Genomic landscapes reveal post-transcriptional modifier disruption in cholangiocarcinoma

Xiangcheng Li (✉ drxcli@njmu.edu.cn)

The First Affiliated Hospital of Nanjing Medical University

Yaodong Zhang

The First Affiliated Hospital of Nanjing Medical University

Zijian Ma

Nanjing Medical University

Changxian Li

University of Hong Kong

Cheng Wang

Nanjing Medical University

Wangjie Jiang

The First Affiliated Hospital of Nanjing Medical University

Jiang Chang

The First Affiliated Hospital of Nanjing Medical University

Sheng Han

the First Affiliated Hospital of Nanjing Medical University

Zefa Lu

The First Affiliated Hospital of Nanjing Medical University

Zicheng Shao

The First Affiliated Hospital of Nanjing Medical University

Yirui Wang

The First Affiliated Hospital of Nanjing Medical University

Hongwei Wang

The First Affiliated Hospital of Nanjing Medical University

Chenyu Jiao

The First Affiliated Hospital of Nanjing Medical University

Dong Wang

The First Affiliated Hospital of Nanjing Medical University

Xiaofeng Wu

The First Affiliated Hospital of Nanjing Medical University

Hongbing Shen

Nanjing Medical University

Xuehao Wang

The First Affiliated Hospital of Nanjing Medical University

Zhibin Hu

Nanjing Medical University

Article

Keywords: Genome Sequencing, N6-Methyladenosine Modification, Post-transcriptional Modification, Tumor Mutation Burden, Driver Genes

Posted Date: December 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-120813/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Communications on June 1st, 2022. See the published version at <https://doi.org/10.1038/s41467-022-30708-7>.

Abstract

Molecular variation of different geographic populations and cholangiocarcinoma (CCA) subtypes indicate CCA's potential genomic heterogeneity and novel genomic features. We analyzed exome-sequencing data of 87 perihilar cholangiocarcinoma (pCCA) and 261 intrahepatic cholangiocarcinoma (iCCA) from 3 Asian centers (including 43 pCCAs and 24 iCCAs from our center). Patients with iCCA presented higher tumor mutation burden and copy number alteration burden (CNAB) than pCCA patients, and CNAB indicated a poorer pCCA prognosis. In our newly identified 12 significantly mutated genes and 5 focal CNA regions, post-transcriptional modification-related novel driver genes *METTL14* and *RBM10* are prone to occur in pCCA. We demonstrated the tumor-suppressing functional role of *METTL14*, a major RNA N6-adenosine methyltransferase (m6A), and its loss-of-function mutation R298H may function through m6A modification on new driver gene *MACF1*. Our results may be valuable for better understanding of post-transcriptional modification effects CCA development, and similarities and differences between pCCA and iCCA.

Introduction

Cholangiocarcinoma (CCA) is a malignant primary hepatobiliary disease originating from every point of the biliary tree, from the canals of Hering to the main bile duct¹. In recent decades, the incidence and mortality of CCA have increased globally, especially in East Asia^{2,3}. In China, the incidence of CCA reached 6–7.55 per 100,000 individuals^{4,5}. The mortality of CCA is almost equal to its incidence due to resistance to common treatments and poor prognosis⁶. The overall five-year survival rate is only 10%⁷. Even for the patients with CCA who are amenable to surgical resection, and the corresponding five-year survival rate is only 25–35%⁸, and high recurrence rates of 50 to 60% persist upon diagnosis, with no effective therapy.

CCA is regarded as a group of different diseases that are further divided into intrahepatic CCA (iCCA), perihilar CCA (pCCA), and distal CCA (dCCA) based on anatomical location⁹. pCCA is the most frequent subtype, accounting for 50–60% of all CCA tumors⁹. In general, the incidence and mortality rates of iCCA was reported rising worldwide in the past decade, whereas the incidences of pCCA seem to be stable or decreasing¹⁰. Besides the burden of HCV infection has been linked with the incidence of iCCA, infection with liver flukes and HBV-related liver diseases have shown a stronger association for iCCA in East Asia^{10,11}. Primary sclerosing cholangitis (PSC) is considered as a predominant risk factor for CCA, especially for pCCA^{10,12}. The pronounced geographical and etiology heterogeneity of iCCA and pCCA indicated the potential diverse cancer-initiating cell in these two subtypes of CCA¹³. In detail, iCCA is considered to originate two different cell of origin (mucin-secreting cholangiocyte or hepatic progenitor cell), due to the significant inter-tumor heterogeneity¹⁴. Unlike iCCA, pCCA only originate from mucin-secreting cholangiocyte^{13,15}.

To date, previous studies of the genomic alterations in a variety of bile duct cancer, especially for iCCA, have demonstrated some known commonalities in mutation of CCA. These studies identified a number of driver genes, but most of them represented the genomic features of iCCA, such as *TP53*, *KRAS*, *SMAD4*, and *ARID1A*^{16–20}. Notably, the latest integrative genomic analysis of Caucasian extrahepatic CCA (eCCA) (including the pCCA and dCCA) revealed distinct molecular characterization of eCCA in western²¹. However, the epidemiological profile of CCA and its subtypes shows enormous geographical variation, indicating underlying genomic heterogeneity across different regions⁵. The prevalence of CCA is highest in East Asia and the subtype distribution is similar. Thus, integration analysis in sequencing data from Asian countries may find new genomic features ignored by previous studies. Here, we performed whole-exome sequencing on 43 pCCA and 24 iCCA, and analyzed an additional 44 pCCA and 237 CCA to investigate the genomic landscapes of pCCA and iCCA, respectively.

Methods And Materials

Detailed information about the clinical sample collection and follow-up; DNA preparation, DNA capture, and sequencing; Next-generation sequencing; Sequencing alignment and detection of somatic variants; Mutational burden and signature analysis; Identification of CCA driver genes and comparison between the iCCA and pCCA; Somatic copy number estimation; Highly amplified/deleted regions identification; Pathway enrichment analysis; Annotation of genomic alterations upon clinical actionability; Cell culture and siRNA transfection; Establishment of stable *METTL14*^{R298H} and *METTL14*^{wt} cells and functional assays; RNA N6-methyladenosine (m⁶A) Quantification; RNA m⁶A immunoprecipitation assay and m⁶A sequencing; M⁶A sequencing analysis; Measurement of RNA lifetime; Western blotting; RNA isolation and quantitative real-time PCR (qPCR); Immunohistochemistry staining; Evaluation of immunostaining; Immunofluorescence analysis and statistical analysis are available in the supplementary material.

Results

Mutational spectrum of pCCA and iCCA

We conducted integration analysis of exome-sequencing data from 87 pCCA cases (including 43 xxx Medical University (xMU) cases and 44 International Cancer Genome Consortium (ICGC) cases) and 261 iCCA cases (including 24 xMU cases, 135 ICGC cases, and 102 cases from Zou et al.'s study¹⁷). The general information and clinical features of patients from the xMU study are listed in **Supplementary Table 1**. The mean sequencing depth of the xMU study was 94.4× in tumor tissues and 91.8× in normal tissues dissected adjacent to the tumor. Detailed sequencing coverage and depth information are listed in **Supplementary Table 2**. We detected a total of 91,995 somatic mutations (**Supplementary Table 2**) in all the cases, and the median number per case was 120.5. The median tumor mutation burden (TMB) was 2.0 megabase per case (/Mb) and the burden was comparable between pCCA (2.0/Mb) and iCCA (2.0/Mb) (Wilcoxon rank-sum test $P = 0.48$, **Supplementary Fig. 1A**). Among these mutations, we defined 28,254 non-synonymous mutations, for which the median was 33 Single Nucleotide Variants (SNVs) and

3 small insertion and deletions (INDELs) per patient. Interestingly, the non-synonymous mutation burden (median 0.61/Mb) of iCCA was significantly higher than that (median 0.47/Mb) of pCCA (Wilcoxon rank-sum test $P = 9.6 \times 10^{-4}$, **Fig. 1A**). Similar to the non-synonymous mutation pattern, iCCA had a higher copy number alteration burden (CNAB) than pCCA (iCCA median = 25.0%, pCCA median = 8.4%, Wilcoxon rank-sum test $P = 2.4 \times 10^{-4}$), no matter the amplification or deletion (**Fig. 1C**).

Next, we investigated the genomic mutational signatures, which provide clues for the carcinogenesis of CCA. Firstly, we classified all the mutations into six classic subtypes. Consistent with previous studies^{17, 18}, C>T/G>A substitutions accounted for the predominant type of all SNVs in all of the patients (**Supplementary Fig. 1B**). The proportion of the mutation subtypes was similar between pCCA and iCCA (Wilcoxon rank-sum test $P = 0.46, 0.09, 0.67, 0.62, 0.51$, and 0.29 respectively for C>A, C>G, C>T, T>A, T>C, and T>G, **Supplementary Fig. 1B**). We further analyzed mutational signatures using R package SignatureEstimation, which accurately reconstructs the mutation profiles of all pCCA and iCCA patients based on a predefined mutational spectrum of 30 COSMIC signatures²². COSMIC signature 1 (Age, 44.71%), signatures associated with mismatch repair deficiency (MMR, Signatures 6, 15, and 26, 31.99%), signatures associated with activated APOBECs (Signatures 2 and 13, 6.48%), and signature 8 (6.41%) were dominant (proportion > 5%) in pCCA patients (**Fig. 1B**). For iCCA patients, COSMIC signature 1 (27.19%) and signature MMR (26.74%) were also identified as dominant, but the proportion was slightly lower than in pCCA patients. In addition, signature 4 related to smoking behavior (10.13%), signature 22 (9.96%) related to aristolochic acid, and signatures specific to liver cancer (Liver, Signatures 12, 16, and 24, 8.7%) were predominant in iCCA. Interestingly, signature Liver were more common in HBV-related iCCA (HBV: 26.0%, non-HBV: 3.8%, **Supplementary Table 3**). Aristolochic acid-related mutations were mainly found in iCCA patients from Zou et al.'s study, but this signature was also observed in pCCA patients (**Supplementary Table 4**). In addition, signature 3, related to homologous recombination deficiency (4.26%), signature 10, related to altered POLE activity (2.44%), and signature 7, related to ultraviolet (UV) light exposure (2.17%), were only identified in iCCA (**Fig. 1B**), which suggested that multiple types of DNA damage occurred during the carcinogenesis progress of CCA.

The results mentioned above revealed the nature of genomic instability in CCA patients. We further investigated the association between TMB/CNAB and the overall survival of patients. We found that the patients carrying higher loads of mutations showed worse outcomes after adjusting for age, sex, and stage at diagnosis (Hazard ratio (HR) = 1.71, $P = 3.9 \times 10^{-4}$, **Fig. 1D**). The association was consistent in pCCA and iCCA patients (**Supplementary Fig. 1C**). When examining CNAB, we noticed that the survival of pCCA patients with higher CNAB was significantly worse than for patients with lower CNAB (HR = 2.70, $P = 3.6 \times 10^{-2}$, **Fig. 1E**); However, there was no significant difference between iCCA patients with distinct CNAB (HR = 0.91, $P = 6.7 \times 10^{-1}$, **Fig. 1F**).

New driver genes

We identified 36 significantly mutated genes (SMGs, **Fig. 2**), also described as mut-drivers, in all patients by MutSig2CV and IntOGen, including 24 reported SMGs and 12 new SMGs not highlighted in previous CCA publications (**Supplementary Table 5**). Among these new SMGs, *MACF1* (5.17%, 18/348) and *AXIN1* (0.86%, 3/348, all of the three patients were with iCCA), which is commonly considered as one of drivers in hepatocellular carcinoma (HCC), are two interactive components of the β -catenin/Wnt signaling pathway (**Fig. 2**). Although the mutation rate of *AXIN1* was low, we noticed that there were multiple non-synonymous mutations in two iCCA patients (**Fig. 2**). The mutations of *PIK3R1*, a component of the *PI3K* pathway, was displayed in 2.59% (9/348) of patients. Post-transcriptional modification genes *RBM10* (3.16%, 11/348) and *METTL14* (0.86%, 3/348) act as well-known alternative splicing regulators and m⁶A writers. Importantly, we identified recurrent mutations (**Supplementary Methods**) in *METTL14*, and all mutations in this gene affected the same amino acid (**Supplementary Fig. 2A**), which were also predicted as potentially deleterious variants by several bioinformatic pathogenicity prediction tools (**Supplementary Table 5**). We observed the same mutation rate (1.44%, 5/348) of chromatin modifiers *SMARCA4* and *WHSC1*. Interestingly, mutations in these genes were prone to co-occur in the same patient (Fisher's exact test OR = 660, $P = 2.29 \times 10^{-6}$, **Fig. 2**). In addition, classic cancer driver genes identified in other cancers were also identified as SMGs in this study, including *ATM* (3.45%, 12/348), *BRCA2* (2.01%, 7/348), and *MLLT4* (2.01%, 7/348). Although *BRCA2*, *EPHA2* and *ATM* were mentioned in previously published work due to their same position recurrent inactivating mutations, our integrated analysis with more sample size firstly identified them as SMGs¹⁸.

We further defined the genes with a distinct mutation rate between pCCA and iCCA (see **Supplementary Methods**). *TP53*, *ARID1A*, *PBRM1*, *MACF1*, *EPHA2*, *ARID2*, *IDH1*, *PTEN*, *RB1*, *BRAF*, *NRAS*, *SLC8A1*, *AXIN*, and *MLLT4* were iCCA-enriched genes, while *RBM10*, *TGFBR2*, *PIK3R1*, *ELF3*, *NACC1*, and *METTL14* were pCCA-enriched genes. The mutation rates of *TP53* and *IDH1* were significantly higher in iCCA (Fisher's exact test $OR_{TP53} = 1.3$, $P_{TP53} = 8.8 \times 10^{-3}$; $OR_{IDH1} = +\infty$, $P_{IDH1} = 4.4 \times 10^{-2}$, **Fig.2**). Importantly, *IDH1* only mutated in iCCA, but not in pCCA. *RBM10*, however, carried significantly more mutations in pCCA than in iCCA (Fisher's exact test $OR_{RBM10} = 3.8$, $P_{RBM10} = 3.2 \times 10^{-2}$).

We further identified 22 regions frequently altered by focal copy number alteration (CNAs) in CCAs (**Fig. 3A, Supplementary Table 6**), including 16 regions overlapping with previously reported peaks, and six new significantly altered regions (amplification at 7q31.2, 22q11.21, and deletion at 1p36.13, 2p24.1, 7q35, and 12q24.33, **Fig.3A**). The cancer related genes located at these peaks were considered to be CNA-driver genes. The gain of a 0.8-Mb region at 7q31.2 occurred in 45 samples (26.6%), which contains the classic

MET oncogene in the RTK/RAS pathway. However, the amplification of 7q31.2 was found rare previously in Caucasian population²³. The loss at 1p36.13 involved tumor suppressor gene *SHDB* and affected 76 samples (45.0%). DNA Polymerase Epsilon (*POLE*) at 12q24.33, where was also considered lost in HCC²⁴, lost one copy in 40 patients (23.7%). In addition, important DNA methylation factor *DNMT3A* (2p24.1) and chromatin methylation factor *EZH2* (7q35) also show loss in 20 (11.8%) and 19 (11.24%) patients (including iCCA and pCCA), respectively. Whereas these two regions were also only considered deleted in iCCA based on a previous single nucleotide polymorphism array(SNP-array) analysis²⁵. We only observed heterogeneous loss in CCA patients (**Fig. 3A**). Among all of the regions, the frequencies of amplification/gain at 8q24.21 (*MYC*), 1q21.3 (*S100A7*), 5p15.33 (*SDHHA*), and 7q31.2 (*MET*) were significantly higher in iCCA (**Supplementary Table 6**). The frequency of high-level amplification was significantly higher in pCCA than in iCCA, although the frequency of amplification/gain was comparable across the subtypes (**Fig.3A**). In addition, we found that copy number loss and nonsynonymous mutations of classic tumor suppressors (*TP53*, *CDKN2A*, *SMAD4*, *PTEN*, and *ATM*) were prone to co-occur in the same patients (Fisher's exact test OR = 8.4, $P = 8.2 \times 10^{-6}$, **Fig. 3B**).

Next, we performed pathway enrichment analysis on all driver genes identified in this study, which included SMGs (mut-drivers) and cancer genes in the frequently altered focal CNA regions (CNA-drivers) described above, and found that the driver genes of CCA were significantly enriched in the RTK-RAS, Wnt, PI3K, Cell Cycle, TP53, TGF-beta, and HIPPO pathways (**Fig.3C**). Including the newly identified *MET* amplification, 34.3% of CCA patients harbored mutations and CNAs in oncogenes from the RTK-RAS pathway, and these alterations occurred in a mutually exclusive manner, as reported in other cancers (**Fig. 3D**).

Functional recurrent mutations in *METTL14*

Among the potential driver genes mentioned above, we identified a new potential driver gene, *METTL14*, the main factor involved in aberrant m6A modification of various cancers, with recurrent and deleterious mutations^{26, 27} (**Supplementary Fig. 2A, Supplementary Table 7**). All three mutations affected the same 298 amino acid residue (*METTL14* p.R298H and p.R298C) and two of them were in the same genomic position (two patients from xMU and ICGC respectively) (**Supplementary Fig. 2A**). We further conducted sanger sequencing in an independent pCCA cohort with extra 40 subjects from xMU cohort and identified an additional p.R298H carrier (**Supplementary Fig. 2B & C**). In the COSMIC database, we also found that the same p.R298H also occurred in three pancreatic ductal carcinomas (**Supplementary Fig. 2A**) and all the three patients were East Asian.

Crystal structures of the *METTL3-METTL14* complex have revealed that p.R298 lies close to the putative RNA-binding groove of the complex which may have a complex role to affect methylation activity²⁸. The recurrent and deletions mutant p.R298H suggested us its possibly positive selection and the need of *METTL14*'s normal action in antitumor progress. However, the relevance of this hotspot mutation and m⁶A mRNA methylation to the CCA has not yet been established. Dysfunction of *METTL14*, the key catalytic protein forming the core m⁶A methyltransferase complex, has shown fundamental biological effects in cancer initiation and progression^{28,29}. Thus, we hypothesized that CCA could be associated with *METTL14* that regulated m⁶A mRNA methylation. Hence, we first examined *METTL14* expression in 69 CCA pairs and matched adjacent normal tissues. The results of qRT-PCR and western blot revealed that expression of *METTL14* was significantly downregulated in tumors ($P < 0.01$; **Fig. 4A & B**). The immunohistochemistry (IHC) staining on tissue microarray also confirmed that *METTL14* staining was decreased in CCA at protein level (**Supplementary Fig.3A**) and its downregulation displayed a marginal significant association with poor cancer-specific survival in CCA (Log-Rank $P = 0.08$; **Supplementary Fig.3B**). Consistently, we found that the m⁶A level of total RNA was significantly decreased in CCA tissues (**Fig. 4C**). Interestingly, we also observed the m⁶A modification level was significantly decreased in *METTL14* low expression CCA group compared with *METTL14* high expression CCA group (**Supplementary Fig. 3C**). These results suggested that *METTL14* and it mediated m⁶A modification were frequently down-regulated or disturbed in CCA.

To determine the “driver” role of *METTL14* and R298H mutations during CCA development, we first confirmed the transfection efficiency of lentiviral constructs expressing *METTL14*^{wt} and *METTL14*^{R298H} in RBE and HCCC9810 cell lines (**Fig. 4D and Supplementary Fig. 3D&E**). In contrast to the cells stably overexpressing *METTL14*^{wt}, cells overexpressing *METTL14*^{R298H} showed significantly decreased overall m⁶A modification ability (**Fig. 4E and Supplementary Fig. 3F**). Subsequently, while overexpression of wild-type *METTL14* decreased cell proliferation, overexpression of the mutation had no noticeable effect on cell proliferation (**Fig. 4F, G & H and Supplementary Fig. 3G, H&K**). In addition, overexpression of *METTL14*^{wt} resulted in a decrease in cell migration and invasion, and *METTL14*^{R298H} remarkably reversed the gene's ability to inhibit migration and invasion in CCA cells (**Fig. 4I & Supplementary Fig. 3I & J**). Taken together, these results provided us evidence for loss of function caused by *METTL14*^{R298H} mutation, suggesting that *METTL14*^{R298H} mutation could reduce the tumor-suppressing effect of *METTL14*^{wt}.

New driver gene *MACF1* served as the target of *METTL14*

We then conducted RNA-Seq and MeRIP-Seq assays in negative control, *METTL14*^{wt}, and *METTL14*^{R298H} cells. A total of 2,601 peaks involving classic transcripts of 1,586 genes were robustly identified by

exomePeak2 in all cells. The most common m⁶A motif, GGAC, was significantly enriched in the m⁶A peaks identified (**Supplementary Fig. 4A**), and the m⁶A peaks were especially enriched in the vicinity of the stop codon. Next, we performed differential m⁶A-methylation analysis between control and *METTL14*^{wt} cells, as well as between *METTL14*^{wt} and *METTL14*^{R298H} cells. Because of the writer role of *METTL14* during the m⁶A methylation modification, we included only m⁶A peaks (712) with increased abundance in *METTL14*^{wt} cells as compared to control cells, and peaks (990) with decreased abundance in *METTL14*^{R298H} as compared to *METTL14*^{wt} cells. A total of 237 peaks were shared by both analyses, including four peaks on the two driver genes mentioned above (*MACF1*, *MET*) (**Fig. 5A & B**). *MACF1* was also the new driver gene identified in this study (**Supplementary Table 5**). We then verified whether m⁶A-modified *MACF1* was susceptible to decay. The lifetime of *MACF1* was prolonged in *METTL14*^{R298H} cells and shortened in *METTL14*^{wt} cells after actinomycin D treatment (**Fig. 5C**). We further immunoprecipitated m⁶A from the RNAs of *METTL14*^{wt} and *METTL14*^{R298H} cells and found that *METTL14*^{R298H} significantly decreased the amount of *MACF1* modified by m⁶A compared to *METTL14*^{wt} (**Fig. 5D**). Because *MACF1* was a very large gene (involving 22Kb of exons), we applied immunofluorescence assays to further confirm that *METTL14*^{wt} mediated *MACF1* degradation, and *METTL14*^{R298H} showed increase expression of *MACF1* compared to *METTL14*^{wt} (**Fig. 5E**). We also found that *METTL14*^{wt} cells decreased the expression of *MACF1*, and no noticeable effect on *MACF1* expression was observed in *METTL14*^{R298H} cells using qRT-PCR (**Supplementary Fig. 4B**).

When we successfully transfected CCA cells with siRNA pools targeting *MACF1* (**Supplementary Fig. 4D**), significant inhibition of tumor metastasis and proliferation was observed (**Fig. 5F, G, H, & I, and Supplementary Fig. 4E & F**). Recent studies reported that *MACF1* was involved in tumor metastasis and cytoskeleton, and that it played crucial roles in the nucleus translocation of β-catenin^{30,31}. In our findings, we first validated *MACF1* upregulation in CCA using qRT-PCR (**Supplementary Fig. 4C**), and found that knockdown of *MACF1* significantly reduced CCA cell proliferation and metastasis *in vitro*. Given the essential role of *MACF1* in regulating the nucleus translocation of β-catenin, immunofluorescence assays showed that the increase of nuclear β-catenin was correlated with the expression of *METTL14*^{R298H} rather than *METTL14*^{wt} (**Fig. 5J & K**). To ascertain the role of *MACF1* in *METTL14*-mediated nucleus translocation of β-catenin, we transfected *MACF1* siRNA in *METTL14*^{R298H}-overexpressing cells, and observed that the nucleus translocation of β-catenin was decreased compared to only *METTL14*^{R298H}-overexpressing cells (**Supplementary Fig. 4G**).

Additionally, *METTL14*^{R298H} resulted in reversing the expression level of E-cadherin, N-cadherin, PCNA, and CyclinD1, which were the downstream targets of β-catenin, relevant to *METTL14*^{wt} (**Fig. 5K**). These

results implied that *METTL14*-mediated m⁶A modification repressed the *MACF1*/ β -catenin pathway in CCA, while *METTL14*^{R298H} mutation disrupted this mechanism.

Discussion

In this study, we included genomic data from 348 CCA samples (including 87 from pCCA and 261 from iCCA) to investigate the genomic landscapes of both pCCA and iCCA. We found the shared and distinct features between these two anatomical subtypes. Currently, there is no effective biomarker that can accurately predict the prognosis of CCA patients^{32,33}. Our study found that TMB has a good predictive effect on the prognosis of East-Asian CCA patients. It is worth noting that pCCA patients with higher CNAB showed worse survival outcomes compared to those with lower CNAB. These findings suggest that we can explore prognostic markers of patients with CCA from a more macroscopic genetic perspective.

In addition to several new driver genes in classic pathways, we identified driver genes participating in post-transcriptional modification (i.e., *METTL14* and *RBM10*), which carried more mutations in pCCA patients. With *in vitro* experiments, we confirmed the role of *METTL14*, as well as the role of this mutation in CCA development. We further determined that m⁶A modification level and expression of new driver gene *MACF1* could be regulated by *METTL14*, which can influence the proliferation and metastasis ability of CCA cells.

The previous study showed that there was lack of evidence that anatomical site determines molecular subtypes according to integration analysis of transcriptomic and methylation data³⁴. Our study identified four new driver genes that were related to chromatin and methylation modification (i.e., mut-drivers: *SMARCA4* and *WHSC1*; CNA-drivers: *DNMT3A* and *EZH2*), further supporting the finding that driver genes cause disruption at in the transcriptional level. Consistent with the previous study's findings, we found comparable frequency between these alterations of driver genes. However, our analysis first identified two genes as drivers (*METTL14* and *RBM10*) and found that they are related to post-transcription modification. Mutations in these two genes were more common in pCCA patients. This evidence suggested that the investigation of the difference between pCCA and iCCA should be extended to the post-transcriptional area. *METTL14* engages in m⁶A modification^{35,36}, which modulates alternative splicing, export, stability, and translation of mRNA³⁷. Interestingly, we found that two new driver genes (*MACF1* and *MET*) could be modulated by the m⁶A alteration caused by *METTL14* mutations. It was worthy of note that all of these genes had a higher mutation rate or copy number altered rate in iCCA than in pCCA. Thus, they were crucial driver genes in both iCCA and pCCA. However, they could be activated through different mechanisms. Although *MACF1* and *MET* could serve as shared therapeutic targets for pCCA and iCCA, we should not ignore the fact that *METTL14* could modulate m⁶A modification of broader genes other than the two drivers, which may lead to unexpected drug resistance or side effects,

and therefore, further investigation in pCCA patients is warranted. Another gene, *RMB10*, is an RNA binding protein and alternative splicing regulator frequently mutated in lung adenocarcinomas³⁸. Similar to lung cancer, the majority of *RBM10* mutations (63.6%) in CCA truncate the protein and may alter downstream splicing of specific genes^{38,39}. Thus, the post-transcriptional events in CCA could be highly disturbed and further study may help determine new molecular subtypes and optimize therapeutic strategy in the clinical setting.

Our study also identified a number of new driver genes, involving multiple important pathways. *MACF1* and *AXIN1* were first identified as SMGs of CCA. They are involved in a complex that also contained *CTNNB1*, *GSK3B*, and *APC*, and can contribute to the activation of β -catenin/*Wnt* signaling pathways^{30,40}. Consistently, the mutations in the two genes were mutually exclusive. CNA-driver *MET* belongs to the classic RTK-RAS signaling pathway, and its inhibitor has been widely used in other cancers^{41,42}. However, *MET* amplification has never been reported in a genomic study of pCCA and reported with rare amplification frequency in Caucasion iCCA²³, suggesting that it may occur more frequently in Chinese patients. According to OncoKB, nearly 106 (30.5%) of all CCA patients carried actionable alterations in the RTK-RAS signaling pathway. *PIK3R1*, phosphatidylinositol 3-kinase regulatory subunit alpha, is the predominant regulatory isoform of PI3K⁴³ and is frequently mutated in multiple cancers^{44,45}. *ATM* served as the activator of the *TP53* tumor suppressor protein and somatic *ATM* mutations or deletions that are commonly found in lymphoid malignancies, pancreatic cancer, and lung adenocarcinoma⁴⁶. Notably, our results suggested that *ATM* could be affected by both truncating mutations and deletion, resulting in the bi-allelic inactivation. Several known tumor suppressors, such as *TP53*, *SMAD4*, *PTEN*, and *CDKN2A*, showed a similar bi-allelic loss pattern. Germline mutations in *BRCA2*, which encodes a *BRCA*-associated protein, were reported in familial CCA cases^{47,48}. Our study also suggested that *BRCA2* could be affected by somatic mutations. *MLLT4*, also known as *AF6*, commonly fused with *KMT2A* in acute myeloid leukemia. It has also been identified as a SMGs in breast cancers⁴⁹. Few genomic alterations has ever been reported in *NACC1*, but it is activated in ovarian serous carcinomas and influences cell apoptosis, senescence, and cytokinesis in cancer cells^{50,51}. Integrated whole-exome analysis of Asian CCAs identified new CCA driver genes and emphasized the similarities and distinctions genomic characteristics between pCCA and iCCA. Importantly, our study also highlighted the effect of underlying post-transcriptional modification in CCA occurrence and development. These results provide a better understanding of the CCA mutational landscape that may drive improvements in clinical practice.

References

1. Global Burden of Disease Cancer, C. *et al.* The Global Burden of Cancer 2013. *JAMA oncology* **1**, 505–527 (2015).
2. Shin, H.R. *et al.* Comparison of incidence of intrahepatic and extrahepatic cholangiocarcinoma—focus on East and South-Eastern Asia. *Asian Pacific journal of cancer prevention: APJCP* **11**, 1159–

1166 (2010).

3. Khan, S.A. *et al.* Rising trends in cholangiocarcinoma: is the ICD classification system misleading us? *Journal of hepatology* **56**, 848–854 (2012).
4. Rizvi, S. & Gores, G.J. Emerging molecular therapeutic targets for cholangiocarcinoma. *Journal of hepatology* **67**, 632–644 (2017).
5. Banales, J.M. *et al.* Expert consensus document: Cholangiocarcinoma: current knowledge and future perspectives consensus statement from the European Network for the Study of Cholangiocarcinoma (ENS-CCA). *Nature reviews. Gastroenterology & hepatology* **13**, 261–280 (2016).
6. Rizvi, S., Khan, S.A., Hallemeier, C.L., Kelley, R.K. & Gores, G.J. Cholangiocarcinoma - evolving concepts and therapeutic strategies. *Nature reviews. Clinical oncology* **15**, 95–111 (2018).
7. Everhart, J.E. & Ruhl, C.E. Burden of digestive diseases in the United States Part III: Liver, biliary tract, and pancreas. *Gastroenterology* **136**, 1134–1144 (2009).
8. Hyder, O. *et al.* A nomogram to predict long-term survival after resection for intrahepatic cholangiocarcinoma: an Eastern and Western experience. *JAMA surgery* **149**, 432–438 (2014).
9. Rizvi, S. & Gores, G.J. Pathogenesis, diagnosis, and management of cholangiocarcinoma. *Gastroenterology* **145**, 1215–1229 (2013).
10. Burak, K. *et al.* Incidence and risk factors for cholangiocarcinoma in primary sclerosing cholangitis. *The American journal of gastroenterology* **99**, 523–526 (2004).
11. Lee, T.Y. *et al.* Hepatitis B virus infection and intrahepatic cholangiocarcinoma in Korea: a case-control study. *The American journal of gastroenterology* **103**, 1716–1720 (2008).
12. Goeppert, B. *et al.* Genomic Characterization of Cholangiocarcinoma in Primary Sclerosing Cholangitis Reveals Therapeutic Opportunities. *Hepatology* (2020).
13. Komuta, M. *et al.* Histological diversity in cholangiocellular carcinoma reflects the different cholangiocyte phenotypes. *Hepatology* **55**, 1876–1888 (2012).
14. Komuta, M. *et al.* Clinicopathological study on cholangiolocellular carcinoma suggesting hepatic progenitor cell origin. *Hepatology* **47**, 1544–1556 (2008).
15. Sato, Y., Harada, K., Sasaki, M. & Nakanuma, Y. Cystic and micropapillary epithelial changes of peribiliary glands might represent a precursor lesion of biliary epithelial neoplasms. *Virchows Archiv: an international journal of pathology* **464**, 157–163 (2014).
16. Li, M. *et al.* Whole-exome and targeted gene sequencing of gallbladder carcinoma identifies recurrent mutations in the ErbB pathway. *Nature genetics* **46**, 872–876 (2014).
17. Zou, S. *et al.* Mutational landscape of intrahepatic cholangiocarcinoma. *Nature communications* **5**, 5696 (2014).
18. Nakamura, H. *et al.* Genomic spectra of biliary tract cancer. *Nature genetics* **47**, 1003–1010 (2015).
19. Farshidfar, F. *et al.* Integrative Genomic Analysis of Cholangiocarcinoma Identifies Distinct IDH-Mutant Molecular Profiles. *Cell reports* **19**, 2878–2880 (2017).

20. Nepal, C. *et al.* Genomic perturbations reveal distinct regulatory networks in intrahepatic cholangiocarcinoma. *Hepatology* **68**, 949–963 (2018).
21. Montal, R. *et al.* Molecular classification and therapeutic targets in extrahepatic cholangiocarcinoma. *Journal of hepatology* **73**, 315–327 (2020).
22. Alexandrov, L.B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
23. Javle, M. *et al.* Biliary cancer: Utility of next-generation sequencing for clinical management. *Cancer* **122**, 3838–3847 (2016).
24. Zhang, Z., Xu, L. & Sun, C. Comprehensive characterization of cancer genes in hepatocellular carcinoma genomes. *Oncol Lett* **15**, 1503–1510 (2018).
25. Sia, D. *et al.* Integrative molecular analysis of intrahepatic cholangiocarcinoma reveals 2 classes that have different outcomes. *Gastroenterology* **144**, 829–840 (2013).
26. Weng, H. *et al.* METTL14 Inhibits Hematopoietic Stem/Progenitor Differentiation and Promotes Leukemogenesis via mRNA m(6)A Modification. *Cell stem cell* **22**, 191–205 e199 (2018).
27. Ma, J.Z. *et al.* METTL14 suppresses the metastatic potential of hepatocellular carcinoma by modulating N(6)-methyladenosine-dependent primary MicroRNA processing. *Hepatology* **65**, 529–543 (2017).
28. Wang, P., Doxtader, K.A. & Nam, Y. Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Molecular cell* **63**, 306–317 (2016).
29. Wang, X. *et al.* Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. *Nature* **534**, 575–578 (2016).
30. Chen, H.J. *et al.* The role of microtubule actin cross-linking factor 1 (MACF1) in the Wnt signaling pathway. *Genes & development* **20**, 1933–1945 (2006).
31. Afghani, N. *et al.* Microtubule actin cross-linking factor 1, a novel target in glioblastoma. *International journal of oncology* **50**, 310–316 (2017).
32. Baudis, M. Genomic imbalances in 5918 malignant epithelial tumors: an explorative meta-analysis of chromosomal CGH data. *BMC cancer* **7**, 226 (2007).
33. Stratton, M.R., Campbell, P.J. & Futreal, P.A. The cancer genome. *Nature* **458**, 719–724 (2009).
34. Jusakul, A. *et al.* Whole-Genome and Epigenomic Landscapes of Etiologically Distinct Subtypes of Cholangiocarcinoma. *Cancer discovery* **7**, 1116–1135 (2017).
35. Wang, Y. *et al.* N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nature cell biology* **16**, 191–198 (2014).
36. Liu, J. *et al.* A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nature chemical biology* **10**, 93–95 (2014).
37. Yang, Y., Hsu, P.J., Chen, Y.S. & Yang, Y.G. Dynamic transcriptomic m(6)A decoration: writers, erasers, readers and functions in RNA metabolism. *Cell research* **28**, 616–624 (2018).

38. Chen, H. *et al.* Genomic and immune profiling of pre-invasive lung adenocarcinoma. *Nature communications* **10**, 5472 (2019).
39. Sun, Y. *et al.* Autoregulation of RBM10 and cross-regulation of RBM10/RBM5 via alternative splicing-coupled nonsense-mediated decay. *Nucleic acids research* **45**, 8524–8540 (2017).
40. Wu, X. *et al.* Skin stem cells orchestrate directional migration by regulating microtubule-ACF7 connections through GSK3beta. *Cell* **144**, 341–352 (2011).
41. Italiano, A. *et al.* Cabozantinib in patients with advanced Ewing sarcoma or osteosarcoma (CABONE): a multicentre, single-arm, phase 2 trial. *The Lancet. Oncology* **21**, 446–455 (2020).
42. Shah, M.A. *et al.* Effect of Fluorouracil, Leucovorin, and Oxaliplatin With or Without Onartuzumab in HER2-Negative, MET-Positive Gastroesophageal Adenocarcinoma: The METGastric Randomized Clinical Trial. *JAMA oncology* **3**, 620–627 (2017).
43. Vallejo-Diaz, J., Chagoyen, M., Olazabal-Moran, M., Gonzalez-Garcia, A. & Carrera, A.C. The Opposing Roles of PIK3R1/p85alpha and PIK3R2/p85beta in Cancer. *Trends in cancer* **5**, 233–244 (2019).
44. Liu, Y. *et al.* The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nature genetics* **49**, 1211–1218 (2017).
45. Clark, V.E. *et al.* Recurrent somatic mutations in POLR2A define a distinct subset of meningiomas. *Nature genetics* **48**, 1253–1259 (2016).
46. Choi, M., Kipps, T. & Kurzrock, R. ATM Mutations in Cancer: Therapeutic Implications. *Molecular cancer therapeutics* **15**, 1781–1791 (2016).
47. Wardell, C.P. *et al.* Genomic characterization of biliary tract cancers identifies driver genes and predisposing mutations. *Journal of hepatology* **68**, 959–969 (2018).
48. Maynard, H. *et al.* Germline alterations in patients with biliary tract cancers: A spectrum of significant and previously underappreciated findings. *Cancer* **126**, 1995–2002 (2020).
49. Pereira, B. *et al.* The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nature communications* **7**, 11479 (2016).
50. Yap, K.L. *et al.* NAC1 is an actin-binding protein that is essential for effective cytokinesis in cancer cells. *Cancer research* **72**, 4085–4096 (2012).
51. Ishikawa, M. *et al.* NAC1, a potential stem cell pluripotency factor expression in normal endometrium, endometrial hyperplasia and endometrial carcinoma. *International journal of oncology* **36**, 1097–1103 (2010).

Figures

Figure 1

Mutational spectrum of patients with CCA. (A) The tumor mutation burden (TMB) of nonsynonymous mutations (substitutions and indels) between pCCAs (presented in red) and iCCAs (presented in blue). P value was obtained from Wilcoxon rank-sum test. (B) Summary composition of COSMIC 30 mutation signatures in pCCAs (red) and iCCAs (blue). COSMIC Signature MMR included signatures 6, 15, and 26; Signature liver included signatures 12, 16, and 24; Signature APOBEC included signatures 2 and 13. (C) Difference of copy number alterations burden (CNAB) no matter amplification or deletion between pCCAs (red) and iCCAs (blue). Wilcoxon rank-sum test was performed to determine the P value. Amp, amplification; Del, deletion. (D-F) Prognostic effect of TMB in all samples. (D) CNAB in 43 pCCAs (E) CNAB in 126 iCCAs. (F) Cox proportional hazard model adjusted for age, gender, and stage was conducted to obtain HR and P value. * $P < 0.05$; ** $P < 0.01$, *** $P < 0.001$

Figure 1

Mutational spectrum of patients with CCA. (A) The tumor mutation burden (TMB) of nonsynonymous mutations (substitutions and indels) between pCCAs (presented in red) and iCCAs (presented in blue). P value was obtained from Wilcoxon rank-sum test. (B) Summary composition of COSMIC 30 mutation signatures in pCCAs (red) and iCCAs (blue). COSMIC Signature MMR included signatures 6, 15, and 26; Signature liver included signatures 12, 16, and 24; Signature APOBEC included signatures 2 and 13. (C) Difference of copy number alterations burden (CNAB) no matter amplification or deletion between pCCAs (red) and iCCAs (blue). Wilcoxon rank-sum test was performed to determine the P value. Amp, amplification; Del, deletion. (D-F) Prognostic effect of TMB in all samples. (D) CNAB in 43 pCCAs (E) CNAB in 126 iCCAs. (F) Cox proportional hazard model adjusted for age, gender, and stage was conducted to obtain HR and P value. * $P < 0.05$; ** $P < 0.01$, *** $P < 0.001$

Figure 2

Mutation landscapes of curated significantly mutated genes (SMGs) and recurrent mutation genes in CCA. Each column represents one CCA patient and each row represents a gene. The gene symbols in red are our newly identified CCA-related genes, and genes in black have been reported in CCA. The bar chart at the left panel indicates the overall mutational rate of curated genes in all CCAs. The right panel is the mutation spectrum and mutation rate for pCCAs and iCCAs, respectively. The asterisks following gene symbol indicate there are recurrent mutations in the relevant genes.

Figure 3

Frequently altered focal CNA regions and characteristics of mutation-driver and CNA-driver genes. (A) Shows 22 curated frequently-altered focal CNA regions involving cancer-related regions. Each column represents one of the CCA patients and each row represents a focal region. The bar chart at the left panel indicates the overall CNA rate in all CCAs. The right panel is the CNA spectrum and rate for pCCAs and iCCAs, respectively. The top panel represents amplification and the bottom represents deletion regions. Region names in red and blue indicate the novel alterations the plus sign following the gene symbol indicates a significant alteration rate difference between pCCAs and iCCAs. Asterisk means the genes in

focal region were also identified in SMGs. (B) The co-occurrence of copy number losses and mutations of relevant tumor suppressor genes. (C) Pathway enrichment results of mutation-driver and CNA- driver genes in TCGA pancancer ten pathways. (D) Characters in CNA and mutation of driver genes in RAS-RTK pathway. *P < 0.05; **P < 0.01, ***P < 0.001

Figure 4

METTL14R298H mutation reduced the METTL14wt tumor-suppressing effect in CCA. (A) The protein level of METTL14 was validated in CCA tissues. (B) Downregulated METTL14 mRNA expression was detected in 69 pairs of iCCA and pCCA tumor tissues as well as the adjacent normal tissues by qRT-PCR. (C) The level of m6A modification in tumor and adjacent normal tissues. (D) The transfection efficiency of lentiviral constructs expressing METTL14wt and METTL14R298H in RBE cell line. (E) METTL14R298H reduced METTL14wt-mediated m6A modification detected by m6A colorimetric quantification in RBE cell lines. (F&G) Proliferation curve of RBE cells with METTL14R298H, METTL14wt or negative control. (H) Colony formation assay of RBE cells with METTL14R298H, METTL14wt, or negative control. The numbers of colonies were counted and presented in the histogram. (I) Representative images (left) and quantification (right) of transwell migration and invasion assays in RBE cells with METTL14wt,

METTL14R298H or negative control. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; T, Tumor tissue, N, normal tissue, R298H, METTL14R298H; WT, METTL14wt; NC, negative control.

Figure 4

METTL14R298H mutation reduced the METTL14wt tumor-suppressing effect in CCA. (A) The protein level of METTL14 was validated in CCA tissues. (B) Downregulated METTL14 mRNA expression was detected in 69 pairs of iCCA and pCCA tumor tissues as well as the adjacent normal tissues by qRT-PCR. (C) The level of m6A modification in tumor and adjacent normal tissues. (D) The transfection efficiency of lentiviral constructs expressing METTL14wt and METTL14R298H in RBE cell line. (E) METTL14R298H reduced METTL14wt-mediated m6A modification detected by m6A colorimetric quantification in RBE cell lines. (F&G) Proliferation curve of RBE cells with METTL14R298H, METTL14wt or negative control. (H) Colony formation assay of RBE cells with METTL14R298H, METTL14wt, or negative control. The numbers of colonies were counted and presented in the histogram. (I) Representative images (left) and quantification (right) of transwell migration and invasion assays in RBE cells with METTL14wt, METTL14R298H or negative control. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; T, Tumor tissue, N, normal tissue, R298H, METTL14R298H; WT, METTL14wt; NC, negative control.

Figure 5

M6A modification mediated expression of MACF1 in β -catenin nucleocytoplasmic transport. (A) Overlap of significantly different m6A modification peaks identified from METTL14R298H compared with METTL14wt group (presented in blue), METTL14wt compared with negative control group (presented in red) and mutation-driver genes (in grey). (B) The log2 scaled fold change of different m6A modification peaks in group WT versus NC (METTL14wt compared with negative controls) (red) and group R298H versus WT (METTL14R298H compared with METTL14wt) (blue). (C) RNA lifetime for MACF1 in RBE cells transfected with METTL14R298H, METTL14wt, or negative control. (D) The m6A modification level of MACF1 was validated in immunoprecipitation (RIP). (E) Representative images of MACF1 immunofluorescence in METTL14wt and METTL14R298H containing in vitro. Scale bars: 50 μ m. The bottom bar plot shows the MACF1 foci number in the METTL14 positive cells. (F) Colony formation assay of RBE cells with MACF1 and si-MACF1. (G) The numbers of colonies were counted and presented in the histogram. (H) Proliferation curve of RBE cells with MACF1 and si-MACF1. (I) Representative images (left) and quantification (right) of transwell migration and invasion assays in RBE cells with

MACF1 and si-MACAF1. (J) Representative images of β -catenin immunofluorescence showed nuclear β -catenin tended to be increased by expression of METTL14R298H compared to METTL14wt. Scale bars: 50 μ m. (K) Western blot analysis was performed to analyze β -catenin in the cytoplasmic and nuclear extracts and expression of E-cadherin, N-cadherin, PCNA, and CyclinD1 in total cell extracts. The P values were calculated using Student's t test. *P < 0.05, **P < 0.01, ***P < 0.001; R298H, METTL14R298H; WT, METTL14wt; NC, negative control.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementary.pdf](#)
- [Supplementary.pdf](#)
- [SupplementaryTableS6.S7.xlsx](#)
- [SupplementaryTableS6.S7.xlsx](#)