Supporting Information

Methods

Sample Preparation and Mass Spectrometry

HepG2-NTCP cells were infected with or without HBV for 7 days. Cell nuclei were extracted with NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Scientific, Waltham, MA) and treated with or without 1% formaldehyde for 10 min. The samples were then lysed and immunoprecipitated with HBc antibody (Dako, Carpinteria, CA) or rabbit IgG (Abcam, Cambridge, United Kingdom) using Pierce Co-IP Kit (Thermo Scientific, Waltham, MA) according to manufacturer's instruction. The resulting samples were washed three times with IP Lysis/ Wash buffer, and then eluted with 50 µl UA buffer (100 mM Tris, 8 M urea).

After elution with UA 1/20th volume of 100 mM DTT (in water) was added and the sample was heated to 37C for 20 minutes. After cooling to room temperature, 3/20th volume of 100 mM Iodoacetamide was added and the sample was incubated for an hour in the dark. Next 1/10th volume of 100 mM beta-mercaptoethanol was added to scavenge excess Iodoacetamide. Water was next added to bring the urea concentration to 4 M and 2 µg of Lys-C (125-05061, Wako, Mountain View, CA) was added and the samples incubated overnight at room temperature. The next morning one volume of 100 mM Tris was added to further dilute the urea to 2 M and 2 µg of trypsin (V5113, Promega, Madison, Wisconsin, United States) was added and the sample mixed and incubated for four hours.

After trypsin digestion samples were diluted with 1 part 1.6% formic acid and chilled on ice. A small aliquot of peptides from a digest of hen lysozyme was added to allow a check of the subsequent quantitation. Next each sample was loaded onto its own stack of two STAGE tips [1] with a C8 empore tip (wide bore 200 μL tip with four gauge 10 C8 cores) on the top and a similarly sized C18 tip on the bottom with adaptors fashioned from the narrower parts of recycled used 1000 μL pipette tips. Column over capacity was intended to address expected polymer contamination. Several loading cycles were used with loading occurring with a reported force of 300G in a chilled centrifuge. After loading was complete each stack was washed serially with (1,2) 100 μL 1.6% formic acid 40 mM ammonium acetate x2, (3) 200 μL 1.6% formic acid, (4) 100 μL 0.16% formic acid (to lower buffering potential of residual solution remaining). These washes were performed using 450G for 5 minutes with complete free solution removal assured with repeated centrifugation when needed. Next the columns were moved to room temperature for ~15 minutes. This results in two pairs (control stack and experiment stack) of column stacks. The control stack from each pair was next treated with 400 μL of 'Light' labeling reagent (normal methyl groups) and the experiment stacks labeled with 400 μL of 'Intermediate' labeling agent (methyl groups with two deuteriums) with reagents prepared and on-columns treatment performed in a manner essentially as described for protocol C [2]. Each reagent was made to flow over the two column beds at room temperature at 100-200G assuring the beds were wet with reagent for a minimum of 30 minutes using recycling of the flow through if needed.

After completion of the labeling cycle the flow through was acidified to pH 2.5 or lower (checked with pH paper with a sacrificial microliter) using 10% formic acid and this solution was chilled. The column sets were also moved to a refrigerator for 15 minutes while the centrifuge chilled down. Next the flow through for each respective set was applied in two cycles to the top of the correct stack with loading occurring at 300G. After this recovery step was complete the serial washing steps, save the last, were repeated. Next the column stack was reversed (C18 on top) and each stack placed over a receiving vessel and the dimethyl labeled peptides eluted first with 100 μL of 0.4% formic acid 40% acetonitrile followed by 100 μL 0.4% formic acid 80% acetonitrile. Next the two samples for each set (now 'marked' with different methyl groups) were mixed thoroughly and distributed equally onto a pair of SCX Stage tips (wide bore 200 μL tip with four gauge 10 SCX cores) with loading at 300G until fully loaded. The columns were washed with 3x200 μL 0.4% formic acid 80% acetonitrile and then with 100 μL 0.4% formic acid 20% acetonitrile. Next the peptides were fractionated into 6 SCX fractions as described by in sample work up options[3], dried down under warm nitrogen gas and resuspended in 20 μL of 0.1% formic acid 2% acetonitrile.

Next 10 μL of the sample was injected onto and LC/MS/MS system comprising a NanoAcquity (Waters) and a (standard) QE-Orbitrap mass spectromter (Thermo Scientific, Waltham, MA). The Nanoacquity was configured with a vented column comprising a trapping column (Symmetry C18, 180 μM x 20 mm, 186003514) and an analytical column (BEH 130 C18 1.7 micron, 75 μM x 250 mm). The column capillary was connected to a conductive union (Upchurch) which provided connection with an emitter with a custom

long ventral coating (New Objective, SilicaTip, FS360-20-10-CEL-20). This was mounted on a pre-Thermo Proxeon Nanospray source. The initial buffer was 0.1% formic acid in water and the organic solution was 0.1% formic acid in acetonitrile.

Ten microliters were each injected onto the vented trap column at a flow rate of 7 μL/minute and acetonitrile 0.3% for 10 minutes. To mitigate the effects of a high background and to improve sampling a very long gradient was performed. After an initial flow rate of 0.1 μL/minute a relatively steep gradient to 2.5% was performed in16.2 minutes which was followed by a 266.9-minute gradient to 23% acetonitrile. This was followed by a 6.9-minute gradient to 80% acetonitrile which was held for 2 minutes prior to an 8-minute drop to 0.3% acetonitrile.

One minute after injection a contact closure signal began collection of data on the QE mass spectrometer. A data dependent analysis was performed with single MS1 spectrum collected from 450-1800 M/Z and then up to 15 signals were targeted for fragmentation data collection with dynamic exclusion for 30 seconds if measured twice in 30 seconds. The raw data were analyzed using MaxQuant 1.5.3.28 [4] and searched against a human proteome downloaded 3/20/16 in addition to a small database including HBV proteins and chicken proteins present in the control. Data were analyzed using a two-channel approach with the two different di-methylation channels used. All other settings used were default settings.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD042733 and 10.6019/PXD042733.

Chromatin isolation for ChIP-qPCR or sequencing

Infected HepG2-hNTCP cells were fixed in freshly prepared 1% formaldehyde in PBS (both ThermoFisher Scientific, Waltham, MA) for 5 min prior to quenching with 125 mM glycine in PBS. Pelleted cells were either used directly for analysis or snap-frozen and stored at -80°C until use. Next, cells were washed with cold lysis buffer [PBS with 0.1% Triton X-100, 0.1% NP-40, 1 mM dithiothreitol (DTT), 50 ng/ml trichostatin A (TSA) to inhibit histone deacetylases and 1x EDTA-free protease inhibitor (Roche, Basel, Switzerland)] and lysed in the same buffer for 10 min on ice. Nuclei were pelleted, resuspended in digestion buffer [H_2O with 50 mM Tris-HCl pH 7.5, 4 mM $MgCl_2$, 1 mM $CaCl_2$, 10% glycerol, 50 ng/ml TSA, and 1x EDTA-free protease inhibitor (Roche, Basel, Switzerland)] and digested with 600 IU/ml micrococcal nuclease (MNase; ThermoFisher Scientific, Waltham, MA) for 12 min at 37°C. Digestion was stopped by addition of 10 mM EDTA. Nuclei were pelleted at 6,500 x g and supernatants collected. The pellet was resuspended in digestion buffer with 10 mM EDTA and 300 mM NaCl and mildly sonicated using a W 375 sonicator (Qsonica, Newtown, CT) at 50% duty cycle and power setting 3. Nuclei were again pelleted, supernatants combined and mixed with an equal amount of sucrose buffer [H_2O with 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM EDTA, 0.01% NP-40, 50 ng/ml TSA, and 1x EDTA-free protease inhibitor (Roche, Basel, Switzerland)]. Samples were concentrated using Amicon Ultra-4 100 kDa centrifugal filter units

(Millipore-Sigma; Merck KGaA, Darmstadt, Germany) and spun on a 5-30% continuous sucrose gradient in sucrose buffer for 4 h at 40,000 x g and 4°C using a SW41Ti rotor (Beckman Coulter Inc., Brea, CA). Chemicals were obtained from Millipore-Sigma unless noted otherwise. Mononucleosome-containing fractions were identified by agarose gel-electrophoresis, pooled, and concentrated to ~500 µl prior to addition of 100 µg/ml bovine serum albumin (ThermoFisher Scientific, Waltham, MA). Mononucleosomes were either stored at -20°C or used immediately for ChIP.

ChIP Sequencing

Antibodies (H3K4me3, Abcam ab8580; HBcAg, DAKO B0586; NCL, Abcam ab13541) were bound to 20 µl/ChIP Dynabeads Protein G (ThermoFisher, Waltham, MA) in ChIP buffer [H_2O with 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, and 200 µg/ml BSA] at 4°C and added to 2 µg of chromatin in ChIP buffer and allowed to bind overnight. Samples were washed six times in LiCl wash buffer [H_2O with 150 mM LiCl, 50 mM Tris-HCl pH 8.0, 1 mM EDTA, 1% NP-40, and 0.7% sodium deoxycholate] and eluted in 100 µl elution buffer [H_2O with 1% SDS and 100 mM $NaHCO_3$]. Crosslinks were reversed by digestion with Proteinase K (ThermoFisher, Waltham, MA) for 5 h at 65°C and then digested for another 30 min with DNase-free RNase A (ThermoFisher, Waltham, MA) at 37°C. DNA was purified using the QIAquick PCR purification kit (Qiagen, Hilden, Germany) according to manufacturer's instructions. 1 ng of DNA was used for sequencing library construction with the KAPA Hyper Prep Kit (Roche) according to manufacturer's instructions. For barcoding, 2 µl of TruSeq RNA adapters stocks were used (Illumina, Inc., San Diego, CA). Up to 6 libraries were pooled per sequencing run

and 500 ng pooled library subject to HBV-specific target enrichment using the xGen Lockdown Reagents Hybridization and Wash Kit (Integrated DNA technologies, Skokie, IL) according to manufacturer's instructions. A custom set of xGen lockdown probes of 60 bp length tiling the entire HBV genome of genotypes A-D was used for target enrichment (Integrated DNA technologies; Dataset S2) together with Dynabeads MyOne T-270 Streptavidin (ThermoFisher, Waltham, MA). The pull-down was amplified for 8 PCR cycles and the product cleaned up using Agencourt AMPure XP beads (Beckman Coulter, Inc.). The product was used to dilute the original library to 20 nM. Success of target enrichment was controlled by qPCR using FAST SYBR Green MasterMix and primers for actin TSS (forward: TAGAAGTCGCAGGACCACACT; reverse: TGGGTAGGTTTGTAGCCTTCAT) and HBV (forward: TTAACAGGCCTATTGATTGGAAA; reverse: TCAACGCAGGATAACCACATT). Typically, 10^{4-5}-fold target enrichment was observed. Sequencing was performed using an Illumina MiSeq with a MiSeq v3 reagent kit for 2x76 bp paired-end reads (Illumina, Inc., San Diego, CA). Raw sequencing files were deposited in gene expression omnibus (GEO) under accession no. GSE123715.

Bioinformatic analysis

Using bowtie2 [5], raw sequencing data in FASTQ format were aligned to a 1.1mer version of an HBV genotype D reference genome V01460.1 (GenBank), shifted by 500 nt compared to the EcoR1 site. The no-discordant, no-mixed, and no-unal options were used and the insert size limited between 120 and 200 bp, as described [6]. Using a custom script, paired reads were merged and the gap between the two mapped reads

represented as “N” with low read quality to represent the entire sequence recovered by sequencing. This file was converted to BEDGRAPH format via BAM using bedtools and samtools, respectively [7,8]. A custom script was used to wrap around reads in the repeated segment of the 1.1mer reference genome to the beginning of the 1.0mer reference genome, again shifted by 500 nt relative to the EcoRI site. HBV-derived reads were normalized to total read counts in the library (alignment to a reference genome combining human (GRCh38.p10), HBV (V01460.1) and phiX (NC_001422.1) genomes. Read counts per position per 10^6 total reads (reads per million, RPM) were saved in WIG format. A custom python script was used to average the read counts across the whole genome and its subregions (core, preS1/S2/S, X). WIG files were visualized using GraphPad Prism 7.04 (GraphPad Software, Inc., La Jolla, CA) and also deposited in gene expression omnibus (GEO) under accession no. GSE123715.

Quantitative real time PCR

For intracellular HBV total DNA and cccDNA quantification, total cellular DNA was isolated from infected cells using NucleoSpin Tissue Tissue kit (Macherey-Nagel, Düren, Germany). HBV total DNA qPCR was performed using rcDNA1745 fw / rcDNA1844 rev primer pair. DNA samples for cccDNA qPCR were treated with 500 U/ml T5 exonuclease (NEB, Ipswich, USA) at 37°C for 30 min. cccDNA specific primer pair cccDNA 92 fw and cccDNA 2251 rev were used for qPCR as described. Extracellular HBV DNA quantification was performed with DNA extraction using DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany). rcDNA1745 fw / rcDNA1844 rev primer pair were used for qPCR. For qRT-PCR, RNA was extracted using RNeasy-kit (Qiagen, Hilden, Germany)

and transcribed into cDNA with Maxima First Strand cDNA Synthesis Kit (Thermo Scientific, Waltham, USA) according to the manufacturer's instructions. qPCR were performed using the LightCycler™ 480 system with SYBR Green Master (Roche, Manheim, Germany) or Probe Master (Roche, Manheim, Germany). A relative unit, defined as the expression ratio of target gene against the reference housekeeping gene (ACTB for RNA, PRNP for DNA), is presented.

Immunofluorescence assay

Cells were washed with precooled PBS and fixed by 4% paraformaldehyde (Invitrogen, USA) for 20 min on ice. Fixed cells were permeabilized with 0.5% Triton X-100 (Sigma-Aldrich, St. Louis, MO, USA) for 10 min at room temperature, following blocked in PBS containing 1% goat serum for 1 h at room temperature. Then, the cells were inoculated with primary antibodies diluted with 1% blocking serum overnight at 4°C. After cells were washed by PBS, the secondary antibodies were added into cells for 1 h in dark at room temperature. Cells were inoculated with Hoechst33258 (Invitrogen, USA) for 5 min at room temperature and images were captured with an Olympus CKX53 FL Microscope (Olympus, Japan). Antibodies: Nucleolin (Cell Signaling Technology, #14574S), Myc (BIOPRIMACY, PMK112M).

Amplification of the three CpG islands and Bisulfite Sequencing

Bisulfite treatment of HBV cccDNA was performed by using Qiagen EpiTect Bisulfite Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. The bisulfite modified DNA was amplified by nested PCR. The primers for amplification of the CpG

islands were used according to Zhang et al 2017[9]. The PCR products of the three CpG islands were cloned into PCR2.1 Vector (Thermo Scientific, Waltham, USA) and subjected to DNA sequencing.

Supplementary Figures

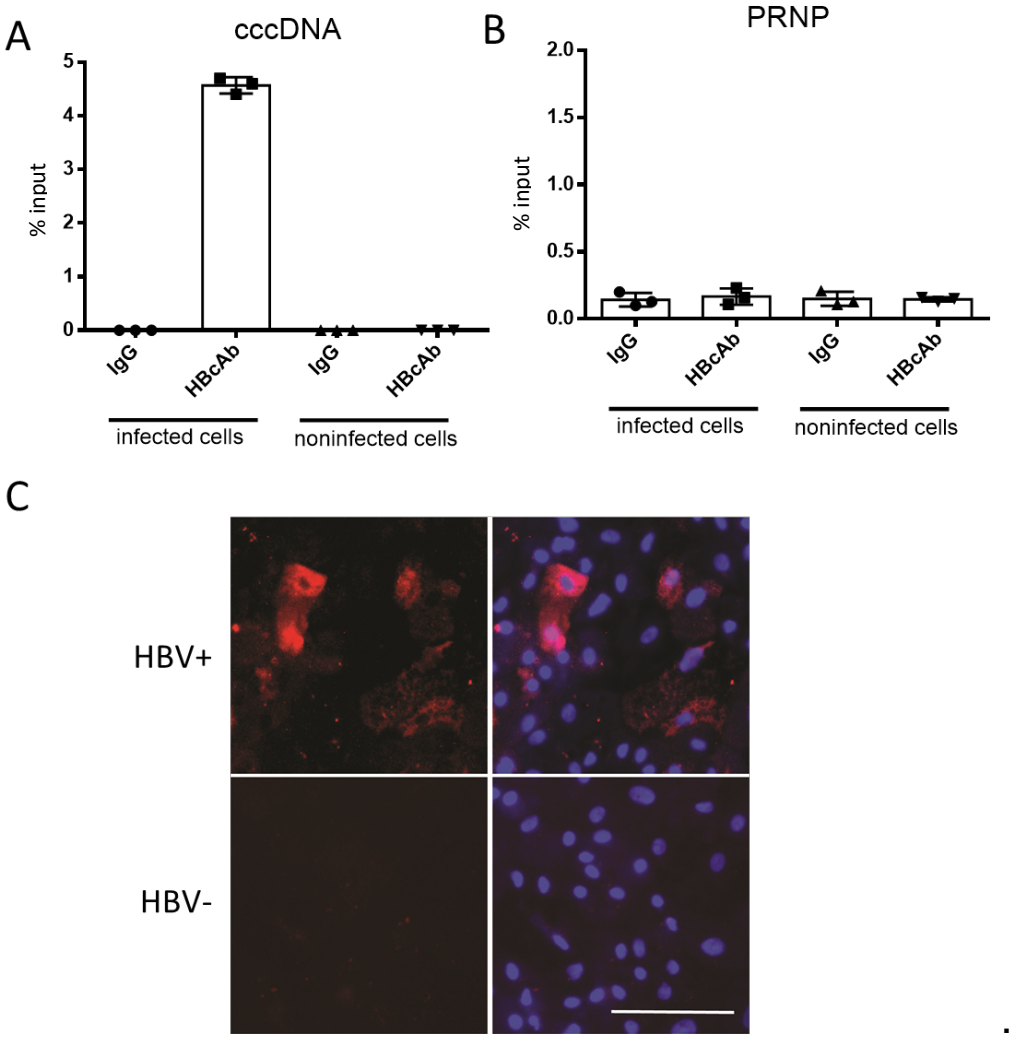

Fig. S1. Immunoprecipitation of cccDNA using anti-core antibody.
HepG2-NTCP cells were infected with or without HBV, and after 7 days. Chromatin immunoprecipitation was performed using lysates of infected or non-infected HepG2-NTCP cells 7 days post-infection. Immunoprecipitation of (A) cccDNA mini-chromosome or (B) host genomic gene PRNP using antibody against HBc or control antibody IgG, was analyzed by qPCR. (C) Cells infected at day 7 were fixed and stained with antibody against HBc. Scale bar = 100 μM.

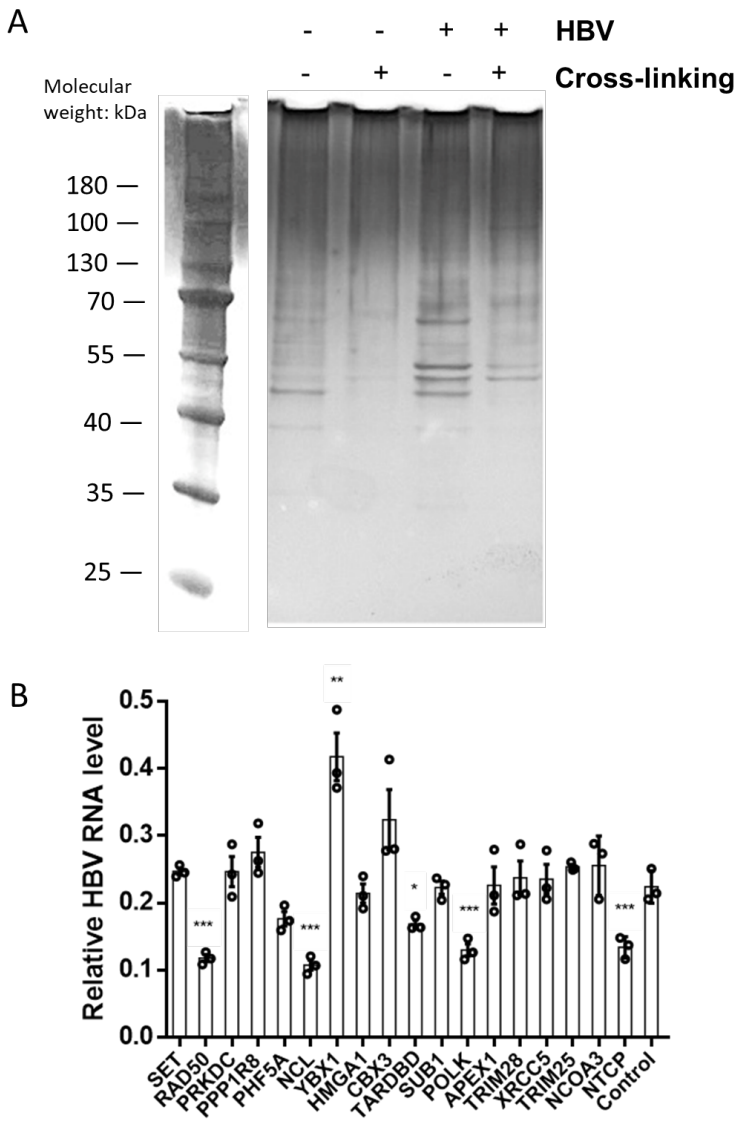

Fig. S2. Identification of HBV cccDNA associated proteins.
(A) HepG2-NTCP cells were infected with HBV, and after 7 days the nuclei were isolated, either with or without cross-linking, and lysed. Anti-HBc or Ig control antibody was used to specifically immunoprecipitate the cccDNA minichromosome. The immunoprecipitated sample was electrophoresed on SDS-PAGE gel and subjected to silver staining. (B) siRNA against selected candidate genes were transfected into HepG2-NTCP cells. After HBV infection, HBV RNA levels were determined by qPCR and expressed as relative units defined as the expression ratio of HBV RNA against the reference housekeeping gene (ACTB). *P< 0.05, **P< 0.01, ***P< 0.001.

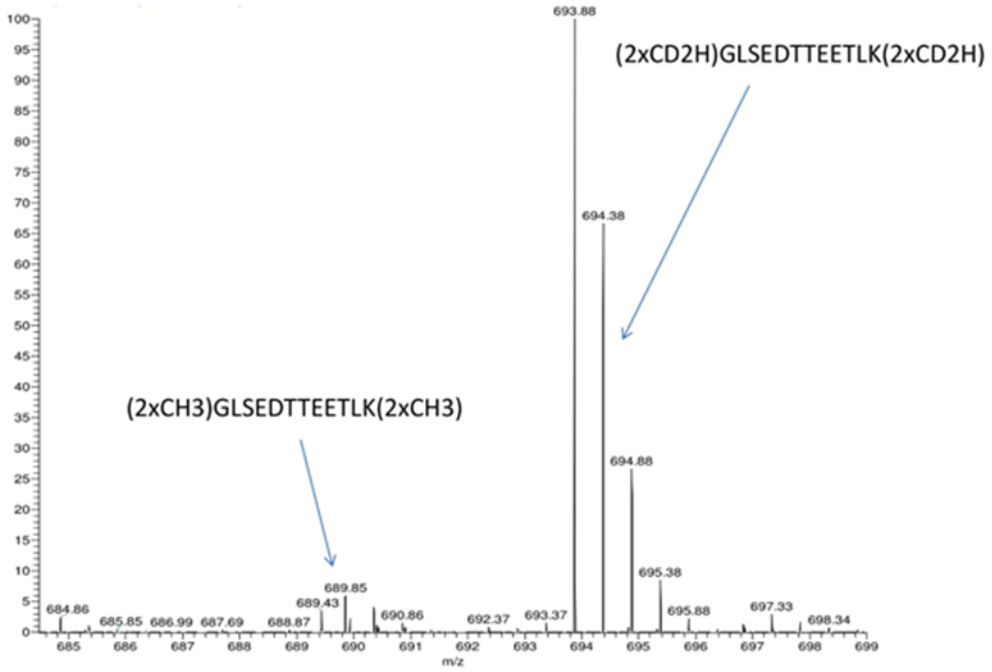

Fig. S3. Identification of Nucleolin by Mass Spectrometry.
The spectrum detail of the enriched NCL peptides showing the GLSEDTTEETLK peptides (nucleolin aa 995-1006) from control Ig (labeled with -CH3) versus anti-HBc pulldown (labeled with -CD2H) is shown here. The spectrum displays the relative overabundance of the peptide, and by inference the protein, in the anti-HBc pulldown experiment.

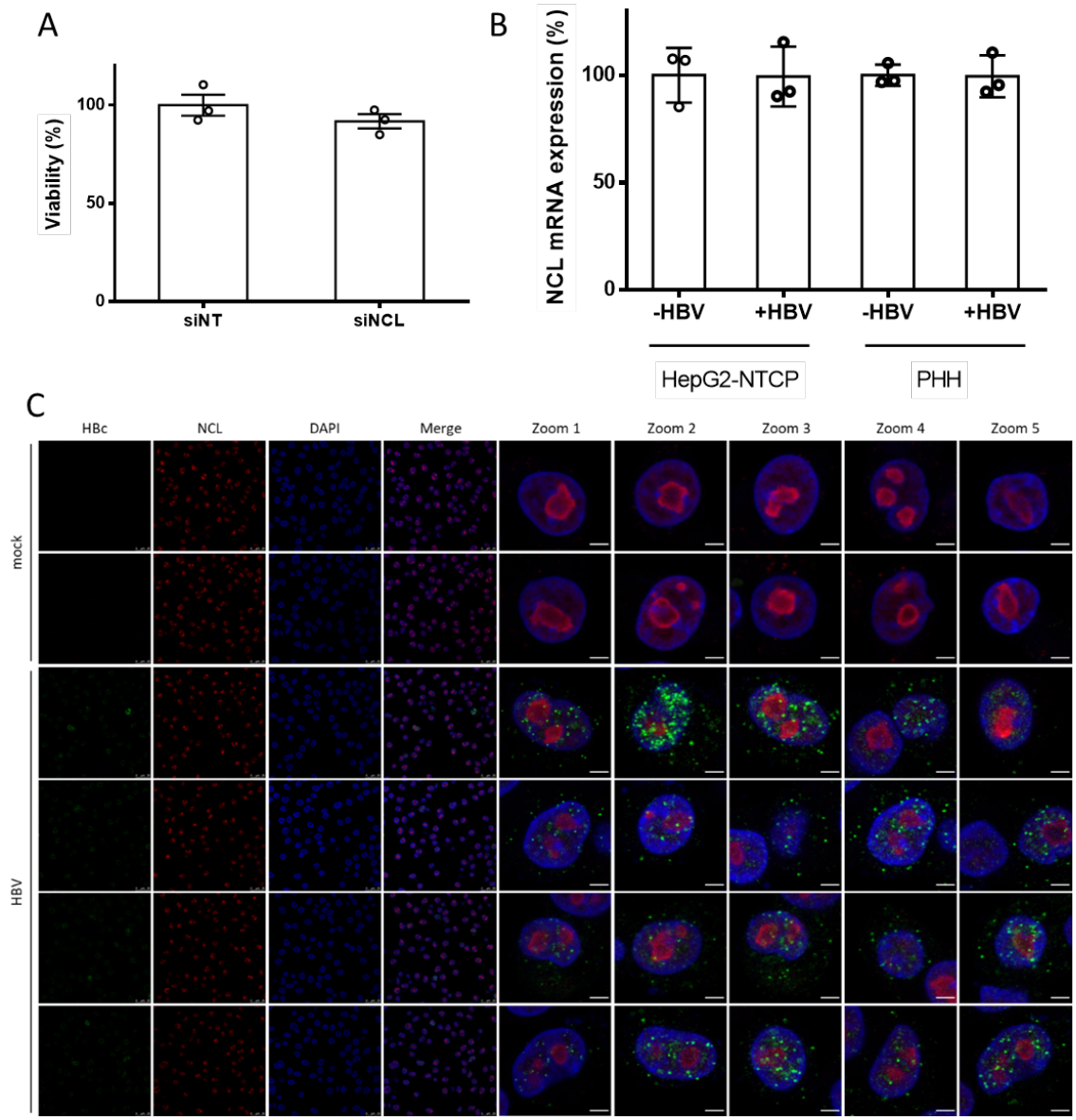

Fig. S4. Nucleolin expression, cell viability and HBV infection.
(A) siRNA against NCL were transfected into HepG2-NTCP cells. After 7 days, cell viability was determined by XTT assay. (B) HepG2-NTCP and PHHs infected with or without HBV. After 7 days, RNA was extracted and NCL mRNA amount was determined by qPCR. (C) HepG2-NTCP cells were infected with HBV, the subcellular distributions of NCL (red) and HBc (green) were determined by immunostaining. Scale bar =2.5 μm.

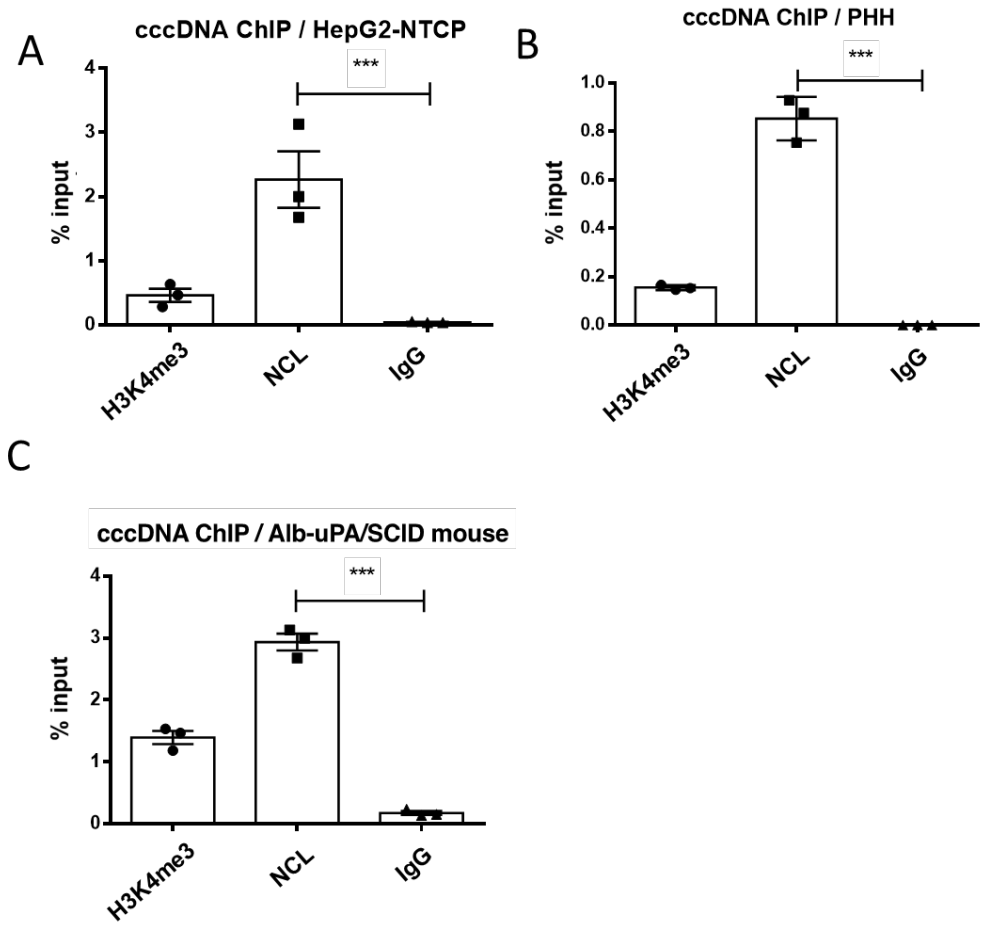

Fig. S5. Chromatin immunoprecipitation of NCL-bound HBV cccDNA in various cell types.

Chromatin immunoprecipitation was performed using lysates of (A) HBV-infected HepG2-NTCP, (B) HBV-infected PHH and (C) liver of HBV infected *Alb-uPA/SCID* mouse. Antibodies against H3K4me3, NCL and control antibody IgG were used. Immunoprecipitated cccDNA was analyzed by qPCR using HBV-specific primers. ****P*< 0.001.

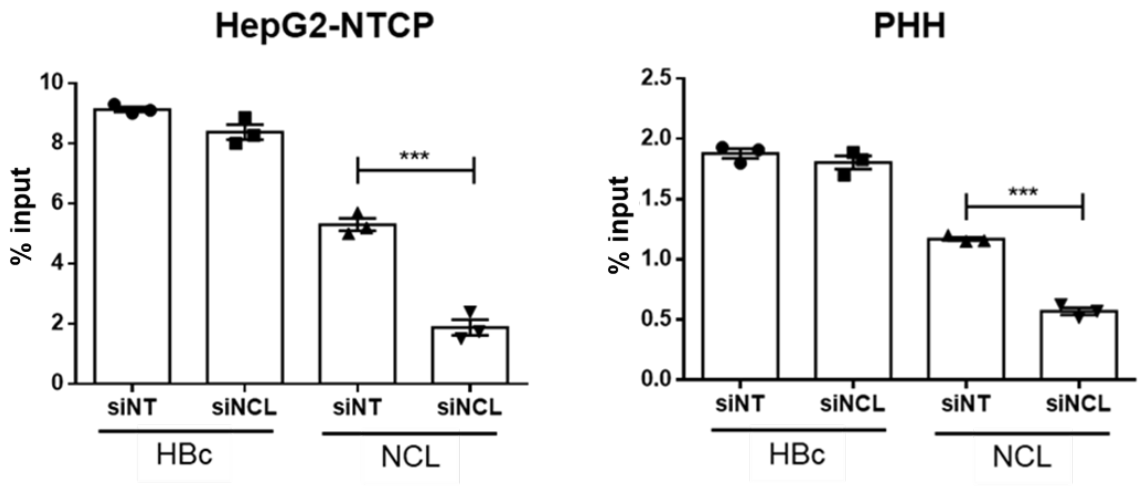

Fig. S6. Binding of HBcAg and nucleolin to HBV cccDNA.
siRNA against NCL or control were transfected into HepG2-NTCP cells and PHH, which were then infected with HBV. Chromatin immunoprecipitation was performed by using antibody against HBcAg or NCL, as indicated. Binding to cccDNA was analyzed by qPCR. % input for IgG control was subtracted from the % input shown for each specific ChIP. ****P*< 0.001.

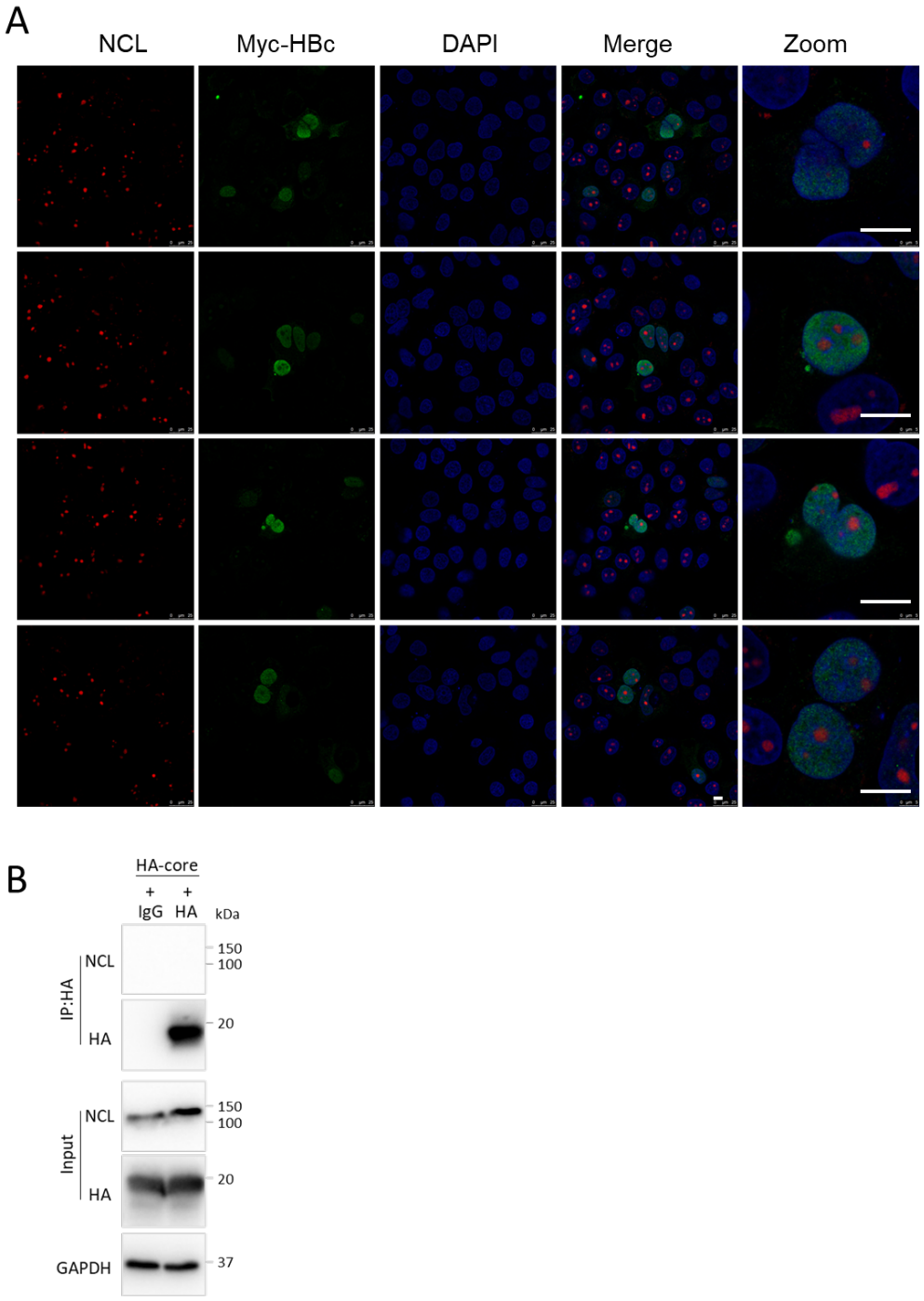

Fig. S7. Investigation of NCL and HBc interaction.
Huh7 cells were transfected with NCL and Myc-HBc or HA-HBc (tagged in the N-terminus) expressing plasmids. Their potential interaction was evaluated by immunostaining (A) and co-IP(B). NCL was stained with anti-NCL and Myc-HBcAg with anti-Myc antibodies. Anti-NCL and anti-HA antibodies were used for immunoprecipitation and Western analysis. Scale bar =10 μm.

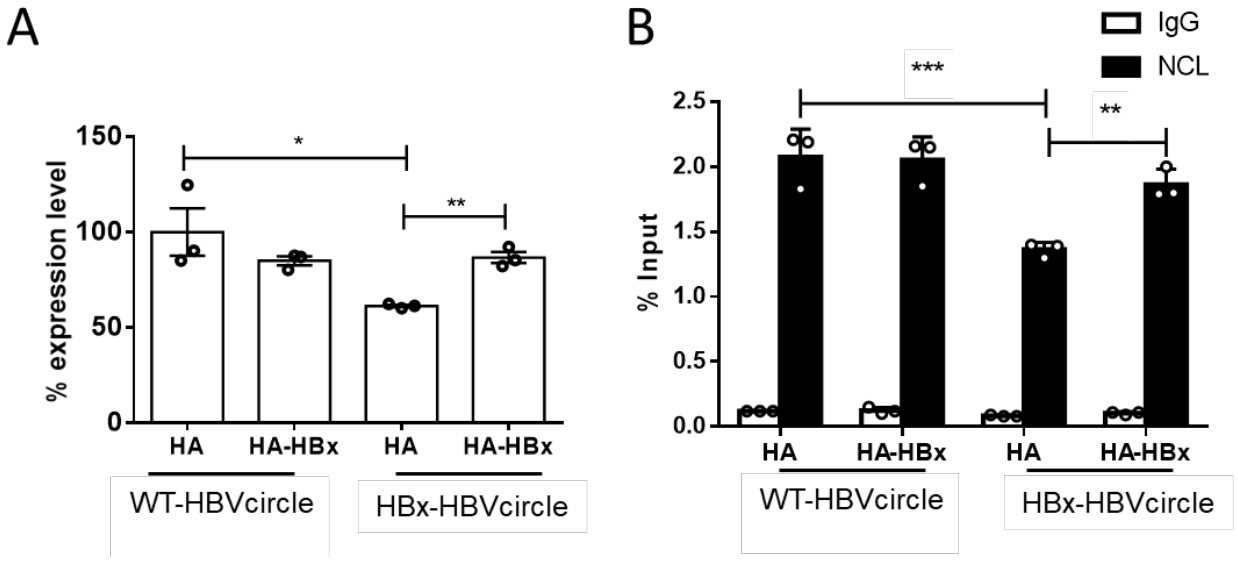

Fig. S8. The role of HBx in NCL mediated transcriptional regulatory effect.
Huh7 cells were transfected with WT-HBVcircle or HBx-HBVcircle, together with HBx expressing or control plasmid. After three days, HBeAg was evaluated by ELISA (A), and ChIP-qPCR of HBVcircle using antibody against NCL was analyzed by qPCR (B). *P< 0.05, **P< 0.01, ***P< 0.001.

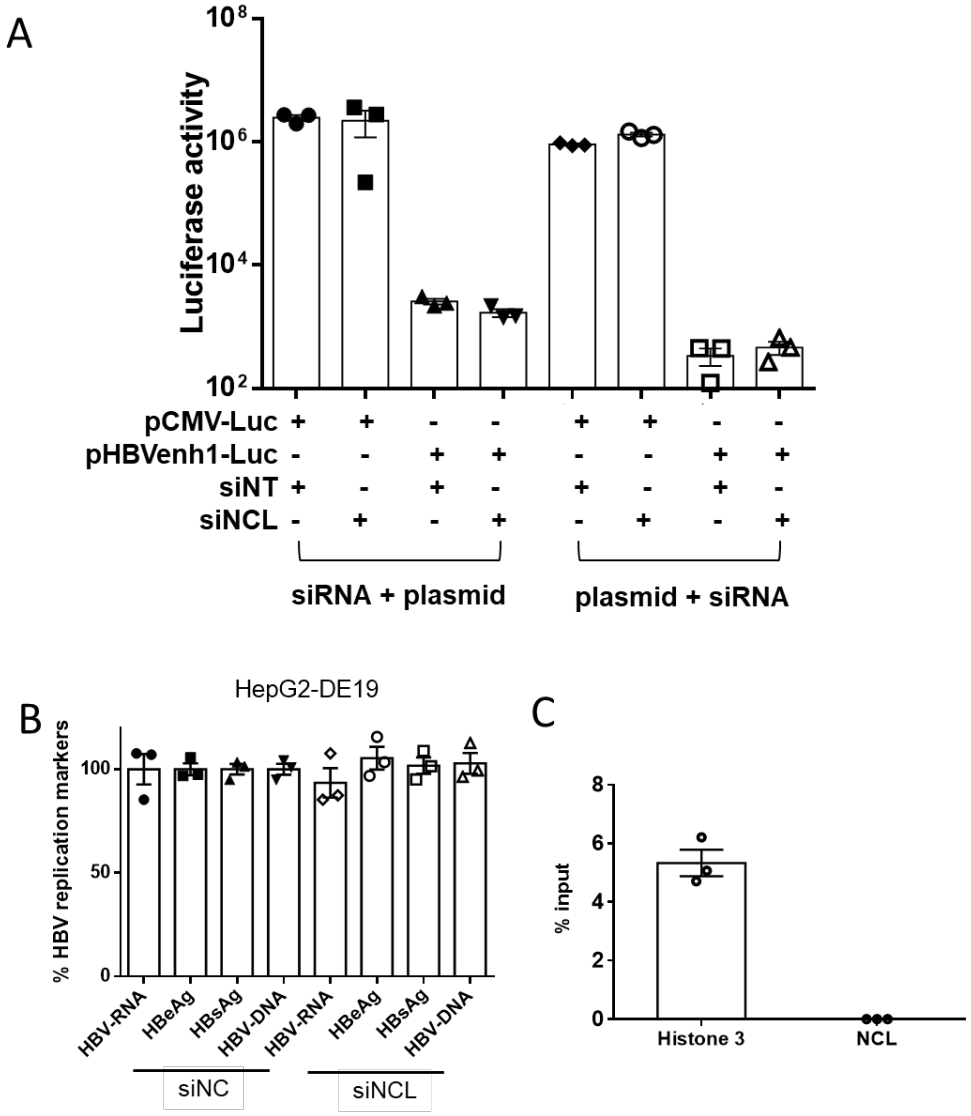

Fig. S9. The role of nucleolin on HBV enhancer and integrated DNA

(A) HepG2-NTCP cells were transfected with siRNA against NCL or control, and then transfected with luciferase plasmid regulated by an CMV promoter or HBV enhancer (siRNA + plasmid). Alternatively, HepG2-NTCP cells were transfected with luciferase plasmid regulated by an CMV promoter or HBV enhancer, and then transfected with siRNA against NCL or control (plasmid + siRNA). Luciferase activities were then determined as described. (B) HepG2-DE19 cells cultured for 7 days in the presence of 1 μg/ml tetracycline to turn off cccDNA synthesis and transfected with siRNA against NCL. Tetracycline was then removed from the media after transfection. 5 days after transfection, various HBV infection markers including HBV RNA, HBeAg, HBsAg and HBV-DNA in cell culture supernatant were determined and levels normalized to the siNT control. (C) HepG2-DE19 cells were cultured for 7 days in the presence of 1 μg/ml tetracycline. ChIP-qPCR was performed by using antibody against histone 3 or NCL. HBV DNA was analyzed by qPCR.

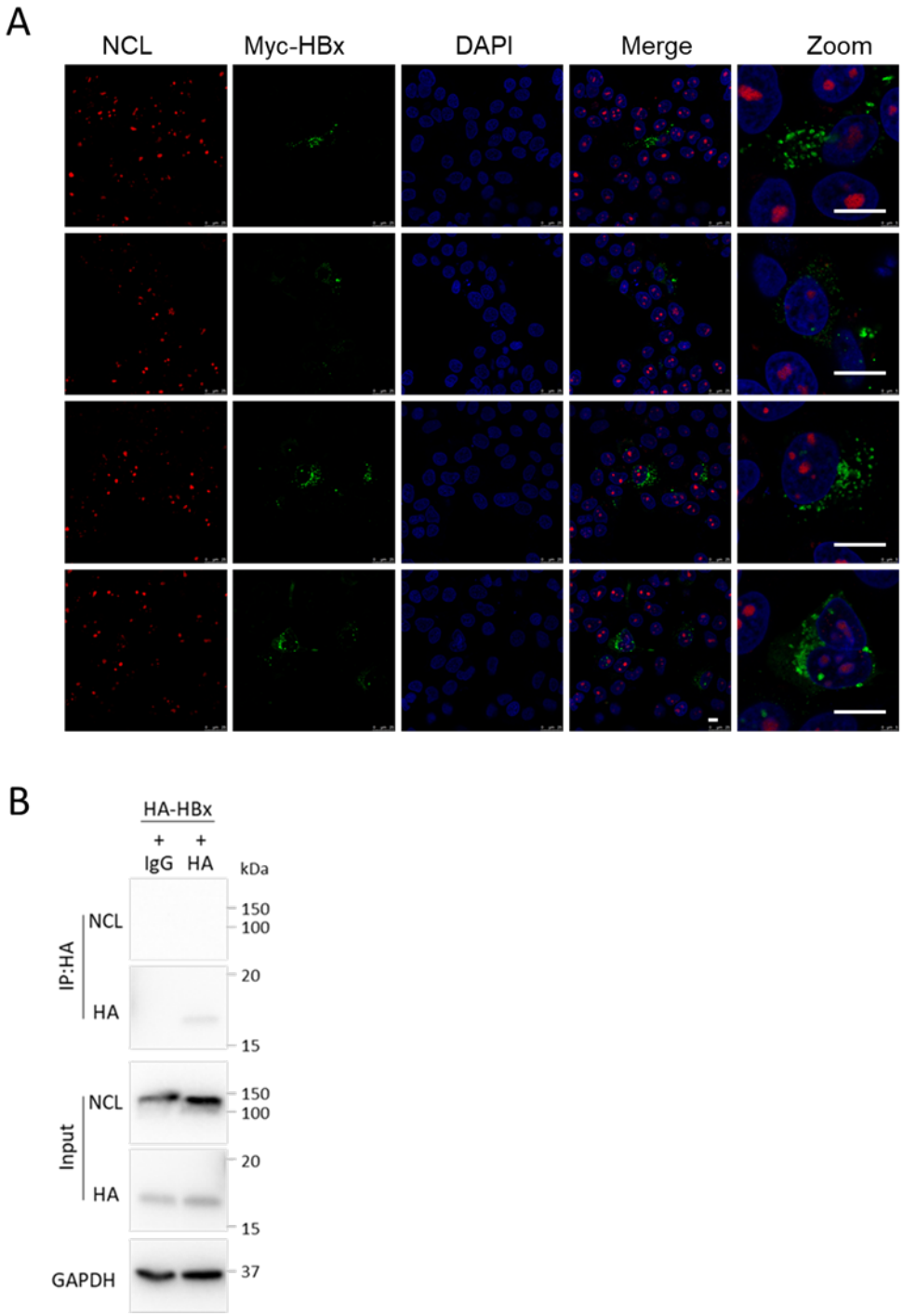

Fig. S10. Investigation of NCL and HBx interaction.
Huh7 cells were transfected with NCL and Myc-HBx or HA-HBx (tagged in the N-terminus) expressing plasmids. Their potential interaction was evaluated by immunostaining (A) and co-IP(B). NCL was stained with anti-NCL and Myc-HBx with anti-Myc antibodies. Anti-NCL and anti-HA antibodies were used for immunoprecipitation and Western analysis. Scale bar =10 μm.

References:

1 Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896-1906, doi:10.1038/nprot.2007.261 (2007).

2 Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc* **4**, 484-494, doi:10.1038/nprot.2009.21 (2009).

3 Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* **11**, 319-324, doi:10.1038/nmeth.2834 (2014).

4 Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372, doi:10.1038/nbt.1511 (2008).

5 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).

6 Tropberger, P. *et al.* Mapping of histone modifications in episomal HBV cccDNA uncovers an unusual chromatin organization amenable to epigenetic manipulation. *Proc Natl Acad Sci U S A* **112**, E5715-5724, doi:10.1073/pnas.1518090112 (2015).

7 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

8 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).

9 Zhang, Y., Mao, R., Guo, H. & Zhang, J. Detection of HBV cccDNA Methylation from Clinical Samples by Bisulfite Sequencing and Methylation-Specific PCR. *Methods Mol Biol* **1540**, 73-84, doi:10.1007/978-1-4939-6700-1_7 (2017).