

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give <i>P</i> values as exact values whenever suitable.</i>
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	WGS (n=494) and RNA sequencing (n=239) libraries were sequenced with Illumina Novaseq (Illumina) with built-in software. For the experimental section, the CCK8 signals for cell proliferation assay were detected with Synergy Neo microplate reader (BioTek). The crystal-violet-stained images for colony formation, cell migration, and invasion assays were scanned by an Olympus IX73 microscope equipped with an DP80 camera (Olympus). IHC slides were scanned by a Leica Aperio AT2. Immunoreactive bands were detected using e-BLOT Touch Imager XLi or Odyssey Sa Infrared Imaging System (LI-COR Biosciences). Electrochemiluminescence (ECL) signals for IL-6 concentration in tumor tissues were recorded on 1300 MESO QuickPlex SQ 120MM instrument (Meso Scale Discovery). RT-qPCR was performed on LightCycler 96 PCR platform (Roche). Phospho-specific protein microarray data was obtained with an Axon Instruments GenePix 4000B Microarray Scanner.
Data analysis	The Linux working environment we used for sequencing data analysis is packed into a Singularity container file and published at Zenodo (https://doi.org/10.5281/zenodo.7260221). The detailed codes for all the software have been deposited at GitHub (https://github.com/ChongJenniferZhang/CLCA_WGS). Statistical analyses were performed using R (version 3.6.0) and GraphPad Prism (version 9.0). FASTP (v0.13.1) https://github.com/OpenGene/fastp BWA (v0.7.12) http://bio-bwa.sourceforge.net/ Sambamba (v0.6.8) https://github.com/biod/sambamba Mutect2 (v4.0.11.0) https://gatk.broadinstitute.org/hc/en-us Strelka (v2.8.4) https://github.com/Illumina/strelka MutSigCV (v1.4) https://www.genepattern.org/modules/docs/MutSigCV dndscv (v0.1.0) https://github.com/im3sanger/dndscv

OncodriveFML (v2.3.0) <https://oncodrivefml.readthedocs.io/en/latest/index.html>
 MutSig2CV_NC (v1.0) https://github.com/broadinstitute/getzlab-PCAWG-MutSig2CV_NC
 ActiveDriverWGS (v1.1.1) <https://cran.r-project.org/web/packages/ActiveDriverWGS/index.html>
 pvalue_combination (v1.0) https://github.com/broadinstitute/getzlab-PCAWG-pvalue_combination
 SigProfilerExtractor (v1.1.0) <https://github.com/AlexandrovLab/SigProfilerExtractor> <https://github.com/AlexandrovLab/SigProfilerSingleSample>
 mSigHdp (v1.1.2) <https://github.com/steverozen/mSigHdp/>
 maftools (v2.6.05) <http://www.bioconductor.org/packages/release/bioc/html/maftools.html>
 Shatterseek (v0.4) <https://github.com/parklab/ShatterSeek>
 ChainFinder (v1.0.1) <https://software.broadinstitute.org/cancer/cga/chainfinder>
 lumpy (v0.2.13) <https://github.com/arq5x/lumpy-sv>
 Sequenza (v2.1.1) <https://cran.r-project.org/web/packages/sequenza/>
 GISTIC (v2.0.23) https://www.genepattern.org/modules/docs/GISTIC_2.0
 purple (v2.34) <https://github.com/hartwigmedical/hmftools/blob/master/purity-ploidy-estimator/README.md>
 AmpliconArchitect (v1.3.r2) <https://github.com/virajbdeshpande/AmpliconArchitect>
 AmpliconClassifier (v0.2.5) <https://github.com/jluebeck/AmpliconClassifier>
 PyClone (v0.13.1) <https://github.com/Roth-Lab/pyclone>
 MutationTimeR (v0.99.3) <https://github.com/gerstung-lab/MutationTimeR>
 PhylogicNDT (v1.0) <https://github.com/broadinstitute/PhylogicNDT>
 Timing_and_Signatures (v1.0) https://github.com/clemencyjolly/PCAWG11-Timing_and_Signatures
 STAR (v2.7.3a) <https://github.com/alexdobin/STAR>
 RSEM (v1.3.3) <https://github.com/deweylab/RSEM>

IHC images were analyzed by Aperio ImageScope v12.4.6(Leica). Band intensity of western blots were assessed by ImageJ 1.53a. RT-qPCR were analyzed with LightCycler 96 SW 1.1 (Roche). Electrochemiluminescence (ECL) signals were analyzed with DISCOVERY WORKBENCH Desktop Analysis Software version 4.0 (Meso Scale Discovery). Phospho-specific protein microarray data was analyzed with GenePix Pro 6.0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The raw sequence data reported in this paper has been deposited in the Genome Sequence Archive in BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under the study accession number PRJCA002666 (<https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA002666>). We also built an interactive website (<http://lifeome.net/database/liver>) for visualizing and analyzing our CLCA data. The data deposited and made public are compliant with the regulations of the Ministry of Science and Technology of China. Other public data used in this study includes, the human reference genome of hg19/GRCh37 (<https://ftp.ensembl.org/pub/grch37/>), PCAWG data (<https://dcc.icgc.org/pcawg/#!>), TCGA-HCC data (<https://portal.gdc.cancer.gov/projects/TCGA-LIHC>), and COSMIC signatures (<https://cancer.sanger.ac.uk/signatures/>).

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	CLCA cohort comprised 427 men (86.4%) and 67 women (13.6%), with a mean age of 56 years (range, 23–84 years). All patients were enrolled from Eastern Hepatobiliary Surgery Hospital and Shanghai Zhongshan Hospital during 2017-2020.
Population characteristics	94.5% of patients had HBV infection. 85.6% of patients were Edmondson-Steiner grades 3 and 4. 26.7% and 36.8% of patients had alcohol drinking and smoking history, respectively. Detailed clinical information was summarized in Supplementary Table 1.
Recruitment	All patients included were diagnosed with hepatocellular carcinoma. No patients received any pre-operative anti-cancer treatment. Each specimen was diagnosed by two senior pathologists. Patients with tissue samples that had sufficient and good-quality DNA were selected.
Ethics oversight	The study protocol was reviewed and approved by the institutional review board at all participating hospitals. This study was performed in accordance with the principles of the Declaration of Helsinki. All participants provided written informed consent. All samples were anonymously coded in accordance with local ethical guidelines. All research participants consent to the publication of research results.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The Chinese Liver Cancer Atlas (CLCA) cohort described in the paper contains 494 HCC patients collected in Eastern Hepatobiliary Surgery Hospital and Shanghai Zhongshan Hospital during 2017-2020. 494 patients with HCC were subjected to sequencing, including WGS (n=494) and messenger RNA sequencing (n=239). Detailed clinical information was summarized in Supplementary Table 1. No sample size calculations were performed for human as the main aim of the study was to build up a resource. For xenograft mice experiments, a minimum of 3 mice for each group of the PLC/PRF/5-sh-Ctrl and PLC/PRF/5-sh-FGA cells are required to reach statistical significance.
Data exclusions	There is no data that were excluded from the WGS and RNA-seq analyses.
Replication	No replication is needed for WGS and RNA-seq samples in our study since they are all clinical samples. For experimental validation of potential drivers of PPP1R12B, KCNJ12, and FGA, dysfunctional cell lines were constructed by either knockdown with two independent short hairpin RNA (shRNA, #1, #2) or knockout with two independent short guide RNA (sgRNA, #1, #2). Then these cell lines were subjected for assessing proliferation, migration, invasion, and self-renewal capacities. Each assay was repeated three times independently and representative images are shown.
Randomization	No randomization was performed for the human tumor samples because this is an observational study. For xenograft models, body weight-matched mice were randomized for subcutaneous injection of PLC/PRF/5-sh-Ctrl and PLC/PRF/5-sh-FGA cells.
Blinding	Our study was not an intervention study and therefore blinding was not required.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Mouse monoclonal anti-β-actin (Cat# AC004, clone AMC0001, RRID:AB_2737399, blot 1:5000; ABclonal) Mouse monoclonal anti-GAPDH (Cat# AC033, clone AMC0062, RRID:AB_2769570, blot 1:5000; ABclonal) Rabbit polyclonal anti-TYK2 (Cat# 9312, RRID:AB_2256719, blot 1:1000; Cell Signaling Technology) Rabbit monoclonal anti-Phospho-Tyk2 (Tyr1054/1055) (D7T8A) (Cat# 68790, clone D7T8A, RRID:AB_2799752, blot 1:1000, staining 1:100; Cell Signaling Technology) Rabbit monoclonal anti-Phospho-Stat3 (Tyr705) (D3A7) XP (Cat# 9145, clone D3A7, RRID:AB_2491009, blot 1:2000; Cell Signaling Technology) Mouse monoclonal anti-Lamin A/C(4C11) (Cat# 4777, clone 4C11, blot 1:2000; Cell Signaling Technology) Mouse monoclonal anti-Fibrinogen α (C-7) (Cat# sc-398806, clone C-7, blot 1:500; Santa Cruz Biotechnology) Rabbit polyclonal anti-Fibrinogen Alpha Chain (Cat# 20645-1-AP, RRID:AB_2878715, staining 1:100; Proteintech) Mouse monoclonal anti-STAT3 (Cat# 60199-1-Ig, clone 3G2D12, RRID:AB_10913811, blot 1:2000; Proteintech) Rabbit polyclonal anti-Ki67 (Cat# ab15580, RRID:AB_443209, staining 1:500; abcam) HRP-conjugated anti-Rabbit (Cat# D-3002; staining 1:1; Supervision) HRP-conjugated Affinipure Goat Anti-Rabbit IgG(H+L) (Cat# SA00001-2, RRID:AB_2722564, blot 1:5000; Proteintech) HRP-conjugated Affinipure Goat Anti-Mouse IgG(H+L) (Cat# SA00001-1, RRID:AB_2722565, blot 1:5000; Proteintech)
-----------------	--

IRDye 800CW Goat anti-Rabbit IgG (H + L) (Cat# 926-32211, RRID:AB_621843, blot 1:20000; LI-COR)
 IRDye 800CW Goat anti-Mouse IgG (H + L) (Cat# 926-32210, RRID:AB_621842, blot 1:20000; LI-COR)

Validation

All antibodies used in this study are commercially available. They are validated by the vendors for the specific assay and species used, with the validation reports available on the vendor's website. All antibodies were titrated to determine the optimal working concentration.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

For the functional validation of three candidate drivers, the human liver cancer cell lines, PLC/PRF/5, PVTT, HepG2, Huh7, SNU387, SNU182, and the normal liver cell line HHL5 were obtained from Shanghai Cell Bank of the Chinese Academy of Sciences. For the validation of AA-related mutational signatures, MCF-10A and HepG2 cells were obtained from the American Type Culture Collection (ATCC).

Authentication

All cell lines used in this study were authenticated by applying short tandem-repeat DNA profiling.

Mycoplasma contamination

We confirm that all cells were tested as mycoplasma negative.

Commonly misidentified lines (See [ICLAC](#) register)

No commonly misidentified cell lines were used.

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals

BALB/c nude mice (5-7 weeks) were obtained from GemPharmatech LLC (JiangSu, China) and used for subcutaneous xenograft. All mice were housed in pathogen free conditions at an ambient temperature 20-26°C and humidity of 30-70% with a 12:12 hour light:dark cycle prior to use. Mice had unrestricted access to regular mouse chow and water. The tumour width (w) and length (l) were measured every 3 days with a caliper, and the diameter of single tumour was < 2cm when sacrificed.

Wild animals

This study did not involve wild animals.

Reporting on sex

Preliminary subcutaneous xenograft experiments were performed on male and female mice, respectively. Similar trends of sh-FGA cells resulted in larger and more aggressive tumours in comparison with those of mice injected with sh-Ctrl cells were observed. To exclude the potential confounding factors of aggression and biting in the male groups, only the female groups were kept and recorded. The tumorigenic role of FGA dysfunction in HCC applies to both sexes.

Field-collected samples

This study did not involve field-collected samples.

Ethics oversight

All mouse experiments were approved by the Animal Care and Use Committee at Eastern Hepatobiliary Surgery Hospital.

Note that full information on the approval of the study protocol must also be provided in the manuscript.