Published in final edited form as:

Hepatology. 2018 September ; 68(3): 949–963. doi:10.1002/hep.29764.

Genomic Perturbations Reveal Distinct Regulatory Networks in Intrahepatic Cholangiocarcinoma

Chirag Nepal^{1,*}, Colm J. O'Rourke^{1,*}, Douglas VNP Oliveira¹, Andrzej Taranta¹, Steven Shema², Prson Gautam³, Julien Calderaro^{4,5,6}, Andrew Barbour⁷, Chiara Raggi⁸, Krister Wennerberg^{1,3}, Xin W. Wang⁹, Anja Lautem¹⁰, Lewis R. Roberts¹¹, and Jesper B. Andersen^{1,#}

¹Biotech Research and Innovation Centre (BRIC), Department of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen N, Denmark ²Center for Cancer Research Genomics Core, National Cancer Institute, NIH Bethesda, Maryland 20892, USA ³Institute for Molecular Medicine Finland (FIMM), University of Helsinki, 00290 Helsinki, Finland ⁴Assistance Publique-Hôpitaux de Paris, Department of Pathology, CHU Henri Mondor, F-94000, Créteil, France ⁵Faculté de Médecine, Université Paris-Est Créteil, Créteil, France ⁶Inserm U955 Equipe 18, Institut Mondor de Recherche Biomédicale, Créteil, France ⁷University of Queensland, Brisbane, Australia ⁸Center for Autoimmune Liver Diseases, IRCCS Istituto Clinico Humanitas, 20089 Rozzano (MI), Italy ⁹Laboratory of Human Carcinogenesis, Center for Cancer Research, National Cancer Institute, Bethesda, Maryland 20892, USA ¹⁰Department of General, Visceral and Transplantation Surgery, University Medical Center Mainz, Mainz, 55131, Germany ¹¹Division of Gastroenterology and Hepatology, Mayo Clinic College of Medicine, Rochester, MN, USA

Abstract

Intrahepatic cholangiocarcinoma (iCCA) remains a highly heterogeneous malignancy that has eluded effective patient stratification to date. The extent to which such heterogeneity can be influenced by individual driver mutations remains to be evaluated. Here, we analyzed genomic (whole-exome sequencing, targeted exome sequencing) and epigenomic data from 496 patients, and used the three most recurrently mutated genes to stratify patients (*IDH*, *KRAS*, *TP53*).

#Corresponding author: Jesper B Andersen, Biotech Research and Innovation Centre (BRIC) Department of Health and Medical Sciences University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N Denmark, Phone: +45 35325834, jesper.andersen@bric.ku.dk.

*Authors contributed equally

Disclosure: The authors have nothing to disclose

Financial & competing interest's disclosure: The authors have no relevant affiliation or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in this manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

DATA AVAILABILITY

Data is accessible at the European Nucleotide Archive with accession number PRJEB14974.

AUTHOR CONTRIBUTIONS

JBA, CN, CJO, DVO, AT conceived the idea, designed and implemented the study; CN, CJO, DVO, SS, AT, JBA, performed experiments and data analyses; CN, CJO, DVO, JBA, interpreted omics data; JC, AB, CR, XWW, AL, LRR, JBA patient recruitment, clinical data collection and pathological assessment; CJO, DVO, AT, PG, KW, drug testing, assay development and performance; JBA, CJO, CN, wrote the manuscript; all authors read, edited and approved the manuscript.

‘undetermined’). Using this molecular dissection approach, each subgroup was determined to possess unique mutational signature preferences, co-mutation profiles and enriched pathways. High-throughput drug repositioning in seven patient-matched cell lines, chosen to reflect the genetic alterations specific for each patient group, confirmed *in silico* predictions of subgroup-specific vulnerabilities linked to enriched pathways. Intriguingly, patients lacking all 3 mutations (‘undetermined’) harbored the most extensive structural alterations while IDH mutant tumors displayed the most extensive DNA methylome dysregulation, consistent with previous findings.

Conclusion—Stratification of iCCA patients based on occurrence of mutations in three classifier genes (*IDH*, *KRAS*, *TP53*) revealed unique oncogenic programs (mutational, structural, epimutational) that influence pharmacologic response in drug repositioning protocols. This genome dissection approach highlights the potential of individual mutations to induce extensive molecular heterogeneity and could facilitate advancement of therapeutic response in this dismal disease.

INTRODUCTION

Intrahepatic cholangiocarcinoma (iCCA) incidence and mortality rates have increased globally over the last three decades, highlighting a growing clinical problem from what was once considered a rare orphan disease (1, 2). The majority of iCCAs are sclerosing adenocarcinomas arising from cholangiocytes lining intra-lobular biliary ductules, as well as peribiliary glands of large intrahepatic bile ducts (1, 3). While the causal biology of this cancer is largely unclear, known risk factors include chronic hepatitis B or C viral infection with cirrhosis (4); chronic liver conditions, such as cholestasis, primary sclerosing cholangitis (5) or liver fluke infestation (6); metabolic disorders, such as diabetes mellitus (DM); and lifestyle or environmental factors, including alcohol consumption and smoking (7). Liver resection remains the only curative treatment, a surgical procedure frequently precluded by early invasion of surrounding tissues. Even among 10–30% patients eligible for resection, 50% recur within the first 12 months. Standard chemotherapy regimens (gemcitabine alongside platin-based compounds) for non-resectable and recurrent iCCA patients are purely palliative, achieving only a limited improvement in survival (8). As most iCCAs are locally advanced or metastatic at presentation, patients have a dismal prognosis with a 5-year survival rate that has persisted below 10% since the 1980s (9). While no targeted therapy has been approved to date, *FGFR2* inhibitors (e.g. BGJ398) may be the first to transform iCCA clinical management, displaying prospective efficacy (though of variable duration) in fusion-positive patients (10). Similarly, future therapeutic advances will likely stem from further molecular stratification of iCCA patients, exposing oncogenic networks against which therapeutics may be strategized.

Recently, several cross-platform integrative analyses have identified broad cholangiocarcinoma subgroups associated with specific clinical and molecular features (11–13). However, inter-study comparison of these subgroups remains difficult given they are enriched by complex molecular characteristics identified in heterogeneous patient cohorts. Identification of single or few genomic alterations which influence downstream oncogenic networks could facilitate refinement of these previously established subgroups, as well as enhance prediction of therapeutic response to pharmacologic intervention against specific molecular backgrounds. As such, mutations in *IDH1* and *IDH2* have already been associated

with unique genomic features (DNA hypermethylation, altered chromatin remodeler expression, increased copy number and expression of mitochondrial genes), as well as sensitivity to dasatinib (14) and PARP inhibitors (15). The potential for individual driver mutations to influence diverse tumor networks is intriguing and remains to be evaluated for other recurrently mutated genes in iCCA, such as *KRAS*, *TP53* and chromatin remodeling enzymes (e.g., *ARID1A* and *BAP1*).

Accordingly, we aimed to deconstruct the molecular programs of iCCA tumors associated with specific individual mutations. To this end, we grouped 496 patients into four subgroups where three were classified based on the most recurrent somatic mutations (*IDH*, *KRAS* and *TP53*) and the remaining based on wild-type status for these genes (referred to as ‘undetermined group’). This stratification approach unveiled distinct mutational signatures and co-mutations, pathway-enrichment and pharmacologic response profiles between subgroups, as well as suggesting an enhanced role of structural and epigenomic alterations in undetermined group tumors. These findings uncover the potential of individual mutations to induce substantial downstream molecular heterogeneity which could facilitate prediction of therapeutic sensitivities for iCCA patients using standard targeted genotyping.

MATERIALS AND METHODS

Patients and samples

Data sets from a series of 496 iCCA patients were obtained for analysis (Supporting Fig S1A). Totally, 165 fresh-frozen iCCAs were collected from surgically resected patients. All samples were immediately frozen in liquid nitrogen and stored at -80°C pending genomic DNA isolation. The Danish Regional Ethics Committee and local authorities at contributing institutions approved the study. Written informed patient consent was obtained in accordance to regional regulation. All patient demographic and clinico-pathological data were anonymized and are summarized in Supporting Tables S1–S2. 150 of these tumors were subjected to targeted exome sequencing (TES) (Table S1), performed using the 48-gene TruSeq Amplicon Cancer Panel (TSACP, Illumina). The remaining 15 cases underwent whole-exome sequencing (WES) (Table S2) and were analyzed alongside 262 publically available WES iCCA samples (Table S3).

SNV calling and analysis

FASTQ files were filtered, subjected to quality control and mapped to the human genome assembly (hg19) using BWA. SNVs and short indels were called using a combination of Varscan (v2.3), somaticSnipper (v1.0.4) and Shimmer. Annovar was used to classify SNVs as synonymous or non-synonymous. Potential impacts of amino acid alterations on protein structure and function were assessed by PolyPhen-2. Motif enrichment analysis was performed by HOMER. Inference of mutational signatures was carried out with SomaticSignatures tool. Pathway enrichment of significantly mutated genes encoding SNVs and indels were assessed using genome MuSiC. Further information regarding bioinformatics analysis can be found as Supporting Information.

***FGFR2* fusion detection**

Totally, 122 of 150 tumors subjected to TES were assessed for known *FGFR2* fusions. RNA isolation was performed and subjected to reverse transcription with SuperScript VILO cDNA Synthesis kit (ThermoFisher Scientific). Fusion products in *FGFR2* were targeted with primer pairs flanking breakpoints of *FGFR2* and previously reported donor genes. Commercially synthesized DNA with *FGFR2* and all known fusion partner sequence fragments (Integrated DNA Technologies) were cloned into plasmid scaffolds and used as positive controls. Fusions were confirmed by Sanger sequencing. Further information regarding fusion detection can be found as Supporting Information.

Cell culture

Cholangiocarcinoma cell lines were purchased from KCLB (SNU-478, SNU-1079, SNU-1196), DSMZ (EGI-1), RIKEN (HuCCT-1, RBE), and obtained from Mayo Clinic (WITT, KMCH). Cells were cultured in RPMI-1640 or DMEM-F12 (ThermoFisher Scientific), all supplemented with 10% heat-inactivated FBS (GE Healthcare), 1% penicillin-streptomycin and 1% L-glutamine (ThermoFisher Scientific). Cells were maintained at 37°C and 5% CO₂.

High-throughput drug repositioning

Drug screening was carried out using a drug library of 525 compounds (FDA-approved or in late stage clinical trials). Cells were seeded into 384-well plates and treated for 72 hours over 5-fold dilution series. Anti-proliferative cellular effects were measured with CellTiter-Glo Luminescent Cell Viability assay (Promega), as per manufacturers' instructions, and quantified using Drug Sensitivity Score (DSS) metric (16). Further information regarding drug screening can be found as Supporting Information.

CNA calling and analysis

Copy Number Alterations (CNAs) were analyzed using VarScan2, DNACopy and GISTIC2 packages. *FGFR2* fusions were predicted using BreakDancer. Further information regarding CNA analysis can be found as Supporting Information.

Genome-wide DNA methylation profiling

In house and publically available Infinium Human Methylation 450k BeadChip (Illumina) data were used to identify an independent patient cohort (n=69) to investigate effects of classifier genes on DNA methylation profiles. Data were processed and analyzed using RnBeads (v1.2.1). Further information regarding methylomics can be found as Supporting Information.

RESULTS

Rational identification of classifier mutations for iCCA patient stratification

To investigate mutational diversity in iCCA, we analyzed 427 patients by targeted exome sequencing ('TES-cohort'; n=150 patients) and whole-exome sequencing ('WES-cohort'; n=277 patients; incorporating 142 tumor/normal matched pairs and 135 tumors) (17–21))

(Supporting Fig. S1A–B). Analysis of TES-cohort of 48 cancer-related genes (Table S1) revealed *IDH1* (n=21/150; 14%), *TP53* (n=19/150; 13%) and *KRAS* (n=18/150; 12%) as the most recurrent mutations (Fig. 1A; Supporting Fig. S2A). Subsequent analysis of WES-cohort (Table S2–3) corroborated our TES-cohort findings where *TP53* (n=57/277; 20%), *KRAS* (n=44/277; 16%) and *IDH1/IDH2* (n=31/277; 11%) were most prevalent (Fig. 1B; Supporting Fig. S2B). Mutations in *IDH1* (R132C/G/L) and *KRAS* (G12A/C/D) were recurrent hotspots, while *TP53* mutations were distributed throughout its entire gene body (Fig. 1C) with prevalent R249S (n=10/45; 22.2%) mutations. Importantly, co-occurrence of *IDH1/KRAS* mutations (TES-cohort: n=1/150; 0.7%; WES-cohort: n=1/277; 0.4%) and *IDH1/TP53* (TES-cohort: n=3/150, 2%; WES-cohort: n=1/277, 0.4%) was rare, suggesting mutual exclusivity (Fig. 1A–B). In comparison, dual mutations in *KRAS* and *TP53* are detected in a subset of patients (TES-cohort: n=4/150; 2.7%; WES-cohort: n=14/277; 5.1%) but are not mutually inclusive ($P=0.07$, Fisher's exact test). Accordingly, we selected to stratify iCCA patients into four groups based on mutational status of these genes, referring to them as IDH-gr (including *IDH1* and *IDH2* mutations), *KRAS*-gr, *TP53*-gr and Udt-gr ('undetermined', wild type for all 3 driver genes).

FGFR2 fusion events have been identified in between 5.5% (21) to 28% (10) of patients, implicating this event as one of the most recurrent genomic alterations in iCCA. We next analyzed the prevalence of fusions in the TES-cohort by performing targeted analysis against known fusion partners, followed by confirmation with Sanger sequencing. We observed that 14.8% (n=18/122) of tumors were *FGFR2-BICC1* fusion positive, where 83.3% (n=15/18) cases fused with *BICC1* in exon 3 and the remaining 16.7% (n=3/18) fused in exon 5 (Supporting Fig. S2A; Supporting Fig. S2C). Clinically, *FGFR2* fusion-positive status was associated with shorter overall survival ($P<0.0001$; Supporting Fig. S2D). The majority of fusion events (n=12/18; 66.7%) fell within Udt-gr. Similarly, in previously reported *FGFR2* fusions from RNAseq data (n=117/277; 42.2%; Japanese patients), *FGFR2* fusions are predominant in Udt-gr (n=3/4; 75%) (Supporting Fig. S2B).

Focusing on tumor/normal matched cases (n=142/277; 51.3%), we annotated 7,921 nonsynonymous SNVs in 5,457 genes (Table S4) and 2,581 synonymous SNVs in 2,229 genes (Table S5) with an average of 2.3 mutations/Mb (Fig. 1D). Apart from the prevalent mutant classifiers (*IDH*, *KRAS* and *TP53*), other significantly mutated genes include chromatin modifiers (*ARID1A*, *BAP1*, *KDM6B*, *SETD2*) and genes newly associated with iCCA (Fig. 1E). We next analyzed known clinical risk factors associated with mutational subgroups and observed hepatitis B virus (HBV) infection was significantly associated with *TP53*-gr (n=26/45; 58%; $P=1.612\times10^{-5}$, Fisher's exact test) (Fig. 1E). Interestingly, HBV was also associated with a large number of Udt-gr tumors (n=21/63; 33%) but did not reach significance when compared to the entire cohort (Fig. 1E). However, no significant association was determined between mutational subgroups and diabetes mellitus (DM), smoking or alcohol consumption.

Mutation-centric tumor stratification identifies diverse oncogenic processes

The mutational signatures defined by immediate 5' - and 3' -sequences flanking SNVs may reflect specific underlying mutagenic processes (22); thus, we investigated mutagenic

signatures associated with the four subgroups. Though C>T transitions and C>A transversions are prevalent overall in iCCA (Supporting Fig. S3A), as previously described (20, 21), mutational signatures are unique between subgroups (Fig. 2A). IDH-gr displayed a specific enrichment of C>A transversions (67%, $P=7.01 \times 10^{-132}$) (signature 18, 28) (Fig. 2A–B). As such, C>A transversions have been associated with oxidative stress (23), complementing previous studies characterizing the physiological role of wild type *IDH1* in protecting hepatocytes from oxidative stress (24). KRAS-gr was characterized by predominant C>T transitions (Fig. 2A) (55%, $P=3.40 \times 10^{-87}$, Fisher's exact test), displaying a preference for purine (G or C) upstream of transitions (Fig. 2B), and suggesting a role for APOBEC (signature 2 and 13) (22). Also, promoters of genes mutated in KRAS-gr were enriched for *FOSL2* motif binding sites (Supporting Fig. S4), in agreement with transcriptional signature derived from *KRAS*-driven tumors (25). TP53-gr displayed preferential unique T>A transversions (26%, $P=2.29 \times 10^{-125}$) with NpT>ApG associated with aristolochic acid (signature 22). Although, a significant enrichment of R249S mutations were detected in TP53-gr (Fig. 1C), aflatoxin signature was not enriched as in hepatocellular carcinoma (HCC). Also, R249S did not display a significant relationship with HBV status ($P=0.48$, Fisher's exact test) or overall survival ($P=0.917$) when comparing with patients manifesting other *TP53*-specific mutations.

Akin to IDH-gr, Udt-gr displayed a preference for C>A transversions (42%, $P=4.69 \times 10^{-10}$). In fact, non-matrix factorization (NMF) clustering of signature contributions across mutation subgroups in patients and CCA cell lines ($n=16$) implicated a close association between IDH-gr and Udt-gr (Supporting Fig. S3B–C). Such similarity was also observed at the transcriptome level in our previously published cohort (26) through hierarchical clustering of significant differentially expressed genes (F-test $P<0.001$; Supporting Fig. S3D). Taken together, these findings suggest that without patient stratification, Udt-gr would have masked the unique genetic patterns elicited by IDH-gr, KRAS-gr and TP53-gr. Furthermore, these data indicate that diversity of oncogenic programs (unique and shared between subgroups) may arise early in tumorigenesis from their initial genetic mutational signatures and possibly transcend different genomic readouts in their consequences.

Inter-group mutational analysis uncovers unique co-mutational profiles

Next, we analyzed the extent of exclusivity of mutated genes among subgroups and observed that the majority of genes encoding nonsynonymous SNVs were unique (Fig. 3A) and co-occurred (2 tumors) within their individual subgroups (Fig. 3B–C; Table S6A–D). Genes encoding somatic synonymous SNVs were also unique to each subgroup (Supporting Fig. S5A; Table S6E), but show minimal overlap (2–4%) (Supporting Fig. S5B) with nonsynonymous SNV genes, suggesting these mutational processes are dissimilar. Furthermore, somatic synonymous SNVs have low recurrence rate and lack significantly mutated genes with iCCA subgroups (Supporting Fig. S5C), which is markedly different from nonsynonymous mutated genes (Fig. 3C).

As expected, *IDH1*, *KRAS* and *TP53* were the most significantly altered genes in their respective groups along with specific enrichment of *BCLAF1* ($P=0.007$) in IDH-gr; *SMAD4* ($P=0.03$) in KRAS-gr; *PTEN* ($P=9.36 \times 10^{-7}$), *RB1* ($P=0.002$) and *LATS2* ($P=0.02$) in the

TP53-gr; and *KDM6B* in the Udt-gr ($P=0.08$) (Fig. 3C; Table S7). Among the previously reported chromatin modifier genes, *ARID1A* mutations were present in IDH-gr, KRAS-gr and Udt-gr but only significantly enriched with Udt-gr tumors ($P=8.86 \times 10^{-8}$), while *BAP1* was preferentially associated with IDH-gr ($P=2.57 \times 10^{-8}$). Notably, no chromatin modifiers were associated with TP53-gr. Association of significantly mutated genes along with the three classifier genes are represented in Fig. 3D.

We next evaluated any prognostic significance across four mutational subgroups. In the WES-cohort of 142 tumor/normal patient samples, KRAS-gr and TP53-gr were significantly associated with poor overall survival ($P=0.003$; Fig. 3E) and faster time to recurrence ($P=0.018$; Fig. 3F). These findings were confirmed in two independent cohorts whereby KRAS-gr was associated with worse overall survival in TES-cohort (N=150) ($P=0.0004$; Fig. 3G) and both KRAS-gr and TP53-gr were correlated with worse overall survival ($P=0.0001$; Fig. 3H) in Nakamura et al. cohort (N=135) (21). Thus, we conclude that these four mutational subgroups have a distinct association of co-mutated genes and different overall survival.

Intrahepatic CCA subgroups harbor unique dysregulated pathways that are reflected in their drug-response profiles

To understand whether genes associated with each subgroup could contribute to different oncogenic programs through tailored pathway dysregulation, we analyzed perturbed pathways associated with each group (Fig. 4A–B; Table S8). IDH-gr was enriched for metabolic pathways, including glutathione metabolism ($P=6.33 \times 10^{-21}$) and citrate cycle ($P=2.37 \times 10^{-14}$), which is consistent with known involvement of *IDH1* in metabolic processes in CCA (27). KRAS-gr was highly enriched for various immune-related pathways, including ErbB ($P=4.39 \times 10^{-17}$), VEGF ($P=2.49 \times 10^{-16}$) and actin cytoskeleton rearrangement ($P=4.7 \times 10^{-10}$). In contrast, MAPK ($P=1.44 \times 10^{-33}$), WNT ($P=2.19 \times 10^{-26}$) and p53 signaling ($P=3.44 \times 10^{-32}$) were enriched in TP53-gr. Since mutations in *KRAS* and *TP53* occur in a small subset of patients, pathway enrichment in dual mutants was independently assessed. Notably, significantly enriched pathways were all present in either KRAS-gr or TP53-gr (Table S8), supporting our previous inclusion of these patients in both KRAS-gr and TP53-gr analyses. Finally, in spite of being the largest group, Udt-gr was exclusively enriched for mTOR signaling ($P=4.76 \times 10^{-3}$), suggesting that Udt-gr is highly heterogeneous with regards to SNVs. To indirectly test the validity of our classifier genes, we selected the next three most recurrent mutated genes as alternative classifiers: *ARID1A*-gr (n=14; 9.9%), *CDC27*-gr (n=12; 8.5%), *BAP1*-gr (n=9; 6.3%), and Remaining-gr (n=110; 77.5%; analogous to Udt-gr) (Supporting Fig. 6). No pathways were uniquely enriched among these surrogate groups, with significant biological processes emerging only in the ‘remaining’ tumors, thus justifying our patient classification.

Next, we sought to determine whether mutational subgroups displayed differential sensitivity to diverse classes of compounds *in vitro* and whether these drug responses reflected *in silico* predictions of enhanced pathway dysregulation. To determine the ‘best fit’ cell models, mimicking the patient groups, we earlier performed NMF clustering of the mutational signatures (genetic substitution patterns) in patients and a panel of 16 CCA cell

lines (Supporting Fig. S3B–C). This confirmed genetic signatures under our mutational classifiers were comparable between cell lines and patients.

Accordingly, we selected 7 cell lines with matched mutational profiles representative of each of the patient subgroups, as well as one mixed mutant cell line (RBE; *KRAS*^{G12V} and *IDH1*^{R132S}) not observed in patients (Table S9). Each cell line was subjected to high-throughput screening (HTS) using a drug library of 525 late-stage trial/FDA-approved compounds at clinically-achievable tissue concentrations (Fig. 4C). We classified compounds as biologically-active (Drug Sensitivity Score (DSS) > 10) or -inactive (DSS < 10) (16, 28) and then linked them to the mutational subgroup in which maximal therapeutic effect was observed (Fig. 4D; Table S9). RNA synthesis inhibitors ($P=0.029$, Fisher's exact test), microtubule-targeting drugs ($P=2.93 \times 10^{-4}$), topoisomerase inhibitors ($P=1.1 \times 10^{-3}$) and PLK1 inhibitors ($P=0.018$), and mTOR pathway modulators ($P=7 \times 10^{-6}$) displayed maximal efficacy in IDH-gr, KRAS-gr, TP53-gr and Udt-gr, respectively. These modalities largely support earlier observed subgroup-enriched pathway dysregulation (Table S8) as therapeutic vulnerabilities in iCCA. The top 21 most effective drugs with specific mutational subgroup-enrichment are depicted in Fig. 4E. Notably, presence of dual mutations appears to abrogate such observed sensitivities in the RBE cell line (mixed). Taken together, these experimental findings support the concept of preferential pathway deregulation, governed by our mutational sub-classification, which may have important implications for exploitation of pharmacological vulnerabilities.

Structural and epigenomic burden vary across iCCA patient subgroups

In SNVs, Udt-gr tumors showed limited enrichment of specific pathways. This suggests either substantial molecular heterogeneity within Udt-gr or that the causality is related to other genome features (gene fusion, recurrent somatic copy number alteration (SCNA) or epigenetics). We analyzed enriched SCNA segments (total counts) and observed IDH-gr has the lowest, followed by KRAS-gr, Udt-gr and TP53-gr (Fig. 5A; Table S10). Frequencies of SNV and SCNA segments exhibit an inverse trend (Fig. 5B), similar to observations in other cancers (29) with enrichment of DNA repair mismatch genes in cases with high mutational burden. As such, KRAS- and IDH-grs have few recurrent focal SCNAs, while TP53- and Udt-gr tumors have comparably higher frequency of recurrent structural variants (and encoded genes) (Supporting Fig. S7A; Table S10). Recurrent deletions were significant across all four groups, while recurrent amplifications were only significant in TP53- and Udt-gr (Fig. 5C). The majority of recurrent SCNA segments were unique to each group (Supporting Fig. S7B; Table S11). The only common cytoband 9p21.3 enriched across all groups was the focal segment that encodes the key tumor suppressor gene, *CDKN2B* (q-value reported by GISTIC2: 0.05 (IDH-gr), 0.08 (KRAS-gr), 4.85×10^{-5} (TP53-gr) 1.52×10^{-7} (Udt-gr)). We then examined whether recurrent SCNA segments were correlated with transposons, microsatellite instability (MSI) or HBV integrations but found no significant correlation (Supporting Fig. S7C–G).

Genes within focal SCNA peaks were highly recurrent (70–90% of patients) in their respective subgroups (Fig. 5D; Table S11), suggesting potential causal roles in their pathogenesis. We performed pathway analysis across each subgroup and identified IDH- and

TP53-grs as largely silent (Fig. 5E; Table S12), suggesting a diminished contribution of structural variants to their pathogenesis. In contrast, KRAS-gr showed enrichment of multiple immune-related processes (NK cytotoxicity, JAK/STAT and cytokine signaling), complementing our findings from SNVs of widespread immune pathway subversion in these patients (Fig. 4A). Similar enrichment of structural alterations in immune pathway genes (related to viral and/or bacterial infection) was also found in Udt-gr, a patient subgroup with a heterogeneous low SNV burden but high SCNA frequency.

To investigate epigenomic contributions to cholangiocarcinogenesis, we classified an independent iCCA patient cohort (n=69; Table S13) (11, 30) into our four mutational subgroups and compared differential DNA methylation profiles relative to surrounding normal tissues. Consistent with our earlier cohorts (Fig. 3E–H), KRAS-gr was again associated with decreased overall survival within 12 months following surgery (Fig. 6A). While all subgroups displayed predominant hypermethylation over hypomethylation across promoters and gene bodies, the total burden of differentially methylated regions (DMRs) varied between subgroups (Fig. 6B; Table S13). Unsurprisingly, given its role in DNA demethylation, IDH-gr displayed the greatest number of differentially methylated regions (DMRs) (n=2499/2646; 94.4% hypermethylation). Notably, Udt-gr harbored the next most extensive DNA methylome alterations. Such findings were also emphasized in pathway enrichment, where KRAS-gr and TP53-gr were largely silent, and Udt-gr was particularly enriched in neuronal-related signaling (Fig. 6C; Table S14). Given the fact that the copy number status has been reported to have little effect on DNA methylation alterations (31, 32), here we demonstrate an additional independent molecular patho-mechanism that may positively affect iCCA subtypes and influence tumor behavior as well as therapeutic response.

The most recurrent focal SCNA segment in Udt-gr includes the gene methyltransferase-like 13 (*METTL13*), which is amplified in 70% of patients (Fig. 5D). Comparison of its structural alteration across 36 tumor types (TCGA) revealed it to be most frequently amplified in CCA (Supporting Fig. S8A), correlating with elevated gene expression and DNA methylation (Supporting Fig. S8B–D). Given *METTL13*'s expected role as a SAM methyltransferase, we tested whether this patient subgroup (*METTL13* amp) was readily distinguishable at the epigenetic level, and established a DNA methylation signature that efficiently distinguishes it from surrounding normal tissues and its parent Udt-gr (Fig. 6D). Patients harboring *METTL13* amplification display greater overall survival ($P=0.00033$; Fig. 6E). KEGG pathway analysis of *METTL13* amplification-unique DMRs revealed enrichment of transcriptional control and kinase signaling pathways (Fig 6F). As such, *METTL13* amplification could potentially contribute to the elusive pathogenicity of Udt-gr tumors (Supporting Fig. S8).

DISCUSSION

Arguably the greatest obstacle in the development of therapeutic strategies for cholangiocarcinoma is its extensive molecular heterogeneity, further confounded by mixed patient cohorts of varied clinical and pathological features. Previous efforts have identified individual genomic events responsible for specific therapeutic sensitivities (e.g., *IDH*

mutations (14, 15) and *FGFR2* fusions (10)) and broad clusters of patients with diverse characteristic pathobiological properties (11–13). What remains unclear are the broader network implications of such individual events and the molecular origins of these clusters. Here, we highlight the potential of the three most recurrent SNVs (*IDH*, *KRAS*, *TP53*) to direct unique oncogenic programs in intrahepatic cholangiocarcinoma and, in turn, expose a subset of their network-based vulnerabilities (Fig. 7). Furthermore, we uncover enhanced roles of structural and epigenomic perturbation mechanisms in patients lacking mutations in these three classifier genes, reinforcing the need for integrative approaches to comprehensively stratify patients.

In both TES- (n=150) and WES-cohorts (n=277), the top three recurrently mutated genes were found to be *TP53*, *KRAS* and *IDH1*, as anticipated (17, 20, 21). Intriguingly, these 3 classifier genes are highly distinct with regards to their role in oncogenesis and, as highlighted in this study, in cholangiocarcinogenesis. Hotspot mutations in *KRAS* (G12A/C/D) play a significant role in iCCA initiation in mouse models (33) and are associated with a proliferation subgroup of patients with poor prognosis (26, 34). Mutational inactivation of tumor suppressor *TP53* contributes to this tumor formation *in vivo* (33) and was previously correlated with hypermutated patient tumors (21). Unlike *KRAS*, no mutational hotspot exists in *TP53* and different mutations in this gene have been shown to generate diverse neomorphic mutants with varied cellular effects (reviewed in (35)), indicating the potential for additional network heterogeneity within this subgroup. Functionally nonsynonymous hotspot mutations in *IDH1* (R132) and *IDH2* (R172) promote cholangiocarcinogenesis through repression of hepatocyte differentiation (36). However, mutation of these metabolic enzymes has broad cellular implications as aberrant downstream metabolites impair function of approximately 20 enzymatic families (only one of which are involved in DNA methylation regulation).

Given the diverse nature of these 3 mutational events, we anticipate diverse network consequences downstream. Prediction of sufficiency of 3 mutations to induce invasive carcinoma (37), however, reinforces the missing molecular pathogenicity garnered through use of single-gene dissection of tumors. Here, we have highlighted the unique co-mutational spectra associated with each of the classifier genes. Interestingly, while *KRAS* and *IDH* mutations are mutually exclusive, *KRAS* and *TP53* mutations coexist in a small subset of patients. Such mutations may be complimentary in a minority of patients given their opposing nature as oncogenes and tumor suppressors, or potentially if they arise at different times during cellular transformation. Indeed, *KRAS*- and *TP53*-gr tumors share the greatest similarity in altered pathways. Nonetheless, these findings remain unaltered if *KRAS/TP53* dual mutants are excluded, suggesting such pathway similarity arises from a common underlying biology (such as cell cycle regulation) rather than inclusion of co-mutants. Additionally, our patient stratification does not take into account, though may subtly reflect, the heterogeneous cellular origins of iCCA. For example, given recent reports of *IDH* mutant (11) and ‘*IDH*-like’ (38) HCC subgroups, this may indicate a putative common origin, further reasoning the continuous spectrum of liver tumors (39). Conversely, *KRAS* mutations are highly recurrent in pancreatic cancer (40) and perihilar CCA (pCCA), potentially suggesting a more likely origin from peribiliary glands.

Initially, a limitation of our genomic stratification appeared to be the relatively large proportion of patients included in Udt-gr (63% TES-cohort, 44% WES-cohort). However, follow-up enquiry of non-mutational mechanisms successfully clarified these cryptogenic tumors. Fusion events involving *FGFR2* are predominant in Udt-gr, identifying a specific fusion-based patient group among these tumors (Fig. 1A–B). More generally, we observed SCNA to inversely correlate with SNV and, therefore, potentially compensate for the lack of SNV-associated genetic and pathway defects in these tumors. This agrees with pan-cancer studies that have identified trade-offs between mutational (‘M-class’) and structural/copy number (‘C-class’) (29), as well as large-scale pan-cancer pharmacogenomics studies that identify different contributions from diverse ‘cancer functional events’ (including mutation, structural alteration and DNA methylation) (41). Specifically, *METTL13* amplification was detected in more than 70% of the Udt-gr tumors with pan-cancer analysis revealing it to be most recurrent in cholangiocarcinoma. Previous studies have reported *METTL13* amplification in HCV-associated hepatocellular carcinoma (42), conflicting with its downregulation in bladder cancer (43). Furthermore, its family member (*METTL3*) was recently uncovered to exert neoplastic properties in acute myeloid leukemia, in part through its role in RNA methylation (44), indicating *METTL13* merits further investigation in iCCA.

To date, clinical trials in cholangiocarcinoma have suffered from both anatomical and molecular heterogeneity reflected in the limited translational and clinical progress made in the past 40 years (45, 46). Here, *in silico* analysis of significantly enriched pathways was successful in identifying broad signaling networks that, through patient stratification, could be exploited to induce enhanced growth arrest *in vitro*. This is further supported by convergent pathway-targeting by different mechanisms (e.g. metabolic subversion by SNV and DNA methylation in IDH-gr, immune insult by SNV and CNA in KRAS-gr). Importantly, lack of specific mutations did not dictate a complete absence of therapeutic response and similarly, co-mutations appeared to be capable of influencing therapeutic benefit (as evident in the dual *KRAS* and *IDH1* mutant RBE cell line). Nonetheless, our stratification successfully identified pathways most responsive to broad classes of repositioned drugs. In conclusion, this study demonstrates the pharmacogenomic potential of mutation- and pathway-driven genetic dissection of tumors. This highlights the potential for driver-gene guided oncogenic network characterization to enable efficient stratification and therapy optimization for intrahepatic cholangiocarcinoma patients, as well as other cancers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project was supported by the Danish Medical Research Council, Danish Cancer Society, Novo Nordisk Foundation, and A.P. Møller Foundation to JBA. CN is supported by the Danish Medical Research Council individual postdoctoral fellowship. CJO and DVO are funded by Marie Skłodowska-Curie postdoctoral fellowships. AT is supported by the Lundbeck Foundation postdoctoral fellowship. XWW was supported by the Intramural Research Program of the Center for Cancer Research, National Cancer Institute (NCI), NIH. JBA thanks the ATC and CCR Genomics Core at the NCI for support with whole-exome and target amplicon sequencing.

References

1. Banales JM, Cardinale V, Carpino G, Marzioni M, Andersen JB, Invernizzi P, Lind GE, et al. Expert consensus document: Cholangiocarcinoma: current knowledge and future perspectives consensus statement from the European Network for the Study of Cholangiocarcinoma (ENS-CCA). *Nat Rev Gastroenterol Hepatol*. 2016; 13:261–280. [PubMed: 27095655]
2. Yao KJ, Jabbour S, Parekh N, Lin Y, Moss RA. Increasing mortality in the United States from cholangiocarcinoma: an analysis of the National Center for Health Statistics Database. *BMC Gastroenterol*. 2016; 16:117. [PubMed: 27655244]
3. Sirica AE, Gores GJ. Desmoplastic stroma and cholangiocarcinoma: clinical implications and therapeutic targeting. *Hepatology*. 2014; 59:2397–2402. [PubMed: 24123296]
4. Razumilava N, Gores GJ. Cholangiocarcinoma. *Lancet*. 2014; 383:2168–2179. [PubMed: 24581682]
5. Claessen MM, Vleggaar FP, Tytgat KM, Siersema PD, van Buuren HR. High lifetime risk of cancer in primary sclerosing cholangitis. *J Hepatol*. 2009; 50:158–164. [PubMed: 19012991]
6. Watanapa P, Watanapa WB. Liver fluke-associated cholangiocarcinoma. *Br J Surg*. 2002; 89:962–970. [PubMed: 12153620]
7. Palmer WC, Patel T. Are common factors involved in the pathogenesis of primary liver cancers? A meta-analysis of risk factors for intrahepatic cholangiocarcinoma. *J Hepatol*. 2012; 57:69–76. [PubMed: 22420979]
8. Valle J, Wasan H, Palmer DH, Cunningham D, Anthoney A, Maraveyas A, Madhusudan S, et al. Cisplatin plus gemcitabine versus gemcitabine for biliary tract cancer. *N Engl J Med*. 2010; 362:1273–1281. [PubMed: 20375404]
9. Everhart JE, Ruhl CE. Burden of digestive diseases in the United States Part III: Liver, biliary tract, and pancreas. *Gastroenterology*. 2009; 136:1134–1144. [PubMed: 19245868]
10. Goyal L, Saha SK, Liu LY, Siravegna G, Leshchiner I, Ahronian LG, Lennerz JK, et al. Polyclonal Secondary FGFR2 Mutations Drive Acquired Resistance to FGFR Inhibition in Patients with FGFR2 Fusion-Positive Cholangiocarcinoma. *Cancer Discov*. 2016
11. Farshidfar F, Zheng S, Gingras MC, Newton Y, Shih J, Robertson AG, Hinoue T, et al. Integrative Genomic Analysis of Cholangiocarcinoma Identifies Distinct IDH-Mutant Molecular Profiles. *Cell Rep*. 2017; 18:2780–2794. [PubMed: 28297679]
12. Chaisaingmongkol J, Budhu A, Dang H, Rabibhadana S, Pupacdi B, Kwon SM, Forgues M, et al. Common Molecular Subtypes Among Asian Hepatocellular Carcinoma and Cholangiocarcinoma. *Cancer Cell*. 2017; 32:57–70. e53. [PubMed: 28648284]
13. Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, Padmanabhan N, Nellore V, et al. Whole-Genome and Epigenomic Landscapes of Etiologically Distinct Subtypes of Cholangiocarcinoma. *Cancer Discov*. 2017
14. Saha SK, Gordan JD, Kleinstiver BP, Vu P, Najem MS, Yeo JC, Shi L, et al. Isocitrate Dehydrogenase Mutations Confer Dasatinib Hypersensitivity and SRC Dependence in Intrahepatic Cholangiocarcinoma. *Cancer Discov*. 2016; 6:727–739. [PubMed: 27231123]
15. Sulkowski PL, Corso CD, Robinson ND, Scanlon SE, Purshouse KR, Bai H, Liu Y, et al. 2-Hydroxyglutarate produced by neomorphic IDH mutations suppresses homologous recombination and induces PARP inhibitor sensitivity. *Sci Transl Med*. 2017;9.
16. Yadav B, Pemovska T, Sz wajda A, Kuleskiy E, Kontro M, Karjalainen R, Majumder MM, et al. Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. *Sci Rep*. 2014; 4:5193. [PubMed: 24898935]
17. Chan-On W, Nairismagi ML, Ong CK, Lim WK, Dima S, Pairojkul C, Lim KH, et al. Exome sequencing identifies distinct mutational patterns in liver fluke-related and non-infection-related bile duct cancers. *Nat Genet*. 2013; 45:1474–1478. [PubMed: 24185513]
18. Gao Q, Zhao YJ, Wang XY, Guo WJ, Gao S, Wei L, Shi JY, et al. Activating mutations in PTPN3 promote cholangiocarcinoma cell proliferation and migration and are associated with tumor recurrence in patients. *Gastroenterology*. 2014; 146:1397–1407. [PubMed: 24503127]

19. Sia D, Losic B, Moeini A, Cabellos L, Hao K, Revill K, Bonal D, et al. Massive parallel sequencing uncovers actionable FGFR2-PPHLN1 fusion and ARAF mutations in intrahepatic cholangiocarcinoma. *Nat Commun.* 2015; 6:6087. [PubMed: 25608663]
20. Zou S, Li J, Zhou H, Frech C, Jiang X, Chu JS, Zhao X, et al. Mutational landscape of intrahepatic cholangiocarcinoma. *Nat Commun.* 2014; 5:5696. [PubMed: 25526346]
21. Nakamura H, Arai Y, Totoki Y, Shirota T, Elzawahry A, Kato M, Hama N, et al. Genomic spectra of biliary tract cancer. *Nat Genet.* 2015; 47:1003–1010. [PubMed: 26258846]
22. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, et al. Signatures of mutational processes in human cancer. *Nature.* 2013; 500:415–421. [PubMed: 23945592]
23. Besaratinia A, Kim SI, Bates SE, Pfeifer GP. Riboflavin activated by ultraviolet A1 irradiation induces oxidative DNA damage-mediated mutations inhibited by vitamin C. *Proc Natl Acad Sci U S A.* 2007; 104:5953–5958. [PubMed: 17389394]
24. Itsumi M, Inoue S, Elia AJ, Murakami K, Sasaki M, Lind EF, Brenner D, et al. Idh1 protects murine hepatocytes from endotoxin-induced oxidative stress by regulating the intracellular NADP(+)/NADPH ratio. *Cell Death Differ.* 2015; 22:1837–1845. [PubMed: 25882048]
25. Vallejo A, Perurena N, Guruceaga E, Mazur PK, Martinez-Canarias S, Zanduetta C, Valencia K, et al. An integrative approach unveils FOSL1 as an oncogene vulnerability in KRAS-driven lung and pancreatic cancer. *Nat Commun.* 2017; 8:14294. [PubMed: 28220783]
26. Andersen JB, Spee B, Blechacz BR, Avital I, Komuta M, Barbour A, Conner EA, et al. Genomic and genetic characterization of cholangiocarcinoma identifies therapeutic targets for tyrosine kinase inhibitors. *Gastroenterology.* 2012; 142:1021–1031. e1015. [PubMed: 22178589]
27. Borger DR, Tanabe KK, Fan KC, Lopez HU, Fantin VR, Straley KS, Schenkein DP, et al. Frequent mutation of isocitrate dehydrogenase (IDH)1 and IDH2 in cholangiocarcinoma identified through broad-based tumor genotyping. *Oncologist.* 2012; 17:72–79. [PubMed: 22180306]
28. Pemovska T, Kontro M, Yadav B, Edgren H, Eldfors S, Sz wajda A, Almusa H, et al. Individualized systems medicine strategy to tailor treatments for patients with chemorefractory acute myeloid leukemia. *Cancer Discov.* 2013; 3:1416–1429. [PubMed: 24056683]
29. Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, Sander C. Emerging landscape of oncogenic signatures across human cancers. *Nat Genet.* 2013; 45:1127–1133. [PubMed: 24071851]
30. Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, Padmanabhan N, Nellore V, et al. Whole-Genome and Epigenomic Landscapes of Etiologically Distinct Subtypes of Cholangiocarcinoma. *Cancer Discov.* 2017; 7:1116–1135. [PubMed: 28667006]
31. Feber A, Guilhamon P, Lechner M, Fenton T, Wilson GA, Thirlwell C, Morris TJ, et al. Using high-density DNA methylation arrays to profile copy number alterations. *Genome Biol.* 2014; 15:R30. [PubMed: 24490765]
32. Houseman EA, Christensen BC, Karagas MR, Wrensch MR, Nelson HH, Wiemels JL, Zheng S, et al. Copy number variation has little impact on bead-array-based measures of DNA methylation. *Bioinformatics.* 2009; 25:1999–2005. [PubMed: 19542153]
33. O'Dell MR, Huang JL, Whitney-Miller CL, Deshpande V, Rothberg P, Grose V, Rossi RM, et al. Kras(G12D) and p53 mutation cause primary intrahepatic cholangiocarcinoma. *Cancer Res.* 2012; 72:1557–1567. [PubMed: 22266220]
34. Sia D, Hoshida Y, Villanueva A, Roayaie S, Ferrer J, Tabak B, Peix J, et al. Integrative molecular analysis of intrahepatic cholangiocarcinoma reveals 2 classes that have different outcomes. *Gastroenterology.* 2013; 144:829–840. [PubMed: 23295441]
35. Freed-Pastor WA, Prives C. Mutant p53: one name, many proteins. *Genes Dev.* 2012; 26:1268–1286. [PubMed: 22713868]
36. Saha SK, Parachoniak CA, Ghanta KS, Fitamant J, Ross KN, Najem MS, Gurumurthy S, et al. Mutant IDH inhibits HNF-4alpha to block hepatocyte differentiation and promote biliary cancer. *Nature.* 2014; 513:110–114. [PubMed: 25043045]
37. Tomasetti C, Marchionni L, Nowak MA, Parmigiani G, Vogelstein B. Only three driver gene mutations are required for the development of lung and colorectal cancers. *Proc Natl Acad Sci U S A.* 2015; 112:118–123. [PubMed: 25535351]

38. Cancer Genome Atlas Research Network. Electronic address wbe, Cancer Genome Atlas Research N. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell*. 2017; 169:1327–1341. e1323. [PubMed: 28622513]
39. Marquardt JU, Andersen JB, Thorgeirsson SS. Functional and genetic deconstruction of the cellular origin in liver cancer. *Nat Rev Cancer*. 2015; 15:653–667. [PubMed: 26493646]
40. Cancer Genome Atlas Research Network. Electronic address aadhe, Cancer Genome Atlas Research N. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma Cancer. *Cell*. 2017; 32:185–203. e113.
41. Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, et al. A Landscape of Pharmacogenomic Interactions in Cancer. *Cell*. 2016; 166:740–754. [PubMed: 27397505]
42. Elsemman IE, Mardinoglu A, Shoaie S, Soliman TH, Nielsen J. Systems biology analysis of hepatitis C virus infection reveals the role of copy number increases in regions of chromosome 1q in hepatocellular carcinoma metabolism. *Mol Biosyst*. 2016; 12:1496–1506. [PubMed: 27040643]
43. Zhang Z, Zhang G, Kong C, Zhan B, Dong X, Man X. METTL13 is downregulated in bladder carcinoma and suppresses cell proliferation, migration and invasion. *Sci Rep*. 2016; 6:19261. [PubMed: 26763933]
44. Vu LP, Pickering BF, Cheng Y, Zaccara S, Nguyen D, Minuesa G, Chou T, et al. The N6-methyladenosine (m6A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. *Nat Med*. 2017; 23:1369–1376. [PubMed: 28920958]
45. Andersen JB, Thorgeirsson SS. A perspective on molecular therapy in cholangiocarcinoma: present status and future directions. *Hepat Oncol*. 2014; 1:143–157. [PubMed: 24772266]
46. Moeini A, Sia D, Bardeesy N, Mazzaferro V, Llovet JM. Molecular Pathogenesis and Targeted Therapies for Intrahepatic Cholangiocarcinoma. *Clin Cancer Res*. 2016; 22:291–300. [PubMed: 26405193]

Fig. 1. Genomic alteration landscape of iCCA patients. **(A)** Distribution of the three most prevalently mutated genes, alongside *FGFR2* fusion events, in targeted sequencing (TS)-cohort (n=150). **(B)** Distribution of the three most prevalently mutated genes and *FGFR2* fusion events in whole exome sequencing (WES)-cohort (n=277). **(C)** Lollipop plot of mutations reveal mutational hotspots in *IDH1* (R132C/G/L) and *KRAS* (G12A/C/D), while mutations in *TP53* are spread across multiple positions. **(D)** Total number of somatic nonsynonymous (SNVs and short indels) and synonymous mutations in WES-cohort (n=142, paired tumor/normal). **(E)** Mutational catalogue of 45 genes along with their significance score (reported by MuSiC), mutation frequency and select risk factors in WES-cohort (n=142, paired tumor/normal). The genes are sorted based on significance score.

Fig. 2.

Mutational spectra in iCCA across four groups. **(A)** Distribution of the 96 substitution patterns, defined by the adjacent 5' - and 3' -sequence of the mutated nucleotide, reveal distinct mutational signatures are preferred across groups. **(B)** Heatmap represents the frequency of 96 substitution patterns for each patient across four mutational subgroups. The numbers of mutations are scaled to 100 and color represents the percentage.

Fig. 3.

Comparison of genes encoding nonsynonymous mutations across four groups. **(A)** A large fraction of genes mutated across four subgroups are unique to each subgroup. **(B)** Distribution of recurrently mutated genes that are unique to each group. **(C)** Significantly mutated genes reported by genome MuSiC. Each bubble represents a gene. The size of each bubble is proportional to the gene's mutational frequency. Genes are sorted alphabetically along x-axis and y-axis indicates $-\log_2(P\text{-value})$. **(D)** Circos plot representing the interaction between three classifier genes and other significant genes (representing potential co-driver and passenger genes). The width of the ribbon represents the interaction count, where the

smallest width represents a single interaction. **(E–H)** Kaplan-Meier analysis reveals differences in recurrence rates and overall survival between mutational subgroups across 3 different patient cohorts.

Fig. 4.

Enrichment of distinct signaling pathways and pharmacogenomic consequences across four mutational subgroups. **(A)** KEGG analysis indicates preferential association of distinct pathways with mutational subgroups. Z-score are computed from MuSiC *P*-values of the pathways by using scale function across groups in R, followed by hierarchical clustering. **(B)** Connectivity of mutated genes across eight pathways. The numbers in boxes represent the percentage of patients in specific mutational subgroups. Different colors represent the four groups. **(C)** Schematic of high-throughput drug repositioning experiments. **(D)** Differential sensitivity of mutational subgroups to diverse compounds. Heatmaps represent mean Drug Sensitivity Scores (DSS) per mutational subgroup, as quantified by viability effects of biologically active compounds (DSS > 10) across 7 CCA cell lines. **(E)** 21 drugs targeting specific pathways and processes established as enriched in specific mutational subgroups.

Fig. 5. Contribution of somatic copy number alterations (SCNAs) across four mutational subgroups. **(A)** TP53-gr and Udt-gr have higher amplification and deletion frequencies compared to IDH-gr and KRAS-gr. **(B)** Frequencies of SNV and SCNA are inversely correlated. **(C)** Significant recurrent amplifications (red) and deletions (blue) reported by GISTIC2. X-axis on the bottom represents q-values of false discovery rate (FDR) and x-axis on top defines G-score (represents amplitude of aberration and frequency of occurrences across samples). Black horizontal line indicates FDR of 0.1. Y-axis represents the chromosomes. **(D)** The frequency of copy number gain and loss of genes across four subgroups. Each bubble represents a gene and the size of bubble indicates the total frequency of gain and loss. **(E)** Inter-group pathway-association of genes encoded in recurrent SCNA segments.

Fig. 6. Genome-wide DNA methylation analysis of mutational subgroups. **(A)** Kaplan-Meier analysis of DNA methylation cohort. **(B)** Hierarchical clustering of top 1% most variable probes. Each mutational subgroup is represented by geometric mean of probe methylation within that group. Number of DMRs are scaled against most epigenetically dysregulated group. **(C)** KEGG pathway analysis of total DMRs by patient subgroup. Significance was set at Q-value < 0.05. **(D)** *METTL13* amplification DNA methylation signature compared to Udt-gr and surrounding normal tissue. Probes are unique to amplification positive patients compared to surrounding normal with minimum methylation difference of 0.2 to Udt-gr. **(E)** Kaplan-Meier analysis of *METTL13* amplification positive and negative patients in TCGA

cohort. **(F)** KEGG pathway analysis of *METTL13* amplification positive DMRs. Black bars indicate Q-value, grey bars indicate P-value. DMR: Differentially Methylated Regions.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 7. Mutation-centric diversity of cholangiocarcinogenesis. Incidence of SNV, recurrent SCNA and differentially methylated regions across groups are shown as a relative ratio (scaled to 1). Integrating 'omic' and pathway analysis, confirmed by *in vitro* drug testing, suggests enhanced activity of specific compounds in specific mutational subgroups.