Am J Cancer Res 2015;5(10):2969-2979
www.ajcr.us /ISSN:2156-6976/ajcr0009727

Original Article

Genetic alteration and misexpression of Polycomb group genes in hepatocellular carcinoma

Shu-Bin Gao[1,2*], Shi-Long Sun[3*], Qi-Lin Zheng[1], Li Zhang[1,2], Yuequan Zhu[4], Guang-Hui Jin[1,2], Li-Xiang Xue[4]

[1]Department of Basic Medical Sciences, Medical College, Xiamen 361102, P.R. China; [2]Fujian Provincial Key Laboratory of Chronic Liver Disease and Hepatocellular Carcinoma, Xiamen University, Chengzhi Building 110, Xiang'an South Road, Xiamen 361102, P.R. China; [3]Ministry of Health Key Laboratory of Radiobiology, Jilin University, 1163 Xinmin Road, Changchun 130021, P.R. China; [4]Medical Research Center, Department of Radiation Oncology, Peking University Third Hospital, Beijing 100191, P.R. China. [*]Equal contributors.

Received April 16, 2015; Accepted June 25, 2015; Epub September 15, 2015; Published October 1, 2015

Abstract: Although the abnormal expression of Polycomb-group (PcG) proteins is closely associated with carcinogenesis and the clinicopathological features of hepatocellular carcinoma (HCC), the genetic mutation profile of PcG genes has not been well established. In this study of human HCC specimens, we firstly discovered a highly conserved mutation site, G553C, in the Polycomb Repressive Complex 2 (PRC2) gene enhancer of zeste homolog 2 (EZH2). This site also harbors a single nucleotide polymorphism (SNP), rs2302427, which plays an important antagonistic role in HCC. Kaplan-Meier survival curves showed that the tumor-free and overall survival of patients with EZH2 G553C were superior to those without the mutation. The G allele frequencies in patients and healthy subjects were 0.2% and 0.122%, respectively, with significant differences in distribution. The individuals carrying the GG and the GC genotypes at rs2302427 showed 3.083-fold and 1.827-fold higher risks of HCC, respectively, compared with individuals carrying the wild-type allele. Furthermore, Immunohistochemical staining revealed that the expression levels of CBX8 (in 53/123 samples) and BMI1 (in 60/130 samples) were markedly increased in human HCC specimens. Importantly, the overall and tumor-free survival rates were significantly reduced in the group of patients who simultaneously expressed PRC1 and PRC2. These results argue that a combination of PRC1 and PRC2 expression has a significant predictive/prognostic value for HCC patients. Taken together, our results indicate the abnormal expression and genetic mutation of PcG members are two independent events; cumulative genetic and epigenetic alterations act synergistically in liver carcinogenesis.

Keywords: Hepatocellular carcinoma, Polycomb, H3K27 trimetylation, SNP

Introduction

Hepatocellular carcinoma (HCC) is the most common form of liver cancer. Numerous studies have greatly advanced our understanding of the pivotal role of specific gene mutations, including *p53, CDKN2A, CTNNB1, AXIN1, HNF1,* and *IRF-2*, in controlling HCC development [1]. Recently, emerging evidence has demonstrated an essential role for epigenetic regulators, such as microRNA or DNA methylation, in HCC [2-5], suggesting that HCC is frequently governed by cumulative genetic and epigenetic alterations [6].

Site-specific histone modifications are the major epigenetic mechanism for controlling a stable gene transcription state, but the importance of histone methylation in HCC development is not well established [7, 8]. One of the best-studied histone modifications required for the maintenance of gene silencing in HCC is the trimethylation of histone 3 lysine 27 (H3K27me3), which is mediated by Polycomb-group (PcG) proteins [4, 8-13]. The PcG comprises at least two distinct complexes. First, the initiation complex, the Polycomb Repressive Complex 2 (PRC2), has a core that in humans consists of the proteins enhancer of zeste homolog 2 (EZH2) and suppressor of zeste 12 (SUZ12). Second, the maintenance complex, PRC1, includes the proteins B-lymphoma Moloney murine leukemia virus insertion region-1 (BMI1), Chromobox homolog 8 (CBX8), and others [14]. The EZH2 SET domain specifically catalyzes the trimethylation of H3K27;

H3K27me3 compresses the chromatin structure and leads to the transcriptional repression of genes such as E-cadherin, RUNX3, and cyclin-dependent kinase inhibitors [15-17]. Recently, we also demonstrated the tumor-promoting activity of PcG, which occurs by direct and indirect regulation in HCC [9]. We found EZH2 occupancy at chromatin coincides with H3K27me3 at promoters and directly silences the transcription of certain tumor suppressors in HCC. The H3K27me3-related target gene network of PcG contains well-established genes (e.g., *CDKN2A*) and previously undescribed genes (e.g., *FOXO3*, *E2F1*, and *NOTCH2*). PcG also represses the expression of the TP53 tumor suppressor in HCC independently of H3K27me3 [9]. These data strongly demonstrate the functional and mechanistic significance of certain gene regulatory networks that are regulated by PcG in HCC.

Compared to PRC2, PRC1 is more complicated and functionally diverse. Rather than catalyzing H3K27me3 methylation like PRC2, the PRC1 components include BMI1 and CBX8 family proteins, which show affinity for H3K27me3. In the prevailing model, PRC2 is recruited to specific genomic loci where it catalyzes H3K27 trimethylation [14]. The trimethylated histones in turn recruit PRC1, which catalyzes H2A ubiquitination and thereby antagonizes RNA polymerase II elongation and represses the transcription of target genes [18]. CBX8 was originally characterized as a transcriptional repressor of *INK4a/ARF* in fibroblasts [19]. We also found that both EZH2 and CBX8 repress *INK4a/ARF* expression in HCC cells through H3K27me3 methylation [9]. The functional significance of PRC1 in the initiation and development of HCC has not been completely deciphered, although its mechanistic importance in mediating the response to H3K27me3 has been addressed. As with EZH2, BMI1 over-expression was detected in 60.5% of primary HCC tumor tissues. The cumulative recurrence rate was significantly higher in BMI1-overexpressing patients than in their BMI1-negative counterparts [12]. The potential biological role of CBX8 in liver cancer remains largely undefined.

Interestingly, as an important epigenetic regulation factor, mutations in *EZH2* gene are also involved in certain types of tumors. For example, sporadic point mutations affecting the Y641 and A677 residues in SET domain of *EZH2* have been identified in lymphoma and myeloid neoplasms [20]. Further studies have indicated that PcG loci are sensitive targets of environmental stress; emerging evidence has shown extensive and frequent somatic PRC2 alterations in early T-cell precursor acute lymphoblastic leukemia, including deletions and sequence mutations in EED, EZH2 and SUZ12 [21]. Because of the clinical importance of PcG in HCC, it is interesting to determine whether the components of PcG are mutated in HCC, and if so, whether these mutations are clinically significant. In the present study, we investigated the clinical significance of PcG members, especially EZH2 and CBX8, and examined genetic mutations of EZH2 and CBX8 in HCC.

Materials and methods

Human HCC samples and immunohistochemistry (IHC)

Informed consent to participate in this study was obtained from each patient, and the protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki as reflected in a priori approval by the Xiamen University Medical Ethics Committee. Paraffin-embedded primary HCC tissues and corresponding adjacent nontumorous liver samples from Han Chinese patients were obtained from the chronic liver disease biological sample bank. Department of Hepatobiliary Surgery, Zhongshan Hospital of Xiamen University [8]. During the same study period, 310 ethnic group-matched individuals who had no self-reported history of cancer at any site were enrolled as controls who underwent physical examination. Personal information and characteristics were collected from the study subjects using interviewer-administered questionnaires. All male and female patients were histopathologically diagnosed with stage I-IV HCC. The demographic data and clinicopathological features are listed in Supplementary Tables 9 and 10. Sections from paraffin-embedded samples were stained with affinity-purified anti-CBX8 or anti-BMI1 antibodies for IHC [8]. The expression of CBX8 or BMI1 in liver tissue as shown by IHC was evaluated by three independent pathologists.

DNA extraction from human HCC samples

DNA from human HCC samples was extracted from fresh tissue or from tissue that was frozen at -80°C after collection by using the E-Z 96 Tissue DNA Kit (Omega) according to the manufacturer's protocol [22].

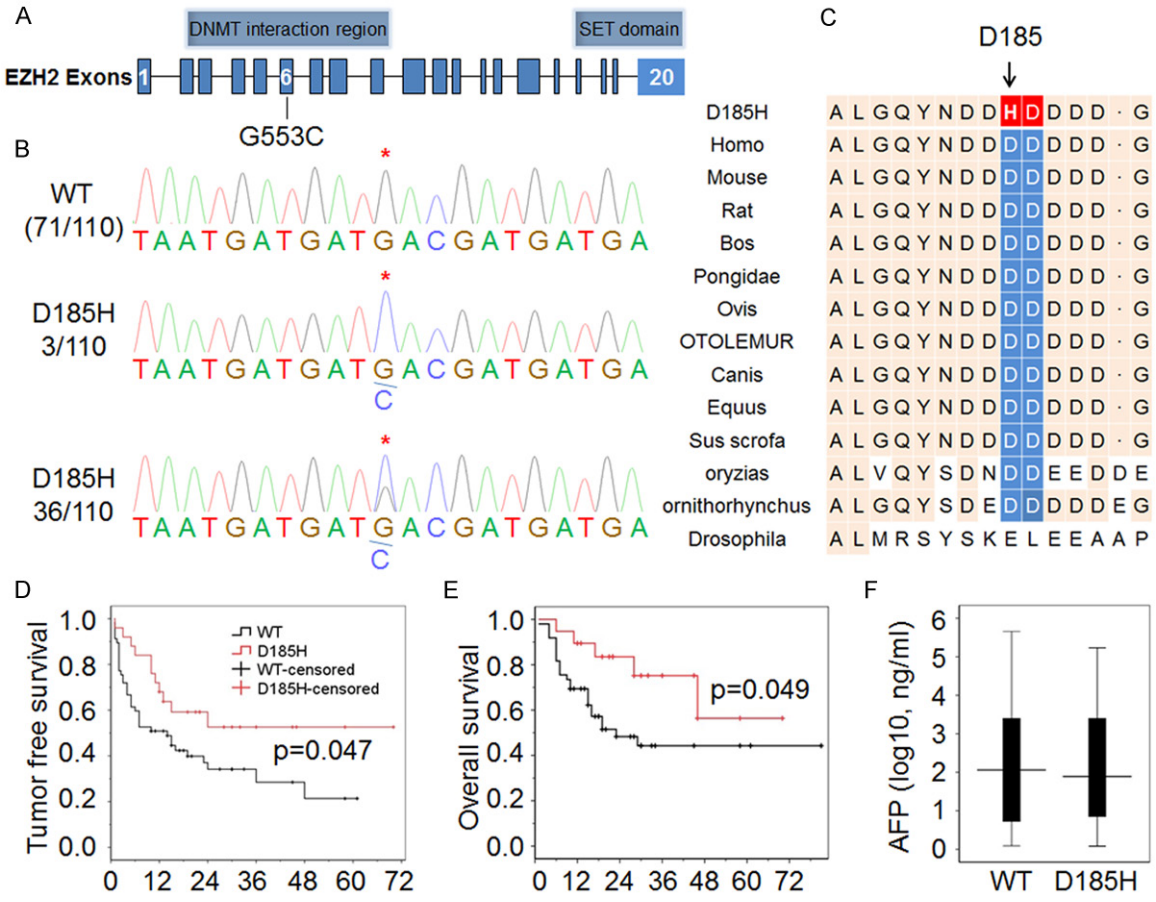


Figure 1. A screening for *EZH2* mutations in HCC. A. Genomic organization of the *EZH2* locus showing alternative exons and the protein domain structure. The location of the mutation affecting Asp185 in exon 6 of the *EZH2* gene is shown. B. Illustration of sequencing results. Of 110 samples, 3 showed distinctly homozygous mutations and 36 showed distinctly heterozygous mutations. The amino acid substitutions in different HCC samples at codon 185 were detected by sequencing. C. A multiple alignment of EZH2 protein sequences from thirteen species; Asp185 is conserved. D and E. Kaplan-Meier curves for tumor-free and overall survival in HCC patients with EZH2 WT and D185H. F. Serum AFP levels of HCC patients with EZH2 WT and D185H. The line in each panel indicates the median.

Exon sequencing

PCR reactions were performed in a volume of 25 μl and contained 50 ng of genomic DNA, 0.175 mM each primer, 0.2 mM dNTP mix, 1.5 mM $MgCl_2$, 10 mM Tris-HCl, 50 mM KCl, and 0.2 μl of Taq DNA Polymerase (TaKaRa). PCR cycling conditions were as follows: 95°C for 5 min, 30 cycles of denaturation at 95°C for 30 s, specified annealing temperature for 30 s, and extension for 30 s at 72°C, and then a final extension step for 10 min at 72°C. After purifying the PCR products, sequencing reactions were performed using a BigDye Terminator v3.1 kit (Applied Biosystems). The primers used for amplification and sequencing reactions, the sizes of the amplified PCR products, and the annealing temperatures for each pair of primers are listed in Supplementary Table 11. The HCC samples were sequenced twice independently to exclude the contamination between samples.

Statistical analysis

Deviations from Hardy-Weinberg equilibrium (HWE) were assessed using the HWSIM program. Differences in the allele and genotype frequencies of *EZH2* and *CBX8* between groups were evaluated using Chi-square tests. Quantitative trait tests were performed using UNPHASED 2.404, which examines the association of genetic polymorphisms with symptom severity. The effects of haplotypes on quantitative trait variation were evaluated using the sub-program QTPHASE of the UNPHASED software, which provided a chi-square value for the association between a particular haplotype

Table 1. Distribution of *EZH2* genotypes and allelic frequencies in the study population

EZH2 (HCC tissue) rs2302427	Genotypes			Allelic Frequency	
	CC	CG	GG	C	G
Patients	72	35	3	0.814	0.186
Controls	222	63	4	0.877	0.122
	χ^2=5.27, d.f.=2, *p*=0.071			χ^2=5.08, d.f.=1, *p*=0.024	
EZH2 (adjacent tissue) rs2302427	Genotypes			Allelic Frequency	
	CC	CG	GG	C	G
Patients	54	28	3	0.8	0.2
Controls	222	63	4	0.877	0.122
	χ^2=6.13, d.f.=2, *p*=0.046			χ^2=6.04, d.f.=1, *p*=0.013	

Table 2. Odds ratio with 95% CI of the *EZH2* gene in HCC patients

Genotypes		Odds Ratio (CI 95%)	*p* value
EZH2 (cancer tissue) rs2302427	G vs C	1.636 (1.074-2.491)	0.021
	CG vs CC	1.713 (1.048-2.8)	0.0377
	GG vs CC	2.313 (0.506-10.58)	0.361
EZH2 (adjacent tissue) rs2302427	G vs C	1.785 (1.138-2.801)	0.011
	CG vs CC	1.827 (1.07-3.121)	0.035
	GG vs CC	3.083 (0.67-14.19)	0.199

and a continuous variable based on the statistical parameters of the QTPHASE sub-program. Other statistical analyses were performed using SPSS software version 17.0 (SPSS, Inc.; Chicago, IL). Survival curves were calculated by the Kaplan-Meier method, and comparisons were performed using the log-rank test. The results for parametric variables are expressed as the means ± SD or the means ± SEM. In all cases, $p<0.05$ was considered statistically significant.

Results

Recurrent G553C mutation in EZH2

To uncover the genetic mutation profile of *EZH2* in HCC, we performed exon sequencing of the entire *EZH2* locus in 110 HCC patients (**Figure 1A**). We focused our analysis on novel sequence changes predicted to affect protein coding. Among these variants, we found a G-to-C mutation in 39 of 110 HCC samples at nucleotide 553 (G553C) within exon 6 of the *EZH2* gene; this transversion corresponds to a D185H amino acid substitution (**Figure 1B**). No other *EZH2* mutations were detected. Homozygous mutations were found in 3 samples, while heterozygous mutations occurred in 36 samples and were thus much more common (**Figure 1B**). The striking recurrence of this mutation suggested that the G553C alteration in *EZH2* is a common feature of HCC. Comparing the amino acid sequence of human *EZH2* with those from other species revealed that the mutation site is highly conserved (**Figure 1C**).

Next, we performed a correlation analysis between the *EZH2* G553C mutation and HCC malignancy in 110 cases with detailed clinical information. Kaplan-Meier survival curves were plotted according to the mutation status. Interestingly, the tumor-free survival (**Figure 1D**, $p=0.047$, log-rank test) and overall survival (**Figure 1E**, $p=0.049$, log-rank test) of patients with the *EZH2* D185H mutation were superior to those of patients without the mutation. This result suggests that the somatic mutation of D185 is most likely a loss-of-function mutation that interferes with the pathogenesis of HCC. No statistically significant correlation was found between G553C mutation and the serum alpha fetoprotein (AFP) level (**Figure 1F**).

The G553C mutation in EZH2 is a single-nucleotide polymorphism (SNP)

Because of the high incidence of *EZH2* G553C mutations in HCC, we assumed that the mutation was a SNP. We thus recruited 311 healthy unrelated people as a control group. Using UNPHASED genetic analysis software (V2.404), we further analyzed the relationship between rs2302427 and HCC. The demographic and clinical characteristics of the selected cases and controls are shown in Supplementary Table 1. The mean ages of the 110 cases with HCC and the 311 controls were 58.3±13.36 and 53.98±12.19 years, respectively. We did not find any significant differences in G553C distribution with respect to age or gender between the HCC and control groups (P>0.05). We analyzed the distribution of rs2302427 in both HCC tissue and in adjacent tissues. In cancer tissue, the three different rs2302427 genotypes CC, CG, and GG showed differences in

Table 3. Quantitative trait analysis of *EZH2* genotypes and alleles

EZH2 (cancer tissue) rs2302427	Genotypes				Allelic Frequency		
	CC	CG	GG	P	C	G	P
Overall survival	13.83	13.51	29.67	0.403	13.77	15.88	0.494
Tumor free survival	12.30	12.89	29.67	0.296	12.42	15.34	0.308
AFP	1.76	1.54	1.44	0.791	1.714	1.523	0.508
EZH2 (adjacent tissue) Rs2302427	Genotypes				Allelic Frequency		
	CC	CG	GG	P	C	G	P
Overall survival	12.65	18.25	30.67	0.181	13.8	20.44	0.075
Tumor free survival	10.22	17.79	30.67	0.049	11.78	20.06	0.019
AFP	1.84	1.53	0.74	0.388	1.776	1.389	0.221

their distribution between HCC patients and the control group; however, the association lacked sufficient statistical significance (**Table 1**, p=0.071). The allelic frequencies of G in patients and healthy subjects were 0.186% and 0.122%, respectively. There were significant differences in the allelic gene frequency distribution, which indicates that the allelic frequency of G was significantly associated with HCC (**Table 1**, OR=1.636, *p*=0.024). Individuals carrying the G allele, the GC genotype or the GG genotype at rs2302427 in the cancer tissue showed a 1.636-fold (95% CI: 1.074-2.491, *p*=0.021), a 1.713-fold (95% CI: 1.048-2.8, *p*=0.0377), or a 2.313-fold (95% CI: 0.506-10.58, *p*=0.361), respectively, higher risk of HCC than patients carrying the wild-type allele (**Table 2**).

In tissues adjacent to the HCC, the numbers of patients who carried the three different genotypes CC, CG, and GG of *EZH2* rs2302427 were 54, 28, and 3, respectively, a distribution that was significantly different compared with control subjects (**Table 1**, p=0.046). The G allele frequencies in patients and healthy subjects were 0.2% and 0.122%, respectively, with significant differences in distribution (**Table 1**, p=0.013). Individuals carrying the GG and GC genotypes at rs2302427 showed a 3.083-fold (95% CI: 0.67-14.19, p=0.199) and a 1.827-fold (95% CI: 1.07-3.121, p=0.035), respectively, higher risk of HCC than individuals carrying the wild-type allele (**Table 2**). The G allele in adjacent tissues was also highly associated with HCC (**Table 2**, OR=1.785, p=0.011).

We further analyzed the distribution of rs2302427 in cancer tissue, para-cancerous tissue, and in healthy control tissue. The distributions in these three groups were in Hardy-Weinberg equilibrium. Both the genotypic and allelic distributions of rs2302427 exhibited differences between the cancer tissue and adjacent tissues (**Tables 1** and **2**). The allelic frequency also differed within the same group. An inconsistent genotype at rs2302427 was observed between the cancer and adjacent tissues in 8 cases (**Tables 1** and **2**). Based on these data, the G allele and the GC genotype act as risk factors for HCC incidence. The G-to-C mutation in both cancerous and adjacent tissues might be a self-protection mechanism for HCC patients. These data indicate that rs2302427 is not only closely associated with HCC SNPs but also a locus that is easily mutated during HCC progression.

Association of the rs2302427 SNP with clinical symptoms

The effect of haplotypes on quantitative trait variation was evaluated using the QTPHASE sub-program of UNPHASED software, which provided a chi-square value for the association between a special haplotype and a continuous variable. Using this software, we examined the association of rs2302427 with phenotype. The tumor-free survival period varied significantly with different genotypes in adjacent tissues as follows: survival for the CC genotype was 10.22 months, survival for the CG genotype was 17.79 months, and survival for the GG genotype was 30.67 months (**Table 3**, p=0.049). The survival for the C allele was 11.78 months, whereas the survival for the G allele was 20.06 months (**Table 3**, p=0.019), which indicated that the G allele may be a protective factor. Although in cancerous tissues there are no significant differences in the survival period with various genotypes, we observed that the survival period in patients with the GG genotype was longer than that for the other 2 genotypes. This may partially confirm the hypothesis that the Gallele might be a protective factor (**Table 3**). The results show that the rs2302427 polymorphism of *EZH2* is significantly associated with the risk of HCC.

The expression of PRC1 is stimulated in HCC and is correlated with poor prognosis

It is unclear whether PRC1 family members are associated with liver carcinogenesis. To deter-

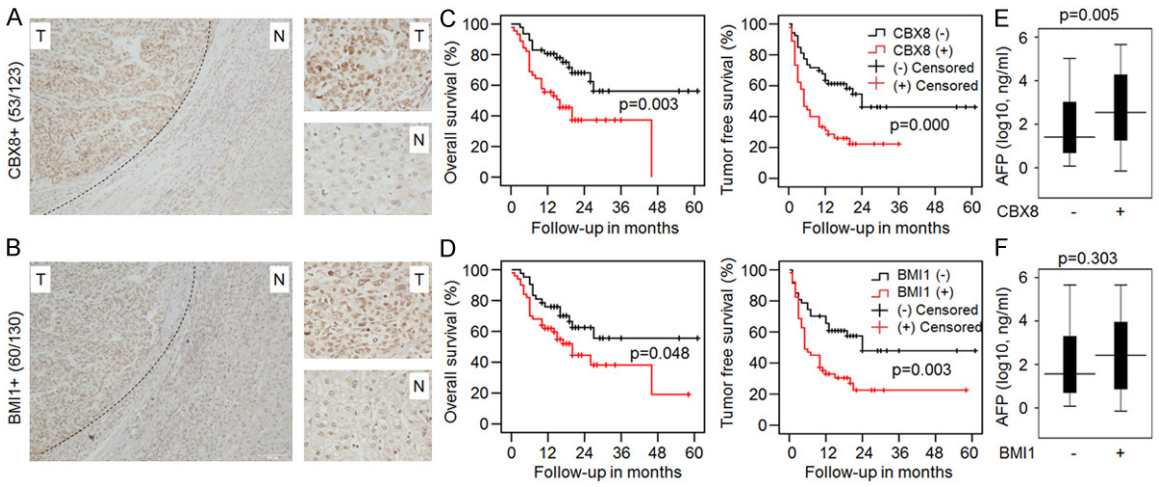


Figure 2. Increased expression of CBX8 or BMI1 is associated with poor prognosis of HCC. A and B. Representative micrographs of high-level CBX8 or BMI1 expression in primary HCC paraffin sections as visualized by IHC staining. The dotted lines indicate junctures between the tumor (T) and adjacent normal tissues (N). The original magnification values are 100× and 400×, respectively. C and D. Kaplan-Meier curves for overall and tumor-free survival in HCC patients with CBX8 or BMI1 hyper-expression (+) or under-expression (-). E and F. Serum AFP levels of HCC patients with CBX8 or BMI1 hyper-expression (+) or under-expression (-). The line in each panel indicates the median.

mine the clinical significance of PRC1 in HCC, we measured the expression of CBX8 and BMI1 in primary HCC samples from patients. Immunohistochemical (IHC) staining revealed that HCC tissues exhibited robust expression of these factors. Moreover, CBX8 staining was exclusively nuclear in HCC tissues but not in adjacent tissues (**Figure 2A** and Supplementary Figure 1A). The expression of CBX8 was markedly increased in 53 out of 123 HCC samples, accounting for 43.1% of the tumors examined. Similarly, the IHC detection revealed overexpression of BMI1 compared to adjacent normal tissues in 60 of 130 HCC tissue samples (**Figure 2B**). We further performed a correlation analysis between PRC1 expression and HCC malignancy. Kaplan-Meier survival curves showed that the 5-year overall survival and tumor-free survival were significantly lower in the CBX8 or BMI1 over-expressing HCC patients than in the under-expressing patients (**Figure 2C** and **2D**, log-rank test). The serum level of AFP was dramatically elevated in the CBX8 over-expressing group compared with the under-expressing group (**Figure 2E**, p=0.005). The hyper-expression of BMI1 was also associated with increased AFP levels; however, the association lacked sufficient statistical significance (**Figure 2F**, p=0.303). These results point to the clinical significance of PRC1 as a biomarker for HCC diagnosis and prognostic evaluation. Further correlation analyses in both cohorts showed that a robust expression level of PRC1 was not associated with the degree of HCC aggressiveness, including differentiation, tumor multiplicity, and neoplasm staging (Supplementary Figure 1B-H).

Simultaneous detection of PRC1 and PRC2 expression is a better predictor of HCC prognosis

To further evaluate the clinical relationship between PRC1 and PRC2 in HCC, we compared PRC1 IHC results to previous EZH2 expression data [8]. A trend toward lower 5-year overall and tumor-free survival rates was observed in groups with stronger expression of either EZH2 or CBX8 compared to under-expressing patients, but the association lacked sufficient statistical significance (**Figure 3A**). Significantly reduced 5-year overall (**Figure 3A**, p=0.000, log-rank test) and tumor-free survival rates (**Figure 3B**, p=0.001, log-rank test) were observed in patients who simultaneously expressed CBX8 and EZH2 compared to CBX8/EZH2 under-expressing patients, especially at the early post-operative stage. Additionally, the serum level of AFP was significantly elevated in the EZH2 and CBX8 co-expressing group compared to the EZH2/CBX8 under-expressing group (**Figure 3C**, p=0.000). Similarly, the overall survival (**Figure 3D**, p=0.015, log-rank test) and tumor-free survival (**Figure 3E**, p=0.009, log-rank test) in the group with elevated expres-

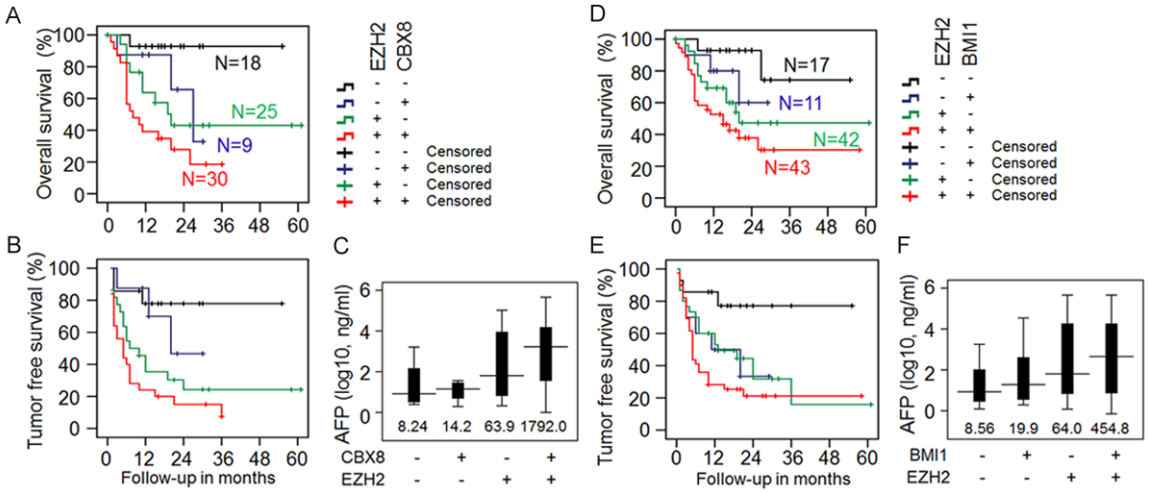


Figure 3. Simultaneous expression of EZH2 and CBX8 is useful for predicting poor HCC prognosis. A and B. Kaplan-Meier curves for overall and tumor-free survival in HCC patients with EZH2 and/or CBX8 hyper-expression or under-expression. C. Serum AFP levels of HCC patients with EZH2 and/or CBX8 hyper-expression (+) or under-expression (-). D and E. Kaplan-Meier curves for overall and tumor-free survival in HCC patients with EZH2 and/or BMI1 hyper-expression or under-expression. F. Serum AFP levels of HCC patients with EZH2 and/or BMI1 hyper-expression (+) or under-expression (-). The line in each panel indicates the median.

sion of both EZH2 and BMI1 was significantly lower than those of patients whose cancerous tissue showed lower expression of EZH2 or BMI1. The expression of AFP with high expression of both EZH2 and BMI1 in cancerous tissue was significantly higher than that with high expression of either EZH2 or BMI1 alone (**Figure 3F**, p=0.029). However, the overall and tumor-free survival showed no significant difference between the group with both BMI1 and CBX8 hyper-expression and the BMI1 and CBX8 hypo-expressing group (Supplementary Figure 2A and 2B). The expression of AFP in these two groups also showed no statistically significant difference (Supplementary Figure 2C). Collectively, our results argue that the combination of PRC1 and PRC2, but not the combination of CBX8 and BMI1, has a significant predictive/prognostic value for HCC patients.

The mutational profile of CBX8 in HCC

As for *EZH2*, we performed a mutation screening of the entire *CBX8* locus in HCC as well by exon-sequencing (**Figure 4A**). Among the 110 examined liver cancer specimens, 1 sample showed a homozygous C-to-T mutation of base 767 in the 5th exon of the *CBX8* gene (**Figure 4B**). In addition, in 13 cases, the 5th exon of CBX8 showed a homozygous G-to-T mutation in base 950 (corresponding to G317V); 56 additional cases were heterozygous for the G950T mutation (**Figure 4C**), which is the functional SNP rs4889891. Comparing the CBX8 amino acid sequence with those of other species revealed that this amino acid is highly conserved (**Figure 4D**). Correlation analysis of the survival curve and of AFP expression with *CBX8* mutation was performed for 110 liver cancer patients. No correlation with tumor-free survival or overall survival was found with *CBX8* mutation in the cancer tissue of the patients (**Figure 4E** and **4F**). AFP expression was clearly higher in patients with *CBX8* mutation than in patients without *CBX8* mutation (**Figure 4G**). This result showed that *CBX8* mutation was not clearly associated with the prognosis of HCC patients but that it had some significance for the diagnosis.

We further analyzed the distribution of the rs4889891 SNP in the 110 HCC cases and in 355 controls (Supplementary Tables 2, 3, 4 and 5). The alleles with the highest distribution frequency at *CBX8* rs4889891 in cancer tissue were homozygous C/C. Individuals carrying the C/A genotype at rs4889891 in the cancer tissue showed a 1.553-fold (95% CI: 0.949-2.539) higher risk of HCC than individuals carrying the wild-type alleles. Quantitative trait analysis showed no significant associations between *CBX8* rs48-89891 and symptom severity, including overall and tumor-free survival (P>0.05).

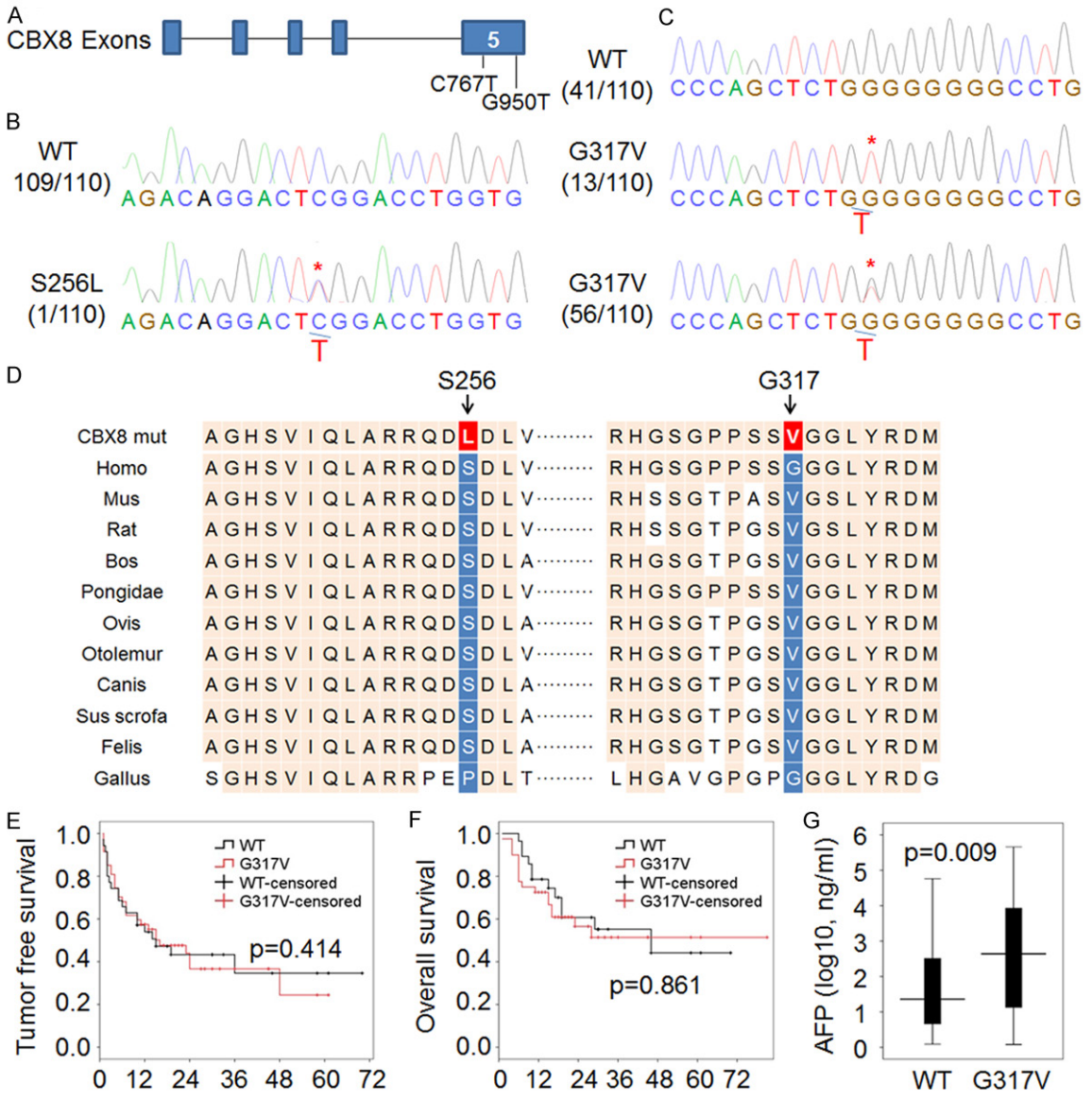


Figure 4. A screening for *CBX8* mutations in HCC. (A) *CBX8* locus showing exons structure. The mutations in exon 5 are shown. Illustration of sequencing results. Of 110 HCC samples, 1 showed a distinctly homozygous mutation in codon 256 causing an amino acid replacement (B) and 13 showed distinct homozygous mutations and 56 showed distinct heterozygous mutations at codon 317, causing an amino acid replacement (C). (D) A multiple alignment of CBX8 from eleven species. (E-G) Kaplan-Meier curves for tumor-free (E) and overall survival (F) and Serum AFP levels (G) in HCC patients with CBX8 WT and G317V. The line in each panel indicates the median.

Discussion

As evidence of human genome evolution, SNPs represent a type of genetic mutation that has become the third generation of genetic markers. In complex diseases such as cancer, some minor alleles in SNPs act as risk factors, whereas others are protective. Unlike SNPs, characterizing somatic mutations and their functional consequences remains a formidable challenge due to their low individual abundance.

Because of the importance of the PcG family in epigenetic regulation, its misexpression is closely associated with liver carcinogenesis [8]. However, the genetic mutation profiles of PcG genes in HCC remain unclear. Yu *et al.* observed that one polymorphic C allele at SNPs rs6950683 and rs3757441 was strongly associated with a lower risk of HCC [10]. In the present study, we discovered another highly conserved mutation site in HCC, G553C, which is also the SNP site rs2302427. Strikingly, we

found that HCC patients carrying the minor G553C allele gain a significant survival advantage compared with those carrying the wild-type allele. The rs2302427 variant was found in both liver cancer and in paracancerous tissues; it was significantly associated with overall and tumor-free survival when found in paracancerous tissues but not when found in liver cancer tissue. This unexpected finding might be related to the distribution of rs2302427. In fact, the distributions of this SNP in cancer tissue and paracancerous tissues were different. It seems the SNP may convert into a mutation site depending on the cellular context. An interesting finding is that the tumor survival rate in G carriers is higher than in wild-type C carriers. Notably, the above results clearly confirm that the G553C point mutation of *EZH2* in liver cancer is a meaningful and protective genetic change that has significant clinical diagnostic and prognostic value. Basically, each single SNP is uniform within an individual. Unexpectedly, we observed a different distribution of rs2302427 in cancerous tissues and adjacent tissues, suggesting that the G553C mutation may be both a SNP and a somatic mutation, a phenomenon that has not been reported in previous studies. It further suggested that the G allele plays an important role in the transition from benign hyperplasia to cancer. Our results show that the rs2302427 site might be a protective SNP that plays an important antagonistic role in the process of liver cancer development; it may thus provide a potential genotype and allele marker to classify patients for different treatment strategies. Whether this mutation initiates consecutive protective mechanisms against tumorigenesis, such as apoptosis, autophagy, and/or senescence, requires further exploration. Furthermore, how the G553C mutation affects the function of the EZH2 protein remains an open question. Our immunoprecipitation (IP) data showed that the G553C mutation did not affect the interaction between EZH2 and DNMT1 despite the mutation site within the DNMT1-interacting region (Supplementary Figure 3). In addition to the PRC2, we also found frequent mutations encoding the G317V (rs4889891) substitution in the CBX8 protein of PRC1 in HCC. Although there was no association between rs4889891 and HCC, and although this mutation had no statistically significant association with prognosis indicators such as tumor-free survival, it did show a strong correlation with the AFP level, which is generally considered as a diagnostic marker of HCC.

In addition to genomic mutations, abnormal expression of PcG proteins is often seen in different types of cancer. In our previous findings, the expression level of the EZH2 protein was markedly increased in HCC, and there was significantly lower overall and tumor-free survival in EZH2-over-expressing HCC patients [8]. In the present study, we found that PRC1 expression levels are also involved in the clinical progression and outcome of HCC patients. As with EZH2, the expression of CBX8 and BMI1 was stimulated in 43.1% and 46.2% of HCC tissues, respectively. A trend toward lower rates of overall and tumor-free survival was observed in cases with the over-expression of CBX8 or BMI1 compared with patients with lower expression levels. Notably, patients with high levels of PRC1 and PRC2 co-expression showed markedly decreased overall survival and tumor-free survival rates, especially at the early post-operative stage, compared to patients with hyper-expression of only one of these proteins. In contrast, over-expression of EZH2 or CBX8 was associated with a very positive outcome for HCC patients. The overall and tumor-free survival rates in such patients were maintained at 90% and 80%, respectively, at 60 months after operation. These data strongly suggest that a combination of PRC1 and PRC2 expression might have a significant prognostic value for HCC patients. It appears that both PRC1 and PRC2 play important roles in promoting HCC but that their target gene networks are relatively independent. According to our previous results, PRC1 (CBX8 and BMI1) and PRC2 (EZH2 and SUZ12) displayed very different gene regulation profiles in HCC [9]. Microarray results indicated that numerous genes are exclusively targeted by either PRC2 or PRC1. Only 220 genes in HCC cells were significantly co-regulated by EZH2, SUZ12, BMI1, or CBX8 [9]. In addition, the PcG family regulates gene expression in both an H3K27me3-dependent and an H3K27me3-independent manner [9]. It appears that both PRC1 and PRC2 play important roles in promoting HCC and that their target gene networks are relatively independent. Multivariate Cox regression analysis also indicated that PRC1 and PRC2 are independent prognostic markers for HCC (Supplementary Tables 5 and 6). Because of their relatively independent pathways, we propose that the

combination of PRC1 and PRC2 could act as an even more precise biomarker for predicting HCC outcomes. The detailed mechanism must be dissected to identify cooperative mechanisms between PRC1 and PRC2 in HCC. Conversely, the combination of CBX8 and BMI1 did not display a significant predictive/prognostic value for HCC. Our previous findings indicated that CBX8 and BMI1 share the vast majority of target genes in HCC [9]. Here, we found that the protein expression levels of CBX8 and BMI1 are significantly correlated in HCC specimens (Supplementary Table 7). Therefore, we confirmed that the PRC1 members CBX8 and BMI1 are mutually overlapping diagnostic targets.

We further wondered whether robust expression of EZH2 could be affected by *EZH2* mutation. However, the results showed that EZH2 overexpression was not significantly associated with *EZH2* mutation (Supplementary Table 8), indicating that abnormal EZH2 expression is not affected by mutation. Combining all these data, altered protein expression and genetic mutation of PcG are two independent events that occur simultaneously in liver carcinogenesis. Therefore, to verify the significance of PcG in cancer progression, not only protein expression but also genetic mutation should be considered.

Acknowledgements

This work was supported by the grants from the National Natural Science Foundation of China (81472020 to S.G, 31071206 to L.X, 91229111 and 81272719 to G.J.).

Disclosure of conflict of interest

None.

Address correspondence to: Li-Xiang Xue, Medical Research Center, Department of Radiation Oncology, Peking University Third Hospital, 49 Huayuan North Road, Beijing 100191, P.R. China. Tel: +86-10-82265123; E-mail: lixiangxue@hsc.pku.edu.cn; Guang Hui Jin, Medical College of Xiamen University, Chengzhi Building 110, Xiang'an South Road, Xiamen, Fujian 361102, P.R. China. Tel: +86-592-218-6173; Fax: +86-592-218-1535; E-mail: ghjin@xmu.edu.cn

References

[1] Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB, Calderaro J, Bioulac-Sage P, Letexier M, Degos F, Clement B, Balabaud C, Chevet E, Laurent A, Couchy G, Letouze E, Calvo F and Zucman-Rossi J. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat Genet 2012; 44: 694-698.

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

[3] Giordano S and Columbano A. MicroRNAs: new tools for diagnosis, prognosis, and therapy in hepatocellular carcinoma? Hepatology 2013; 57: 840-847.

[4] Au SL, Wong CC, Lee JM, Fan DN, Tsang FH, Ng IO and Wong CM. Enhancer of zeste homolog 2 epigenetically silences multiple tumor suppressor microRNAs to promote liver cancer metastasis. Hepatology 2012; 56: 622-631.

[5] Huang J, Wang Y, Guo Y and Sun S. Down-regulated microRNA-152 induces aberrant DNA methylation in hepatitis B virus-related hepatocellular carcinoma by targeting DNA methyltransferase 1. Hepatology 2010; 52: 60-70.

[6] Herceg Z and Paliwal A. Epigenetic mechanisms in hepatocellular carcinoma: how environmental factors influence the epigenome. Mutat Res 2011; 727: 55-61.

[7] Xu B, Li SH, Zheng R, Gao SB, Ding LH, Yin ZY, Lin X, Feng ZJ, Zhang S, Wang XM and Jin GH. Menin promotes hepatocellular carcinogenesis and epigenetically up-regulates Yap1 transcription. Proc Natl Acad Sci U S A 2013; 110: 17480-17485.

[8] Gao SB, Xu B, Ding LH, Zheng Q, Zhang L, Zheng QF, Li SH, Feng ZJ, Wei J, Yin ZY, Hua X and Jin GH. The functional and mechanistic relatedness of the EZH2 and menin in hepatocellular carcinoma. J Hepatol 2014; 61: 832-839.

[9] Gao SB, Zheng QF, Xu B, Pan CB, Li KL, Zhao Y, Zheng QL, Lin X, Xue LX and Jin GH. EZH2 represses target genes through H3K27-dependent and H3K27-independent mechanisms in hepatocellular carcinoma. Mol Cancer Res 2014; 12: 1388-1397.

[10] Yu YL, Su KJ, Hsieh YH, Lee HL, Chen TY, Hsiao PC and Yang SF. Effects of EZH2 polymorphisms on susceptibility to and pathological development of hepatocellular carcinoma. PLoS One 2013; 8: e74870.

[11] Cai MY, Tong ZT, Zheng F, Liao YJ, Wang Y, Rao HL, Chen YC, Wu QL, Liu YH, Guan XY, Lin MC, Zeng YX, Kung HF and Xie D. EZH2 protein: a promising immunomarker for the detection of hepatocellular carcinomas in liver needle biopsies. Gut 2011; 60: 967-976.

[12] Yonemitsu Y, Imazeki F, Chiba T, Fukai K, Nagai Y, Miyagi S, Arai M, Aoki R, Miyazaki M, Nakatani Y, Iwama A and Yokosuka O. Distinct expression of polycomb group proteins EZH2 and BMI1 in hepatocellular carcinoma. Hum Pathol 2009; 40: 1304-1311.

[13] Sasaki M, Ikeda H, Itatsu K, Yamaguchi J, Sawada S, Minato H, Ohta T and Nakanuma Y. The overexpression of polycomb group proteins Bmi1 and EZH2 is associated with the progression and aggressive biological behavior of hepatocellular carcinoma. Lab Invest 2008; 88: 873-882.

[14] Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS and Zhang Y. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 2002; 298: 1039-1043.

[15] Cao Q, Yu J, Dhanasekaran SM, Kim JH, Mani RS, Tomlins SA, Mehra R, Laxman B, Cao X, Kleer CG, Varambally S and Chinnaiyan AM. Repression of E-cadherin by the polycomb group protein EZH2 in cancer. Oncogene 2008; 27: 7274-7284.

[16] Fujii S, Ito K, Ito Y and Ochiai A. Enhancer of zeste homologue 2 (EZH2) down-regulates RUNX3 by increasing histone H3 methylation. J Biol Chem 2008; 283: 17324-17332.

[17] Kotake Y, Cao R, Viatour P, Sage J, Zhang Y and Xiong Y. pRB family proteins are required for H3K27 trimethylation and Polycomb repression complexes binding to and silencing p16INK4alpha tumor suppressor gene. Genes Dev 2007; 21: 49-54.

[18] Zhou W, Zhu P, Wang J, Pascual G, Ohgi KA, Lozach J, Glass CK and Rosenfeld MG. Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. Mol Cell 2008; 29: 69-80.

[19] Dietrich N, Bracken AP, Trinh E, Schjerling CK, Koseki H, Rappsilber J, Helin K and Hansen KH. Bypass of senescence by the polycomb group protein CBX8 through direct binding to the INK4A-ARF locus. EMBO J 2007; 26: 1637-1648.

[20] Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, Paul JE, Boyle M, Woolcock BW, Kuchenbauer F, Yap D, Humphries RK, Griffith OL, Shah S, Zhu H, Kimbara M, Shashkin P, Charlot JF, Tcherpakov M, Corbett R, Tam A, Varhol R, Smailus D, Moksa M, Zhao Y, Delaney A, Qian H, Birol I, Schein J, Moore R, Holt R, Horsman DE, Connors JM, Jones S, Aparicio S, Hirst M, Gascoyne RD and Marra MA. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet 2010; 42: 181-185.

[21] Zhang J, Ding L, Holmfeldt L, Wu G, Heatley SL, Payne-Turner D, Easton J, Chen X, Wang J, Rusch M, Lu C, Chen SC, Wei L, Collins-Underwood JR, Ma J, Roberts KG, Pounds SB, Ulyanov A, Becksfort J, Gupta P, Huether R, Kriwacki RW, Parker M, McGoldrick DJ, Zhao D, Alford D, Espy S, Bobba KC, Song G, Pei D, Cheng C, Roberts S, Barbato MI, Campana D, Coustan-Smith E, Shurtleff SA, Raimondi SC, Kleppe M, Cools J, Shimano KA, Hermiston ML, Doulatov S, Eppert K, Laurenti E, Notta F, Dick JE, Basso G, Hunger SP, Loh ML, Devidas M, Wood B, Winter S, Dunsmore KP, Fulton RS, Fulton LL, Hong X, Harris CC, Dooling DJ, Ochoa K, Johnson KJ, Obenauer JC, Evans WE, Pui CH, Naeve CW, Ley TJ, Mardis ER, Wilson RK, Downing JR and Mullighan CG. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. Nature 2012; 481: 157-163.

[22] Wu Y, Feng ZJ, Gao SB, Matkar S, Xu B, Duan HB, Lin X, Li SH, Hua X and Jin GH. Interplay between menin and K-Ras in regulating lung adenocarcinoma. J Biol Chem 2012; 287: 40003-40011.

Supplementary materials and methods

Cell culture and gene transfection

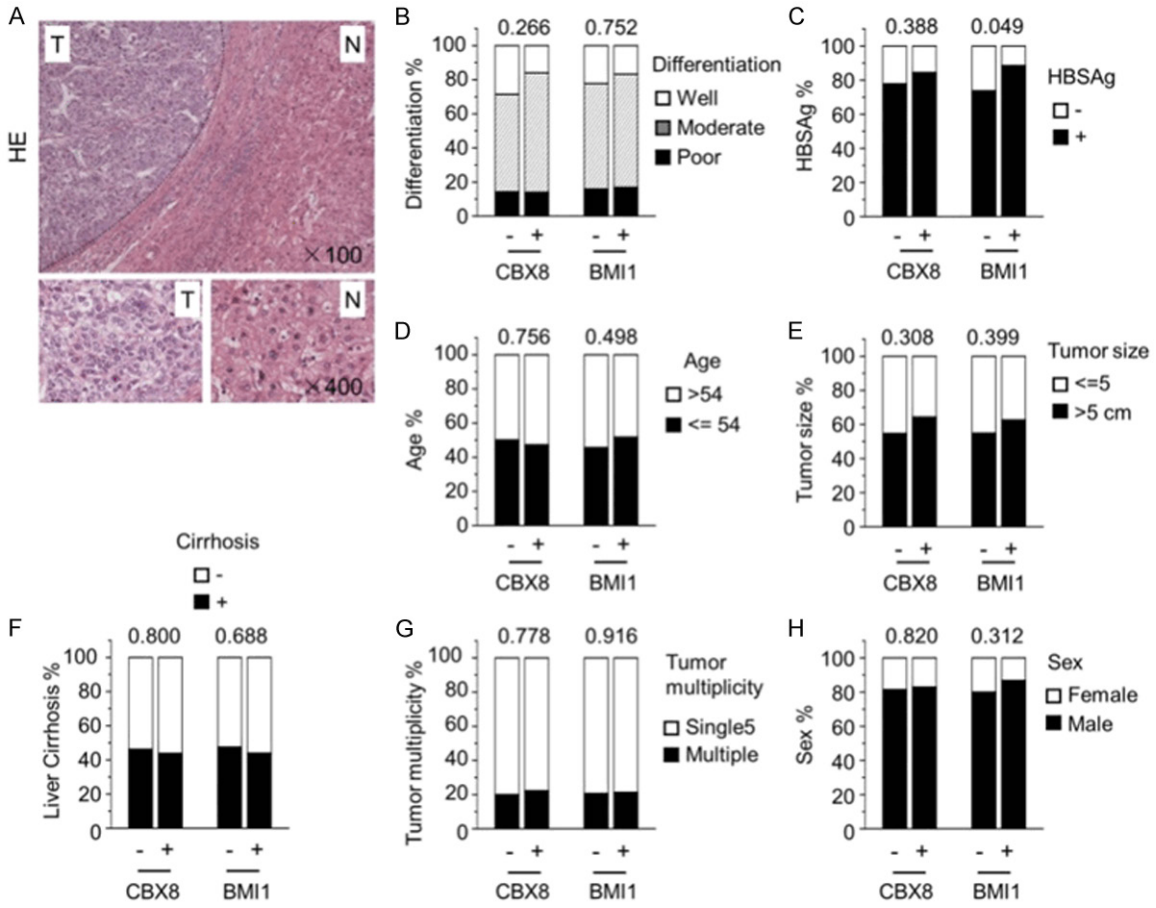
HepG2 was obtained from the Type Culture Collection of the Chinese Academy of Sciences (Shanghai, China). Cell line was cultured in Dulbecco's modified Eagle medium (DMEM; Gibco, Invitrogen) supplemented with 10% fetal bovine serum (FBS; Hyclone) and 1% penicillin-streptomycin. EZH2, EZH2-D185H, EZH2-Y641N were constructed with the pLNCX2 retroviral vector (631510, Clontech). Recombinant retroviruses were packaged using GP2-293 cells according to the BD Retro-XTM Universal Packaging System protocol (BD Biosciences Clontech). Add 1 mL recombinant retroviruses and 1 mL DMEM to each corresponding 60 mm dish. After the virus has been added to the dish, then add polybrene to a final concentration of 10 μg/mL. The transfected cells were selected with either 2 μg/ml puromycin (Hyclone) for 1 week, and selected stable cells were continuously cultured until harvest for analysis.

Western blotting and IP assays

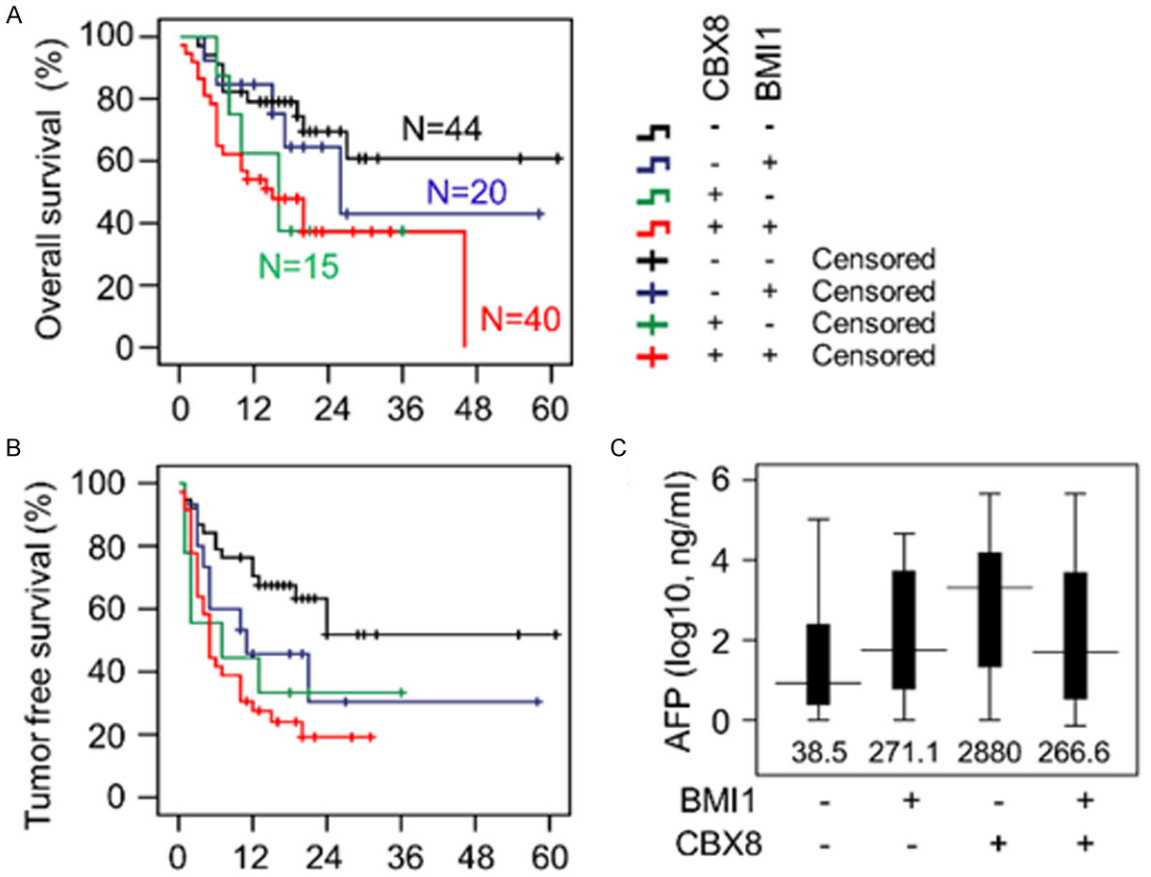
IP assays were performed essentially as previously described [1]. Briefly, the harvested cells were lysed and incubated with a primary antibody or with normal mouse IgG (Millipore). Protein A-agarose beads (Santa Cruz) were then added. Bound proteins were collected by centrifugation and analyzed by Western blotting. Western blotting was performed as previously described [1]. The antibodies are listed in Supplementary Table 12.

Supplementary reference

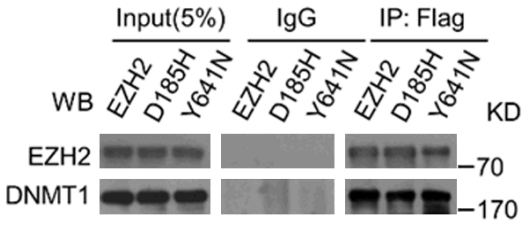
[1] Wu Y, Feng ZJ, Gao SB, Matkar S, Xu B, Duan HB, Lin X, Li SH, Hua X, Jin GH. Interplay between menin and K-Ras in regulating lung adenocarcinoma. J Biol Chem 2012; 287: 40003-11.



Supplementary Figure 1. Correlation analysis of PRC1 expression and clinicopathological features in HCC. A. H&E staining in HCC paraffin sections. The dotted lines indicate juncture of tumor (T) and adjacent tissue(N). B and C. The proportion of differentiation and HBSAg in HCC patients with CBX8 or BMI1 hyper-expression (+) and under-expression (-). D and E. The proportion of age and tumor size in HCC patients with CBX8 or BMI1 hyper-expression (+) and under-expression (-). F-H. The proportion of liver cirrhosis, tumor multiplicityand sex in HCC patients with CBX8 or BMI1 hyper-expression (+) and under-expression (-).



Supplementary Figure 2. Correlation analysis of PRC1 expression andHCC survival. A and B. Kaplan-Meier curves for overall and tumor-free survival in HCC patients with CBX8 and/or BMI1 hyper-expression and under-expression. C. Serum AFP levels of HCC patients with CBX8 and/or BMI1 hyper-expression (+) and under-expression (-). In each panel, the line indicates the median.



Supplementary Figure 3. IP assays for interaction between DNMT1 and EZH2. EZH2-WT, EZH2-D185H and EZH2-Y641N overexpressed HepG2 cells lysates were immunoprecipitated for flag, and the immunoprecipitates wereanalyzed by Western blotting (WB) for the presence of EZH2 and DNMT1.

Supplementary Table 1. Demographic Characteristics of Controls and patients with HCC

Variable	Controls (N=311)	Patients (N=110)	*p* Value
Age	Mean ± SD	Mean ± SD	
	58.3±13.63	53.98±12.19	>0.05
Gender	%	%	
Male	40.84%	78.43%	<0.01
Female	59.16%	21.57%	
Stage			
I		3 (2.94%)	
II		61 (59.80%)	
IIIa		18 (17.65%)	
IIIb		14 (13.73%)	
IV		6 (5.88%)	

Supplementary Table 2. Distribution of CBX8 Genotypes and Allelic Frequencies in HCC and control

CBX8 rs4889891	Genotypes			Allelic Frequency	
	CC	CA	AA	C	A
HCC	33	59	13	0.59	0.41
Controls	119	137	49	0.61	0.39
	χ^2=3.99, d.f.=2, *p*=0.135			χ^2=0.25, d.f.=1, *p*=0.61	

Supplementary Table 3. Odds Ratio with 95% CI of CBX8 Gene in HCC Patients

Genotypes		Odds Ratio (CI 95%)	*p* value
CBX8 rs4889891	A vs C	1.085 (0.787-1.495)	0.617
	CA vs CC	1.553 (0.949-2.539)	0.0461
	AA vs CC	0.956 (0.464-1.971)	0.363

Supplementary Table 4. Quantitative traits analysis of CBX8 Genotypes and Allelic

CBX8 rs4889891	Genotypes				Allelic Frequency		
	CC	CA	AA	P	C	A	P
Overall survival	17.58	11.86	17.0	0.284	14.88	13.44	0.562
Tumor free survival	15.22	11.82	12.23	0.637	13.62	11.95	0.468
AFP	1.25	1.90	2.02	0.161	1.556	1.933	0.117

Supplementary Table 5. The multivariate Cox regression model for overall survival rates

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
EZH2	-0.992	0.62	2.564	1	0.109	0.371	0.11	1.249
CBX8	-1.239	0.474	6.841	1	0.009	0.29	0.114	0.733
LogAFP	0.339	0.144	5.556	1	0.018	1.403	1.059	1.86
Differentiation	0.781	0.405	3.71	1	0.054	2.184	0.986	4.834
HBSAG	-0.136	0.528	0.067	1	0.796	0.873	0.31	2.456
Age	0.697	0.503	1.921	1	0.166	2.008	0.749	5.379
Sex	-1.492	0.808	3.408	1	0.065	0.225	0.046	1.096

Supplementary Table 6. The multivariate Cox regression model for tumor free survival rates

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
EZH2	1.482	0.509	8.493	1	0.004	4.403	1.625	11.932
CBX8	0.72	0.359	4.033	1	0.045	2.054	1.017	4.148
LogAFP	0.249	0.125	3.952	1	0.047	1.283	1.004	1.639
Differentiation	0.951	0.316	9.081	1	0.003	2.589	1.395	4.807
HBSAG	-0.129	0.439	0.087	1	0.768	0.879	0.372	2.076
Age	-0.209	0.352	0.351	1	0.554	0.812	0.407	1.619
Sex	1.444	0.626	5.312	1	0.021	4.236	1.241	14.458

Supplementary Table 7. Correlation analysis of CBX8 and BMI1 protein expression in HHC

		BMI1		
		-	+	Total
CBX8	-	44	20	64
	+	15	40	55
Total		59	60	119

Chi-Square: P=0.000.

Supplementary Table 8. Correlation analysis between rs2302427 and EZH2 protein expression in HHC

EZH2 rs2302427	Genotypes			Allelic Frequency	
	CC	CG	GG	C	G
EZH2 (-)	10	8	1	0.737	0.263
EZH2 (+)	37	16	2	0.818	0.182
	χ^2=6.13, d.f.=1.306, p=0.52			χ^2=1.156, d.f.=1, p=0.28	

Supplementary Table 9. Correlation analysis of CBX8 expression and clinicopathological features in HCC

Variable	CBX8 expression			
	All cases	Negative	Positive	*p* Value*
Age (years)				0.756
≤53.13	60	35 (58.33%)	25 (41.67%)	
>53.13	63	35 (55.56%)	28 (44.44%)	
Sex				0.820
Male	101	57 (56.44%)	44 (43.56%)	
Female	22	13 (59.09%)	9 (40.91%)	
HBSAg				0.388
Yes	87	49 (56.32%)	38 (43.68%)	
No	21	14 (66.67%)	7 (33.33%)	
A-Fetoprotein (ng/ml)				0.018
≤20	40	29 (72.5%)	11 (27.5%)	
>20	67	33 (49.25%)	34 (50.75%)	
Liver cirrhosis				0.800
Yes	62	35 (56.45%)	27 (43.55%)	
No	51	30 (58.82%)	21 (41.18%)	
Tumor size (cm)				0.308
≤5	45	29 (64.44%)	16 (35.56%)	
>5	64	35 (54.69%)	29 (45.31%)	
Tumor multiplicity				0.778
Single	87	52 (59.77%)	35 (40.23%)	
Multiple	23	13 (56.52%)	10 (43.48%)	
Differentiation				0.266
Well	26	18 (69.23%)	8 (30.77%)	
Moderate	71	36 (50.70%)	35 (49.30%)	
Poor	16	9 (56.25%)	7 (43.75%)	
Relapse				0.001
Yes	63	26 (41.27%)	37 (58.73%)	
No	39	29 (74.36%)	10 (25.64%)	

*χ^2 test.

Supplementary Table 10. Correlation analysis of BMI1 expression and clinicopathological features in HCC

Variable	BMI1 expression			
	All cases	Negative	Positive	*p* Value*
Age (years)				0.498
≤53.39	63	32 (50.79%)	31 (49.21%)	
>53.39	67	38 (56.72%)	29 (44.44%)	
Sex				0.312
Male	108	56 (51.85%)	52 (48.15%)	
Female	22	14 (63.64%)	8 (36.36%)	
HBSAg				0.049
Yes	91	45 (49.45%)	46 (50.55%)	
No	22	16 (72.73%)	6 (27.27%)	
A-Fetoprotein (ng/ml)				0.459
≤20	41	24 (58.54%)	17 (41.46%)	
>20	74	38 (51.35%)	36 (48.65%)	
Liver cirrhosis				0.688
Yes	64	32 (50.00%)	32 (50.00%)	
No	54	29 (53.70%)	25 (46.30%)	
Tumor size (cm)				0.399
≤5	49	28 (57.14%)	21 (42.86%)	
>5	69	34 (49.28%)	35 (50.72%)	
Tumor multiplicity				0.916
Single	94	50 (53.19%)	44 (46.81%)	
Multiple	25	13 (52.00%)	12 (48.00%)	
Differentiation				0.752
Well	23	14 (60.87%)	9 (39.13%)	
Moderate	75	39 (52.00%)	36 (48.00%)	
Poor	19	10 (52.63%)	9 (47.37%)	
Relapse				0.003
Yes	66	26 (39.39%)	40 (60.61%)	
No	42	29 (69.05%)	13 (30.95%)	

*χ^2 test.

Supplementary Table 11. The list of primers used in the study

Primer	Forward primer	Reverse primer	PCR prod. (bp)	Annealing temperature (°C)
EZH2 P1	CACAGGTGATCATATTCAGGCT	TGAGGTCAATGATTTCCTCCCA	980	58
EZH2 P2	GCTACAGCTTAAGGTTGTCCT	ACTGTCTTGATTCACCTTGACA	270	58
EZH2 P3	AGAGTTTGCTCTCTGACTAGT	ACCTTAGCCTCCCAAGTGCT	140	55
EZH2 P4	GCTTCCCAGTGCTCTTAAAGC	TGTAGTGGCTCATCCGCTAC	430	58
EZH2 P5	GTAGCGGATGAGCCACTACA	CAGAGCAATCCTCAAGCAACA	796	58
EZH2 P6	TTCAGTAAGAGCCTGAAGGA	ATCACCTCCACCAAAGTGCA	310	58
EZH2 P7	GAACTTTGCCCTGATGTTGAC	ACCAACAACAGCCCTTAGGA	1500	55
EZH2 P8	GTTAAGCAGCTCTCTGTTGG	CAAGGGTGCATTACCCAGAG	1740	55
EZH2 P9	GTCTACTTTGTCCCCAGTCC	ACCCTCGTTTCTGAACACTC	1535	55
EZH2 P10	CCCTGAAGAACTGTAACCAG	CACTGGTGTCAGTGAGCATG	310	58
EZH2 P11	CATGCTCACTGACACCAGTG	GTCCAGAGTTCACAATCCAG	163	58
EZH2 P12	GTCTCAGCACATGTTGGATG	CTGGGTAAGCCCTCAGGATG	605	58
CBX8 P1	GAGTATTATGGTCTGTCGTGCG	GCATCGCTGCAAGTCTAAGG	325	58
CBX8 P2	TCTTGTATCCTCGTCATGCG	GCAAGTGGGACTCTAGTCCAG	435	58
CBX8 P3	CTGGACTAGAGTCCCACTTGC	GTTGGTGCTGCTACTGACTTG	295	58
CBX8 P4	CCACTCTTGAGACCATCTTCTC	GCCAGCTGGATCACACTGTG	535	58
CBX8 P5	CCTCGGAGAGTACCTCAAGG	CACTATGCCAAGCTCACGCTC	570	58

Supplementary Table 12. The list of antibodies used in the study

Antibodies			
Western blotting, IF			
Antigen	Host	Company	Cat#/code
GAPDH	rabbit	Epitomics	#2251-1
EZH2	rabbit	Cell signaling	#5246
SUZ12	rabbit	Cell signaling	#3737
CBX8	rabbit	Sigma	R30545
BMI1	rabbit	Epitomics	#5590-1
DNMT1	rabbit	Epitomics	#3147-1