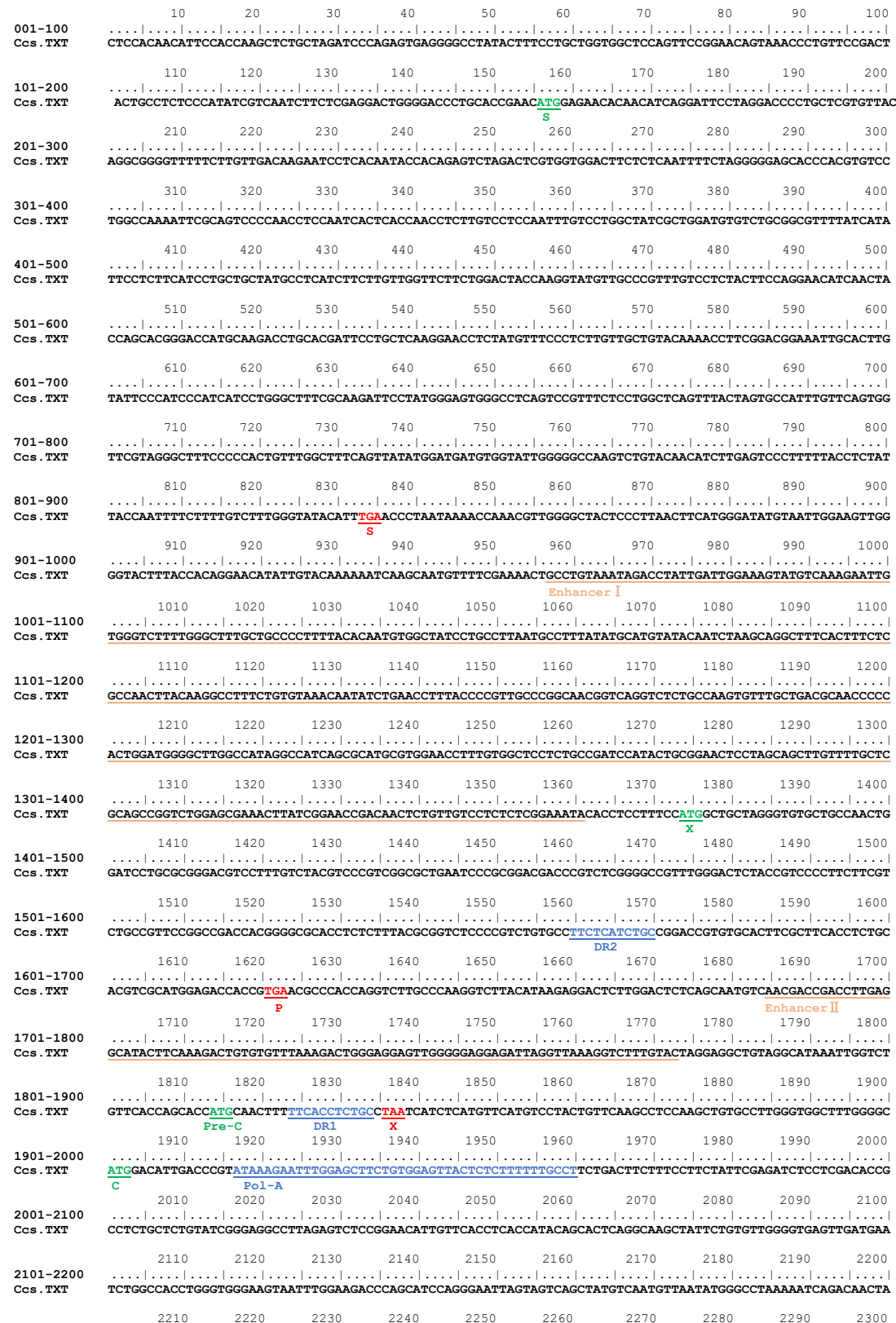


Supplementary Figure

Supplementary Figure S1



2201-2300
Ccs. TXT

2301-2400
Ccs. TXT

2401-2500
Ccs. TXT

2501-2600
Ccs. TXT

2601-2700
Ccs. TXT

2701-2800
Ccs. TXT

2801-2900
Ccs. TXT

2901-3000
Ccs. TXT

3001-3100
Ccs. TXT

3101-3200
Ccs. TXT

3201-3215
Ccs. TXT

Pre-S1

Pre-S2

Figure S1. The established consensus sequence of HBV genotype C. The full-length consensus genome is composed of 3215 nucleotides. Start codon was marked with green. Stop codon was marked with red. DR1, DR2, Pol-A were marked with blue. Two enhancers were marked with orange. Four promoters (core promoter, CP, nt1613-1849; PreS1 promoter, SP I, nt2718-2808; PreS2 promoter, SP II, nt2983-3210; X promoter, XP, nt1171-1361).

Supplementary Tables

Supplementary Table S1. The translated amino acid sequences of the four open reading frames (*P*, *preS/S*, *preCore/Core*, *X*) from the consensus sequence of genotype C.

Title	Gene locus	Amino Acid sequences
Polymerase (843 aa)	nt(2307..3215,1..1623)	1 MPLSYQHFRK LLLLDDEAGP LEEELPRLAD EGLNRRVAED LNLGNLNSVI PWTHKVG NFT 61 GLYSSTVPVF NPEWQTPSFP HIHLQEDIIN RCQQYVGPLT VNEKRRLKLI MPARFYPNLT 121 KYLPLDKGIK PYYPEHAVNH YFKTRHYLHT LWKAGILYKR ETTRSASFQ SPYSWEQELQ 181 HGRLVFQST RHGDESFCQ SSGILSRSPV GPCVRSQ LKQ SRLGLPQQG SLARGKSGRS 241 GSIRARVHPT TRRSFGVEPS GSGHIDNSAS STSSCLHQSA VRKTAYSHLS TSKRQSSSGH 301 AVELHNIPPS SARSQSEGPI LSCWWLQFRN SKPCSDYCLS HIVNLLEDWG PCTEHGEHNI 361 RIPRTPARVT GGVFLVDKNP HNTTESRLVV DFSQFSRGST HVSWPKF AVP NLQSLTNLLS 421 SNLSWLSLDV SAAFYHIPLH PAAMPHLLVG SSGLPYRVAR LSSTSRNIN YQHGTMQDLHD 481 SCSRNLVYSL LLLYKTFGRK LHLYSHPIIL GFRKIPMGVG LSPFLLAQFT SAICSVVRR 541 FPHCLAFSYM DDVVLGAKSV QHLESFTSI TNFLLSLGIH LNPNTKRWG YSLNFMGYVI 601 GSWGTLPEQH IVQKIKQCFR KLPVNRPIDW KVCQRIVGLL GFAAPFTQCG YPALMPLYAC 661 IQSKQAFTFS PTYKAF LCKQ YLNLYPVARQ RSLG LCVFAD ATPGWGLAI GHQMRGTFV 721 APLPIHTAEL LAACFARSRS GAKLIGTDNS VVLSRKYTSF PWLLGCAANW ILRGTSFVYV 781 PSALNPADDP SRGRLGLYRP LLRLPFRPTT GRTSLYAVSP SVP SHLPDRV HFASPLHVAW 841 RPP

Large S protein (400 aa)	nt(2848..3215,1..835)	1MGGWSSKPRQ GMCTNLSVPN PLGFFPDHQL DPAFGANSNN PDWDFNPNKD HWPEANQVGA 61 GAFGPGFTTP HGGLLGWSPQ AQGILTTPVA APPPASTNRQ SCRQPTPISP PLRDSHPQAM 121 QWNSTTFHQALDPRVRGLY FPAGGSSGT VNPVPTTASP ISSIFSRTGD PAPNMENTTS 181 GFLGPLLVQL AGFFLLTRIL TIPQSLDSWW TSLNFLGGAP TCPGQNSQSP TSNHSPSTSC 241 PICPGYRWMCLRRFIIFLFI LLLCLIFLLV LLDYQGMLPV CPLLPSTSTT STGPCKTCTI 301 PAQGTSMFPS CCCTKPSDGN CTCIPISSW AFARFLWEWA SVRFSWLSLL VPFVQWVFG 361 SPTVWLSVIW MMWYWGPSLY NILSPFLPL PIFFCLWVYI
middle S protein (281 aa)	nt(3205..3215,1..835)	1MQWNSTTFHQ ALLDPRVRGL YFPAGGSSGT TVNPVPTTAS PISSIFSRTG DPAPNMENTT 61 SGFLGPLLVQL QAGFFLLTRI LTIPQSLDSW WTSNLFLLGA PTCGQNSQS PTSNHSPTSC 121 PPICPGYRWM CLRRFIIFLF ILLCLIFLL VLLDYQGMLPV VCPLLPSTST TSTGPCKTCT 181 IPAQGTSMFP SCCCTKPSDG NCTCIPISS WAFARFLWEW ASVRFSWLSL VPFVQWVFG 241 LSPTVWLSVI WMMWYWGPSL YNILSPFLPL PIFFCLWVYI
S protein (226 aa)	nt(155..835)	1MENTTSGFLG PLLVLQAGFF LLTRILTIQ SLDSWWTSLN FLGGAPTCG QNSQSPTSNNH 61 SPTSCPPICP GYRWMCLRRF IIFLLILLCL LIFLLVLLDY QGMLPVCPLL PGTSTTSTGP 121 CKTCTIPAQG TSMFPSCCCT KPSDGNCTCI PIPSSWAFAR FLWEWASVRF SWLSLLVPFV 181 QWVGLSPTV WLSVIWMMWY WGPSLYNILS PFLPLPIFF CLWVYI
X protein (154 aa)	nt(1374..1838)	1MAARVCCQLD PARDVLCLRP VGAESRGRPV SGPFGLTLP SSSAVPADHG AHLRLRGLPV 61 CAFSSAGPCA LRFTSARRME TTVNAHQVLP KVLHKRTLGL SAMSTTDLEA YFKDCVFKDW 121 EELGFEIRLK VFLGGCRHK LVCSAPPCNF FTSA
precore/core protein (212 aa)	nt(1814..2452)	1MQLFHLCLII SCSCPTVQAS KLCGLWLGWM DIDPYKEFGA SVELLSFLPS DFFPSIRDLL 61 DTASALYREA LESPEHCSPH HTALRQAILC WGELMNLATW VGSNLEDPAS RELVVSYNV 121 NMGLKIRQLL WFHISCLTFG RETVLEYLVS FGVWIRTPPA YRPPNAPILS TLPETTIVRR 181 RGRSPRRRTP SPRRRRSQSP RRRRSQSRES QC
Core protein (183 aa)	nt(1901..2452)	1MDIDPYKEFG ASVELLSFLP SDFPISIRD LDTASALYRE ALESPEHCSP HHTALRQAIL 61 CWGELMNLAT VVGSNLEDP SRELVSYSYN VNMGLKIRQL LWFHISCLTF GRETVLEYLV 121 SFGVWIRTPP AYRPPNAPIL STLPETTIVR RGRSPRRRT PSRRRSQS PRRRRSQSRE 181 SQC

Supplementary Table 2

Positional frequency report of 138 HBV genotype C genomes.

The detailed data were provided in Excel.

Supplementary Table S3. Serum data of the mice recipients injected with pHBV1.3C.

Marker (Unit)	BALB/c mouse	Different kinetics of serum markers at the indicated days with pHBV1.3C hydrodynamic injection					
		1	4	7	14	28	42
HBsAg (IU/mL)	#1	12.78	173.95	190.00	1.24	*	*
	#2	11.67	85.83	131.17	0.85	0.15	*
	#3	4.65	39.40	68.36	0.14	/	/
	#4	11.72	65.56	111.10	58.08	19.18	4.7
	#5	3.98	40.61	86.23	/	/	/
	#6	21.32	/	/	/	/	/
HBeAg (COI)	#1	10.67	37.62	25.13	0.31	0.31	0.29
	#2	5.15	10.21	14.55	3.80	0.88	0.13
	#3	2.74	4.97	4.62	0.59	/	/
	#4	4.37	7.00	9.06	11.61	5.44	2.57
	#5	4.25	3.63	6.36	/	/	/
	#6	2.67	/	/	/	/	/
ALT (IU/mL)	#1	694	2	3	2	1	1
	#2	107	2	1	1	2	3
	#3	160	6	3	2	/	/
	#4	107	2	2	1	1	2
	#5	181	5	4	/	/	/
	#6	65	/	/	/	/	/
HBV DNA (IU/mL)	#1	211274.63	17900.70	43779.01	N	N	N
	#2	827022.20	44660.18	714040.90	1026.92	N	N
	#3	602078.80	42710.64	102043.20	24852.68	/	/
	#4	61631.79	37125.55	122997.11	221303.11	19422.89	2958.24
	#5	2718457.00	28256.99	328867.31	/	/	/
	#6	39420.21	/	/	/	/	/

The serum data of #4 recipient is provided in the Figure 3B of the manuscript.

/ There were no data collected.

N The serum was not enough for qPCR analysis of HBV DNA.

* The value was below the lower limit detection (0.05 IU/mL).

The cutoff of HBeAg ≥ 1.0 is positive.

The value of ALT > 50 IU/mL is positive.

Mouse #6 at day two, #5 at day seven, #3 at day 14 was euthanized to collect hepatic tissue sample, respectively.