver fluke infection and prognosis.³¹ Tumors in CCA cluster 1 were predominantly liver fluke-positive and harbored *ERBB2* amplification and CpG island hypermethylation. Cholangiocarcinoma cluster 4 was enriched with liver fluke-negative cases and *BAP1* and *IDH1/2* mutations and *FGFR* alterations, which are similar to Nakamura cluster 3. Cases in CCA cluster 4 also showed CpG shore hypermethylation, partly due to *IDH1* mutations, and better prognosis. Tumors in CCA cluster 1 showed a significantly higher contribution of COSMIC signature 1 and intratumoral heterogeneity. They proposed that chronic inflammation by liver fluke infection might continuously induce CpG island methylation, which resulted in accumulation of COSMIC signature 1-associated mutations and tumor heterogeneity in CCA cluster 1 cases, whereas tumors in CCA cluster 4 acquired strong driver mutations with rapid growth. The Cancer Genome Atlas CCA group more clearly isolated a subtype enriched with *IDH1/2*-mutated tumors, which showed a unique methylation pattern and upregulation of mitochondrial gene expression signatures.²⁸ Their data could separate *IDH1/2* mutated cases from *BAP1/FGFR2* altered cases based on methylation signature.

Current integrative studies have already identified stable molecular subtypes (possibly three) in HCC, and four or more subtypes including *FGFR2* fusions, *IDH1/2* and *BAP1*-mutated group in CCA, however, further collection of trans-omics data of clinical samples with different ethnic and epidemiological backgrounds should be

(A) TCGA HCC (mixed ethnicity) Hoshida's (mixed ethnicity) TIGER (Thailand cases)



(B) Nakamura BTC (Japanese) TCGA (US) ICGC (mixed ethnicity) TIGER (mixed ethnicity)

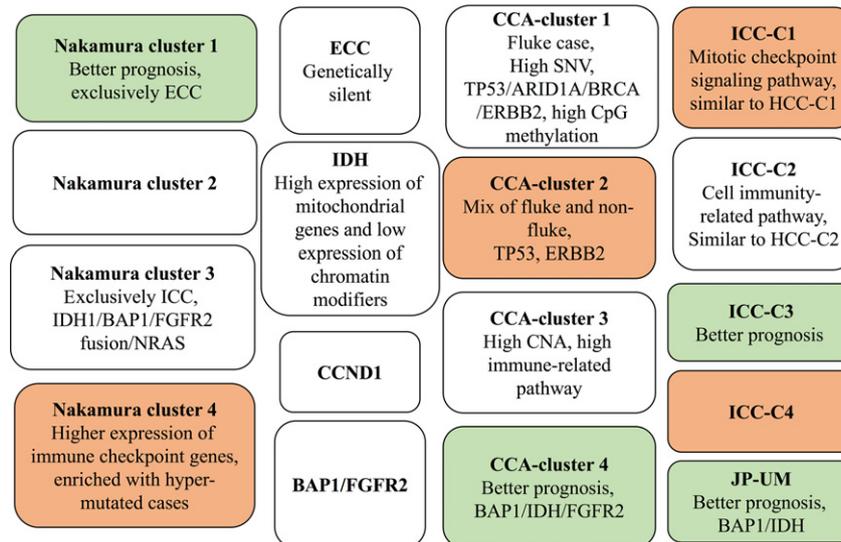


FIGURE 6 Molecular classifications of hepatocellular carcinoma (HCC) and biliary tract cancer (BTC) by integrative molecular analyses. (a) Clinical and molecular features in subtypes of HCC reported in three studies^{8,42,43} are summarized. TPM, *TERT* promoter mutation. (b) Clinical and molecular features in subtypes of BTC or cholangiocarcinoma (CCA) reported in four studies^{24,27,30,43} are summarized. Subtypes colored green showed better prognosis; those colored orange showed poorer prognosis. ECC, extrahepatic cholangiocarcinoma; ICC, intrahepatic cholangiocarcinoma; JP-UM, Japanese-specific ICC subtype; TCGA, The Cancer Genome Atlas

required. There would exist some overlap in molecular signatures of subtypes between HCC and ICC despite distinctive core driver gene sets between the two tumor types.

8 | FUTURE PERSPECTIVES

Comprehensive molecular and genomic analyses of large cohorts of hepatobiliary cancers have now uncovered landscapes of driver genes, characteristic mutational signatures, and molecular classifications of these tumor types. To apply this knowledge for better

treatment, diagnosis, and prevention of these intractable cancers, genetic screening of each tumor sample, so-called clinical sequencing, for optimized and individualized choice of treatment, clinical trials, and translational research of new drugs, and novel approaches for early diagnosis, such as liquid biopsy, are expected. Furthermore, deeper and broader genomic and basic research, together with international collaborations, are required to understand the biological roles of each or combinations of identified driver genes in the carcinogenesis processes of these tumor types, and to uncover intratumoral, interpatient, and inter-ethnic diversities of genomic profiles in more detail.

ACKNOWLEDGMENTS

This work was supported in part by the Practical Research for Innovative Cancer Control from the Japan Agency for Medical Research and Development (AMED), National Cancer Center Research and Development Funds (29-A-6).

CONFLICT OF INTEREST

The authors have no conflict of interest.

ORCID

Tatsuhiko Shibata  <http://orcid.org/0000-0002-0477-210X>

REFERENCES

- Yang JD, Roberts LR. Hepatocellular carcinoma: a global view. *Nat Rev Gastroenterol Hepatol*. 2010;7:448-458.
- Shibata T, Aburatani H. Exploration of liver cancer genomes. *Nat Rev Gastroenterol Hepatol*. 2014;11:340-349.
- Rizvi S, Gores GJ. Pathogenesis, diagnosis, and management of cholangiocarcinoma. *Gastroenterology*. 2013;145:1215-1229.
- Zheng S, Zhu Y, Zhao Z, Wu Z, Okanurak K, Lv Z. Liver fluke infection and cholangiocarcinoma: a review. *Parasitol Res*. 2017;116:11-19.
- Totoki Y, Tatsuno K, Covington KR, et al. Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. *Nat Genet*. 2014;46:1267-1273.
- Schulze K, Imbeaud S, Letouze E, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genet*. 2015;47:505-511.
- Fujimoto A, Furuta M, Totoki Y, et al. Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. *Nat Genet*. 2016;48:500-509.
- Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell*. 2017;169:1327-1341.
- Bell RJ, Rube HT, Xavier-Magalhães A, et al. Understanding TERT promoter mutations: a common path to immortality. *Mol Cancer Res*. 2016;14:315-323.
- Nault JC, Mallet M, Pilati C, et al. High frequency of telomerase reverse-transcriptase promoter somatic mutations in hepatocellular carcinoma and preneoplastic lesions. *Nat Commun*. 2013;4:2218.
- Nault JC, Calderaro J, Di Tommaso L, et al. Telomerase reverse transcriptase promoter mutation is an early somatic genetic alteration in the transformation of premalignant nodules in hepatocellular carcinoma on cirrhosis. *Hepatology*. 2014;60:1983-1992.
- Chiba K, Lorbeer FK, Shain AH, et al. Mutations in the promoter of the telomerase gene *TERT* contribute to tumorigenesis by a two-step mechanism. *Science*. 2017;357:1416-1420.
- St Pierre R, Kadoch C. Mammalian SWI/SNF complexes in cancer: emerging therapeutic opportunities. *Curr Opin Genet Dev*. 2017;42:56-67.
- Sun X, Wang SC, Wei Y, et al. Arid1a has context-dependent oncogenic and tumor suppressor functions in liver cancer. *Cancer Cell*. 2017;32:574-589.
- Taniguchi K, Roberts LR, Aderca IN, et al. Mutational spectrum of beta-catenin, AXIN1, and AXIN2 in hepatocellular carcinomas and hepatoblastomas. *Oncogene*. 2002;21:4863-4871.
- Sankar TS, Wastuwidyaningtyas BD, Dong Y, et al. The nature of mutations induced by replication-transcription collisions. *Nature*. 2016;535:178-181.
- Abbas W, Kumar A, Herbein G. The eEF1A proteins: at the crossroads of oncogenesis, apoptosis, and viral infections. *Front Oncol*. 2015;5:75.
- Sung WK, Zheng H, Li S, et al. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nat Genet*. 2012;44:765-769.
- Zhao LH, Liu X, Yan HX, et al. Genomic and oncogenic preference of HBV integration in hepatocellular carcinoma. *Nat Commun*. 2016;7:12992.
- Nault JC, Datta S, Imbeaud S, et al. Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. *Nat Genet*. 2015;47:1187-1193.
- Park KJ, Lee J, Park JH, et al. Adeno-associated virus 2-mediated hepatocellular carcinoma is very rare in Korean patients. *Ann Lab Med*. 2016;36:469-474.
- Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500:415-421.
- Letouze E, Shinde J, Renault V, et al. Mutational signatures reveal the dynamic interplay of risk factors and cellular processes during liver tumorigenesis. *Nat Commun*. 2017;8:1315.
- Ng AWT, Poon SL, Huang MN, et al. Aristolochic acids and their derivatives are widely implicated in liver cancers in Taiwan and throughout Asia. *Sci Transl Med*. 2017;9. <https://doi.org/10.1126/scitranslmed.aan6446>.
- Nakamura H, Arai Y, Totoki Y, et al. Genomic spectra of biliary tract cancer. *Nat Genet*. 2015;47:1003-1010.
- Fujimoto A, Furuta M, Shiraishi Y, et al. Whole-genome mutational landscape of liver cancers displaying biliary phenotype reveals hepatitis impact and molecular diversity. *Nat Commun*. 2015;6:6120.
- Saha SK, Parachoniak CA, Ghanta KS, et al. Mutant IDH inhibits HNF-4 α to block hepatocyte differentiation and promote biliary cancer. *Nature*. 2014;513:110-114.
- Farshidfar F, Zheng S, Gingras MC, et al. Integrative genomic analysis of cholangiocarcinoma identifies distinct IDH-mutant molecular profiles. *Cell Rep*. 2017;18:2780-2794.
- Borger DR, Tanabe KK, Fan KC, et al. Frequent mutation of isocitrate dehydrogenase (IDH)1 and IDH2 in cholangiocarcinoma identified through broad-based tumor genotyping. *Oncologist*. 2012;17:72-79.
- Saha SK, Gordan JD, Kleinstiver BP, et al. Isocitrate dehydrogenase mutations confer dasatinib hypersensitivity and SRC dependence in intrahepatic cholangiocarcinoma. *Cancer Discov*. 2016;6:727-739.
- Jusakul A, Cutcutache I, Yong CH, et al. Whole-genome and epigenomic landscapes of etiologically distinct subtypes of cholangiocarcinoma. *Cancer Discov*. 2017;7:1116-1135.
- Li M, Zhang Z, Li X, et al. Whole-exome and targeted gene sequencing of gallbladder carcinoma identifies recurrent mutations in the ErbB pathway. *Nat Genet*. 2014;46:872-876.
- Gu TL, Deng X, Huang F, et al. Survey of tyrosine kinase signaling reveals ROS kinase fusions in human cholangiocarcinoma. *PLoS ONE*. 2011;6:e15640.
- Wu YM, Su F, Kalyana-Sundaram S, et al. Identification of targetable FGFR gene fusions in diverse cancers. *Cancer Discov*. 2013;3:636-647.
- Arai Y, Totoki Y, Hosoda F, et al. Fibroblast growth factor receptor 2 tyrosine kinase fusions define a unique molecular subtype of cholangiocarcinoma. *Hepatology*. 2014;59:1427-1434.
- Wang Y, Ding X, Wang S, et al. Antitumor effect of FGFR inhibitors on a novel cholangiocarcinoma patient derived xenograft mouse model endogenously expressing an FGFR2-CCDC6 fusion protein. *Cancer Lett*. 2016;380:163-173.
- Borad MJ, Champion MD, Egan JB. Integrated genomic characterization reveals novel, therapeutically relevant drug targets in FGFR and

- EGFR pathways in sporadic intrahepatic cholangiocarcinoma. *PLoS Genet.* 2014;10:e1004135.
38. Javle M, Lowery M, Shroff RT, et al. Phase II study of BGJ398 in patients with FGFR-altered advanced cholangiocarcinoma. *J Clin Oncol.* 2018;36:276-282.
39. Goyal L, Saha SK, Liu LY, et al. Polyclonal secondary FGFR2 mutations drive acquired resistance to FGFR inhibition in patients with FGFR2 fusion-positive cholangiocarcinoma. *Cancer Discov.* 2017;7:252-263.
40. Kubo S, Nakanuma Y, Takemura S, et al. Case series of 17 patients with cholangiocarcinoma among young adult workers of a printing company in Japan. *J Hepatobiliary Pancreat Sci.* 2014;21:479-488.
41. Benbrahim-Tallaa L, Lauby-Secretan B, Loomis D, et al. Carcinogenicity of perfluorooctanoic acid, tetrafluoroethylene, dichloromethane, 1,2-dichloropropane, and 1,3-propane sultone. *Lancet Oncol.* 2014;15:924-925.
42. Mimaki S, Totsuka Y, Suzuki Y, et al. Hypermutation and unique mutational signatures of occupational cholangiocarcinoma in printing workers exposed to haloalkanes. *Carcinogenesis.* 2016;37:817-826.
43. Hoshida Y, Nijman SM, Kobayashi M, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res.* 2009;69:7385-7392.
44. Chaisaingmongkol J, Budhu A, Dang H, et al. Common molecular subtypes among asian hepatocellular carcinoma and cholangiocarcinoma. *Cancer Cell.* 2017;32:57-70.
45. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B.* 1995;57:289-300.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Shibata T, Arai Y, Totoki Y. Molecular genomic landscapes of hepatobiliary cancer. *Cancer Sci.* 2018;109:1282-1291. <https://doi.org/10.1111/cas.13582>