Identification of Slug and Sox7 as transcriptional repressors binding to the Hepatitis B Virus Core Promoter

Hui Ling Ko[1], Tze Hau Lam[1], Huijin Ng[1,2], Jiaying Toh[1,3],

Liang Wei Wang[1,4], Ee Chee Ren[1,5*]

[1] Singapore Immunology Network, 8A Biomedical Grove, #03-06 Immunos, Singapore 138648.

[2] Current address: Oxford Center for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Headington OX3 7LE

[3] Current address: Department of Microbiology & Immunology, Stanford University, 300, Palo Alto, CA94304

[4] Current address Division of Medical Sciences, Virology Program, Harvard Medical School, 260 Longwood Ave, Boston, MA02115.

[5] Department of Microbiology & Immunology, Yong Loo Lin School of Medicine, National University of Singapore, 5 Science Drive 2, Singapore 119260

*Corresponding Author:

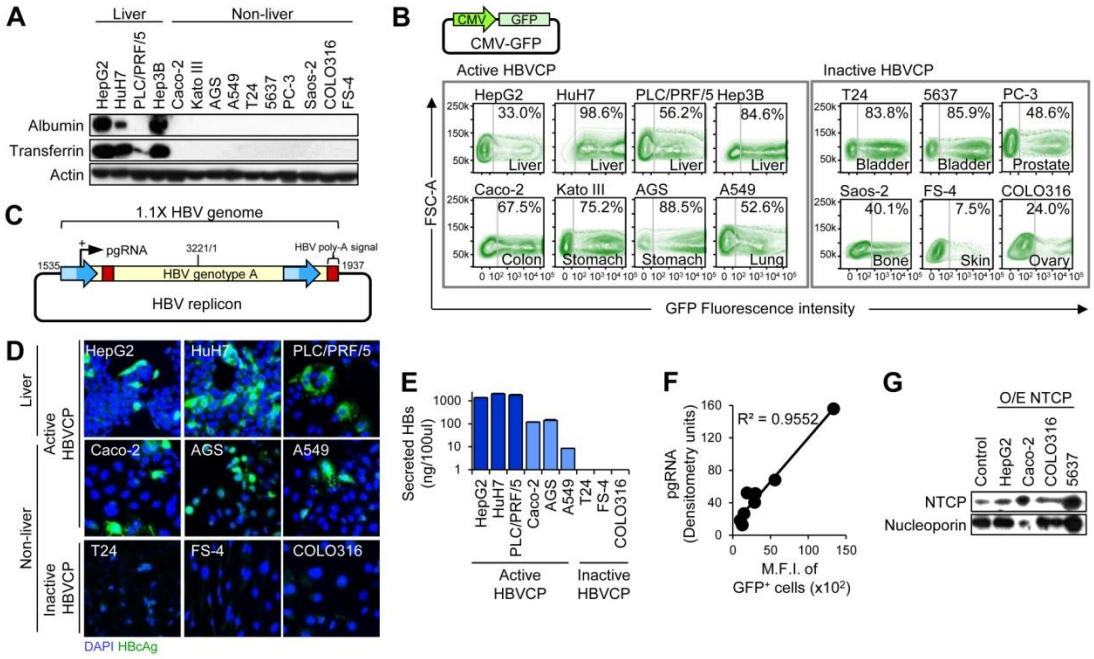
Ee Chee REN,
8A, Biomedical Grove, Immunos #03-06, Singapore 138648
Tel: 65-64070004; Fax: 65-64642056
Email: ren_ee_chee@immunol.a-star.edu.sg

Keywords: HBV, Slug, Sox7, Transcription, HNF4α, pgRNA, cccDNA, promoter, repression.

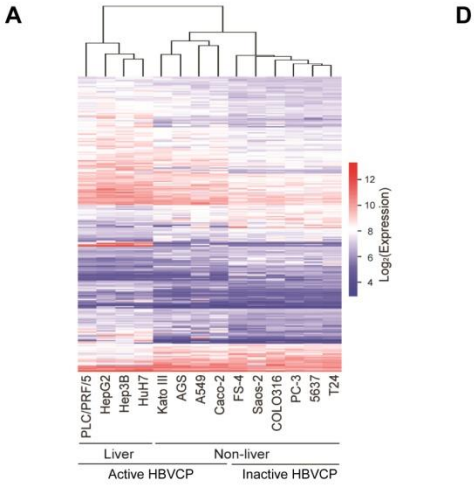
Electronic Word count: 5410 words
Number of figures and tables: 6 Figures 0 Tables
Conflict of Interest Statement: The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.
Financial support
This study was financially supported by a Singapore Immunology Network core grant C07014.
Authors' contributions
H.L.K. and E.C.R. contributed to experimental design, functional analyses and writing of manuscript. T.H.L. contributed to informatics and statistical analyses. H.L.K., H.N., J.T and L.W.W. contributed to data collection and experimentation. E.C.R. is the senior and corresponding author who designed and led the project.

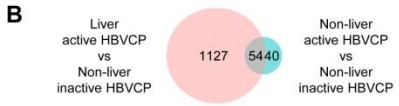
Supplementary Data

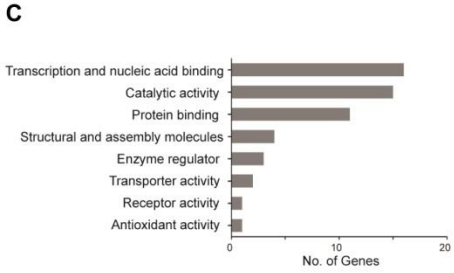
Supplementary Figures



Supplementary Fig. 1. HBVCP is active only in liver and selected non-liver cells.
(A) Expression of liver-specific markers in cell lines by western blot. Only cell lines of liver origin express albumin and/or transferrin. Actin was used as loading control. (B) Cell lines were tested for transfection efficiency using CMV-GFP reporter construct. (C) The HBV replicon. 1.1x full-length HBV from genotype A (nt 1535-1937) was cloned into pcDNA3.1+ vector. pgRNA synthesis is controlled at the HBVCP (blue arrow), terminating at the HBV poly-adenylation (poly-A) signal after initial read-through transcription. Since all cell lines possess CMV promoter activity (Supplementary Fig. 1B) but only selected cell lines generate pgRNA (Fig. 1B), HBcAg (Supplementary Fig. 1D) and HBs (Supplementary Fig. 1E), the CMV promoter has little effect on HBV transcript synthesis. (D) Cytoplasmic staining for HBV capsid protein HBcAg (green) is seen only in cell lines with active HBVCP. Cells were transfected with HBV replicon for 72 hours. Nuclei are stained with DAPI. (E) Only cells with active HBVCP support HBV replication and secrete HBs into culture media after transfection with HBV replicon for 72 hours. (F) HBVCP transcription activity by reporter assay correlates well with pgRNA expression (Fig. 1B) in cells with active HBVCP. (G) Overexpression of NTCP for HBV infection in cell lines 48 hours post-transfection (Fig. 1C). Mock-transfected HepG2 cells were used as control for NTCP expression. Membrane protein nucleoporin p62 was used as loading control.





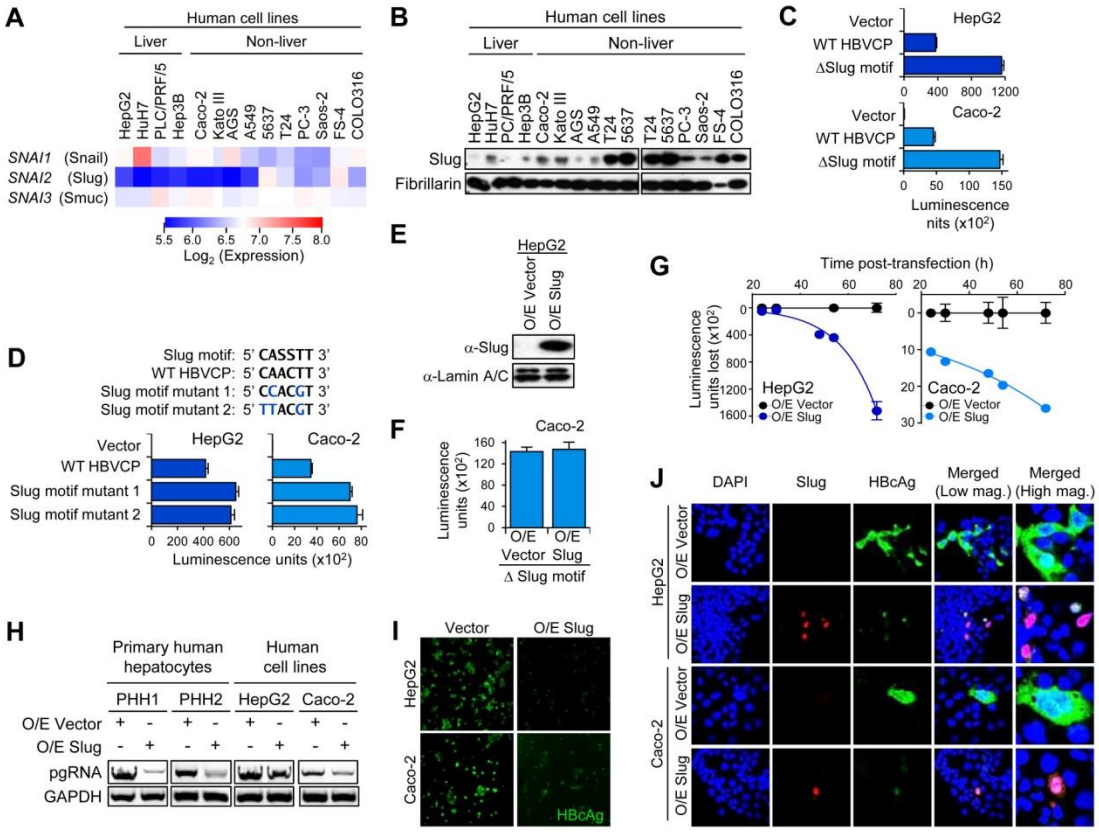


D

Gene_ID	Gene Name
Transcription and nucleic acid binding factors	
C1orf131	Chromosome 1 open reading frame 131
ETS1	Protein C-ets-1
FSTL1	Follistatin-related protein 1
HNF4A	Hepatocyte nuclear factor 4-alpha
INHBA	Inhibin beta A chain
KCTD12	BTB/POZ domain-containing protein KCTD12
MAK16	MAK16 homolog (S. cerevisiae)
NOSTRIN	Nostrin
PNPT1	Polyribonucleotide nucleotidyltransferase 1
PTCD3	Pentatricopeptide repeat domain 3
SEMA4G	Semaphorin-4G
SNAI2	Snail homolog 2 (Drosophila)
SNRPD1	Small nuclear ribonucleoprotein Sm D1
SUSD1	Sushi domain-containing protein 1
WDR43	WD repeat-containing protein 43
ZFHX4	Zinc finger homeobox protein 4
Catalytic activity	
ALDH2	Aldehyde dehydrogenase
ALDH5A1	Succinate-semialdehyde dehydrogenase
B3GNT2	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 2
CHUK	Inhibitor of nuclear factor kappa-B kinase subunit alpha
DDX18	ATP-dependent RNA helicase
EPT1	Ethanolaminephosphotransferase 1
ERMP1	Endoplasmic reticulum metallopeptidase 1
GCNT3	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase 3
GSR	Glutathione reductase
HMGCS1	Hydroxymethylglutaryl-CoA synthase
NAE1	NEDD8-activating enzyme E1 regulatory subunit
PDSS1	Decaprenyl-diphosphate synthase subunit 1
PPID	Peptidyl-prolyl cis-trans isomerase D
SMURF2	E3 ubiquitin-protein ligase
TANC2	Protein TANC2
Protein binding	
BCAR3	Breast cancer anti-estrogen resistance 3
C16orf70	Chromosome 16 open reading frame 70
C9orf100	Chromosome 9 open reading frame 100
EPB41L5	Erythrocyte membrane protein band 4.1 like 5
HSPA14	Heat shock 70kDa protein 14
HSPA9	Heat shock 70kDa protein 9 (mortalin)
LRP12	Low density lipoprotein receptor-related protein 12
NSMCE4A	Non-SMC element 4 homolog A (S. cerevisiae)
NUP35	Nucleoporin 35kDa
PSMD11	Proteasome (prosome, macropain) 26S subunit, non-ATPase, 11
RNF43	Ring finger protein 43
Structural and assembly molecules	
EXOC6	Exocyst complex component 6
FUZ	Fuzzy homolog (Drosophila)
STOML2	Stomatin-like protein 2
WASF2	Wiskott-Aldrich syndrome protein family member 2
Enzyme regulator	
ARHGAP12	Rho GTPase-activating protein 12
MYO9B	Unconventional myosin-Ixb
TBC1D14	TBC1 domain family member 14
Transporter activity	
FLVCR1	Feline leukemia virus subgroup C receptor-related protein 1
SLC39A14	Zinc transporter ZIP14
Antioxidant activity	
GPX2	Glutathione peroxidase 2
Receptor activity	
IDH1	Isocitrate dehydrogenase 1 (NADP+), soluble
Unknown	
FAM35A	Family with sequence similarity 35, member A

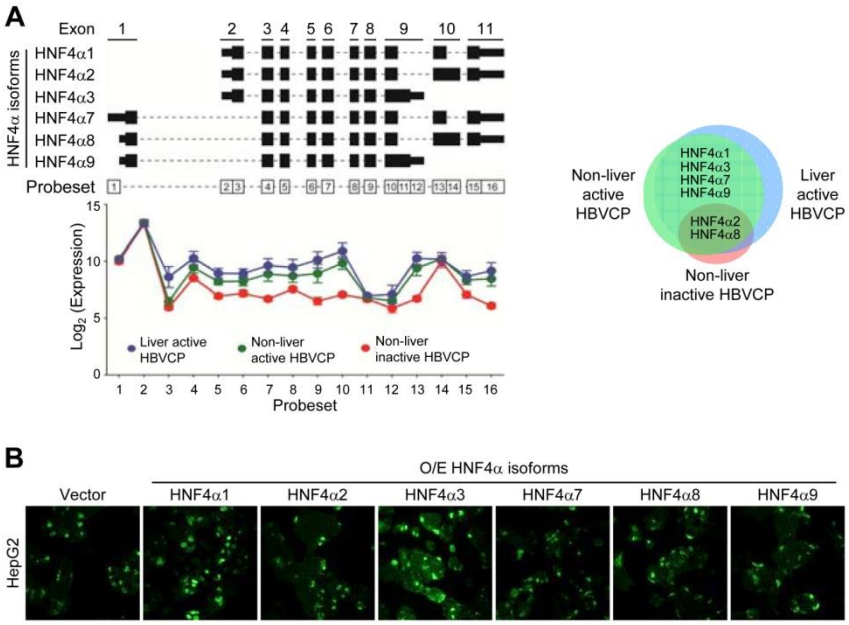
Supplementary Fig. 2. Human transcriptome array (HTA).

(A) Unsupervised hierarchical clustering of cells where HBVCP is active or inactive by relative expression of differentially expressed protein coding transcripts. (B) Number of genes differentially expressed between active HBVCP and inactive HBVCP cell clusters. (C) Genes differentially expressed between cells with active HBVCP and inactive HBVCP categorized by biological functions defined using GO terms. (D) Genes in each GO term category.



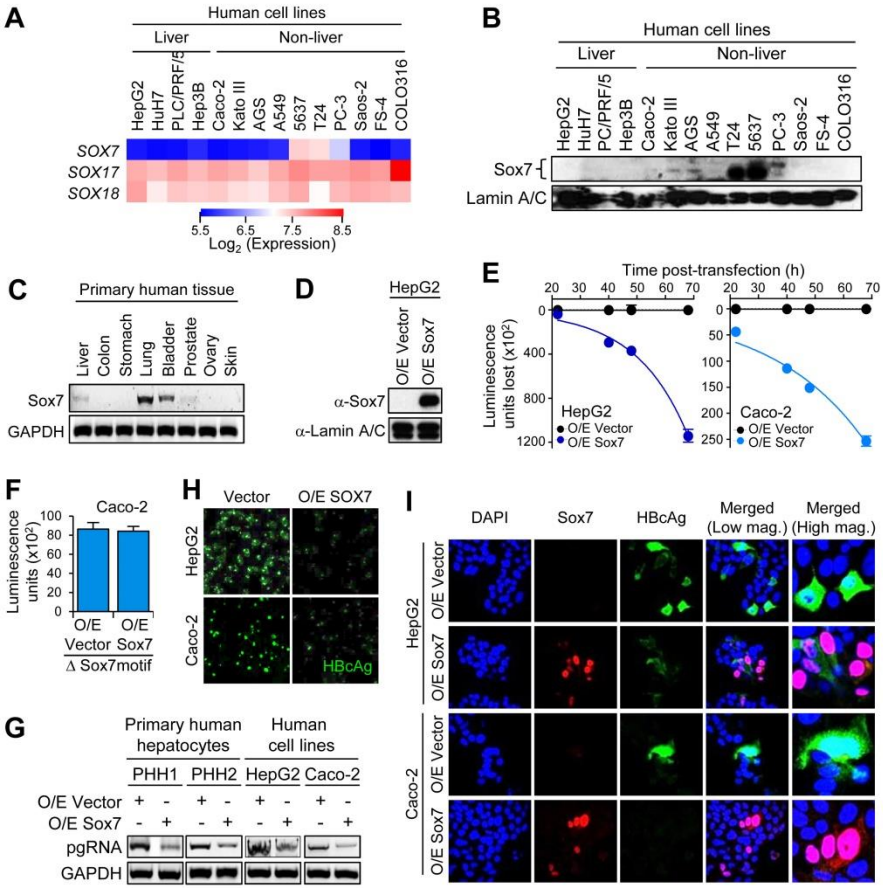
Supplementary Fig. 3. Slug is a specific transcription repressor at the HBVCP.

(A) *SNAI2* gene expression by HTA compared with other Snail family genes *SNAI1* and *SNAI3*. (B) Slug protein expression in the 14 cell lines by western blotting. Slug is highly expressed in cells with inactive HBVCP. Expression of fibrillarin was used as loading control. (C) Slug exerts its repressive effect on HBVCP transcription in a motif-dependent manner, as motif deletion increased transcription at the HBVCP (P<0.001, Student's T-test). (D) Slug motif mutations (blue text) increase HBVCP activity in HepG2 and Caco-2 (P<0.001, Student's T-test), suggesting that the repressive activity of Slug is motif sequence-dependent. "S" denotes either the base "C" or "G". (E) Western blot of nuclear lysates showing overexpressed Slug in HepG2 cells 48 hours post-transfection. Lamin-A/C was used as loading control. (F) Slug motif deletion nullifies effect of overexpressed Slug on HBVCP by luciferase reporter assay. (G) Time-course showing continued HBVCP transcription suppression with Slug overexpression. (H) pgRNA expression 72 hours after replicon transfection in PHH and cell lines with active HBVCP overexpressing Slug. GAPDH was used as loading control. (I, J) Representative images of cells with active HBVCP overexpressing Slug (red) have negligible staining for HBcAg (green). Cell nuclei are stained with DAPI.



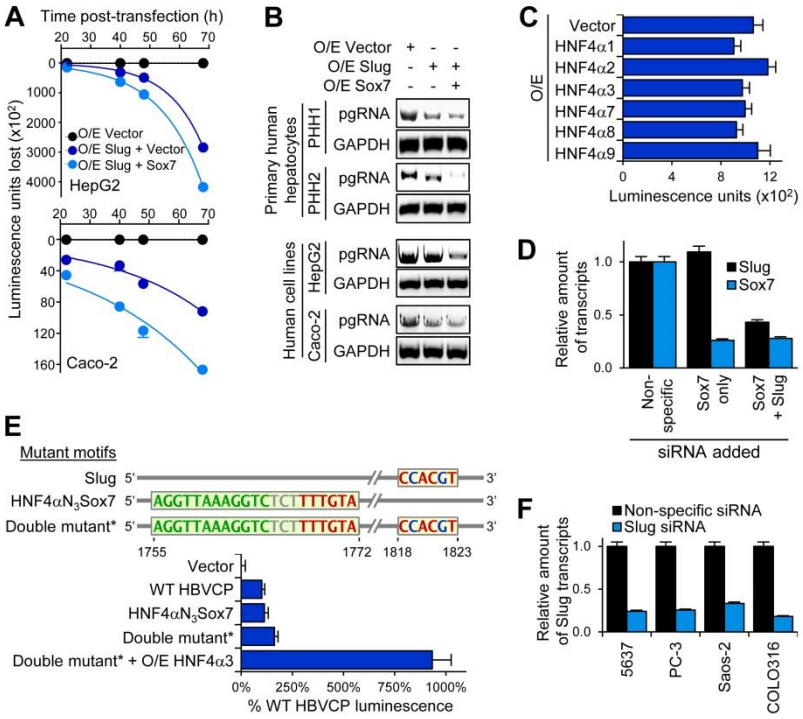
Supplementary Fig. 4. HNF4α isoforms are differentially expressed between cells with active and inactive HBVCP.

(A) Differential expression of HNF4A gene exons in cells by HTA. Exons are indicated as bars, with protein coding sequences represented as thicker bars than the 5'UTRs and 3'UTRs. The signal for probeset 3 detecting for coding sequence of exon 2 utilized in HNF4α1, HNF4α2 and HNF4α3 is higher in liver cells than non-liver cells, which is consistent with western blot data. The signal for probeset 13 detecting for HNF4α1, HNF4α2, HNF4α7 and HNF4α8 is high in HBV permissive cells, but probeset 14 detecting for exon 10 in HNF4α2 and HNF4α8 is not differentially expressed between cells with active or inactive HBVCP, together suggesting that HNF4α1 is preferentially expressed in cells with active HBVCP. There is marginally higher signal for probeset 12 which detects for exon 9 utilized in HNF4α3 and HNF4α9 in cells with active HBVCP. The Venn diagram summarizes the exon array data for HNF4A gene suggesting that in contrast to HNF4α2 and HNF4α8, HNF4α1/3/7/9 correlate with efficiency for HBVCP transcription activity. (B) Representative images at low magnification showing different intensities of staining for HBcAg (green) when HNF4α isoforms are overexpressed in HepG2.



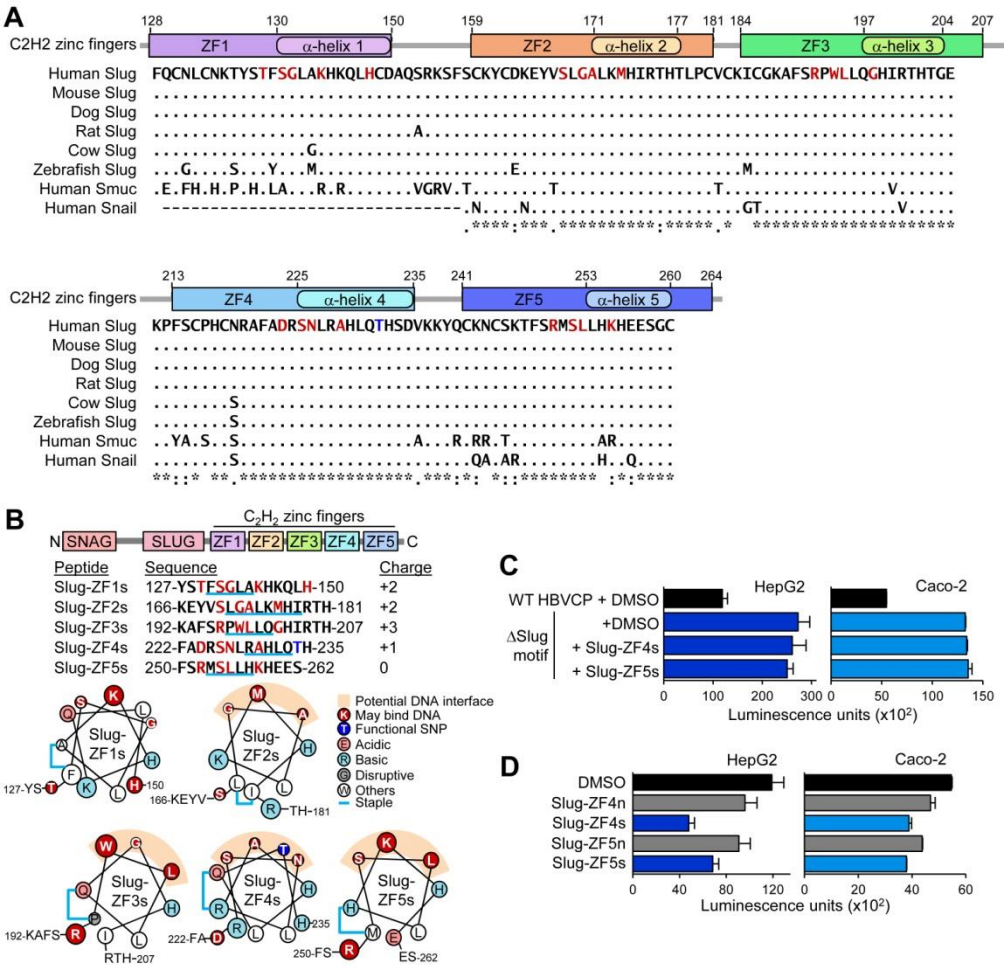
Supplementary Fig. 5. Sox7 is a specific transcription repressor at the HBVCP.

(A) *SOX7* gene expression by HTA compared with other Sox group F family genes. (B) Expression of Sox7 in human cell lines by western blotting. Expression of Lamin A/C was used as loading control. (C) Expression of full-length Sox7 in primary human tissues by RT-PCR. Expression of GAPDH was used as loading control. (D) Overexpressed Sox7 in HepG2 nuclear lysates by western blot 48 hours post-transfection. Expression of Lamin A/C was used as loading control. (E) Time-course showing Sox7 suppresses HBVCP activity when overexpressed in HepG2 and Caco-2. (F) Sox7 motif deletion abrogates the repressive effects of overexpressed Sox7 on HBVCP transcription. (G) pgRNA expression is reduced in PHH and cell lines co-transfected with Sox7 and HBV replicon for 72 hours. (H, I) Representative images of cells with active HBVCP overexpressing Sox7 (red) have negligible staining for HBcAg (green). Cell nuclei are stained with DAPI.



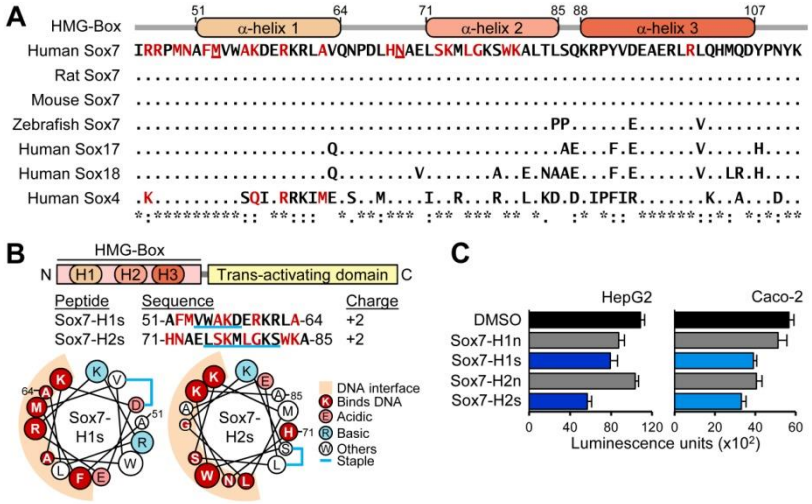
Supplementary Fig. 6. Combined effect of Slug and Sox7 on HBVCP inhibition.

(A) Time-course showing suppression of HBVCP activity with Slug overexpression alone, which is inhibited to greater extent with further addition of overexpressed Sox7. (B) Reduced pgRNA expression with overexpressed Slug alone further diminishes with when Sox7 is also overexpressed in PHH and cells with active HBVCP co-transfected for 72 hours with HBV replicon. Expression of GAPDH is used as loading control. (C) HBVCP remains inactive with overexpression of HNF4α3 alone Slug^{+} Sox7^{+} PC-3 cells. (D) Slug and Sox7 transcript expression is decreased (P<0.001, Student's T-test) by quantitative real-time PCR in PC-3 cells treated with Slug or Sox7 specific siRNA. (E) Inactive HBVCP in Slug^{+} Sox7^{+} 5637 non-liver cells becomes activated in a HNF4α-dependent manner by Slug motif mutation (blue text) together with increasing the spacer (grey text) between HNF4α and Sox7 motifs (P<0.001, Student's T-test). *Double mutant: HBVCP-Luc construct doubly mutated at the Slug motif and carrying the HNF4αN_3Sox7 mutation. (F) Slug transcripts are downregulated (P<0.001, Student's T-test) in cells with inactive HBVCP transfected with Slug-specific siRNA.



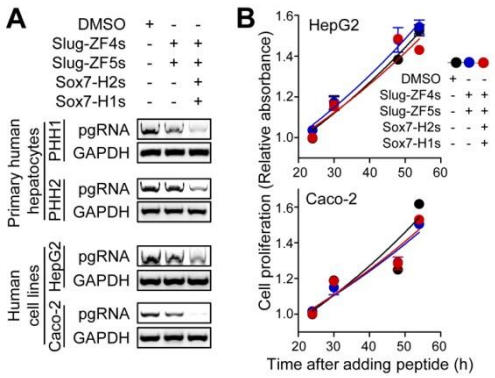
Supplementary Fig. 7. Slug stapled peptides.

(A) Multiple sequence alignments of Slug orthologues and human Slug family members for the five C_2H_2 zinc fingers (ZF1-5). As C_2H_2 zinc fingers bind DNA through residues -1, +2, +3 and +6 of the α-helix, these residues of predicted α-helical regions are highlighted in red text. The variable Slug residue encoded by a functional SNP is highlighted in blue text. (B) Slug functional domains showing five C_2H_2 zinc fingers (ZF1-5), each may bind DNA. Potential DNA binding residues are indicated in red text, site of hydrocarbon staples are underlined in blue. Helical wheel diagrams show predicted DNA-binding interfaces of Slug stapled peptides. ZF1 is deduced to have negligible HBVCP-binding. (C) Luciferase reporter assay showing slug stapled peptides act specifically on HBVCP as Slug motif deletion abrogates HBVCP transcription repression in the presence of Slug-ZF4s and Slug-ZF5s. (D) Hydrocarbon stapling to preserve helicity and function of α-helical peptides derived from Slug zinc fingers 4 and 5 (Slug-ZF4n and Slug-ZF5n) achieved enhanced Slug mimetic function (P<0.01, Student's T-test), resulting in further transcription inhibition at the HBVCP.



Supplementary Fig. 8. Sox7 stapled peptides.

(A) Multiple sequence alignments of Sox7 orthologues and human Sox group F family members for the N-terminal DNA-binding HMG-Box. Residues of α-helical regions corresponding to DNA-binding residues in Sox17-DNA crystal structure are highlighted in red text. As α-helix 3 only has 1 residue that interacts with DNA, its role in DNA-binding is negligible compared to α-helices 1 and 2. (B) Sox7 functional domains showing the DNA-binding HMG-box containing 3 α-helices (H1-H3). Only H1 and H2 are predicted to bind DNA strongly. DNA binding residues are in red text, site of hydrocarbon staples are underlined in blue. Helical wheel diagrams show predicted DNA binding interface of Sox7 stapled peptides. (C) Hydrocarbon-stapled peptides (Sox7-H1s and Sox7-H2s) derived from Sox7 HMG-box helices 1 and 2 (Sox7-H1n and Soc7-H2n) show enhanced Sox7 mimetic function (P<0.01, Student's T-test) to further reduce transcription at the HBVCP.



Supplementary Fig. 9. Combined effect of Slug and Sox7 stapled peptides.

(A) PHH and cells with active HBVCP were transfected with HBV replicon for 24 hours and treated with Slug stapled peptides show decreased pgRNA synthesis 72 hours post-transfection. When the cells were also concurrently treated with Sox7 stapled peptides, pgRNA expression was further diminished. (B) The stapled peptides are not cytotoxic as no difference in cell proliferation by WST-1 assay was observed in cells treated with the peptides compared with DMSO control.

Supplementary Table 1. **Amplification and cloning primers for HNF4α isoforms**

Primer	**Sequence**
NotI-HNF4α123-F	5’ TGAGCGGCCGCGATATGCGATCTC 3’
NotI-HNF4α789-F	5’ TGAGCGGCCGCGATATGGTCAGCG 3’
HNF4α1278-XmaI-R	5’ ATTCCCGGGATAACTTCCTGCTTGGTG 3’
HNF4α39-XmaI-R	5’ ATTCCCGGGAGCAACTTGCCCAAAGCG 3’
HNF4α17F	5’ AACGGACAGATGTCCACCCCTGAGACC 3’
HNF4α17R	5’ GGTCTCAGGGGTGGACATCTGTCCGTT 3’
HNF4α28F	5’ AGCAACGGACAGATGTGTGAGTGGCC 3’
HNF4α28R	5’ GGCCACTCACACATCTGTCCGTTGCT 3’

Cloning strategy

HNF4α isoform	**Step 1: Generate fragments with complementary ends**		**Step 2: Anneal fragments and amplify**	
	5’ fragment	**3’ fragment**	**Forward primer**	**Reverse primer**
HNF4α1	A	B	NotI-HNF4α123-F	HNF4α1278-XmaI-R
HNF4α2	C	D	NotI-HNF4α123-F	HNF4α1278-XmaI-R
HNF4α7	F	B	NotI-HNF4α789-F	HNF4α1278-XmaI-R
HNF4α8	G	D	NotI-HNF4α789-F	HNF4α1278-XmaI-R
HNF4α3	E (Does not require annealing of separate fragments)			
HNF4α9	H (Does not require annealing of separate fragments)			

Fragment	**Forward primer**	**Reverse primer**
A	NotI-HNF4α123-F	HNF4α17-R
B	HNF4α17-F	HNF4α1278-XmaI-R
C	NotI-HNF4α123-F	HNF4α28-R
D	HNF4α28-F	HNF4α1278-XmaI-R
E	NotI-HNF4α123F	HNF4α39-XmaI-R
F	NotI-HNF4α789-F	HNF4α17-R
G	NotI-HNF4α789-F	HNF4α28-R
H	NotI-HNF4α789-F	HNF4α39-R