SUPPLEMENTAL INFORMATION

Supplemental Information includes:

- Materials and Methods section
- Modeling and Docking section
- Legends to Supplementary Figures
- 9 Supplementary figures
- 8 Supplementary tables
- Supplemental Information References

MATERIALS AND METHODS

Cell cultures and drug treatment. Human hepatoma HepG2 cells, the HepG2-derived clones HepG2.2.15 and HepAD38 and the NTCP-HepG2 cells were cultured in supplemented Dulbecco's modified Eagle's medium (DMEM) as described [1] and maintained in a 5% CO2 humidified incubator at 37°C.

Primary cultures of human hepatocytes and HBV infections. Primary human hepatocytes (PHHs) (provided by Prof. M. Rivoire, Centre Leon Bérard, Lyon, France) were prepared from HBV, HCV and HIV negative adult patients undergoing lobectomy or segmental liver resection for medically required purposes unrelated to this research program. PHHs were prepared using the protocol described in [2] with minor modifications. Liver samples were first perfused in Solution I (NaCl 58,44 M, KCl 74,56 M, Na2HPO4, 2H2O 177,99 M, Hepes 0,6%, EGTA 0,5 mM; pH7,4) and then in Solution II (NaCl 58,44 M, KCl 74,56 M, Na2HPO4, 2H2O 177,99 M, Hepes 0,6%, CaCl2 110 M) containing 0,4 mg/ml of collagenase from Clostridium histolyticum (Sigma-Aldrich, #C5138). Cells were seeded at $2{,}5x10^5$ cells/cm² on collagen type IV pretreated plates (Corning, #354236) and cultured overnight in William's medium (Life Technologies, #22551-089) supplemented with 10% Fetal Clone II (GE Healthcare), 1% penicillin/streptomycin (Invitrogen, #15140122), 1% Glutamax X100 (Invitrogen, #35050038), 5 µg/ml insulin (Sigma–Aldrich, #I9278) and 5×10^{-7} M hydrocortisone (Upjohn Laboratories). PHHs were then extensively washed in serum-free medium, kept in serum-free medium for 24 h to counter-select the growth of contaminating fibroblast and endothelial cells and then plated in complete William's medium. PHHs were treated with 2% of DMSO (Sigma–Aldrich, #D2650) for 24 h before HBV infection and then incubated for 16 h with a wt HBV inoculum produced in HepAD38 cells at a multiplicity of infection of 1000 vge (virus genome equivalents)/cell, in presence of 4% PEG-8000 (polyethylene glycol, Sigma–Aldrich, #1546605). In the case of the HBV (x-) inoculum cells are incubated for 16 h at a multiplicity of infection of 500 vge (virus genome equivalents)/cell. For inhibition of DLEU2 in PHHs, mixed LNA/DNA oligonucleotides (Gapmers) against human DLEU2 ncRNA were purchased from Exiqon (Cat#300600). DLEU2

specific Gapmers and scrambled control Gapmers (CTL) were delivered to cells by Gymnosis at a final concentration of 50 nM, 24h before HBV infection (MOI 100).

HBV viral inoculum preparation and infection. The wild type and the HBx defective (HBV (x-) HBV inocula used in this study are concentrated as previously described [3-4] from the supernatant of HepAD38 and HepG2 H1.3∆x cells, respectively. HepAD38 is an HepG2-derived stable cell line carrying a 1.3 HBV transgene (genotype D, serotype ayw) under the control of a tet-off promoter. The HepG2 H1.3∆x cells carry a stable integration of a 1.3 fold HBV genome (Genotype D, subtype ayw) with premature stop codon mutations in both the 5' and 3' HBx ORFs. HBV (x-) [5]. HepAD38 and HepG2 H1.3∆x cells were cultured in HYPERFlask® (Corning, #10020) coated with type IV collagen (Corning, #354236) in 550 ml of complete DMEM-F12 medium supplemented with 10% decomplemented FBS (Gibco, #10270-106), 1% penicillin/streptomycin (Invitrogen, #15140122) and 1% sodium-pyruvate (Invitrogen, #11360039). Supernatants (550 ml/flask) were harvested twice a week for 3-4 months, clarified through 0.45 µm and then 0.22 µm filters (Millipore, #10785534) and precipitated overnight at 4°C with 8% PEG-8000 (Sigma–Aldrich, #1546605). The precipitates were centrifuged at 3500 g for 1 h and the pellets resuspended in Opti-MEM (Invitrogen, #31985070) to achieve a 50- to 100-fold concentration. Each HBV inoculum, after DNA extraction (QIAmp Ultrasens Virus kit, Qiagen), was titered by qPCR using serial dilutions of an HBV plasmid to build a standard curve. Primers are detailed in Table S8. All viral preparations were tested to confirm the absence of endotoxins (Lonza Verviers, Belgium).

Patient samples

The HCC samples (tumor and non-tumor) analysed in Figure 1g are from the 'Biobanque INSERM U1052 - CRCL Hépatologie (French IRB 'CPP Sud-Est IV' approval #11/040 / 2011). Written informed consent was obtained from each patient.

Transient transfection of full-length HBV DNA genomes. Monomeric linear full-length wild-type (WT) and HBx mutant (HBx mt) HBV genomes were released from the pCR.HBV.A.EcoRI and the

pCR.HBXmt.A.EcoRI plasmids [5] using EcoRI-PvuI (New England Biolabs). Linear HBV monomers were transfected into HepG2 cells using the Mirrus Bio trans IT-LT1 reagent (Mirrus, Cat#MIR2305) as previously described [1]. For inhibition of DLEU2 in HepG2 cells, mixed LNA/DNA oligonucleotides (Gapmers) were generated (Exiqon, Cat#300600) against human DLEU2 ncRNA. DLEU2 specific Gapmers and scrambled control Gapmers (CTL) were transfected at a final concentration of 40 nM cells using the Lipofectamine Plus reagent (Invitrogen Cat#11514015 and Cat#18324020).

Western blot. For Western blot analysis, cells were lysed and sonicated in ice-cold RIPA buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol, 0,2% NP-40) supplemented with protease inhibitors (Roche). For immunoprecipitation experiments, cells were lysed in NET lysis buffer (50 mM Tris-HCl pH 7,5, 150 mM NaCl, 1 mM $MgCl_2$, 0,5% NP-40, 0,01% NaN_3) with protease inhibitors (Roche). Immunoprecipitations were performed on 1 mg of protein extracts using anti-HBx antibody (ThermoFisher, Cat#MA1-081) over night at 4 °C. G Plus agarose beads (50 µl slurry for IP, Pierce cat. #22851) were added for 1 h at 4 °C. Immunocomplexes were washed 3 times with NET buffer 2 (50 mM Tris-HCl pH 7,5, 150 mM NaCl, 1 mM $MgCl_2$, 0,1% NP-40, 0,01% NaN_3) and eluted in Laemmli sample buffer. Protein lysates were separated on sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis gels and transferred onto nitrocellulose membranes. After probing with the different primary antibodies (anti-HBx ThermoFisher, cat#MA1-081 or Abcam cat#ab39716, anti EZH2 CST cat#5246S, anti Suz12 CST cat#3737S, anti-actin CST cat#4970S) and horseradish peroxydase-coupled secondary antibodies, chemoluminescence signals were captured with the ChemiDoc Imaging System (Bio-Rad).

***ChIP assays*.** Cells were resuspended in 1-2 ml of ChIP lysis buffer (50 mM Tris HCL, pH 8, 0,5% NP40, 1 mM EDTA, 100 mM NaCl) and incubated 10 min at 4°C. The lysates were centrifuged at 10.000 g for 2 minutes to pellet the nuclei. The supernatants were removed, and the nuclei fixed in 1% formaldehyde for 30 min at 4°C. Isolated cross-linked nuclei were extracted with a 20 mM Tris, pH 8, 3 mM $MgCl_2$, 20 mM KCl buffer containing protease inhibitors, pelleted by microcentrifugation,

and lysed by incubation in SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-chloride, pH 8,1) containing protease inhibitors. The chromatin solution was then sonicated for 5 pulses of 45 s at 80% power to generate 300- to 1000-bp DNA fragments using a Bioruptor Sonicator (Diagenode Inc). One hundred µl of Dynabeads Protein G (Invitrogen, Cat#10003D) were added to each 1 ml chromatin preparation and incubated on a rotator for 14–16 h at 4°C. The antibodies used were anti-HBx (ThermoFisher, Cat#MA1-081), anti-Pol II (Diagenode, Cat#C15200004), anti-AcH4 (Millipore, Cat#06-866), anti-Histone H3 (trimethyl K27) (Abcam, Cat#6002), anti-CTCF (Diagenode, Cat# C15410210-50); anti-EZH2 (Active Motif, Cat#39002) anti-SUZ12 (CST cat#3737S). Immunoprecipitations with nonspecific immunoglobulins (Santa Cruz Biotechnology Inc.) were included in each experiment as negative controls. After the reverse cross-linking, immunoprecipitated chromatin was purified by phenol/chloroform (1:1) extraction associated to Phase Lock Gel (5 Prime, Cat#2302820) and ethanol precipitation. ChIPed chromatin was analyzed by real-time PCR using either primers (NCC1 and CCCAS) and probes (FL and Red) specific for the HBV cccDNA or specific primers for each gene promoter (Table S8).

HBV and cellular mRNAs. Total RNA was extracted from using the TRIzol reagent (Invitrogen) as recommended by the manufacturer. Nuclear and cytoplasmic extraction was performed using the PARIS KIT (Ambion, Cat#AM1921). RNA samples were treated with RQ1 RNase-Free DNase (Promega, Cat#M6101) for 30 min at 37°C and stored until use. RNA quality and quantity were monitored by ethidium bromide staining and UV absorbance. For HBV pgRNA analysis, 2 µg of DNase-treated RNA was reverse transcribed and amplified using the ThermoScript RT-PCR System (Invitrogen, Cat #11146016). Then 2 µl of each cDNA was quantified by real-time PCR analysis (Light Cycler; Roche Diagnostics) using the pgRNA-specific primers and probes indicated in Table S4. The human Beta Actin housekeeping gene Light Cycler Set (Roche Diagnostics) was used to normalize the RNA samples. Cellular mRNAs levels were assayed, after reverse transcription (ThermoScript, Invitrogen, Inc., Carlsbad, US-CA) using random primers, by PCR using MyTaq™ HS DNA Polymerase (Bioline, Cat#BIO-21111) and by quantitative qPCR using the specific TaqMan

FAM-probes (Applied Biosystems, Inc.) and the specific primers listed in Table S8. Results were normalized to human Beta Actin and human GAPDH.

RNA immunoprecipitation (RIP). RNA immunoprecipitation was performed using 10^6 HepG2 or HepG2-NTCP cells or primary human hepatocytes per antibody condition**.** Nuclear pellets were fixed in 1% formaldehyde for 15 min, resuspended in 1 ml freshly prepared RIP buffer (50 mM KCl, 25 mM Tris pH 7.4, 5 mM EDTA, 0.5 mM DTT, 0.5% NP40, 100 U/ml RNAse inhibitor SUPERASin, Protease Inhibitor Cocktail) and split into two fractions of 500 μl each (noA and IP). Mechanic shearing was then performed using a dounce homogenizer with 15–20 strokes. After centrifugation at 13.000 rpm for 10 min, supernatants were transferred into new Eppendorf tubes and 1% of the total volume was removed and collected as input RNA. To couple the antibodies to the Dynabeads™ Protein G (Invitrogen), 10 ng of anti-HBx or anti-EZH2 were added to 50 µl of beads resuspended in 1ml of Antibody coupling buffer (3% BSA, Protease Inhibitor Cocktail (Roche)) overnight at 4°C, one day prior to the RIP assay. To remove the uncoupled antibodies, beads were washed thrice with the Antibody coupling buffer and equilibrated in 100µl of Antibody coupling buffer. Pre-cleared nuclear lysates were next incubated overnight at 4°C with the beads coupled to the indicated specific antibody or non-specific antibody. Flow through and 5% of the beads was stored for SDS analysis. After three washes in RIP buffer, the beads were subjected to Proteinase K treatment in the Reverse cross-linking buffer [100 mM NaCl, 10 mM Tris-HCl pH 7.0, 1 mM EDTA, 0.5% SDS, 1 mg/ml Proteinase K (Euromedex)] at 50°C for 1 hour.

Finally, co-precipitated RNAs and the input were extracted by TRIzol and eluted with nuclease-free water. Reverse transcription of DNAse-treated RNA was performed with Velo (Invitrogen, Inc., Carlsbad, US-CA) according to manufacturer's instructions. cDNA was used to amplify the DLEU2 regions and MALAT1 and Beta Actin.

RNA pull-down. RNA pull-down was performed using PierceTM Magnetic RNA-Protein Pull-Down Kit (Thermo Scientific; cat.no 20164) using 100 pmol of 3'-end-desthiobiotinylated DLEU2 RNAs (PierceTM RNA 3' End Desthiobiotinylation Kit. Thermo Scientific cat. 20163), 0.5, 0.75, 1 or 1.5 μg

of HBx recombinant purified protein (Prospec cat. HBV-271) and 1.2, 2, 2.5 or 7.5 μg of EZH2 recombinant purified protein (Origene cat. TP302054) following manufacturer's instructions. RNAs containing or not the binding sites of HuR protein, provided in the kit, were used as controls.

Chromatin Isolation by RNA Purification (ChIRP). Antisense DNA probes were designed against DLEU2 isoforms using an online designer at https://www.biosearchtech.com/support/ tools/design-software/chirp-probe-designer. Forty-eight hours after transfection with linear HBV monomers, HepG2 nuclear pellets were crosslinked in presence of 1% glutaraldehyde in PBS for 10 min at room temperature by shaking and then quenched with 0.125 M glycine for 5 min. The pellets were washed in PBS, freezed in liquid nitrogen and stored at -80°C. Pellets were resuspended in 1 ml lysis buffer (50 mM Tris HCl pH 7, 10 mM EDTA, 1% SDS, 1mM PMSF, 100 U/ml RNAase inhibitor (Sigma, Cat #AM2696), protease inhibitor cocktail (PIC, Sigma, Cat #P8340). Chromatin was sonicated by Bioruptor (Diagenode) in a cold room using the following parameters: pulse interval 30 s ON and 45 s OFF, until DNA was in the size range of 100–500 bp. From each condition 10 μl were taken and preserved as DNA/RNA input. Probe hybridization to chromatin was performed using the hybridization buffer (750 mM NaCl, 50 mM Tris-HCl 7.0, 1 mM EDTA, 1% SDS, 15% formamide, PMSF, protease inhibitor cocktail, and RNAse inhibitor). DLEU2 probes (100 pmol) were added to chromatin, mixed by rotation at 37°C overnight. One hundred μl of washed streptavidin magnetic C1 beads (Invitrogen) were added per 100 pmol of probes and incubated for another 30 min at 37°C. Beads/biotin-probes/RNA/chromatin complexes were collected, washed 5 times with wash buffer (2X SSC, 0.5% SDS and PMSF). RNA was extracted after treatment by proteinase K buffer (100mM NaCl, 10mM Tris HCl pH7.5, 1mM EDTA, 0.5% SDS and PMSF, protease inhibitor cocktail, and RNAse inhibitor) with phenol:chloroform extraction and ethanol precipitation. The RNA was analyzed by quantitative reverse-transcription PCR (qRT-PCR). Bound DNA was purified by phenol:chloroform (1:1) extraction associated to Phase Lock Gel (5 Prime, Cat#2302820) and ethanol precipitation and analyzed by real-time PCR amplification using specific primers and LightCycler 480 SYBR Green I Master (Roche) (Table S8).

NanoString nCounter lncRNA gene expression assay and data analysis. NanoString nCounter assays were performed using 100 ng of purified RNA following manufacturer's instructions (NanoString Technologies). The long non-coding RNA probe set was selected from the custom probe set. Sample preparation and hybridization reactions were performed according to manufacturer's instructions (NanoString Technologies). All hybridization reactions were incubated at 65°C for a minimum of 16 h. Hybridized probes were purified and counted on the nCounter Prep Station and Digital Analyzer (NanoString Technologies) following the manufacturer's instructions. For each assay, a high-density scan (600 fields of view) was performed. Data analysis was performed using the nSolver analysis software (NanoString Technologies) (https://www.nanostring.com/products/analysis-software/nsolver) and housekeeping genes were used for data normalization.

Digital droplet PCR (ddPCR). A 22-µL reaction mixture was prepared comprising 11 µL of 2X ddPCR Supermix™ for probes (no dUTP) (Bio-Rad), 1.1 µL of primers and probe mix, and 5 µL of cDNA or DNA. Nucleic acid inputs were adjusted to have acceptable negative events: 10ng for pgRNA duplex PCR, 50ng for cccDNA singleplex PCR and 100ng for DLEU2 duplex PCR. The several PCR in duplex were done combining DLEU2 gene (intron or exon) with the human control GUSB gene (#Hs99999908_m1, Thermofischer) for cDNA, HBV (Pa03453406_s1, Thermofischer) with the human control HBB gene (#Hs00758889_s1, Thermofischer) for DNA, pgRNA (Forward primer ggagtgtggattcgcactcct, reverse primer agattgagatcttctgcgac and probe aggcaggtcccctagaagaagaactcc) with the human control GUSB gene (#Hs99999908_m1, Thermofischer) for cDNA and the singleplex cccDNA (forward primer ccgtgtgcacttcgcttca, reverse primer gcacagcttggaggcttga and probes catggagaccaccgtgaacgccc). Droplet formation was carried out using a QX100 droplet generator. Subsequent amplification was performed in the C1000 Touch™deep-well thermal cycler (Bio-Rad) with a ramp rate of 2 °C/s and the lid heated to 105 °C, according to the Bio-Rad recommendations. First, the enzyme was activated at 95 °C for 10 min followed by 40 cycles of denaturation at 94 °C for 30 s and 60 °C for one minute. The enzyme was deactivated at 98 °C for 10 min and the reaction was kept at 4 °C.

In silico modeling. **a) Structure.** The FASTA sequence of DLEU2 intron 1 (selected region of 439 bp), as given by GenomeBrowser (https://genome.ucsc.edu/cgi-bin/hgGateway), was used to obtain the secondary structure via DotKnot [6,7]. The FASTA sequence and the secondary structure were used to obtain the tertiary structure via the RNAcomposer tool [8]. Secondary structure predictions for HBx, based on amino acid sequence, have been performed with IUPred [9] (http://iupred.enzim.hu/), PONDR [10] (http://www.pondr.com/) and DISOPRED2 [11] (http://bioinf.cs.ucl.ac.uk/web_servers/disopred/disopred_overview/). The HBx structure has been modelled by I-TASSER [12-14] (https://zhanglab.ccmb.med.umich.edu/I-TASSER/), including in the modelling the constraints on cysteine linkages (default parameters values for the homology modeling has been used; pairs of constrained cysteines are C7-C69, C17-C143, C61-C115, C78 -C137). **b) Interaction.** The interaction probabilities for HBx-DLEU2, based on the nucleotide sequence, have been evaluated with the methods RPISeq [15] (http://pridb.gdcb.iastate.edu/RPISeq/) and catRAPID [16,17] (http://s.tartaglialab.com/page/catrapid_group). RNA-binding residues in HBx were predicted via RNABindRPlus [18] (http://ailab1.ist.psu.edu/RNABindRPlus/), based on the amino acid sequence. HBx protein and RNA models were docked by using HEX [19] (http://hex.loria.fr/) and NPDock [20-22] (http://genesilico.pl/NPDock) for the preliminary screening and determination of the interface residues, and by using HADDOCK [23,24] for a more refined, interacting-residues driven docking.

Statistical Analysis

Data were imported from TCGA Firebrowse APIs (http://www.firebrowse.org/api-docs/). Transcripts per Millions (TPMs) were calculated from RSEM matrixes with the formula TPM = scaled, estimate*10^6. Visualizations were built via custom R scripting (ggboxplot, ggpaired functions of gplots package). Boxplots follow standard representation, with the main box representing the Interquartile range and the middle line representing median values. TPM values of single RNAs in Figure 5 (a,c) and in Figure S7 were computed via unpaired t-tests. For the Venn Diagram in Fig. 5b, the genes whose expression was positively correlated with the expression of DLEU2 or EZH2 in

the TCGA-LIHC dataset were extracted via the R2 platform (https://r2.amc.nl) with the function "Find correlated Genes with a Single Gene", then intersecting the resulting lists via custom R scripting with the list of genes identified as HBx direct target genes by ChIP-Seq. Kaplan-Meier curves and survival fits were computed via the R2 web portal (http://r2.amc.nl), by setting the highest survival censorship to 60 months, and partitioning the dataset by each RNA median value.

For the rest of the data, the statistical analyses were performed using GraphPad Prism version 6.00 GraphPad Software, La Jolla California USA, (https://www.graphpad. com/scientific-software/prism/). The data are presented as average ± SD. P values were determined using unpaired t-test. P values ≤0.05 were considered statistically significant. Additional details can be found in the Legends to figures.

MODELING AND DOCKING

(a) RNA structure modelling

Experimental structures of lncRNAs are scarce [25] and methods to predict *de novo* RNA/lncRNA structures are subject of intense and challenging investigation [26-28], at the edge of the RNA research.

We modeled a portion of intron 1 (a selected region of 439 bases) starting from the FASTA sequence (https://genome.ucsc.edu/cgi-bin/hgGateway). We first obtained the secondary structure via DotKnot [6,7], a hybrid heuristic algorithm, based on sequence matching and free energy minimization. The algorithm is overall faster than methods based on free energy minimization only, that can become unaffordable for long sequences [29]. It performs a preliminary detection of pseudoknots, used in a second step during the secondary structure prediction. Pseudoknot structures can become important for long RNAs, in particular in relationship with transcription and regulatory activities [30] (Figure S5).

The FASTA sequence and the secondary structure were used to obtain the 3D structure of intron 1 (Figure 3) via the RNAcomposer tool [8], a method based on a machine translation principle. RNAcomposer has been tested for high resolution large RNA structure prediction [27,28] starting from a user-defined secondary structure. It allows to model up to 500 residues, therefore it is suited for the large intron considered here.

(b) HBx structure modelling

HBx is a small protein (154 aa), with subdomains devoted to different functions [31]. It is a partially disordered protein (hybrid type Intrinsically Unstructured Protein, IUP), as it contains some 10% of helices and some beta turns (in the globular part), coexisting with disordered coils. HBx disordered nature makes challenging to determine its structure via NMR or X-rays, therefore few information on the three-dimensional structure is available. It is known that the presence of nine cysteines in the amino acid sequence lead to the formation of four disulfide bonds, which role is to stabilize and to control the protein function [32]. In this work, the HBx structure has been modelled by I-TASSER [12-14], including in the modelling the constraints on cysteine linkages. Due to the partial IU nature of HBx, scores for the homology structures obtained are not high (from -4.25 to -5). We selected the

most stable structure to be used for docking (Figure S4a); however, the best five structures modelled by I-TASSER share the same overall secondary structure and the presence of the globular and the disordered subdomains.

The 3D structure obtained via I-TASSER compares well with the degree of intrinsic disorder predictions (Figure S4b) based on the primary amino acid sequence, as given by IUPred [9], and PONDR [10], that predicts disorder only in the first part of the sequence. Results from I-TASSER modelling are also in very good agreement with prediction of secondary structure elements (Figure S4c) from DISOPRED2 [11], while our model is different from a previous 3D model [33], not including constraints on cysteines.

(c) Sequence-based interaction methods

We estimated the interaction probabilities for the HBx-lncRNAs with various tools based on primary sequences analysis.

Based on the HBx protein sequence only, the RNABindRPlus [17] method identifies residues with higher interacting propensity for RNA, highlighted on the modelled HBx tertiary structure (Figure S5a). Residues with the highest interaction scores are located in the disordered region, with some exception in the loops in the globular part.

We used catRAPID [16,17], a method providing information on protein-RNA interaction, to study HBx interaction with intron 1. According to catRAPID, a large portion of intron 1 (~19000 nucleotides, including the primers) interacts with both HBx (global score of 0.98 over 1.00) and with EZH2 (global score of 1.00). To assess positive and negative controls, we investigated the interaction of MALAT1 (full and partial sequence) with HBx and EZH2. The complete MALAT1 sequence (8775 basis) is strongly interacting with EZH2 (global score of 0.95 over 1.00) in full agreement with experiments [17]. We used it as positive control for catRAPID results. HBx is found to be interacting with MALAT1 only if the full MALAT1 is considered (0.83 global score). When a portion of ~1100 bases, containing the primers used in our experiments, is investigated with catRAPID, a much lower interaction is found (0.41 over 1.00) corresponding to a non-interacting case (the threshold is global score > 0.5 for interacting systems), in agreement with our experimental findings.

(d) Docking predictions

It has been recently proven, with a high-throughput docking experiment, that different complexes have different overall signals in their docking scores [34], and that standard docking algorithms scores are able to distinguish real interactors from non-interacting cases. We therefore used here docking results to establish the existence of protein-RNA interactions and, with some caution, we relied on the score rankings from HEX [19] and HADDOCK [23,24] as relative indicators of the strength of the interaction to evaluate the propensity of proteins to bind DLEU2 intron 1. HEX uses a shape complementarity scoring function, i.e. it performs rigid docking and spatial shape fitting, providing a very efficient configurations sampling. HADDOCK includes the most refined description of the interaction, but, as it requires information on the interaction interface, we used preliminary docking results from HEX to define it. In particular, the HADDOCK docking protocol requires the determination of the list of *active* residues. We assigned them as the protein residues (intron 1 nucleotides) with at least one atom within 5 Å of intron 1 nucleotides (protein residues) in the complexes produced by HEX. These amino acids (bases) are used as putative binding sites at the protein/RNA interface. *Passive* residues, also required by HADDOCK protocol, are defined automatically around the active ones.

By docking of HBx on DLEU2 intron 1 with HEX (by including also the electrostatic contributions in the scoring function, beside the volumetric terms), interestingly we found two loci on the intron 1 with nearly equal probability to be occupied by HBx (see Figure 3a, panels 1-2). Also in the case of the interaction of the PRC2 protein EZH2 with DLEU2 intron 1, HEX calculations provide two loci on the DLEU2 intron 1 with nearly equal probability to be occupied by EZH2 (Figure 3b, panels 1-2). The most stable configurations of HBx (EZH2)-DLEU2 intron 1 complex have been further refined with HADDOCK, confirming the presence of two distinct configurations with very similar scores, corresponding to HBx (EZH2) bound to one or the other loci on RNA already found via HEX.

HEX calculations have been carried out to investigate the ternary EZH2 - DLEU2 intron 1 - HBX complexes, and HADDOCK results confirm the most stable configurations of the proteins - DLEU2 intron 1 triplet complexes (Figure 3c, panels 1-8).

For studying the triplet complexes, we used at first the DLEU2 intron 1 - EZH2 binary complex as reference term. We focus on the two identified loci on DLEU2 intron 1, close to those potentially targeted by HBx, displaying a nearly equal probability to be occupied by EZH2. The most stable ternary complexes obtained are shown in Figure 3c (panels 5-8). HBx can occupy either one or the other of the two most probable EZH2 loci on the DLEU2 intron 1 (see Figure 3c, panels 5 and 8) or even the same locus on DLEU2 intron 1 (see Figure 3c, panels 6 and 7). The opposite mechanism was also investigated, i.e. we used as docking targets for EZH2 the two binary HBx- DLEU2 intron 1 complexes shown in Figure 3a, obtaining the ternary complexes shown in Figure 3c, panels 1-4. Also in this case, we could find configurations in which the two proteins (HBx and EZH2) bind the intron 1 of DLEU2 in the same region (Figure 3c, panels 1 and 4) or in which EZH2 occupies one or the other HBx loci on the DLEU2 intron 1 (Figure 3c, panels 2 and 3). We verified that a similar mechanism holds in the case of exon 6 (unpublished), where HBx and EZH2 share the same preferential binding site. If the exon 6 locus is occupied by one of the two proteins, the other one arranges to bind close by.

Our *in silico* model cannot take into account conformational changes induced by HBx binding on the lncRNA or on the protein structure (even though refined docking methods performs some atomic local relaxation), and therefore we did not investigate the possible effects of HBx on the lncRNA ability to bind other proteins upon structural modifications induced by HBx interaction.

e) Protein-protein and protein-lncRNAs control docking

We assessed the quality of both our HBx model and our docking protocol by using some experimental information on other HBx complexes and RNA structures.

X-rays data on DDB1-HBx short stretch complex [35] (PDB entry: 3I7H) show that HBx interacts with DDB1 via a short helix structure (residues 88-100). We docked (via HADDOCK) our HBx model and DDB1; the resulting score value is slightly higher (by 5%) than in the HBx- DLEU2 intron 1 case. We verified that i) the interaction is driven by the same short helix (aa 88-100), and ii) the short HBx helix sits at the same DDB1 site, as experimentally observed. Further docking results for our HBx model and protein HBXIP (PDB entry: 3MS6) identify the interface aa 137-140, as proposed in [33].

We also relied on some control dockings performed using other RNAs, looking for a negative control

case, by choosing a small portion of MALAT1, a lncRNA that, according to our experimental data, has no interaction with HBx. In particular, docking HBx to the MALAT1 triple helix (4PLX.pdb), returns an interaction score lower by the 25% with respect to HBx- DLEU2 intron 1. This can be used as negative control for HBx-RNA interactions.

On the other hand, given the strong interaction of HBx and DDB1, one could conceive that DDB1 alone, or even its complex with HBx, would interact with the lncRNA considered in this work. Results from our docking protocols show that neither DDB1 nor the HBX-DDB1 complex significantly interact with the lncRNA, providing a second case of negative control for protein-lncRNA interaction. In other terms, our docking procedure is able to discriminate among real or just supposed protein-RNA interactors.

Legends to Supplementary Figures

Figure S1. DLEU2 expression in hepatocellular carcinoma (a) lncRNAs expression from RNA-Seq data [36], tumor and adjacent normal tissues from 60 patients. T= Tumor. ANT= Adjacent Normal Tissues. **(b)** Expression of selected lncRNAs in hepatocellular carcinoma (HCC) (microarray dataset in 81 patients) [37].

Figure S2. HBx modulates DLEU2 expression in HepG2, NTCP-HepG2 and HepAD38 cell lines. (a) HBx occupancy and H4 histone acetylation on the promoter region of DLEU2. Crosslinked chromatin from mock, wild-type HBV (HBV wt) or HBV HBx mutant (HBV mt HBx) replicating HepG2 and induced HepAD38 cells was immuno-precipitated with specific anti-HBx, anti-AcH4 antibodies or relevant IgG controls, and then analyzed by real-time qPCR using specific primer pairs. The detection of HBV cccDNA using specific primers in the ChIPed DNA from HBV wt-infected or HBV wt-replicating cells (cccDNA-ChIP) served as a technical positive control for the ChIP procedure. ChIP results are expressed as fold induction (FI) of the % of input with respect to mock. **(b)** Real-time qPCR of DLEU2 mRNA in mock, HBV wt-infected HepG2-NTCP cells (7 days) and HepG2 (48 h) and HepAD38 (48 h) HBV-replicating cells. Results are expressed as values relative to endogenous human GAPDH mRNAs. **(c)** PCR detection (*upper panels*) and relative densitometry analysis (*lower panel*s) of exons 2-4 and β-Actin in mock and HBV wt-infected HepG2-NTCP cells (7 days), HepG2 (48 h) and HepAD38 (48 h) HBV-replicating cells. Results are expressed as FI relative to the mock after normalization to endogenous human β-Actin mRNAs. **(d)** PCR detection (*upper panels*) and relative densitometry analysis (*lower panel*) of the DLEU2 exons 2-4, intron 1 and β-Actin in the nuclear fraction (N) of mock and HBV wt-replicating HepG2 cells. Results are expressed as in S2c. **(e)** Real-time qPCR of DLEU2 exons 9-10 (*lower left panel*), PCR detection (*lower right panel*) and the relative densitometry analysis (*upper right panels)* of intron 1 in mock and in the nuclear fraction of HBx-HA transfected HepG2 cells. Results are expressed as in S2c. **(f)** Real-time qPCR of DLEU1 mRNA in mock and HBV wt-replicating HepG2 cells. Results are expressed as in S2b. In panels (a) to (f) * = p-value < 0.05; Mann-Whitney test.

Figure S3. HBx binding to DLEU2. (a) Lysates from HBV wt-replicating HepG2 cells were subjected to RIP using IgG control or anti-HBx antibody. DLEU2 exons 2-4, intron 1, β-Actin and MALAT1 were detected by PCR and analyzed by relative densitometry. **(b)** Detection of HBx recombinant protein (17 kDa) by Western blot with the anti-HBx antibody (ThermoFisher antibody Cat#MA1-081). The 10 KDa faster migrating band represents a degradation product of the recombinant protein. **(c)** Immunoprecipitation (IP) of HBx with the anti-HBx antibody from ThermoFisher (Cat#MA1-081) and detection by immunoblotting using the same anti-HBx antibody in HBV wt-replicating HepG2 cells (*right*) and by the anti-HBx from Abcam (cat#Ab39716) in HBV-infected HepG2-NTCP cells (*left)*. **(d)** Immunoblot analysis of RNA pull-down eluates showing the interaction between DLEU2 RNA (100 pmol) and HBx (0.5 μg). In (a) * = p-value < 0.05, Mann-Whitney test

Figure S4. Modeling HBx protein. (a) HBx 3D structure, with color highlighted secondary structure. Coil regions (mainly in the first half of the protein sequence) are in white and cyan, alpha helices in violet, beta-sheets in yellow. The homology modeling within I-TASSER has been performed by using default parameter values. Constraints on four disulfide bonds [30] among pairs of cysteines (C7-C69, C17-C143, C61-C115, C78 -C137) have been used. The 3D structure is different from a previous model not including constraints on cysteines [31]. Residues predicted to interact with RNA, according to RNABindRPlus, are explicitly shown, colored in orange (interacting score P > 0.9 over a maximum score of 1.0), green (interacting score 0.9 > P > 0.7) and blue (interacting score 0.7 > P > 0.5. Our 3D model is in good agreement with both the **(b)** IUPRED results that predict the most disordered regions in the initial part of the sequence, and the **(c)** DISOPRED results on the secondary structure.

Figure S5. DLEU2 RNA secondary structure. Secondary structure of DLEU2 intron 1 portion (439 nucleotides) is predicted with DotKnot and plotted with Pseudowiever. Pseudoknots are highlighted in bright yellow. Parts interacting with HBx (as obtained from docked HBx-DLEU2 intron 1 complex)

are highlighted, in particular the two adsorption loci on DLEU2 intron 1 are identified as light red and bright red zones.

Figure S6. DLEU2 depletion by specific locked nucleic acid (LNA) longRNA Gapmers in HBV-replicating hepatic cell lines. (a) Real-time PCR of DLEU2 RNA levels in mock and HBV-infected PHH (4 days) (*left panel*) or HBV wt-replicating HepG2 cells (*middle panel*) or NTCP-HepG2 cells (4 days) (*right panel*) in the presence of scrambled Gapmers (CTL) or DLEU2 Gapmers pools. Results are expressed as fold induction (FI) relative to the mock after normalization to endogenous human GAPDH mRNAs. **(b)** Real time PCR of HBV pgRNA levels in HBV wt-eplicating HepG2 cells (48h) and in HBV-infected HepG2-NTCP cells (4 days) in the presence of scrambled Gapmer (CTL) or DLEU2-specific Gapmer pools. Results are expressed as in S6a. Data in panels (a) and (b) represent means ± SD from at least three independent experiments performed in duplicate. In (a-b) * = p-value < 0.05, Mann-Whitney test

Figure S7. (a) TRIM13, CCNB2, DNMT1, PRC1, POLE2 and ZBTB34 mRNA levels (transcriptomic data from the TCGA-LIHC,dataset) in non tumor (NT) and in tumor liver tissues (T) from 371 HCC patients. **(b)** HCC patients survival according to TRIM13, CCNB2, DNMT1, PRC1, POLE2 and ZBTB34 expression levels. Kaplan-Meier curves and survival fits were computed via the R2 web portal (http://r2.amc.nl) by setting the highest survival censorship to 60 months and partitioning the dataset by each RNA median value.

Figure S8. DLEU2 and HBx cooperate in the transcriptional regulation of the TRIM13 host gene. (a) Real time PCR of TRIM13 mRNA in mock, HBV-infected HepG2-NTCP cells (7 days); HBV wt-replicating HepG2 cells (48 h) and non-induced (Tet+) or induced (Tet-) HepAD38 cells (48 h). Results are expressed as values relative to endogenous human GAPDH mRNAs. **(b)** H4 acetylation at the TRIM13 promoter region. Cross-linked chromatin from mock, HBV wt- or HBV mt HBx-transfected HepG2 cells and induced HepAD38 cells (48 h) was immunoprecipitated with an anti-AcH4 antibody or relevant control IgG and then analyzed by real-time qPCR using specific

primer pairs. ChIP results are expressed as fold induction (FI) of the % of input with respect to mock and analyzed as in Figure 4c. **(c)** Real Time PCR of TRIM13 mRNA in HBV wt-infected HepG2-NTCP cells (4 days) in the presence or absence of scrambled Gapmer (CTL) and DLEU2 Gapmers. Results are expressed as values relative to endogenous human GAPDH mRNAs. **(d)** Real time PCR of TRIM13 mRNA in mock, HBV wt or HBV mt HBx transfected HepG2 cells (48 h) in the presence or absence of scrambled Gapmers (CTL) and DLEU2 Gapmers. Results are expressed as fold induction (FI) relative to the mock after normalization to endogenous human GAPDH mRNAs. **(e)** H4 acetylation at the TRIM13 promoter in HBV-replicating HepG2 cells (HBV) transfected with scrambled Gapmers (CTL) or DLEU2 Gapmers. ChIP results are expressed as fold induction (FI) of the % of input relative to mock. **(f)** EZH2 occupancy (left) and H3meK27 mark (*right)* on TRIM13 promoter by ChIP assay in HBV wt or HBV mt HBx-replicating HepG2 cells (48 h). ChIP results are expressed as % of input and as FI of the % of input with respect to mock. Data in panels (a) to (f) represent means ± SD from at least three independent experiments performed in duplicate. In (a) to (f) * = p-value < 0.05; Mann-Whitney test.

Figure S9. EZH2 and SUZ12 protein levels in HBV infected cells. Western blot analysis (*left*) with the indicated antibodies and relative densitometry analysis (*right*) of total protein extracts from **(a)** HBV-infected PHHs (7 days) and from **(b)** HBV-infected HepG2-NTCP (7 days). Histograms in panels (a) and (b) represent means ± SD from at least three independent experiments.

Figure S1

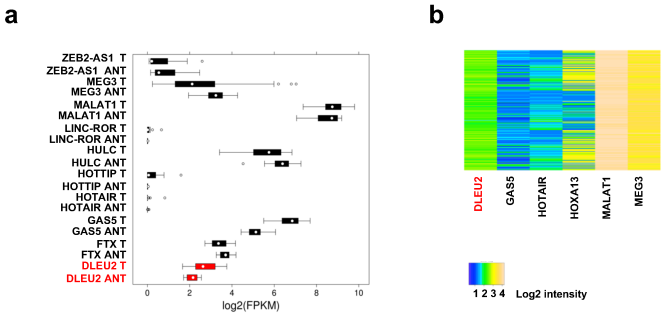

a
ZEB2-AS1 T
ZEB2-AS1 ANT
MEG3 T
MEG3 ANT
MALAT1 T
MALAT1 ANT
LINC-ROR T
LINC-ROR ANT
HULC T
HULC ANT
HOTTIP T
HOTTIP ANT
HOTAIR T
HOTAIR ANT
GAS5 T
GAS5 ANT
FTX T
FTX ANT
DLEU2 T
DLEU2 ANT
0
2
4
6
8
10
log2(FPKM)
b
DLEU2
GAS5
HOTAIR
HOXA13
MALAT1
MEG3
1 2 3 4 Log2 intensity

Figure S2

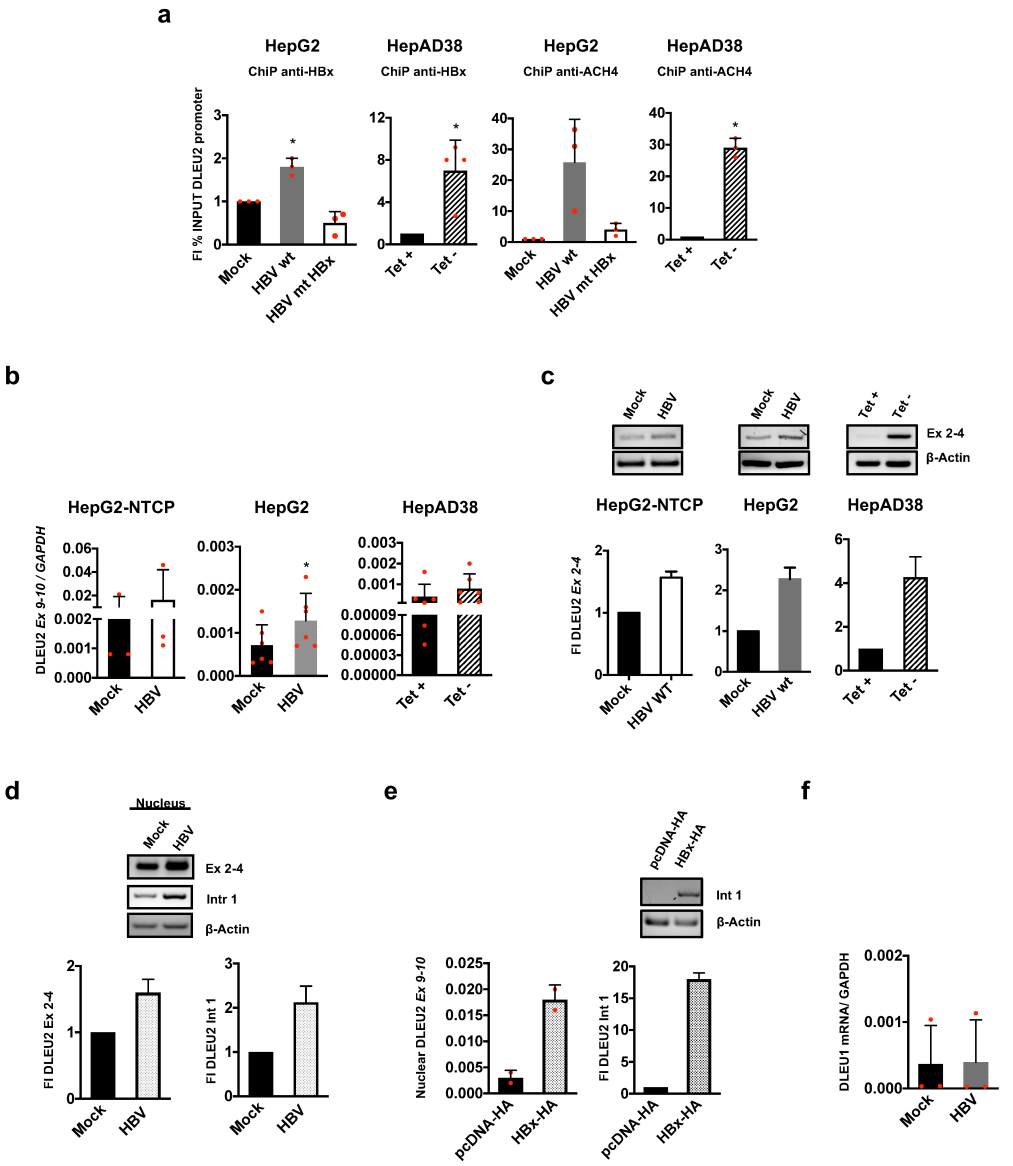

a
HepG2
ChiP anti-HBx
HepAD38
ChiP anti-HBx
HepG2
ChiP anti-ACH4
HepAD38
ChiP anti-ACH4
Fi % INPUT DLEU2 promoter
Mock
HBV wt
HBV mt HBx
Tet +
Tet -
b
HepG2-NTCP
HepG2
HepAD38
DLEU2 Ex 9-10 / GAPDH
Mock
HBV
Tet +
Tet -
c
Mock
HBV
Tet +
Tet -
Ex 2-4
β-Actin
HepG2-NTCP
HepG2
HepAD38
FI DLEU2 Ex 2-4
Mock
HBV WT
HBV wt
Tet +
Tet -
d
Nucleus
Mock
HBV
Ex 2-4
Intr 1
β-Actin
FI DLEU2 Ex 2-4
FI DLEU2 Int 1
Mock
HBV
e
pcDNA-HA
HBx-HA
Int 1
β-Actin
Nuclear DLEU2 Ex 9-10
FI DLEU2 Int 1
pcDNA-HA
HBx-HA
f
DLEU1 mRNA/ GAPDH
Mock
HBV

Figure S3

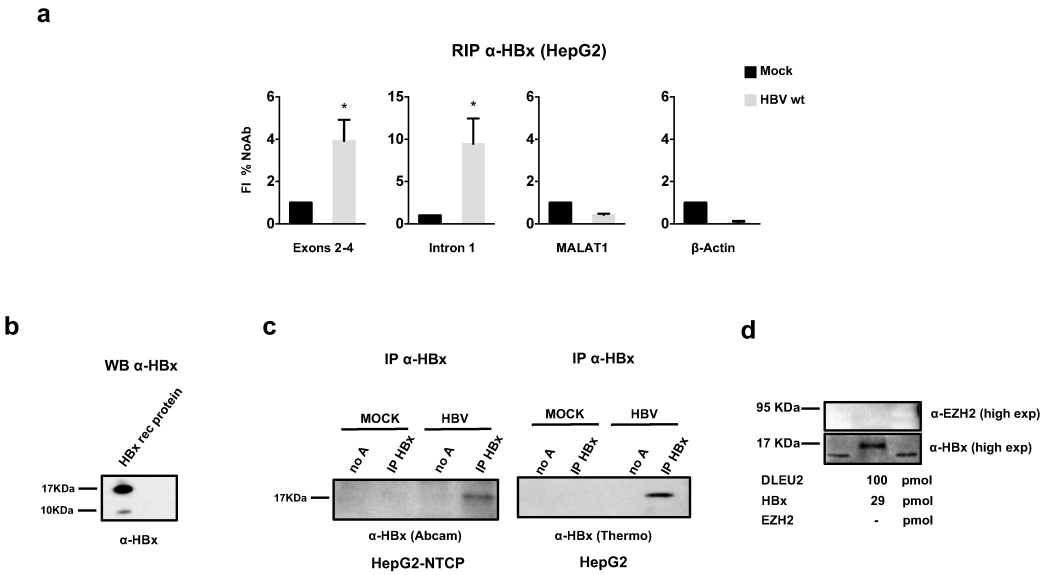

a
RIP α-HBx (HepG2)
Mock
HBV wt
FI % NoAb
Exons 2-4
Intron 1
MALAT1
β-Actin
b
WB α-HBx
HBx rec protein
17KDa
10KDa
α-HBx
c
IP α-HBx
IP α-HBx
MOCK
HBV
MOCK
HBV
no A
IP HBx
no A
IP HBx
no A
IP HBx
no A
IP HBx
17KDa
α-HBx (Abcam)
α-HBx (Thermo)
HepG2-NTCP
HepG2
d
95 KDa
17 KDa
α-EZH2 (high exp)
α-HBx (high exp)
DLEU2 100 pmol
HBx 29 pmol
EZH2 - pmol

Figure S4

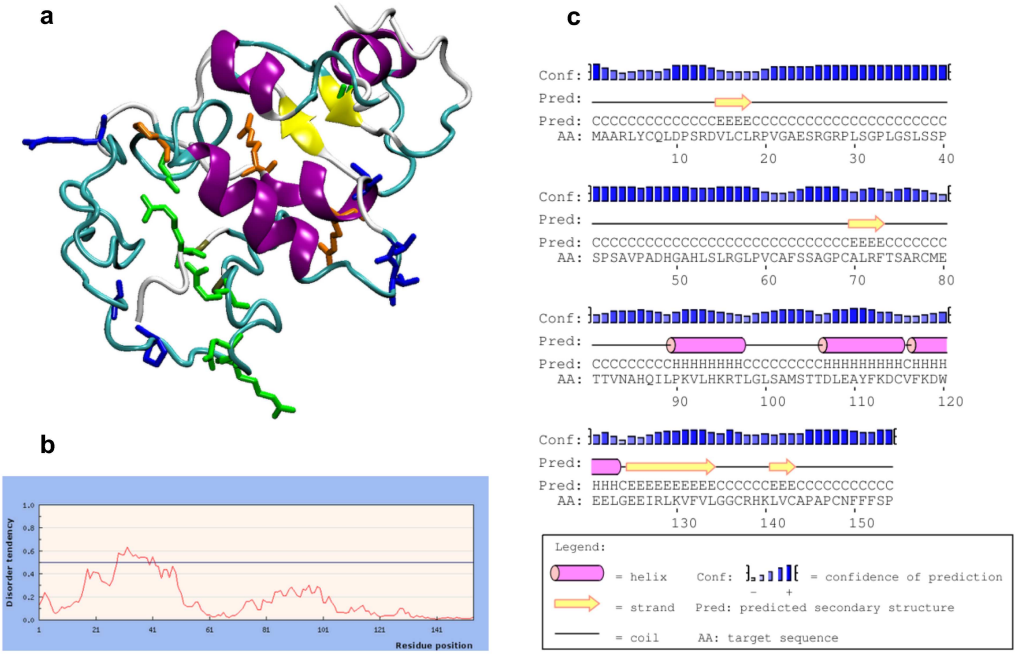

a
b
c
Disorder tendency
Residue position
Conf:
Pred:
AA:
Legend:
= helix
= strand
= coil
= confidence of prediction
Pred: predicted secondary structure
AA: target sequence

Figure S5

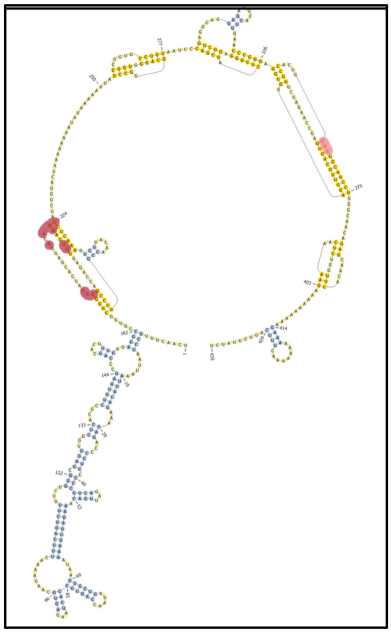
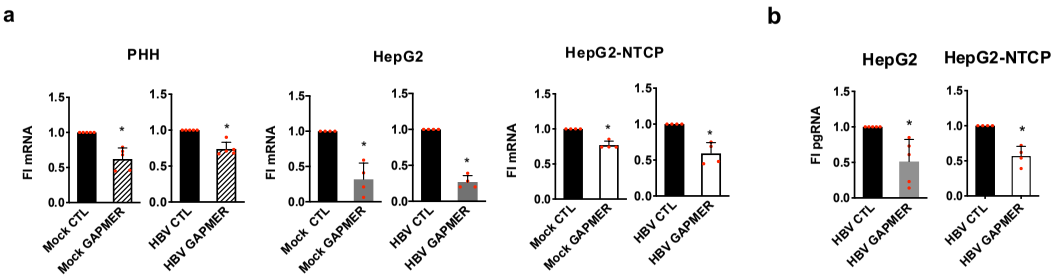


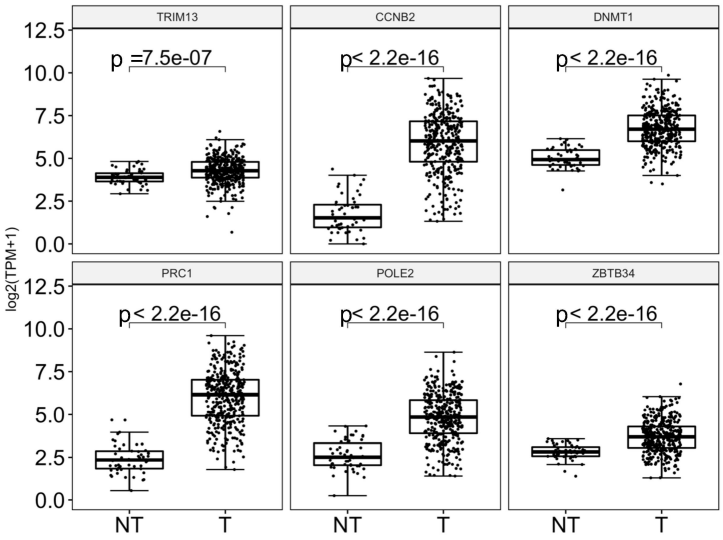
Figure S6



Salerno D, *et al*. *Gut* 2020; 69:2016–2024. doi: 10.1136/gutjnl-2019-319637

Figure S7

a


TRIM13
CCNB2
DNMT1
PRC1
POLE2
ZBTB34
p =7.5e-07
p< 2.2e-16
log2(TPM+1)
NT
T

b

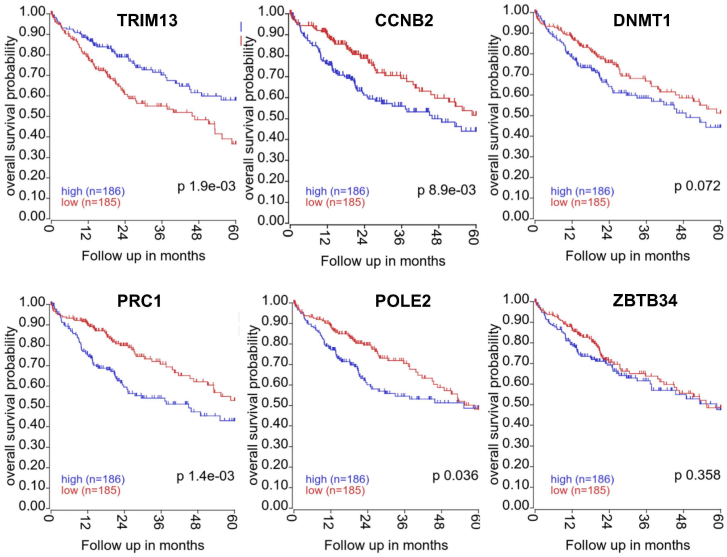

TRIM13
CCNB2
DNMT1
PRC1
POLE2
ZBTB34
overall survival probability
high (n=186)
low (n=185)
p 1.9e-03
p 8.9e-03
p 0.072
p 1.4e-03
p 0.036
p 0.358
Follow up in months

Figure S8

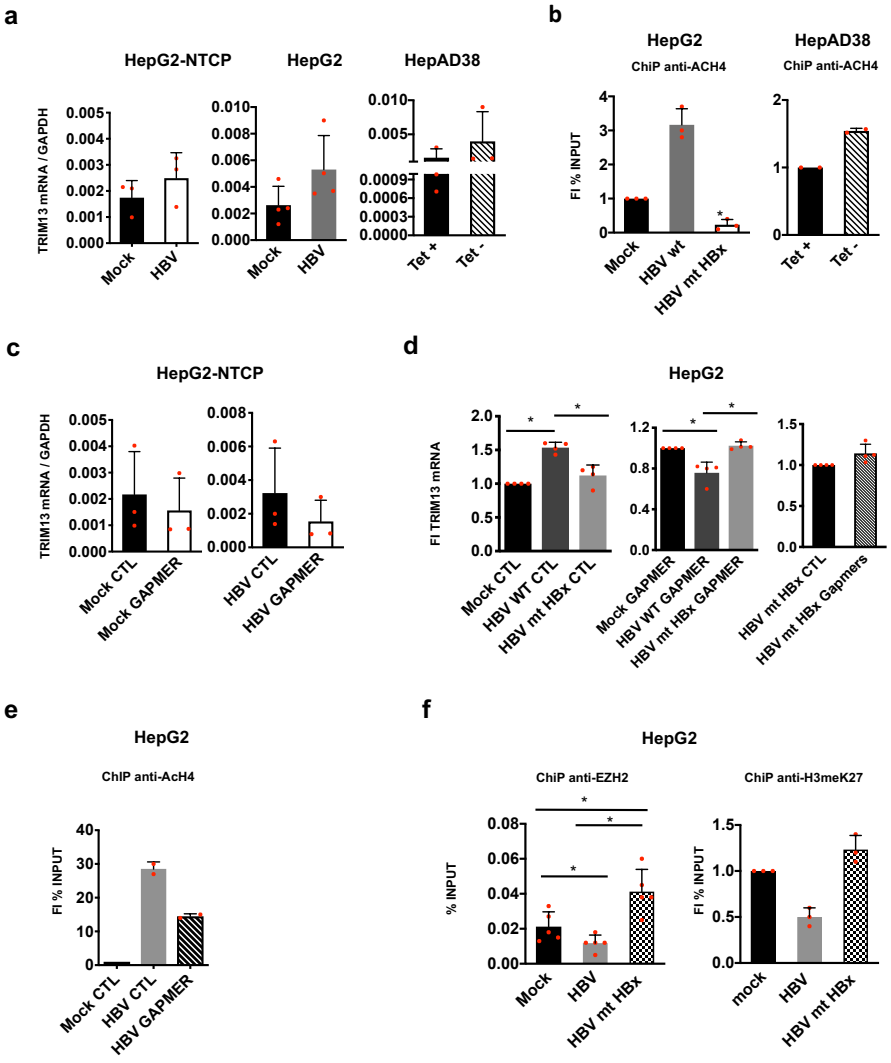

a
HepG2-NTCP
HepG2
HepAD38
TRIM13 mRNA / GAPDH
Mock
HBV
Tet +
Tet -
b
HepG2
ChiP anti-ACH4
HepAD38
ChiP anti-ACH4
FI % INPUT
Mock
HBV wt
HBV mt HBx
Tet +
Tet -
c
HepG2-NTCP
TRIM13 mRNA / GAPDH
Mock CTL
Mock GAPMER
HBV CTL
HBV GAPMER
d
HepG2
FI TRIM13 mRNA
Mock CTL
HBV WT CTL
HBV mt HBx CTL
Mock GAPMER
HBV WT GAPMER
HBV mt HBx GAPMER
HBV mt HBx CTL
HBV mt HBx Gapmers
e
HepG2
ChIP anti-AcH4
FI % INPUT
Mock CTL
HBV CTL
HBV GAPMER
f
HepG2
ChiP anti-EZH2
ChiP anti-H3meK27
% INPUT
FI % INPUT
Mock
HBV
HBV mt HBx
mock
HBV
HBV mt HBx

Figure S9

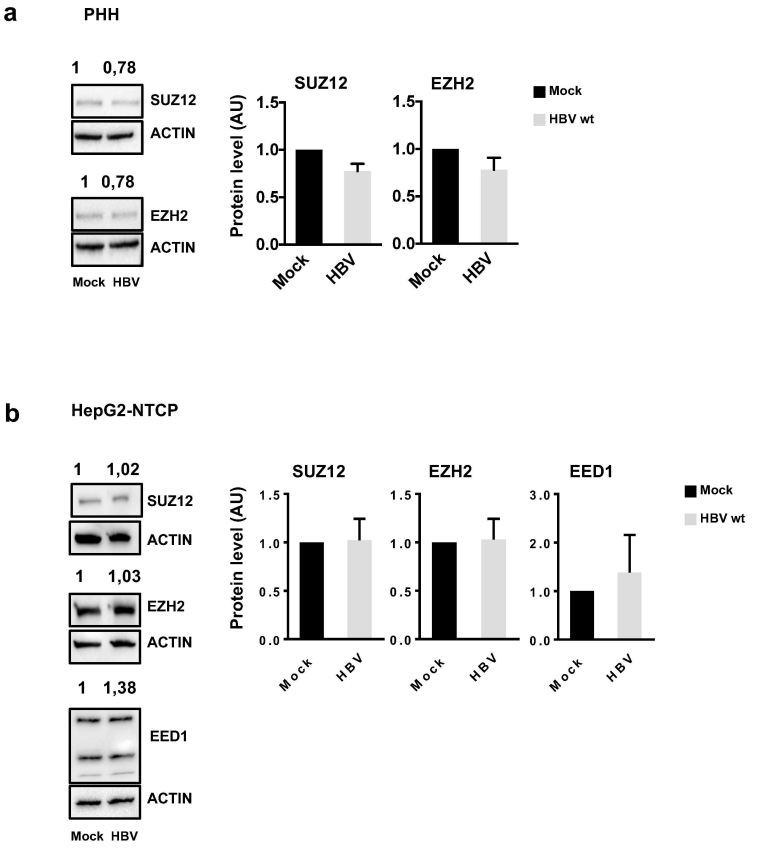

a
PHH
1 0,78
SUZ12
ACTIN
1 0,78
EZH2
ACTIN
Mock HBV
SUZ12
EZH2
Protein level (AU)
1.5
1.0
0.5
0.0
Mock
HBV
Mock
HBV wt
b
HepG2-NTCP
1 1,02
SUZ12
ACTIN
1 1,03
EZH2
ACTIN
1 1,38
EED1
ACTIN
Mock HBV
SUZ12
EZH2
EED1
Protein level (AU)
3.0
2.0
1.0
Mock
HBV wt

Table S1

LncRNA	HBx peak position
1. C1orf126	Promoter
2. CASC2	Promoter
3. LINC00271	Promoter/ In Gene
4. LINC00277	Promoter
5. LINC00299	Promoter/ In Gene
6. LINC00441	Promoter
7. LINC00478	Promoter/ In Gene
8. LINC00486	Promoter/ In Gene
9. LINC00487	Promoter
10. LINC00521	Promoter
11. LINC00526	Promoter/ In Gene
12. MIR4500HG	Promoter
13. PLK1S1	Promoter/ In Gene
14. RBM26-AS1	Promoter
15. DLEU2	Promoter/ In Gene
16. CHODL-AS1	In Gene
17. HCG18	In Gene
18. LINC00303	In Gene
19. LINC00305	In Gene
20. LINC00330	In Gene
21. LINC00340	In Gene
22. LINC00467	In Gene
23. LINC00473	In Gene
24. LINC00476	In Gene
25. LINC00511	In Gene
26. LINC00535	In Gene
27. LINC00536	In Gene
28. MIR210HG	In Gene
29. MIR31HG	In Gene
30. PVT1	In Gene
31. SKINTL	In Gene
32. SNHG12	In Gene
33. SOX2-OT	In Gene
34. TMEM72-AS1	In Gene

Table S2

Probeset	HUGO	R.value	R.pvalue	Hbx peak
ABCC10_89845	ABCC10	0.455	0.0012271767	
ABHD3_171586	ABHD3	0.408	0.0068224833	
ACIN1_22985	ACIN1	0.45	0.0015427303	promoter
ACVR2B_93	ACVR2B	0.418	0.0049164901	
AFP_174	AFP	0.502	0.0001550709	
AGBL2_79841	AGBL2	0.436	0.0026407528	
ALS2_57679	ALS2	0.4	0.0089852837	in gene
ANGEL1_23357	ANGEL1	0.401	0.0086182898	
ANKAR_150709	ANKAR	0.422	0.0043772188	in gene
ANKRD13A_88455	ANKRD13A	0.474	0.0005596312	in gene
ANKRD36_375248	ANKRD36	0.397	0.009773233	
ARHGEF1_9138	ARHGEF1	0.429	0.0034266939	in gene
ARID3A_1820	ARID3A	0.479	0.0004490759	
ARID3B_10620	ARID3B	0.486	0.0003308236	in gene
ARL6IP6_151188	ARL6IP6	0.417	0.0051661087	
BLM_641	BLM	0.512	9,46E+09	
BRD1_23774	BRD1	0.407	0.0071231974	
C12orf76_400073	C12orf76	0.587	1,12E+08	in gene
C13orf23_80209	PROSER1	0.543	1,71E+09	
C13orf34_79866	BORA	0.534	2,91E+08	
C14orf101_54916	TMEM260	0.546	1,46E+09	
C14orf139_79686	LINC00341	0.397	0.0097835723	
C15orf23_90417	KNSTRN	0.422	0.0042686153	
C15orf29_79768	KATNBL1	0.439	0.0023084816	
C15orf42_90381	TICRR	0.457	0.0011670061	
C15orf56_644809	C15orf56	0.428	0.0034501421	
C21orf45_54069	MIS18A	0.501	0.0001608295	
C22orf39_128977	C22orf39	0.524	5,03E+09	in gene

C2orf15_150590	C2orf15	0.533	3,02E+09	promoter
C2orf54_79919	C2orf54	0.41	0.0065608719	
C2orf60_129450	TYW5	0.484	0.0003656675	
C5orf13_9315	NREP	0.453	0.0013541265	
C5orf34_375444	C5orf34	0.424	0.0040856975	
C6orf146_222826	FAM217A	0.402	0.0083679336	
C8ORFK29_340393	TMEM249	0.422	0.0042791582	
C9orf45_81571	MIR600HG	0.443	0.0019990112	
CASP3_836	CASP3	0.473	0.0005892718	
CBFA2T2_9139	CBFA2T2	0.463	0.0009167023	promoter
CCDC150_284992	CCDC150	0.528	3,95E+09	
CCNB2_9133	CCNB2	0.408	0.0070015017	promoter
CCNL1_57018	CCNL1	0.417	0.0050950341	promoter
CCNYL1_151195	CCNYL1	0.486	0.0003343846	
CDC42BPG_55561	CDC42BPG	0.442	0.0020909356	in gene
CDCA4_55038	CDCA4	0.424	0.0039743006	
CG030_116828	N4BP2L2-IT2	0.454	0.0012994755	
CHEK2_11200	CHEK2	0.451	0.0014816451	
CHST13_166012	CHST13	0.432	0.0030254254	
CKAP2_26586	CKAP2	0.406	0.0073251417	
CLDN1_9076	CLDN1	0.475	0.0005415688	in gene
CMYA5_202333	CMYA5	0.437	0.0025517349	in gene
COMMD2_51122	COMMD2	0.464	0.0008766404	
CPSF7_79869	CPSF7	0.446	0.0017876323	
CREB1_1385	CREB1	0.409	0.0066069983	in gene
CROT_54677	CROT	0.47	0.0006676271	
CSNK1G1_53944	CSNK1G1	0.417	0.0051250316	in gene
CTDSPL2_51496	CTDSPL2	0.41	0.0064647978	
CYP20A1_57404	CYP20A1	0.434	0.002799887	in gene
CYP4F22_126410	CYP4F22	0.437	0.0025133981	promoter
DBF4_10926	DBF4	0.455	0.001241255	

DDX12_440081	DDX12P	0.448	0.0016463493	
DGCR8_54487	DGCR8	0.401	0.0086304395	
DICER1_23405	DICER1	0.464	0.0008566198	
DIP2A_23181	DIP2A	0.423	0.004122365	in gene
DLEU2_8847	DLEU2	1	0	
DNA2_1763	DNA2	0.425	0.0038439883	
DNAH8_1769	DNAH8	0.465	0.0008440995	in gene
DNAI1_27019	DNAI1	0.4	0.0088225478	
DNAJC9_23234	DNAJC9	0.429	0.0033236753	
DNM1P35_100128285	DNM1P35	0.398	0.0093192565	
DNMT1_1786	DNMT1	0.396	0.0098359714	promoter
DNMT3B_1789	DNMT3B	0.459	0.0010458062	
DOPEY1_23033	DOPEY1	0.441	0.0021524537	promoter
DOT1L_84444	DOT1L	0.403	0.0080884726	in gene
DQX1_165545	DQX1	0.425	0.0039067797	
DSCAML1_57453	DSCAML1	0.403	0.0081457314	in gene
ELF1_1997	ELF1	0.447	0.001712932	in gene
ENPP5_59084	ENPP5	0.397	0.009763942	
ERCC6_2074	ERCC6	0.464	0.0008779451	
ESCO1_114799	ESCO1	0.397	0.0097832822	
EXOSC8_11340	EXOSC8	0.448	0.0016683451	in gene
FAM118A_55007	FAM118A	0.404	0.0077435074	
FAM119A_151194	METTL21A	0.478	0.0004748156	in gene
FAM48A_55578	SUPT20H	0.506	0.0001283992	
FANCD2_2177	FANCD2	0.441	0.002141698	
FBXO11_80204	FBXO11	0.432	0.0029851349	
FBXO5_26271	FBXO5	0.412	0.006109282	
FER1L5_90342	FER1L5	0.47	0.0006916937	
FNBP4_23360	FNBP4	0.466	0.0008028974	
FZD5_7855	FZD5	0.445	0.0018925188	
GALNT3_2591	GALNT3	0.462	0.0009519416	

GCFC1_94104	PAXBP1	0.483	0.0003808331	
GEN1_348654	gen-01	0.455	0.0012543617	
GJB6_10804	GJB6	0.442	0.0020489573	promoter
GLDC_2731	GLDC	0.396	0.0099818199	promoter
GOLGA6L5_374650	GOLGA6L5P	0.416	0.0053922015	
GOLGA6L9_440295	GOLGA6L9	0.408	0.0068256581	
GPC5_2262	GPC5	0.516	7,57E+09	in gene
GPR137C_283554	GPR137C	0.446	0.0018081281	
GPR143_4935	GPR143	0.399	0.0092774872	in gene
GPRC5D_55507	GPRC5D	0.407	0.0070697978	
GRHL1_29841	GRHL1	0.414	0.0056943325	promoter
GTSE1_51512	GTSE1	0.411	0.0062364712	
H19_283120	H19	0.398	0.0093272043	
HAUS2_55142	HAUS2	0.442	0.002066981	
HELLS_3070	HELLS	0.431	0.0031836956	
HIC2_23119	HIC2	0.559	6,48E+08	
HNRNPA3P1_10151	HNRNPA3P1	0.445	0.0018950979	
HNRNPA3_220988	HNRNPA3	0.511	9,74E+08	
IGSF1_3547	IGSF1	0.468	0.0007487614	
IL1RAP_3556	IL1RAP	0.463	0.0009106159	in gene
ILF3_3609	ILF3	0.463	0.0009139349	
INTS4_92105	INTS4	0.398	0.0094034768	
INTS6_26512	INTS6	0.419	0.0048687667	
JUB_84962	AJUBA	0.414	0.0056923278	
KBTBD6_89890	KBTBD6	0.465	0.0008257055	
KCNJ3_3760	KCNJ3	0.443	0.0019946214	in gene
KIAA0528_9847	C2CD5	0.399	0.0090697882	
KIAA1704_55425	GPALPP1	0.466	0.0008012055	
KIF20B_9585	KIF20B	0.423	0.0041784896	
KIF6_221458	KIF6	0.489	0.000282736	in gene
KIFC1_3833	KIFC1	0.411	0.006258148	

KLRAQ1_129285	PPP1R21	0.536	2,51E+09	
LCTL_197021	LCTL	0.409	0.0066503591	
LGR4_55366	LGR4	0.397	0.0095742902	in gene
LIG1_3978	LIG1	0.443	0.0019760096	in gene
LMNB1_4001	LMNB1	0.401	0.0086278164	in gene
LOC284100_284100	YWHAEP7	0.415	0.0055508857	
LOC641367_641367	LOC641367	0.421	0.004433561	
LOC642846_642846	LOC642846	0.418	0.0049021408	
LOC647121_647121	EMBP1	0.49	0.0002685365	
LRRC66_339977	LRRC66	0.422	0.0043779091	
LSR_51599	LSR	0.443	0.0020398516	promoter
MAL2_114569	MAL2	0.441	0.0022045714	
MAPK6_5597	MAPK6	0.398	0.0094050714	
MBTD1_54799	MBTD1	0.441	0.0021892319	
MCCD1_401250	MCCD1	0.402	0.0083286665	
MED4_29079	MED4	0.481	0.000407935	
MEP1A_4224	MEP1A	0.47	0.0006672635	
MEST_4232	MEST	0.517	7,20E+09	promoter
MIR17HG_407975	MIR17HG	0.424	0.0040664072	
MITD1_129531	MITD1	0.433	0.0029103399	promoter
MOSC1_64757	MOSC1	0.428	0.0034476678	
MTRF1_9617	MTRF1	0.534	2,77E+08	In gene
MXD1_4084	MXD1	0.429	0.0034303004	
MYO5C_55930	MYO5C	0.433	0.0029686082	
NAA16_79612	NAA16	0.646	1,45E+06	In gene
NAALAD2_10003	NAALAD2	0.397	0.0096692249	In gene
NAALADL1_10004	NAALADL1	0.399	0.0091477041	
NARG2_79664	ICE2	0.426	0.0037588748	
NCAPD2_9918	NCAPD2	0.4	0.008974431	in gene
NCRNA00183_554203	JPX	0.5	0.0001683939	
NEK8_284086	NEK8	0.423	0.0041994108	

NFYA_4800	NFYA	0.406	0.0074596907	
NKTR_4820	NKTR	0.431	0.0031921644	
NUDT15_55270	NUDT15	0.568	3,93E-03	
NUSAP1_51203	NUSAP1	0.406	0.0073554139	
OVGP1_5016	OVGP1	0.46	0.0010207866	
PAN3_255967	PAN3	0.417	0.005090931	
PAQR9_344838	PAQR9	0.431	0.0030979177	
PASK_23178	PASK	0.566	4,31E-03	
PBX2_5089	PBX2	0.404	0.0078179648	in gene
PDCD7_10081	PDCD7	0.446	0.0017903705	
PEG10_23089	PEG10	0.431	0.0031957443	
PHIP_55023	PHIP	0.411	0.0063480032	
PI4KAP1_728233	PI4KAP1	0.44	0.0022841645	
PIF1_80119	PIF1	0.416	0.0052630723	
PIK3R2_5296	PIK3R2	0.514	8,11E+09	promoter
PLAGL2_5326	PLAGL2	0.474	0.0005800511	
PM20D2_135293	PM20D2	0.42	0.0046127674	
POLE2_5427	POLE2	0.497	0.0001992556	promoter
POLQ_10721	POLQ	0.445	0.0018604781	
PRC1_9055	PRC1	0.415	0.0055869022	promoter
PRDM4_11108	PRDM4	0.403	0.0081383553	
PRKCI_5584	PRKCI	0.415	0.0055049625	
PRKRA_8575	PRKRA	0.397	0.0095865302	
PROM2_150696	PROM2	0.396	0.0098476661	
PRPF39_55015	PRPF39	0.446	0.0018157957	promoter
PSPH_5723	PSPH	0.427	0.0036385118	in gene
RAPH1_65059	RAPH1	0.497	0.0001990583	in gene
RAPSN_5913	RAPSN	0.41	0.0064447137	
RARS2_57038	RARS2	0.472	0.0006181429	
RBM9_23543	RBFOX2	0.401	0.0086841314	in gene
REP15_387849	REP15	0.438	0.0023784206	promoter

RFXAP_5994	RFXAP	0.424	0.0039702655	in gene
RGL3_57139	RGL3	0.445	0.0018595519	
RNF138_51444	RNF138	0.472	0.0006202891	
RNF219_79596	RNF219	0.487	0.0003157979	promoter
RNF39_80352	RNF39	0.458	0.0010932958	
RPIA_22934	RPIA	0.424	0.004068626	
RTKN_6242	RTKN	0.462	0.0009390725	in gene
RTTN_25914	RTTN	0.401	0.0086935479	
SCARNA12_677777	SCARNA12	0.408	0.0068274723	promoter
SEPT7P2_641977	SEPT7P2	0.468	0.0007333111	
SERPINA5_5104	SERPINA5	0.404	0.0078947955	promoter
SETDB2_83852	SETDB2	0.436	0.0026476466	in gene
SFI1_9814	SFI1	0.499	0.0001764059	promoter
SFRS18_25957	PNISR	0.4	0.0088325713	
SFRS8_6433	SFSWAP	0.4	0.0089654211	promoter
SKA3_221150	SKA3	0.42	0.0046566904	
SLC12A9_56996	SLC12A9	0.426	0.0037821212	
SLC13A2_9058	SLC13A2	0.438	0.0023811158	
SLCO4C1_353189	SLCO4C1	0.467	0.0007569245	in gene
SMARCA1_6594	SMARCA1	0.464	0.0008577814	
SMC4_10051	SMC4	0.494	0.0002277301	
SMPD4_55627	SMPD4	0.397	0.0095823554	
SNHG1_23642	SNHG1	0.406	0.0074668787	
SPATS2L_26010	SPATS2L	0.567	4,03E+08	promoter
SPOPL_339745	SPOPL	0.397	0.0097024923	
STRBP_55342	STRBP	0.411	0.0063695564	
STX16_8675	STX16	0.453	0.0013624245	
SUGT1L1_283507	SUGT1P3	0.496	0.0002081501	
SUV39H2_79723	SUV39H2	0.414	0.0057901786	
SVIP_258010	SVIP	0.458	0.0011089868	
TDG_6996	TDG	0.419	0.0048025396	in gene

TET1_80312	TET1	0.52	6,15E+09	
TIA1_7072	TIA1	0.467	0.0007739849	in gene
TLK1_9874	TLK1	0.497	0.0001929891	promoter
TMEM194B_100131211	TMEM194B	0.428	0.0034664565	
TSGA10_80705	TSGA10	0.525	4,75E-02	promoter
TTC21B_79809	TTC21B	0.458	0.0011145171	in gene
TTC3_7267	TTC3	0.414	0.0056417702	promoter
TTK_7272	TTK	0.445	0.0018703907	in gene
USP37_57695	USP37	0.44	0.0022729087	in gene
USP39_10713	USP39	0.435	0.0027285043	
USPL1_10208	USPL1	0.4	0.0088403475	
VN1R1_57191	VN1R1	0.427	0.0036413598	
VRK1_7443	VRK1	0.404	0.0077553996	
WBP11_51729	WBP11	0.409	0.006653993	promoter
WBP4_11193	WBP4	0.412	0.0062122489	promoter
WDR33_55339	WDR33	0.493	0.000234489	in gene
XG_7499	XG	0.438	0.0023825217	
XPO1_7514	XPO1	0.439	0.0023355258	
XRCC3_7517	XRCC3	0.495	0.000210528	
ZBED4_9889	ZBED4	0.43	0.0032556014	
ZBTB34_403341	ZBTB34	0.407	0.0071104348	promoter
ZBTB5_9925	ZBTB5	0.439	0.0023516224	
ZCCHC9_84240	ZCCHC9	0.399	0.0091696669	in gene
ZFP14_57677	ZFP14	0.463	0.0009152051	
ZNF100_163227	ZNF100	0.452	0.0013998744	in gene
ZNF107_51427	ZNF107	0.485	0.0003485831	promoter
ZNF10_7556	ZNF10	0.402	0.0082108049	
ZNF253_56242	ZNF253	0.411	0.0062591469	
ZNF273_10793	ZNF273	0.526	4,47E+07	
ZNF345_25850	ZNF345	0.423	0.0042284615	in gene
ZNF37B_100129482	ZNF37BP	0.473	0.0006031227	

ZNF431_170959	ZNF431	0.406	0.007331712	
ZNF445_353274	ZNF445	0.405	0.0074946793	
ZNF493_284443	ZNF493	0.398	0.0092735168	
ZNF502_91392	ZNF502	0.396	0.0098156796	
ZNF529_57711	ZNF529	0.457	0.0011560126	
ZNF567_163081	ZNF567	0.432	0.0030429442	
ZNF620_253639	ZNF620	0.467	0.0007580301	
ZNF675_171392	ZNF675	0.414	0.0056319478	
ZNF708_7562	ZNF708	0.455	0.0012716317	
ZNF70_7621	ZNF70	0.479	0.0004491961	
ZNF77_58492	ZNF77	0.432	0.0030380501	
ZNF823_55552	ZNF823	0.402	0.0082885405	
ZNF92_168374	ZNF92	0.457	0.0011492729	
ZNF93_81931	ZNF93	0.478	0.0004735888	promoter
ZWILCH_55055	ZWILCH	0.49	0.0002700487	

Table S3. Functional analysis of the genes co-regulated with DLEU2 RNA in HCC HBV-related patients (r>0,3) (TCGA, https://cancergenome.nih.gov/)

NAME	P-value	Adjusted p-value	Description
miRNA Regulation of DNA Damage Response_Homo sapiens_WP1530	0.0001429	0.01045	WikiPathways 2016
Cytosine methylation_Homo sapiens_WP3585	0.0001741	0.01045	WikiPathways 2016
DNA Damage Response_Homo sapiens_WP707	0.0002594	0.01096	WikiPathways 2016
Gene Expression_Homo sapiens_R-HSA-74160	0.000003057	0.001541	Reactome 2016
Cell Cycle_Homo sapiens_R-HSA-1640170	0.00001853	0.004668	Reactome 2016
DNA metabolic process (GO:0006259)	0.00001814	0.01870	GO Biological Process 2018
mRNA cis splicing, via spliceosome (GO:0045292)	0.00007509	0.03871	GO Biological Process 2018
G-quadruplex DNA unwinding (GO:0044806)	0.0001172	0.04028	GO Biological Process 2018

Table S4. Functional analysis of DLEU2 co-regulated HBx target genes

NAME	P-value	Adjusted p-value	Description
mRNA cis splicing, via spliceosome (GO:0045292)	1.849e-8	0.000003605	GO Biological Process 2018
mRNA 5'-splice site recognition	0.00005516	0.005378	GO Biological Process 2018
DNA Damage Response	0.005619	0.1528	WikiPathways 2016
G2/M DNA replication checkpoint	0.008224	0.2295	Reactome 2016

Table S5. Functional analysis of the genes co-regulated with EZH2 RNA in HCC HBV-related patients (r>0,3) (TCGA, https://cancergenome.nih.gov/)

NAME	P-value	Adjusted p-value	Description
DNA Replication	0.0001707	0.003672	WikiPathways 2016
Cell Cycle	0.00052488	0.00752333	KEGG
G2/M Checkpoints_Homo sapiens_R-HSA-69481	0.0000057	0.000198	Reactome 2016
Cell Cycle, Mitotic_Homo sapiens_R-HSA-69278	0.000000010	0.000001057	Reactome 2016
M Phase_Homo sapiens_R-HSA-68886	0.00001511	0.0003985	Reactome 2016
Epigenetic regulation of gene expression_Homo sapiens_R-HSA-212165	0.000003012	0.0001837	GO Biological Process 2018
positive regulation of histone methylation (GO:0031062)	0.0000563	0.001414	GO Biological Process 2018
mitotic recombination (GO:0006312)	0.0001044	0.0041421	GO Biological Process 2018

Table S6. Functional analysis of EZH2 co-regulated HBx target genes

NAME	P-value	Adjusted p-value	Description
DNA Replication	0.0000525900	0.0028	WikiPathways 2016
Cell Cycle	0.0003112	0.006536	WikiPathways 2016
G1 to S cell cycle control	0.000274914	0.006536	WikiPathways 2016
Cell Cycle, Mitotic	0.0000000000928	0.00000001096	Reactome 2016
Mitotic G2-G2/M phases_Homo sapiens_R-HSA-453274	0.00002466	0.000582	Reactome 2016
mitotic cell cycle phase transition (GO:0044772)	0.000000011	0.00000703	GO Biological Process 2018
G1/S transition of mitotic cell cycle (GO:0000082)	0.00001362	0.0010135	GO Biological Process 2018
mitotic recombination (GO:0006312)	0.0001044	0.0041421	GO Biological Process 2018

Table S7

	Entrez Gene	OMIN	Description
TRIM13	10206	605661	- Tripartite Motif Containing 13 (TRIM13) ER resident RING E3 ligase - Regulates NEMO (negative regulator of NF-κB signaling) ubiquitination[1] - Putative tumor suppressor[1] - Regulates initiation of autophagy during ER stress[2] - Negative regulator of MDA5-mediated type I IFN production[3]
CCNB2	9133	602755	- G2/Mitotic-Specific Cyclin-B2 - Negatively associated with the overall survival of patients with HCC[4,5,6]
DNMT1	1786	126375	- DNA Methyltransferase 1 - Upregulated by HBx[7] - Pharmacological targeting of DNMT1 inhibits HCC cell growth[8]
PRC1	9055	603484	- Protein Regulator of Cytokinesis 1 - High expressions associated with poorer overall survivals in HCC patients[9] - Identified as a hub gene, implicated in tumor (T) stage and histologic grade of HCCs[9] - Associated with early tumor recurrence in HBV-relate HCC patients[10]
POLE2	5427	602670	- DNA polymerase epsilon 2 accessory subunit (POLE2) - Highly expressed in lung cancer cell lines[11] - POLE mutation causes polymerase proofreading-associated polyposis and is associated to risk of colorectal cancer[12]
ZBTB34	403341	611692	- Zinc Finger and BTB Domain-Containing Protein 34 - Nuclear localization and putative transcriptional repressor[13]

1. Tomar D et al. Cell Signal 2014;12:2606-13
2. Tomar D et al. Biochim Biophys Acta 2012;1823:316-26
3. Narayan K et al. J Virol 2014;8:10748-10757
4. Zhang Q et al. Mol Med Rep 2019;19:2479-2488
5. Gao X et al. Biosci Rep 2018;38:Pii BSR20181441
6. Wu M et al. World J Surg Oncol 2019;17:77
7. Zhao Z et al. Oncol Rep 2017;37:2811-2818
8. Bárcena-Varela M et al. Hepatology 2019;69:587-603
9. Li B et al. J Cell Biochem 2019, Feb 1; doi: 10.1002/jcb.28420 [Epub ahead of print]
10. Wang SM et al. Clin Cancer Res 2011;17:6040-51
11. Li J et al. Oncol Rep. 2018;40:2477-2486
12. Chubb D et al. Nat Commun. 2016;7:11883
13. Qi J et al. Mol Cell Biochem 2006;290:159-67

TABLE S8

ChIP primers	**Forward (5'-3')**	**Reverse (5'-3')**
DLEU2 promoter	GCTGATAACCAGTGCCACTAA	CCTCTCAAAGTGCTGGGATTA
TRIM13 promoter	ACCCAAACTTCCTCAACTGG	GGAATGGCTCCTCCAGAATTTA
CCNB2 promoter	AACCCCAACACACCAGAAGA	TTGGGAAAGCAGACGTAGGA
PCR primers	**Forward (5'-3')**	**Reverse (5'-3')**
DLEU2 exons 2-4	ACCTGTAGCAGAGAACCAATT	TTCCTTGCAGTACACCTTTCA
DLEU2 intron1	CTGTAATCCCAGCACTTTGAGA	GGTTCCTGTTAGTGCAACTACT
MALAT1	AGGCGTTGTGCGTAGAGGA	GGATTTTTACCAACCACTCGC
β-Actin	GCACTCTTCCAGCCTTCC	AGGTCTTTGCGGATGTCC
CCNB2	ACCTACTGCTTCTGTCAAACCA	GCAGAGCAAGGCATCAGAAA
Real Time probes	**Probe**	
DLEU1	Hs00705554_s1 (Applied)	
DLEU2 exons	Hs00863925_m1 (Applied)	
DLEU2 intron 1	ID APXGUY2 (Applied, probe custom)	
EZH2	Hs00544833_m1 (Applied)	
TRIM13	Hs00328634_s1 (Applied)	
β-Actin	cat.no. 05532957001 (Roche)	
GAPDH	Hs02758991_g1	
GUS-b	Hs99999908_m1	
HBV primers and probes	**Forward (5'-3')**	**Reverse (5'-3')**
pgRNA	GCCTTAGAGTCTCCTGAGCA	GAGGGAGTTCTTCTTCTAGG
cccDNA	CTCCCCGTCTGTGCCTTCT	GCCCCAAAGCCACCCAAG
	Probe	
pgRNA FRET	AGTGTGGATTCGCACTCCTCCAGC-FL	
pgRNA Red640	ATAGACCACCAAATGCCCCTATCTTATCAAC	
cccDNA FRET	GTTCACGGTGGTCTCCATGCAACGT-FL	
cccDNA Red640	AGGTGAAGCGAAGTGCACACGGACC	

Supplemental Information References

1. Guerrieri F, Belloni L, D'Andrea D, *et al.* Genome-wide identification of direct HBx genomic targets. *BMC Genomics* 2017;**18**:184.
2. Lecluyse EL & Alexandre E. Isolation and culture of primary hepatocytes from resected human liver tissue. *Methods Mol Bio.* 2010;**640**:57-82.
3. Hantz O, Parent R, Durantel D, *et al.* Persistence of the hepatitis B virus covalently closed circular DNA in HepaRG human hepatocyte-like cells. *J Gen Virol* 2009;**90:**127-135.
4. Lucifora J, Arzberger S, Durantel D, *et al.* Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. *J Hepatol* 2011;**55**:996-1003.
5. Belloni L, Pollicino T, De Nicola F, et al. Nuclear HBx binds the HBV minichromosome and modifies the epigenetic regulation of cccDNA function. Proc. Natl. Acad. Sci. 2009;106:19975–19979.
6. Sperschneider J & Datta A. Dot knot: Pseudoknot prediction using the probability dot plot under a refined energy model. *Nucleic Acids Res* 2018;**38**: e103.
7. Sperschneider J, Datta A, Wise MJ. Heuristic RNA pseudoknot prediction including intramolecular kissing hairpins. *RNA* 2011;**17**:27-38.
8. Popenda M, Szachniuk M, Antczak M, *et al.* Automated 3D structure composition for large RNAs. *Nucleic Acids Res.* 2012;**40**:e112.
9. Dosztányi Z, Csizmok V, Tompa P, *et al.* IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 2005;**21**:3433-3434.
10. Xue B, Dunbrack RL, Williams RW. *et al*. PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys Acta - Proteins Proteomics* 2010;**1804**:996-1010.
11. Ward JJ, McGuffin LJ, Bryson K, *et al.* The DISOPRED server for the prediction of protein disorder. *Bioinformatics* 2004;**20**:2138-2139.
12. Yang J, Yan R, Roy A *et al.* The I-TASSER suite: Protein structure and function prediction. *Nature Methods* 2015;**12**:7–8.
13. Roy A, Kucukural A, Zhang Y. I-TASSER: A unified platform for automated protein structure and

function prediction. *Nat Protoc* 2010;**5**:725–738.

14. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 2008;**9**:40.
15. Muppirala UK, Honavar VG, Dobbs D. Predicting RNA-Protein Interactions Using Only Sequence Information. *BMC Bioinformatics* 2011;**12**:489.
16. Bellucci M, Agostini F, Masin M, *et al.* Predicting protein associations with long noncoding RNAs. *Nature Methods* 2011;**8**:444–445.
17. Cirillo D, Blanco M, Armaos A, *et al.* Quantitative predictions of protein interactions with long noncoding RNAs. *Nature Methods* 2017;**14**:5-6.
18. Walia RR, Xue LC, Wilkins K *et al.* RNABindRPlus: A predictor that combines machine learning and sequence homology-based methods to improve the reliability of predicted RNA-binding residues in proteins. *PLoS One* 2014;**9**:e97725.
19. Ritchie DW, Kozakov D & Vajda S. Accelerating and focusing protein-protein docking correlations using multi-dimensional rotational FFT generating functions. *Bioinformatics* 2008;**24:**1865-1873.
20. Puton T, Kozlowski L, Tuszynska I, *et al.* Computational methods for prediction of protein-RNA interactions. *J. Struct. Biol.* 2012;**179**:261-268.
21. Tuszynska I, Matelska D, Magnus M, *et al.* Computational modeling of protein-RNA complex structures. *Methods* 2014;**65**:310-319.
22. Tuszynska I, Magnus M, Jonak K, *et al.* NPDock: A web server for protein-nucleic acid docking. *Nucleic Acids Res* 2015;**43:**W425–W430.
23. Dominguez C, Boelens R, Bonvin AMJ. HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* 2003;**125**:1731–1737.
24. Van Zundert GCP, Rodrigues JPGLM, Trellet M, *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol* 2016;**428:**720-725.
25. Liu F, Somarowthu S, Pyle AM. Visualizing the secondary and tertiary architectural domains of lncRNA RepA. *Nature Chemical Biology* 2017;**13**:282–289.
26. Yan K, Arfat Y, Li D, *et al*. Structure Prediction: New Insights into Decrypting Long Noncoding RNAs. *Int J Mol Sci* 2016;**17**:132.

27. Miao Z, Adamiak RW, Blanchet MF, *et al*. RNA-Puzzles Round II: assessment of RNA structure prediction programs applied to three large RNA structures. *RNA* 2015;**21**:1066-84.
28. Miao Z, Adamiak RW, Antczak M, *et al*. RNA-Puzzles Round III: 3D RNA structure prediction of five riboswitches and one ribozyme. *RNA* 2017;**23**:655-672.
29. Sperschneider J, Datta A. KnotSeeker: heuristic pseudoknot detection in long RNA sequences. *RNA* 2008;**14**:630-640.
30. Brierley I, Pennell S, Gilbert RJC. Viral RNA pseudoknots: versatile motifs in gene expression and replication. *Nature Reviews Microbiology* 2007;**5**,598–610.
31. Lee SH, Cha EJ, Lim JE, *et al*. Structural characterization of an intrinsically unfolded mini-HBx protein from Hepatitis B Virus. *Mol Cells* 2012;**34**,165-169.
32. Sidhu K, Kumar S, Reddy VS, *et al*. Mass spectrometric determination of disulfide bonds in the biologically active recombinant HBx protein of hepatitis B virus. *Biochemistry* 2014;**22**:4685-95.
33. van Hemert FJ, van de Klundert MAA, Lukashov VV, *et al*. Protein X of Hepatitis B virus: origin and structure similarity with the central domain of DNA glycosylase. *PLoS ONE* 2011;**6**:e23392.
34. Wass MN, Fuentes G, Pons C, *et al*. Towards the prediction of protein interaction partners using physical docking. *Molecular Systems Biology* 2011;**7**:469.
35. Li T, Robert EI, van Breugel PC, *et al*. A promiscuous alpha-helical motif anchors viral hijackers and substrate receptors to the CUL4-DDB1 ubiquitin ligase machinery. *Nat Struct Mol Biol* 2010;**17**:105-11.
36. Yang Y, Chen L, Gu J, *et al*. Recurrently deregulated lncRNAs in hepatocellular carcinoma. *Nat Commun* 2017;**8**:14421.
37. Villa E, Critelli R, Lei B *et al*. Neoangiogenesis-related genes are hallmarks of fast-growing hepatocellular carcinomas and worst survival. Results from a prospective study. Gut. 2016;**65**:861–869.