HHS Public Access

Author manuscript

Oncogene. Author manuscript; available in PMC 2016 March 24.

Published in final edited form as:

Oncogene. 2015 September 24; 34(39): 5095–5104. doi:10.1038/onc.2014.438.

Integrative Genomics Identifies YY1AP1 as an Oncogenic Driver in EpCAM+ AFP+ Hepatocellular Carcinoma

Xuelian Zhao[1,†], **Sonya Parpart**[1], **Atsushi Takai**[1], **Stephanie Roessler**[1,†], **Anuradha Budhu**[1], **Zhipeng Yu**[1], **Michael Blank**[2,†], **Ying E. Zhang**[2], **Hu-Liang Jia**[4], **Qing-Hai Ye**[4], **Lun-Xiu Qin**[4], **Zhao-You Tang**[4], **Snorri S. Thorgeirsson**[3], and **Xin Wei Wang**[1,*]

[1]Laboratory of Human Carcinogenesis, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892, USA

[2]Laboratory of Cellular and Molecular Biology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892, USA

[3]Laboratory of Experimental Carcinogenesis, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892, USA

[4]Liver Cancer Institute, Fudan University, Shanghai, China

Abstract

Identification of key drivers and new therapeutic targets is important given the poor prognosis for hepatocellular carcinoma (HCC) patients, particularly those ineligible for surgical resection or liver transplant. However, the approach to identify such driver genes is facing significant challenges due to the genomically heterogenous nature of HCC. Here, we tested whether the integrative genomic profiling of a well-defined HCC subset that is classified by an extreme $EpCAM^+$ AFP^+ gene expression signature and associated with poor prognosis, all attributes of a stem cell-like phenotype, could uncover survival-related driver genes in HCC. Following transcriptomic analysis of the well-defined HCC cases, a Gene Set Enrichment Analysis (GSEA) coupled with genomic copy number alteration assessment revealed that YY1-associated protein 1 (YY1AP1) is a critical oncoprotein specifically activated in $EpCAM^+$ AFP^+ HCC. YY1AP1 silencing eliminates oncogene addiction by altering the chromatin landscape and triggering massive apoptosis *in vitro* and tumor suppression *in vivo*. YY1AP1 expression promotes HCC proliferation and is required for the maintenance of stem cell features. We revealed that YY1AP1 cooperates with YY1 to alter the chromatin landscape and activate transcription of stemness regulators. Thus, YY1AP1 may serve as a key molecular target for $EpCAM^+$ AFP^+ HCC subtype. Our results demonstrate the feasibility and power of a new strategy by utilizing well-defined

*Correspondence: Dr. Xin Wei Wang, National Cancer Institute, Building 37, Room 3044A, 37 Convent Drive, Bethesda, MD 20892, USA. xw3u@nih.gov.

†Current addresses: XZ: Department of Molecular Pharmacology, Albert Einstein College of Medicine, Bronx, New York 10461; S.R.: Institute of Pathology, University Hospital Heidelberg, Germany; M.B.: Laboratory of Molecular and Cellular Cancer Biology, Faculty of Medicine, Bar-Ilan University, Safed 1321504 Israel.

Conflict of interest:

The authors declare no conflict of interest.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc)

patient samples and integrative genomics to uncover critical pathways linked to HCC subtypes with prognostic impact.

Keywords

YY1AP1; EpCAM$^+$ AFP$^+$; hepatocellular carcinoma; heterogeneity; somatic copy number alteration

Introduction

Similar to many other solid tumors, hepatocellular carcinoma (HCC), the major form of primary liver cancer, is considered incurable despite recent progress in diagnosis and the development of new treatment modalities[1]. A major contributing factor to poor outcome is the presence of extensive inter-tumor heterogeneity as revealed in numerous genome-wide investigations[2–9]. Genomic heterogeneity results in varying degrees of clinical presentation and tumor biology, which impedes treatment options and poses a significant challenge to cancer management. For example, nearly half of patients who are identified early and undergo surgical resection for small HCC (<5 cm) develop recurrent tumors[10]. Although targeted molecular therapies such as Sorafenib are promising new treatment modalities for a select group of patients with advanced HCC, their effectiveness has been suboptimal[11]. The overall improvement of cancer mortality is modest, and HCC remains one of the most difficult-to-cure malignancies, with a 5-year survival rate lingering at 16% in the United States[12]. Thus, the key to a fundamental improvement of patients' survival is to accurately identify homogeneous cancer subgroups and decipher the unique set of molecular events that drive their tumors.

Because human malignancies share a set of common cancer hallmarks, key genes responsible for each hallmark are, in principle, ideal druggable targets for targeted therapies[13]. However, recent genome-wide investigations indicate that each cancer type, including HCC, has a unique set of compromised genomic loci[14,15]. Theoretically, each tumor has a variation of altered cancer hallmarks that cause molecular networks to converge into a unique node essential for survival. However, it is a challenge to identify such a distinct molecular node essential for the growth of a particular type of tumor.

Accumulating evidence indicates that HCC heterogeneity can emanate from cancer stem cells (CSC), or through the selection of cells from different origins that reflect hepatic lineage, under the influence of various etiological factors such as hepatitis B or C virus (HBV or HCV) infection, alcohol consumption, or obesity[16,17]. We have previously revealed an aggressive subgroup of HCC that expresses high levels of EpCAM and AFP (EpCAM$^+$ AFP$^+$), which is associated with HCC metastasis and poor prognosis. This subgroup of HCC expresses stem cell-like gene expression traits and displays CSC characteristics such as self-renewal, lineage-specific differentiation, tumor initiation and high invasiveness, hence the name: hepatic stem cell-like HCC (HpSC HCC)[8,18–20]. Here, we coupled transcriptomic profiling of well-defined HCC specimens that represent extreme end points in the spectrum of cell lineage and tumor aggressiveness with gene set enrichment and somatic copy number analyses to identify novel oncogenic drivers

responsible for the EpCAM$^+$ AFP$^+$ HCC subgroup. Our study demonstrates that YY1AP1 transcriptionally regulates key stemness genes. The fact that its expression is essential for the maintenance of cancer cells makes YY1AP1 a potential therapeutic target for EpCAM$^+$ AFP$^+$ HCC.

Results

A unique transcriptomic profile in a well-defined subgroup of EpCAM$^+$ AFP$^+$ HCC

Using global transcriptomic analysis, we previously identified several HCC subgroups including a subgroup of aggressive HCC with stem cell-like traits and elevated levels of both EpCAM and AFP (EpCAM$^+$ AFP$^+$)[8]. However, heterogeneous expression of EpCAM and AFP is evident in clinical specimens as the number of HCC cells that express EpCAM and AFP proteins range from 5–70%[8,20]. We hypothesized that well-defined HCC subgroups representing the extreme ends of the marker expression spectra (dually positive or dually negative) may show homogeneity in their molecular changes and allow us to identify key driver genes in each HCC subgroup by employing an integrative genomic approach. To test this strategy, we reanalyzed publically available transcriptomic datasets (NCBI GEO accession number: GSE14520, GSE1898 and GSE4024) that contain gene expression data derived from two cohorts totally comprised of 380 HCC patients. Among the 241 HCC cases from the LHC cohort, we identified 38 *x*-HpSC HCC cases (i.e., >1000 ng/ml of serum AFP and in the top tertile of EpCAM expression) and 40 *x*-MH (mature hepatocyte-like subtype) HCC cases (i.e., <20 ng/ml serum AFP and in the bottom tertile of EpCAM expression) (Figure. 1 and Supplementary Table. S1).

Supervised class comparison analysis revealed that 1,573 statistically significant genes (p<0.001) could discriminate the *x*-HpSC and *x*-MH groups. Since *x*-HpSC tumors are associated with poor outcome, we reasoned that the above subgroup-related stem cell-like genes should be associated with patient survival. Survival risk prediction analysis showed that these genes were significantly associated with overall survival of the *x*-HpSC and *x*-MH cases (log-rank p=0.002; permutation p<0.01) (Figure. 1B). Using a log-rank test in the same patients (n=78), 932 survival-related genes were revealed to significantly link to overall survival, a majority of which overlap with the stem cell-like genes (hypergeometric p< 2.2E-16) (Figure. 1C). The association with survival of the identified signature was also validated in the remaining LHC cohort (n=163, log-rank p=0.016, permutation p=0.07) and an independent LEC cohort (n=139, log-rank p=0.0005, permutation p=0.02) (Supplementary Figure. S1). These results indicated that the *x*-HpSC and *x*-MH subgroups have a vast difference in gene expression, which is associated with HCC survival.

Integrative genomic analysis to define driver genes in EpCAM$^+$ AFP$^+$ HCC

To examine the critical signaling pathways associated with CSC features, we performed GSEA by comparing the gene expression profiles of *x*-HpSC and *x*-MH subgroups. Among 11 significantly enriched gene sets targeted by different transcriptional factors, three sets were annotated as transcriptional targets of transcription factor YY1 (yin yang 1) (NOM p<0.0001, FDR q=0.043) (Figure. 2A and Supplementary Table S2). Furthermore, we found both YY1 and the coactivator YY1AP1 are elevated in *x*-HpSC revealed by microarray data

(Figure. 2B and Supplementary Figure. S2A), suggesting a selective activation of YY1 transcriptional network in *x*-HpSC. Using arrayCGH data, We further analyzed the somatic copy number alterations (SCNA) of YY1 signaling related genes, YY1AP1 (1q22), YY1 (14q32) and YY2 (23p22.12-p21.3) (arrayCGH data accession number: GSE14322)[21]. We found the copy number of YY1AP1, but not YY1 or YY2, displayed a preferential gain frequency in HpSC cases (70%) compared to it does in MH case (24%) (Figure. 2C).

Amplification of chromosome 1q is common in human cancers and this region potentially contains oncogenes. By further investigating this region in the HpSC cases, we revealed 3 minimal commonly amplified regions, namely R1, R2 and R3, at chromosome 1q21.1-22 (Supplementary Figure. S2B), which harbor a total of 173 genes. We then searched for those genes that have correlated gene copy number and gene expression level using Spearman correlation analysis. Seven genes have significant correlation between copy number and their gene expression, but only one gene, YY1AP1, is significantly associated with patient survival determined by Cox regression analysis (Figure. 2D). We validated that YY1AP1 gene copy number obtained by quantitative copy number PCR can successfully predict patient survival in 76 HCC cases (p=0.033, $pTrend$=0.014) (Figure. 2E).

Oncogene addiction of YY1AP1 in HCC cells

The above results suggested YY1AP1 as a driver gene on which the growth of the EpCAM$^+$ AFP$^+$ HCC may depend on. However, by examining 20 different human normal tissues, we found that normal liver tissue had the lowest YY1AP1 expression, which suggested a functional silencing of YY1AP1 in normal liver (Supplementary Figure. S3A). Consistently, expression of YY1AP1 and GON4L, a paralog of YY1AP1[22], were also very low in either fetal or adult livers, suggesting that YY1AP1 may also not be essential in embryonic liver development (Supplementary Figure. S3B) and YY1AP1 activation might be an acquired event during pathogenesis of liver cancer. For further functional study of YY1AP1 in hepatocarcinogenesis, we screened 5 cell lines with different EpCAM and AFP status, including Huh 1 and Huh 7 (5–15% EpCAM$^+$ AFP$^+$), Hep3B and HepG2 (100% EpCAM$^+$ AFP$^+$, HepG2, a hepatoblastoma/HCC-like cell line), and SK-Hep-1 (dually negative). Among these cell lines, HepG2 expressed the highest protein level of YY1AP1 (Figure. 3A). Considering the dually positive EpCAM and AFP, which may most closely reflect the properties of EpCAM$^+$ AFP$^+$ HCC, HepG2 may be an ideal cell line model to investigate the oncogenic molecular basis for the hepatocarcinogenesis in YY1AP1-signaling altered HCC. Five shRNA constructs targeting YY1AP1 were tested in HepG2 cells. Consistent with the sequence specificity, shRNA C2 and C6 were able to knockdown YY1AP1 efficiently with no effect on GON4L, which shares 97% sequence identity with YY1AP1 and could potentially interfere with YY1AP1 (Figure. 3B and Supplementary Figure. S3C). Colony formation assays were carried out with shRNA C2 targeting YY1AP1. Consistent with the hypothesis, the viability of Hep3B and HepG2 (100% EpCAM$^+$ AFP$^+$) other than Huh1 and Huh7 (5–15% positive EpCAM and AFP) or SK-Hep-1 (dually negative EpCAM and AFP) cells exhibited oncogene addiction on YY1AP1 presence (Figure. 3C and Supplementary Figure. S3D). To rule out any off-target effect, we confirmed the effect with shRNA C6 that targets a different region of YY1AP1. As shown in Supplementary Figure. S3E, both shRNA C2 and C6 could inhibit cell growth dramatically. It is noted that Hep3B,

which displayed a cell survival dependence on YY1AP1, has a relatively low level of YY1AP1 despite its $EpCAM^+$ AFP^+ status like HepG2. To clarify this survival dependence, we overexpressed YY1AP1 in Hep3B and assessed the cell proliferation. The result showed that overexpression of YY1AP1 significantly increased the proliferation rate of Hep3B (Figure. 3D). These results showed that YY1AP1 expression plays a vital role for the survival of $EpCAM^+$ AFP^+ cells.

Since silencing of YY1AP1 has a lethal effect on $EpCAM^+$ AFP^+ cells immediately after shRNA was introduced, we developed a Doxycycline (Dox)-inducible shRNA-C2 cell line (referred as HepG2-Tet-C2) in which the expression of C2 and GFP (reporter for shRNA induction) were induced by addition of Dox (Figure. 4A). We observed that both clonogenecity (2D) and spheroid formation (3D) of HepG2-Tet-C2 cells were significantly inhibited upon shRNA-C2 induction, which demonstrated that the viability of HepG2 cells is YY1AP1-dependent (Figure.4B and 4C and Supplementary Figure. S4A). We also revealed that the cell death triggered by YY1AP1 was through apoptosis pathway since elevated cleaved caspase-3 is observed and pretreatment of cells with caspase inhibitor Z-VAD was able to block the cell death (Figure.4D and 4E and Supplementary Figure. S4B).

Impact of YY1AP1 in tumorigenicity

We next examined the tumorigenic potential of YY1AP1 in nude mice model. According to literatures, 5×10^6 HepG2 cells per site for tumor formation analysis in nude mice has average tumor latency in about 10 days, and as little as 5×10^5 HepG2 cells were able to form tumors within 4 weeks[23–25]. In our study, we injected 2×10^6 HepG2-Tet-C2 cells subcutaneously into flanks of athymic nude mice (9 animals, 18 injection sites per experimental group; 2 independent sets of experiments were set on 2 consecutive days). One group of mice (n=6) was fed with Dox diet immediately after the injection and maintained in the same diet for 3 weeks. The rest of the animals (n=12) were first kept on a control diet for 2 weeks. Two weeks later, after all animals developed measurable tumors, animals were split into two subgroups randomly and one subgroup (n=6) was switched to a Dox diet for YY1AP1 shRNA induction (Figure. 5A). We found that YY1AP1 silencing induced by Dox diet resulted in a significant decrease in both tumor incidence and tumor volume (Figure. 5B and Supplementary Figure. S5A). We also found that Dox treatment starting at day 14 after tumors reached a detectable size could significantly slow down the tumor growth compared to the control group (Supplementary Figure. S5B). Further histological analysis and TUNEL staining of the resected tumor samples revealed that tumor areas from Dox-fed mice were highly apoptotic as compared to tumors from the Dox-free group (Figure. 5C and Supplementary Figure S5C). Thus, tumor initiation and maintenance of HepG2 cells in nude mice depend on the presence of YY1AP1.

Role of YY1AP1 in regulating the global chromatin landscape

As evidenced in the spheroid formation assay, YY1AP1 silencing led to reduced EpCAM expression and spheroid formation (Figure. 6A). According to the literature, YY1AP1 functionally acts as an coactivator of YY1[26]. YY1, a ubiquitously expressed transcription factor, functions either as a transcriptional activator or repressor depending on its interactors and specific sites on regions of chromatin that contain modified histones[27]. We

hypothesized that YY1AP1 may be involved in epigenetic regulation of HCC transcriptional networks. One clue is from the observation that multiple genes with unrelated promoters such as EpCAM, Nanog and OSKM genes (i.e., Oct-4, Sox-2, Klf4, c-Myc) showed decreased expression upon YY1AP1 knockdown (Figure. 6B and Supplementary Figure. S6A). Consistent with HepG2 cells, a similar result was obtained by examining Hep3B cells. Knockdown of YY1AP1 resulted in decreased c-Myc, Nanog, AFP and EpCAM expressions, but no change in two differentiation genes UGT2B7 and CYP3A4 (Supplementary Figure. S6B). To determine if YY1AP1 is involved in histone modification, a correlation analysis was first carried out between multiple histone modifiers and YY1AP1 using EpCAM as a control. We found that several HDMs, HMTs and HDACs showed a significant positive linear relationship with YY1AP1 (Supplementary Table S3). We then determined post-translational modifications of all 4 core histones (H2A, H2B, H3 and H4) of HepG2-Tet-C2 cells with or without Dox treatment. We found specific downregulations of H2B ubiquitination, H3K4 trimethylation, H3K79 dimethylation and H3K36 dimethylation upon YY1AP1 depletion (Figure. 6C). Since H2B ubiquitination, H3K4 trimethylation and H3K79 dimethylation occur as tandem events during transcriptional initiation[28–31] and H3K36 dimethylation is linked to transcriptional elongation[32], it appears that YY1AP1 is linked to both transcriptional initiation and elongation. In contrast, we did not observe any change in histone acetylation (Supplementary Figure. S6C). To determine if histone demethylases (HDMs) are functionally linked to YY1AP1, we treated cells with HDM inhibitors that target various members of the HDM family[33]. We found that all 3 demethylase inhibitors, 2,4-PDCA, Daminozide and tranylcypromine, were able to partially rescue cell death upon YY1AP1 knockdown (Supplementary Figure. S6D). These results indicate that YY1AP1 is involved in epigenetic regulation of gene expression and HDMs contribute to YY1AP1 transcriptional networks in maintaining the viability of $EpCAM^+$ AFP^+ cells.

By searching the promoter region of EpCAM, we found 4 putative YY1 binding sites. Thus, we tested whether YY1 functionally interacts with YY1AP1 to regulate its expression. An *in vivo* chromatin immunoprecipitation (ChIP) assay and luciferase assay were carried out. We found that endogenous YY1 binds to an EpCAM promoter and its binding is diminished upon YY1AP1 knockdown (Figure. 6D and Supplementary Table. S4). Consistent data were obtained when an H3K4Me3 ChIP analysis was performed (Supplementary Figure S7). These results indicate that the binding of YY1 to an EpCAM promoter is facilitated by YY1AP1. To determine if YY1 and YY1AP1 regulates EpCAM expression, a luciferase reporter under the control of an EpCAM promoter was used. Consistently, YY1AP1 silencing resulted in suppression of increased EpCAM promoter activity induced by ectopic expression of YY1 (Figure. 6E). *In vivo* interaction between YY1 and YY1AP1 protein can be observed using a Duolink assay (Figure. 6F). To investigate if the YY1AP1-YY1 link in regulating EpCAM expression could also be observed in other HCC cells, we carried out the EpCAM promoter luciferase assays in Huh-7 cells. We found that EpCAM expression in Huh-7 was also regulated by YY1 and YY1AP1 expression (Supplementary Figure. S8). Taken together, our results indicate that YY1AP1 cooperates with YY1 to modulate gene transcription.

Discussion

Considering the vast inter-tumor heterogeneity in solid tumors, one of the major goals in cancer research is to identify specific druggable cancer driver genes whose function are essential for the fitness of cancer cells within a defined tumor subgroup. It is known that genome-wide investigations may be effective in defining tumor subgroups since all cancers arise as a result of somatically acquired changes in the DNA of cancer cells[34]. However, as a result of tumor evolution, each solid tumor carries hundreds and thousands of somatic genome alterations accumulated over time as documented in the COSMIC database and elsewhere, and a majority of mutations may be noncontributing passenger mutations whose functions are not essential to tumor cells at the time a tumor specimen is procured and analyzed[35,36]. The presence of considerable somatic changes found in solid tumors may explain the increased genomic intra-tumor heterogeneity[37]. Recently, integrative genomics, through the combination of exonic mutations and SCNA data, have shown promise in revealing candidate drivers linked to colorectal and lung cancer[38,39]. We also explored various integrative “omics” approaches to define liver tumor subgroups and to identify key functional genes unique to each subgroup[21,40,41]. These integrative approaches are designed to identify the “Achilles heel” of cancer as the basis to develop targeted cancer therapeutics.

In this study, we applied an integrated genomics strategy to a well-defined $EpCAM^+$ AFP^+ HCC subgroup with top-tier expression patterns of EpCAM and AFP that may represent a homogeneous group of patients with similar tumor biology. $EpCAM^+$ AFP^+ HCC is considered one of the most aggressive HCC subgroups which associated with aggressive tumor growth, metastasis, treatment resistance and poor prognosis[17]. Selecting well-defined tumor specimens may reduce biological heterogeneity and thus provide a better chance to identify key altered signaling pathways as druggable targets. Encouragingly, this strategy allowed us to successfully identify YY1AP1 whose functions are essential for the fitness of $EpCAM^+$ AFP^+ HCC cells. $EpCAM^+$ AFP^+ cells have developed a dependence on YY1AP1, a cancer phenomenon known as oncogene addiction[42]. Silencing of YY1AP1 in $EpCAM^+$ AFP^+ HCC cells resulted in a profound cell death phenotype. Interestingly, YY1AP1 is expressed at a very low level in normal liver cells, suggesting that YY1AP1 activation is an acquired event during HCC development. Moreover, silencing of YY1AP1 resulted in an alteration of the chromatin landscape along with a reduced expression of many CSC-related genes such as EpCAM, AFP, c-Myc, Sox2 and TCF4, whose activities are essential for the maintenance of HCC cells. Taken together, YY1AP1 represents a promising therapeutic target for $EpCAM^+$ AFP^+ HCC.

Our results are consistent with the literature describing that YY1AP1 may work together with YY1 to alter the chromatin landscape and regulate gene expression. It is an interesting observation that YY1AP1 suppression leads to ubH2B, H3K36Me2 and H3K4Me3 changes, suggesting a transcriptional activation mediated by YY1AP1. This could be a direct effect of YY1AP1 or an indirect effect from downregulation of other genes. More extensive biochemical experiments are needed to address whether and how YY1AP1 is involved in such epigenetic changes. In addition, we also found that YY1AP1 expression is positively correlated with HDM expression in HCC clinical specimens, and that YY1AP1 silencing mediated cell death, which could be partially blocked by HDM inhibitors, Our results are

consistent with the hypothesis that survival of EpCAM$^+$ AFP$^+$ HCC cells may have adapted to YY1AP1 activation, in which an elevation of YY1AP1 expression may promote the recruitment of YY1 and other histone modifying proteins such as HMT to the gene promoters containing YY1 binding sites. Such recruitment would result in alteration of the chromatin landscape by increasing histone H3K4 methylation thereby activating stem cell-related genes that are essential for maintaining HCC growth. Elimination of YY1AP1 may result in the removal of the PcG complex-mediated HMT activity, leading to hypo-methylation of H3K4 due to increased HDM activities in cancer cells. Such YY1AP1 dependence in cancer cells could trigger apoptosis upon withdrawal (see working model in Figure 7). The results above indicate that YY1AP1 functions as a driver gene in HCC. Though such an effect may be coupled with YY1, as many previous paradigmatic studies suggested, it is not limited to YY1 regulatory machinery. Furthermore, it is possible that YY1AP1 is also involved in regulating other chromatin regulatory complexes and acts as key player in driving HCC development.

In this study, we have revealed that EpCAM is transcriptionally regulated by YY1AP1, which is dependent on YY1 to bind to EpCAM promoter and activate its transcription. Previous studies have demonstrated that EpCAM is downstream gene of Wnt-β-catenin signaling and p53 (summarized in Figure. 3A)[19,43]. Based on our current data that not all HCC cells are sensitive to YY1AP1 silencing, we speculate that both YY1AP1 and β-catenin play roles in EpCAM expression when p53 signaling is silenced (wild type), which leads to YY1AP1 oncogene addition. However, when p53 signaling is activated (mutated), the YY1AP1 dependency may be compromised. Additional experiments may be needed to clarify how YY1AP1 works together with β-catenin and/or p53 in regulating EpCAM expression.

One limitation of this study is that most of the subsequent cell culture/mouse xenograft studies were carried out in a single hepatoblastoma/HCC-like cell line, HepG2 cells, which is likely to be closely related to the properties of EpCAM$^+$ AFP$^+$ primary liver cancer. We showed Hep3B, an EpCAM$^+$ AFP$^+$ HCC cell line, also displays YY1AP1 dependence in cell viability, which indicates that the profound lethal effect upon YY1AP1 silencing is not restricted to a particular cell line. The conclusion addressed from our mechanistic studies including EpCAM regulation and chromatin remodeling is applicable to the cell line tested before further investigation. Overall, our results indicate that YY1AP1 as an oncogenic driving force may serve as the ‘Achilles heel’ and a druggable molecular target in a subset of HCC with YY1AP1 activation.

Material and Methods

Clinical Specimens, Expression Microarray, ArrayCGH Platforms

We identified HCC samples as well as disease-free normal liver samples with available genomic and clinical data from our previously published studies[5,21,44]. A total of 380 HCC cases had available genomic data; 256 cases with Affymetrix U133A2.0 array data, 139 cases with LEC oligoarray data, and 76 with Agilent Human-Genome-CGH-105A array data (these 76 patients also have Affymetrix U133 A2.0 array data). All data is publically available in the NCBI GEO Repositories (accession numbers: GSE14520, GSE14322). The

array platforms are described in detail in previous publications[5,21,44]. Additional publically available microarray data for HCC was also analyzed from NCBI GEO Repositories (accession numbers: GSE1898 and GSE4024).

Short Hairpin RNA Sequences

shYY1AP1 C2, CAGATAGCGAAGGAACTATTT; C3, CCCTTAATTGTTTCTGGCAAT; C4, GCTCCTGACAACATCATTAAA; C5, GCATCTGTTATCTTCACTGTT; C6, GTTGTCAAGATGGAACCTGAA.

Clonogenicity, 3D Spheroid Formation Assays, Cell Proliferation Assay

Huh 1, Huh 7, HepG2, Hep3B or SK-Hep-1 cells were transiently infected with lentiviruses and selected with 2ug/ml puromycin for 3 days. 500 Huh 1, Huh 7 and Hep3B cells, 4 500 HepG2 cells, or 500 SK-Hep-1 cells were seeded into 6-well plates in triplicate and cultured with puromycin for 10 days. Colonies were then fixed and stained for quantification. For 3D cultures, cells were spread onto 48-well plate coated with MatrigelTM Basement Membrane Matrix (BD Bioscience, Bedford, MA). 80μl of Matrigel was distributed to each well and incubated at 37°C for 30 min, 1 000 HepG2 Tet-C2 cells were seeded onto the gelled Matrigel. The cells were covered with complete media containing 2% Matrigel and incubated at 37°C in 5% CO2. 500ng/ml of doxycycline was added at day 1. The spheroid number and size were evaluated at day 5 with the IX51 Inverted Microscope (Olympus, Tokyo, Japan). Spheroid numbers were counted at 20 randomly selected fields for each group. Sixty spheroids from each group were randomly selected for diameter measurement. The xCELLigence RTCA DP device (ACEA Biosciences, CA) was used for real-time cell proliferation assay. A total of 5,000 Hep3B cells transfected with control or YY1AP1 overexpression vector were seeded in E-plate 16 (ACEA Biosciences) per well. Cell index that reflects the integrated measure of cell number was monitored for 40 hours with programmed signal detection every 30 minutes. Data acquisition and analysis were performed with RTCA software 2.0 (ACEA Biosciences).

Duolink Fluorescence Assay

The Duolink II fluorescence assay was used to analyze the interaction between YY1AP1 and YY1 in the HepG2-Tet-C2 cells. Cells were fixed with 4% paraformaldehyde and permeabilized with 0.5% Triton X-100 for 10 minutes, then blocked and incubated with mouse anti-YY1AP1 Ab (1:100 dilution; Sigma) and rabbit anti-YY1 Ab (1:100 dilution; Abcam) overnight at 4°C. The Duolink II Proximity Ligation Assay (PLA) was then performed according to manufacturer's manual (Olink Bioscience, Uppsala, Sweden). A reporter substrate is formed if the PLA® probes are in very close proximity to one another indicating an interaction between the proteins examined. The reporter substrate appears as a red dot under a standard microscope. After mounting, the cells were visualized using a multiphoton confocal laser-scanning microscope (Carl Zeiss). Ten images per condition were acquired at 100× magnification with an average of ten cells per image. The number of red dots per cell was quantified from the average number of red dots per cell in ten images.

ChIP Assays

ChIP assays were performed as per the manufacturer's protocol using a SimpleChIP® Plus Enzymatic IP kit (Cell Signaling Techonology). ChIP grade antibodies specific to YY1 (Abcam, ab12132; 1μg/reaction) or H3K4Me3 (Cell Signaling Technology, #9751) were used for immunoprecipitations with rabbit IgG as control. Purified chromatin was examined by qRT-PCR. Four primer sets were used to detect different regions of the EpCAM promoter. Primer sequences are shown in Supplementary Table S4. Cycle threshold (Ct) values were normalized to the 2% input sample (Percent Input = $2\% \times 2^{(C[T]\ 2\%Input\ Sample - C[T]\ IP\ Sample)}$).

Luciferase Assays

pGL3-EpCAM containing firefly luciferase (described previously in[19]) was cotransfected or infected with a pRL-null vector containing renilla luciferase, YY1 or a control plasmid (RC207198 and PS100001, respectively, from Origene), and shYY1AP1-C2 or control sheGFP lentivirus. Twenty-four hours after transfection/infection, luciferase activities were measured using Dual-Luciferase Reporter Assay (Promega). Firefly luciferase activity was normalized to renilla.

Mice study

Six-week-old female athymic nude mice (Crl: NU-Foxn1nu) were purchased from Charles River Laboratories (Wilmington, MA). The research involving animals complied with protocols approved by the National Cancer Institute Bethesda Animal Care and Use Committee (protocol number: LHC-003). HepG2-Tet-C2 cells were suspended in 200μl 1:1 mixture of media and Matrigel (BD Biosciences) and subcutaneously injected into flanks of mouse. Subcutaneous nodules with a diameter of at least 0.5cm that persist for at least 3 days were considered the presence of a tumor. Tumor size was measured by caliper and recorded by *in vivo* fluorescence imaging. Alfa-alfa free control global 16% protein diet and 200ppm doxycycline diet were purchased from Harlan Laboratories, Inc., Madison, WI.

Reagents

Antibodies to YY1AP1 (HPA006986) from Sigma; antibody to EpCAM (MAB960) from R&D Systems; anti-DDK monoclonal antibody (TA50011-100) from Origen; antibody to p53 from Santa Cruz Biotechnology (sc-126); antibodies to TCF4 (#2953), E-Cadherin (#5296), AFP (#2137), cleaved caspase 3 (Asp175, #9961), c-Myc (#5605), Sox2 (#3579), β-catenin (#8480), Histone H2B (#12364), ub-Histone H2B (#5546), Methyl-Histone H3 (#9847) and Acetyl-Histone H3 (#9933) antibody sample kits, Tri-Mehyl-Histone H3 (Lys4) (#9751) and Tri-Methyl-Histone H3 (Lys27) (#9733) were from Cell Signaling Technology. Tet-system approved FBS and Doxycycline were from Clontech Laboratories. Caspase inhibitor Z-VAD was from Calbiochem. Daminozide was purchased from Sigma. Tranylcypromine and 2,4-PDCA were purchased from ENZO Life Sciences.

Statistical Analysis

GSEA was performed at http://www.broadinstitute.org/gsea/. Unsupervised hierarchical clustering analysis and class comparison analyses were performed using BRB Array Tool

software 3.3. We used multivariate survival risk prediction algorithm with 10-fold cross validation and 100 random permutations of the class label to test the association of stem cell-related genes with survival. In this analysis, 90% of the samples were randomly chosen to build a classifier containing all 1 573 *x*-HpSC-related genes, which was then used to predict the remaining 10% of the cases. The accuracy of the prediction, reflected by permutation *p*-values, was calculated after 100 repetitions of this random partitioning process. To identify oncogenic driver genes, which have correlated gene expression and gene copy numbers that are attributed to patient survival, we calculated the Spearman correlation coefficients (*r*) of the gene expression value and the gene copy number of 173 amplified genes in the chromosomal 1q22 region. Kaplan-Meier survival analysis was used to compare patient survival using GraphPad Prism. The statistical *p*-value was generated by permutation analysis and the Cox-Mantel log-rank test. Student's t-test (two-tailed) was used for statistical analysis of comparative data between two groups.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Dr. Curtis Harris for critical reading of the manuscript, Dr. Zheng-Gang Liu for advice on apoptosis study, Dominic Esposito at the Advanced Technology Program of SAIC-Frederick, Inc., for the help in designing the Tet-C2 inducible construct and Liver Tissue Cell Distribution System (LTCDS) at University of Minnesota (Minneapolis, MN) for generous donation of normal liver specimens. The authors also thank Karen Yarrick for the bibliographic assistance. This work was supported by the Intramural Research Grants of the Center for Cancer Research and the U.S. National Cancer Institute Z01 BC 010313, Z01 BC 010877 and Z01 BC 010876.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011; 61:69–90. [PubMed: 21296855]
2. Laurent-Puig P, Legoix P, Bluteau O, Belghiti J, Franco D, Binot F, et al. Genetic alterations associated with hepatocellular carcinomas define distinct pathways of hepatocarcinogenesis. Gastroenterology. 2001; 120:1763–1773. [PubMed: 11375957]
3. Ye QH, Qin LX, Forgues M, He P, Kim JW, Peng AC, et al. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. Nat Med. 2003; 9:416–423. [PubMed: 12640447]
4. Lee JS, Chu IS, Heo J, Calvisi DF, Sun Z, Roskams T, et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology. 2004; 40:667–676. [PubMed: 15349906]
5. Lee JS, Heo J, Libbrecht L, Chu IS, Kaposi-Novak P, Calvisi DF, et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. Nat Med. 2006; 12:410–416. [PubMed: 16532004]
6. Budhu A, Forgues M, Ye QH, Jia LH, He P, Zanetti KA, et al. Prediction of venous metastases, recurrence and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. Cancer Cell. 2006; 10:99–111. [PubMed: 16904609]
7. Boyault S, Rickman DS, de Reynies A, Balabaud C, Rebouissou S, Jeannot E, et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. Hepatology. 2007; 45:42–52. [PubMed: 17187432]
8. Yamashita T, Forgues M, Wang W, Kim JW, Ye Q, Jia H, et al. EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. Cancer Res. 2008; 68:1451–1461. [PubMed: 18316609]

9. Hoshida Y, Nijman SM, Kobayashi M, Chan JA, Brunet JP, Chiang DY, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. Cancer Res. 2009; 69:7385–7392. [PubMed: 19723656]
10. Zhou XD, Tang ZY, Yang BH, Lin ZY, Ma ZC, Ye SL, et al. Experience of 1000 patients who underwent hepatectomy for small hepatocellular carcinoma. Cancer. 2001; 91:1479–1486. [PubMed: 11301395]
11. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med. 2008; 359:378–390. [PubMed: 18650514]
12. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin. 2013; 63:11–30. [PubMed: 23335087]
13. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011; 144:646–674. [PubMed: 21376230]
14. Stephens PJ, Tarpey PS, Davies H, Van LP, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012; 486:400–404. [PubMed: 22722201]
15. Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB, et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat Genet. 2012; 44:694–698. [PubMed: 22561517]
16. Ji J, Wang XW. Clinical implications of cancer stem cell biology in hepatocellular carcinoma. Semin Oncol. 2012; 39:461–472. [PubMed: 22846863]
17. Yamashita T, Wang XW. Cancer stem cells in the development of liver cancer. J Clin Invest. 2013; 123:1911–1918. [PubMed: 23635789]
18. Ji J, Yamashita T, Budhu A, Forgues M, Jia HL, Li C, et al. Identification of microRNA-181 by genome-wide screening as a critical player in EpCAM-positive hepatic cancer stem cells. Hepatology. 2009; 50:472–480. [PubMed: 19585654]
19. Yamashita T, Budhu A, Forgues M, Wang XW. Activation of hepatic stem cell marker EpCAM by Wnt-β-catenin signaling in hepatocellular carcinoma. Cancer Research. 2007; 67:10831–10839. [PubMed: 18006828]
20. Yamashita T, Ji J, Budhu A, Forgues M, Yang W, Wang HY, et al. EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features. Gastroenterology. 2009; 136:1012–1024. [PubMed: 19150350]
21. Roessler S, Long EL, Budhu A, Chen Y, Zhao X, Ji J, et al. Integrative genomic identification of genes on 8p associated with hepatocellular carcinoma progression and patient survival. Gastroenterology. 2012; 142:957–966. [PubMed: 22202459]
22. Kuryshev VY, Vorobyov E, Zink D, Schmitz J, Rozhdestvensky TS, Munstermann E, et al. An anthropoid-specific segmental duplication on human chromosome 1q22. Genomics. 2006; 88:143–151. [PubMed: 16545939]
23. Guo M, House MG, Hooker C, Han Y, Heath E, Gabrielson E, et al. Promoter hypermethylation of resected bronchial margins: a field defect of changes? Clin Cancer Res. 2004; 10:5131–5136. [PubMed: 15297416]
24. Yang EB, Tang WY, Zhang K, Cheng LY, Mack PO. Norcantharidin inhibits growth of human HepG2 cell-transplanted tumor in nude mice and prolongs host survival. Cancer Lett. 1997; 117:93–98. [PubMed: 9233837]
25. Zhang K, Loong SL, Connor S, Yu SW, Tan SY, Ng RT, et al. Complete tumor response following intratumoral 32P BioSilicon on human hepatocellular and pancreatic carcinoma xenografts in nude mice. Clin Cancer Res. 2005; 11:7532–7537. [PubMed: 16243828]
26. Wang CY, Liang YJ, Lin YS, Shih HM, Jou YS, Yu WC. YY1AP, a novel co-activator of YY1. J Biol Chem. 2004; 279:17750–17755. [PubMed: 14744866]
27. Cuddapah S, Schones DE, Cui K, Roh TY, Barski A, Wei G, et al. Genomic profiling of HMGN1 reveals an association with chromatin at regulatory regions. Mol Cell Biol. 2011; 31:700–709. [PubMed: 21173166]
28. Sun ZW, Allis CD. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature. 2002; 418:104–108. [PubMed: 12077605]

29. Laribee RN, Krogan NJ, Xiao T, Shibata Y, Hughes TR, Greenblatt JF, et al. BUR kinase selectively regulates H3 K4 trimethylation and H2B ubiquitylation through recruitment of the PAF elongation complex. Curr Biol. 2005; 15:1487–1493. [PubMed: 16040246]
30. Lee JS, Shukla A, Schneider J, Swanson SK, Washburn MP, Florens L, et al. Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. Cell. 2007; 131:1084–1096. [PubMed: 18083099]
31. Kim J, Guermah M, McGinty RK, Lee JS, Tang Z, Milne TA, et al. RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. Cell. 2009; 137:459–471. [PubMed: 19410543]
32. Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol Cell Biol. 2005; 25:3305–3316. [PubMed: 15798214]
33. Rotili D, Mai A. Targeting Histone Demethylases: A New Avenue for the Fight against Cancer. Genes Cancer. 2011; 2:663–679. [PubMed: 21941621]
34. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009; 458:719–724. [PubMed: 19360079]
35. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, et al. The landscape of somatic copy-number alteration across human cancers. Nature. 2010; 463:899–905. [PubMed: 20164920]
36. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res. 2011; 39:D945–D950. [PubMed: 20952405]
37. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012; 366:883–892. [PubMed: 22397650]
38. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012; 487:330–337. [PubMed: 22810696]
39. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012; 489:519–525. [PubMed: 22960745]
40. Oishi N, Kumar MR, Roessler S, Ji J, Forgues M, Budhu A, et al. Transcriptomic profiling reveals hepatic stem-like gene signatures and interplay of miR-200c and epithelial-mesenchymal transition in intrahepatic cholangiocarcinoma. Hepatology. 2012; 56:1792–1803. [PubMed: 22707408]
41. Budhu A, Roessler S, Zhao X, Yu Z, Forgues M, Ji J, et al. Integrated metabolite and gene expression profiles identify lipid biomarkers associated with progression of hepatocellular carcinoma and patient outcomes. Gastroenterology. 2013; 144:1066–1075. [PubMed: 23376425]
42. Weinstein IB. Cancer. Addiction to oncogenes--the Achilles heal of cancer. Science. 2002; 297:63–64. [PubMed: 12098689]
43. Sankpal NV, Willman MW, Fleming TP, Mayfield JD, Gillanders WE. Transcriptional repression of epithelial cell adhesion molecule contributes to p53 control of breast cancer invasion. Cancer Res. 2009; 69:753–757. [PubMed: 19141643]
44. Roessler S, Jia HL, Budhu A, Forgues M, Ye QH, Lee JS, et al. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. Cancer Research. 2010; 70:10202–10212. [PubMed: 21159642]

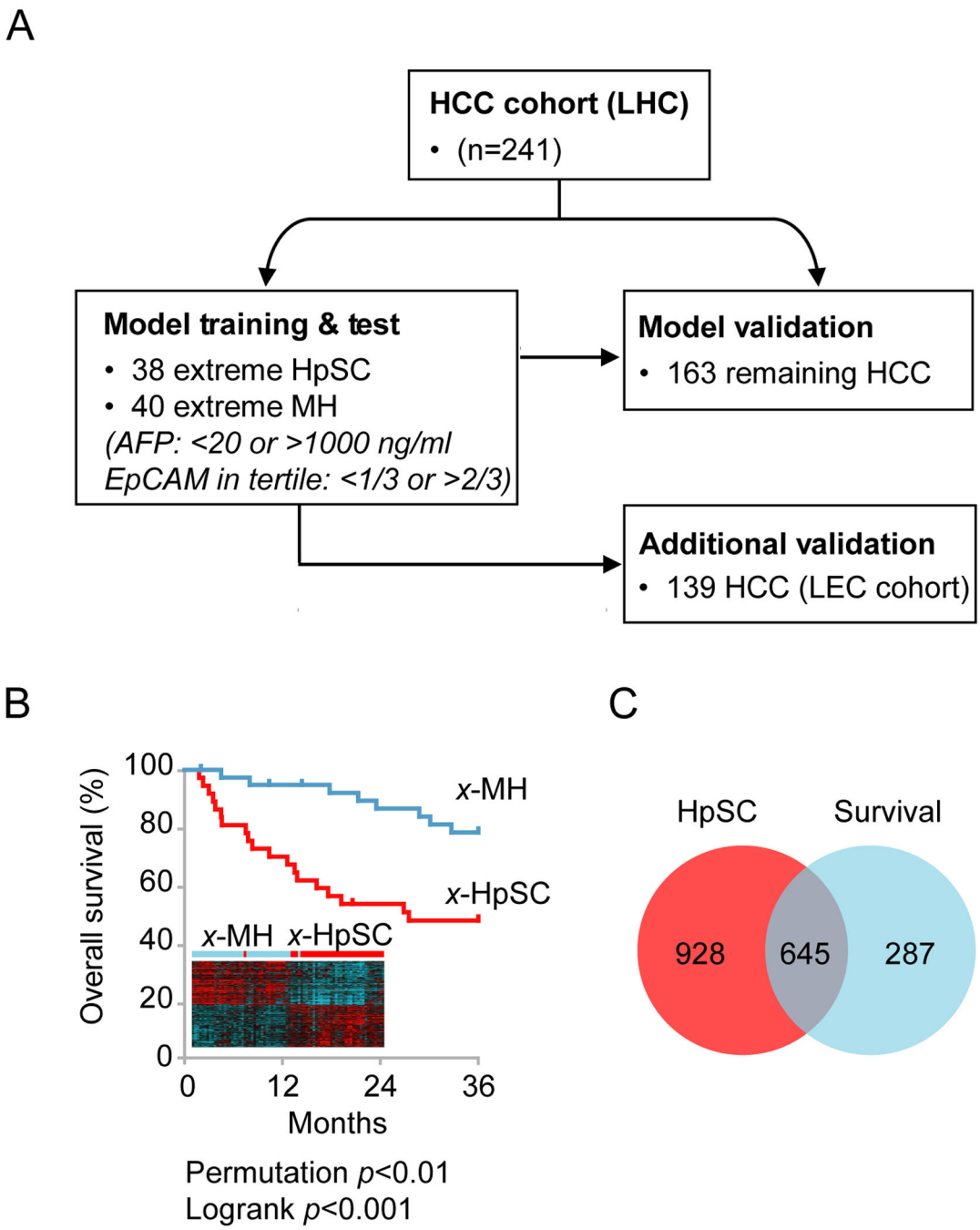
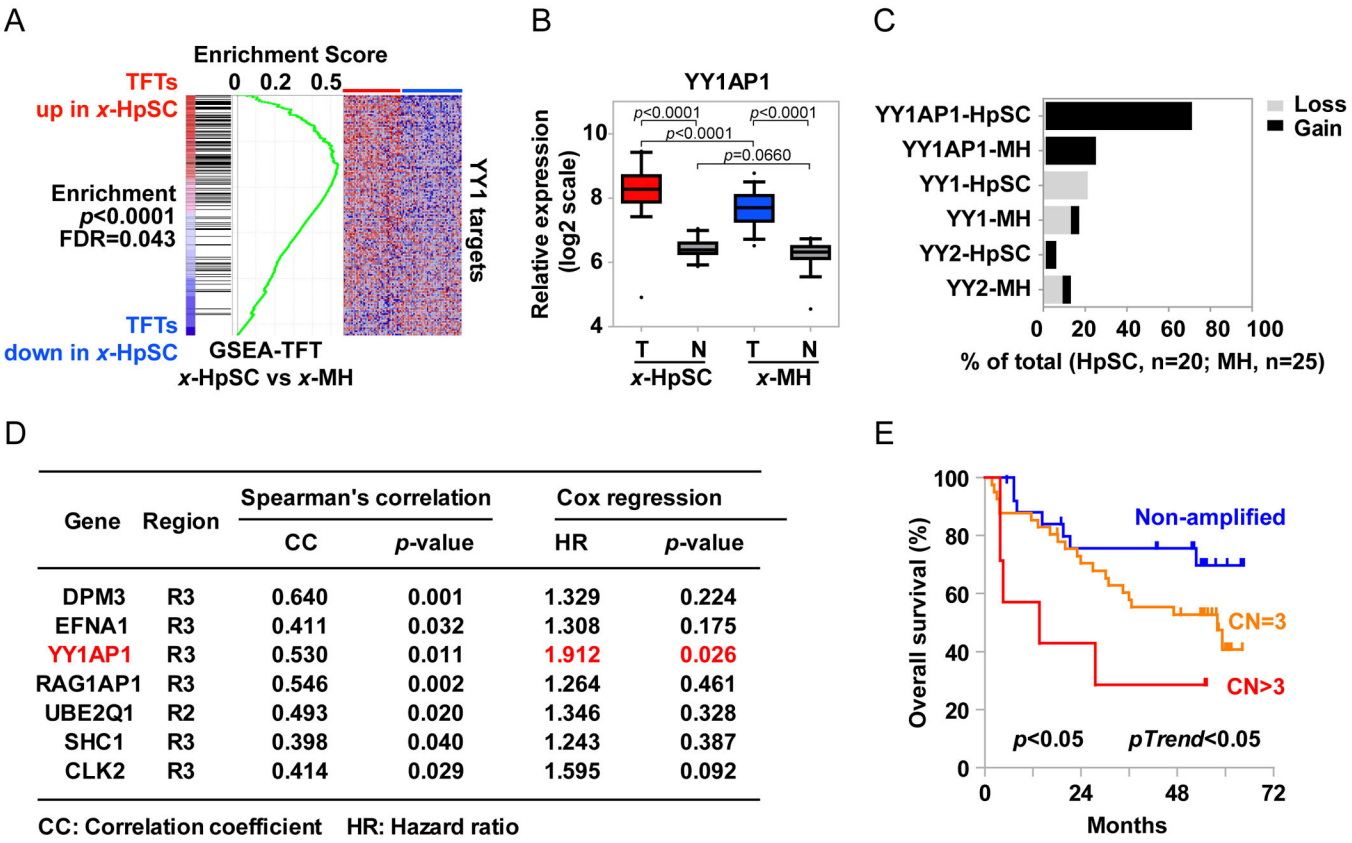


Figure 1.

Identification of a survival related, stem cell-like gene signature in HCC. (*A*) Overview of the approach used to develop and validate a gene signature that differentiates *x*-HpSC and *x*-MH subgroups. (*B*) Stem cell-like genes are significantly associated with overall survival in *x*-HpSC and *x*-MH cases (n=78). (*C*) Venn diagram showing the overlap between the survival- and stem cell-related (*x*-HpSC vs. *x*-MH) genes.



D

Gene	Region	Spearman's correlation		Cox regression	
		CC	*p*-value	HR	*p*-value
DPM3	R3	0.640	0.001	1.329	0.224
EFNA1	R3	0.411	0.032	1.308	0.175
YY1AP1	R3	0.530	0.011	1.912	0.026
RAG1AP1	R3	0.546	0.002	1.264	0.461
UBE2Q1	R2	0.493	0.020	1.346	0.328
SHC1	R3	0.398	0.040	1.243	0.387
CLK2	R3	0.414	0.029	1.595	0.092

CC: Correlation coefficient HR: Hazard ratio

Figure 2.

Activated YY1-YY1AP1 signaling in EpCAM$^+$ AFP$^+$ HCC. (*A*) GSEA analysis of genes that are enriched in the *x*-HpSC subgroup. (*B*) Mean gene expression levels of YY1AP1 in paired tumor and non-tumor tissue samples from *x*-HpSC or *x*-MH subgroups (*x*-HpSC, n=20; *x*-MH, n=25). (*C*) Loss/gain frequency of YY1AP1, YY1 and YY2 in *x*-HpSC and *x*-MH subgroups detected by arrayCGH. A threshold of 0.5 was used. (*D*) Seven genes that have significantly correlated gene expression and copy number changes within 3 amplicons. Cox regression analysis was used for survival risk prediction analysis (HR=1.912, $p<0.05$). (*E*) Kaplan-Meier overall survival based on the quantitative YY1AP1 gene copy numbers determined by qPCR in all 76 HCC patients used in the arraryCGH study.

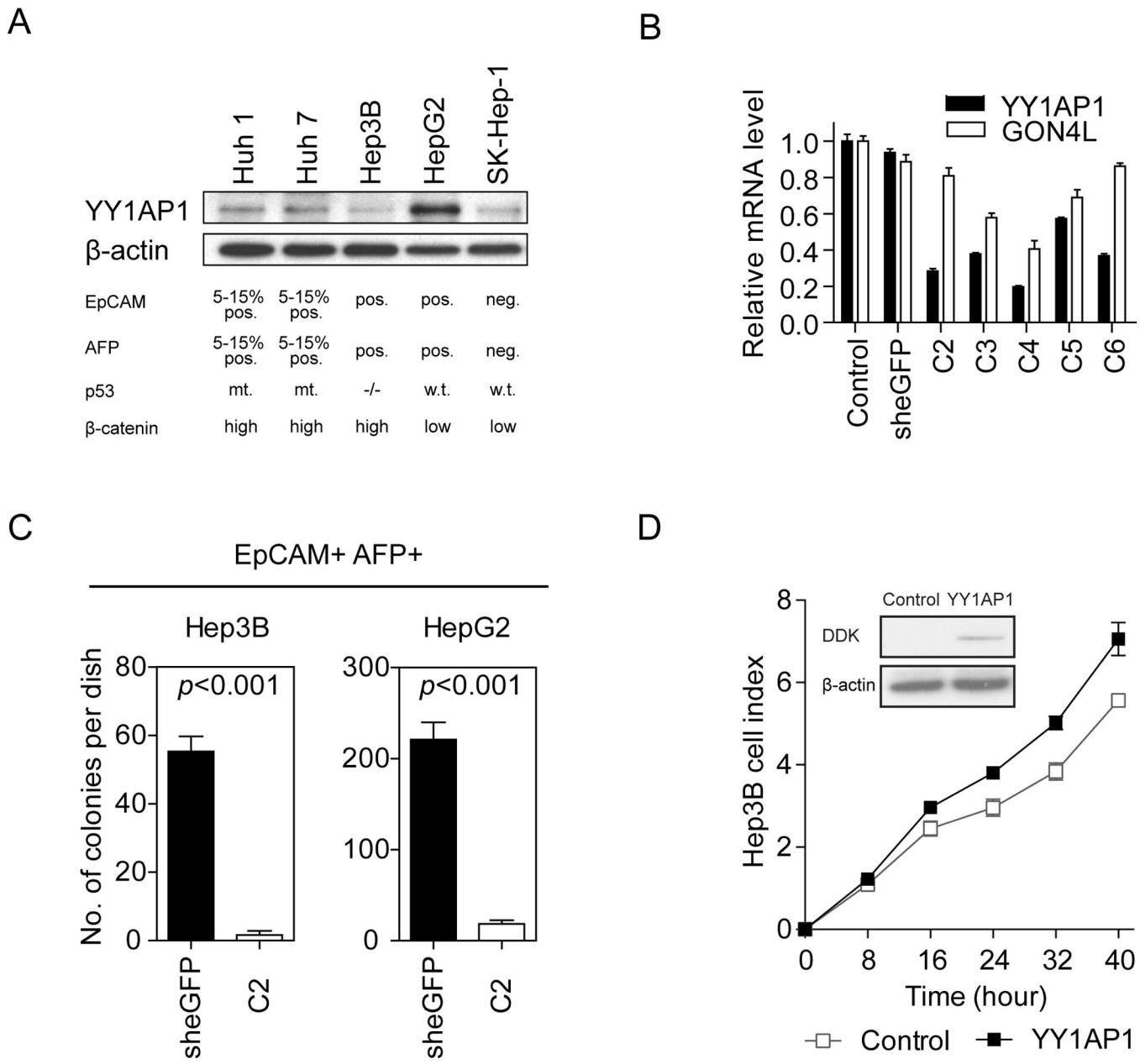


Figure 3.

YY1AP1 characterization and silencing *in vitro*. (*A*) YY1AP protein levels in 5 cell lines and cell line backgrounds. (*B*) Efficiency of YY1AP1mRNA silencing using lentiviral constructs containing YY1AP1 shRNAs that target various regions of the transcript. Samples were collected for analysis 2 days after YY1AP1 knockdown. (*C*) Colony formation assays for shRNA C2 in Hep3B and HepG2. *(D)* Real-time cell proliferation of Hep3B cells with or without YY1AP1 expression.

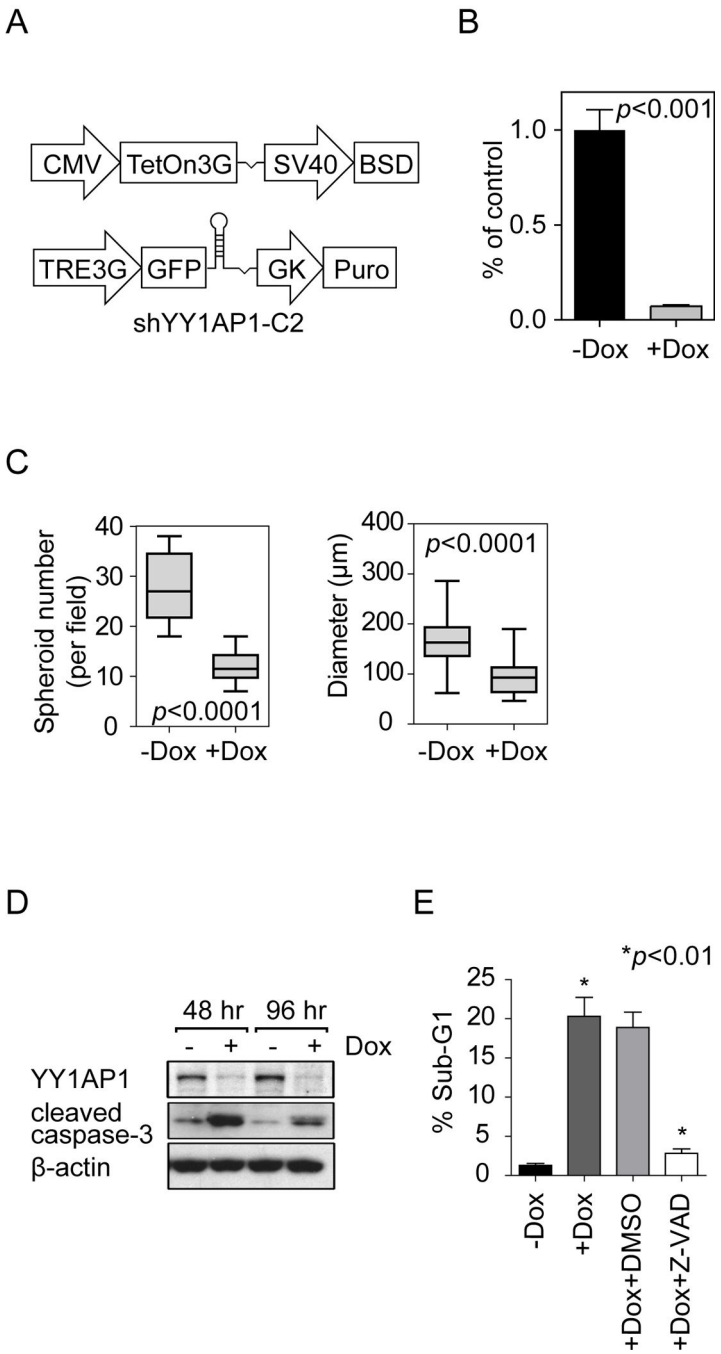


Figure 4.

YY1AP1 silencing leads to apoptosis. (*A*) Schematic of Tet-On 3G doxycycline-inducible constructs. (*B*) Quantification of colony formation assay from 2D cell culture. (*C*) Quantification of spheroid formation assay and the diameter of spheroids from 3D cell culture. (*D*) Immunoblot of YY1AP1 and active caspase 3 in the HepG2-Tet-C2 cells after 48 and 96 hours exposure to doxycycline. (*E*) Cell cycle analysis of HepG2-Tet-C2 cells with or without Z-VAD treatment.

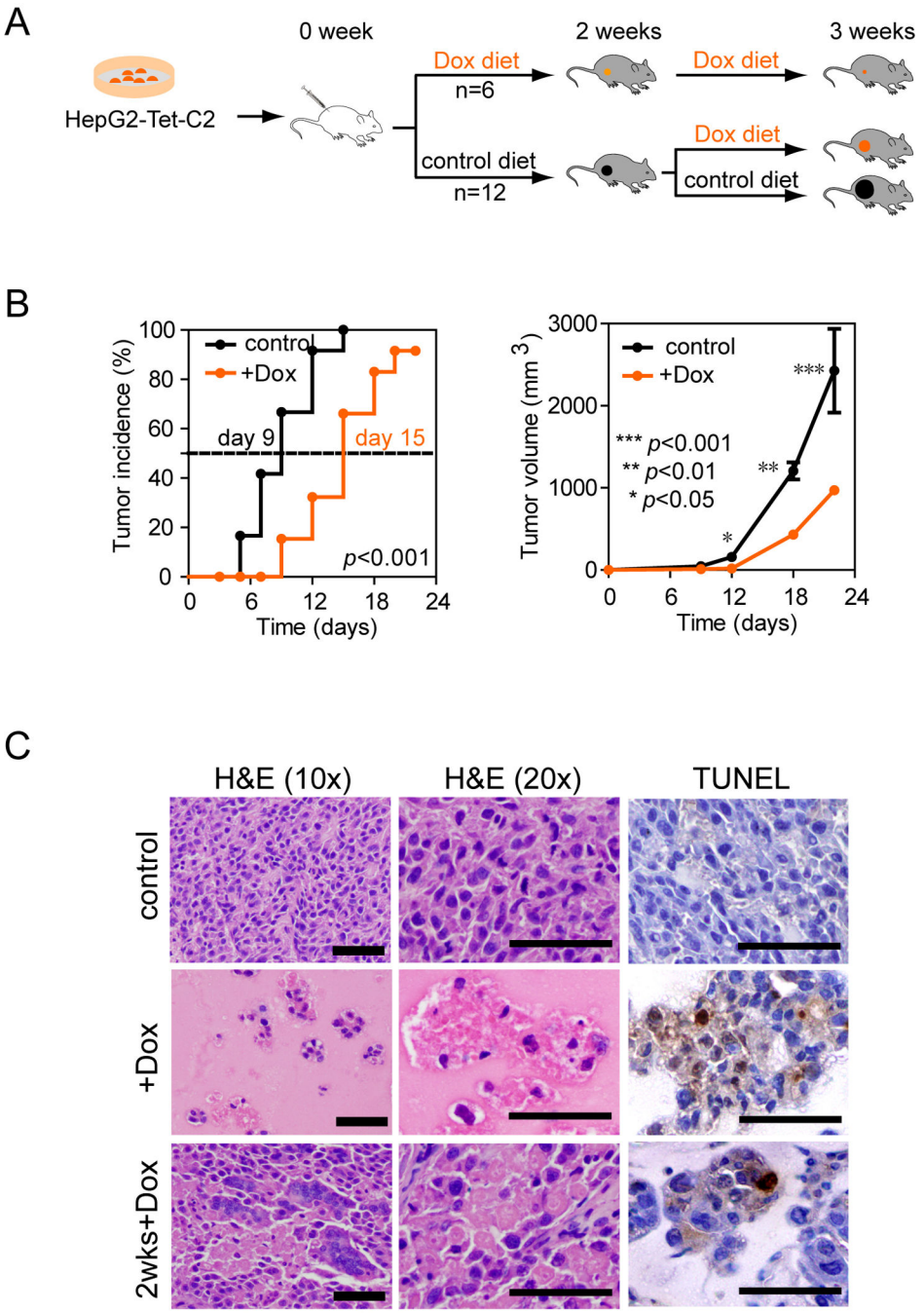


Figure 5.

Depletion of YY1AP1 suppresses tumor formation *in vivo*. (*A*) Experimental protocol of Dox-induced YY1AP1 silencing system. (*B*) Tumor incidence and size of control or Dox dieted mice. Data are presented as mean ± SEM (n=6/group). (*C*) Representative H&E and TUNEL staining images of tumors from mice with control diet (top panel), 3 weeks Dox diet (middle panel), and 1 week Dox diet after 2 weeks control diet (bottom panel). Black bar indicates 50 μm.

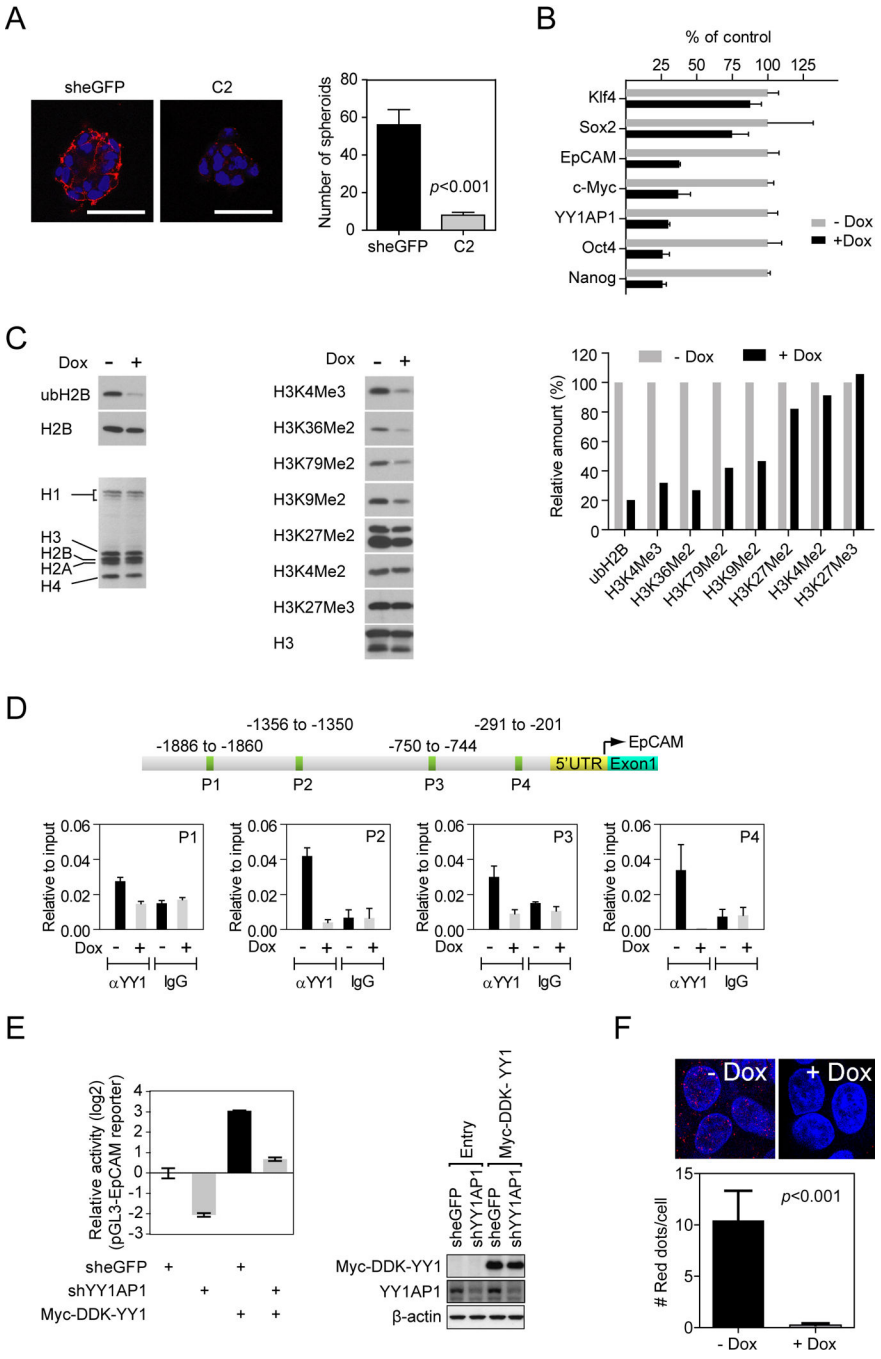


Figure 6.

YY1AP1 modifies histones and changes the chromatin landscape. (*A*) YY1AP1 silencing resulted in a reduction of EpCAM protein in spheroids. (*B*) qRT-PCR analysis after 2 days Dox treatment of HepG2-Tet-C2 cells. Data are normalized to actin and shown as relative to control cells (-Dox) and represented as mean ± SD, n=3. (*C*) Screening of ubH2B and various histone methylation marks after YY1AP1 depletion in HepG2-Tet-C2 cells. (*D*) Endogenous YY1 chromatin immunoprecipitation (ChIP) assays for EpCAM promoter region. (*E*) Luciferase reporter assay for EpCAM promoter activity in the presence of YY1

and YY1AP1. (*F*) Duolink assay showing *in vivo* interactions between YY1AP1 and YY1 in HepG2-Tet-C2 cells.

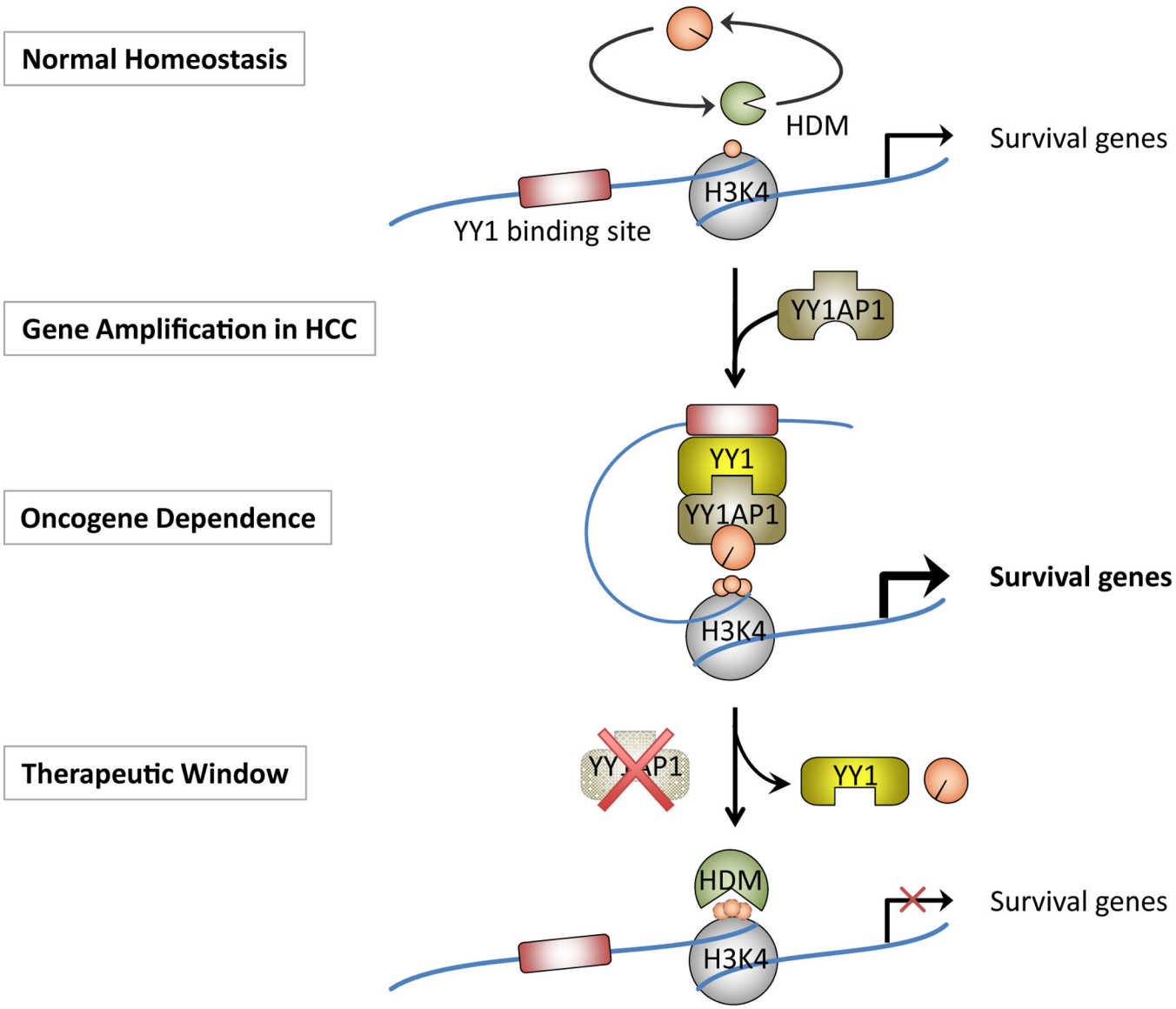


Figure 7.

Schematic describing a potential mechanism of cancer cells to acquire oncogene addition to YY1AP1 and a proposed therapeutic approach to reverse the disease state.