Deep learning-based phenotyping reclassifies combined hepatocellular-cholangiocarcinoma

Julien Calderaro  
*Université Paris Est Créteil*  
Pooja Navale  
*Washington University School of Medicine in St. Louis*  
et al.

Follow this and additional works at: https://digitalcommons.wustl.edu/oa_4

Part of the Medicine and Health Sciences Commons

Please let us know how this document benefits you.

**Recommended Citation**

https://digitalcommons.wustl.edu/oa_4/2965

This Open Access Publication is brought to you for free and open access by the Open Access Publications at Digital Commons@Becker. It has been accepted for inclusion in 2020-Current year OA Pubs by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Deep learning-based phenotyping reclassifies combined hepatocellular-cholangiocarcinoma

Primary liver cancer arises either from hepatocytic or biliary lineage cells, giving rise to hepatocellular carcinoma (HCC) or intrahepatic cholangiocarcinoma (ICCA). Combined hepatocellular-cholangiocarcinomas (cHCC-CCA) exhibit equivocal or mixed features of both, causing diagnostic uncertainty and difficulty in determining proper management. Here, we perform a comprehensive deep learning-based phenotyping of multiple cohorts of patients. We show that deep learning can reproduce the diagnosis of HCC vs. ICCA with a high performance. We analyze a series of 405 cHCC-CCA patients and demonstrate that the model can reclassify the tumors as HCC or ICCA, and that the predictions are consistent with clinical outcomes, genetic alterations and in situ spatial gene expression profiling. This type of approach could improve treatment decisions and ultimately clinical outcome for patients with rare and biphenotypic cancers such as cHCC-CCA.
on a self-supervised feature extractor\textsuperscript{11} with an attention-MIL aggregation model\textsuperscript{12–14} (Fig. 1B) to distinguish pure HCCs (785 WSIs from \( n = 424 \) patients) from pure ICCAs (239 WSIs from \( n = 167 \) patients) (Methods, Supplemental Tables 1 and 2). In this cohort ("Discovery cohort", Fig. 1C), the model achieved a cross-validated area under the receiver operator characteristic curve (AUROC) of 0.99 [ ± 0.01], corresponding to an almost perfect separability of the classes (Fig. 2A), reaching a sensitivity of 97.9% and specificity of 97.6%. As another piece of evidence for the plausibility of the model’s predictions, we subsequently evaluated the model on another patient cohort, the publicly available TCGA cohort, which was composed of \( n = 333 \) HCCs (TCGA-LIHC) and \( n = 27 \) ICCAs (TCGA-CHOL). The labels of the TCGA cohort were not seen by the model during training, however the training was exposed to some TCGA image data during self-supervised pretraining, which might affect an intermediary result but not the subsequent results. We found that the model reached an AUROC of 0.94 [ ± 0.05], representing a very good generalizability to this additional dataset (Fig. 2B). Next, we asked which tissue structures were used by the model to make its prediction and found that the model placed a high attention to areas with an ICCA-like phenotype (glandular structures and fibrous stroma) (Supplemental Fig. 1). Together, these data show that the AI model can robustly distinguish pure HCC from pure ICCA tumors (Fig. 2C).

We used this model as the starting point for our subsequent experiments.

**AI model application on cHCC-CCA samples**

Subsequently, we applied the trained model to a large multicentric cohort of cases which were initially diagnosed as cHCC-CCA (Supplemental Table 3). We investigated the spatial prediction maps and found that, generally, regions with HCC-like morphology were assigned a high "HCCness" by the model, while regions with ICCA-like morphology were assigned a high "ICCAness" by the model (Fig. 3A). For tumors with a significant proportion of equivocal or intermediate features, it was however much more difficult for pathologists to determine the morphology and proportion of areas with high "HCCness" or "ICCAness". As region-specific predictions are not clinically actionable, we further investigated the patient-level prediction scores. We found that these scores followed a bimodal distribution, with a subset of cases peaking at a high HCC prediction and the remainder of the cases peaking at a high ICCA prediction (Fig. 3B). Importantly, there was no association between the predictions and the clinical centers managing the patients (\( p = 0.62 \)). Together, these data show that the AI model can process tissue samples of cHCC-CCA cases and re-classify them as HCC or CCA. We then sought to determine if a simple pathological reclassification of cHCC-CCA tumors by
microscopic examination was associated with AI predictions. To this aim, all cases were reviewed in a blinded way by an expert liver pathologist (JC) and cHCC-CCA were reclassified as HCC or ICCA according to the more abundant morphological component (Methods). Interestingly, only a slight concordance was observed between the pathological analysis and the model’s predictions (Cohen’s Kappa 0.19, Supplemental Fig. 2), indicating that the AI model does not simply assess the more abundant tissue component in the way a human pathologist would.

Clinical outcomes based on AI-based reclassification

Next, we investigated the potential clinical and biological relevance of the AI-based reclassification of cHCC-CCA. One of the major differences between HCC and ICCA is their clinical outcome, with worse overall 5-year survival rates for patients with ICCA. We thus aimed to determine if our reclassification had an impact on the prognosis. Indeed, patients with cHCC-CCA reclassified as ICCA had a shorter median survival (29 months) than patients with a tumor reclassified as HCC (median survival not reached, \( p = 0.052 \), hazard ratio = 1.76 95% CI 0.98-3.14, Fig. 3C and Supplemental Table 4). Survival prediction is particularly relevant in patients who receive a liver transplant, as donor organs should be prioritized for patients with a good prognosis. A diagnosis of cHCC-CCA is currently considered a contra-indication to this therapeutic modality which remains a curative option for patients with HCC. As observed for resection, patients with cHCC-CCA reclassified as HCC showed a prolonged 5-year overall survival (76.4% survival rate, median survival not reached), which is similar to that usually observed in patients with conventional HCC (Fig. 3D and Supplemental Table 5). Transplanted patients with an incidental diagnosis of cHCC-CCA at transplant, which were re-classified as ICCA by the AI model, had a poor median survival of 48 months and a 5-year overall survival of only 45.4% (hazard ratio = 2.69, 95% CI 1.15-6.31). As opposed to resected patients, the prognostic impact remained significant on multivariate analysis. This observation may be explained by the fact

![Fig. 2](image-url) Development of a deep-learning model for HCC/ICCA classification.

A Receiver operator curve (ROC) for the internal validation of binary classification of HCC and ICCA cases. B ROC curve for the external validation of binary classification (HCC vs. ICCA) task on TCGA dataset. The error band shows the 1000 fold bootstrapped 95% confidence interval. C H&E slide of two randomly selected cases for HCC and ICCA. Attention map of the model and the class prediction scores are used as explainability methods to check the capability of the trained model in detecting the correct features within the WSI. The class prediction heatmap is weighted by the attention. Source data are provided as a Source Data file. This analysis was repeated independently with similar results five times.
that the competing mortality from cirrhosis and the risk of tumor recurrence are minimized by liver transplantation, as the whole diseased liver is replaced by the intervention. We also observed differences in patient characteristics according to the 2 different therapeutic modalities: transplanted patients were more frequently male \((p=0.039)\), cirrhotic \((p<0.001)\), with a higher frequency of alcohol consumption \((p<0.001)\) and a lower rate of HBV infection \((p=0.004)\) (Supplemental Table 6).

We further aimed to determine if the conventional reclassification of cHCC-CCA (according to the more abundant contingent assessed by a blinded pathologist) yielded any prognostic value, but observed that it had not a significant impact on survival in either resected \((p=0.16)\) or transplanted \((p=0.32)\) patients (Supplemental Figs. 3 and 4). In order to assess the degree of inter-observer variability of this histological reclassification, slides were also reviewed by another pathologist. The overall agreement was only fair (Cohen’s Kappa of 0.37), supporting the use of a more standardized and reproducible system such as our model. Altogether, these data suggest that our AI-based reclassification of cHCC-CCA allows us to make more clinically relevant predictions about disease outcomes than a classical pathological assessment.

**AI-based reclassification and genomic alterations**

We next investigated if the model’s predictions were concordant with known genetic differences of HCC and ICCA. We performed targeted next-generation sequencing with a panel that includes all major genes involved in HCC or ICCA development for \(n=104\) randomly selected cases. We identified several cases with alterations in TERT promoter, CTNNB1 and NFE2L2, which typically occur in HCCs, and several cases with FGFR2 fusions and IDH1/2, KRAS, NRAS, BRAF and HER2 mutations, which typically occur in ICCAs. We found that all genetic alterations in HCC-specific genes occurred in the tumor subset which the AI model had re-classified as HCC (Fig. 3E). Eleven out of 16 genetic alterations which are typical for ICCA occurred in tumors that were re-classified as CCA. In other words, the AI predictions match the genomic alterations of cHCC-CCA \((p=0.0009)\) in Fisher’s exact test, suggesting that the model detects patterns directly linked to the genetic defects identified by genomic profiling of the tumor tissue.

**Spatial transcriptomics analysis and AI predictions**

To gain further insights into the in situ relationships between the models’ predictions and the underlying biology, we performed spatial transcriptomics on tissue sections obtained from formalin-fixed, paraffin embedded blocks of 6 randomly selected cHCC-CCA cases. We then applied our model on the corresponding WSI and matched the prediction heatmaps with the gene expression profiling data. We investigated, within each case, the differences between the 100 image tiles most highly associated with a prediction of ICCA and the 100 tiles most highly associated with a prediction of HCC. We observed that the model’s predictions matched the underlying in situ gene expression profiles of the tumors. For Sample #A4 (Fig. 4A and B), areas predicted as ICCA-like were indeed associated with increased expression of genes related to cholangiocyte differentiation (e.g. EPCAM, HNF1B and KRT7) and decreased expression of well-known hepatocytic markers.
(ALB, FABP1 and APOB) (Fig. 4B and Supplementary Data 1). Similar findings were obtained for 2 additional cases (#A1 and #A2), while the results for remaining samples were less clear with few significantly dysregulated genes (#A3, #A5 and #A6) (Supplementary Data 1), possibly due to the constraints of performing spatial transcriptomics analyses on formalin-fixed paraffin-embedded material.

**Discussion**

In summary, our study shows that AI-based reclassification of cHCC-CCA into one of the “pure” HCC or CCA categories could improve prognostication, which is critical given the therapeutic implications, and also help to determine if a given cHCC-CCA tumor is genetically more similar to HCC or ICCA. A diagnosis of cHCC-CCA remains a formidable challenge for physicians and little or no evidence is available to guide the treatment options for the patients. Hence, oncologists often recommend treatment according to HCC or CCA therapeutic strategies, but the responses are often poor and their outcome dire. Recent large scale molecular studies of cHCC-CCA have failed to demonstrate any specific genetic alterations, and most cases have a similar gene expression profile to that of HCC or ICCA. This reclassification could be performed by genetic profiling, however these approaches are not universally available and are lengthy and costly. By definition, routine histopathological slides are available for every single one of these patients as histopathological evaluation is needed to make a diagnosis of cHCC-CCA in the first place.

Here, we have shown that an AI system can make a clear call for either HCC or ICCA. Reclassifying tumors as ICCA could be clinically useful as some of the associated alterations, including FGFR2 fusions and BRAF and IDH1 mutations, can be targeted by specific drugs. Measurable antitumor activity has indeed been reported with pemigatinib (FGFR inhibitors), futibatinib (FGFR inhibitors), ivosidenib (IDH1 inhibitor) or neratinib (pan-HER tyrosine kinase inhibitor). The standard of care for patients with advanced disease is also different between HCC (atezolizumab plus bevacizumab or durvalumab plus tremelimumab) and ICCA (durvalumab plus chemotherapy), and prospective clinical trials, although very challenging to carry out, are needed to determine if patients with cHCC-CCA may benefit from our reclassification approach to be allocated the systemic treatment that fits with the predicted class.

![Image](https://example.com/image.png)

**Fig. 4 | Combination of deep-learning heatmaps with spatial transcriptomics unravels the gene expression profile of areas that markedly impact the predictions. A** Example of a case processed by spatial transcriptomics: the HE section and its corresponding prediction heat map are presented, with the upper left area being considered as ICCA-like. This analysis was repeated independently with similar results for two other cases, as mentioned in the section “Spatial Transcriptomics Analysis and AI Predictions”. **B** Predictions matches the in situ gene expression profile with the ICCA like area showing upregulation of biliary/cholangiocyte genes (EPCAM, HNF1B and KRT7) and downregulation of hepatocytic genes (ALB, FABP1 and APOB). Raw data are available online as described in the “Data Availability” Section.
A limitation of our study is that by the nature of this problem, there is no ground truth for our proposed reclassification. We rely on a combination of clinical and genomic markers to demonstrate the plausibility and utility of our proposed reclassification scheme. We further observe that as for any biomarker, model predictions close to the decision cutoff are associated with a higher ambiguity. For such cases falling in the mid-range of the AI score, pathologists could integrate other factors such as clinical probability or imaging results to enhance the prediction confidence. Our software is currently suitable for research use only, and regulatory approvals will be needed to ensure its reliability and efficacy. The sharing of our source code however encourages further development and application by the wider research community.

The next step for the implementation of such models will be their validation on biopsies, as they are the only type of samples that can be obtained before surgery (resection or transplantation) or in patients with advanced disease not amenable for curative therapies. It may be challenging as biopsy is rarely performed due to the existence of non-invasive HCC diagnostic criteria. There is however a renewed interest in biopsy, in particular in the context of clinical trials, and this validation process could be undertaken in the near future. This could also be an opportunity to determine whether the addition of biological features (AFP/CA 19-9) or radiological findings may improve the classification.

In conclusion, our study demonstrates that AI may be useful for tumors that do not fit into common nosological frameworks. Developing evidence-based guidelines for such rare and challenging entities is indeed difficult (if not impossible). Our method could be applied to other cancer subtypes with mixed or biphenotypic differentiation that present a therapeutic challenge, such as combined adenocarcinomas/neuroendocrine tumors or adenosquamous carcinomas. We also believe that the combination of deep-learning heatmaps with spatial transcriptomics is a useful approach to provide insights into the molecular profile of highly predictive areas, and thus demonstrates that AI can be used as a tool for understanding tumor tissue in a research context.

Methods

Ethics statement
This study reports a retrospective analysis of tissue samples of archival tissue of primary liver tumors which was collected in a multicentric way. The protocol was approved by the review board of Université Paris Est Creteil, France (ID n° APHP22012), conducted in accordance with the Declaration of Helsinki and the legislations of each participating center. In this international multicentric cohort informed consents were obtained from patients when required by local regulations. Centers with informed written consent obtained: Hamburg, Barcelona, Mondor, Chinese University of Hong Kong, Beaujon, Paul Brousse. Centers with waiver of consent after IRB approval: University of Texas Southwestern, Stanford, Aachen, Pitié-Salpêtrière, Michigan University, Chennai, Rouen, Saint Antoine, Lille, Angers, Milano, Amiens, Hong Kong, Poitiers, St Louis University, Seoul National University College of Medicine, Prince of Songkhla, Montpellier, Brest, Reims, Yale School of Medicine, Bachmai Hospital, Mayo Clinic Rochester, Regensburg.

Patients and samples
Slides used for training (pure HCC and iCCA) were obtained from the archives of five pathology departments. Inclusion criteria were as follows: (1) patients with HCC or iCCA treated by surgical resection, (2) lack of preoperative antitumor treatment and (3) available WSI and baseline clinical, biological and pathological features. They were scanned using a Hamamatsu Nanozoomer S360 (ndpi encoding format) or a Leica Aperio (svs encoding format) scanning device. For the validation cohort, we used TCGA-LIHC and TCGA-CHOL cohorts (HCC \(n=333\) and iCCA \(n=27\)). Slides of cHCC-CCA were obtained from European (\(n=18\)), American (\(n=6\)) and Asian (\(n=6\)) liver centers. Inclusion criteria were: patients treated by surgical resection or liver transplantation, diagnosis of cHCC-CCA as defined by the World Health Organization and available histological slides and baseline clinical data.

Pathological reviewing
For all cases, an expert liver pathologist (JC) reviewed the cHCC-CCA histological slides and quantified each contingent (HCC, ICCA, and intermediate/equivocal). To compare the AI model’s prediction with a conventional morphological reclassification, cHCC-CCA cases were reclassified as HCC if the HCC contingent was more abundant that the ICCA contingent or as ICCA if the ICCA was more abundant that the HCC contingent.

Development and validation of the deep learning model
Processing of WSIs was performed according to a pre-defined protocol\(^{14}\). Digitized WSI were preprocessed by tessellation into non-overlapping small patches of size 224 × 224 × 3 pixels at an edge length of 256 µm. Background and blurry tiles were removed in order to provide the deep learning model with clean and informative input. As described before, the Canny Edge detector module with a threshold of 2 from OpenCV was used\(^{15}\). The primary analysis was carried out using the raw image tiles, and was repeated after color-normalization with the Macenko method to investigate potential batch effects\(^{16}\). Then, we used our previously published pipeline “Marugoto” for supervised Deep Learning\(^{17,18}\). The pipeline consists of a feature extraction module which transforms each tile into a feature vector of size 1 × 2048. Feature vectors of all tiles for each slide are subsequently processed by an aggregation module which outputs a single score for a given WSI. As the feature extraction module, we used a resnet50 which was pre-trained in a self-supervised way with the RetCCL method in a previous study\(^{19}\). The pretraining included TCGA image data but no labels. As the aggregation module, we used a custom-built attention-based multiple instance learning (attML)\(^{20}\). AttML incorporates an attention mechanism that involves two fully connected layers that compute an attention score for each tile, resulting in a bag-level feature vector obtained by scaling the embeddings of each tile using the softmax of its attention score, and adding them up\(^{20}\). This bag-level feature vector is then transformed into a final classification through another fully connected layer. To train the model on multiple patients in each batch, a subset of tiles from each patient is considered sufficient, and in each epoch, the tiles are re-sampled\(^{20,22}\). We trained the model for 100 epochs with a patient of 16 for an early-stopping callback. We used a batch size of 64 for the training subset and defined a fixed bag size of 512. For every training epoch, the instances for each bag have been randomly re-sampled. We used all the tiles of the WSIs for prediction in the test and validation cohorts, with a batch size of 1. In order to evaluate the model’s internal performance, we performed a 5-fold cross-validation at the level of patients.

Identification of molecular alterations
Tumor areas were first macro-dissected from formalin-fixed, paraffin embedded tissue blocks, and mRNA and DNA extractions were performed using the Maxwell RSC Plus DNA FFPE Kit iVD and the Maxwell RSC RNA FFPE Kit iVD (Promega, France). They were further quantified using a Qubit fluorometer in combination with the Qubit\(^{TM}\) dsDNA HS Assay Kit and Qubit\(^{TM}\) RNA HS Assay Kit (ThermoFisher Scientific). RNA was reverse transcribed to cDNA using SuperScript IV Vilo Master Mix (ThermoFisher Scientific). The Oncomine Comprehensive Library Assay v3C was used to amplify 50 nanograms of DNA and RNA (as measured by fluorometry). Amplicons were digested, barcoded and
amplified with the Ion AmpliSeq Library and Ion Xpress barcode adapter kits (ThermoFisher Scientific). After quantification, 50 pM of each library were multiplexed and clonally amplified on ion-sphere particles using a Ion Chef instrument (ThermoFisher Scientific). The ISP templates were loaded onto an Ion-540 chip and sequenced using an Ion SS device and the Ion 540™ Kit–Chef. The Ion Reporter Software was used to assess performance and analyze sequencing data using specific stringent filters (allele frequency between 5 and 90%; only exonic location, read depth >300X).

Spatial transcriptomics
Samples were first screened for RNA quality (DV200 scores > 50%, Tapestation), as recommended on the Visium Tissue Preparation Guide (10X Genomics). Visium spatial gene expression slides and reagents kits were used according to manufacturer instructions.

Five-micrometer thick tissue sections were cut from the FFPE tumor block and placed within the fiducial frames (n = 4) of the FFPE Visium Spatial Gene Expression Slide. Each capture area has ~5000 gene expression spots that include a partial read 1 sequencing primer (Illumina TruSeq Read 1), 16 nt spatial barcode, a 12 nt unique molecular identifier (UMI) and a 30 nt poly(dT) sequence (captures ligation product). Spots provide a resolution of ~5–10 cells. Slides were deparaffinized and stained. They were further coverslipped and scanned at 40X resolution using a Hamamatsu S360 scanning device. Coverslips were removed, and a decrosslinking step was performed.

Probes were hybridized using the Visium Hybridization Mix. After a post-hybridization wash (FPF Post-Hyb Wash and SSC Buffers), a ligase is added to seal the junction between the probe pairs that have hybridized to RNA, forming a ligation product. The ligation products were released from the tissue section upon RNase treatment and permeabilization, and further captured on the slide. Probes were extended by the addition of UMI, partial read 1 and spatial barcode. We then obtained spatially barcoded products for library preparation.

A qPCR was performed to determine the cycle numbers, and the ligated and spatially barcoded products underwent indexing via Sample Index PCR. Sequencing of libraries was performed on a Next-Seq 2000 instrument with a P3 flow cell (100 cycles, Illumina, CA, USA).

Survival analyses
Statistical analysis and visualization were performed using R software version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org) and Bioconductor packages (version 3.4). Overall survival was defined by the interval between surgical resection/liver transplantation and death or last follow-up. Survival curves were represented using the Kaplan-Meier method compared with log-rank statistics. Univariate analysis was performed using the Cox proportional-hazards regression model with variables with a P-value < 0.05 selected for multivariate analysis. All tests were two-tailed and a P-value < 0.05 was considered significant. For patients treated by surgical resection, inclusion criteria were lack of pre-operative treatment, lack of metastatic or macroscopic residual disease at the time of surgery, and uninodal tumors. For liver transplantation, all patients with available clinical follow-up were included.

Statistics, reproducibility and other tools
Measurements were taken from distinct samples, i.e., the same sample was never measured repeatedly. For supervised classification experiments, the primary endpoint was the area under the receiver operating characteristic curve (AUROC) with 95% confidence intervals obtained by 1000x bootstrapping. The MI-CLAIM checklist is provided in Supplemental Table 7. In accordance with the COPE Committee on Publication Ethics) position statement of 13 February 2023 (https://publicationethics.org/cope-position-statements/ai-author), the authors hereby disclose the use of the following artificial intelligence models during the writing of this article. GPT-4 (OpenAI) for checking spelling and grammar.

Reporting summary
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability
Some of the data that support the findings of this study are publicly available, and some are proprietary datasets provided for this analysis under collaboration agreements. All data (including histological images) from the TCGA database are available at https://portal.gdc.cancer.gov. Raw sequencing data for the proprietary cohorts have been uploaded to the European Nucleotide Archive (ENA) (accession number PRJEB62487). All other histopathology image data with accompanying metadata are under controlled access according to the local ethical guidelines and can only be requested directly from the respective study groups that independently manage data access for their study cohorts. The central data collection was managed by JC to whom sharing requests can be directed and will be responded to within 4 weeks. Source data for figures are provided with this paper.

Code availability
The data was analyzed using custom-developed open-source software. Our deep learning methods use Python with h5py v3.6, numpy v1.22, openpyxl v3.0, pandas v1.4, torch v1.8, fastai v2.5, fire v0.4. The transcriptomics analysis and statistical analysis methods use R v4.1.2, Seurat v4.1.1 for differential gene expression analysis, Seurat v4.3.0 for visualization, glmGamPoi v1.6.0, dplyr v1.0.1, crayon v1.3.5, ggplot2 v3.4.0, gridExtra v2.3, MAST v1.20.0, ggrepel v0.9.2, readxl v1.4.3, survminer v0.4.9, rms v6.4-1, survival v3.3-5, ComplexHeatmap v2.13.1, irr v0.84.1, vcd v1.4-11. All source codes are publicly available: https://github.com/KatherLab/preprocessing-ng for WSI tessellation, https://github.com/KatherLab/preProcessing for color-normalization and https://github.com/KatherLab/marugo for model training and deployment and https://github.com/qinghezeng/ST_cHCC-CCA for spatial transcriptomics analysis.

References


Acknowledgements

J.C. is supported by Institut National du Cancer (PRT-K22-146 and TRANSCAN TANGERINE project), Bayer, Fondation Bristol Myers Squibb for the Recherche en Immuno-Oncologie, Cancéropole Ile-de-France Emergence, Fondation ARC, Ipsen, Ligue Contre le Cancer, and Fondation de l’Avenir. J.N.K. is supported by the Max-Eder-Programme of the German Cancer Aid (grant #70113864), the German Federal Ministry of Education and Research (PEARL, 01KD21014C; CAMINO, 01EO2101; SWAG, 01KD2215A), the German Academic Exchange Service (SECAI, 57616814), the German Federal Joint Committee (Transplant.KI, 01VS21048) the European Union’s Horizon Europe and innovation programme (ODELIA, 101057091; GENIAL, 101096312) and the National Institute for Health and Care Research (NIHR, NRHR213331) Leeds Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. J.N.K and TL received funding from the Federal Ministry of Health under grant number 2520DAT111 (Deep Liver) and from the German Federal Ministry of Education and Research under grant number O3IL03128 (TransformLiver). S.C. is supported by Fondation ARC and Les Entreprises contre le Cancer (Gefluc). H.K. is supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2022R1A2C2010348). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. H.K. is supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2022R1A2C2010348). Q.Z. is funded by the China Scholarship Council (Grant n°201908070052). A.D. is supported by grant from Instituto de Salud Carlos III (PI22/ 01427), A.F. Grant from Instituto de Salud Carlos III (PI18/00542), M.R. is supported by grant from Instituto de Salud Carlos III (PI15/00145, PI18/0358, PI22/ 01427 and PMP22/00054).CIBERehd is funded by the Instituto de Salud Carlos III. Some of the authors of this article are members of the European Reference Network (ERN) RARE-LIVER. This work was funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Author contributions


Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

J.N.K. declares consulting services for Owkin, France; DoMore Diagnostics, Norway, Panakeia, UK and Histofy, UK; furthermore he holds shares in StratifiAIGmbh and has received honoraria for lectures by AstraZeneca, Bayer, Eisai, MSD, BMS, Roche, Pfizer and Fresenius. JC reports consulting for Crosscope. HR reports consulting service for Boston. The remaining authors declare no competing interests.

Additional information

Supplementary material available at https://doi.org/10.1038/s41467-023-43749-3.

Correspondence

Requests for materials should be addressed to Julien Calderaro or Jakob Niklas Kather.

Peer review information

Nature Communications thanks the anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available at https://doi.org/10.1038/s41467-023-43749-3.

Reprints and permissions information

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher’s note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Julien Calderaro1,2,3,4, Narmin Ghaffari Laleh5,6, Qinghe Zeng7,8, Pascale Maille1,2,3, Loetitia Favre1,2,3, Anaïs Pujals1,2,3, Christophe Klein7,9, Céline Bazille10, Lara R. Hej11,12, Arnaud Uguen13,14, Tom Luedde15, Luca Di Tommaso16, Aurélie Beaufraître17,18, Augustin Chatain3, Delphine Gastineau3, Cong Trung Nguyen19, Hięp Nguyen-Canh20,21, Khuyen Nguyen Thi22, Viviane Gneumni23,24, Rondell P. Graham25, Frédéric Charlotte26, Dominique Wendum27, Mukul Vij28, Daniela S. Allende29,30, Federico Aucejo31, Alba Diaz32, Benjamin Rivière33, Astrid Herrero34, Katja Evert35, Diego Francesco Calvisi36, Jérémy Augustin1,2,3, Wei Qiang Leow36, Howard Ho Wai Leung37, Emmanuel Boleslawski38, Mohamed Rela39, Arnaud François40, Anthony Wing-Hung Cha37, Alejandro Forner39,40, Arnaud Uguen31, Christophe Tourmigand41, Emmanuelle Kempf42, Haeryoung Kim53, Massih Ningarhari54, Sophie Michalak-Provost55, Purva Gopal56, Raffaele Brustia57, Eric Vibert58, Kornelius Schulze59, Darius F. Rüther60, Sören A. Weidemann60, Rami Rhaiem61, Jean-Michel Pawlotsky1,3,38, Xuchen Zhang62, Alain Luciani63, Sébastien Mulé64, Alexis Laurent57, Giuliana Amaddeo64, Hélène Regnault64, Eleonora De Martin65, Christine Sempoux66, Pooja Navale67, Maria Westerhoff68, Regina Cheuk-Lam Lo69,70, Jan Bednarsch71, Annette Gouw71, Catherine Guettier11,71,72, Marie Lequoy73, Kenichi Harada21, Pimsiri Sripongpun74, Poowadon Wetwittayaklang75, Nicolas Loménie18, Jarukit Tantipisit76, Apichat Kaewdech31,18, Jeanne Shen76,77, Valérie Paradis17,18, Stefano Caruso1,2,3,4, and Jakob Nikolos Kather7,58

1Université Paris Est Créteil, INSERM, IMRB, F-94010 Créteil, France. 2Assistance Publique-Hôpitaux de Paris, Henri Mondor-Albert Chenevier University Hospital, Department of Pathology, Créteil, France. 3Inserm, U955, Team 18, Créteil, France. 4European Reference Network (ERN) RARE-LIVER, Créteil, France. 5Else Kroener Fresenius Center for Digital Health, Medical Faculty Carl Gustav Carus, Technical University Dresden, Dresden, Germany. 6Department of Medicine III, University Hospital RWTH Aachen, RWTH Aachen university, Aachen, Germany. 7Centre d’Histologie, d’Imagerie et de Cytométrie (CHIC), Centre de Recherche des Cordeliers, Paris, France. 8Laboratoire d’Informatique Paris Descartes (LIPADE), Université Paris Cité, Paris, France. 9INSERM, Sorbonne Université, Université Paris Cité, Paris, France. 10Caen University Hospital, Department of Pathology, Caen, France. 11Department of Surgery and Transplantation, University Hospital RWTH Aachen, Aachen, Germany. 12Institut de Pathologie, University Hospital RWTH Aachen, Aachen, Germany. 13CHRU Brest, Department of Pathology, Brest 29220, France. 14Univ Brest, Inserm, CHU de Brest, LBAI, UMR1227 Brest, France. 15Clinic for Gastroenterology, Hepatology and Infectious Diseases, University Hospital Düsseldorf, Medical Faculty of Heinrich Heine University Düsseldorf, Düsseldorf, Germany. 16Department of Pathology, Humanitas University, Humanitas Clinical and Research Center, IRCCS, Rozzano, Milan, Italy. 17Assistance Publique–Hôpitaux de Paris, Beaujon University Hospital, Department of Pathology, F-92110 Clichy, France. 18Université de Paris, Inflammation Research Center, Inserm, U1149, CNRS, ERL8125, F-75018 Paris, France. 19Department of Pathology, E Hospital, Hanoi Medical University, Hanoi, Vietnam. 20Pathology Center, Bachmai Hospital, Hanoi, Vietnam. 21Department of Human Pathology, Kanazawa University Graduate School of Medicine, Kanazawa, Japan. 22Pathology and Molecular biology Center, National Cancer Hospital, Hanoi, Vietnam. 23University Lille, UMR9020-U1277, Cancer Heterogeneity Plasticity and Resistance to Therapies, Lille, France. 24CHU Lille, Institute of Pathology, Lille, France. 25Department of Laboratory Medicine and Pathology, Mayo Clinic Rochester, Rochester, MN, USA. 26Assistance Publique-Hôpitaux de Paris, Pitié-Salpêtrière University Hospital, Department of Pathology, Paris, France. 27Assistance Publique-Hôpitaux de Paris, Saint-Antoine University Hospital, Department of Pathology, Paris, France. 28Department of Pathology, Dr Rela Institute and Medical Centre, Bharath Institute of Higher Education and Research, Chennai, India. 29Department of Hepatobiliary Pathology, Cleveland Clinic Foundation, Cleveland, OH, USA. 30Robert J. Tomisch Pathology & Laboratory Medicine Institute, Cleveland Clinic, 9500 Euclid Avenue, L25, Cleveland, OH 44195, USA. 31Department of Gastrointestinal and Hepatobiliary Surgery, Cleveland Clinic Foundation, Cleveland, OH, USA. 32Barcelona Clinic Liver Cancer (BCLC) Group, Department of Pathology, Hospital Clinic de Barcelona, Universitat de Barcelona, Barcelona, Spain. 33Department of Pathology, Gui-de-Chaluiac University Hospital, 80, avenue Augustin-Fliche, 34295 Montpellier, France. 34Department of Digestive and Hepatobiliary Surgery, Gui-de-Chaluiac University Hospital, 80, avenue Augustin-Fliche, 34295 Montpellier, France. 35Institute of Pathology, University of Regensburg, Franz-Josef-Strauß-Allee 11, 93053 Regensburg, Germany. 36Department of Anatomical Pathology, Singapore General Hospital, Singapore, Singapore. 37Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong, Shatin, Hong Kong. 38CHU Lille, Department of Digestive and Hepatobiliary Surgery, Lille, France. 39Dr Rela Institute and Medical Centre, Bharath Institute of Higher Education and Research, Chennai, India. 40Rouen University Hospital, Department of Pathology,