scientific reports

Check for updates

OPEN

A live single-cell reporter system reveals drug-induced plasticity of a cancer stem cell-like population in cholangiocarcinoma

Krittiyabhorn Kongtanawanich[1], Sunisa Prasopporn[1,2], Supawan Jamnongsong[1,2], Nontaphat Thongsin[3], Tongchai Payungwong[1,2], Seiji Okada[4], Marianne Hokland[5], Methichit Wattanapanitch[3] & Siwanon Jirawatnotai[1,2,4,6]✉

Cancer stem cells (CSC) play an important role in carcinogenesis and are acknowledged to be responsible for chemoresistance in cholangiocarcinoma (CCA). Studying CCA CSC has been challenging, due to lack of consensus CSC markers, and to their plastic nature. Since dual expression of the core pluripotent factors SOX2/OCT4 has been shown to correlate with poor outcome in CCA patients, we selected the SOX2/OCT4 activating short half-life GFP-based live reporter (SORE6-dsCopGFP) to study CSC dynamics at the single-cell level. Transduction of five human CCA cell lines resulted in the expression of 1.8–13.1% GFP-positive (SORE6^{POS}) cells. By live imaging, we found that SORE6^{POS} CCA cells possess self-renewal capacity and that they can be induced to differentiate. Significantly, the SORE6^{POS} cells were highly tumorigenic, both in vitro and in vivo, thus implicating the characteristics of primary CSCs. When we then analyzed for selected CSC-related markers, we found that the majority of both CD133^{+}/CD44^{+}, and CD133^{+}/LGR5^{+} CCA cells were SORE6^{POS} cells. Exposing transduced cells to standard CCA chemotherapy revealed higher growth rate inhibition at 50% (GR_{50}s) for SORE6^{POS} cells compared to GFP-negative (SORE6^{NEG}) ones indicating that these CSC-like cells were more resistant to the treatment. Moreover, the chemotherapy induced SORE6^{POS} from SORE6^{NEG} cells, while retaining the existing SORE6^{POS} population. Finally, treatment of transduced cells with CDK4/6 inhibitors in vitro for 3 days resulted in a lowered CSC number in the culture. Thus, applying a live reporter system allowed us to elucidate the stem cell diversity and drug-induced plasticity of CCA CSCs. These findings have clear implications for future management of such patients.

Keywords Cholangiocarcinoma, Drug resistance, Cancer stem cell, CDK4/6 inhibitor, Live reporter

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
5-FU	5-Fluorouracil; IUPAC: 5-fluoro-1H-pyrimidine-2,4-dione
AB	Abemaciclib; IUPAC: N-[5-[(4-ethylpiperazin-1-yl)methyl]pyridin-2-yl]-5-fluoro-4-(7-fluoro-2-methyl-3-propan-2-ylbenzimidazol-5-yl)pyrimidin-2-amine
ANPEP	Alanyl aminopeptidase, membrane
APC	Allophycocyanin
ARRIVE	Animal research: reporting of in vivo experiments
ATRA	All-trans retinoic acid; IUPAC name: (2E,4E,6E,8E)-3,7-dimethyl-9-(2,6,6-trimethylcyclohex-en-1-yl)nona-2,4,6,8-tetraenoic acid
AUC	Area under the curve

[1]Department of Pharmacology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. [2]Siriraj Center of Research Excellence for Precision Medicine and Systems Pharmacology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. [3]Siriraj Center for Regenerative Medicine, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. [4]Division of Hematopoiesis, Joint Research Center for Human Retrovirus Infection, Kumamoto University, Kumamoto, Japan. [5]Department of Biomedicine, Aarhus University, Aarhus, Denmark. [6]Faculty of Pharmacy, Silpakorn University, Nakhon Pathom, Thailand. ✉email: siwanon.jir@mahidol.ac.th

AVMA	American Veterinary Medical Association
Balb/c-RJ	BALB/c mice with Rag2 and Jak3 double knockout
CCA	Cholangiocarcinoma
CCA CSC	Cholangiocarcinoma stem cell
CDK	Cyclin-dependent kinase
CHOL	Cholangiocarcinoma
Cis	Cisplatin; IUPAC: azane, dichloroplatinum
COA	Certificate of analysis
CSC	Cancer stem cell
cyTOF	Cytometry by time-of-flight
DMEM	Dulbecco's modified Eagle's medium
DMF	Dimethylformamide, IUPAC name: N,N-dimethylformamide
DMSO	Dimethyl sulfoxide; IUPAC name: methylsulfinylmethane
EDTA	Ethylenediaminetetraacetic acid; IUPAC: 2-[2-[bis(carboxymethyl)amino]ethyl-(carboxymethyl)amino]acetic acid
EPCAM	Epithelial cell adhesion molecule
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
Gem	Gemcitabine; IUPAC: 4-amino-1-[(2R,4R,5R)-3,3-difluoro-4-hydroxy-5-(hydroxymethyl)oxolan-2-yl]pyrimidin-2-one
Gem+Cis	Gemcitabine plus cisplatin
GFP	Green fluorescence protein
GR	Growth rate inhibition
GR_{40}	Growth rate inhibition at 40 percent
GR_{50}	Growth rate inhibition at 50 percent
GR_{90}	Growth rate inhibition at 90 percent
HEK293T	Human embryonic kidney 293T
HTS	High-throughput screening
iPS	Induced pluripotent stem
IHC	Immunohistochemistry
JCRB	Japanese Collection of Research Bioresources
LGR5	Leucine-rich repeat-containing G-protein coupled receptor 5
mCMV	Minimal cytomegalovirus
MU-IACUC	Mahidol University-Institute Animal Care and Use Committee
NANOG	Nanog homeobox
NIH	National Institutes of Health
NIR	Near infrared
NIS	Nikon Imaging Software
OCT4	Octamer-binding transcription factor 4
OS	Overall survival
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PD	Palbociclib; IUPAC: 6-acetyl-8-cyclopentyl-5-methyl-2-[(5-piperazin-1-ylpyridin-2-yl)amino]pyrido[2,3-d]pyrimidin-7-one
PE	Phycoerythrin
PROM1	Prominin 1
RPMI 1640	Roswell Park Memorial Institute 1640 Medium
rtPCR	Reverse transcriptase polymerase chain reaction
$SORE6^{NEG}$	GFP-negative expression
$SORE6^{POS}$	GFP-positive expression
$SORE6^{WEAK}$	GFP-weak expression
SOX2	SRY-Box Transcription Factor 2
SOX9	SRY-Box Transcription Factor 9
ssGSEA	Single-sample gene set enrichment method
TCGA	The Cancer Genome Atlas Program
THY1	Thy-1 cell surface antigen
tSNE	t-distributed stochastic neighbor embedding
ZFP206	Zinc finger and SCAN domain containing 10

Cholangiocarcinoma (CCA) is an aggressive tumor of the bile duct epithelia with a poor prognosis, primarily due to drug resistance[1,2]. Recent evidence point to cancer stem cells (CSCs) as one of the potential sources of drug resistance[3,4]. These cells belong to a minor cancer cell population with special characteristics, such as self-renewal, apoptosis resistance, slow proliferation, and generation of new cancer cells[5,6]. Dynamic remodeling and phenotype plasticity of the intrahepatic biliary CSCs have been shown to be to play key roles in tissue regeneration[7]. In agreement with this, CCA has been identified to contain high levels of CSCs[8].

Currently, CCA CSCs are intensively studied, mainly by approaches that rely on CSC-associated surface markers including EpCAM, CD44, CD133, LGR5[9–11]. However, results from these studies often appear discordant, and at times contrasting, partly due to lack of consensus on the CCA CSC markers[11–15].

CSCs have been shown to instigate heterogeneity in tumors by existing in reversible states of dormancy, quiescence, aneuploidy, and stemness[16,17], and cancer microenvironmental signals have been shown to influence the heterogeneity and number of stem cells in the tumor. Thus, stemness is a plastic state that could be acquired and lost based on the microenvironment and signaling cues[18,19]. There have been discussions on the possibility of targeting CSC for cancer therapy. Since, CSCs can cause cancer recurrence, metastasis, heterogeneity, and treatment resistance, an effective therapeutic strategy should include attacking CSCs[20]. However, more studies are necessary to advance the CSC-targeting drug discovery, especially in CCA[8]. Thus, an experimental set-up, by which CSCs can be studied is required to understand the CSC plasticity and its hierarchical staging.

Thus, we developed a novel surface marker-independent system to identify CSCs in commercially available CCA cell lines to study CSC number and plasticity in real-time and at the single-cell level. The system relies on a live GFP reporter regulated by the binding of two putative pluripotent transcription factors SOX2 and OCT4 to regulatory elements (SORE6 sequence), from the NANOG promoter[21]. Using this GFP reporter system we were able to identify CCA cells with CSC properties. Importantly, apart from a high degree of plasticity, we demonstrated the plastic nature of CCA CSCs during drug treatment, thus helping to understand the role of CCA CSCs in drug resistance.

Results

Dual expression of core pluripotent factors SOX2/OCT4 is associated with poor clinical outcome in CCA

Since CSC has been implicated as a driver for disease severity in CCA, we investigated the association between key proteins in CCA CSC and disease prognosis. Overall survival analyses were performed based on data from The Cancer Genome Atlas (TCGA) by focusing on the roles of core pluripotency proteins. We found that, as a single marker, none of the factors and prominent CSC markers (*NANOG*, *SOX2*, *OCT4*, *SOX9*, *CD44*, *EPCAM*, *LGR5*, *PROM1* [CD133], *ANPEP* [CD13], and *THY1* [CD90]) significantly associated with overall survival in CCA patients (Supplementary Fig. 1). SOX2 and OCT4 are known to co-regulate the pluripotency gene expression program, and promote expression of other key stem cell factors, including *nanog*, and *ZFP206*[21–23]. Interestingly, dual expression of *SOX2* and *OCT4* has been identified to play an important role in carcinogenesis and progression of several types of cancer[24–27]. In line with this we found that as a pair, the dual expression of *SOX2* and *OCT4* was strongly associated with overall survival in CCA patients. SOX9 has been implicated as a CSC factor in certain types of cancer such as colon, prostate, and liver[28,29]. However, the role of SOX9 in other tissues is less clear. From the TCGA database, we found that high *SOX9* expression appeared to correlate with a longer survival (Supplementary Fig. 1d). *SOX9* mRNA levels were relatively low in induced pluripotent stem cell, compared to CCA cells (Supplementary Fig. 1k, l), supporting the notion that SOX9 plays a role in cholangiocyte lineage commitment[30]. From these results, we believe that *SOX9* expression does not have discriminative power to distinguish CCA CSCs from CCA cells. Since the high combined expression of *SOX2* and *OCT4* transcripts significantly correlated with poor survival (Fig. 1a), the dual expression of *SOX2* and *OCT4* might be a key driver for the stem cell-mediated tumorigenesis in CCA and thus, the co-expression of SOX2/OCT4 may be a used as a specific marker of CCA CSC.

SOX2/OCT4 express and are functional in a subpopulation of CCA cells

Given these results, we used the previously published SOX2/OCT4-responsive reporter system SORE6-dsCopGFP[31], for further investigation of the CSC population in CCA. To this end we transduced five CCA cell lines, KKU-055, KKU-213B, SSP-25, TFK-1 and KKK-D068 with the SORE6-dsCopGFP (Fig. 1b) and analyzed by fluorescence microscopy (Fig. 1c) as well as by flow cytometry (Fig. 1d, Supplementary Fig. 2a) the temporal expression of dsCopGFP thus indicating SOX2/OCT4 activity in these cells. Under standard culture conditions, the GFP-positive cells (called as SORE6^{POS} cells) ranged from 1.84 to 13.10% of the cells in the CCA cell lines (Fig. 1e). Furthermore, to characterize the isolated SORE6^{POS} and SORE6^{NEG} cells, we used RT-PCR to analyze the expression of *SOX2* and *OCT4*. We found that SORE6^{POS} cells generally displayed enhanced expression for both stemness genes compared to SORE6^{NEG} cells (Supplementary Fig. 2b).

Differentiation and dedifferentiation potential of SORE6^{POS} CCA cells

Since the SORE6^{POS} cells have been reported as CSC-like cells[27,31,32], we then studied, whether our SORE6^{POS} CCA cells were CSCs. We treated them with a differentiation inducer all-*trans* retinoic acid (ATRA), which has been shown to induce differentiation in CSCs[33–35]. Three days of ATRA treatment induced a dose-dependent reduction in the percentage of SORE6^{POS} cells (Fig. 2a). Importantly, applying the same experimental setup on fluorescence-activated cell sorting (FACS)-sorted SORE6^{POS} cells (Supplementary Fig. 3), the number of SORE6^{POS} cells decreased while the SORE6^{NEG} cells increased, indicating differentiation of SORE6^{POS} to SORE6^{NEG} cells (Fig. 2b).

Then, we examined the potential of SORE6^{POS} cells to spontaneously differentiate, which has been described as one of the hallmarks of CSCs[31,36]. FACS-sorted SORE6^{POS} TFK-1 and KKK-D068 cells were grown under standard culture conditions and the GFP expression following each passage was analyzed (Fig. 2c, d). After sorting, all of the TFK-1 and KKK-D068 cells were SORE6^{POS} (Fig. 2c, P.0, dark line compared to Fig. 2c, pre-sort, dark line). Following passage 1, number of cells with high GFP reduced (Fig. 2c, P.1). The GFP signal further reduced after 2 passages (Fig. 2c, P. 2). Similar results were observed for the KKK-D068 cells (Fig. 2d). Single cell analyses of a SORE6^{POS} cell over a period of 36 h showed that a SORE6^{POS} cell divided into daughter cells with reduced GFP expression (Fig. 2e, and Supplementary video 1). To validate the spontaneous differentiation potential of the CCA CSCs in the more physiological condition we grew sorted SORE6^{POS} cells in three-dimensional (3D) spheroids. Although, we started the 3D culture with nearly pure SORE6^{POS} cells, we found

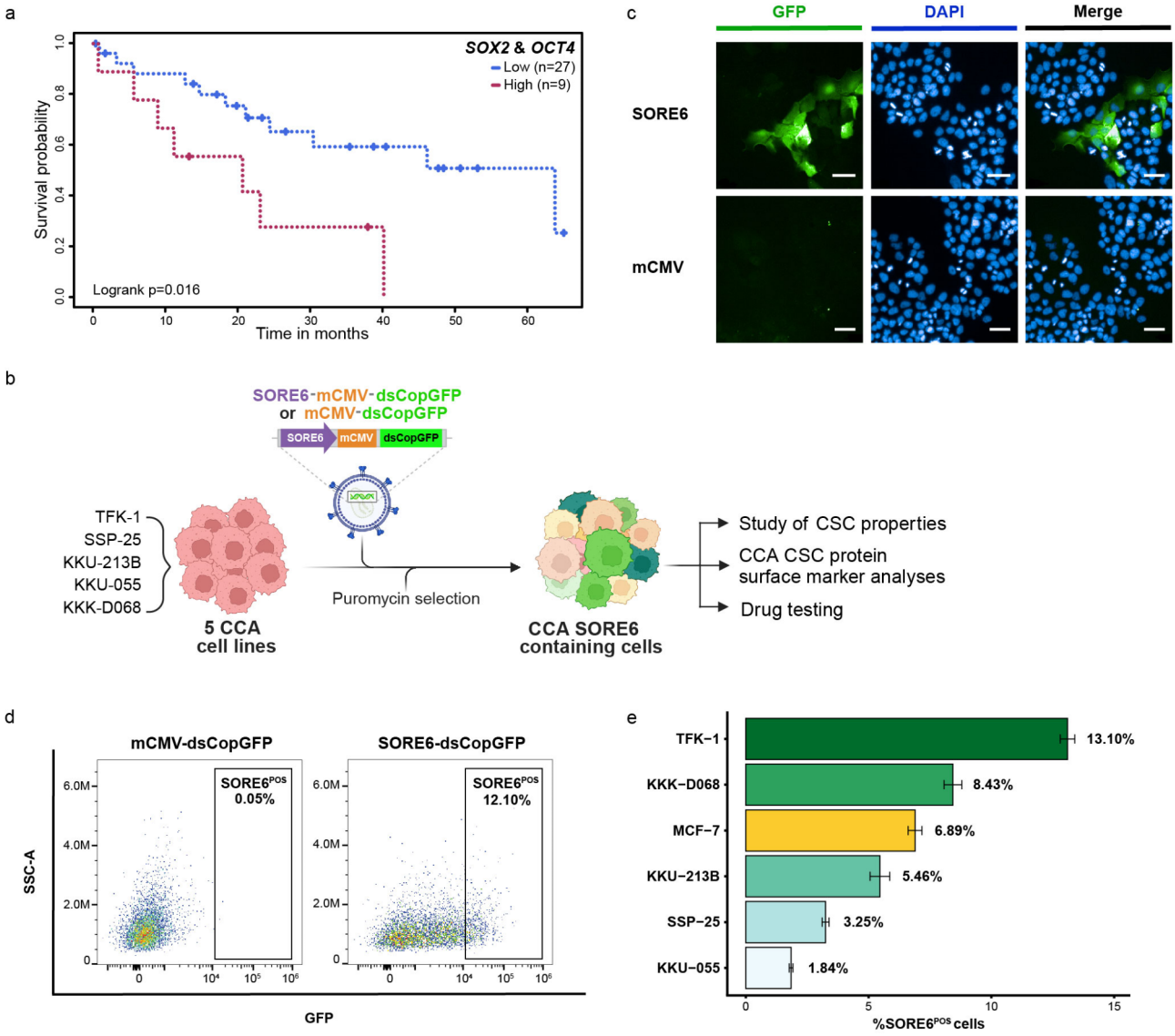


Fig. 1. SOX2 and OCT4 are relevant drivers for stem cell characteristics in CCA. (**a**) Kaplan-Meier survival analyses show an overall survival of patients based on combined expression levels (low; blue vs. high; red line) of *SOX2* and *OCT4* transcripts. (**b**) Constructions of CCA cell lines containing the GFP-based live reporter (SORE6-dsCopGFP). Created by BioRender.com / Mahidol University. (**c**) Spatial expressions of GFP in TFK-1 cells containing the SORE6-dsCopGFP (SORE6) under standard culture conditions. TFK-1 cells transduced with mCMV-dsCopGFP (mCMV) reporter were used as controls. Scale bar = 50 µm. The images were taken using a 10x objective lens. (**d**) Identification of GFP-expressing subpopulation ($SORE6^{POS}$ cells) (right panel) by flow cytometry. The same cell line transduced with mCMV-dsCopGFP (left panel) was used to identify the background. (**e**) Averaged percentages of the $SORE6^{POS}$ cells in different cell lines were determined by flow cytometry. The data represent mean ± s.d.; $n = 3$.

that, at day 14 the majority of the cells in the 3D culture no longer expressed GFP (Fig. 2f), thus consistent with the results in the two-dimensional (2D) culture setup.

It has been proposed that, under permissive conditions, CSCs may be induced from non-CSCs by dedifferentiation[20,37]. To address this, we FACS-sorted GFP-negative ($SORE6^{NEG}$) CCA cells and cultured these cells for two passages. We found that starting from passage 0 post-sorting $SORE6^{NEG}$ TFK-1 cells, comprising 99.82% $SORE6^{NEG}$cells, increasing numbers of $SORE6^{POS}$ cells were detected at later passages (1.28% and 5.20% $SORE6^{POS}$ cells emerged in passage 1 and 2, respectively) (Fig. 3a, b, and Supplementary Fig. 3). The number of $SORE6^{POS}$ cells following passage 2 was comparable to the number of $SORE6^{POS}$ cells in the pre-sort TFK-1 cells (5.71%). Similar results were observed for another CCA cell line, KKK-D068 (Fig. 3c, d). To confirm these findings, live-cell imaging was performed in TFK-1 cell line to demonstrate the dedifferentiation from $SORE6^{NEG}$ to $SORE6^{POS}$ cells (Fig. 3e, and Supplementary video 2). Specifically, we observed $SORE6^{POS}$ cells emerging from $SORE6^{NEG}$ cells. We also found that $SORE6^{POS}$ cells can be generated from $SORE6^{POS}$ cells,

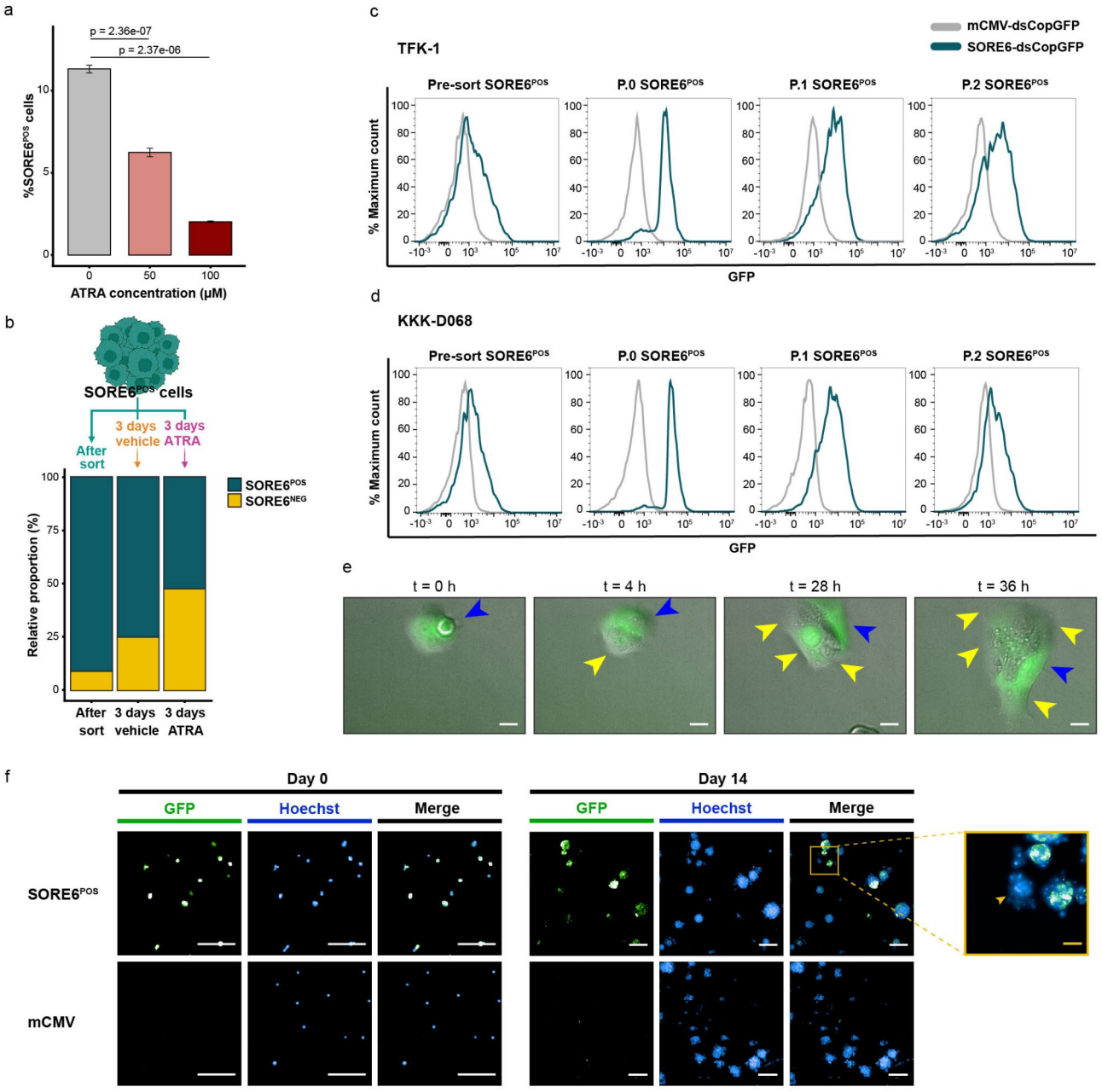


Fig. 2. SOX2/OCT4 activity in SORE6^{POS} cells is transient and is suppressed by ATRA. (**a**) Suppression of SORE6^{POS} TFK-1 cells following 3-days treatment with ATRA. The data represent mean ± s.d.; $n = 3$. (**b**) Percent SORE6^{POS} and SORE6^{NEG} cells of TFK-1 cells after 3-days treatment with 50 μM ATRA or vehicle control (0.5% DMSO) of the FACS sorted SORE6^{POS} cells. A representative of three independent experiments analyzed by flow cytometry is shown. The data represent mean ± s.d.; $n = 3$. Created by BioRender.com / Mahidol University. (**c**) Spontaneous differentiation of SORE6^{POS} to SORE6^{NEG} cells. Sorted SORE6^{POS} TFK-1 and (**d**) KKK-D068 cells were followed over 2 passages. FACS analyses were performed to determine the numbers of SORE6^{POS} cells. P.0, immediately after sorting; P.1, first passage after sorting; P.2, second passage after sorting. (**e**) Live cell imaging was performed to observe spontaneous differentiation of a SORE6^{POS} TFK-1 cell. Scale bar = 20 μm. Yellow arrowheads indicate daughter cells, and blue arrowheads indicate the mother cell. (**f**) Emerging of SORE6^{NEG} cells from sorted SORE6^{POS} cells in TFK-1 3D culture after 14 days of culture. White scale bar = 200 μm, yellow scale bar = 50 μm. The images were taken by a 20x air objective lens. ATRA: all-*trans* retinoic acid.

resembling the self-renewal process in stem cells (Fig. 3f, and Supplementary video 3). Thus, by our live-reporter system we demonstrated that a subset of CCA cells temporarily express functional core pluripotent factors SOX2/OCT4, and that GFP-expressing SORE6^{POS} cells can be generated from SORE6^{NEG} cells and vice versa. In addition, SORE6^{POS} cells can also be regenerated from other SORE6^{POS} cells. Taken together these results point

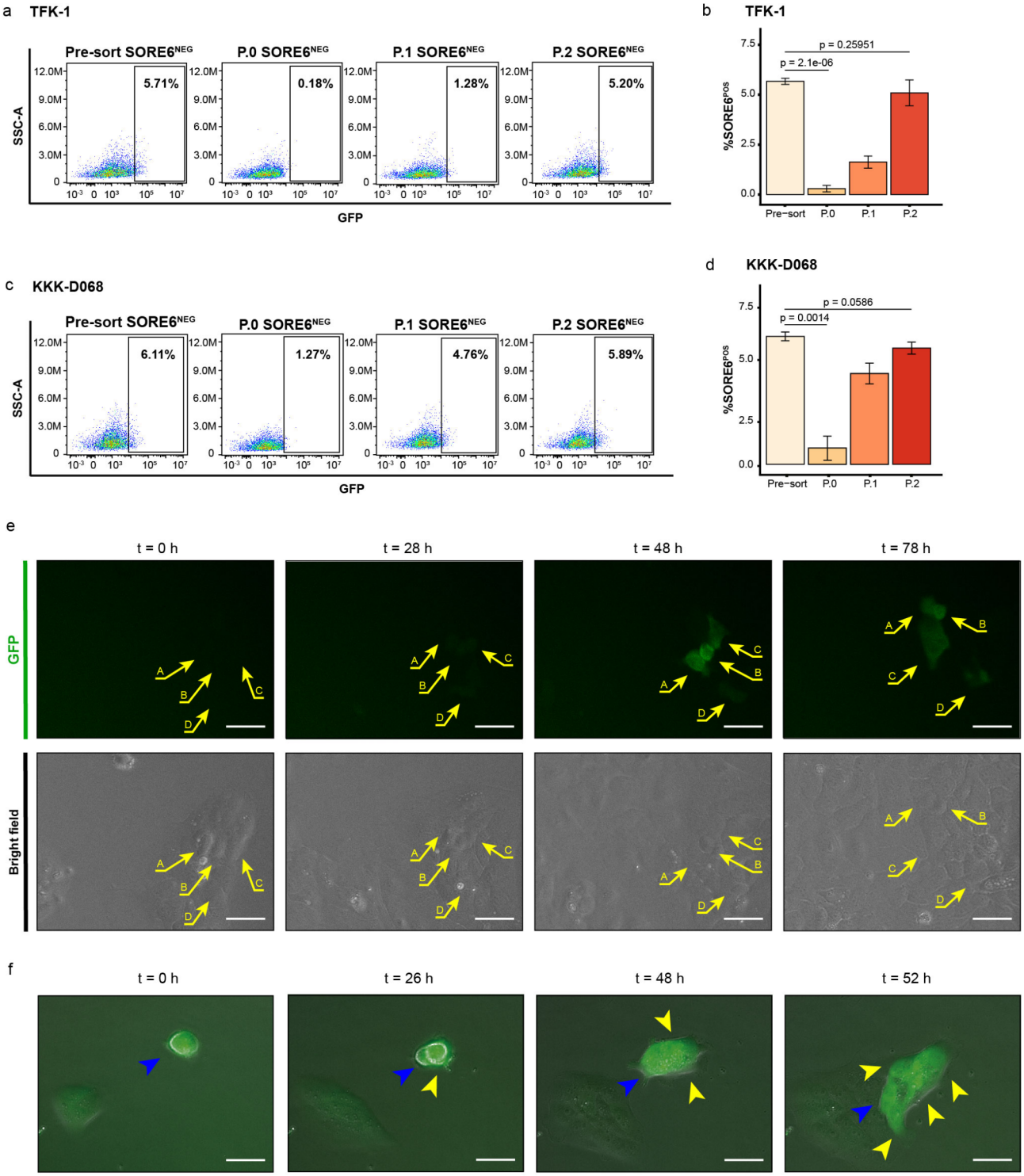


Fig. 3. SORE6^{POS} cells emerge from SORE6^{NEG} CCA cells. FACS-sorted SORE6^{NEG} TFK-1 (**a**,**b**) and KKK-D068 cells (**c**,**d**) were grown as monolayer cultures. Percentages of SORE6^{POS} cell in indicated passages were determined by flow cytometry. The bar graphs represent mean ± s.d.; $n = 3$ from the independent experiments. P.0, immediately after sorting; P.1, first passage after sorting; P.2, second passage. 2D live cell imaging of TFK-1 cells revealed spontaneous dedifferentiation (**e**) and self-renewal (**f**). Scale bar = 50 µm. For (**e**), A, B, C, and D indicate different cells. For (**f**), yellow arrowheads indicate daughter cells, and blue arrowheads indicate the mother cell.

to the existence of a steady state level of a CSC cell population in CCA, and that this equilibrium is maintained due to the plasticity of the CCA CSC cells.

SORE6^{POS}cells show increased tumorigenic capability in both in vitro cultures and in vivo xenograft models

To assess the tumorigenicity of SORE6^{POS} CCA cells, both in vitro and in vivo tumorigenesis assays were performed (Fig. 4a). The in vitro tumorigenic assay was performed by evaluating the potential of cells to grow independently from the surface attachment (anchorage-independent growth). We found that FACS-sorted SORE6^{POS} TFK-1 cells grew at significantly higher numbers of spheroids of larger size, compared to SORE6^{NEG} cells (Fig. 4b-e, and Supplementary Fig. 4a). Furthermore, signal intensities were positively correlated with the size of the spheroids (Fig. 4f). To assess the tumorigenic potential of the SORE6^{POS} cells in vivo, mouse xenograft experiments were performed. The SORE6^{POS} TFK-1 and KKK-D068 cells showed significantly higher growth rates, than the SORE6^{NEG} counterparts (Fig. 4g, h). We analyzed SOX2, and OCT4 expressions from the mouse xenografts collected at the endpoint and found that both SORE6^{POS} (initially implanted with SORE6^{POS} cells) and SORE6^{NEG} (initially implanted with SORE6^{NEG} cells) groups express comparable levels of SOX2, and OCT4 protein, by IHC (Supplementary Fig. 4b). We explained that the plasticity of the CCA CSC caused the number of CSC in the tumor from both groups to adjust to a steady level, after a long period of tumor implantation in the mice.

These results indicated that the SORE6^{POS} population comprised increased tumorigenic potential compared to SORE6^{NEG} cells. Thus, these results indicated that the SORE6^{POS} CCA cells possess several properties characteristic of CSCs.

SORE6^{POS} CCA characterizes a rare population which coincides with CD133^{+} CCA cells

Ongoing research focuses on identifying specific markers for CCA CSCs; the most prominent candidates of which include CD44, CD133, LGR5, and EpCAM[11–14]. However, no consensus on the putative CSC marker(s) has been made as yet. We therefore explored the activities of SOX2/OCT4 in CCA cells expressing CD44, CD133, LGR5, and EpCAM by multiplex flow cytometry. t-distributed Stochastic Neighbor Embedding (t-SNE) analyses show the cells, positive for these CSC markers together with SORE6^{POS} cells in 5 different CCA cell lines (Fig. 5a). Each CCA cell line expressed the surface proteins differently, i.e. expression of EpCAM was found in most of the CCA cells in all cell lines (percentages of marker-positive cells are indicated in the t-SNE maps). CD44 expressions were found in most of the CCA cells, except KKU-055, in which CD44 was found in only a small subpopulation (around 1%). EpCAM and CD44 alone did not appear to have discriminative power to distinguish CCA CSCs. These results were consistent with previous reports[11]. LGR5^{+}, a proposed marker for CSC in gastrointestinal cancer[38] expressed differently in each cell line. Interestingly, CD133^{+} expressed as clusters in all CCA cell lines. These data suggest that a single marker does not hold enough discriminative power to classify CCA CSCs.

To investigate whether our SORE6^{POS} cells expressed CSC-associated surface markers we found that only a few of the CCA cells positive for CSC markers expressed a strong GFP signal (Supplementary Fig. 5a-b, dark blue). However, when CCA cells with weaker GFP signal (SORE6^{WEAK}) were included, (Supplementary Fig. 5a-b, dark blue + green), we observed that a high percentage of CD133^{+} CCA cells was SORE6^{POS} (ranging from 95.36% in TFK-1 to 25.29% in KKU-055; mean ± s.d. = 72.08% ± 27.30, Fig. 5b-e). CCA cells with other CSC markers contained lower numbers of SORE6^{POS}, i.e. 39.62% ± 15.75 of EpCAM^{+} CCA cells were SORE6^{POS} cells (ranging from 52.23% in TFK-1 to 14.06% in KKU-055 cells, Fig. 5b-e).

When co-expressions of multiple CSC markers were analyzed, we found that most of the CCA cells expressing double markers, CD133^{+}CD44^{+}, and CD133^{+}LGR5^{+}, were SORE6^{POS} (Fig. 5f-h), averaging around 80% of the double-marker-positive populations across the 5 cell lines (Fig. 5g). These sub-populations constituted around 0.74% in average (ranging from 0.11 to 1.45%) of the CCA cells in culture (Fig. 5i). When subsets of CCA cells with triple or quadruple CSC markers were analyzed, the CD133^{+}CD44^{+}EpCAM^{+}, and CD133^{+}CD44^{+}LGR5^{+}EpCAM^{+} showed very high percentages of cells expressing GFP (Supplementary Fig. 6a-b, d-e). However, these triple and quadruple-positive cells constituted a very small portion of the CCA in culture (0.27% in average, 0.02–1.28%) (Supplementary Fig. 6c, f). These results point to CD133-positive cells representing a CCA population with active SOX2/OCT4. The results also suggest that there might be subpopulations of CSCs with different surface markers that constitute the CCA CSCs, and that more than one surface marker are required to delineate subpopulations of CCA CSCs.

CSC-like SORE6 ^{POS} cells are more resistant to chemotherapy

Since CSCs have been proposed as a key driver for chemoresistance in CCA, we investigated the impact of standard CCA-chemotherapies (gemcitabine, cisplatin, and 5-fluorouracil [5-FU]) in our CCA cell lines containing the live reporter. By high-content imaging, dose-response curves, GR values, and SORE6^{POS} percentages were analyzed (Fig. 6a-c, and Supplementary Fig. 7a). We found that as concentrations of gemcitabine, gemcitabine + cisplatin, and 5-FU increased, the viability of CCA in general decreased, whereas the percent of SORE6^{POS} cells increased. We found that GR_{50}s of SORE6^{POS} cells were higher than SORE6^{NEG} for all the drugs tested except for cisplatin, for which SORE6^{POS} and SORE6^{NEG} had comparable GR_{50}s (Fig. 6d-f, Supplementary Fig. 4a, 7b). At the end of the treatment period, SORE6^{POS} populations were enriched in all cultures (Fig. 6g), indicating that the CSC-like SORE6^{POS} cell population is more resistant to certain chemotherapies, and increases over the time.

We next asked how these CSC-like population increase under the chemotherapy. To address this, SORE6^{POS} and SORE6^{NEG} populations were FACS-sorted and treated with 5-FU (Fig. 6h, i and Supplementary Fig. 3). At day 3 after sorting, 25% of the FACS-sorted SORE6^{POS} population treated with vehicle control spontaneously loosed the GFP signal, and became SORE6^{NEG} (Fig. 6h, vehicle). Interestingly, 5-FU treatment inhibited the

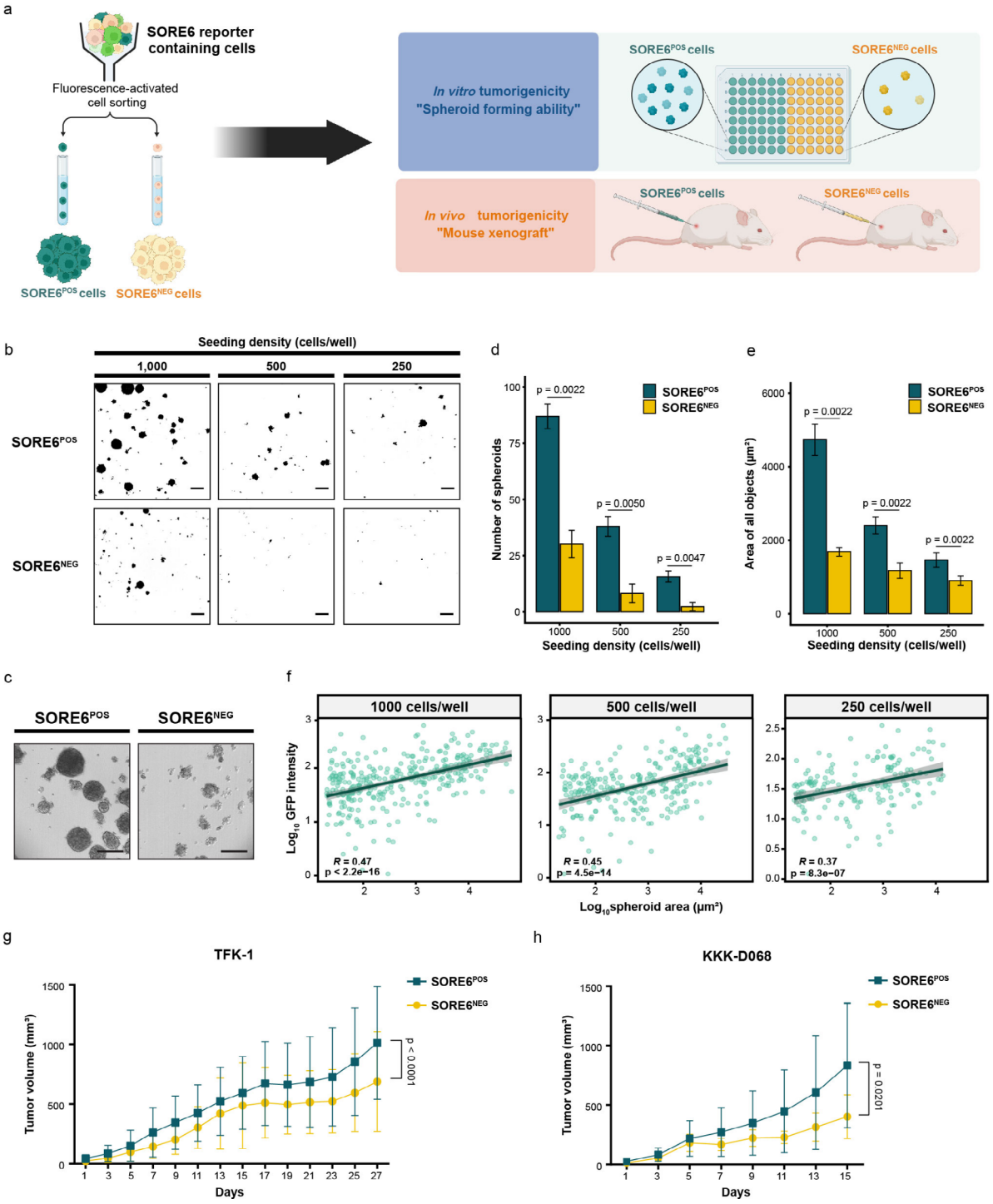


Fig. 4. High tumorigenic potency in SORE6^{POS} cell. (**a**) Schematic diagram of the in vitro and in vivo tumorigenic assay. Created by BioRender.com / Mahidol University. (**b**) In vitro tumorigenicity was tested by growing the FACS-sorted SORE6^{POS} or SORE6^{NEG} TFK-1 cells into spheroids. Scale bar = 300 μm. Representative images of the spheroid morphology from initial seeding density 1,000 cells/well shown in (**c**). Scale bar = 200 μm. Numbers and sizes of the spheroid were quantified by image analysis and shown in the bar graphs (**d**,**e**). The data represent mean ± s.d.; $n = 6$. P-values for in vitro tumorigenicity were analyzed by the Wilcoxon rank sum test. (**f**) Positive correlations between $\log_{10}$ spheroid area vs. $\log_{10}$ GFP intensity. The regression line (black line) was estimated by linear regression of Pearson correlation with 95% confidence interval (gray band). R = Pearson's correlation coefficient, p = p-value. The data was acquired from image analysis. (**g**, **h**) the sorted SORE6^{POS} and SORE6^{NEG} cells were implanted into severe immunocompromised mice. Line graphs indicate growth ± s.d. of the TFK-1 (**g**) and KKK-D068 (**h**) tumors. $n = 10$ for TFK-1 tumors, and $n = 8$ for KKK-D068. P-values for in vivo data were analyzed by paired t-test.

transition of GFP-positive to GFP-negative population (Fig. 6h, 5-FU). On the other hand, around 10% of SORE6^{NEG} cells spontaneously transformed to SORE6^{POS} (Fig. 6i, vehicle) and here 5-FU treatment promoted the transformation of SORE6^{POS} to 25% (Fig. 6i, 5-FU). We observed similar results in 3D cultured CCAs, i.e., 5-FU treatment increased the SORE6^{POS} cell population (Fig. 6j). These results indicate that 5-FU treatment may promote CSC by stabilizing the SORE6^{POS} population and facilitating SORE6^{NEG} to SORE6^{POS} transformation.

CDK4/6 inhibitors suppress CSC-like CCA populations

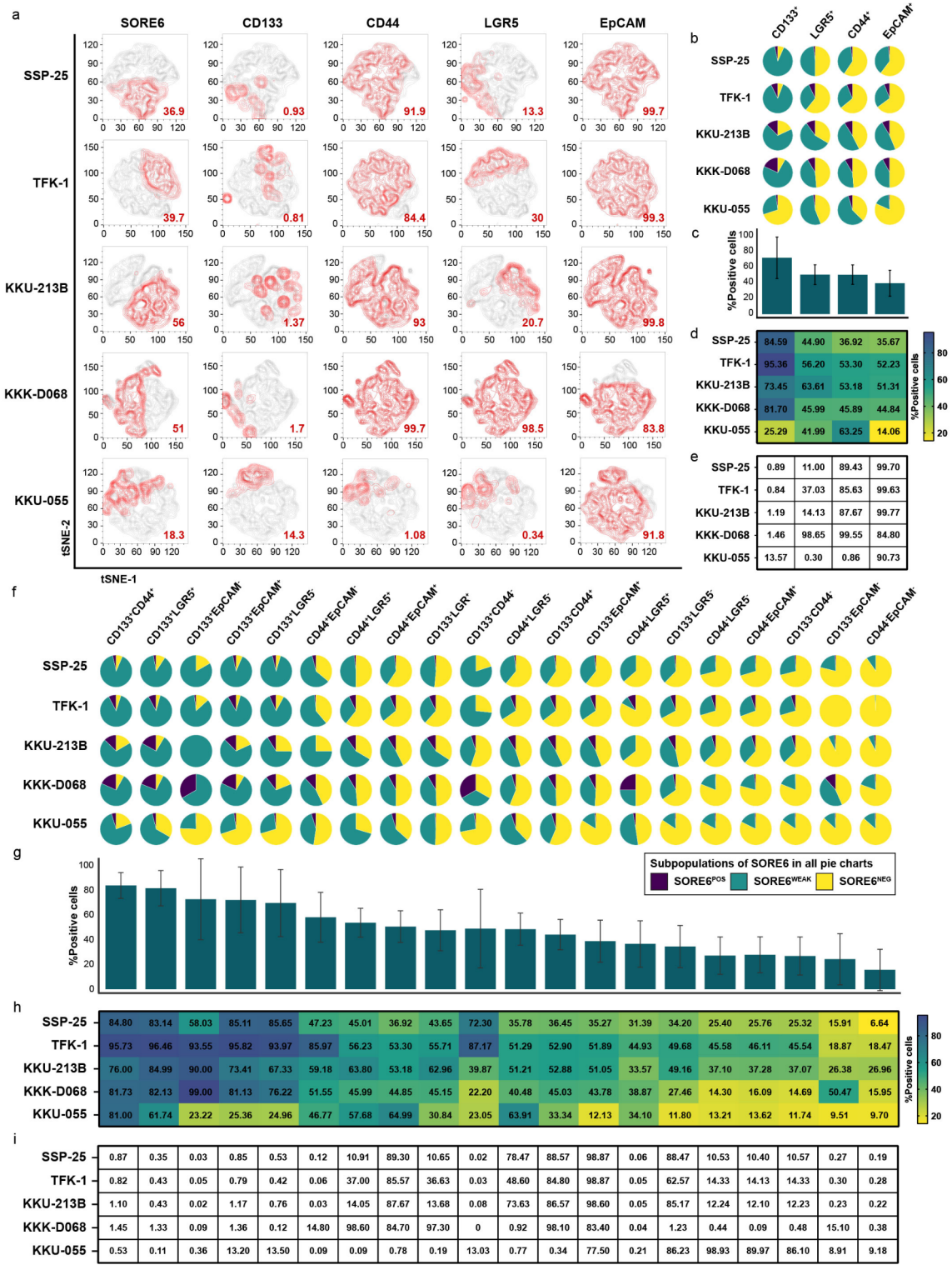
CDK4/6 inhibitors have been tested in CCA as either mono- or combination-therapy[39–43] and interestingly, CDK4/6 inhibitors were shown to suppress SOX2 and OCT4 levels by means of protein destabilization[44]. Using our live reporter system, we examined the effect of CDK4/6 inhibitors on the CCA CSCs. We found that the CDK4/6 inhibitors, abemaciclib (AB) and palbociclib (PD), suppressed the number of SORE6^{POS} cells in a dose-dependent manner (Fig. 7a-b), and following 3 days of treatment with CDK4/6 inhibitors, the percentage of SORE6^{POS} cell in the culture decreased from 12 to 4% (Fig. 7c). To track the reduction of SORE6^{POS} cells, we analyzed FACS-sorted SORE6^{POS} and SORE6^{NEG} CCA cells in culture during CDK4/6 inhibitor treatment (Fig. 7d, e and Supplementary Fig. 3). We found that 25% of SORE6^{POS} CCA cells spontaneously lost their GFP expression and became SORE6^{NEG} after 3 days in culture (Fig. 7d, vehicle). Addition of PD further increased the SORE6^{NEG} population (Fig. 7d, PD). Moreover, we found that in the SORE6^{NEG} population, CDK4/6 inhibition suppressed the spontaneous transition of SORE6^{NEG} to SORE6^{POS} cells (Fig. 7e, PD). Lastly, we compared the CSC-like SORE6^{POS} number for all treatments, using area under the curves (AUCs) from the normalized SORE6^{POS} percentages. Chemotherapies (Fig. 7f, Gem, 5-FU, Gem + Cis) appeared to increase AUC, while CDK4/6 inhibitors decreased the AUC (Fig. 7f, AB and PD). In conclusion, these data show that applying a live reporter system we have identified CDK4/6 inhibitors as a potential CSC reduction therapy in CCA.

Discussion

The study of CCA CSCs helps to advance our understanding on CCA carcinogenesis and management. However, CSCs are scarce, plastic, and exhibiting different characteristics and levels of differentiation and rendering bulk or snapshot-based study approaches ineffective. In addition, although the functional characterization implied that CSCs may comprise different hierarchical groups resembling that of the stem cell hierarchy, it has been difficult to physically isolate and study these cells because of the lack of specific markers. Here, the live biosensor SORE6-dsCopGFP reporter, which responds to SOX2 and OCT4 transcriptional function has been employed to study CCA CSCs in five different CCA cell lines. It is well documented that SOX2 and OCT4 are high in the stemness hierarchy and have been shown to be drivers in several types of CSC[45,46]. In addition, from the TCGA-CHOL data set, dual expression of *SOX2* and *OCT4* significantly correlated with the poor survival of CCA patients. Due to the limited number of patients in the cohort, the long-term survivals of the patients beyond 60 months was not available. In addition, in the *CD44*-High and *LGR5*-High CCA, the data were only available until around 50 months, possibly due to the loss of follow-up. Short half-life of dsCopGFP was designed to detect the temporal change in the SOX2/OCT4 activities[31]. Wild-type GFP protein has a half-life of approximately of 26 h, however, addition of peptides from cyclin B1 destructive box and PEST sequence from the mouse ornithine decarboxylase gene reduced the half-life of dsCopGFP to 5.5 h, especially suitable for detecting activity of transcription factor such as SOX2 and OCT4[47]. While we were working on our project, an independent group designed the same SOX2/OCT4-based reporter to study interaction between macrophages and CCA CSC. They reported that that SORE6^{POS} cells generally displayed enhanced expression for both stemness genes and CSC genes compared to SORE6^{NEG} cells. Thus, these data indicate that the SORE6-dsCopGFP reporter can be used to identify cells with stem-like characteristics[48].

We demonstrated that the SORE6^{POS} cells showed several genuine CSC characteristics, e.g. resistance to chemotherapy, increased spheroid formation, self-renewal, and high tumorigenesis capability in immunocompromised animals. Thus, we suggested that the SORE6^{POS} cells were CSC cells. Interestingly, applying the SORE6 reporter system, we were able to demonstrate plasticity in these cells, and moreover they were able to self-renew and dedifferentiate which might be important to maintain the stem cell number in the culture. This observation might also explain one mechanism by which the cancer stem cell population is maintained in tumors.

Several surface markers have been proposed to identify CCA CSCs. When analyzing for these markers we showed - as expected - that SORE6^{POS} cells were enriched in CCA cells expressing CD133^{+}. When multiple markers were analyzed, we found a higher enrichment of SORE6^{POS} cells in both the CD133^{+}CD44^{+} and the CD133^{+}LGR5^{+} subpopulation. Thus, our results suggest that there may be several minor subpopulations of CCA CSCs in the culture, as proposed by a previous study[49]. Interestingly, we found highly differential expression of CSC markers among our 5 CCA model cell lines. These results implicate existence of different types of CSCs and the heterogeneity of CCA CSC in the cell lines. These results were more evident, especially when CCA cells with weak GFP expression (SORE6^{WEAK}) were included in the study (Fig. 5). These cells with low SOX2/OCT4 activities may implicate different stages of progenitor or progenitor-like cells within the CSC population[50]. We hypothesize that the SORE6^{WEAK} CCA cells contained lower levels of SOX2/OCT4 compared to the SORE6^{POS} cells, and that SORE6^{WEAK} and SORE6^{POS} may be distinctive stem cell populations. It has been demonstrated that variable expression levels of SOX2 and OCT4 regulates the hierarchy of stem cells[51,52]. For example, in the ventricular zone of the developing mouse cortex, SOX2 is more highly expressed in the slowly dividing, multipotent, radial glia cells than in the rapidly dividing intermediate progenitor cells, which are committed to neurogenesis[52,53]. Whether or not, the SORE6^{WEAK} and SORE6^{POS} cells may be of CSCs with different potential requires further investigation. However, low numbers of cells in the minor subpopulations prevented us from isolating and growing these for further studies. It will be informative when these single cells with varying levels of SOX/OCT4 activities are studied in conjunction with other single cell technologies, such as single-cell DNA/



RNA-sequencing, and mass cytometry (CyTOF). Anatomical subtypes of CCA, i.e. extrahepatic, perihilar, and intrahepatic, show clinical and biological differences[54]. Likewise, the expression of a given CCA CSC marker can be varied between the subtypes[55]. Since, our cell line models, namely SSP-25, KKU-213B, KKK-D068, KKU-055 are of intrahepatic origin, except for TFK-1 which is derived from extrahepatic CCA, the results from this study should need further validation, especially for cell lines of extrahepatic and perihilar origin.

Our live reporter CCA cells allowed us to study the role of CSC following drug treatments[5]. We showed that under standard CCA chemotherapies, CCA CSCs are relatively resistant. Thus, enrichment of CSCs was observed after only 3 days of treatment. However, we also demonstrated that CSC enrichment was not necessarily observed following all treatments. Cisplatin was found not to enrich SORE6^{POS} cells. We showed that unlike chemotherapy, CDK4/6 inhibition depleted CCA CSCs. Supportive roles of G1 CDKs in embryonic and adult stem cells have been established by several groups[44,56,57]. Recent works showed that CDK4/6 inhibition

◀**Fig. 5**. Expression of selected CSC-associated markers on SORE6 cell subsets in five CCA cell lines. (**a**) t-SNE diagrams of SORE6^{POS} cell immuno-stained for putative surface protein markers as indicated. Red and grey areas represent marker positive or negative, respectively. The red numbers show the percentage of surface marker-positive cells. (**b**) Pie charts show percentages of SORE6 cell subsets in single surface marker-defined CCA subpopulations, as indicated on the top of the panel. Strong GFP expression (SORE6^{POS}; dark blue), weak GFP expression (SORE6^{WEAK}; green), and no GFP expression (SORE6^{NEG}; yellow). (**c**) The bar graphs show the percentage of SORE6 cells (strong and weak signals) in single surface marker-defined CCA subpopulations as indicated in (b). (**d**) Heatmap shows averages of percent SORE6 cells (strong and weak signals) in single surface marker-defined CCA subpopulations as indicated in (b), broken down in detail for each cell line. (**e**) Percentages of cells in single surface marker-defined subpopulations. (**f**) Pie charts show percentages of SORE6 cell subsets in double surface marker-defined CCA subpopulations, as indicated at the top of the panel. Strong GFP expression (SORE6^{POS}; dark blue), weak GFP expression (SORE6^{WEAK}; green), and no GFP expression (SORE6^{NEG}; yellow). (**g**) The bar graphs show the percentage of SORE6 cells (strong and weak signals) in double surface marker-defined CCA subpopulations as indicated in (f). (**h**) Heatmap shows averages of percent SORE6 cells (strong and weak signals) in double surface marker-defined CCA subpopulations as indicated in (f), broken down in detail for each cell line. (**i**) Percentages of cells in double surface marker-defined subpopulations. All pie charts were generated from a representative experiment. Each bar in all bar graphs showed average number from 5 CCA cell lines ± s.d. from three biological replicates ($n = 15$).

mitigate expressions of stem cell factors and CSCs in several types of cancer, such as breast cancer, esophageal cancer, acute myeloid leukemia[58–62]. We demonstrated here that inhibition of CDK4/6 activities by palbociclib or abemaciclib not only blocks CCA, but also depletes the CCA CSC population. This information can be used to design novel therapy that target both CCA and the CSCs.

In conclusion, by applying a live reporter-based system to study CSCs in different CCA cell lines we have been able to delineate the plasticity as well as drug-induced dynamics of CSCs. Since CCA is a highly cytoreduction resistant disease these findings have clear implications for the management of the patients and could prove useful in the development and efficacy screening of novel drugs.

Materials and methods

Cell culture

Four CCA cell lines of intrahepatic origin; KKU-055, KKK-D068, KKU-213B, SSP-25, together with an extrahepatic cell line TFK-1[42,63,64] were obtained from the Japanese Collection of Research Bioresources (JCRB, Osaka, Japan). For population types, 5 CCA cell lines 3 were from Thai CCA (KKU-055, KKK-D068, and KKU-213B) and 2 Japanese CCA (SSP-25 and TFK-1). For anatomical classification, KKU-055, KKK-D068, KKU-213B, and SSP-25 are intrahepatic CCA, while TFK-1 are extrahepatic CCA[64]. About drug sensitivity, all CCA cell lines were sensitive to palbociclib[39]. KKK-D068, KKU-055, TFK-1, and SSP-25 were sensitive to most of chemotherapy including standard treatments[64]. KKU-055, KKK-D068, KKU-213B, and HEK293T (American Type Culture Collection) were maintained in Dulbecco's modified Eagle's medium (DMEM), while Roswell Park Memorial Institute 1640 Medium (RPMI 1640) was used for SSP-25 and TFK-1. Each culture medium was supplemented with 10% fetal bovine serum (FBS) (Thermo Fisher Scientific) and 1% penicillin/streptomycin. For routine subculturing, the cells were washed once with phosphate-buffered saline (PBS), then 0.25% trypsin-EDTA (Thermo Fisher Scientific) to detach the cells. The culture medium was added onto the plate as volume ratio by 9 culture medium to 1 trypsin-EDTA to inhibit trypsinization. All cell lines were maintained at 37 °C in a humidified atmosphere containing 5% CO_2 and routinely tested for mycoplasma by PCR.

Lentivirus production, cell transduction, and construction of CCA cell lines with live reporter for CSC

The lentiviral-based stem-cell reporter system[31] was a kind gift from Dr. Lalage M. Wakefield, NCI, USA. In the system, six tandem repeats of a composite OCT4/SOX2 response element (SORE6) coupled with minimal cytomegalovirus (mCMV) promoter are used to drive expression of the green fluorescent protein (GFP) reporter. The destabilized copepod GFP was used in this construct to ensure the temporal resolution (SORE6-dsCopGFP). A similar construct without SORE6 (mCMV-dsCopGFP) was used as flow cytometric gating control and background control of fluorescence microscope. For more detail of the construct please refer to reference[31]. To generate lentiviral reporters, either SORE6-dsCopGFP or mCMV-dsCopGFP plasmid was co-transfected with the second-generation packaging plasmids composed of psPAX2 and pMD2.G into HEK293T cells, by Lipofectamine™ LTX (Invitrogen,) (1 µg DNA: 1.5 µL lipofectamine ratio) in serum- and antibiotic-free Opti-MEM medium (Thermo Fisher Scientific), for 8 h. Then, the medium was replaced with fresh HEK293T culture medium and incubated for 48 h. The viral supernatant was collected and filtrated (0.45 µM filter), then mixed with 5 µg/mL polybrene and fresh culture medium in 1:1 ratio. To transduce the viral reporter, the cells were exposed to the viral mixture for 24 h. Puromycin selection was performed for 5 days in the sub-confluent CCA cultures.

Reverse transcription polymerase chain reaction (rtPCR)

RNA extraction, cDNA conversion, and RT-qPCR were performed as previously described[65]. Briefly, RNA was extracted using the Total RNA Purification Kit (GeneMark), and then 1 µg of RNA was converted to cDNA using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). The qPCR was performed

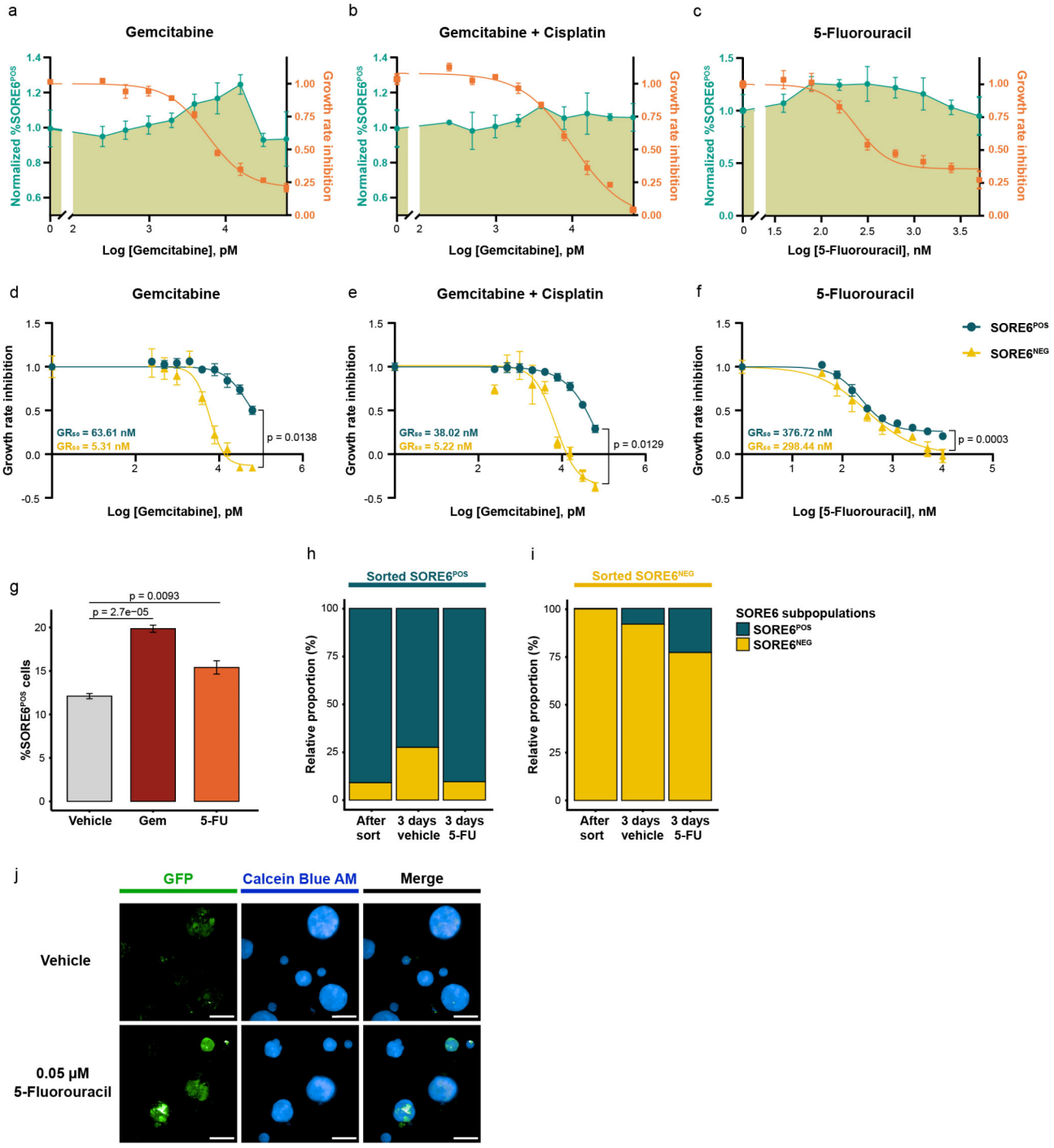


Fig. 6. CSC mediated-drug resistance following chemotherapy. (**a**–**c**) The normalized percentages of SORE6^{POS} TFK-1 cells treated with increasing doses of drugs (green lines, left y-axis). Dose response of total CCA cell numbers are shown in orange lines (right y-axis). The data represent mean ± s.d.; $n = 3$. (**d**–**f**) Dose-response curves for sorted SORE6^{POS} and sorted SORE6^{NEG} TFK-1 cells following chemotherapy. GR_{50}s of each line is shown in the graph. The data represent mean ± s.d.; $n = 3$. The data was acquired from image analysis. (**g**) Percent SORE6^{POS} cells after 3 days of 0.063 μM gemcitabine or 0.313 μM 5-FU. The data represent mean ± s.d.; $n = 3$. Percent subpopulations of sorted SORE6^{POS} (**h**) and SORE6^{NEG} cells (**i**) treated with 0.313 μM 5-FU for 3 days, then quantified by flow cytometry. The data represent mean ± s.d.; $n = 3$. (**j**) TFK-1 multi-spheroids under 0.05 μM 5-FU for 3 days. Scale bar = 100 μm. The data in (**a**–**f**) were acquired from image analysis, while (**g**–**i**) were from flow cytometry. 5-FU: 5-fluorouracil, Gem: gemcitabine.

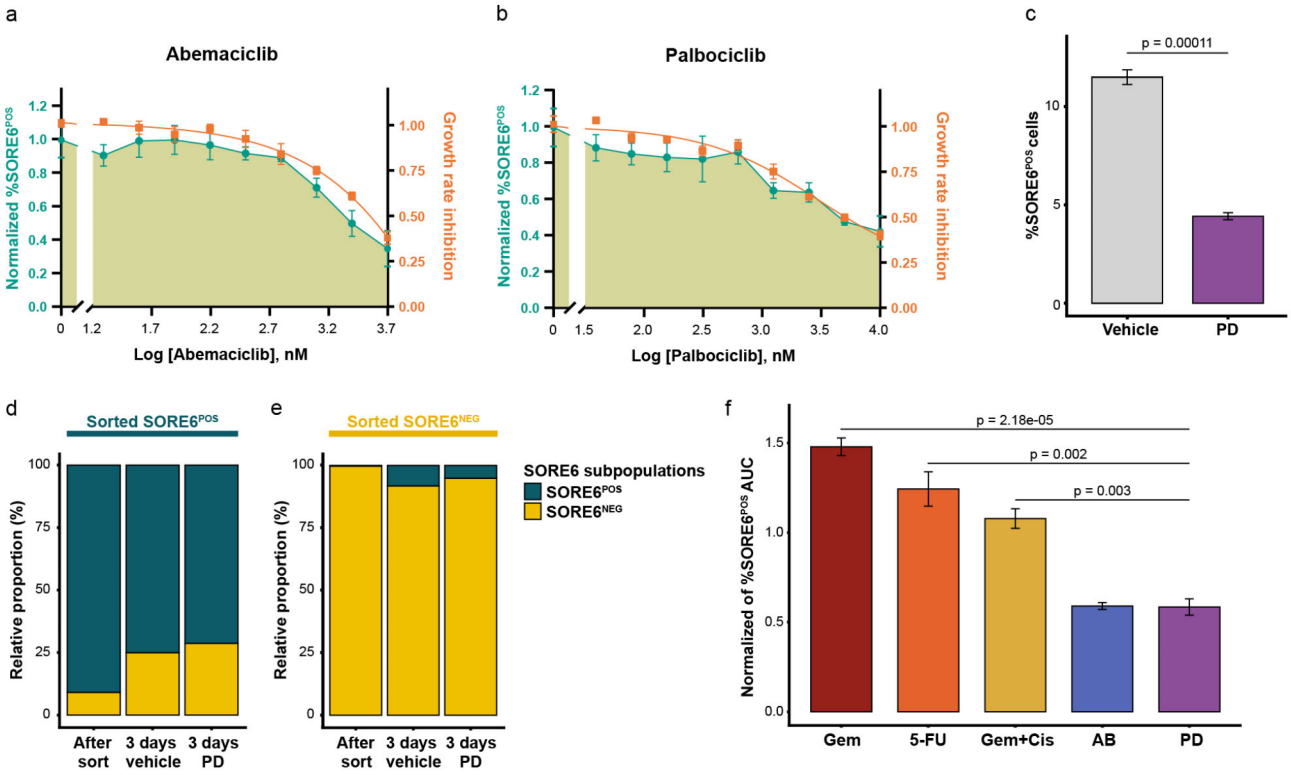


Fig. 7. Effect of CDK4/6 inhibitors on the SORE6^{POS} population. (**a**,**b**) The normalized percentages of SORE6^{POS} TFK-1 cells treated with increasing doses of AB and PD (green lines, left y-axis). Total CCA cell numbers (orange graphs, right y-axis). The data represent mean ± s.d.; $n = 3$. (**c**) Percent SORE6^{POS} cells following 3 days of 5 μM PD treatment. Data represent mean ± s.d.; $n = 3$. Percent subpopulations of sorted SORE6^{POS} (**d**) and SORE6^{NEG} (**e**) cells treated with 5 μM PD for 3 days, then quantified by flow cytometric analysis. The data represented mean ± s.d.; $n = 3$. (**f**) Inductions and suppressions of SORE6^{POS} cells by chemotherapy and CDK4/6 inhibitors. Normalized AUC of percent SORE6^{POS} TFK-1 cells following indicated treatments. The data represent mean ± s.d.; $n = 3$. The data in (**a**,**b**,**f**) were acquired from image analysis, while (**c**–**e**) were from flow cytometry. 5-FU: 5-fluorouracil, AB: abemaciclib, Gem: gemcitabine, Gem + Cis: gemcitabine and cisplatin, PD: palbociclib.

using KAPA™ SYBR® FAST qPCR Master Mix Kit (Kapa Biosystems) on the CFX96™ Real-Time PCR detection system (BioRad). Primers: GAPDH-F 5′-GTCAACGGATTTGGTCGTATTG-3′; GAPDH-R 5′-CATGGGTG GAATCATATTGGAA-3′, SOX9-F 5′- GCTCTGGAGACTTCTGAACGA-3′; SOX9-R 5′-CCGTTCTTCACC GACTTCCT-3′, SOX2-F 5′- GAGCTTTGCAGGAAGTTTGC-3′; SOX2-R 5′- GCAAGAAGCCTCTCCTTG AA-3′, OCT4-F 5′- TCGAGAACCGAGTGAGAGG-3′; OCT4-R 5′- GAACCACACTCGGACCACA-3′. Gene expressions were analyzed relative to induced pluripotent stem (iPS) cells.

Flow cytometry analysis and FACS

The reporter-containing cells were detached from the plate by routine trypsinization, and collected by centrifugation. Then the cell pellets were washed three times with PBS supplemented with 3% FBS and resuspended in the same solution. GFP and Live/Dead Zombie NIR (BioLegend) detection was performed on Cytoflex™ flow cytometer (Beckman Coulter). BD LSR Fortessa™ Cell Analyzer (BD Biosciences) was used for multicolor detection. For cell surface protein analyses, the trypsinization duration was done as short as possible and depending on each cell line to prevent cell surface marker protein destruction. The following fluorescence-conjugated antibodies to cell surface markers for CCA CSC were used; mouse PE anti-human CD133 (#372803), rat PerCP-Cy5.5 anti-mouse/human CD44 (#103031), mouse PE-Cy7 anti-human CD326 (EpCAM) (#324221) all from BioLegend, USA, and rat APC anti-human LGR5, #130-100-85 (Miltenyi Biotec). Following washing with 3% FBS in PBS, cells were blocked by Odyssey blocking buffer (Li-Cor) for 30 min on ice, then stained with the antibody cocktail together with Live/Dead Zombie NIR for 15 min in the dark. Isotype-matched antibodies were used to estimate the non-specific binding; mouse PE IgG1, κ isotype control antibody (#400111), rat PerCP/Cyanine5.5 IgG2b, κ isotype control antibody (#400631), mouse PE/Cyanine7 IgG2b, κ isotype control antibody (#400325), rat APC IgG2b, κ Isotype control antibody (#400611) (all from BioLegend). UltraComp™ eBeads (Invitrogen) were used for compensation. For FACS, the cells were detached and washed as described above. The cells were passed through a cell strainer and kept on ice in the dark not more than 2 h until sorting was performed on a BD FACS Aria™ III Cell Sorter. To increase the purity of GFP-expressing cell, the sorting was repeated in some experiments.

For data analysis, the data from FACS and cell analyzer were analyzed by FlowJo software (version 10.8.1). Cell lines transduced with mCMV-dsCopGFP-transduced were used as controls. To determine SORE6^{POS} cells, the SORE6- dsCopGFP-transduced cells with signal exceeded 99.95% of the cell line match mCMV-dsCopGFP cells will be accepted as SORE6^{POS} cells. The cells that expressed GFP at an intensity between negative and positive populations were indicated as SORE6^{WEAK} cells. In some experiments.

High content imaging

High content imaging was performed at 5x, 10x, and 20x using High-throughput screening (HTS) microplate reader Operetta CLS™ (PerkinElmer). Data were processed and analyzed by Columbus Image Data Storage and Analysis System (PerkinElmer) and R software (version 4.2.2). For image analysis in the confocal mode, the maximum projection was applied. For 2D live cell imaging, Nikon Inverted Microscope ECLIPSE Ti-1 was used at 20x. The image was taken every 2 h for 3–5 days. Nikon Imaging Software (NIS-Elements) was used for imaging processing and analysis.

2D drug testing

Cells were seeded into a 6-well plate (100,000 cells/well), 12-well plate (60,000 cells/well) or 384-well plate (1,000 cells/well) in culture medium, and allowed to grow for 24 h, before treating with gemcitabine (Selleckchem, #S1714), 5-FU (Selleckchem, #S1209), cisplatin (Medchem, #HY-17394), AB (Selleckchem, S7158), palbociclib (Selleckchem, #S1579), or ATRA (Tokyo Chemical Industry, #R0064) at various concentrations for 3 days. 0.5% of Dimethyl sulfoxide (DMSO) or Dimethylformamide (DMF) in culture medium matched the vehicle of the drug.

For flow cytometry, cells were stained with Live/Dead Zombie NIR and analyzed on a Cytoflex™ flow cytometer. Data were shown as cell percentages.

For high-content imaging analysis, cells were fixed and stained with 4% paraformaldehyde (PFA) and Hoechst 33342 (Abcam), respectively. The high-content imaging was performed as stated in the previous section. The number of nuclei was determined by Hoechst 33342. The dose-response curves and growth rate inhibition (GR) were constructed and calculated as described[66]. GFP intensity of the single cell in each well was measured from the images to determine for percentage of SORE6^{POS} cell. Briefly, GFP intensities were measured from nuclear extension areas to avoid incorrect cell segmentation. mCMV-dsCopGFP cells were used as control to determine SORE6^{POS} cells. Normalized %SORE6^{POS} cells were calculated as the percentages of SORE6^{POS} cells of the treatment groups divided by the percentages of cells treated with vehicle. To compare %SORE6^{POS} cells across treatments, AUC of normalized %SORE6^{POS} cells from concentrations between GR_{90} to GR_{40} was used. The normalization was performed by dividing with the AUC of cells treated with vehicle.

In vitro tumorigenesis and 3D drug testing.

The cells were seeded at densities of 1,000, 500, 250 cells/well in a 96 well-plate ultra-low attachment flat bottom (Corning) containing medium supplemented with 10% FBS, and 0.7% StemXVivo methylcellulose (R&D systems). Cells were fixed, and then nuclei were stained on day 14 using 4% PFA for 30 min and 10 µg/mL Hoechst 33342 for overnight, respectively. By using maximum projection, nuclei area > 8000 µm^2 from all fields in each well were counted as spheroids. The GFP intensities of the spheroids were measured from the spheroid area.

For 3D drug testing, 4,000 TFK-1 cells were seeded and allowed to form spheroids for 3 days, and then 5-FU was added for 5 days. 0.5% of DMSO in culture medium was used as vehicle control. 10 mM Calcein Blue AM (Abcam) was used to visualize live cells in spheroids. The summation of the Calcein Blue AM area called "Spheroid viability index" from every sample was calculated. The data were used to draw the dose-response curves. The relative CSC content was calculated from the means of the GFP intensity from viable spheroid normalized to the vehicle control.

Mouse xenograft

To access tumorigenicity in vivo, an animal study was performed under a protocol (COA no. 010/2564) approved by the Mahidol University-Institute Animal Care and Use Committee (MU-IACUC), according to Assessment and Accreditation of Laboratory Animal Care guidelines. All protocols complied with the NIH Guidelines and were reported in accordance with ARRIVE guidelines (https://arriveguidelines.org). Animals were housed in standard laboratory cages. Food and water were given ad libitum. Either 5×10^5 of SORE6^{POS} and SORE6^{NEG} cells (for TFK-1), or 1×10^6 cells (for KKK-D068) were mixed with 10% Matrigel® (Corning), before inoculated subcutaneously into the flank of the anesthetized Balb/c-RJ mice (BALB/c: Rag2/Jak3 dKO). Tumor growth was measured daily. The tumor volumes were calculated using the equation: Tumor Volume = ½ (Length × Width^2). After the tumor reached the experimental endpoint, both groups in the same cell line were euthanasia using lethal anesthesia (I.P. injection of 300 mg/kg ketamine: 30 mg/kg xylazine), following the AVMA Guidelines.

Immunohistochemistry

Immunohistochemistry (IHC) was performed according to standard protocol[67]. Stained mouse tissues were stained with SOX2 (clone: SP76) and OCT4 (clone: MRQ-10) antibodies.

Survival analysis

To determine the association between CCA CSC-related genes with clinical outcome, the overall survival analysis was performed on the bulk RNA-seq of primary CHOL tumors from the TCGA dataset[68] using Survival Genie (https://bhasinlab.bmi.emory.edu/SurvivalGenie/)[69]. All primary cholangiocarcinoma (CHOL) tumors in the TCGA dataset with available transcriptome profile and clinical information ($n = 36$) were included in

the survival analysis. To compare the overall survival (OS) in the Kaplan-Meier plot between patients with high and low gene expression we used the survfit function from the survival package in R software[70]. The patients with high and low gene expression in both individual and combination gene analyses were categorized by *Cutp* option from survMisc R package[71]. Therefore, individual genes *(SOX2*, *OCT4*, *NANOG*, *SOX9*, *CD44*, *PROM1* (encoding CD133), *EPCAM*, *LGR5*, *ANPEP* (encoding CD13), and *THY1* (encoding CD90)) and a combination of genes (*SOX2* + *OCT4*) were analyzed using the Survival Genie gene-based option[69]. Log-rank *p*-value lower than 0.05 were considered statistically significant.

Statistical analysis

The graphs and each statistical significance were determined and created with GraphPad Prism software and R (http://www.R-project.org/, version 4.2.2). Statistical significance was determined between groups using two-tailed t-test. If specific statistical determination was used, the test was in-text stated. A p-value < 0.05 was considered statistically significant for all tests.

Data availability

The data generated during this study are available from the corresponding author upon reasonable request.

Received: 4 April 2024; Accepted: 18 September 2024
Published online: 30 September 2024

References

1. Brown, Z. J., Ruff, S. M. & Pawlik, T. M. Developments in FGFR and IDH inhibitors for cholangiocarcinoma therapy. *Expert Rev. Anticancer Ther.***23**, 257–264. https://doi.org/10.1080/14737140.2023.2176846 (2023).
2. Zheng, Q., Zhang, B., Li, C. & Zhang, X. Overcome Drug Resistance in Cholangiocarcinoma: New Insight into mechanisms and Refining the preclinical experiment models. *Front. Oncol.***12**, 850732. https://doi.org/10.3389/fonc.2022.850732 (2022).
3. Batlle, E. & Clevers, H. Cancer stem cells revisited. *Nat. Med.***23**, 1124–1134. https://doi.org/10.1038/nm.4409 (2017).
4. Phan, T. G. & Croucher, P. I. The dormant cancer cell life cycle. *Nat. Rev. Cancer*. **20**, 398–411. https://doi.org/10.1038/s41568-020-0263-0 (2020).
5. Phi, L. T. H. et al. Cancer Stem Cells (CSCs) in Drug Resistance and their Therapeutic Implications in Cancer Treatment. *Stem Cells Int.* 5416923. (2018). https://doi.org/10.1155/2018/5416923 (2018).
6. Dean, M., Fojo, T. & Bates, S. Tumour stem cells and drug resistance. *Nat. Rev. Cancer*. **5**, 275–284. https://doi.org/10.1038/nrc1590 (2005).
7. Kamimoto, K. et al. Heterogeneity and stochastic growth regulation of biliary epithelial cells dictate dynamic epithelial tissue remodeling. *Elife*. **5**. https://doi.org/10.7554/eLife.15034 (2016).
8. Tamai, K., Fujimori, H., Mochizuki, M. & Satoh, K. Cancer Stem cells in Intrahepatic Cholangiocarcinoma; their molecular basis, and therapeutic implications. *Front. Physiol.***12**, 824261. https://doi.org/10.3389/fphys.2021.824261 (2021).
9. McGrath, N. A., Fu, J., Gu, S. Z. & Xie, C. Targeting cancer stem cells in cholangiocarcinoma (review). *Int. J. Oncol.***57**, 397–408. https://doi.org/10.3892/ijo.2020.5074 (2020).
10. Wu, H. J. & Chu, P. Y. Role of Cancer Stem cells in Cholangiocarcinoma and therapeutic implications. *Int. J. Mol. Sci.***20**. https://doi.org/10.3390/ijms20174154 (2019).
11. Cardinale, V. et al. Profiles of cancer stem cell subpopulations in cholangiocarcinomas. *Am. J. Pathol.***185**, 1724–1739. https://doi.org/10.1016/j.ajpath.2015.02.010 (2015).
12. Wang, M. et al. Isolation and characterization of tumorigenic extrahepatic cholangiocarcinoma cells with stem cell-like properties. *Int. J. Cancer*. **128**, 72–81. https://doi.org/10.1002/ijc.25317 (2011).
13. Kawamoto, M. et al. Combined Gemcitabine and Metronidazole is a Promising Therapeutic Strategy for Cancer stem-like Cholangiocarcinoma. *Anticancer Res.***38**, 2739–2748. https://doi.org/10.21873/anticanres.12516 (2018).
14. Raggi, C. et al. Dysregulation of Iron Metabolism in Cholangiocarcinoma stem-like cells. *Sci. Rep.***7**, 17667. https://doi.org/10.1038/s41598-017-17804-1 (2017).
15. Tang, D. G. Understanding cancer stem cell heterogeneity and plasticity. *Cell. Res.***22**, 457–472. https://doi.org/10.1038/cr.2012.13 (2012).
16. Milanovic, M. et al. Senescence-associated reprogramming promotes cancer stemness. *Nature*. **553**, 96–100. https://doi.org/10.1038/nature25167 (2018).
17. Lytle, N. K., Barber, A. G. & Reya, T. Stem cell fate in cancer growth, progression and therapy resistance. *Nat. Rev. Cancer*. **18**, 669–680. https://doi.org/10.1038/s41568-018-0056-x (2018).
18. Poleszczuk, J. & Enderling, H. Cancer Stem Cell Plasticity as Tumor Growth Promoter and Catalyst of Population Collapse. *Stem Cells Int.* 3923527. (2016). https://doi.org/10.1155/2016/3923527 (2016).
19. Richard, V., Kumar, T. R. S. & Pillai, R. M. Transitional dynamics of cancer stem cells in invasion and metastasis. *Transl Oncol.***14**, 100909. https://doi.org/10.1016/j.tranon.2020.100909 (2021).
20. Perez-Gonzalez, A., Bevant, K. & Blanpain, C. Cancer cell plasticity during tumor progression, metastasis and response to therapy. *Nat. Cancer*. **4**, 1063–1082. https://doi.org/10.1038/s43018-023-00595-y (2023).
21. Kuroda, T. et al. Octamer and Sox elements are required for transcriptional cis regulation of nanog gene expression. *Mol. Cell. Biol.***25**, 2475–2485. https://doi.org/10.1128/MCB.25.6.2475-2485.2005 (2005).
22. Rodda, D. J. et al. Transcriptional regulation of nanog by OCT4 and SOX2. *J. Biol. Chem.***280**, 24731–24737. https://doi.org/10.1074/jbc.M502573200 (2005).
23. Wang, Z. X. et al. Oct4 and Sox2 directly regulate expression of another pluripotency transcription factor, Zfp206, in embryonic stem cells. *J. Biol. Chem.***282**, 12822–12830. https://doi.org/10.1074/jbc.M611814200 (2007).
24. Swain, N., Thakur, M., Pathak, J. & Swain, B. SOX2, OCT4 and NANOG: the core embryonic stem cell pluripotency regulators in oral carcinogenesis. *J. Oral Maxillofac. Pathol.***24**, 368–373. https://doi.org/10.4103/jomfp.JOMFP_22_20 (2020).
25. Vaddi, P. K., Stamnes, M. A., Cao, H. & Chen, S. Elimination of SOX2/OCT4-Associated prostate Cancer stem cells blocks Tumor Development and enhances therapeutic response. *Cancers (Basel)*. **11**. https://doi.org/10.3390/cancers11091331 (2019).
26. You, L., Guo, X. & Huang, Y. Correlation of Cancer stem-cell markers OCT4, SOX2, and NANOG with Clinicopathological features and prognosis in operative patients with rectal Cancer. *Yonsei Med. J.***59**, 35–42. https://doi.org/10.3349/ymj.2018.59.1.35 (2018).
27. Li, J. et al. Identification and characterization of Cancer Stem-Like cells in ALK-Positive anaplastic large cell lymphoma using the SORE6 reporter. *Curr. Issues Mol. Biol.***43**, 543–557. https://doi.org/10.3390/cimb43020041 (2021).
28. Liang, X. et al. An enhancer-driven Stem Cell-Like Program mediated by SOX9 blocks intestinal differentiation in Colorectal Cancer. *Gastroenterology*. **162**, 209–222. https://doi.org/10.1053/j.gastro.2021.09.044 (2022).

29. Panda, M., Tripathi, S. K. & Biswal, B. K. SOX9: an emerging driving factor from cancer progression to drug resistance. *Biochim. Biophys. Acta Rev. Cancer*. **1875**, 188517. https://doi.org/10.1016/j.bbcan.2021.188517 (2021).
30. Shang, T. et al. Diverse functions of SOX9 in liver development and homeostasis and hepatobiliary diseases. *Genes Dis.***11**, 100996. https://doi.org/10.1016/j.gendis.2023.03.035 (2024).
31. Tang, B. et al. A flexible reporter system for direct observation and isolation of cancer stem cells. *Stem Cell. Rep.***4**, 155–169. https://doi.org/10.1016/j.stemcr.2014.11.002 (2015).
32. Padua, D. et al. A SOX2 reporter system identifies gastric Cancer Stem-Like cells sensitive to Monensin. *Cancers (Basel)*. 12. https://doi.org/10.3390/cancers12020495 (2020).
33. Brown, G. Targeting the retinoic acid pathway to Eradicate Cancer Stem cells. *Int. J. Mol. Sci.***24**. https://doi.org/10.3390/ijms24032373 (2023).
34. Yao, W. et al. All-trans retinoic acid reduces cancer stem cell-like cell-mediated resistance to gefitinib in NSCLC adenocarcinoma cells. *BMC Cancer*. **20**, 315. https://doi.org/10.1186/s12885-020-06818-0 (2020).
35. Young, M. J. et al. All-trans retinoic acid downregulates ALDH1-mediated stemness and inhibits tumour formation in ovarian cancer cells. *Carcinogenesis*. **36**, 498–507. https://doi.org/10.1093/carcin/bgv018 (2015).
36. Meacham, C. E. & Morrison, S. J. Tumour heterogeneity and cancer cell plasticity. *Nature*. **501**, 328–337. https://doi.org/10.1038/nature12624 (2013).
37. Morral, C. et al. Zonation of ribosomal DNA transcription defines a stem cell hierarchy in Colorectal Cancer. *Cell. Stem Cell*. **26**, 845–861. e812 (2020).
38. Cao, W. et al. LGR5 marks targetable tumor-initiating cells in mouse liver cancer. *Nat. Commun.* 11, (1961). https://doi.org/10.1038/s41467-020-15846-0 (2020).
39. Sitthithumcharee, G. et al. Dependency of Cholangiocarcinoma on cyclin D-Dependent kinase activity. *Hepatology*. **70**, 1614–1630. https://doi.org/10.1002/hep.30704 (2019).
40. Al Baghdadi, T. et al. Palbociclib in patients with pancreatic and biliary Cancer with CDKN2A alterations: results from the targeted Agent and profiling utilization Registry Study. *JCO Precis Oncol.***3**, 1–8. https://doi.org/10.1200/PO.19.00124 (2019).
41. Arora, M. et al. Synergistic combination of cytotoxic chemotherapy and cyclin-dependent kinase 4/6 inhibitors in biliary tract cancers. *Hepatology*. **75**, 43–58. https://doi.org/10.1002/hep.32102 (2022).
42. Suppramote, O. et al. The Acquired vulnerability caused by CDK4/6 inhibition promotes Drug Synergism between Oxaliplatin and Palbociclib in Cholangiocarcinoma. *Front. Oncol.***12**, 877194. https://doi.org/10.3389/fonc.2022.877194 (2022).
43. Song, X. et al. Combined CDK4/6 and Pan-mTOR inhibition is synergistic against Intrahepatic Cholangiocarcinoma. *Clin. Cancer Res.***25**, 403–413. https://doi.org/10.1158/1078-0432.CCR-18-0284 (2019).
44. Liu, L. et al. G1 cyclins link proliferation, pluripotency and differentiation of embryonic stem cells. *Nat. Cell. Biol.***19**, 177–188. https://doi.org/10.1038/ncb3474 (2017).
45. Zhang, S., Xiong, X. & Sun, Y. Functional characterization of SOX2 as an anticancer target. *Signal. Transduct. Target. Ther.***5**, 135. https://doi.org/10.1038/s41392-020-00242-3 (2020).
46. Zhang, Q., Han, Z., Zhu, Y., Chen, J. & Li, W. The role and specific mechanism of OCT4 in Cancer Stem cells: a review. *Int. J. Stem Cells*. **13**, 312–325. https://doi.org/10.15283/ijsc20097 (2020).
47. Corish, P. & Tyler-Smith, C. Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Eng.***12**, 1035–1040. https://doi.org/10.1093/protein/12.12.1035 (1999).
48. Wang, X., Golino, J. L., Hawk, N. V. & Xie, C. Reciprocal Interaction of Cancer Stem cells of Cholangiocarcinoma with Macrophage. *Stem Cell. Rev. Rep.***19**, 2013–2023. https://doi.org/10.1007/s12015-023-10557-7 (2023).
49. Bian, J. et al. Characterization of immunogenicity of malignant cells with stemness in Intrahepatic Cholangiocarcinoma by single-cell RNA sequencing. *Stem Cells Int.***2022** (3558200). https://doi.org/10.1155/2022/3558200 (2022).
50. Mohan, A., Raj Rajan, R., Mohan, G., Kollenchery Puthenveettil, P. & Maliekal, T. T. Markers and reporters to reveal the Hierarchy in Heterogeneous Cancer Stem cells. *Front. Cell. Dev. Biol.***9**, 668851. https://doi.org/10.3389/fcell.2021.668851 (2021).
51. Xiong, L. et al. Oct4 differentially regulates chromatin opening and enhancer transcription in pluripotent stem cells. *Elife*. **11**. https://doi.org/10.7554/eLife.71533 (2022).
52. Hagey, D. W., Bergsland, M. & Muhr, J. SOX2 transcription factor binding and function. *Development*. **149**. https://doi.org/10.1242/dev.200547 (2022).
53. Hutton, S. R. & Pevny, L. H. SOX2 expression levels distinguish between neural progenitor populations of the developing dorsal telencephalon. *Dev. Biol.***352**, 40–47. https://doi.org/10.1016/j.ydbio.2011.01.015 (2011).
54. Dar, F. S. et al. National guidelines for the diagnosis and treatment of hilar cholangiocarcinoma. *World J. Gastroenterol.***30**, 1018–1042. https://doi.org/10.3748/wjg.v30.i9.1018 (2024).
55. Nevi, L. et al. DCLK1, a putative stem cell marker in human cholangiocarcinoma. *Hepatology*. **73**, 144–159. https://doi.org/10.1002/hep.31571 (2021).
56. Purba, T. S. et al. CDK4/6 inhibition mitigates stem cell damage in a novel model for taxane-induced alopecia. *EMBO Mol. Med.***11**, e11031. https://doi.org/10.15252/emmm.201911031 (2019).
57. He, S. et al. Transient CDK4/6 inhibition protects hematopoietic stem cells from chemotherapy-induced exhaustion. *Sci. Transl Med.* 9. https://doi.org/10.1126/scitranslmed.aal3986 (2017).
58. Tsao, A. N., Chuang, Y. S., Lin, Y. C., Su, Y. & Chao, T. C. Dinaciclib inhibits the stemness of two subtypes of human breast cancer cells by targeting the FoxM1 and hedgehog signaling pathway. *Oncol. Rep.***47**. https://doi.org/10.3892/or.2022.8316 (2022).
59. Oh, S. J. et al. Targeting cyclin D-CDK4/6 sensitizes Immune-Refractory Cancer by blocking the SCP3-NANOG Axis. *Cancer Res.***78**, 2638–2653. https://doi.org/10.1158/0008-5472.CAN-17-2325 (2018).
60. Kishino, E. et al. Anti-cell growth and anti-cancer stem cell activity of the CDK4/6 inhibitor palbociclib in breast cancer cells. *Breast Cancer*. **27**, 415–425. https://doi.org/10.1007/s12282-019-01035-5 (2020).
61. Li, F. et al. YAP1-Mediated CDK6 activation confers Radiation Resistance in Esophageal Cancer - Rationale for the combination of YAP1 and CDK4/6 inhibitors in Esophageal Cancer. *Clin. Cancer Res.***25**, 2264–2277. https://doi.org/10.1158/1078-0432.CCR-18-1029 (2019).
62. Xie, X. et al. Abemaciclib drives the therapeutic differentiation of acute myeloid leukaemia stem cells. *Br. J. Haematol.***201**, 940–953. https://doi.org/10.1111/bjh.18735 (2023).
63. Prasopporn, S. et al. Combining the SMAC mimetic LCL161 with gemcitabine plus cisplatin therapy inhibits and prevents the emergence of multidrug resistance in cholangiocarcinoma. *Front. Oncol.***12**, 1021632. https://doi.org/10.3389/fonc.2022.1021632 (2022).
64. Jamnongsong, S. et al. Comprehensive drug response profiling and pan-omic analysis identified therapeutic candidates and prognostic biomarkers for Asian cholangiocarcinoma. *iScience*. **25**, 105182. https://doi.org/10.1016/j.isci.2022.105182 (2022).
65. Netsrithong, R. et al. Multilineage differentiation potential of hematoendothelial progenitors derived from human induced pluripotent stem cells. *Stem Cell. Res. Ther.***11**, 481. https://doi.org/10.1186/s13287-020-01997-w (2020).
66. Clark, N. A. et al. GRcalculator: an online tool for calculating and mining dose-response data. *BMC Cancer*. **17**, 698. https://doi.org/10.1186/s12885-017-3689-3 (2017).
67. Jirawatnotai, S. et al. Cdk4 is indispensable for postnatal proliferation of the anterior pituitary. *J. Biol. Chem.***279**, 51100–51106. https://doi.org/10.1074/jbc.M409080200 (2004).
68. Cancer Genome Atlas Research. The Cancer Genome Atlas Pan-cancer analysis project. *Nat. Genet.***45**, 1113–1120. https://doi.org/10.1038/ng.2764 (2013).

69. Dwivedi, B., Mumme, H., Satpathy, S., Bhasin, S. S. & Bhasin, M. Survival genie, a web platform for survival analysis across pediatric and adult cancers. *Sci. Rep.***12**, 3069. https://doi.org/10.1038/s41598-022-06841-0 (2022).
70. Package, A. for Survival Analysis in R (2021).
71. survMisc Miscellaneous Functions for Survival Data (2018).

Acknowledgements
K.K. was supported by Chulabhorn Foundation and the Royal Golden Jubilee Ph.D. Scholarship Program (RGJ-PHD) (Grant No. PHD/0060/2561) under Thailand Research Fund (TRF) and National Research Council of Thailand (NRCT). The project is funded by NRCT, N41A640162, and supported by Mahidol University (Basic Research Fund: fiscal year 2022). M.W. is supported by Chalermphrakiat Grant, Faculty of Medicine Siriraj Hospital, Mahidol University. S. Ji. is supported by the R.E.D program, Faculty of Medicine, Siriraj Hospital, Mahidol University. We also thank BioRender for the figure preparation. We would like to thank Siriwal Suwanpitak, Mahidol University, and Professor Nilay Sethi, Dana-Farber Cancer Institute for their technical support and discussion.

Author contributions
K.K. and S.Ji. conceptualized the manuscript. K.K., S.P., T.P., and S.Ja performed the experiments. S.Ja., N.T., M.H., M.W., and S.Ji. provided critical comments, experimental framework, and experimental design. K.K. performed data analysis and data visualization. K.K., M.H., and S.Ji. wrote the original draft. S.Ji. funding acquisition. S.Ji. and S.O. provided resources. All authors have read and approved the final manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-73581-8.

Correspondence and requests for materials should be addressed to S.J.

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

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2024