# ORIGINAL RESEARCH

In Vivo Genome-Wide CRISPR Activation Screening Identifies Functionally Important Long Noncoding RNAs in Hepatocellular Carcinoma

Lok-Sze Wong,[1,*] Lai Wei,[1,*] Gengchao Wang,[1] Cheuk-Ting Law,[1] Felice Ho-Ching Tsang,[1] Wai-Ching Chin,[1] Irene Oi-Lin Ng,[1] and Chun-Ming Wong[1]

[1]The State Key Laboratory of Liver Research, Department of Pathology, Li-Ka Shing Faculty of Medicine, The University of Hong Kong
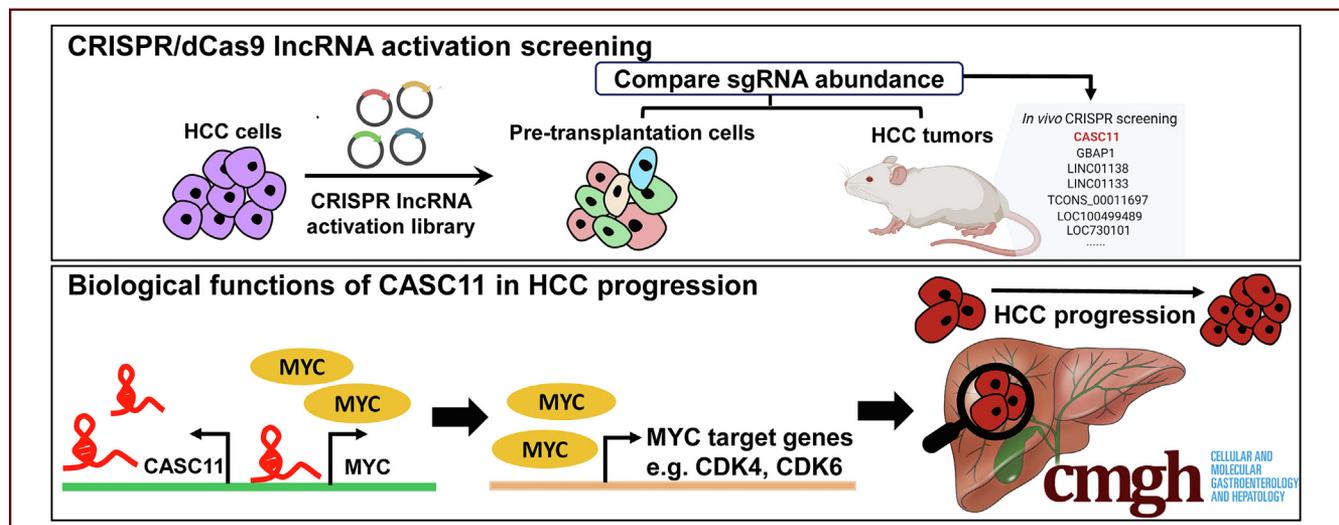
SUMMARY

Our integrated transcriptomic data and functional CRISPR–based screening identified functional long noncoding RNAs that are frequently overexpressed in hepatocellular carcinoma. Cancer Susceptibility 11 (CASC11), the top candidate in the library screening, is shown to regulate the expression of MYC proto-oncogene in a *cis*-regulatory manner, which consequently dysregulates a subset of cell-cycle regulators and drives hepatocellular carcinoma tumor growth.

BACKGROUND & AIMS: Long noncoding RNAs (lncRNAs) are found to have profound impacts on diverse cellular processes. Although high-throughput sequencing studies have shown the differential lncRNA expression profiles between hepatocellular carcinoma (HCC) and nontumor livers, the functional impacts of lncRNAs on HCC development await further investigation. Herein, we sought to address the functional roles of lncRNAs in HCC pathogenesis by in vivo functional screening.

METHODS: We performed genome-wide clustered regularly interspaced short palindromic repeats (CRISPR)/dead CRISPR-associated protein 9 (dCas9) lncRNA activation screening in HCC xenografts. We characterized the clinical relevance of positively selected lncRNAs using transcriptomic data sets. We

used CRISPR-based gene activation and knockdown approaches to show the functional roles of positively selected lncRNAs including Cancer Susceptibility 11 (CASC11) in HCC. RNA sequencing and chromatin isolation by RNA purification sequencing were used to investigate the molecular mechanisms of CASC11 in HCC progression.

RESULTS: The in vivo functional screening identified 1603 positively selected lncRNAs, 538 of which were overexpressed in HCC patients. Systematic transcriptomic data analysis and clinical investigation showed that patients with high expression of these lncRNA candidates correlated with aggressive tumor behaviors. Overexpression of these lncRNAs aggravated HCC cell growth. Detailed characterization of a lncRNA candidate, CASC11, showed its pivotal role in cell proliferation and tumor growth. Mechanistically, chromatin isolation by RNA purification sequencing showed that CASC11 was bound to the CASC11/MYC proto-oncogene shared promoter region on chromosome 8q24. CASC11 modulated the transcriptional activity of MYC in a *cis*-regulatory manner, which affected the expression of MYC downstream target genes, consequently promoting G1/S progression.

CONCLUSIONS: Our study showed the power of in vivo CRISPR screening, which comprehensively investigated the functionality of lncRNAs in HCC progression, providing a rationale for targeting these lncRNAs clinically. *(Cell Mol*

Hepatocellular carcinoma (HCC) is a highly aggressive primary liver malignancy that frequently occurs in patients with liver cirrhosis. HCC causes more than 800,000 deaths annually as a result of delayed diagnosis and lack of effective therapeutic interventions.[1] Therefore, a better understanding of the underlying mechanisms of HCC development may provide valuable insights into the development of new predictive biomarkers and treatment strategies for HCC patients.

HCC development traditionally has been associated with the mutations of oncogenes and tumor-suppressor genes. However, beyond the protein-coding genome, the importance of the noncoding genome in HCC development has not been investigated sufficiently. The power of high-throughput sequencing technology has made a serendipitous discovery of a vast landscape of regulatory elements in the human genome.[2] More than 80% of the human genome is actively transcribed into a myriad of noncoding RNAs. Among these, long noncoding RNAs (lncRNAs) represent the largest class of noncoding RNA species in the human transcriptome. LncRNAs, once considered as transcriptional noise, have shown their tremendous biological importance in chromatin remodeling, gene transcription, and RNA processing.[3] Aberrant expression of lncRNAs has been shown to have profound effects on cancer hallmarks, including sustaining proliferative signaling; inhibiting cell death and growth suppressors; inducing immortality, angiogenesis, invasion, and metastasis; rewriting the metabolic program; and evading the immune system.[4]

Although the differential expression of lncRNAs in human cancers has been accomplished by RNA-sequencing (RNA-seq) or microarray studies, such approaches do not provide evidence suggesting which lncRNAs are functional, let alone what roles they play in cancer development. Recent genome-wide functional screening approaches such as clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated protein 9 (Cas9) knockout screening have become a popular tool to identify protein-coding genes that are crucial for cancer cell growth and resistance to chemotherapeutic drugs in a high-throughput manner.[5–9] However, the knockout approach attained by generating frameshift mutations within exons is not applicable to lncRNAs lacking open reading frames. For this reason, recent studies have exploited the power of CRISPR/ dead Cas9 (dCas9)-based systems to manipulate the transcription of lncRNAs in in vitro functional screening.[10,11] Previous studies used a cell culture system to identify functional lncRNAs promoting drug resistance. However, the expression of lncRNAs is influenced strongly by various factors where cancer cells reside, including the culture condition, the components of extracellular matrix, as well as the crosstalk between tumor cells and other cells in a tumor.[12] To our knowledge, there currently is no systematic study showing the functions of lncRNAs at a large scale in HCC. Here, to precisely infer the functional roles of lncRNAs in HCC development, we took the advantage of genome-wide CRISPR activation screening in a xenograft mouse model. We identified a group of lncRNA candidates whose up-regulation significantly promoted HCC cell growth. Further gain- and loss-of-function experiments confirmed the oncogenic roles of Cancer Susceptibility 11 (CASC11) in HCC progression by positively regulating the transcription of MYC proto-oncogene in a *cis*-regulatory manner. CASC11 depletion evoked global expression changes of genes involved in MYC signaling and cell-cycle progression. Overall, our study provides a systematic framework that integrates the clinical transcriptomic data with functional CRISPR activation screening to uncover functional lncRNAs that influence HCC tumor growth.

Results

In Vivo Genome-Wide CRISPR Activation Screening Identifies a Subset of HCC Promoting lncRNAs

To identify the functional lncRNAs that promote HCC development, we performed a genome-wide CRISPR/dCas9 lncRNA activation screening in xenograft mouse model (Figure 1*A*). MHCC97H cells expressing dCas9/VP64 and MS2 RNA-binding domain fused to the transcription activators p65 and heat shock factor 1 (MS2-p65-HSF1) were transduced with the human lncRNA activation library, which contains 96,458 single-guide RNAs (sgRNAs) targeting 10,504 lncRNAs, with 10 sgRNAs tiling the 800-bp upstream region of each lncRNA transcriptional start site. The successful infected clones were injected into both flanks of each mouse at a density of 2×10^6 cells subcutaneously. Considering the tumor viability, tumors from 20 mice were pooled for amplicon sequencing to achieve the $800\times$ library representation. sgRNA abundance in tumors and

Figure 1. In vivo genome-wide CRISPR activation screening. (*A*) Schematic representation of genome-wide CRISPR activation screening. (*B*) Scatterplot showing the enrichment of sgRNAs targeting the positively selected lncRNAs (FDR, <0.05) compared with nontargeting sgRNAs in tumors vs pretransplantation cells. (*C*) Pipeline for the identification of functional lncRNA candidates. FPKM, fragments per kilobase per million; MOI, multiplicity of infection.

pretransplantation cells was evaluated by high-throughput sequencing.

More than 90% of sgRNA library constructs were retained in all samples, collectively suggesting that our library sequencing had sufficient library coverage and read depth. We also evaluated the screen performance based on the distribution of nontargeting sgRNA controls, which were clustered mostly at the center of the scatterplot, suggesting that they were neither enriched nor depleted throughout the library screening (Figure 1B). We used the Model-based Analysis of Genome-wide CRISPR/Cas9 Knockout (MAGeCK) algorithm with additional filtering steps to identify positively selected lncRNAs in our library screening (Figure 1C). We hypothesized that if a lncRNA is beneficial for HCC growth in a harsh tumor microenvironment, its corresponding sgRNAs would be enriched in tumors as compared with the pretransplantation cells. A total of 7211 sgRNAs (log_2 fold change > 1, false discovery rate [FDR] < 0.05) were significantly enriched in tumors, while only 3 sgRNAs were significantly depleted in tumors. These 7211 enriched sgRNAs targeted the promoter regions of 5283 lncRNAs, among which 1603 lncRNAs were targeted by at least 2 enriched sgRNAs (FDR < 0.05) (Supplementary Table 1).

To identify high-confidence positively selected lncRNAs, we evaluated the expression levels of lncRNA candidates by using RNA-seq data of HCC/non-tumor (NT) liver samples. Because some lncRNAs were not available in The Cancer Genome Atlas (TCGA)–HCC RNA-seq data, we also evaluated the expression profiles of positively selected lncRNAs in our RNA-seq from the University of Hong Kong (HKU) cohort. After filtering out lncRNAs with overall negligible fragments per kilobase per million (FPKM) values, we observed that most positively selected lncRNAs showed distinct differential expression patterns between tumors and nontumor tissues in both the TCGA–HCC and HKU cohort (Figure 2A). More than 60% of positively selected lncRNAs were overexpressed in HCC, whereas less than 10% of positively selected lncRNAs were down-regulated in HCC clinical samples (Figure 2B). For further analysis, we ranked the positively selected lncRNA candidates by calculating the average log_2 fold change of normalized sgRNA read counts between xenograft samples and pretransplantation cells. As a result, 538 lncRNAs of 1603 positively selected lncRNAs were overexpressed significantly in human HCC (Figure 2C). Interestingly, a subset of positively selected lncRNAs were reported previously to be oncogenic drivers in HCC, inclusive of LINC01138 and RAET1K, which served as positive controls of our in vivo library screening. In addition, we identified a subset of lncRNA candidates whose regulatory functions in HCC require more rigorous investigation, such as CASC11, Glucosylceramidase Beta Pseudogene 1 (GBAP1), LOC730101, LOC100499489, and LINC01133 (Figure 2D). Intriguingly, there was no correlation between the level of sgRNA enrichment in xenograft tumors versus pretransplantation cells and the genomic locations of sgRNAs, suggesting that sgRNAs targeting closer to transcriptional start sites (TSSs) did not necessarily achieve greater sgRNA enrichment (data not shown).

Clinical Relevance of Top lncRNA Candidates in HCC Patients

We compiled a list of positively selected lncRNAs in the library screening for further characterization (Supplementary Table 1). Before the extensive experimental validation of our library screening, we first evaluated the clinical relevance of positively selected lncRNAs. We selected 7 lncRNAs among the top candidates who showed higher expression in tumors relative to nontumor tissues for further validation (Figure 3A). Importantly, Kaplan–Meier analysis showed that higher expression of lncRNAs, namely CASC11, GBAP1, LOC730101, LOC100499489, LINC01138, and LINC01133, was associated with poor overall survival in HCC patients (Figure 3B). To evaluate the hazard ratios associated with the expression of each positively selected lncRNA in 2 independent patient cohorts, we constructed a multivariate Cox regression model. Interestingly, CASC11, LINC01138, LOC730101, LOC100499489, and TCONS_00011697 appeared to have high a hazard ratio in relation to other clinical covariates, suggesting that they could be independent prognostic biomarkers for the risk assessment of HCC patients (Figure 3B). To predict the biological pathways associated with the up-regulation of the selected lncRNA candidates, we stratified HCC patients from TCGA-HCC and HKU-HCC cohorts into high and low expression groups based on the median expression level of selected lncRNA candidates in tumors, and performed gene set enrichment analysis using hallmark annotations retrieved from Molecular Signature Database. We identified positive enrichment of biological pathways associated with more aggressive tumor behaviors (Figure 3C). Notably, patients with higher expression of CASC11, GBAP1, LINC01138, LOC730101, and TCONS_00011697 were associated with biological pathways that modulated cancer cell proliferation and cell-cycle progression, including E2F transcription factors (E2F) targets and Growth 2 phase/mitotic phase (G2/M) checkpoints ($P < .05$). Moreover, patients with higher expression of LINC01133 and LINC01138 were more likely to be associated with epithelial mesenchymal transition ($P < .05$). Patients with increased expression of CASC11, LOC730101, and LINC01133 showed higher enrichment of genes that modulated inflammatory response, cytokine production, and immune signature, suggesting that these lncRNAs may act as critical players in regulating the dynamics and plasticity of inflammatory networks and tumor immunity. To further explore the functional importance of our positively selected lncRNA candidates, we individually expressed the top enriched sgRNAs targeting the representative lncRNA candidates, showing that most enriched sgRNAs significantly induced the expression of their corresponding lncRNAs. In addition, transcriptional activation of selected lncRNA candidates significantly augmented HCC cell proliferation (Figure 4). Collectively, our clinical investigations and experimental validation reify our library screening as a powerful platform to identify clinically relevant lncRNAs that foster multiple cancer hallmark functions in HCC.
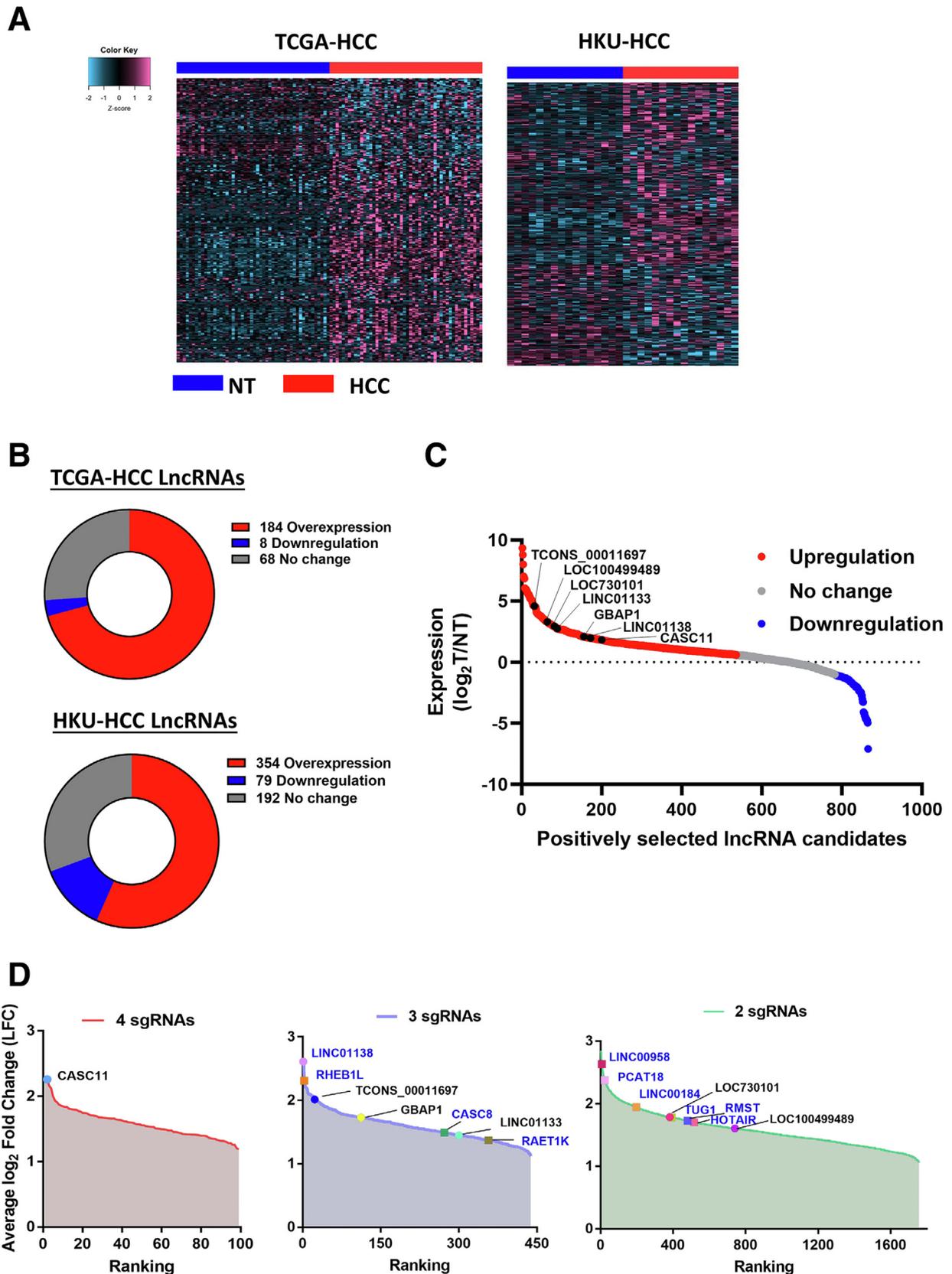
Figure 2. Expression levels of positively selected lncRNAs in HCC tumors. (*A*) Differential expression of positively selected lncRNAs between tumor and nontumor tissues. (*B*) Proportion of overexpressed lncRNAs as shown in the TCGA–HCC and HKU–HCC cohort. (*C*) Expression level of positively selected lncRNAs in tumors vs nontumors. Seven positively selected lncRNAs are indicated in the plot. (*D*) Rank of positively selected lncRNA candidates according to the average log$_2$ fold change (FC) of their corresponding enriched sgRNAs (FDR, <0.05). NT, non-tumor; T, tumor.
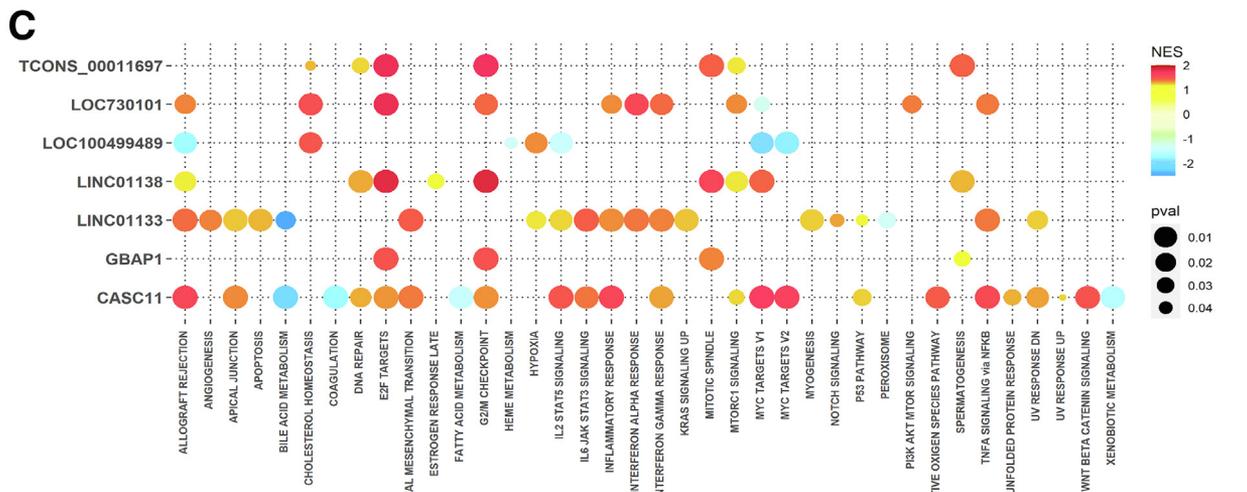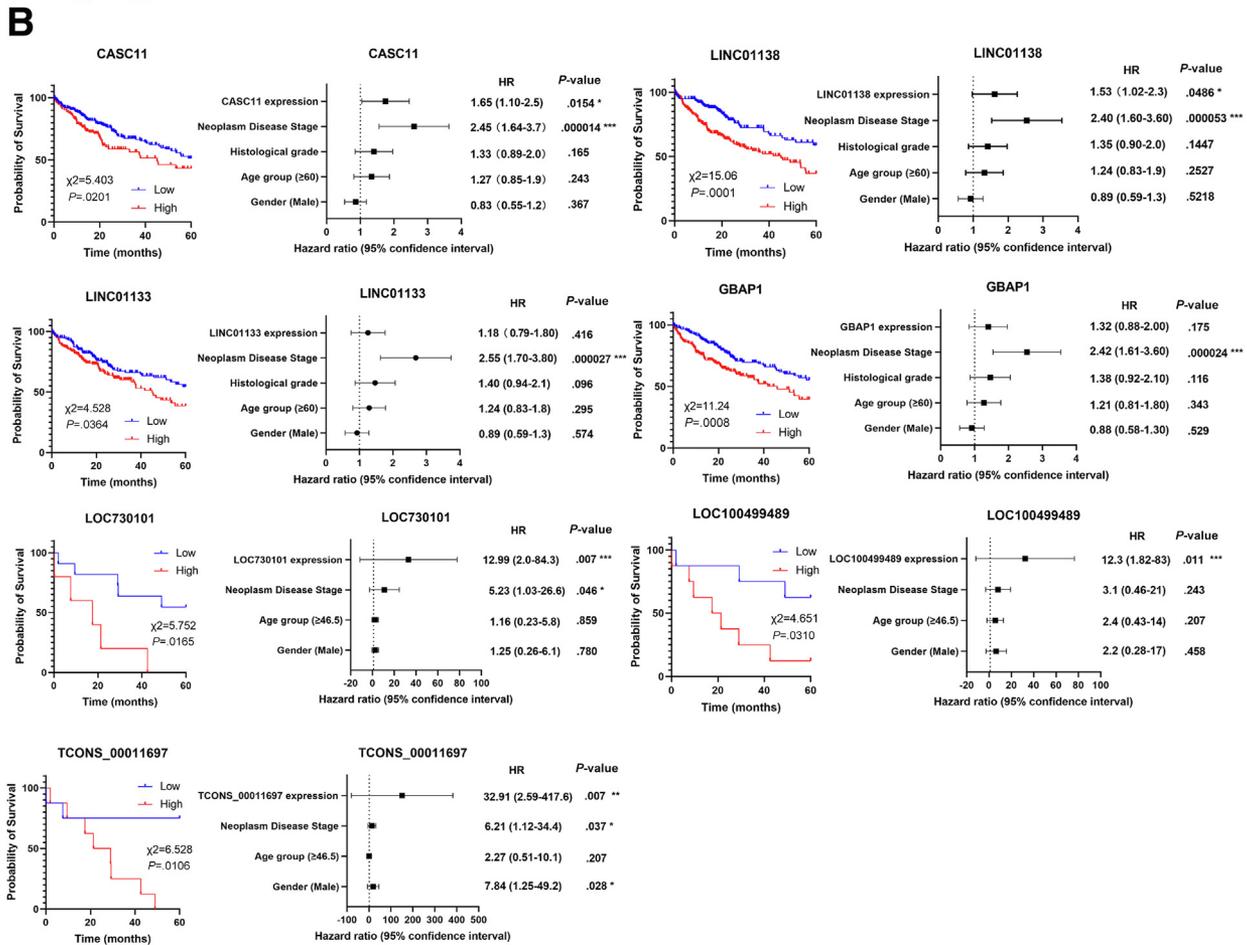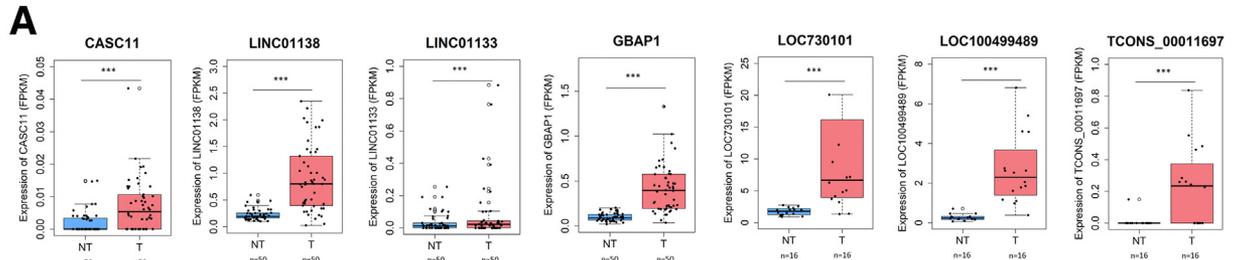
CASC11 with 4 significantly enriched sgRNAs hits ($\log_2$ fold change > 1, FDR < 0.05) was identified as a top-ranked positively selected lncRNA in the library screening. However, little is known about the functional roles of CASC11 in liver carcinogenesis. We confirmed the noncoding nature of CASC11 by using the ORF finder software from the National Centre for Biotechnology Information (Bethesda, MD), PhyloCSF codon substitution frequency analysis (Cambridge, MA), and our ribosome-sequencing data[13,14] (data not shown). We observed higher expression of CASC11 in human HCC than nontumor livers, which was detected in 54% and 42% of HCC patients with paired HCC and NT liver samples in the TCGA-HCC (n = 50) and HKU-HCC (n = 63) cohorts, respectively (Figure 5A). Notably, hepatitis B virus and hepatitis C virus carriers were shown to have higher expression of CASC11 compared with patients with no viral infection background, suggesting that viral infection may influence the transcription of CASC11 (Figure 5B). The upregulation of CASC11 was associated with advanced neoplasm histologic grades and the presence of vascular invasion in HCC patients (Figure 5C and D), indicating that CASC11 may serve as a prognostic indicator for the risk assessment of HCC patients.

CASC11 Drives HCC Cell Proliferation and Cell Progression In Vitro and In Vivo

To further corroborate the functional roles of CASC11 in HCC progression, we knocked down CASC11 expression in MHCC-97H and HepG2 cells by specific small hairpin RNAs (shRNAs) and locked nucleic acid (LNA) (Figure 6A). Knockdown of CASC11 significantly impaired HCC cell proliferative and colony-forming abilities (Figure 6B and C). We further examined the effects of CASC11 on tumor growth by introducing CASC11 overexpressing cells and CASC11 knockdown cells into nude mice subcutaneously. Consistent with our in vitro findings, CASC11 depletion impeded HCC growth in nude mice (Figure 6D). On the other hand, overexpression of CASC11 further aggravated HCC tumor growth in the CASC11-overexpressing group compared with the nontargeting control group (Figure 6E). Together, our in vitro and in vivo data collectively suggest that CASC11 is a prominent driver for HCC progression.

ChIRP-Seq Shows Significant Enrichment of CASC11 Binding on CASC11/MYC Shared Promoter

The oncogenic nature of CASC11 prompted us to investigate the mechanistic details of CASC11 in HCC further. We performed chromatin isolation of RNA purification

sequencing (ChIRP-seq) to investigate the genomic binding region of CASC11. Two nonoverlapping ChIRP probe sets (ie, A and B pools) targeting the full length of CASC11 transcript were used to pull down CASC11 interacting DNA. ChIRP-seq signals that were enriched concordantly in both samples were considered as high-confidence CASC11 binding sites. Of the 2789 high-confidence peaks across the genome, more than 50% of peaks were mapped to genic regions, with 12.6% of peaks annotated to gene promoters (Figure 7A). Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis showed that genes whose promoters were bound by CASC11 were associated with multiple cancer-related pathways, leading to the transcriptional misregulation in cancer (Figure 7B). Importantly, we found that CASC11 was enriched dramatically at the CASC11/MYC shared promoter region on chromosome 8q24 (Figure 7C and D, and Supplementary Table 2). To evaluate the binding efficacy and specificity of CASC11 on its target chromatin regions, we quantified CASC11 binding at MYC promoter by ChIRP–quantitative polymerase chain reaction (qPCR), using glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as the promoter control. CASC11 probes successfully retrieved more than 70% of CASC11 transcripts, whereas probes tiling β-galactosidase (LacZ) retrieved neither CASC11 nor GAPDH, further confirming the pull-down efficacy of CASC11 probes (Figure 7E). Subsequent qPCR quantification showed that CASC11 binding was detected at the CASC11/MYC shared promoter (-1500 to -2000 and +100 to +400 relative to MYC TSS), whereas the GAPDH promoter did not show any CASC11 binding signals (Figure 7F). Collectively, our data showed that CASC11 is bound preferentially to the CASC11/MYC shared promoter region, indicating a potential role of CASC11 on cis-regulation of MYC transcription.

CASC11 Modulates MYC Transcription

CASC11 is localized mainly in the nucleus (Figure 8A). CASC11 lies next to MYC in a head-to-head manner, with its promoter 1467-bp apart from MYC (Figure 9A). To predict the functional connection between CASC11 and MYC, we looked into the RNA-seq profile from the TCGA–HCC cohort, which showed a strong correlation between CASC11 and MYC expression (Figure 8B). Notably, patients with MYC amplification had higher expression of CASC11 ($P < .001$) (Figure 8C). HCC samples with MYC gain/MYC amplification appeared to have significant expression correlation with CASC11, whereas MYC wild-type HCC samples did not show the expression correlation in MYC and CASC11 (Figure 8D).

Considering the strong enrichment of CASC11 at the CASC11/MYC shared promoter, we next asked whether CASC11 exerts its influence on HCC growth through trans

---

Figure 3. (*See previous page*). **Clinical relevance of top lncRNA candidates in in vivo CRISPR library screening.** (A) Expression levels of selected lncRNA candidates in HCC tissues and adjacent nontumor liver tissues as determined by transcriptome sequencing in the HKU and TCGA cohorts. (B) *Left*: Kaplan–Meier plots of the top lncRNA candidates indicating overall survival of patients from the HKU and TCGA cohorts. The *P* value is indicated in the plots. The statistical significance was evaluated by the log-rank test. *Right*: Forest plot showing the multivariate analysis (Cox proportional hazard analysis). The impact of each top lncRNA candidate and significant clinical covariates on the overall survival of HCC patients. The hazard ratio (HR) and the *P* value were calculated by the Wald test. (C) Pathway analysis of enriched lncRNA candidates. Hallmark gene sets were used for this analysis. NES, normalized enrichment score; NT, non-tumor; T, tumor.
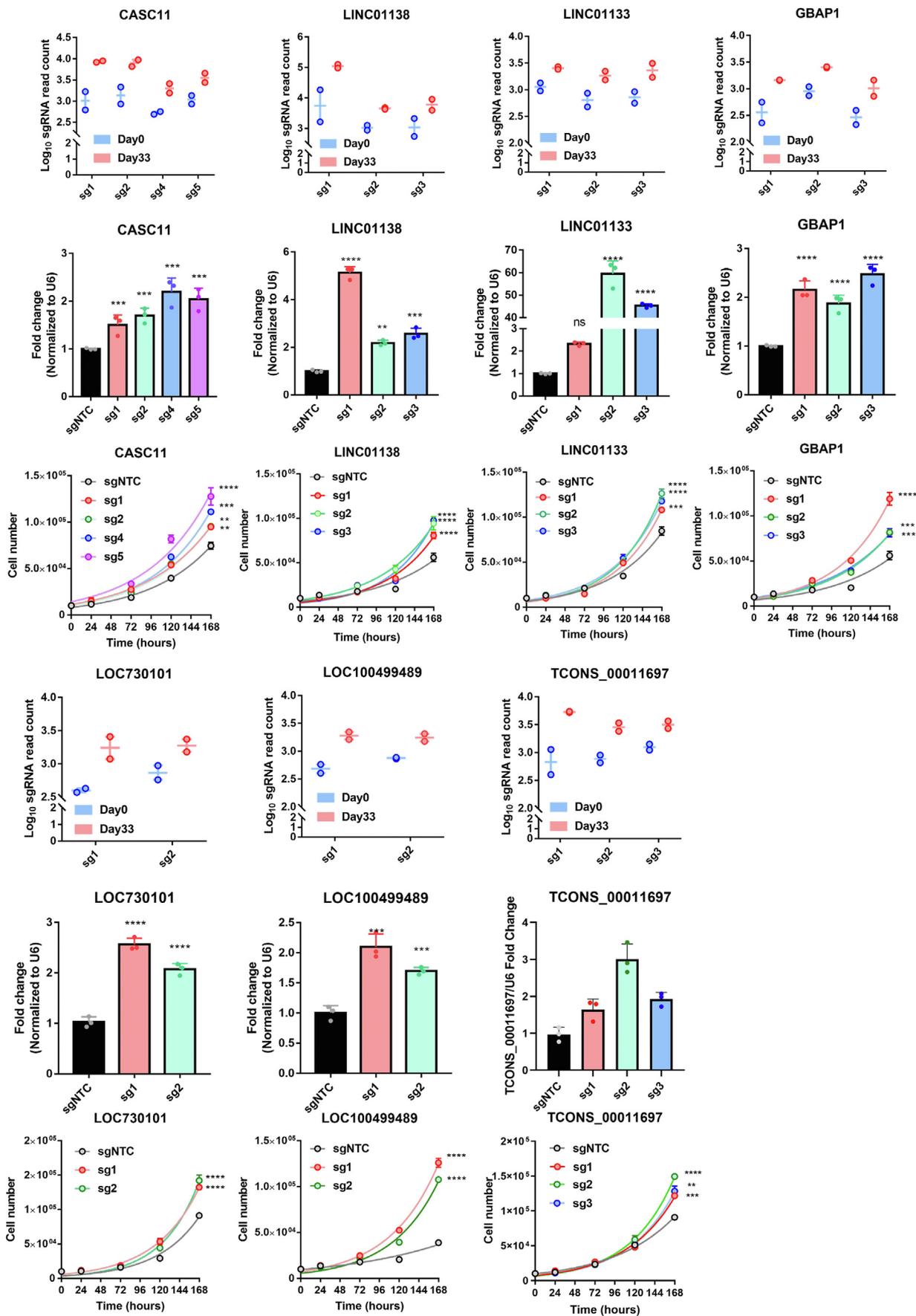
(nonlocal) or *cis*- (local) regulatory function. To determine whether CASC11 contributed to HCC progression through *trans*-regulatory function, we overexpressed complementary DNA (cDNA), encoding the full length of CASC11 in a panel of HCC cells. However, we did not observe any significant changes in MYC expression at transcriptional or translational levels (data not shown). We also found no significant difference in cell proliferation and colony-forming capacity in CASC11-overexpressing cells compared with the empty vector, suggesting that CASC11 is unlikely to act in *trans* (data not shown). To test whether CASC11 promoted HCC growth through regulating MYC in *cis*, we evaluated the expression of MYC upon synergistic activation mediator (SAM) targeting on CASC11 promoter. The majority of positively enriched sgRNAs targeting CASC11 promoter in our library screening concordantly led to a significant up-regulation in MYC messenger RNA expression (Figure 9*B*). However, one of the major challenges of CRISPR activation (CRISPRa) -induced overexpression of lncRNA is the risk of off-target effects arising from the perturbation of neighboring genes via their shared promoter.[15] To determine whether the simultaneous CRISPRa-induced activation of CASC11 and MYC expression was the outcome of the *cis*-regulatory mechanism or was the direct transcriptional activation of the neighboring gene, we infected HCC cells with 2 independent LNAs to deplete nuclear-enriched CASC11, which showed a remarkable reduction in MYC expression (Figure 9*C*). To investigate whether CASC11 is required for MYC transcription, we silenced CASC11 transcript by LNA in the context of CASC11 overexpression, which strikingly showed significant down-regulation in both CASC11 and MYC expression (Figure 9*D*). Given the potential influence of CASC11 on MYC transcription, we silenced MYC by 2 independent small interfering RNAs (siRNAs) in CASC11 overexpression cells to determine whether CASC11 exerted its biological functions through MYC. We noted that MYC silencing compromised the cell proliferative potential induced by CASC11 overexpression (Figure 9*E*). In addition, we generated MYC knockout subclones by infecting CRISPR/Cas9-carrying MHCC-97H cells with 3 independent sgRNAs targeting MYC, finding that the effect of MYC knockout on cell growth phenocopied that of CASC11 silencing in MHCC-97H cells (data not shown).

## CASC11 Drives HCC Progression Through a Subset of Cell-Cycle–Related Genes

Our in vitro and in vivo data, as well as the mechanistic studies, shows the important role of CASC11 in promoting HCC progression. To decipher CASC11 downstream pathways, we performed transcriptome sequencing analysis in MHCC-97H cells treated with LNA targeting CASC11. We identified 683 differentially expressed genes upon CASC11 knockdown, of which 292 genes were down-regulated and 391 genes were up-regulated ($P < .05$) (Figure 10*A* and Supplementary Table 3). Gene set enrichment analysis showed that gene sets, namely G2M checkpoint and MYC targets from the Molecular Signatures Database Hallmark collection, were dysregulated significantly in CASC11 knockdown cells (Figure 10*B* and Supplementary Table 4). Consistently, we observed in the TCGA-HCC cohort that the expression of MYC target genes was enriched significantly in the CASC11–high-HCC group compared with the CASC11–low-HCC group (Figure 10*C*). MYC target genes and cell-cycle–associated genes were enriched positively in patients with high expression of CASC11, in concordance with the findings from the RNA-seq of CASC11 knockdown cells (Figure 10*D*).

To further analyze the functional relevance of CASC11 in MYC-driven HCC, we knocked down CASC11 in HCC cell lines with or without MYC overexpression (Figure 11*A*). We observed that MYC and CASC11 were overexpressed in HepG2 and MHCC97H, but had a low expression in Huh7 cells (Figure 11*B*). Knockdown CASC11 induced a pronounced reduction in MYC expression at both the transcriptional and translational levels in all tested cell lines (Figure 11*C* and *D*). We also investigated the status of several cell-cycle master regulators, such as cyclin-dependent kinase (CDK4), CDK6, DNA binding protein inhibitor 1 (ID1), DNA binding protein inhibitor 2 (ID2), and E2F transcription factor 1 (E2F1) in CASC11-depleted MHCC97H, HepG2, and Huh7 cells. These genes were down-regulated significantly at the transcriptional level upon CASC11 depletion. However, we found the reduced protein expression of CDK4 and CDK6 only in MHCC97H and HepG2 cells, but not in Huh7. Conversely, CRISPRa-induced overexpression of CASC11 profoundly up-regulated the expression of a subset of cell-cycle–related genes (Figure 11*E*). These findings collectively suggest CASC11 is the main driver of activated MYC signaling with MYC amplification. Previous studies have suggested that CDK4 and CDK6 were direct transcriptional targets of MYC.[16] Supporting this, we observed significant MYC peaks on the promoter regions of CDK4 and CDK6 in the chromatin immunoprecipitation sequencing (ChIP-seq) profiles of HepG2 retrieved from the Encyclopedia of DNA Elements (Figure 12*A*). In addition, we observed that MYC and CDK4 were highly correlated with the expression of CASC11 (Figure 12*B*). We reasoned that CASC11 would modulate CDK4 and CDK6 expressions via MYC. To explore the functional connection between CASC11 and MYC, we silenced MYC and CASC11 expression, which consistently decreased the expression of MYC-regulating cell-cycle genes (Figure 12*C*). Pathway enrichment analysis showed that

---

Figure 4. (*See previous page*). **In vivo CRISPRa screening identifies a subset of positively selected lncRNAs that contribute to HCC cell growth.** *Upper*: Read count of enriched sgRNAs targeting the top hit candidates. *Middle*: Validation of overexpression efficiency of enriched sgRNAs as determined by qPCR analysis. All expression levels were normalized to U6. *Lower*: Overexpression of selected lncRNA candidates promoted cell proliferation in MHCC-97H cells. Results represent the mean values of 3 independent experiments and the statistical significance was evaluated by 1-way analysis of variance with the Dunnett multiple comparison test. Data are presented as means ± SD, **$P < .01$, ***$P < .005$, and ****$P < .001$.
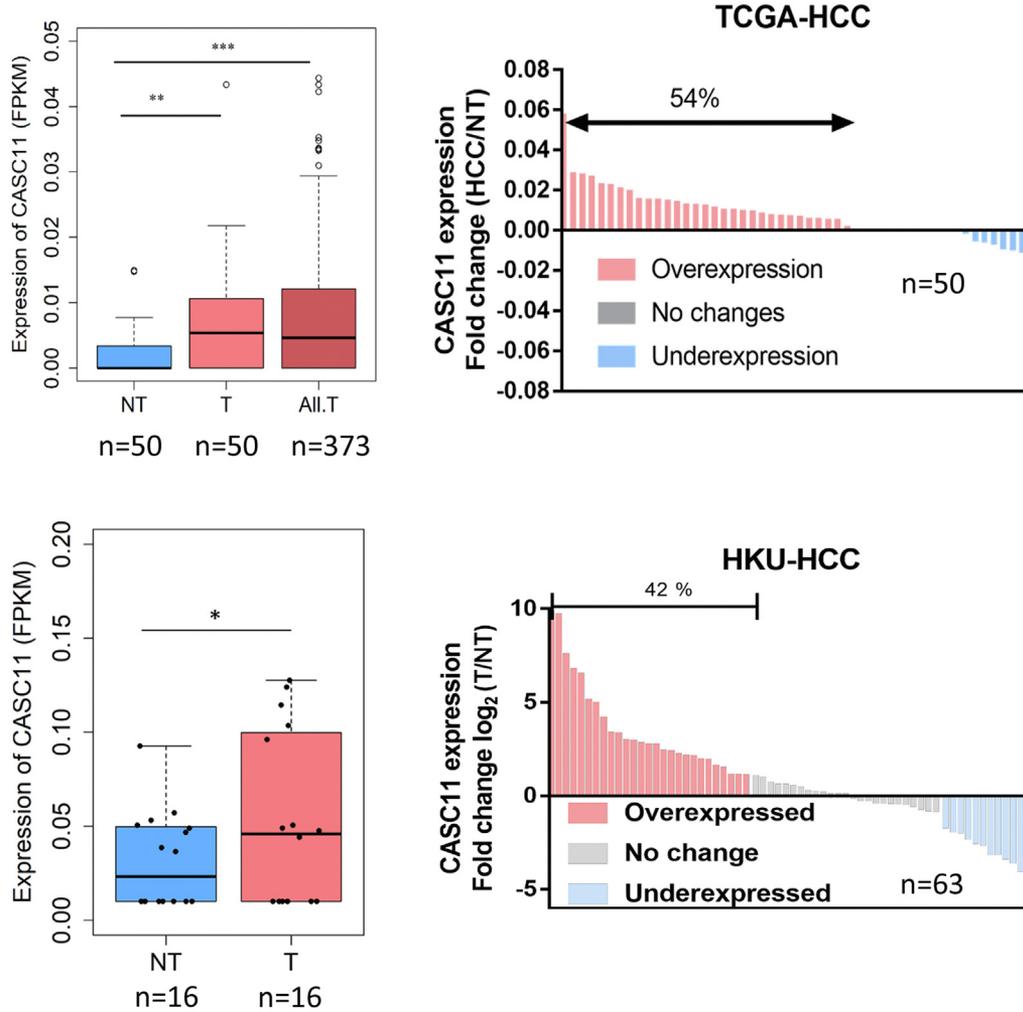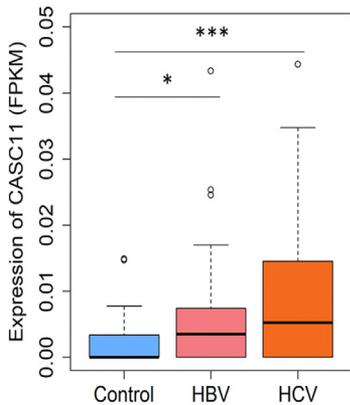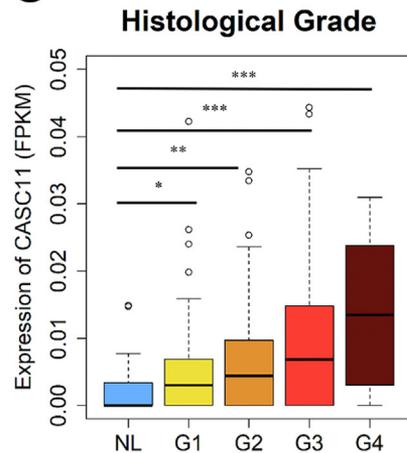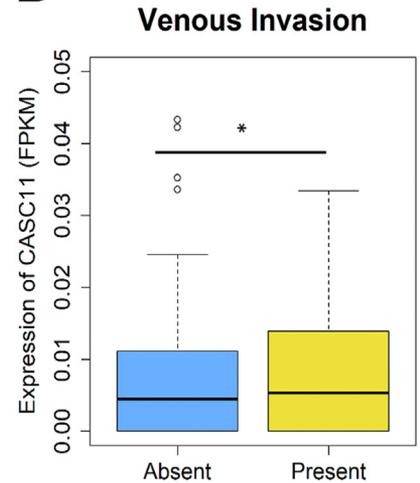
Figure 5. Frequent up-regulation of CASC11 in HCC patients. (*A*) Relative expression of CASC11 in HCC and nontumor liver tissues in TCGA cohort and HKU cohort. Data were presented as FPKM. (*B*) Expression of CASC11 in control (non–hepatitis B virus [HBV] or hepatitis C virus [HCV] carriers), HBV carriers, and HCV carriers (1-way analysis of variance with the Dunnett test). (C) Expression of CASC11 was associated positively with more advanced neoplasm histologic grade (1-way analysis of variance with the Dunnett test) and (*D*) vascular invasion in HCC patients (Student *t* test). All T, all tumor; NT, non-tumor; T, tumor; FPKM: fragments per kilobase per million.
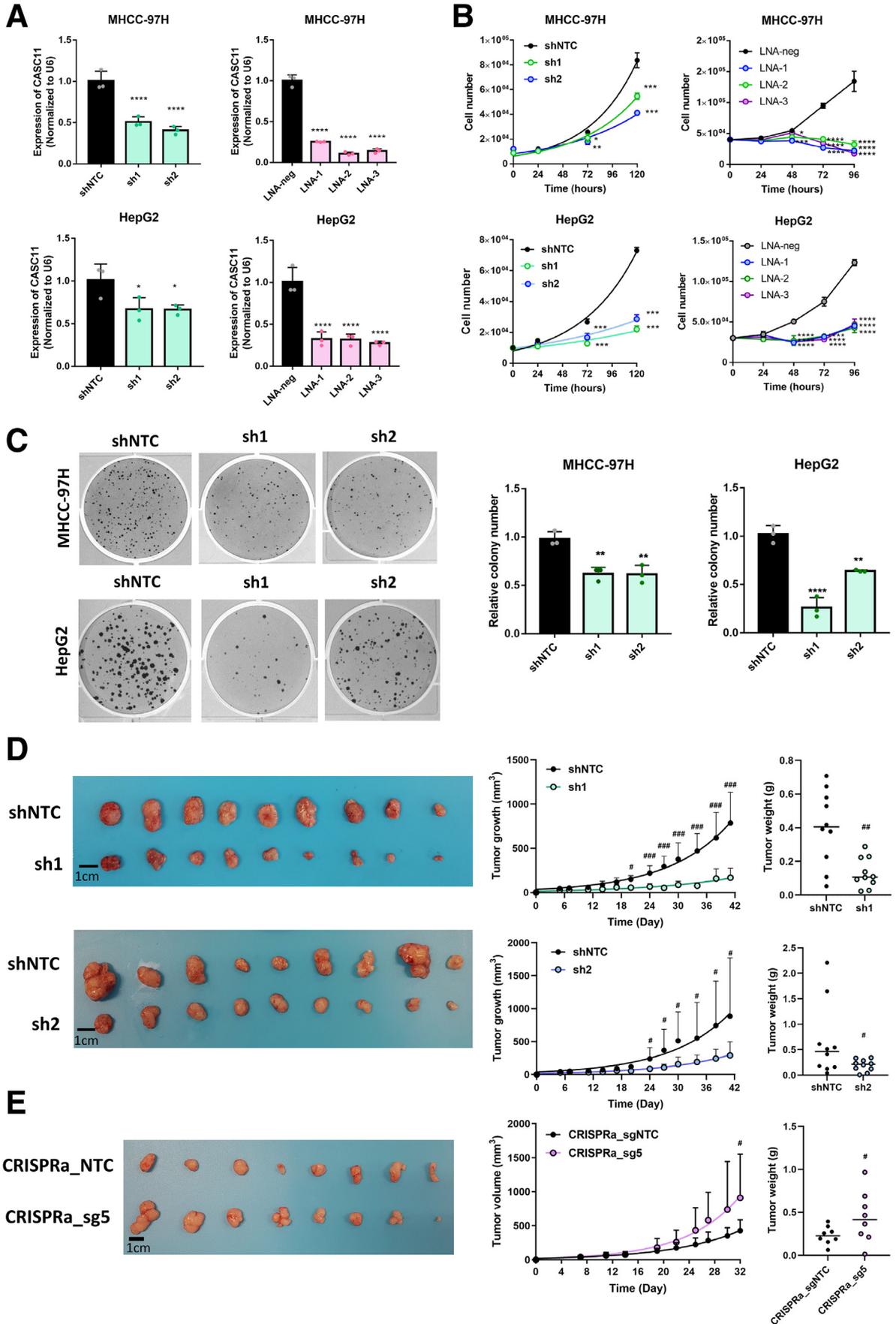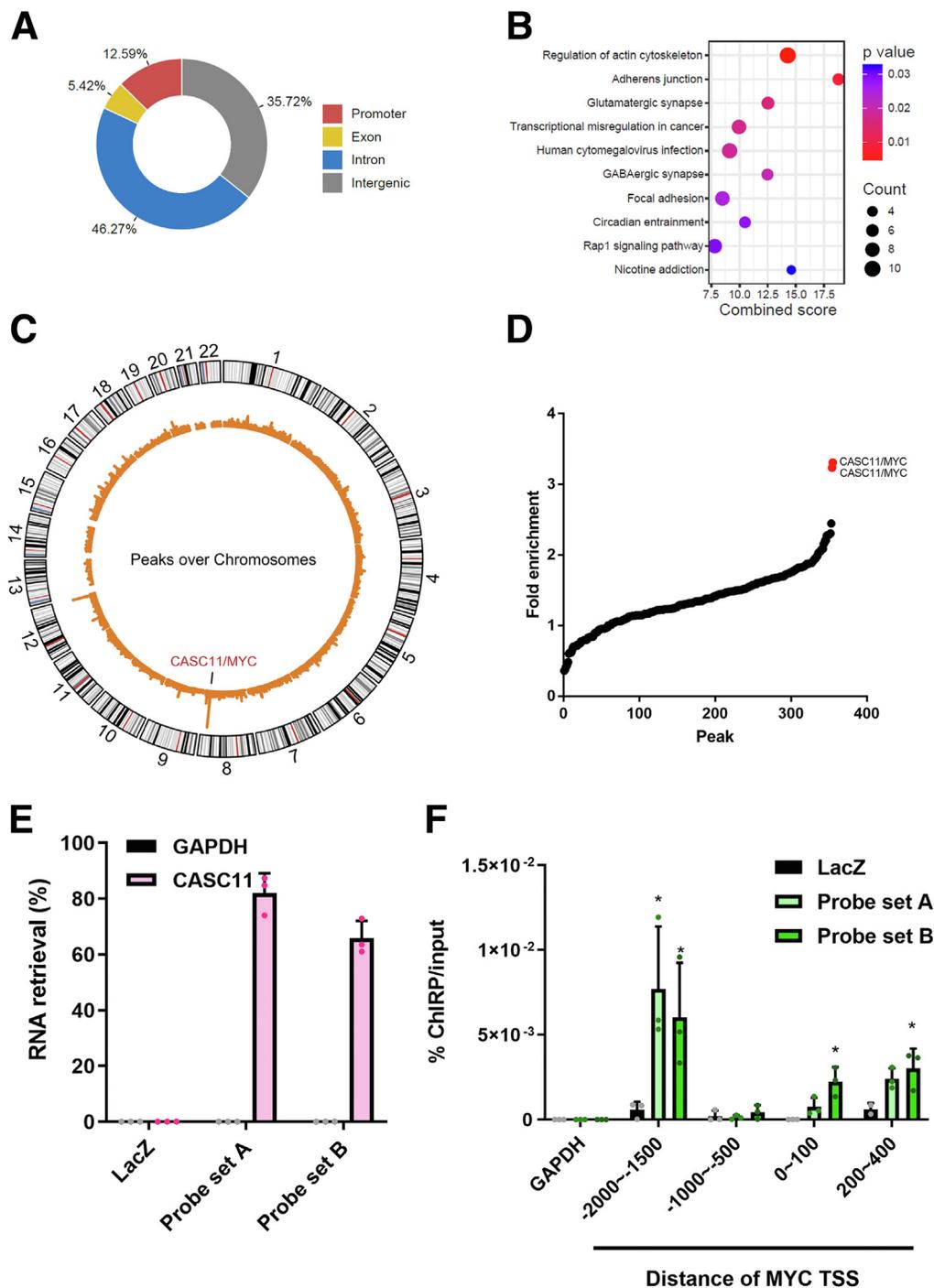
Figure 7. ChIRP-seq shows genome-wide CASC11 binding sites. (*A*) Distribution of CASC11 binding sites at different genomic regions. Promoter regions are defined as ±3000 bp of the nearest gene's TSS. (*B*) Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis of promoter-binding ChIRP-seq peaks. (*C*) Circos plot showing genome-wide CASC11 binding sites. (*D*) ChIRP-seq signals (fold enrichment over input) of CASC11-bound peaks at promoter regions. (*E*) Percentage of total CASC11 RNA retrieved by biotinylated CASC11 antisense probes. The LacZ probe serves as a negative control. (*F*) ChIRP qPCR analysis measuring CASC11 enrichment across different regions of the CASC11/MYC shared promoter. GAPDH serves as a negative control. Results represent mean values of 3 independent experiments and the statistical significance was evaluated by 1-way analysis of variance with the Dunnett multiple comparison test. Data are plotted as means ± SD, *$P < .05$. GABA, gamma-aminobutyric acid; LacZ, β-galactosidase.

Figure 6. (*See previous page*). Knockdown (KD) of CASC11 significantly inhibits HCC cell proliferation in vitro and in vivo. (*A*) Knockdown efficiency of CASC11 in MHCC-97H and HepG2 transfected with shRNAs or LNAs targeting CASC11 was measured by RT-qPCR. Relative expression was normalized to U6. (*B*) Relative cell viability of MHCC-97H and HepG2 after CASC11 KD. (*C*) The number of colonies was assessed after 2 weeks postseeding of stable CASC11 KD MHCC-97H and HepG2 cells. (*D*) Tumor growth of xenografts injected with stable CASC11 KD MHCC-97H cells or nontargeting cells (n = 10). Tumor volume and tumor weight were measured in the nontargeting control (NTC) and CASC11-KD group. (*E*) Tumor volume and tumor weight were measured in NTC and CASC11-activating groups (n = 8). Representative images of xenografts are included. (*A–C*) Mean values of 3 independent experiments Statistical significance was evaluated by 1-way analysis of variance with the Dunnett multiple comparison test. Data are pre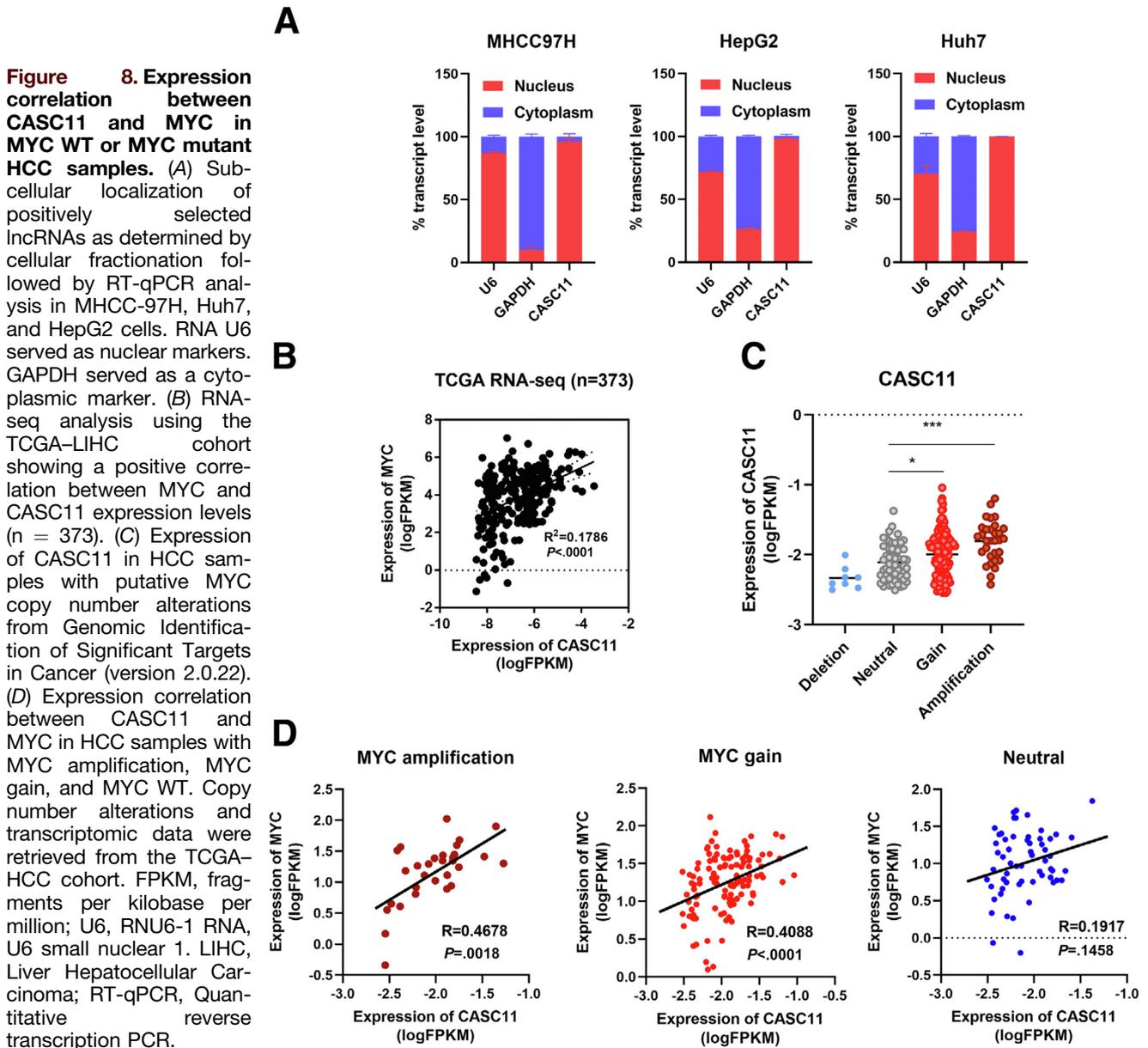sented as means ± SD, *$P < .05$, **$P < .01$, ***$P < .005$, and ****$P < .001$. (*D* and *E*) presented as means ± SD, #$P < .05$, ##$P < .01$, ### $P < 0.01$ by Student t test. RT-qPCR, Quantitative reverse transcription PCR.

Figure 8. Expression correlation between CASC11 and MYC in MYC WT or MYC mutant HCC samples. (*A*) Subcellular localization of positively selected lncRNAs as determined by cellular fractionation followed by RT-qPCR analysis in MHCC-97H, Huh7, and HepG2 cells. RNA U6 served as nuclear markers. GAPDH served as a cytoplasmic marker. (*B*) RNA-seq analysis using the TCGA–LIHC cohort showing a positive correlation between MYC and CASC11 expression levels (n = 373). (*C*) Expression of CASC11 in HCC samples with putative MYC copy number alterations from Genomic Identification of Significant Targets in Cancer (version 2.0.22). (*D*) Expression correlation between CASC11 and MYC in HCC samples with MYC amplification, MYC gain, and MYC WT. Copy number alterations and transcriptomic data were retrieved from the TCGA–HCC cohort. FPKM, fragments per kilobase per million; U6, RNU6-1 RNA, U6 small nuclear 1. LIHC, Liver Hepatocellular Carcinoma; RT-qPCR, Quantitative reverse transcription PCR.

MYC-knockout and CASC11-silencing cells shared many common downstream target genes that were associated with cell-cycle G1/S phase transition, DNA-dependent DNA replication, and nucleic acid metabolic processes (Figure 12*D*).

We examined the functional impact of CASC11 on cell-cycle progression. In concordance with differentially expressed gene analysis, CASC11 depletion induced the accumulation of cells at G1-phase and S-phase defects, whereas overexpression of CASC11 accelerated G1/S progression (Figure 11*F* and *G*). Intriguingly, the delay of G1/S progression as a result of MYC silencing was reversed by the ectopic expression of CASC11, indicating that CASC11 may play its role in cell-cycle control through MYC (Figure 11*H*). Taken together, CASC11 may promote cell-cycle progression through the modulation of MYC.

Discussion

In view of the high mortality rate of HCC, we are now endeavoring to understand the underlying mechanism of HCC pathogenesis. The exquisite complexity of the human genome portends a daunting task for HCC treatment. It becomes apparent that dysregulation of the noncoding genome has profound effects on cancer phenotypes.[2] Despite extensive annotations of lncRNAs, the functional mechanisms of lncRNAs in HCC remain largely elusive, accentuating the necessity to devise an integrated pipeline that helps uncover functional lncRNAs in a high-throughput manner. One of the greatest challenges in studying human lncRNAs is the genomic complexity of lncRNAs and a lack of effective tools that can define their physiological functions in a high-throughput manner.[17] CRISPR/Cas9 knockout screening has been used widely to identify essential protein-
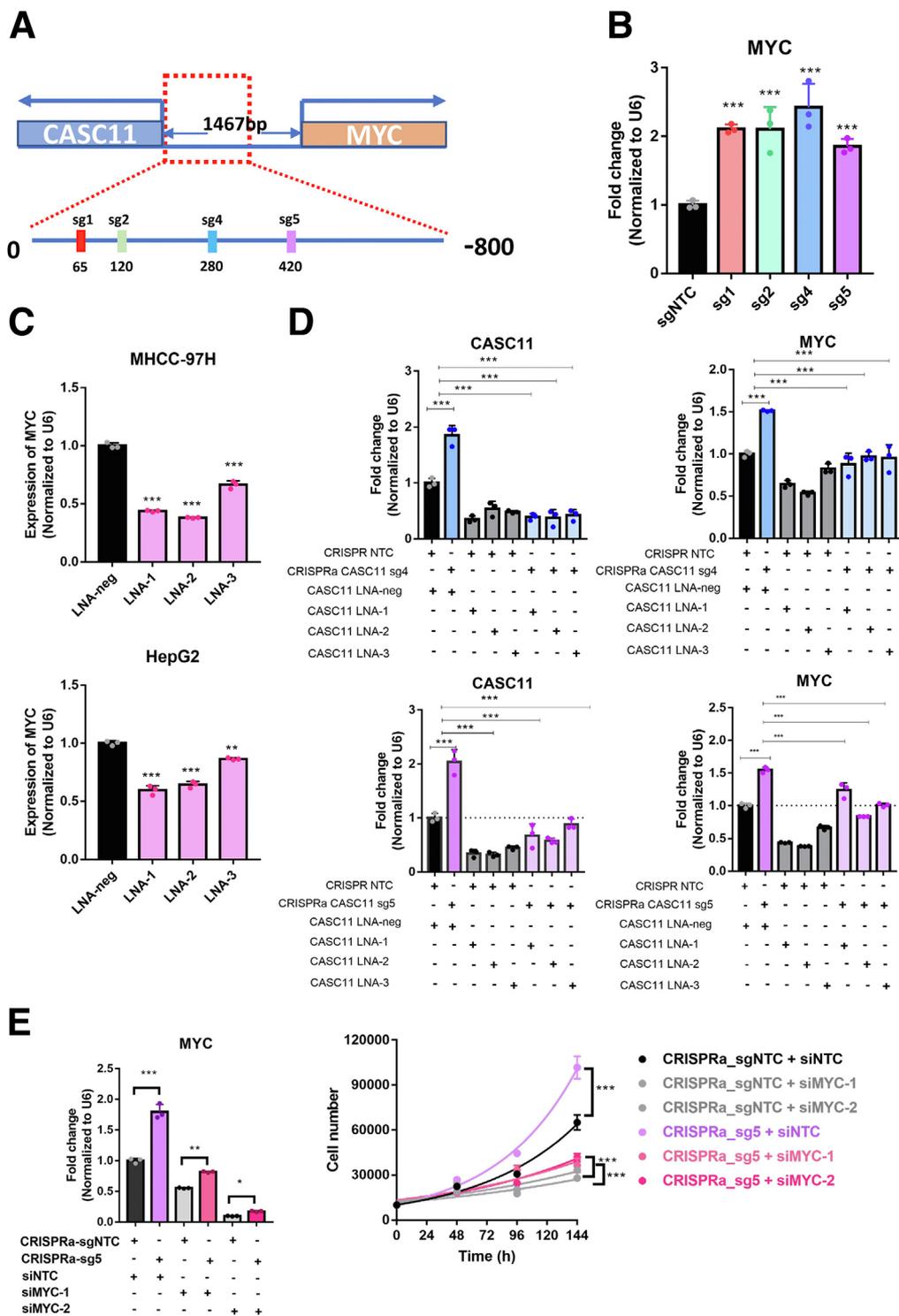
Figure 9. CASC11 modulates MYC expression and its transcriptional activity. (*A*) Schematic representation of enriched sgRNAs targeting the shared promoter region of CASC11 and MYC. (*B*) Up-regulation of MYC upon CASC11 activation in MHCC-97H cells determined by RT-qPCR. (*C*) Down-regulation of MYC in MHCC-97H and HepG2 upon LNA-mediated CASC11 knockdown (KD). (*D*) Expression of CASC11 and MYC at 48 hours after transfection with LNAs targeting CASC11 upon CASC11 overexpression. (*E*) *Left*: RT-qPCR analysis of MYC expression in CASC11-overexpressing cells after transfection with siRNAs targeting MYC for 48 hours. *Right*: Proliferation capacity of CASC11-overexpressing cells upon MYC silencing for 48 hours. (*B–E*) Mean values of 3 independent experiments. Statistical significance was evaluated by 1-way analysis of variance with the Dunnett multiple comparison test. Data are presented as means ± SD (*$P < .05$, **$P < .01$, ***$P < .005$). RT-qPCR, Quantitative reverse transcription PCR.

coding genes that promote tumor growth and metastasis in xenograft mouse models, but an in vivo CRISPR activation screening targeting lncRNAs has not yet been reported.[7,17–19]

A variety of local and environmental factors including hypoxic conditions, proliferating tumor cells, infiltrating inflammatory cells, vascular system, stromal cells, and extracellular matrix create a unique environment that strongly influences molecular and cellular events, thereby promoting tumor growth.[12] Considering the unique expression and tissue-specificity of lncRNAs, an in vivo screening has a greater potential to discover HCC-promoting lncRNAs in mouse models than in vitro cell cultures. Here, we conducted a CRISPR/dCas9 library activation screening in the xenograft mouse model to address the functional roles of lncRNAs in HCC progression. In contrast to CRISPR
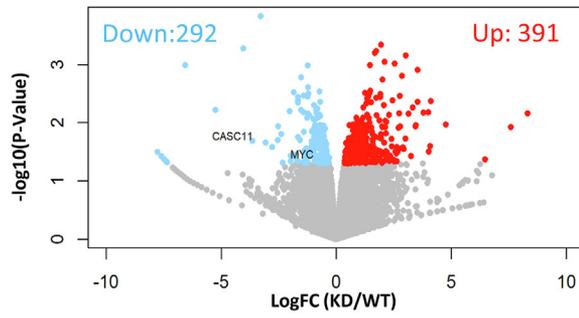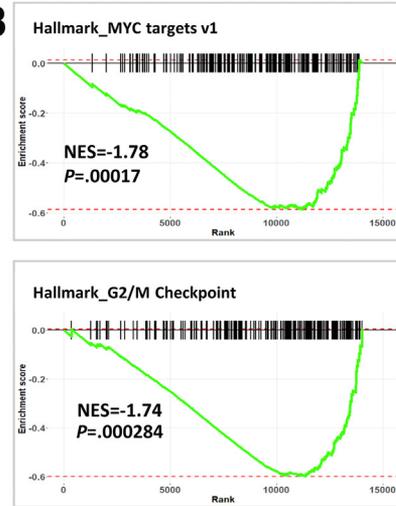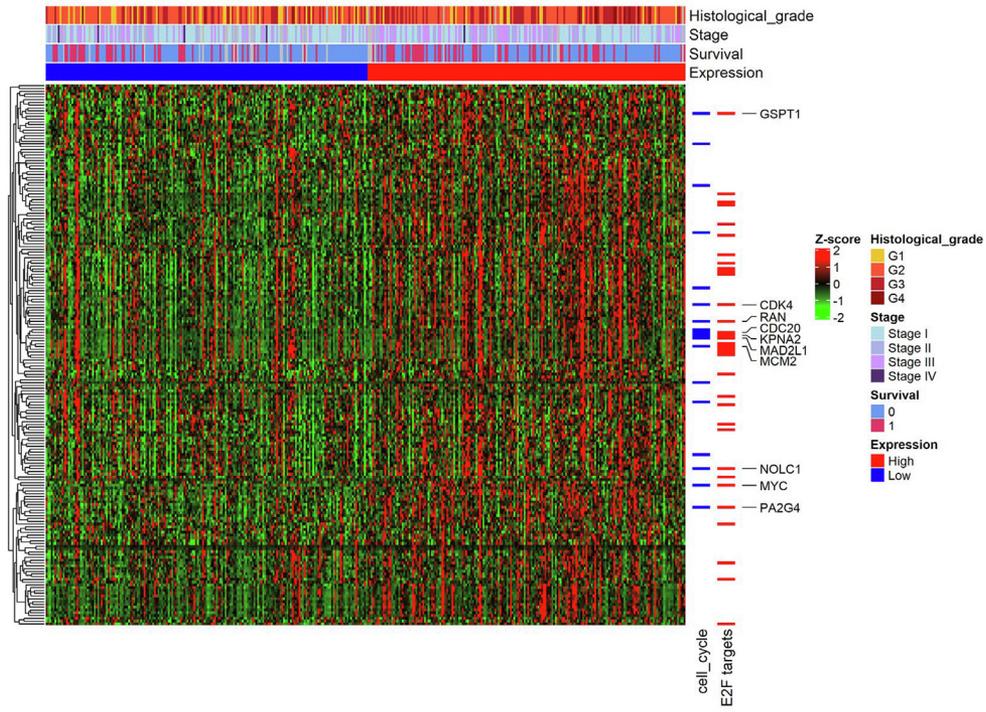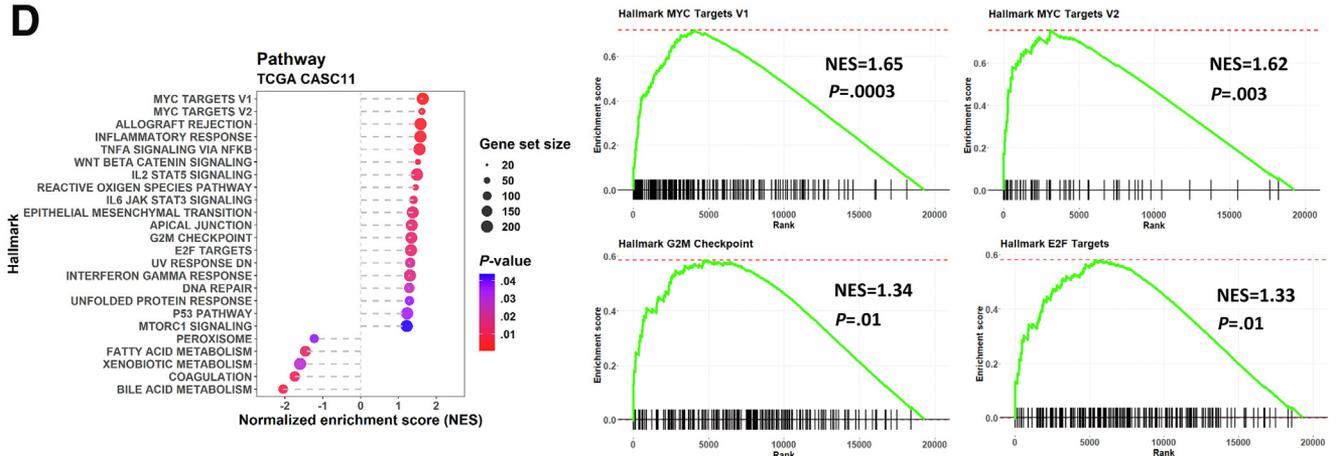
knockout screening, in which the consistent enrichment of multiple sgRNAs indicates high-confidence screening hits, not every sgRNA can induce the transcriptional activation of target lncRNA loci. CRISPRa acts effectively within a small window around the targeted lncRNA TSS, in which the efficacy of targeted sgRNA is influenced strongly by several genomic features including chromatin accessibility, transcription factor binding sites, and the position of sgRNAs relative to the TSS. sgRNAs with low efficacy would greatly impede the overall statistical power in the gene-level analysis as designed for CRISPR knockout screening. Hence, we performed sgRNA-level analysis by comparing the relative abundance of individual sgRNA between tumors and pre-transplantation cells. Considering the potential of false-positive results from the enriched singleton sgRNA, only the lncRNA loci that were targeted by at least 2 corresponding sgRNAs were considered positively selected in tumors vs pretransplantation cells. As a result, we identified 1603 positively selected lncRNAs in our library screening, and the majority of them showed distinct expression patterns in tumors as indicated in transcriptomic data sets in the TCGA–HCC and HKU–HCC cohorts. In addition to the well-characterized lncRNA oncogenic drivers in HCC progression, we also identified a wealth of novel lncRNAs whose expression levels were comparatively higher in tumors, among which LINC01133, LOC100499489, LOC730101, GBAP1, TCONS_00011697, and CASC11 promoted HCC cell proliferation upon CRISPRa overexpression, warranting efforts for more detailed functional characterization. Hence, the functional CRISPR screening facilitates the discovery of a wealth of novel predictive biomarkers. Importantly, overexpression of these positively selected lncRNAs promoted HCC growth, further confirming the reliability of our library screening and highlighting the importance of using an appropriate experimental system with the relevant physiological condition to define the functionality of lncRNAs.

For a better understanding of the modes of actions of lncRNAs in promoting HCC pathogenesis, we furthered our investigation on the functional importance of CASC11. Strikingly, our in vitro and in vivo studies showed the cell proliferation and tumorigenesis-promoting effects of CASC11. CASC11 is located at the 8q24.21 genomic region, where many cancer-associated single-nucleotide polymorphisms alter the transcription of lncRNAs, consequently affecting the transcriptional activities of MYC.[20] MYC, regulating as many as 15% of genes in the human genome, acts as an important hub for orchestrating a broad spectrum of biological functions.[21] Although scientists have made valiant efforts on studying the roles of MYC in various human cancer diseases for decades, it still is challenging to inhibit

cancer growth by direct MYC inhibition owing to its potential detrimental consequences for normal cell development.[22] Hence, there is a voracious appetite for strategies that indirectly interfere with MYC transcription. Indeed, several cis-acting lncRNAs have been reported to regulate MYC transcription, through facilitating the activity of MYC promoter-binding transcription factors and the formation of higher-order chromatin architecture.[23–25] For example, MYC-modulating lncRNA (MYMLR) was found to regulate the MYC promoter activity by binding to MYC enhancer.[25] Although MYMLR shows some overlap with CASC11, they are transcribed from different promoters and they showed different regulatory cascades, indicating the genomic complexity of this locus.

Realizing that CASC11 is transcribed divergently near the MYC promoter, we pursued our investigation on the complex interplay between MYC and CASC11. Notwithstanding multiple lines of evidence showing the oncogenic role of CASC11 in different cancers, to our knowledge, there is a paucity of studies on the functional mechanism of the CASC11/MYC axis in HCC progression.[26–29] Considering an increasing body of evidence showing that lncRNAs regulate adjacent protein-coding genes, we speculated that CASC11 functions through MYC, consequently dysregulating its downstream targets. ChIRP-seq analysis showed that CASC11 preferentially is bound on the CASC11/MYC promoter region. In addition, our functional assays unequivocally suggested that the perturbation of CASC11 expression drastically affected MYC expression through cis-regulatory function. The explanation could be that CASC11 exerts its local function—either in the act of its transcription or in its own transcript—to modulate the chromatin-associated processes and influence the transcription of its neighboring gene MYC. These findings also highlight the importance of using the CRISPR activation system to activate the target lncRNA transcription at the endogenous level, thereby capturing the local functions of lncRNAs and recapitulating all lncRNA isoforms at the targeted loci, whereas this is not the case in traditional cDNA overexpression.

Although we showed the cis-regulatory role of CASC11 on MYC transcription, several questions regarding the complex regulatory circuitry of CASC11 and MYC await further investigation. Increasing evidence has shown that antisense lncRNA transcript mediates R-loop formation in the local chromatin that favors the binding of the transcriptional machinery, thereby inducing the transcriptional activation of its neighboring gene.[30] Toward this, we observed an enriched DNA:RNA immunoprecipitation sequencing signal in the CASC11/MYC shared promoter. The involvement of CASC11 in MYC transcriptional activity also may be pertinent to a burgeoning concept of liquid–liquid

Figure 10. (*See previous page*). **CASC11 regulates a subset of MYC target genes and cell-cycle–related genes.** (*A*) Volcano plot showing differential gene expression (*P* < .05) upon LNA-mediated CASC11 knockdown for 48 hours. (*B*) Gene set enrichment analysis showing top cancer hallmarks regulated by CASC11. (*C*) Heatmap showing the expression patterns of MYC-associated genes in CASC11-low and CASC11-high groups from the TCGA–LIHC data set. Clinical samples were dichotomized based on the median cut-off expression level of CASC11. Genes associated with cell-cycle progression and E2F targets are indicated. (*D*) Gene set enrichment analysis identified cancer hallmarks associated with CASC11 expression in HCC patients. FC, fold change ; LIHC, Liver Hepatocellular Carcinoma; E2F, E2F transcription factors
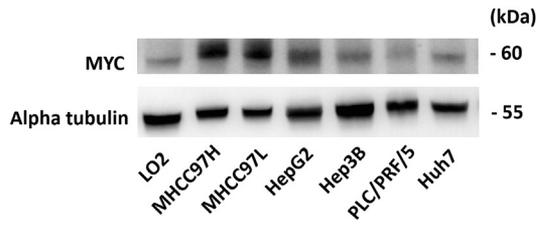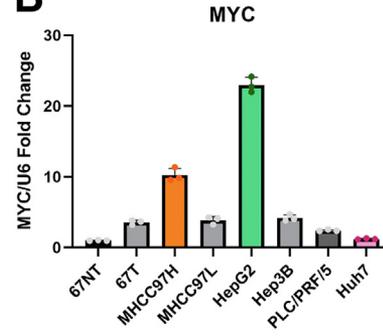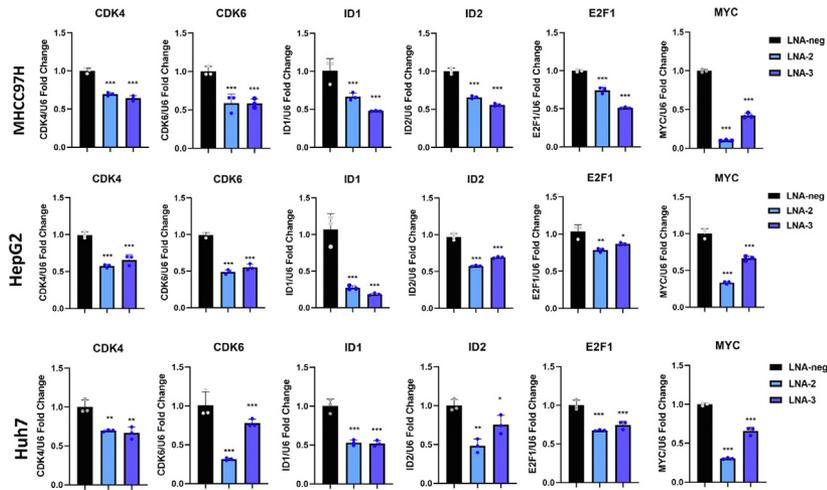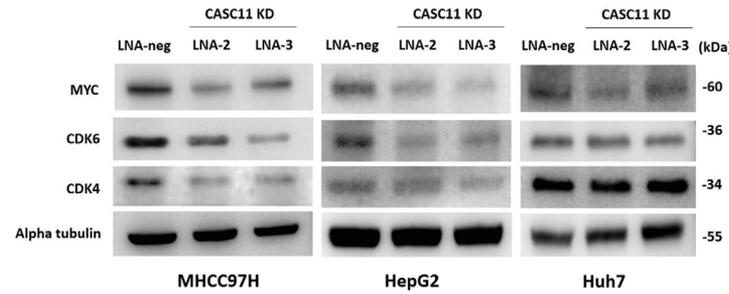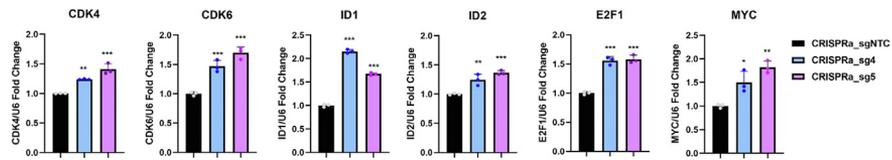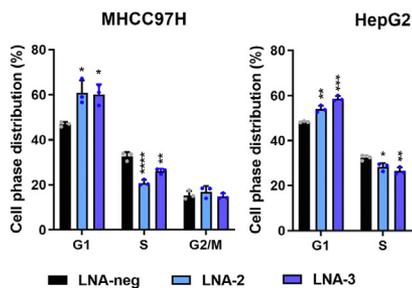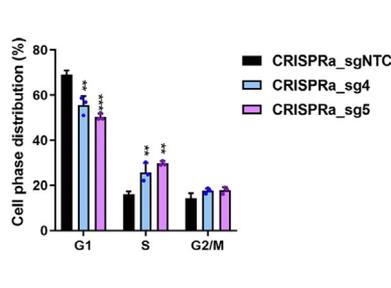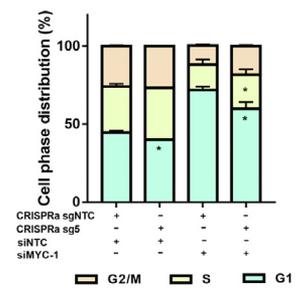
phase separation, in which lncRNAs facilitate the assembly of RNA binding proteins to form a spatial compartment in the nucleus for gene transcription.[31]

Cell-cycle dysregulation is a central hallmark of cancer progression.[32] We showed that CASC11 promoted HCC cell growth by dysregulating cell-cycle pathways and MYC-associated pathways. However, because our ChIRP-seq data did not show significant CASC11 enrichment on the CDK4 and CDK6 promoters, we speculated that the down-regulation of CDK4 and CDK6 expression in CASC11-depleted cells presumably was caused by the down-regulation of MYC. Interestingly, CASC11 knockdown and MYC knockout cells showed consistent enrichment in many common downstream targets, which were particularly associated with G1/S-phase progression and DNA replication. Importantly, MYC silencing decelerated the rate of G1/S progression driven by CASC11 overexpression, lending support to the idea that CASC11 exerts its functional role through MYC.

Conclusions

In summary, our study has unveiled the power of genome-wide CRISPR activation screening in tumor xenografts to uncover oncogenic lncRNAs, providing a rationale to target these lncRNAs clinically. We integrated computational analysis of clinical transcriptomic data sets with functional CRISPR activation screening, which systematically showed the physiological relevance of lncRNAs in promoting HCC progression. As proof-of-principle, we showed that CASC11 is a prominent driver in HCC progression through regulating the expression of MYC and its downstream targets. Further investigations are required to translate the experimental findings into clinical application by developing functional lncRNA candidates as potential biomarkers and therapeutic targets for HCC patients.

Materials and Methods
Clinical Specimens

HCC and their corresponding nontumorous liver tissues were obtained from 16 patients with surgical resection at Queen Mary Hospital between 1997 and 2007. The use of clinical specimens was approved by the Institutional Review Board of the University of Hong Kong and the Hong Kong Hospital Authority.

Cell Culture

The HCC cell line MHCC97H was obtained from Dr Z. Y. Tang (Fudan University, Shanghai, China). Huh-7 was a gift from Dr H. Nakabayashi (Hokkaido University, Hokkaido, Japan). 293FT was obtained from American Type Culture Collection (Manassas, VA). All of these cell lines were cultured in Dulbecco's modified Eagle medium (Gibco, Grand Island, NY) supplemented with 10% fetal bovine serum (Life Technologies) and 100 U/mL penicillin/streptomycin (Invitrogen). MHCC97H cells were supplemented with 1 mmol/L sodium pyruvate.

Cell Line Authentication

Human HCC cell lines were authenticated by the AuthentiFiler PCR Amplification Kit (Life Technologies). To evaluate whether the HCC cell lines used in this study resembled the primary HCC tumors, a comprehensive pan-cancer analysis was performed by comparing the transcriptomic profiles of our cell lines with more than 9753 primary tumors from TCGA and 1378 cell lines from the Cancer Cell Line Encyclopedia encompassing 32 cancer types. Reference-Based Single-Cell RNA-seq Annotation (Single R, San Francisco, CA) was used to infer the cell of origin of our tested cell lines by using the aforementioned reference transcriptomic data sets.

Pooled Library Amplification and Transduction

Human CRISPR 3–plasmid lncRNA activation pooled library (1000000106; Addgene, Watertown, MA) was a gift from Professor Feng Zhang. Briefly, the Refseq noncoding RNAs (release 69) catalog and the Broad Institute lncRNA catalog were combined to filter for all major classifications of lncRNAs no less than 200-bp long and their TSSs $\geq$50-bp away from the neighboring gene[33,34]. This results in a library composed of 95,958 sgRNAs targeting 10,504 lncRNAs. Each lncRNA TSS was targeted by 10 sgRNAs. Each sgRNA was individually cloned into BsmB1 sites of lenti sgRNA (MS2) zeocin (zeo) backbone (61427; Addgene). Five hundred nontargeting sgRNAs were included as nontargeting controls. sgRNA library cloning was described in a protocol by Joung et al.[10] The sgRNA library was amplified at 50–100 ng/μL using Endura Electrocompetent cells (60242; Lucigen, Middleton, WI) according to the manufacturer's instructions.

Figure 11. (***See previous page***). **CASC11 regulates cell-cycle progression through MYC.** (*A*) Western blot showing the protein expression of MYC in a panel of HCC cell lines and 2 immortalized normal liver cell lines. α-tubulin served as housekeeping normalization. (*B*) Transcript level of MYC and CASC11 in a panel of HCC cell lines and 1 pair of HCC clinical samples as determined by RT-qPCR. Data were normalized with U6 and compared with the nontumor (NT) liver sample (67 NT). (*C*) RT-qPCR validation of cell-cycle–related genes upon LNA-mediated CASC11 knockdown for 48 hours. Relative expression was normalized to U6. (*D*) Western blot showing protein expression of CDK4, CDK6, MYC, and α-tubulin upon CASC11 silencing in MHCC-97H, HepG2, and Huh7 cells. (*E*) RT-qPCR measurement showing the up-regulation of CASC11-regulated genes as a result of CASC11 overexpression. (*F*) Cell-cycle profiles of MHCC-97H and HepG2 cells by flow cytometry using propidium iodide (PI) staining upon CASC11 knockdown for 72 hours. (*G*) Cell-cycle profiles of MHCC-97H cells by flow cytometry using PI staining after activation of CASC11. (*H*) Cell-cycle profiles of CASC11-overexpressing cells infected with MYC targeting siRNAs. Result is presented with respect to the control. (*A–H*) Results represent mean values of 3 independent experiments. Statistical significance was evaluated by 1-way analysis of variance with the Dunnett multiple comparison test. Data are presented as means $\pm$ SD, *$P < .05$, **$P < .01$, ***$P < .005$, and ****$P < .001$. RT-qPCR, Quantitative reverse transcription PCR.

Figure 12. The regulatory role of CASC11 and MYC on cell-cycle–associated genes. (A) ChIP-seq profiles from the Encyclopedia of DNA Elements, showing enriched signals of MYC on promoter regions of CDK4 and CDK6. (B) Spearman correlation between CASC11 and MYC or CDK4 or CDK6 in HCC clinical samples (n = 50). The gene expression in TCGA transcriptomic data was added with a pseudocount representing the lowest FPKM value and expressed as $\log_2$-transformed FPKM. Spearman correlation was used to evaluate the expression of data. (C) RT-qPCR analysis showing the expression changes of cell-cycle–related genes at 48 hours post-transfection with LNA targeting CASC11 or siRNA targeting MYC, respectively. Data are presented as mean ± SD from three independent experiments (D) Gene Ontology Pathway analysis showing the enrichment of common downstream targets as a result of CASC11 knockdown and MYC knockout. FC, fold change; FPKM, fragments per kilobase million; RT-qPCR, Quantitative reverse transcription PCR .
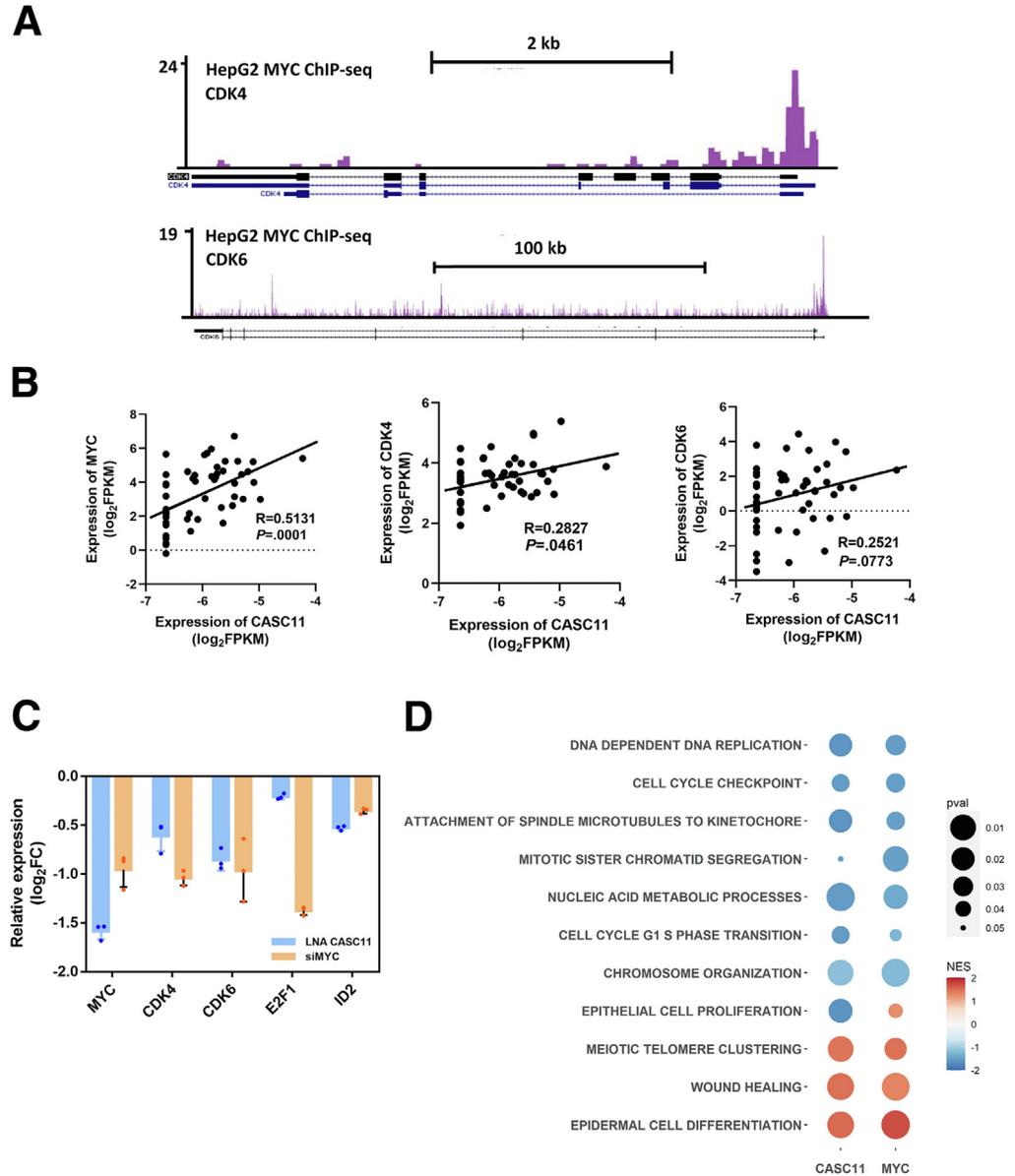
For transduction, plasmids were packaged into lentivirus with the appropriate lentiviral packaging plasmid system (pPACKH1-GAG, pPACKH1-REV, and pVSV-G; System Biosciences, Palo Alto, CA) in HEK293FT cells using Lipofectamine 3000 (Life Technologies) according to the manufacturer's instructions. CRISPR SAM is a 3-vector system: dCas9–VP64–blast (61425; Addgene), MS2–P65–HSF1 activator complex with a 2A Hygromycin resistance marker (MPHv2, 89308; Addgene), and lenti-sgRNA (MS2) zeo (61427; Addgene). MHCC97H cells were transduced with dCas9–VP64 and MPHv2. The successfully transduced HCC cells were selected by blasticidin (10 μg/mL) (Life Technologies) and hygromycin (300 μg/mL) (Life Technologies) for 5 days.

Before the sgRNA library transduction, the concentration of the selection reagents (ie, zeocin) was determined by a kill curve (Life Technologies). Lentiviral titer was calculated by adding different doses of lentiviral supernatant to a fixed number of cells and determining the cell viability after complete selection for 7 days. The multiplicity of infection (MOI) was calculated as the number of transduced cells under antibiotic selection divided by the number of cells without antibiotic selection. MHCC-97H cells expressing SAM were transduced with lentiviral sgRNA library at a low MOI (MOI, <0.3) to make sure each cell contained only 1 sgRNA. Transduced cells were selected in 300 ug/mL zeocin (Life Technologies) for 7 days. Cells were maintained at more than 500 cells per sgRNA during passing to ensure sufficient library coverage. After a week of zeocin selection, successful clones were divided into pretransplantation cells and cells for in vivo screening. For sgRNA validation, sgRNAs were individually cloned into the BsmB1 site of

lenti-sgRNA (MS2) optimized puro backbone. The cloned sgRNAs were delivered in SAM complex expressing MHCC97H cells, followed by 2 μg/mL puromycin selection for 5 days.

In Vivo Library Screening

All animal works were performed under the guidelines of the Animal Experimentation Ethics Committee of the University of Hong Kong. A total of 8×10^7 sgRNA library transduced MHCC97H cells were used to ensure library representation. Cells were resuspended in phosphate-buffered saline and Matrigel (BD Biosciences, Franklin Lakes, NJ) (1:1 ratio) and injected into both flanks of 4- to 6-week-old male BALB/c nude mice at a density of 2×10^6 cells per flank. At 4 weeks after transplantation, mice were killed.

Genomic DNA Extraction

To maintain the full library representation, 330 μg genomic DNA was used. Genomic DNA was extracted from tumors and pretransplantation cells by salt precipitation as previously described by Chen et al.[7] Briefly, tumor tissues were frozen, ground, and incubated with lysis buffer (50 mmol/L Tris, 50 mmol/L EDTA, 1% sodium dodecyl sulfate [SDS], pH 8) and proteinase K (Life Technologies) at 55ºC overnight. The lysate mixture was incubated with RNaseA (Life Technologies), 7.5 mol/L ammonium acetate, isopropanol, and ethanol for genomic DNA precipitation.

PCR Amplification of sgRNA Regions and sgRNA Readout by Deep Sequencing

Before deep-sequencing analysis, sgRNA regions were amplified from genomic DNA using NEBnext High Fidelity 2× Master Mix (New England Biolabs, Ipswich, MA) with 24 cycles of PCR reactions. U6 primers for sgRNA amplification are listed in Supplementary Table 5. A total of 330 μg genomic DNA was used per sample to maintain the library representation. PCR products were purified by gel electrophoresis at 100 V for 50 minutes, followed by visualization under UV illumination in Alphalmager 2200 (Alpha Innotech, San Leandro, CA), followed by gel purification (Qiagen) and ethanol precipitation. Massive parallel amplicon pair-ended sequencing analysis was performed by Novogene Technology to evaluate sgRNA abundance (Beijing, China). Results from 2 independent deep-sequencing results were analyzed by the MAGeCK algorithm (0.5.7).

Computational Analysis of CRISPR Activation Screening

We used the MAGeCK algorithm (0.5.7) to identify positively selected lncRNAs in the library screening. Briefly, reads from Illumina sequencing (San Diego, CA) were demultiplexed by Trim-Galore (version 0.6.5) (RRID:SCR_011847). The MAGeCK algorithm count was used to annotate the sequencing reads and adjust the read count distribution and sequencing depth of all samples using median normalization. sgRNA sequences were filtered

further based on the following criteria: duplicate entries of the same sgRNA sequence; and sgRNA sequence representing more than 1 lncRNA loci. The average fold change of normalized sgRNA read counts, negative binomial P value, and FDR of each sgRNA across independent experimental replicates were calculated using the MAGeCK test function. lncRNAs were considered to be positively selected if they were targeted by at least 2 corresponding sgRNAs with a FDR of less than 5%.

cDNA Overexpression

A lentiviral expression Cytomegalovirus vector (CMV) with multiple cloning sites (MCS) carrying Flag and HA tags (pCDH–CMV–MCS–3X–Flag–3X–HA–puro vector) was used to express the full length of CASC11 cDNA. Briefly, pCDH–CMV–MCS–3X–Flag–3X–HA–puro vector was digested with Not1 and Xba1to remove HA and Flag sequence. The full length of CASC11 cDNA was cloned into the digested pCDH vector. The confirmed clone was verified by sequencing. MHCC97H cells and the Alexander hepatoma (PLC/PRF/5) cells were transfected with either pCDH–CASC11 vector or pCDH empty vector using the Lipofectamine 3000 transfection kit (Life Technologies) at a MOI < 0.5 and subsequently selected with puromycin for 3 days. The transcript levels were measured by qPCR. The sequence of cDNA overexpression is listed in Supplementary Table 5.

Nuclear and Cytoplasmic Fractionation

Nuclear and cytoplasmic fractionation of MHCC97H cells were prepared using the RNA subcellular Isolation Kit (Active Motif, Carlsbad, CA) as described by the manufacturer's protocol. GAPDH glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as a cytoplasmic endogenous control, while U6 small nuclear 1 (U6) was used as the nuclear endogenous control. The efficiency of fractionation was determined by qPCR, followed by 2% agarose gel electrophoresis in 1× Tris–borate–EDTA buffer.

Cell Proliferation Assay and Colony Formation Assay

For the cell proliferation assay, cells were seeded at a density of 10,000 cells/well in a 24-well plate in triplicate and incubated in a 37ºC humidified CO_2 incubator. The number of cells was quantified for 6 days by a Z1 Coulter Counter Cell and Particle Counter (Beckman Coulter, Brea, CA). For colony formation assays, 1000 cells were seeded in a 6-well plate until colonies formed in complete medium (approximately 14 days). Colonies were fixed in methanol and stained with 0.1% crystal violet (Sigma-Aldrich, Burlington, MA) for visualization. ImageJ (National Institutes of Health, Bethesda, MD) was used to count the number of colonies.

Cell-Cycle Analysis

Cells were fixed with 75% ethanol at 4ºC overnight. The next day, cells were treated with 10 μL of 2 mg/mL

propidium iodide (Sigma-Aldrich) and 10 mg/mL RNase-A (Life Technologies) at 37°C for 30 minutes. Analysis was performed on a FACSCanto II Flow Cytometer (BD Biosciences).

Reverse-Transcription and Real-Time PCR

Total RNA was extracted by TRIzol reagent (Invitrogen). First-strand cDNA was synthesized from 1 μg total RNA using Fast PrimeScript Reverse Transcription (RT) Master Mix (TaKaRa, San Jose, CA). qPCR analysis was performed using SYBR green master mix (Life Technologies) on an Applied Biosystem StepOnePlus Real-time PCR system (Life Technologies). Sequences for qRT-PCR primers are listed in Supplementary Table 5. Hypoxanthine-guanine phosphor-ibosyltransferase (HPRT) or U6 served as an internal control. Relative RNA levels were calculated using comparative cycle threshold (CT) ($2^{-\Delta\Delta CT}$).

Western Blot

Cells were lysed in RIPA buffer (50 mmol/L Tris-HCl, 150 mmol/L NaCl, 0.1% SDS, 5 mmol/L EDTA, and 0.5% sodium deoxycholate) containing proteinase and phosphatase inhibitors (Roche, Basel, Switzerland). Protein samples were separated on SDS–polyacrylamide gel electrophoresis, transferred to a hydrophobic polyvinyl iodine difluoride membrane (GE Healthcare, Chicago, IL), and probed with primary antibodies at 4°C overnight and secondary antibodies at room temperature for 1 hour. Antibodies used for Western blot were obtained from Cell Signalling Technology and Santa Cruz: anti-MYC antibody (9402; Cell Signalling Technology, Danvers, MA), anti-CDK6 antibody (D4S8S; Cell Signalling Technology), anti-CDK4 antibody (sc-23896; Santa Cruz, Santa Cruz, CA), and anti–α-tubulin antibody (2144S; Cell Signalling Technology).

RNA Interference

CASC11 targeting shRNA sequences were cloned into the pLKO.1 puro plasmid. siRNAs targeting MYC were designed by Integrated DNA Technology (Coralville, IA). GapmeR antisense (Exiqon, Skelstedet, Vedbaek) oligonucleotides (LNA) were designed by Qiagen. siRNAs and LNAs were transfected at 500 nmol/L and 50 nmol/L using RNAiMAX transfection reagent (Life Technologies) or Lipofectamine 3000 (Life Technologies) according to the manufacturer's protocol. Sequences of shRNAs, siRNAs, and LNAs are listed in Supplementary Table 5.

ChIRP-seq

ChIRP-qPCR assay was performed as described previously.[35] A total of 11 antisense biotinylated probes against CASC11 was designed by LGC Biosearch Technologies (Petaluma, CA). The probes were split into 2 pools (probe set A and B). Two independent ChIRP-seq runs with 2 pools were performed separately. Briefly, 1 × 10^7 cells were cross-linked with 1% glutaraldehyde for 10 minutes (Sigma-Aldrich) and quenched with 0.125 mol/L glycine (Sigma-Aldrich) for 5 minutes. Cells were lysed in complete lysis

buffer (50 mmol/L Tris-HCl, pH 7.0, 10 mmol/L EDTA, 1% SDS, Protease inhibitor Cocktail III (Calbiochem, San Diego, CA) and RNase inhibitor (Life Technologies) and sonicated with a Bioruptor sonicator (Bioruptor; Diagenode, Liège, Belgium) to an average length of 100–500 bp. Chromatin was hybridized with 100 pmol of CASC11 antisense probes and complete hybridization buffer (Protease inhibitor Cocktail and RNase inhibitor) and incubated at 37°C overnight. The RNA-chromatin hybrid was captured by Strep-tavidin C1 magnetic beads (Life Technologies) and washed with wash buffer (2× saline sodium citrate, 0.5% SDS, and Protease inhibitor Cocktail). Ten percent of bead samples were purified for RNA isolation and 90% of bead samples for DNA isolation. For RNA isolation, bead samples were subjected to proteinase K and RNA elution buffer (Tris, pH 7.0, 1% SDS) and incubated at 95°C for 15 minutes, followed by TRIzol-chloroform extraction and the miRNeasy Mini column purification (Qiagen). One-step, real-time, reverse-transcription qPCR was performed to quantify the enriched transcripts in eluted RNA. For DNA isolation, bead samples were resuspended in complete DNA elution buffer (50 mmol/L) and incubated at 37°C for 30 minutes with gentle rotation. Then, samples were subjected to RNaseA, RNaseH, and proteinase K treatment and incubated at 50°C for 45 minutes, followed by phenol:chloroform:isoamyl:alcohol extraction (Sigma). Eluted DNA was quantified by qPCR or sequencing library construction. For ChIRP-qPCR, bound chromatin was assessed using primers targeting the MYC promoter region. ChIRP pull-down of GAPDH served as a negative control. For ChIRP-seq, the library preparation and deep sequencing were conducted by the Centre for PanorOmic Sciences, HKU. ChIRP-seq raw reads were uniquely mapped to the human genome (GRCh38) using Bowtie2 (version 2.4.1) (RRID:SCR_016368), followed by a peak calling algorithm using MACS2 (version 2.2.7) (RRID:SCR_013291) with a q-value cut-off of 0.01. Fold enrichment of chromatin association of CASC11 was calculated by normalizing the common peak signals of 2 pools to input. Peak signals of probe set A and probe set B samples were normalized to input. Probes and primer sequences are listed in Supplementary Table 5.

Clinical Investigation of Positively Selected lncRNAs in HCC Clinical Samples

Positively selected lncRNAs were dichotomized into high- and low-expression groups based on either mean or median expression cut-off level, whichever gave the greatest degree of discrepancy. The overall survival rate in HCC patients above and below the cut-off expression level was calculated using the Kaplan–Meier method. To investigate the effect of positively selected lncRNAs on survival, lncRNA expression in relation to clinical parameters including tumor stage (tumor stages 1 and 2 vs tumor stages 3 and 4), histologic grade (histologic grades 1 and 2 vs histologic grades 3 and 4), gender (male vs female), and age (< median age vs > median age) were assessed in a multivariate Cox regression model. Hazard ratios were calculated to evaluate the prognostic effect of lncRNAs.

RNA-Seq and Bioinformatics Analysis

Transcriptome sequencing (RNA-seq) was performed in 16 pairs of HCC samples and the corresponding nontumor samples (Bioproject Accession ID: 294031). Library preparation and RNA-seq of CASC11 knockdown cells and MYC knockout cells were performed as described previously.[5] Differential expression analysis was performed using EdgeR (RRID:SCR_012802) with default parameters. Gene set enrichment analysis was performed using fast gene set enrichment analysis R package (version 3.10) (RRID:SCR_001905) with default parameters. Pathway annotations of Hallmark gene sets were retrieved from the Molecular Signatures Database and Gene Ontology.

ChIP-Seq Analysis

ChIP-seq of MYC was obtained from the Encyclopedia of DNA Elements. Read alignment was performed by Bowtie2 and peak calling by MACS2.

Pathway Enrichment Analysis

Gene set enrichment analysis was performed to evaluate genes that are over-represented in the experimental condition vs the control by fast gene set enrichment analysis R package (version 3.10) with default parameters. Annotations of different gene sets were retrieved from the Molecular Signatures Database.

TCGA Data

Clinical information of HCC patients including RNA-seq expression profiles of HCC patients (FPKM value), disease-free and overall survival, as well as histologic grades were downloaded from TCGA via the Broad Institute (http://gdac.broadinstitute.org). Gene set enrichment analysis on TCGA data was performed using gene set enrichment analysis version 4.0.3 (https://www.gsea-msigdb.org/gsea/downloads.jsp).

Prediction of Protein-Coding Capacity of lncRNA

The protein-coding potential of lncRNAs was predicted using CPAT (http://lilab.research.bcm.edu/cpat) and ORF finder (https://www.ncbi.nlm.nih.gov/orffinder). Homo sapiens HOX transcript antisense RNA (HOTAIR) served as the noncoding RNA control while GAPDH served as coding control.

Statistical Analysis

All statistical analysis was performed using RStudio (version 3.6.1) and GraphPad PRISM software (version 9.2; GraphPad Software, San Diego, CA). Data are presented as means ± SEM. The Student t test or 1-way analysis of variance was used to determine the difference between 2 groups or more than 2 groups, followed by the Dunnett multiple comparison test. A Kaplan–Meier plot was used to analyze patient survival data. In all comparisons, $P < .05$ was considered statistically significant. At least 3 biological replicates were performed. Graphical abstract was created with Biorender.com.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre L, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68:394–424.
2. Strausberg RL, Riggins GJ. Navigating the human transcriptome. Proc Natl Acad Sci U S A 2001; 98:11837–11838.
3. Wong CM, Tsang FH, Ng IO. Non-coding RNAs in hepatocellular carcinoma: molecular functions and pathological implications. Nat Rev Gastroenterol Hepatol 2018;15:137–151.
4. Schmitt AM, Chang HY. Long noncoding RNAs in cancer pathways. Cancer Cell 2016;29:452–463.
5. Wei L, Lee D, Law CT, Zhang MS, Shen J, Chin DW, Zhang A, Tsang FH, Wong CL, Ng IO, Wong CC, Wong CM. Genome-wide CRISPR/Cas9 library screening identified PHGDH as a critical driver for sorafenib resistance in HCC. Nat Commun 2019;10:4681.
6. Joung J, Konermann S, Gootenberg JS, Abudayyeh OO, Platt RJ, Brigham MD, Sanjana NE, Zhang F. Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. Nat Protoc 2017;12:828–863.
7. Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, Scott DA, Song J, Pan JQ, Weissleder R, Lee H, Zhang F, Sharp PA. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. Cell 2015; 160:1246–1260.
8. Song CQ, Li Y, Mou H, Moore J, Park A, Pomyen Y, Hough S, Kennedy Z, Fischer A, Yin H, Anderson DG, Conte D Jr, Zender L, Wang XW, Thorgeirsson S, Weng Z, Xue W. Genome-wide CRISPR screen identifies regulators of mitogen-activated protein kinase as suppressors of liver tumors in mice. Gastroenterology 2017; 152:1161–1173.e1.
9. Yau EH, Kummetha IR, Lichinchi G, Tang R, Zhang Y, Rana TM. Genome-wide CRISPR screen for essential cell growth mediators in mutant KRAS colorectal cancers. Cancer Res 2017;77:6330–6339.
10. Joung J, Engreitz JM, Konermann S, Abudayyeh OO, Verdine VK, Aguet F, Gootenberg JS, Sanjana NE, Wright JB, Fulco CP, Tseng YY, Yoon CH, Boehm JS, Lander ES, Zhang F. Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. Nature 2017;548:343–346.
11. Bester AC, Lee JD, Chavez A, Lee YR, Nachmani D, Vora S, Victor J, Sauvageau M, Monteleone E, Rinn JL, Provero P, Church GM, Clohessy JG, Pandolfi PP. An integrated genome-wide CRISPRa approach to functionalize lncRNAs in drug resistance. Cell 2018; 173:649–664.e20.
12. Whiteside TL. The tumor microenvironment and its role in promoting tumor growth. Oncogene 2008; 27:5904–5912.

13. Lin MF, Jungreis I, Kellis M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics 2011;27:i275–i282.

14. Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res 2013;41:e74.

15. Esposito R, Bosch N, Lanzos A, Polidori T, Pulido-Quetglas C, Johnson R. Hacking the cancer genome: profiling therapeutically actionable long non-coding RNAs using CRISPR-Cas9 screening. Cancer Cell 2019;35:545–557.

16. Akbari Dilmaghnai N, Shoorei H, Sharifi G, Mohaqiq M, Majidpoor J, Dinger ME, Taheri M, Ghafouri-Fard S. Non-coding RNAs modulate function of extracellular matrix proteins. Biomed Pharmacother 2021;136:111240.

17. Wong LS, Wong CM. Decoding the roles of long non-coding RNAs in hepatocellular carcinoma. Int J Mol Sci 2021;22:3137.

18. Manguso RT, Pope HW, Zimmer MD, Brown FD, Yates KB, Miller BC, Collins NB, Bi K, LaFleur MW, Juneja VR, Weiss SA, Lo J, Fisher DE, Miao D, Van Allen E, Root DE, Sharpe AH, Doench JG, Haining WN. In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. Nature 2017;547:413–418.

19. Kodama M, Kodama T, Newberg JY, Katayama H, Kobayashi M, Hanash SM, Yoshihara K, Wei Z, Tien JC, Rangel R, Hashimoto K, Mabuchi S, Sawada K, Kimura T, Copeland NG, Jenkins NA. In vivo loss-of-function screens identify KPNB1 as a new druggable oncogene in epithelial ovarian cancer. Proc Natl Acad Sci U S A 2017;114:E7301–E7310.

20. Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, Beckwith CA, Chan JA, Hills A, Davis M, Yao K, Kehoe SM, Lenz HJ, Haiman CA, Yan C, Henderson BE, Frenkel B, Barretina J, Bass A, Tabernero J, Baselga J, Regan MM, Manak JR, Shivdasani R, Coetzee GA, Freedman ML. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet 2009; 41:882–884.

21. Meyer N, Penn LZ. Reflecting on 25 years with MYC. Nat Rev Cancer 2008;8:976–990.

22. Chen H, Liu H, Qing G. Targeting oncogenic Myc as a strategy for cancer treatment. Signal Transduct Target Ther 2018;3:5.

23. Jiang Y, Jiang YY, Xie JJ, Mayakonda A, Hazawa M, Chen L, Xiao JF, Li CQ, Huang ML, Ding LW, Sun QY, Xu L, Kanojia D, Jeitany M, Deng JW, Liao LD, Soukiasian HJ, Berman BP, Hao JJ, Xu LY, Li EM, Wang MR, Bi XG, Lin DC, Koeffler HP. Co-activation of super-enhancer-driven CCAT1 by TP63 and SOX2 promotes squamous cancer progression. Nat Commun 2018;9:3619.

24. Ling H, Spizzo R, Atlasi Y, Nicoloso M, Shimizu M, Redis RS, Nishida N, Gafa R, Song J, Guo Z, Ivan C, Barbarotto E, De Vries I, Zhang X, Ferracin M, Churchman M, van Galen JF, Beverloo BH, Shariati M, Haderk F, Estecio MR, Garcia-Manero G, Patijn GA, Gotley DC, Bhardwaj V, Shureiqi I, Sen S, Multani AS, Welsh J, Yamamoto K, Taniguchi I, Song MA, Gallinger S, Casey G, Thibodeau SN, Le Marchand L, Tiirikainen M, Mani SA, Zhang W, Davuluri RV, Mimori K, Mori M, Sieuwerts AM, Martens JW, Tomlinson I, Negrini M, Berindan-Neagoe I, Foekens JA, Hamilton SR, Lanza G, Kopetz S, Fodde R, Calin GA. CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res 2013;23:1446–1461.

25. Kajino T, Shimamura T, Gong S, Yanagisawa K, Ida L, Nakatochi M, Griesing S, Shimada Y, Kano K, Suzuki M, Miyano S, Takahashi T. Divergent lncRNA MYMLR regulates MYC by eliciting DNA looping and promoter-enhancer interaction. EMBO J 2019;38:e98441.

26. Cheng N, Wu J, Yin M, Xu J, Wang Y, Chen X, Nie Z, Yin J. LncRNA CASC11 promotes cancer cell proliferation in hepatocellular carcinoma by inhibiting miRNA-188-5p. Biosci Rep 2019;39:BSR20190251.

27. Han Y, Chen M, Wang A, Fan X. STAT3-induced upregulation of lncRNA CASC11 promotes the cell migration, invasion and epithelial-mesenchymal transition in hepatocellular carcinoma by epigenetically silencing PTEN and activating PI3K/AKT signaling pathway. Biochem Biophys Res Commun 2019;508:472–479.

28. Song H, Liu Y, Li X, Chen S, Xie R, Chen D, Gao H, Wang G, Cai B, Yang X. Long noncoding RNA CASC11 promotes hepatocarcinogenesis and HCC progression through EIF4A3-mediated E2F1 activation. Clin Transl Med 2020;10:e220.

29. Zhang Z, Zhou C, Chang Y, Zhang Z, Hu Y, Zhang F, Lu Y, Zheng L, Zhang W, Li X, Li X. Long non-coding RNA CASC11 interacts with hnRNP-K and activates the WNT/beta-catenin pathway to promote growth and metastasis in colorectal cancer. Cancer Lett 2016; 376:62–73.

30. Boque-Sastre R, Soler M, Oliveira-Mateos C, Portela A, Moutinho C, Sayols S, Villanueva A, Esteller M, Guil S. Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. Proc Natl Acad Sci U S A 2015;112:5785–5790.

31. Fox AH, Nakagawa S, Hirose T, Bond CS. Paraspeckles: where long noncoding RNA meets phase separation. Trends Biochem Sci 2018;43:124–135.

32. Bertoli C, Skotheim JM, de Bruin RA. Control of cell cycle transcription during G1 and S phases. Nat Rev Mol Cell Biol 2013;14:518–528.

33. Frankish A, Diekhans M, Jungreis I, Lagarde J, Loveland Jane E, Mudge JM, Sisu C, Wright JC, Armstrong J, Barnes I, Berry A, Bignell A, Boix C, Carbonell Sala S, Cunningham F, Di Domenico T, Donaldson S, Fiddes Ian T, García Girón C, Gonzalez JM, Grego T, Hardy M, Hourlier T, Howe KL, Hunt T, Izuogu OG, Johnson R, Martin FJ, Martínez L, Mohanan S, Muir P, Navarro FCP, Parker A, Pei B, Pozo F, Riera FC, Ruffier M, Schmitt BM, Stapleton E, Suner M-M, Sycheva I, Uszczynska-Ratajczak B, Wolf MY, Xu J, Yang Yucheng T, Yates A, Zerbino D, Zhang Y, Choudhary Jyoti S, Gerstein M, Guigó R, Hubbard TJP, Kellis M, Paten B, Tress ML, and Flicek P. GENCODE 2021. Nucleic Acids Research. 2020. 49(D1): p. D916-D923.

34. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badretdin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C, Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, and Pruitt KD. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Research. 2015. 44(D1): p. D733-D745.

35. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell 2011; 44:667–678.

Correspondence
Address correspondence to: Chun-Ming Wong, MD, State Key Laboratory of Liver Research and Department of Pathology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong. e-mail: jackwong@pathology.hku.hk; fax: (852) 22185227.

CRediT Authorship Contributions
   Lok Sze Wong (Conceptualization: Lead; Investigation: Lead; Methodology: Lead; Software: Lead; Writing – original draft: Lead)
   Lai Wei (Conceptualization: Lead; Data curation: Lead; Methodology: Equal)
   Geng-Chao Wang (Data curation: Supporting; Resources: Supporting)
   Cheuk-Ting Law (Resources: Equal; Software: Equal)
   Ho-Ching Tsang (Data curation: Supporting; Methodology: Supporting)
   Wai-Ching Chin (Resources: Supporting; Validation: Supporting)
   Irene OL Ng (Funding acquisition: Lead; Supervision: Supporting)
   Chun-Ming WONG (Conceptualization: Lead; Funding acquisition: Lead; Writing – review & editing: Lead)

Ethical Approval and Consent to Participate
The use of clinical specimens was approved by the institutional review board of the University of Hong Kong and the Hong Kong Hospital Authority. Animal experiments were performed with the approval of the Committee on the Use of Live Animals in Teaching and Research of the University of Hong Kong. All experimental procedures strictly followed the animals (Control of Experiments) ordinance of Hong Kong.

Data Availability Statement
The authors confirmed that the data supporting the findings of the study are available in the article and Supplementary materials. Supplementary tables are provided as Microsoft Excel tables (Redmond, WA). The genome-wide transcriptomic sequencing and chromatin isolation by RNA purification sequencing are available in the Sequence Read Archive (SRA) database: PRJNA786081. This article does not generate original code. Any additional information required to re-analyze the data reported in this article is available from the lead contact upon request.

Conflicts of interest
The authors disclose no conflicts.