

Computational Analysis of Cholangiocarcinoma Phosphoproteomes Identifies Patient-Specific Drug Targets



Shirin Elizabeth Khorsandi^{1,2,3}, Arran D. Dokal^{4,5,6}, Vinothini Rajeeve^{4,5}, David J. Britton^{4,5,6}, Megan S. Illingworth^{2,3}, Nigel Heaton¹, and Pedro R. Cutillas^{4,5,6,7}

ABSTRACT

Cholangiocarcinoma is a form of hepatobiliary cancer with an abysmal prognosis. Despite advances in our understanding of cholangiocarcinoma pathophysiology and its genomic landscape, targeted therapies have not yet made a significant impact on its clinical management. The low response rates of targeted therapies in cholangiocarcinoma suggest that patient heterogeneity contributes to poor clinical outcome. Here we used mass spectrometry-based phosphoproteomics and computational methods to identify patient-specific drug targets in patient tumors and cholangiocarcinoma-derived cell lines. We analyzed 13 primary tumors of patients with cholangiocarcinoma with matched nonmalignant tissue and 7 different cholangiocarcinoma cell lines, leading to the identification and quantification of more than 13,000 phosphorylation sites. The phosphoproteomes of cholangiocarcinoma cell lines and patient tumors were significantly correlated. MEK1, KIT, ERK1/2, and several cyclin-dependent kinases were among the protein kinases

most frequently showing increased activity in cholangiocarcinoma relative to nonmalignant tissue. Application of the Drug Ranking Using Machine Learning (DRUML) algorithm selected inhibitors of histone deacetylase (HDAC; belinostat and CAY10603) and PI3K pathway members as high-ranking therapies to use in primary cholangiocarcinoma. The accuracy of the computational drug rankings based on predicted responses was confirmed in cell-line models of cholangiocarcinoma. Together, this study uncovers frequently activated biochemical pathways in cholangiocarcinoma and provides a proof of concept for the application of computational methodology to rank drugs based on efficacy in individual patients.

Significance: Phosphoproteomic and computational analyses identify patient-specific drug targets in cholangiocarcinoma, supporting the potential of a machine learning method to predict personalized therapies.

Introduction

Cholangiocarcinoma is the second most common primary hepatic malignancy, originating from cholangiocytes in the biliary tract (1). Global incidences and mortality rates are growing annually with prognosis remaining poor; most patients die within a year of diagnosis, irrespective of treatment modality (2). At present, the only chance of “curative” treatment is surgical resection when the disease is diagnosed

at an early stage, but even then, 5-year survival is less than 20% (3). Progress has been made in understanding the mutational landscape of cholangiocarcinoma, but the translation of this new knowledge into effective therapies has been slow (4). Mutations are most frequently found on genes that are common to other solid tumors, such as receptor tyrosine kinases (e.g., FGFR1–3, ERBB2, c-MET, EGFR), intracellular signaling molecules (e.g., KRAS, BRAF/ARAF, PI3KCA), tumor suppressors (e.g., TP53, SMAD4, CDKN2A/B), and epigenetic modifiers [e.g., ARID1A, BAP1, MLL3, isocitrate dehydrogenase 1/2 (IDH1/2); refs. 4–8].

Chemotherapies such as gemcitabine plus cisplatin are widely accepted as the standard of care (2, 9). Clinical trials of targeted therapies have shown that some patients respond well to such treatments, but most do not, even after stratification using genetic markers. As a recent example, the pan-FGFR inhibitor BGJ398 (Infigratinib) showed a 15% response rate (RR) in an unselected population compared with 19% in patients with alterations in their FGFR2 gene. Even for those patients that initially responded, acquired resistance limited efficacy (10, 11). Similarly, Pemigatinib, a recently approved selective inhibitor of FGFR1, 2, and 3 achieved 35.5% RR in patients with FGFR2 fusions/rearrangement (12). Other examples of genetic-based stratification of targeted therapies include trials that assess IDH inhibitors, CDK4/6 inhibitors, BH3 mimetics (targeting MCL1), PARP/ATM inhibitors, immune checkpoint inhibitors, and epigenetic modifiers (4–8, 13–16). While it is clear that stratification using genetic markers can improve RRs in the selected population, more needs to be done to improve understanding of the molecular alterations in each subtype, in order to reveal more effective drug targets and identify better predictive biomarkers for selection of the most effective targeted therapy for a given patient (6–8, 17).

Protein kinases are a large enzyme family of the human genome consisting of approximately 520 genes (18, 19). By catalyzing reversible

¹Institute of Liver Studies, Kings College Hospital, London, United Kingdom. ²The Roger Williams Institute of Hepatology, Foundation for Liver Research, London, United Kingdom. ³Faculty of Life Sciences & Medicine, King's College London, London, United Kingdom. ⁴Cell Signaling & Proteomics Group, Centre for Genomics & Computational Biology, Barts Cancer Institute, Queen Mary University of London, London, United Kingdom. ⁵Mass Spectrometry Laboratory, Barts Cancer Institute, Queen Mary University of London, London, United Kingdom. ⁶Kinomica Ltd, Cheshire, United Kingdom. ⁷The Alan Turing Institute, The British Library, London, United Kingdom.

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

Corresponding Authors: Pedro R. Cutillas, Cell Signaling & Proteomics Group, Centre for Genomics & Computational Biology, Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, United Kingdom. Phone: 207-882-5555; E-mail: p.cutillas@qmul.ac.uk; and Shirin Elizabeth Khorsandi, The Roger Williams Institute of Hepatology, 111 Coldharbour Lane, London SE5 9NT, United Kingdom. E-mail: shirin.khorsandi@kcl.ac.uk

Cancer Res 2021;81:5765–76

doi: 10.1158/0008-5472.CAN-21-0955

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 International (CC BY-NC-ND).

©2021 The Authors; Published by the American Association for Cancer Research

posttranslational phosphorylation of protein amino acid residues [mainly on serine (Ser), threonine (Thr) or tyrosine (Tyr)] protein kinases are able to regulate key signaling pathways that control cell replication, growth, metabolism, and death (20, 21). Kinase activity is commonly deregulated in cancer, through a combination of cellular events, for which the genetic component is important but not fully deterministic (22–24). Other molecular events, in addition to genetic alterations, such as epigenetic regulation of gene expression, signals from the microenvironment, and the activation status of phosphatases also contribute to kinase, and hence, oncogenic pathway activation. Consequently, stratification of kinase therapies based on genetic alterations has not been as effective as originally expected (25, 26). The hierarchy and global contribution of each kinase activity in primary tumors is yet to be fully established, but it is acknowledged that kinases determine cancer cell phenotype (27). A number of kinase inhibitors are clinically in use for cholangiocarcinoma or are in development (28) and identifying tumors addicted to the target kinase (by assessing not only its expression, but also its activation status), is expected to provide a means to stratify patients for therapy with greater accuracy than currently possible (29, 30).

Phosphoproteomics is an arm of proteomics that focuses on the characterization of phosphorylation events. The phosphorylation status of a protein represents the balance of kinases and phosphatases activities. Thus, computational methods have been devised to harness the information inherent in phosphoproteomics data to impute read-outs of net activation of kinase-driven oncogenic signaling pathways (31, 32). LC/MS-MS phosphoproteomics is now a routine technique in most proteomics laboratories with these methods enabling quantification of more than 10,000 phosphorylation sites per experiment (33, 34). The recent emergence of big data platforms, with the development of computational approaches that can infer responses to drug therapy such as kinase activities from phosphoproteomic data, is starting to present a means by which therapies can be recommended (35–38).

Clinical evidence indicates that it may be difficult to identify universal protein drug targets that may be suitable for all patients with cholangiocarcinoma (2, 4). Therefore, the aim of this work was to identify patient-specific drug targets in cholangiocarcinoma. To achieve this, we analyzed phosphoproteomic data from a panel of representative cholangiocarcinoma cell lines and surgically resected cholangiocarcinoma tissue. Computational methods were then used to analyze the phosphoproteomic data in order to identify overactive kinases and to apply a clinically relevant machine learning (ML) model of cholangiocarcinoma *in silico*. These computational models allowed us to identify kinase inhibitors and other drug classes as potential therapeutics for individual patients with cholangiocarcinoma. The present results illustrate the potential of ML in recommending therapies based on the cholangiocarcinoma phosphoproteome.

Materials and Methods

Clinical samples

Tumor (T) tissue samples were taken from the cholangiocarcinoma tumor mass, while matched background (B) liver tissue samples were distant to the tumor mass ($n = 13$). All tissue samples were snap frozen (-80°C) within 30 minutes of surgical resection. All patients gave written informed consent. The study was institutionally approved and samples were accessed via King's Liver BioBank (NHS Health Research Authority, Research Ethics Committee 08/H0704/117). Additional nonconfidential clinical infor-

mation such as subtype, tumor stage, gender, and recurrence are summarized in Supplementary Table S1.

Cell lines, compounds, and viability assay

Extrahepatic cholangiocarcinoma (EH-CCA) cell lines EGI-1 (CVCL_1193), TFK-1 (CVCL_2214), and CCC-5 (CVCL_LM83) were purchased from the DSMZ-German Collection. MMNK-1 (CVCL_M266) a benign cholangiocyte cell line, IH-cholangiocarcinoma (IH-CCA) cell lines KKU-213 (CVCL_M261) and OZ (CVCL_3118) were purchased from the Japanese Collection of Research Bioresources Cell Bank. IH-CCA cell lines HuCCT1 (CVCL_0324) & HuH-28 (CVCL_2955), and gallbladder cancer cell lines TGBC1TKB (CVCL_1769) & TGBC24TKB (CVCL_1770) were purchased from Cell Bank, RIKEN BioResource Research Centre. MCF10a (CVCL_0598) a benign epithelial cell line was obtained from G.Ficz (Barts Cancer Institute, QMUL). Before an experimental run *Mycoplasma* testing was performed (MycAlert, Lonza). Specific culture conditions are detailed in Supplementary Table S2.

CDK1 and CDK2 inhibitor AZD5438 and JAK2 inhibitor AZD1480 were from AstraZeneca. CDK4/6 inhibitor LY2835219 (Abemaciclib) was from Eli Lilly and Company and the MEK1/2 inhibitor trametinib from GlaxoSmithKline. Otherwise, all test compounds were purchased from Selleckchem, solubilized in DMSO and stored at -80°C . Controls for all experiments was vehicle (DMSO) alone. Cellular viability was determined using a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT; Sigma-Aldrich) colorimetric assay.

Tissue/cell preparation for mass spectrometry: lysis and protein digestion

Lysis buffer was prepared fresh on the day of tissue pulverization/lysis and consisted of 8 mol/L urea in 20 mmol/L HEPES (pH 8) supplemented with 1 mmol/L Na_3VO_4 , 1 mmol/L sodium, 1 mmol/L sodium β -glycerol phosphate, and 2.5 mmol/L $\text{Na}_2\text{H}_2\text{P}_2\text{O}_7$. Frozen tissue specimens weighing 0.05 to 0.8 g were pulverized on dry ice and liquid nitrogen using a MultiSample BioPulverizer (catalog no. 59012MS, BioSpec). Ground tissue samples were transferred to 2 ml Lo-bind Eppendorf tubes containing equal amounts of glass beads (Sigma G8772–100G) and 1,000 μL lysis buffer on ice. Samples were vortexed, then rotated at 4°C for 30 minutes to homogenize the samples. After lysis, samples were centrifuged at 4°C for 5 minutes at $15,000 \times g$. The lysate was transferred to a fresh 2 mL Lo-bind tube for sonification (Diagenode Bioruptor Plus) at 50% intensity, for 3 bursts of 15 seconds, resting 15 seconds between pulses. Finally, samples were spun at $15,000 \times g$, 4°C for 10 minutes, and the sample transferred to fresh 2 mL Lo-bind tubes. Protein concentration was determined using the Pierce bicinchoninic acid (BCA) assay (Thermo Fisher Scientific).

In the dark, 0.5 mg protein was reduced and alkylated by sequential incubation in 4 mmol/L dithiothreitol (DTT) and 8 mmol/L iodoacetamide (IAA), each for 30 minutes, in a heat block set to 25°C (shaking at 1,200 rpm). The urea concentration was reduced to 2 mol/L by the addition of 20 mmol/L HEPES (pH 8) for protein digestion. Then 100 μL of trypsin beads [50% slurry of TLCK-trypsin (Thermo-Fisher Scientific; catalog no. 20230)] conditioned with 3 washes of 20 mmol/L HEPES (pH 8) were added and the samples incubated for 16 hours at 37°C with agitation. Trypsin beads were removed by centrifugation at $2,000 \times g$ for 5 minutes at 5°C . The resultant peptide solutions were desalted on 10 mg Oasis HLB cartridges (Waters). For proteomics analysis peptides were eluted with 500 μL of acetonitrile (ACN) solution [30% ACN, 0.1% trifluoroacetic acid (TFA)], dried in a speed

vacuum (RVC 2–25, Martin Christ Gefriertrocknungsanlagen GmbH) and stored at -80°C .

Phosphorylated peptide enrichment

Enrichment of phosphorylated peptides was performed with TiO_2 (GL Sciences) as previously described, with some modifications (39). In brief, peptide eluents were normalized to 0.5 mL with glycolic acid buffer 2 (1 mol/L glycolic acid, 5% TFA, 80% ACN) and incubated with 25 μL of TiO_2 slurry (50% slurry in 1% TFA) for 5 minutes with end over end rotation, at room temperature and then centrifugation for 30 seconds at $1,500 \times g$. For each sample, 80% of the supernatant was transferred to a fresh tube and stored on ice. The remaining 20% was used to resuspend the bead pellets that were loaded into empty prewashed PE-filtered spin-tips (Glygen Corp) and packed by centrifugation at $1,500 \times g$ for 3 minutes. Spin tips were then sequentially washed with 100 μL of glycolic acid buffer 2, ammonium acetate buffer (100 mmol/L ammonium acetate in 25% ACN) and 10% ACN by room temperature centrifugation for 3 minutes at $1,500 \times g$. For phosphopeptide recovery, the addition 50 μL of 5% ammonium water followed by centrifugation for 5 minutes at $1,500 \times g$ was repeated four times. Eluents were snap frozen on dry ice, dried in a speed vacuum, and the peptide pellets stored at -80°C .

LC/MS-MS

For phosphoproteomics, peptide pellets were resuspended in 20 μL of reconstitution buffer (20 fmol/ μL of yeast enolase in 3% ACN, 0.1% TFA), sonicated in a water bath and then 5 μL of each sample loaded onto an LC/MS-MS system consisting of a Dionex UltiMate 3000 RSLC directly coupled to an Orbitrap Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific). For proteomics, pellets were resuspended in reconstitution buffer (0.5 $\mu\text{g}/\mu\text{L}$) and 2 μL were injected into the mass spectrometry (MS) equipment. The LC system used mobile phases A (3% ACN; 0.1% FA) and B (100% ACN; 0.1% FA). Peptides were trapped on a μ -precolumn (catalog no. 160454) at 10 μL per minute flow rate and separated on Thermo Scientific EASY-Spray LC Column (PepMapTM RSLC C18, 2 μm , 100 \AA , 75 $\mu\text{m} \times 50 \text{ cm}$; P/N ES803, S/N 10637404) set to a temperature of 40°C . The following parameters were used: 3% to 23% B gradient for 60 minutes and a flow rate of 0.3 $\mu\text{L}/\text{minute}$. Samples were run in the LC/MS-MS system in a randomized manner by shuffling samples before loading. As they eluted from the nano-LC system, peptides were infused into the online connected Q-Exactive Plus system operating with a 2.1 second duty cycle. Acquisition of full scan survey spectra (m/z 375–1,500) with a 70,000 full width at half maximum (FWHM) resolution was followed by data-dependent acquisition in which the 20 most intense ions were selected for higher energy collisional dissociation (HCD) and MS/MS scanning (200–2,000 m/z) with a resolution of 17,500 FWHM. A 30-second dynamic exclusion period was enabled with an exclusion list of a 10 parts per million (ppm) mass window. Overall duty cycle generated chromatographic peaks of approximately 30 seconds at the base, which allowed the construction of extracted ion chromatograms (XIC) with at least 10 data points.

Computational MS

Mascot Daemon 2.5.0 (Matrix Science Ltd) was used to automate peptide identification from MS data. Peak list files [Mascot Generic Format files (MGF)] from RAW data were generated with Mascot Distiller v2.5.1.0 and loaded into the Mascot search engine (v2.5) to match MS/MS data to peptides (40). The searches were performed against the SwissProt Database (uniprot_sprot_2014_08.fasta for phosphoproteomic analysis) with an FDR of approximately 1% and

the following parameters: 2 trypsin missed cleavages, mass tolerance of ± 10 ppm for the MS scans and ± 25 mmu for the MS/MS scans, carbamidomethyl Cys as a fixed modification, PyroGlu on N-terminal Gln, and oxidation of Met as variable modifications. For phosphoproteomic experiments phosphorylation on Ser, Thr, and Tyr was also included as variable modifications. Using the in-house developed Peak Statistic Calculator (Pescal), XIC for all the peptides identified across all samples were constructed with ± 7 ppm and ± 2 minutes mass and retention time windows, respectively. Peak areas from all XIC were calculated. Undetectable peptides were given an intensity value of 0. Values of three analytic replicates per sample were averaged and intensity values for each peptide were normalized to total sample intensity.

Bioinformatics and ML approaches

The first computational model applied, was Kinase Substrate Enrichment Analysis (KSEA) where activation of a given kinase pathway is inferred from phosphoproteomic data as previously described (27). While the second approach was Drug Ranking Using Machine Learning (DRUML; ref. 41). KSEA is an enrichment method based on statistical hypothesis tests that identifies kinase activities (measured as the phosphorylation of validated kinase substrates) significantly overrepresented in a given dataset compared with a background/reference dataset. In contrast, DRUML is an algorithm designed for prediction and generates a ranked list of drugs based on their efficacy to reduce cancer-cell proliferation. Of note, DRUML uses internally normalized distance metrics of drug response as input feature and is built from an ensemble of ML models trained on in house proteomic/phosphoproteomic data of cancer cells ($n = 48$) that have been validated with external data (4).

Hierarchical clusters were constructed within the R statistical computing environment (3.2.3) using the Euclidean distance metric in the heatmap2 package. EC_{50} were calculated in R from the cell viability data using the dose response curves package (W1.4 function). Statistical analysis was performed in R (version 3.2.3) or Microsoft Excel 2013. A paired, two-tailed Student t test was applied to assess significance in phosphoproteomic and proteomic data between matched samples. Normalized intensity values were assumed to follow a normal distribution. Groups were assumed to present similar variance and P values were adjusted for multiple testing using Benjamini–Hochberg procedures, while hypergeometric testing was used to calculate enrichment in KSEA analysis.

Results

Characterization of the cholangiocarcinoma phosphoproteome

We performed a phosphoproteomic analysis of cholangiocarcinoma cell lines and patient tumors using label-free LC/MS-MS (Fig. 1A). The study included 13 primary human T samples with matched adjacent B tissues (Fig. 1B) and 9 human cancer cell lines derived from cholangiocarcinoma ($n = 7$) and gallbladder cancers ($n = 2$; Fig. 1C). Clinical features of patients are shown in Supplementary Table S1. For comparison, 2 epithelial nonmalignant cell lines derived from human breast and cholangiocyte cells were also analyzed. The tumor and adjacent tissue samples were analyzed in triplicate, whereas cell lines were analyzed from two independent cultures in technical triplicate ($n = 2 \times 3$). Overall, 14,119 and 13,749 phosphorylation sites were identified (FDR < 0.02) in primary tumors and cell lines, respectively, leading to the generation of 2,008,716 quantitative data points. The principal component analysis (PCA) plots demonstrated grouping of samples according to cell of origin or anatomic

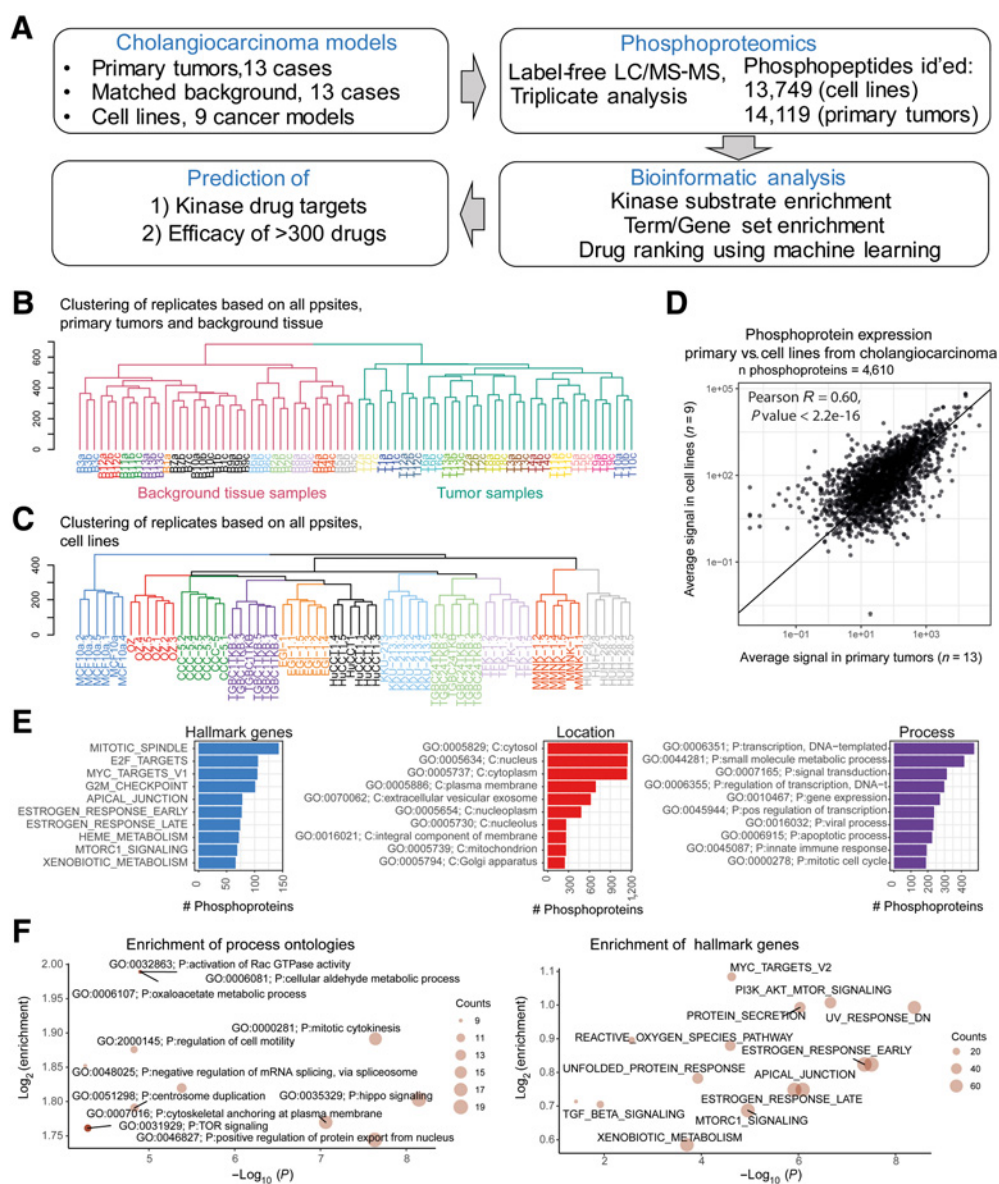


Figure 1.

Phosphoproteomics analysis of cholangiocarcinoma cell lines and patient tumors. **A**, Summary of workflow analysis for tumors and cell lines of patients with cholangiocarcinoma. **B**, Unsupervised hierarchical clustering based on normalization according to mean across all samples for protein phosphorylation sites (ppsites) identified in patient resected tissue. Dendrogram demonstrating clustering according to whether samples are replicates or from primary T/B liver. **C**, Unsupervised hierarchical clustering based on normalization according to mean across all samples for protein phosphorylation sites identified in cell lines. Dendograms demonstrating clustering according to whether sample replicate and cell type. **D**, Pearson correlation plot of phosphoprotein signals from cholangiocarcinoma cell lines ($n = 9$) versus patient cholangiocarcinoma tissue ($n = 13$) illustrating a high level of similarity between samples sets. **E**, Number of phosphoproteins in cholangiocarcinoma phosphoproteomes belonging to the named hallmark genes, gene ontology (GO) cellular location and process. **F**, Enrichment of gene ontology cellular process and hallmark genes in cholangiocarcinoma phosphoproteomes relative to the human Swissprot Uniprot proteome.

location of cholangiocarcinoma as illustrated in Supplementary Fig. S1A and S1B.

Unsupervised hierarchical clustering of all samples based on normalized phosphopeptide signals showed a distinct separation between the T and B primary cholangiocarcinoma tissues, with samples clustering by pathology (T or B, not by patient) and by technical replicate (Fig. 1B). Similarly, for the analysis of phosphoproteomes derived from cholangiocarcinoma cell lines, replicates grouped together by cell line identity in unsupervised hierarchical clustering analysis

(Fig. 1C). These results indicate that the label-free phosphoproteomics approach produced highly reproducible and high-quality quantitative data.

To compare the phosphoproteomes of cell lines and primary tissue samples, we summed the intensities of phosphopeptides derived from given proteins to generate quantitative values for 4,610 phosphoproteins. Comparison between averaged phosphoprotein signals derived from tumors and those generated from cell lines showed a strong correlation between the datasets (Pearson $R = 0.60$, $P = 1.2 \times 10^{-16}$), thus

revealing a surprisingly high similarity between the phosphoproteomes of cell lines and primary cholangiocarcinoma tissue samples (Fig. 1D). Mapping the identified cholangiocarcinoma phosphoproteome to hallmark genes, pathways, and ontologies highlighted that cholangiocarcinoma-derived phosphoproteins have roles in fundamental biological processes including signal transduction, mitosis, and gene expression among others and are annotated to be present in subcellular locations including the cytosol, plasma membrane, and nucleus (Fig. 1E). In order to identify highly represented pathways in the cholangiocarcinoma phosphoproteome, the data was compared with the Swissprot Uniprot proteome by enrichment analysis. Several pathways and gene sets were found to be enriched relative to the total proteome (at FDR-adjusted $P < 0.05$ by hypergeometric test) including known oncogenic processes such as Hippo signaling, mitotic cytokinesis, Myc targets, and PI3K/mTOR signaling (Fig. 1F). These results show that our phosphoproteomics data is of high quality; we provide these data as a community resource in the PRIDE repository of proteomics data (PXD027329).

Kinase substrate and pathway enrichment analysis of phosphoproteomics data identifies active kinases in cholangiocarcinoma

To identify differentially regulated kinases and pathways in cholangiocarcinoma relative to background tissue, we analyzed the phosphoproteomics data using pathway enrichment analysis and KSEA methods (27). KSEA considers validated kinase substrates or downstream targets to derive values of enrichment associated to the activation of kinases. This analysis led to the identification of 55 kinases with at least 10 substrates in tumor samples and 66 in cell lines. Of these, frequently activated kinases in tumors relative to background and in cancer cell lines included MEK1 (gene name MAP2K1), KIT, ROCK1/2, ERK1/2 (MAPK1/3), ABL2, and several cyclin-dependent kinases such as CDK1, CDK2, CDK4, and CDK5 (Fig. 2A and B). Enrichment analysis against pathways and hallmark genes further showed an increase in oncogenic processes such as JAK/STAT signaling, angiogenesis, RAS signaling, and mitotic gene sets relative to noncancer cells (Fig. 2C and D). Representation of the data in heatmap where pathways, kinase activities, and patients are separated by hierarchical clustering, uncovered differences in the extent of pathway activation across patients and cell lines (Fig. 2E and F), indicating that there is a high degree of biochemical heterogeneity across patients.

To confirm the results of KSEA and ontology enrichment analysis, the data was interrogated for known markers of kinase activation by comparing absolute phosphorylation signals (normalized to total phosphorylation signal) across sample groups. Thus, these quantitative values of protein phosphorylation, which are here termed ppIndex, represent the proportion of phosphorylation signal relative to total measured phosphorylation within a sample, expressed in ppm units. Relative phosphorylation across individual samples is shown in Supplementary Fig. S2. Phosphorylation of ERK1/2 (gene names MAPK1 and MAPK3) and MAP2K2 (MEK2) were increased in primary tumors and cell lines relative to background (Fig. 3A), in agreement with the result of KSEA. Additionally, phosphorylation of PRAS40 (AKT1S1) at S183, a marker of mTOR signaling activity, was increased, as were some phosphorylation sites on 4EBP1 (Fig. 3B). While the JAK substrate STAT3 at Y705 was decreased in tumors relative to background (Fig. 3B). Also, consistent with KSEA, phosphorylation of CDK1, CDK7, and their substrates were increased in tumors and cell lines relative to background (Fig. 3C). Thereby, our analysis uncovered kinases commonly activated in cholangiocarcinoma (Fig. 3A–C) as

well as patient-specific overactive signaling pathways in cholangiocarcinoma primary tumors and cell line models (Fig. 2E and F).

Identification of synergistic kinase inhibitor combinations

To confirm that the activation of kinase-driven pathways in cholangiocarcinoma cells is functionally relevant, we next investigated whether the observed increase in CDK1/2, CDK4/6, and MAPK activities across the cholangiocarcinoma primary tumors and cell lines over background tissue, translates to targeted drug sensitivity. To this end, we assessed the viability of a panel of cell lines treated with specific inhibitors of CDK1/2 (AZD5438), CDK4/6 (LY2835219), or MEK1/2 (trametinib) with a range of concentrations (10, 50, 100, 5000, 1,000 nmol/L) for 48 hours. The MEK inhibitor (MEKi) was the most potent compound in reducing cell viability for 3 of the 4 cell lines tested (Fig. 3D). KLU-213 viability was affected by CDK1/2 inhibition at concentrations above 1,000 nmol/L (Fig. 3E), whereas EGI-1 cells were the most sensitive to the CDK4/6 inhibitor (Fig. 3F). However, overall, the cholangiocarcinoma cell lines tested showed modest sensitivity to single kinase inhibitor treatment (Fig. 3D–F). Notably, we observed that these cells activated several pathways in parallel (Fig. 2) and showed elevated levels of phosphorylation in known substrates of the kinases that the selected compounds target (Fig. 3A). We therefore postulated that multiple pro-survival nodes within cancer cells may be providing compensatory signals that allow cells to escape treatments with single drugs, as observed in other cancer models (42).

To test this hypothesis, EGI-1 cells were treated with a panel of compounds as single agents or in combinations with selection based on the KSEA output. We also tested the JAK2 inhibitor AZD1480 as the JAK/STAT pathway was elevated in cholangiocarcinoma tissue (Fig. 2), although we also noted that the levels of the JAK substrate STAT3 pY705 in tumor tissues was lower relative to background tissue. We found that each of the treatments had morphologic effects on EGI-1 cells following 48 hours of treatment compared with control cells (Supplementary Fig. S3).

Viability assays confirmed that trametinib (MEKi, at 30 nmol/L) is the only kinase inhibitor that reduces EGI-1 viability at low dose. However, all combination treatments that included trametinib enhanced EGI-1 cell death compared with trametinib monotherapy (Fig. 3G). The combination of LY2835219 (CDK4/6i, at 50 nmol/L) with AZD5438 (CDK1/2i, 100 nmol/L) and AZD5438 (100 nmol/L)/trametinib (30 nmol/L) were particularly synergistic (combination index 0.5 and 0.55, respectively) (Fig. 3H), as assessed by the Bliss independence model (43). Additionally, triple inhibition with trametinib, LY2835219, and AZD5438 was synergistic with a combination index of 0.3. While the JAK inhibitor was ineffective in reducing cell viability by itself or in combination with trametinib, these results suggest that inhibition of several CDKs in combination with the MAPK pathway may be required for full inhibition of cholangiocarcinoma cell viability.

DRUML predicts drug efficacy in cholangiocarcinoma

In order to identify drugs, other than kinase inhibitors, that may be efficacious for cholangiocarcinoma, we reanalyzed the phosphoproteomic data using DRUML. DRUML is a recently developed algorithm, trained on the responses of cancer cell lines ($n = 48$) to 411 drugs with different modes of action, using an ensemble of ML models (41). The highest ranking drugs (i.e., predicted to be more efficacious by DRUML) for both tumors of patients with cholangiocarcinoma and cell lines included leptomycin b (an antibiotic), ouabain (Na/K-ATPase inhibitor), obatoclox mesylate (Bcl-2 inhibitor), panobinostat [HDAC inhibitor (HDACi)], belinostat (HDACi) and

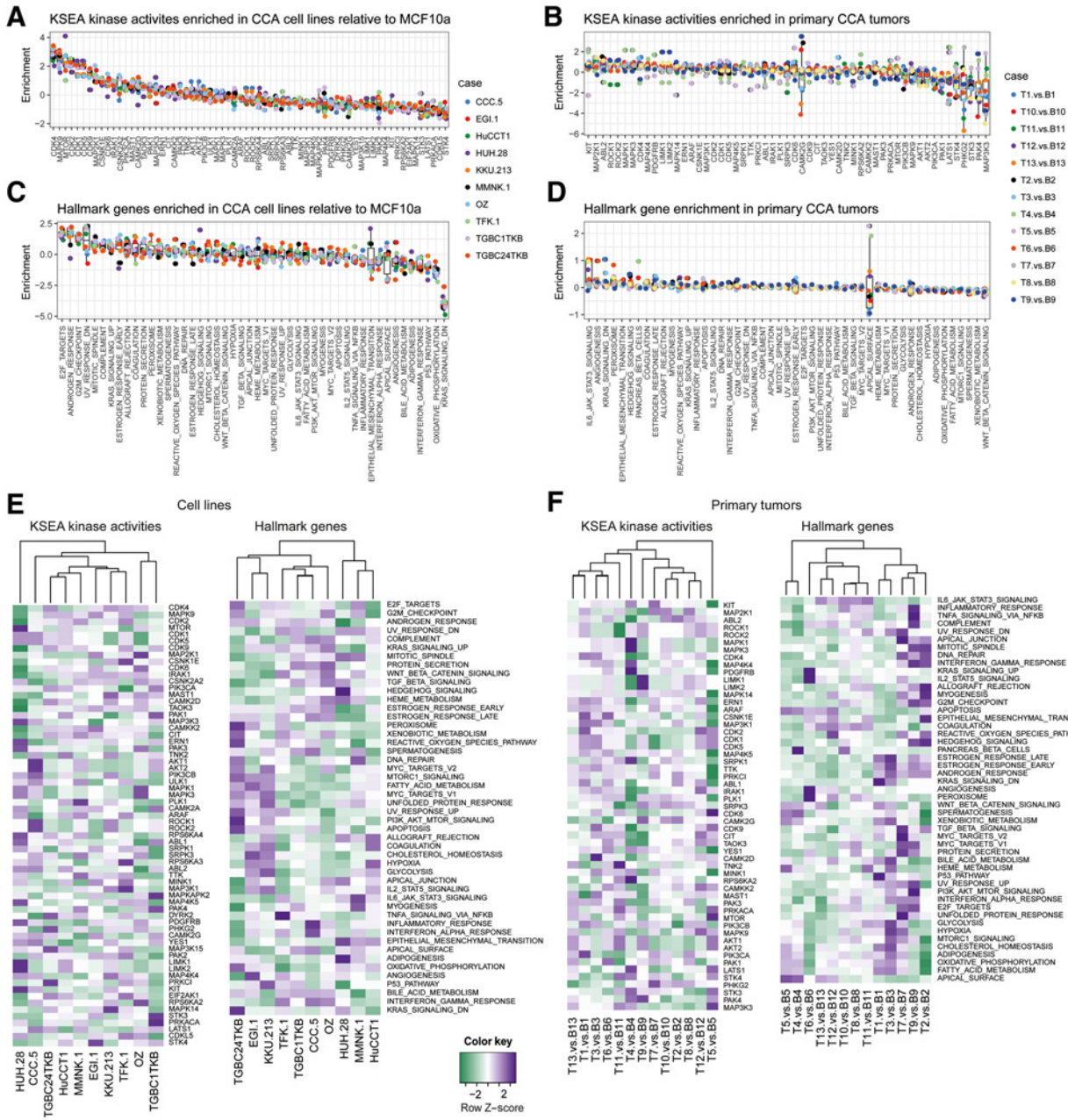


Figure 2. Heterogeneity of kinase activation and hallmark gene enrichment in cholangiocarcinoma cell lines and primary tumors. **A** and **B**, KSEA of phosphoproteomes from cholangiocarcinoma (CCA) cell lines relative to MCF10a (**A**) or from primary cholangiocarcinoma tissue relative to background (**B**). **C** and **D**, Pathway enrichment analysis of hallmark genes from cholangiocarcinoma cell lines relative to MCF10a (**C**) or from primary cholangiocarcinoma tissue relative to background (**D**). **E** and **F**, Hierarchical clustering and heat maps for KSEA inferred kinase activities (**E**) and pathway enrichment for hallmark genes (**F**) in the cholangiocarcinoma cell lines and cholangiocarcinoma primary tumors.

dinaciclib (CDK inhibitor; **Fig. 4A** and **B**). The predicted efficacy of drugs in cells lines and primary tumors correlated (**Fig. 4C**), suggesting that cell lines can be an appropriate model for ranking drugs based on DRUML predicted efficacy.

Toxicity to noncancer cells is a key factor that severely limits the clinical efficacy of drugs. We therefore explored the utility of identifying drugs with potentially a high therapeutic window by distinguishing those with high predicted efficacy in primary cholangiocarcinoma

relative to background tissue. We selected drugs with an average predicted area above the drug dose response curve (AAC) more than 0.2 in tumors and an overall greater efficacy in tumors relative to background (**Fig. 5A**). A total of 36 drugs passed these criteria (**Fig. 5B**). Among the drugs with greater difference in predicted efficacy were two HDACi (belinostat and CAY10603), metabolic inhibitors (fluvastatin, lovastatin, ML210), tubulin inhibitors (vinblastine and vinorelbine), and several kinase inhibitors including

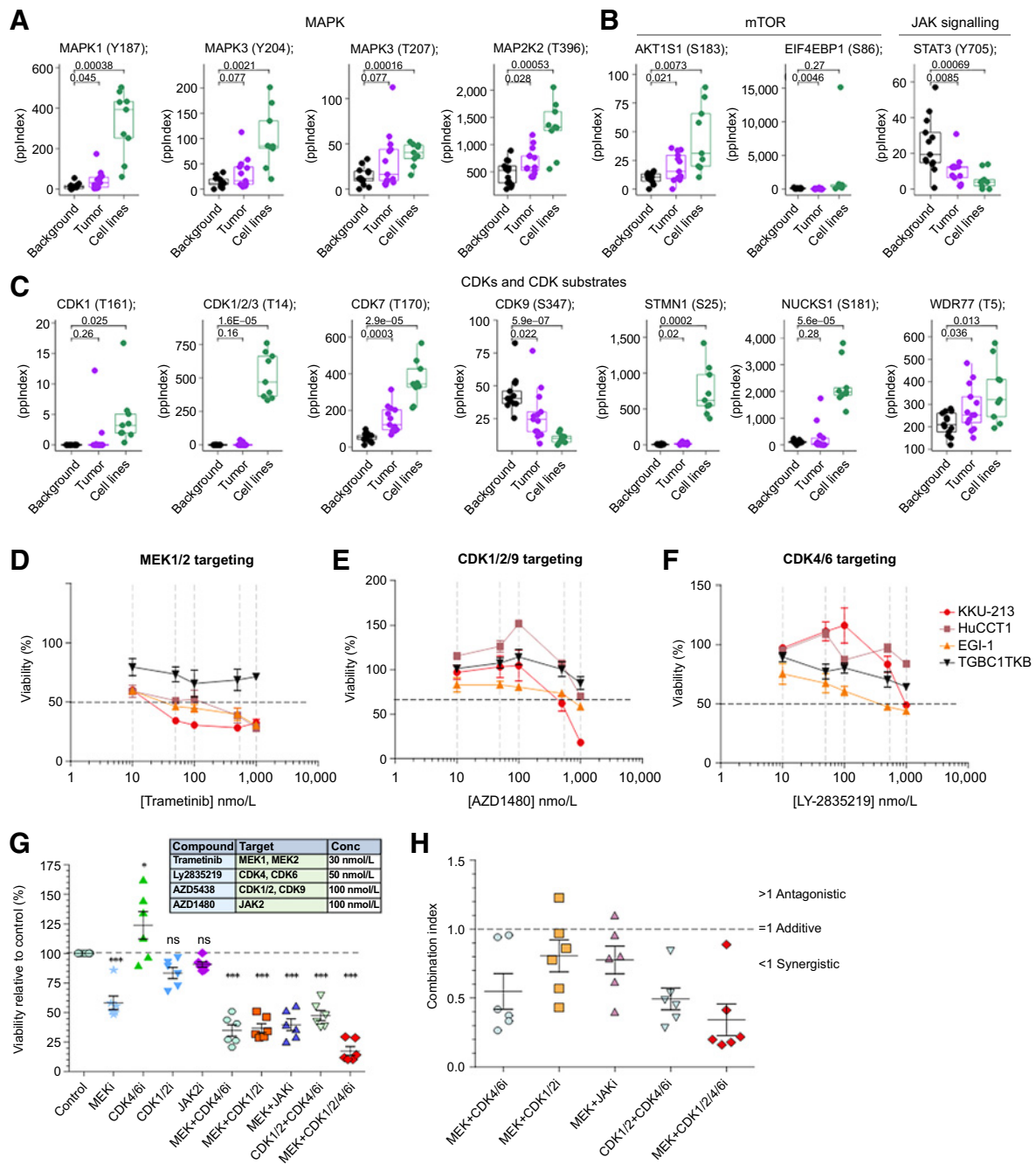


Figure 3. Markers of kinase pathway activation in cholangiocarcinoma and kinase inhibitor synergy. **A-C**, Selected markers of kinase activation in primary cholangiocarcinoma tissue, background tissue, and cholangiocarcinoma cell lines. ppindex, proportion of phosphorylation signal relative to total measured phosphorylation within a sample, expressed in ppm. **D-F**, Low-dose targeted combination therapies effectively induce cell death in EGI-1 cells. Cell survival assay in cholangiocarcinoma panel treated with 10, 50, 100, 500, 1,000 nmol/L of (i) MEK1/2 inhibitor trametinib, (ii) CDK4/6 inhibitor LY2835219, (iii) CDK1 inhibitor AZD5438. Y-axis represents percentage of viability. Viability was assessed using MTT assay and x-axis represents dosing of drug compound in nmol/L; logarithmic scale was used. Data points are mean \pm SEM ($n = 3$ per condition and duplicated). **G**, Cell survival assay in EGI-1 cells treated with low-dose single agents or treatment combinations as detailed in inserted table. Viability assessed using MTT assay; $n = 6$ per condition. ***, $P < 0.001$; ns, nonsignificant. **H**, Combination indexes for combination treatments, summarizing additivity, synergy, or antagonism of drug combination as defined by Bliss independence.

dasatinib, midostaurin (a multi-targeted kinase inhibitor), trametinib (a MEK inhibitor), and momelotinib (JAK inhibitor). Inhibitors of the PI3K/MTOR/AKT pathway were particularly well represented in this

set, with three drugs targeting PI3K rated in the top-ranking drugs for cholangiocarcinoma. Other PI3K/ATK/MTOR pathway inhibitors in this selected set included GSK2636771 (a PIK3CB inhibitor),

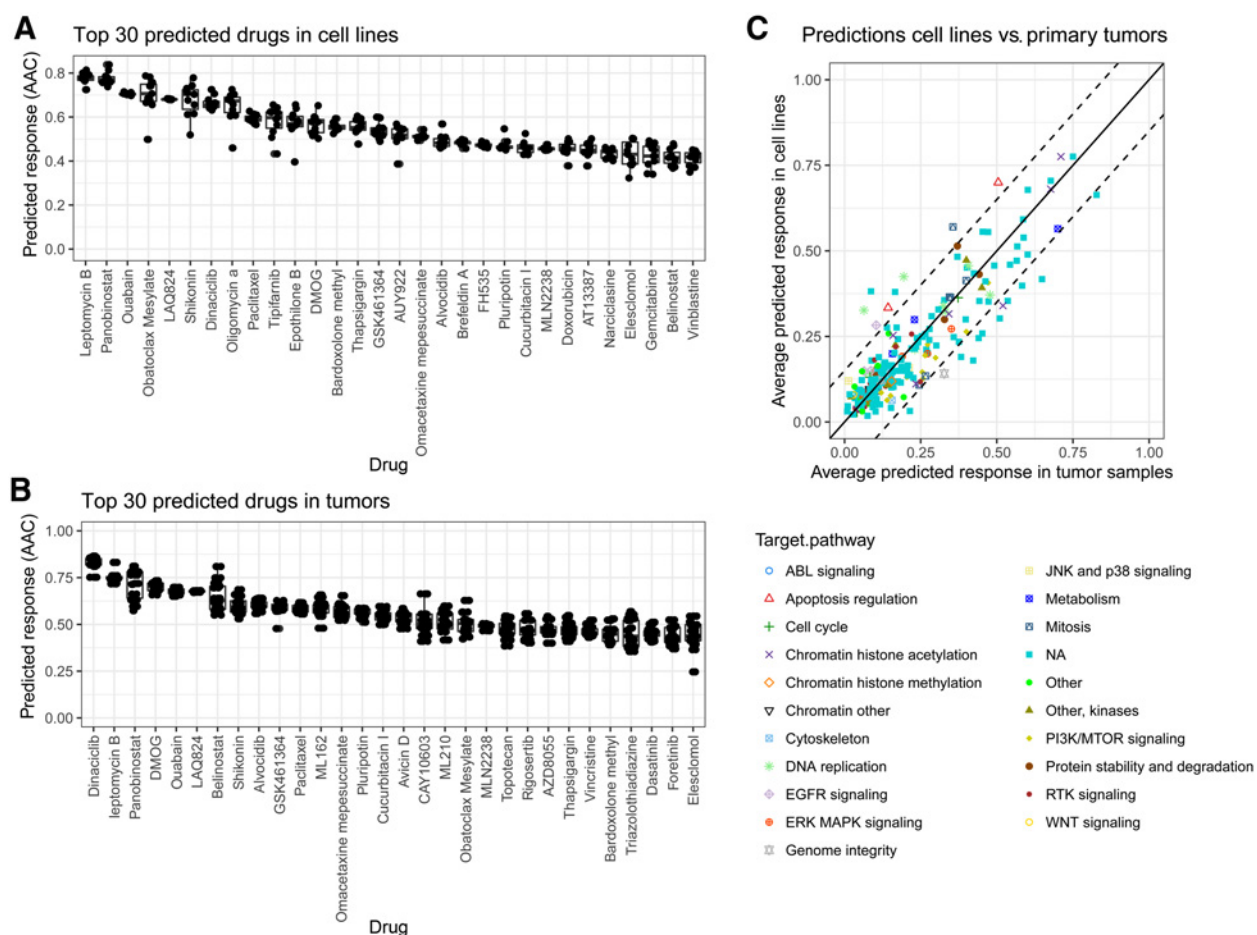


Figure 4. Drug ranking using machine learning (DRUML) prediction of drug efficacy in cholangiocarcinoma primary tumors and cell lines. **A**, Top 30 drugs predicted by DRUML to be efficacious in cholangiocarcinoma cell lines. **B**, Top 30 drugs predicted by DRUML to be efficacious in patient cholangiocarcinoma tissue. **C**, Correlation plot between DRUML-predicted highest ranking drugs in cholangiocarcinoma cell lines versus patient cholangiocarcinoma tissue, demonstrating a good correlation. AAC, average predicted area above the drug dose response curve.

rapamycin (which targets mTORC1), AS601245 (targets GSK3B), and AT6867 (whose intended targets are AKT and S6K; **Fig. 5B**). Ranking of drugs per patient showed that the HDACi belinostat was the highest ranked drug for all patients (**Fig. 5C**). CAY10603, another HDACi, was the second ranking drug in patients 1 and 5, while avicin D, a plant-derived triterpenoid known to induce apoptosis in cancer cells, was second ranking in 5 out of the 13 patient samples tested (**Fig. 5C**). Hierarchical clustering of the predicted drug response data highlighted patterns of drug responses specific for each patient. Together, these results uncover potential patient-specific drugs for cholangiocarcinoma.

Functional validation of DRUML-predicted drug sensitivity rankings

To validate the DRUML-predicted drug responses for the panel of cholangiocarcinoma cell lines, the predicted responses were compared with experimental drug responses obtained from repositories of drug-response data. There was a remarkably high correlation between predicted drug responses and experimental drug responses for cell lines ($n = 5$) present in the PharmacDB repository of drug-response data (accessed Oct 2020) for the selected drugs (**Fig. 6A**) with the

exception of the noncancer cell line MCF10A. The association was also significant for approved drugs and for those in different stages of clinical development (**Fig. 6B**). Overall, the correlation between predicted and measured drug responses was statistically significant when considering all drugs and those at different stages of clinical development (phase I–III; P values ranked from $1.6e-3$ to $1.3e-61$ by Spearman rank correlation; **Fig. 6C**). In addition, to further validate our DRUML rankings, we measured the viability of EH-CCA (EGI-1) and IH-CCA (HuCC1) cell lines treated with compounds found to be high ranked by DRUML (HDACi: belinostat and CAY10603) relative to a low ranked compound (Quizartinib, a tyrosine kinase inhibitor). In agreement with the correlation between predicted and actual responses across a large number of drugs (**Fig. 6**), the results were consistent with DRUML-derived rankings (Supplementary Fig. S4A–S4C) thus supporting the utility of DRUML in identifying drugs with high efficacy in decreasing cholangiocarcinoma proliferation.

Discussion

In addition to genomics, several other OMIC techniques now exist to enable oncologists, at least in principle, to make truly informed

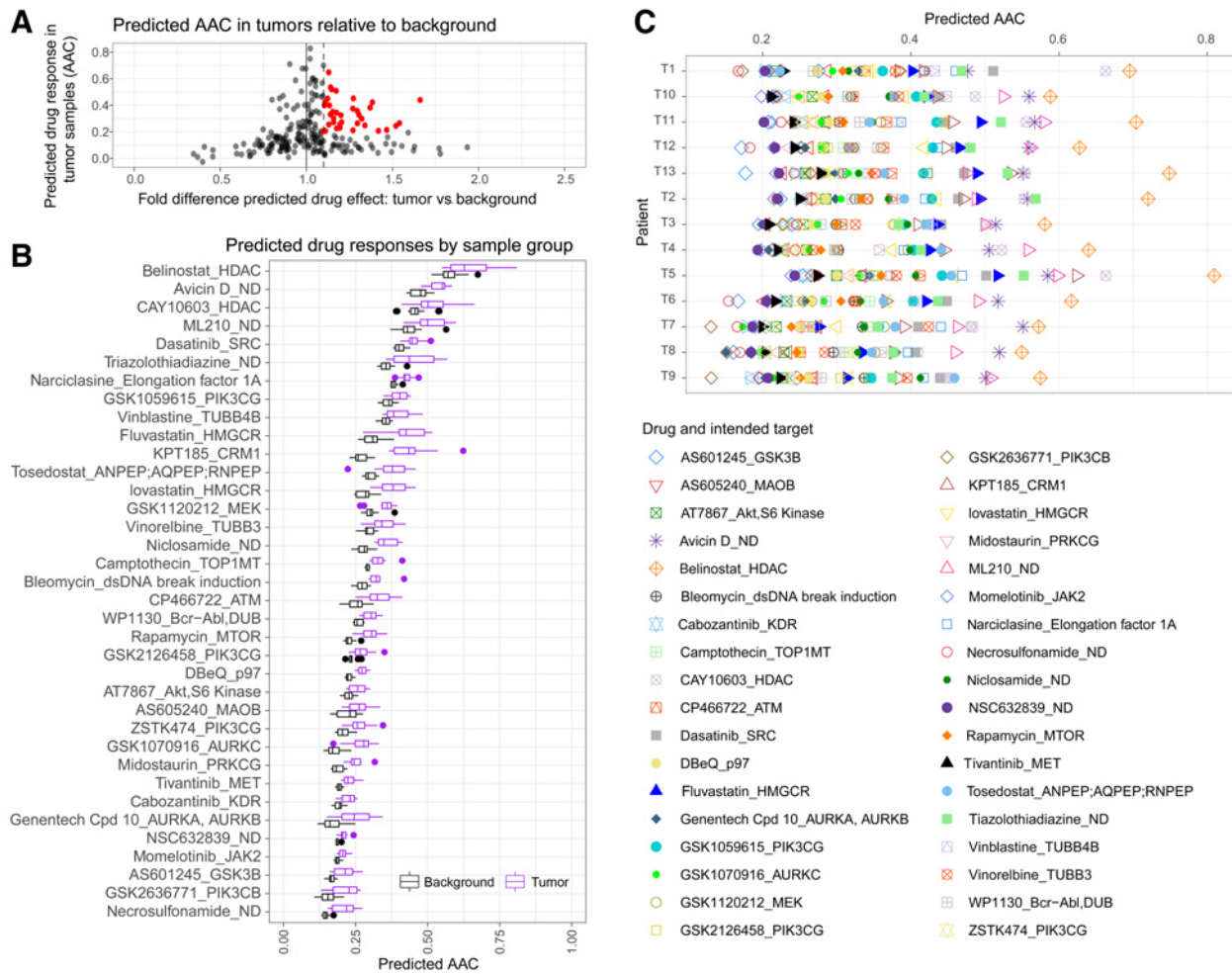


Figure 5. DRUML identification of drugs with predicted high therapeutic window. **A**, Comparison of predicted drug responses in cholangiocarcinoma and background tissue. Drugs with an average drug response (AAC) > 0.2 in the primary tumor and an overall greater efficacy in tumor tissue relative to background are shown in red. **B**, Drugs ($n = 36$) found by DRUML to have a predicted AAC > 0.2 in the tumor tissue, with at least 1.1-fold greater efficacy in the tumor tissue compared with background tissue. Boxplots show the median and interquartile ranges of predicted responses for the 13 patient samples. **C**, DRUML ranking of drugs according to each patient.

therapeutic decisions based on a patient’s cancer biology. These tools, which include proteomics and phosphoproteomics, have the potential to provide oncologists with large volumes of molecular data. However, finding actionable targets from OMIC datasets is challenging, as the link between molecular (genetic or proteomic) markers, phenotype, and pathophysiology is complex and not necessarily linear (27). To overcome this problem, computational methods based on statistics and ML have been developed in order to gain biological information from these complex datasets, as well as enable prediction of drug responses (27, 41, 44). Ultimately, the combination of OMIC technology with ML will advance the development of artificial intelligence tools for personalizing treatments (45).

Treating cholangiocarcinoma is challenging, with patient outcomes remaining poor despite the improvement in the understanding of the cellular mechanisms that underlie this complex disease (2, 4). This is in part related to the molecular heterogeneity of cholangiocarcinoma, meaning single-target therapeutics are mainly ineffective, even after patient stratification based on genetic markers (10–12). A deeper understanding of the cholangiocarci-

noma OMIC landscape will help develop novel biomarkers for disease treatment. This study both adds to the limited data on the cholangiocarcinoma phosphoproteome, as well as simultaneously exploring the clinical utility of computational algorithms, such as KSEA and DRUML, to recommend therapeutic strategies for the individual patient. This approach could potentially prove to be more effective than therapeutic stratification based on genetic alterations that are presently accepted clinically (25, 26). This view is consistent with the increasing realization of the contribution that nongenetic mechanisms make to disease progression and to therapeutic drug resistance in cancer (46).

Our work demonstrates that label-free MS phosphoproteomics is able to reliably distinguish cholangiocarcinoma samples according to pathology, whether that be patient-specific derived tumor tissue or cancer cell line (Fig. 1). To further explore the kinase pathways of therapeutic interest in cholangiocarcinoma we applied the bioinformatic pipeline KSEA (27), where validated kinase substrates or targets are used to infer state of kinase activation. Using this computational approach, the kinases that we identified as of importance in

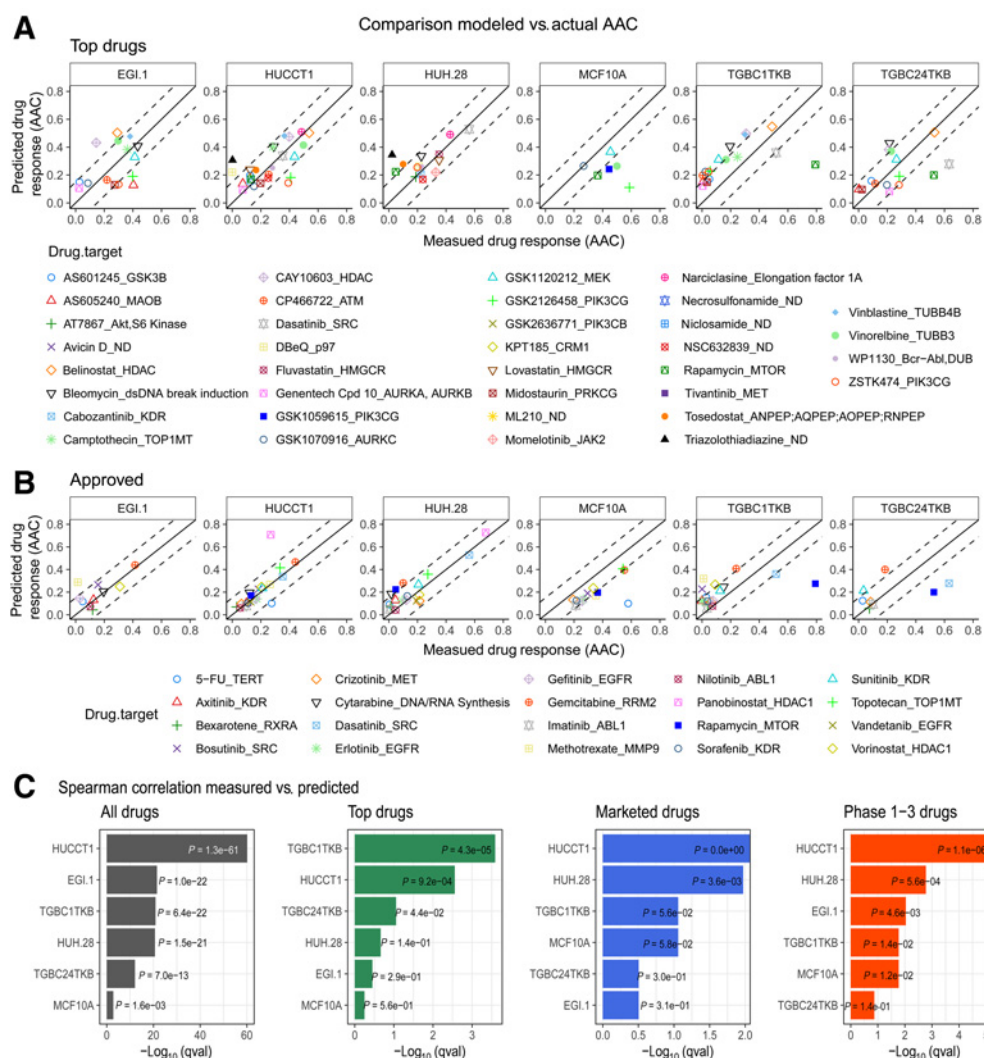


Figure 6. Validation of DRUML for drug prioritization in cholangiocarcinoma cellular models. **A**, Correlation plots demonstrating concordance between DRUML-predicted responses and all recorded drug responses in the PharmacDB repository. **B**, Correlation plots for DRUML-predicted responses and recorded drug responses in the PharmacDB repository for approved drugs. **C**, Spearman ranked correlation demonstrating concordance between DRUML predictions and recorded drug responses from PharmacDB as a function of drugs' stage of development. AAC, average predicted area above the drug dose response curve.

cholangiocarcinoma included MEK1 (gene name MAP2K1), KIT, ROCK1/2, ERK1/2 (MAPK1/3), ABL2, and several cyclin-dependent kinases such as CDK1, CDK2, CDK4, and CDK5 (Fig. 2). Although a number of these kinases have been cited as relevant in cholangiocarcinoma carcinogenesis, genomic assessments for treatment stratification have not always recapitulated the same importance for these druggable molecules (24).

The therapeutic benefit of directly targeting dysregulated kinase pathways was further assessed using cholangiocarcinoma cell lines treated with inhibitors for highly active kinases as selected by KSEA. Cholangiocarcinoma cell lines, despite their recognized limitations, have previously been demonstrated in genomic analysis to be robust cholangiocarcinoma disease models that recapitulate the genomic landscape of the tumors from which they originate (47). Similarly, we observed a high concordance between the phosphoproteomes of cholangiocarcinoma cell lines and patient cholangiocarcinoma tumor

samples (Fig. 1D), suggesting that cholangiocarcinoma cell lines reproduce cholangiocarcinoma biology and provide reasonable models for use in drug prediction validation.

Cell viability assays indicated that trametinib (MEK1/2 inhibitor) was the most effective drug across 3 of the 4 cholangiocarcinoma cell lines tested (Fig. 3). However, as prior studies have shown, inhibitors when used independent of other therapeutics have modest results, with trametinib currently moving to clinical trial in combination with hydroxychloroquine or dabrafenib (48). In our hands, kinase inhibitors were not effective on their own at inhibiting cholangiocarcinoma cell proliferation. As expected, an increase in treatment efficacy was seen when a combination of therapeutics was used, particularly the combination of abemaciclib/trametinib/AZD5438. These results suggest that several kinase-driven proliferative pathways exist in cholangiocarcinoma cells with compensatory mechanisms limiting the efficacy of treatment that target single pathways.

Due to the apparent limitation of kinase inhibitors and to further explore other therapeutic options for cholangiocarcinoma (including nonkinase inhibitors) we reanalyzed the cholangiocarcinoma phosphoproteomic data (patient tissue samples and cholangiocarcinoma cell lines) using a recently developed DRUML algorithm (41). An advantage of DRUML, in comparison with KSEA, is that the model was directly trained on drug response data and does not require comparison to a control or reference sample, as it internally normalizes distance metrics of a drug response. In contrast, KSEA infers kinase activities and assumes the extent of kinase activation to determine sensitivity of inhibitors against such kinase. Also, the association between kinase activation and therapeutic sensitivity to the respective kinase inhibitor may not hold true in tumors that activate several kinase-driven pathways in parallel with the potential to compensate for each other.

On the basis of DRUML (Fig. 5), the highest-ranking drugs predicted to have the greatest therapeutic effect included HDACi (belinostat and CAY10603), which have recently been reviewed as a future therapeutic for cholangiocarcinoma (49), metabolic inhibitors (fluvastatin, lovastatin, ML210), tubulin (vinblastine and vinorelbine), and several kinase inhibitors such as midostaurin, imatinib, and dasatinib, the latter being a drug that is currently in trial for cholangiocarcinoma (2). Notably, DRUML's predictions correlated with recorded drug responses in PharmacDB (Fig. 6). In addition, our cell viability assays demonstrated that the HDACi were effective (Supplementary Fig. S4). These initial findings are supportive of DRUML's ability to use phosphoproteomic data to recommend therapeutic suggestions of clinical relevance.

The application of computational approaches to make treatment recommendations based on the biological data of a given cancer holds great promise. Limitations of the algorithms used in the present work are mainly related to KSEA-derived kinase activities in tumors were measured relative to those in background liver, and several drugs in the DRUML algorithm have unknown modes of action, thus limiting the mechanistic information that maybe derived from these inhibitors. Furthermore, expansion of the training data for DRUML to include other clinically relevant cholangiocarcinoma drugs, as well as modeling the effect of stromal cells, will increase the utility of these *in silico* approaches for therapeutic decision making in cholangiocarcinoma. Future work will test the predictive power of these algorithms in other *in vitro* systems and *in vivo* models such as patient derived xenografts. Ultimately, clinical trials will be required to assess the power of this

technology as a multi-analyte companion diagnostic for multidrug prioritization.

In summary, treatment of cholangiocarcinoma is suboptimal. The computational approaches and data presented here, provide both insight into the heterogeneity of biochemical pathway activation in cholangiocarcinoma and proof-of-principle analysis of computational predictive modeling to advance personalized oncology in this disease.

Authors' Disclosures

S.E. Khorsandi reports grants from King's College Hospital Charity, Barts & The London Charity during the conduct of the study and grants from King's College Hospital Charity outside the submitted work. A.D. Dokal reports personal fees from Kinomica Ltd. outside the submitted work. D.J. Britton reports other support from Kinomica Ltd. outside the submitted work; in addition, D.J. Britton has a patent for DRUML pending and a patent for KSEA patent issued and licensed to Kinomica Ltd. M. Illingworth reports grants and personal fees from King's College Hospital Charity during the conduct of the study. N. Heaton is a shareholder in Kinomica Ltd. specializing in precision medicine research and diagnostics and cell signaling. P.R. Cutillas reports grants from Cancer Research UK during the conduct of the study; grants from Blood Cancer UK; and personal fees from Kinomica Ltd. outside the submitted work; in addition, P.R. Cutillas has a patent for DRUML pending. No disclosures were reported by the other authors.

Authors' Contributions

S.E. Khorsandi: Conceptualization, resources, funding acquisition, investigation, writing—original draft, project administration, writing—review and editing. A.D. Dokal: Data curation, formal analysis, validation, investigation, visualization, methodology, writing—original draft, writing—review and editing. V. Rajeev: Resources, investigation, methodology, project administration. D.J. Britton: Conceptualization, supervision, investigation, methodology, project administration, writing—review and editing. M.S. Illingworth: Data curation, validation, writing—review and editing. N. Heaton: Resources, funding acquisition, writing—review and editing. P.R. Cutillas: Conceptualization, resources, data curation, software, formal analysis, supervision, investigation, writing—original draft, writing—review and editing.

Acknowledgments

This work was supported by a donation Chris Lane/Ricoh Europe Challenge donation via King's College Hospital Charity and Barts and The London Charity.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 26, 2021; revised August 11, 2021; accepted September 20, 2021; published first September 22, 2021.

References

- Rizvi S, Gores GJ. Pathogenesis, diagnosis, and management of cholangiocarcinoma. *Gastroenterology* 2013;145:1215–29.
- Banales JM, Marin JGG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. *Nat Rev Gastroenterol Hepatol* 2020;17:557–88.
- Blechacz B. Cholangiocarcinoma: current knowledge and new developments. *Gut Liver* 2017;11:13–26.
- Valle JW, Lamarca A, Goyal L, Barriuso J, Zhu AX. New horizons for precision medicine in biliary tract cancers. *Cancer Discov* 2017;7:943–62.
- Chan-On W, Nairismägi ML, Ong CK, Lim WK, Dima S, Pairojkul C, et al. Exome sequencing identifies distinct mutational patterns in liver fluke-related and non-infection-related bile duct cancers. *Nat Genet* 2013;45:1474–8.
- Churi CR, Shroff R, Wang Y, Rashid A, Kang HC, Weatherly J, et al. Mutation profiling in cholangiocarcinoma: prognostic and therapeutic implications. *PLoS One* 2014;9:e115383.
- Farshidfar F, Zheng S, Gingras MC, Newton Y, Shih J, Robertson AG, et al. Integrative genomic analysis of cholangiocarcinoma identifies distinct IDH-mutant molecular profiles [published correction appears in *Cell Rep*. 2017 Jun 27;19:2878–2880]. *Cell Rep* 2017;18:2780–94.
- Nakamura H, Arai Y, Totoki Y, Shirota T, Elzawahry A, Kato M, et al. Genomic spectra of biliary tract cancer. *Nat Genet* 2015;47:1003–10.
- Valle JW, Furuse J, Jital M, Beare S, Mizuno N, Wasan H, et al. Cisplatin and gemcitabine for advanced biliary tract cancer: a meta-analysis of two randomised trials. *Ann Oncol* 2014;25:391–8.
- Javle M, Lowery M, Shroff RT, Weiss KH, Springfield C, Borad MJ, et al. Phase II study of BGI398 in patients with FGFR-altered advanced cholangiocarcinoma. *J Clin Oncol* 2018;36:276–82.
- Goyal L, Saha SK, Liu LY, Siravegna G, Leshchiner I, Ahronian LG, et al. Polyclonal secondary *FGFR2* mutations drive acquired resistance to FGFR inhibition in patients with *FGFR2* fusion-positive cholangiocarcinoma. *Cancer Discov* 2017;7:252–63.
- Abou-Alfa GK, Sahai V, Hollebecque A, Vaccaro G, Melisi D, Al-Rajabi R, et al. Pemigatinib for previously treated, locally advanced or metastatic cholangiocarcinoma: a multicentre, open-label, phase 2 study. *Lancet Oncol* 2020;21:671–84.

13. Ross JS, Wang K, Gay L, Al-Rohil R, Rand JV, Jones DM, et al. New routes to targeted therapy of intrahepatic cholangiocarcinomas revealed by next-generation sequencing. *Oncologist* 2014;19:235–42.
14. Nakagawa S, Okabe H, Sakamoto Y, Hayashi H, Hashimoto D, Yokoyama N, et al. Enhancer of zeste homolog 2 (EZH2) promotes progression of cholangiocarcinoma cells by regulating cell cycle and apoptosis. *Ann Surg Oncol* 2013;20: S667–75.
15. Martin-Liberal J, Ochoa de Olza M, Hierro C, Gros A, Rodon J, Taberero J. The expanding role of immunotherapy. *Cancer Treat Rev* 2017;54:74–86.
16. El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet* 2017;389:2492–502.
17. Hezel AF, Deshpande V, Zhu AX. Genetics of biliary tract cancers and emerging targeted therapies. *J Clin Oncol* 2010;28:3531–40.
18. Manning G, Plowman GD, Hunter T, Sudarsanam S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem Sci* 2002;27:514–20.
19. Hunter T. 1001 protein kinases redux—towards 2000. *Semin Cell Biol* 1994;5: 367–76.
20. Cohen P. Protein kinases—the major drug targets of the twenty-first century?. *Nat Rev Drug Discov* 2002;1:309–15.
21. Hunter T. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* 1995;80:225–36.
22. Lahiry P, Torkamani A, Schork NJ, Hegele RA. Kinase mutations in human disease: interpreting genotype-phenotype relationships. *Nat Rev Genet* 2010;11: 60–74.
23. Torkamani A, Schork NJ. Prediction of cancer driver mutations in protein kinases. *Cancer Res* 2008;68:1675–82.
24. Wan PT, Garnett MJ, Roe SM, Lee S, Niculescu-Duvaz D, Good VM, et al. Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* 2004;116:855–67.
25. McLeod HL. Cancer pharmacogenomics: early promise, but concerted effort needed. *Science* 2013;339:1563–6.
26. Relling MV, Evans WE. Pharmacogenomics in the clinic. *Nature* 2015;526: 343–50.
27. Casado P, Rodriguez-Prados JC, Cosulich SC, Guichard S, Vanhaesebroeck B, Joel S, et al. Kinase-substrate enrichment analysis provides insights into the heterogeneity of signaling pathway activation in leukemia cells. *Sci Signal* 2013; 6:rs6.
28. Ferguson FM, Gray NS. Kinase inhibitors: the road ahead. *Nat Rev Drug Discov* 2018;17:353–77.
29. Bhullar KS, Lagarón NO, McGowan EM, Parmar I, Jha A, Hubbard BP, et al. Kinase-targeted cancer therapies: progress, challenges and future directions. *Mol Cancer* 2018;17:48.
30. Sawyers CL. Opportunities and challenges in the development of kinase inhibitor therapy for cancer. *Genes Dev* 2003;17:2998–3010.
31. Chen WG, White FM. Proteomic analysis of cellular signaling. *Expert Rev Proteomics* 2004;1:343–54.
32. Cutillas PR, Jørgensen C. Biological signalling activity measurements using mass spectrometry. *Biochem J* 2011;434:189–99.
33. Ficarro SB, McClelland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, et al. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol* 2002;20:301–5.
34. Thingholm TE, Jensen ON, Larsen MR. Analytical strategies for phosphoproteomics. *Proteomics* 2009;9:1451–68.
35. Casado P, Hijazi M, Britton D, Cutillas PR. Impact of phosphoproteomics in the translation of kinase-targeted therapies. *Proteomics* 2017 March;17 [Epub ahead of print].
36. Dermit M, Dokal A, Cutillas PR. Approaches to identify kinase dependencies in cancer signalling networks. *FEBS Lett* 2017;591:2577–92.
37. Wilkes E, Cutillas PR. Label-free phosphoproteomic approach for kinase signaling analysis. *Methods Mol Biol* 2017;1636:199–217.
38. Wirbel J, Cutillas P, Saez-Rodriguez J. Phosphoproteomics-based profiling of kinase activities in cancer cells. *Methods Mol Biol* 2018;1711:103–32.
39. Montoya A, Beltran L, Casado P, Rodriguez-Prados JC, Cutillas PR. Characterization of a TiO₂ enrichment method for label-free quantitative phosphoproteomics. *Methods* 2011;54:370–8.
40. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999;20:3551–67.
41. Gerdes H, Casado P, Dokal A, Hijazi M, Akhtar N, Osuntola R, et al. Drug ranking using machine learning systematically predicts the efficacy of anti-cancer drugs. *Nat Commun* 2021;12:1850.
42. Alcolea MP, Casado P, Rodriguez-Prados JC, Vanhaesebroeck B, Cutillas PR. Phosphoproteomic analysis of leukemia cells under basal and drug-treated conditions identifies markers of kinase pathway activation and mechanisms of resistance. *Mol Cell Proteomics* 2012;11:453–66.
43. Ryall KA, Tan AC. Systems biology approaches for advancing the discovery of effective drug combinations. *J Cheminform* 2015;7:7.
44. Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, et al. A landscape of pharmacogenomic interactions in cancer. *Cell* 2016;166: 740–54.
45. Goecks J, Jalili V, Heiser LM, Gray JW. How machine learning will transform biomedicine. *Cell* 2020;181:92–101.
46. Marine JC, Dawson SJ, Dawson MA. Non-genetic mechanisms of therapeutic resistance in cancer. *Nat Rev Cancer* 2020;20:743–56.
47. Lau DK, Mouradov D, Wasenang W, Luk IY, Scott CM, Williams DS, et al. Genomic profiling of biliary tract cancer cell lines reveals molecular subtypes and actionable drug targets. *iScience* 2019;21:624–37.
48. Wang P, Song X, Utpatel K, Shang R, Yang YM, Xu M, et al. MEK inhibition suppresses K-Ras wild-type cholangiocarcinoma in vitro and in vivo via inhibiting cell proliferation and modulating tumor microenvironment. *Cell Death Dis* 2019;10:120.
49. Pant K, Peixoto E, Richard S, Gradilone SA. Role of histone deacetylases in carcinogenesis: potential role in cholangiocarcinoma. *Cells* 2020;9:780.