学位論文

Establishment and characterization of a novel cancer stem-like cell of cholangiocarcinoma

（胆管癌由来の新規癌幹細胞の樹立と性状解析）

パナワン オラサ
PANAWAN ORASA

Department of Tumor Genetics and Biology, Medical Science Major,
Doctoral Course of the Graduate School of Medical Sciences,
Kumamoto University

Academics advisor

Professor OIKE Yuichi
Associate Professor ARAKI Norie

Department of Tumor Genetics and Biology, Medical Sciences Major
Doctoral Course of the Graduate School of Medical Sciences,
Kumamoto University

September 2023

学　位　論　文

Title of Thesis : **Establishment and characterization of a novel cancer stem-like cell of cholangiocarcinoma**
（胆管癌由来の新規癌幹細胞の樹立と性状解析）

Name of Author : パナワン オラサ
PANAWAN Orasa

Name of supervisor : Professor OIKE Yuichi
Associate professor ARAKI Norie
Department of Tumor Genetics and Biology, Medical Sciences Major, Doctoral Course of the Graduate School of Medical Sciences

Name of examiner : Prof. NAKAO Mitsuyoshi (Department of Medical Cell Biology)
Prof. SASHIDA Goro (Department of Trancecriptional Regulation in Leukemogenesis)
Associate Prof. BABA Masaya (Department of Cancer Metabolism)
Research Associate Prof. BABA Yoshifumi (Department of Gastroenterological Surgery)

September 2023

Establishment and characterization of a novel cancer stem-like cell of cholangiocarcinoma

Authors:

Orasa Panawan[1,2] ,Atit Silsirivanit[1,2,3], Chih-Hsiang Chang[1] , Siyaporn Putthisen[2], Piyanard Boonnate[4], Taro Yokota[1], Yuki Nishisyama-Ikeda[1] , Marutpong Detarya[1,2,3], Kanlayanee Sawanyawisuth[2,3] , Worasak Kaewkong[5] , Kanha Muisuk[6], Sukanya Luang[2,3], Kulthida Vaeteewoottacharn[2,3,4], Ryusho Kariya[4], Hiromu Yano[7], Yoshihiro Komohara[7], Kunimasa Ohta[8], Seiji Okada[4], Sopit Wongkham[2,9], Norie Araki[1]

Affiliations:

[1]Department of Tumor Genetics and Biology, Graduate School of Medical Sciences, Faculty of Life Sciences, Kumamoto University, Kumamoto, Japan

[2]Department of Biochemistry, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

[3]Cholangiocarcinoma Research Institute, Khon Kaen University, Khon Kaen, Thailand

[4]Division of Hematopoiesis, Joint Research Center for Human Retrovirus Infection, Kumamoto University, Kumamoto, Japan

[5]Department of Biochemistry, Faculty of Medical Sciences, Naresuan University, Phitsanulok, Thailand

[6]Department of Forensic Medicine, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

[7]Department of Cell Pathology, Faculty of Life Sciences, Kumamoto University, Kumamoto, Japan

[8]Department of Stem Cell Biology, Faculty of Arts and Science, Kyushu University, Fukuoka, Japan

[9]Center for Translational Medicine, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

*Corresponding Authors:

Atit Silsirivanit (Email: atitsil@kku.ac.th)

Norie Araki (Email: nori@gpo.kumamoto-u.ac.jp)

Abstract

Recent evidence suggests the impact of cancer stem cells (CSC) on the therapeutic resistance of cancer cells. Compared with other cancer types, the knowledge about CSC in cholangiocarcinoma (CCA) is limited due to the lack of a CSC model. In this study, we successfully established a sphere-forming CCA stem-like cell, KKU-055-CSC, from a cancer cell line. The KKU-055-CSC noticeably exhibits the CSC characteristics, including 1) the ability to grow stably and withstand continuous passage for a long period of culture under the stem cell medium, 2) high expression of stem cell markers, 3) low responsiveness to standard chemo-drugs, 4) multi-lineage differentiation, and 5) faster and constant tumor formation in xenograft mouse models. To identify the CCA-CSC associated pathway, we have performed a global proteomics and functional cluster/network analysis. Global proteomics identified the 6000 proteins in total, and the significantly upregulated proteins in CSC compared with FCS-induce differentiating CSC and its parental cells were extracted. Network analysis revealed that high mobility group A1 (HMGA1) and Aurora-A signaling via the STAT-3

pathways were enriched in KKU-055-CSC. Knockdown of HMGA1 in KKU-055-CSC suppressed the expression of stem cell markers, increased cell proliferation, and enhanced drug sensitivity, especially Aurora-A inhibitors. *In silico* analysis demonstrated that the expression of HMGA1 was correlated with Aurora-A expressions and poor survival of CCA patients. In conclusion, we have established a unique CCA stem-like cell model, potentially used as a tool for studying CSC biology of CCA. Using this model, we identified the HMGA1 and Aurora-A signaling as the novel essential pathway for CSC-CCA.
Keywords: cancer stem cell, liver, bile duct, cholangiocarcinoma

Introduction

Cholangiocarcinoma (CCA) is the most common primary liver cancer arising from biliary epithelial cells [1]. Although it is considered a rare tumor, the number of cases and deaths are increasing each year worldwide; especially, the highest incidence and mortality rates of CCA have been reported in Northeastern Thailand [2]. The aggressive malignant phenotype with high metastasis and therapeutic resistance led to a poor prognosis for CCA patients. The survival rate of CCA patients is generally short and varies from 12-24 months [3]. Recently, surgical resection at the early stage has been reported as the only effective treatment for CCA patients, offering 15-40% of 5-year survival [3, 4]. Chemotherapy with gemcitabine, cisplatin, and 5-FU is necessary for patients with advanced and unresected tumors [5, 6]. However, resistance to chemotherapeutic drugs and tumor recurrence remains significant problem in CCA patients [7, 8]. Nowadays, no standard regimen for curative treatment of CCA is available [6].

Cancer stem cell (CSC), a subpopulation of cancer cells having self-renewal and multi-lineage differentiating capacity, was reported as an important factor in promoting tumor progression, recurrence, and therapeutic resistance in many types of cancer [9, 10]. Targeting CSCs is purposed to be a promising strategy for achieving curative treatment [9, 11]. Several approaches have been proposed to target CSCs, including activation of chemo/radio-sensitivity, specific pathway interference, differentiation/mobilization therapy, ablation through prospective factors, disruption of immune evasion pathways, and antiangiogenic treatments, *etc*. [12]. However, the molecular mechanism of CSC-specific pathways for stemness maintenance or differentiation is poorly understood. To target CSC for cancer treatment, it is necessary to isolate CSCs and analyze their specific properties. Currently, several *in vitro* and *in vivo* experimental protocols have been proposed to isolate and assess the cancer stem-like cells through such as tumor-sphere formation assay, a single cell colony formation assay, and cell sorting based on their specific cell surface markers by flow cytometry [13, 14]. The isolated cells were found to exhibit CSC phenotypes, such as drug resistance, multi-lineage differentiation, and tumor formation [15, 16]. The sources of this minor but useful cell subpopulation could be either primary tumor tissues or the established cancer cell lines [17-19]. Many cancer stem cell markers have been identified in CCA, such as, aldehyde dehydrogenase (ALDH), epithelial cell adhesion molecule (EpCAM), CD133, CD44, CD44v9CD90, CD13, and SRY-box transcription factor 2 (SOX2). [1, 20, 21]. The accumulated evidence suggests that the tumor tissue of CCA is rich in the CSC population and implicates CCAs as stem cell-based diseases [1, 22]. However, the information regarding the biology of CSC in CCA is very limited due to the lack of

cholangiocarcinoma-cancer stem cell (CCA-CSC) models to be characterized.

In this study, a novel stable/long-life CCA stem-like cell, KKU-055-CSC, has been isolated from the CCA cell line that previously established from the patient tissue. *In vitro* and *in vivo* experiments revealed that the cells exhibit the features of CSCs, suggesting its potential to be a representative CSC clone for CCA. Using this clone, we sought to identify the CCA-CSC-associated pathways and their key molecules as candidates to be targeted, by the global proteome analysis. Functional studies and *in silico* studies have been conducted to explore the roles of the selected CCA-CSC-associated molecules.

Materials and Methods

2.1 Cell line and cell culture

KKU-055, an adherent CCA cell line established from poorly differentiated primary CCA tissue of a 56-year-old Thai male, was obtained from the Japanese Collection of Research Bioresources Cell Bank, Osaka, Japan, through Cholangiocarcinoma Research Institute, Khon Kaen University, Thailand. The cell was maintained in Dulbecco's Modified Eagle's Medium (DMEM, Thermo Fisher Scientific, Waltham, MA) supplemented with 10% fetal calf serum (FCS, MP Biomedicals, Santa Ana, CA). At 80% confluence, the cells were sub-passaged using trypsin-EDTA approximately twice a week. KKU-055 cholangiocarcinoma stem-like cell (KKU-055-CSC) was maintained in a stem cell culture medium: serum-free DMEM supplemented with B-27 (1:50; GIBCO/Invitrogen), heparin (5 mg/ml; Sigma-Aldrich, St. Loius, MO), human fibroblast growth factor (FGF, 20 ng/ ml; PeproTech Inc, East Windsor, NJ), epidermal growth factor (EGF, 20 ng/ml; PeproTech Inc), and insulin (10 ng/ml, (Sigma-Aldrich, St. Loius, MO). FGF, EGE, and insulin were added to the cells every 2 days. The CSC were sub-cultured once a week using Accumax (Innovative Cell Technologies, San Diego, CA). All cell lines were cultured in a 37°C incubator with 5% CO_2. Since the adherent KKU-055 cell was used to establish KKU-055 CSC, it would be referred as "parental KKU-055".

2.2 Establishment of CCA stem-like cells from a CCA cell line

The CCA stem-like cells, KKU-055-CSCs, were derived/reprogramed from the adherent parental KKU-055 CCA cell line using previously reported protocols [17, 23]. Briefly, the parental KKU-055 cells were trypsinized and washed with phosphate buffer saline (PBS) to remove the FCS. The $1x10^6$ cells were resuspended in a stem cell culture medium and seeded to form spheres in an ultra-low attachment 100-mm culture dish (Corning, Kennebunk, ME). FGF, EGF, and insulin were added every 2 days. Cells were continuously cultured and sub-passaged once a week for more than 3 months with $1x10^6$ cells/dish, >15 passages, to assure that the cells formed spheres/spheroids stably during their culture. After 15 passages, the cells were stably cultured in the normal cell culture dish (TC-treated Cell Culture Dish, Falcon®).

2.3 Cell proliferation

The KKU-055-CSC and parental KKU-055 cells were seeded in a 96-well plate at 1,000 cells/100 µl/well. Cell number was measured at day 0 (1 h after the seeding), day1 (24 h), day2 (48 h), and day 3 (72 h), followed by incubating the cells with 10 µl Cell Counting kit-8 (WST-8, Dojindo,

Kumamoto, Japan) for 4 h, and analyzing their absorbances at 450 nm. The study with 5 replicates of the seeding wells for each condition, and three separated independent experiments were performed.

2.4 Drug sensitivity assay

The KKU-055-CSC or parental KKU-055 cells were seeded in a 24-well plate (2×10^4 cells/well). After 24 h incubation, the cells were treated with the desired concentrations of 5-fluorouracil (5-FU), cisplatin, and gemcitabine. PBS (for gemcitabine) or equal DMSO concentration (for cisplatin or 5-FU) were used as the controls. After 72 hrs of incubation, the cells were trypsinized and mixed with an equal volume of Trypan Blue solution (Gibco, Grand Island, NY). The viable cells were counted using a hemocytometer under the microscope with phase contrast mode. The relative cell viability was calculated as the percent (%) viable cells in the treatment conditions compared with the control. Each treatment was performed in 3-5 replicates, and those experiments were repeated at least twice.

2.5 Multi-lineage differentiation

KKU-055-CSC (1.0×10^6 cells) were pre-seeded and allowed to form the small spheres in stem cell cultured medium for 2 days in a 100-mm cell culture dish, then 500 µl each was transferred to each well of a 24-well plate. For FCS-induced differentiation, 55 µl of FCS was added into the spheres to obtain the final concentration of approximately 10% FCS. For the lineage-specific differentiation, the cell culture supernatant was replaced by StemPro™ Osteogenesis and Adipogenesis Differentiation media (Thermo Fisher Scientific, Grand Island, NY), incubated for 14 days independently, as previously described [24]. On day 14, Alizarin Red S staining was used to determine the osteocyte differentiation, represented by calcium granule accumulation. The adipocyte differentiation was identified by Oil Red O staining, which stains the lipid droplets, as previously described [24].

2.6 Western blot analysis

Cells were lysed using urea lysis buffer (8 M of urea, 2 M Thiourea, 0.5 mM NaCl) and phase-transfer surfactant (PTS) solution (12 mM SDC, 12 mM SLS, in 100 mM Tris-HCl (pH 9), supplemented with phosphatase inhibitor cocktail 2 and cocktail 3 (PhosStop; Roche Diagnostics, Mannheim, Germany) and protease inhibitor cocktail (Complete Mini; Roche Diagnostics, Mannheim, Germany). Protein concentration was determined using Quick Start™ Bradford Protein Assay (Bio-rad, Hercules, CA) for urea lysates and Pierce ™ BCA Protein assay kit (Thermofisher Scientific, Rockford, IL) for PTS lysates.

Cell lysates were mixed with sample solubilizing medium (0.5M Tris-HCl pH 6.8, 4% (w/v) SDS, 10% (v/v) glycerol and 0.1% (w/v) bromophenol blue) and incubated at 37°C for 1 hr. Equal amounts of protein were loaded into each well, and the electrophoresis was performed in SDS-PAGE running buffer (0.025M Tris base, 0.192M Glycine, 0.1% SDS) with a constant 20 mA/gel for 1 hour. After electrophoresis, the proteins were transferred to PVDF membrane using transferring buffer (40 mM Tris base, 20 mM sodium acetate, 2 mM EDTA, pH 7.4, 20% (v/v) methanol, and 0.05% (w/v) SDS with constant 20 voltage for 1 hour. Non-specific reactivity was blocked by incubating the membrane with 5% skim milk or 5% bovine serum albumin (BSA) for 1 h at room temperature. After that, the membranes were probed overnight with the desired concentrations of primary antibodies

(supplementary data). After complete incubation, the membranes were incubated with 1:10,000 HRP-conjugated anti-mouse or anti-rabbit antibodies. The chemiluminescent signal was developed by ECLTM Prime Western Blotting Detection (GE Healthcare) and detected by ImageQuant LAS 600 image analyzer. The signal intensities were analyzed using ImageJ software. Beta-actin was used as an internal control.

2.7 Immuno-cytochemistry staining

KKU-055-CSC and KKU-055 parental cells were seeded in a 24-well plate (2 x 10^4 cells/well) and cultured for 72 h. For the FCS-induced differentiated form of KKU-055-CSC (KKU-055-CSC-DIFF), after cell dissociation, the cells were resuspended in 10% FCS containing DMEM. After 72 h incubation, the culture media was gently removed and the cells were washed twice with PBS, followed by fixation in 90% cold methanol for 15 min. After fixation, the cells were washed twice with PBS, followed by blockage of the non-specific reactivity with 1% BSA for one h. Then, the cells were incubated overnight with the desired concentration of primary antibodies at 4°C. After the overnight incubation, the cells were washed with PBS and incubated with 1:1,000 Alexa-547-conjugated anti-mouse or Alexa 488-conjugated anti-rabbit antibody for 1 h at room temperature. The DAPI diluted 1:1,000 (Invitrogen corporation, Eugene, OR) was used for nuclear staining.

2.8 RNA extraction and real-time reverse transcriptase polymerase chain reaction

Total RNA was extracted using the RNeasy Mini kit (Vivantis Technology, Selangor, Malaysia). RNA concentration was measured using OD at 260 nm, and quality was assessed by agarose gel (1%) electrophoresis. Two micrograms of total RNA were used for cDNA synthesis by a High-capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster, CA) according to the manufacturer's recommendation. The converted cDNA was diluted with nuclease-free water to obtain a final concentration of 20 ng/µl and used as the template for real-time polymerase chain reaction (PCR). The Light Cycler 480® SYBR green I master (Roche Diagnostics, Mannheim, Germany) was used for the PCR in the Light Cycler 480® system (Roche Diagnostics) with the condition of 95°C for 5 min, followed by 40 cycles of 95°C for 10 sec, 60°C for 10 sec and 72°C for 3 sec and the final step of 40°C for 10 sec. Expression of the target genes was quantified ($2^{-\Delta\Delta Cp}$) relative to *beta-actin* as the reference gene. All primers are listed in the **supplement data.**

2.9 *In vivo* tumor transplantation

The *in vivo* experiments were performed in the BALB/c Rag-2null/Jak3null (BRJ) mice [25] under the guidelines of the Animal Experiment Committee of Kumamoto University. The KKU-055-CSC or parental KKU-055 cells were subcutaneously injected into 8 to 12-week-old BRJ mice by varying cell numbers as Group 1) $3x10^4$ cells/nodule, 2) $1x10^5$ cells/nodule, 3) $3x10^5$ cells/nodule, and 4) $1x10^6$ cells/nodule. Each group was performed in 5 replicates. The animal conditions were monitored daily, and the tumor size was measured every 3 days using a caliper. Tumor volume (V) was calculated from length (L, the long part of the tumor and width (W, the short part of the tumor) using the following formular.

$$V = (W^2 \times L)/2$$

The tumors were harvested and weighted when the diameter reached 2 cm at most in the group of the xenografts

2.10 Immunohistochemistry of tumors from mouse xenografts

Tumor samples were fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS), embedded in paraffin, and cut into 4 μm sections; HE staining was performed for checking the cell morphology in the tumor. For immunostaining, the sections were deparaffinized using xylene and rehydrated with a graded series of ethanol. Antigen retrieval was performed by heat-induced epitope retrieval in citrate buffer. The sections were incubated with 1% bovine serum albumin in Tris-buffered saline with 0.1% NaN3 for blocking, and stained with primary antibodies against EpCAM (Abcam, ab71916), and Ki67 (Abcam, ab16667). The sections were incubated with horseradish peroxidase (HRP)-conjugated anti-mouse secondary antibodies (Nichirei). Positive signals were visualized using HistoGreen substrate (Cosmo Bio) for staining with diaminobenzidine (DAB; Nichirei).

2.11 Small interference RNA (siRNA) treatment

The siRNA against HMGA1 (siHMGA1) was transfected into KKU-055-CSC using an electroporation system. KKU-055-CSC ($2x10^5$ cells) were resuspended in 10 μl of serum-free Opti-MEM (Thermo Fisher Scientific, Waltham, MA) containing 50 pmole of siHMGA1. The electroporation was performed by pulsing 1,100 volts, 30 pulse-width, 2 times in MicroPorator MP-100 (Digital bio, Thermo Fisher, Algonquin, IL) using 10 μl Neon™ tips (Invitrogen, Carlsbad, CA). After the electroporation, the cells were transferred into 24-well or 6-well plates and incubated for 48-72 h in the 5% CO2 incubator at 37oC. The synthesized sequences of siHMGA1 used in this experiment were 5'-GUGCCAACACCUAAGAGACCUTT-3' (sense) and 5'-AGGUCUCUUAGGUGUUGGCACTT-3'(antisense) [26]. Cells were treated with Silencer Negative Control siRNA (Ambion, Carlsbad, Canada) as a control siRNA**.**

2.12 Cytogenetic and karyotype analysis

Parental KKU-055 and KKU-055-CSC cells were subjected to a conventional chromosomal analysis as in the protocol previously described [6]. Briefly, after the fixation, the chromosomes in the cells were trypsin-Giemsa-banded to identify individual metaphase chromosomes. Representative non-overlapped chromosome figures were imaged. The karyotype analysis was based on criteria of the International System for Human Cytogenetic Nomenclature (ISCN) [27]. The modal chromosome number was determined in at least 20 cells from each parental KKU-055 and KKU-055-CSC.

2.13 Short tandem repeat (STR) analysis

The DNAs from KKU-055-CSC and parental KKU-055 cells were extracted using the genomic extraction kit (QIAGEN, Hilden, Germany). A standard STR profile was obtained with 15 loci and the gender marker (amelogenin) which were amplified by using the AmplFLSTR PCR Amplification kit and analyzed by Applied Biosystems 3130xl Genetic Analyzer (Applied Biosystem, Foster City, CA). Obtained data were analyzed using the GeneMapper ID-X Software Version 1.4 (Applied Biosystems).

2.14 Cellular protein preparation for proteomics

The cellular proteins were extracted from parental KKU-055, KKU-055-CSC, and its differentiated form, and prepared according to the protocol previously described [28]. Cells were cultured up to 80% confluences in 10-cm diameter dishes for 72 h. The cell lysate was harvested

using PTS lysis buffer (12 mM SDC, 12 mM SLS, in 100 mM Tris buffer (pH 9) containing protease inhibitor cocktails (Sigma-Aldrich, St. Loius, MO), and sonicated in ice bath using Bioruptor UCD-250 (Bio sonic, Kanazawa, Japan) with 5 s ON/ 1 s OFF pulse for 5 min x 5 times. The 400 μg of cell lysates was reduced with 10 mM dithiothreitol at 37°C for 30 min, followed by carbamidomethylated with 55 mM iodoacetamide at 37°C for 30 min, under the dark condition. The alkylated protein solution was 5-fold diluted with 50 mM ammonium bicarbonate, digested with 4 ug (Lys-C:substrate proteins, 1:100) of Lys-C (Wako Chemicals) for 3 hr at 37°C, and then digested by 4 ug (trypsin:substrate proteins, 1:100) of trypsin (MS grade, Promega,) overnight at 37°C. After the digestion, an equal volume of ethyl acetate was added to remove the detergents by phase transfer after the acidification with 0.5% trifluoroacetic acid (final concentration). The mixture was vortexed and centrifuged at 15,700 x g for 2 min; the aqueous layer was collected, dried, rehydrated with 0.1% TFA, and desalted using StageTips with SDB-XC disk membranes. The peptides were kept in SDB elution solution (80% ACN and 0.1% TFA) at −80 °C until the MS analysis. Four replicate samples were prepared for each group.

2.15 nanoLC-MS/MS analysis

The LC/MS/MS analyses were performed in an Orbitrap Fusion Tribrid mass spectrometer (Thermofisher Scientific), connected to an EASY-nLC™ 1200 (Thermofisher Scientific). Peptides were first loaded onto an EASY-SprayTM C18 Trap column (Thermofisher Scientific, P/N 164946, 3 μm, 75 μm ID× 20 mm length), and then separated on C18 packed emitter column, 3 μm, 100 Å, 75 μm ID × 120 mm length from Nikkyo Technos Co., Ltd (Tokyo, Japan). The injection volume was 10 μL, and the flow rate was 300 nL/min. The mobile phases were (A) 0.1% formic acid and (B) 0.1% formic acid/80% ACN. Multiple-linear gradient elution was performed as follows: 2-5%B in 1 min, 5－40% B in 239 min, 40-100% B in 5 min, and 100% B for 5 min. Spray voltages of 2300 V, the mass scan range; m/z 350－1650, a maximum injection time;50 ms, and detection at a mass resolution;120,000 at m/z 200 in the orbitrap analyzer. A maximum cycle time;3 seconds for full MS scan and MS2 scans. The precursor ions with a 1.6 m/z isolation window, a maximum injection time;75 ms, and a mass resolution;30,000. Dynamic exclusion;60 sec,the normalized HCD;30, a mass resolution;15,000 at m/z 200 in the Orbitrap analyzer. A lock mass 391.28429 was used.

2.16 Proteomics data analysis Mass spectrometry data were processed and searched against the UniProtKB/SwissProt (2021_09) human database using Protein Discoverer 2.4. The following parameters were applied for peptide and protein identification: a precursor mass tolerance of 4.5 ppm, a fragment ion mass tolerance of 20 ppm, and a minimum peptide length of 7 amino acids. Fixed modification; Cys Carbamidomethylation and variable modifications;methionine oxidation and protein N-terminal acetylation. FDR <1% at the PSM and protein level.

Quantitative proteomics data were analyzed by Perseus (v1.6.15.0) using a Label Free Quantitation (LFQ) intensity (log2) data [29]. The protein identifications from each group (KKU-055-CSC, -DIF, -Parental) of 4 replicates were filtered to require 4 valid values in at least 1 group. Then, for the relative quantitation, the imputation of the missing values in the samples data was subjected with random numbers drawn from a normal distribution with a downshifting of 1.8 and a standard deviation of 0.3 using Perseus 1.6.14.0 [30, 31].

For cluster analysis, based on the expression patterns of each cellular group, fuzzy c-means clustering was applied to categorize the proteins that are most down-regulated or up-regulated. The definition of the clusters was created by the k-means algorithm. For differential analysis between the two groups, the t-test was applied, and the results were shown in the form of a volcano plot, and the significant data points in the volcano plot were extracted after a permutation-based FDR calculation (FDR=0.05).

Network analysis was performed using KeyMolnet (2022) (KM-Data, Japan). The lists of differentially expressed protein were imported into KeyMolnet, and the "start points and end points" network search algorithm were used to generate the molecular interaction networks. The specific activated networks in KKU-055-CSCs including the direct/indirect activation, transcriptional regulation, and complex formation were extracted and illustrated in the KeyMolnet viewer.

2.17 Statical analysis of cellular phenotypes

The statistical analyses and graph construction of the cellular proliferation, survival, tumor sizes of the xenograft, RT-PCR, and Western blotting were performed in GraphPad Prism 9.0 software (GraphPad Software, Inc., La Jolla, CA). Data were presented as the mean ± standard deviation (SD). The student's t-test was used to compare the data between groups, *p*-value *< 0.05, **< 0.01 was considered statistically significant.

3. Results

3.1 KKU-055 CCA stem-like cells (KKU-055-CSC), stably established from the KKU-055 CCA cell line, expresses stem cell markers.

To establish KKU-055-CSC, the parental KKU-055 CCA cells ($1x10^6$ cells) were seeded onto an ultra-low attachment culture dish and continuously cultured in a stem cell culture medium. The cells were able to form the floating three-dimensional clusters, which were continuously grew and long-term serial sub-passaged. After sub-passage, reformation of the small spheres (< 20 μm) could be observed within 48 h and continuously enlarge to reach approximately 100 μm within 7 days (**Figure 1A**). After 3 months, the cells could propagate as a floating-sphere without attaching in the normal TC-treated cell culture dishes. Over the past 3 years, the cells were able to be serial sub-passaged and re-formed the spheres constantly for more than 100 passages. In our study, the cells were continuously cultured up to 20 passages then kept in the -80 ˚C in the stem cell banker, all experiments were performed within 20 passages of the refreshed stock cells. As the cells can survive to undergo self-renewal and maintain their ability to form similar shape of spheres constantly over long-term culture under the stem cell culture condition, we therefore firstly named these cells as KKU-055-sphere cells (KKU-055-SC).

To validate whether KKU-055-SCs exhibited stem cell characteristics in general, we examined the expression of stem cell markers in KKU-055-SCs compared with parental KKU-055 cells. By Western blotting, the reported CCA stem markers such as SOX2, CD44, CD147, and EpCAM [21] were found to be highly expressed in KKU-055-CSC but less expressed or undetectable in parental KKU-055 cells **(Figure 1B)**. High expression of SOX2, CD44, EpCAM, and CD44v9 in KKU-055-SCs was also demonstrated by the immunofluorescence staining **(Figure 1C)**. In addition,

the mRNA expression of *SOX2, c-myc,* OLG2, *Oct3/4, CD44,* and *EpCAM* was assessed by real-time RT-PCR. The results showed that these stem cell-related transcription factors were highly expressed in KKU-055-SCs compared with parental KKU-055 cells **(Figure 1D)**. These results demonstrate that KKU-055-SCs have stem cell-like properties with the expression of stem cell markers.

3.2 KKU-055-SC exhibited stem cell phenotypes.

3.2.1 Cellular proliferation

CSC has been defined as a quiescent cell with a low proliferative rate; however, it possesses a high tumorigenic ability. The CCK-8 cell proliferation assay shows that KKU-055-SCs exhibit a significantly lower proliferation rate than parental KKU-055 cells (**Figure 2A**).

3.2.2 Sensitivity against anti-cancer drugs

Another important CSC property is the gain of chemoresistance; therefore, we have analyzed the sensitivities of KKU-055-SCs against standard chemotherapeutic drugs for CCA treatment. Compared with parental KKU-055 cells, the KKU-055-SC exhibited higher resistance to 5-FU, cisplatin, and gemcitabine (**Figure 2B)**. The half-maximal inhibitory concentration (IC50) at 72 h of 5-FU in KKU-055-SC (31.6 μM) was significantly higher than parental cells (24.5 μM). For cisplatin, the IC50 in KKU-055-SC was 5.6 μM, which is significantly higher than that of parental KKU-055 cells (3.4 μM). The IC50 of gemcitabine in KKU-055-SC was 29.8 nM, while in parental KKU-055 cells was 13.8 nM.

3.2.3 Multi-lineage differentiation

To demonstrate the ability in multi-lineage differentiation of KKU-055-SC, the cells were treated with FCS, adipogenesis differentiation medium, and osteogenesis differentiation medium (**Figure 2C**). After the treatment of 10% FCS for 72 h, KKU-055-SC became adherent cells that resemble parental cells. Moreover, the KKU-055-SC could differentiate into adipocyte and osteocyte by adipogenesis and osteogenesis differentiation media after 14 days, respectively. Lipid droplets accumulated in the cytoplasm of the induced adipocytes was shown by the positive staining with Oil Red O. The osteocyte differentiation was confirmed with calcium staining by Alizarin Red.

3.2.4 Tumor initiation ability in the mouse transplantation model

Xenotransplantation in a mouse model was used to evaluate the *in vivo* tumorigenicity of KKU-055-SCs and parental KKU-055. The cells were subcutaneously injected into BRJ mice by varying cell numbers. The injection and observation schedule were shown in **Figure 3A**. The ability in tumor formation of KKU-055-SC was significantly higher than parental KKU-055 derived tumor (**Figure 3B-D**). At week 4, the obvious tumor formation was observed in mice injected with $3x10^5$ and $1x10^6$ cells/nodule; therefore, the mice were euthanatized, and the tumors were excised at day 31. With $1x10^6$ cells/nodule, the weight of tumors formed from KKU-055-SC was 0.84 ± 0.24 g, while that from parental KKU-055 cells was 0.14 ± 0.05 g. For the transplantation with $3x10^5$ cells/nodule, the weight of tumors formed from KKU-055-CS was 0.39 ± 0.26 g, while that by parental KKU-055 was 0.11 ± 0.06 g. At week 5, the tumor nodules transplanted with $1x10^5$ cells of KKU-055-SCs were observed and allowed to grow until they reached a 2 cm diameter at day 46 (approximately 6.5 weeks). Of those transplanted with $1x10^5$ cells/nodule, the tumor weight of KKU-055-SCs was 0.57 ± 0.32 g,

and that of parental KKU-055 was 0.13 ± 0.08 g. The immunohistochemistry showed that all tumors arisen from KKU-055-SCs showed the stronger positive staining signal of a proliferative marker–KI-67 and CCA-stem cell marker–EpCAM compared with those of parental-KKU-055 cells **(Figure 3F)**. These results indicated that KKU-055-SC possess higher *in vivo* tumorigenicity compared with the parental KKU-055 cancer cell line.

Collectively, our cell line derived CCA sphere cells, exhibited the cancer stem cell phenotypes, including stem cell marker expression, self-renewal, drug resistance, multi-lineage differentiation, and higher tumor formation ability. Therefore, we defined the KKU-055-SC as a CCA-stem like cells (CSC) and renamed KKU-055-CSC.

3.3 KKU-055-CSC) showed the similar chromosome phenotypes to a parental KKU-055 cell line.

Karyotype analysis revealed that KKU-055-CSCs show similar chromosome patterns of parental KKU-055 cells, exhibiting similar aneuploidy karyotypes with marked structural abnormalities of the chromosomes. The representative G-banded karyotypes are demonstrated in **Figure 4**. The numbers of chromosomes among 20 cells of each were illustrated an equal modal chromosome number of 54 (53-56 chromosomes of parental KKU-055 cells and 52-55 chromosomes of KKU-055-CSCs). In addition, several chromosome markers (M, mar) were noted similarly in both, such as for parental KKU-055: 53+XX; +1, +3, +6, +7, +8, +10, +12, -13, -17, +21, +mar and for KKU-055-CSC: 53+XX; +1, +3, +4, +7, +8, +10, +12, -13, -17, +21, +mar. To verify the origin of KKU-CSC, STR analysis was performed comparing with the information of parental KKU-055 and its original patient tissue available in the JCRB database (https://cellbank.nibiohn.go.jp/~cellbank/en/search_res_det.cgi?ID=7070). As shown in **Table 1,** the STR profile of the KKU-055-CSCs were > 95% similar to those of the parental KKU-055 cells and original patient tissue, supporting the same origin of the KKU-055-CSCs and parental KKU-055 cells. In addition, after long-term passages over 3 years, the KKU-055-CSCs have not changed their chromosomal phenotypes.

3.4 Proteome profiling differentially expressed in KKU-055-CSC, compared with KKU-055-DIF, and KKU-055-parental cells.

To explore the molecular characteristics of CSC, we performed the global proteome analysis to compare KKU-055-CSC with its FCS-induced differentiated form (KKU-055-DIF) and the parental KKU-055. Four biological replicated samples from each of 1) KKU-055-CSC, 2) KKU-055-DIF, and 3) parental KKU-055, were prepared and subjected for the proteomics analysis as a workflow in **Figure 4A**. The results of global LFQ proteomics showed that 42,706 unique peptides corresponding to 5,925 proteins (FDR<1%) were identified in total. The identified proteins in each sample revealed a high correlation between the biological quadruplicates in the Pearson correlation analysis (**supplementary data).** To obtain an overview of the global proteome changes arising between three different stages of the cells, a hierarchical clustering analysis was performed, as shown in **Figure 4B**. Cluster analysis showed that CSCs clustered closer to DIF than to parental among the

three different stages of the cell. The *k*-mean clustering revealed 5 discrete protein expression patterns (**Figure 4 B-C**). Clusters 1 and 2 represented the 936 and 456 proteins that low expressed in CSC and increased upon differentiation and maintained high expression in parental cells. Cluster 3 is a group of proteins (n = 508) that are highly expressed in the differentiation stage and low expressed in CSC and parental cancer cells. Cluster 4 contained 972 proteins that is highly expressed in CSC and decreased in differentiation stage and parental cells. Cluster 5 is defined by the group of proteins (n = 655) that highly expressed in CSC and parental cells and low in differentiated form. In addition, we observed the proteome expression levels of stem cell markers in all of the cells, and found that the specific stem cell markers such as CD44, CD38, EpCAM, Kif4, and multi-drug resistance proteins such as ALDH3A1, ALDH1A1, and ABCB1 are overexpressed in KKU-055-CSC and gradually downregulated during the differentiation and the parental cells, suggesting that KKU-055-DIF may be intermediated stage of transitioning from stem cell stage to be parental cancer cells (**supplementary data**).

3.5 Network analyses of proteins significantly expressed in KKU-055-CSC

Next, to characterize the overall functional view of differential protein expression in each group, we compared the differential protein data in the two groups between KKU-055-CSC versus parental KKU-055, and KKU-055-CSC versus KKU-055-DIF as shown in the volcano plot (**Figure 5A**). Compared with KKU-055-DIF, we identified 1,988 up-regulated and 555 down-regulated proteins in KKU-055-CSC. Compared with parental KKU-055, there are 1,816 proteins up-regulated and 1,100 proteins down-regulated in KKU-055-CSC, significantly. Using these protein data, we tried to focus on proteins up/down-regulated in KKU-055-CSC with overlapping between DIF and parental cells (**Figure 5B**). We found 1,123 proteins and 319 proteins in up- and down-regulated in common respectively in KKU-055-CSC.

These KKU-055-CSC specific proteins obtained above were further analyzed with pathway-based characterization using KeyMolnet software. As a result, the top 20 and 10 of the over-representative pathways which are highly enriched in up- and down-regulated, respectively, in the KKU-055-CSC are significantly identified and listed in **Table 2.** The "Condensin signaling pathway" (*p-value* =2.41 x 10^{-13}), "Aurora signaling pathway" (*p-value* =1.08 x 10^{-8}), "BET family signaling pathway" (*p-value* =3.69 x 10^{-8}), and "transcription regulation by High Mobility Group Protein (HMGP)" (*p-value* =8.68 x 10^{-7}) *etc.* were found to be highly enriched in KKU-055-CSC. In contrast, we found "Lipoprotein metabolism" (*p-value* =7.53 x 10^{-10}), "Integrin family" (*p-value* =2.47 x 10^{-8}), "Sphingolipid signaling pathway" (*p-value* =5.37 x 10^{-7}), "Integrin signaling pathway" (*p-value* =8.25 x 10^{-6}), *etc.* were highly down-regulated in KKU-055-CSC. Additionally, we noted that "Spliceosome assembly pathway" was enrich as high score in both up- and down- regulated in CSC.

To extract the abnormal signaling pathway controlling stemness of KKU-055-CSC, we conduct a network analysis focused on up-regulated protein in CSC (1,123 molecules) by KeyMolnet. Using the network search algorithm "start points and end points", up-regulated proteins was set as "start point" and cancer stem cell pathway was set as "End point". The results revealed a complex network of targets with the most statistically significant relationships **(Figure 6A).** Among them, we

extracted a particularly specific network focusing on Aurora A; which existed in the highly enriched pathway in CSC, HMGA1; which showed the highest differential expression, and STAT3, which is a significant signaling pathway controlling many molecules in this network **(Figure 6B)**. The extracted network indicated that HMGA1 modulates the expression of CD44 which might function as a transcriptional factor. In addition, HMGA1 is also associated with stem cell factors including CD133, BRG, Oct4, NANOG, SWI/SNF, c-myc, and CD44 via activation of STAT3, which has been reported as a critical downstream target of HMGA1 [32]. Moreover, we also found that HMAG1 has interaction with Aurora-A via STAT3 which was also identified as a specific protein significantly up-regulated in CSC by the proteomics (**supplementary data**). These network analyses strongly suggest that HMGA1-STAT3-Aurora-A is a significant network associated with stemness maintenance in cholangiocarcinoma cancer stem cells via controlling stem cell factors.

3.6 HMGA1 signaling pathway regulates stemness maintenance and chemoresistance of CCA-CSC.

Our proteomics data suggested the association of HMGA1 with stemness maintenance of CCA-CSC. Western blot analysis and immunofluorescence were used to validate the expression of HMGA1 in KKU-055-CSC, its differentiated form, and the parental KKU-055. The result showed that HMGA1 is highly expressed in KKU-055-CSC and drastically reduced in KKU-055-DIF and parental KKU-055 cells **(Figure 7A).** To investigate the role of HMGA1 in the regulation of CCA-CSC, we silenced HMGA1 expression in KKU-055-CSC using specific siRNA. After 72 h of siHMGA1 treatment, the expression of HMGA1 in KKU-055-CSC was significantly decreased compared with that of siControl **(Figure 7B).** After the knockdown of HMGA1, the morphology of KKU-055-CSC was slightly changed by increasing the adherent activity to the surface of the culture dish (**supplementary data**). Cell proliferation ability of KKU-055-CSC was dramatically increased after the suppression of HMGA1 (**Figure 7C**). Moreover, we found that the knockdown of HMGA1 leads to the reduction of stem cell markers such as SOX2 and CD44 in immunocytochemistry **(Figure 7B)**. These results suggested that the suppression of HMGA1 may facilitate the differentiation of KKU-055-CSC and reflect its proliferation ability. Furthermore, it is known that the differentiation of CSCs affects their drug resistance [17, 23, 33, 34]; therefore, we have determined the chemosensitivity of KKU-055-CSC after the knockdown of HMGA1. As we expected, the results showed that KKU-055-CSC treated with siHMGA1 becomes more sensitive significantly against 5-FU, cisplatin, and gemcitabine than siControl treated cells (**Figure 7D)**. These results suggested the role of HMGA1 in stemness maintenance and chemoresistance of CCA-CSC.

3.7 Aurora signaling pathway involves cell viability of CCA-CSC.

In addition to HMGA1, our proteomics data showed that the Aurora signaling pathway is highly enriched in KKU-055-CSC (**Table 2**). Western blot analysis and immunofluorescence confirmed that Aurora A is highly expressed in KKU-055-CSC compared to the KKU-055-DIF and parental KKU-055 (**Figure 7E)**. To investigate the role of the Aurora signaling pathway, US-FDA-approved five drugs targeting Aurora-A activity were used to treat KKU-055-CSC compared with

parental cells for 72 h. Aurora-A inhibitors showed higher inhibitory effects on cell viability of KKU-055-CSC than the parental cell (**Figure 7F-G)**. Moreover, we have analyzed the combined effects of the knockdown of HMGA1 and Aurora-A inhibitors in KKU-055-CSC and parental cells. After the pre-treatment of siHMGA1 or siControl for 24 h, CSC and parental cells were treated with the desired concentration of Aurora-A inhibitors for 48 h. The results showed that knockdown of HMGA1 could sensitize the effect of Aurora-A inhibitors in KKU-055-CSC, while this sensitized effect was not observed in parental cells (**Figure 7F-G)**. These findings strongly suggest the potential of combined suppression of HMGA1 and Aurora-A for improvement of CCA treatment via targeting CCA-CSC.

3.7 *In silico* statistical analysis of *HMGA1*and *AurolaA* expressions in CCA tissues

The mRNA expression levels of *HMGA1* and *AURKA* in human CCA were analyzed using The Cancer Genome Atlas (TCGA) dataset with GEPIA analysis (http://gepia.cancer-pku.cn/). The expression of in *HMGA1* and *AURKA* in human CCA tissues (n=36) were compared with normal counterparts (n=9). A significantly higher level of *HMGA1* more than normal tissue was identified in CCA patients (**supplementary data**). Kaplan-Meier survival analysis clearly showed that the higher expression of *HMGA1* in human CCA tissues is correlated to the poor overall survival of patients **(supplementary data).** Also, *AURKA* levels are found to be positively correlated to *HMGA1* mRNA levels in TCGA human CCA with the highest Pearson index (R^2) of 0.73 (p-value 1.7 x 10^{-8}) (**supplementary data**). These data suggest that both expressions of HMGA1 and AURKA, found to be the most significant factors of CCA-CSC, are correlated to the CCA malignancy in CCA patients.

Discussion

Cancer stem cells (CSC) are a minor population of tumor cells discovered to be one of the main factors contributing to tumor development and progression, which is proposed to be an effective target for cancer treatment. Here, we successfully establish a cell line-derived CCA stem-like cell, which could represent the CSC clone that usefulness for further *in vitro* and in *vivo* studies. The established CSC clone–KKU-055-CCA stem-like cells (KKU-055-CSC) were derived from an adherent KKU-055 CCA cell line. The KKU-055-CSC exhibited CSC phenotypes, including expression of stem cell markers, drug resistance, multi-lineage differentiation, and tumor formation. In addition, the proteome analysis followed by functional assays revealed the importance of HMGA1 and Aurora-A for stemness maintenance, chemoresistance, and survival of CCA-CSC.

In CCA, little is known about CSC biology; this might be due to the lack of the CSC model of CCA. In fact, the cancer stem-like cells could be obtained from several sources by many methods [15, 35-38]. In this study, we uniquely isolated the cancer stem-like cells from the established cancer cell line using the protocol modified from our previous studies in glioma [17, 23]. This protocol can provide a practical and uncomplicated method for culturing the primary cancer cells or cell lines in specific stem cell media to produce the stem-like cells which can maintain their characters for long-term under the specific culture condition. Using this method, for the first time we successfully established a long-term culturable CCA stem-like cell (KKU-055-CSC), as a valuable tool for studying the biology of CCA stem cells.

To authenticate that the isolated cells could be a representative CSC model, several CSC phenotypes should be shown through many *in vitro* and *in vivo* experiments [33, 39-41]. In our study, the KKU-055-CSC could exhibit CSC phenotypes, such as expression of CSC markers, multi-lineage differentiation, chemoresistance and *in vivo* tumor formation.

CSC factors/markers regulate stem cell phenotypes [21] and suppression of them could weaken the capacity of CSCs [42-44]. Many CSC markers have been reported for CCA, including ALDH, CD13, CD44, CD44v9, CD90, CD133, CD147, EpCAM, SOX2, *etc.* [1, 20, 21]. High expression of these CSC markers is associated with poor clinical outcomes in CCA patients [45]. A high level of ALDH expression was associated with the high migratory capacity and gemcitabine resistance of CCA cells [43]. Overexpression of CD44, CD44v6, CD44v8-10 and EpCAM increases the predictability of post-operative CCA recurrence [20]. CD44v9 was demonstrated to induce stem-like phenotypes and enhance CCA progression via Wnt/β-catenin signaling pathway [42]. In our study, KKU-055-CSC showed a high expression of CSC markers, including CD44, CD147, SOX2, EpCAM, *Oct3/4, c-myc,* and *OLG2*. Our proteomics data also confirmed that KKU-055-CSC expressed a higher level of CD44, CD38, EpCAM, and KLF4 than its differentiated form and the parental cancer cell **(supplementary data)**. This information suggested that these markers could be the potential CSC markers applicable as a target for treating CCA in the future.

The functional analysis showed that KKU-055-CSC gains the ability on chemoresistance. Supporting this function, the data from proteomics in this study showed that KKU-055-CSC expressed a higher level of proteins related to multi-drug resistance, including ALDH3A1, ALDH1A1, and ATP binding cassette (ABC) transporters, compared with its differentiated form and the parental cancer cell **(supplementary data).**

Collectively, our data showed that the KKU-055-CSC is potentially a representative CCA stem-like cell. However, it is known that the population of CSCs are heterogenous; the different CSC clones with different degree of stemness may exhibit different phenotypes and aggressiveness. Based on this fact, it is suggested that KKU-055-CSC could represent a particular sub-group of CCA-CSCs. For more understanding of CSC of CCA, the more CSC-like clones with different phenotypes and degrees of stemness may need to be isolated and characterized.

A comprehensive comparison of protein expression profiles by mass spectrometry is a powerful tool for analyzing the complexity of cellular molecular signaling. Our global proteome analysis revealed the distinct protein expression pattern of KKU-055-CSC compared with its differentiating form (DIF) and the parental cell (Par). We found the up-regulated integrin signaling pathway in KKU-055-DIF and the parental KKU-055, suggesting the involvement of this signaling pathway during CSC differentiation. Our finding agrees with the previous reports including ours in glioma that the integrin signaling is enhanced during CSC-differentiation and may serve as a niche for CSC differentiation [17, 19]. Unlike normal stem cells, CSC has the potential to differentiate into the cancer cell and attenuate the aggressive behavior [46]. In this study, we used FCS to induce CSC differentiation, and found that protein expression patterns in differentiating cancer stem cells were similarly and reciprocally transitioned between cancer stem cells and cancer cells, suggesting that differentiating cancer stem cells are in an intermediate stage of differentiation into cancer cells from

cancer stem cells.

Based on pathway analysis using KeyMolnet, the transcriptional regulation pathway by HMGA protein was extracted as one of the most enhanced network in CSC, and was dramatically declined during differentiation. This information suggested the role of HMGA signaling in the stemness maintenance of CCA stem cells. HMGA proteins, including HMGA1 and HMGA2, are chromatin architectural proteins that interact with the transcriptional types of machinery to change chromatin structure and regulate the transcription of many genes [47]. Many studies showed that HMGA proteins are highly expressed in embryonic stem cells and some adult stem cells, such as hematopoietic stem cells and intestinal stem cells, indicating their fundamental role in self-renewal and differentiation in the stem cells [47]. Recently, HMGA proteins were found to be overexpressed in many types of cancer [48] including cholangiocarcinoma. In our *in-silico* study, HMGA1 expression profile of TCGA CCA tissues showed that HMGA1 was overexpressed in human CCA tissues and high expression of HMGA1 was correlated with poor survival of CCA patients **(supplementary data).** The previous studies demonstrated that HMGA1 facilitates stemness maintenance and inhibits the differentiation of human embryonic stem cells. Overexpression of HMGA1 could enhance the expression of *OCT4, NANOG, SOX,* and *C-myc* [49, 50]. In breast cancer, HMGA proteins were found to play a critical role in the epithelial-mesenchymal transition (EMT) [51]. However, the function of HMGA protein in CSC is not yet clearly understood. Based on our network analysis, HMGA1 and stem cell factor-related proteins (SOX2, CD44, CD133, EpCAM, KLF4, Nanog, OCT4, OCT1, SWI/SNF, etc.) are associated through STAT3 activation **(Figure 7B)**. In a previous study, it has been demonstrated that STAT3 is a critical downstream target of HMGA1. It also showed that HMGA1 binds to a conserved region of the *STAT3* promoter *which can* increase the expression level of STAT3 in leukemia cells from *HMGA1a* transgenic mice. Belton, A. *et. al.* found a STAT3 consensus DNA-binding site (TTN5AA) in the HMGA1 promoter region that is conserved in humans and mice and could mediate STAT3 dimer binding and transactivation [52]. Our results demonstrate that the expression of STAT3 is higher in CSC conditions compared with DIF while not expressed in parental cells (**supplementary data).** The results from the extracted network in this study suggested that STAT3 regulates the expression of HMGA1 in a feed-forward manner, controlling the expression of STST3 itself via HMGA1 regulation which could drive a stem-like cell state in CCA-CSCs.

According to our experimental results, HMGA1 was found to regulate cell survival and drug resistance of CCA-CSC (**Figure 8**). We demonstrated that suppression of HMGA1 by siRNA contributes to growth enhancement of CCA-CSC and inhibits stemness maintenance resulting in decreased expression of stemness-related molecules, such as CD44 and SOX2. Similar findings are reported in those of ovarian cancer [53], suggesting HMGA1 playing a critical role in the transcriptional control of stemness-related genes in CSCs. The knockdown of HMGA1 also led to increased sensitivity against chemotherapeutic reagents, including 5-FU, cisplatin, and gemcitabine in CCA-CSC. HMGA1-related chemotherapy resistances were also found in colon cancer cells [53, 54], glioblastoma stem cells [55], suggesting that HMGA1 is a key regulator of drug resistance. In this study, we demonstrated for the first time that the suppression of HMGA1 can reduce stemness

maintenance, induce cell differentiation in CSC, and enhance the chemosensitivity of CCA-CSC.

Besides, the Aurora signaling pathway was also found to be highly enriched in CSCs. Aurora kinase A (Aurora-A) is a serine/threonine kinase that play an important role in many steps in cell biology during meiosis [56]. In glioma stem cells, Aurora-A regulates the self-renewal and tumorigenicity by interaction with AXIN and disruption of the AXIN/GSK3b/β-catenin destruction complex, resulting in activation of Wnt signaling [57]. Moreover, Aurora-A can translocate to the nucleus which enhances breast cancer stem cell (BCSC) phenotype by activation of MYC promoter [58]. We showed that the treatment of Aurora-A inhibitors can reduce the cell viability of both CCA-CSC and parental CCA cells. Interestingly, the suppression of HMGA1 can enhance the Aurora-A inhibitor effects, which are specifically higher in CSC, suggesting the important role of HMGA1 and Aurora A in CCS-CSC. However, the information of HMGA1 associated with Aurora A is very limited. Our study here could be the first report to present the relationship between HMGA1 and Aurora A in CSC. We tried to validate this relationship by using global mRNA profiles of CCA generated in TCGA. The result showed a strong correlation between *HMGA1* and *AURKA* mRNA levels in human CCA with the highest Pearson index (R^2) of 0.73 (p-value 1.7 x 10^{-8}) **(supplementary data).** Here, we first demonstrate the strong correlation between *HMGA1* and *AURKA in CCA*.

In conclusion, our findings show that HMGA1-SYAT3-AuroraA related pathways are the key regulator in the maintenance of CSC-like features in CCA stem cells. Targeting HMGA1-(STST3)-Aurora-A may be a promising new strategy for CSC-eradication therapeutics, although further study will be needed to clarify the indeed mechanism.

References

1. Cardinale V, Renzi A, Carpino G, Torrice A, Bragazzi MC, Giuliante F, DeRose AM, Fraveto A, Onori P, Napoletano C *et al*: **Profiles of cancer stem cell subpopulations in cholangiocarcinomas**. *Am J Pathol* 2015, **185**(6):1724-1739.
2. Sripa B, Pairojkul C: **Cholangiocarcinoma: lessons from Thailand**. *Curr Opin Gastroenterol* 2008, **24**(3):349-356.
3. Kodali S, Shetty A, Shekhar S, Victor DW, Ghobrial RM: **Management of Intrahepatic Cholangiocarcinoma**. *J Clin Med* 2021, **10**(11).
4. DeOliveira ML, Cunningham SC, Cameron JL, Kamangar F, Winter JM, Lillemoe KD, Choti MA, Yeo CJ, Schulick RD: **Cholangiocarcinoma: thirty-one-year experience with 564 patients at a single institution**. *Ann Surg* 2007, **245**(5):755-762.
5. Lee KJ, Yi SW, Cha J, Seong J, Bang S, Song SY, Kim HM, Park SW: **A pilot study of concurrent chemoradiotherapy with gemcitabine and cisplatin in patients with locally advanced biliary tract cancer**. *Cancer Chemother Pharmacol* 2016, **78**(4):841-846.
6. Sripa B, Seubwai W, Vaeteewoottacharn K, Sawanyawisuth K, Silsirivanit A, Kaewkong W, Muisuk K, Dana P, Phoomak C, Lert-Itthiporn W *et al*: **Functional and genetic characterization of three cell lines derived from a single tumor of an Opisthorchis viverrini-associated cholangiocarcinoma patient**. *Hum Cell* 2020, **33**(3):695-708.

7. Luvira V, Eurboonyanun C, Bhudhisawasdi V, Pugkhem A, Pairojkul C, Luvira V, Sathitkarnmanee E, Somsap K, Kamsa-ard S: **Patterns of Recurrence after Resection of Mass-Forming Type Intrahepatic Cholangiocarcinomas**. *Asian Pac J Cancer Prev* 2016, **17**(10):4735-4739.
8. Kawamoto M, Umebayashi M, Tanaka H, Koya N, Nakagawa S, Kawabe K, Onishi H, Nakamura M, Morisaki T: **Combined Gemcitabine and Metronidazole Is a Promising Therapeutic Strategy for Cancer Stem-like Cholangiocarcinoma**. *Anticancer Res* 2018, **38**(5):2739-2748.
9. Klonisch T, Wiechec E, Hombach-Klonisch S, Ande SR, Wesselborg S, Schulze-Osthoff K, Los M: **Cancer stem cell markers in common cancers - therapeutic implications**. *Trends Mol Med* 2008, **14**(10):450-460.
10. Batlle E, Clevers H: **Cancer stem cells revisited**. *Nat Med* 2017, **23**(10):1124-1134.
11. Klemba A, Purzycka-Olewiecka JK, Wcislo G, Czarnecka AM, Lewicki S, Lesyng B, Szczylik C, Kieda C: **Surface markers of cancer stem-like cells of ovarian cancer and their clinical relevance**. *Contemp Oncol (Pozn)* 2018, **22**(1A):48-55.
12. Murar M, Vaidya A: **Cancer stem cell markers: premises and prospects**. *Biomarkers in Medicine* 2015, **9**(12):1331-1342.
13. Pastrana E, Silva-Vargas V, Doetsch F: **Eyes Wide Open: A Critical Review of Sphere-Formation as an Assay for Stem Cells**. *Cell Stem Cell* 2011, **8**(5):486-498.
14. Jariyal H, Gupta C, Bhat VS, Wagh JR, Srivastava A: **Advancements in Cancer Stem Cell Isolation and Characterization**. *Stem Cell Reviews and Reports* 2019, **15**(6):755-773.
15. Rao GH, Liu HM, Li BW, Hao JJ, Yang YL, Wang MR, Wang XH, Wang J, Jin HJ, Du L *et al*: **Establishment of a human colorectal cancer cell line P6C with stem cell properties and resistance to chemotherapeutic drugs**. *Acta Pharmacologica Sinica* 2013, **34**(6):793-804.
16. Masciale V, Grisendi G, Banchelli F, D'Amico R, Maiorana A, Sighinolfi P, Stefani A, Morandi U, Dominici M, Aramini B: **Isolation and Identification of Cancer Stem-Like Cells in Adenocarcinoma and Squamous Cell Carcinoma of the Lung: A Pilot Study**. *Front Oncol* 2019, **9**:1394.
17. Niibori-Nambu A, Midorikawa U, Mizuguchi S, Hide T, Nagai M, Komohara Y, Nagayama M, Hirayama M, Kobayashi D, Tsubota N *et al*: **Glioma initiating cells form a differentiation niche via the induction of extracellular matrices and integrin alphaV**. *PLoS One* 2013, **8**(5):e59558.
18. Matsuda K, Sato A, Okada M, Shibuya K, Seino S, Suzuki K, Watanabe E, Narita Y, Shibui S, Kayama T *et al*: **Targeting JNK for therapeutic depletion of stem-like glioblastoma cells**. *Scientific Reports* 2012, **2**.
19. Narushima Y, Kozuka-Hata H, Koyama-Nasu R, Tsumoto K, Inoue J, Akiyama T, Oyama M: **Integrative Network Analysis Combined with Quantitative Phosphoproteomics Reveals Transforming Growth Factor-beta Receptor type-2 (TGFBR2) as a Novel**

Regulator of Glioblastoma Stem Cell Properties. *Mol Cell Proteomics* 2016, **15**(3):1017-1031.

20. Padthaisong S, Thanee M, Namwat N, Phetcharaburanin J, Klanrit P, Khuntikeo N, Titapun A, Sungkhamanon S, Saya H, Loilome W: **Overexpression of a panel of cancer stem cell markers enhances the predictive capability of the progression and recurrence in the early stage cholangiocarcinoma**. *J Transl Med* 2020, **18**(1):64.
21. McGrath NA, Fu J, Gu SZ, Xie C: **Targeting cancer stem cells in cholangiocarcinoma (Review)**. *Int J Oncol* 2020, **57**(2):397-408.
22. Correnti M, Raggi C: **Stem-like plasticity and heterogeneity of circulating tumor cells: current status and prospect challenges in liver cancer**. *Oncotarget* 2017, **8**(4):7094-7115.
23. Putthisen S, Silsirivanit A, Panawan O, Niibori-Nambu A, Nishiyama-Ikeda Y, Main P, Luang S, Ohta K, Muisuk K, Wongkham S *et al*: **Targeting alpha2,3-sialylated glycan in glioma stem-like cells by Maackia amurensis lectin-II: A promising strategy for glioma treatment**. *Exp Cell Res* 2021:112949.
24. Ohta K, Kawano R, Ito N: **Lactic acid bacteria convert human fibroblasts to multipotent cells**. *PLoS One* 2012, **7**(12):e51866.
25. Okada S, Vaeteewoottacharn K, Kariya R: **Application of Highly Immunocompromised Mice for the Establishment of Patient-Derived Xenograft (PDX) Models**. *Cells* 2019, **8**(8).
26. Akaboshi S, Watanabe S, Hino Y, Sekita Y, Xi Y, Araki K, Yamamura K, Oshima M, Ito T, Baba H *et al*: **HMGA1 is induced by Wnt/beta-catenin pathway and maintains cell proliferation in gastric cancer**. *Am J Pathol* 2009, **175**(4):1675-1685.
27. Shaffer LG MG-JJ, Shmid M: **An international system for human cytogenetic nomenclature**. *Karger Medical and Scientific Publishers* 2013.
28. Masuda T, Ishihama, Y.: **Sample Preparation for Shotgun Proteomics by Using Phase Transfer Surfactants**. *Proteome Letters* 2016, **1**.
29. Rudolph JD, Cox J: **A Network Module for the Perseus Software for Computational Proteomics Facilitates Proteome Interaction Graph Analysis**. *J Proteome Res* 2019, **18**(5):2052-2064.
30. Beck S, Michalski A, Raether O, Lubeck M, Kaspar S, Goedecke N, Baessmann C, Hornburg D, Meier F, Paron I *et al*: **The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics**. *Mol Cell Proteomics* 2015, **14**(7):2014-2029.
31. Yang P, Humphrey SJ, Cinghu S, Pathania R, Oldfield AJ, Kumar D, Perera D, Yang JYH, James DE, Mann M *et al*: **Multi-omic Profiling Reveals Dynamics of the Phased Progression of Pluripotency**. *Cell Syst* 2019, **8**(5):427-445 e410.
32. Hillion J, Dhara S, Sumter TF, Mukherjee M, Di Cello F, Belton A, Turkson J, Jaganathan S, Cheng L, Ye Z *et al*: **The high-mobility group A1a/signal transducer and activator of transcription-3 axis: an achilles heel for hematopoietic malignancies?** *Cancer Res* 2008, **68**(24):10121-10127.

33. Alisi A, Cho WC, Locatelli F, Fruci D: **Multidrug resistance and cancer stem cells in neuroblastoma and hepatoblastoma**. *Int J Mol Sci* 2013, **14**(12):24706-24725.
34. Colamaio M, Tosti N, Puca F, Mari A, Gattordo R, Kuzay Y, Federico A, Pepe A, Sarnataro D, Ragozzino E *et al*: **HMGA1 silencing reduces stemness and temozolomide resistance in glioblastoma stem cells**. *Expert Opin Ther Targets* 2016, **20**(10):1169-1179.
35. Rao GH, Liu HM, Li BW, Hao JJ, Yang YL, Wang MR, Wang XH, Wang J, Jin HJ, Du L *et al*: **Establishment of a human colorectal cancer cell line P6C with stem cell properties and resistance to chemotherapeutic drugs**. *Acta Pharmacol Sin* 2013, **34**(6):793-804.
36. Palmini G, Zonefrati R, Mavilia C, Aldinucci A, Luzi E, Marini F, Franchi A, Capanna R, Tanini A, Brandi ML: **Establishment of Cancer Stem Cell Cultures from Human Conventional Osteosarcoma**. *J Vis Exp* 2016(116).
37. Jariyal H, Gupta C, Bhat VS, Wagh JR, Srivastava A: **Advancements in Cancer Stem Cell Isolation and Characterization**. *Stem Cell Rev Rep* 2019, **15**(6):755-773.
38. Ghanei Z, Jamshidizad A, Joupari MD, Shamsara M: **Isolation and characterization of breast cancer stem cell-like phenotype by Oct4 promoter-mediated activity**. *J Cell Physiol* 2020, **235**(11):7840-7848.
39. Raggi C, Correnti M, Sica A, Andersen JB, Cardinale V, Alvaro D, Chiorino G, Forti E, Glaser S, Alpini G *et al*: **Cholangiocarcinoma stem-like subset shapes tumor-initiating niche by educating associated macrophages**. *J Hepatol* 2017, **66**(1):102-115.
40. Wang M, Xiao J, Shen M, Yahong Y, Tian R, Zhu F, Jiang J, Du Z, Hu J, Liu W *et al*: **Isolation and characterization of tumorigenic extrahepatic cholangiocarcinoma cells with stem cell-like properties**. *Int J Cancer* 2011, **128**(1):72-81.
41. Phi LTH, Sari IN, Yang YG, Lee SH, Jun N, Kim KS, Lee YK, Kwon HY: **Cancer Stem Cells (CSCs) in Drug Resistance and their Therapeutic Implications in Cancer Treatment**. *Stem Cells Int* 2018, **2018**:5416923.
42. Suwannakul N, Ma N, Midorikawa K, Oikawa S, Kobayashi H, He F, Kawanishi S, Murata M: **CD44v9 Induces Stem Cell-Like Phenotypes in Human Cholangiocarcinoma**. *Front Cell Dev Biol* 2020, **8**:417.
43. Chen MH, Weng JJ, Cheng CT, Wu RC, Huang SC, Wu CE, Chung YH, Liu CY, Chang MH, Chen MH *et al*: **ALDH1A3, the Major Aldehyde Dehydrogenase Isoform in Human Cholangiocarcinoma Cells, Affects Prognosis and Gemcitabine Resistance in Cholangiocarcinoma Patients**. *Clin Cancer Res* 2016, **22**(16):4225-4235.
44. Yan Y, Li Z, Kong X, Jia Z, Zuo X, Gagea M, Huang S, Wei D, Xie K: **KLF4-Mediated Suppression of CD44 Signaling Negatively Impacts Pancreatic Cancer Stemness and Metastasis**. *Cancer Res* 2016, **76**(8):2419-2431.
45. Gu MJ, Jang BI: **Clinicopathologic significance of Sox2, CD44 and CD44v6 expression in intrahepatic cholangiocarcinoma**. *Pathol Oncol Res* 2014, **20**(3):655-660.
46. Arima Y, Nobusue H, Saya H: **Targeting of cancer stem cells by differentiation therapy**. *Cancer Sci* 2020, **111**(8):2689-2695.

47. Parisi S, Piscitelli S, Passaro F, Russo T: **HMGA Proteins in Stemness and Differentiation of Embryonic and Adult Stem Cells**. *Int J Mol Sci* 2020, **21**(1).
48. Sepe R, Piscuoglio S, Quintavalle C, Perrina V, Quagliata L, Formisano U, Terracciano LM, Fusco A, Pallante P: **HMGA1 overexpression is associated with a particular subset of human breast carcinomas**. *J Clin Pathol* 2016, **69**(2):117-121.
49. Shah SN, Kerr C, Cope L, Zambidis E, Liu C, Hillion J, Belton A, Huso DL, Resar LM: **HMGA1 reprograms somatic cells into pluripotent stem cells by inducing stem cell transcriptional networks**. *PLoS One* 2012, **7**(11):e48533.
50. Yanagisawa BL, Resar LM: **Hitting the bull's eye: targeting HMGA1 in cancer stem cells**. *Expert Rev Anticancer Ther* 2014, **14**(1):23-30.
51. Pegoraro S, Ros G, Piazza S, Sommaggio R, Ciani Y, Rosato A, Sgarra R, Del Sal G, Manfioletti G: **HMGA1 promotes metastatic processes in basal-like breast cancer regulating EMT and stemness**. *Oncotarget* 2013, **4**(8):1293-1308.
52. Belton A, Xian L, Huso T, Koo M, Luo LZ, Turkson J, Page BD, Gunning PT, Liu G, Huso DL *et al*: **STAT3 inhibitor has potent antitumor activity in B-lineage acute lymphoblastic leukemia cells overexpressing the high mobility group A1 (HMGA1)-STAT3 pathway**. *Leuk Lymphoma* 2016, **57**(11):2681-2684.
53. Kim DK, Seo EJ, Choi EJ, Lee SI, Kwon YW, Jang IH, Kim SC, Kim KH, Suh DS, Seong-Jang K *et al*: **Crucial role of HMGA1 in the self-renewal and drug resistance of ovarian cancer stem cells**. *Exp Mol Med* 2016, **48**:e255.
54. D'Angelo D, Mussnich P, Rosa R, Bianco R, Tortora G, Fusco A: **High mobility group A1 protein expression reduces the sensitivity of colon and thyroid cancer cells to antineoplastic drugs**. *BMC Cancer* 2014, **14**:851.
55. Puca F, Tosti N, Federico A, Kuzay Y, Pepe A, Morlando S, Savarese T, D'Alessio F, Colamaio M, Sarnataro D *et al*: **HMGA1 negatively regulates NUMB expression at transcriptional and post transcriptional level in glioblastoma stem cells**. *Cell Cycle* 2019, **18**(13):1446-1457.
56. Li M, Gao K, Chu L, Zheng J, Yang J: **The role of Aurora-A in cancer stem cells**. *Int J Biochem Cell Biol* 2018, **98**:89-92.
57. Xia Z, Wei P, Zhang H, Ding Z, Yang L, Huang Z, Zhang N: **AURKA governs self-renewal capacity in glioma-initiating cells via stabilization/activation of beta-catenin/Wnt signaling**. *Mol Cancer Res* 2013, **11**(9):1101-1111.
58. Zheng F, Yue C, Li G, He B, Cheng W, Wang X, Yan M, Long Z, Qiu W, Yuan Z *et al*: **Nuclear AURKA acquires kinase-independent transactivating function to enhance breast cancer stem cell phenotype**. *Nat Commun* 2016, **7**:10180.

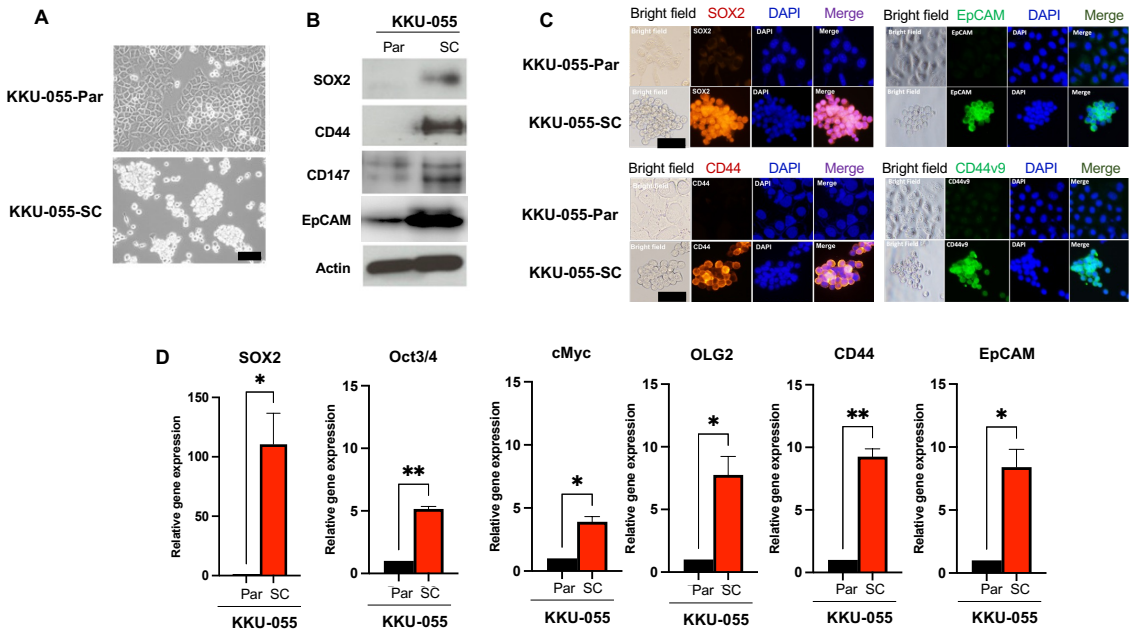

Figure 1. The phenotypic character of cell-line derived cancer stem-like cell of cholangiocarcinoma and its parental cell line. (A) Phase contrast microscopic analysis of cell morphologies of KKU-055-Parental cells and KKU-055-CSCs at >50 passages. KKU-055-Parental cells show the adherent polygonal-like shapes while KKU-055-CSCs exhibit the sphere forming shapes (non-adherent forms). Scale bar = 50 um. The expression of stem cell markers in KKU-055-CSCs comparing with KKU-055-Parental cells. Expression of stem cell markers were determined by **(B)** Western Blotting for the identification of SOX2, CD44, CD147, and EpCAM. The b-actin was used for an internal marker, **(C)** Immunocytofluoresence analyses for the expression and cellular localization of SOX2, CD44, CD44 v9, and EpCAM using those antibodies. DAPI was used for nuclear staining. (D) Real time-PCR for quantitation of *Sox2, C-myc, OLG-2*, *Oct 3/4*, *CD44,* and *EpCAM*. All stem cell markers were significantly up-regulated in KKU-055-CSCs comparing with KKU-055-Parental cells. (*p-value* < 0.05)

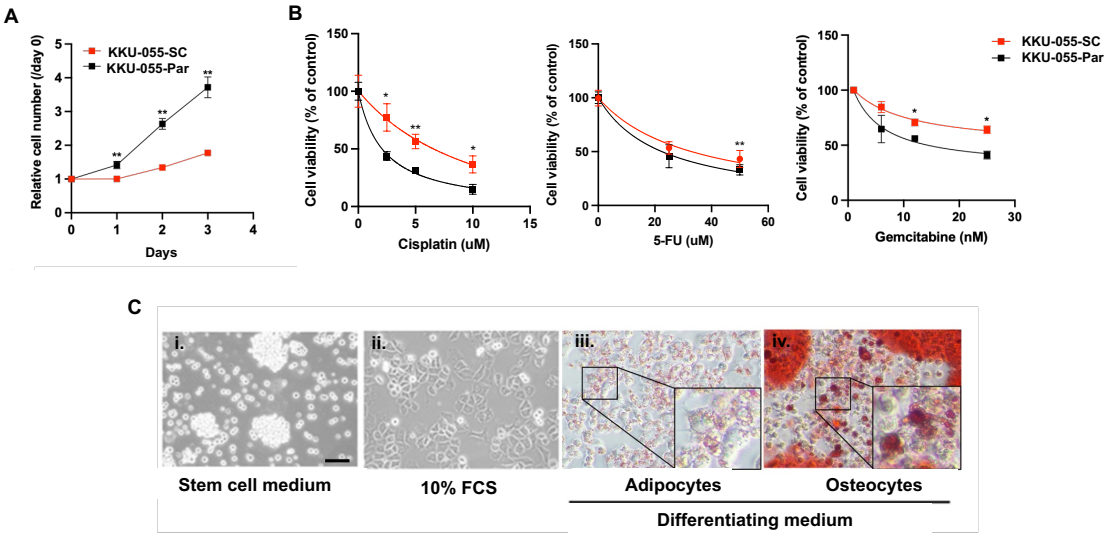

Figure 2. Characterization of KKU-055-CSCs phenotypes (A) Proliferation of KKU-055-CSCs was compared to that of the Parental cells in a cell proliferation assay using the CCK-8 Cell Counting Kit. **(B)** Dose-response

curves of KKU-055-CSCs and their parental cells against 5-FU, Cisplatin, and Gemcitabine treatments. The cell viabilities were analyzed by Trypan blue methods as described in the material and methods. **(C)** Phase contrast microscope analyses of the cell morphologies of KKU-055-CS during culturing in stem cell medium (i), inducing the differentiation to CCA by the addition of 10% FCS for 3 days (ii), inducing the specific adipocyte differentiation by incubating with StemPro™ Adipogenesis Differentiation medium for 14 days followed by Oil Red O staining to detect the fat-droplet (red) (iii), and inducing the specific osteocyte differentiation by incubating with StemPro™ Osteogenesis Differentiation medium for 14 days, followed by Alizarin Red S staining of calcium granule detection (red)(iv). (scale bar = 50 um, *p-$value$ < 0.05, ** p-$value$ < 0.01)

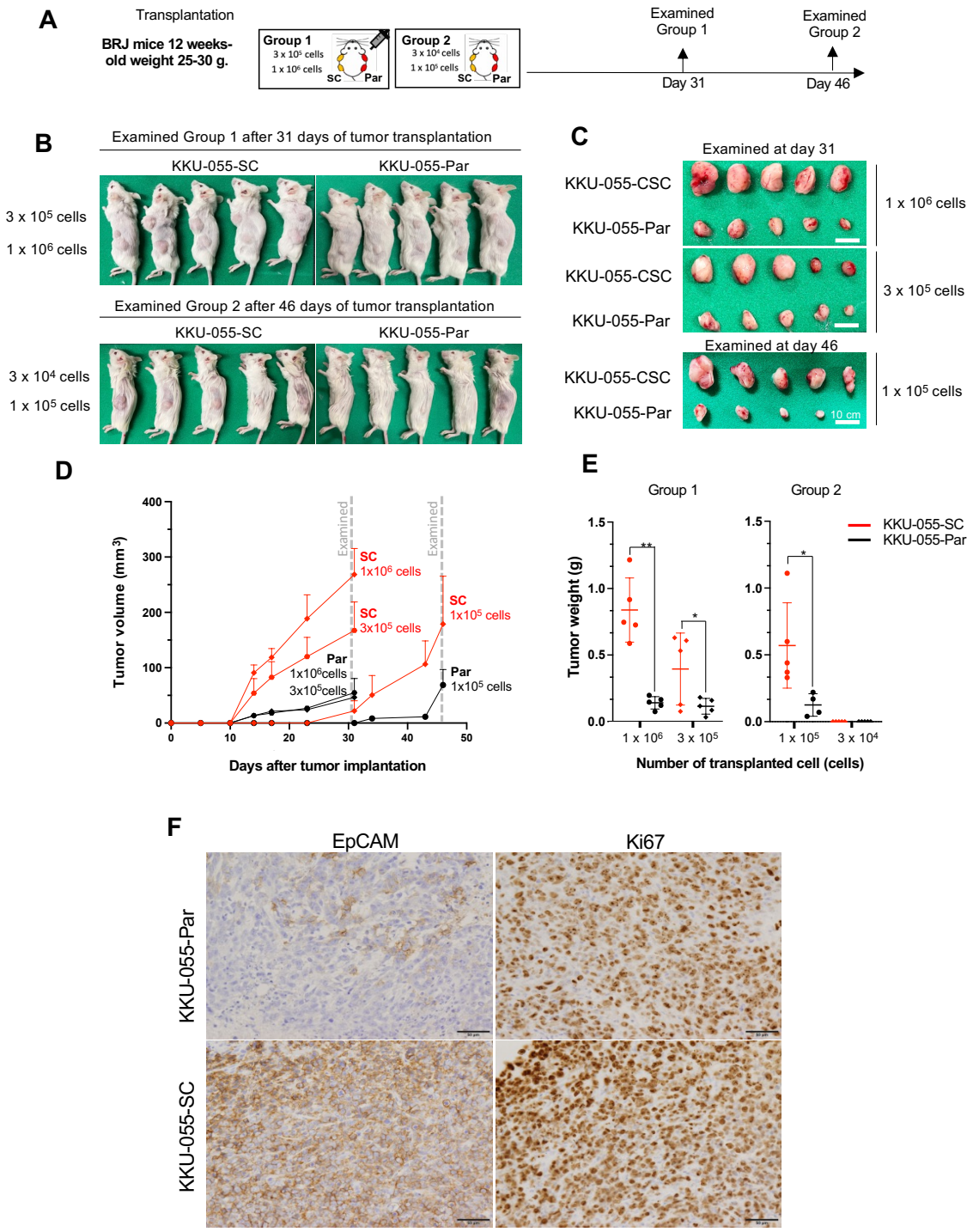

Figure 3. Tumor initiating ability of KKU-055-CSCs in the mouse transplantation model Tumor initiating ability of KKU-055-CSCs was compared to that of the Parental cells after the subcutaneous transplantations into immunodeficient mice (KKU-055-CSC; n=15, Parental KKU-055-CSC; n=15). **(A)** The BRJ mice with 8 to 12-week-old were separated into 2 groups. KKU-055-CSSCs and KKU-055-Parental cells were injected in Group 1; $1x10^6$ cells/nodule (n=5) and $3x10^5$ cells/nodule (n=5), in group 2; $1x10^5$ cells/nodule (n=5). The mice in group 1 were sacrificed at 31 days after injection and in group 2 were sacrificed at 46 days after injection. The tumor sizes, tumor volume, and weight were shown in **(B, C), (D),** and **(E),** respectively. **(F)** Immunohistochemistry of the tumors excised from the mice transplanted the KKU-055-CSCs and -Parental cells, using antibodies against EpCAM

and Ki67. KKU-055-CSCs formed expanded malignant tumors with a high proliferative index and high expression of EpCAM and Ki67 compared to those of parental cells.

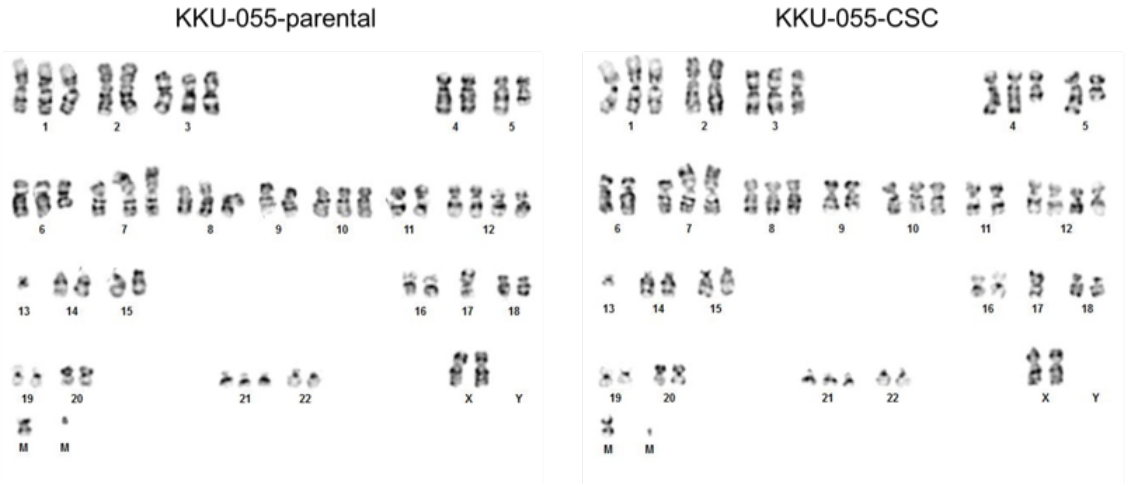

Figure 4 Karyotypes of KKU-055-Parental cell (53+XX) and KKU-055-CSC (53+XX) The numbers of chromosomes of 20 cells of each KKU-055-Parental cell and KKU-055-CSC were showed similar chromosomal abnormalities. The equal modal chromosome number of 54 (53-56 chromosomes of parental KKU-055 cells and 52-55 chromosomes of KKU-055-CSCs). Several chromosome markers (M, mar) were similar in both, such as for parental KKU-055: 53+XX; +1, +3, +6, +7, +8, +10, +12, -13, -17, +21, +mar and for KKU-055-CSC: 53+XX; +1, +3, +4, +7, +8, +10, +12, -13, -17, +21, +mar.

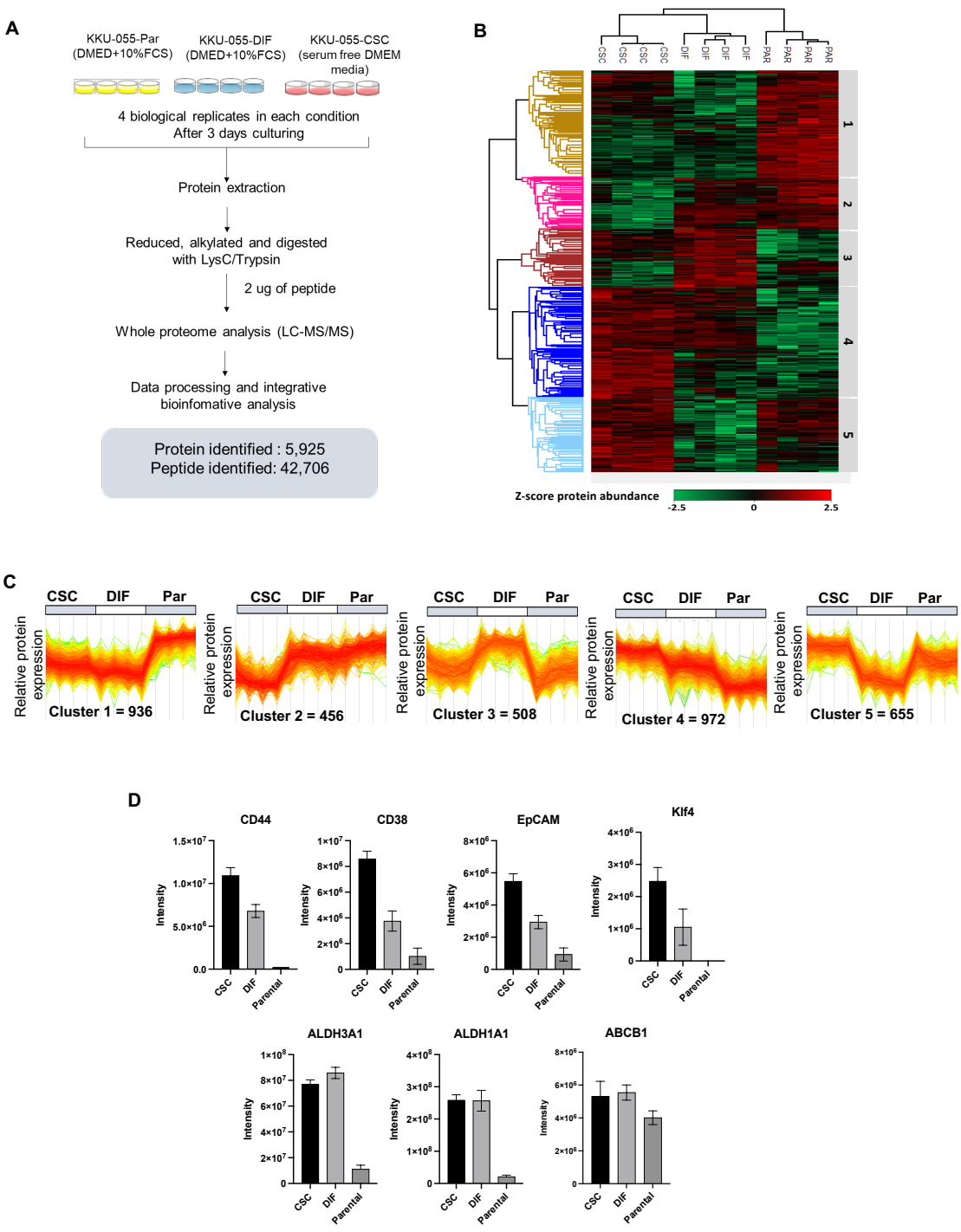

Figure 5. Proteome profiling of KKU-055-CSC, KKU-055-DIF and KKU-055-parental (A) Schematic diagram of the experimental protocol for proteome profiling by mass spectrometry. Three groups of the cells:KKU-055-CSC, KKU-055-DIF(differentiating CSCs induced by the FCS) and KKU-055-parentals with 4 replicates each were used. **(B)** Cluster analysis of the quantitatively identified proteins in the three different stages of cell transition. Each horizontal column represents a sample map of four replicates in each condition, and each vertical row represents an individual protein, with relative expression values displayed as an expression matrix (heat map) using z-score of

log_2 protein abundances ranging from -2.5 (green) to 2.5 (red). Red and green colors indicate increased and decreased protein expression, respectively. **(C)** *k*-mean clustering revealed 5 discrete protein expression patterns and pathway enrichment of each cluster were shown. **(D)** Global proteomics identified the expression of cancer stem cell markers which are members of cluster four from figure 5-C being quantitatively higher in of KKU-055-CSC compared with KKU-055-DIF (after 72 h of differentiation) and KKU-055-Parental cells.

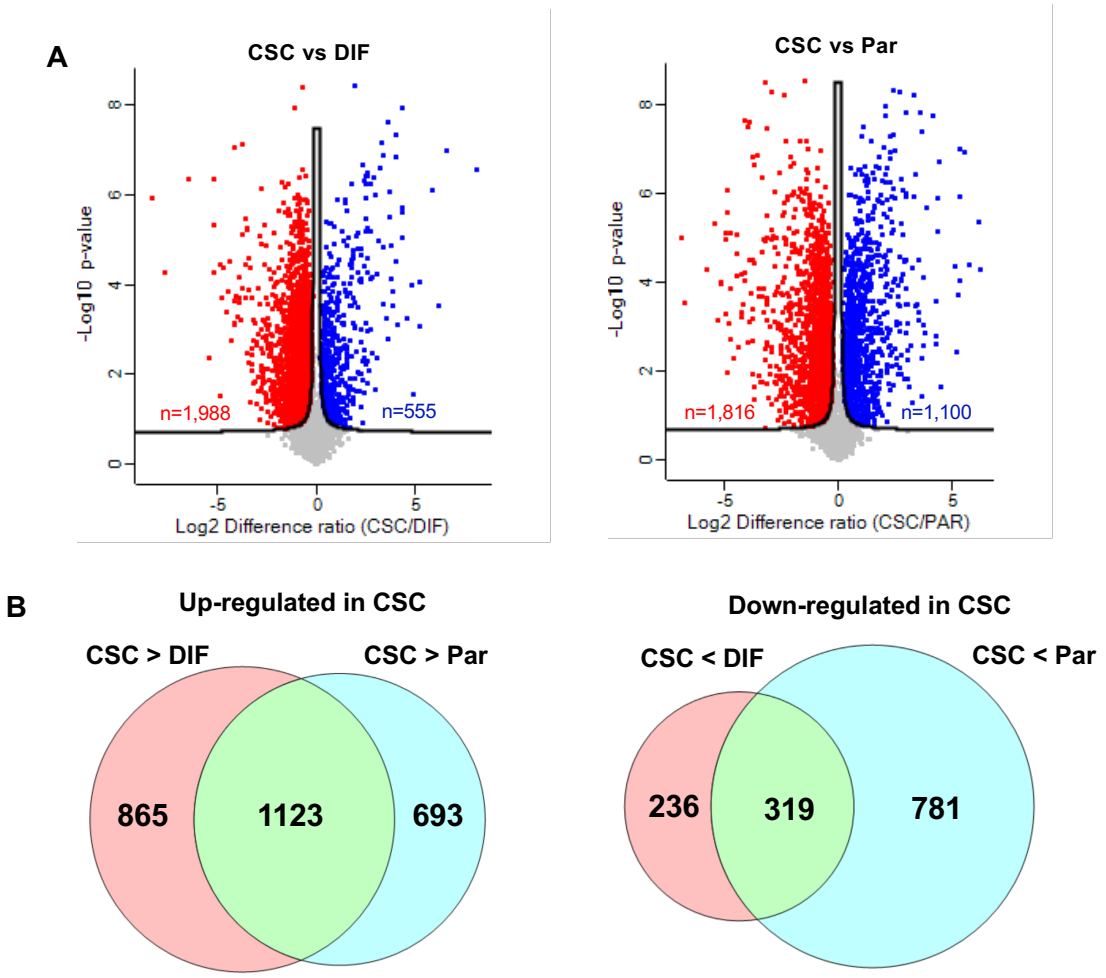

Figure 6. Differential protein expression (A) Volcano Plot for differential protein expression, KKU-055-CSC vs KKU-055-Parental cells and KKU-055-CSC vs KKU-055-DIF cells. Scattered points represent protein: the x-axis is the Log_2 fold change for the ratio the comparison, whereas the y-axis is the Log_{10} *p-value*. Red dots are protein significantly up-regulated in CSC, and Blue dots are those protein significantly regulated in KKU-055-Parental cells and KKU-055-DIF. (FDR=0.05) **(B)** Ven diagram of up/down regulated protein in CSC between KKU-055-CSC vs KKU-055-Parental cells and KKU-055-CSC vs KKU-055-DIF cells.

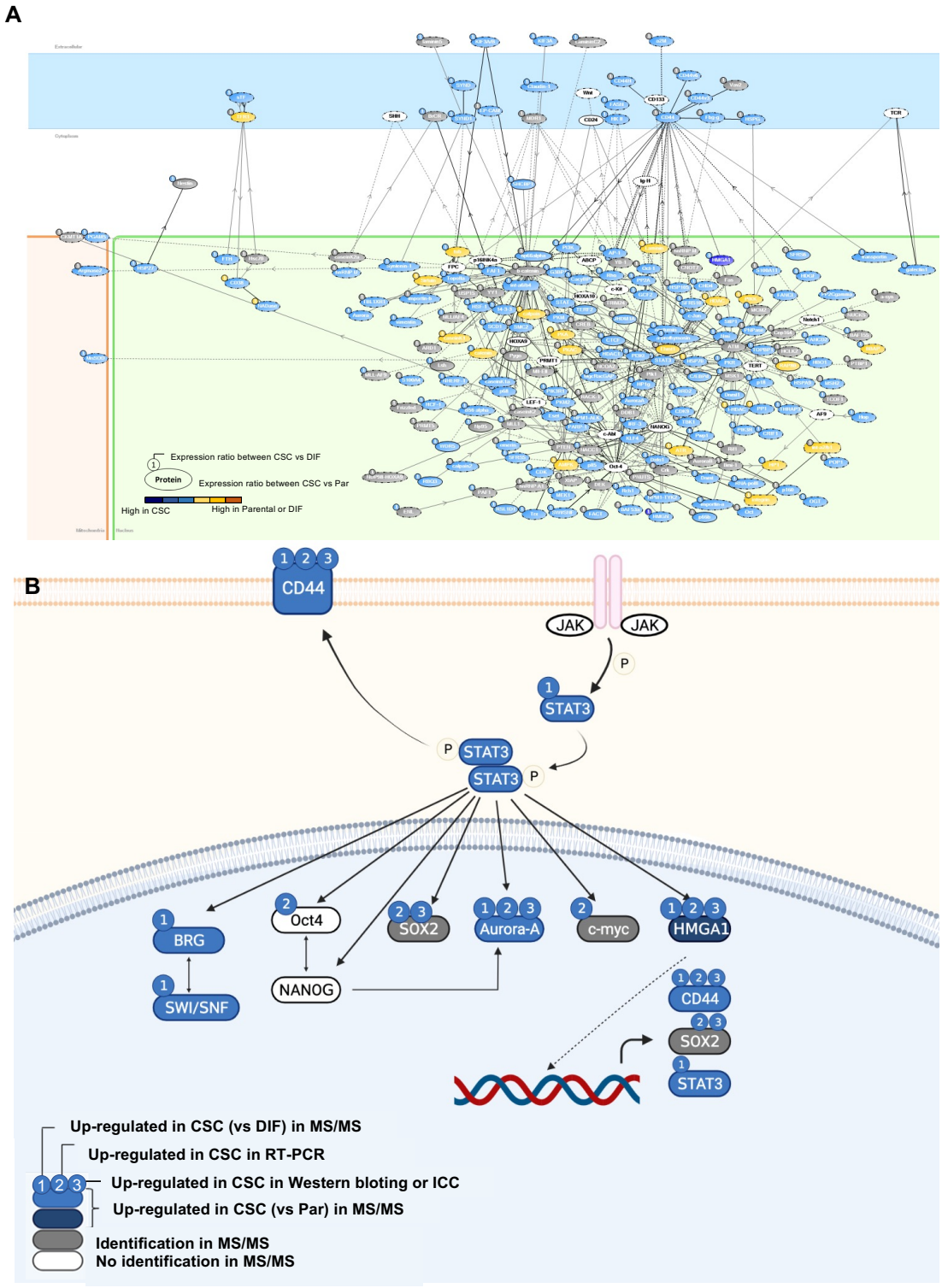

Figure 7. Molecular network analysis of up-regulated proteins in KKU-055-CSC (A) The list of up-regulated proteins in KKU-055-CSC was imported into KeyMolnet by using the "start points and end points" network search algorithm, KeyMolnet reveal a highly complex network of targets with the most statistically significant relationships. **(B)** Extraction of the abnormal molecular network including HMGA1, and related molecules from the molecular network shown in (A). Color set point is described in right panel. The molecular relationships are indicated by a solid line with arrow (direct binding or activation), solid line with arrow and stop (direct inactivation), solid line

without arrow (complex formation), dashed line with arrow (transcriptional activation), and dashed line with arrow and stop (transcriptional repression).

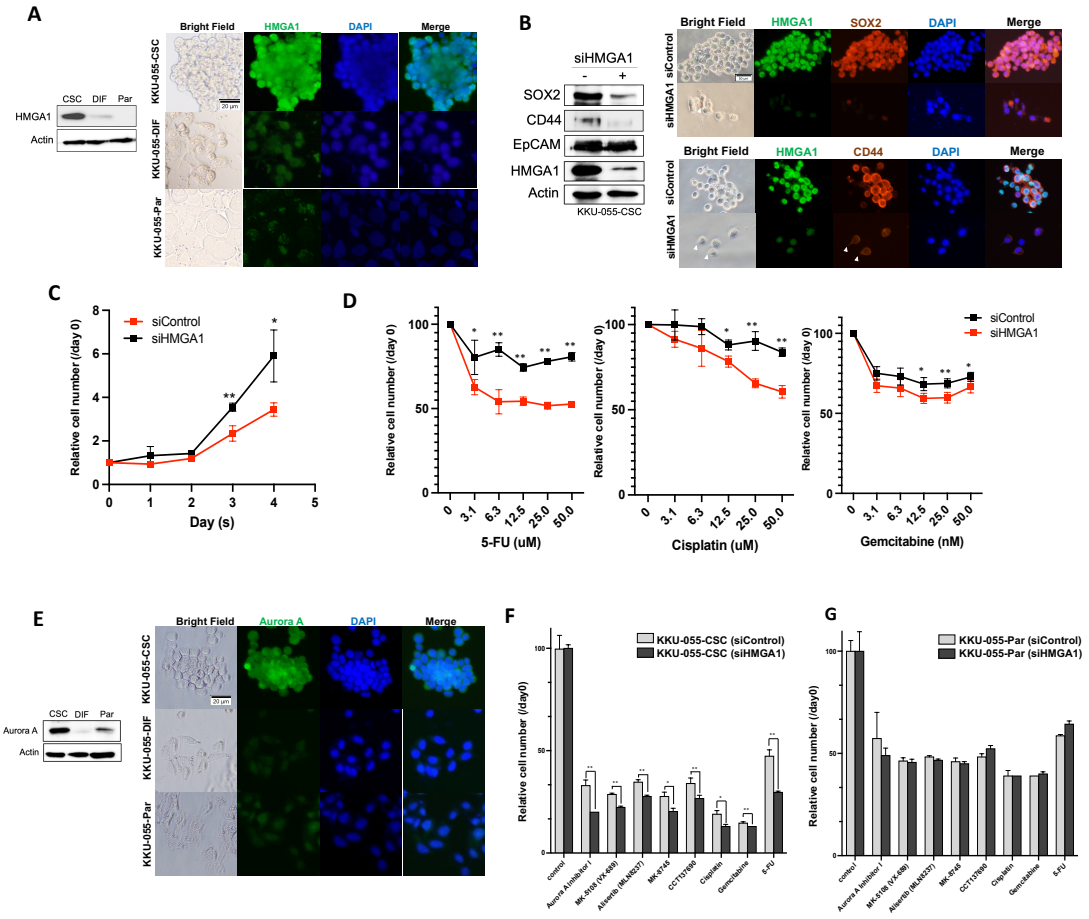

Figure 8. Experimental investigation of HMGA1 signaling pathways associated in cholangiocarcinoma stem cell regulation. (A) Expression of HMGA1 comparing with KKU-055-CSC, -DIF, and –Parental cells was determined by western Blotting. The b-actin was used for an internal marker. Immunocytofluoresence analysis for the nuclear expression of HMGA1 was up-regulated in KKU-055-CSC comparing with KKU-055-DIF and -Parental cells. **(B)** KKU-055-CSC were treated with siHMGA1 (50 pmole) for 72 hr. HMGA1, SOX2 and CD44 expression was depleted after treatment with siHMGA1. **(C)** Cell viability of KKU-055-CSC after treat with siHMGA1 was decrease when compare with control, the cell viability was determined using CCK-8 cell counting kit. **(D)** Cell viability of KKU-055-CSC were determined after combination of siHMGA1 and cancer drug (5-FU, cisplatin, gemcitabine) were analyzed by CCK-8 cell counting kit. **(E)** Expression of Aurora-A compared with KKU-055-CSC, -DIF, and -Parental cells was determined by western Blotting and Immunocytofluoresence. Cell viability of **(F)** KKU-055-CSC and **(G)** KKU-055-Parental were determined after the combination of siHMGA1 and cancer drug (5-FU, cisplatin, gemcitabine) and Aurora kinase inhibitors were analyzed by using CCK-8 cell counting kit. (**p-value* < 0.05, ** *p-value* < 0.01)

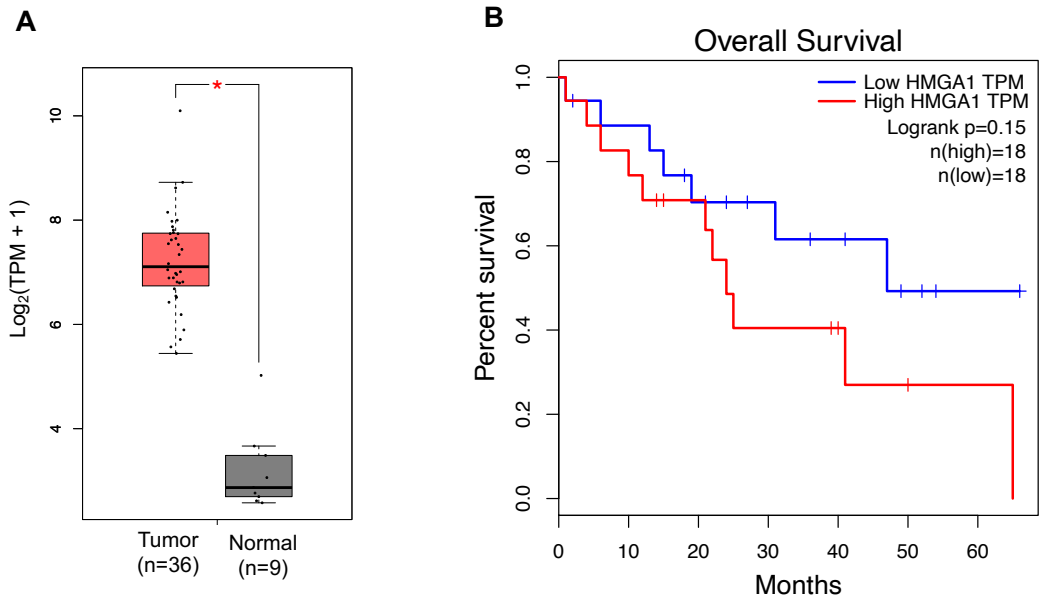

Figure 9. Higher expression of *HMGA1* in human TCGA CCA tissues (CHOL) are correlated with overall survival of patients (A) *HMGA1* mRNA is highly expressed in human TCGA CCA tissues. (tissue number red: glay: more detail explanation) **(B)** Kaplan-Meier survival analysis of low *HMGA1* (Blue n=18) and *HMGA1* high (Red n=18) expression. The cut-off point was median expression of *HMGA1*.

Table 1. Analysis of short tandem repeats (STR) in KKU-055-parental and KKU-055-CSC

Locus	KKU-055 (JCRB1551)	KKU-055-parental	KKU-055-CSC
D8S1179	ND	14, 16, 17	14, 16, 17
D21S11	ND	29, 31	29, 31
D7S820	7, 8	7, 8	7, 8
CSF1PO	10, 11	10, 11	10, 11
D3S1358	ND	16, 17	16, 17
TH01	9, 9	9, 9	9, 9
D13S317	11, 11	11, 11	11, 11
D16S539	10,13	10, 13	10, 13
D2S1338	ND	22, 22	22, 22
D19S433	ND	14, 14	14, 14
vWA	17, 19	17, 19	17, 19
TPOX	8, 8	8, 8	8, 8
D18S51	ND	16, 17	16, 17
D5S818	11, 11	11, 11	11, 11
FGA	ND	26, 26	25, 26
Amelogenin	X	X	X

Table 2. Pathway-based characterization of up- and down-regulated proteins identified in KKU-055-CSC. Pathway enrichment analysis was performed to list up the most enriched pathway categories in the up-**(A)** or down-**(B)** regulated molecules in KKU-055-CSC, using KeyMolnet software.

A. Up-regulated pathways.

Rank	KeyMolNet pathway	score	p-value
1	Spliceosome assembly	54.838	3.11E-17
2	Condensin signaling pathway	41.917	2.41E-13
3	Aurora signaling pathway	26.553	1.02E-08
4	BET family signaling pathway	24.69	3.69E-08
5	Transcriptional regulation by High mobility group protein	20.135	8.68E-07
6	BRCA signaling pathway	19.992	9.59E-07
7	Transcriptional regulation by SREBP	19.086	1.80E-06
8	Kinesin family signaling pathway	17.975	3.88E-06
9	Transcriptional regulation by RB/E2F	17.461	5.54E-06
10	Nucleophosmin signaling pathway	16.426	1.14E-05
11	Arginine methylation	16.23	1.30E-05
12	ATM/ATR signaling pathway	15.604	2.01E-05
13	PARP signaling pathway	14.984	3.09E-05
14	ING signaling pathway	14.815	3.47E-05
15	TopoII signaling pathway	14.63	3.94E-05
16	TACC signaling pathway	14.252	5.13E-05
17	Sirtuin signaling pathway	14.225	5.22E-05
18	CDK signaling pathway	13.901	6.54E-05
19	Transcriptional regulation by Kaiso	13.76	7.21E-05
20	HAT signaling pathway	13.384	9.36E-05

B. Down-regulated pathways.

Rank	KeyMolNet pathway	score	p-value
1	Lipoprotein metabolism	30.307	7.53E-10
2	Spliceosome assembly	25.517	2.08E-08
3	Integrin family	25.273	2.47E-08
4	Sphingolipid signaling pathway	20.827	5.37E-07
5	Integrin signaling pathway	16.887	8.25E-06
6	ARF family signaling pathway	16.845	8.50E-06
7	Sphingoglycolipid degradation	15.776	1.78E-05
8	HDAC signaling pathway	14.763	3.60E-05
9	VHL signaling pathway	13.398	9.27E-05
10	2-Hydroxyglutarate signaling pathway	12.321	1.96E-04
11	Nicotinic Acetylcholine Receptor signaling pathway	11.877	2.66E-04

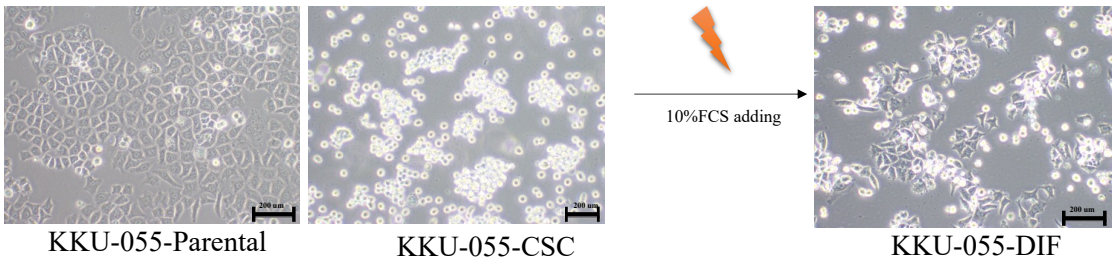

Figure S1. Cell morphologies of KKU-055-Parental cell, KKU-055-CSC and its differentiation form after the induction with 10% FCS for 3 days.

Supplementary data

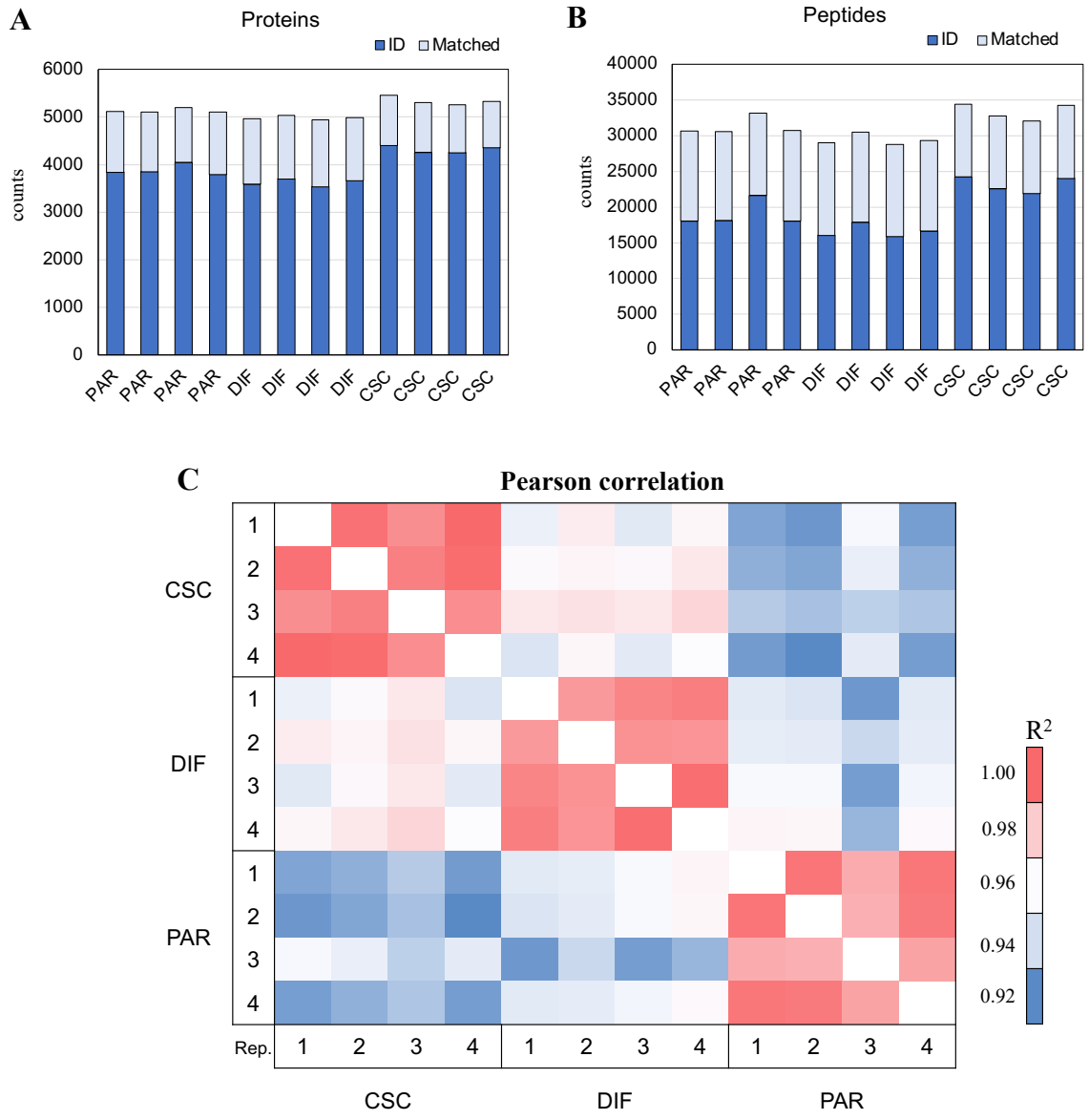

Figure S2. The numbers of **(A)** proteins and **(B)** peptides identified in the 4 separated biological samples from each of KKU-055-CSC, KKU-055-DIF, and KKU-055-Parental cells. Solid column indicates the number of identifications, white blue column indicates the identified numbers in the match between run. (**C**) Correlations of proteins LFQ intensities between the 4 biological replicates from each of KKU-055-CSC, KKU-055-DIF, and KKU-055-Parental cells. Pearson correlation showed the highly reproducible identification of each replicate in the same group. Y axis X axis are labelled as the 4 biological replicates (as a number of 1, 2, 3, 4) from each of KKU-055-CSC, KKU-055-DIF, and KKU-055-Parental.

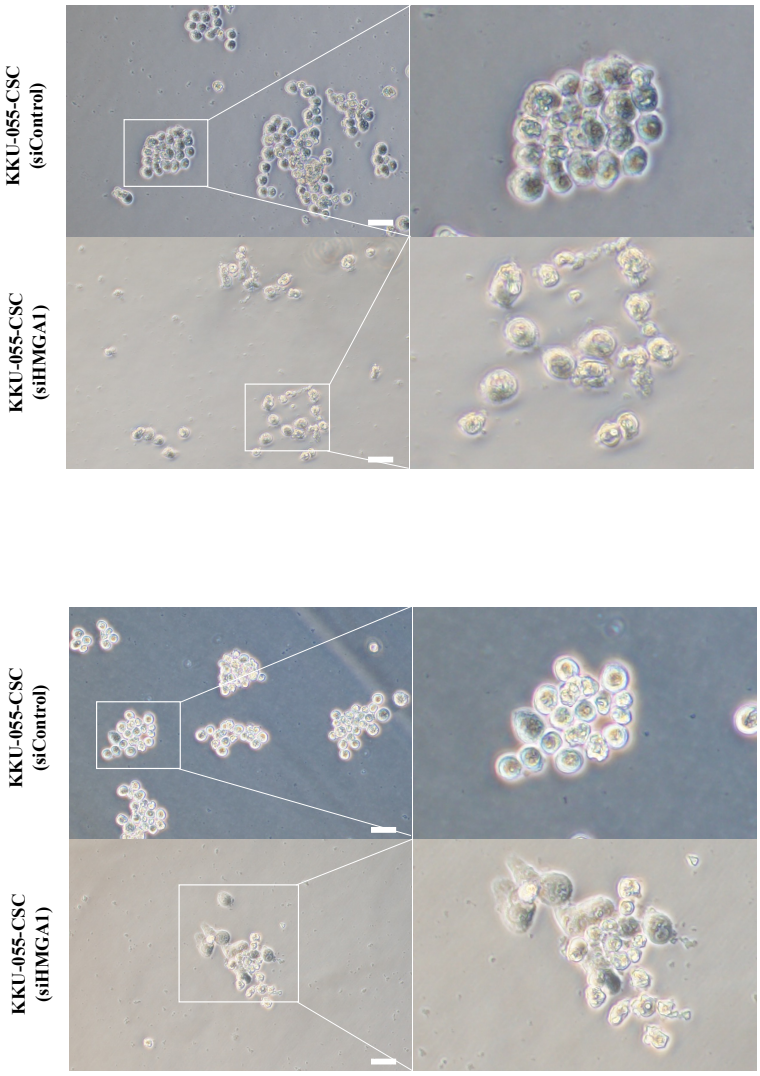

Figure S3. Cell morphologies of KKU-055-CSC treated with siControl and siHMGA1. The cell after siHMGA1 treatment induced the adherent to the cell culture dish.

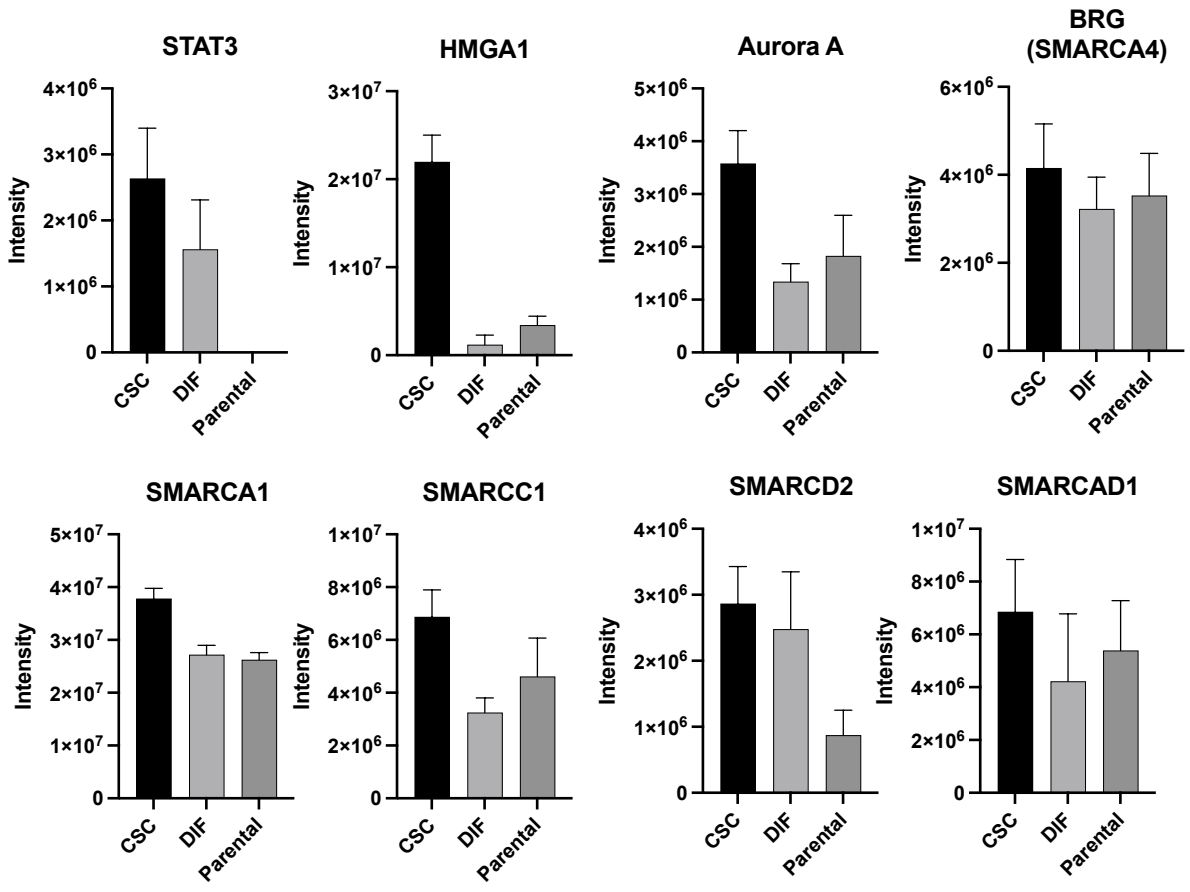

Figure S4. Proteomics identified the expression of STAT3, HMGA1, Aurora A, BRG and other SWI/SNF-related protein (noted) being overexpressed in KKU-055-CSC compared with KKU-055-DIF or -Par.

Noted:

SMARCA5:SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5

SMARCC1: SWI/SNF complex subunit SMARCC1

SMARCD2:SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 2

SMARCAD1:SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A containing DEAD/H box 1

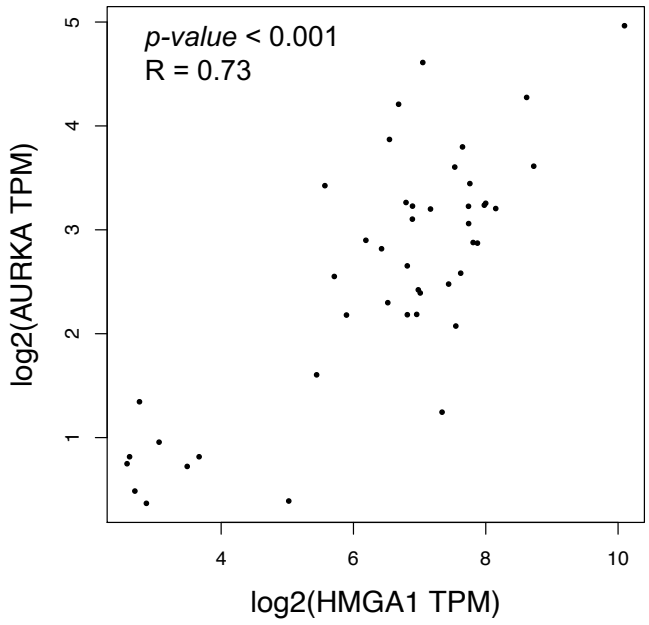

Figure S5. The positive correlation analysis between *HMGA1* and *AURKA* mRNA level in TCGA human cholangiocarcinoma with the highest Pearson index (R^2) of 0.73 (p-value 1.7 x 10^{-8})

Supplementary Table S1. Information of antibodies

Anti-body	Company	cat. Number	Dilution		
			WB	ICC	IHC
anti-SOX2	Cell signaling technology	3579S	1:1000	1:200	
anti-CD44	Cell signaling technology	3570S	1:1000	1:200	
anti-CD147	Abcam	ab666	1:5000		
anti-EpCAM	Abcam	ab71916	1:1000		
anti-HMGA1	Abcam	ab129153	1:1000	1:500	
anti-Aurora	Cell signaling technology	3092	1:1000	1:200	
anti-CD44v9	Cosmo bio	LKG-M001	1:1000	1:300	
Anti-Ki67	Abcam	ab16667	1:100		
anti-Actin	Sigma	A5441	10,000		
anti-mouse IgG -HRP	Thermo Scientific	62-6520	10,000		
anti-rabbit IgG-HRP	Cell signaling	7074F2	10,000		
anti-Rb Alexa 488	Invitrogen	A311034		1:300	
anti-Ms Alexa 568	Invitrogen	A11031		1:300	
anti-Rat Alexa 488	Invitrogen	A11006		1:300	
anti-Rb Alexa 568	Invitrogen	A11036		1:300	

Supplementary table 2. Information of PCR primer for real-time PCR

Primer	Gene	Forward	Reward	produce size (bp)	Annealing temperature (°C)
1	c-myc	ACCACCAGCAGCGACTCTGA	GGGCTGTGAGGAGGTTTGCT	143	60
2	Oct3/4	AGTGCCCGAAACCCACACTG	ACCACACTCGGACCACATCCT	81	60
3	SOX2	GAGGGGTGCAAAAGAGGAGA	CGTGAGTGTGGATGGGATTG	139	60
4	OLIG2	ACACAAATGGTAAACTCCTCCA	ACACGGCAGACGCTACAAA	98	60
5	CD44	TAAGGACACCCCAAATT CCA	ACTGCAATGCAAACTGCAAG	85	60
6	EpCAM	GTGCTGGTGTGTGAACACTG	CGCGTTGTGATCTCCTTCTG	130	60
7	β-actin	TCGTGCGTCACATTAAGGAC	GAAGGAAGGCTGGAAGAGTG	176	60