Liver Cancer

Manuscript:	LIC-2018-2-2/R1 RESUBMISSION
Title:	Heterogeneity of Epigenetic and Epithelial Mesenchymal Transition Marks in Hepatocellular Carcinoma with Keratin 19 Proficiency
Authors(s):	Naosuke Yokomichi (Co-author), Naoshi Nishida (Co-author), Yuzo Umeda (Co-author), Fumitaka Taniguchi (Co-author), Kazuya Yasui (Co-author), Toshiaki Toshima (Co-author), Yoshiko Mori (Co-author), Akihiro Nyuya (Co-author), Takeshiro Tanaka (Co-author), Takeshi Yamada (Co-author), Yoshiyuki Yamaguchi (Co-author), Takahito Yagi (Co-author), Toshiyoshi Fujiwara (Co-author), Ajay Goel (Co-author), Masatoshi Kudo (Co-author), Takeshi Nagasaka (Corresponding Author)
Keywords:	Basic Science, Biomarkers, Carcinogenesis, DNA Methylation, Hepatocellular Carcinoma, K19, KRT19, LINE-1
Type:	Original Paper

Kawasaki Medical School
Kawasaki Medical School Hospital

Takeshi Nagasaka, MD, PhD.
Associate Professor/Director
Department of Clinical Oncology
Department of Clinical Genetics

577 Matsushima, Kurashiki City
Okayama 701-0192, Japan
Phone: +81-86-462-1111
Fax: +81-86-464-1134
E-mail: takeshin@med.kawasaki-m.ac.jp

March 24, 2018

Prof. Masatoshi Kudo
Editor-in-Chief
Liver Cancer

Manuscript No. LIC-2018-2-2: Heterogeneity of Epigenetic and Epithelial Mesenchymal Transition Marks in Hepatocellular Carcinoma with Keratin 19 Proficiency

Dear Prof. Masatoshi Kudo:

My colleagues and I would like to thank you and the reviewers again for their insightful comments. We have addressed the reviewers concerns and have improved the manuscript according to the critique we received. We look forward to a positive response of our manuscript from the editorial staff based upon these responses.

All changes are ***highlighted in red*** in the revised manuscript **with changes marked.**

Since regards,

Takeshi Nagasaka

Kawasaki Medical School
Kawasaki Medical School Hospital

Takeshi Nagasaka, MD, PhD.
Associate Professor/Director
Department of Clinical Oncology
Department of Clinical Genetics

577 Matsushima, Kurashiki City
Okayama 701-0192, Japan
Phone: +81-86-462-1111
Fax: +81-86-464-1134
E-mail: takeshin@med.kawasaki-m.ac.jp

Reviewer #1 report:

This manuscript presents very important data for Keratin 19 (K19) in hepatocellular carcinoma (HCC) in vitro with various factors and in clinical course of HCC patients who underwent hepatic resection.
As mentioned by the authors, the present data suggest not only proficiency in HCC may be useful as a biomarker for identifying patients who have a poor prognosis with extrahepatic recurrence, but also that K19-proficient HCCs likely arise from hepatocytes or HCCs via epigenetic reprogramming, leading to EMT features. I have enjoyed reviewing your manuscript, which is a very food contribution to the elucidation for the heterogeneity underlying tumor development.

I recommend that this paper be accepted.

Minor point:

1. K19 is commonly expressed cholangiocarcinoma (CCC) and combined HCC and CCC. K19 Was there difference for extra-hepatic metastatic organs (e.g. lymph node) compared with reported that of typical HCC?

>> We thank for bringing up this point. As far as we examined, histological findings of extra-hepatic metastatic organs with H-E staining were similar to those of the original tumor.

2. I understand that K19 is a useful biomarker in clinical course of HCC. Although pathological findings are revealed as a typical HCC, K19 proficient HCC showed character differ from typical HCC. Was there no elevation of the levels of serum CEA or CA19-9 in K19 proficient HCC? If you have the data, please show them.

>> We thank for this comment. Unfortunately, we did not check serum CEA or CA19-9 routinely for our cohorts because our samples were evaluated retrospectively.

Kawasaki Medical School
Kawasaki Medical School Hospital

Takeshi Nagasaka, MD, PhD.
Associate Professor/Director
Department of Clinical Oncology
Department of Clinical Genetics

577 Matsushima, Kurashiki City
Okayama 701-0192, Japan
Phone: +81-86-462-1111
Fax: +81-86-464-1134
E-mail: takeshin@med.kawasaki-m.ac.jp

Reviewer #2 report:

Yokomichi et al examine the relationship between keratin 19 (K19) expression and KRT promoter/ long interspersed nucleotide element-1 (LINE-1) methylation status in human liver cancer cell lines and HCC tissues. They find lower KRT19 promoter methylation level and genome-wide LINE-1 hypermethylation in K19-proficient HCCs. In addition, K19 expression was related with an increase of EMT features and poor prognosis. This paper revealed the novel mechanism of K19 expression in HCC cells. The methods are sophisticated and the results are very interesting. I have some questions about pathological evaluation.

1. The reviewer is not familiar with the term 'K19-proficient' and 'K19-deficient' HCCs. I would guess most of HCCs usually do not express K19. Should we use 'K19-deficient' for such HCCs? Please explain why the authors used these terms, but not 'K19-positive' and 'K19-negative'.

>> We thank for this comment. As the reviewer mentioned, the terms such as 'K19-positive' and 'K19-negative' are common in the K19 studies. Simply 'K19-positive' and 'K19-negative' are easy to follow. However, several studies used 'deficient' in Keratin expression, for example, Hesse M et al (EMBO J. 2000 Oct 2; 19(19): 5060–5070.). So we used the term, 'K19-deficient', and next adapted the term 'K19-proficient' as an opposite word for 'K19-deficient'. The etymology of 'proficient' is that having or marked by an advanced degree of competence, as in an art, vocation, profession, or branch of learning. This contains complexity. As you know, K19-positive is varied 5%-100% in the stain. However, if the reviewer considered that 'K19-positive' and 'K19-negative' is the better for use in this study, we will change.

2. Page 6, immunohistochemistry analysis, lines 1-5 from the bottom. The authors define 0 when the percentage of the tumor cells with positive staining of K19, K7, NOTCH-1, and vimentin is 0-5%. However, in the next line, they considered as positive ≥ 05; (5 or more than 5) of tumor cells were stained. The 2 sentences are inconsistent. The same question is applied on HepPar-1, arginase-1 and E-cadherin. ≥50 (50 or more) is correct for the interpretation?

Kawasaki Medical School
Kawasaki Medical School Hospital

Takeshi Nagasaka, MD, PhD.
Associate Professor/Director
Department of Clinical Oncology
Department of Clinical Genetics

577 Matsushima, Kurashiki City
Okayama 701-0192, Japan
Phone: +81-86-462-1111
Fax: +81-86-464-1134
E-mail: takeshin@med.kawasaki-m.ac.jp

>> We agree with the reviewer's comment, and have corrected the number. Because this was a simple mistake in writing the manuscript, the results of this study did not change by this correction.

3. Page 6, immunohistochemistry analysis, lines 1-5 from the bottom. The reviewer wonder how the authors determined the ranges of scores 0-4. This should be clearly mentioned.

>> We agree with the reviewer's comment, and have clarified how we determined the ranges of scores 0-4. As we have added in the manuscript, we used these original ranges to obtain data. As many studies used the different ranges as below (yellow highlighted is considered to be positive);

Score	0	1	2	3	4	
This study	<5%	5 to 10%	11 to 20%	21 to 50%	51% or more	
K19	0	<1 to 24%	25to 49%	50 to 74%	75% or more	ref. 4
K19*	<1%	weak staining in ≥1%	moderate staining in ≥1%	strong staining in ≥1%		ref. 5
K19	<5%	5 to 24%	25 to 74%	75% or more		ref. 7
K19	<5%	5 to 50%	51% of more			ref. 8

* For cohort 2, K19 positivity was defined as membranous and/or cytoplasmic expression in > 5% of tumor cells with moderate or strong intensity

Therefore, to examine association among the markers, we uniformed number of score such as 0-4.

4. Page 6, immunohistochemistry analysis, lines 1-5 from the bottom. There is a sentence that "HepPar-1, arginase-1, and E-cadherin were considered as positive when ≥ 50% of tumor cells showed membranous stain". I would think HepPar-1 and arginase-1 positive-reaction is usually observed diffusely in the cytoplasm, but not as membranous. Please clarify this point.

Kawasaki Medical School
Kawasaki Medical School Hospital

Takeshi Nagasaka, MD, PhD.
Associate Professor/Director
Department of Clinical Oncology
Department of Clinical Genetics

577 Matsushima, Kurashiki City
Okayama 701-0192, Japan
Phone: +81-86-462-1111
Fax: +81-86-464-1134
E-mail: takeshin@med.kawasaki-m.ac.jp

>> We agree with the reviewer's comment, and we have corrected the words. We changed the sentence as below.

HepPar-1, arginase-1, and E-cadherin were considered as positive when ≥50% of tumor cells showed membranous stain.

>> HepPar-1, arginase-1, and E-cadherin were considered as positive when ≥51% of tumor cells were stained

5. Figure 3 is not convincing. In the figures of the score 1, 2, and 3, the number of positive cells looks smallest in 2, and the numbers of score 1 and 3 are not very different. I have a question about this scoring system. Is the expression scored according to the percentage of positive tumor cells/all tumor cells in HCC nodules? If so, I feel that showing the 5 pictures of different scores in Figure 3a is of little significance. I would think such figures are important when comparing the intensity of the positive expression. The authors should show more convincing immunostain pictures of low power view clearly showing difference of positive cell rate in the whole nodules or should just show a picture of representative positive staining. The same is true on Supplementary Figure S4. The pictures do not always match their scores, e.g., score 3 and 4 in E-cadherin.

>> We thank for the reviewer's comment. Following the reviewer's recommendation, we have added immunostain pictures of low power view in Figure 3 and Supplementary Figure S4B. We scored according to the percentage of positive tumor cells/all tumor cells in one slice.

Reviewer #3 report:

The manuscript is well written and the contents are informative.

"transarterial chemoembolization" should be changed to "transcatheter arterial chemoembolization".

>> We thank for this comment. We changed the word according to the reviewer's suggestion in P5 and P13, highlighted in red.

Heterogeneity of Epigenetic and Epithelial Mesenchymal Transition Marks in Hepatocellular Carcinoma with Keratin 19 Proficiency

Naosuke Yokomichi, M.D.,[a] Naoshi Nishida, M.D., PhD.,[b] Yuzo Umeda, M.D., PhD.,[a] Fumitaka Taniguchi, M.D.,[a] Kazuya Yasui, M.D.,[a] Toshiaki Toshima, M.D.,[a] Yoshiko Mori, M.D.,[a] Akihiro Nyuya, M.S.,[a, c] Takehiro Tanaka, M.D., PhD.,[d] Takeshi Yamada, M.D., PhD.,[e] Takahito Yagi, M.D., PhD.,[a] Toshiyoshi Fujiwara, M.D., PhD.[a] Yoshiyuki Yamaguchi, M.D.,PhD.[c] Ajay Goel, PhD.,[f] Masatoshi Kudo, M.D., Ph.D.,[b] and Takeshi Nagasaka, M.D., PhD.,[a, c*]

[a]Department of Gastroenterological Surgery, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama 700-0914 Japan

[b]Department of Gastroenterology and Hepatology, Kindai University Faculty of Medicine, Osaka 589-8511 Japan

[c]Department of Clinical Oncology, Kawasaki Medical School, Kurashiki 701-0192 Japan

[d]Department of Pathology, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama 700-0914 Japan

[e]Department of Gastrointestinal and Hepato-Biliary-Pancreatic Surgery, Nippon Medical School, Tokyo 113-8603 Japan

[f]Center for Gastrointestinal Cancer Research, Center for Epigenetics, Cancer Prevention and Cancer Genomics, Baylor Research Institute and Charles A Sammons Cancer Center, Baylor University Medical Center, Dallas, TX 75246, USA

Running title: Hepatocellular Carcinoma with Keratin 19 Proficiency

Corresponding Author:

Takeshi Nagasaka, M.D., Ph.D (takeshin@med.kawasaki-m.ac.jp), Department of Clinical Oncology, Kawasaki Medical School, Kurashiki City, Okayama, 701-0192 Japan

Phone: +81-86-462-1111; Fax: +81-86-464-1134

E-mail Address:

Naosuke Yokomichi: n-yokomichi@sis.seirei.or.jp

Naoshi Nishida: naoshi@med.kindai.ac.jp

Yuzo Umeda: y.umeda@d9.dion.ne.jp

Fumitaka Taniguchi: fmtktngch@gmail.com

Kazuya Yasui: kazu57790852@gmail.com

Toshiaki Toshima: tossy2020@yahoo.co.jp

Yoshiko Mori: yoshikomori1@gmail.com

Akihiro Nyuya: n.akihiro0406@gmail.com

Takehiro Tanaka: takehiro@md.okayama-u.ac.jp

Takeshi Yamada: y-tak@nms.ac.jp

Takahito Yagi: liver@md.okayama-u.ac.jp

Toshiyoshi Fujiwara: toshi_f@md.okayama-u.ac.jp

Yoshiyuki Yamaguchi: shogo@med.kawasaki-m.ac.jp

Ajay Goel: Ajay.Goel@bswhealth.org

Masatoshi Kudo: m-kudo@med.kindai.ac.jp

Abstract

Objective: Keratin 19 (K19) expression is a potential predictor for poor prognosis in patients with hepatocellular carcinoma (HCC). To clarify the feature of K19-proficient HCC, we traced epigenetic footprints in cultured cells and clinical materials.

Patients and Methods: *In vitro*, *KRT19* promoter methylation was analyzed and 5-aza-dC with trichostatin A (TSA) treatment was performed. Among 564 surgically resected HCCs, the clinicopathological relevance of K19-proficent HCCs was performed in comparison with hepatocytic (HepPar-1 and arginase-1), epithelial-mesenchymal transition (E-cadherin and vimentin), biliary differentiation-associated (K7 and NOTCH-1) markers, and epigenetic markers (*KRT19* promoter/long interspersed nucleotide element-1 [*LINE-1*] methylation status).

Results: *KRT19* promoter methylation was clearly associated with K19 deficiency and 5-aza-dC with trichostatin A treatment stimulated K19 re-expression, implicating DNA methylation as a potential epigenetic process for K19 expression. After excluding HCCs with recurrence, TNM stage as IIIB or greater, preoperative therapy, transplantation, and combined hepatocellular-cholangiocarcinoma, we assessed 125 from 564 HCC cases. In this cohort, K19 expression was found in 29 HCCs (23.2%), and corresponded with poor survival following surgery (P = 0.025) and extrahepatic recurrence free survival (P = 0.017). Compared with K19-deficient HCCs, lower *KRT19* promoter methylation level was observed in K19-proficient HCCs (P < 0.0001). Conversely, HCC with genome-wide *LINE-1* hypermethylation was frequently observed in K19-proficient HCCs (P = 0.0079). Additionally, K19 proficiency was associated with K7 proficiency (P = 0.043), and reduced E-cadherin and HepPar-1 expression (P = 0.043 and < 0.0001, respectively).

***Conclusions*:** K19-proficient HCC exhibited poor prognosis owing to extrahepatic recurrence, with molecular signatures differing from those in conventional cancer stem cells, providing novel insights of the heterogeneity underlying tumor development.

***Keywords*:** K19, KRT19, Hepatocellular carcinoma, Methylation, LINE-1,

Introduction

Hepatocellular carcinoma (HCC) constitutes the sixth most common neoplasm and the third leading cause of cancer deaths worldwide [1-3]. Surgical resection and liver transplantation comprise the only curative treatments for patients with early-stage HCC. However, the high rate of recurrence or metastases leads to worse prognosis in HCC after curative resection [2].

The positive expression of keratin 19 (K19), a marker for biliary or hepatic progenitor cells and early hepatoblasts, has been significantly associated with poor prognosis along with stemness-related and epithelial-mesenchymal transition (EMT) features in HCC [4-8]. In addition, K19 proficiency in patients with HCC has frequently been associated with vascular invasion, poorly differentiated tumors, and tumor recurrence after resection, radiofrequency ablation, or transplantation [9]. However, to date the cell origin of K19-proficient HCC has remained unclear [6]. Multiple studies suggested that HCC cells with K19 proficiency were originated from hepatic progenitor cells as such cells express progenitor cell markers, have invasive potential, and exhibit chemoresistance [6,9-11]. Conversely, others indicated that the expression of K19 in human HCCs may result from the dedifferentiation of malignant hepatocytes during continuous mutagenesis [12-14]. Furthermore, studies of liver regeneration have also found difficulties in identifying the cell origin of liver cancers including K19 proficient HCCs [15,16].

Although the mechanisms contributing to hepatocarcinogenesis remain unclear, it is widely accepted that HCC exhibits numerous genetic abnormalities, such as chromosomal alterations, gene amplifications, and mutations, as well as epigenetic alterations [17]. For example, increased DNA methylation levels of tumor suppressor genes correlate positively with HCC development and progression [18,19]; accordingly, a genome-wide methylation analysis has identified tumor suppressor genes in HCC [20]. However, although the gene encoding K19 (*KRT19*), located on chromosome 17 [21], contains a CpG island within its promoter region [22], to date there have been no reports regarding *KRT19* promoter methylation in any malignancies. Thus, although K19 expression regulation mechanisms in HCC have not been fully elucidated, the presence of a promoter CpG island suggests DNA methylation as a potential epigenetic process in this malignancy.

In the present study, we attempted to reveal the features of K19-proficient HCC by tracing epigenetic footprints in cultured cells and clinical materials. Firstly, we examined epigenetic alterations and underlying molecular mechanisms of K19-positive HCC cell lines. Next, from a panel of 564 surgically resected HCCs, we clarified the clinicopathological relevance of K19-proficent HCCs by analyzing robust methylation in the *KRT19* promoter region and *LINE-1* elements in comparison with other cholangiocytic (K7), hepatocytic (HepPar-1 and arginase-1), EMT (E-cadherin and vimentin), and biliary differentiation-associated (NOTCH-1) markers.

Patients and Methods

Patients

We retrospectively analyzed consecutive patients with initial HCC who received surgical resection at Okayama University Hospital from January 2000 to December 2010. The histopathological diagnosis of HCC was based on the World Health Organization criteria. The clinical history, pathological reports, and hematoxylin and eosin stained slides for all cases were reviewed to confirm the diagnosis. We excluded patients with a diagnosis of combined hepatocellular cholangiocarcinoma, with recurrent HCC, major vascular invasion, and with rupture or other organ invasion. We additionally excluded patients who received transplantation, non-curative resection, and preoperative therapy; e.g., transcatheter arterial chemoembolization. The tumor size was measured macroscopically after the removal of the tumor. Microvascular invasion was graded based on histopathological evaluation. The histological grade of tumor differentiation was determined according to the classification of Edmondson and Steiner (ES differentiation grade) [23]. Fibrosis of the non-neoplastic parenchyma was classified based on the Meta-analysis of Histologic Data in Viral Hepatitis (METAVIR) scoring system that assesses the degree of fibrosis ranging from F0 (no fibrosis) to F4 (cirrhosis) [24], Tumor grade was classified according to the 7th edition UICC/AJCC TNM staging system [25]. Survival time was determined from the date of surgical resection. Local or remote recurrence of disease after surgery was investigated by clinical assessment and regular abdominal ultrasonography or computed tomography. The study was approved by the institutional review

board of the Okayama University Hospital.

Cell lines

K19-proficient HCC cell lines: HepG2, HuH7, and PLC/PRF/5; K19-deficient HCC cell lines: HLE and HLF; and a colon cancer cell line HT29 as a K19 positive control were purchased from American Type Culture Collection (Manassas, VA) [26]. All cell lines were cultured using Dulbecco's modified Eagle's medium (Sigma Aldrich, St. Louis, MO) supplemented with 10% fetal bovine serum, penicillin (100 IU/mL), and streptomycin (100 μg/mL) at 37°C in a humidified incubator with 5% CO_2.

Immunohistochemistry (IHC) analysis

IHC was performed using formalin-fixed, paraffin-embedded tissue sections of surgically resected liver specimens. After deparaffinization and blocking of the endogenous peroxidase, antigen retrieval was performed by microwaving with citrate buffer (pH 6.0) or TRIS-ethylenediaminetetraacetic acid (pH 9.0). The sections were incubated with primary antibodies against K19 (1:200, mouse monoclonal, Cell Signaling Technology, Danvers, MA), keratin 7 (K7, 1:100, mouse monoclonal; Daco), HepPar-1 (1:100, mouse monoclonal; Daco, Glostrup, Denmark), arginase-1 (1:5000, rabbit polyclonal; Sigma-Aldrich), E-cadherin (1:1, mouse monoclonal; Daco), and vimentin (1:200, mouse monoclonal; Abcam, Cambridge, UK). Then, the sections were incubated with secondary antibodies conjugated to peroxidase-labeled polymer, using the EnVision system (Dako). Color development was performed using 3, 3-diaminobenzidine and the sections were counterstained with hematoxylin. Negative controls were carried out by substitution of the primary antibodies with non-immunized serum, resulted in no signal detection. IHC results were interpreted by pathologists blinded to the corresponding clinicopathological data. K19, K7, NOTCH-1, and cytoplasmic vimentin were scored according to the percentage of tumor cells with positive staining as follows: 0 = <5%, 1 = 5–10%, 2 = 11–20%, 3 = 21–50%, and 4 = 51–100%, and considered as positive when ≥5% of tumor cells were stained, as reported previously [4-8]. HepPar-1, arginase-1, and E-cadherin were considered as positive when ≥51% of tumor cells were stained [27], and were scored using the percentage of tumor cells with positive staining as follows: 0 = 0–10%, 1 = 11–20%, 2 = 21–50%, 3 = 51–80%, and 4 = 81–100%.

Western blotting

Protein was separated by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis and electroblotted onto Immun-Blot polyvinylidene fluoride membranes (BIO-RAD, Hercules, CA). After treating with 5% fat-free dried milk in 1x TBST for 1 hour at room temperature, the membranes were incubated with mouse anti-human CK19 (1:1000; Cell Signaling Technology) for 1 hour at room temperature, followed by secondary antibody for 1 hour at room temperature. Actin was used as an internal positive control. Target proteins were detected using enhanced chemiluminescence (ECL; BIO-RAD).

Bisulfite modification of DNA

Genomic DNA from the cell lines and formalin-fixed, paraffin-embedded specimens was extracted using a QIAamp DNA Mini Kit (Qiagen, Valencia, CA) and TaKaRa DEXPAT Kit (TaKaRa Bio Inc., Otsu, Japan), respectively. Approximately 1 μg DNA was subsequently modified using the EZ DNA Methylation Kit (Zymo Research, Orange, CA).

Bisulfite sequencing

Polymerase chain reaction products from the *KRT19* promoter were amplified by a set of primers (**Supplementary Table S1**) for bisulfite DNA cloning and sequencing. The *KRT19* promoter was cloned from bisulfite-treated DNA into the pCR2.1TOPO vector using the TOPO-TA cloning system (Life Technologies, Carlsbad, CA), followed by automated DNA sequencing with both the forward (F) and reverse (R) primers using an ABI 310-Avant NA sequencer (Applied Biosystems, Foster City, CA).

Methylation analysis of KRT19 and LINE-1

We performed quantitative methylation analysis for the promoter CpG islands of *KRT19* and *LINE-1* in tumors, their matched corresponding non-tumor liver tissues, and cell lines. For *KRT19* analysis, we used the high-sensitive assay for bisulfite DNA (Hi-SA), a modified combined bisulfite restriction analysis method, by which fluorescence labeled DNA fragments are detected using a genetic analyzer.[28] The primer sequences are summarized in **Supplementary Table S1**, whereas the details for the remaining assays have been described previously [28,29]. We analyzed two regions in the *KRT19* promoter; termed region 1 and region 2. Region 2 contained two sites that can be restricted by *Hha*I. The methylation ratio was

calculated by the ratio of the peak of methylated and unmethylated bands (examples are shown in **Supplementary Figure S1**). We used primers for *LINE-1* methylation analysis as described previously [30] with assay modification by adding a fluorescent dye to measure methylation ratio using an ABI 310-Avant NA sequencer (**Supplementary Table S1**). Additionally, we analyzed association between methylation and expression for KRT19 of HCCs by the data obtained from cBioportal for cancer genomics (http://www.cbioportal.org/).

Demethylation analysis

HLF, HLE, PLC/PRF/5, HepG2, and HuH7 cells were cultured with 5-aza-2'-deoxycitidine (5-Aza-dC, Wako, Japan). 5-Aza-dC was dissolved with phosphate buffered saline and diluted with medium. The concentrations of 5-Aza-dC were set as 1μM according to the nadir of promoter methylation rate in both region 1 and 2 of *KRT19* (**Supplementary Figure S2**). After 5-Aza-dC treatment from day1 to day3, cells were treated by 100μM trichostatin A (TSA, MERCK, Darmstadt, Germany), a potent histone deacetylase (HDAC) inhibitor, for 24h. DNA and RNA were extracted after TSA treatment for determining methylation status and expression status by a microarray analysis (SurePrint G3 Human Gene Expression 8X60K v2, Agilent, Santa Clara, CA).

Statistical analysis

All statistical analyses were performed using JMP software (version 10.0; SAS Institute, Inc, Cary, NC). We compared K19 expression status with various clinicopathological features and the results of other immunostaining using the Fisher's exact test. Methylation levels in regions 1 and 2 of the *KRT19* promoter were analyzed as both continuous and categorical variables (positive, methylation level in both region 1 and 2 > 10%). Methylation levels in *LINE-1* were categorized according to the mean value (≥ 55% methylation in *LINE-1* defined as hypermethylation and < 55% defined as hypomethylation). Categorical variables were compared using Fisher's exact test. Differences between continuous variables were determined using the analysis of variance test (ANOVA). Correlation coefficient between continuous variables was nonparametrically determined (Spearman's correlation coefficient [ρ]). Overall survival (OS) was calculated from the date of surgical resection to the date of death owing to HCC or last follow-up for

censored patients. Recurrence free survival was calculated from the date of surgical resection to the date of the first documentation of local, regional, or distant relapse, appearance of a second primary lesion by computed tomography, and/or magnetic resonance imaging routinely performed every 6 months. Extrahepatic-recurrence free survival was calculated from the date of surgical resection to the date of the first documentation of appearance of extrahepatic recurrences by computed tomography and/or magnetic resonance imaging routinely performed every 6 months. OS, recurrence free survival, and extrahepatic-recurrence free survival were estimated using the Kaplan-Meier method. Then, a multivariate analysis for OS was performed using a Cox-proportional hazards model. All *P*-values reported were calculated using two-sided tests and values <0.05 were considered statistically significant.

Results

Association of KRT19 promoter methylation with K19 expression in cultured cell lines

To examine methylation status in the *KRT19* promoter, which contains dense CpG sites, the promoter region of *KRT19* was divided into two regions (region 1 and region 2, **Figure 1a**). Methylation status in the discrete regions was analyzed as a continual variable. Cloning and bisulfite sequencing were used to precisely confirm this methylation status, providing validation and further evidence that the K19-deficient HLF cells showed dense methylation throughout the *KRT19* promoter CpGs whereas the K19-proficient HuH7 cells exhibited no methylation (**Figure 1b**). Both regions in the *KRT19* promoter were hypermethylated in K19-deficient cell lines (HLE and HLF) but were hypomethylated in K19-proficient cell clines (HepG2 and HuH7, **Figure 1c**). The mean methylation ratio of K19-deficient and K19-proficient cells in region 1 was 31.7% (95% confidence interval [95%CI]; 11.6–51.6%) and 6.7% (−6.0–19.4%), respectively ($P = 0.0426$) whereas those of K19-deficient and K19-proficient cells in region 2 were 96.8% (68.2–125.4%) and 6.7% (−4.9–31.2%), respectively ($P = 0.0014$). Thus, methylation levels in both regions of *KRAT19* were significantly inversely associated with K19 expression. We next treated five HCC cell lines (HepG2, HuH7, HLEE, HLF, and PLC/PRF/5) with a demethylating agent, 5'-Aza-dC, and a potent HDAC inhibitor, TSA. Among them, the K19-deficient HLF and HLE cells with higher methylation rate in the *KRT19* promoter showed that, although 5'-Aza-dC alone treatment caused demethylation in the *KRT19* promoter but failed to recover *KRT19* expression, 5'-Aza-dC following TSA treatment successfully recovered

expression level of *KRT19* by the microarray analysis, indicating that K19 expression would be regulated by promoter methylation and histone modification (**Figure 1d**).

We also examined genome global methylation level using *LINE-1* retrotransposons, which constitute a substantial portion (approximately 17%) of the human genome and are regarded as a surrogate marker of global DNA methylation. Although the mean methylation ratio of *LINE-1* of K19-proficient cells (36.5% [95%CI; 2.5–60.5%]) was higher than that of K19-deficient cells (26.8% [−11.2–64.8%]), the difference was not statistically significant (P = 0.6, **Figure 1b**).

Clinicopathological features in patients with K19-proficient HCC

To assess the precise clinical landscapes of HCC with K19 proficiency, we retrospectively examined patients with HCC who underwent surgical resection. Among 564 patients recruited, 125 met the inclusion and exclusion criteria of this study (**Figure 2**). The clinicopathological features of 125 patients are summarized in **Supplementary Table S2.** Of these, 113 patients (90.4%) were categorized in TNM stage I or II and no patients exhibited major vascular invasion. IHC analyses revealed that K19 expression was detected in 29 patients (23.2%) and that patients with K19-proficient HCC were significantly younger than those with K19-deficient HCC (P = 0.020, **Figure 3a** and **Table 1**). Although a high level of serum alpha-fetoprotein and microvascular invasion were more frequently observed in patients with K19-proficient HCC (P = 0.021 and 0.019, respectively), K19 proficiency had no association with TNM stage, tumor size, tumor number, or differentiation. Patients with K19-proficient HCC showed poorer survival after surgery (P = 0.025) and poorer extra-hepatic metastasis-free survival (P = 0.017, **Figure 3b-d**). Multivariate analysis demonstrated that K19 acted as an independent prognostic factor for survival after surgery (**Table 2**).

K19 expression and promoter methylation status in the KRT19 gene

Next, we investigated the methylation levels of discrete regions in the *KRT19* promoter in the 125 HCCs. Initially, methylation levels in the discrete regions obtained from fluorescent Hi-SA were analyzed as a continuous variable. In region 1, the mean methylation level was 2.3% [95%CI; 1.2–3.5%] among HCC tissues with K19 proficiency but 8.7% (95%CI; 5.4–11.9%) among HCC tissues with K19 deficiency (P =

0.0315, ANOVA; **Figure 4a**). In region 2, the mean methylation levels were 2.7% [95%CI; 0.9–4.5%] among HCC tissues with K19 proficiency and 15.7% (95%CI; 9.6–21.8%) among HCC tissues with K19 deficiency (*P* = 0.0228, ANOVA; **Figure 4b**). To define the threshold of methylation levels in region 1 and region 2 of the *KRT19* promoter, we examined the methylation levels in the discrete regions of adjacent normal liver tissues. The mean methylation level in 123 adjacent normal liver tissues was 4.3% [95%CI; 2.5–6.1%] in region 1 and 8.3% in region 2. Therefore, we defined *KRT19* methylation in both regions at 10% or more as a continuous variable [i.e., ≥10% methylation as methylation-positive (methylated) and <10% methylation as methylation-negative (unmethylated)]. Using this criterion, we found that 20 (16.0%) and 28 (22.4%) of 125 HCCs could be categorized as methylated in region 1 and region 2 of the *KRT19* promoter, respectively. All of the 20 HCCs (100%) categorized as methylated in region 1 showed K19 deficiency whereas all of the 29 K19-proficient HCCs (100%) were unmethylated in region 1 (*P* = 0.0038, **Figure 4a**. With respect region 2, 25 of 28 HCCs (89.3%) with region 2 methylation showed K19 deficiency whereas 26 of 29 HCCs (89.7%) with K19-proficient HCCs were unmethylated (*P* = 0.12, **Figure 4b**). We also examined association between methylation and expression for KRT19 of 442 HCCs by the data obtained from cBioportal for cancer genomics (**Supplementary Figure S3**). Although K19-deficient HCCs had various methylation levels in KRT19, most of K19-profcient HCCs showed the lower methylation levels. Thus, HCCs with K19 proficiency demonstrated unmethylation in the discrete promoter regions of the *KRT19* gene. Together with the results of our *in vitro* study, these findings suggest that the mechanism of K19 expression may be partially regulated by promoter methylation and histone modification in the *KRT19* gene.

***Association between LINE-1* methylation and *KRT19* promoter methylation in K19-proficient HCCs**

Evaluation of the *LINE-1* methylation levels in the 119 HCC tissues demonstrated that the mean methylation level was 54.8% (95%CI; 50.7–59.0%, **Figure 4c**). When we categorized *LINE-1* methylation levels in the two groups according to the threshold of 55% (≥ 55% methylation defined as hypermethylation and < 55% defined as hypomethylation), hypermethylation in *LINE-1* was frequently observed in K19-proficient (19 of 26 [73.1%]) compared with K19-deficient HCCs (40 of 93 [43.0%], *P* = 0.0079) (**Figure 4c**).

We then asked whether an association exists between genome-wide methylation level and the local regions in the *KRT19* promoter. For this, we performed multivariate correlation by comparing the methylation levels of *LINE-1* and regions 1 and 2 in the *KRT19* promoter from K19- proficient and deficient HCCs, respectively (**Figure 4d, e**). Regardless of K19 expression status, positive correlations were observed in the methylation level between regions 1 and 2 in the *KRT19* promoter (Spearman's correlation coefficients [ρ] = 0.2941, *P* = 0.0043 in K19-deficient HCCs; ρ = 0.3455, *P* = 0.0644 in K19-proficient HCCs). In contrast, although there was no association between *LINE-1* methylation and *KRT19* promoter methylation levels in K19-deficient HCCs, an inverse correlation was observed in K19-proficient HCCs (ρ = −0.3353, *P* = 0.094 between *LINE-1* and region 1; ρ = −0.4424, *P* = 0.0236 between *LINE-1* and region 2).

Markers associated with expression profiles of EMT, hepatocytic and biliary differentiation.

We next clarified features of HCC in relation to K19 expression status in comparison with EMT markers (E-cadherin and vimentin), hepatocytic markers (HepPar-1 and Arginase-1), and markers associated with biliary differentiation (K7 and NOTCH-1). Examples of scoring of IHC staining are shown in **Supplementary Figure S4**. K19-proficient HCCs demonstrated increased EMT features with loss of E-cadherin (*P* = 0.043) and gain of vimentin expression (*P* = 0.084, **Table 3**). With respect to organ signature, findings of hepatocytic or cholangiocytic markers contrasted; although K19-proficient HCCs exhibited decreased number of cases with positive expression of the hepatocytic markers HepPar-1 and Arginase-1, they showed at least one positive result of staining for HepPar-1 or Arginase-1. Thus, K19 expression correlated with an increase of EMT features accompanying K7 positive expression and a reduction of hepatocytic features. By multivariate correlations (**Figure 5**), in K19-deficient HCCs, hepatocytic markers showed inverse associations with vimentin and NOTCH-1. Conversely, for HCCs with K19-proficiency, although expression of E-cadherin was decreased and that of K7 was increased, the positive association of HepPar-1 and arginase-1 expression [ρ = 0.5981, *P* = 0.0006]) indicated that hepatocytic features were conserved.

Discussion

Herein, we have shown the biological significance of HCC with K19 proficiency with relation to methylation status in the promoter region of *KRT19*, genome-wide methylation level of *LINE-1*, and EMT features. Results from clinical samples as well as our *in vitro* study indicated that K19-deficient HCCs demonstrated higher methylation level in the promoter region of *KRT19* compared with K19-proficient HCCs, suggesting that K19 expression might be regulated by the density of its promoter methylation. Our demethylation analysis demonstrated re-expression of *KRT19* required not only application of a demethylating agent, 5'-Aza-dC, but also a HDAC inhibitor, TSA, in K19-deficient HLF and HLE cells, suggesting that epigenetic mechanisms including histone modification play an important role in K19 expression.

Consistent with previous studies [4-8], we confirmed the utility of K19 as a prognostic biomarker for HCC. Notably, we extracted 125 HCCs from 564 surgically resected HCCs according to exclusion criteria that excluded HCCs with major vascular invasion pathologically, which is the strongest prognostic factor for this malignancy. We also excluded recurrent HCCs for accurate survival analysis, and HCCs with preoperative transcatheter arterial chemoembolization because such treatment might influence the tumor characteristics including K19 expression on the membrane of the cancer cells. Thus, the 125 HCCs analyzed in this study were, as far as possible, homogeneous in both clinical and pathological settings. Of this cohort, K19 proficiency was observed in 23.2% (29/125) HCCs and was more frequently observed in younger, female patients. Other clinical features of K19-proficient HCCs included higher serum alpha-fetoprotein levels, positive for pathological microvascular invasion, and poor survival after surgery.

In this study, K19-proficient HCCs exhibited EMT features with loss of E-cadherin and gain of vimentin expression. This association between EMT features and K19 proficiency was consistent with a prior study reporting that K19-proficient HCC exhibited significantly increased EMT-related protein and mRNA expression [5]. EMT has been shown to be a pivotal mechanism contributing to cancer invasion and metastasis including HCC [31,32]. Consistent with this hypothesis, patients with K19-proficient HCC in this and other studies [7,33] showed poorer survival and more frequently exhibited extra-hepatic metastasis.

Generally, primary liver cancers are classified into the following subtypes; HCCs, cholangiocarcinomas (CCAs), combined HCC-CCAs, hepatoblastomas, and fibrolamellar hepatocellular carcinomas [34]. K19 is commonly expressed in two types of liver cancers, CCAs and combined HCC-CCAs; however, their cell origins and means of development are not yet sufficiently understood. Recent studies using a mouse model of hepatocyte fate tracing have revealed that CCAs may be originated from fully differentiated hepatocytes via NOTCH signaling activation [35,36]. Combined HCC-CCAs are pathologically diagnosed based on the classical type of combined HCC-CCAs as areas of typical HCC and CCA mixed within the tumor, with the latest edition of this classification proposing a subtype with stem cell features [37].

Kawai et al. reported that K19-proficient HCCs possessed cancer stem cell features, such as EMT features and the activation of the TGFb/Smad signal cascade [38]. Zhang et al. reported that HCCs expressed a stem cell marker CD133, considered to be a marker for cancer stem cells of HCC, and showed hypomethylation of the global DNA methylation marker *LINE-1* [39]. Therefore, we next examined the association between K19-proficent HCCs and *LINE-1* methylation levels associated with cancer stem cell features. Notably, in our cohort, K19-proficent HCCs exhibited *LINE-1* hypermethylation that was significantly correlated with demethylation in the *KRT19* promoter. This evidence obtained from clinical samples was also supported by our *in vitro* study that K19-proficient cell lines were more likely show increased *LINE-1* methylation level (**Figure 1b**). As Kim et al. demonstrated that only 1.5% of HCCs showed both CD133 and K19 expression [5], it is reasonable to presume that the K19-proficient HCCs were not equivalent to the CD133-proficient HCCs that are associated with *LINE-1* hypomethylation. With respect to IHC staining, our cohort of HCCs showed inverse correlation in expression status between NOTCH-1 and hepatocytic markers irrespective of K19 expression status. Although K19-proficient HCCs demonstrated increased EMT and cholangiocytic features and reduced number of cases with positive expression of hepatocytic markers, strongly conserved hepatocytic features were still observed.

This study has some limitations. For example, analyzed samples were obtained from a retrospective cohort in a single hospital. However, this study describes the novel features observed in HCCs with K19 proficiency; increased EMT features and decreased mature signatures of hepatocyte cells; increased genome-wide DNA methylation levels and reduced promoter methylation density in the *KRT19* gene.

Additionally, *in vitro* analysis revealed that expression of K19 was regulated under promoter methylation and histone modification. Our clinical data suggest that not only detection of K19 proficiency in HCC may be useful as a biomarker for identifying patients who have a poor prognosis with extrahepatic recurrence, but also that K19-proficient HCCs likely arise from hepatocytes or HCCs via epigenetic reprogramming, leading to EMT features. Thus, our findings provide novel insights regarding the heterogeneity underlying tumor development.

Declarations

Acknowledgements

The authors would like to thanks Mr. Toru Nakai and Mrs. Tae Yamanishi for the immunohistochemistry techniques.

Funding

This work was supported by grants from MEXT/JSPS KAKENHI (20590572, 25860409, 26462016, and 15H03034 to TN).

Availability of data and materials

The datasets obtained and/or analyzed during the current study are available from the corresponding author in reasonable request.

Authors' contributions

NY performed experiments and analysed data supported by YU, FT, KY, TT, YM, and AY; NN and MK performed all microarray analysis; TT and TakeY performed all immunohistochemistry analysis. TakaY and TF prepared all samples and summrized clinical information; AL and SL recruited patients for analysis; YY and AG analysed data and wrote the paper; TN obtained funding, designed the work, analysed data, and wrote the paper.

Ethics approval and consent to participate

The human tissues used in this study were approved by the institute ethical committee of Okayama University.

Consent for publication

Not applicable.

Competing interests

The authors have no conflicts of interest that pertain to this work.

References

1 Boyle P, Levin B, International Agency for Research on Cancer., World Health Organization.: World cancer report 2008. Lyon
Geneva, International Agency for Research on Cancer ;
Distributed by WHO Press, 2008.

2 Forner A, Llovet JM, Bruix J: Hepatocellular carcinoma. Lancet 2012;379:1245-1255.

3 Jemal A, Center MM, DeSantis C, Ward EM: Global patterns of cancer incidence and mortality rates and trends. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2010;19:1893-1907.

4 Wu PC, Fang JW, Lau VK, Lai CL, Lo CK, Lau JY: Classification of hepatocellular carcinoma according to hepatocellular and biliary differentiation markers. Clinical and biological implications. The American journal of pathology 1996;149:1167-1175.

5 Kim H, Choi GH, Na DC, Ahn EY, Kim GI, Lee JE, Cho JY, Yoo JE, Choi JS, Park YN: Human hepatocellular carcinomas with "Stemness"-related marker expression: keratin 19 expression and a poor prognosis. Hepatology 2011;54:1707-1717.

6 Govaere O, Komuta M, Berkers J, Spee B, Janssen C, de Luca F, Katoonizadeh A, Wouters J, van Kempen LC, Durnez A, Verslype C, De Kock J, Rogiers V, van Grunsven LA, Topal B, Pirenne J, Vankelecom H, Nevens F, van den Oord J, Pinzani M, Roskams T: Keratin 19: a key role player in the invasion of human hepatocellular carcinomas. Gut 2014;63:674-685.

7 Uenishi T, Kubo S, Yamamoto T, Shuto T, Ogawa M, Tanaka H, Tanaka S, Kaneda K, Hirohashi K: Cytokeratin 19 expression in hepatocellular carcinoma predicts early postoperative recurrence. Cancer science 2003;94:851-857.

8 Durnez A, Verslype C, Nevens F, Fevery J, Aerts R, Pirenne J, Lesaffre E, Libbrecht L, Desmet V, Roskams T: The clinicopathological and prognostic relevance of cytokeratin 7 and 19 expression in hepatocellular carcinoma. A possible progenitor cell origin. Histopathology 2006;49:138-151.

9 Villanueva A, Hoshida Y, Battiston C, Tovar V, Sia D, Alsinet C, Cornella H, Liberzon A, Kobayashi M, Kumada H, Thung SN, Bruix J, Newell P, April C, Fan JB, Roayaie S, Mazzaferro V, Schwartz ME, Llovet JM: Combining clinical, pathology, and gene expression data to predict recurrence of hepatocellular carcinoma. Gastroenterology 2011;140:1501-1512 e1502.

10 Villanueva A, Hoshida Y, Toffanin S, Lachenmayer A, Alsinet C, Savic R, Cornella H, Llovet JM: New strategies in hepatocellular carcinoma: genomic prognostic markers. Clinical cancer research : an official journal of the American Association for Cancer Research 2010;16:4688-4694.

11 Andersen JB, Loi R, Perra A, Factor VM, Ledda-Columbano GM, Columbano A, Thorgeirsson SS: Progenitor-derived hepatocellular carcinoma model in the rat. Hepatology 2010;51:1401-1409.

12 Santos NP, Oliveira PA, Arantes-Rodrigues R, Faustino-Rocha AI, Colaco A, Lopes C, Gil da Costa RM: Cytokeratin 7/19 expression in N-diethylnitrosamine-induced mouse hepatocellular lesions: implications for histogenesis. International journal of experimental pathology 2014;95:191-198.

13 Ezzoukhry Z, Louandre C, Trecherel E, Godin C, Chauffert B, Dupont S, Diouf M, Barbare JC, Maziere JC, Galmiche A: EGFR activation is a potential determinant of primary resistance of hepatocellular carcinoma cells to sorafenib. International journal of cancer Journal international du cancer 2012;131:2961-2969.

14 Yoneda N, Sato Y, Kitao A, Ikeda H, Sawada-Kitamura S, Miyakoshi M, Harada K, Sasaki M, Matsui O, Nakanuma Y: Epidermal growth factor induces cytokeratin 19 expression accompanied by increased growth abilities in human hepatocellular carcinoma. Laboratory investigation; a journal of technical methods and pathology 2011;91:262-272.

15 Itoh T, Miyajima A: Liver regeneration by stem/progenitor cells. Hepatology 2014;59:1617-1626.

16 Yanger K, Zong Y, Maggs LR, Shapira SN, Maddipati R, Aiello NM, Thung SN, Wells RG, Greenbaum LE, Stanger BZ: Robust cellular reprogramming occurs spontaneously during liver regeneration. Genes & development 2013;27:719-724.

17 Nishida N, Goel A: Genetic and epigenetic signatures in human hepatocellular carcinoma: a systematic review. Current genomics 2011;12:130-137.

18 Nishida N, Nagasaka T, Nishimura T, Ikai I, Boland CR, Goel A: Aberrant methylation of multiple tumor suppressor genes in aging liver, chronic hepatitis, and hepatocellular carcinoma. Hepatology 2008;47:908-918.

19 Nishida N, Kudo M, Nagasaka T, Ikai I, Goel A: Characteristic patterns of altered DNA methylation predict emergence of human hepatocellular carcinoma. Hepatology 2012;56:994-1003.

20 Revill K, Wang T, Lachenmayer A, Kojima K, Harrington A, Li J, Hoshida Y, Llovet JM, Powers S: Genome-wide methylation analysis and epigenetic unmasking identify tumor suppressor genes in hepatocellular carcinoma. Gastroenterology 2013;145:1424-1435 e1421-1425.

21 Schweizer J, Bowden PE, Coulombe PA, Langbein L, Lane EB, Magin TM, Maltais L, Omary MB, Parry DA, Rogers MA, Wright MW: New consensus nomenclature for mammalian keratins. The Journal of cell biology 2006;174:169-174.

22 Zody MC, Garber M, Adams DJ, Sharpe T, Harrow J, Lupski JR, Nicholson C, Searle SM, Wilming L, Young SK, Abouelleil A, Allen NR, Bi W, Bloom T, Borowsky ML, Bugalter BE, Butler J, Chang JL, Chen CK, Cook A, Corum B, Cuomo CA, de Jong PJ, DeCaprio D, Dewar K, FitzGerald M, Gilbert J, Gibson R, Gnerre S, Goldstein S, Grafham DV, Grocock R, Hafez N, Hagopian DS, Hart E, Norman CH, Humphray S, Jaffe DB, Jones M, Kamal M, Khodiyar VK, LaButti K, Laird G, Lehoczky J, Liu X, Lokyitsang T, Loveland J, Lui A, Macdonald P, Major JE, Matthews L, Mauceli E, McCarroll SA, Mihalev AH, Mudge J, Nguyen C, Nicol R, O'Leary SB, Osoegawa K, Schwartz DC, Shaw-Smith C, Stankiewicz P, Steward C, Swarbreck D, Venkataraman V, Whittaker CA, Yang X, Zimmer AR, Bradley A, Hubbard T, Birren BW, Rogers J, Lander ES, Nusbaum C: DNA sequence of human chromosome 17 and analysis of rearrangement in the human lineage. Nature 2006;440:1045-1049.

23 Edmondson HA, Steiner PE: Primary carcinoma of the liver: a study of 100 cases among 48,900 necropsies. Cancer 1954;7:462-503.

24 Ichida F, Tsuji T, Omata M, Ichida T, Inoue K, Kamimura T, Yamada G, Hino K, Yokosuka O,

Suzuki H: New Inuyama classification; new criteria for histological assessment of chronic hepatitis. International Hepatology Communications 1996;6:112-119.

25 Edge S, Byrd D, Compton C, Fritz A, Greene F, Trotti A: AJCC cancer staging manual, ed 7th. New York, Springer, 2010.

26 Wu F, Nishioka M, Fujita J, Murota M, Ohtsuki Y, Ishida T, Kuriyama S: Expression of cytokeratin 19 in human hepatocellular carcinoma cell lines. International journal of oncology 2002;20:31-37.

27 Yan BC, Gong C, Song J, Krausz T, Tretiakova M, Hyjek E, Al-Ahmadie H, Alves V, Xiao SY, Anders RA, Hart JA: Arginase-1: a new immunohistochemical marker of hepatocytes and hepatocellular neoplasms. The American journal of surgical pathology 2010;34:1147-1154.

28 Nagasaka T, Tanaka N, Cullings HM, Sun DS, Sasamoto H, Uchida T, Koi M, Nishida N, Naomoto Y, Boland CR, Matsubara N, Goel A: Analysis of fecal DNA methylation to detect gastrointestinal neoplasia. Journal of the National Cancer Institute 2009;101:1244-1258.

29 Yoshida K, Nagasaka T, Umeda Y, Tanaka T, Kimura K, Taniguchi F, Fuji T, Shigeyasu K, Mori Y, Yanai H, Yagi T, Goel A, Fujiwara T: Expansion of epigenetic alterations in EFEMP1 promoter predicts malignant formation in pancreatobiliary intraductal papillary mucinous neoplasms. Journal of cancer research and clinical oncology 2016;142:1557-1569.

30 Lertkhachonsuk R, Paiwattananupant K, Tantbirojn P, Rattanatanyong P, Mutirangura A: LINE-1 Methylation Patterns as a Predictor of Postmolar Gestational Trophoblastic Neoplasia. BioMed research international 2015;2015:421747.

31 Moody SE, Perez D, Pan TC, Sarkisian CJ, Portocarrero CP, Sterner CJ, Notorfrancesco KL, Cardiff RD, Chodosh LA: The transcriptional repressor Snail promotes mammary tumor recurrence. Cancer cell 2005;8:197-209.

32 Ma CQ, Yang Y, Wang JM, Du GS, Shen Q, Liu Y, Zhang J, Hu JL, Zhu P, Qi WP, Qian YW, Fu Y: The aPKCiota blocking agent ATM negatively regulates EMT and invasion of hepatocellular carcinoma. Cell death & disease 2014;5:e1129.

33 Zhuang PY, Zhang JB, Zhu XD, Zhang W, Wu WZ, Tan YS, Hou J, Tang ZY, Qin LX, Sun HC: Two pathologic types of hepatocellular carcinoma with lymph node metastasis with distinct prognosis on the basis of CK19 expression in tumor. Cancer 2008;112:2740-2748.

34 Oikawa T: Cancer Stem cells and their cellular origins in primary liver and biliary tract cancers. Hepatology 2016;64:645-651.

35 Sekiya S, Suzuki A: Intrahepatic cholangiocarcinoma can arise from Notch-mediated conversion of hepatocytes. The Journal of clinical investigation 2012;122:3914-3918.

36 Fan B, Malato Y, Calvisi DF, Naqvi S, Razumilava N, Ribback S, Gores GJ, Dombrowski F, Evert M, Chen X, Willenbring H: Cholangiocarcinomas can originate from hepatocytes in mice. The Journal of clinical investigation 2012;122:2911-2915.

37 Bosman FT, World Health Organization., International Agency for Research on Cancer.: WHO classification of tumours of the digestive system, ed 4th. Lyon, International Agency for Research on Cancer,

2010.

38 Kawai T, Yasuchika K, Ishii T, Katayama H, Yoshitoshi EY, Ogiso S, Kita S, Yasuda K, Fukumitsu K, Mizumoto M, Hatano E, Uemoto S: Keratin 19, a Cancer Stem Cell Marker in Human Hepatocellular Carcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research 2015;21:3081-3091.

39 Zhang C, Xu Y, Zhao J, Fan L, Jiang G, Li R, Ling Y, Wu M, Wei L: Elevated expression of the stem cell marker CD133 associated with Line-1 demethylation in hepatocellular carcinoma. Annals of surgical oncology 2011;18:2373-2380.

Figure Legends

Figure 1. Methylation and expression analyses of K19. (**a**) Schematic representation of the location of discrete *KRT19* gene promoter regions and the results of *KRT19* methylation assessment by highly-sensitive fluorescence assay for bisulfite DNA (Hi-SA). The gray squares denote the coding exon regions of the *KRT19* gene. The blue and green squares represent the restriction sites for HhaI; vertical lines indicate CpG sites; thick horizontal blue and green lines represent polymerase chain reaction fragments; arrows represent primers; and black arrows point out methylated alleles cleaved by the restriction enzyme. (**b**) Cloning and sequencing of *KRT19* regions 1 and 2. Polymerase chain reaction products that were amplified by primer sets for bisulfite cloning were cloned into a TOPO cloning vector and sequenced. For the two cell lines, at least 12 clones were sequenced. Empty circles indicate unmethylated CpG sites. Filled circles represent methylated CpG sites. (**c**) Expression of K19 protein in cell lines and methylation rates in the *KRT19* gene and *LINE-1*. (**d**) Expression ratio (log2 ratio) of *KRT19* messenger RNA in the five cell lines (HepG2, HuH7, HLE, HLF, and PLC/PRF/5). Expression ratio denotes log2 ratio obtained from the signal intensity of *KRT19* messenger RNA from the cells treated with 5-aza-dC and TSA divided by the signal intensity of *KRT19* messenger RNA from the cells untreated by a microarray analysis.

Figure 2. STROBE diagram of the patient cohort.

Figure 3. Immunohistochemistry (IHC) for K19 and clinical outcomes of 125 HCCs. (**a**) Examples of IHC staining of K19 in HCCs classified into five groups according to the percentage of tumor cells with positive staining. (**b**) Kaplan-Meier survival curves for disease-free survival (DFS), overall survival (OS), and extrahepatic metastasis (EHM) recurrence-free survival (EFS) according to K19 expression status.

Figure 4. Methylation analyses of *KRT 19* and *LINE-1* in 125 HCCs. (**a**) Results of methylation analysis in *KRT19* region 1. Average methylation level (left panel) and methylation frequency (right panel) of *KRT19*

region 1 in HCCs with or without K19 proficiency. (**b**) Results of methylation analysis in *KRT19* region 2. Average methylation level (left panel) and methylation frequency (right panel) of *KRT19* region 2 in HCCs with or without K19 proficiency. (**c**) Results of methylation analysis in *LINE-1*. Average methylation level (left panel) and methylation frequency (right panel) of *LINE-1* in HCCs with or without K19 proficiency. Left panel; in the box plot diagrams, the horizontal line within each box represents the median; the limits of each box are the interquartile ranges, and the whiskers are the maximum and minimum values. The *P* values were based on Kruskal-Wallis 1-way analysis of variance on ranks and represent the statistical differences in average methylation between K19-deficient and proficient HCCs. Right panel; the numbers in each box (right panel) denote the number of cases. The *P* values were based Fisher's exact test. (**d**) Scatter-plot matrix demonstrating the pairwise Spearman's correlation coefficient [ρ] between three analyzed loci in the cohort of 125 HCCs. The Y and X-axes denote methylation rates.

Figure 5. Correlation between six markers examined by immunohistochemical (IHC) staining in125 HCCs. (A,B) Scatter-plot matrices demonstrating the pairwise Spearman's correlation coefficient [ρ] between six analyzed markers in the cohort of 96 K19-deficient HCCs (**a**) and of 29 K19-proficient HCCs (**b**). The Y and X-axes denote IHC score.

Supplementary files

Supplementary Table 1. Primer Sequences.

Supplementary Table 2. Patient's characteristics.

Supplementary Fig. S1. Examples of *KRT19* methylation assessment by highly-sensitive fluorescence assay for bisulfite DNA (Hi-SA).

Supplementary Fig. S2. Demethylation analysis of *KRT19* in HLF cell lines. Change of methylation rate in the *KRT19* promoter (A) and expression analysis of K19 (B) 5 days after treatment with various concentrations of the 5-aza-2-dC agent alone.

Supplementary Fig. S3. Methylation (HM450) beta-values for genes in 442 HCC samples. For genes with multiple methylation probes, the probe most anti-correlated with expression.

Supplementary Fig. S4. Immunohistochemistry (IHC) for K7, NOTCH-1, vimentin, E-cadherin, arginase-1, and HepPar-1. Examples of IHC staining of the six markers in HCCs classified into five groups according to the percentage of tumor cells with positive staining. (A) by high power focus. (B) by low power focus.

Heterogeneity of Epigenetic and Epithelial Mesenchymal Transition Marks in Hepatocellular Carcinoma with Keratin 19 Proficiency

Naosuke Yokomichi, M.D.,[a] Naoshi Nishida, M.D., PhD.,[b] Yuzo Umeda, M.D., PhD.,[a] Fumitaka Taniguchi, M.D.,[a] Kazuya Yasui, M.D.,[a] Toshiaki Toshima, M.D.,[a] Yoshiko Mori, M.D.,[a] Akihiro Nyuya, M.S.,[a, c] Takehiro Tanaka, M.D., PhD.,[d] Takeshi Yamada, M.D., PhD.,[e] Takahito Yagi, M.D., PhD.,[a] Toshiyoshi Fujiwara, M.D., PhD.[a] Yoshiyuki Yamaguchi, M.D.,PhD.[c] Ajay Goel, PhD.,[f] Masatoshi Kudo, M.D., Ph.D.,[b] and Takeshi Nagasaka, M.D., PhD.,[a, c*]

[a]Department of Gastroenterological Surgery, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama 700-0914 Japan

[b]Department of Gastroenterology and Hepatology, Kindai University Faculty of Medicine, Osaka 589-8511 Japan

[c]Department of Clinical Oncology, Kawasaki Medical School, Kurashiki 701-0192 Japan

[d]Department of Pathology, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama 700-0914 Japan

[e]Department of Gastrointestinal and Hepato-Biliary-Pancreatic Surgery, Nippon Medical School, Tokyo 113-8603 Japan

[f]Center for Gastrointestinal Cancer Research, Center for Epigenetics, Cancer Prevention and Cancer Genomics, Baylor Research Institute and Charles A Sammons Cancer Center, Baylor University Medical Center, Dallas, TX 75246, USA

Running title: Hepatocellular Carcinoma with Keratin 19 Proficiency

Corresponding Author:

Takeshi Nagasaka, M.D., Ph.D (takeshin@med.kawasaki-m.ac.jp), Department of Clinical Oncology, Kawasaki Medical School, Kurashiki City, Okayama, 701-0192 Japan

Phone: +81-86-462-1111; Fax: +81-86-464-1134

E-mail Address:

Naosuke Yokomichi: n-yokomichi@sis.seirei.or.jp

Naoshi Nishida: naoshi@med.kindai.ac.jp

Yuzo Umeda: y.umeda@d9.dion.ne.jp

Fumitaka Taniguchi: fmtktngch@gmail.com

Kazuya Yasui: kazu57790852@gmail.com

Toshiaki Toshima: tossy2020@yahoo.co.jp

Yoshiko Mori: yoshikomori1@gmail.com

Akihiro Nyuya: n.akihiro0406@gmail.com

Takehiro Tanaka: takehiro@md.okayama-u.ac.jp

Takeshi Yamada: y-tak@nms.ac.jp

Takahito Yagi: liver@md.okayama-u.ac.jp

Toshiyoshi Fujiwara: toshi_f@md.okayama-u.ac.jp

Yoshiyuki Yamaguchi: shogo@med.kawasaki-m.ac.jp

Ajay Goel: Ajay.Goel@bswhealth.org

Masatoshi Kudo: m-kudo@med.kindai.ac.jp

Abstract

Objective: Keratin 19 (K19) expression is a potential predictor for poor prognosis in patients with hepatocellular carcinoma (HCC). To clarify the feature of K19-proficient HCC, we traced epigenetic footprints in cultured cells and clinical materials.

Patients and Methods: *In vitro*, *KRT19* promoter methylation was analyzed and 5-aza-dC with trichostatin A (TSA) treatment was performed. Among 564 surgically resected HCCs, the clinicopathological relevance of K19-proficent HCCs was performed in comparison with hepatocytic (HepPar-1 and arginase-1), epithelial-mesenchymal transition (E-cadherin and vimentin), biliary differentiation-associated (K7 and NOTCH-1) markers, and epigenetic markers (*KRT19* promoter/long interspersed nucleotide element-1 [*LINE-1*] methylation status).

Results: *KRT19* promoter methylation was clearly associated with K19 deficiency and 5-aza-dC with trichostatin A treatment stimulated K19 re-expression, implicating DNA methylation as a potential epigenetic process for K19 expression. After excluding HCCs with recurrence, TNM stage as IIIB or greater, preoperative therapy, transplantation, and combined hepatocellular-cholangiocarcinoma, we assessed 125 from 564 HCC cases. In this cohort, K19 expression was found in 29 HCCs (23.2%), and corresponded with poor survival following surgery (P = 0.025) and extrahepatic recurrence free survival (P = 0.017). Compared with K19-deficient HCCs, lower *KRT19* promoter methylation level was observed in K19-proficient HCCs (P < 0.0001). Conversely, HCC with genome-wide *LINE-1* hypermethylation was frequently observed in K19-proficient HCCs (P = 0.0079). Additionally, K19 proficiency was associated with K7 proficiency (P = 0.043), and reduced E-cadherin and HepPar-1 expression (P = 0.043 and < 0.0001, respectively).

***Conclusions*:** K19-proficient HCC exhibited poor prognosis owing to extrahepatic recurrence, with molecular signatures differing from those in conventional cancer stem cells, providing novel insights of the heterogeneity underlying tumor development.

***Keywords*:** K19, KRT19, Hepatocellular carcinoma, Methylation, LINE-1,

Introduction

Hepatocellular carcinoma (HCC) constitutes the sixth most common neoplasm and the third leading cause of cancer deaths worldwide [1-3]. Surgical resection and liver transplantation comprise the only curative treatments for patients with early-stage HCC. However, the high rate of recurrence or metastases leads to worse prognosis in HCC after curative resection [2].

The positive expression of keratin 19 (K19), a marker for biliary or hepatic progenitor cells and early hepatoblasts, has been significantly associated with poor prognosis along with stemness-related and epithelial-mesenchymal transition (EMT) features in HCC [4-8]. In addition, K19 proficiency in patients with HCC has frequently been associated with vascular invasion, poorly differentiated tumors, and tumor recurrence after resection, radiofrequency ablation, or transplantation [9]. However, to date the cell origin of K19-proficient HCC has remained unclear [6]. Multiple studies suggested that HCC cells with K19 proficiency were originated from hepatic progenitor cells as such cells express progenitor cell markers, have invasive potential, and exhibit chemoresistance [6,9-11]. Conversely, others indicated that the expression of K19 in human HCCs may result from the dedifferentiation of malignant hepatocytes during continuous mutagenesis [12-14]. Furthermore, studies of liver regeneration have also found difficulties in identifying the cell origin of liver cancers including K19 proficient HCCs [15,16].

Although the mechanisms contributing to hepatocarcinogenesis remain unclear, it is widely accepted that HCC exhibits numerous genetic abnormalities, such as chromosomal alterations, gene amplifications, and mutations, as well as epigenetic alterations [17]. For example, increased DNA methylation levels of tumor suppressor genes correlate positively with HCC development and progression [18,19]; accordingly, a genome-wide methylation analysis has identified tumor suppressor genes in HCC [20]. However, although the gene encoding K19 (*KRT19*), located on chromosome 17 [21], contains a CpG island within its promoter region [22], to date there have been no reports regarding *KRT19* promoter methylation in any malignancies. Thus, although K19 expression regulation mechanisms in HCC have not been fully elucidated, the presence of a promoter CpG island suggests DNA methylation as a potential epigenetic process in this malignancy.

In the present study, we attempted to reveal the features of K19-proficient HCC by tracing epigenetic footprints in cultured cells and clinical materials. Firstly, we examined epigenetic alterations and underlying molecular mechanisms of K19-positive HCC cell lines. Next, from a panel of 564 surgically resected HCCs, we clarified the clinicopathological relevance of K19-proficent HCCs by analyzing robust methylation in the *KRT19* promoter region and *LINE-1* elements in comparison with other cholangiocytic (K7), hepatocytic (HepPar-1 and arginase-1), EMT (E-cadherin and vimentin), and biliary differentiation-associated (NOTCH-1) markers.

Patients and Methods

Patients

We retrospectively analyzed consecutive patients with initial HCC who received surgical resection at Okayama University Hospital from January 2000 to December 2010. The histopathological diagnosis of HCC was based on the World Health Organization criteria. The clinical history, pathological reports, and hematoxylin and eosin stained slides for all cases were reviewed to confirm the diagnosis. We excluded patients with a diagnosis of combined hepatocellular cholangiocarcinoma, with recurrent HCC, major vascular invasion, and with rupture or other organ invasion. We additionally excluded patients who received transplantation, non-curative resection, and preoperative therapy; e.g., transcatheter arterial chemoembolization. The tumor size was measured macroscopically after the removal of the tumor. Microvascular invasion was graded based on histopathological evaluation. The histological grade of tumor differentiation was determined according to the classification of Edmondson and Steiner (ES differentiation grade) [23]. Fibrosis of the non-neoplastic parenchyma was classified based on the Meta-analysis of Histologic Data in Viral Hepatitis (METAVIR) scoring system that assesses the degree of fibrosis ranging from F0 (no fibrosis) to F4 (cirrhosis) [24], Tumor grade was classified according to the 7th edition UICC/AJCC TNM staging system [25]. Survival time was determined from the date of surgical resection. Local or remote recurrence of disease after surgery was investigated by clinical assessment and regular abdominal ultrasonography or computed tomography. The study was approved by the institutional review

board of the Okayama University Hospital.

Cell lines

K19-proficient HCC cell lines: HepG2, HuH7, and PLC/PRF/5; K19-deficient HCC cell lines: HLE and HLF; and a colon cancer cell line HT29 as a K19 positive control were purchased from American Type Culture Collection (Manassas, VA) [26]. All cell lines were cultured using Dulbecco's modified Eagle's medium (Sigma Aldrich, St. Louis, MO) supplemented with 10% fetal bovine serum, penicillin (100 IU/mL), and streptomycin (100 μg/mL) at 37°C in a humidified incubator with 5% CO_2.

Immunohistochemistry (IHC) analysis

IHC was performed using formalin-fixed, paraffin-embedded tissue sections of surgically resected liver specimens. After deparaffinization and blocking of the endogenous peroxidase, antigen retrieval was performed by microwaving with citrate buffer (pH 6.0) or TRIS-ethylenediaminetetraacetic acid (pH 9.0). The sections were incubated with primary antibodies against K19 (1:200, mouse monoclonal, Cell Signaling Technology, Danvers, MA), keratin 7 (K7, 1:100, mouse monoclonal; Daco), HepPar-1 (1:100, mouse monoclonal; Daco, Glostrup, Denmark), arginase-1 (1:5000, rabbit polyclonal; Sigma-Aldrich), E-cadherin (1:1, mouse monoclonal; Daco), and vimentin (1:200, mouse monoclonal; Abcam, Cambridge, UK). Then, the sections were incubated with secondary antibodies conjugated to peroxidase-labeled polymer, using the EnVision system (Dako). Color development was performed using 3, 3-diaminobenzidine and the sections were counterstained with hematoxylin. Negative controls were carried out by substitution of the primary antibodies with non-immunized serum, resulted in no signal detection. IHC results were interpreted by pathologists blinded to the corresponding clinicopathological data. K19, K7, NOTCH-1, and cytoplasmic vimentin were scored according to the percentage of tumor cells with positive staining as follows: 0 = <5%, 1 = 5–10%, 2 = 11–20%, 3 = 21–50%, and 4 = 51–100%, and considered as positive when ≥5% of tumor cells were stained, as reported previously [4-8]. HepPar-1, arginase-1, and E-cadherin were considered as positive when ≥51% of tumor cells were stained [27], and were scored using the percentage of tumor cells with positive staining as follows: 0 = 0–10%, 1 = 11–20%, 2 = 21–50%, 3 = 51–80%, and 4 = 81–100%.

Western blotting

Protein was separated by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis and electroblotted onto Immun-Blot polyvinylidene fluoride membranes (BIO-RAD, Hercules, CA). After treating with 5% fat-free dried milk in 1x TBST for 1 hour at room temperature, the membranes were incubated with mouse anti-human CK19 (1:1000; Cell Signaling Technology) for 1 hour at room temperature, followed by secondary antibody for 1 hour at room temperature. Actin was used as an internal positive control. Target proteins were detected using enhanced chemiluminescence (ECL; BIO-RAD).

Bisulfite modification of DNA

Genomic DNA from the cell lines and formalin-fixed, paraffin-embedded specimens was extracted using a QIAamp DNA Mini Kit (Qiagen, Valencia, CA) and TaKaRa DEXPAT Kit (TaKaRa Bio Inc., Otsu, Japan), respectively. Approximately 1 μg DNA was subsequently modified using the EZ DNA Methylation Kit (Zymo Research, Orange, CA).

Bisulfite sequencing

Polymerase chain reaction products from the *KRT19* promoter were amplified by a set of primers (**Supplementary Table S1**) for bisulfite DNA cloning and sequencing. The *KRT19* promoter was cloned from bisulfite-treated DNA into the pCR2.1TOPO vector using the TOPO-TA cloning system (Life Technologies, Carlsbad, CA), followed by automated DNA sequencing with both the forward (F) and reverse (R) primers using an ABI 310-Avant NA sequencer (Applied Biosystems, Foster City, CA).

Methylation analysis of KRT19 and LINE-1

We performed quantitative methylation analysis for the promoter CpG islands of *KRT19* and *LINE-1* in tumors, their matched corresponding non-tumor liver tissues, and cell lines. For *KRT19* analysis, we used the high-sensitive assay for bisulfite DNA (Hi-SA), a modified combined bisulfite restriction analysis method, by which fluorescence labeled DNA fragments are detected using a genetic analyzer.[28] The primer sequences are summarized in **Supplementary Table S1**, whereas the details for the remaining assays have been described previously [28,29]. We analyzed two regions in the *KRT19* promoter; termed region 1 and region 2. Region 2 contained two sites that can be restricted by *Hha*I. The methylation ratio was

calculated by the ratio of the peak of methylated and unmethylated bands (examples are shown in **Supplementary Figure S1**). We used primers for *LINE-1* methylation analysis as described previously [30] with assay modification by adding a fluorescent dye to measure methylation ratio using an ABI 310-Avant NA sequencer (**Supplementary Table S1**). Additionally, we analyzed association between methylation and expression for KRT19 of HCCs by the data obtained from cBioportal for cancer genomics (http://www.cbioportal.org/).

Demethylation analysis

HLF, HLE, PLC/PRF/5, HepG2, and HuH7 cells were cultured with 5-aza-2'-deoxycitidine (5-Aza-dC, Wako, Japan). 5-Aza-dC was dissolved with phosphate buffered saline and diluted with medium. The concentrations of 5-Aza-dC were set as 1μM according to the nadir of promoter methylation rate in both region 1 and 2 of *KRT19* (**Supplementary Figure S2**). After 5-Aza-dC treatment from day1 to day3, cells were treated by 100μM trichostatin A (TSA, MERCK, Darmstadt, Germany), a potent histone deacetylase (HDAC) inhibitor, for 24h. DNA and RNA were extracted after TSA treatment for determining methylation status and expression status by a microarray analysis (SurePrint G3 Human Gene Expression 8X60K v2, Agilent, Santa Clara, CA).

Statistical analysis

All statistical analyses were performed using JMP software (version 10.0; SAS Institute, Inc, Cary, NC). We compared K19 expression status with various clinicopathological features and the results of other immunostaining using the Fisher's exact test. Methylation levels in regions 1 and 2 of the *KRT19* promoter were analyzed as both continuous and categorical variables (positive, methylation level in both region 1 and 2 > 10%). Methylation levels in *LINE-1* were categorized according to the mean value (≥ 55% methylation in *LINE-1* defined as hypermethylation and < 55% defined as hypomethylation). Categorical variables were compared using Fisher's exact test. Differences between continuous variables were determined using the analysis of variance test (ANOVA). Correlation coefficient between continuous variables was nonparametrically determined (Spearman's correlation coefficient [ρ]). Overall survival (OS) was calculated from the date of surgical resection to the date of death owing to HCC or last follow-up for

censored patients. Recurrence free survival was calculated from the date of surgical resection to the date of the first documentation of local, regional, or distant relapse, appearance of a second primary lesion by computed tomography, and/or magnetic resonance imaging routinely performed every 6 months. Extrahepatic-recurrence free survival was calculated from the date of surgical resection to the date of the first documentation of appearance of extrahepatic recurrences by computed tomography and/or magnetic resonance imaging routinely performed every 6 months. OS, recurrence free survival, and extrahepatic-recurrence free survival were estimated using the Kaplan-Meier method. Then, a multivariate analysis for OS was performed using a Cox-proportional hazards model. All *P*-values reported were calculated using two-sided tests and values <0.05 were considered statistically significant.

Results

Association of KRT19 promoter methylation with K19 expression in cultured cell lines

To examine methylation status in the *KRT19* promoter, which contains dense CpG sites, the promoter region of *KRT19* was divided into two regions (region 1 and region 2, **Figure 1a**). Methylation status in the discrete regions was analyzed as a continual variable. Cloning and bisulfite sequencing were used to precisely confirm this methylation status, providing validation and further evidence that the K19-deficient HLF cells showed dense methylation throughout the *KRT19* promoter CpGs whereas the K19-proficient HuH7 cells exhibited no methylation (**Figure 1b**). Both regions in the *KRT19* promoter were hypermethylated in K19-deficient cell lines (HLE and HLF) but were hypomethylated in K19-proficient cell clines (HepG2 and HuH7, **Figure 1c**). The mean methylation ratio of K19-deficient and K19-proficient cells in region 1 was 31.7% (95% confidence interval [95%CI]; 11.6–51.6%) and 6.7% (−6.0–19.4%), respectively ($P = 0.0426$) whereas those of K19-deficient and K19-proficient cells in region 2 were 96.8% (68.2–125.4%) and 6.7% (−4.9–31.2%), respectively ($P = 0.0014$). Thus, methylation levels in both regions of *KRAT19* were significantly inversely associated with K19 expression. We next treated five HCC cell lines (HepG2, HuH7, HLEE, HLF, and PLC/PRF/5) with a demethylating agent, 5'-Aza-dC, and a potent HDAC inhibitor, TSA. Among them, the K19-deficient HLF and HLE cells with higher methylation rate in the *KRT19* promoter showed that, although 5'-Aza-dC alone treatment caused demethylation in the *KRT19* promoter but failed to recover *KRT19* expression, 5'-Aza-dC following TSA treatment successfully recovered

expression level of *KRT19* by the microarray analysis, indicating that K19 expression would be regulated by promoter methylation and histone modification (**Figure 1d**).

We also examined genome global methylation level using *LINE-1* retrotransposons, which constitute a substantial portion (approximately 17%) of the human genome and are regarded as a surrogate marker of global DNA methylation. Although the mean methylation ratio of *LINE-1* of K19-proficient cells (36.5% [95%CI; 2.5–60.5%]) was higher than that of K19-deficient cells (26.8% [−11.2–64.8%]), the difference was not statistically significant (P = 0.6, **Figure 1b**).

Clinicopathological features in patients with K19-proficient HCC

To assess the precise clinical landscapes of HCC with K19 proficiency, we retrospectively examined patients with HCC who underwent surgical resection. Among 564 patients recruited, 125 met the inclusion and exclusion criteria of this study (**Figure 2**). The clinicopathological features of 125 patients are summarized in **Supplementary Table S2.** Of these, 113 patients (90.4%) were categorized in TNM stage I or II and no patients exhibited major vascular invasion. IHC analyses revealed that K19 expression was detected in 29 patients (23.2%) and that patients with K19-proficient HCC were significantly younger than those with K19-deficient HCC (P = 0.020, **Figure 3a** and **Table 1**). Although a high level of serum alpha-fetoprotein and microvascular invasion were more frequently observed in patients with K19-proficient HCC (P = 0.021 and 0.019, respectively), K19 proficiency had no association with TNM stage, tumor size, tumor number, or differentiation. Patients with K19-proficient HCC showed poorer survival after surgery (P = 0.025) and poorer extra-hepatic metastasis-free survival (P = 0.017, **Figure 3b-d**). Multivariate analysis demonstrated that K19 acted as an independent prognostic factor for survival after surgery (**Table 2**).

K19 expression and promoter methylation status in the KRT19 gene

Next, we investigated the methylation levels of discrete regions in the *KRT19* promoter in the 125 HCCs. Initially, methylation levels in the discrete regions obtained from fluorescent Hi-SA were analyzed as a continuous variable. In region 1, the mean methylation level was 2.3% [95%CI; 1.2–3.5%] among HCC tissues with K19 proficiency but 8.7% (95%CI; 5.4–11.9%) among HCC tissues with K19 deficiency (P =

0.0315, ANOVA; **Figure 4a**). In region 2, the mean methylation levels were 2.7% [95%CI; 0.9–4.5%] among HCC tissues with K19 proficiency and 15.7% (95%CI; 9.6–21.8%) among HCC tissues with K19 deficiency (*P* = 0.0228, ANOVA; **Figure 4b**). To define the threshold of methylation levels in region 1 and region 2 of the *KRT19* promoter, we examined the methylation levels in the discrete regions of adjacent normal liver tissues. The mean methylation level in 123 adjacent normal liver tissues was 4.3% [95%CI; 2.5–6.1%] in region 1 and 8.3% in region 2. Therefore, we defined *KRT19* methylation in both regions at 10% or more as a continuous variable [i.e., ≥10% methylation as methylation-positive (methylated) and <10% methylation as methylation-negative (unmethylated)]. Using this criterion, we found that 20 (16.0%) and 28 (22.4%) of 125 HCCs could be categorized as methylated in region 1 and region 2 of the *KRT19* promoter, respectively. All of the 20 HCCs (100%) categorized as methylated in region 1 showed K19 deficiency whereas all of the 29 K19-proficient HCCs (100%) were unmethylated in region 1 (*P* = 0.0038, **Figure 4a**. With respect region 2, 25 of 28 HCCs (89.3%) with region 2 methylation showed K19 deficiency whereas 26 of 29 HCCs (89.7%) with K19-proficient HCCs were unmethylated (*P* = 0.12, **Figure 4b**). We also examined association between methylation and expression for KRT19 of 442 HCCs by the data obtained from cBioportal for cancer genomics (**Supplementary Figure S3**). Although K19-deficient HCCs had various methylation levels in KRT19, most of K19-profcient HCCs showed the lower methylation levels. Thus, HCCs with K19 proficiency demonstrated unmethylation in the discrete promoter regions of the *KRT19* gene. Together with the results of our *in vitro* study, these findings suggest that the mechanism of K19 expression may be partially regulated by promoter methylation and histone modification in the *KRT19* gene.

***Association between LINE-1* methylation and *KRT19* promoter methylation in K19-proficient HCCs**

Evaluation of the *LINE-1* methylation levels in the 119 HCC tissues demonstrated that the mean methylation level was 54.8% (95%CI; 50.7–59.0%, **Figure 4c**). When we categorized *LINE-1* methylation levels in the two groups according to the threshold of 55% (≥ 55% methylation defined as hypermethylation and < 55% defined as hypomethylation), hypermethylation in *LINE-1* was frequently observed in K19-proficient (19 of 26 [73.1%]) compared with K19-deficient HCCs (40 of 93 [43.0%], *P* = 0.0079) (**Figure 4c**).

We then asked whether an association exists between genome-wide methylation level and the local regions in the *KRT19* promoter. For this, we performed multivariate correlation by comparing the methylation levels of *LINE-1* and regions 1 and 2 in the *KRT19* promoter from K19- proficient and deficient HCCs, respectively (**Figure 4d, e**). Regardless of K19 expression status, positive correlations were observed in the methylation level between regions 1 and 2 in the *KRT19* promoter (Spearman's correlation coefficients [ρ] = 0.2941, *P* = 0.0043 in K19-deficient HCCs; ρ = 0.3455, *P* = 0.0644 in K19-proficient HCCs). In contrast, although there was no association between *LINE-1* methylation and *KRT19* promoter methylation levels in K19-deficient HCCs, an inverse correlation was observed in K19-proficient HCCs (ρ = −0.3353, *P* = 0.094 between *LINE-1* and region 1; ρ = −0.4424, *P* = 0.0236 between *LINE-1* and region 2).

Markers associated with expression profiles of EMT, hepatocytic and biliary differentiation.

We next clarified features of HCC in relation to K19 expression status in comparison with EMT markers (E-cadherin and vimentin), hepatocytic markers (HepPar-1 and Arginase-1), and markers associated with biliary differentiation (K7 and NOTCH-1). Examples of scoring of IHC staining are shown in **Supplementary Figure S4**. K19-proficient HCCs demonstrated increased EMT features with loss of E-cadherin (*P* = 0.043) and gain of vimentin expression (*P* = 0.084, **Table 3**). With respect to organ signature, findings of hepatocytic or cholangiocytic markers contrasted; although K19-proficient HCCs exhibited decreased number of cases with positive expression of the hepatocytic markers HepPar-1 and Arginase-1, they showed at least one positive result of staining for HepPar-1 or Arginase-1. Thus, K19 expression correlated with an increase of EMT features accompanying K7 positive expression and a reduction of hepatocytic features. By multivariate correlations (**Figure 5**), in K19-deficient HCCs, hepatocytic markers showed inverse associations with vimentin and NOTCH-1. Conversely, for HCCs with K19-proficiency, although expression of E-cadherin was decreased and that of K7 was increased, the positive association of HepPar-1 and arginase-1 expression [ρ = 0.5981, *P* = 0.0006]) indicated that hepatocytic features were conserved.

Discussion

Herein, we have shown the biological significance of HCC with K19 proficiency with relation to methylation status in the promoter region of *KRT19*, genome-wide methylation level of *LINE-1*, and EMT features. Results from clinical samples as well as our *in vitro* study indicated that K19-deficient HCCs demonstrated higher methylation level in the promoter region of *KRT19* compared with K19-proficient HCCs, suggesting that K19 expression might be regulated by the density of its promoter methylation. Our demethylation analysis demonstrated re-expression of *KRT19* required not only application of a demethylating agent, 5'-Aza-dC, but also a HDAC inhibitor, TSA, in K19-deficient HLF and HLE cells, suggesting that epigenetic mechanisms including histone modification play an important role in K19 expression.

Consistent with previous studies [4-8], we confirmed the utility of K19 as a prognostic biomarker for HCC. Notably, we extracted 125 HCCs from 564 surgically resected HCCs according to exclusion criteria that excluded HCCs with major vascular invasion pathologically, which is the strongest prognostic factor for this malignancy. We also excluded recurrent HCCs for accurate survival analysis, and HCCs with preoperative transcatheter arterial chemoembolization because such treatment might influence the tumor characteristics including K19 expression on the membrane of the cancer cells. Thus, the 125 HCCs analyzed in this study were, as far as possible, homogeneous in both clinical and pathological settings. Of this cohort, K19 proficiency was observed in 23.2% (29/125) HCCs and was more frequently observed in younger, female patients. Other clinical features of K19-proficient HCCs included higher serum alpha-fetoprotein levels, positive for pathological microvascular invasion, and poor survival after surgery.

In this study, K19-proficient HCCs exhibited EMT features with loss of E-cadherin and gain of vimentin expression. This association between EMT features and K19 proficiency was consistent with a prior study reporting that K19-proficient HCC exhibited significantly increased EMT-related protein and mRNA expression [5]. EMT has been shown to be a pivotal mechanism contributing to cancer invasion and metastasis including HCC [31,32]. Consistent with this hypothesis, patients with K19-proficient HCC in this and other studies [7,33] showed poorer survival and more frequently exhibited extra-hepatic metastasis.

Generally, primary liver cancers are classified into the following subtypes; HCCs, cholangiocarcinomas (CCAs), combined HCC-CCAs, hepatoblastomas, and fibrolamellar hepatocellular carcinomas [34]. K19 is commonly expressed in two types of liver cancers, CCAs and combined HCC-CCAs; however, their cell origins and means of development are not yet sufficiently understood. Recent studies using a mouse model of hepatocyte fate tracing have revealed that CCAs may be originated from fully differentiated hepatocytes via NOTCH signaling activation [35,36]. Combined HCC-CCAs are pathologically diagnosed based on the classical type of combined HCC-CCAs as areas of typical HCC and CCA mixed within the tumor, with the latest edition of this classification proposing a subtype with stem cell features [37].

Kawai et al. reported that K19-proficient HCCs possessed cancer stem cell features, such as EMT features and the activation of the TGFb/Smad signal cascade [38]. Zhang et al. reported that HCCs expressed a stem cell marker CD133, considered to be a marker for cancer stem cells of HCC, and showed hypomethylation of the global DNA methylation marker *LINE-1* [39]. Therefore, we next examined the association between K19-proficent HCCs and *LINE-1* methylation levels associated with cancer stem cell features. Notably, in our cohort, K19-proficent HCCs exhibited *LINE-1* hypermethylation that was significantly correlated with demethylation in the *KRT19* promoter. This evidence obtained from clinical samples was also supported by our *in vitro* study that K19-proficient cell lines were more likely show increased *LINE-1* methylation level (**Figure 1b**). As Kim et al. demonstrated that only 1.5% of HCCs showed both CD133 and K19 expression [5], it is reasonable to presume that the K19-proficient HCCs were not equivalent to the CD133-proficient HCCs that are associated with *LINE-1* hypomethylation. With respect to IHC staining, our cohort of HCCs showed inverse correlation in expression status between NOTCH-1 and hepatocytic markers irrespective of K19 expression status. Although K19-proficient HCCs demonstrated increased EMT and cholangiocytic features and reduced number of cases with positive expression of hepatocytic markers, strongly conserved hepatocytic features were still observed.

This study has some limitations. For example, analyzed samples were obtained from a retrospective cohort in a single hospital. However, this study describes the novel features observed in HCCs with K19 proficiency; increased EMT features and decreased mature signatures of hepatocyte cells; increased genome-wide DNA methylation levels and reduced promoter methylation density in the *KRT19* gene.

Additionally, *in vitro* analysis revealed that expression of K19 was regulated under promoter methylation and histone modification. Our clinical data suggest that not only detection of K19 proficiency in HCC may be useful as a biomarker for identifying patients who have a poor prognosis with extrahepatic recurrence, but also that K19-proficient HCCs likely arise from hepatocytes or HCCs via epigenetic reprogramming, leading to EMT features. Thus, our findings provide novel insights regarding the heterogeneity underlying tumor development.

Declarations

Acknowledgements

The authors would like to thanks Mr. Toru Nakai and Mrs. Tae Yamanishi for the immunohistochemistry techniques.

Funding

This work was supported by grants from MEXT/JSPS KAKENHI (20590572, 25860409, 26462016, and 15H03034 to TN).

Availability of data and materials

The datasets obtained and/or analyzed during the current study are available from the corresponding author in reasonable request.

Authors' contributions

NY performed experiments and analysed data supported by YU, FT, KY, TT, YM, and AY; NN and MK performed all microarray analysis; TT and TakeY performed all immunohistochemistry analysis. TakaY and TF prepared all samples and summrized clinical information; AL and SL recruited patients for analysis; YY and AG analysed data and wrote the paper; TN obtained funding, designed the work, analysed data, and wrote the paper.

Ethics approval and consent to participate

The human tissues used in this study were approved by the institute ethical committee of Okayama University.

Consent for publication

Not applicable.

Competing interests

The authors have no conflicts of interest that pertain to this work.

References

1 Boyle P, Levin B, International Agency for Research on Cancer., World Health Organization.: World cancer report 2008. Lyon
Geneva, International Agency for Research on Cancer ;
Distributed by WHO Press, 2008.

2 Forner A, Llovet JM, Bruix J: Hepatocellular carcinoma. Lancet 2012;379:1245-1255.

3 Jemal A, Center MM, DeSantis C, Ward EM: Global patterns of cancer incidence and mortality rates and trends. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2010;19:1893-1907.

4 Wu PC, Fang JW, Lau VK, Lai CL, Lo CK, Lau JY: Classification of hepatocellular carcinoma according to hepatocellular and biliary differentiation markers. Clinical and biological implications. The American journal of pathology 1996;149:1167-1175.

5 Kim H, Choi GH, Na DC, Ahn EY, Kim GI, Lee JE, Cho JY, Yoo JE, Choi JS, Park YN: Human hepatocellular carcinomas with "Stemness"-related marker expression: keratin 19 expression and a poor prognosis. Hepatology 2011;54:1707-1717.

6 Govaere O, Komuta M, Berkers J, Spee B, Janssen C, de Luca F, Katoonizadeh A, Wouters J, van Kempen LC, Durnez A, Verslype C, De Kock J, Rogiers V, van Grunsven LA, Topal B, Pirenne J, Vankelecom H, Nevens F, van den Oord J, Pinzani M, Roskams T: Keratin 19: a key role player in the invasion of human hepatocellular carcinomas. Gut 2014;63:674-685.

7 Uenishi T, Kubo S, Yamamoto T, Shuto T, Ogawa M, Tanaka H, Tanaka S, Kaneda K, Hirohashi K: Cytokeratin 19 expression in hepatocellular carcinoma predicts early postoperative recurrence. Cancer science 2003;94:851-857.

8 Durnez A, Verslype C, Nevens F, Fevery J, Aerts R, Pirenne J, Lesaffre E, Libbrecht L, Desmet V, Roskams T: The clinicopathological and prognostic relevance of cytokeratin 7 and 19 expression in hepatocellular carcinoma. A possible progenitor cell origin. Histopathology 2006;49:138-151.

9 Villanueva A, Hoshida Y, Battiston C, Tovar V, Sia D, Alsinet C, Cornella H, Liberzon A, Kobayashi M, Kumada H, Thung SN, Bruix J, Newell P, April C, Fan JB, Roayaie S, Mazzaferro V, Schwartz ME, Llovet JM: Combining clinical, pathology, and gene expression data to predict recurrence of hepatocellular carcinoma. Gastroenterology 2011;140:1501-1512 e1502.

10 Villanueva A, Hoshida Y, Toffanin S, Lachenmayer A, Alsinet C, Savic R, Cornella H, Llovet JM: New strategies in hepatocellular carcinoma: genomic prognostic markers. Clinical cancer research : an official journal of the American Association for Cancer Research 2010;16:4688-4694.

11 Andersen JB, Loi R, Perra A, Factor VM, Ledda-Columbano GM, Columbano A, Thorgeirsson SS: Progenitor-derived hepatocellular carcinoma model in the rat. Hepatology 2010;51:1401-1409.

12 Santos NP, Oliveira PA, Arantes-Rodrigues R, Faustino-Rocha AI, Colaco A, Lopes C, Gil da Costa RM: Cytokeratin 7/19 expression in N-diethylnitrosamine-induced mouse hepatocellular lesions: implications for histogenesis. International journal of experimental pathology 2014;95:191-198.

13 Ezzoukhry Z, Louandre C, Trecherel E, Godin C, Chauffert B, Dupont S, Diouf M, Barbare JC, Maziere JC, Galmiche A: EGFR activation is a potential determinant of primary resistance of hepatocellular carcinoma cells to sorafenib. International journal of cancer Journal international du cancer 2012;131:2961-2969.

14 Yoneda N, Sato Y, Kitao A, Ikeda H, Sawada-Kitamura S, Miyakoshi M, Harada K, Sasaki M, Matsui O, Nakanuma Y: Epidermal growth factor induces cytokeratin 19 expression accompanied by increased growth abilities in human hepatocellular carcinoma. Laboratory investigation; a journal of technical methods and pathology 2011;91:262-272.

15 Itoh T, Miyajima A: Liver regeneration by stem/progenitor cells. Hepatology 2014;59:1617-1626.

16 Yanger K, Zong Y, Maggs LR, Shapira SN, Maddipati R, Aiello NM, Thung SN, Wells RG, Greenbaum LE, Stanger BZ: Robust cellular reprogramming occurs spontaneously during liver regeneration. Genes & development 2013;27:719-724.

17 Nishida N, Goel A: Genetic and epigenetic signatures in human hepatocellular carcinoma: a systematic review. Current genomics 2011;12:130-137.

18 Nishida N, Nagasaka T, Nishimura T, Ikai I, Boland CR, Goel A: Aberrant methylation of multiple tumor suppressor genes in aging liver, chronic hepatitis, and hepatocellular carcinoma. Hepatology 2008;47:908-918.

19 Nishida N, Kudo M, Nagasaka T, Ikai I, Goel A: Characteristic patterns of altered DNA methylation predict emergence of human hepatocellular carcinoma. Hepatology 2012;56:994-1003.

20 Revill K, Wang T, Lachenmayer A, Kojima K, Harrington A, Li J, Hoshida Y, Llovet JM, Powers S: Genome-wide methylation analysis and epigenetic unmasking identify tumor suppressor genes in hepatocellular carcinoma. Gastroenterology 2013;145:1424-1435 e1421-1425.

21 Schweizer J, Bowden PE, Coulombe PA, Langbein L, Lane EB, Magin TM, Maltais L, Omary MB, Parry DA, Rogers MA, Wright MW: New consensus nomenclature for mammalian keratins. The Journal of cell biology 2006;174:169-174.

22 Zody MC, Garber M, Adams DJ, Sharpe T, Harrow J, Lupski JR, Nicholson C, Searle SM, Wilming L, Young SK, Abouelleil A, Allen NR, Bi W, Bloom T, Borowsky ML, Bugalter BE, Butler J, Chang JL, Chen CK, Cook A, Corum B, Cuomo CA, de Jong PJ, DeCaprio D, Dewar K, FitzGerald M, Gilbert J, Gibson R, Gnerre S, Goldstein S, Grafham DV, Grocock R, Hafez N, Hagopian DS, Hart E, Norman CH, Humphray S, Jaffe DB, Jones M, Kamal M, Khodiyar VK, LaButti K, Laird G, Lehoczky J, Liu X, Lokyitsang T, Loveland J, Lui A, Macdonald P, Major JE, Matthews L, Mauceli E, McCarroll SA, Mihalev AH, Mudge J, Nguyen C, Nicol R, O'Leary SB, Osoegawa K, Schwartz DC, Shaw-Smith C, Stankiewicz P, Steward C, Swarbreck D, Venkataraman V, Whittaker CA, Yang X, Zimmer AR, Bradley A, Hubbard T, Birren BW, Rogers J, Lander ES, Nusbaum C: DNA sequence of human chromosome 17 and analysis of rearrangement in the human lineage. Nature 2006;440:1045-1049.

23 Edmondson HA, Steiner PE: Primary carcinoma of the liver: a study of 100 cases among 48,900 necropsies. Cancer 1954;7:462-503.

24 Ichida F, Tsuji T, Omata M, Ichida T, Inoue K, Kamimura T, Yamada G, Hino K, Yokosuka O,

Suzuki H: New Inuyama classification; new criteria for histological assessment of chronic hepatitis. International Hepatology Communications 1996;6:112-119.

25 Edge S, Byrd D, Compton C, Fritz A, Greene F, Trotti A: AJCC cancer staging manual, ed 7th. New York, Springer, 2010.

26 Wu F, Nishioka M, Fujita J, Murota M, Ohtsuki Y, Ishida T, Kuriyama S: Expression of cytokeratin 19 in human hepatocellular carcinoma cell lines. International journal of oncology 2002;20:31-37.

27 Yan BC, Gong C, Song J, Krausz T, Tretiakova M, Hyjek E, Al-Ahmadie H, Alves V, Xiao SY, Anders RA, Hart JA: Arginase-1: a new immunohistochemical marker of hepatocytes and hepatocellular neoplasms. The American journal of surgical pathology 2010;34:1147-1154.

28 Nagasaka T, Tanaka N, Cullings HM, Sun DS, Sasamoto H, Uchida T, Koi M, Nishida N, Naomoto Y, Boland CR, Matsubara N, Goel A: Analysis of fecal DNA methylation to detect gastrointestinal neoplasia. Journal of the National Cancer Institute 2009;101:1244-1258.

29 Yoshida K, Nagasaka T, Umeda Y, Tanaka T, Kimura K, Taniguchi F, Fuji T, Shigeyasu K, Mori Y, Yanai H, Yagi T, Goel A, Fujiwara T: Expansion of epigenetic alterations in EFEMP1 promoter predicts malignant formation in pancreatobiliary intraductal papillary mucinous neoplasms. Journal of cancer research and clinical oncology 2016;142:1557-1569.

30 Lertkhachonsuk R, Paiwattananupant K, Tantbirojn P, Rattanatanyong P, Mutirangura A: LINE-1 Methylation Patterns as a Predictor of Postmolar Gestational Trophoblastic Neoplasia. BioMed research international 2015;2015:421747.

31 Moody SE, Perez D, Pan TC, Sarkisian CJ, Portocarrero CP, Sterner CJ, Notorfrancesco KL, Cardiff RD, Chodosh LA: The transcriptional repressor Snail promotes mammary tumor recurrence. Cancer cell 2005;8:197-209.

32 Ma CQ, Yang Y, Wang JM, Du GS, Shen Q, Liu Y, Zhang J, Hu JL, Zhu P, Qi WP, Qian YW, Fu Y: The aPKCiota blocking agent ATM negatively regulates EMT and invasion of hepatocellular carcinoma. Cell death & disease 2014;5:e1129.

33 Zhuang PY, Zhang JB, Zhu XD, Zhang W, Wu WZ, Tan YS, Hou J, Tang ZY, Qin LX, Sun HC: Two pathologic types of hepatocellular carcinoma with lymph node metastasis with distinct prognosis on the basis of CK19 expression in tumor. Cancer 2008;112:2740-2748.

34 Oikawa T: Cancer Stem cells and their cellular origins in primary liver and biliary tract cancers. Hepatology 2016;64:645-651.

35 Sekiya S, Suzuki A: Intrahepatic cholangiocarcinoma can arise from Notch-mediated conversion of hepatocytes. The Journal of clinical investigation 2012;122:3914-3918.

36 Fan B, Malato Y, Calvisi DF, Naqvi S, Razumilava N, Ribback S, Gores GJ, Dombrowski F, Evert M, Chen X, Willenbring H: Cholangiocarcinomas can originate from hepatocytes in mice. The Journal of clinical investigation 2012;122:2911-2915.

37 Bosman FT, World Health Organization., International Agency for Research on Cancer.: WHO classification of tumours of the digestive system, ed 4th. Lyon, International Agency for Research on Cancer,

2010.

38 Kawai T, Yasuchika K, Ishii T, Katayama H, Yoshitoshi EY, Ogiso S, Kita S, Yasuda K, Fukumitsu K, Mizumoto M, Hatano E, Uemoto S: Keratin 19, a Cancer Stem Cell Marker in Human Hepatocellular Carcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research 2015;21:3081-3091.

39 Zhang C, Xu Y, Zhao J, Fan L, Jiang G, Li R, Ling Y, Wu M, Wei L: Elevated expression of the stem cell marker CD133 associated with Line-1 demethylation in hepatocellular carcinoma. Annals of surgical oncology 2011;18:2373-2380.

Figure Legends

Figure 1. Methylation and expression analyses of K19. (**a**) Schematic representation of the location of discrete *KRT19* gene promoter regions and the results of *KRT19* methylation assessment by highly-sensitive fluorescence assay for bisulfite DNA (Hi-SA). The gray squares denote the coding exon regions of the *KRT19* gene. The blue and green squares represent the restriction sites for HhaI; vertical lines indicate CpG sites; thick horizontal blue and green lines represent polymerase chain reaction fragments; arrows represent primers; and black arrows point out methylated alleles cleaved by the restriction enzyme. (**b**) Cloning and sequencing of *KRT19* regions 1 and 2. Polymerase chain reaction products that were amplified by primer sets for bisulfite cloning were cloned into a TOPO cloning vector and sequenced. For the two cell lines, at least 12 clones were sequenced. Empty circles indicate unmethylated CpG sites. Filled circles represent methylated CpG sites. (**c**) Expression of K19 protein in cell lines and methylation rates in the *KRT19* gene and *LINE-1*. (**d**) Expression ratio (log2 ratio) of *KRT19* messenger RNA in the five cell lines (HepG2, HuH7, HLE, HLF, and PLC/PRF/5). Expression ratio denotes log2 ratio obtained from the signal intensity of *KRT19* messenger RNA from the cells treated with 5-aza-dC and TSA divided by the signal intensity of *KRT19* messenger RNA from the cells untreated by a microarray analysis.

Figure 2. STROBE diagram of the patient cohort.

Figure 3. Immunohistochemistry (IHC) for K19 and clinical outcomes of 125 HCCs. (**a**) Examples of IHC staining of K19 in HCCs classified into five groups according to the percentage of tumor cells with positive staining. (**b**) Kaplan-Meier survival curves for disease-free survival (DFS), overall survival (OS), and extrahepatic metastasis (EHM) recurrence-free survival (EFS) according to K19 expression status.

Figure 4. Methylation analyses of *KRT 19* and *LINE-1* in 125 HCCs. (**a**) Results of methylation analysis in *KRT19* region 1. Average methylation level (left panel) and methylation frequency (right panel) of *KRT19*

region 1 in HCCs with or without K19 proficiency. (**b**) Results of methylation analysis in *KRT19* region 2. Average methylation level (left panel) and methylation frequency (right panel) of *KRT19* region 2 in HCCs with or without K19 proficiency. (**c**) Results of methylation analysis in *LINE-1*. Average methylation level (left panel) and methylation frequency (right panel) of *LINE-1* in HCCs with or without K19 proficiency. Left panel; in the box plot diagrams, the horizontal line within each box represents the median; the limits of each box are the interquartile ranges, and the whiskers are the maximum and minimum values. The *P* values were based on Kruskal-Wallis 1-way analysis of variance on ranks and represent the statistical differences in average methylation between K19-deficient and proficient HCCs. Right panel; the numbers in each box (right panel) denote the number of cases. The *P* values were based Fisher's exact test. (**d**) Scatter-plot matrix demonstrating the pairwise Spearman's correlation coefficient [ρ] between three analyzed loci in the cohort of 125 HCCs. The Y and X-axes denote methylation rates.

Figure 5. Correlation between six markers examined by immunohistochemical (IHC) staining in125 HCCs. (A,B) Scatter-plot matrices demonstrating the pairwise Spearman's correlation coefficient [ρ] between six analyzed markers in the cohort of 96 K19-deficient HCCs (**a**) and of 29 K19-proficient HCCs (**b**). The Y and X-axes denote IHC score.

Supplementary files

Supplementary Table 1. Primer Sequences.

Supplementary Table 2. Patient's characteristics.

Supplementary Fig. S1. Examples of *KRT19* methylation assessment by highly-sensitive fluorescence assay for bisulfite DNA (Hi-SA).

Supplementary Fig. S2. Demethylation analysis of *KRT19* in HLF cell lines. Change of methylation rate in the *KRT19* promoter (A) and expression analysis of K19 (B) 5 days after treatment with various concentrations of the 5-aza-2-dC agent alone.

Supplementary Fig. S3. Methylation (HM450) beta-values for genes in 442 HCC samples. For genes with multiple methylation probes, the probe most anti-correlated with expression.

Supplementary Fig. S4. Immunohistochemistry (IHC) for K7, NOTCH-1, vimentin, E-cadherin, arginase-1, and HepPar-1. Examples of IHC staining of the six markers in HCCs classified into five groups according to the percentage of tumor cells with positive staining. (A) by high power focus (B) by low power focus..

Table 1 Comparison of Clinicopathologic Features Between CK19-proficient and deficient HCC

Variable	All (n=125)	K19-deficient HCC (n = 96)	K19-proficient HCC (n = 29)	*P* value*
Age – y, mean (SD)	65.1 (9.9)	66.2 (8.76)	61.3 (12.6)	0.020
Sex – no. (%)				
Female	31 (25)	19 (20)	12 (41)	0,027
Male	94 (75)	77 (80)	17 (59)	
ECOG performance status – no. (%)				
0	114 (91)	87 (91)	27 (93)	1.0
1-2	11 (9)	9 (9)	2 (7)	
HCV Antibody – no. (%)				
Positive	61 (50)	48 (52)	13 (46)	0,67
Negative	60 (50)	45 (48)	15 (54)	
Missing	4	3	1	
HBs Antigen – no. (%)				
Positive	32 (27)	22 (24)	10 (36)	0,33
Negative	86 (73)	68 (76)	18 (64)	
Missing	7	6	1	
AST (IU/l) – no. (%)				
≥ 60	37 (31)	26 (28)	11 (41)	0,24
< 60	84 (69)	68 (72)	16 (59)	
Missing	4	2	2	
Total bilirubin (mg/dl) – no. (%)				
≥ 1.5	9 (7)	5 (5)	4 (15)	0,11
< 1.5	112 (93)	89 (95)	23 (85)	
Missing	4	2	2	
Alpha-fetoprotein (ng/ml) – no. (%)				
≥ 1000	14 (12)	7 (8)	7 (26)	0,021
< 1000	100 (88)	80 (92)	20 (74)	
Missing	11	9	2	
PIVKA-II (mAU/ml) – no. (%)				
≥ 200	52 (46)	38 (44)	14 (52)	0,51
< 200	62 (54)	49 (56)	13 (48)	
Missing	11	9	2	
TNM stage – no. (%)				
I	71 (57)	58 (60)	13 (45)	0,31
II	42 (34)	30 (31)	12 (41)	
IIIA	12 (10)	8 (8)	4 (14)	
Tumor size (cm) – no. (%)				
≥5	35 (28)	23 (24)	12 (41)	0.10
<5	90 (72)	73 (76)	17 (59)	
Number of tumor – no. (%)				
Multiple	31 (25)	23 (24)	8 (28)	0,81
Single	94 (75)	73 (76)	21 (72)	
Histology – no. (%)				
Poorly	15 (12)	9 (9)	6 (21)	0,11
Well/Moderately	110 (88)	87 (91)	23 (79)	
Microvascular invasion – no. (%)				
Present	37 (30)	23 (24)	14 (48)	0,019
Absent	88 (70)	73 (76)	15 (52)	
Fibrosis stage † – no. (%)				
F0-3	85 (68)	65 (68)	20 (69)	1.0
F4	40 (32)	31 (32)	9 (31)	

* ANOVA or Fisher exact test.
† New Inuyama classification which assesses the degree of fibrosis ranging from F0 (no fibrosis) to F4 (cirrhosis).
Abbreviations: SD, standard deviation; HCV, hepatitis C virus; HBs, hepatitis B surface; AFP, alpha-fetoprotein.

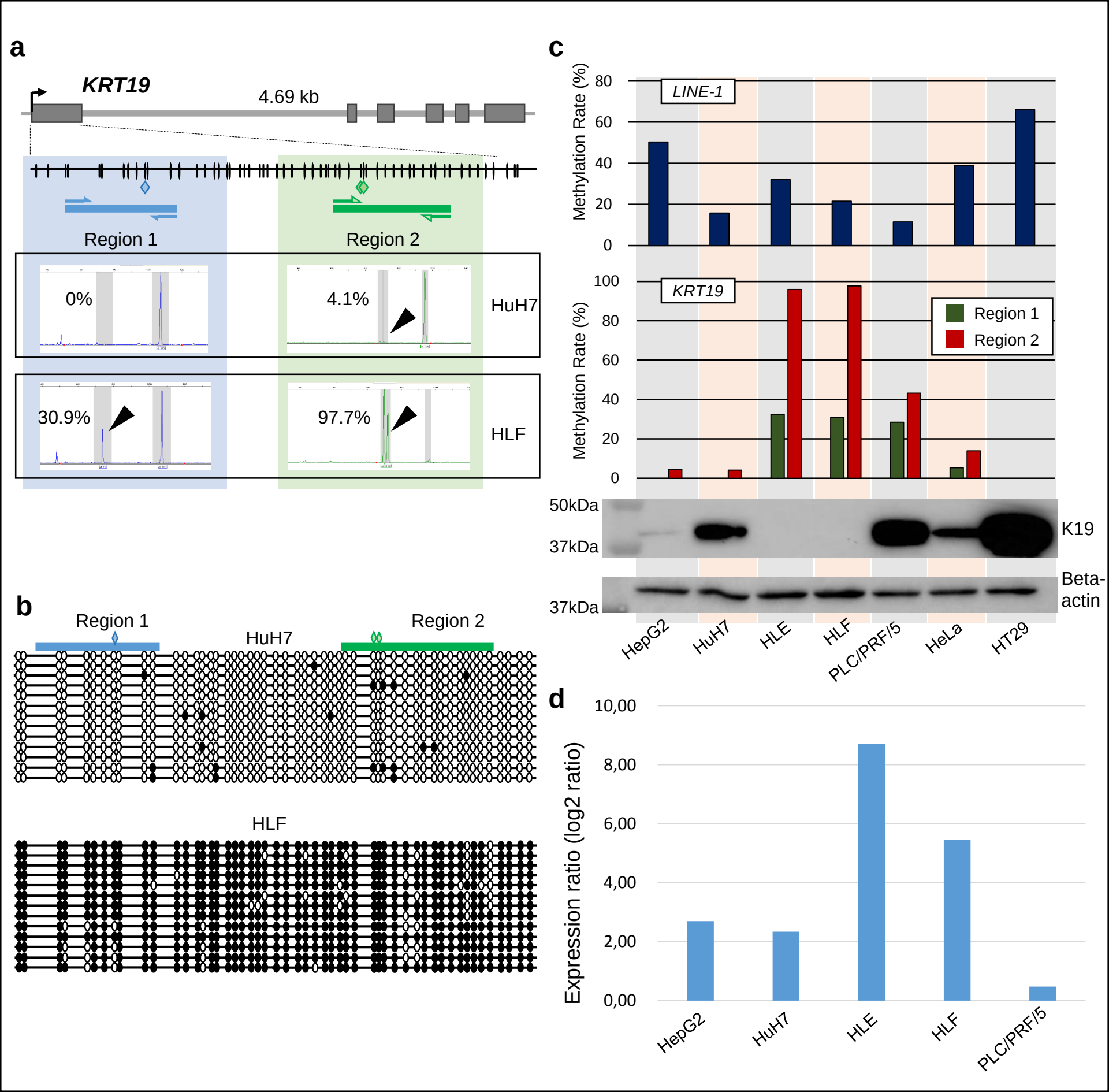
Table 2 Univariate and multivariate analysis for overall survival

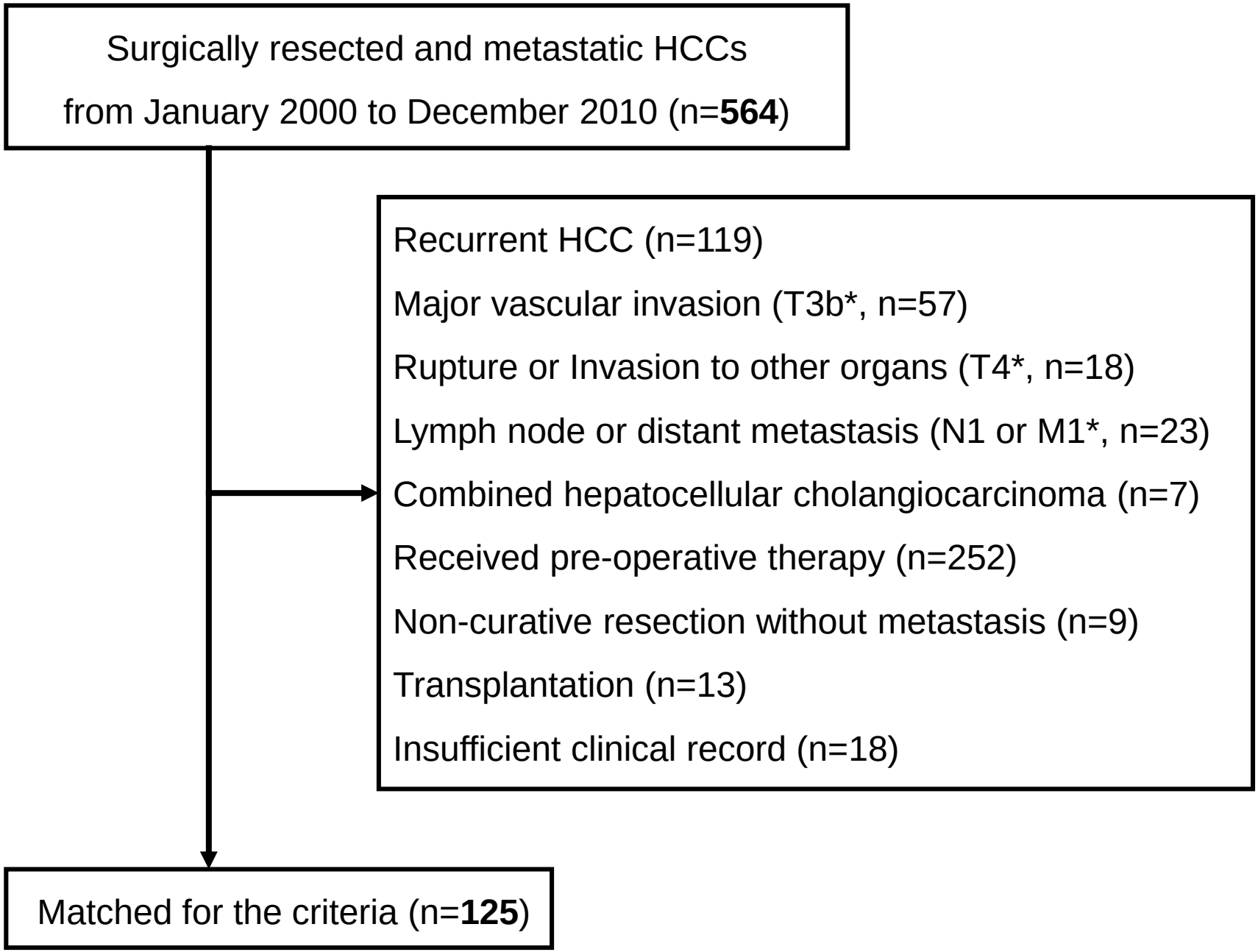
Variables	Univariate analysis*			Multivariate analysis*		
	HR	95% CI	*P* value	HR	95% CI	*P* value
K19 proficiency (vs K19 deficiency)	2,1	1.05 – 4.18	0,036	2,9	1.21 – 6.41	0,018
Age ≥ 65 (vs < 65)	1,6	0.83 – 3.37	0,16	1,7	0.75 – 4.00	0,21
Female (vs Male)	1,3	0.62 – 3.03	0,52	1,5	0.60 – 4.06	0,41
ECOG PS ≥ 1 (vs PS 0)	3,7	1.07 – 9.65	0,04	2.70	0.66 – 8.96	0,15
AFP ≥ 1000 (vs <1000)	0,7	0.16 – 1.90	0.50	0.50	0.10 – 1.84	0,31
PIVKA-II ≥ 200 (vs <200)	1,5	0.74 – 2.93	0,27	1	0.45 – 2.35	0,92
Number of tumor, multiple (vs single)	4,2	2.15 – 8.16	< 0.0001	3	1.26 – 6.87	0,013
Tumor size ≥ 5cm (vs <5cm)	1,5	0.72 – 3.06	0,25	1,5	0.61 – 3.38	0,37
Poor Differentiation (vs others)	1,1	0.38 – 2.63	0,82	0,4	0.10 – 1.05	0,063
Microvascular invasion, positive (vs negative)	2.70	1.36 – 5.23	0,0051	2,9	1.27 – 6.65	0,012
Fibrosis stage F4† (vs F0-3)	2	1.03 – 3.88	0.040	2,1	0.94 – 4.55	0.070

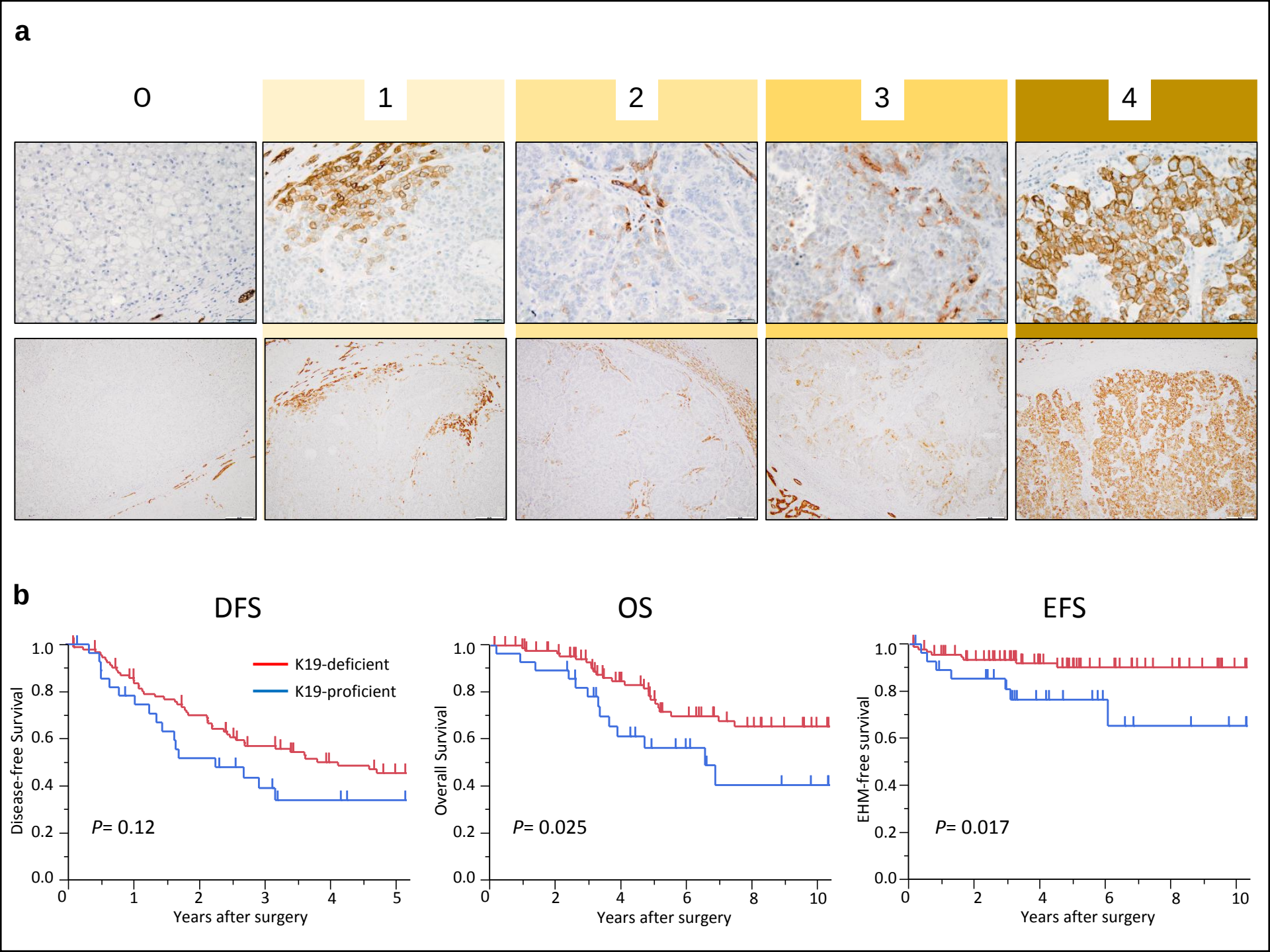
* Cox proportional hazard model.
† New Inuyama classification which assesses the degree of fibrosis ranging from F0 (no fibrosis) to F4 (cirrhosis).
Abbreviations: HR, hazard ratio; CI, confidence interval; ECOG, Eastern Cooperative Oncology Group; PS, performance status; AFP, alpha-fetoprotein; PIVKA-II, protein induced by vitamin K absence-II.

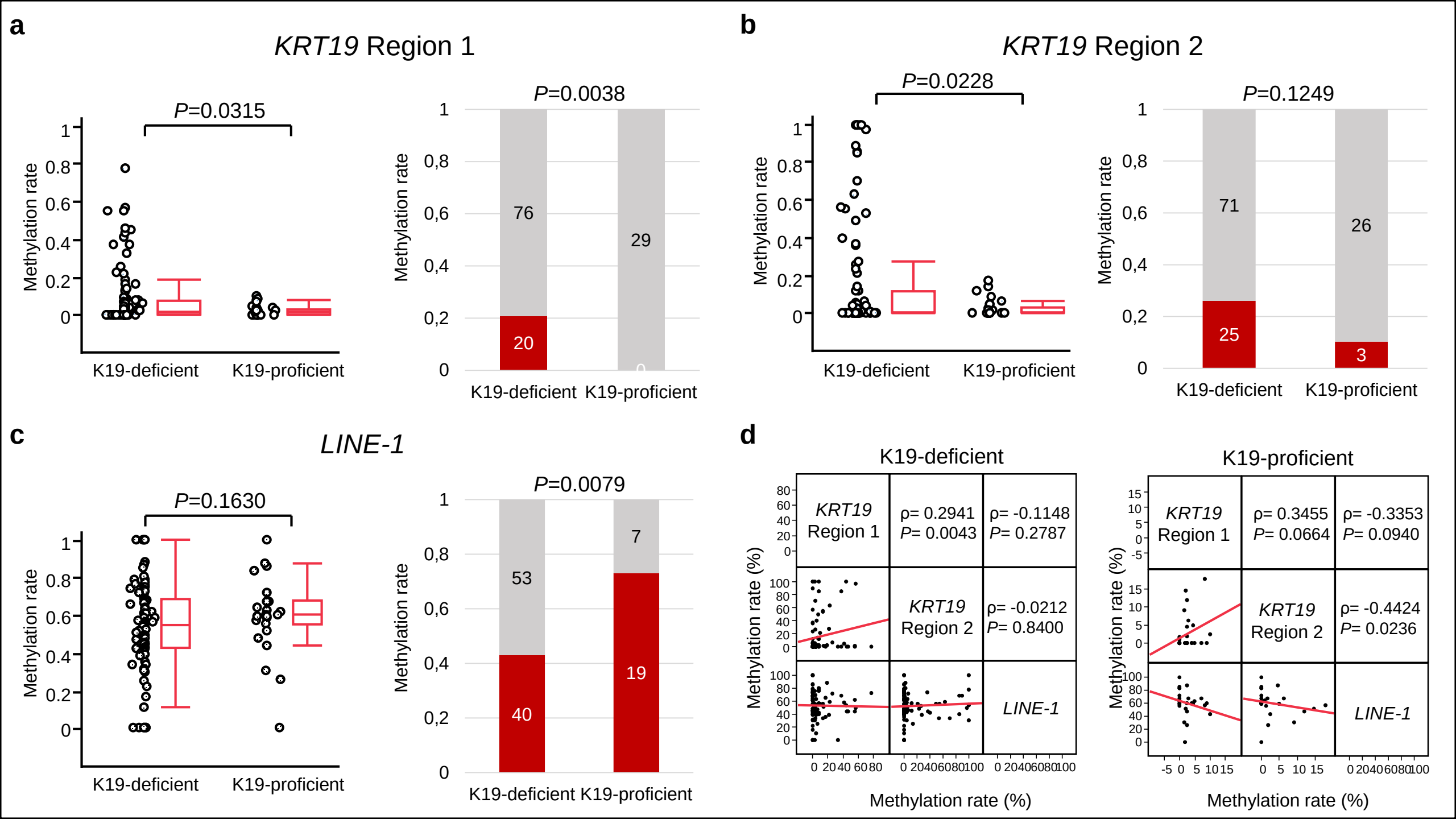
Table 3. Association between K19 expression status and other markers

Variable		All (n=125) no. (%)	K19-deficient HCC (n=96) no. (%)	K19-proficient HCC (n=29) no. (%)	*P* value
E-cadherin					
	Proficient	106 (85)	85 (89)	21 (72)	0.043
	Deficient	19 (15)	11(12)	8 (28)	
Vimentin					
	Proficient	8 (6)	4 (4)	4 (14)	0.084
	Deficient	117 (94)	92 (96)	25 (86)	
K7					
	Proficient	43 (34)	28 (29)	15 (52)	0.043
	Deficient	82 (66)	68 (71)	14 (48)	
NOTCH-1					
	Proficient	72 (58)	55 (57)	17 (59)	1.00
	Deficient	53 (42)	41 (43)	12 (41)	
HepPar-1					
	Proficient	90 (72)	74 (77)	16 (55)	0.033
	Deficient	35 (28)	22 (23)	13 (45)	
Arginase-1					
	Proficient	87 (70)	76 (79)	11 (34)	<0.0001
	Deficient	38 (30)	20 (21)	18 (66)	
LINE-1 methylation ratio					
	Mean % (95%CI)		53.3 (48.6-58.1)	60.3 (52.0-68.6)	
	55% or more	59 (50)	40 (43)	19 (73)	0.0079*
	under 55%	60 (50)	53 (57)	7 (27)	
	Not analysed	6	3	3	

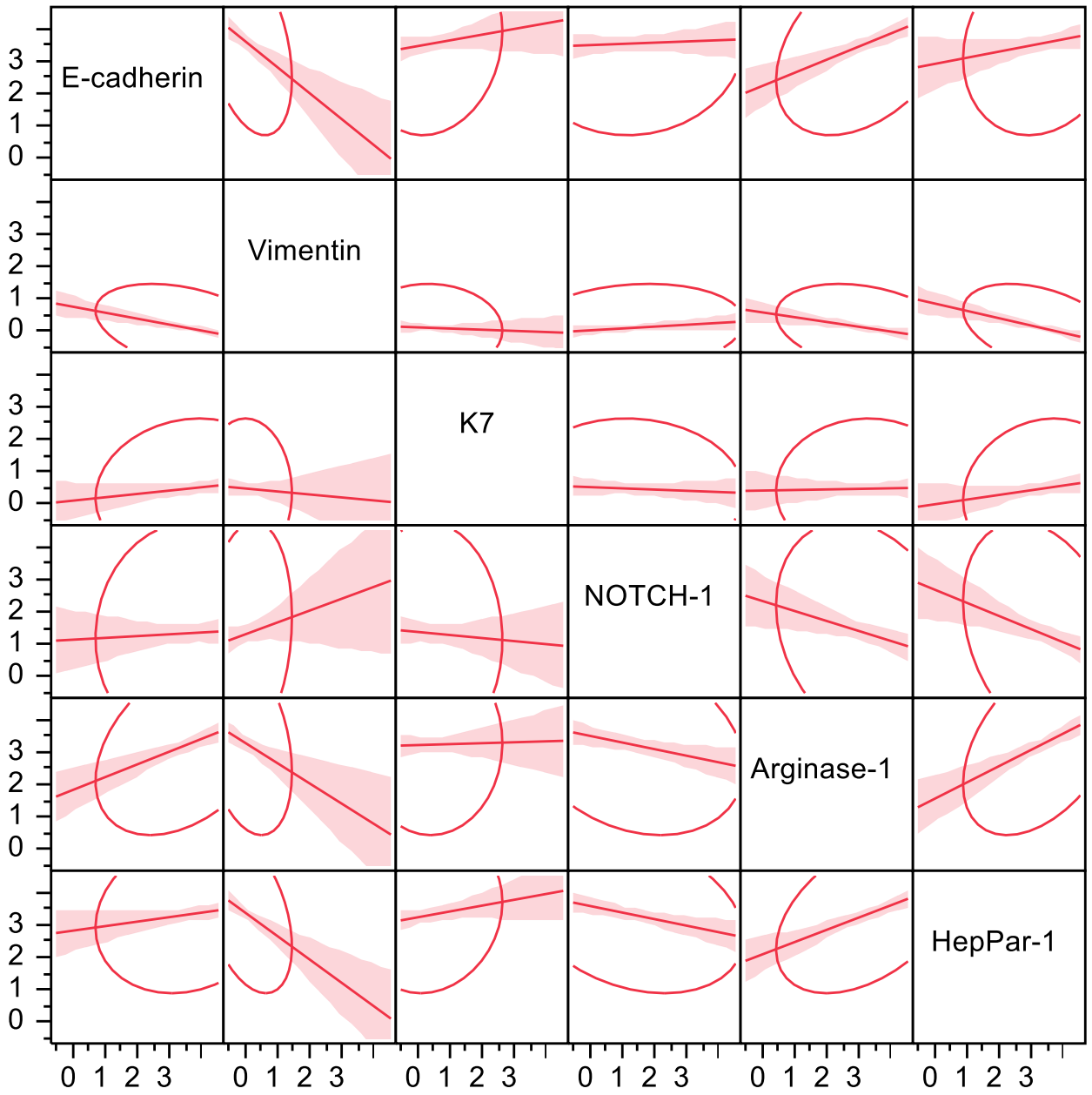

a
KRT19
4.69 kb
Region 1
Region 2
0%
4.1%
HuH7
30.9%
97.7%
HLF
b
Region 1
Region 2
HuH7
HLF
c
LINE-1
Methylation Rate (%)
80
60
40
20
0
KRT19
Methylation Rate (%)
100
80
60
40
20
0
Region 1
Region 2
50kDa
37kDa
K19
37kDa
Beta-actin
HepG2
HuH7
HLE
HLF
PLC/PRF/5
HeLa
HT29
d
Expression ratio (log2 ratio)
10,00
8,00
6,00
4,00
2,00
0,00
HepG2
HuH7
HLE
HLF
PLC/PRF/5


Surgically resected and metastatic HCCs
from January 2000 to December 2010 (n=**564**)
Recurrent HCC (n=119)
Major vascular invasion (T3b*, n=57)
Rupture or Invasion to other organs (T4*, n=18)
Lymph node or distant metastasis (N1 or M1*, n=23)
Combined hepatocellular cholangiocarcinoma (n=7)
Received pre-operative therapy (n=252)
Non-curative resection without metastasis (n=9)
Transplantation (n=13)
Insufficient clinical record (n=18)
Matched for the criteria (n=**125**)


a
0
1
2
3
4
b
DFS
K19-deficient
K19-proficient
Disease-free Survival
P= 0.12
Years after surgery
OS
Overall Survival
P= 0.025
Years after surgery
EFS
EHM-free survival
P= 0.017
Years after surgery

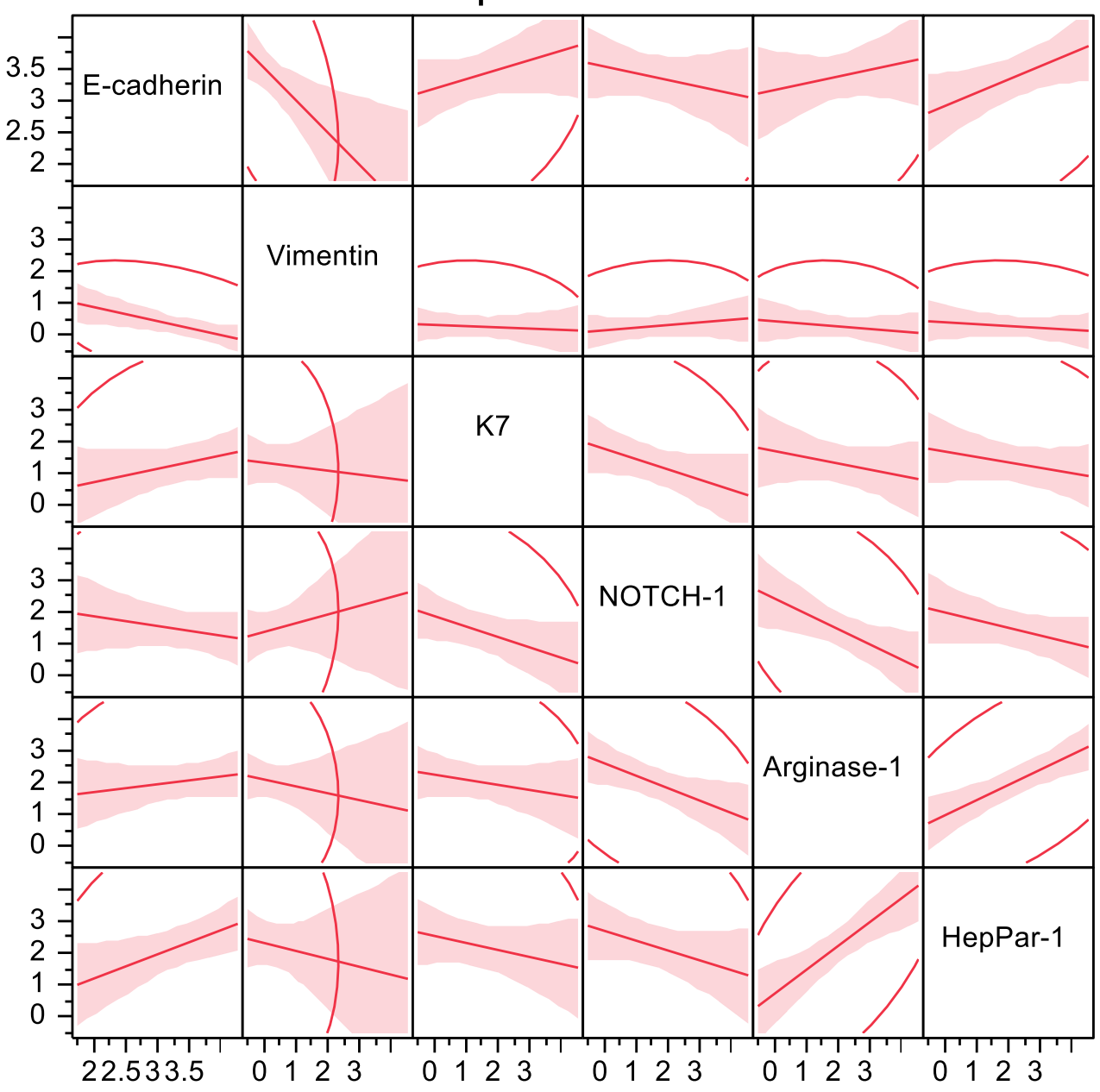

a
KRT19 Region 1
P=0.0315
Methylation rate
K19-deficient
K19-proficient
P=0.0038
76
20
29
0
b
KRT19 Region 2
P=0.0228
P=0.1249
71
25
26
3
c
LINE-1
P=0.1630
P=0.0079
53
40
7
19
d
K19-deficient
K19-proficient
KRT19 Region 1
KRT19 Region 2
LINE-1
ρ= 0.2941
P= 0.0043
ρ= -0.1148
P= 0.2787
ρ= -0.0212
P= 0.8400
ρ= 0.3455
P= 0.0664
ρ= -0.3353
P= 0.0940
ρ= -0.4424
P= 0.0236
Methylation rate (%)

a



Variable	Spearman (ρ)	*P* value
HepPar-1 vs Vimentin	**-0.3248**	**0.0012**
HepPar-1 vs NOTCH-1	**-0.2979**	**0.0032**
Arginase-1 vs Vimentin	**-0.2787**	**0.006**
Arginase-1 vs NOTCH-1	**-0.274**	**0.0069**
Vimentin vs E-cadherin	**-0.2128**	**0.0374**
K7 vs Vimentin	-0.0317	0.7594
NOTCH-1 vs K7	-0.0317	0.7593
NOTCH-1 vs E-cadherin	-0.0091	0.9298
Arginase-1 vs K7	0.0505	0.6251
HepPar-1 vs E-cadherin	0.0839	0.4162
K7 vs E-cadherin	0.1158	0.2612
HepPar-1 vs K7	0.197	0.0544
NOTCH-1 vs Vimentin	0.1988	0.0521
Arginase-1 vs E-cadherin	**0.3038**	**0.0026**
HepPar-1 vs Arginase-1	**0.3794**	**0.0001**

b



Variable	Spearman (ρ)	*P* value
HepPar-1 vs Vimentin	-0.2951	0.1202
HepPar-1 vs NOTCH-1	-0.2651	0.1646
Arginase-1 vs Vimentin	-0.1913	0.3201
Arginase-1 vs NOTCH-1	**-0.4079**	**0.028**
Vimentin vs E-cadherin	**-0.5325**	**0.0029**
K7 vs Vimentin	-0.0978	0.6139
NOTCH-1 vs K7	-0.2944	0.1211
NOTCH-1 vs E-cadherin	-0.136	0.4819
Arginase-1 vs K7	-0.2944	0.1211
HepPar-1 vs E-cadherin	-0.1873	0.3306
K7 vs E-cadherin	0.2843	0.135
HepPar-1 vs K7	-0.2048	0.2866
NOTCH-1 vs Vimentin	0.3058	0.1066
Arginase-1 vs E-cadherin	0.151	0.4343
HepPar-1 vs Arginase-1	**0.5981**	**0.0006**

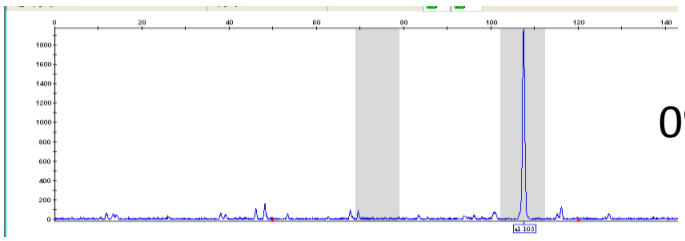
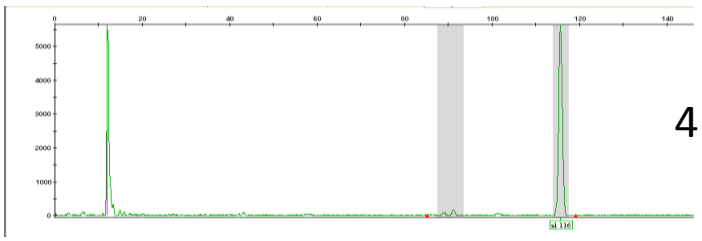
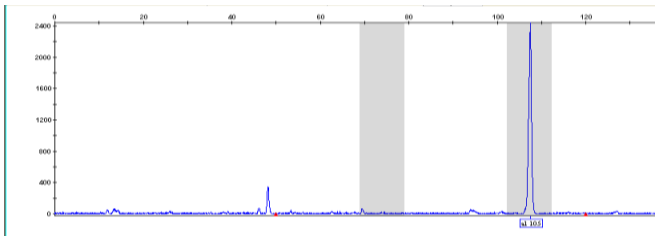
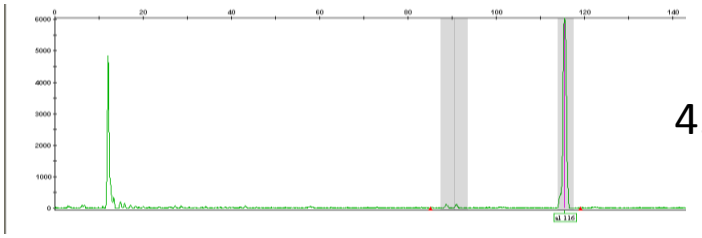
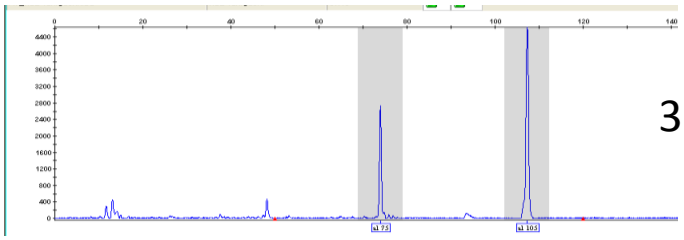
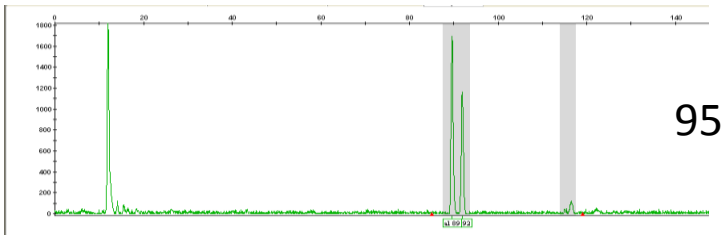
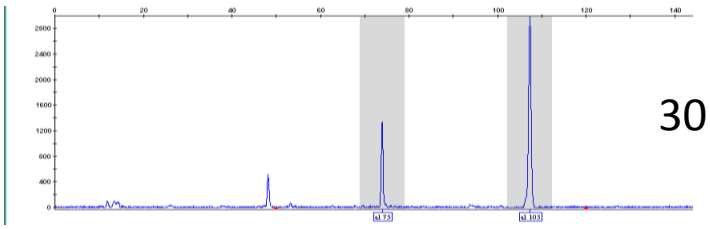
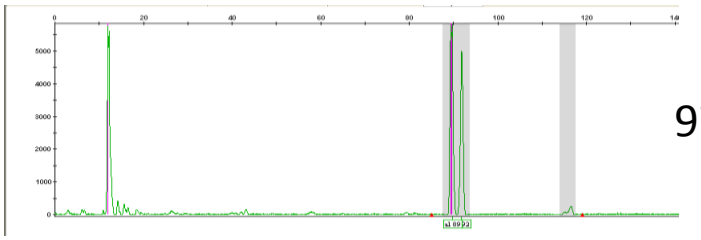
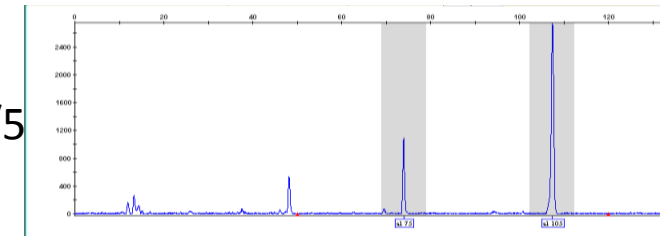
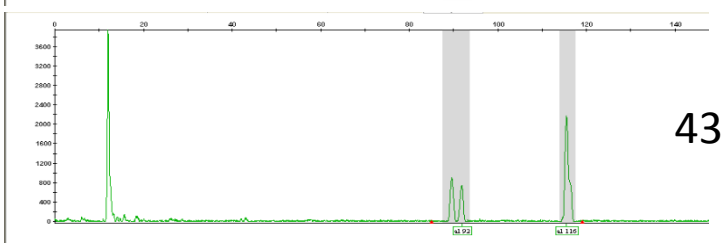
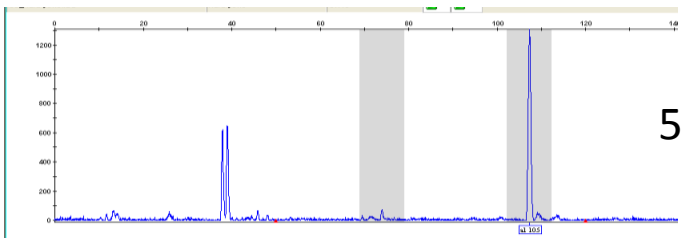
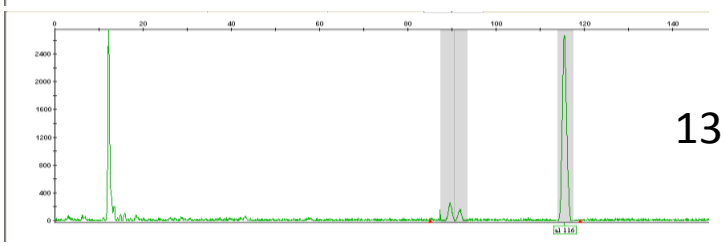
SupplementaryTable 1. Primer Sequences

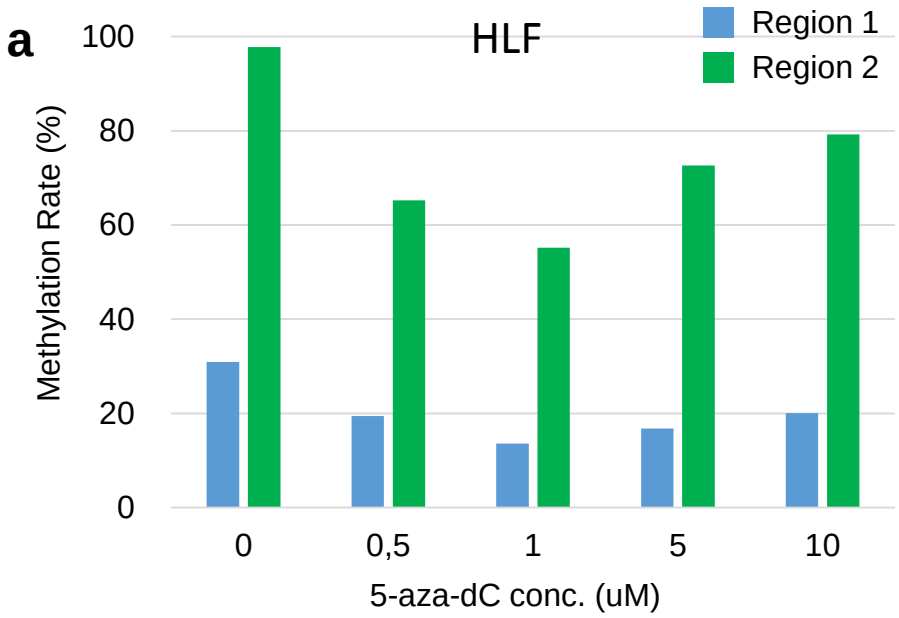
Name	Sequence	Product Size
K19-Region1-F	FAM- GGGAGGGTTTAGGTTTTTGTTTTT	108bp
K19-Region1-R	ACCRCCTAACCTCCTACCTAAA	
K19-Region2-F	GTYGGAGTTTTYGTGAATGTTG	121bp
K19-Region2-R	VIC-AACTTCCTACAACTATCRCCAATC	
K19-cloning-F	GGGTTTTYGTTTATTTTGTTTYGT	694bp
K19-cloning-R	CRAATCRCAACTTCTAAAACCAA	
LINE1-F	GYGTAAGGGGTTAGGGAGTTTTT	160bp
LINE1-R	FAM-RTAAAACCCTCCRAACCAAATATAAA	

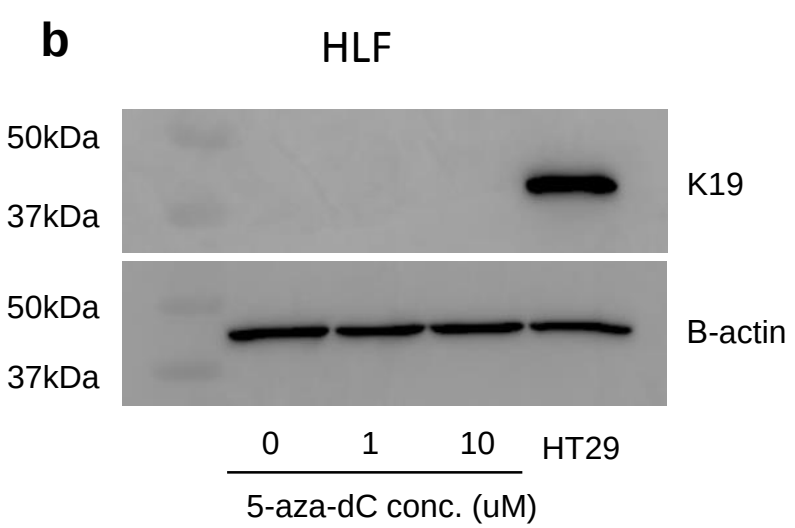
Supplementary Table 2 Patients characteristics

Variable		n=125
Age – y, mean ± SD		65.1±9.9
Sex – no. (%)		
	Female	31 (25)
	Male	94 (75)
EOCG performance status – no. (%)		
	0	114 (91)
	1	10 (8)
	2	1 (1)
Laboratory data		
	HCV Antibody positive – no. (%)	61 (50)
	HBs Antigen positive – no. (%)	32 (27)
	AST – IU/l, median (range)	43 (15-289)
	Albumin – g/dl, median (range)	4.0 (3.0-4.9)
	Total bilirubin - mg/dl, median (range)	0.73 (0.33-9.8)
	AFP - ng/ml, median (range)	14 (1.1-60054)
	PIVKA-II – mAU/ml, median (range)	104.5 (10 - 141380)
Child-Pugh class – no. (%)		
	A	120 (96)
	B	2 (2)
	missing	3
TNM stage – no. (%)		
	I	71 (57)
	II	42 (34)
	IIIA	12 (10)
Tumor size – cm, median (range)		3.2 (0.8–17.5)
Number of tumor – no. (%)		
	Single	94 (75)
	Multiple	31 (25)
Microvascular invasion – no. (%)		
	Absent	88 (70)
	Present	37 (30)
Histology– no. (%)		
	Well	22 (18)
	Moderately	88 (70)
	Poorly	15 (12)
Fibrosis stage – no. (%) *		
	F0	10 (8)
	F1	29 (23)
	F2	23 (18)
	F3	22 (18)
	F4	40 (32)

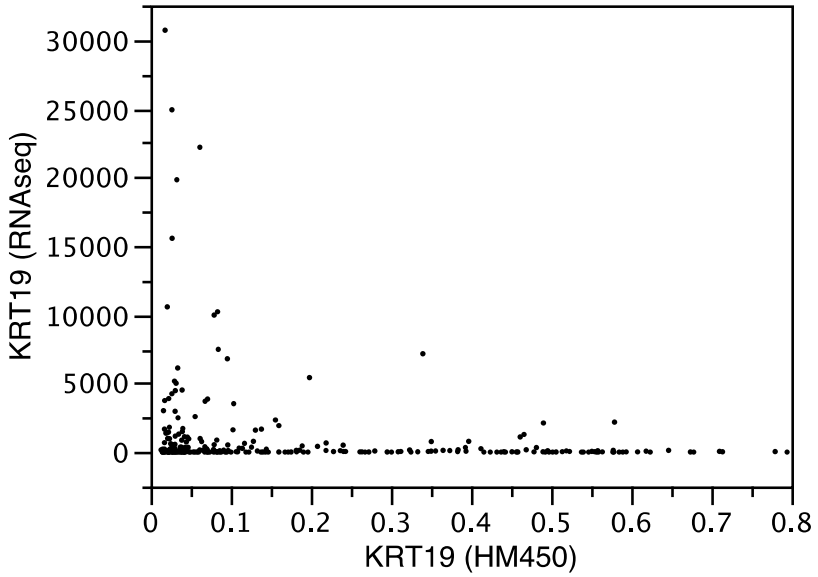
* New Inuyama classification which assesses the degree of fibrosis ranging from F0(no fibrosis) to F4(cirrhosis).
Abbreviations: SD, standard deviation; ECOG, Eastern Cooperative Oncology Group; HCV, hepatitis C virus; HBs, hepatitis B surface; AST, aspartate aminotransferase; AFP, alpha fetoprotein; PIVKA-II, protein induced by vitamin K absence-II.

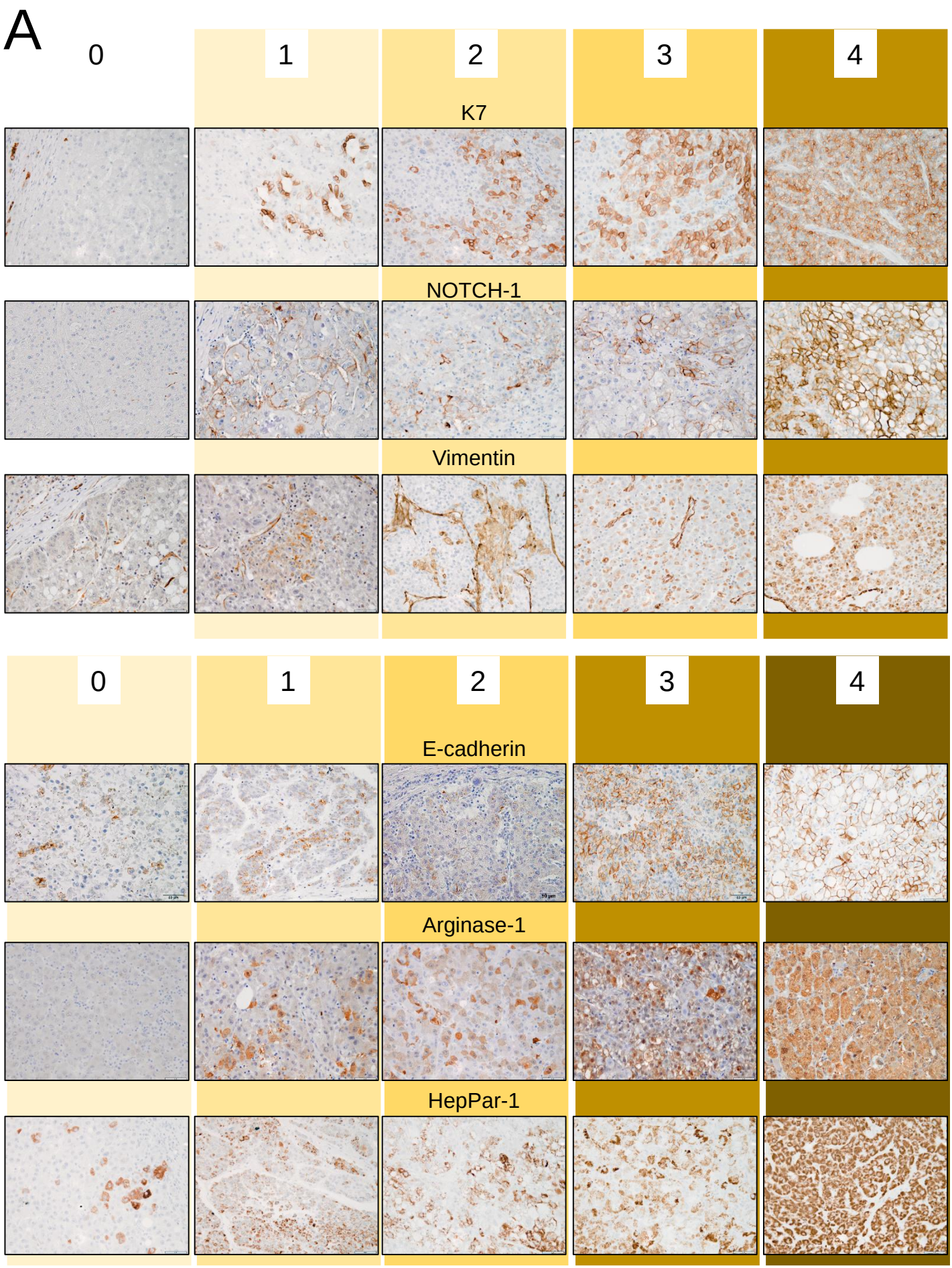
	Region 1		Region 2	
HepG2		0%		4.5%
HuH7		0%		4.1%
HLE		32.5%		95.9%
HLF		30.9%		97.7%
PLC/PRF/5		28.4%		43.2%
HeLa		5.3%		13.9%


a
HLF
Region 1
Region 2
Methylation Rate (%)
100
80
60
40
20
0
0
0,5
1
5
10
5-aza-dC conc. (uM)


b
HLF
50kDa
37kDa
K19
50kDa
37kDa
B-actin
0
1
10
HT29
5-aza-dC conc. (uM)

HCC (TCGA, Provisional)
442 samples




A
0
1
2
3
4
K7
NOTCH-1
Vimentin
0
1
2
3
4
E-cadherin
Arginase-1
HepPar-1


B
0
1
2
3
4
K7
NOTCH-1
Vimentin
0
1
2
3
4
E-cadherin
Arginase-1
HepPar-1