



HAL
open science

Full-length 5'RACE identifies all major HBV transcripts in HBV-infected hepatocytes and patient serum

Bernd Stadelmayer, Audrey Diederichs, Fleur Chapus, Michel Rivoire, Gregory Neveu, Antoine Alam, Laurent Fraisse, Kara Carter, Barbara Testoni, Fabien Zoulim

► To cite this version:

Bernd Stadelmayer, Audrey Diederichs, Fleur Chapus, Michel Rivoire, Gregory Neveu, et al.. Full-length 5'RACE identifies all major HBV transcripts in HBV-infected hepatocytes and patient serum. *Journal of Hepatology*, 2020, 73, pp.40 - 51. 10.1016/j.jhep.2020.01.028 . hal-03490308

HAL Id: hal-03490308

<https://hal.science/hal-03490308>

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Full-length 5'RACE identifies all major HBV transcripts in HBV-infected**
2 **hepatocytes and patient serum**

3 Bernd Stadelmayer^{1,2}, Audrey Diederichs^{1,2}, Fleur Chapus^{1,2}, Michel Rivoire³,
4 Gregory Neveu⁴, Antoine Alam⁴, Laurent Fraisse⁴, Kara Carter⁴, Barbara Testoni^{1,2}
5 and Fabien Zoulim^{1,2,5}

6 ¹ INSERM U1052, CNRS UMR-5286, Cancer Research Center of Lyon (CRCL),
7 Lyon, 69008, France

8 ² University of Lyon, Université Claude-Bernard (UCBL), 69008 Lyon, France

9 ³ INSERM U1032, Centre Léon Bérard (CLB), 69008 Lyon, France

10 ⁴ Evotec, 1541 Avenue Marcel Mérieux, 69280 Marcy l'Etoile, France

11 ⁵ Hospices Civils de Lyon (HCL), 69002 Lyon, France

12

13 **Corresponding author**

14 Bernd Stadelmayer

15 INSERM U1052, CNRS UMR-5286, Cancer Research Center of Lyon (CRCL), Lyon,
16 69008, France

17 bernd.stadelmayer@inserm.fr

18

19 **Keywords**

20 HBV transcripts, HBV viral particles, serological HBV RNAs, Hepatitis B, HBV
21 transcripts

22 **Electronic word count manuscript: 6390**

23 **Number of figures: 7**

24

25 **Conflict of interest:**

26 This work was part of a collaborative research agreement between INSERM and
27 Evotec.

28

29 **Financial Support**

30 Evotec and ANRS grants to FZ and BT.

31

32 **Authors' contribution**

33 B.S. conceived the HBV full-length 5'RACE assay. B.S. performed 5'RACEs, Western
34 blots, qPCRs, ELISA, Iodixanol gradient sedimentations, infections of HepG2-NTCP
35 cells and PHHs, cloning of cDNAs, analyzed and interpreted the data. A.D. performed
36 infections of PHHs, qPCRs and analyzed the data. F.C. helped in performing cell
37 culture experiments. M.R. provided liver resections. FZ provided serum samples from
38 patients. G.N., A.A., L.F., K.C., B.T. and F.Z. interpreted the data and supervised the
39 work. B.S. wrote the manuscript with editing input from all authors.

40

41

42

43 **Abstract**

44 **Background and aims:** Covalently closed circular DNA (cccDNA) is the episomal
45 form of the Hepatitis B virus (HBV) genome that stably resides in the nucleus of
46 infected hepatocytes. cccDNA is the template for the transcription of six major viral
47 RNAs, i.e. preC- pg-, preS1/2-, S- and HBx-RNA. All viral transcripts share the same
48 3' end and are all to various degrees subsets of each other. Especially under infection
49 conditions, it has been difficult to study in depth the transcription of the different viral
50 transcripts. We thus wanted to develop a method with which we could easily detect the
51 full spectrum of viral RNAs in any lab.

52 **Methods:** We set-up the HBV full-length 5'RACE (rapid amplification of cDNA ends)
53 method with which we measured and characterized the full spectrum of viral RNAs in
54 cell culture and in chronically infected patients.

55 **Results:** In addition to canonical HBx transcripts coding for full-length X, we identified
56 shorter HBx transcripts potentially coding for short X proteins. We showed that
57 Interferon β treatment leads to a strong reduction of preC- and pg-RNAs but only has
58 a moderate effect on the other viral transcripts. We found pgRNA, one spliced pgRNA
59 variant and a variety of HBx transcripts associated with viral particles generated by
60 HepAD38 cells. The different HBx RNAs are both capped and uncapped. Lastly, we
61 identified 3 major categories of circulating RNA species in patients with chronic HBV
62 infection: pgRNA, spliced pgRNA variants and HBx.

63 **Conclusions:** The HBV full-length 5'RACE method should significantly contribute to
64 the understanding of HBV transcription during the course of infection and therapy and
65 may inform the development of novel therapies aimed at targeting cccDNA.

66 **Electronic word count abstract:** 274

67 **Lay Summary**

68 Especially under infection conditions, it has been difficult to study in depth the
69 transcription of the different HBV transcripts. This study introduces the HBV full-length
70 5'RACE method that can be used in any standard lab to discriminate all viral transcripts
71 in cell culture and in the serum of patients.

72

73 **Introduction**

74 Hepatitis B remains worldwide a major public health problem despite the availability of
75 prophylactic vaccination and anti-viral treatments [1]. The Hepatitis B virus (HBV)
76 persists as a covalently closed circular DNA (cccDNA) of approximately 3.2 kb that is
77 embedded into the chromatin of hepatocytes as an episomal entity [2]. The amount of
78 cccDNA and its transcriptional activity vary along the natural course of infection and
79 are a main determinant of viral persistence and reactivation. From 4 promoters, 6 major
80 viral RNAs are expressed: preC, pgRNA, preS1, preS2, S and HBx mRNAs (Fig.1A).
81 preC and pgRNA transcripts are characterized by a 100bp redundancy at their 3' ends
82 and consequently contain two epsilon stem loop structures at their 5' and 3' ends
83 (Fig.1A). Every viral RNA is translated into proteins that are essential for the viral life
84 cycle. In this regard, the pgRNA is particular in the sense that apart from encoding the
85 two viral proteins C (core) and P (polymerase), it serves as the template for the viral
86 DNA genome. Genomic viral DNA synthesis is mediated by the reverse transcriptase
87 activity of the viral P protein once pgRNA and the P protein are encapsidated into the
88 nucleocapsids formed by the C protein [3,4]. preC RNA codes for the HBe protein that
89 has a potential immunoregulatory function [5]. preS1/2 and S mRNAs encode the 3
90 viral surface proteins L, M and S, respectively, that build up the viral envelope. The

91 shortest of the 6 transcripts is the HBx transcript that gives rise to a 154 amino-acid
92 long cytoplasmic and nuclear X protein [6–10]. X directly interacts with a multitude of
93 host proteins explaining its diverse trans-activating effects in host cells [11–14].
94 The organization of the HBV genome is highly condensed and all transcripts are to
95 various degrees subsets of each other (Fig.1A). Indeed, all HBV transcripts share the
96 same 3' end and, thus, the HBx sequence constitutes the 3' end of every viral transcript
97 (Fig1A). Therefore, most HBV RNAs are indistinguishable by quantitative real time
98 PCR (qPCR). At present, only the larger viral RNAs can be differentiated by Northern
99 blotting during the course of infection. Most probably because of low expression levels,
100 HBx transcripts are undetectable by Northern blotting under infection conditions.
101 Accordingly, not much is known about HBx transcription during viral infection.
102 Moreover, there is a knowledge gap about the viral RNA species, per se, found in cell
103 culture, but also in viral particles and in chronically HBV infected patients.
104 To this end, we set up a strategy called HBV full-length 5'RACE with which we
105 characterized all major intra- and extracellular HBV RNAs during viral infection of
106 cultured hepatocytes. With the HBV full-length 5' RACE approach, we followed the
107 expression of the different viral RNAs over the course of infection and additionally
108 studied how they are modulated by various drug treatments. Furthermore, we identified
109 the viral RNAs that are associated with viral particles produced by HepAD38 cells.
110 Finally, we analyzed the full spectrum of circulating HBV RNAs in patients with chronic
111 HBV infection.

112

113 **Materials and methods**

114

115 **Cell lines, viral inoculum and infection conditions**

116 For infection purpose, HBV particles were concentrated from the supernatant of
117 HepAD38 (HBV genotype D) cells by filtering and PEG precipitation as described
118 previously [15]. The HepAD38 cell line was a kind gift of Dr C. Seeger (Fox Chase
119 Cancer Center, Philadelphia, USA) [16]. The HepG2-NTCP cell line was a kind gift of
120 Dr S. Urban (Heidelberg University, Germany) [17]. Primary human hepatocytes (PHH)
121 were isolated from surgically removed liver sections, cultured, and infected as
122 described previously [15]. HepG2-NTCP cells and PHHs were infected at a M.O.I. of
123 100 vg/cell in culture media supplemented with 4% PEG (polyethylene glycol 8000).

124

125 **Serum samples**

126 Serum samples were collected in the Lyon Biobank after written informed consent from
127 the patients and the protocol was approved by the Lyon institutional Ethic Committees
128 (#DC-2008-235, CPP 11/040 20110530 A11-168, CNIL n°1789480v0).

129

130 **RNA ligase-mediated 5'RACE**

131 RNAs were isolated using a guanidinium thiocyanate–phenol–chloroform extraction
132 protocol (TRI reagent Sigma; [18]). 5'RACE was essentially done as described in the
133 GeneRacer Kit manual (ThermoFisher Scientific) except Tobacco Acid
134 Pyrophosphatase was substituted by RNA 5' Pyrophosphohydrolase (New England
135 Biolabs) and SuperScript reverse transcriptase III by SuperScript reverse transcriptase
136 IV (ThermoFisher Scientific). Reverse transcription reaction was performed using 3'
137 HBV specific Gsp1 primer 5'-GGTGCGCAGACCAATTTATG-3'. For the 5'RACE PCR
138 reaction Platinum™ SuperFi™ DNA Polymerase (ThermoFisher Scientific),
139 GeneRacer 5' primer and HBV specific nested primer Gsp2 5'-
140 GTGCACACGGTCCGGCAGATG-3' were used (Fig.1B). 5'RACE PCR was run in a

141 iTOUCH 1000 BIORAD thermocycler using the following touch down PCR program:
142 Initial denaturation step 98°C 3min > 5x(98°C 10s; 72°C 3min) > 5x(98°C 10s; 70°C
143 3min) > 25x(98°C 10s; 64,4 20s; 72°C 3min) > 72°C 10min. pg and preC specific PCRs
144 were run using primer Gsp2.2 5'-GCTTCCCGATACAGAGCTGAGG-3' in combination
145 with either pg ACTGAAGGAGTAGAAAACTTTTTTACCTCTG or preC
146 ATGGACTGAAGGAGTAGAAAATAAATTGGTCTGCG specific primer in a iTOUCH
147 1000 BIORAD thermocycler using the following 2-step PCR program: Initial
148 denaturation step 98°C 3min > 34x(98°C 10s; 72°C 30s) > 72°C 10min.

149

150 **Cloning of 5'RACE amplicons**

151 5'RACE amplicons were A-tailed and cloned into the pGEM-T vector (Promega) using
152 standard procedures. Clones were verified by digestion and sequenced by Sanger
153 sequencing at Eurofins Genomics.

154

155 **Drug treatments of culture hepatocytes**

156 Interferon beta (PBL Assay Science) was used at 500 IU/ml, preS1-myr (Myrcludex-
157 like) peptide (sequence: GTNLSVPNPLGFFPDHQLDPAFRANSNNPDWDFNPN
158 KDHWPEANKVG, synthesized by GeneScript) was used at 500nM, Tenofovir (Sigma
159 Aldrich) was used at 100µM. Drug regimens are detailed in the corresponding figures.

160

161 **ELISA for viral antigens**

162 Prior to ELISAs supernatants were 10 times diluted. ELISAs for HBeAg detection in
163 cell supernatants were performed according to the manufacturer's protocol using the
164 CLIA kit from Autobio Diagnostic.

165

166 **Protein isolation and Western blotting**

167 Whole cell proteins were isolated by a first lysis with buffer A (20mM Tris-HCL pH7.5;
168 0.5mM EDTA; 0.1% TritonX100; 100mM NaCl; 10% Glycerol; 2mM MgCl₂; 10mM
169 mercaptoethanol; 0.5 mM PMSF (phenylmethylsulfonyl fluoride)) for 30 min at 4°C.
170 Subsequently, the lysate was centrifuged for 5 min at 4000rpm using a tabletop
171 centrifuge. Supernatant was transferred to a fresh tube and the pellet was incubated
172 with 5 volumes of buffer B (20mM Tris-HCL pH7.5; 0.5mM EDTA; 0.1% TritonX100;
173 400mM NaCl; 10% Glycerol; 2mM MgCl₂; 10mM mercaptoethanol; 0.5 mM PMSF) for
174 30 min on 4°C and then centrifuged for 10 min at 13000 rpm. Both supernatants were
175 pooled. 1 OD per condition determined by Bradford assay was diluted in Laemmli-
176 buffer and resolved on an SDS-PAGE gel.

177 For the detection of surface proteins in gradient fractions, 10µl of fractions were directly
178 diluted in Laemmli buffer and resolved on an SDS-PAGE gel. Proteins were
179 electroblotted onto a nitrocellulose membrane using the Transblot system (Biorad).
180 Surface proteins L/M/S were detected by the monoclonal antibody H166 [19], a kind
181 gift of Abbott Laboratories.

182

183 **Iodixanol gradient sedimentation analysis**

184 To analyze viral particles, 200ml of HepAD38 cell supernatant was centrifuged for 5min
185 at 500g to sediment cell debris. This was followed by an additional centrifugation step
186 for 45min at 2000g. Supernatant was laid over a 20% sucrose cushion and centrifuged
187 at 90000g for 2h using a Beckmann 32Ti rotor to pellet viral particles. Pellet was
188 resuspended in 1ml PBS and material was loaded onto a 10 ml 10% to 50% Iodixanol
189 gradient in PBS and centrifuged at 175000g for 14h in a Beckmann SW41 rotor. 900µl
190 fractions were collected from the top of the gradient.

191

192 **MNase digestion of HepAD38 supernatants**

193 To isolate encapsidated viral RNAs we performed MNase (Micrococcal Nuclease)
194 digestions of the HepAD38 supernatants. MNase is a DNA and RNA endonuclease
195 that degrades double- and single-stranded DNA and RNA. 200ml of HepAD38
196 supernatant was centrifuged for 5min at 500g to sediment cell debris. This was
197 followed by an additional centrifugation step for 45min at 2000g. The supernatant was
198 laid over a sucrose cushion and centrifuged at 90000g for 2h using a Beckmann 32Ti
199 rotor to pellet viral particles. Viral particles were resuspended in MNase buffer, 2 µg of
200 a non-relevant plasmid (pcDNA3 vector + insert) was added to monitor digestion
201 (Suppl.Fig.5). MNase digestion was performed for 25min at 37°C using 30 gel units
202 (New England Biolabs) in a 150µl reaction. Reactions were stopped by Trizol.

203

204 **Real-time PCRs**

205 RNA or DNA was extracted using the MasterPure complete kit (Epicentre) following
206 manufacturer's instructions. Real-time PCRs were performed essentially as described
207 in Lebossé et al. [20]. 3kb RNAs were detected by the primer pair
208 ggagtgtggattcgcactcct (A) and agattgagatcttctcgcac (B) together with the Taqman
209 hybridization probe ([6FAM]-AGGCAGGTCCCCTAGAAGAAGAACTCC-[BHQ1]) [21].
210 Total RNAs were measured using primers and Taqman probe Pa03453406_s1
211 (ThermoFisher Scientific). Signals were normalized by the housekeeping gene GUSB
212 detected with primers and Taqman probe Hs99999908_m1 (ThermoFisher Scientific).
213 Total DNA was measured using primers and Taqman probe Pa03453406_s1 (Life
214 Technologies). Serial dilutions of a plasmid that contains a HBV monomer
215 (pHBVEcoR1) served as an external quantification standard. The amount of total DNA

216 was normalized to the cell number using primers and Taqman probe Hs00758889_s1
217 (ThermoFisher Scientific) that are specific for the β -globin locus. PCRs were performed
218 in duplicates. Amplification was carried out on the QuantStudio 7 Flex with the following
219 PCR conditions: Initial denaturation step 95°C 20s > 40x(95°C 1s; 60°C 20s).

220

221 **Accession code for sequencing data**

222 All Sanger sequencing data is available at the European Nucleotide Archive (ENA) at
223 <http://www.ebi.ac.uk/ena/data/view/PRJEB36101> (accession code PRJEB36101).

224

225 **Results**

226

227 **Detection of all major viral RNAs by HBV full length 5'RACE**

228 In the context of the HBV genome, transcription of 5' positioned genes obscures the
229 measurement of transcripts starting further downstream because all transcripts share
230 the same 3' ends (Fig.1A). Currently, the measurement of viral RNAs relies on 3
231 techniques: Northern blotting, quantitative RT PCR and RNA-next generation
232 sequencing (NGS). Northern blotting allows the resolution of viral RNAs according to
233 their molecular weights on an agarose gel but misses the HBx transcript under infection
234 conditions [22]. Quantitative RT PCR cannot distinguish between the different viral
235 transcripts and hence it is commonly used exclusively for measuring total RNAs vs 3.5
236 kb RNAs. Similarly, standard short reads RNA-NGS approaches cannot distinguish the
237 different viral transcripts as soon as more than one viral transcript are present. To our
238 knowledge, there is only one publication in which the authors tried to discriminate all
239 the different viral transcripts expressed in cell culture and in patients by a strategy
240 called cap analysis of gene expression (CAGE) [23]. However, the major limitation of

241 this study was that only the first 50 nucleotides of the viral transcripts were sequenced
242 and thus no conclusions could be drawn about viral full-length transcripts and
243 especially pgRNA splice variants.

244 Thus, to specifically detect and characterize the different intra- and extracellular HBV
245 RNAs in full length we set up the HBV full-length 5'RACE strategy. The technique
246 follows a standard RNA ligase-mediated 5'RACE protocol including initial
247 dephosphorylation and decapping steps to ensure the amplification of only full-length
248 transcripts with respect to 5' ends. Truncated, uncapped RNA species are eliminated
249 from the subsequent 5'RACE amplification process because they miss the 5'
250 phosphate group that is essential for anchor ligation (Fig.1B). After performing the HBV
251 full-length 5'RACE, the different HBV transcripts are resolved on an agarose gel
252 according to their molecular weights.

253 The initial experiment was carried out in HepG2-NTCP cells that had been infected for
254 8 days. As shown in Fig.1C we found all major transcripts including HBx migrating at
255 their expected sizes on the gel. Their relative sizes are reduced in length by
256 approximately 338 bp (base pairs) with respect to template viral RNAs because of the
257 use of gene specific 3' primer 2 (Gsp2) for PCR. PCR products are specific to viral
258 RNAs as shown by performing the 5'RACE without the reverse transcription reaction
259 (Suppl.Fig.1A) and by their absence in the non-infected condition (Fig.1C). In addition,
260 PCR products are specific to capped viral RNAs as demonstrated in Suppl.Fig.1C and
261 further below. Of note, the proportional quantities of the different viral RNAs are not
262 directly translated by the 5'RACE PCR reaction. This is because the viral transcripts
263 exhibit an important heterogeneity in their respective lengths and, hence, are subject
264 to different amplification efficiencies during the 5'RACE PCR reaction, i.e., gel band

265 intensities of the 5'RACE PCR product do not provide an absolute information on the
266 original proportional quantities of the different viral RNA species.

267 We then investigated the expression kinetics of the different viral RNA species during
268 the course of infection. In both HepG2-NTCP cells and PHHs we found HBx to be the
269 first transcript detectable by full-length 5'RACE as early as 8h post-infection (Fig.1D).
270 The remaining transcripts are detected later during infection, in HepG2-NTCP cells
271 from 24 hours on and in PHHs after a longer delay (not earlier than 48h, see additional
272 experiment Suppl.Fig.2). The variance in the kinetics of viral RNAs observed for the
273 two experiments shown for PHHs (Fig.1D and Suppl.Fig.2) probably resulted from the
274 heterogeneity of cells retrieved from two different donors. However, these results are
275 similar to the data published in Niu et al. where the authors followed viral RNAs in
276 PHHs during the course of infection by an RNA-NGS approach [22].

277

278 **Short versions of HBx transcripts are produced in HepG2-NTCP cells and PHHs**

279 To confirm their identity, the obtained amplicons were fully characterized by subcloning
280 and Sanger sequencing (Fig.1E and PRJEB36101). Sequencing of the 5'RACE
281 amplicons revealed an important diversity in transcript start sites (TSSs) for the HBx
282 transcripts expressed in HepG2-NTCP cells and PHHs (red bars in Fig.1E). In contrast,
283 all sequenced clones for the preS2 transcripts were identical and share the same 5'
284 end in PHHs and HepG2-NTCP cells (TSS nucleotide (nt) 3159 with respect to
285 reference sequence GenBank U95551.1, blue bars in Fig.1E). The same was true for
286 pgRNA (TSS nt1820) and preC (TSS nt1791). Surprisingly, we were not able to detect
287 spliced pgRNA variants, most likely because of their low proportional copy numbers
288 with respect to the other viral RNA species in HepG2-NTCP cells and PHHs.

289 In HepG2-NTCP cells, HBx transcripts can be subdivided into two categories: i.
290 “canonical” transcripts initiating closely upstream (TSSs nt1243 to nt1338) of the first
291 ATG of HBx (light red) coding for the full-length X protein (154 amino acids) and ii.
292 transcripts starting after the first ATG of HBx (dark red) presumably coding for a short
293 version of X potentially using 2 ATGs further downstream as start codons
294 (Met79/Met103). In PHHs, we additionally sequenced a third category of longer HBx
295 transcripts (TSS nt1065 and nt1151, very light red Fig.1E) that we did not detect in
296 HepG2-NTCP cells. The longer transcripts include ATGs from the P protein open
297 reading frame and could potentially code for its RNaseH subdomain. The number of
298 sequenced clones positive for shorter and longer HBx transcripts was much smaller as
299 compared to clones harboring canonical HBx (see Fig.1E), indicating that shorter and
300 longer HBx transcripts are most likely less abundant in cells.

301

302 **Tenofovir treatment primarily increases the pgRNA transcript in PHHs and** 303 **HepG2-NTCP cells**

304 Subsequently, to further validate our approach, we investigated the effect of treatment
305 interventions on viral transcripts in HBV infected HepG2-NTCP cells and PHHs using
306 the full-length 5'RACE assay. We first exposed the cells to a Myrcludex-like peptide
307 (hereafter Myrcludex) and Tenofovir, two functionally well-characterized drugs
308 interfering with the viral life cycle at different stages. Myrcludex is a viral entry inhibitor
309 peptide binding to the HBV receptor sodium tauro-cholate co-transporting polypeptide
310 (NTCP) and thus efficiently prevents viral infection [17,24]. Tenofovir, on the other
311 hand, is a nucleotide analog (NA) that interferes with DNA chain elongation during the
312 reverse transcription reaction and, accordingly, inhibits viral genome replication [25].
313 For HepG2-NTCP experiments, cytotoxicity of drug regimens was excluded by visual

314 inspection. For experiments in PHHs, effect of drug regimens on cell viability was
315 measured by the CellTiter-Glo approach (Suppl.Fig.3). As shown in Fig2.A we
316 measured viral RNAs by 5'RACE on 4, 8 and 15 days post infection. Myrcludex
317 treatment was initiated 1 day before infection to saturate NTCP receptors, Tenofovir
318 treatment was started 3 days post infection (see timescale in Fig.2A). After 8 days of
319 infection, conventional qPCRs were performed to monitor their effect on 3kb RNAs (pg-
320 and preC-RNA), total RNAs and total DNA (Fig.2B). In the Myrcludex condition, no viral
321 RNAs and DNAs were detectable by qPCRs. Tenofovir treatment, on the other hand,
322 led to a ~1.5 fold increase in 3kb RNAs and total RNAs and to a strong decrease in
323 viral DNAs.

324 Concordantly, full length 5'RACEs showed no specific signal when Myrcludex blocks
325 viral entry (Fig.2A upper gel). The bands visible in the Myrcludex lane 8 days post
326 infection are non-specific, as verified by cloning and sequencing (data not shown).
327 However, Tenofovir treatment lead to an increase in 3kb 5'RACE signals indicating
328 more preC/pgRNA as compared to the non-treated control conditions (Fig.2A lower
329 gel). Tenofovir administration had no effect on the S and HBx transcripts, especially
330 from day 8 on. Since the experiment is exclusively controlled by using equal amounts
331 of input RNA for the different conditions, we normalized the quantity of the 3kb band
332 (preC/pgRNA) in each lane by its corresponding HBx band. This procedure produces
333 the 3kb ratio (intensity 3kb band/intensity HBx band) and shows that, for each time-
334 point, the 3kb ratios are elevated, i.e. more preC/pgRNAs are found in Tenofovir
335 treated as compared to untreated cells (Fig.2C). Moreover, we observed an increase
336 of the 3kb ratio over time supporting the concept that viral full-length transcripts
337 accumulate with Tenofovir treatment duration in HepG2-NTCP cells.

338 Comparable to the experiments in HepG2-NTCP cells, no 5'RACE signals are detected
339 in PHHs treated with Myrcludex (Fig.3A/B). In PHHs too, Tenofovir treatment causes
340 an exclusive increase in the 3kb band as compared to the other 5'RACE products (i.e.
341 S transcripts and HBx transcripts) over the time of infection (Fig.3A/B/C). Next, in order
342 to elucidate whether the increase in 3kb transcripts reflects an augmentation of pgRNA,
343 preC-RNA or both, we designed primers that are specific for pg- or preC-RNA using
344 full-length 5'RACE cDNAs (for specificity and outline of strategy see Suppl.Fig.4).
345 pg/preC-RNA specific 5'RACE PCRs indicate that the effect of Tenofovir is much more
346 pronounced for pgRNA than for preC RNA on day 8 post infection, an observation that
347 is consistent with the mode of action of Tenofovir as a reverse transcription inhibitor
348 (Fig.3A gels pg and preC). To support the results, we quantified HBeAg by ELISA as
349 another read out for preC using the supernatants of treated and untreated cells.
350 Consistently, by ELISA we did not detect any effect of Tenofovir on extracellular HBeAg
351 (Fig.3D). Furthermore, we show by Western blotting that cellular S protein levels are
352 not affected by Tenofovir administration (Fig.3E). In summary, these results validated
353 the HBV full length 5'RACE technique and further confirmed that Tenofovir treatment
354 primarily increases pgRNA but has no effect on preC-RNA, S transcripts and HBx RNA.

355

356 **Interferon β treatment reduces both pg- and preC-transcripts without affecting**
357 **other viral transcripts in PHHs**

358 Patients with chronic Hepatitis B are treated with Interferon alpha and nucleos(t)ides
359 analogues leading to the suppression of viral replication [25]. In HepG2 or chicken
360 hepatoma cell lines that express viral RNAs from transfected or integrated plasmids,
361 Interferon alpha efficiently down-regulates viral RNA transcription by reducing active
362 chromatin marks like H3K9ac and H3K27ac from HBV and DHBV chromatin [26,27].

363 We wanted to take advantage of the full-length 5'RACE and investigate how Interferon
364 β affects the different viral RNAs under infection conditions in PHHs. Interferon β is a
365 type I interferon which utilizes the same receptor as IFN alpha for which it has been
366 shown that it inhibits HBV genome replication in hepatocyte culture [28,29]. Interferon
367 β treatment was started on day 3 post infection, 5'RACEs (Fig.4A) and control qPCRs
368 (Fig.4B) were carried out on day 8 post infection. Surprisingly, we found that Interferon
369 β treatment predominantly reduces viral 3kb transcripts i.e. preC- and pgRNA without
370 significantly affecting the other viral transcripts. This prominent effect is observed by
371 full-length 5'RACE and preC/pg-RNA specific 5'RACE PCRs (Fig.4A).
372 To complement the 5'RACE analysis, we show by Western blotting that the S protein
373 levels are only moderately affected in the Interferon condition (Fig.3E). In contrast,
374 extracellular HBeAg is strongly reduced, as determined by ELISA (Fig.3D). We omitted
375 measuring secreted HBsAg in supernatants by ELISA since it had been shown that
376 Interferons inhibit HBV virion secretion by inducing tetherin expression [30].

377

378 **Viral particles produced in HepAD38 cells contain different HBx transcripts, at**
379 **least one spliced pgRNA variant and pgRNA**

380 Recently it was shown that viral particles produced in cell culture or circulating in the
381 blood of chronically infected patients may harbor viral RNA species [22,31,32]. To gain
382 more insight into the type of viral RNAs that may be contained in these particles, we
383 performed 5'RACEs with material derived from supernatants of HepAD38 cells (See
384 Fig.5A for purification steps). We focused our work on viral particles produced by the
385 HepAD38 cell line since HepG2-NTCP cells and PHHs in infection assays (Fig.1-Fig.4)
386 were infected with viral inocula generated by the same cell line.

387 Prior to full-length 5'RACE analysis, the pelleted material was digested with MNase to
388 detect exclusively protected RNAs (see Suppl.Fig.5 for digestion efficacies). As shown
389 in Fig.5B/C in the supernatant of HepAD38 cells, we found pgRNA, one pgRNA spliced
390 variant (known as splice variant 6 (SP6) [33], splice donor site nt2471/nt2472; splice
391 acceptor site nt488/nt499, Fig.5C) and HBx transcripts with different lengths (long,
392 canonical, and short) that are protected from MNase digestion. For their identification,
393 we characterized all RNAs that were protected from MNase degradation (Fig.5B) by
394 cloning and Sanger sequencing (Fig.5C and PRJEB36101). Neither S transcripts nor
395 preC-RNA were detected by sequencing. Furthermore, by pgRNA and preC RNA
396 specific 5'RACE analysis, we confirmed that protected RNAs did not include preC RNA
397 (Fig.5B). Importantly, the 5'RACE products were specific to viral RNAs since there was
398 no signal in the negative control reactions without RT.

399 To confirm that the MNase-protected RNAs are contained in viral particles, we
400 performed Iodixanol gradient sediment analysis. Overall, 12 fractions were collected
401 from the top out of which fractions 2 to 11 were characterized in detail with respect to
402 viral DNA (Fig.5D), S antigens (HBsAg, Fig.5E) and viral 5'RACE RNA content
403 (Fig.5F). Viral DNA sedimented between fractions 4 and 8 with a sharp peak in fraction
404 7. All three HBsAg co-sedimented between fractions 5 and 9 with the strongest
405 intensities found in fractions 6 and 7.

406 By full-length 5'RACE we could detect the different HBx transcripts similar to the
407 MNase approach. The majority of HBx transcripts was detected in fractions 7 and 8
408 co-fractionating with viral DNA and HBsAg. Most probably because of RNA quantity
409 issues, full-length 5'RACE neither detected pgRNA nor the pgRNA splice variant found
410 by the MNase approach. However, by performing pgRNA specific 5'RACE analysis

411 using the pgRNA specific primer, pgRNA was detected in gradient fractions 6 to 9 with
412 the highest abundance in fraction 7 similar to viral DNA.

413 Taken together, we found pgRNA, one pgRNA derived splice variant (SP6) and
414 different HBx transcripts co-fractionating with HBAGs in gradient fractions where viral
415 DNA is present (Fig.5C). HBx transcripts are of different lengths: long (TSSs nt1065 to
416 nt1198, canonical (TSSs nt1243 to nt1338bp) and short (TSSs nt1418bp to nt1533).

417

418 **HBx RNAs in viral particles are both capped and un-capped**

419 Capped RNAs are protected against de-phosphorylation, the first step of the full-
420 length-5'RACE protocol that is necessary to exclude degraded or un-capped RNAs
421 from the full-length 5'RACE (Fig.1B). The interpretation of our data, i.e. that the
422 different RNAs detected by the 5'RACE in viral particles are indeed capped, relies on
423 the efficiency of this de-phosphorylation step. An inefficient de-phosphorylation step
424 could leave un-capped RNAs phosphorylated and thus could produce false positive
425 signals during the full-length 5'RACE. However, un-capped/degraded viral RNAs that
426 are generated by the RNase H activity of the P protein during cDNA synthesis must
427 be present in viral particles, as discussed by Bai et al. [31].

428 To clarify these issues, we carried out control 5'RACE experiments that specifically
429 detect un-capped RNAs using again MNase digested material from HepAD38 cells
430 derived supernatants. The un-capped RNA specific 5'RACE protocol lacks the de-
431 capping and de-phosphorylation steps from the standard full-length 5'RACE and thus
432 capped RNAs are not ligated to the anchor sequence and consequently are not
433 amplified by the final PCR reaction (Fig.6A). To control de-phosphorylation efficacy,
434 we included or not a de-phosphorylation step in the un-capped RNA specific 5'RACE
435 protocol (compare lane 4 with lane 5 in Fig.6B). Simultaneously, we performed the

436 standard full-length 5'RACE (capped RNA specific) on the same material (Fig.6B lane
437 3) and on cellular RNAs derived from PHHs (Fig.6B lane 1).

438 Collectively, the data in Fig.6B demonstrates the authenticity of the capped RNAs
439 found in viral particles since all uncapped RNAs contained in viral particles are
440 dephosphorylated during the capped RNA specific full-length 5'RACE approach
441 (compare lane 4 with lane 5 in Fig.6B). However, we also confirm that un-capped RNA
442 species up to 800bp exist in the MNase digested material (Fig.6B lane 5). Sub-cloning
443 and sequencing identified the uncapped RNAs as long, canonical and short HBx
444 transcripts (data not shown). We detected neither pgRNA nor the pgRNA splice variant
445 as uncapped RNAs (Fig.6B lane 5).

446

447 **Full length 5'RACE identifies 3 major categories of circulating RNA species in** 448 **patients with chronic HBV infection**

449 An accurate monitoring of intrahepatic cccDNA levels and activity during patient
450 management is limited by the need for invasive liver biopsy procedures. So far, no
451 single serum parameter has been shown to accurately reflect the transcriptional activity
452 of the cccDNA pool in the liver [34]. Moreover, it is still an open question, which
453 serological viral RNA species exist in the plasma/serum of patients. Indeed, pgRNA
454 has been detected in sera of patients, spliced pgRNA variants and HBx transcripts as
455 well, but there is not much information about whether and how these transcripts co-
456 exist [22,32,35–38]. To gain deeper insight into these questions, we characterized the
457 major circulating viral RNA species found in three highly viremic chronically infected
458 patients (Fig.7A) by full-length 5'RACE (Fig.7B). All three patients were HBe-positive
459 and patient #2 had elevated alanine aminotransferase (ALT) levels. Figure 7B shows
460 the different 5'RACE profiles for the three patients and visual inspection revealed a

461 differential pattern between the patients. Importantly, for all patients 5'RACE signals
462 were specific to RNAs (Fig.7B/-RT).

463 To characterize the RNA species found in plasma, we sub-cloned and sequenced the
464 5'RACE products of patient 1 and 2 (Fig.7C and PRJEB36101). We did not detect any
465 S and preC transcripts in the plasma samples. For both patients 1 and 2, we found co-
466 circulating pgRNA and pgRNA splice variants (known as SP1/2/3/5 [33], see Fig.7C
467 for splice donor/acceptor sites). We again identified an important heterogeneity for the
468 HBx transcripts including long (TSSs nt991-nt1197), canonical (TSSs nt1278-nt1350)
469 and short (TSSs nt1397-nt1558) HBx transcripts similar to the HBx transcripts
470 associated with *in vitro* produced viral particles. Interestingly, the pgRNA splice
471 variants observed in patients #1 and #2 were different (see Fig.7C for alignments of
472 transcripts in individual patients). In summary, HBV RNAs found in plasma samples of
473 patients belong to one of the three categories: 1. pgRNA, 2. pgRNA splice variant, 3.
474 HBx transcript (long, canonical and short HBx transcripts). Considering the sequencing
475 data together with the patient specific 5'RACE profiles, the results presented here open
476 up the possibility that patients exhibit qualitative and/or quantitative differences in
477 circulating HBV RNAs.

478

479 **Conclusions**

480 We set up the HBV full-length 5'RACE approach to overcome the current limitations
481 for the measurement and discrimination of the different HBV RNAs. Our results show
482 that the HBV full-length 5'RACE detects all major viral RNAs including pgRNA splicing
483 variants and most importantly, HBx. The full-length 5'RACE is a qualitative approach
484 though sensitive enough to demonstrate changes during the course of infection (Fig.1)

485 and in experimental conditions where cultured hepatocytes were treated with Tenofovir
486 (Fig.2A/3A) or Interferon β (Fig.4A).

487 Contrary to preC/pg-RNA and the S transcripts, we found a variety of transcripts for
488 HBx that are characterized by heterogeneous TSSs (Fig.1D). Of special interest are
489 the long and short HBx transcripts, that were found in cells, viral particles produced by
490 cultured hepatocytes but also in plasmas of patients, indicating their biological
491 relevance. Some of the long and short HBx transcripts could potentially code for other
492 proteins than full-length canonical X. Shorter than canonical HBx transcripts and short
493 X proteins have been discussed in the literature [23,39,40]. Short X starting from Met
494 103 of the canonical X would lack the DDB1-binding domain of the full-length protein
495 [41] and thus would have lost its co-regulatory activity of the Cul4A E3 ubiquitin ligase
496 complex [12,41,42]. In addition, the C-terminal domain of full-length X is required for
497 p53 interaction [13] and consequently, short X could be a protein that predominantly
498 interacts with p53.

499 On the other hand, long HBx transcripts have not yet been described. We could detect
500 them in plasma, viral particles produced by the HepAD38 cell line and in cell extracts
501 from PHHs, but surprisingly not in HepG2-NTCP cells. Long HBx transcripts could code
502 for the RNaseH domain of the P protein since the transcripts include ATGs upstream
503 of HBx. Whether the long transcripts code for RNaseH and which function the RNaseH
504 domain separated from its RT domain has, needs further investigation. Interestingly, it
505 has been shown that the HBV RNaseH domain is a stable protein without the RT
506 domain [43–46], a characteristic that is not true for all the RNaseH domains of other
507 RTs [47,48].

508 Besides coding for proteins, there is also the possibility that the short and long HBx
509 RNAs function as regulatory non-coding RNAs (ncRNAs). Many DNA and RNA viruses

510 express ncRNAs that target either host or viral factors [49]. Especially long non-coding
511 RNAs (lncRNAs) are an interesting class of ncRNAs that regulate host or viral gene
512 expression at the level of chromatin, transcription, post-transcription and translation.
513 Interestingly, for the Kaposi's sarcoma-associated herpesvirus (KSHV), Japanese
514 encephalitis virus (JEV), dengue virus, and West Nile Virus (WNV) it has been
515 shown that lncRNAs antagonize the antiviral Interferon response [50–55].
516 The capacity of the full-length 5'RACE to discriminate all the major viral transcripts
517 allowed us to show that viral RNAs can be differentially regulated. We demonstrated
518 that Interferon β reduced preC-RNA and pgRNA without affecting the other viral
519 transcripts (Fig.4A). In overexpression assays [26,27,56] and mouse model infection
520 systems [57], this differential phenotype could not be differentiated from a general
521 down-regulation of viral transcripts. Moreover, in combination with the qPCR assays
522 (Fig.4B), the full length 5'RACE data also indicates that preC- and pgRNAs constitute
523 the majority of viral transcripts in infected PHHs confirming CAGE-NGS results recently
524 published by Altinel et al. [23].
525 It will be important to elucidate how Interferon β induces the specific down-regulation
526 of preC/pg-RNAs. One possible scenario could be that Interferon β treatment induces
527 the specific degradation of the full-length viral transcripts in PHHs. Interestingly,
528 Interferon stimulated gene 20 (ISG20) degrades HBV viral RNAs by binding to the
529 epsilon stem-loop structure [58]. If the targeting/degradation activity of ISG20 in PHHs
530 would be specific for the epsilon structure at the 5' end of an RNA molecule, only full-
531 length viral RNAs (preC/pg-RNAs) would be degraded. However, Liu et al.
532 demonstrated by over expression assays that ISG20 degrades viral RNAs irrespective
533 whether the epsilon structure is at the 5' or 3' end of the transcripts. The authors
534 additionally demonstrated that ISG20 does not act alone but requires co-factors for its

535 target specificity. Thus, either another ISG displaying different targeting characteristics
536 is induced upon Interferon β treatment or, the specificity of ISG20 is modulated by
537 different co-factors in PHHs.

538 Another possibility could be that Interferon β specifically reduces the activity of the HBV
539 core promoter regulating preC/pg-RNAs expression. It has been shown that the core
540 HBV promoter is negatively regulated by the NF- κ B factor p65 [59,60], another rapidly
541 induced ISG, which could preferentially down-regulate the HBV core promoter once
542 activated by Interferon β . Alternatively, the ISG TRIM22 can suppress HBV core
543 promoter activity in HepG2 cells. TRIM22 is a member of the tripartite motif (TRIM)
544 superfamily whose members have been shown to be expressed in response to
545 Interferons to restrict viral activity, especially with respect to retroviruses [61].

546 We also characterized the full spectrum of viral RNAs associated with viral particles
547 and gained a detailed overview of their composition (Fig.5/6). Our results integrate
548 earlier published results, but at the same time provide new insights into the diversity of
549 particle associated viral RNAs. For instance, in addition to pgRNA we also detect
550 spliced pgRNA variants in viral particles. Spliced pgRNA variants have been first
551 described in cell lines and infected livers [62–65] and later have been shown to be
552 incorporated into the nucleocapsids and then reverse transcribed into HBV DNA to
553 generate defective HBV particles [33,66–69]. Interestingly, it also has been shown that
554 pgRNA splice variants can code for additional viral proteins [70,71]. Furthermore, our
555 results substantiate the notion that viral particles contain HBx RNA that may have an
556 important role for the establishment of infection as proposed by Niu et al. [22]. In this
557 regard, we found that different HBx RNAs exist as capped and un-capped molecules
558 in viral particles, an observation that raises the question whether one is the product of
559 the other (Fig.6B). As depicted in Figure 1A, the six viral RNAs are transcribed from

560 four different promoters. After productive initiation of Polymerase II, nascent transcripts
561 are co-transcriptionally capped in the nucleus by different capping enzymes. Capping
562 of transcripts promotes translation and protects against RNA degradation by cellular
563 exonucleases. The capped non-canonical HBx RNAs are puzzling with respect to their
564 5'ends since, to our knowledge, promoters that initiate transcription from these
565 positions in the HBV genome have not yet been molecularly characterized. In addition,
566 transcripts that lack the 5' epsilon stem-loop structure should be excluded from viral
567 capsids [72–75]. Therefore, the question arises whether the capped HBx transcripts in
568 viral particles could be produced rather post- than co-transcriptionally. The hypothetical
569 post-transcriptional generation of encapsidated HBx RNAs would require 2 major
570 steps: i. an endolytic cut to generate the 5'end of the RNA molecule and ii. the addition
571 of the cap structure. A scenario would be that the RNase H domain of the P protein
572 produces specific endolytic cuts in the pgRNA during reverse transcription in viral
573 particles giving rise to the 5' ends of the different HBx RNAs. Interestingly, the RNase
574 H domain of the avian myeloblastosis virus (AMV) reverse transcriptase e.g. initially
575 cuts every 100-200 nucleotides during reverse transcription. Whereas for HIV-1 and
576 MoMLV, the cleavage frequency centers around 100–120 nucleotides [48,76].
577 Consequently, it will be important to determine the cutting frequency of the HBV RNase
578 H domain and whether it correlates with the molecular sizes of the different HBx
579 transcripts (400-800nts). Furthermore, the post-transcriptional generation of capped
580 non-canonical HBx RNAs would also require that the capping of un-capped 5' ends
581 takes place inside viral particles. Since HBV does not code for its own capping
582 enzymes, viral particles would have to contain host-capping enzymes that fulfill this
583 task. Notably, for RNA alphaviruses it has been shown that viral particles that contain
584 non-capped viral RNAs induce innate immune responses in host cells and therefore

585 are less infectious than viral particles carrying capped RNAs [77]. It is intriguing to
586 speculate that HBV might have found a way to cap the different RNAs arising during
587 the maturation of viral particles to hide them from the innate immune system of host
588 cells. Importantly, it is very likely that other pgRNA-derived RNA species are generated
589 during the reverse transcription process in addition to the different HBx RNAs
590 characterized in this study. However, the HBV full-length 5'RACE method is specific to
591 RNA molecules that are intact until nt1810 (Gsp1 annealing point) of the HBV genome
592 (reference sequence GenBank U95551.1). Our approach is blind to RNA species that
593 do not include this distal 3' sequence, i.e. to all RNAs that have been released from
594 the pgRNA by two cuts upstream of this reference point during reverse transcription.
595 Whether these RNA species are also capped, similar to the HBx transcripts that were
596 identified in this study, remains to be investigated.

597 In recent years, many studies evaluated circulating pgRNA as a biomarker for disease
598 development and treatment efficacy. A good correlation between serological and intra-
599 hepatic pgRNA levels, as well as between circulating pgRNA and the ratio between
600 intracellular pgRNA and cccDNA have been observed. Therefore, serological pgRNA
601 could represent a non-invasive read out for monitoring intra-hepatic cccDNA
602 transcriptional activity, even during NA treatment [37,78]. To get a deeper
603 understanding of the complexity of the circulating viral RNAs we characterized the
604 majority of serological viral RNAs in three patients using the HBV full-length 5'RACE
605 (Fig.7). Similar to the RNA species associated with viral particles, circulating RNAs
606 belong to one of the three categories: 1. pgRNA, 2. pgRNA splice variant, 3. HBx
607 transcripts (long, canonical and short). Likewise, we detected neither preC- nor S-
608 transcripts in the plasmas of patients. Considering the possibility that the different HBx
609 RNAs are post-transcriptionally generated inside viral particles these results support

610 the concept that viral RNAs found in the blood circulation are all derived from packaged
611 pgRNAs and protected against degradation by nucleases in the blood of patients.
612 Consequently, all viral RNAs that are not packaged into viral particles, i.e. preC- and
613 S-transcripts, are not detected by the full-length 5'RACE. At this point, it should be
614 emphasized that the capped RNA specific 5'RACE approach only detects RNAs that
615 are non-degraded. It excludes all the partially digested viral RNAs that might be present
616 in the blood of patients, especially of those with high ALT levels. However, with our
617 limited sub-cloning and sequencing approach, we cannot categorically exclude the
618 presence of S- or preC-RNAs in the plasmas of patients. Yet, they might represent
619 minor viral RNA species since we readily detected them by the same cloning and
620 sequencing approach in infected cells.

621 At last, the 5'RACE profiles of the different patients exhibit differential patterns of PCR
622 bands. The differential PCR patterns together with the identification of specific pgRNA
623 splice variants in different patients demonstrate that the composition of the viral RNA
624 species can vary among patients. Our observations open new avenues of research to
625 determine the biological and clinical relevance of circulating viral RNAs as novel
626 biomarkers for treatment efficacy and their potential to decrease the pool of
627 intrahepatic cccDNA.

628

629 **Acknowledgements**

630 The authors would like to thank Maud Michelet, Jennifer Molle, Anaëlle Dubois, and
631 Océane Floriot, for their help in the isolation of primary human hepatocytes, as well as
632 Prof. Michel Rivoire's surgical staff for providing liver resections.

633

634 **References**

- 635 [1] Zoulim F, Lebossé F, Levrero M. Current treatments for chronic hepatitis B virus infections.
636 *Curr Opin Virol* 2016;18:109–16. <https://doi.org/10.1016/j.coviro.2016.06.004>.
- 637 [2] Moreau P, Cournac A, Palumbo GA, Marbouty M, Mortaza S, Thierry A, et al.
638 Tridimensional infiltration of DNA viruses into the host genome shows preferential contact with
639 active chromatin. *Nat Commun* 2018;9:4268. <https://doi.org/10.1038/s41467-018-06739-4>.
- 640 [3] Nassal M. HBV cccDNA: viral persistence reservoir and key obstacle for a cure of chronic
641 hepatitis B. *Gut* 2015;64:1972–84. <https://doi.org/10.1136/gutjnl-2015-309809>.
- 642 [4] Seeger C, Mason WS. Molecular biology of hepatitis B virus infection. *Virology* 2015;479–
643 480:672–86. <https://doi.org/10.1016/j.virol.2015.02.031>.
- 644 [5] Mitra B, Wang J, Kim ES, Mao R, Dong M, Liu Y, et al. Hepatitis B Virus Precore Protein
645 p22 Inhibits Interferon-alpha Signaling by Blocking STAT Nuclear Translocation. *J Virol*
646 2019;JVI.00196-19, jvi;JVI.00196-19v1. <https://doi.org/10.1128/JVI.00196-19>.
- 647 [6] Cha M-Y, Ryu D-K, Jung H-S, Chang H-E, Ryu W-S. Stimulation of hepatitis B virus genome
648 replication by HBx is linked to both nuclear and cytoplasmic HBx expression. *J Gen Virol*
649 2009;90:978–86. <https://doi.org/10.1099/vir.0.009928-0>.
- 650 [7] Dandri M, Petersen J, Stockert RJ, Harris TM, Rogler CE. Metabolic Labeling of Woodchuck
651 Hepatitis B Virus X Protein in Naturally Infected Hepatocytes Reveals a Bimodal Half-Life and
652 Association with the Nuclear Framework. *J VIROL* 1998;72:6.
- 653 [8] Doria M, Klein N, Lucito R, Schneider RJ. The hepatitis B virus HBx protein is a dual
654 specificity cytoplasmic activator of Ras and nuclear activator of transcription factors. *EMBO J*
655 1995;14:4747–57. <https://doi.org/10.1002/j.1460-2075.1995.tb00156.x>.
- 656 [9] Korniyev D, Ramakrishnan D, Voitenleitner C, Livingston CM, Xing W, Hung M, et al.
657 Spatiotemporal Analysis of Hepatitis B Virus X Protein in Primary Human Hepatocytes. *J Virol*
658 2019;93:e00248-19, /jvi/93/16/JVI.00248-19.atom. <https://doi.org/10.1128/JVI.00248-19>.
- 659 [10] Murakami S. Hepatitis B virus X protein: a multifunctional viral regulator. *J Gastroenterol*
660 2001;36:651–60. <https://doi.org/10.1007/s005350170027>.
- 661 [11] Benhenda S, Ducroux A, Riviere L, Sobhian B, Ward MD, Dion S, et al. Methyltransferase
662 PRMT1 Is a Binding Partner of HBx and a Negative Regulator of Hepatitis B Virus Transcription. *J*
663 *Virol* 2013;87:4360–71. <https://doi.org/10.1128/JVI.02574-12>.
- 664 [12] Decorsière A, Mueller H, van Breugel PC, Abdul F, Gerossier L, Beran RK, et al. Hepatitis B
665 virus X protein identifies the Smc5/6 complex as a host restriction factor. *Nature* 2016;531:386–380.
666 <https://doi.org/10.1038/nature17170>.
- 667 [13] Elmore LW, Hancock AR, Chang S-F, Wang XW, Chang S, Callahan CP, et al. Hepatitis B
668 virus X protein and p53 tumor suppressor interactions in the modulation of apoptosis. *Proc Natl Acad*
669 *Sci* 1997;94:14707–12. <https://doi.org/10.1073/pnas.94.26.14707>.
- 670 [14] Truant R, Antunovic J, Greenblatt J, Prives C, Cromlish JA. Direct Interaction of the Hepatitis
671 B Virus HBx Protein with p53 Leads to Inhibition by HBx of p53 Response Element-Directed
672 Transactivation. *J VIROL* 1995;69:9.
- 673 [15] Lucifora J, Arzberger S, Durantel D, Belloni L, Strubin M, Levrero M, et al. Hepatitis B virus
674 X protein is essential to initiate and maintain virus replication after infection. *J Hepatol* 2011;55:996–
675 1003. <https://doi.org/10.1016/j.jhep.2011.02.015>.
- 676 [16] Ladner SK, Otto MJ, Barker CS, Zaifert K, Wang GH, Guo JT, et al. Inducible expression of

677 human hepatitis B virus (HBV) in stably transfected hepatoblastoma cells: a novel system for
678 screening potential inhibitors of HBV replication. *Antimicrob Agents Chemother* 1997;41:1715–20.
679 <https://doi.org/10.1128/AAC.41.8.1715>.

680 [17] Ni Y, Lempp FA, Mehrle S, Nkongolo S, Kaufman C, Fälth M, et al. Hepatitis B and D
681 Viruses Exploit Sodium Taurocholate Co-transporting Polypeptide for Species-Specific Entry into
682 Hepatocytes. *Gastroenterology* 2014;146:1070-1083.e6. <https://doi.org/10.1053/j.gastro.2013.12.024>.

683 [18] Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium
684 thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987;162:156–9.
685 [https://doi.org/10.1016/0003-2697\(87\)90021-2](https://doi.org/10.1016/0003-2697(87)90021-2).

686 [19] Peterson DL, Paul DA, Lam J, Tribby II, Achord DT. Antigenic structure of hepatitis B
687 surface antigen: identification of the “d” subtype determinant by chemical modification and use of
688 monoclonal antibodies. 1984:9.

689 [20] Lebossé F, Testoni B, Fresquet J, Facchetti F, Galmozzi E, Fournier M, et al. Intrahepatic
690 innate immune response pathways are downregulated in untreated chronic hepatitis B. *J Hepatol*
691 2017;66:897–909. <https://doi.org/10.1016/j.jhep.2016.12.024>.

692 [21] Volz T, Lutgehetmann M, Wachtler P, Jacob A, Quaas A, Murray JM, et al. Impaired
693 Intrahepatic Hepatitis B Virus Productivity Contributes to Low Viremia in Most HBeAg-Negative
694 Patients. *Gastroenterology* 2007;133:843–52. <https://doi.org/10.1053/j.gastro.2007.06.057>.

695 [22] Niu C, Livingston CM, Li L, Beran RK, Daffis S, Ramakrishnan D, et al. The Smc5/6
696 Complex Restricts HBV when Localized to ND10 without Inducing an Innate Immune Response and
697 Is Counteracted by the HBV X Protein Shortly after Infection. *PLOS ONE* 2017;12:e0169648.
698 <https://doi.org/10.1371/journal.pone.0169648>.

699 [23] Altinel K, Hashimoto K, Wei Y, Neuveut C, Gupta I, Suzuki AM, et al. Single-Nucleotide
700 Resolution Mapping of Hepatitis B Virus Promoters in Infected Human Livers and Hepatocellular
701 Carcinoma. *J Virol* 2016;90:10811–22. <https://doi.org/10.1128/JVI.01625-16>.

702 [24] Lempp FA, Urban S. Inhibitors of Hepatitis B Virus Attachment and Entry. *Intervirology*
703 2014;57:151–7. <https://doi.org/10.1159/000360948>.

704 [25] Zoulim F, Locarnini S. Hepatitis B Virus Resistance to Nucleos(t)ide Analogues.
705 *Gastroenterology* 2009;137:1593-1608.e2. <https://doi.org/10.1053/j.gastro.2009.08.063>.

706 [26] Belloni L, Allweiss L, Guerrieri F, Pediconi N, Volz T, Pollicino T, et al. IFN- α inhibits HBV
707 transcription and replication in cell culture and in humanized mice by targeting the epigenetic
708 regulation of the nuclear cccDNA minichromosome 2012. <https://doi.org/10.1172/JCI58847>.

709 [27] Liu F, Campagna M, Qi Y, Zhao X, Guo F, Xu C, et al. Alpha-Interferon Suppresses
710 Hepadnavirus Transcription by Altering Epigenetic Modification of cccDNA Minichromosomes.
711 *PLoS Pathog* 2013;9:e1003613. <https://doi.org/10.1371/journal.ppat.1003613>.

712 [28] Lucifora J, Durantel D, Testoni B, Hantz O, Levrero M, Zoulim F. Control of hepatitis B virus
713 replication by innate response of HepaRG cells. *Hepatology* 2010;51:63–72.
714 <https://doi.org/10.1002/hep.23230>.

715 [29] Shen F, Li Y, Wang Y, Sozzi V, Revill PA, Liu J, et al. Hepatitis B virus sensitivity to
716 interferon- α in hepatocytes is more associated with cellular interferon response than with viral
717 genotype. *Hepatology* 2018;67:1237–52. <https://doi.org/10.1002/hep.29609>.

718 [30] Yan R, Zhao X, Cai D, Liu Y, Block TM, Guo J-T, et al. The Interferon-Inducible Protein
719 Tetherin Inhibits Hepatitis B Virus Virion Secretion. *J Virol* 2015;89:9200–12.

- 720 <https://doi.org/10.1128/JVI.00933-15>.
- 721 [31] Bai L, Zhang X, Kozlowski M, Li W, Wu M, Liu J, et al. Extracellular Hepatitis B Virus
722 RNAs Are Heterogeneous in Length and Circulate as Capsid-Antibody Complexes in Addition to
723 Virions in Chronic Hepatitis B Patients. *J Virol* 2018;92:e00798-18, /jvi/92/24/e00798-18.atom.
724 <https://doi.org/10.1128/JVI.00798-18>.
- 725 [32] Wang J, Shen T, Huang X, Kumar GR, Chen X, Zeng Z, et al. Serum hepatitis B virus RNA is
726 encapsidated pregenome RNA that may be associated with persistence of viral infection and rebound.
727 *J Hepatol* 2016;65:700–10. <https://doi.org/10.1016/j.jhep.2016.05.029>.
- 728 [33] Günther S, Sommer G, Iwanska A, Will H. Heterogeneity and Common Features of Defective
729 Hepatitis B Virus Genomes Derived from Spliced Pregenomic RNA. *Virology* 1997;238:363–71.
730 <https://doi.org/10.1006/viro.1997.8863>.
- 731 [34] Testoni B, Lebossé F, Scholtes C, Berby F, Miaglia C, Subic M, et al. Serum hepatitis B core-
732 related antigen (HBcrAg) correlates with covalently closed circular DNA transcriptional activity in
733 chronic hepatitis B patients. *J Hepatol* 2019;70:615–25. <https://doi.org/10.1016/j.jhep.2018.11.030>.
- 734 [35] Hacker HJ, Zhang W, Tokus M, Bock T, Schröder CH. Patterns of Circulating Hepatitis B
735 Virus Serum Nucleic Acids during Lamivudine Therapy. *Ann N Y Acad Sci* 2004;1022:271–81.
736 <https://doi.org/10.1196/annals.1318.042>.
- 737 [36] Lam AM, Ren S, Espiritu C, Kelly M, Lau V, Zheng L, et al. Hepatitis B Virus Capsid
738 Assembly Modulators, but Not Nucleoside Analogs, Inhibit the Production of Extracellular
739 Pregenomic RNA and Spliced RNA Variants. *Antimicrob Agents Chemother* 2017;61:e00680-17,
740 /aac/61/8/e00680-17.atom. <https://doi.org/10.1128/AAC.00680-17>.
- 741 [37] Liu S, Zhou B, Valdes JD, Sun J, Guo H. Serum Hepatitis B Virus RNA: A New Potential
742 Biomarker for Chronic Hepatitis B Virus Infection. *Hepatology* 2019;69:1816–27.
743 <https://doi.org/10.1002/hep.30325>.
- 744 [38] Wang J, Sheng Q, Ding Y, Chen R, Sun X, Chen X, et al. HBV RNA virion-like particles
745 produced under nucleos(t)ide analogues treatment are mainly replication-deficient. *J Hepatol*
746 2018;68:847–9. <https://doi.org/10.1016/j.jhep.2017.10.030>.
- 747 [39] Kwee L. Alternate Translation Initiation on Hepatitis B Virus X mRNA Produces Multiple
748 Polypeptides That Differentially Transactivate Class II and III Promoters. *J VIROL* 1992;66:8.
- 749 [40] Zheng Y. Novel Short Transcripts of Hepatitis B Virus X Gene Derived from Intragenic
750 Promote. *J Biol Chem* n.d.
- 751 [41] Hodgson AJ, Hyser JM, Keasler VV, Cang Y, Slagle BL. Hepatitis B virus regulatory HBx
752 protein binding to DDB1 is required but is not sufficient for maximal HBV replication. *Virology*
753 2012;426:73–82. <https://doi.org/10.1016/j.virol.2012.01.021>.
- 754 [42] Leupin O, Bontron S, Schaeffer C, Strubin M. Hepatitis B Virus X Protein Stimulates Viral
755 Genome Replication via a DDB1-Dependent Pathway Distinct from That Leading to Cell Death. *J*
756 *Virol* 2005;79:4238–45. <https://doi.org/10.1128/JVI.79.7.4238-4245.2005>.
- 757 [43] Choi J, Kim EE, Park YI, Han YS. Expression of the active human and duck hepatitis B virus
758 polymerases in heterologous system of *Pichia methanolica*. *Antiviral Res* 2002;55:279–90.
759 [https://doi.org/10.1016/S0166-3542\(02\)00023-2](https://doi.org/10.1016/S0166-3542(02)00023-2).
- 760 [44] Potenza N, Salvatore V, Raimondo D, Falanga D, Nobile V, Peterson DL, et al. Optimized
761 expression from a synthetic gene of an untagged RNase H domain of human hepatitis B virus
762 polymerase which is enzymatically active. *Protein Expr Purif* 2007;55:93–9.

763 <https://doi.org/10.1016/j.pep.2007.04.005>.

764 [45] Tavis JE, Cheng X, Hu Y, Totten M, Cao F, Michailidis E, et al. The Hepatitis B Virus
765 Ribonuclease H Is Sensitive to Inhibitors of the Human Immunodeficiency Virus Ribonuclease H and
766 Integrase Enzymes. *PLoS Pathog* 2013;9:e1003125. <https://doi.org/10.1371/journal.ppat.1003125>.

767 [46] Wei X, Peterson DL. Expression, Purification, and Characterization of an Active RNase H
768 Domain of the Hepatitis B Viral Polymerase. *J Biol Chem* 1996;271:32617–22.
769 <https://doi.org/10.1074/jbc.271.51.32617>.

770 [47] Champoux JJ, Schultz SJ. Ribonuclease H: properties, substrate specificity and roles in
771 retroviral reverse transcription: Retroviral RNases H. *FEBS J* 2009;276:1506–16.
772 <https://doi.org/10.1111/j.1742-4658.2009.06909.x>.

773 [48] Schultz SJ, Champoux JJ. RNase H activity: Structure, specificity, and function in reverse
774 transcription. *Virus Res* 2008;134:86–103. <https://doi.org/10.1016/j.virusres.2007.12.007>.

775 [49] Tycowski KT, Guo YE, Lee N, Moss WN, Vallery TK, Xie M, et al. Viral noncoding RNAs:
776 more surprises. *Genes Dev* 2015;29:567–84. <https://doi.org/10.1101/gad.259077.115>.

777 [50] Bidet K, Dadlani D, Garcia-Blanco MA. G3BP1, G3BP2 and CAPRIN1 Are Required for
778 Translation of Interferon Stimulated mRNAs and Are Targeted by a Dengue Virus Non-coding RNA.
779 *PLoS Pathog* 2014;10:e1004242. <https://doi.org/10.1371/journal.ppat.1004242>.

780 [51] Chang R-Y, Hsu T-W, Chen Y-L, Liu S-F, Tsai Y-J, Lin Y-T, et al. Japanese encephalitis
781 virus non-coding RNA inhibits activation of interferon by blocking nuclear translocation of interferon
782 regulatory factor 3. *Vet Microbiol* 2013;166:11–21. <https://doi.org/10.1016/j.vetmic.2013.04.026>.

783 [52] Qiu L, Wang T, Tang Q, Li G, Wu P, Chen K. Long Non-coding RNAs: Regulators of Viral
784 Infection and the Interferon Antiviral Response. *Front Microbiol* 2018;9:1621.
785 <https://doi.org/10.3389/fmicb.2018.01621>.

786 [53] Rossetto CC, Pari G. KSHV PAN RNA Associates with Demethylases UTX and JMJD3 to
787 Activate Lytic Replication through a Physical Interaction with the Virus Genome. *PLoS Pathog*
788 2012;8:e1002680. <https://doi.org/10.1371/journal.ppat.1002680>.

789 [54] Rossetto CC, Pari GS. Kaposi's Sarcoma-Associated Herpesvirus Noncoding Polyadenylated
790 Nuclear RNA Interacts with Virus- and Host Cell-Encoded Proteins and Suppresses Expression of
791 Genes Involved in Immune Modulation. *J Virol* 2011;85:13290–7. <https://doi.org/10.1128/JVI.05886-11>.

793 [55] Yang X, Li H, Sun H, Fan H, Hu Y, Liu M, et al. Hepatitis B Virus-Encoded MicroRNA
794 Controls Viral Replication. *J Virol* 2017;91:e01919-16, [/jvi/91/10/e01919-16.atom](https://doi.org/10.1128/JVI.01919-16).
795 <https://doi.org/10.1128/JVI.01919-16>.

796 [56] Rang A, Günther S, Will H. Effect of interferon alpha on hepatitis B virus replication and gene
797 expression in transiently transfected human hepatoma cells. *J Hepatol* 1999;31:791–9.
798 [https://doi.org/10.1016/S0168-8278\(99\)80279-7](https://doi.org/10.1016/S0168-8278(99)80279-7).

799 [57] Pasquetto V, Wieland SF, Uprichard SL, Tripodi M, Chisari FV. Cytokine-Sensitive
800 Replication of Hepatitis B Virus in Immortalized Mouse Hepatocyte Cultures. *J Virol* 2002;76:5646–
801 53. <https://doi.org/10.1128/JVI.76.11.5646-5653.2002>.

802 [58] Liu Y, Nie H, Mao R, Mitra B, Cai D, Yan R, et al. Interferon-inducible ribonuclease ISG20
803 inhibits hepatitis B virus replication through directly binding to the epsilon stem-loop structure of viral
804 RNA. *PLOS Pathog* 2017;13:e1006296. <https://doi.org/10.1371/journal.ppat.1006296>.

- 805 [59] Lin Y-C, Hsu E-C, Ting L-P. Repression of hepatitis B viral gene expression by transcription
806 factor nuclear factor-kappaB. *Cell Microbiol* 2009;11:645–60. [https://doi.org/10.1111/j.1462-](https://doi.org/10.1111/j.1462-5822.2008.01280.x)
807 5822.2008.01280.x.
- 808 [60] Quasdorff M, Protzer U. Control of hepatitis B virus at the level of transcription. *J Viral Hepat*
809 2010;17:527–36. <https://doi.org/10.1111/j.1365-2893.2010.01315.x>.
- 810 [61] Gao B, Duan Z, Xu W, Xiong S. Tripartite motif-containing 22 inhibits the activity of
811 hepatitis B virus core promoter, which is dependent on nuclear-located RING domain. *Hepatology*
812 2009;50:424–33. <https://doi.org/10.1002/hep.23011>.
- 813 [62] Chen P-J, Chen C-R, Sung J-L, Chen D-S. Identification of a Doubly Spliced Viral Transcript
814 Joining the Separated Domains for Putative Protease and Reverse Transcriptase of Hepatitis B Virus. *J*
815 *VIROL* 1989;63:7.
- 816 [63] Su T-S, Lui W-Y, Lin L-H, Han S-H, P'eng F-K. Analysis of hepatitis B virus transcripts in
817 infected human livers. *Hepatology* 1989;9:180–5. <https://doi.org/10.1002/hep.1840090203>.
- 818 [64] Su T-S, Lai C-J, Huang J-L, Lin L-H, Yauk Y-K, Chang C, et al. Hepatitis B Virus Transcript
819 Produced by RNA Splicing. *J VIROL* 1989;63:8.
- 820 [65] Suzuki T, Masui N, Kajino K, Saito I, Miyamura T. Detection and mapping of spliced RNA
821 from a human hepatoma cell line transfected with the hepatitis B virus genome. *Proc Natl Acad Sci*
822 1989;86:8422–6. <https://doi.org/10.1073/pnas.86.21.8422>.
- 823 [66] Brechot C. Defective Hepatitis B Virus Particles Are Generated by Packaging and Reverse
824 Transcription of Spliced Viral RNAs In Vivo. *J VIROL* 1991;65:5.
- 825 [67] Kock J, Nassal M, Deres K, Blum HE, von Weizsacker F. Hepatitis B Virus Nucleocapsids
826 Formed by Carboxy-Terminally Mutated Core Proteins Contain Spliced Viral Genomes but Lack Full-
827 Size DNA. *J Virol* 2004;78:13812–8. <https://doi.org/10.1128/JVI.78.24.13812-13818.2004>.
- 828 [68] Lam AM, Ren S, Espiritu C, Kelly M, Lau V, Zheng L, et al. Hepatitis B Virus Capsid
829 Assembly Modulators, but Not Nucleoside Analogs, Inhibit the Production of Extracellular
830 Pregenomic RNA and Spliced RNA Variants. *Antimicrob Agents Chemother* 2017;61:e00680-17,
831 /aac/61/8/e00680-17.atom. <https://doi.org/10.1128/AAC.00680-17>.
- 832 [69] Rosmorduc O, Petit M-A, Pol S, Capel F, Bortolotti F, Berthelot P, et al. In vivo and in vitro
833 expression of defective hepatitis B virus particles generated by spliced hepatitis B virus RNA.
834 *Hepatology* 1995;22:10–9. <https://doi.org/10.1002/hep.1840220103>.
- 835 [70] Huang H-L, Jeng K-S, Hu C-P, Tsai C-H, Lo SJ, Chang C. Identification and Characterization
836 of a Structural Protein of Hepatitis B Virus: A Polymerase and Surface Fusion Protein Encoded by a
837 Spliced RNA. *Virology* 2000;275:398–410. <https://doi.org/10.1006/viro.2000.0478>.
- 838 [71] Soussan P, Garreau F, Zylberberg H, Ferray C, Brechot C, Kremsdorf D. In vivo expression of
839 a new hepatitis B virus protein encoded by a spliced RNA 2000. <https://doi.org/10.1172/JCI8098>.
- 840 [72] Bartenschlager R, Schaller H. Hepadnaviral assembly is initiated by polymerase binding to the
841 encapsidation signal in the viral RNA genome. *EMBO J* 1992;11:3413–20.
842 <https://doi.org/10.1002/j.1460-2075.1992.tb05420.x>.
- 843 [73] Hirsch RC, Loeb DD, Pollack JR, Ganem D. cis-Acting Sequences Required for
844 Encapsidation of Duck Hepatitis B Virus Pregenomic RNA. *J VIROL* 1991;65:8.
- 845 [74] Junker-Niepmann M, Bartenschlager R, Schaller H. A short cis-acting sequence is required for
846 hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. *EMBO J*

847 1990;9:3389–96. <https://doi.org/10.1002/j.1460-2075.1990.tb07540.x>.

848 [75] Knaus T, Nassal M. The encapsidation signal on the hepatitis B virus RNA pregenome forms
849 a stem-loop structure that is critical for its function. *Nucleic Acids Res* 1993;21:3967–75.
850 <https://doi.org/10.1093/nar/21.17.3967>.

851 [76] DeStefano JJ, Buiser RG, Mallaber LM, Myers TW, Bambara RA, Fay PJ. Polymerization and
852 RNase H activities of the reverse transcriptases from avian myeloblastosis, human immunodeficiency,
853 and Moloney murine leukemia viruses are functionally uncoupled. *J Biol Chem* 1991;266:7423–31.

854 [77] Sokoloski KJ, Haist KC, Morrison TE, Mukhopadhyay S, Hardy RW. Noncapped Alphavirus
855 Genomic RNAs and Their Role during Infection. *J Virol* 2015;89:6080–92.
856 <https://doi.org/10.1128/JVI.00553-15>.

857 [78] Wang J, Yu Y, Li G, Shen C, Meng Z, Zheng J, et al. Relationship between serum HBV-RNA
858 levels and intrahepatic viral as well as histologic activity markers in entecavir-treated patients. *J*
859 *Hepatol* 2018;68:16–24. <https://doi.org/10.1016/j.jhep.2017.08.021>.

860

861

862 **Fig.1. Set-up of the HBV full-length 5'RACE.**

863 (A) Illustration of the Hepatitis B viral genome with its four promoters (arrows) and viral
864 transcripts. preCore RNA (preC), pregenomic RNA (pg), Large surface protein RNA
865 (preS1), Middle surface protein RNA (preS1), Small surface protein RNA (S), X protein
866 RNA (HBx) and covalently closed circular DNA (cccDNA).

867 (B) Overview of the capped RNA specific HBV full-length 5'RACE protocol. In a first
868 step, RNAs are dephosphorylated to assure that degraded and truncated RNAs are
869 excluded from the 5'RACE assay. 5' capped full-length RNAs are protected from the
870 dephosphorylation reaction. Subsequently, the 5' cap structure (m^7Gppp) is removed
871 from the full-length RNAs by a de-capping reaction resulting in a free phosphate group
872 (p) at their 5' ends to which an RNA anchor (green box) can be ligated in the following
873 step. Then, the viral RNAs are reverse transcribed using the gene specific primer Gsp1
874 (black arrow) integrating the anchor sequence into the cDNA. Finally, the different viral
875 cDNAs are amplified by PCR using the second gene specific primer Gsp2 (black arrow)
876 and an anchor specific primer (green arrow) in the same tube.

877 (C) Transcript-specific 5'RACE amplicons using HepG2-NTCP cells 8 days post
878 infection (p.i.) as compared to non-infected (n.i.) cells. Arrow indicates HBx. Individual
879 amplicons are reduced in length by approximately 337 bp with respect to template viral
880 RNAs because of the usage of gene specific primers. All amplicons were sequenced
881 by sub-cloning and Sanger sequencing (Suppl. Table 1 and Fig.1E for overview).

882 (D) HBx is the first viral transcript detected after infection in HepG2-NTCP cells and
883 PHH primary human hepatocytes (PHHs).

884 (E) Alignment of sequenced amplicons with respect to HBV genome genotype D
885 (reference sequence GenBank U95551.1). pgRNA and preC RNA in orange. preS1 in
886 dark blue, preS2 in blue and S in light blue. Long HBx transcripts in very light red,
887 canonical HBx transcripts in light red and short HBx transcripts in red. To the right, total
888 number of sequenced clones per transcript category.

889

890 **Fig.2. Tenofovir treatment increases full-length transcripts in HepG2-NTCP cells.**

891 (A) Full-length 5'RACE amplicons using HepG2-NTCP cells after 4, 8 and 15 days of
892 infection. Myrcludex drug regime was initiated 1 day before infection. Tenofovir
893 treatment was started 3 days p.i.. No specific signals are detected by 5'RACE in the
894 Myrcludex condition as verified by sequencing (stars: non-specific signals). Tenofovir
895 leads to an increase in full-length transcripts as compared to non-treated cells.

896 (B) Quantitative PCRs measuring 3kb RNAs (preC and pgRNA), total RNAs and total
897 DNA in treated versus control cells.

898 (C) 3kb ratio (intensity 3kb band/intensity HBx band/lane) is increased and further
899 increases over time in the Tenofovir condition as compared to untreated cells. One
900 representative experiment is shown.

901 Experiments were performed in triplicates. Student's t-test p-values with respect to no
902 treatment: ***<0.001, ****<0.0001

903

904 **Fig.3. Tenofovir treatment primarily augments the pgRNA transcript without**
905 **affecting the other viral transcripts in PHHs.**

906 (A) Full-length 5'RACE and preC/pg RNA-specific 5'RACE amplicons using PHHs after
907 4 and 8 days of infection. Myrcludex drug regime was initiated 1 day before infection.
908 Tenofovir treatment was started 3 days p.i.. No specific signals are detected by both
909 the 5'RACEs in the Myrcludex condition. Tenofovir primarily increases the pgRNA
910 transcript (see preC/pg RNA-specific 5'RACEs) without affecting the other viral
911 transcripts (see full-length 5'RACEs).

912 (B) Quantitative PCRs measuring 3kb RNAs (preC and pgRNA), total RNAs and total
913 DNA in treated versus control cells.

914 (C) 3kb ratio (intensity 3kb band/intensity HBx band/lane) is increased and further
915 increases over time in the Tenofovir condition as compared to untreated cells. One
916 representative experiment is shown.

917 (D) ELISAs against HBeAg with supernatants of PHHs treated with Tenofovir and
918 Interferon β as compared to n.i. and untreated PHHs. Tenofovir treatment does not
919 change HBeAg levels whereas Interferon β strongly reduces HBeAg.

920 (E) Western blot against surface proteins (HBsAgs) using antibody H166 (Abbott)
921 recognizing a conserved epitope in the antigenic loop after 8 days of Myrcludex,
922 Tenofovir or Interferon β treatment. Tenofovir treatment does not affect surface protein
923 levels. Interferon β treatment moderately affects surface protein levels.

924 Experiments were performed in triplicates. Student's t-test p-values with respect to no
925 treatment: ***<0.001, ****<0.0001

926

927 **Fig.4. Interferon β treatment reduces both pg- and preC-transcripts without**
928 **affecting the other viral transcripts in PHHs.**

929 (A) Full-length 5'RACE and preC/pg RNA-specific 5'RACE amplicons using PHHs after
930 8 days of infection. Interferon β treatment was initiated 3 days p.i.. Interferon β
931 treatment decreases both pg- and preC- transcripts (see preC/pg RNA-specific
932 5'RACEs) without affecting the other viral transcripts, like HBx (see full-length
933 5'RACE).

934 (B) Quantitative PCRs measuring 3kb RNAs (preC and pgRNA), total RNAs and total
935 DNA in treated versus control cells.

936 Experiments were performed in triplicates. Student's t-test p-values with respect to no
937 treatment: **<0.01, ***<0.001, ****<0.0001

938

939 **Fig.5. Viral particles produced in HepAD38 cells contain different HBx**
940 **transcripts, at least one spliced pgRNA variant and pgRNA.**

941 (A) Schematic diagram of procedures used to obtain viral particles.

942 (B) Full-length 5'RACE and preC/pg RNA-specific 5'RACE amplicons using HepAD38
943 supernatant precipitated through sucrose cushion, with/without micrococcal nuclease
944 (MNase) digestion, with/without reverse transcriptase (RT). All amplicons were
945 sequenced by sub-cloning and Sanger sequencing (Suppl. table 2). Different HBx
946 RNAs (long, normal and short), pgRNA and one pgRNA splice variant were identified.
947 Neither preC RNA nor S transcripts were detected.

948 (C) Alignment of sequenced amplicons with respect to HBV genome genotype D
949 (reference sequence GenBank U95551.1). To the right, total number of sequenced
950 clones per transcript category.

951 (D) Total viral DNA in gradient fractions of Iodixanol gradient. Viral DNA fractionated
952 between fractions 4 and 8 with a sharp peak in fraction 7.

953 (E) HBsAgs in Iodixanol gradient. All three HBsAgs co-sedimented between fractions
954 5 and 9 with the strongest intensities found in fractions 6 and 7 as shown by Western-
955 blotting.

956 (F) Full-length 5'RACE detects the majority of long, normal and short HBx transcripts
957 in gradient fractions 7 and 8 co-fractionating with viral cDNA and S proteins. pgRNA-
958 specific 5'RACE detects pgRNA in gradient fractions 6 to 9 with the highest abundance
959 in fraction 7 similar to cDNA.

960

961 **Fig.6. HBx RNAs in viral particles are both capped and un-capped.**

962 (A) Overview of the un-capped RNA specific HBV full-length 5'RACE protocol. The un-
963 capped RNA specific 5'RACE protocol lacks the de-capping and de-phosphorylation
964 steps from the standard HBV full-length 5'RACE. In a first step, isolated RNAs are
965 digested by DNase to remove residual DNA molecules. In a next step, the RNA anchor
966 (green box) is ligated to the free phosphate group of un-capped RNAs. Then, the viral
967 RNAs are reverse transcribed using the gene specific primer Gsp1 (black arrow)
968 integrating the anchor sequence into the cDNA. Finally, the different viral cDNAs are
969 amplified by PCR using the second gene specific primer Gsp2 (black arrow) and an
970 anchor specific primer (green arrow) in the same tube.

971 (B) Capped RNA specific 5'RACE using total RNAs of PHHs (lane 1). Capped and un-
972 capped RNA specific 5'RACEs using MNase-digested supernatant of HepAD38 cells
973 (lane 3-5). Presence of different un-capped HBx RNA (amplicons sequenced by sub-
974 cloning and Sanger sequencing, data not shown) in viral particles (lane 5). The de-
975 phosphorylation step of the 5'RACE protocol is very effective (compare lane 4 with

976 lane 5) demonstrating the authenticity of the capped RNA species identified by the
977 HBV full-length 5'RACE in viral particles and in cells (lanes 1 and 3). L (1kb ladder
978 (NEB)).

979

980 **Fig.7. Full length 5'RACE identifies 4 major categories of circulating RNA**
981 **species in plasma and sera of patients.**

982 (A) Clinical characteristics associated with the blood samples of chronic Hepatitis B
983 patients used in this study.

984 (B) Full-length 5'RACE amplicons using patient plasmas (patient #1 and #2) and serum
985 (patient #3) plus minus RT. All amplicons for patients 1 and 2 were sequenced by sub-
986 cloning and Sanger sequencing (Suppl. Table 3A/B). Sequences of serological RNAs
987 belong to one of the three categories: 1. pgRNA 2. pgRNA splice variant 3. HBx
988 transcript (long, normal and short HBx transcripts).

989 (C) Alignment of sequenced amplicons in patient #1 and #2 with respect to HBV
990 genome genotype D. To the right, total number of sequenced clones per transcript
991 category.

992

Figure 1

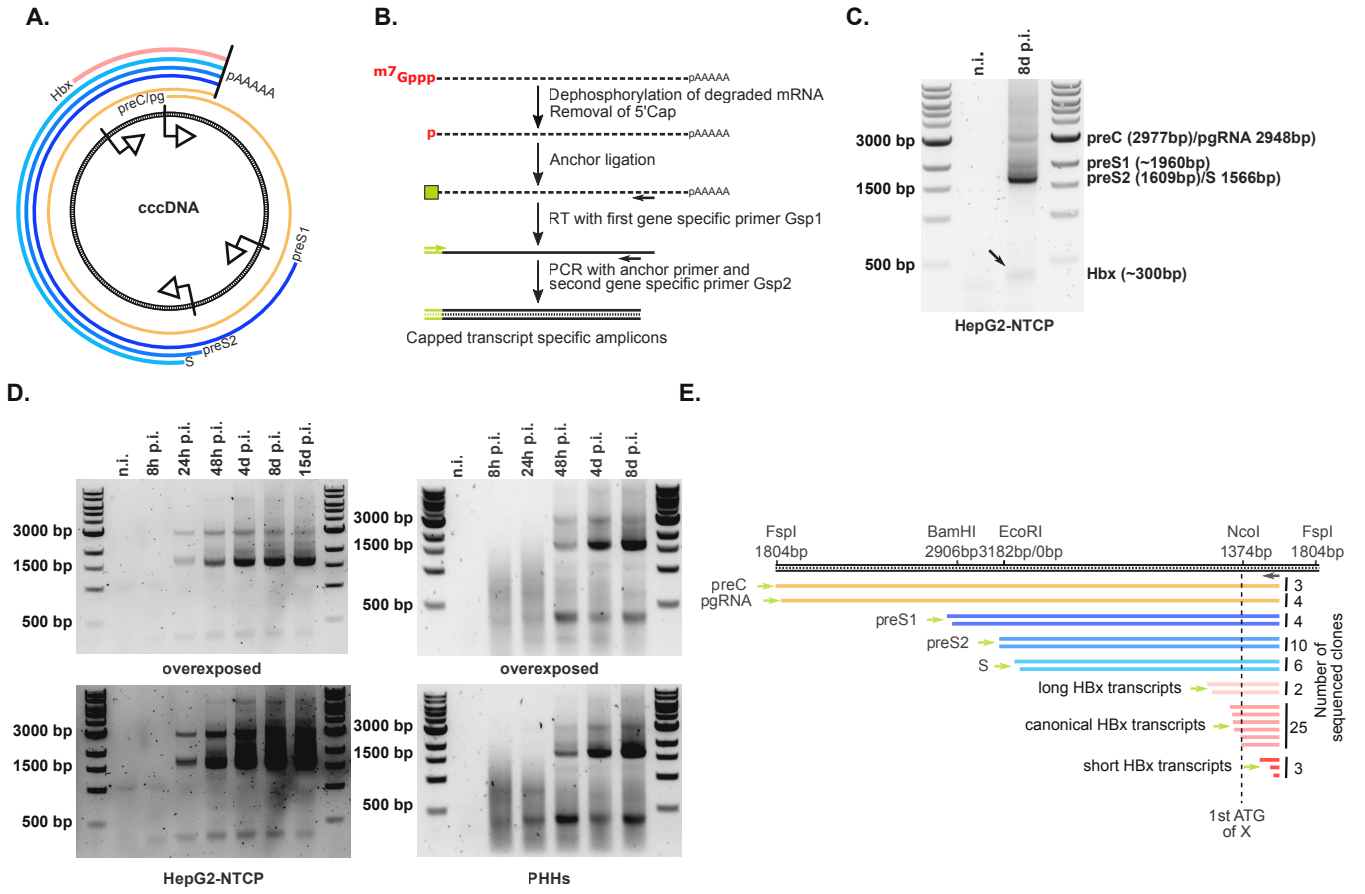
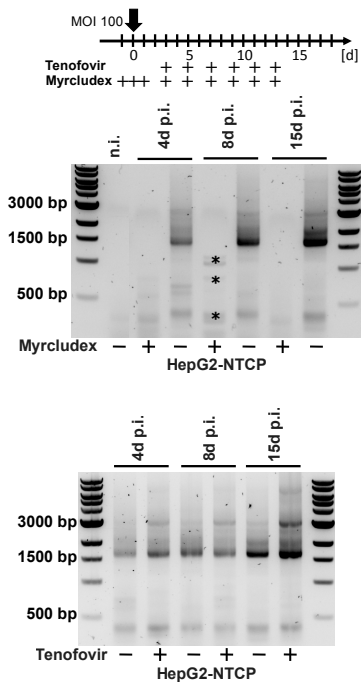
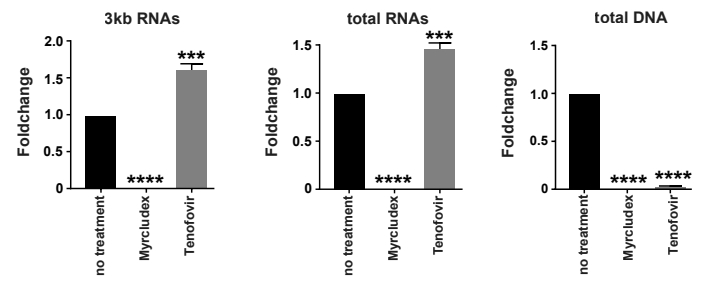


Figure 2

A.



B.



C.

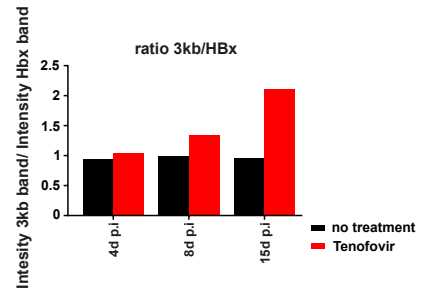
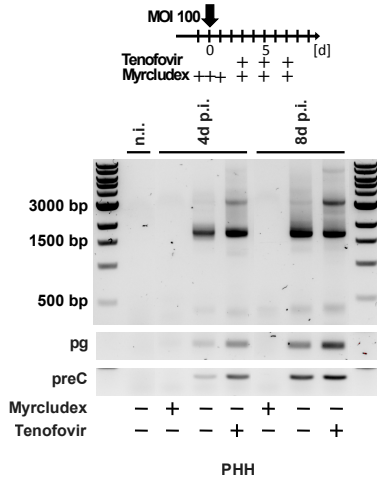
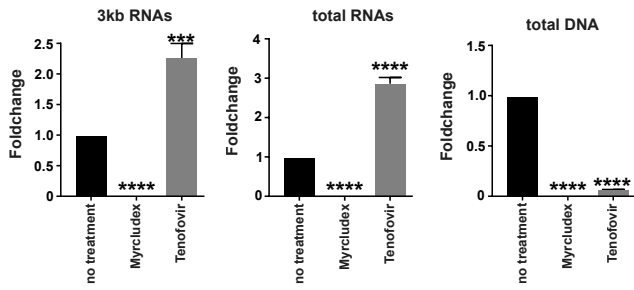


Figure 3

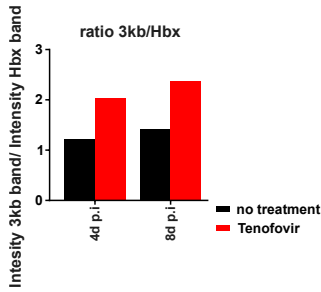
A.



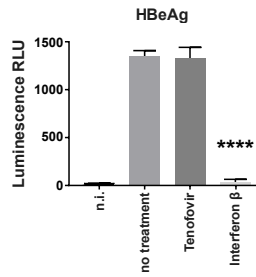
B.



C.



D.



E.

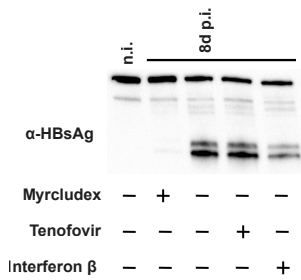
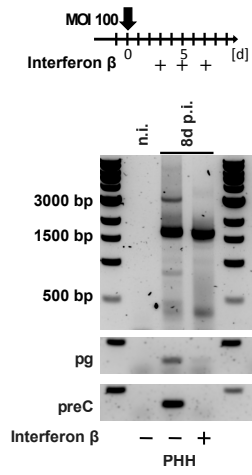


Figure 4

A.



G.

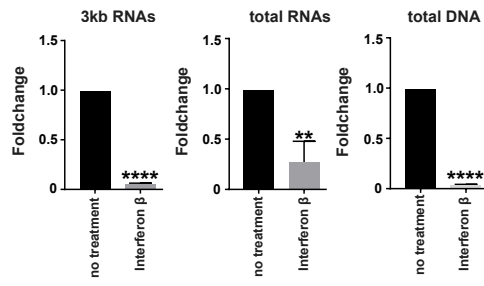


Figure 5

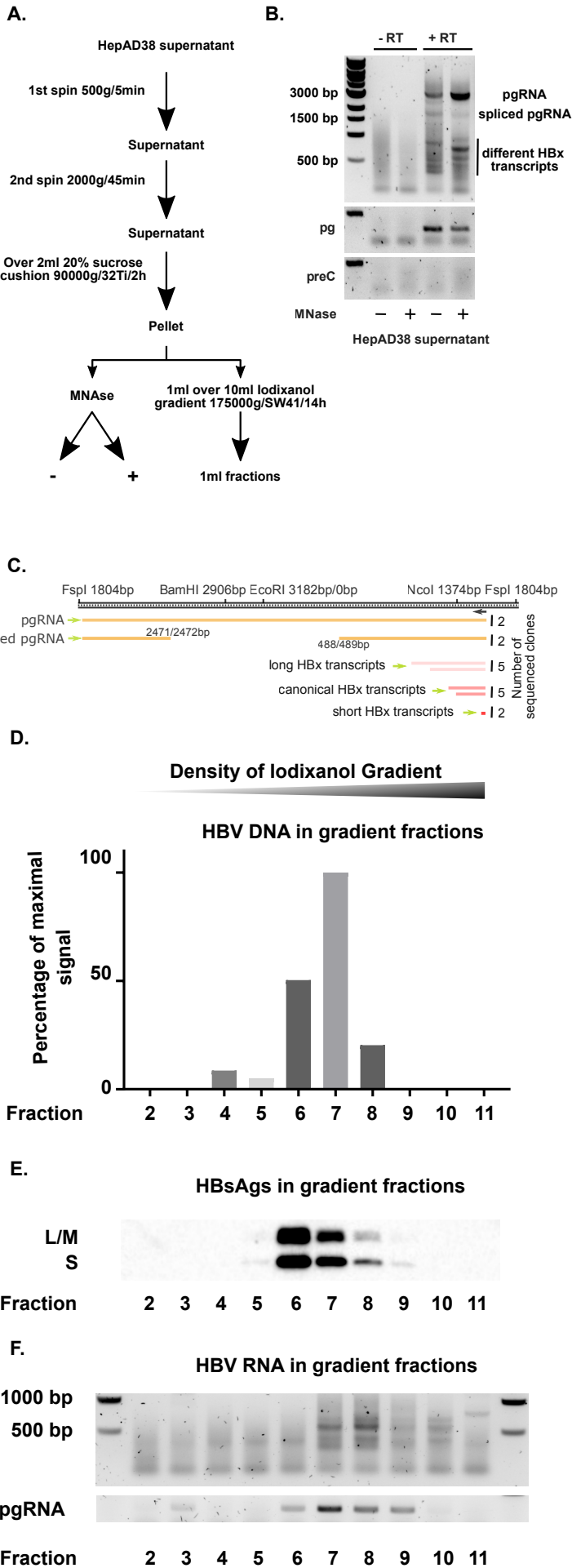
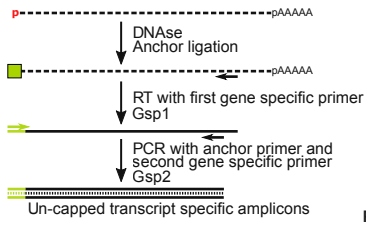


Figure 6

A.



B.

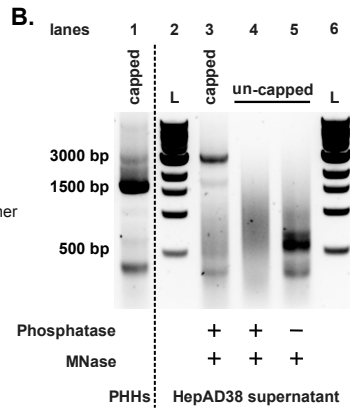
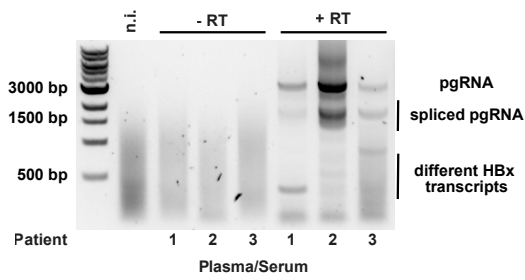


Figure 7

A.

Patient	HBe	ALT	DNA
1	+	normal	8 log IU/ml
2	+	244 IU/ml	8.9 log IU/ml
3	+	normal	8.2 log IU/ml

B.



C.

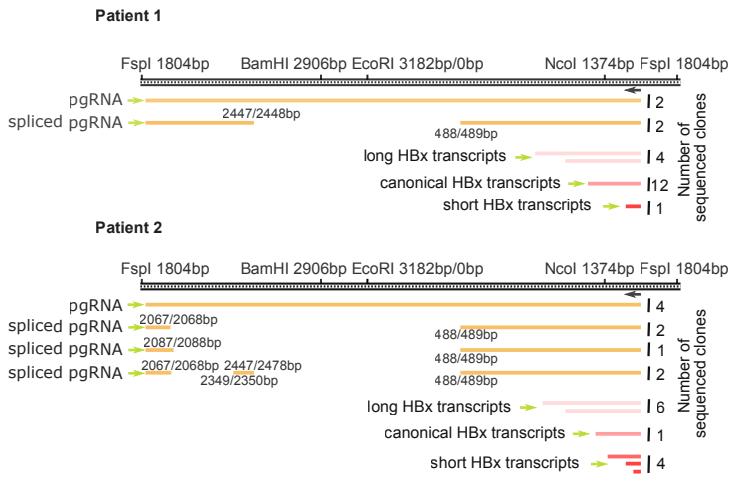
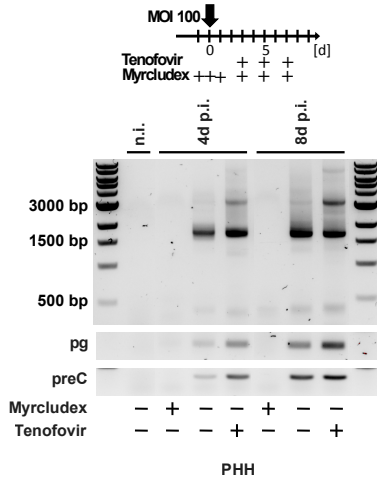
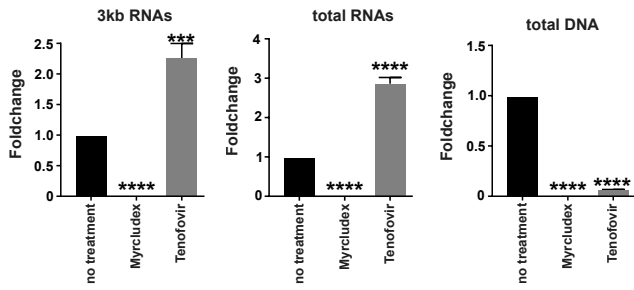


Figure 3

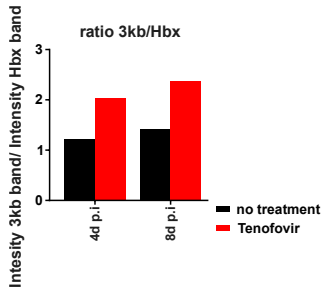
A.



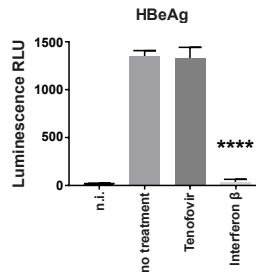
B.



C.



D.



E.

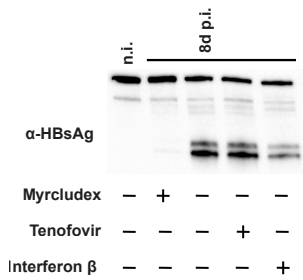


Figure 4

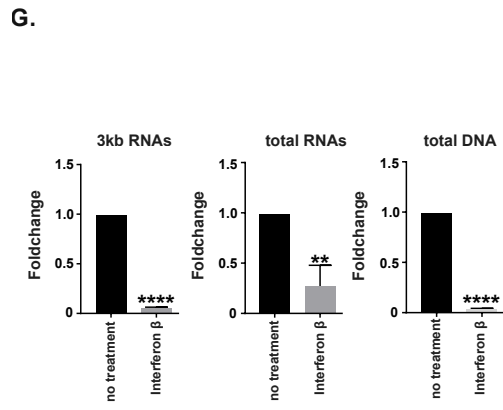
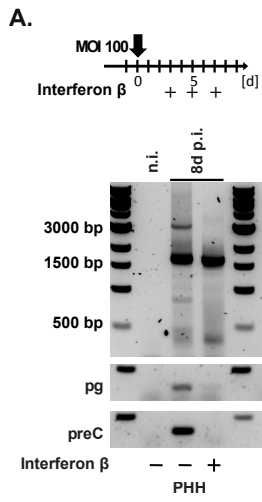


Figure 5

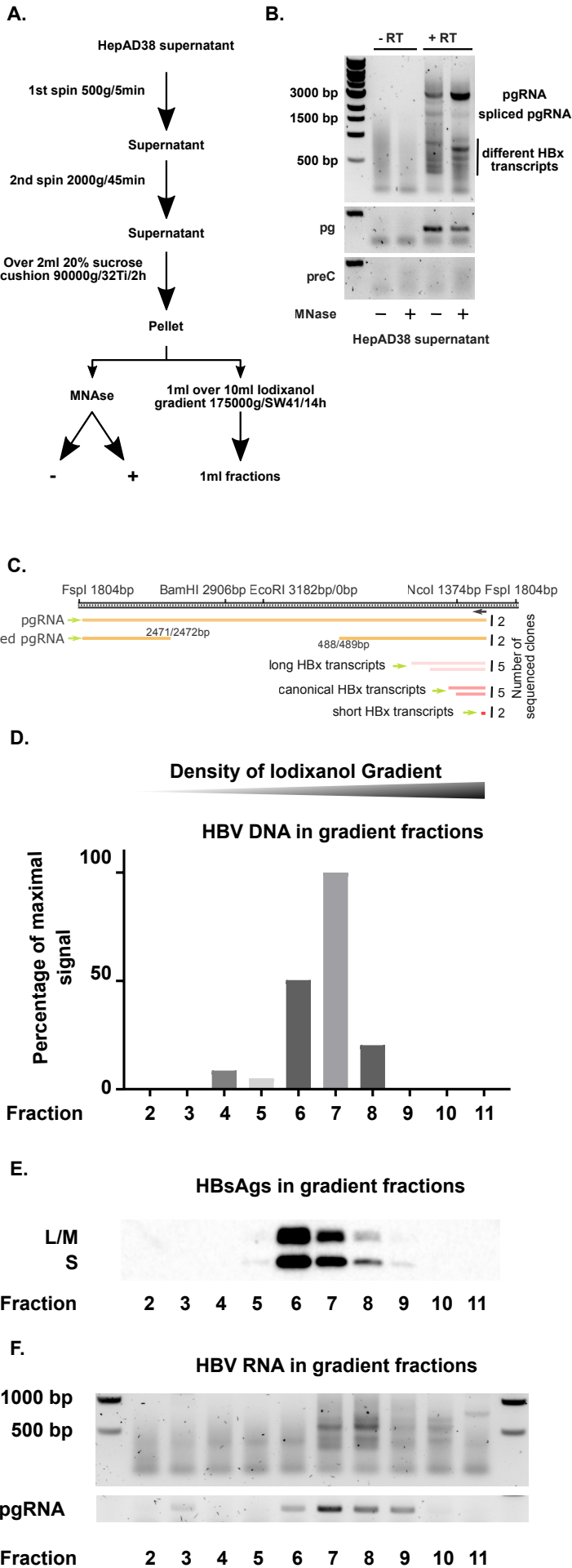
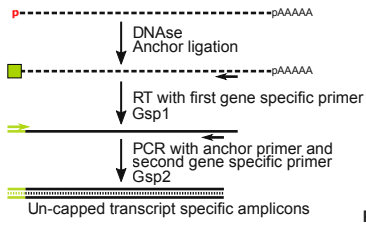


Figure 6

A.



B.

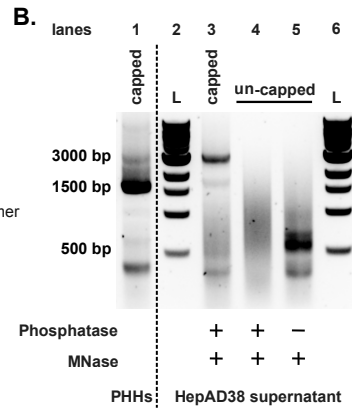
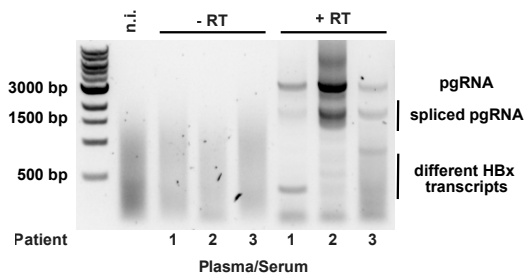


Figure 7

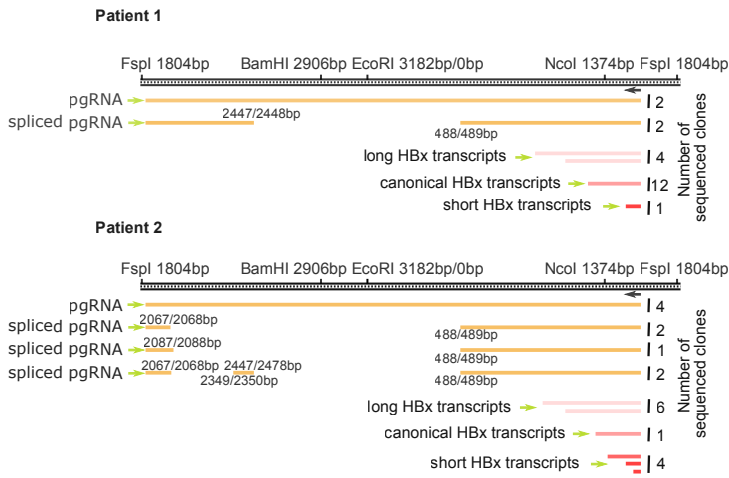
A.

Patient	HBe	ALT	DNA
1	+	normal	8 log IU/ml
2	+	244 IU/ml	8.9 log IU/ml
3	+	normal	8.2 log IU/ml

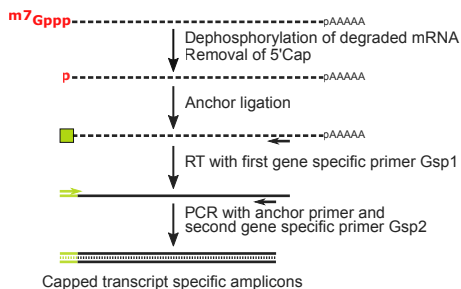
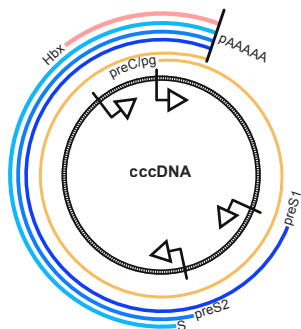
B.



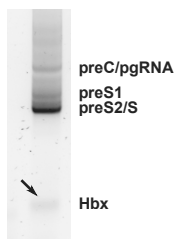
C.



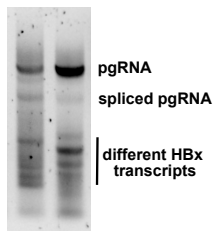
Full length HBV 5'RACE



Intracellular viral RNAs



Encapsidated viral RNAs



Circulating viral RNAs

