Caractérisation des ARN VHB circulants dans l'infection chronique par le VHB

Alexia Paturel

N°d'ordre NNT : 2022LYSE1053

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N°340
Biologie Moléculaire Intégrative et Cellulaire (BMIC)

Spécialité de doctorat : Virologie et Biologie Cellulaire
Discipline : Infectiologie

Soutenue à huis clos le 17/03/2022, par :
Alexia PATUREL

Characterisation of circulating HBV RNAs in chronic HBV patients

Caractérisation des ARN VHB circulants dans l'infection chronique par le VHB

Devant le jury composé de :

M. Pr. ZOULIM Fabien, PU-PH, CRCL HCL, Lyon	Président
M. Dr. SOUSSAN Patrick, MCU-PH, AP-HP, CRSA, Paris	Rapporteur
Mme. Dr. DANDRI Maura, PU, UKE, DZIF, Hamburg (Allemagne)	Rapporteur
Mme. Pr. POLLICINO Teresa, PU-PH, Université de Messine (Italie)	Examinateur
Mme Dr. HEIL Marintha, PhD, Roche Molecular Systems, Pleasanton (USA)	Examinateur
M. Pr. LEVRERO Massimo, PU-PH, CRCL HCL, Lyon	Directeur de Thèse

Université Claude Bernard – LYON 1

Président de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	Mme Céline BROCHIER
Vice-Président de la Commission de Recherche	M. Petru MIRONESCU
Directeur Général des Services	M. Pierre ROLLAND

COMPOSANTES SANTE

Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT
Faculté d'Odontologie	Doyenne : Mme Dominique SEUX
Faculté de Médecine et Maïeutique Lyon Sud - Charles Mérieux	Doyenne : Mme Carole BURILLON
Faculté de Médecine Lyon-Est	Doyen : M. Gilles RODE
Institut des Sciences et Techniques de la Réadaptation (ISTR)	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques (ISBP)	Directeur : M. Claude DUSSART

COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE

Département Génie Electrique et des Procédés (GEP)	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
Ecole Supérieure de Chimie, Physique, Electronique (CPE Lyon)	Directeur : Gérard PIGNAULT
Institut de Science Financière et d'Assurances (ISFA)	Directeur : M. Nicolas LEBOISNE
Institut National du Professorat et de l'Education	Directeur : M. Pierre CHAREYRON
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Polytechnique Lyon	Directeur : Emmanuel PERRIN
UFR Biosciences	Administratrice provisoire : Mme Kathrin GIESELER
UFR des Sciences et Techniques des Activités Physiques et Sportives (STAPS)	Directeur : M. Yannick VANPOULLE
UFR Faculté des Sciences	Directeur : M. Bruno ANDRIOLETTI

« La chute n’est pas un échec. L’échec c’est de rester là où on est tombé. »
Socrate

« Falling down is not a failure. Failure comes when you stay where you have fallen »
Socrate

RESUME

L'infection chronique par le virus de l'hépatite B (VHB) reste un problème majeur de santé publique au niveau mondial avec 240 millions de porteurs chroniques. Les traitements actuels de l'infection chronique par le VHB ne permettent pas d'obtenir une guérison fonctionnelle, définie par la perte de l'AgHBs et permettant l'arrêt du traitement ainsi qu'une diminution du risque de complications hépatiques. Un grand nombre de nouvelles molécules thérapeutiques antivirales ciblant directement ou indirectement l'ADNccc nucléaire, l'intermédiaire de réplication servant de matrice pour la transcription de tous les ARNm viraux, afin de réduire le réservoir d'ADNccc et son activité transcriptionnelle et d'obtenir une guérison fonctionnelle avec une durée de traitement limitée, sont en cours d'évaluation clinique.

La suppression persistante de l'ADN du VHB dans le sérum et la perte des antigènes AgHBs sont actuellement les critères d'efficacité thérapeutique pour toutes nouvelles molécules mais ne sont pas appropriés pour les patients déjà traités par des analogues nucléotidiques diminuant l'ADN viral. En outre, bien que les niveaux d'AgHBs sont prédictifs de la perte de l'AgHBs, la cinétique de réduction des taux d'AgHBs est trop lente pour être utilisée dans le contexte d'un essai clinique de courte durée. De plus, l'AgHBs peut provenir des séquences intégrées du génome viral et ne pas refléter ni le nombre de molécules d'ADNccc, ni son activité transcriptionnelle. En revanche, l'ARN VHB circulant reflète l'activité transcriptionnelle de l'ADNccc dans le foie et serait donc un bon

biomarqueur non invasif pour le suivi de la maladie et l'évaluation de nouvelles molécules thérapeutiques.

L'établissement d'un standard d'ARN VHB reste un objectif important pour le développement de tous tests de quantification de l'ARN VHB circulant. Dans ce but, nous avons conçu plusieurs lignées cellulaires clonales portant des génomes mutés du VHB. La lignée cellulaire clonale (Huh7-3D29) sécrétant des niveaux élevés d'ARN du VHB et de très faibles quantités d'ADN du VHB a été générée et caractérisée.
Nous avons pu montrer que, comme déjà observé dans les sérums de patients VHB chroniques, les ARN-VHB sécrétés par le clone Huh7-3D29 sont principalement, mais pas exclusivement, des ARN de 3,5 kb et, que les ARN-VHB peuvent être aussi trouvés dans les exosomes. Des expériences de dilution en série et de reproductibilité utilisant une RT-qPCR en une étape conçue par Roche, ont montré que les surnageants des cellules Huh7-3D29 peuvent être utilisés avec les tests de quantification de l'ARN du VHB basés sur la PCR et constituent une source illimitée de standard ARN du VHB.

Bien qu'il existe des preuves solides à l'appui de l'utilité clinique de la quantification des ARN VHB circulants, l'impact des transcrits chimériques virus-hôte à partir de séquences intégrées du VHB sur les niveaux d'ARN VHB circulant reste à étudier. De plus, pour comprendre l'effet des thérapies VHB sur les séquences intégrées du VHB et sur leur transcription, il sera important de quantifier à la fois les ADN et les transcrits chimériques virus-hôte. Nous avons mis en place un protocole de capture suivi par séquençage à haut débit pour l'analyse des ADN et des ARN VHB. Les résultats obtenus jusqu'à présent montrent que notre protocole de capture/séquençage peut détecter et quantifier les intégrations du VHB dans des lignées cellulaires de CHC portant des séquences intégrées du VHB (par exemple, PLC/PRF/5 et Hep3B) ainsi que dans le foie d'un porteur chronique du VHB.

Dans l'ensemble, nos résultats contribuent au développement, à l'établissement et à l'interprétation de la quantification de l'ARN VHB circulant en tant que nouveau biomarqueur de l'infection chronique par le VHB.

SUMMARY

Hepatitis B virus (HBV) infection remains a major global public health problem with 240 million chronic carriers. Current treatments for chronic HBV infection are unable to achieve a functional cure, defined by the loss of the HBsAg and allowing the treatment cessation as well as a decreased risk of liver complications. Several new therapeutic molecules that target the cccDNA, the nuclear replicative intermediate that acts as a template for the transcription of all viral mRNAs, with the aim of silencing or reducing the pool of cccDNA to achieve a functional cure with finite treatment duration, are at the preclinical or early clinical stage of evaluation.

Persistent suppression of serum HBV DNA and HBsAg negativization are considered the most important primary efficacy endpoints for novel molecular entities in clinical development. However, the suppression of HBV DNA in serum is not an appropriate efficacy endpoint to evaluate new HBV treatments in patients treated with DNA-decreasing nucleotide analogues. In addition, although HBsAg levels may predict HBsAg loss, the kinetic of HBsAg decline is too slow in most patients to be useful to predict functional cure within the short duration of early clinical trials. Finally, HBsAg antigens may originate from HBV sequences integrated into the host genome and do not reflect the number of cccDNA molecules and their transcriptional activity.

Circulating HBV RNA have been shown to reflect the transcriptional activity of cccDNA in the liver and would thus be a good non-invasive biomarker for the functional inactivation of cccDNA or the reduction of the cccDNA reservoir.

The establishment of an HBV RNA standard remains an important objective for the development of all circulating HBV RNA quantification assays. To this aim, we have generated several clonal cell lines carrying mutated HBV genomes. A clonal cell line (Huh7-3D29) secreting high levels of HBV RNAs and very low amounts of HBV DNAs was generated and fully characterized. We showed that, similar to what observed in patients' sera, HBV-RNAs secreted from the Huh7-3D29 clone are predominantly, but not exclusively, 3.5 kb RNAs and that HBV-RNAs can be found in exosomes. Serial dilution and reproducibility experiments using a one-step RT-qPCR designed by Roche have shown that Huh7-3D29 supernatants are suitable for PCR-based HBV RNA quantification assays and constitutes an unlimited source of HBV RNA standard.

Although there is strong evidence supporting the clinical utility of circulating HBV RNAs quantification, the impact of virus-host chimeric transcripts from integrated HBV sequences on the levels of circulating HBV RNAs remains to be studied. Moreover, to understand the effect of HBV therapies on integrated HBV sequences and on their transcription, it will be important to quantify at the same time chimeric virus-host DNAs and transcripts. We have set up a capture/NGS sequencing workflow for both HBV DNA and RNA analysis. The results obtained so far show that our capture/sequencing workflow can detect and quantify HBV integrations in HCC cell lines (e.g., PLC/PRF/5 and Hep3B) carrying HBV integrated sequences as well as in the human liver of a chronic HBV carrier.

Altogether, our results contribute to the development, establishment and interpretation of serum HBV RNA as a new biomarker for chronic HBV infection.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

aa	Amino acids
AASLD	American association for the study of the liver disease
ALT	Alanine Aminotransferase
BCP	Basal core promoter
CAMs	Capsid assembly modulators
cccDNA	Covalently closed circular DNA
ChHBV	chimpanzee hepatitis B virus
CLIA	Chemiluminescence immunoassay
CTD	C-terminal domain
CURS	Core Upstream Regulatory Sequence
ddPCR	Digital droplet PCR
DHBV	Duck hepatitis B virus
DHSC	Department of health and social care
DNA	Deoxyribonucleic acid
DR1	Direct repeat 1
DR2	Direct repeat 2
dslDNA	Double-strand linear DNA
EASL	European association for the study of the liver
ECLIA	Electrochemiluminescence immunoassay
EGFR	Epidermal growth factor receptor
EMT	Epithelio-mesanchymal transition
EN	Enhancer
ER	Endoplasmic reticulum
ESCRT	Endosomal sorting complexes required for transport

FDA	Food and Drug Administration
GiHBV	Gibbon hepatitis B virus
GRE	Glucocorticoid-responsive element
GSHV	Ground squirrel hepatitis virus
HAP	Heteroaryldihydropyrimidines
HBcAg	HBc antigen
HBeAg	HBe antigen
HBsAg	HBs antigen
HBSP	Hepatitis B spliced protein
HBV	Hepatitis B virus
HBx	Hepatitis B x protein
HCC	Hepatocellular carcinoma
HCV	Hepatits C virus
HDV	Hepatits Delta virus
HIV	human immunodeficiency virus
HNF1	Hepatocyte nuclear factor 1
Hsp	Heat-shock protein
HSPG	Heparan sulfate proteoglycans
HTS	High troughput sequencing
IBB	Importin β binding
IFN	Interferons
invPCR	Inverse nested PCR
IS	International standard
ISRE	
IU	International units
kb	Kilobase
LHBs	Large hepatitis B virus surface protein
LINE	Long interspersed nuclear element
lncRNA	Long non-coding RNA
LTR	Long terminal repeat
MHBs	Medium hepatitis B virus surface protein
MMEJ	Microhomology-mediated end joining
MVBs	Multivesicular bodies
NA or Nucs	Nucleos(t)ide analogues
NAP	Nucleic acid polymers
NER	Nucleotide excision repair
NES	Nuclear export signal
NFκB	Nuclear factor kappa B
NG	N-glycosylation
NGS	Next generation sequencing
NHEJ	Non-homologous end-joining
NIBSC	National Institute for Biological Standards and Control
NK	Natural killer
NLS	Nuclear localization signals

NPC	Nuclear pore complex
NRE	Negative regulatory element
nt	Nucleotides
NTCP	Na+ taurocholate cotransporting polypeptide
NTD	N-terminal domain
OG	O-glycosylation
ORF	Open reading frame
OuHBV	Orangutan hepatitis B virus
PAGE	Polyacrylamide gel electrophoresis
PAMP	Pathogen Associated Molecular Pattern
PAS	Polyadenylation signal
PCR	Polymerase chain reaction
PF-rcDNA	Protein free-rcDNA
pgRNA	Pregenomic RNA
Pol	Polymerase
Poly or PAS	Polyadenylation signal
PPA	Phenylpropenamide
PRE	Post-transcriptional regulatory element
PreC	Precore
PRR	Pattern Recognition Receptors
PTB	Polypyrimidine track binding protein
r	Terminal redundancy
RACE	Rapid amplification of complementary DNA (cDNA)-ends
rcDNA	Relaxed circular DNA
RNA	Ribonucleic acid
RT	Reverse transcriptase
SBA	Sulfamoylbenzamide
SHBs	Small surface protein of the hepatitis B virus
SINE	Short interspersed nuclear element
SP1	Spliced protein 1
SSA	Strand annealing
ssDNA	Single-stranded DNA
TDP2	Tyrosil DNA phosphodiesterase 2
TNF-α	Tumor Necoris Factor alpha
TSS	Transcription starting site
UDG	Uracil DNA glycosylase
URR	Upper regulatory region
VCJ	Virus-cell junction
WES	Whole exome sequencing
WGS	Whole genome sequencing
WHO	World health organization
WHV	Woodchuck hepatitis virus
WMHBV	Woolly monkey hepatitis B virus

INTRODUCTION

CHAPTER I: THE HEPATITIS B VIRUS

1.1 Hepatitis B Virus discovery

In 1963, the American physician/geneticist Baruch Samuel Blumberg, who worked on the variability of responses and susceptibility to disease at the Fox Chase Cancer Center in Philadelphia, discovered what he called the "Australia Antigen" (HBsAg) in the serum of an Australian Aboriginal person (Blumberg et al., 1965). In 1968, this protein was found to be part of the virus that causes "serum hepatitis" (hepatitis B) by virologist Alfred Prince. In 1970, David Dane observed for the first time the hepatitis B virus (HBV) by electron microscopy, leading to the naming of HBV viral particles as Dane particles (Dane et al., 1970). In 1976, Blumberg won the Nobel Prize in Physiology or Medicine for his work on hepatitis B.

HBsAg became soon the serological marker for hepatitis B infection and was subsequently purified from the plasma of chronic carriers to prepare the first vaccines against HBV. During the next few years, a series of observations in humans and primates by scientists including Maurice Hilleman, Saul Krugman, Robert Purcell in the US, and Philippe Maupas in France provided additional support for the vaccine. Maupas' team obtained the first results of vaccination against HBV in 1975 and in 1980 the results of the first field trial were published by W. Szmuness and his colleagues in New York City. (Figure 1)

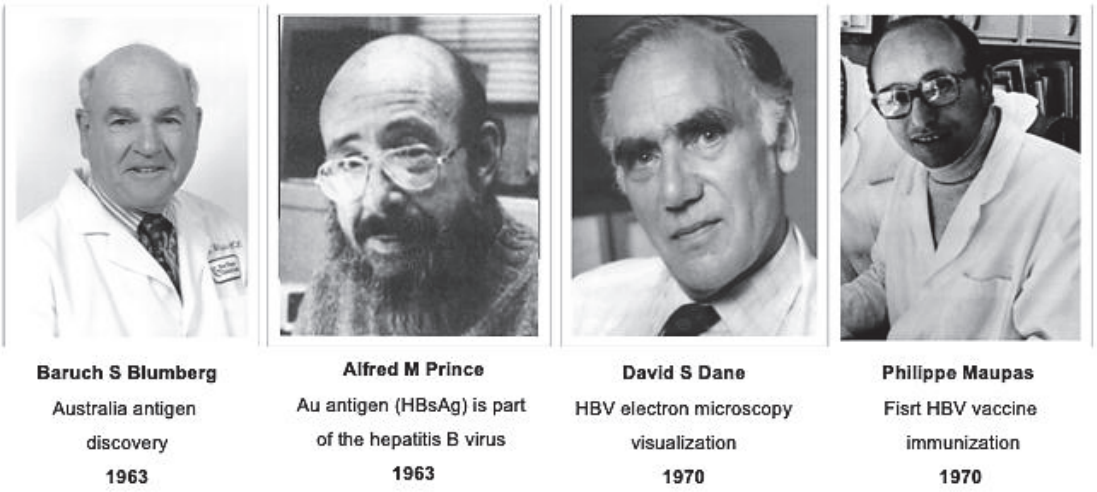


Figure 1 - First HBV research actors

1.2 Virus classification

HBV belongs to the category of para-retroviruses or class VII of David Baltimore's classification and more specifically to the *Hepadnaviridae* family. This includes hepatotropic viruses infecting mammal (*Orthohepadnaviridae*) or avian species (*Avihepadnaviridae*).

HBV infects humans and chimpanzees. Other *Orthohepadnaviruses* infect groundhogs (WHV or woodchuck hepatitis virus), squirrels (GSHV or ground squirrel hepatitis virus), and monkeys (WMHBV or Woolley Monkey hepatitis virus, ChHBV or Chimpanzee hepatitis B virus, OuHBV or Orangutan hepatitis B virus, and GiHBV Gibbon hepatitis B virus). *Avihepadnaviruses* infect the Peking duck (DHBV or Duck hepatitis virus) and heron (HHV or heron hepatitis B virus) (Lefkowitz et al., 2018; Rasche et al., 2016).

1.3 HBV Structure

The HBV is a small enveloped (42nm in diameter) and encapsidated (27 nm in diameter) virus that contains a 3.2 kb relaxed circular double-stranded DNA genome (rcDNA) (Landers et al., 1977; Summers et al., 1975). The viral envelope consists of the surface glycoproteins S (small), M (medium) and L (large) in a 4:1:1 ratio, embedded in a lipid bilayer from the endoplasmic reticulum and the multivesicular bodies (Sheu and Lo, 1992) (Figure 2). The envelope covers an icosahedral capsid

formed by core protein dimers that protect the viral genome. There are two forms of capsids: the so-called T=3 form, composed of 90 dimers of core proteins, and the more common T=4 form, composed of 120 dimers (Zlotnick et al., 2015).

The HBV DNA genome is a double stranded and contains four overlapping open reading frames (ORF) (P, PreC/C, PreS/S and X) with every base coding for at least one ORF. The nucleotide numbering of the HBV genome is based, for all HBV genotypes with the exception of genotype G, on the unique EcoRI restriction enzyme site. The HBV genome displays six initiation codons, four promoters (preS1, preS2, core and X) and two enhancers (ENI and ENII) located upstream of the core promoter (Quarleri, 2014). Additional elements that regulate viral transcription, such as the glucocorticoid-responsive element (GRE), the negative regulatory element (NRE) and the IFN-stimulated response element (ISRE) element are present in the HBV genome and will be further described in section 1.4.4.1.4 Transcriptional regulation: HBV enhancers and promoters.

Seven viral proteins (polymerase, HBe, SHBs, MHBs, LHBs, core and X) are translated from the six major polyadenylated and capped messenger RNAs (pgRNA, 3.5 kb, 2.4 kb, 2.1 kb, and 0.7 kb RNAs) (Nassal and Schaller, 1993). HBV transcripts and proteins will be further described in sections 1.4.4 Transcription and 1.4.5 Translation.

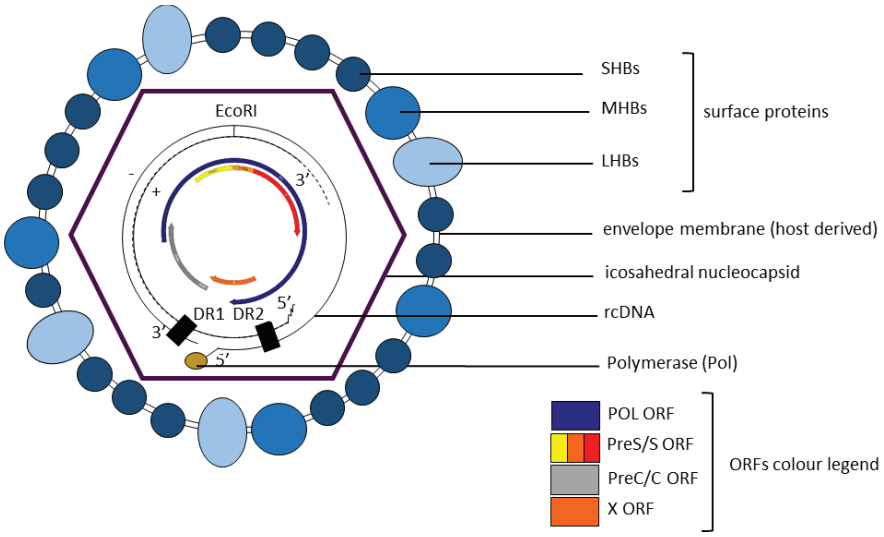


Figure 2 - Schematic representation of the HBV virion structure

1.4 HBV replication cycle

This section will address the replication of the viral cycle and will be based on Figure 3. Throughout the text, the major stages of this cycle will be denoted by chronological order depicted by numbers in brackets and shown in Figure 3.

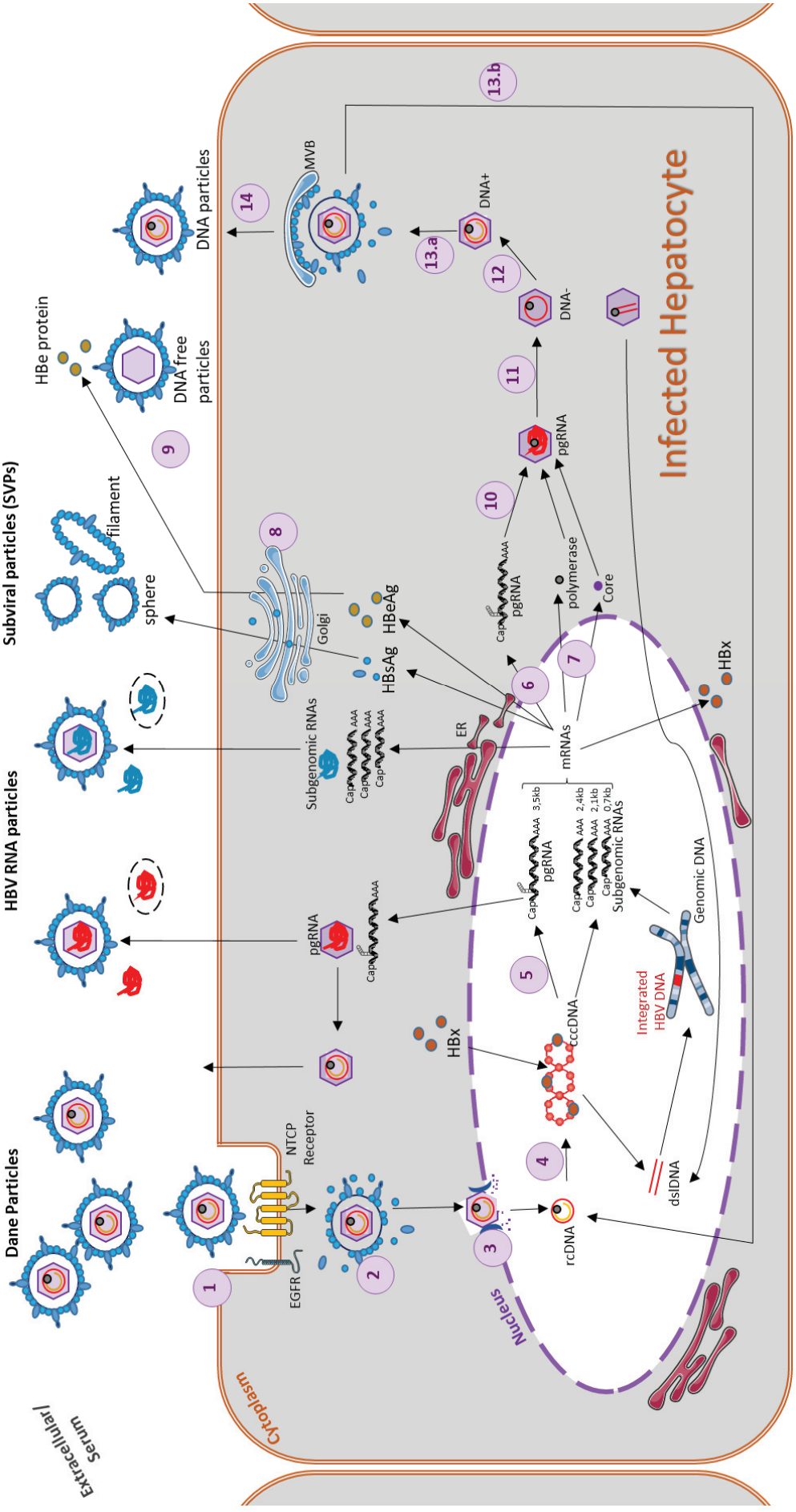


Figure 3 - HBV replication cycle
RE is from smart.servier.com website
1- Recognition of the hepatocyte by virion
2- Internalization of virion by endocytosis
3- Capsid transport to the nucleus
4- Conversion of rcDNA in cccDNA
5- RNA transcription
6- RNA transport to the cytoplasm
7 - Translation of viral proteins
8- Envelope proteins directed to the ER
9- HBeAg and subviral particles secretion
10- Encapsidation of pregenomic RNA and the polymerase
11- Reverse transcription of the long strand of DNA -
12- Short strand (+) synthesis
13.a- The nucleocapsids are directed to the ER
13.b- Nucleocapsids recycled to the nucleus
14- Secretion of viral particles

1.4.1 Viral entry and intracellular transport

The hepatitis B virus can infect lymphocytes, pancreatic cells, macrophages and hepatocytes (Mason et al., 1993; Seeger and Mason, 2000) but HBV replication occurs only in hepatocytes. The recent evidence that HBV replicates also in lymphocytes remains to be confirmed (Yan et al., 2016).

The viral replication cycle can be subdivided into different steps.

(1) First the virion recognizes the hepatocyte by:

- A low affinity interaction between the a-determinant of the 3 HBs proteins (common to SHBs, MHBs and LHBs proteins) within the virion envelope and the Heparan sulfate proteoglycans (HSPG) present on the surface of the hepatocyte (Sureau and Salisse, 2013). Among heparin sulfate family, the glypican 5 was demonstrated to specifically mediate HBV entry into hepatocyte (Verrier et al., 2016).
- A higher affinity interaction between the myristoylated PreS1 domain of the L envelope protein (Schulze et al., 2007) and the sodium taurocholate co-transporting polypeptide (NTCP). This bile acid transporter has been identified as an HBV receptor (Yan and Li, 2015). Interestingly, internalization of HBV virion through NTCP receptor is potentiate by epidermal growth factor receptor (EGFR) (Iwamoto et al., 2019). Thus, it is not surprising to found undetectable EGFR expression by western blot in HepG2 cell line that have a low rate of HBV infection (Zhao et al., 2013).

(2) The virion is then internalized within the cell by clathrin-dependent endocytosis leading to the release in the cytoplasm of the capsid containing the viral genome (Herrscher et al., 2020a). HBV genetic material can act as a Pathogen Associated Molecular Pattern (PAMP) and be recognized by the Pattern Recognition Receptors (PRR) in infected hepatocytes leading to the activation of the innate immune response. To prevent PRR recognition the incoming viral genomes are maintained within the viral capsids until they are transported into the nucleus.

(3) The virions are transported to Rab5-containing early endosomes and then by the Rab7-containing late endosomes (Figure 4). After, capsid release into the cytoplasm, they are translocated to the

nuclear pore complex (NPC) by retrograde active transport via microtubules (Jiang and Hildt, 2020; Macovei et al., 2013; Rabe et al., 2006).

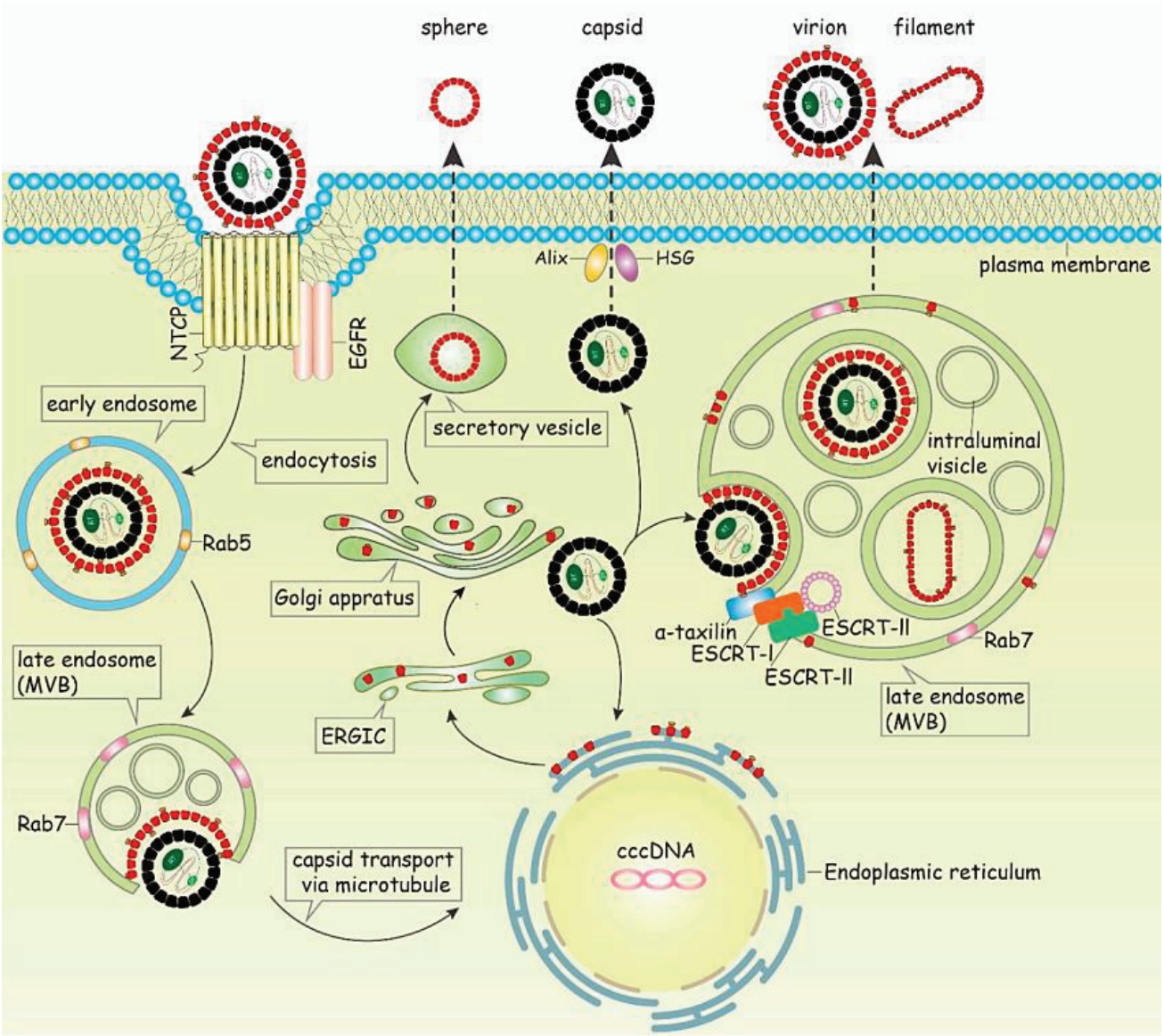


Figure 4 - Viral entry and HBV particles release
From (Jiang and Hildt, 2020)

Once at the NPC, the capsid dissociates and allow the release of the Pol-linked rcDNA. The size of the polymerase limits Pol-linked rcDNA simple diffusion to the nucleus, that is restricted to <42kDa and suggests an active transport relying on the polymerase bipartite nuclear localization signals (NLS) that is exposed by casein kinase II (CKII)-dependent phosphorylation (Lupberger et al., 2013)

HBc proteins from the capsids disassembled at the NPC can dimerize again to bind and regulate the functions of the nuclear cccDNA minichromosome (see below, section 1.4.2 cccDNA formation and regulation) (Bock et al., 2001) and to form empty capsids in the nucleus, whose function is still unclear.

The C-terminal domain of HBc contains two NLS and one importin β binding (IBB) domain. The importin α and β are karyopherin-like proteins that allow import from the outside to the inside of the nucleus. HBc NLS binding to importin β requires importin-α while the IBB domain allows the direct interaction to importin β (Gallucci and Kann, 2017). Importantly, the interaction between the C-terminal domain of the HBc protein and importins is highly dependent on the exposure of the C-terminal domain and the maturation status of the capsid and only unassembled HBc proteins can expose the C-terminal IBB domain.

The HBV genome, present in circulating viral particles and in HBV capsids, transported to the NPC after infection to be released in the nucleus, is the double-stranded relaxed circular DNA (rcDNA) composed of a sense (-) and an antisense (+) strand joined by a cohesive overlap of 244 bp (Blumberg, 1997). The 5'-end region of the sense strand contains 8 nucleotides redundancy that is covalently attached to the tyrosine residue 63 of the viral polymerase. The positive strand is incomplete at the 3' from 600 to 2100 nt (Summers et al., 1975). Two 10-12nt repeat sequences designated DR1 and DR2 stabilize the partially double stranded rcDNA (Figure 5).

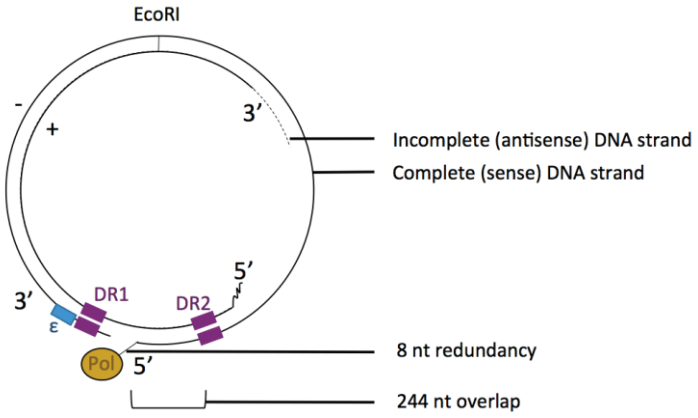


Figure 5 - Schematic representation of the HBV relaxed circular DNA (rcDNA), HBV genome contained in Dane particles

1.4.2 cccDNA formation and regulation

4) In the nucleus, the rcDNA is converted into a minichromosome, the covalently closed circular DNA (cccDNA). This step is crucial as the rcDNA is not competent for viral mRNA synthesis and requires the conversion into a chromatinized cccDNA. The rcDNA has a notched ring structure similar to damaged cellular DNA leading to its repair by intracellular DNA repair mechanisms.

The rcDNA has four distinct features that need to be taken care of to allow the conversion into a cccDNA:

1) the HBV polymerase covalently bound to the minus strand of the rcDNA,

2) the Pol-protein linked sequence,

3) a 5'capped RNA primer,

4) the gap on the rcDNA plus strand.

Thus, cccDNA is a complex process that involves several actors and multiple steps.

a) First, the viral polymerase, which is covalently attached to the rcDNA is removed. Several mechanisms have been proposed (Wei and Ploss, 2021)

1. the host tyrosyl DNA phosphodiesterase (TDP2) enables the phosphodiester bond to be cleaved to generate a protein free-rcDNA (PF-rcDNA) devoid of HBV polymerase (Königer et al., 2014). However, TDP2 is not essential for the formation of cccDNA as when it is depleted, HBV infection can still occur (Cui et al., 2015).
2. the removal of HBV polymerase and the Pol protein-linked sequence (of about 8 nt), also called terminal redundancy (r) sequence by DNA endonucleases such as FEN1;
3. the removal of the viral polymerase by a serine protease digestion step;

b) second, the cleavage of the r sequence is carried out by the FEN1 protein followed by ligation of the generated nick by LIG1 and LIG3 (Kitamura et al., 2018).

c) finally, the repair of the incomplete positive strand is performed. It requires the interaction of PCNA with Pol delta that will synthesize together the 3'end of the plus strand and displace the RNA primer for further degradation. The RNA primer is removed from the positive DNA strand by FEN1 activity or enzymes with RNAse H activity such as HBV polymerase. This generates a nick that is sealed by LIG1 and LIG3.

However, the completion of the positive strand can also be performed by other cellular polymerases such as Pol κ, Pol ι and Pol η.

Once formed the cccDNA supercoils and assembles with cell histones and non-histone proteins to form a minichromosome (Figure 6).

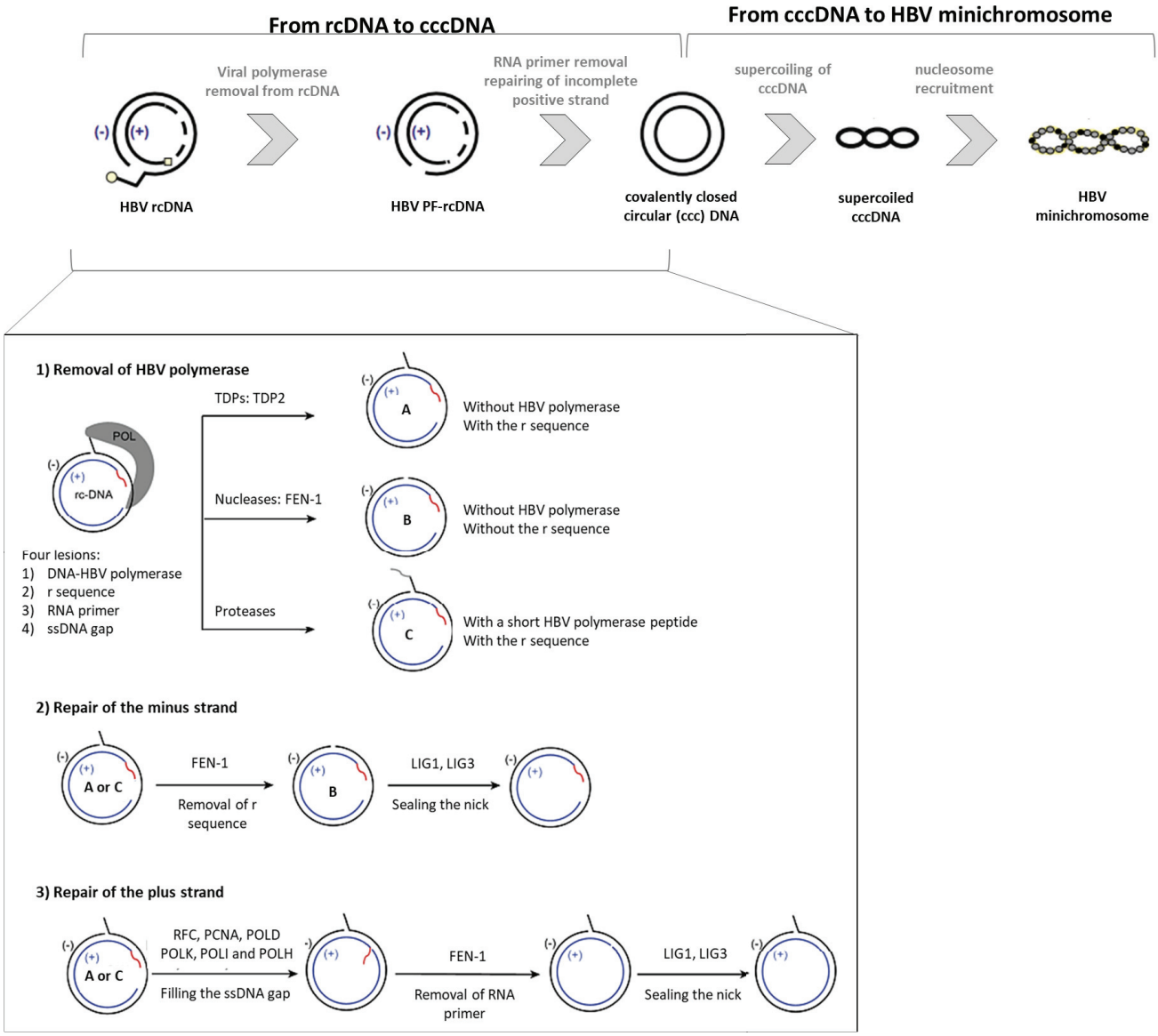


Figure 6 - Schematic representation of the conversion of rcDNA into the viral cccDNA minichromosome adapted from (Levrero et al., 2009) and from (Wei and Ploss, 2021)

This process occurs rapidly in HBV infected cells as the cccDNA is detected 24 hours after infection (Qi et al., 2016). The cccDNA half-life is still the object of some controversy but it appears that cccDNA is very stable in quiescent hepatocytes and may remain intact throughout the life of the hepatic cell. The cccDNA reservoir varies from 1 to 50 copies per infected hepatocyte: averaging up to 17 copies in the duck model (Zhang et al., 2003), up to 50 copies in the woodchuck model (Moraleda et al., 1997) but much less in the human model (up to 10 copies). Notably, the number of cccDNA may vary from one infected cell to another in each model (Köck et al., 2010).

(5) The cccDNA minichromosome serves as a template for the transcription of viral messenger RNAs. There are two types of viral RNAs: subgenomic RNAs, encoding the envelope and the HBx proteins, and the 3.5 kb RNAs that are longer than the genome 3.2 kb genome and include the precore RNA, translated into the precore precursor of the viral HBeAg, and the pregenomic RNA (pgRNA) that is translated into the viral core and Pol proteins and serves as replicative intermediate. (6) The viral RNAs will be transported into the cytoplasm and translated to produce the viral proteins (7). The transcription of viral RNAs will be further detailed in section 1.4.4 Transcription. Several cellular transcriptional factors, including HNF3 and HNF4, the RXRα, PPARα, FXR nuclear receptors, C/EBP and CREB, have been shown to bind the cccDNA minichromosome. Several histone modifying enzymes, including the CBP/p300, PCAF/GCN5, and HAT1 acetylases, the HDAC1, HDAC4, SIRT1 and SIRT3 deacetylases, the SETDB1, PRMT1, PRMT5, Set1A methyltransferases and the LSD1 demethylase, also bind the cccDNA minichromosome (Alarcon et al., 2016; Belloni et al., 2009; Rivière et al., 2015; Zhang et al., 2017). cccDNA transcription appears to be dependent on the acetylation of histones H3 and H4 bound to cccDNA in cells, humanized mouse HBV infection models (Pollicino et al. 2006; Belloni et al 2019; Belloni et al 2012) and in chronic hepatitis B patients (Pollicino et al., 2006). In addition to host transcription factors and chromatin modifiers, the HBx and possibly the HBc viral proteins also play a role in the regulation of HBV minichromosome transcriptional activity (Belloni et al., 2009; Lucifora et al., 2011).

The pgRNA is then encapsidated with the viral polymerase (10). Viral genome synthesis proceeds within the capsids to generate rcDNA HBV genomes through multiple steps that are detailed in the section 1.4.3 Viral genome replication. Mature rcDNA-containing nucleocapsids are either shuttled to the nucleus to amplify and maintain the cccDNA pool (*recycling pathway*) or assembled with mature envelope proteins to be secreted outside of cells (*egress pathway*).

1.4.3 Viral genome replication

Hepadnaviruses replicate their circular rcDNA genome by the reverse transcription of the 3.5 kb pregenomic RNA replicative intermediate. The pgRNA contains a stem-loop structure of 60 nucleotides in 5', also called ε (epsilon) loop. This ε sequence is also found at the 3’ end of all HBV transcripts, but unlike the 5’ ε loop, its function is still unclear. The ε loop is highly conserved in hepadnaviruses and harbors, from 5’ to 3’, a lower stem, a central bulge, an upper stem, and an apical

loop. The pgRNA also contains the DR1 and DR2, 10-12 nucleotides repeated sequences. These repeated sequences are thought to be involved in maintaining a circular structure, which is essential for the replication.

Replication is initiated by the binding of the polymerase to the ε sequence of the pgRNA, followed by the recruitment of core proteins/dimers to form the viral capsid. Thus, the polymerase is encapsidated with the pgRNA by interacting with the encapsidation signal ε (Bartenschlager et al., 1990; Hirsch et al., 1990; Knaus and Nassal, 1993). This process is facilitated by the Hsp90, Hsp70, Hsp40 chaperones and their respective co-chaperones as well as by the heat stress cognate 70 (Hsc70) protein (Hu and Seeger, 1996a; Hu et al., 2002). The RNA-binding motif protein 24 (RBM24) also interacts with HBV Pol and the ε loop on the pgRNA to mediate the encapsidation process (Yao et al., 2019). Together with the pgRNA, cellular factors are also included in the capsid such as eukaryotic translation initiation factor 4 (eIF4E), DEAD-box RNA helicase DDX3 and the APOBEC3G protein (Kim et al., 2010; Nguyen and Hu, 2008; Wang et al., 2009).

The ε sequence serves as a template to synthetize a short oligomer (GAA or GTAA for DHBV) (Wang and Seeger, 1992) that is covalently bound to the tyrosine in position 63 (Y63) of the viral polymerase and is translocated, after its synthesis, from the N-terminal DR1 domain to the C-terminal DR1 domain (Lanford et al., 1997; Zoulim and Seeger, 1994). This allows the polymerase to change its matrix and synthesize the negative strand of the rcDNA using the translocated oligomer as a primer (11). In parallel, pgRNA is degraded by the RNase H activity of the polymerase with the exception of 11 to 18nt of the DR1 sequence in 5’ (Summers and Mason, 1982) that are protected from digestion probably by the steric hindrance exerted by the RT/RNase H domain. The undegraded DR1 sequence is then transferred to the DR2 repeated sequence and serves as a primer for the viral positive strand synthesis. The synthesis of the incomplete/positive strand is made possible by the redundancy at the 5' and 3' ends of the negative DNA (12) (Will et al., 1987). Two possible explanations for the incomplete synthesis of the HBV DNA positive strand have been proposed: a) a lack of a sufficient amount of dNTP trapped in the capsid and b) a lack of space in the capsid. This latter hypothesis is questionable since the capsid has been show to have a considerable structural plasticity (Böttcher et al., 2006).

If the translocation of the undegraded DR1 primer does not occur, the positive strand is still synthesized but the circularization does not occur, resulting in the generation of a linear double stranded DNA (dslDNA) (Staprans et al., 1991) (Figure 7). Mature capsids containing a dslDNA can be

recycled to the nucleus and the dslDNA converted to cccDNA by non-homologous end-joining (NHEJ) but the cccDNA molecules derived from a dslDNA are defective for rcDNA synthesis (Bill and Summers, 2004). The dslDNA corresponds to 3-26% of the total viral DNA with important patient to patient variations (Zhao et al., 2016b)

Overall, it can be estimated that nucleocapsids particle contain in 90% of cases the rcDNA, in approximately 10% of cases the dslDNA and in 1% unspliced HBV RNA (Iannacone and Guidotti, 2021).

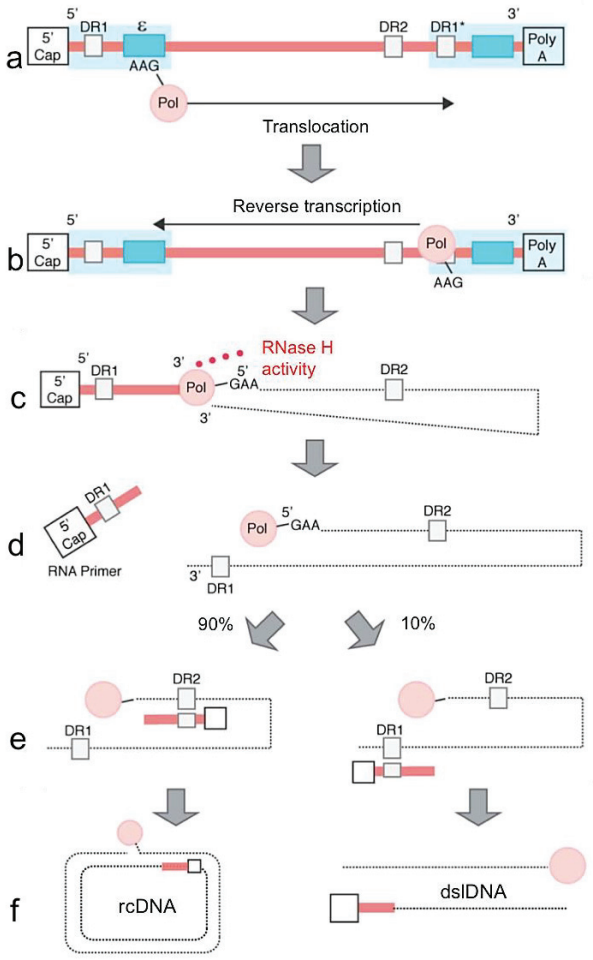


Figure 7 - HBV reverse transcription process from (Tu et al., 2017)

a. Binding between the polymerase and the ε loop of the pgRNA,

Synthesis of 3nt from the ε template

b. Reverse transcription: synthesis of the negative strand of the rcDNA

c. pgRNA degradation by the RNase H activity of the polymerase

d. 11-18nt of the pgRNA is not degraded

e and **f**. In 90%, the remaining nt are transferred to the DR2 to form rcDNA.

In 10%, the nt transfer does not occur and forms the dslDNA

As already mentioned, a small, variable proportion of newly formed mature capsids containing rcDNA can be recycled in the nucleus (13.b) to increase the cccDNA reservoir whereas the majority is enveloped to form infectious particles. Mature rcDNA-containing capsids acquire the host-derived lipid bilayer in which viral envelope proteins are embedded by budding into late endosomes or multi-vesicular bodies (MVBs) (Jiang et al., 2015) (13.a). Budding to MVBs occurs through the endosomal sorting complexes required for transport (ESCRT) I, II and III and is mediated by alpha-taxilin, gamma 2-adaptin, vps4 and Nedd4 (Jiang and Hildt, 2020; Lambert et al., 2007; Patient et al., 2009) (14). Conversely, the extracellular release of naked capsids involves Alix and HSG proteins.

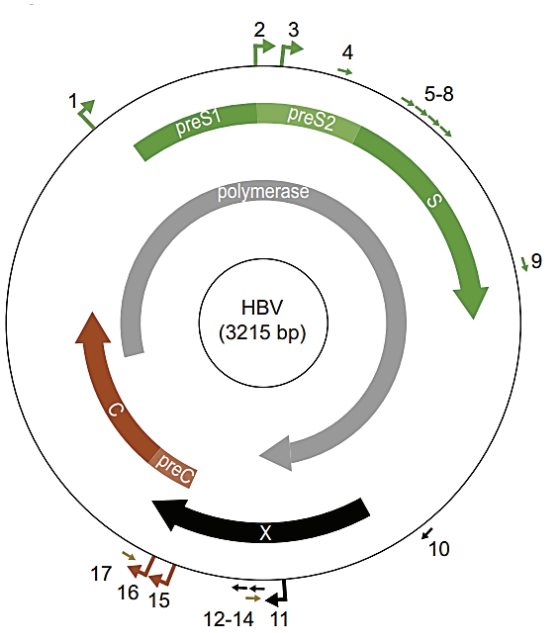
Virion production varies from patient to patient according to the different phases of the HBV infection/ disease and can be as high as 10^{11} virions produced per day, a rate higher than that of the human immunodeficiency virus (HIV). HBV virions average half-life in plasma has been estimated to be about 1-3 days (Quarleri, 2014).

1.4.4 Transcription

The cccDNA minichromosome is used as a template by the cellular RNA polymerase type II for the synthesis of all viral messenger RNAs, classified, according to their size, into subgenomic mRNAs and 3.5 kb RNA. Six major RNAs are transcribed: 3.5 kb RNAs (pgRNA and precore RNA), one 2.4 kb RNA (PreS1 RNA), two 2.1 kb RNAs (PreS2 RNA and S RNA) and one 0.7 kb RNA (HBx). These RNAs are translated into the different viral proteins (*Table 1*) that are further described in the section 1.4.5 Translation. All transcripts have a 5' cap and a 3' polyadenylation, making them indistinguishable from cellular RNAs. The polyadenylation signal (PAS) is common to all viral RNAs.

In addition to the 6 transcription starting site (TSS) for the 6 major HBV transcripts listed above, 11 novel TSS have been identified in HCC/nontumor liver pairs from HBV positive patients and in cell models of HBV infection using the cap analysis of gene expression (CAGE) technique (Altinel et al. 2016). Interestingly, the transcription profile in HCC and nontumor liver were similar with some variability in the quantitative analysis. From the highest to the lowest, the levels of HBV transcripts in the liver of HBV patients is as follows: (1) PreS2, (2) pgRNA, (3) S, (4) PreS1, (5) PreC and (6) X (Figure 8). Notably, the main transcript detected by CAGE in eight HBV-positive blood sample is the pgRNA (Altinel et al., 2016). The relative proportion of the different viral RNAs might change in response to therapy and, using a 5-RACE method, it has been shown that interferon-β leads to a

strong reduction of preC and pgRNAs but has only a moderate effect on the other viral transcripts (Stadelmayer et al., 2020).



Major peaks (>1,000 raw tags)

#	Gene	Start	End	Total count	Strand	Model
1	PreS1	2802	2807	12,748	+	Y
2	PreS/S	3183	19	201,883	+	Y
3	S	25	121	19,004	+	Y
11	X	1523	1524	1,013	+	Y
15	PreC/C	1745	1746	1,106	+	N
16	pgRNA	1816	1819	112,452	+	Y

Minor peaks (100 ~ 1,000 raw tags)

#	Gene	Start	End	Total count	Strand	Model
4	S	132	135	248	+	N
5	S	203	207	527	+	Y
6	S	258	262	282	+	N
7	S	284	285	528	+	Y
8	S	302	303	118	+	N
9	S	759	760	729	+	Y
10	X	1210	1211	124	+	Y
12	AS	1557	1610	157	-	Y
13	X	1572	1573	117	+	N
14	X	1655	1656	316	+	N
17	AS	1822	1834	207	-	N

Figure 8 - Quantification of HBV transcripts and comprehensive TSS map of HBV in chronically infected non tumor livers and HCC using CAGE Assay from Altinel et al study. Model column on the right indicates by "N" for no or "Y" for yes whether the TSS peak is also found in HBV replicating model.

1.4.4.1 Unspliced HBV RNAs

1.4.4.1.1 3.5 kb RNAs

pgRNA and precore RNA 3.5 kb RNAs are both 1.1-fold longer than the HBV genome and contains redundant sequences, notably two DR1 regions and one DR2 region. 3.5 kb RNAs are also characterized by the presence of two ε loops at the 5' and 3'end. The 5' ε loop of pgRNA allows the binding of the polymerase during rcDNA synthesis. Conversely, the 5' ε loop of the precore RNA does not function in encapsidation due to the translation of the precore protein that melts the RNA structure (Chen and Brown, 2012). Of note, the 3' ε loop is shared by all the viral transcripts and is not sufficient to mediate encapsidation (Rieger and Nassal, 1996).

The precore RNA transcript is 30 bp longer than the pregenomic RNA and serves as a template for the translation of the HBeAg protein. The pgRNA, in addition to serve as template to be encapsidated and reverse-transcribed to generate rcDNA new viral genomes, is a bicistronic RNA that encodes for the Pol and core/HBc proteins. Bicistronic RNAs are rare in eukaryotic organism because ribosomes

usually bind to the 5'-cap of the mRNA and scan the mRNA from 5' to 3' end until the first AUG where translation will start. Downstream ORFs are usually very poorly translated unless specific mechanisms such as leaky scanning, internal initiation and reinitiation occur to allow translation initiation at internal AUG (Ryabova et al., 2002). In the case of the HBV pgRNA, the core protein is more translated than the Pol protein.

The pgRNA and precore RNA transcription is directed by two distinct promoters within the core promoter region (Yu and Mertz, 1996).

1.4.4.1.2 2.4 kb and 2.1 kb RNAs

The 2.1 kb and 2.4 kb RNAs code for the small (S), medium (M) and large (L) envelope/surface proteins. The 2.4 kb RNA transcription is controlled by the PreS1 promoter and encodes for the LHBs protein. The transcription of the two 2.1 kb RNAs (PreS2 RNA and S RNA) is controlled by the PreS2 promoter and encode respectively for the MHBs and the SHBs proteins.

1.4.4.1.3 HBx RNAs

The 0.7 kb transcript codes for the HBx protein. Another smaller and highly represented HBx transcript, detected by CAGE, likely encodes a shorter form of the X protein (Altinel et al., 2016). Shorter HBx transcripts have been also identified using the 5'-RACE technique (Stadelmayer et al., 2020).

1.4.4.1.4 Transcriptional regulation: HBV enhancers and promoters

Four viral promoters (preS1, preS2, core and X) and 2 enhancers (EN1 and EN2) as well as positive and negative regulatory elements distributed all over the viral genome have been identified (Figure 10) (Moolla et al., 2002). Except for the preS1 promoter, HBV promoters do not contain the classical TATA-box sequence.

The **core promoter** (nt 1575 to nt 1849) controls transcription initiation of the precore RNA and pgRNA (Yaginuma and Koike, 1989). It consists of a basal core promoter (BCP) and an upper regulatory region (URR) (Figure 9). **The BCP** (nt 1743 to nt 1849) is sufficient for accurate initiation of both precore mRNA and pgRNA transcription *in vivo*. The BCP domain has two distinct non-canonical TATA boxes specific for pgRNA and precore RNA transcription, and several distinct cis-acting elements for ubiquitous or hepatic nuclear factors that independently direct transcription of pgRNA and precore RNA (Chen et al., 1995; Yu and Mertz, 1996; Yuh et al., 1992).

The **URR** (nt 1613 to nt 1742) includes the core upstream regulatory sequence (CURS) that contains two domains (CURS-A and CURS-B). The 2 CURS domains act independently and in concert as a positive regulator of the BCP (Moolla et al., 2002). Upstream of the CURS element, is located the negative regulatory element (NRE) that is divided into 3 subregions: NRE-α, NRE-β and NRE-γ. The NRE represses the core promoter when bound by NRE-binding protein (Sun et al., 2001).

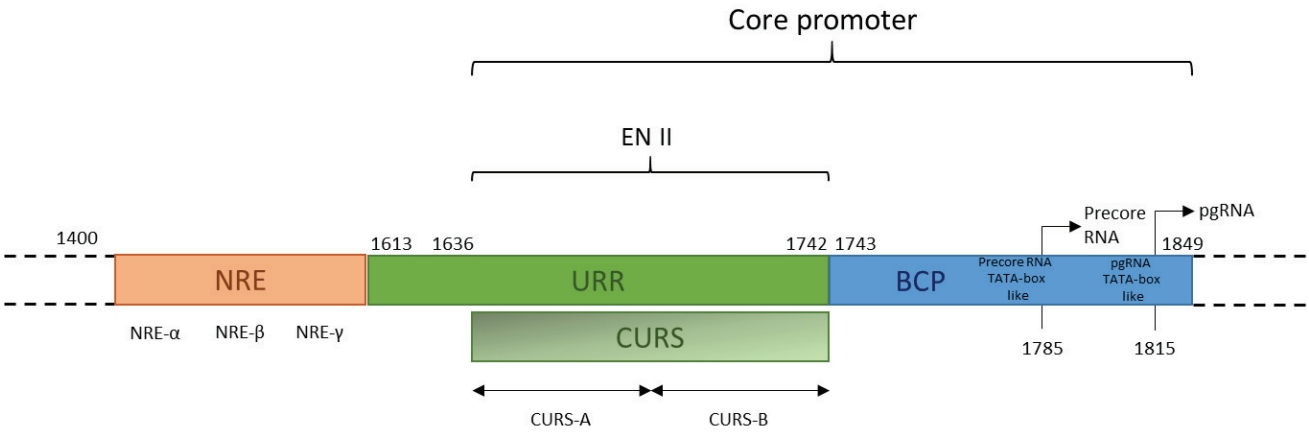


Figure 9 - HBV core promoter regulates precore RNA and pgRNA transcription

The **preS1 promoter** controls the transcription of L surface protein RNA. It comprises a classical TATA box and a hepatocyte nuclear factor 1 (HNF1) binding site (Raney et al., 1994). The preS1 promoter is negatively regulated by sequences downstream of the preS2/S promoter (Moolla et al., 2002).

The **preS2/S promoter** directs the transcription of the 2 mRNAs specific for the M and S surface proteins. It does not contain a classical TATA box, but contains seven regulatory elements and a CCAAT regulatory element (Zhou and Yen, 1991) that inhibit preS1 RNA transcription and stimulate the preS2/S RNAs transcription (Moolla et al., 2002). This is further detected at the protein level with a higher proportion of SHBs and MHBs proteins (Moolla et al., 2002).

The **X promoter** is located 140 nt upstream of the first X TSS and overlap with the EN1 (Quarleri, 2014). It has no classical TATA box and contains sites for ubiquitous or liver-specific transcription factors such as C/EBP, ATF, AP1/Jun-Fos, NF1, p53 (Moolla et al., 2002).

The **EN1** is a 270 nt sequence located between the S and X ORFs (Shaul et al., 1985; Tognoni et al., 1985). It is composed of different domains: the 5' modulator element, the core central domain and the 3' domain that overlaps with the X promoter (Fukai et al., 1997). It potentiates the transcription of the precore RNA, pgRNA and X RNA that are under the control of the core and X promoters, but has only a minor effect on surface protein RNAs transcription. EN1 contains binding sites for ubiquitous transcription factors (NF1, AP1 and NFκB) and liver specific factors (LRH-1/hB1F, HNF1, HNF3b, HNF4 and the C/EBP family) (Quarleri, 2014).

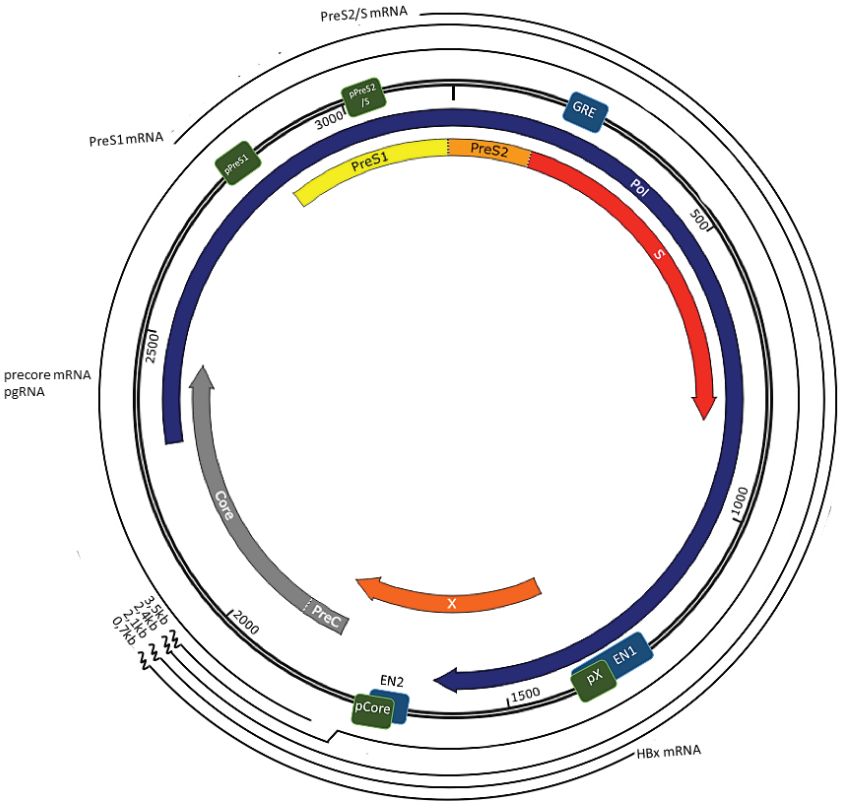


Figure 10 - Genomic organization of hepatitis B virus: transcripts and transcriptional regulation adapted from (Quarleri, 2014)

The **EN2** is a 105 nt sequence located upstream of the BCP promoter (Yee, 1989) that increases the transcriptional activity of the preS1, preS2 and X promoters as well as pgRNA and precore RNA transcription. EN2 contains two distinct regions called IIA and IIB that cooperate to form a powerful enhancer. EN2 contains binding sites for the cellular transcription factors C/EBP, RXR, PPAR, HNF4, HNF3, FTF (Cai et al., 2003).

The **post-transcriptional regulatory element (PRE)** overlaps the ORF X (nucleotides 1151 to 1684 in genotype D). The PRE is a multifunctional element that acts both as a silencer and an enhancer of splicing. The PRE protects HBV unspliced viral RNAs from splicing and allows their export from the nucleus to the cytoplasm. The ability of PRE to inhibit splicing was initially linked to the polypyrimidine tract binding protein (PTB) binding sites. More recent data have implicated the intronic splicing silencer (ISS) element and its ability to fold in a double hairpin structure as responsible of splicing silencing. Destruction of the ISS hairpin structure abolishes the splicing silencer activity (Huang et al., 2011). Conversely, deletion of the splicing regulatory element (SRE-1) in the PRE reduces pgRNA splicing in cell culture model, thus identifying the SRE-1 as a splicing enhancer (Heise et al., 2006).

The glucocorticoid responsive element (GRE) is composed of two 5'-TGTTCCT-3' repeated sequences at positions 341 and 370 in the S gene. GRE cooperates with EN1 (Tur-Kaspa et al., 1986) and stimulates viral gene expression and viral replication (Chou et al., 1992).

The IFN-stimulated response element (ISRE) is located in the enhancer 1/X gene promoter region of the HBV genome. IFN-alpha inhibits HBV replication by reducing cccDNA transcription (Belloni et al., 2012). The IFN-alpha response is mediated by the ISRE and dependent of the HBV (Guo et al., 2019).

A major role in the regulation of HBV transcription and HBV replication is played by the **HBx protein**, whose activities are detailed in section 1.4.5 Translation.

1.4.4.2 HBV Spliced variants

Splicing is required to generate most of functional eukaryotic mRNAs in which introns are removed and exons are joined and allows to increase mRNAs and protein diversity. The HBV splicing process is mediated by the spliceosome, a large host ribonucleoprotein complex.

The first spliced HBV RNA was identified in 1989 and to date at least 22 distinct viral spliced RNAs have been described (Figure 11) (Chen et al., 2015; Günther et al., 1997; Lee et al., 2008). The identification of these spliced HBV RNA was performed originally by northern blot and RT-PCR but nowadays, new spliced RNA can be identified by RNA-seq analysis. So far, only single and double spliced isoforms have been identified. Five splice donor sites have been described (at position 2067, 2087, 2447, 2471 and 2985 according to HBV genotype D nucleotides numbering) and 6 spliced acceptor sites (at position 2236, 2350, 2902, 3169, 282 and 489 according to HBV genotype D nucleotides numbering). The spliced donor site 2447 and the spliced acceptor site 489 are the most frequently used in HBV RNA splicing.

The high level of conservation of the splice sites suggests a critical role of alternative splicing in HBV infection. Overall, 22 spliced HBV RNAs have been identified so far in the liver, blood, cell cultures and cell supernatants (Figure 11). Some spliced RNAs have been identified in circulating defective HBV DNA particles both in patients and cell culture samples (Kremsdorf et al., 2021).

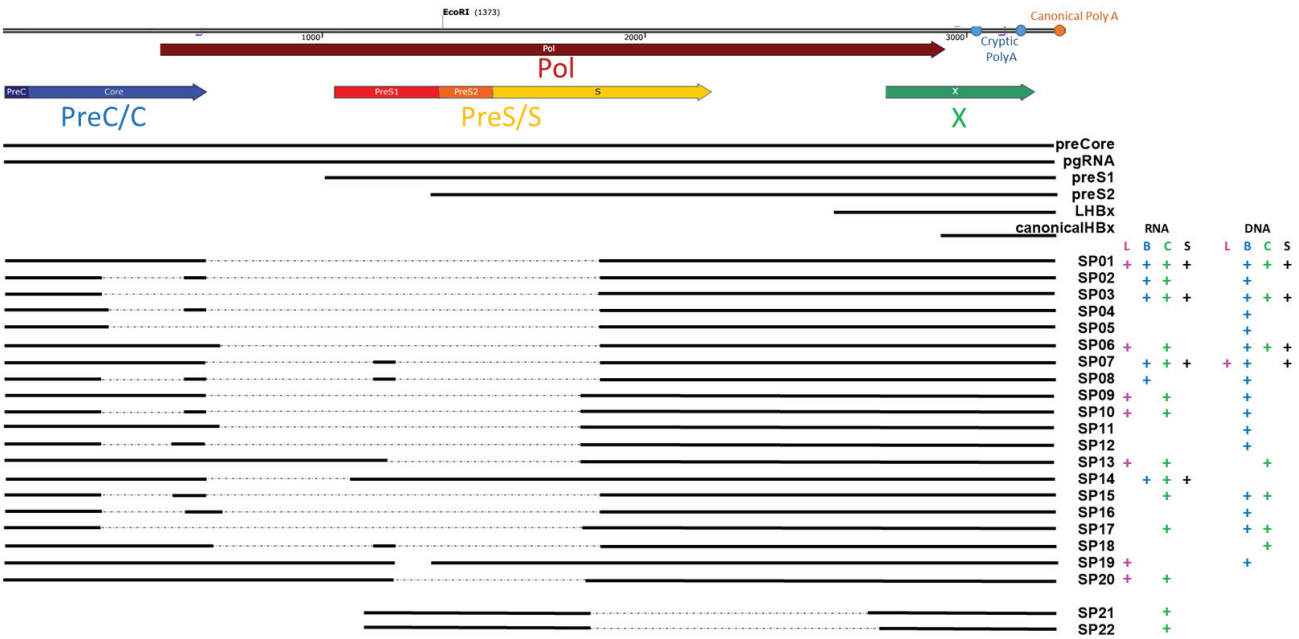


Figure 11 - Scheme of the distinct HBV RNAs including spliced RNAs
Adapted from (Kremsdorf et al., 2021)

Spliced variants SP1, SP2 and SP3 account for 70% of the variants with the most common spliced RNA being the SP1, a 2.2 kb single spliced RNA generated by the removal of a 1.3 kb intron from the pgRNA that accounts for 30% of the pgRNAs in transfected cells and up to 60% in patient samples (Candotti and Allain, 2017; Chen et al., 2015). The vast majority of spliced variants (20 out 22 spliced HBV RNAs) are derived from pgRNA and only a minority (2 out 22 spliced RNAs) from PreS2/S RNAs (Candotti and Allain, 2017).

All pgRNA derived spliced variants contain the packaging signal ε. The presence of the ε sequence allows the packaging of these transcripts, when the HBc and Pol proteins are present, and their retro-transcription can generate defective viral particles. In the absence of transcomplementation with a wild-type virus, these viral particles are replication defective (Günther et al., 1997; Rosmorduc et al., 1995). The SP1 isoform generates up to 70% of whole HBV DNA circulating particles. The proportion of SP1 defective viral particles in patients is related to liver disease progression (Soussan et al., 2008). Spliced variants have been associated to a reduced response to IFNs treatment (Chen et al., 2015).

HBV spliced variants can be translated into non-canonical viral proteins such HBSP from SP1 transcript, HBDSP from SP7, Pol-HBs fusion protein from SP14, and the splicing PreS1 deletion protein (sp1PS1) from SP19 (Chen et al., 2015; Kremsdorf et al., 2021). HBSP protein is the main protein identified as originating from spliced pgRNA. It is composed of 99 to 113 amino acids depending on the viral genotype (Soussan et al., 2000). HBSP has been described to impact on cell proliferation and cell migration (Chen et al., 2012) as well as on viral replication, liver disease and liver carcinogenesis (Bayliss et al., 2013; Soussan et al., 2008). Anti-HBSP antibodies have been detected in one-third of sera samples from chronic HBV infection patients (Soussan et al., 2000). It has been proposed but not proved that the HBx potentially translated from all the spliced variants migth contribute to viral replication.

1.4.5 Translation

Seven major proteins from 4 ORFs are encoded by the viral genome: HBeAg, core, polymerase, SHBsAg, MHBsAg, LHBsAg and X (*Table 1*).

RNA size	RNAs	Proteins	RNA Length
Pregenomic	Precore RNA	Precore protein (HBeAg)	3.5 kb
	pgRNA	Viral polymerase (Pol)	3.5 kb
		Core protein (HBcAg)	
Subgenomic	PreS1 RNA	Envelope protein L	2.4 kb
	PreS2/S RNA	Envelope proteins M and S	2.1 kb
	X RNA	HBx protein	0.7-0.9 kb

Table 1 - Hepatitis B virus RNA and the related proteins

1.4.5.1 The polymerase protein

The protein P or polymerase (Pol) is the largest viral protein (90 kDa and 845 aa). The interaction in the cytoplasm with heat shock proteins ensures its correct conformation (Hu and Seeger, 1996b). The HBV Polymerase is a multifunctional protein with RNA-dependent DNA polymerase/reverse transcriptase, DNA-dependent DNA polymerase and RNAaseH activities. HBV Pol protein structure comprises 4 domains from the N-terminal to C-terminal domain: the template domain (TP), the spacer domain, the reverse transcription domain (RT) and the RNase H domain (Figure 12) (Radziwill et al., 1990).

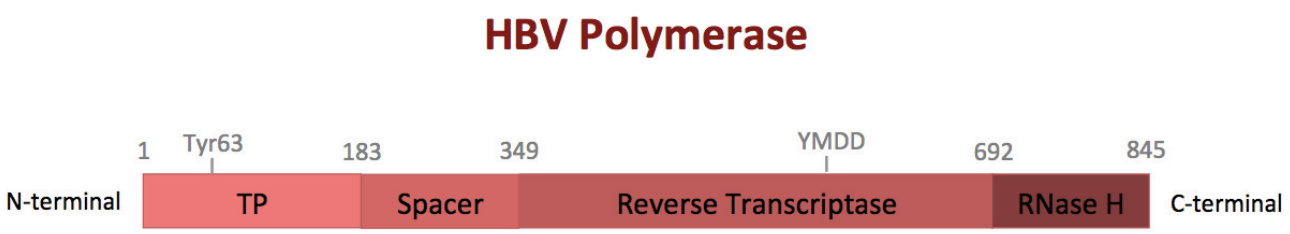


Figure 12 - HBV Polymerase protein structure
adapted from **(Tong and Revill, 2016)**

RT and RNase H domains are found in other viral reverse transcriptases, but the **TP domain** is unique to Hepadnaviruses. The TP domain is required for pgRNA binding, for pgRNA packaging in the nucleocapsid and for initiation of reverse transcription since the TP domain contains a tyrosine residue (Y63) that provides the hydroxyl group needed for the covalent binding of the first nucleotide of the negative strand DNA to the viral polymerase (Clark et al., 2017; Lanford et al., 1997; Zoulim and Seeger, 1994).

The **Spacer domain** links the TP and RT domains. Its role is still unclear. The sequence is poorly conserved and may have insertions and deletions that may confer flexibility for conformational changes to respond to immune pressure (Chen et al., 2013).

The **RT domain** carries the reverse transcription and DNA polymerase activities. The YMDD (tyrosine, methionine, aspartate) sequence, conserved within the different HBV genotypes, is the catalytic site necessary for viral replication (Jeong et al., 1996).

While the **RNAse H domain** is essential for pgRNA degradation after retro-transcription of the minus DNA strand. The RNAse activity of the HBV Pol stops before complete digestion of the pgRNA thus leaving few nucleotides that serve as a primer for the synthesis of the DNA positive strand.

Due to the lack of any host homologue and its key role in viral replication, HBV Pol is a good therapeutic target (Clark et al., 2017).

1.4.5.2 The HBx protein

The HBx protein is the smallest viral protein of 17 kDa and 154 aa (Figure 13). It has no homology with known mammalian genes but is highly conserved in mammalian *hepadnaviruses* and absent in *Avihepadnaviridae* (Sitterlin et al., 1997).

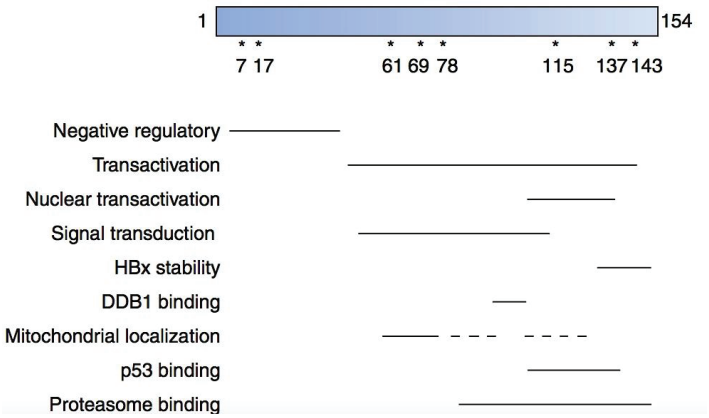


Figure 13 - HBx protein structure from (Slagle and Bouchard, 2016)

After infection, HBx interacts with the DDB1 host E3 ubiquitin ligase to ubiquitinate and tag for proteasomal degradation the Smc5/6 complex, a restriction factor against episomal DNAs that silences cccDNA transcription (Decorsière et al., 2016; Murphy et al., 2016). HBx also modulates the recruitment of several chromatin-modifying enzymes onto the cccDNA and affects the acetylation, methylation, and phosphorylation of cccDNA-associated histones to favour cccDNA transcription whereas, in the absence of HBx, the cccDNA is epigenetically silenced (Alarcon et al., 2016; Belloni et al., 2009; Luo et al., 2013; Rivière et al., 2015)

HBx also modulates transcription of cellular genes. A ChIP-Seq of HBx-bound chromatin identified the genome wide repertoire of direct HBx target genes. In addition to genes potentially involved in cell growth and cancer development, HBx also activates transcription of genes that potentiate endocytosis and autophagy to promote viral replication and represses the expression of miRNAs that target pgRNA, thus protecting it from degradation and again promoting viral replication (Guerrieri et al., 2017).

HBx has been the object of many studies and several additional *cytoplasmic* functions have been assigned to it. HBx has been shown to modulate apoptosis, cell cycle and oncogenesis (Wu et al., 2010). As an example, HBx can bind to the tumor-suppressor a protein p53 and interfere with its role in the nucleotide excision repair (Jia et al., 1999) and also down-regulates the expression of the p53 activators ASPP1 and ASPP2 (Ali et al., 2014).

1.4.5.3 The surface proteins (HBsAg)

There are three surface proteins that are collectively called HBsAg: the small (SHBs) of 27 kDa, the medium (MHBs) of 31 kDa and a large (LHBs) surface protein of 42 kDa. They form the envelope of the virus, and the small protein is the most represented envelope protein. The 3 proteins are derived from a single ORF of 389 aa which contains three domains:

- PreS1 (108 to 119 aa depending on the genotype), present only in the LHBs protein,
- PreS2 (55 aa) present in both LHBs and MHBs
- S (226 aa), common to all three surface proteins.

As already mentioned, LHBs is translated from the 2.4 kb preS1 RNA, MHBs from the 2.1 kb PreS2 RNA and SHBs from the 2.1 kb S RNA. Translation happen in a 1:1:4 ratio (Sheu and Lo, 1992) and SHBs is the most represented envelope protein. Each envelope protein has a distinct ATGs but they share the same stop codon (Glebe and Urban, 2007). SHBs, MHBs and LHBs also share four trans-membrane domains located in the S-domain. They are synthesized in the ER and then maturate into the Golgi Apparatus. The S domain also includes a cytosolic loop and an antigenic loop (Figure 14) (Alfaiate et al., 2015).

In addition to be present in the envelope of infectious virions, the 3 surface proteins also form subviral particles (SVPs), comprising spheres and filaments, which are secreted by infected cells. SHBs protein is found in virions, spheres and filaments, while the L protein is mostly found in virions and filaments and barely detectable in the spheres (Heermann et al., 1984). Thus, SHBs and LHBs are required to produce HBV particles, whereas the function of MHBs is less clear. The PreS2 domain is conserved through all *orthohepadnaviruses*, but the MHBs protein it is not required for virions secretion, subviral particles formation and HBV infectivity (Glebe and Urban, 2007).

Surface proteins play a major role in hepatocyte-virus recognition. The S domain contains an antigenic loop with an a-determinant. This is highly conserved in all viral strains and is a conformational epitope that can bind to the heparan sulfate on the hepatocytes surface. This low affinity interaction favours the high affinity interaction between the myristoylated (Myr) domain of PreS1 and the sodium taurocholate cotransporting polypeptide (NTCP) receptor. The PreS1 domain, only present in the LHBs protein, is myristoylated on the N-terminal Gly in position 2 and adopts two conformations (Macrae et al., 1991). The e-preS conformation that is involved in virus–hepatocyte interaction, and the i-preS conformation that permits the interaction with the nucleocapsid in the cytoplasm.

A N-glycosylation (NG) site is located in the S domain at the amino acid Asn-146. Glycosylation is never complete with a glycosylated/non-glycosylated ratio of 1:1. The non-glycosylated S domain on N146 is essential for HBV virion production while the glycosylated form of N146 is involved in HBV secretion (Julithe et al., 2014). A second NG site is located at the amino acid Asn-4 in the PreS2 domain. This second NG site is modified only in the MHBs protein but not in the LHBs protein (Heermann et al., 1984). Finally, there is an O-glycosylation (OG) site at the level of Thr-37 (Tolle et

al., 1998). This OG site is absent in genotype A where Thr-37 is replaced by a Asp-37 (Schmitt et al., 2004).

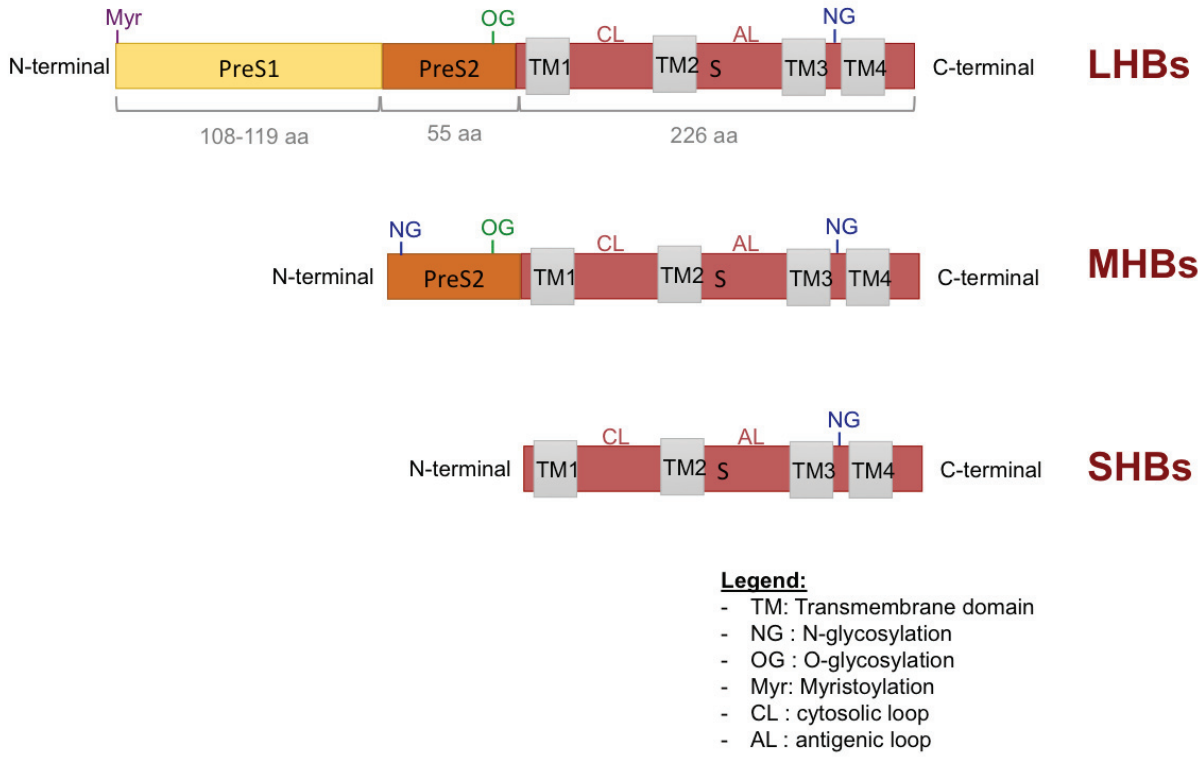


Figure 14 - HBV Surface Proteins
adapted from (Alfaiate, Dény, and Durantel 2015; Glebe and Urban 2007)

1.4.5.4 The Core protein (HBc)

The precore and core proteins are derived from the same ORF PreC/C but have two distinct ATGs and are translated from 2 independent RNAs (Bruss, 2007). The Precore protein, that is maturated into the HBeAg, is translated from the Precore RNA, while the core protein is translated from the pgRNA.

The 21kDa core protein (HBc) contains between 183 and 185 amino acids, depending on the genotype, with an N-terminal (NTD) and a C-terminal (CTD) domain or protamine domain which is rich in arginines (Chain and Myers, 2005). This arginine rich domain is involved in the interaction between core proteins and pgRNA (Nassal, 1992). The protamine domain also contains a nuclear localization signal (NLS) and a nuclear export signal (NES), used in HBc shuttling (Li et al., 2010).
Core proteins assembly into dimers at cysteine 61 (Nassal et al., 1992) and form the icosahedral capsid in which the pgRNA and viral polymerase are encapsidated. There are two forms of capsids:

the so-called T=3 form (30 nm), composed of 90 dimers of core proteins, and the most common, the T=4 form (34nm), with 120 dimers (Zlotnick et al., 2015).

The HBc protein is also a structural component of the HBV minichromosome (Bock et al., 2001) and may also be involved in transcriptional regulation (Chong et al., 2017). It has been recently shown that the HBc derived from incoming HBV virions participate at an early stage of cccDNA formation and transcription (Lucifora et al., 2021).

1.4.5.5 The Precore (PreC) protein and HBeAg

The Precore (PreC) protein is translated from the 3.5 kb precore RNA. The Precore protein p25 is very similar to the Core protein, but differs at its N-terminal domain by 29 additional amino acids. It is then translocated to the ER where it will be truncated by 19 amino acids at the N-terminus to form the shorter Precore protein p22 (Messageot et al., 2003). The remaining 10 amino acids preC specific domain prevents the rest of the protein from taking a conformation similar to the core protein (Messageot et al., 2003). The p22 protein is then cleaved at its C-terminal domain in the Golgi apparatus to generate the 15-18 kDa HBeAg. HBeAg is a non-structural protein and is not essential for viral replication (Figure 15). HBeAg has been described to interfere with innate immunity pathways and to be tolerogenic, possibly due to its structure similar to HBc (Milich and Liang, 2003; Walsh and Locarnini, 2012).

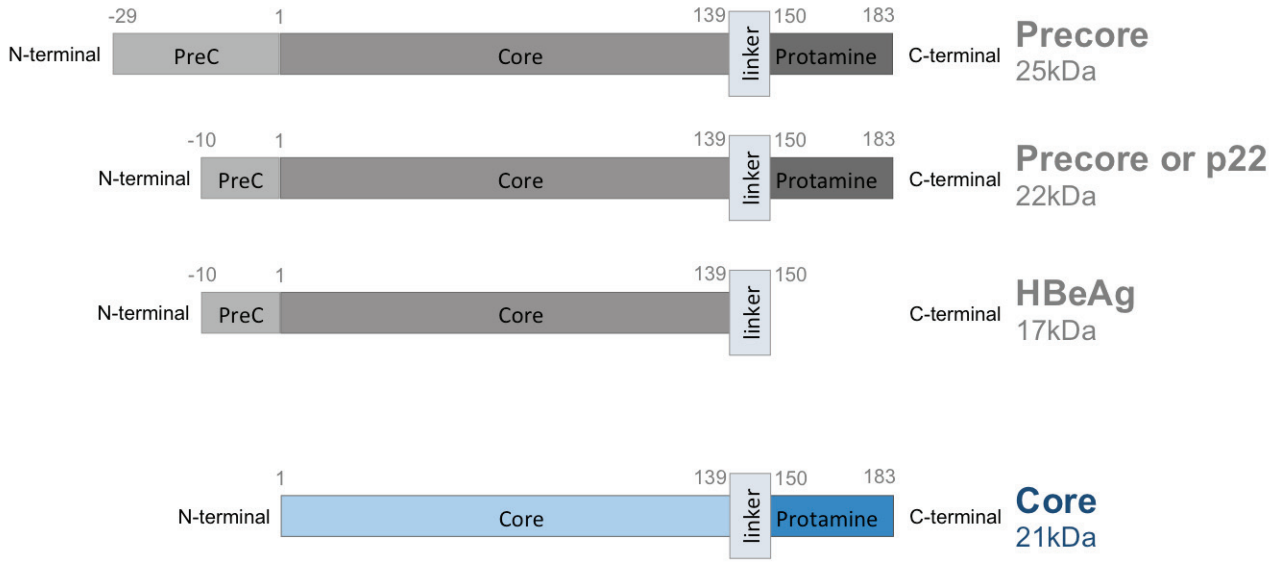


Figure 15 - Precore, core and HBeAg protein structure adapted from (Steven et al., 2005)

1.4.6 Secreted HBV particles

Secreted viral particles comprise both complete viral particles and sub-viral particles. Complete viral particles are enveloped and contain a nucleocapsid e.g. the combination of a capsid and encapsidated nucleic acids. Immature nucleocapsids contain either pgRNA or single stranded DNA (ssDNA). Mature nucleocapsids contain rcDNA. Complete virions (enveloped DNA particles containing a mature capsid) are secreted into the extracellular space and infect new hepatocytes. Incomplete (HBV RNA particles with an immature pgRNA capsid, DNA free particles with empty capsids) and subviral particles (naked capsids, HBs filaments and spheres) are also generated and secreted (Hu and Liu, 2017).

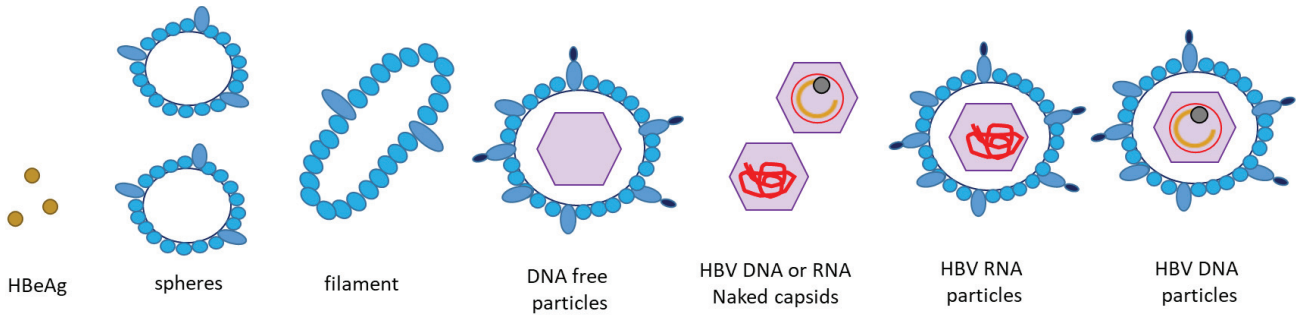


Figure 16 - Complete and incomplete Hepatitis B virus particles

1.4.6.1 Complete virion

As stated above, the 42nm HBV virion consists of a lipid envelope containing the surface glycoproteins L, M, and S. The envelope covers the icosahedral capsid formed by the core protein dimers that protect the rcDNA genome of 3.2 kb (Figure 16). The number of complete virions found in the blood varies according to the different phases of HBV infection and disease.

1.4.6.2 Circulating HBV RNAs

In 1996, it was described for the first time, using the rapid amplification of complementary DNA (cDNA)-ends (RACE) technique, that in addition to HBV DNA also polyadenylated viral RNAs are present in the serum of patients (Köck et al., 1996). HBV RNA levels in the serum varies according to the phase of infection and disease, usually 1 to 2 log_{10} lower than the HBV DNA viral load (van Bömmel et al., 2015; Huang et al., 2015; Rokuhara et al., 2006; Wang et al., 2016).

The type of circulating HBV RNA (polyadenylated pgRNAs, subgenomic RNAs and spliced RNAs; polyA-free species) and their compartimentalization (free RNA, exosome RNA, enveloped particles, subviral particles) have been the object of intense research efforts but remain still debated. In 2016, a team showed that serum HBV RNA is pgRNA contained in virus-like particles (Wang et al., 2016). A second study, in the same year, confirmed that the majority of circulating HBV RNAs were pgRNAs, either polyadenylated or poly-A free (Altinel et al., 2016). Several studies have confirmed that circulating HBV RNAs are full-length transcripts ending at the polyadenylation signal “TATAAA” motif (van Bömmel et al., 2015; Butler et al., 2018; Jansen et al., 2016; Wang et al., 2016). This does not exclude the possible presence of circulating HBV RNAs ending at the cryptical polyadenylation signal “CATAAA” (van Bömmel et al., 2015), reflecting either the use of these alternative polyadenylation signals in the cccDNA (van Bömmel et al., 2015) or transcription from HBV integrated sequence (Hilger et al., 1991).

Several pgRNA spliced variants have also been detected as circulating HBV RNAs (Lam et al., 2017a). Although, there is growing consensus that pgRNA or viral spliced RNAs derived from pgRNA are the major circulating HBV RNA species, HBx mRNA has also been detected using high throughput RNA sequencing (Niu et al., 2017). Using a 5'RACE PCR, pgRNA, spliced pgRNA variants and HBx RNAs have been detected in the sera of patients with chronic HBV infection (Stadelmayer et al., 2020).

The majority of circulating HBV RNA appears to be in enveloped RNA particles ((Jansen et al., 2016; Rokuhara et al., 2006; Wang et al., 2016) Kim et al., manuscript in preparation). However, HBV RNAs could also be secreted in CD81+ exosomes from HepG2 cells (Kouwaki et al., 2016); Kim et al., manuscript in preparation) and have been found in circulating exosome in chronic hepatitis B patients (Kim et al., manuscript in preparation). RNA-containing naked capsids were also detected in the plasma of 4 different patients with chronic hepatitis B (Bai et al., 2018)

1.4.6.3 Empty (genome-free) virions (enveloped capsids)

Empty virions or genome-free virions with enveloped capsids that do not contain any genetic material can be also found in HBV infected individuals. The existence of these particles has been suspected for a long time, since Dane particles deficient in viral nucleic acids have been identified (Gerin et al., 1975). Notably, these empty DNA-free virions are predominant over complete virions (Ning et al., 2011).

1.4.6.4 Naked Capsids

Naked capsids (non enveloped nucleocapsids) are found both in the supernatant of HBV-replicating cell lines as well as in the blood of HBV patients (Bai et al., 2018; Sun and Nassal, 2006). Naked capsids carrying HBV DNA or RNA have been detected in patient sera, mainly as capsid-antibody-complexes (Bai et al., 2018). It has been suggested that these HBV particles might contribute to HBV genome transmission due to their efficient internalization through clathrin-mediated endocytosis (Cooper and Shaul, 2006).

1.4.6.5 Subviral particles: spheres and filaments

HBsAg filaments and spheres are composed only of HBV surface proteins. They differ in diameter: the spheres are 22 nm in diameter and the filaments are 18 nm wide, but of variable lengths between 80 and 400nm. These are the particles that were discovered by Baruch Blumberg and initially named Australia or Au antigens (Blumberg et al., 1965). HBsAg proteins are produced in large excess by HBV-infected hepatocytes and circulate in large quantities in the blood. HBsAg is thought to be secreted at 1 000 to 100 000 higher levels than virions (Blumberg, 1977; Chai et al., 2008; Désiré et al., 2015; Ganem and Prince, 2004). The large amounts of spheres and filaments released have multiple negative effects on host immune response : a) the chronic exposure of HBV specific T cells to high HBsAg loads contributes to T cell exhaustion (Bertoletti and Ferrari, 2016); b) the excess of subviral HBs particles act as a decoy binding anti-virus antibodies thus competing for complete virions neutralization (Nassal and Schaller, 1993).

HBV SVPs have a high content in lipids (25% by weight of SVPs) and glycans. The main lipids in SVPs are phosphatidylcholine (~60%), cholesteryl ester (~14%), cholesterol (~15%), and triglycerides (~3%).

They also present important differences in HBs proteins composition. Spheres are composed of SHBs (90%) and MHBs proteins (10%), whereas filaments, like complete virions, contain SHBs (80%), MHBs (10%), and LHBs proteins (10%) (Herrscher et al., 2020b). Virions and filaments appear to share the same mode of secretion, being released from hepatocytes by the endosomal sorting complexes required for transport (ESCRT) via multivesicular bodies (MVB). Spheres, they are secreted by the ER-Golgi constitutive secretion pathway (Jiang et al., 2015; Watanabe et al., 2007). Future spheres assemble initially in the lumen of the ER and are then packaged into structures that transported to the ER-Golgi intermediate compartment (ERGIC) where they are converted into spherical particles for secretion (Patient et al., 2007, 2009).

1.5 Genetic diversity of HBV

The nucleotide substitution rate of the viral polymerase is estimated to be between 10^{-4} - 10^{-5} per nucleotide per year (Orito et al., 1989; Park et al., 2003). This error rate is ten times higher than the polymerase of other DNA viruses (Croagh et al., 2015). These natural mutations have led to the apparition of the different HBV genotypes and mutants.

Initially, the genetic diversity was classified based on the antigenic properties of HBsAg into 10 serotypes (ayw1, ayw2, ayw3, ayw4, ayr, adw2, adw4q-, adw4q+, adrq+ and adrq-) (- Molecular basis of HBV serotypes and their corresponding genotypes) (*Table 2*) (Kay and Zoulim, 2007). Then, with new technologies and sequencing, a classification of viral diversity by genotype was adopted.

Serotype/HBsAg aa sequence	122	127	134	159	160	177	178	Corresponding genotype
ayw1	R	P	F	A	K			A, B, C, or D
ayw2	R	P			K			D or E
ayw3	R	T			K			D
ayw4	R	L			K			D, E, or F
ayr	R				R			C
adw2	K	P			K			A, B, C, F or G
adw3	K	T			K			B
adw4q-	K	L			K		Q	F or H
adrq+	K				R	V	P	C
adrq-	K				R	A		C

Table 2 - Molecular basis of HBV serotypes and their corresponding genotypes adapted from (Kay and Zoulim, 2007)

1.5.1 The different genotypes and subgenotypes

There are 9 different HBV genotypes (designated by the letters A to I) and one additional putative HBV genotype (HBV J), with a genetic divergence of more than 8% between each genotype. Each genotype can be subdivided into more than 30 subgenotypes with nucleotide differences ranging from 4 to 8% within a given genotype (Bousali et al., 2021). Thus, a 12% overall divergence between

all genotypes exists with the most divergent genotypes being F and H (Olinger et al., 2008; Schaefer, 2007; Tatematsu et al., 2009).

The 9 HBV genotypes are distinct in terms of geographical distribution (Figure 17): genotype D has the highest prevalence after genotype C, and they are ubiquitous; genotype A is found mainly in the United States and North-Western Europe; genotype B is more commonly found in North and Southeast Asia; genotype C is prevalent in Asia-Pacific; genotype E is mainly found in Africa; genotype F is predominantly found in South and Central America; genotype G is present in France and the United States and genotype H is only found in North of America (Zehender et al., 2014).
Five genotypes are responsible for 97% of chronic hepatitis B infections: C is the most common (26%), followed by D (22%), E (18%), A (17%) and B (14%). Finally, genotypes F to I represent only a minority of infections with a prevalence of 2% (Velkov et al., 2018).

Genotypes have also been associated with the pathogenesis of HBV infection. Patients infected with genotype A and D appear to be more prone to develop a chronic infection compared to patients infected with genotype B and C viruses (Fletcher et al., 2019). Genotype C is associated with more severe liver pathogenicity with higher rates of development of cirrhosis or hepatocellular carcinoma (Chan et al., 2009). HBV genotype can also influence response to treatments. For instance, patients infected with HBV genotype A or B respond better to interferons than those infected with genotype C or D (Erhardt et al., 2005; Wai et al., 2002).

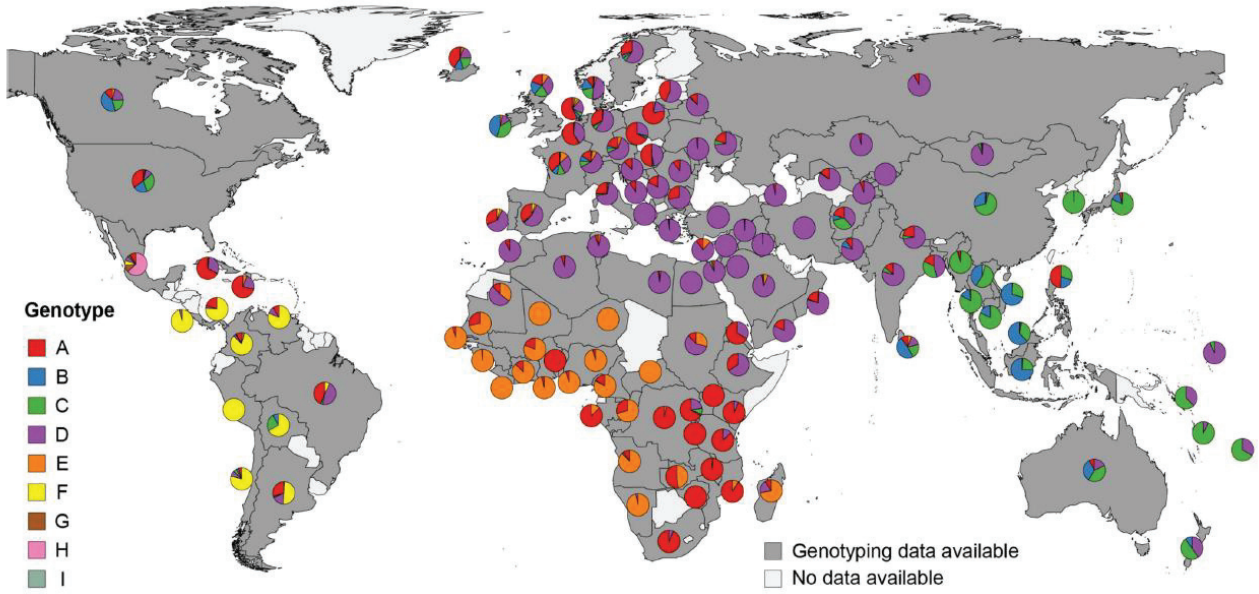


Figure 17 - HBV genotype distribution over the world
from **(Velkov et al., 2018)**

1.5.2 HBV genome mutation leading to phenotypic variations

The errors made by the HBV polymerase can also lead to mutations that are positively selected and become prevalent in the viral progeny at the single patient level and at the population level. Some are well known, such as the mutation inserting a stop codon in the precore domain leading to the loss of the HBe antigen production. Additional well known and prevalent mutations are those found in the core promoter, which are associated with higher viral replication, mutations in the PreS2 and S ORF that affect the HBsAg or the mutations in the transcriptase/polymerase that confer resistance to antiviral treatments.

1.5.2.1 The Precore variants

Precore mutations, one of the first clinically recognized HBV mutations, result in the loss of HBe antigen expression. Mutations affecting the production of HBe antigens are widespread throughout the world and affect the Mediterranean basin, Africa, Asia and Europe. In France, it is estimated that 50% of HBV-infected subjects carry a pre-core mutation (Grandjacques et al., 2000).

The most common precore mutation is a guanosine (G) to adenine (A) substitution at nucleotide 1896 (G1896A), which prevents the production of HBeAg by introducing a premature stop codon at codon 28 of the Precore open reading frame. The Precore G1896A mutation is more common among

patients with genotype D (65-75%) than genotype A (9-18%). The mutation is located within the ε-loop. To stabilize the ε structure, the nucleotide at position 1896 is paired with the nucleotide at position 1858, which naturally is a thymidine (T) in genotypes B, D, E, and G and a cytosine (C) in genotype A. Because of the presence of a C at position 1858, the 1896A mutation makes the ε-loop unstable and negatively impacts viral replication. Infection with the precore 1896A mutant virus is often accompanied by the presence of mutations in the core region. Among people chronically infected with HBV, each year 10% of them undergo a spontaneous HBeAg to anti-HBe seroconversion that is accompanied by a sharp decrease in HBV replication. HBeAg loss due to a mutation in the Precore domain does not necessarily indicate cessation of viral replication and is associated with continuous liver pathogenicity and a more severe chronic hepatitis (Carman et al., 1995). Patients infected with a Precore mutant HBV have a higher risk to develop a fulminant hepatitis, but Precore mutants are also found in asymptomatic carriers or acute self-limited hepatitis (Baumert et al., 2007).

1.5.2.2 The basal core promoter (BCP) variants

Some patients carry HBV mutated in the core promoter. Most core promoter mutants comprise a double mutation in the BCP at nucleotides 1765 and 1767 with a conversion from A to T and G to A, respectively. Notably, these changes affect the overlapping HBx ORF and result in a double amino acid substitution at HBx codons 130 and 131 that have been associated with a stronger HBx pro-oncogenic activity. These mutations are also associated with a decrease in the transcription of precore RNAs with a decrease in the expression HBe antigen or a loss of HBe antigen expression (Li et al., 1999).

Several mutations at other sites in the core promoter have been identified (Liu et al., 2009) and in one study in Chinese patients the most frequent mutation was the substitution of a thymine by a cytosine at position 1758 of the precore region (T1758C) (Liu et al., 2010).

1.5.2.3 The HBsAg variants

HBV variants carrying mutations and deletion in the PreS regions are quite frequent and may have relevant pathobiological and clinical implications: a) an imbalance in the synthesis of LHBs, MHBs and SHBs may lead to retention within the ER and an ER stress leading to the induction of oxidative DNA damage and genomic instability (Hsieh et al., 2004; Pollicino et al., 2014) b) the development of fulminant hepatitis and fibrosing cholestatic hepatitis (Pollicino et al., 2014); c) an increased risk to develop cirrhosis and hepatocellular carcinoma (Cohen et al., 2020; Pollicino et al., 2014).

Several mutants in the S region have been identified that impact vaccine response or immune and diagnostic escape (Coppola et al., 2015). The change of glycine 145 of the S ORF into an arginine leads to an antigenic modification of the site targeted by the vaccine (Schilling et al., 2003) and this mutant escapes immune surveillance and can infect the host even in the presence of anti-HBs antibodies. Some S mutants may acquire an oncogenic potential. This is the case for the SW172 stop mutation which is associated with HBs protein retention and an increased HCC risk in patient with HBV genotype C (Lai et al., 2016).

1.5.2.4 DNA polymerase mutations

Numerous mutants appear following continuous treatment with some nucleotide analogues (NUCs) that inhibit the HBV Pol enzymatic activities. This is particularly true with lamivudine. After 3 years, 53% of patients will have a mutation in the YMDD catalytic site of the polymerase and after 4 years 70% will develop a YMDD mutation (Lai et al., 2003). The most frequent is the rtM204V/I mutation that alters the spatial conformation of the catalytic domain leading to resistance to the antiviral activity of lamivudine but also to a reduced replication. The appearance and selection of the compensatory mutations rtL180M or rtV173L restores the viral replication (Delaney et al., 2003). Patients that develop resistance to lamivudine will also be cross-resistant to other NUCs such as telbivudine and will develop resistance mutations for entecavir, that has a very genetic barrier in naive patients (He et al.).

CHAPTER II: HEPATITIS B VIRUS INFECTION

2.1 Epidemiology

It is estimated that 2 billion people have been infected with HBV at some point in their lives and that over 240 million individuals are chronic HBV carriers. Of note, only 10% of HBV infections have been diagnosed and only 1% of chronic hepatitis B patients receive a treatment worldwide (Polaris Observatory Collaborators, 2018). About 40% of people infected with HBV develop cirrhosis, liver failure and hepatocellular carcinoma overtime and 25% of these infected people will die from one of the complications associated with the virus. Altogether, hepatitis B virus is considered one of the most pathogenic human viruses. Importantly, the global public health burden of HBV is increasing. The sum of all the costs generated by an HBV infection for an individual rose from $2,670 in 2006 to $74,274 in 2015 (Chang and Nguyen, 2017).

Based on seroprevalence, HBV infection can be classified into three broad classes of endemicity: high (prevalence >7%), intermediate (prevalence between 2-7%) and low (prevalence lower than 2%). The seroprevalence of HBV is estimated at 3.61% of the world population with the highest prevalence being found in Africa with levels reaching 8.83% of the population (Figure 18) (Schweitzer et al., 2015). China, South-East Asia, Indonesia, and sub-Saharan Africa have also been classified among the highly endemic areas with seroprevalences of more than 8%. Three-quarters of patients with chronic hepatitis are found in Asia-Pacific zones (Song and Kim, 2016). A study performed in 2014 indicates

that the Chinese prevalence of hepatitis B virus surface antigen (HBsAg) in persons aged 1-59 years is 7.18% (Zhang et al., 2015). In China, the HBV is a real health scourge because 60% of cirrhosis cases and 80% of hepatocellular carcinomas are caused by HBV infection (Fanning et al., 2019).

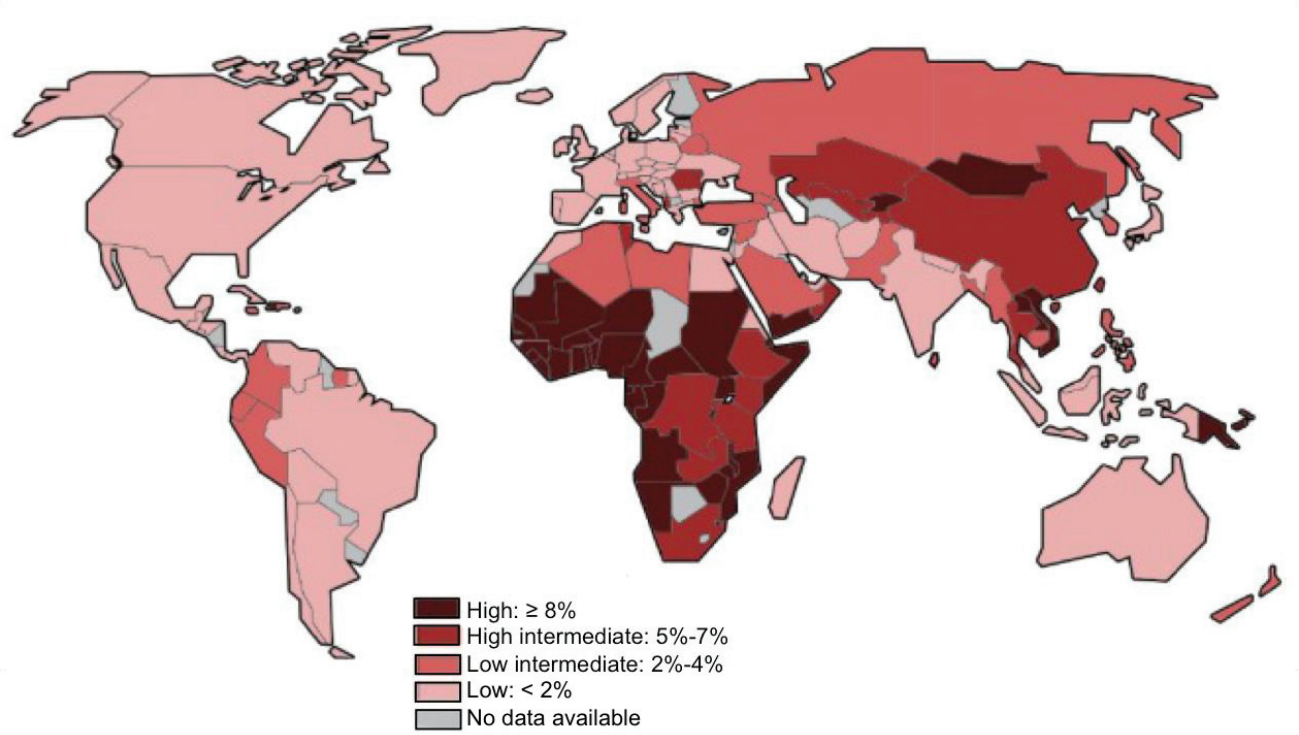


Figure 18 - World map of hepatitis B distribution from (Jefferies et al., 2018)

Medium endemic areas include South America, South-West Asia, Eastern and Western Europe. The prevalence of HBV in these regions is estimated to be between 2 and 7% (Song and Kim, 2016). HBV prevalence in Europe is much lower, around 0.5% and 0.7% (Blachier et al., 2013). Data for Latin American countries is scarce and does not allow an exact prevalence to be established, but it is estimated that there are between 7 and 12 million people with chronic hepatitis in this region (Alvarado-Mora and Pinho, 2013). In the United States, 2.2 million people are carriers of chronic HBV infection despite vaccination efforts (Kowdley et al., 2012). This prevalence of 0.3% has not changed since 1999 due to the immigration of individuals from regions where HBV is endemic (Chang and Nguyen, 2017).

2.2 Transmission and symptoms

HBV is highly contagious. It is estimated that the risk of blood contamination is 100 times higher for HBV than for HIV (Alter, 2006). Transmission occurs either vertically through mother-to-child transmission or horizontally from one individual to another by body fluids contact (e.g., blood and blood products transfusion, needle exchange, organ donation, unprotected sex, tattoos and other percutaneous transmission). The risk of vertical transmission in HBsAg+ and HBeAg+ mothers is 70-90%, whereas in HBsAg+ and HBeAg- mothers this risk is lower ranging between 10 and 40%. Mother-to-child transmission happens through intrauterine transmission, transmission during delivery or by a closed contact between the mother and infant (Gentile and Borgia, 2014). In highly endemic areas, peripartum transmission is the predominant mode of transmission, while in low endemic areas, drug injection and sexual contact are the main transmission modes. The earlier the infection occurs in the life of the subject, the greater is the risk of chronic infection (Lok, 2002).

HBV can be found in all body fluids: blood, saliva, urine, nasopharyngeal secretions, tears, vaginal secretions, menstrual blood, perspiration, breast milk and semen (Karayiannis et al., 1985; Lavanchy, 2004; Scott et al., 1980). Following contact, the virus crosses the bloodstream to reach the liver where it can infect up to 100% of hepatocytes, the common cells likely to be targeted for infection.

The incubation period varies from 1 to 4 months after infection. Hepatitis B is often asymptomatic but few patients develop symptoms including nausea and vomiting, asthenia, fever, and eventually black urine or jaundice (Kappus and Sterling, 2013; Liang, 2009). If only non-specific symptoms are present they give little guidance to the doctor on the diagnosis to be made.

The infection of hepatocytes is non-cytopathic, and the virus does not cause acute metabolic, biochemical, and morphological alterations of the cell. Therefore, the extent of histopathology is thought to be the consequence of the adaptive immune response to infection. Rarely, the immune reaction is disproportionate, resulting in a fulminant hepatitis (see section 2.4 Natural history of HBV infection). The adaptive immune system controls the infection, with the elimination of most infected hepatocytes and the loss of HBsAg (see section 2.3 Diagnosis of hepatitis B infection) within 6 months in 95% of adults, 70% of children between 1 and 5 years old and only 5 to 10% of new-borns (Chang et al., 2014). Thus, the reservoir of chronic carriers is mainly composed of individuals infected during

childhood. Low levels of inactive / silenced HBV cccDNA can persist in a variable but limited number of hepatocytes in patients that have lost HBsAg and are clinically recovered after the acute HBV infection. The adaptive and innate immune system are thought to play a key role in maintaining the control of this residual virus.

2.3 Diagnosis of hepatitis B infection

HBV infection screening is problematic as most of chronically infected patients are asymptomatic until they develop late-stage disease (cirrhosis or hepatocellular carcinoma). In many cases, is an elevation of the ALAT and ASAT aminotransferases that leads to a differential diagnosis of the diseases that may cause liver damage, including HBV serological markers (HBsAg, HBeAg, anti-HBc and anti-HBe antibodies and by serum HBV-DNA quantification) (*Table 3*).

The first level tests include anti-HBc and HBsAg detection. Anti-HBc indicates a contact with the virus and HBsAg detection confirms on-going HBV infection. HBsAg is detectable 3 weeks before clinical signs are observed and is detected up to one month after recovery. The persistence of this marker after 6 months with a negative IgM anti-HBc implies the transition to a chronic HBV infection (Charre et al., 2019; Kao, 2008).

Several techniques and diagnostic kits are available for the detection and quantification of HBs antigens including chemiluminescence (CLIA), electrochemiluminescence (ECLIA) techniques and less expensive enzyme-linked immunoassay (ELISA) (Sommese et al., 2014). Screening for HBV can also be done through so-called *point of care* screening. A *point of care* is a medical laboratory diagnostic test intended to be performed in close proximity to the patient and delivering a result in a short period of time (30 to 60 minutes) (Hirzel et al., 2015). In 2020, three HBsAg rapid tests (Determine HBsAg 2, Alere Medical Co., Japan; VIKIA HBsAg, bioMérieux, France and SD Bioline WB, Abbott Korea, Korea) had met WHO prequalification criteria (https://www.who.int/diagnostics_laboratory /evaluations/PQ_list/en/). In France only the Biomerieux "Vikia® HBsAg test" is approved and allows the qualitative detection of the HBV surface antigen in human serum, plasma or whole blood by immuno-chromatography (Bottero et al., 2013). The quantification of HBs antigens is given in International Units per milliliter (IU/mL) and in most quantitative assays, the lower limit of quantification is 0.05 IU/mL.

In the case of HBsAg positivity, the presence of HBeAg and anti-HBe antibodies is evaluated and HBV DNA quantification is performed to determine whether the patient is HBeAg+ or HBeAg- and to define the clinico-virological phase of chronic hepatitis B (see section 2.4 Natural history of HBV infection). In all HBsAg positive patients an HDV coinfection should be evaluated by testing for anti-HDV antibodies, and in case of positivity an HDV RNA quantification should be performed. The anti-HBc-IgM measurement is carried out in the case of a suspicion of an acute HBV infection.

In selected cases with a negative HBsAg test and positive anti-HBc and a suspicion of acute HBV infection an HBV viremia titration should be performed. This is because in the early phase of acute HBV infection the levels of HBs antigens can be so low that even with very sensitive tests they may not be detected.

Finally, HBV serological markers also include anti-HBs antibodies that indicate long-term immunity and are positive, together with anti-HBc antibodies, in resolved HBV infections. Anti-HBs antibodies are the only serologic marker found in HBV vaccinated individuals.

SEROLOGICAL MARKERS				
	HBsAg	**anti-HBs**	**anti-HBc**	**anti-HBc IgM**
Uninfected, unvaccinated patient	-	-	-	-
After vaccination	-	+	-	-
Healed infection, patient protected	-	+	+	-
Old cured infection (spontaneous clearance of anti-HBs)	-	-	+	-
Acute infection	+	-	+	+
Chronic Infection	+	-	+	-

Table 3 - Hepatitis B infection Serological Markers adapted from **(Hadziyannis and Laras, 2018)**

2.4 Natural history of HBV infection

2.4.1 Acute HBV infection

Following infection, there is a latency period before an immune response can be detected (Ferrari et al., 2003). At the serological level, the first antibodies produced are anti-HBcAg antibodies. These antibodies indicate past or current infection. Antibodies directed against HBeAg antigens can also subsequently be detected. Within 6 months, an HBs seroconversion follows with the disappearance

of HBsAg and the appearance of anti-HBsAg antibodies that confer protective immunity to the patient and represent a clinical marker of successful virus clearance / control (*Table 3*).

The majority of patients are asymptomatic and only a few develop jaundice or flu-like symptoms. However, when the immune response is too active against infected hepatocytes, this can lead massive hepatocytes death and liver failure (fulminant hepatitis). Only 1% of patients with HBV infection develop fulminant hepatitis. Although infrequent, it is a serious condition as it is fatal in 50-80% of cases without a liver transplant.

2.4.2 Chronic HBV infection

In 5 to 10% of unvaccinated adults and in 90% of unvaccinated children, the combination of innate immune system failure in hepatocytes, defective anti-HBV T cell responses and the intrinsically immunosuppressive liver microenvironment, can lead to a chronic hepatitis B infection. This is diagnosed if HBsAg (and HBV DNA) are still detected 6 months after acute infection.

The natural history of the chronic carrier status goes through five major phases, defined by several clinical and virological parameters (transaminases levels, viral DNA, HBeAg and HBsAg in the serum): HBeAg-positive chronic infection (phase1), HBeAg-positive chronic hepatitis (phase 2), HBeAg-negative chronic infection (phase 3), HBeAg-negative chronic hepatitis (phase 4) and HBsAg-negative phase (phase 5) (*Table 4*) (Lampertico et al., 2017; Zoulim et al., 2018)

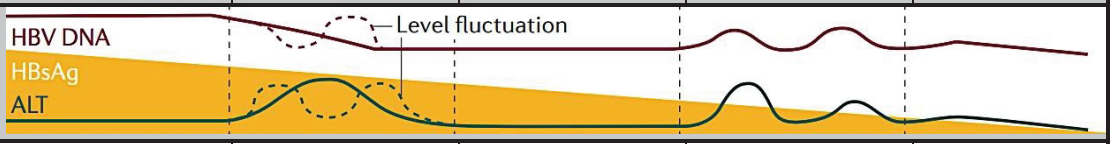
Serologic Markers of chronic hepatitis phases						
	ACUTE INFECTION	CHRONIC HEPATITS				
		PHASE 1	PHASE 2	PHASE 3	PHASE 4	PHASE 5
		HBeAg-positive chronic infection	HBeAg-positive chronic hepatitis	HBeAg-negative chronic infection	HBeAg-negative chronic hepatitis	Resolution of HBV - Occult infection
Old Terminology	/	Immune Tolerant	Immune reactive HBeAg+	Inactive Carrier	HBeAg- chronic hepatitis	/
Schematic level of HBV DNA, HBsAg and ALT	/					
HBsAg	Positive	High	High/ Intermediate	Low	Intermediate	Negative
HBeAg	Positive	Positive	Positive	Negative	Negative	Negative
Viral load	Positive	> 10^7 IU/ml	10^4-10^7 IU/ml	< 2,000 IU/ml	> 2,000 IU/ml	Negative
Anti-HBs	-	-	-	-	-	+
Anti-HBc	+	+	+	+	+	+
Anti-HBe		-	↑	++	↓	+
ASAT/ALAT	Normal	Normal	Elevated	Normal	Elevated	Normal
Liver Disease		None or minimal	Moderate to severe	Moderate to severe	Moderate to severe	None or sequellae (Fibrosis and HCC)
Treatment indicated	-	-	+	-	+	-

Table 4 – Hepatitis B natural history

adapted from (Zoulim et al., 2018)

2.4.2.1 HBeAg-positive chronic infection (immune tolerant)

Patients in the *HBeAg-positive chronic infection* phase, previously named *immune tolerance* phase, have been long considered to be in a disease-free phase with absent or minimal necrotico-inflammatory activity in the liver, low hepatocyte mortality and often normal levels of transaminases (Papatheodoridis et al., 2012). This phase is characterized by high serum HBV DNA levels (up to 10^{10} copies of HBV/mL of serum) as well as high HBeAg and HBsAg titers. The immune response is limited (Hui et al., 2007) and patients have low levels of activated HBV-specific CD8+ T cells. Patients generally remain between 10 and 30 years in a chronic HBeAg positive infection (Burns and Thompson, 2014). Due to the low level of liver disease the treatment of these patients is not recommended by current clinical guidelines from liver scientific societies (EASL, AASLD, APASL). However, these patients are very contagious because they have high viraemia. Moreover, recent data indicate that HBV DNA integration and clonal expansion of hepatocytes carrying HBV integrations, a condition that may contribute in the long term to the development of a liver cancer, are already detected in this early phase of chronic infection (Mason et al., 2016).

2.4.2.2 HBeAg-positive chronic hepatitis (immune reactive HBeAg-positive)

After variable intervals of time, that can be up to decades, patients with HBeAg-positive chronic infection experience, for reasons and with mechanisms that are still poorly defined, have an activation of HBV-specific immune cells with liver inflammation and progress to the immune reactive HBeAg-positive chronic hepatitis phase. In this phase, HBV DNA levels can fluctuate but tend to decrease (Lampertico et al., 2017; Liu et al., 2006a) and repeated hepatic flares caused by immune-mediated liver damage are observed (Lampertico et al., 2017; Wang et al., 2010). These patients also have necrotico-inflammatory lesions in the liver and faster fibrosis progression than in other phases (Liaw, 2009). HBeAg level decreases and when HBeAg becomes negative and anti-HBe antibodies can be detected (HBeAg seroconversion) the disease activity subsides. Treatment with anti-HBV nucleos(t)ides (NUCs) is indicated in HBeAg-positive chronic hepatitis patients (Lampertico et al., 2017) to increases the rate of HBe/anti-HBe seroconversion. After HBe seroconversion, that is also accompanied in few cases by a loss of HBsAg and anti-HBs seroconversion, NUCs treatment can be stopped.

2.4.2.3 HBeAg-negative chronic infection (inactive HBV carrier)

After seroconversion to the anti-HBeAg positive state, the elimination of many HBV infected hepatocytes during the HBeAg-positive chronic hepatitis phase, allows many patients to progress into an inactive HBV carrier phase defined as HBeAg-negative chronic infection. Both disease activity and viral replication are low in these patients with normal transaminases, HBV DNA less than 2000 IU/mL or even undetectable (Liu et al., 2016) and the progression of liver damage is slow compared to the previous phase. A patient is confirmed as an inactive carrier after a minimum of one-year follow-up during which transaminases and HBV DNA are analyzed every 3 to 6 months. Inactive HBV carriers continue to be monitored for potential relapses. The long-term prognosis is globally favorable but depends, in the single patient, of the level of fibrosis accumulated during the immune reactive phase before seroconversion (Invernizzi et al., 2016).

2.4.2.4 HBeAg-negative chronic hepatitis

This phase may develop rapidly after the HBe seroconversion or after several years of inactive phase. The chronic hepatitis HBeAg-negative phase is characterized by a reactivation of virus replication leading again to chronic liver inflammation. The patient undergoes periodic reactivations with variations in HBV DNA and transaminases levels (European Association for the Study of the Liver, 2012). Many patients in the chronic hepatitis HBeAg-negative phase carry HBV mutants that cannot produce HBeAg (see section 1.5.2.1 The Precore variants) (Hadziyannis and Vassilopoulos, 2001). These patients often present a more progressive liver disease and low remission rates (Lampertico et al., 2017) and should be treated long-term with NUCs (Lampertico et al., 2017).

2.4.2.5 HBsAg-negative phase / functional cure / occult HBV infection

Each year, only 1-2% of patients in the HBeAg-negative chronic hepatitis phase progress to a so-called HBsAg-negative or occult infection phase (Chu and Liaw, 2010). This phase is defined as the presence of replication-competent HBV DNA (i.e. episomal HBV covalently closed circular DNA [cccDNA]) in the liver and/or HBV DNA in the blood of people who test negative for hepatitis B surface antigen (HBsAg) by currently available assays. These patients may be or not positive for anti-HBV antibodies (anti-HBc and less frequently anti-HBc and anti-HBs) and, in the absence of other causes of liver damage, have normal ALT levels (Raimondo et al., 2019). Viral replication is kept suppressed by the adaptive immune system and in most cases the prognosis is benign. However, if another cause of chronic liver disease is present, occult HBV may confer an accelerated progression to cirrhosis and HCC development (Raimondo et al., 2019). Finally, since cccDNA is still present in the liver, immunosuppressive and anti-neoplastic treatments (targeted therapies, chemotherapy and radiotherapy) may lead to reactivation of viral replication (Chemin and Trépo, 2005; Raimondo et al., 2019).

2.4.3 Prevention of HBV infection: prophylactic vaccine

There are several commercial vaccines containing as immunogen HBsAg proteins obtained by genetic recombination in the yeast Saccharomyces cerevisiae. These HBsAg vaccine induce an induction of T-helper lymphocytes to produce protective antibodies against the virus.

In the U.S., 3 different vaccines are produced: Recombivax HB® produced by Merck (U.S.), Engerix-B® produced by SmithKline Beecham Biologicals (Belgium) (Centers for Disease Control and Prevention (CDC), 2000) Twinrix®, a combined hepatitis A (HAVRIX®) and B vaccine (Engerix-B®) is produced by GlaxoSimthKline (Rendi-Wagner et al., 2001). To facilitate administration in childhood vaccination programs, polyvalent vaccines including HBV are also available for injection in children.

In France, 4 monovalent vaccines are marketed: Engerix®, Genhevac B®, Fendrix® and HBVaxPro® (Barril and Teruel, 2008). Other HBV vaccines available in France include: a) the Sci-B-Vac® vaccine includes all the 3 envelope proteins of the virus (SHBs, MHBs and LHBs) (Alon et al., 2017); b) the Twinrix® bivalent vaccine for hepatitis A and B, the Infanrix Hexa® (GlaxoSimthKline); c) the hexavalent vaccine for infants for vaccination against six diseases: diphtheria, tetanus, pertussis, hepatitis B, polio and invasive Haemophilus influenzae type b infections (Van Der Meeren et al., 2012).

WHO recommends the early vaccination of newborns and of individuals at high risk of infection including individuals travelling to endemic areas, sharing the household with an infected CHB patient, health and prison professionals, individuals with health problems, drug users, individuals fond of tattooing or piercing and individuals with multiple sexual partners. Three (0, 1, and 6 months) or four (0, 1, 2, and 12 months) doses of the vaccine are given intramuscularly and the patient is considered protected when serology shows antibody levels greater than 10 IU/L. The HBV vaccines are effective in 90% of adults and 95% of children. Efficacy may decrease in immunosuppressed individuals, obese, smokers (Mast et al., 2006). About 75% of patients who respond to the vaccine maintain their anti-HBs positive for 60 years. For patients who do not respond to the vaccine, a new dose can be administered and leads to immunity in 25 to 50% of individuals and a triple dose leads to immunity in 44 to 100% of cases (Chang and Nguyen, 2017). Of note, Recombivax HB® and Engerix® vaccines do not contain Thiomersal, a preservative that has been claimed to induce vaccination side effects (Geier et al., 2014). An association between multiple sclerosis and vaccination against hepatitis B virus has been suggested. One study showed an increased risk of developing multiple sclerosis with

an odds ratio of 3.1 and a 95% confidence interval of 1.5- 6.3 (Hernán et al., 2004), but it has not been confirmed in subsequent studies.

The best way to reduce HBV endemicity in all countries is to prevent mother-child infection as well as infection in childhood (Jourdain et al., 2019). To this aim newborns should be vaccinated within 24 hours of delivery from HBV-positive mother. It is estimated that vaccination at birth with a 90% coverage would prevent 84% of HBV-related deaths. The number of countries with early childhood immunization policies has increased from 31 countries in 1992 to 184 countries in 2014. In 2014, 82% of children worldwide received the vaccine but only 38% of newborns were vaccinated at birth. Moreover, in 2015 only 59% of countries vaccinated 90% of babies. In the U.S., lower HBV vaccine completion was associated with ethnic minority groups and inadequate health insurance coverage (Chang and Nguyen, 2017).
Other measures to reduce infection risks include immunoprophylaxis that may be administered to people who have not been vaccinated after exposure to the virus to prevent infection. Newborns whose mothers are infected with HBV may also receive these immunoglobulins (Lo et al., 1985).

2.4.4 Antiviral therapies

Despite the availability of prophylactic vaccines and antiviral therapies that suppress viral replication, no curative treatment has been developed to date. International initiatives have emerged, such as the ICE-HBV or the WHO World Hepatitis Day, to promote the global efforts towards *HBV cure*.

In May 2016, the World Health Organization (WHO) adopted a global hepatitis strategy with the goal of eliminating viral hepatitis as a public health threat by 2030 (www.who.int/hepatitis/publications/hep-elimination-by-2030-brief/en/). The targets to be achieved by 2030 are: a 90% reduction in new cases of chronic hepatitis B and C; a 65% reduction in mortality due to hepatitis B and C; a 80% of treatment-eligible persons with chronic hepatitis B and C infections being treated.

The ultimate goals of hepatitis B therapy are to improve the quality of life and survival of HBV infected patients. To this end, treatments aim to: a) suppress viral replication and liver inflammation; b) delay

or decrease fibrosis progression and cirrhosis development; c) prevent liver failure, hepatocellular carcinoma other complications of cirrhosis (Lampertico et al., 2017).

Current therapeutic agents include nucleos(t)ide analogs (NA or NUCs) and interferons. HBV treatment with these drugs is recommended for patients with virological activity and liver inflammation (e.g., high viral replication and high levels of ALT). The goal of therapy is to achieve one or more of the virological endpoints that are considered relevant and associated with the clinical endpoints listed above: a) significant reduction or negativization of viremia; b) HBeAg loss and anti-HBe seroconversion; c) HBsAg loss with or without anti-HBs seroconversion. Very few patients maintain viral suppression off-treatment and even fewer loses HBsAg. Loss of HBs antigens with or without the production of anti-HBs antibodies allows to stop treatment safely and is associated with a long-term improvement in the clinical outcome (Fanning et al., 2019; Lampertico et al., 2017).

Mechanistically, four strategies can be used to treat chronic hepatitis patients:
- inhibit HBV replication
- restore defective host innate and adaptive immune responses
- sensitize selectively infected hepatocytes to immune-mediated elimination
- target the cccDNA pool to :
    a) decrease cccDNA (inhibit virus entry and *de novo* infection; block rcDNA to cccDNA conversion)
    c) silence cccDNA (permanent suppression of cccDNA transcription)
    b) eliminate cccDNA (cccDNA destabilization/degradation or destruction of infected cells) (Fanning et al., 2019)

The American Association for the Study of the Liver Diseases (AASLD) and the European Association for the Study of the Liver (EASL) have defined four possible levels of HBV cure after HBV treatment, ***beyond the suppression of HBV replication with continuous treatment*** (*Table 5*) (Cornberg et al., 2020).

	Serum		Liver		
	HBV DNA (serum)	**HBsAg (serum)**	**cccDNA transcription**	**cccDNA**	**Integrated HBV DNA**
Complete sterilizing cure	Undetectable	Undetectable	Absent	Undetectable	Undetectable
Complete cure	Undetectable	Undetectable	Absent	Undetectable	Detectable
Functional cure	Undetectable	Undetectable	Suppressed	Low	Detectable
Partial cure	Undetectable	Detectable	Low or suppressed	Low	Detectable

Table 5 - HBV cure definitions and treatment outcomes

Functional cure is defined by a sustained off-treatment loss of HBsAg that is associated with an improved long-term clinical outcome (remission of liver inflammation, decrease in the risks of cirrhosis and HCC and low rates of disease reactivation) (Cornberg et al., 2020). From the virological point of view *functional cure* requires a sustained shutdown of viral replication with undetectable viral load and HBsAg in serum. In the liver low levels of transcriptionally inactive cccDNA with various degrees of epigenetic silencing can persist. Innate and adaptive antiviral immune responses control this residual reservoir of viral minichromosome reservoir (Fanning et al., 2019). The presence and persistence of integrated viral sequences can contribute to maintain the exhaustion of immune responses though the expression of viral antigens and HBV oncogenic events (Pollicino and Caminiti, 2021). ***Partial cure*** is defined as a sustained undetectable serum HBV DNA *off treatment* with low levels of HBsAg (Cornberg et al., 2020) and reflects HBsAg expression from integrated sequences despite a sustained control of cccDNA activity (Fanning et al., 2019). The new therapies that are being developed aim to achieve more optimistic goals: a) ***complete cure*** with a complete eradication of cccDNA and no residual risk of viral reactivation and b) ***sterilizing cure*** with the elimination of all intrahepatic viral sequences including cccDNA and integrated sequences and a further reduction of HCC risk.

2.4.4.1 Current anti-HBV treatments

Eight treatments have been approved by the Food and Drug Administration (FDA) in the United States: two formulations of interferons and 6 Nucleos(t)ide analogues (Chang et al., 2014)(*Table 6*). Availability of the different treatments varies from region to region and in different countries within the same region according to approvals and reimboursability rules by local authorities.

Substance	Date and classification	Abbreviations	Trade Name	Percentage of resistant patients	log_{10} reduction	HIV Treatment
Lamivudine	1995 Nucleoside Analogue	3TC	Epivir, Zeffix, Heptovir	20% and 70% after 2 and 4 years	5.5 log_{10}	Yes
Adefovir Dipivoxil	2002 Nucleotide Analogue	PMEA	Hepsera, Preveon	3%, 9%, 18%, and 28% after 2, 3, 4, and 5 years	4.9 log_{10}	No
Entecavir	2005 Nucleoside Analogue	ETV	Baraclude	0.8% after 3 years (or 25% after 3 years for patients who have failed lamivudine therapy)	6.9 log_{10}	No
Tenofovir disoproxil fumarate	2001 Nucleotide Analogue	TDF	Viread	0% after 6 years	4-5.5 log_{10}	Yes
Tenofovir alafenamide	2001 Nucleotide Analogue	TAF	Vemlidy	0% after 2 years	4-5.5 log_{10}	Yes
Telbivudine	2006 Nucleoside Analogue	LdT	Sebivo	5% after 1 year	6.3 log_{10}	No
Emtricitabine	2004 Nucleoside Analogue	FTC	Emtriva Coviracil Racivir	18% after 2 years	4-5.5 log_{10}	No

Table 6 - HBV Nucleos(t)ide Analogue HBV treatments

adapted from (Lampertico et al., 2017; Lim, 2017)

2.4.4.1.1 Nucleos(t)ide analogues

Nucleos(t)ide Analogues (NA or NUCs) are currently the treatment of choice. They inhibit viral replication through the inhibition of the viral polymerase. NA inhibit polymerisation of the negative or positive strand by their incorporation into a neo-synthesized DNA chain as they lack a 3' hydroxyl group, which prevents the downstream addition of a new nucleotide (De Clercq, 1999).

Lamivudine was one of the first antiviral treatments used. Despite the 3 to 5 $\log_{10}$ copies/mL viral load decrease after one year of treatment, nearly 70% of patients after 4 years developed a resistance found in or around the active Tyr-Met-Asp-Asp site of the polymerase. Other treatments also have high resistance rates such as adefovir with a resistance rate of 30% after 5 years of treatment (Tillmann, 2007) or Telvibudine, particularly when treatment was started in patients with very high viral load (Lampertico et al., 2017). Current preferred nucleos(t)ide analogues, entecavir (ETV), tenofovir disoproxil fumarate (TDF) and tenofovir alafenamide (TAF), offer highly effective suppressive therapy with very low rates of resistance with long-term therapy (Table 6).

Third generation NUCs efficiently suppress HBV DNA replication with a decrease in serum HBV DNA by up to 6 $\log_{10}$ after one year of treatment (Valaydon and Locarnini, 2017) but not all patients on therapy achieve and maintain an undetectable viral load (Lampertico et al., 2017).

NUCs do not have direct effects on production, transcription, or stability of cccDNA. However, NUCs can indirectly reduce the cccDNA pool by limiting mature capsids recycling to the nucleus (Fanning et al., 2019) and induce a partial restoration of HBV specific adaptive immune responses (Bertoletti and Ferrari, 2016). The impact on cccDNA is slow, with only a 0.8 $\log_{10}$ of cccDNA reduction after 1 year of treatment (Werle-Lapostolle et al., 2004). Clinically this translates into a low rate of HBsAg loss (Cao et al., 2014) and virological relapse occurs in the vast majority of patients even when treatment is stopped after many years of undetectable serum HBV DNA. Clinical trials of entecavir or tenofovir have shown only 1-3% HBsAg loss after one year of treatment, with minimal increase to 2.5-4.9% after 7-10 years of continuous (Jeng and Lok, 2021). Higher rates of HBsAg loss have been reported among patients who were HBeAg positive pretreatment likely due to preponderance of integrated HBV DNA as the source of HBsAg in HBeAg negative patients (Lampertico et al., 2017).

2.4.4.1.2 Interferons

Patients with HBV chronic infection are also treated with interferons (IFN), either standard interferons or, since the 1990s, pegylated alpha 2a interferons, administered parenterally. Pegylation increases the half-life of the molecule and reduces the administration to once a week for 48 weeks (Cooksley et al., 2003).

Class I interferons have multiple antiviral activities (reduction of cccDNA transcription, cccDNA deamination and partial degradation, accelerated capsids decay) (Fanning et al., 2019) and immunomodulatory effects (activation of NK cells and dendritic cells) (Bertoletti and Ferrari, 2016).

Only 30-40% of chronically infected patients respond to Interferons (Liu et al., 2006b) they are more effective for genotypes A and B but much less effective for genotypes C and D (Lampertico et al., 2017). Interferon treatment is accompanied in a substantial proportion of patients by systemic adverse effects (flu-like symptoms, headache, fatigue), hematological toxicity (leukopenia and thrombocytopenia), local reactions at the injection site, thyroid disorders, mood changes and psychiatric manifestations (Janssen et al., 2005) and is contraindicated in a number of cases (decompensated cirrhosis and pregnancy).

Compared to NA, interferons result in a greater decrease in HBe and HBs antigens (especially in genotype A patients). Interferons are less effective in suppressing viral replication but response are more durable over time (Lok et al., 2017). Various strategies combining NUC and pegylated-IFN have been evaluated to increase the rate of HBsAg loss. One study found that de novo combination of tenofovir DF and pegylated-IFN for 48 weeks resulted in an overall higher HBsAg loss rate 24 weeks post-treatment (average 8.1%; 37.5% vs. 7% in genotype A vs. non-A) as compared to pegylated-IFN monotherapy (2.9%) (Marcellin, 2016). A recent meta-analysis of 60 studies found that de novo combination of NUC and IFN improved HBsAg loss rate compared to NUC monotherapy but not IFN monotherapy (Liu, 2020). In patients already treated with NUCs switching to IFN appeared to have a more marked effect on increasing HBsAg loss compared to adding IFN.

2.4.4.2 Novel antiviral therapies

The barriers to achieve a ***cure*** for HBV are related to the lack of drugs that target directly the cccDNA and its long half-life of cccDNA and/or can restore the defects of the innate and adaptive immune system in chronic hepatitis B patients. Different antiviral and immunomodulatory strategies are explored with the final goal to reduce, silence, or eliminate cccDNA (Figure 19). Some of these new molecular entities (NMEs) are still in their preclinical development but the last 24 months have seen a rapid increase in NMEs with potent antiviral or immunomodulator effects *in vitro* entering clinical development.

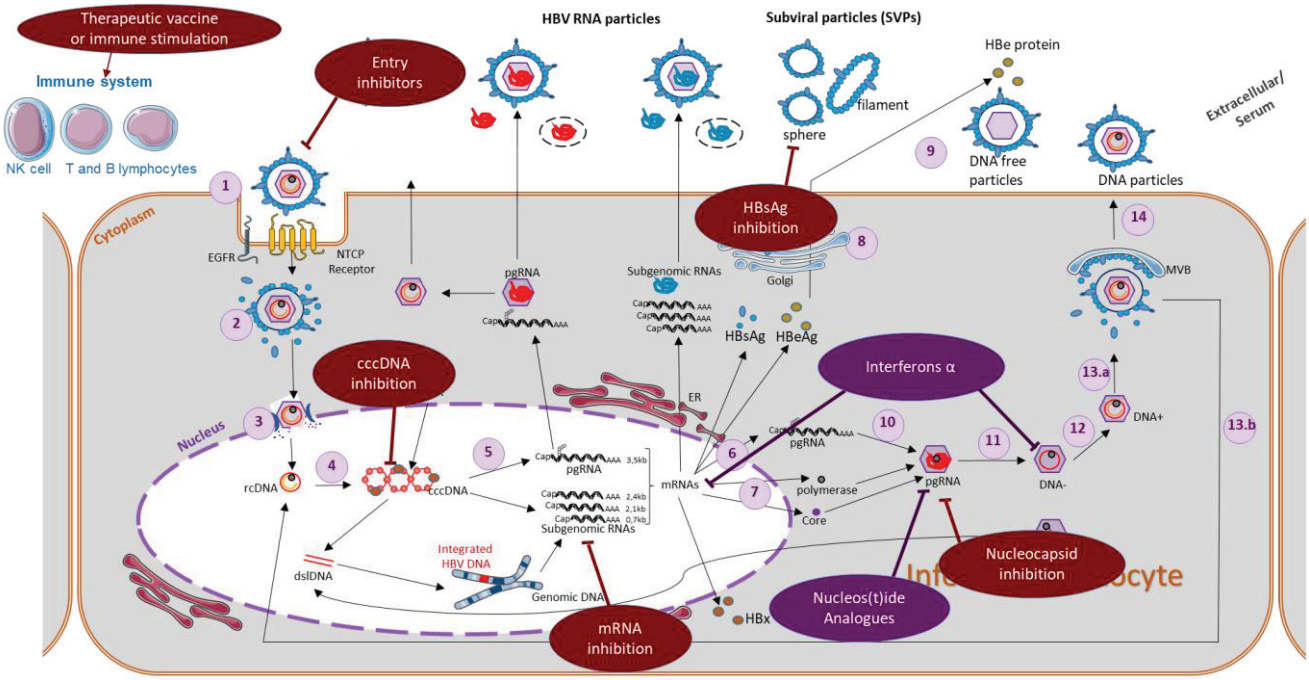


Figure 19 - Targets of current and Novel antiviral Therapies adapted from (Fanning et al., 2019; Lok et al., 2017)

In purple are represented the current antiviral therapies and in red are represented the novel antiviral therapies.

RE and immune cells pattern are from smart.servier.com website

2.4.4.2.1 HBV entry inhibitors

Inhibition of virus entry into the hepatocyte by the NTCP receptor may prevent infection of new hepatocytes and, indirectly, decrease the cccDNA reservoir, decrease replication and virus release, several strategies exist to inhibit virus entry:

- **Irreversible NTCP inhibitors -** allosteric NTCP inhibitors. These molecules have a very long half-life but also block the passage of bile salts and any other molecules taken up by the NTCP receptor. Myrcludex B and cyclosporin A belong to this class (Lok et al., 2017).
- **NTCP substrates** - all molecules transportable by NTCP (taurocholate, ezetimibe). Treatment requires high concentrations and these molecules have a very short half-life which limits clinical application.
- **Attachment inhibitors** - positively or negatively charged drugs that bind to the virus or proteoglycans. This approach can be effective but it is non-specific.
- **Neutralizing antibodies** - targeting the PreS1 domain. Antibodies are highly specific but parenteral administration and large amounts of antibodies are required to neutralize viral particles.

2.4.4.2.2 Direct targeting cccDNA

Strategies that directly target the cccDNA pool include the block of rcDNA conversion to cccDNA, the silencing of cccDNA with a permanent suppression of its transcription, the elimination of cccDNA by destabilization/degradation.

- Cellular enzymes involved in ***cccDNA formation*** such as TDP2, FEN1, POLK are potential therapeutic targets. However, as for other host targeting drugs the potential for *off target* effects and functional redundancy that would limit the impact of inhibiting only a given target must be carefully considered.
- **Cytokines** such as IFN-α or IFN-γ, lymphotoxin-β agonist receptors, and Tumor Necrosis Factor alpha (TNF-α) can up-regulate APOBEC3A or APOBEC3B deaminases and target to cccDNA

degradation. However, these cytokines allow only a partial degradation of the cccDNA pool (Lucifora et al., 2014; Xia et al., 2016)

- **DNA cleaving enzymes**: once the cccDNA is formed, it can be targeted by different types of *molecular scissors* such as endonucleases, meganucleases, CRISPR-associated nucleases. The CRIPR-Cas9 approach uses guide RNAs to specifically target viral DNA. This strategy has already shown promise in preclinical experiments with 90% of viral DNA being cleaved by this enzyme and only 7% of cleaved viral genomes generating "mutants" that are still functional (Seeger and Sohn, 2016). Thus, the use of a combination of several guide RNAs could decrease the percentage of functional mutants and lead to a potential therapy (Martinez et al., 2021).

2.4.4.2.3 cccDNA silencing

Interferon-a (IFN-a) and interleukin-6 (IL-6) have both been shown to decrease cccDNA-bound histone acetylation and thereby transcriptional activity (Belloni et al., 2012; Liu et al., 2013; Palumbo et al., 2015; Tropberger et al., 2015). Many of the host chromatin modifying enzymes that bind the cccDNA (eg., the acetyltransferases CBP/p300, PCAF/CGN5, the histone deacetylases HDAC1, HDAC4, SIRT1 and SIRT3, the methyltransferases SETDB1, PRMT1, PRMT5, Set1A and the demethylases LSD1) are druggable and their modulation affects cccDNA transcription and HBV replication in cell culture HBV infection models but the intrinsic redundancy of epigenetic pathways and the potential for unwanted *off target* effects on host target genes are potential limits for their therapeutic use.

Viral proteins that associate with the cccDNA are potentially of greater interest. Whereas the contribution of cccDNA bound HBc to cccDNA maintenance and transcription is still a matter of investigation, the essential role of HBx in cccDNA transcription and HBV replication and is well established. HBx promotes cccDNA transcription by hijacking the DDB1-CUL4-ROC1 E3 ligase complex to degrade the cellular protein SMC5/6, a restriction factor against episomal DNAs that silences cccDNA transcription (Decorsière et al., 2016). HBx-DDB1 interaction is therefore a potential therapeutic target. Nitazoxanide, a thiazolide-type anti-infective agent used in helminthic, protozoal, and viral infections, has been shown to inhibit HBx-DDB1 interaction and to restore SMC5 levels in cultured hepatocytes (Sekiba et al., 2019a). Nitazoxanide has been tested in a pilot uncontrolled clinical trial recruiting a limited number of patients and showed some antiviral activity against HBV,

but intrahepatic viral parameters were not evaluated. Patients were rolled over to standard NUC therapy and the impact of Nitazoxanide off therapy could not be evaluated (Rossignol and Bréchot, 2019). Proteasome degradation of the SMC5/6 complex also requires the activation of the NEDD8 protein complex and the inhibition of the NEDD8-activating enzyme (NAE) restores levels of SMC5/6 and prevents cccDNA transcription (Sekiba et al., 2019b). Based on these observations the NAE inhibitor pevonedistat is currently in a Phase Ib clinical trial (Lockhart et al., 2019).

2.4.4.2.4 Translation inhibitors

Reduction of HBV antigens could contribute to HBV cure through direct inhibition of virion and subviral particle production and indirect boosting host innate and HBV-specific immune responses. Approaches to reduce HBV antigen burden include, in addition to the inhibitors of cccDNA transcription described above, mRNA destabilisers and translation inhibitors (small interfering (si)RNAs and antisense oligonucleotides (ASOs) (Fanning et al., 2019). Drug development of mRNA destabilizers has been hampered by toxicities inherent with targeting host pathways. In contrast, translation inhibitors appear to be safe and well-tolerated. Translation inhibitors include small interfering RNAs targeting viral transcripts for degradation by the RISC/Ago2 complex and antisense oligonucleotides that bind and degrade them by recruiting RNAseH. Both the Gal-Nac conjugated siRNAs and the naked ASOs are administered through subcutaneous injection.

The first-generation siRNA ARC-520 which targets all viral RNAs decreased in viral replication and the HBsAg production was due to its design failed to target HBs transcripts transcribed from integrated HBV DNA (Yuen et al., 2019). All 4 second generation siRNAs in Phase 2 clinical development (JNJ-3989; VIR-2218; RG-6346; AB-729) achieve predictable *on-treatment* HBsAg responses, which appear to be durable for many months after the final dose. Monthly dosing of siRNAs results in stepwise reductions in HBsAg. Indeed, a profound antigen reduction is thought to be needed to achieve restoration of host immune responses, as a prerequisite for off-treatment immune control. The preliminary results from the first Phase II studies of siRNA plus NUCs show that HBsAg reduction plateau after 4-6 months and patients do not achieve HBsAg loss. The absence of ALT flares excludes the risk of hepatotoxicity from off-target binding but also suggests a lack of immune restoration and the need to combine with an immune modulator. Conversely, the ASO GSK-836 study has shown up

to 4 log_{10} reductions in HBsAg, associated with ALT flares. Current GSK-836 studies will determine whether ASOs can achieve sustained HBsAg loss and functional cure.

2.4.4.2.5 Capsid assembly modulators

Capsid Assembly Modulators (CAMs) have two specific mechanisms of action: (i) inhibition of HBV replication by interfering with HBV capsid assembly and encapsidation of pgRNA and (ii) inhibition of de-novo cccDNA replenishment, by interfering with capsid disassembly (Fanning et al., 2019; Ko et al., 2019; Schinazi et al., 2018).

There are two classes of CAMs:

- **The heteroaryldihydropyrimidines (HAP) family or class I capsid inhibitors.** They induce the formation of aggregates and aberrant capsid structure. From the prototype Bay 41-4109, several molecules have been synthesized and entered clinical development: GLS4 (Fanning et al., 2019; Wu et al., 2013) and RO7049389 (Yuen et al., 2021).
- **The phenylpropenamide (PPA) and sulfamoylbenzamide (SBA) families or class II capsid inhibitors.** These CAMs accelerate the formation of capsid particles with normal or quasi-normal structure.

There are currently more than 20 CAMs in preclinical or clinical development. Second generation CAMs, JNJ-379, RO7049389, ABI-H0731 (vebicorvir) and ABI-2158 all show improved antiviral potency. In Phase II studies, CAMs are combined with a NUC to prevent virologic breakthrough from selection of CAM-resistant variants and to intensify the *on-treatment* HBV DNA suppression. Notably, the association of vebicorvir plus NUC study failed to significantly reduce HBsAg levels and the subgroup of patients that stopped treatment relapsed within 4 months, suggesting that the combination of 2 antivirals is not sufficient to achieve functional cure.

A third generation of CAMs (ABI-H3733, AB-836, ALG-000184, VNRX-9945) with 2-3 log_{10} greater *in vitro* antiviral potency are entering clinical development.

2.4.4.2.6 Inhibitors of HBsAg particles formation/egress

Several compounds have been shown to have effects on HBsAg particle formation and/or egress.

The Benzimidazole derivative BM601 interferes with the aggregation of surface proteins in the Golgi apparatus leading to a decrease in the secretion of HBsAg and complete virions (Xu et al., 2014).

Nucleic acid polymers (NAP) efficiently reduce HBs antigens levels in the hepatocytes and in serum by a mechanism that is still poorly understood. A second mode of action of NAPs is preventing virus entry. Clinical trials of REP 2055 and REP 2139 NAPs used as monotherapy followed by a combination with IFN or thymosin alpha 1 have led to a significant decrease in the level of circulating HBsAg, viremia and anti-HBs seroconversion in a significant proportion of patients. Similar results were obtained in preliminary studies with REP 2139 and REP 2165 in triple combination with tenofovir and IFN (Lok et al., 2017).

Monoclonal antibodies directed against HBsAg have been proposed to deplete or neutralize HBsAg particles in the bloodstream. However, their impact on viral replication is still unproven and probably they would not be effective unless coupled with other approaches (Golsaz Shirazi et al., 2014).

Finally, the triazol-o-pyrimidine derivative HBF-0259, targeting both SCCA1 (involved in HBsAg secretion) and cyclophylin A (CypA) (involved in post-NTCP steps of virus entry) reduces sub-viral particles and DNA virions secretion (Dougherty et al., 2007; Mohebbi et al., 2016).

2.4.4.2.7 Immunomodulators

The host immune response to HBV infection is a major determinant of clinical outcome following acute infection. Broad HBV antigen specific CD8+ T cell responses are required to clear HBV infection. Approaches to activate antiviral immunity against HBV include (i) stimulating antiviral effector T and B lymphocytes and antigen presenting cells with toll-like receptor agonists; (ii) rescuing the T-cell exhaustion which accompanies chronic HBV infection; (iii) generating “new” T-cells with T-cell therapeutic vaccines (Fanning et al., 2019).

The RIG-I agonist SB-9200 induces the production of IFN and other cytokines (Korolowicz et al.) and inhibits HBV replication (Li et al., 2018a). SB-9200 development has been discontinued due to toxicity in the Phase II trial of combination with NUCs.

TLR-7 and 8 agonists have potent antiviral activity in woodchucks and chimpanzees (Boni et al., 2018) but in patients with chronic hepatitis B, these agents have shown modest antiviral activity and have been associated with dose limiting toxicities (Gane et al., 2021; Janssen et al., 2018). The roles of TLR-7 and TLR-8 agonists in future combination therapies are currently being evaluated in Phase II studies. PD-1/L1 blockade should reverse T-cell exhaustion associated with life-long HBV infection. A small pilot study demonstrated that a single dose nivolumab (1/10 dose approved for cancer therapy) was safe and well-tolerated and led to HBsAg decline in 20/22 patients with one patient achieving sustained HBsAg loss (add Gane 2019). Multiple Phase 2 studies are evaluating the role of anti-PD1 mAbs and siRNAs in combination with other new antiviral and immunomodulatory agents.

Several strategies to develop HBV therapeutic vaccines have been tested protein-based vaccines, DNA-based vaccines, live vector-based vaccines, peptide-based vaccines and cell-based therapies (Li et al., 2017). Therapeutic vaccine development in chronic hepatitis B has been often described as a graveyard. High dose HBsAg vaccines as monotherapy have shown no conclusive effect. Other viral proteins (HBc in the DV-601, ABX203 vaccines or HBx in the GS-4774 vaccine) have been tested (Alonso et al., 2017; Spellman and Martin, 2011) without any consistent clinical effect. Similarly, DNA-based vaccines, such as the INO-1800 vaccine, containing a plasmid coding for HBsAg and HBcAg (Manzoor et al., 2015) or HB-100, which codes for surface, core, Pol and HBx proteins (Ha et al., 2002; Yang et al., 2006) and live vector-based vaccines have shown variable capacity to induce some anti-HBV immune responses but no clinically relevant impact on virological parameters.

Recent advances in vaccine technology combined with the availability of other novel agents which may restore HBV-specific immunity (siRNAs, PD-1/L1 blockage) have provided the conceptual and practical frame for new immunomodulatory combinatory approaches (Michler et al., 2020).

CHAPTER III: NEW HBV BIOMARKERS

3.1 HBV infection biomarkers

HBV remains a major global public health problem with over 240 million chronic carriers. As previously discussed, with current treatments for HBV infection (nucleotide analogues and interferons). However, only a small proportion of patients achieve a functional cure, defined by the loss of the HBsAg, thus allowing the treatment to be stopped. The development of new antiviral therapies targeting the cccDNA reservoir is a major objective towards HBV cure. Several new therapeutic molecules (new antivirals and immuno-modulators) are at the preclinical or clinical evaluation stage with the aim of inducing a cure with a time-limited treatment.

To quantify cccDNA levels and its transcriptional activity, it is important not only to evaluate the size and the activity of the cccDNA pool in the different phases of the disease but also to measure the impact of new treatments. The best indicator of cccDNA transcriptional activity would be the ratio of intrahepatic cccDNA / 3.5 kb RNAs, requiring a liver biopsy. However, liver biopsy is an invasive procedure that is, nowadays, rarely performed outside clinical studies/trials. In addition, since liver infection is heterogeneously distributed in the liver, it is not certain that the small amount of liver

tissue taken from the biopsy would reflect the whole liver. Furthermore, despite technological advances and harmonization efforts, cccDNA quantification remains a difficult technique limited to specialized laboratories and did not enter clinical virology routine. Indeed, most samples also contain large amounts of rcDNA, which is from the analytical point of view very similar to cccDNA, making it even more difficult to quantify (Nassal, 2015).

Current biomarkers used to evaluate treatment activity are the normalization of ALTs, HBV DNA non-detectability, HBeAg seroconversion (in HBeAg positive patients), HBsAg decrease and HBsAg loss. An improvement in inflammation and fibrosis at the histological level is also expected. ALTs alone cannot be relied upon to predict patient outcomes since certain comorbidities may interfere with ALT levels, in particular hepatic steatosis (Enomoto et al., 2016). HBV DNA in serum is the best indicator of the antiviral activity of a new treatment in naive patients but the suppression of HBV DNA in serum is not an appropriate efficacy endpoint for patients already treated with NUCs (Figure 20). HBsAg loss cannot be considered a biomarker of the activity of new therapeutic molecules since it is at the same time the desired endpoint and part of the definition of cure. Moreover, each year only 1-2% of patients treated with NA lose HBsAg (Tseng et al., 2012) and HBsAg antigens may derive from the integration of the viral genome into the host genome and not fully reflect the transcriptional activity of cccDNA (Wooddell et al., 2017). The predictive value of different thresholds of HBsAg decrease remains to be validated and the kinetic of HBsAg decrease is often too slow to allow early outcome prediction and the conduction in clinical trials (Zoulim et al., 2015). Altogether, there is an urgent need to find new biomarkers that, by reflecting cccDNA pool size and activity, are capable to predict outcomes with the new treatments. These new biomarkers would also be useful: a) to better characterize patients in the different phases of the HBV infection and disease; b) to predict response to IFN at an early stage allowing to extend the treatment only in patients with high probability to reach a functional cure; and c) would allow to identify patients that could discontinue NUCs without risk of relapse (Charre et al., 2019; Fan et al., 2020).

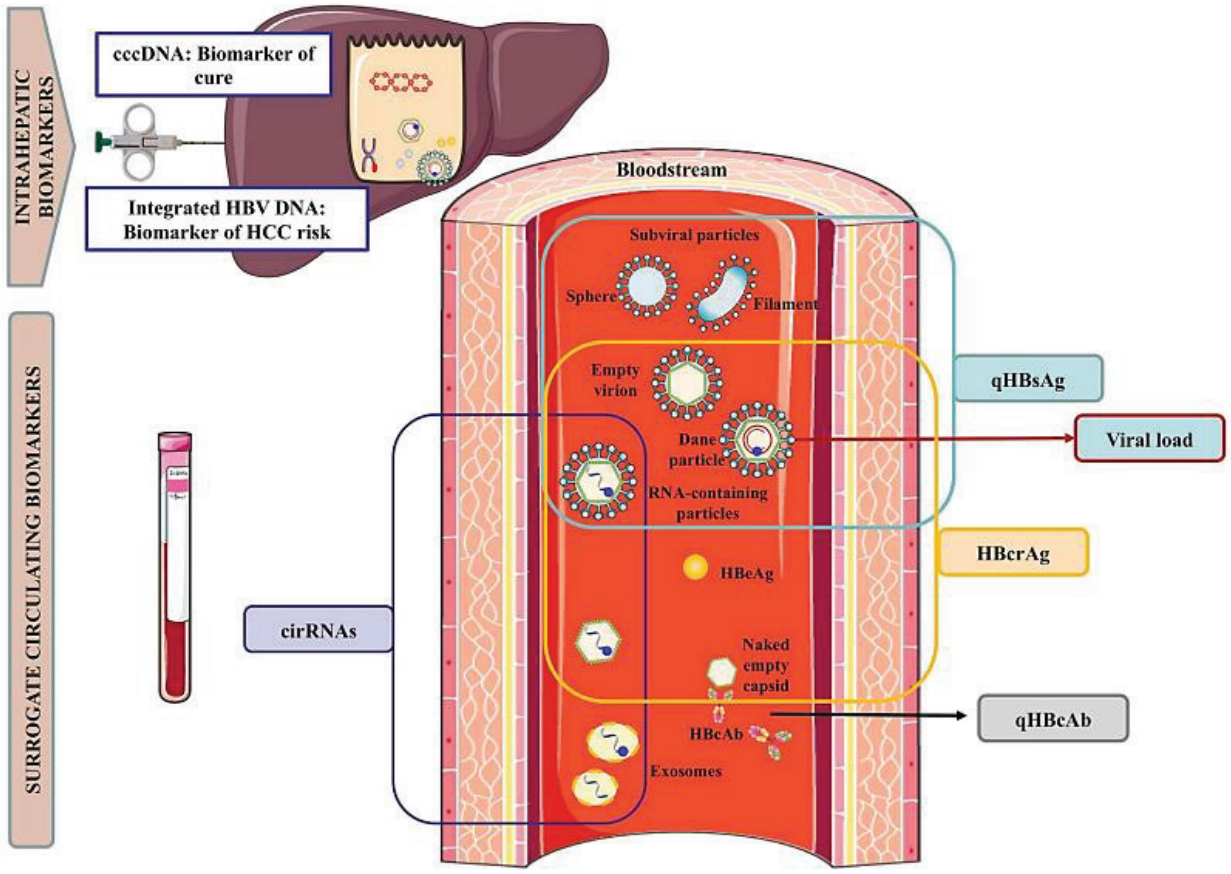


Figure 20 - HBV surrogate biomarkers
from **(Charre et al., 2019)**

3.2 New Biomarker candidates

3.2.1 HBcrAg

Hepatitis B core related antigen or HBcrAg is a composite biomarker that measures HBeAg, HBcAg, and the pre-core p22 protein levels, all derived from the HBV precore and core and sharing 149 amino acids. Sample pretreatment with a detergent allows the release of all core-related antigens from immune complexes or enveloped particles (empty virions, Dane particles, RNA-containing virions) and naked capsids. The immunoassay that has been developed to quantify HBcrAg uses a mixture of monoclonal antibodies generated in HBV core antigen-immunized mice and was first described in 2002 (Kimura et al., 2002). Since 2014 it is available as a fully automated quantitative assay (Lumipulse G1200 CLEIA from Fujirebio, Tokyo, Japan) (Charre et al., 2019).

Numerous studies have evaluated the potential of HBcrAg as a biomarker of HBV activity in the liver.

As expected, since HBeAg is one of the components detected by the assay (Maasoumy et al., 2015; Seto et al., 2014), higher serum HBcrAg levels were found in HBeAg+ patients than in HBeAg- patients. HBcrAg levels correlate with both viral replication and the level of intrahepatic cccDNA (Wang et al., 2019). Patients with HBcrAg lower than 1,000 U/mL (3 $\log_{10}$) also have lower amounts of cccDNA, a reduced transcriptional activity of cccDNA and a lower fibrosis and necro-inflammatory activity score (Testoni et al., 2019). Thus, HBcrAg might also help to distinguish anti-HBe chronic infection from anti-HBe chronic hepatitis. HBcrAg levels decrease in patients treated with NA or pegylated-interferons (Lam et al., 2017b; Ma et al., 2016).

HBcrAg assay has also important limitations related to the fact that: a) it detects multiple antigens; b) in HBeAg most of HBcrAg reactivity is due to HBeAg; c) it has a low sensitivity that need to be improved (Inoue and Tanaka, 2019). Only 30 to 50% of HBeAg-negative patients are currently identified as positive for HBcrAg (Testoni et al., 2019) and even more NUC treated patients are negative.

3.2.2 Circulating HBV RNAs

Circulating HBV RNAs have been already described in Chapter I, section 1.4.6.2. Their quantification has been proposed as a biomarker for disease monitoring, personalized therapeutic management, and the evaluation of new therapeutic molecules. Several studies have suggested that circulating HBV RNAs are mostly composed of 3.5 kb/pgRNA and might therefore reflect the transcriptional activity of cccDNA in the liver and be a good non-invasive biomarker for the inactivation of cccDNA or a reduction of the cccDNA reservoir.

Several studies have shown that circulating HBV RNAs may help monitor CHB infection (Lin et al., 2020; Liu et al., 2019; Luo et al., 2020). Circulating RNA levels have been evaluated in treatment naive patients in the different phases of the disease and the highest levels were found in HBeAg+ chronic infection followed by HBeAg+ chronic hepatitis, HBeAg- chronic hepatitis and the lowest level in HBeAg- chronic infection (Ghany et al., 2021; Wang et al., 2018). Similar results were obtained in HIV-HBV coinfected patients (Chung et al., 2021) and pregnant women (Patel et al., 2019).

In patients treated with NAs, serum RNAs can predict the emergence of YMDD mutants (Hatakeyama et al., 2007). HBV RNA quantification in the serum of patients treated with NAs or pegylated interferons predicts HBeAg seroconversion (Luo et al., 2019; Zhang et al., 2020). A rapid decrease in

circulating HBV RNA in chronic HBeAg- patients has been correlated with long-term loss of HBsAg (Farag et al., 2020; Limothai et al., 2019).

Circulating HBV RNA and HBcrAg have also been compared as biomarkers. A study on 127 HBeAg+ patients has shown that circulating HBV RNA was a better biomarker of HBe seroconversion (Ma et al., 2020). Both HBcrAg and circulating HBV RNA predict severe ALT flares and HBV DNA reactivation after treatment withdrawal (Carey et al., 2019). In patients with HBsAg loss after a long-term NUC suppressive therapy, the positivity of circulating HBV RNA was associated with viral reactivation (Carey et al., 2019). These results suggest that circulating HBV RNA could be a biomarker of choice. However, one study has shown that HBcrAg correlates better with cccDNA levels than HBV RNAs (Chen et al., 2019) and in another study loss of circulating HBV RNA would be achieved in many patients before the loss of HBcrAg (Liao et al., 2019).

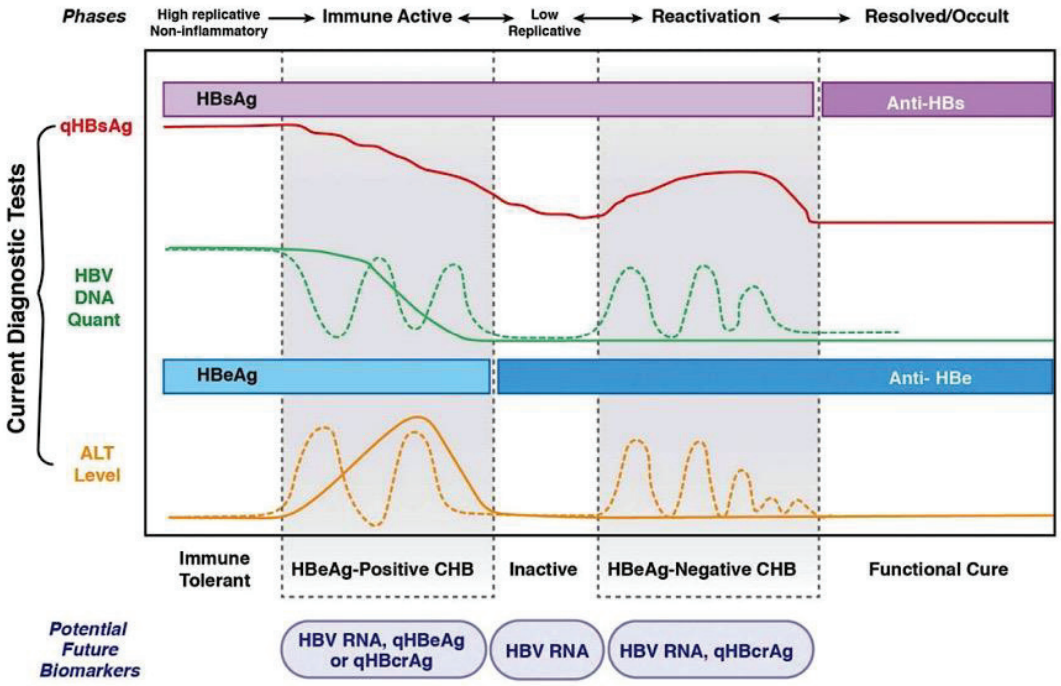


Figure 21 - Potential future biomarkers and their potential use according to the phase of HBV infection disease from (Coffin et al., 2019)

3.2.3 Circulating HBV RNA detection and quantification

A major limitation for the correct interpretation of the results obtained in the different studies and the adoption of circulating HBV RNA as a biomarker is the lack of standardized protocols (e.g., how to extract the RNAs from serum/plasma, how to retro-transcribe them, and finally how to choose primers for amplification: single pairs or multiple pairs, genotype inclusivity) and of aa recognized standard to calibrate their quantification (see below).

Different strategies have been used: primer pairs at the 3' end (X region) would amplify all HBV transcripts, irrespective of whether they are derived from cccDNA transcription or integrated HBV sequences; primer pairs targeting the canonical polyA signal site would detect all transcripts derived from cccDNA whereas primer pairs targeting the cryptic polyA signal site would theoretically mainly detect transcripts derived from HBV integrations; primers targeting the 5' end should amplify the 3.5 kb transcripts as well as pgRNA spliced transcripts (Charre et al., 2019).

Circulating HBV RNA quantification using RT-qPCR-based quantitative methods have shown a lower limit of quantification (LLOQ) of 10^3 copies/mL (Charre et al., 2019). The use of digital PCR allows the reduction of this threshold down to 100 copies/mL (Limothai et al., 2020).

Several companies are currently developing investigational "research only" kits to quantify circulating HBV RNAs (Otsuka and Koike, 2020). The Abbott m2000 RNA assay, that utilizes 2 primer pairs in the 5' and in the 3', claims a LOD of 1.65 logU/mL (1U HBV RNA = 1 U HBV DNA) (Abbott Laboratories, Abbott Park, IL) (Butler et al., 2018). The Roche cobas® HBV RNA investigational assay targets the canonical polyA signal and displays a highly sensitive and reproducible measurement of circulating HBV RNA with a LOD less than 5 copies/mL (Scholtès et al. 2022). A third commercial assay is proposed by Rendu Biotechnology HBV-SAT assay (Rendu Biotechnology, Shanghai, China).

The amount of circulating HBV RNA in treatment naive patients is usually 1-2 $\log_{10}$ copies/mL lower than the HBV DNA levels (Butler et al., 2018) (Scholtès et al. 2022). Circulating HBV RNAs decrease rapidly in patients treated with CAMs and siRNAs or ASOs reflecting, due to the mode of action of these classes of drugs a target engagement rather than the reduction of the cccDNA pool size and activity. In patients treated with NUCs circulating HBV RNA levels remain initially stable or there might be even an accumulation, followed over time by a slow decrease. This has been mechanistically explained by an increase in cccDNA transcription following the decrease in viral load or by the

accumulation of immature capsids containing pgRNA that cannot be retrotranscribed to DNA and are secreted as such (Lam et al., 2017a).

Although the potential of circulating RNAs as biomarker no longer seems to be open for discussion, many aspects remain to be confirmed or better defined. These include the exact nature of these circulating RNAs, their compartmentalization (extracellular vescicles/exosomes; virions, naked capsids, protein-bound or free), the proportion of circulating RNAs present in the different compartments in the different phases of the disease and the impact of HBV RNAs transcribed from HBV integrated sequences.

3.2.4 World Health Organization International Standard for hepatitis B virus Nucleic Acid

World Health Organization (WHO) International Standards (IS) for HBV DNA has already been established. The use of circulating HBV RNA as a biomarker would greatly benefit from the use of a specific RNA standard for its quantification. As a guide for the future establishment of an HBV RNA standard, this chapter describes how the HBV DNA standard has been developed and validated by the scientific community.

The HBV DNA IS is routinely used in nucleic acid amplification technique-based assays for HBV DNA quantification. Four standards have been validated over time since the standard material is limited in quantity. All standards were derived from the same plasma donor (genotype A2, HBsAg subtype adw2). This plasma was stored by the Eurohep Pathobiology Group as R1 and quantified at 2.7 x 10^9 HBV DNA molecules /mL (Heermann et al., 1999). In 2016, the IS was sequenced and the sequence is available under the GenBank accession number KY003230 (Jenkins et al., 2017).

The UK National Institute for Biological Standards and Control (NIBSC) plays a major role in assuring the quality of biological medicines through the development of standards and reference materials. NIBSC is a division of the Medicines and Healthcare products Regulatory Agency and part of the UK Department of Health and Social Care (DHSC). After validation by the WHO expert committee, the NIBSC offers online the latest validated standards and the corresponding technical data sheet.

Concerning HBV, it is possible to buy online the 4th WHO International Standard for HBV DNA for 132£ within the limit of 5 standards aliquots per customer and per year.

The first HBV DNA standard was validated by a WHO collaborative study group in 1999 (Saldanha et al., 2001). It is referenced in articles under the code of 97/746 or AA. The original material was diluted 1:500 with pooled plasma negative for HIV RNA, HCV RNA, HBV DNA and parvovirus B19 DNA. It was then aliquoted in 2000 vials of 0.5 mL and 4 vials were sent to each laboratory participating in the study. The concentration of the standard vials was established as 1 x 10^6 IU/mL according to the quantification performed by 22 laboratories from 9 different countries. Based on the non diluted quantification of 2.7 x 10^9 genome equivalent /mL, it can be estimated that 1 IU is equivalent to 5.4 genome equivalent. Notably, although the theoretical concentration of this standard should be 5.4 x 10^6 genome equivalent /mL the titer observed in the study is 2.0-2.6 x 10^6 genome equivalent /mL. This difference has been explained by a loss of material during lyophilization or extraction (some laboratories performed extractions without proteinase K digestion). Finally, the long-term stability under different storage temperatures was studied and the IS is stable at +4°C and +20°C for at least 18 weeks and at 37°C for at least 8 weeks.

The second standard was validated by the WHO collaborative study group in 2006, and it is referenced as 97/750 or BB. It was processed in the same way as the first standard. Following quantification by 6 laboratories, the vial concentration of the standard was established at 1 x 10^6 IU/mL. The long-term storage was also evaluated and it was found that no degradation was observed after 51 weeks stored at -20°C or +4°C (Baylis et al., 2008).

The third standard was validated by the WHO collaborative study group in 2011 and is coded by the NIBSC as 10/264. The quantification of this third standard is 8.5 x 10^5 IU/mL (Fryer et al., 2017).

The fourth standard was validated in 2016 by the WHO expert committee on biological standardization and the corresponding publication is under preparation. Its NIBSC code is 10/266 and it was assigned a potency of is 9.55 x 10^5 IU/mL (Organization and Standardization, 2016).

Very likely a 5th study will be organized soon to replace the 4th IS due to the limited amount of IS material.

All this highlights the key points for the development, synthesis and establishment of a WHO international standard for HBV RNA. Essential characteristics are:

- stability over time

- to be produced in sufficient amount to last for several years as a reference standard. Degradation can be prevented by freeze-dried.
- validation of its quality and its quantification by a large number of laboratories.

CHAPTER IV: HBV INTEGRATION

4.1 Integration of HBV viral genome

4.1.1 Double stranded linear DNA

The dslDNA, discovered in 1991, is a minor product of viral replication (Loeb et al., 1991; Staprans et al., 1991). Figure 22 summarizes how dslDNA is generated. The HBV polymerase has three basic functions during replication: primer synthesis, DNA synthesis from the RNA and DNA strands, and RNA degradation through its RNaseH function. After binding to the ε loop of the pgRNA, the HBV polymerase synthesizes 3 nucleotides (GAA) using the epsilon loop as a template. These nucleotides serve as a trinucleotide primer and are covalently attached to the polymerase. When the polymerase translocates to the direct repeat 1 (DR1) region in the 3' of the pgRNA it brings the primer with it. The polymerase then synthesizes viral (-) DNA strand from the pgRNA using the oligonucleotide primer. At the same time, the polymerase degrades the pgRNA template using its RNaseH function, but it leaves 11-18 nt of the matrix. Since the polymerase is covalently attached to the newly synthesized negative DNA strand, a nucleic loop is formed. In 90% of cases, the translocation of the 18-nucleotide primer to the DR2 sequence takes place, leading to rcDNA synthesis (Lee et al., 2004). However, in 10% of the cases the RNA primer is not translocated, and reverse transcription takes

place from the DR1 sequence leading to the synthesis of dslDNA. Subsequently, these dslDNAs can be enveloped and secreted like virions (Figure 22).

In cellular models, 30% of virions contain dslDNA (Tu et al., 2018). In patients, dslDNA represents between 3 and 26% of total DNA with a median of 7.24%. These proportions seem to increase with the progression of the disease and dslDNA proportion in serum is 14% and 20% in patients with cirrhosis and HCC. Serum dslDNA levels are also increased in patients treated with IFN-α (Zhao et al., 2016b).

dslDNA, like rcDNA, can either be secreted or recycled to the nucleus where it will be converted into cccDNA (Yang and Summers, 1995). However, this cccDNA is slightly different from the cccDNA originating from rcDNA since it contains a 16nt redundancy. This leads to a cccDNA that is defective for pgRNA and Precore mRNAs, the synthesis of rcDNA and the expression of viral polymerase and capsid protein. One study has estimated that 7 to 20% of the cccDNA pool was derived from dslDNA. The formation of cccDNA from dslDNA is thought to occur through nonhomologous recombination mechanisms involving Ku80 (Guo et al., 2012; Yang et al., 1996). dslDNA can also integrate into double-strand breaks into the host cell DNA, probably by nonhomologous end joining (Bill and Summers, 2004). A second, larger form of dslDNA has been described (Figure 23) (Yang and Summers, 1995; Yang et al., 1996). This longer dslDNA is believed to be derived from rcDNA by strand displacement DNA synthesis through the cohesive overlap region of the rcDNA. The longer dslDNA has a large terminal redundancy encompassing the cohesive overlap region (Mason et al., 2021). In chimpanzees, 75% of the dslDNA was the short form against 25% for the larger form (Mason et al., 2009).

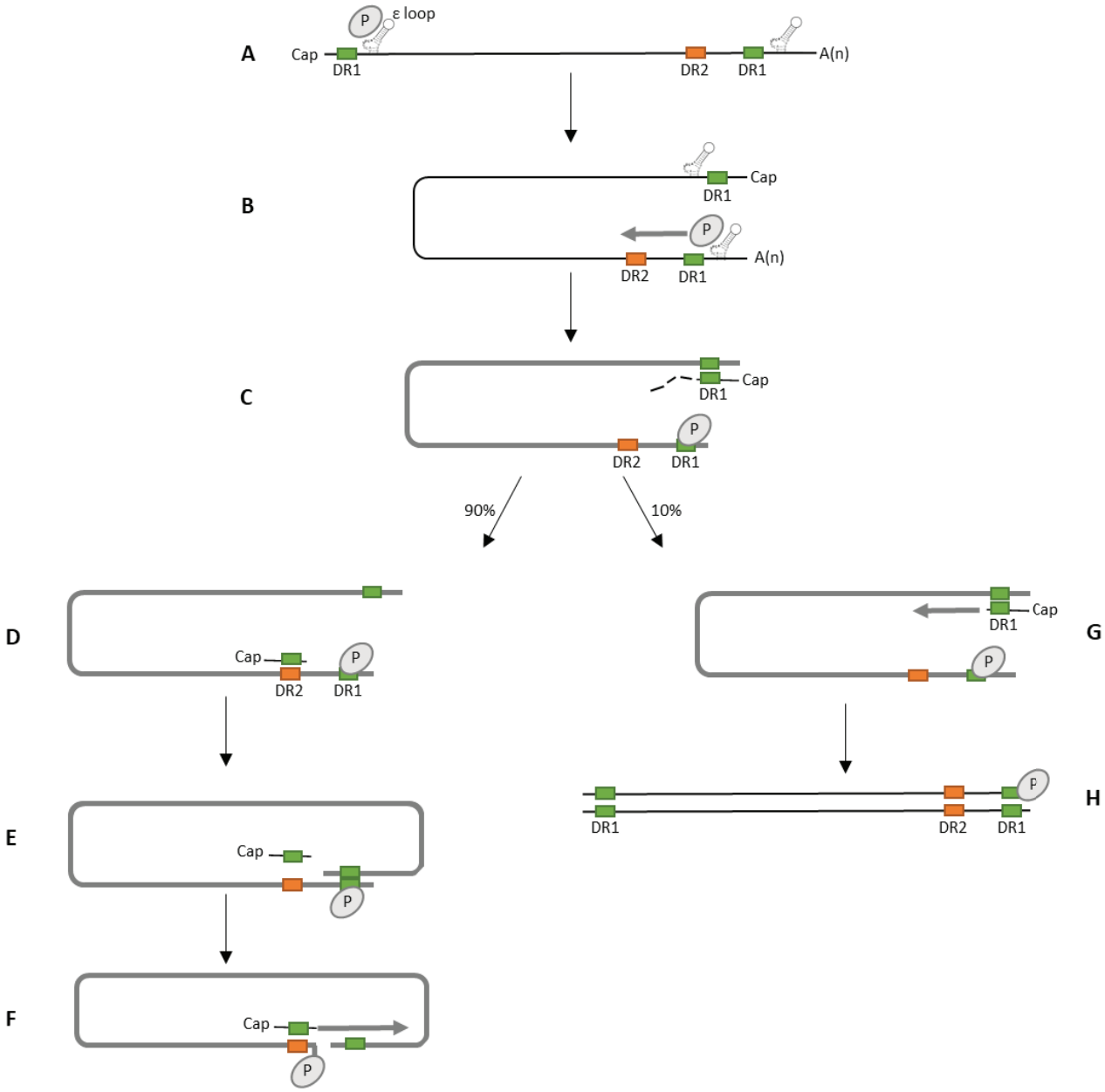


Figure 22 - Synthesis of rcDNA and dslDNA
Adapted from (Lee et al., 2004)
A. Binding of the polymerase to the ε loop of the pgRNA,
Synthesis of 3nt from the ε template
B. Polymerase switch from ε loop to DR1 and Reverse transcription: synthesis of the negative strand of the rcDNA
C. pgRNA degradation by the RNase H activity of the polymerase, 11-18nt of the pgRNA is not degraded

rcDNA SYNTHESIS
D. Translocation of the plus-strand DNA primer to the DR2 sequence
E. Circularization of the template by DR1 hybridisation and initiation of plus-strand DNA synthesis
F. Generation of rcDNA

DSL DNA SYNTHESIS
G. No plus-strand primer translocation and synthesis of the plus-strand DNA
H. Generation of dslDNA

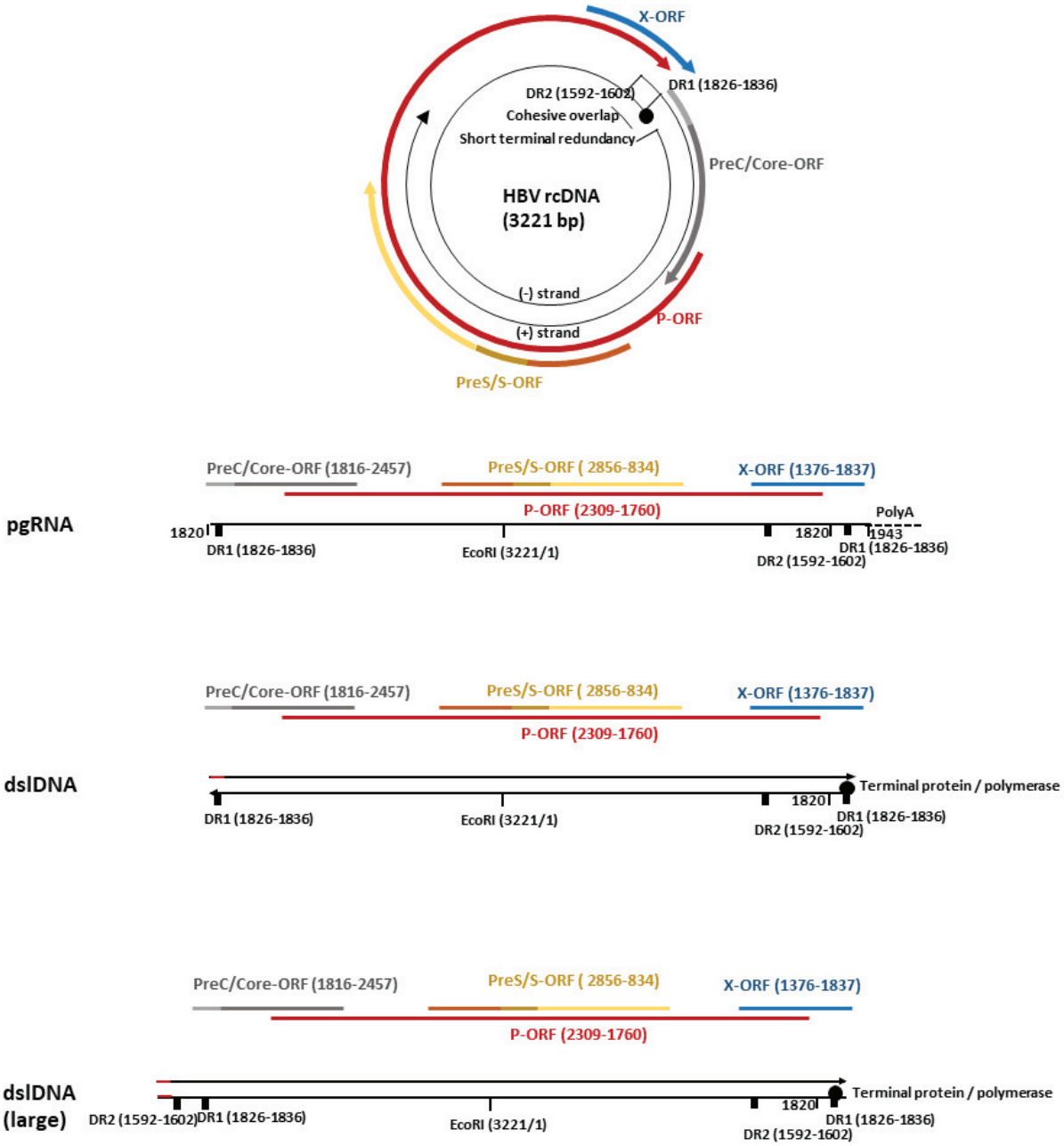


Figure 23 - HBV replication intermediates including the two HBV dslDNA forms from (Mason et al., 2021)

4.1.2 HBV DNA integration

HBV DNA integration is thought to come from the dslDNA and to occur preferentially at double-strand breaks in cellular DNA. Three studies support this hypothesis. After insertion of a single restriction site for the enzyme I-SceI in the genome of a chicken hepatoma cell line and the expression

of I-SceI from a co-transfected plasmid, an increase in the integration frequency at double-strand breaks was observed (Bill and Summers, 2004). The second study shows that after induction of DNA damage by H_20_2 treatment in HepG2.2.15, the frequency of HBV integration increases by 50%. Similarly, with an inhibitor of the Poly(ADP-ribosyl)ation polymerase that impedes DNA damage repair, the frequency of HBV integration also increases by 50% (Dandri et al., 2002). The third study confirms the increased number of HBV integrations in HepG2.2.15 cells after H_20_2 treatment using invPCR for HBV integration detection (Ruan et al., 2020).

The integration of HBV does not require viral proteins and is dependent on cellular repair pathways (Tu et al., 2021a). Double-strand breaks can be repaired by homologous recombination (gene conversion), usually involving strand invasion of a sister chromatid, or by error-prone mechanisms such as single-strand annealing and in particular non-homologous end joining (NHEJ) (Bill and Summers, 2004). NHEJ pathway repairing double-stranded DNA breaks is non-conservative and it does not restore the initial DNA sequence. Repair is done by ligation of both ends of the break without going through strand repair. The analysis of the integrated HBV fragments confirms the presence of deletions and insertions at the sites of end joining (Bill and Summers, 2004; Tu et al., 2021a). The relative contribution of canonical sequence-independent NHEJ and microhomology-mediated end joining (MMEJ), also called non classical NHEJ, to integration events is still unclear (Mladenov et al., 2016).

Integration can occur very early during HBV infection. In HBV-infected Huh7 cells viral integrations have been detected already 3 days after infection (Tu et al., 2018) and within few hours post-HBV-infection in HepaRG cells by invPCR and clonal sequencing (Chauhan et al., 2017). *In vivo*, integration has also been observed in patients with acute hepatitis (Kimbi et al., 2005) and in congenitally infected 5-month-old infants (Scotto et al., 1983).

The analysis of 5331 integration events collected in 7 published studies has shown that most HBV integrations occur between the viral Core/X gene and introns of the host genome (2789 vs 2420 in gene regions) (Lee et al., 2019). Two studies have shown that up to 40% of HBV integration breakpoints are found at nucleotides 1400-1900 of the viral genome (Lee et al., 2019; Sung et al., 2012; Zhao et al., 2016a). About a quarter of these intergenic integrations occurs in repeated elements and particularly in long interspersed nuclear element (LINE) repeats. Besides LINE, long

terminal repeat (LTR) and short interspersed nuclear elements (SINE) are also involved in HBV integration. 13 human genes are recurrently targeted by HBV. More frequently the telomerase (TERT), the histone H3K4 methyltransferase (MLL4 _ mixed-lineage leukaemia 4, also called KMT2B), Fibronectin 1 (FN1), albumin (ALB), the cyclin A2 (CCNA2), the cyclin E1 (CCNE1) and less frequently CPS1, ESRRG, RAI1, ATRNL1, ERBB4, GRID2, LRP1B, PRKDC, LINC00486, GLI2, PRKN, FAM157A, SOX5, GTF2I, PDE3A and CHRM3 (Bousali et al., 2021; Lee et al., 2019). These recurrent sites of HBV integration are found more frequently in tumor compared to non-tumor tissues. HBV integrations (SOX5, GTF2I, CHRM3, ALB, and CSP1) can also be found primarily in non-tumor tissues (You et al., 2019).

Many integrations have been found preferentially in liver tumor tissues. In non-tumor tissues integration occurs in a more dispersed manner and no preferential sites or similarities were found between patients (Sung et al., 2012; Zhao et al., 2016a).

Finally, entry inhibitors, but not interferon-alpha, polymerase inhibitors and capsid inhibitors, block integration. In one *in vitro* study, MyrB decreases the integration rate from 1 integration per 300 cells to 1 integration per 4,000 cells (Tu et al., 2018).

4.1.3 Expression of HBV DNA integrated sequences

Transcription from the dslDNA differs from cccDNA. First, RNAs transcribed from cccDNA mostly terminate at the canonical polyadenylation signal (poly A) located in the core ORF whereas RNAs transcribed from dslDNA mostly terminate at a non-canonical cryptic poly A signal due to the fact that in the dslDNA the canonical poly A signal is located at the 5’ end (Tu et al., 2021a). Second, the expression of all ORFs is altered except for the HBsAg ORF (Figure 24). The Pol, Core and PreS1/S2/S ORFs are intact but only PreS1/S2/S maintains regulation by its native promoter (Tu et al., 2017).
In addition to the missing HBV S promoter, and even if a host cell promoter is used, the transcripts started at the pgRNA TSS from the dslDNA and HBV integrations cannot be functional as the premature termination removes structural elements from the 3’end (notably the DR1 sequence) needed for pgRNA reverse transcription (Tu et al., 2021a). In addition, even in the absence of core

promoter, some transcripts containing part of the HBe/HBcAg ORFs have been detected in HCC derived cell lines (Lee et al., 1999). This might be explained by active cellular promoters upstream.

PLC/PRF/5 and Hep3B cell lines, that contain multiple HBV integrations express only the HBs antigens (Freitas et al., 2014), supporting the notion that only the PreS1/S2/S ORF is functional. HDV virions assembled in PLC/PRF/5 were able to infect primary human hepatocytes while it was not the case for HDV virions assembled in Hep3B.

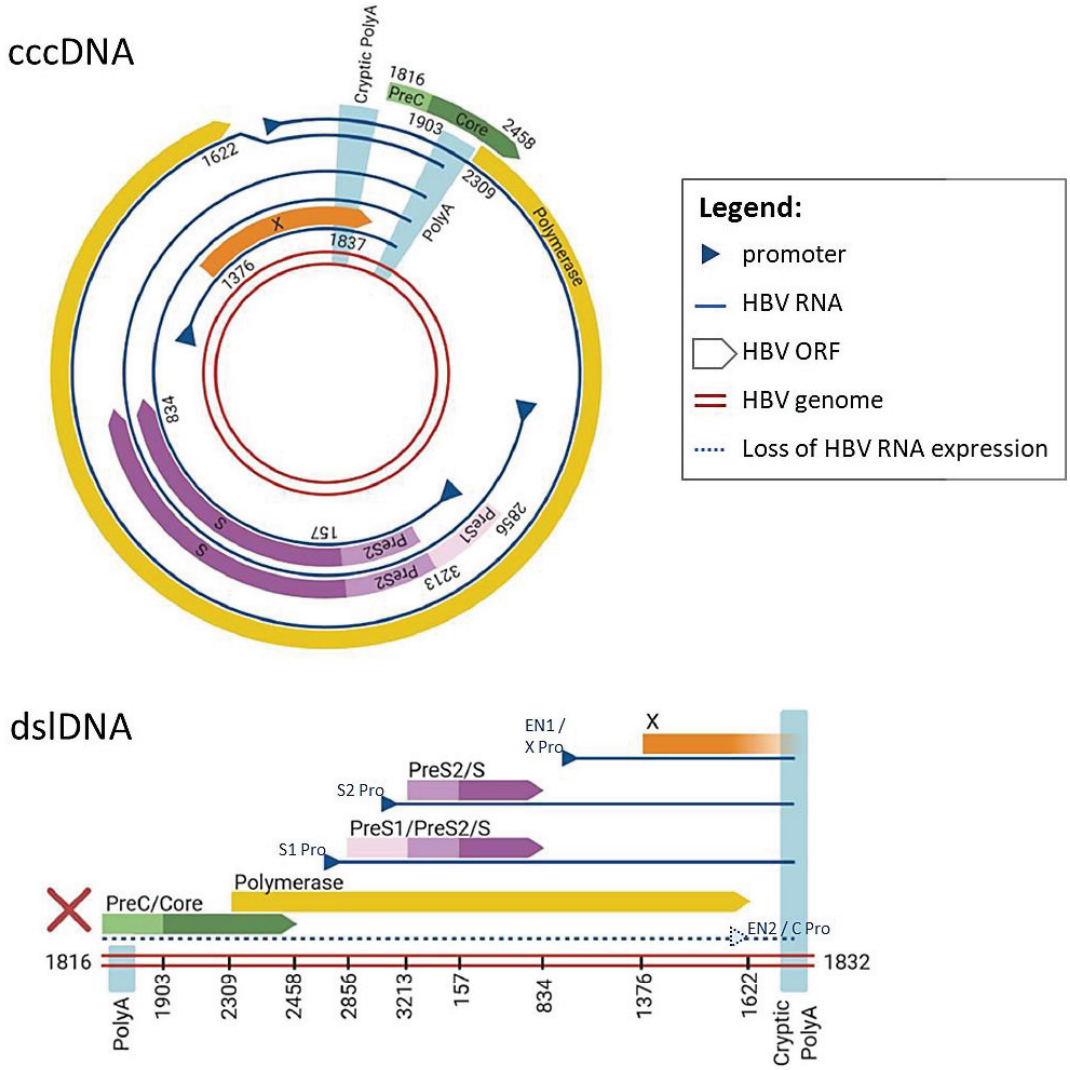


Figure 24 - cccDNA and dslDNA structures adapted from (Tu et al., 2021a)

Enhancer 1 appears to be active in the integrated form of HBV (Shamay et al., 2001) and HBx transcripts are also detected. However, they are truncated by at least 3 amino acids (Schlüter et al.,

1994). Several studies have shown that the HBx protein remains functional as a transcriptional transactivator even if truncated by up to 14 amino acids (Kumar et al., 1996). Notably, the truncated HBx proteins seems to be predominantly nuclear (Chami et al., 2000; Sirma et al., 1999). Moreover, the loss of the HBx stop codon in the integrated HBV sequences may lead to the production of chimeric HBx protein from the fusion transcripts between the HBV and cellular genome.

4.1.4 The role of integration in HBV-associated HCC

Integrated HBV sequences are found in 85-90% of tissue samples from patients with HBV-related HCC (Jiang et al., 2012; Murakami et al., 2005; Pollicino et al., 2011; Zhong et al., 2000). Following the integration of HBV into cellular chromosomes, two potential mechanisms can contribute to hepatocarcinogenesis. First, a cis-mediated mechanism which corresponds to integrations modulating the expression of proximal cellular genes, particularly genes involved in cell replication, and cell death (Larsson et al., 2018). These genetic changes contribute to carcinogenesis by causing an increase in oncogenic or anti-apoptotic signals and ultimately to the selection of liver clones with a growth advantage (Figure 25) (Bonilla Guerrero and Roberts, 2005). And secondly, a trans-mediated mechanisms leading to the chronic expression of particular viral antigens (Toh et al., 2013).

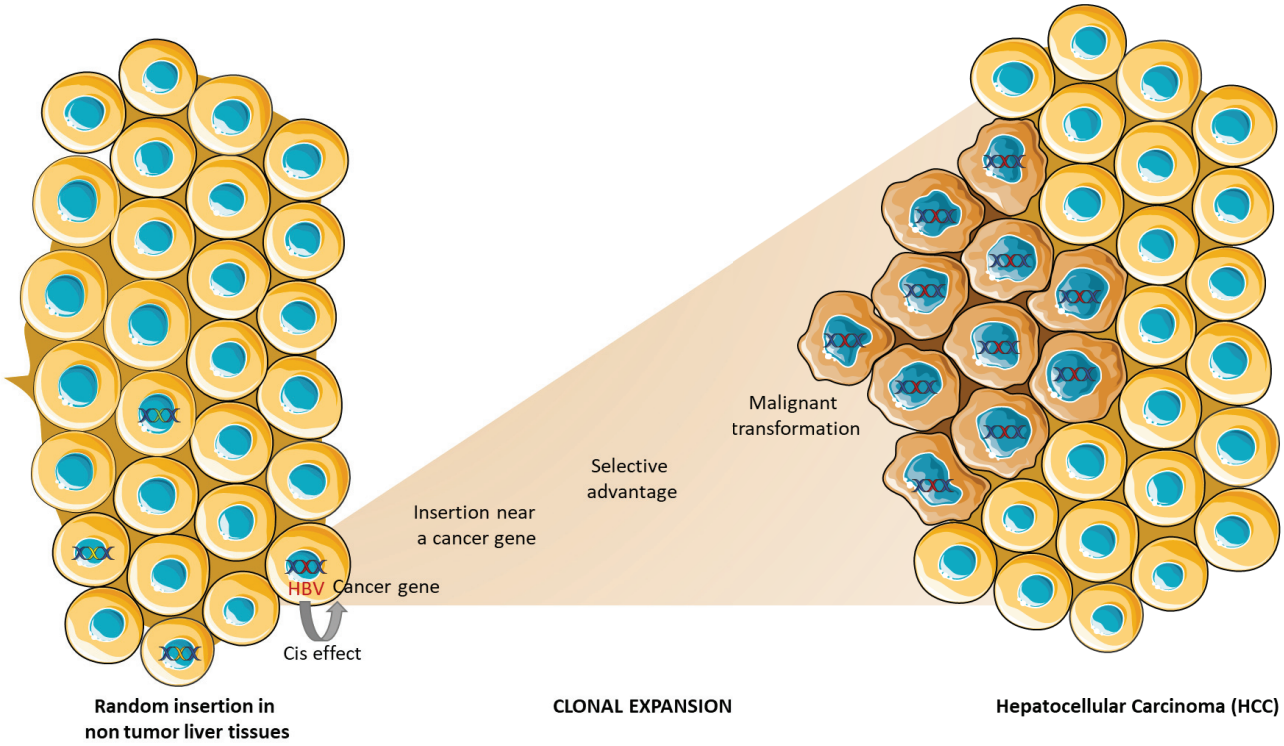


Figure 25 - HBV integration cis-mediated insertional mutagenesis can lead to carcinogenesis

More clonal HBV insertions (e.g., variable number of hepatocytes carrying the same HBV integrated sequences) were identified in tumors compared to the non-tumor liver tissue (Péneau et al., 2021). Concordantly, out of 3,320 integration sites detected by He et al., 2,602 were present in the tumor (1,397) or peri-tumor (1,205), and 718 in the normal liver tissue (He et al., 2020). Conversely, active HBV replication was more frequently detected in non-tumor tissues and associated with a higher number of non-clonal integrations (Péneau et al., 2021).

In the next two paragraphs the mechanisms induced by integration and involved in the development of HCC are described in detail. It is important to underline that a causal relationship between the integration event and HCC development is not established in all HCC cases. Integration is an early event in the natural history of HBV infection and continues throughout the chronic HBV infection whereas the emergence of HCC occurs only few decades after infection. This latency can be explained by the fact that integration is a random event and hepatocytes transformation needs a combination of factors, including chromosomal DNA damage resulting from inflammation, HBV DNA integration and hepatocyte death with compensatory proliferation induced by immune-mediated killing of infected cells.

4.1.4.1 Cis-mediated mechanisms

Studies comparing HCCs to non-tumor tissue suggest that a greater number and frequency of integrations into coding promoters and coding regions is found in the HCCs. Virus integrations occur randomly but when integration confers a growth advantage, it promotes clonal expansion leading ultimately to tumorgenesis. It is therefore coherent to find preferential sites of HBV integration in tumor and especially around genes regulating survival or proliferation, including TERT, MLL4, FN1, CCNA2 and CCNE1.

The TERT gene encodes the telomerase enzyme. Telomeres protect chromosomes from degradation. At each cell division the telomeres are reduced and when their size is critically low the hepatocyte enters apoptosis. The telomerase postpone apoptosis by adding at each cell division cycle a repeated DNA fragment at the end of the chromosome (Yuan et al., 2019). Alterations of the TERT gene by viral integration are the most frequent in all series. In one study, out of 95 patients with HCC, 34 had HBV integration into the TERT gene enhancer. This has been associated with greater TERT expression and a more aggressive form of HCC (Sze et al., 2020).

MLL4 is a methyltransferase that methylates the lysine 4 (K4) in H3 histone. The H3K4 methylation is correlated to gene activation and development (Crump and Milne, 2019). HBV integration near MLL4 gene leads to its overexpression (Dong et al., 2015; Sung et al., 2012).

Another gene frequently targeted by HBV integration is FN1. This gene encodes for two proteins: the soluble plasma fibronectin-1 and the insoluble cellular fibronectin-1 proteins. Hepatocytes produce and release in the bloodstream the soluble plasma fibronectin-1 that acts in blood cotting, wound healing and in extracellular matrix formation (Grinnell, 1984). FN1 deregulation is believed to promote an environment favorable to invasion and metastasis (Zhang et al., 2016).

HBV integration frequently impacts the cell cycle transition genes CCNA2 and CCNE1. Both proteins promote the S phase entry and progression. Specifically, cyclin A2 is involved in the G1/S and G2/M transition through cyclin-dependent kinase (CDK2). Cyclin E1 levels are high at G1/S and decrease across S phase transition. HBV DNA integration disrupts both genes. A chimeric HBV-CCNA2 transcript is also detected as a result of a substitution of 152 aa at the N-terminus of cyclin A by 156 aa of the MHBs protein (Pollicino et al., 2011). HBV-CCNA2 transcripts are detected in up to 12.5% of patients

with HCC. No HBV-CCNA2 transcript has been detected in cirrhotic patients without HCC. The HBV-CCNA2 chimeric protein promotes cell cycle progression and is associated to an aggressive HCC (Bayard et al., 2018; Chiu et al., 2016).

Rarer HBV DNA integrations can lead to overexpression of PDE3A in HCC tissues compare to non-tumoral tissue (Lu et al., 2018). This gene encodes a member of the cGMP-inhibited cyclic nucleotide phosphodiesterase (cGI-PDE) family and the protein hydrolyze both cAMP and cGMP, playing a critical role in many cellular processes. This protein has been associated with other cancers such as breast cancer, lung adenocarcinoma, cervical carcinoma, melanoma and HCC (Hao et al., 2020; Lu et al., 2018; de Waal et al., 2016).

HBV DNA integration can also target host transcription factor genes. The GLI2 gene encodes for a transcription factor that belongs to the Gli family of C2H2-type zinc finger proteins. GLI2 is up-regulated in HCC as well as in prostate and breast cancer and is associated with poor prognosis (Kim et al., 2007; Zhang et al., 2013).

Although in tumors many HBV integration events are found in genes, most integrations still take place in intergenic regions. These integrations can generate chimeric transcripts from the fusion between the viral genome and the human genome. The HBx-LINE1 RNA is a long non-coding RNA (lncRNA) found in HCC tumor tissues. It contains 6 binding sites for miR-122. As a result of this binding, it induces miR-122 depletion leading to an increased expression of the APOBEC2 gene and increased epithelio-mesenchymal transition (EMT) mediated by the activation of the Wnt/β-catenin signalling pathway and a reduction in E-cadherin, and finally an increase in cell migration (Li et al., 2019; Liang et al., 2016). However, the consequence of this alteration remains controversial. In one study, it was found in 23.3% of HCC and associated with a worse prognosis (Lau et al., 2014) while in another study it was not found in the 50 HCC patients studied with a different technology. Recently, a third study also reported the absence of this lncRNA in the 119 HCC tumors studied despite using similar techniques (Trung et al., 2019).

4.1.4.2 Trans-mediated mechanisms

HBV integration can allow the synthesis of HBsAg and truncated or chimeric HBx.
Integration can contribute to the accumulation of the HBs proteins in the ER leading to cellular stress or ER stress and the formation of ground glass hepatocytes. The detection of ground glass hepatocytes is correlated with an inflammatory activity favorable to pre-neoplastic lesions (Ringelhan and Protzer, 2015; Wu et al., 2014). Moreover, the potential pro-oncogenic activity of some truncated S proteins has been confirmed in transgenic mice: overexpression of envelope proteins led to liver cell adenoma after 8 months and to HCC after 20 months in 100% of mice (Dunsford et al., 1990).

Integration can also lead to the synthesis of the C-terminally truncated HBx protein. These proteins have been shown to confer stem cell-like properties, e.g. self-renewal, tumorigenicity and increased cell migration mediated by a STAT3/NANOG signalling pathway (Ching et al., 2017). In addition, truncated HBx proteins may play a role in metastasis development through the activation of c-Jun/AP-1 and through the association with the Myc-associated zinc finger protein (Li et al., 2016; Sze et al., 2013). Truncated HBx proteins have been also shown to induce chemoresistance, particularly to doxorubicin in Huh7 hepatocellular carcinoma cells (Jegal et al., 2019).
Finally, truncated HBx transcripts have been shown to down-regulate miR-15a/miR-16-1, that are tumor suppressor microRNAs in multiple human cancers (Wang et al., 2013).

4.1.5 Remaining questions about HBV integration

It is not yet clear which host cell DNA damage repair pathway(s) is/are involved in viral genome integration. These could be the classical NHEJ pathway, the MMEJ pathway or the single strand annealing (SSA) pathway. Knowing which cellular DNA repair pathways are involved in HBV integration would allow a better understanding of the cell characteristics that may favor the integration process. If cells containing specific dysfunctions of the DNA repair pathways are more susceptible to viral integration, their identification may have a predictive value and be exploited as a biomarker of HCC risk. The cellular proteins that promote dslDNA synthesis, allow DNA binding to

recruit DNA repair pathways, or alter chromosomal patterning and stability are not yet known. Similarly, the viral factors impacting the integration of dslDNA are not fully understood.

Although it is presumed that the integration comes from dslDNA, it is also possible that single-stranded DNA (ssDNA) and splice variants of HBV play a role in integration (Figure 26). Since these nucleic acids have the same terminal sequence as dslDNA and are present in the same form as dslDNA due to the lack of cis-acting circularization, they are potential substrates for viral integration in the hepatocyte genome (Tu et al., 2017).

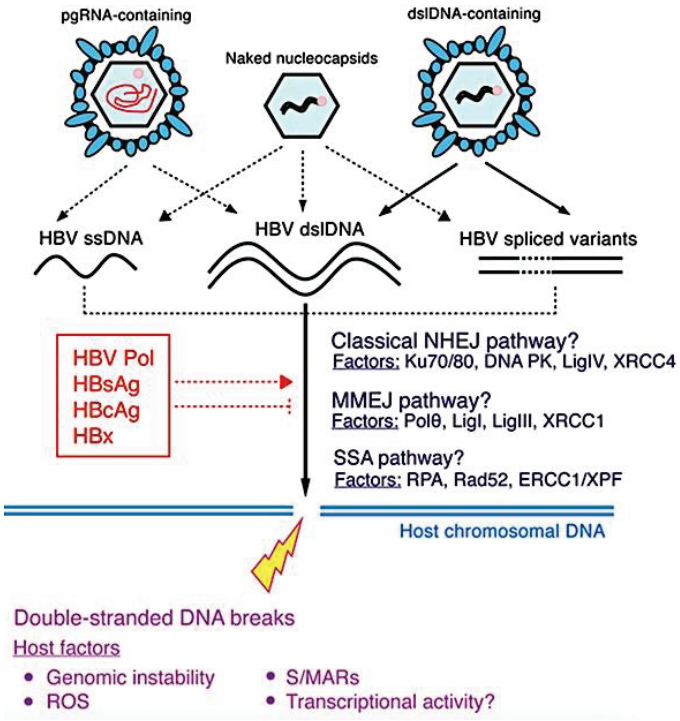


Figure 26 - Potential HBV integration mechanisms from (Tu et al., 2017)

AIM

As already mentioned, several new therapeutic molecules are at the preclinical or clinical evaluation stage with the aim of silencing or reducing the pool of cccDNA to achieve a *functional cure* with finite treatment duration.

Circulating RNAs from HBV have been shown to reflect the transcriptional activity of cccDNA in the liver and would thus be a good non-invasive biomarker for the inactivation of cccDNA or reduction of the cccDNA reservoir. The quantification of circulating HBV RNAs may contribute also to disease monitoring and therapeutic decisions in patients treated with the currently available drugs (NUCs and IFNs), in particular by guiding the decision to stop NUCs or identify patients who are likely to respond to IFNs.

In 2017, the French National Agency (ANR) (https://www.cirb-rna.fr/main/fr_FR/) funded the CirB-RNA RHU project that is conducted in Pr. Zoulim's and Pr. Levrero's teams, in collaboration with Roche diagnostic (Pleasanton, California) as industrial partner.

The goals of this project are the validation of circulating HBV RNAs as an innovative, clinically relevant, non-invasive biomarker for patients with chronic hepatitis B to predict a functional cure, and the development of the related diagnostic test.

My project thesis is part of the larger CirB-RNA project and focuses on 3 specific objectives:

1) **Generation of a molecular standard for circulating HBV RNA detection and quantification in CHB patients**

   The establishment of an HBV RNA standard remains a challenge and is crucial for the development of a circulating HBV RNA quantification assay. Although the current WHO HBV DNA standard may be an alternative for the quantification of serum HBV RNA, it has been challenged whether an HBV RNA unit is equivalent to an international HBV DNA unit. More importantly the WHO HBV DNA is derived from a clinical sample, which implies the presence of HBV RNA. A viral RNA standard, stable over time and produced in large quantities could be a good candidate as an international reference standard.

   The results of my efforts to generate a stable clonal cell line secreting HBV RNA and its characterization are presented in the submitted manuscript:

   Alexia Paturel, Francesca Casuscelli di Tocco, Delphine Bousquet, Marie-Laure Plissonnier, Xavier Grand, Hyoseon Tak, Françoise Berby, Caroline Scholtès, Barbara Testoni, Fabien Zoulim, Massimo Levrero. (2022). Generation of a molecular standard for circulating HBV RNA detection and quantification assays in CHB patients

2) **Generation of a capture high throughput viral DNA and RNA sequencing workflow for the study of HBV integration impact on circulating HBV RNA from CHB patients**

   Although the predictive nature of circulating RNAs no longer seems to be open for discussion, the impact of the integrated HBV into the human genome on circulating RNAs is not well defined.

   During my thesis, I have worked on the development of a robust a HBV capture-sequencing workflow for viral DNA and RNA sequencing using a multiplexed short reads NGS approach that allows DNA host-virus junctions and RNA virus-host fusion transcripts to be detected and quantified.

   The results of these efforts, that I shared with Dr Francesca Casuscelli di Tocco, are presented in the second section of the Results.

3) Characterization of circulating Hepatitis B virus RNAs *in vitro* and chronic hepatitis B patients

The use of circulating HBV RNA as a biomarker requires a deep knowledge on the exact nature of circulating RNAs (3.5 kb species, subgenomic viral RNAs), their compartmentalization (enveloped viral particles, naked capsids, subviral HBs particles, extracellular vesicles) and their distribution in the different compartment in the different stages of the disease. Multiple approaches have been used for this characterization, including density gradients, ddPCR, 5' RACE PCR and single molecule long read sequencing using the Nanopore technology.

Some of these results are presented as part of the manuscript on the generation of the HBV RNA standard and in the section on the detection of HBV integration and the characterization of their transcription.

I have also collaborated to a more sistematic approach to the characterization of circulating Hepatitis B virus RNAs *in vitro* and chronic hepatitis B patients whose results are presented in the manuscript in preparation:

Doohyun Kim, Delphine Bousquet, Marie-Laure Plissonnier, Hyoseon Tak, Xavier Grand, Chloé Goldsmith, Françoise Berby, Isabelle Bordes, Alexia Paturel, Aaron Hamilton, Marintha Heil, Massimo Levrero, Barbara Testoni and Fabien Zoulim. Circulating Hepatitis B Virus (HBV)-RNA associates with extracellular vesicles in supernatant of HBV-infected hepatocytes and serum of chronically infected patients.

MATERIALS AND METHODS

This section intends to provide background information on the cellular models and the different molecular biology and nucleic acids sequencing techniques that have been used in the 3 studies presented in the Results section.

1. Cell Lines

1.1 Huh7 cells

Huh7 cells have been used in this PhD thesis to derive polyclonal cell lines and individual clones carrying wild type and genetically modified versions of 1.1x HBV genomes.

Huh7 (JCRB0403) cell line was established in 1982 in Japan from a 57-year-old Japanese male with well differentiated hepatocellular carcinoma (Figure 27) (Nakabayashi et al., 1982). Huh7 cells have a human karyology and an heteroploid chromosome pattern with a mean modal of 60 (range of 55-63). Huh7 cells have an epithelial morphology, a doubling time of about 1.5-2.0 days (Kasai et al., 2018) and they have the ability to proliferate continuously with only Na2SeO3 supplemented medium without serum, although they are routinely cultivated in 10% FBS. Huh7 cells express produce the following human proteins: albumin, prealbumin, α1-antitrypsin, ceruloplasmin, fibrinogen, fibronectin, haptoglobin, hemopexin, β-lipoprotein, α2-macroglobulin, β2-microglobulin,

transferrin, complement components 3 and 4, and α1-fetoprotein. They also express a wide range of liver-specific transcription factors and genes (Nakabayashi et al., 1984, 1985).

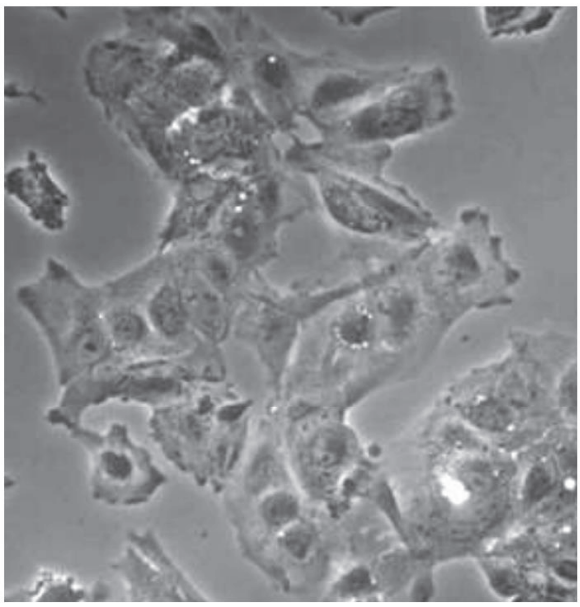


Figure 27 - Micrograph of Huh7 cells

Wild-type Huh7 cells cannot be infected by HBV inocula but Huh7 clones carrying human NTCP have been developed with this aim. Transfection of >1 length HBV genomes bypass the viral entry step and initiate HBV replication. Huh7 cells can be infected with HCV (Lohmann et al., 1999) and the Huh7.5 subclone, dysplaying higher susceptibility to infection, serves as an efficient HCV infection model. Injection of Huh7 cells subcutaneously into athymic mice leads to the development of tumors similar to those developed by the original patient. CD133, a stemness marker, is detected in a sub-population of Huh7 cells, suggesting that this cell line harbours a sub-population of cancer stem-like cells (Ma et al., 2008).

1.2 Cell Lines used for the study of HBV integration

In vitro model systems of HBV DNA integration include tumor-derived cell lines (PLC/PRF/5 and Hep3B), engineered HBV-producing cell lines (HepG2.2.15 and HepAD38), HBV transfection models (with ectopic introduction of viral DNA or RNA), HBV infection of cells expressing the viral receptor NTCP (Huh7-NTCP, HepG2-NTCP, HepaRG, HepaRG-NTCP, and Primary human hepatocytes (PHH) (*Table 7*).

Model system	Type	Production of infectious virus	Formation of new integration events	References
PLC/PRF/5	Tumor-derived cell line	No	No	(Alexander et al., 1976a; Dejean et al., 1983)
Hep3B	Tumor-derived cell line	No	No	(Aden et al., 1979; Knowles et al., 1980)
HepG2.2.15	Engineered HBV-producing cell line	Yes	Yes	(Dandri et al., 2002; Ruan et al., 2020)
HepAD38	Engineered HBV-producing cell line	Yes	Unknown	(Ladner et al., 1997)
Transfection of HBV over-length constructs	HBV transfection	Yes	Unknown	(Acs et al., 1987; Sells et al., 1988; Sureau et al., 1986, 1988; Will et al., 1982)
Transfection of monomeric HBV DNA	HBV transfection	Yes	Unknown	(Günther et al., 1998; Guo et al., 2016; Mutz et al., 2018)
Transfection of HBV virions DNA	HBV transfection	Yes	Yes	(Tu et al., 2019)
Transfection of *in vitro* transcribed pgRNA	HBV transfection	Yes	Unknown	(Yu et al., 2019)
Huh7-NTCP	HBV infection	Yes	Yes	(Tu et al., 2018, 2021b)
HepG2-NTCP	HBV infection	Yes	Yes	
HepaRG	HBV infection	Yes	Yes	
HepaRG-NTCP	HBV infection	Yes	Yes	
Primary Human Hepatocytes	HBV infection	Yes	Yes	

Table 7 - The different *in vitro* model systems used for HBV DNA integration study adapted from (Tu et al., 2021a)

The models used in a given study depend on several criteria such as ease of use, reproducibility, and similarity to the integrations found in HBV infected patients. In our study, we used PLC/PRF/5 and Hep3B, which are tumor-derived cell lines and therefore have integrations that were established in a native setting. The disadvantage of these cell lines is that they do not allow to study the mechanisms

underlying integration since they do not generate new integration events. To study the mechanisms of integration dynamically models based on either infection through NTCP receptors or direct insertion of HBV genome by transfection should be used. These models also have some drawbacks in that transfected cells generate a large number of new integrations but do not reflect native infection that leads to a low number of integration events. Infections mediated by the NTCP receptor are more like what happens in patients. However, the low number of integrations generated makes their detection difficult. A specific section is dedicated to the discussion of the "Techniques to detect viral integration".

1.2.1 PLC/PRF/5 (Alexander) cells

PLC/PRF/5, also known as Alexander, is a cell line derived in 1973 from the primary liver carcinoma of a Shangaan male aged of 24 years in Mozambique (Alexander et al., 1976a; Edman et al., 1980) (Figure 28*)*. Their morphology is similar to cultured hepatocytes and they have biosynthetic capabilities of normal liver parenchymal cells (Knowles et al., 1980). Plating efficiency is low (40-50%) and doubling time is 35-40 hours. At confluence, cells continue to divide, do not pile up but detach and die (Alexander et al., 1976b). They form tumors in the nude mice (Knowles et al., 1980).

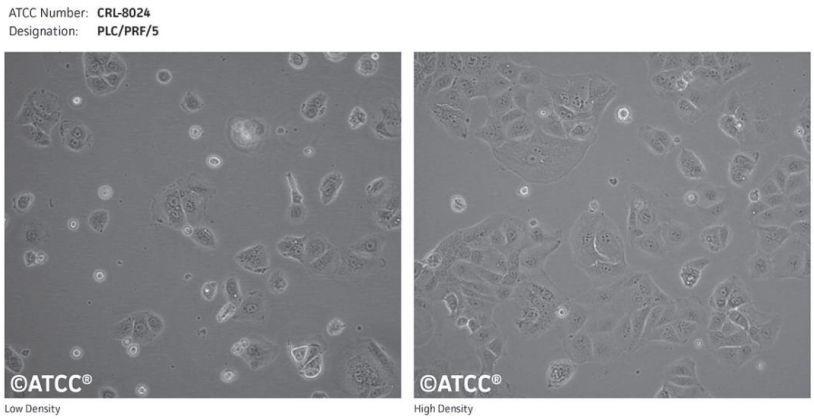


Figure 28 - Cell micrograph of PLC/PRF/5
from ATCC website (https://www.lgcstandards-atcc.org/en/Products/Cells_and_Microorganisms.aspx)

The PLC/PRF/5 have a human karyology. However, the chromosome pattern is heteroploid with a mean modal of 56 (range of 50-60). Chromosomes 8, 9, 10, 17 and 21 are monosomal while chromosomes 2, 6, 7 and 11 are trisomal. The number of chromosomes can go up to 4 in the case of chromosome 15. Despite these anomalies these cells are very stable: 10 years after their generation

and after 110 passages, PLC/PRF/5 they still present the same karyotype (Pinto et al., 1985). PLC/PRF/5 cells carry a p53 mutation at codon 249 with a substitution of a serine for an arginine that is linked to aflatoxin B1 exposure (Hsu et al., 1993).

A recent study, using HBV-capture followed by short reads next-generation sequencing, has reconstructed the HBV sequences integrated in the PLC/PRF/5 genome (NCBI reference LC533934). Thirteen virus/human junctions were found in 8 different chromosomes including chromosomes 3, 4, 5, 8, 11, 12, 13 and 17. The integrations present in chromosomes 3, 11 and 17 were subsequently sequenced by Sanger (Ishii et al., 2020). A second study used a 3rd generation long-reads sequencing technique and found 84 HBV fragments inserted in 7 chromosomes. Differently from the study by Ishii et al., an additional HBV integration was found in chromosome 16 whereas the integrations in chromosomes 5 and 11 were not detected (Chen et al., 2021). A third study coupled HBV-capture with 3rd generation long-reads sequencing and confirmed HBV integration in chromosomes 5, 11 and 16, found an additional HBV integration in chromosome 1 but failed to detect the HBV integration in chromosome 4 that was found in the first 2 studies (Ramirez et al., 2021). *Table 8* summarizes the HBV integrations found in the 3 most recent studies using high-throughput sequencing. Collectively, 10 integrations have been found but none of the 3 studies detected all the 10 integrations. The 3 studies used high-throughput sequencing but different sequencing technologies preceded or not by an oligonucleotide capture step.

	Chr 1	Chr 3	Chr 4	Chr 5	Chr 8	Chr 11	Chr 12	Chr 13	Chr 16	Chr 17
Ishii et al HBV capture + Short Read Sequencing	X	✓	✓	✓	✓	✓	✓	✓	X	✓
Chen et al Long Read Sequencing	X	✓	✓	X	✓	X	✓	✓	✓	✓
Ramirez et al HBV capture + Long Read Sequencing	✓	✓	X	✓	✓	✓	✓	✓	✓	✓

Table 8 – HBV is integrated in differents chromosomes of the human PLC/PRF/5 cell line.
Depending on workers and sequencing technologies HBV integration found in PLC/PRF/5 cell line vary

With exception of one study based on Southern Blot analysis, reporting the integration of a complete genome (Edman et al., 1980), all other studies have shown multiple partial HBV integrations and 2 chromosomal translocations (Brechot et al., 1980; Chakraborty et al., 1980; Edman et al., 1980; Marion et al., 1980; Ramirez et al., 2021; Shamay et al., 2001; Shaul et al., 1984; Ziemer et al., 1985) PLC/PRF/5 secrete HBsAg (Aden et al., 1979; Ishii et al., 2020; Monjardino and Crawford, 1979) with a rate estimated at 500 ng/day/10^6 cells (MacNab et al., 1976). PLC/PRF/5 envelope proteins support the formation and secretion of infectious HDV virions (Freitas et al., 2014).

Sequencing data indicate that the HBV fragments integrated in chromosomes 12, 13, 16 and 17 have the potential to express S transcripts (Chen et al., 2021). Chromosomes 11 can express both PreS1 and PreS2/S RNAs, while the HBV integration on chromosome 13 can express only the PreS2/S RNA. HBV sequences integrated in chromosomes 1, 8, and 3 are not transcriptionally active. The vast majority of transcripts in PLC/PRF/5 cells are chimeric and use either downstream cellular polyadenylation signal or the viral cryptical polyadenylation signal (Ramirez et al., 2021).

Complete virions, empty virions and core particles as well as HBs or HBe proteins have not been detected in PLC/PRF/5 or in the cell supernatants (Edman et al., 1980; Ishii et al., 2020; Malmström et al., 2012). Northern Blot results showed only HBs transcripts (Edman et al., 1980; Ou and Rutter, 1985). Whether HBx and the Polymerase are expressed in PLC/PRF/5 cells is less clear. Two studies have reported the presence of the HBx protein in PLC/PRF/5, based upon the results of HBx perturbation experiments (Chan and Ng, 2006; Li et al., 2014). The long read PacBio IsoSeq RNA sequencing performed by Chen et al. indicates that HBV integrations in chromosomes 13 and 16 have the potential to express a truncated HBx (Chen et al., 2021). Two studies detected a transcript containing sequences that are unique to the pgRNA and the paper from Ishii and coll. observed by immunofluorescence the expression of a protein recognized by anti-HBV Pol antibodies in Huh7 cells transfected with the integrated sequences obtained by HBV capture (Ishii et al., 2020; Malmström et al., 2012; Park et al., 2008). The most recent paper published using the HBV targeted PacBio technology for long-reads high-throughput DNA and RNA sequencing could not detect any pgRNA or HBx mRNA (Ramirez et al., 2021).

1.2.2 Hep3B cells

Hep3B cells, also referred to Hep3B2.1-7, has been established in 1976 from a US 8 years old black boy with HCC and lung metastases (Aden et al., 1979).

Hep3B cells have a human-like karyology with an average of 60 chromosomes per cell (Qiu et al., 2015), biosynthetic capabilities of liver parenchymal cells (Knowles et al., 1980), epithelial cell morphology (Figure 29) and a doubling time of 40-50 hours.

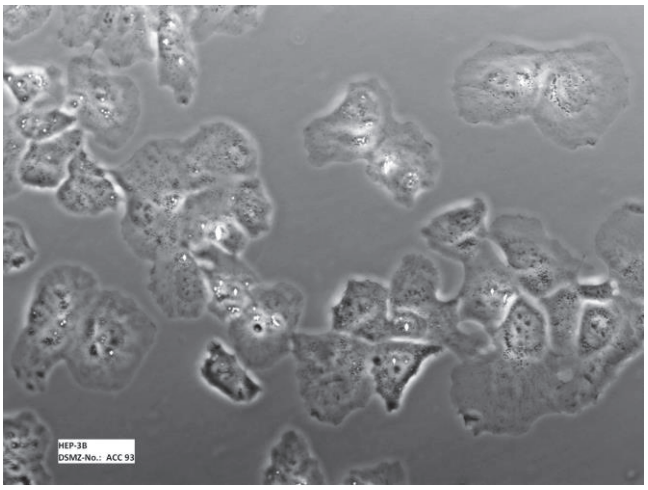


Figure 29 - Cell micrograph of Hep3B cell line from DSMZ (https://www.dsmz.de/fileadmin/downloads/ACC/cells_in_culture/HEP-3B_Bild_1.jpg)

After injection in athymic (nude) mice, aggressive tumors histologically similar to HCC can grow and develop metastases in the lungs (Knowles et al., 1980). Retinoblastoma protein expression is altered and p53 alleles are deleted in Hep3B cells (Puisieux et al., 1993).

Hep3B have two copies of genotype A HBV DNA integrated into the host genome (Twist et al., 1981). Sanger sequencing showed that HBV sequences in Hep3B are identical to the HBV genome with the GenBank accession number GQ477489.1 (Freitas et al., 2014). Northern blot analysis showed five HBV RNAs (4.0, 3.3, 2.9, 2.6 and 2.2 kb). Using HBV probes specific for different HBV regions, PreS/S and X regions were detected (Su et al., 1986). Hep3B cells secrete variable amounts of HBsAg (Aden et al., 1979). The PreS2/S ORF appears to be intact with an active promoter allowing the synthesis of surface antigens. Consequently, Hep3B are able to support HDV assembly and secretion of HDV virions but the HDV particles produced by Hep3B appeared to be non-infectious (Freitas et al., 2014). The latest study using the targeted PacBio long read DNA sequencing and the PacBio Iso-seq RNA sequencing, confirmed the presence of 2 HBV insertions linked with 2 translocations (chromosome 16 to chromosome Y and chromosome 4 to chromosome 13) with HBV insertions in the middle

(Ramirez et al., 2021). This study has also shown that HBV insertion on chromosome 16/Y is transcriptionally silent, and that the hybrid chromosome 4/13 allows the transcription of PreS2/S RNA (Ramirez et al., 2021). This result explains why hepatitis D virus particles produced in Hep3B can not be infectious due to the absence of PreS1 region, needed for the HBV entry into hepatocytes.
HBx sequence is truncated (Freitas et al., 2014; Ramirez et al., 2021; Su et al., 1998). Although enhancer I appears to be present and potentially active, PreC/pgRNA transcripts have not been detected and only one study has suggested that HBx is expressed, based upon indirect evidence from perturbation experiments (Li et al., 2014; Shamay et al., 2001).

2. Droplet digital PCR (ddPCR)

The concept of digital PCR was developed already in 1992 (Sykes et al., 1992) but the first publication using digital PCR dates back to 1999 and the method was not widely used until the first commercial instruments became available in 2007. Digital PCR is based on the same biochemistry as qPCR but differs in how the quantitation is performed. Digital PCR is based upon the division/dilution of the PCR mix into smaller reactions that run individually. The number of positive/amplified reactions are counted and, using Poisson's law, the exact number of templates can be calculated. Digital PCR can quantify low concentrations of DNA and is used for the detection of pathogens, gene mutations, gene copy number variation, DNA modifications and mRNA expression levels.

Several dilution strategies can be used: the mix can be divided into microwell plates, capillaries or water-in-oil droplet emulsions. The latest, is referred to as digital droplet PCR (ddPCR) and it has been commercialized by Bio-Rad®. Several digital PCR machines from Bio-Rad®, Life Technologies®, Fluidigm®, Formulatrix®, or Stilla® are widely used in hospitals and research laboratories. The QX200 from Biorad has steadily become the most widely used ddPCR platform across all research fields and in 2017, 67% of total number published articles mentioning a digital PCR were using ddPCR (Rutsaert et al., 2018). The technology used in the QX200 was spun out originally by QuantaLife, a startup created by Lawrence Livermore Lab that was subsequently acquired by Biorad.
Figure 30 describes the ddPCR workflow: after PCR reaction mix preparation a large number of droplets (20,000 in the case of the QX200) are generated by manual partitioning, immiscible liquid chemistry, or atomization. The ideal droplet contains a single target DNA molecule or none at all.

Afterwards, the partitions are amplified individually through thermal cycling. Finally, positive partitions with the target sequence and negative partitions without the target sequence are detected and their number is calculated using the Poisson's Law in order to correct for the possibility that multiple copies of target DNA could have been included in the same droplets. Positive droplets have a higher fluorescence than negative droplets, but the presence of negative droplets is important for the proper quantification of samples. It is recommended to have at least 100 negative droplets in the reaction to allow a correct application of the Poisson's law and to be able to generate a robust threshold between negative and positive droplets. The analysis of the results can be done either in one dimension, where each droplet is plotted on a graph showing the fluorescence intensity as a function of the number of total droplets generated, or in two dimensions. This type of analysis is performed when several targets are amplified in the same sample (duplex quantification) and droplets are plotted on a graph showing the fluorescence intensity of channel 1 (FAM) versus the fluorescence intensity of channel 2 (VIC).

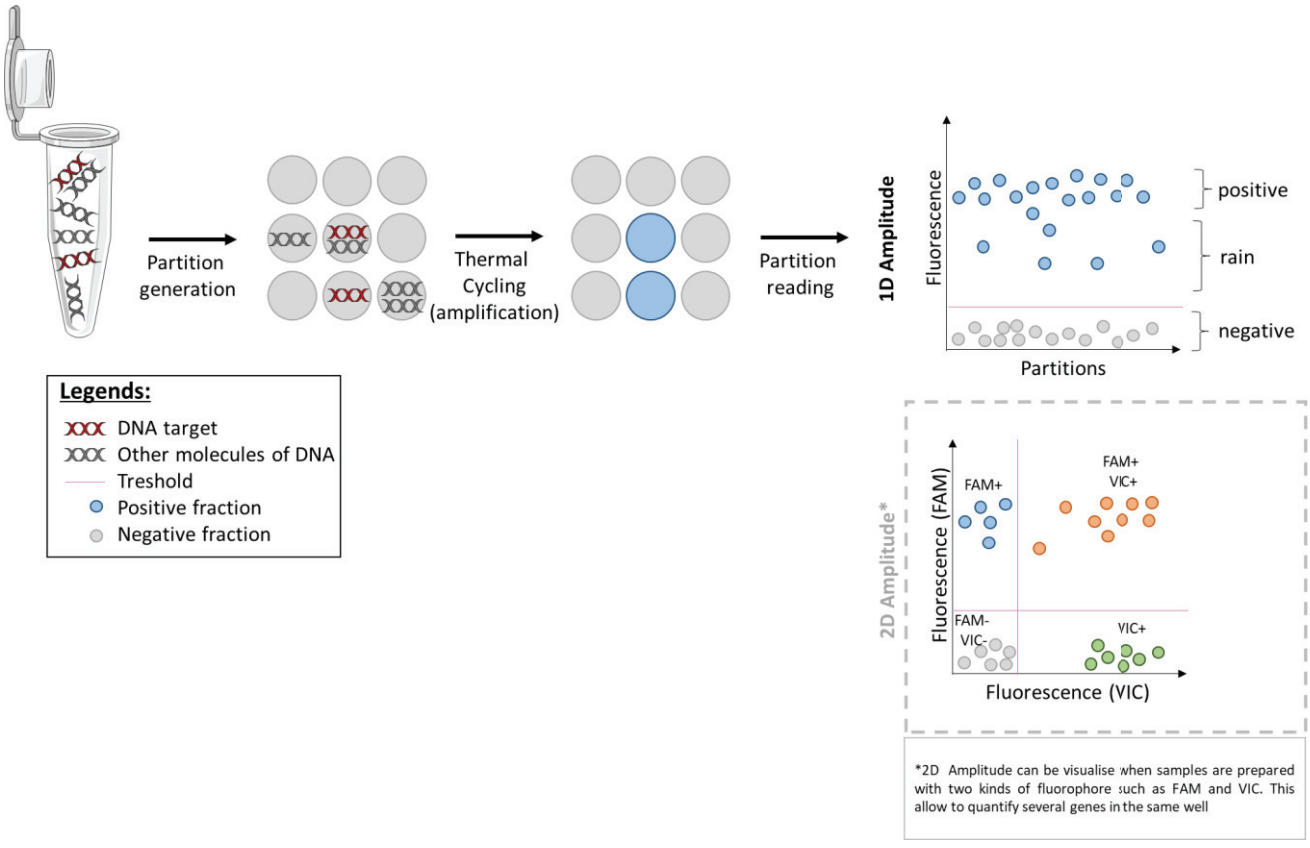


Figure 30 - Digital PCR concept adpted from (Rutsaert et al., 2018)

ddPCR is a valuable technique with several advantages. First, ddPCR is more sensitive than qPCR. Second, unlike qPCR that requires a standard curve to establish the sample concentration, ddPCR is an absolute quantification method which does not require any standard. Third, ddPCR allows the detection of rare targets by increasing their abundance within an individual compartment. Lastly, ddPCR tolerates better PCR inhibitory substances such as heparin, SDS and other common inhibitors (Hoshino and Inagaki, 2012; Sedlak et al., 2014).

ddPCR also has some limitations. In comparison to qPCR, ddPCR is longer to perform as it requires a pre-PCR partitioning and a post-PCR reading steps. The quantification of a 96 samples plate takes 5 hours with the QX200. In addition, without including the cost of the instruments, ddPCR is slightly more expensive in terms of reagents per sample. On average, it costs $5 per sample, whereas qPCR costs $2 per sample. This cost can be decreased since it is not necessary to load a standard and costs are expected to come down over time. The final drawback of ddPCR is that the droplet containing the amplicons are lost and cannot be used for an eventual future sequencing. It is therefore essential to prepare duplicate samples: one for quantification and the second for amplicon sequencing. Moreover, before using the amplicons generated by ddPCR it is necessary to break the droplet emulsion (Basu, 2017).

False-positive droplets can be generated by ddPCR but with no more than 0.16-0.22 copies/well and can be easily eliminated. They can arise from contaminations or, more frequently, disturbed droplets that merge together leading to a droplet with a higher baseline fluorescence that is miscalled as positive (Kiselinova et al., 2014).

ddPCR has been used for HBV research, particularly for cccDNA and 3.5 kb / pgRNA quantification. ddPCR improved the limit detection of cccDNA and HBV DNA requiring only 0.54 to 0.59 copies for accurate as compared to at least 10^2 copies of cccDNA or HBV DNA with the conventional quantitative PCR (Li et al., 2018b). ddPCR has also been shown to be 10-100 fold more sensitive than qPCR for the detection of intrahepatic HBV cccDNA with a lower limit of detection (LLOD) and quantitation (LLOQ) estimated at 0.8 copies/10^5 cells and 3.8 copies/10^5cells, respectively (Caviglia et al., 2018). The sensitivity of RT-ddPCR is also increased compared to the sensitivity of RT-qPCR (10^2 vs. 10^3 copies/mL) (Limothai et al., 2020).

3. Oxford Nanopore sequencing

Nanopore sequencing is a new technology for single molecule, long reads nucleic acids (DNA/RNA) sequencing. There are two kinds of nanopores: biological and solid-state nanopores. In 2005 Hagan Bayley and Gordon Sanghera developed at Oxford Nanopore Technologies (ONT) a sequencing workflow based upon biological nanopores to sequence DNA. Their first device commercialized in 2014 was the MinION (Figure 31) with a capacity to generate 50 Gb per run. The access to this new technology was quite affordable, with an initial cost of $1000 including the MinION device and a set of reagents. In 2017 ONT introduced the compact desktop sequencer GridION that combines up to 5 MinION flow cells and generates up to 250 Gb of data. In 2019, the PromethION sequencer, that can use up to 48 PromethION flow cells simultaneously with a capacity of 290 Gb per flow cell, was made available for large-scale projects requiring ultra-high throughput. Promethion and Gridion are less accessible financially with costs of $49,955 and $195,455 respectively and less accessible in terms of size and portability. As far as portability is concerned, ONT is developing the SmidgION sequencer that will allow to use the same nanopore technology but with smartphones or other low power mobile devices.

3.1 The MinION device

Unlike the more common Illumina technology (see paragraph 4.3.2 Illumina sequencing), ONT sequencers do not rely on a sequencing by synthesis chemistry. Nanopores of 1.8nm in length are embedded in a lipid membrane. The MinION flow cell has 2,048 individual protein nanopore and the software uses 512 of them at the time. The lipid membranes in which the nanopores are embedded act as an electric resistant barrier to prevent the current from going back and forward. Current passage through the pore is unidirectional. Single strand DNA or RNA are pulled on through the pore and depending on which base is passing the change in the current is measured and a base calling is performed.

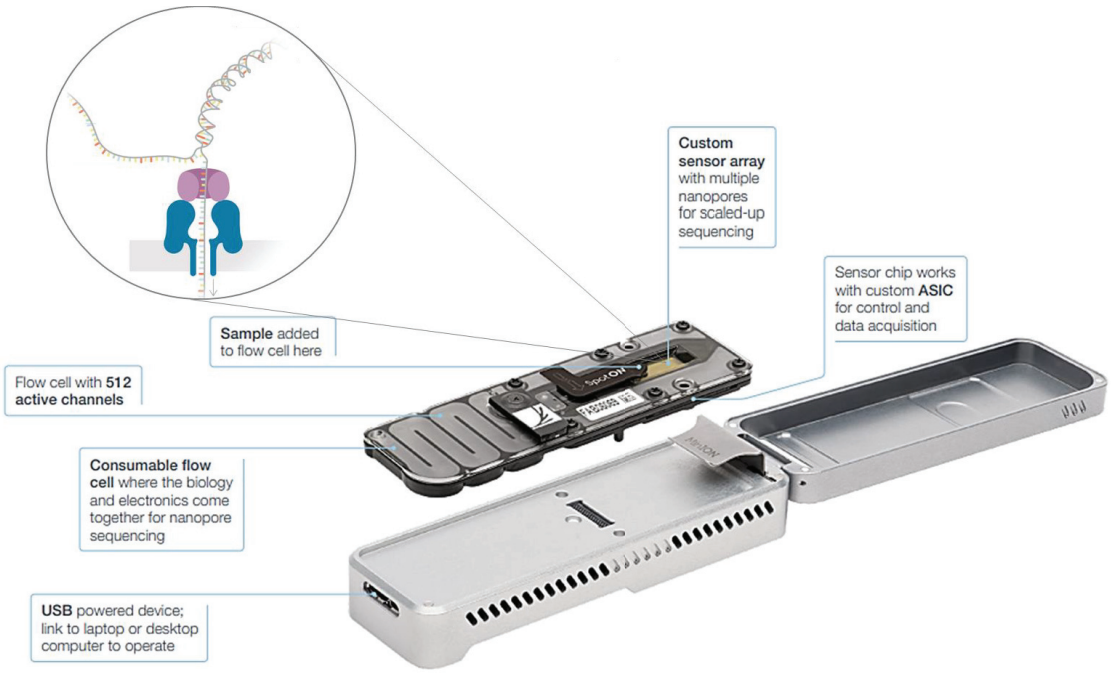


Figure 31 - The MinION sequencing device
From (Lu et al., 2016)

The technology developed by ONT has several advantages. First, it allows the single molecule sequencing of long reads with a maximum length up to a few hundred thousand base pairs. The longest reads sequenced to date reached almost 1 megabase (882kb) (Jain et al., 2018). The read length is limited only by the length of the nucleic acid molecule presented to the ONT device and not by any instrument limitation *per se* (Laver et al., 2015). Long read sequencing allows to easily assemble genomes *de novo* and facilitates the sequencing of regions with long repetitive elements. The nanopore sequencing also allows to observe base modifications such as methylation (Goldsmith et al., 2021). Direct RNA and protein sequencing protocols are under development. Since the technology does not rely on PCR amplification during sequencing it avoids all PCR associated bias. Moreover, it does not require an imaging equipment allowing to scale down the size to get the most portable sequencer available. A MinION device measures only 10 x 3 x 2 cm and weights less than 100g. The informatics requirements are also low: a Windows 7 or 8 computer with more than 8 Gb of RAM and more than 128GB of hard disk space is sufficient (Lu et al., 2016). The device can be plugged directly into a standard USB3 port on the computer. The MinKNOW software allows an easy device control and data acquisition. ONT has also developed the Flongle adapter for the MinION Flow Cells to enable direct, real-time sequencing, on smaller, single-use Flongle Flow Cells, which can generate up to 2.8 Gb.

The principal limitation of the technology is its relatively high error rate around 5% for the most widely used flow cell R9.4.1 (FLO-MIN106D) that has decreases, according to the ONT's website information, to 3% for the newest R10.3 flowcell. The error rate can go up to 20% depending on the flow cell and the library preparation (Kono and Arakawa, 2019). This is directly linked to the reading mode, which does not proceed base by base. The size of the nanopores currently is such that the length of DNA present inside the pore corresponds to five or six bases and the signal observed is therefore characteristic of a sequence of bases more than each single one.

3.2 Oxford Nanopore sequencing of HBV

Only few published papers on HBV sequencing using the ONT technology are available. The first published HBV Nanopore sequencing was performed in 2018 on previously characterized HBV-infected clinical samples. They succeeded in recovering the complete virus sequence of 3,200 nt and identified recombinant variants. Despite an error rate of 12%, the HBV consensus genome was consistent with the Sanger sequencing (Sauvage et al., 2018). The study was performed with the R7.3 flow cell and would probably have had a lower error rate with the R9.4.1 flow cell. McNaughton and coll. sequenced viral DNA after isothermal rolling-circle amplification to enrich HBV DNA and generate concatemers of amplicons containing multiple successive copies of the same genome. The same libraries were sequenced using the ONT and the Illumina technology. The amplification step improved sequencing accuracy to 99.7% (McNaughton et al., 2019). Nanopore sequencing has been applied also to sequence HBV DNA in dried serum spots using an extraction-free DNA technique (Astbury et al., 2020). Nanopore sequencing on a MinION also allowed the detection of HBV integrations in tumor and non-tumor liver tissue from an HBV positive hepatocellular carcinoma (Smits et al., 2021). Finally, Nanopore sequencing has been used to study 5mCpG distribution in native HBV sequences (Goldsmith et al., 2021). Up to date, no published data are available on RNA sequencing (cDNA-PCR amplicons or direct RNA sequencing) on MinION devices.

4. Detection and sequencing of HBV integrations

4.1 Southern Blot

In the 1980s, several studies investigated liver samples from HCC patients by Southern Blot using probes specific for HBV DNA. Integrated HBV DNA is detected as high molecular weight bands (Figure 32) that does not generate a 3.2 kb band after digestion with restriction enzymes cutting only once in the viral genome (typically EcoRI) (Fowler et al., 1986; Shafritz et al., 1981). Southern blot is not very sensitive and can detect between 1 to 10 pg of HBV DNA (i.e. 10^3-10^5 copies of HBV) and does not allow to estimate the quantity of these integrations.

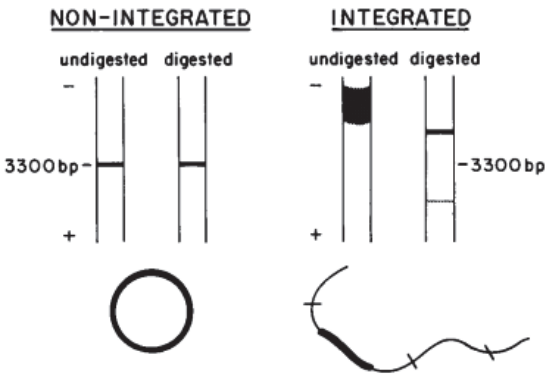


Figure 32 - Southern Blot result of HBV DNA samples from (Shafritz et al., 1981)

4.2 PCR-based approaches

A much better sensitivity is obtained by using methods based on PCR and on the detection of cell-virus junctions. Each integration forms a unique junction between the cellular genome and the viral genome. It is this unique fusion that allows researchers to study integration. This fusion also enables researchers to study the role of integration in hepatocarcinogenesis and to trace the clonal origin of a hepatocyte. These cells are a self-renewing cell population with a very small contribution from stem cells. Thus, the daughter cells of a hepatocyte, containing a particular virus/cell junction, also contain the same fusion.

By coupling HBV-specific primers and primers annealing to Alu sequences, Short-Interspersed Nuclear Elements (SINEs) found in the number of 1 million per human genome, (Figure 33) to analyze HCC samples it has been shown that HBV integrations are preferentially found near the ORFs of genes associated with tumor progression (Grover et al., 2003; Larsson et al., 2018). However, this study is controversial because Alu sequences are concentrated in the proximity of ORFs and not evenly distributed in the genome (Grover et al., 2003; Larsson et al., 2018)

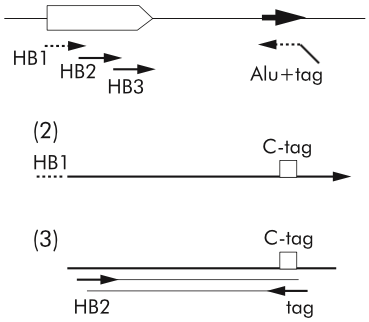


Figure 33 - Alu-PCR technique to detect HBV DNA integrations
from **(Murakami et al., 2005)**

The primer HB1 is synthesized with dUTP instead of dTTP and can be denatured by uracil DNA glycosylase (UDG) treatment. After a first PCR, the products are flanked by a HBV sequence and a new tag sequence. After UDG treatment the HB1 sequence is degraded. A second and third PCR are performed using primers HB2 and HB3 with the primer tag.

The inverse nested PCR (invPCR) technique (Figure 34) was developed in the 90' to detect fragments of DHBV integrated in the liver of DHBV-infected ducks. DNA containing cell and virus sequences is first cleaved with restriction enzymes that cut once in the virus and multiple times in the host genome, sujected to circularization by DNA ligation, and then cleaved again with restriction enzymes expected to cut once in the viral sequences. These steps generate an inverted product in which the unknown cellular sequence is now flanked by the integrated virus sequences. Subsequently, PCR reactions using primers specific for the viral sequence allows the amplification, and eventually the sequencing, of the unknown cellular sequence (Yang and Summers, 1999).

This technique was used by Mason and coll. to detect the integrated WHV DNA in WHV-infected woodchuck (Mason et al., 2010). Although this technique has a good simplicity/cost ratio, invPCR can only detect approximately 10% of total integrations. This may be due to the efficiency of fragment inversion during reverse PCR, the proportion of infected cells and the lack of restriction sites required for reverse PCR (Tu et al., 2018).

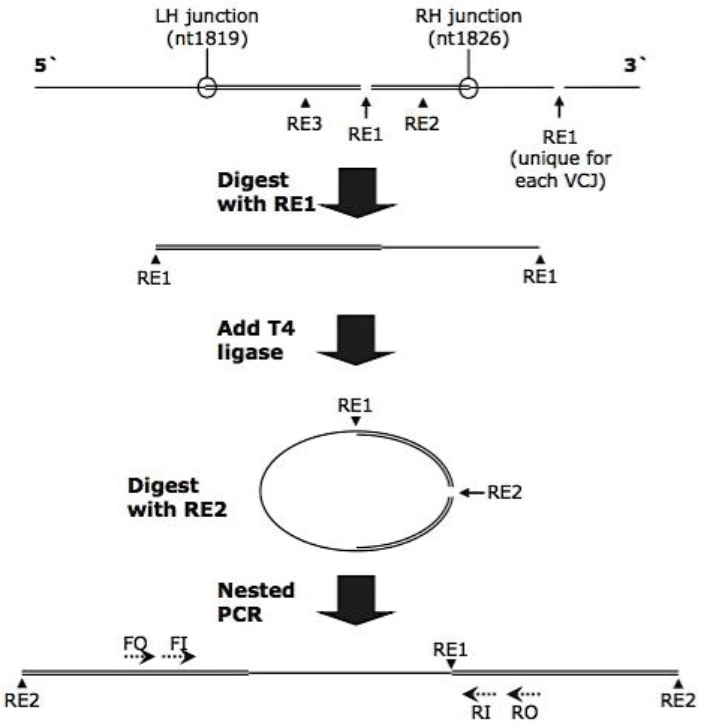


Figure 34 - Inverse nested PCR to detect HBV DNA integrations
from **(Tu and Jilbert, 2017)**

The virus-cell junction (VCJ) is excised by RE1 digestion. The digestion product is then circularized using a T4 ligase. Subsequently, an RE2 excises the viral sequence to linearized it. This forms an inverted product flanked by viral sequences. Then a nested PCR is performed.

4.3 Sequencing methods

4.3.1 Sanger sequencing

Frederick Sanger and his colleagues introduced the "plus and minus" system in 1975 to determine the order of the nucleotides in a nucleic acid sequence. This method is based on the insertion of radiolabelled nucleotides at the end of the nucleotide chain during *in vitro* DNA replication. The plus and minus technique used DNA polymerase to synthesize from a primer incorporating radiolabelled nucleotides, before performing two second polymerisation reactions: a "plus" reaction where only one type of radiolabelled nucleotide is present thus all extensions will end with that base, and "minus" reaction, in which three unlabeled nucleotides are used, which produces sequences up to the position before the next missing nucleotide. Sequence analysis was then performed by X-ray Polyacrylamide Gel Electrophoresis (PAGE) comparing the 8 reactions lanes (4 "minus" and 4 "plus") (Sanger and Coulson, 1975). A similar method was also developed by Allan Maxam and Walter Gilbert

(Maxam and Gilbert, 1977). Sanger and Gilbert's sequencing methods are known as first-generation sequencing and was rewarded with the Nobel Prize in Chemistry in 1980. Using this plus and minus method, they sequenced the first DNA sequence of a bacteriophage φX174. Known today as PhiX, this sequence is used as a spike-in control in many laboratories (Sanger et al., 1977a).

In 1977 Sanger's team developed the "chain-termination' method" or dideoxy technique (Sanger et al., 1977b). After denaturation of the double strand DNA, primers, dNTPs and radiolabeled ddNTPs in low concentration are added in the 4 extension reactions (one for each base). ddNTPs lack the 3' hydroxyl group that is required for extension of DNA chains, and therefore cannot form a bond with the 5' phosphate of the next dNTP (Figure 35a). The polymerization fragments generated are then visualized via electrophoresis on a high-resolution polyacrylamide gel and the sequence is inferred by reading 'up' the gel, as the shorter DNA fragments migrate faster (Figure 35b).

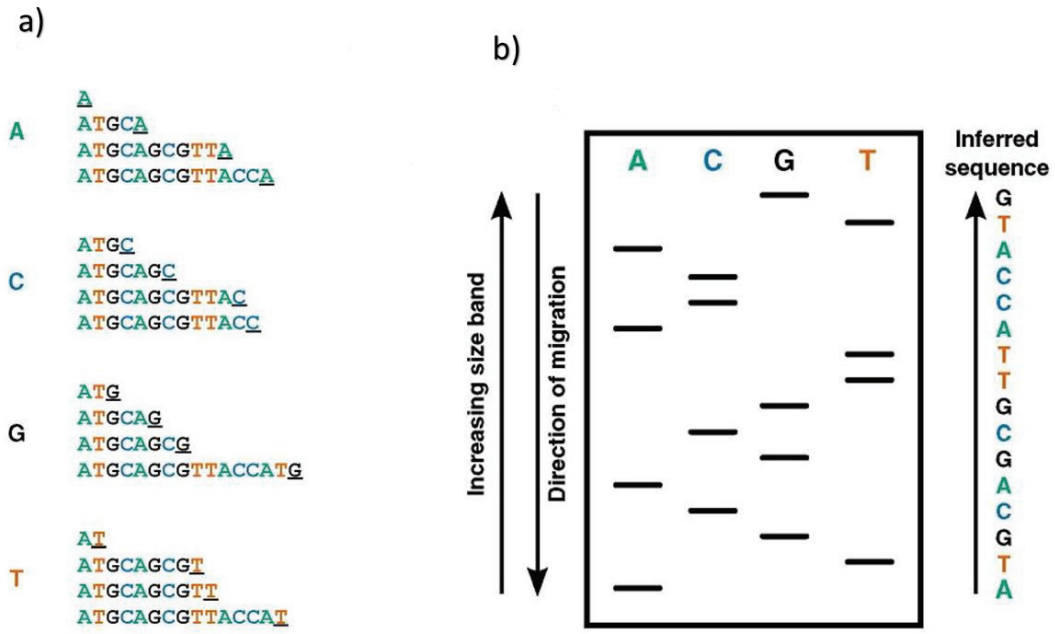


Figure 35 - Sanger sequencing principle adapted from (Heather and Chain, 2016)

A number of improvements were made to Sanger sequencing in the following years (Sanger et al., 1977b; Tipu and Shabbir, 2015): a) the replacement of radiolabelled with fluorometric base terminators carrying different fluorescent group for each of the four bases (yellow for A, blue for G, red for C and green one for T), allowing the 4 reactions can be run in one lane instead of four. Results are presented as a chromatogram on which the different fluorescence peaks corresponding to the

DNA sequence can be read (Figure 36); b) improved detection through "capillary based electrophoresis".

Sanger DNA sequencers generate reads between 600 and 900 bp in length with accuracies exceeding 99.97% (Wommack et al., 2008).

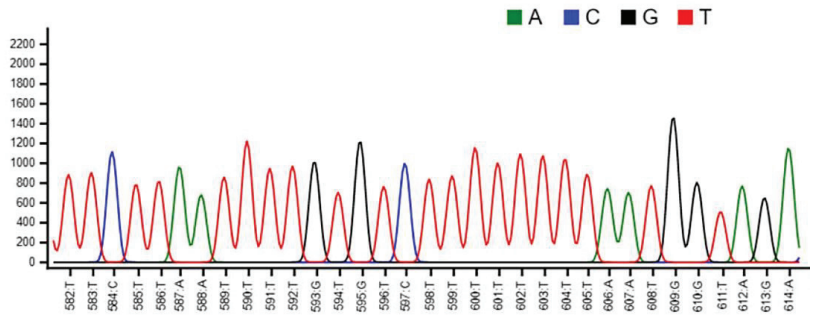


Figure 36 - Example of Sanger Sequencing Chromatogram adapted from (Rausch et al., 2020)

These improvements contributed to the development of increasingly automated DNA sequencing the paved the way to the complete genome sequencing in the human genome project (Sanger et al., 1977b; Tipu and Shabbir, 2015).

4.3.2 Illumina sequencing

The search for more rapid sequencing and for technologies that could lower the cost / base has driven the research efforts for the development of Next-Generation Sequencing (NGS) or High-Throughput Sequencing (HTS).

The Illumina® technology appeared in 2006, initially under the name of Solexa. It was acquired by Illumina® in 2007. This method is based on sequencing by synthesis (SBS). The DNA to be sequenced is randomly fragmented and then adapters are attached to each end. The collection of all these fragments coupled to the adapters is called "library". The libraries, after few rounds of amplification and accurate quantification, are loaded onto the sequencing instrument containing a flow cell with oligonucleotides complementary to the adapters. In the next step, each of the fragments attached to the flow cell is amplified by a "bridge PCR" to generate up to 1,000 identical copies in proximity called a cluster (Figure 37). Clusters density on the flow cell is around ten million single-molecule clusters / cm^2. The chemistry for next generation sequencing on Illumina platforms is similar to the

fluorescent terminator method used in Sanger sequencing with some differences. The clusters are sequenced using reversible fluorescently labeled nucleotides. During each sequencing cycle, a single dNTP is added to the nucleic acid chain. The dNTP is used as a terminator of the polymerisation so that at each incorporation the fluorescent dye is imaged to identify the base. Then a chemical cleavage and the restoration of the hydroxyl group allows the incorporation of the next nucleotide. As all four nucleotides are present in each ring, natural competition minimizes bias incorporation.

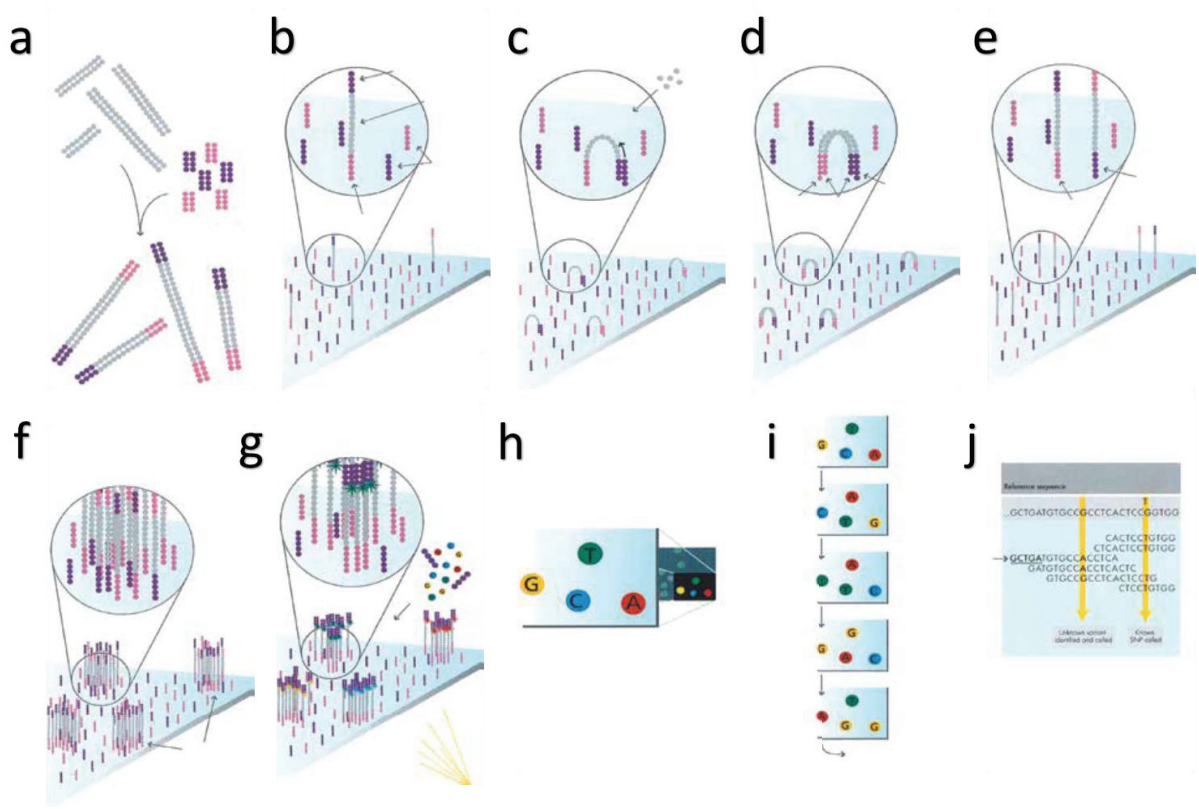


Figure 37 – **SBS Technology workflow from Illumina®**

- **a. Prepare genomic DNA sample :** Randomly fragmentation of genomic DNA and ligation of adapters to both ends of the fragments**.**
- **b. Attach DNA to the flow cell surface :** Bind single-stranded fragments randomly to the inside surface of the flow cell channels.
- **c. Bridge amplification :** Add unlabeled nucleotides and enzyme to initiate bridge amplification.
- **d. Fragments become double stranded :** The enzyme incorporates nucleotides to build double-stranded bridges on the flow cell.
- **e. Denature the double-stranded molecules :** Denaturation leaves single-stranded templates anchored to the substrate
- **f. Complete amplification** : Several million dense clusters of double stranded DNA are generated in each channel of the flow cell.
- **g. Determine first base :** The first sequencing cycle begins by adding four labeled reversible terminators, primers and DNA polymerase.
- **h. Image First Base :** After laser excitation, the emitted fluorescence from each cluster is captured and the first base is identified.
- **i. Sequencing over multiple chemistry cycles:** The sequencing cycles are repeated to determine the sequence of bases in a fragment, one base at a time (repeat step g and h over and over).
- **j. Align data :** The data are aligned and compared to a reference, and sequencing differences are identified.

The Illumina® SBS technology can be used as a 2 channel or 4 channel methods for base detection (Figure 38). The 4 channels method is based on taking 4 distinct images at each cycle of the sequencing run whereas the 2 channels method is based on taking only 2 images and therefore on a mix of dyes that are distinguished by using a red filter band for C bases and a green filter band for T bases. The clusters that are flagged by green and red filters are A bases while G bases are unflagged. This reduction in the number of image acquisitions allowed to reduce significantly the sequencing time. The MiniSeq, NextSeq 500/550 and NovaSeq 6000 systems use the 2-channel method while the MiSeq and HiSeq use the 4-channel method. According to Illumina®, since both methods are based on the same SBS technology, results from sequencers using either method are comparable.

More recently Illumina has introduced a one channel detection method that couples the SBS technology with a metal-oxide semiconductor (CMOS) technology. The technology is based on the acquisition of two images. In the first image, only the A and T bases have a colour (the same colour). The A base has a cleavable bond between the fluorescent molecule and the base. Base C, on the other hand, is modified to allow the attachment of a fluorescent molecule can then attach to it. In the second image, after the chemical cleavage/attachment, the C and T will have a colour. The iSeq 100 systems use this technology.

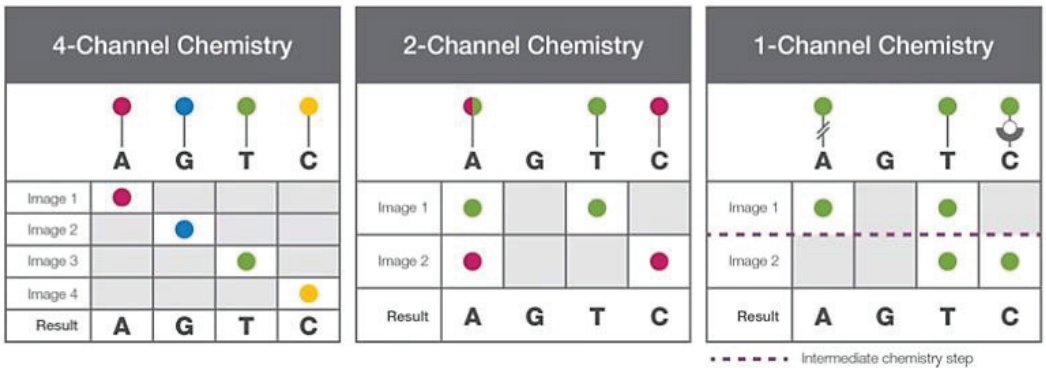


Figure 38 - Illumina® Chanel chemistry methods from Illumina® website

Illumina sequencing can be done as a single read or as a paired end read, i.e. both ends of the DNA fragment in the library are sequenced (forward and reverse reads are considered as paired reads). Paired end read sequencing generates twice the number of reads (better coverage), allows a more accurate alignment of the reads and the detection of indels (insertion or deletion) which is not possible with single read sequencing. Most recent studies are done in paired-end sequencing, except for small RNAs (miR-Seq) sequencing.

An important characteristic of NGS is the multiplexing (Figure 39), i.e. the possibility of pooling a large number of sequenced libraries simultaneously in a single sequencing run. The addition of a unique index sequence on each fragment to be sequenced makes it possible to identify each read and to sort them before the final analysis (demultiplexing). Paired-end sequencing coupled with multiplexing has considerably reduced the "time to data" for multisample studies.

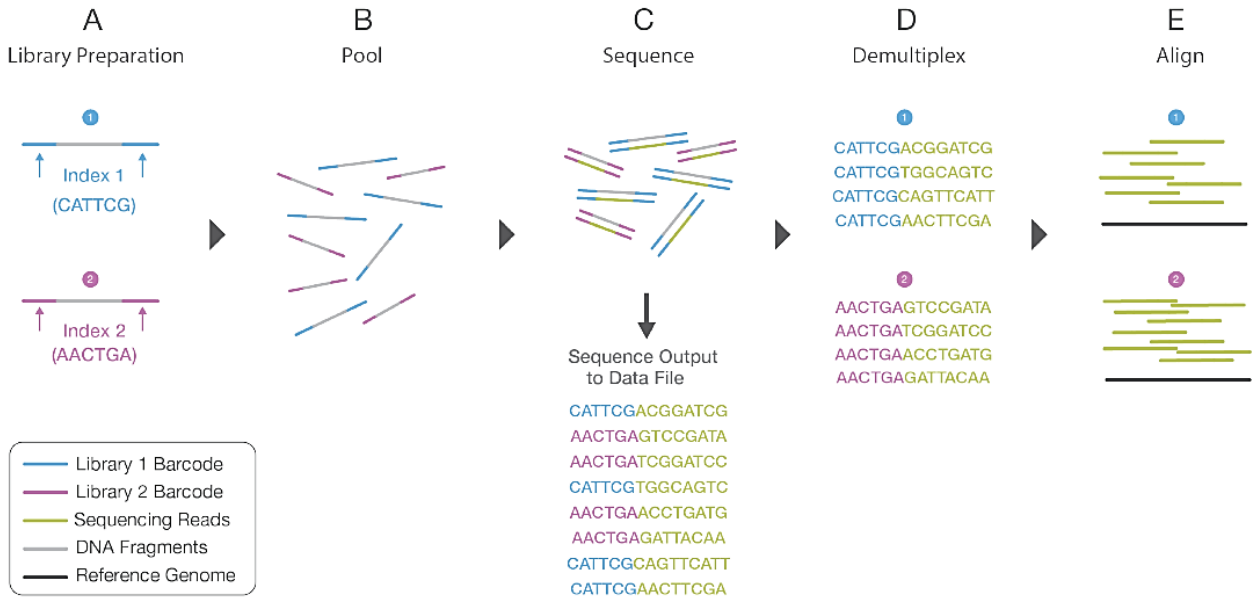


Figure 39 - Library muliplexing overview from Illumina® website

There is a wide range of Illumina® NGS instruments from the MiniSeq™ with an output between 1.8 and 7.5 Gb for targeted sequences, to the NovaSeqTM 6000 system which can generate between 6 Tb and 20 B reads in 2 days for population-scale studies.

The MiSeq sequencer was first released in 2011. It is very compact sequencer and allows both single- and paired-end runs with reads ranging from 1 x 36 bp to 2 x 300 bp. A sequencing run can provide up to 15 Gb output data and 25M single reads or 50 M paired end reads. The MiSeq is cost-effective for amplicon sequencing, target enrichment sequencing, metagenomics. With the MiSeq, the user has access to the BaseSpace® platform which is an onsite or web-based platform for real-time data uploading, data analysis and run monitoring (Ravi et al., 2018).

Base calling accuracy is evaluated by the Phred quality score also called Q score. Whether Phred assigns a Q score of 30 (Q30) to a base, this is equivalent to the probability of an incorrect base call 1 in 1000 times. Most of the time, sequencing quality will be expressed as % bases Q ≥30, e.g., the percentage of bases with a Phred quality score greater than or equal to 30.
The quality of the data obtained after sequencing depends on several factors such as library preparation, library quantification, denaturation and dilution of the libraries, clonal cluster density, loading concentration and sequencing run parameters.

Genomic/DNA sequencing include whole-genome sequencing (WGS), whole-exome sequencing (WES), de novo sequencing and targeted sequencing. Transcriptomic/RNA-Seq include total and mRNA sequencing, targeted RNA sequencing, small RNAs and noncoding RNAs sequencing.
WGS allows the sequencing of the entire genome, exome sequencing allows the sequencing of the exonic part of the genome which corresponds to less than 2% of the total genome but contains most of the known disease-causing variants. De novo sequencing corresponds to the sequencing of a new genome for which no reference sequence exists. The analysis is done by assembling the reads to obtain one or more contig sequences (the quality of the sequencing and the number of contig sequences will depend on the number of gaps between reads). Targeted sequencing allows the sequencing of a subset of genes or regions of interest of the genome allowing to decrease the cost and time of sequencing and to increase the coverage from 30-50x to 500-1000x. Two types of workflows are available for targeted sequencing: either by enrichment or by amplicon generation (Figure 40).

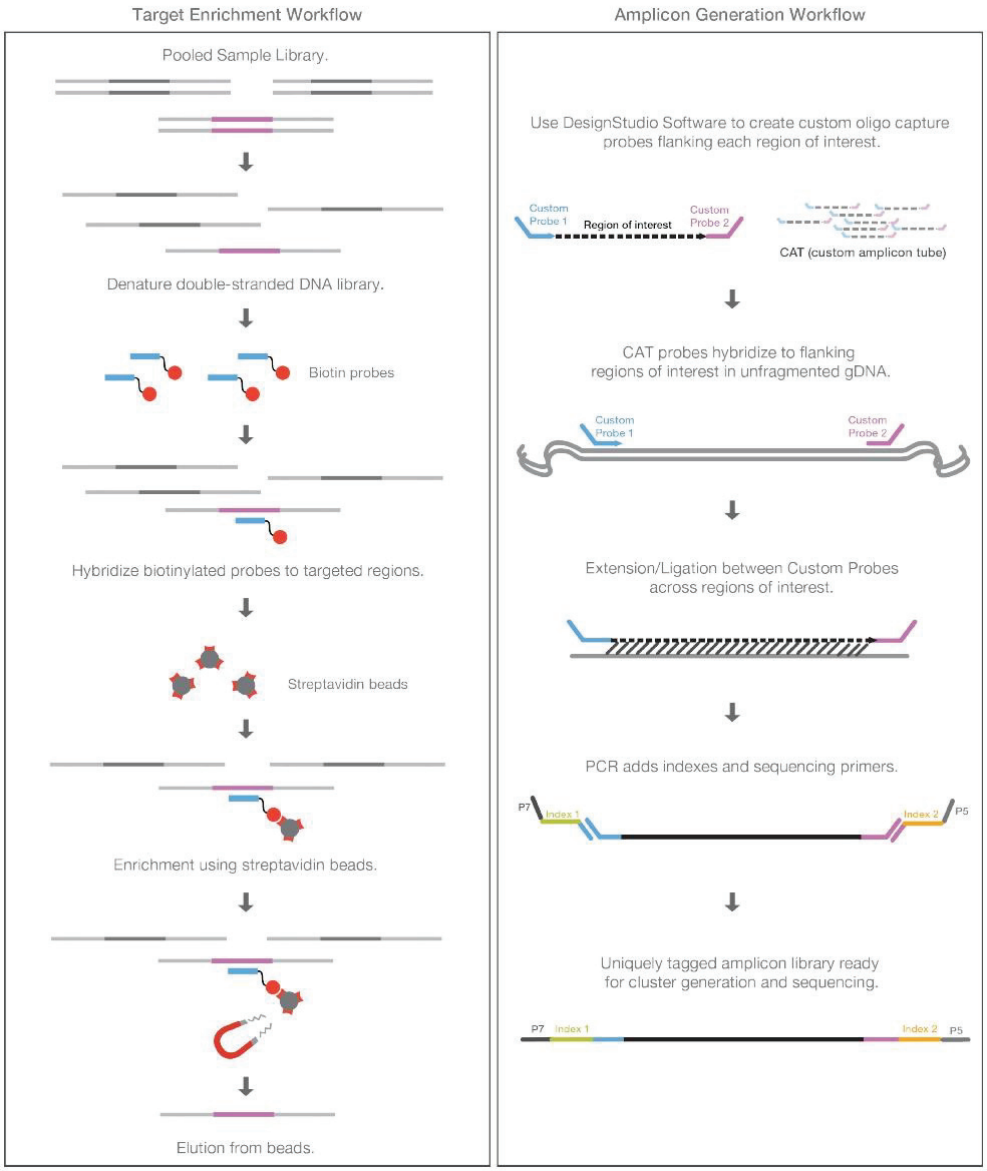


Figure 40 - Targeted sequencing workflows from Illumina®

4.3.3 The SeqCap EZ HyperCap method for HBV capture/enrichment

The workflow of the Nimblegene-Roche SeqCap EZ HyperCap method includes sample extraction, DNAse treatment, adapters ligation, libraries preparation with the HyperPrep RNA or HyperPlus DNA kits, target capture enrichment by hybridization to Nimblegen SeqCap EZ specific probes, PCR amplification (few cycles) with the KAPA HiFi LM-PCR Master Mix, followed by paired-end sequencing on a MiSeq Illumina platform (Figure 41). One advantage of the SeqCap EZ HyperCap method is its ability to work for both RNA and DNA capture.

The SeqCap EZ HyperCap method was chosen to capture HBV-host fusion sequences and to study the contribution of integrated HBV-DNA to circulating HBV RNAs. To optimize the capture of HBV sequences, high density tiling of 105mer biotinylated oligonucleotides covering the entire HBV genome of >340 HBV isolates and including all HBV genotypes and subtypes, have been used (Figure 41).

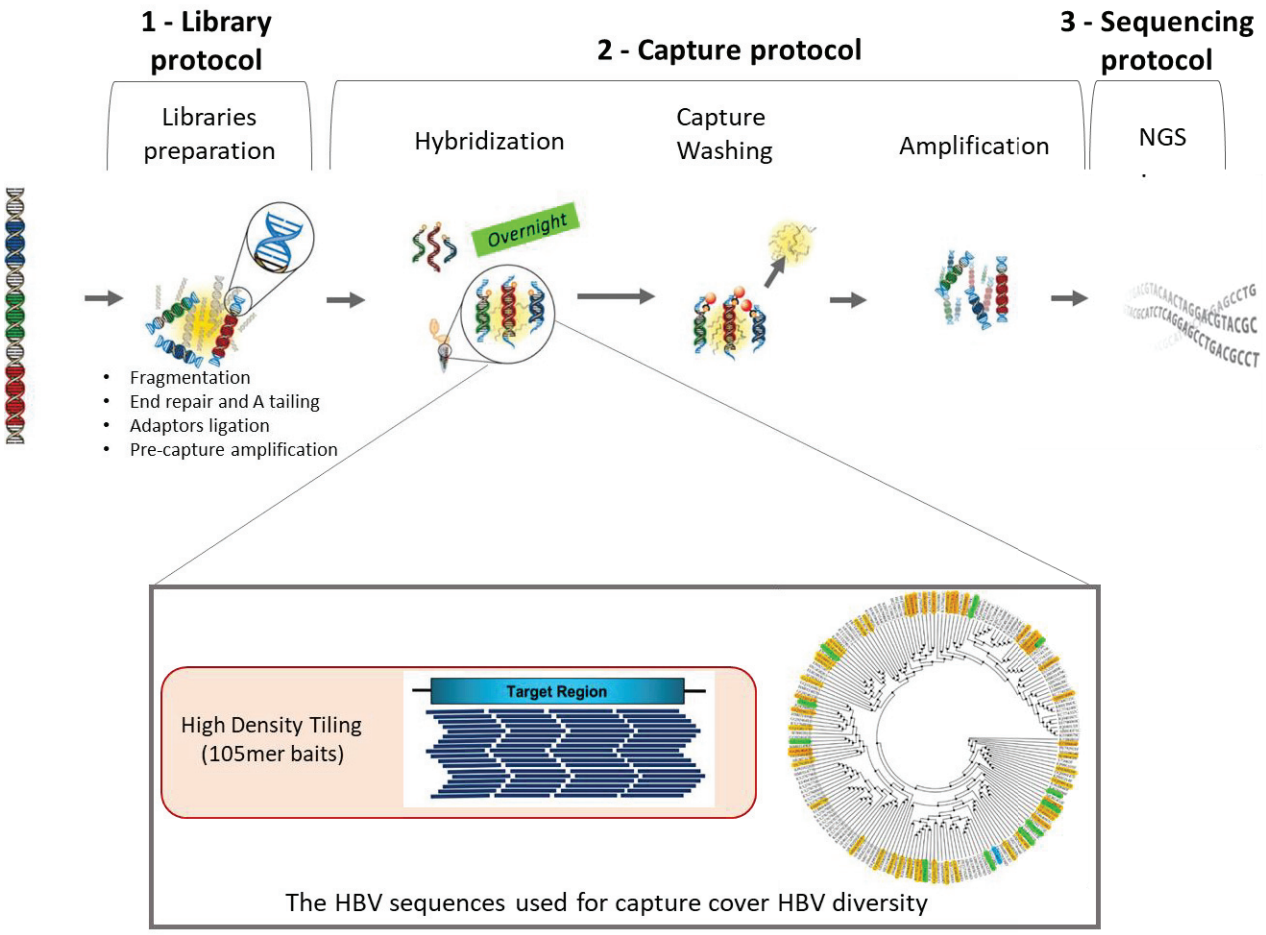


Figure 41 - HBV DNA and RNA Capture sequencing workflow

RESULTS

Generation of a molecular standard for circulating HBV RNA detection and quantification in CHB patients

Generation of a molecular standard for circulating HBV RNA detection and quantification assays in CHB patients

Alexia PATUREL[1,2], Francesca CASUSCELLI di TOCCO[1,2], Delphine BOUSQUET[1,2], Marie-Laure PLISSONNIER[1], Xavier GRAND[1,2], Françoise BERBY[1,2], Caroline SCHOLTÈS[1,2,4], Barbara TESTONI[1], Fabien ZOULIM[1,2,3], Massimo LEVRERO[1,2,3,5]

Affiliations

[1] Cancer Research Center of Lyon (CRCL), UMR Inserm U1052 / CNRS 5286, Lyon, France
[2] University of Lyon, University Claude Bernard Lyon 1, 69008 Lyon, France
[3] Department of Hepatology, Croix Rousse hospital, Hospices Civils de Lyon, France
[4] Laboratoire de Virologie, Institut des Agents Infectieux, Hospices Civils de Lyon, Lyon, France
[5] Department of Internal Medicine, SCIAC and the IIT Center for Life Nanoscience, Sapienza University, Rome, Italy

 # Corresponding author:
Prof. Massimo Levrero
Equipe “Epigenetics, microenvironment and liver cancer”
Cancer Research Center of Lyon (CRCL) - INSERM U1052
151 cours Albert Thomas
69424 Lyon Cedex 03
E-mail: massimo.levrero@inserm.fr

Abstract

Background: The quantification of serum HBV RNA levels has been proposed as a useful biomarker to evaluate HBV treatment activity and to predict patient outcome. In untreated patients, serum HBV RNA levels reflect the cccDNA transcriptional activity. Different assays have been designed and 2 PCR-based research use only (RUO) investigational assays (IA) have been developed. Quantification was calibrated using the WHO HBV DNA standard or a synthetic armored RNA (arRNA), respectively. Here we have designed and generated a stable clonal cell line producing an RNA-based standard for the calibration of circulating HBV RNA assays.

Methods: HBV RNA producing Huh7-derived stable cell lines were generated by transfecting pTriEX plasmids containing 1.1 length HBV DNA genomes carrying mutations in the catalytic site (YMAA mutation), the TP-domain (Y63F) of the polymerase, the ε-loop of the pgRNA (mutation A1G).

Results: The best clonal cell line (Huh7-3D29) secreting HBV RNAs carried a double YMAA and Y63F mutation. Huh7-3D29 maintained a high ratio RNA versus DNA phenotype over several passages in culture with negligible residual secreted HBV DNA. Density gradient centrifugation showed that most of the secreted HBV RNA was detected in naked capsid followed by virions-like particles and only a minority in exosomes. Nanopore sequencing showed that the majority of secreted HBV RNAs are full length transcripts and pgRNA derived spliced RNAs and all terminate at the HBV canonical poly-adenylation site RNAs. 5'RACE analysis confirmed the secretion of 3.5 kb RNA species. Finally, Huh7-3D29 cells showed a high and upscalable secreted RNA yield (1,300 standard curves in 9 days from 1 flask).

Conclusion: We generated a clonal cell line that produces high amounts of HBV RNAs with very low quantities of contaminating HBV DNAs, representing a stable source of RNA standard for HBV RNA assays calibration.

Key words: hepatitis B virus (HBV); chronic hepatitis B (CHB); Biomarker, pregenomic RNA (pgRNA); standard

Introduction

Hepatitis B virus (HBV) remains a major public health problem worldwide, despite the availability of an effective vaccine and antiviral therapies. 296 million people were living with chronic hepatitis B infection in 2019, with 0.5-1 million dying annually because of HBV-related advanced liver disease. Long term treatment with anti-HBV nucleos(t)side analogues (NUCs) suppresses viral replication, liver inflammation and disease progression (Lampertico et al., 2017). However, only a small proportion of patients achieve a *functional HBV cure*, defined by a sustained off-treatment loss of HBsAg that is associated with an improved long-term clinical outcome and can stop treatment with very low rates of disease reactivation (Cornberg et al, 2020). Innate and adaptive antiviral immune responses control the residual cccDNA reservoir (Fanning et al., 2019). A *complete HBV cure* with eradication of all cccDNA in the liver and no residual risk of viral reactivation, is rarely, if at all, achieved with NUCs. The development of new antiviral therapies targeting the cccDNA reservoir is a major objective towards *HBV cure* (Fanning et al., 2019). Several new molecular entities (NMEs) with direct antiviral or immuno-modulatory activities, are at the preclinical or clinical evaluation stage with the aim of inducing a cure with a finite duration treatment. Direct assessment of cccDNA levels and activity requires liver biopsy and cccDNA quantification remains technically challenging (Nassal, 2015; Charre et al., 2019). To assist the clinical development of these NMEs, there is an urgent need for the development of non-invasive biomarkers that adequately reflect the size and transcriptional activity of intrahepatic cccDNA. HBV DNA and HBsAg, the currently used biomarkers to define HBV endpoints have several limitations in this setting: i) serum HBV DNA is the best indicator of the antiviral activity of a new treatment in naive patients but it is already suppressed in NUC treated patients; ii) HBsAg loss cannot be considered a biomarker of the activity of new therapeutic molecules since it is at the same time the desired endpoint and part of the definition of cure; iii) the predictive value of different thresholds of HBsAg decrease remains to be validated and the kinetic of HBsAg decrease is often too slow to allow early outcome prediction in clinical trials. In addition, HBsAg may be expressed from viral genome integrated in the host genome and thus may not fully reflect the transcriptional activity of cccDNA (Wooddell et al., 2017). These new biomarkers would also be useful to: a) better characterize patients in the different phases of the HBV infection and disease; b) predict response to IFN therapy early after its initiation to extend treatment only in patients with high probability to reach a functional cure; and c) identify patients that could discontinue NUCs without risk of relapse (Charre et al., 2019; Fan et al., 2020).

Several studies have identified circulating HBV RNAs as a biomarker reflecting the transcriptional activity of cccDNA (Giersch et al., 2019; Liu et al., 2019) and have shown that circulating HBV RNA may help monitor CHB infection (Carey et al., 2019; Fan et al., 2020; Ma et al., 2020). Although the potential relevance of circulating HBV RNAs as a biomarker is well established, a major limitation is the lack of standardized protocols (e.g., sample extraction, retro-transcription, choice of primers for amplification, genotype inclusivity, specificity for RNA). RUO IA kits to quantify circulating HBV RNAs have been developed by industry (Butler et al., 2018)(Scholtès et al. 2022). Another important limitation is the lack of a recognized standard to calibrate circulating HBV RNA quantification.

A World Health Organization (WHO) International Standard (IS) for HBV DNA has been established to use as a calibrator for HBV nucleic acid amplification technique (NAT)-based assays. Due to their limited amount, four standards, derived from the same plasma donor (genotype A2, HBsAg subtype adw2, quantified at 2.7 x 10^9 HBV DNA molecules /mL (Heermann et al., 1999), GenBank accession number KY003230 (Jenkins et al., 2017)) have been validated and distributed over time (Baylis et al., 2008; Fryer et al., 2017; Organization and Standardization, 2016; Saldanha et al., 2001).The WHO HBV DNA standard has been used as a calibrator for the quantification of serum HBV RNA (Butler et al., 2018) but its appropriateness has been challenged. Indeed, it is questionable whether an HBV RNA unit is equivalent to an international HBV DNA unit defined by the plasma derived WHO HBV DNA standard that also contains circulating HBV RNA (Liu et al., 2020).

Here we describe the generation of a stable clonal cell line producing an RNA-based standard for the calibration of circulating HBV RNA assays.

Materials and methods

Cell lines

The human Huh7, Hep3B and PRF/PLC/5 hepatocarcinoma cell lines were cultured at 37°C in a humidified atmosphere containing 5% CO2 in Dulbecco's modified Eagle medium supplemented with 10% Hyclone fetal clone II serum 1% GlutaMax (Gibco), 1% Penicillin / Streptomycin (Gibco), 1% sodium pyruvate (Gibco), 1% non-essential amino acids (NEAA, Gibco).

Plasmids

The HBV RNA producing cell lines was generated by transfecting a pTriEX plasmid derivative containing 1.1 units length wild type and mutated HBV DNA genomes in Huh7 cells. In total, 3 mutations were selected: YMAA and Y63F in the polymerase ORF, and A1G in the ε-loop region. Six HBV mutated genomes of genotype D have been generated and inserted into the pTriEX-Bsd vector: a plasmid containing the WT 1.1 units HBV genome, 2 plasmids containing the 1.1 units HBV genome carrying the YMAA mutation or the Y63F mutation and 3 plasmids containing the 1.1 units HBV genome carrying the following combinations of mutations: YMAA + Y63F, A1G + YMAA, and A1G + Y63F. All plasmids include the blasticidin resistance gene for cell culture selection. Wild type and mutated 1.1 units HBV genomes were synthetized and the plasmids produced and controlled by sequencing by GenScript®.

Cell transfection and culture

Huh7 cells were seeded at 60–80% confluency and were allowed to adhere overnight. Cells were then transfected with the indicated amounts of total plasmid DNA with the TransIT®-2020 kit (Mirus) in serum-free Opti-MEM medium (Life Technologies), following the manufacturer's instructions. After 12h, the transfection medium was removed and was replaced with fresh medium. Transfected cells were grown and amplified in selective growth media (Huh7 medium described above) supplemented with 4 µg/mL of Blasticidin).

Clonal cell line isolation and expansion

Colony formation was obtained after serial dilution of Huh7 polyclonal cell population from $1x10^4$ cells to 100 cells per $150cm^2$ petri dish. Cell plates were incubated at 37°C with 5% CO_2 for 20 days to allow colony formation and growth. Using single channel pipettor, single colonies were transferred

to 1mL PBS. Cells were pelleted by centrifugation for 5 min at 900rpm and resuspended in 50 µL of 0.25% trypsin-EDTA (Gibco) and kept for 5 min at 37°C. Cells were then plated in 24 wells plates with 1mL fresh Huh7 media. Clonal cell lines were amplified under blasticidin selection. Clones to be further carried on were chosen based on cell growth, HBV RNA / HBV DNA ratio, and HBV RNA productivity.

High scale supernatant production

To test the stability over time of the HBV RNA secretory phenotype and to perform higher scale supernatant collection to produce the HBV RNA standard, cells were seeded (25 x 10^6 cells / 175 cm^2 flask) and cultured overnight before change of media containing 2.5% DMSO. To collect large quantities of HBV RNAs, cells were kept under standard culture conditions for 9 days. Every 3 days, supernatant was collected, and fresh media was added. After 9 days, all collected supernatants were pooled and centrifuged for 5 min at 1,500 rpm to remove any cellular debris.

Ultracentrifugation, DNA and RNA isolation

Pooled cell culture supernatants were ultracentrifuged for 5 hours at 25,000 rpm over a 4 mL 20 % w/v sucrose cushion. After ultracentrifugation, the supernatant was removed, and the virus-containing pellet was resuspended in 600 µL of PBS 1X. 28.5 µL aliquots (28.5 µL out of 600 µL) of each resuspended pellet was completed to 200 µL with PBS 1X. DNA and RNA from resuspended pellets were extracted using High Pure Viral Nucleic Acid (Roche, Diagnostics) according to the manufacturers' instructions. The same final volume of extracted nucleic acid elution (40 µL) was used for DNA or RNA analysis. RNA samples were treated with RQ1 RNase-Free DNase (Promega, Cat#M6101) for 30 min at 37°C and stored until use.

Reverse transcription (RT) and droplet digital PCR (ddPCR)

For HBV RNA analysis, 4 µL of DNase-treated RNA was reverse transcribed and amplified using the SuperScript™ IV VILO™ Master Mix (Invitrogen, Cat # 11766500). A 22 µL reaction mixture was prepared comprising 11 µL of 2X ddPCR Supermix™ for probes (no dUTP) (Bio-Rad), 1.1 µL of primers and probe mix, and 5 µL of cDNA or DNA. Nucleic acid inputs were adjusted to have acceptable rates of negative events: samples were diluted 1:5. Probes and primers include: the HBV Pa03453406_s1 set (Thermofischer) for total HBV RNA quantification and the forward primer ggagtgtggattcgcactcct, reverse primer agattgagatcttctgcgac and probe aggcaggtcccctagaagaagaactcc for 3.5 kb HBV RNA

species quantification. Droplet formation was carried out using a QX100 droplet generator. Subsequent amplifications were performed in the C1000 Touch™deep-well thermal cycler (Bio-Rad) with a ramp rate of 2°C/s and the lid heated to 105°C, according to the Bio-Rad recommendations. First, the enzyme was activated at 95°C for 10 min followed by 40 cycles of denaturation at 94°C for 30 s and 60°C for one minute. The enzyme was deactivated at 98°C for 10 min and the reaction was kept at 4°C.

5'RACE analysis

5'RACE 5'RACE was performed as previously described in (Stadelmayer et al., 2020). Briefly, RNAs were isolated using a guanidinium thiocyanate–phenol–chloroform extraction protocol (TRI reagent (Sigma). 5'RACE was essentially performed as described in the GeneRacer Kit manual (ThermoFisher Scientific) except Tobacco Acid Pyrophosphatase was substituted by RNA 5' Pyrophosphohydrolase (New England Biolabs) and SuperScript reverse transcriptase III by SuperScript reverse transcriptase IV (ThermoFisher Scientific). The reverse transcription reaction was performed using 3' HBV specific Gsp1 primer 5'-TTAGGCAGAGGTGAAAAAAGTTG-3'. For the 5'RACE PCR reaction Prime Star super mix DNA Polymerase (TAKARA bio), GeneRacer 5' primer and HBV specific nested primer Gsp1 5'-TTAGGCAGAGGTGAAAAAAGTTG -3' were used. 5'RACE PCR was run in a C100 Touch thermocycler (Biorad) using the following PCR program: Initial denaturation step 98°C 3 min >5x (98°C 10 s; 72°C 3 min) >5x (98°C 10 s; 70°C 3 min) >25x (98°C 10 s; 64,4 20 s; 72°C 3min) >72°C 10min.

Quantification of HBV RNA with the Roche Manual Workflow (MWF) investigational assay

The Roche Manual Workflow (MWF) is a manual version developed for research purposes of the cobas® 6800/8800 Automated Investigational Assay (IA) for the detection and quantification of circulating HBV RNAs in chronic HBV patients (Scholtes et al., 2022). MWF uses identical chemistry and PCR primers/probe sets as the cobas® 6800/8800 automated IA but does not require high throughput and automation. Results obtained with this manual assay were highly correlated with those obtained with cobas HBV RNA (Scholtes et al., 2022). HBV RNA assay characteristics are the following: i) linear results between 10 and 10^7 copies/mL in clinical samples of several HBV genotypes; ii) precision and reproducibility with standard deviation below 0.15 $\log_{10}$ copies/mL and coefficients of variation below 5% throughout the linear range; minimal impact (<0.3 $\log_{10}$ copies/mL) of HBV DNA on HBV RNA quantification at DNA:RNA ratios of up to approximately 10^6. HBV RNA quantification was calibrated using a synthetic armored RNA (arRNA) containing 435 bp derived from

the 3’ end of HBV pgRNA, packaged in MS2-phage (kindly provided by Roche Diagnostics, Pleasanton, CA) arRNA was quantified by ddPCR (BioRad) using primers and a probe in the precore/core region. HBV RNA quantification with the MWF was performed according to the manufacturer's protocol.

Iodixanol/sucrose density gradient centrifugation

300 mL of supernatants were collected and centrifuged at 1,500 xg for 15 min at room temperature. Cellular debris were removed by filtration through a 0.22 µm filter (Merck Millipore, KGaA, Darmstadt, Germany). Viral particles were concentrated using Amicon® Pro Purification System with 100kDa filter. Concentrated supernatants were then ultracentrifuged at 110,000 ×g for 2h at 4˚C. The pellets were washed with 8 mL of PBS and a second ultracentrifugation was performed at 110,000 ×g for 2h at 4˚C. Pellets were resuspended in 2.5 mL of 10% iodixanol solution. 10%, 20%, 30% and 40% iodixanol solutions were prepared by mixing Optiprep™ (Axis Shield) with buffer containing 0.25 M sucrose, 10 mM Tris at pH 8.0, and 1 mM EDTA, with final pH 7.4. Resuspended EVs pellets were layered on the top of the gradient and then subjected to ultracentrifugation in a SW41-Ti Rotor tube (Beckman) for 6 hours at 4˚C at 110,000 ×g. Twelve fractions of 1mL were recovered and analyzed separately.

Viral protein detection

ELISA tests for HBeAg and HBsAg detection in cell supernatants were performed according to the manufacturer's protocol using the CLIA kits from Autobio Diagnostic.

Viral proteins were detected in the density gradient fractions by Western blot.

Fractions were mixed with Laemmli buffer and heated at 95°C for 5 min. Proteins were migrated in 4-20% mini-PROTEAN@ TGX stain-FreeTM Precast Gel (Bio-Rab Laboratories) and transferred onto a nitrocellulose membrane (Bio-Rab Laboratories). Membranes were blocked 1 hour with 5% milk or BSA (Sigma) in TBS (1 x Tris Buffer Saline (Sigma) and stained with primary antibodies in blocking buffer overnight at 4°C and stained with HRP-conjugated secondary antibodies (1/50,000) for 1 hour at room temperature. Detection was performed using Clarity or Clarity Max Western ECL and the ChemiDoc XRS system (Biorad).

Libraries Preparation and MinION Sequencing

RNA from supernatants was extracted using the QIAamp® MinElute® Virus Spin kit (Qiagen). Samples were prepared individually following the recommendations of the Nanopore cDNA-Seq protocol for the SQK-PCS109 kit. cDNA synthesis was performed using 9µl of ultracentrifuged extracted RNA from cell supernatant. RT and strand-switching primers were provided by Oxford Nanopore Technologies (ONT) with the SQK-PCS109 kit. Following RT, PCR amplification was performed using the LongAmp Taq 2X Master Mix (New England Biolabs, Ipswich, MA, USA) and the following cycling conditions: 1 cycle (95°C for 30 s), 14 cycles (95°C for 15 s, 62°C for 15 s, and 65°C for 3 min), and 1 last cycle (65°C for 6 min). PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA). The cDNA sequencing libraries were prepared using a total of 100 fmol of cDNA each. Nanopore libraries were the sequenced using a MinION Mk1B sequencing device with R9.4 flow cells. Sequencing was controlled and data were generated using ONT MinKNOW software (v3.4.12). Runs were terminated after 48 h and FAST5 files were generated.

Data processing and Bioinformatic Analysis

Base calling was performed with Guppy (v. 5.0.14, ONT) and simultaneously filtered for base called reads with PHRED quality > 7. Further quality controls of reads before and after mapping were performed with pycoQC (v. 2.5.2). Median PHRED score is 11.455 for passed reads and 99.6 % of reads are conserved with acceptable quality. Reads containing amplification adapters sequences were extracted with seqkit (v. 2.1.0). The *grep* function permitted to extract reads containing Forward adapter or Reverse adapter, allowing 4 mismatches due to error rate of Oxford MinION Nanopore sequencing. Then, the *rmdup* function was used to remove duplicated reads. Reads that passed all filters were mapped to each known HBV transcript species (preC, pgRNA, preS1, preS2, HBs, L-HBx, canonical HBx, s-HBx), to splice-variants (SP01-20) derived from HBV genome (strain ayw, NC_003977.2), and to the genomic sequence corresponding to the full-length preCore RNA sequence (from TSS to the canonical HBV poly-A) with Minimap2 (v. 2.21-r1071) using the spliced long reads option (*-ax splice*). *Sam* files were filtered, binarized and sorted with Samtools (v. 1.7). HBV transcript species and splice variants were quantified using Salmon (v. 1.6.0) based on transcriptomic alignment. Start positions were extracted from genomic alignment using Samtools and graphical representations were produced using R scripts (v. 4.1.2).

Statistics

Statistical tests were performed using GraphPad Prism v7.05 sofware. Pairwise comparisons were analyzed using the Wilcoxon signed-rank test for non-parametric data (Fig. 3B) and paired t test for parametric data (Fig. 3Aii). Shapiro-Wilk normality test was performed to assess the normality of the parameter's distribution (Graphpad prism). Differences were considered significant at confidence levels greater than 95% ($P \leq 0.05$).

Results

Strategies to generate an HBV RNA standard

From a theoretical point of view an HBV RNA standard could be produced by: i) extracting HBV RNA from patients' sera/plasma, but residual HBV DNA present in serum could be a confounding parameter, ii) by *in vitro* transcription but this would lack the structural components of RNA containing viral particles and would not allow to standardize the extraction step, iii) by engineering HBV replicating cell lines to exclusively or predominantly produce RNA containing viral particles similar to those observed in patient sera. These different approaches must balance quantitative yield, convenient production setting and costs with the need to avoid / minimize the impact of residual HBV DNA on the HBV RNA assay. We have chosen to generate a stable cell line carrying an integrated HBV genome with mutation in the Pol ORF leading to the expression of a non-functional polymerase that would still allow pgRNA encapsidation but prevent its reverse transcription into rcDNA (Figure 1a). This would result in an inversion of the HBV DNA/RNA ratio with a higher proportion of pgRNA than rcDNA in cell supernatants.

Three candidate mutations (YMAA, Y63F and A1G) were selected according to the published data and their expected impact on pgRNA encapsidation, reverse transcription, rcDNA formation and global viral production (Figure 1b).

The YMDD motif is in the catalytic site of the HBV Pol and mutations in this motif are involved in resistance to lamivudine (Hatakeyama et al., 2007; Lai et al., 2003; Ono-Nita et al., 1999; Zhu et al., 2011). Many mutations in the YMDD domain also result in a 20 to 90% decrease in viral DNA synthesis. The specific mutations selected for this study change the YMDD motif in YMAA (D540A/D541A).

The Polymerase tyrosine 63 residue (Y63) is required for the priming of reverse transcription by providing the hydroxyl group necessary for the covalent binding between the first nucleotide of the negative strand of DNA and the viral polymerase (Clark et al., 2017; Lanford et al., 1997; Zoulim and Seeger, 1994). Mutation of tyrosine 63 to a phenylalanine (Y63F) results in a strong increase in HBV RNA virions containing pgRNA (Wang et al., 2016).

The third mutation that was selected is A1G or Q18R in the pgRNA stem loop structure of the encapsidation signal. It was shown to induce a drastic decrease in viral DNA synthesis by interfering with the *in vitro* priming step (Feng et al., 2013).

A wild type (WT) and 5 mutant 1.1 units length HBV genotype D genomes were synthesized and cloned into a pTriEX vector plasmid. After transfection and blasticidin selection, 6 Huh7 polyclonal cell lines (WT, (1) YMAA; (2) Y63F; (3) YMAA + Y63F; (4) A1G + YMAA and (5) A1G + Y63F). HBV DNA and HBV RNA analysis in cell extracts and cell supernatants showed that the polyclonal cell lines carrying the (3) YMAA + Y63F and (5) A1G + Y63F displayed the desired secretory phenotype with an inversion of the HBV DNA/RNA ratio in cell supernatant and were selected for the generation of clonal cell lines (Supplementary Figure 1).

The Huh7 polyclonal cell lines WT, (3) YMAA + Y63F and (5) A1G + Y63F were plated at low density for colony formation. After 20 days of blasticidin selection 24 clones for each polyclonal cell line were isolated, expanded and phenotyped for HBV DNA/RNA secretion. Clones Huh7-WT18, Huh7-5D21, Huh7-3D29 and Huh7-3D35 were selected based on cell growth, HBV RNA / HBV DNA ratio and HBV RNA production (data not shown) and carried on for further characterization and expansion.

Secretory phenotype of selected clonal cells lines.

At 9 days post plating, Huh7-5D21, Huh7-3D29 and Huh7-3D35 clonal cell lines showed high HBV RNA levels in cell supernatant as compared to HBV DNA levels (Fig. 2A). Huh7-WT18 cells secreted 15 folds more HBV DNA than HBV RNA (1.06×10^7 copies of HBV DNA / mL vs 0.7×10^6 copies of 3.5 kb HBV RNA / mL). Conversely, the three selected mutant cell lines displayed an inversion of the secreted HBV DNA/HBV RNA ratio. 3.5 kb HBV RNA secretion levels were 1.03×10^7 copies / mL in Huh7-3D35 cells, 1.76×10^7 copies / mL in Huh7-5D21 cells and 6.09×10^7 copies / mL in Huh7-3D29 cells. All the 3 mutant clonal cell lines secreted more 3.5 kb HBV RNA than the Huh7-WT18 wild type clone: 92 times higher in Huh7-3D29, 27 times higher in Huh7-5D21 cells and 16 times higher in Huh7-3D35 cells. Notably, lamivudine treatment for 9 days resulted in a sharp reduction of secreted HBV DNA in Huh7-WT18 wild type cells and a further reduction, but not a complete abrogation, of secreted HBV DNA Huh7-3D29 cells (Supplementary Figure 2).

After first series of freezing/thawing (3 passages of amplification) all the Huh7-5D21 and Huh7-3D29 mutant cell lines maintained their RNA secretory phenotype (Fig. 2B). A decrease in all HBV nucleic acids secretion was observed over the passages with a 35% reduction of 3.5 kb HBV RNA secretion for Huh7-3D29 cells but without a significant change in the HBV RNA/DNA ratio. Huh7-3D29 was confirmed to be the best candidate clonal cell line for HBV RNAs secretion.

The inverted HBV DNA/RNA ratio was confirmed also after a second cycle of freezing/thawing (3-4 passages of amplification) (Fig. 2C). A decrease of 40% of secreted HBV 3.5 kb RNAs was observed in Huh7-3D29 cell line (3.99 x 10^7 vs 2.44 x 10^7 copies/mL) and the HBV RNA secretory phenotype was conserved.

3D29 HBV RNA standard validation using the HBV RNA Roche MWF

Ten fold dilution (from 10^5 to 10^1 copies/mL) of HBV armored RNA (arRNA, Roche Diagnostics, Pleasanton, CA) and Huh7-3D29 supernatants, processed as described in Materials and Methods, have been quantified using the HBV RNA Roche MWF kit. The amplification curves obtained after RT-qPCR from the MWF showed very similar profiles (Fig. 3Ai and Aiii). Furthermore, no significant difference was found between CTs of the 10-fold diluted arRNA and Huh7-3D29 cell supernatants.

HBV RNA quantification with and without DNase treatment on supernatants from Huh7-3D29 cells collected and processed at three different passages showed no significant difference (Fig. 3B). The Roche MWF IA used here is a manual version of the cobas® 6800/8800 automated IA for HBV RNA quantification that does not include a DNase treatment. The lack of differences in RNA quantification with the MWF between +/- DNase treatments, thus validates the use of Huh7-3D29 supernatants as an RNA standard for the HBV RNA cobas® 6800/8800 automated assay.

Characterization of the HBV RNAs secreted by the Huh7-3D29 clonal cell line

Huh7-WT18, Huh7-3D29, and Huh7-3D35 cells supernatants were further analyzed by iodixanol/sucrose density gradient ultracentrifugation followed by HBsAg quantification by ELISA, HBV DNA, and total HBV RNA quantification by ddPCR, HBc and CD9 detection by Western blot (Fig. 4).

In Huh7-WT18 cell supernatants (Fig. 4A) HBsAg was detected in fractions 4-8 of the density gradient, with one major peak in fraction 7 (Fig. 4A-i), and HBc was detected from fractions 7 to 9 (Fig. 4A-ii). HBV DNA was found in almost all the fractions except in fractions 1 and 2, with a peak in fraction 8 (Fig. 4A-i). HBV RNA was found in fractions 7-8 (Fig. 4A-i). Fractions 3 and 4 contain exosomes, identified by the detection of the exosomal membrane protein CD9 (Klymiuk et al., 2019). In Huh7-WT18 cell supernatant exosomes contained only HBV DNA. The detection of both HBsAg and HBc identifies viral particles, while the detection of HBc only identifies naked capsids. Fractions 7 and 8, positive for both HBsAg and HBc, were positive for both HBV DNA (virions) and HBV RNA (RNA

particles). Fraction 9, positive for HBc only, contained only low levels of HBV RNA compared to HBV DNA, indicating that naked capsids mostly carry HBV DNA in WT cells.

In the HBV mutant clonal cell lines Huh7-3D29 and Huh7-3D35, the HBV RNA was found in RNA particles (fraction 7) and in naked capsids (fractions 8 and 9) while the residual HBV DNA peak was found in virions (fraction 7). As compared to the Huh7-WT18 cell, a larger proportion of HBV RNA was found in naked capsids. This observation is consistent with the reduction of HBV DNA synthesis in the immature capsid imposed by the HBV mutations and the resulting blockade of capsids maturation. Regarding the profile of HBsAg, HBc and CD9, no differences were noticed as compared to the WT clone.

Next, to further characterize the RNA species secreted by the Huh7-WT18 and mutant Huh7-3D29 clone we have performed single molecule long reads Nanopore sequencing. The sequencing results show that: a) the secreted HBV RNA is identical to the HBV genome used for the establishment of the different cell lines; and b) the majority of HBV RNAs detected are the full-length transcript and pgRNA derived spliced RNAs whereas only a minority of PreS/S RNAs, PreS/S derived spliced RNA and HBx RNAs were detected (Fig. 5a).

Finally, 5'RACE reactions (Stadelmayer et al., 2020) were performed on: a) cell extracts from PLC/PRF/5 and Hep3B cells, known to express mostly S RNAs and, possibly, in the case of the Hep3B, HBx transcripts; b) cell extracts and the cell supernatants of Huh7-WT18 and Huh7-3D29 cell lines as well as of tet-off induced HepAD38 cells as a positive control. As expected, we detected S transcripts in PLC/PRF/5 and Hep3B cells and HBx transcripts in Hep3B cells (Fig. 5b). As shown in Figure 5c (right panel), 3.5 kb RNA species were detected only in the supernatant of Huh7-3D29 cells, thus confirming the Nanopore sequencing results.

Altogether, the Nanopore and the 5' RACE results support the use of the Huh7-3D29 RNA standard for the 2 research use only HBV RNA investigational assays (Butler et al., 2018; Scholtes et al, 2022) as well as most in house PCR-based HBV RNA assays.

Discussion

Several clonal cell lines transduced with HBV mutated genomes were generated to produce large quantities of secreted viral RNAs and develop an RNA standard to be used to calibrate PCR-based circulating HBV RNA quantification assays. The best clonal cell line displaying the desired RNA secretory phenotype (e.g., inversion of the secreted HBV DNA/RNA ratio) carried a double mutation in the catalytic site of the polymerase (YMAA mutation) and the TP-domain (Y63F) of the polymerase. Although, HBV RNA secretion slightly decreased after several freeze-thawing and amplification cycles, Huh7-3D29 cells maintained their RNA secretory phenotype with a high RNA/DNA ratio in cell supernatants and minimal residual DNA (~5 $\log_{10}$). If one considers the performing characteristics of the Roche MWR used in this study (or the related HBV RNA cobas® 6800/8800 automated investigational assay) that tolerates ~10^6 DNA:RNA ratios without losing specificity and linearity, the HBV RNA standard produced and secreted by the Huh7-3D29 cell line fullfill this criterion. Notably, HBV RNA quantification of Huh7-3D29 supernatants with the Roche MWF showed no significant difference with and without DNase treatment, supporting the use of the Huh7-3D29 derived HBV RNA standard with both automated and manual assays. Additional observations validate the use of Huh7-3D29 cell supernatants as an HBV RNA standard. First, the quantification of HBV RNAs with the Roche MWF of standard curves obtained from 10-fold dilution of Huh7-3D29 supernatants and a synthetic HBV armored RNA showed the same profile. Second, the CTs obtained with both standards showed no significant differences. Third, the RNA production and secretion in Huh7-3D29 cells show a good degree of concordance to what observed *in vivo* in chronic HBV carriers (Kim et al., 2022).

The results of iodixanol/sucrose gradients indicate that the majority of secreted HBV RNAs are encapsidated and are found in naked capsid and in virions-like particle. Single molecule long reads Nanopore sequencing indicates that the majority of the Huh7-3D29 secreted HBV RNAs are full length or spliced 3.5 kb species and most, if not all, RNAs terminate at the canonical HBV polyA site. These results support the potential use of Huh7-3D29 supernatants as a standard for all HBV RNA investigational assays and *in house* PCR-based HBV RNA assays.

According to our results, the HBV RNA quantities produced from one 175cm^2 flask of Huh7-3D29 clonal cell line in 9 days provides enough material for 1,300 standard curves using 20 µL of 10^6 copies/mL preparations. Therefore, the Huh7-3D29 cell line constitutes a potentially unlimited source of HBV RNA standard. The decrease in HBV RNA secretion we observed after several passages, even if the inverted DNA/RNA ratio is well conserved, may raise a concern about the long-term

stability of the RNA secretory phenotype. To overcome this potential limitation, we have *stored* over 20 very early passage aliquots frozen and conserved in liquid nitrogen. Moreover, each Huh7-3D29 HBV RNA standard production can be scaled up. As a comparison, the HBV Nucleic Acids WHO international standard produced consisted of 2000 aliquots of 0.5 mL with a concentration of 1×10^6 IU/mL and the second, third and fourth HBV WHO standard productions were in the same range (Baylis et al., 2008; Fryer et al., 2017; Saldanha et al., 2001).

The requirement for a WHO international standard is a lyophilized preparation that has proven long-term stability and which is available in adequate quantities (to last several years) (Saldanha et al., 2001). Until now, all viral nucleic acids WHO IS were obtained from serum preparation. Therefore, the production of an international HBV RNA standard from a clonal cell line would be a novelty and would guarantee a sustained access. Next steps for developing the Huh7-3D29 HBV RNA standard into an internationally established standard or a WHO endorsed standard Huh7-3D29 material will include i) the lyophilization of the standard material,ii) the evaluation of its performance by a large panel of laboratories, iii) the analysis of its performance by the independent expert committee on biological standardization of the WHO (Organization and Standardization, 2016).

Legends to Figures

Figure 1. Strategies to generate an HBV RNA standard

(A) Strategies to produce an HBV RNA standard. (B) HBV mutations described in the literature and their impact on HBV replication. (C) Mutations selected for the generation of HBV RNA secreting cell lines and their described effects on pgRNA encapsidation, rcDNA production and viral replication.

Figure 2. DNA and RNA from WT and mutant Huh7 derived clonal cell lines: impact of freezing/thawing and amplification. ddPCR quantification of HBV secreted nucleic acid from Huh7 clonal cell line at different cell passages and freezing/thawing cycles. (A) Passages 7-8 (according to the cell line); (B) Passages 10-11 (according to the cell line); (c) Passages 13-15 (according to the cell line).

Figure 3. 3D29 HBV RNA standard validation using the HBV RNA Roche MWF IA

(A) Quantification of ten folds dilution (from 10^5 to 10^1 copies/mL) of HBV armored RNA (arRNA, Roche Diagnostics, Pleasanton, CA) (left panel (i)) and Huh7-3D29 supernatants (right panel (iii)), processed as described in Materials and Methods, using the HBV RNA Roche MWF kit. Ct values and corresponding concentration in copies/mL are given in the middle panel (ii). No significant difference was found between CT of 10-fold diluted armored RNA and Huh7-3D29 RNA using a paired t test (P-value = 0.54). Validation of samples as parametric by Shapiro-Wilk test. (B) 3D29 HBV RNA quantification with and without DNAse treatment on 3 different passages of 3D29. No significant difference was found between DNAse treatment and no DNAse treatment using a Wilcoxon test (P-value = 0.25). Validation of samples as nonparametric by Shapiro-Wilk test.

Figure 4. HBV DNA and HBV RNA compartimentalization in Huh7-3D29 cell supernatant.

Analysis of Huh7-WT18 (A), Huh7-3D35 (B), and Huh7-3D29 (C) cells supernatants by iodixanol/sucrose density gradient ultracentrifugation. (i) HBsAg quantification by ELISA, HBV DNA, and total HBV RNA quantification by ddPCR in gradient fractions. (ii) Western Blot detection of viral HBc protein and CD9 exosome marker.

Figure 5 – Characterization of HBV RNAs from Huh7-3D29 cell culture supernatant. (A) cDNA-PCR Nanopore sequencing of HBV RNAs secreted in Huh7-WT18 and mutant Huh7-3D29 supernatants. The majority of HBV RNAs detected are the full length transcript and the pgRNA derived spliced RNAs

and only a minority of PreS/S RNAs, PreS/S derived spliced and HBx RNAs were detected. (B) 5'RACE assay on PLC/PRF/5 and Hep3B, known as cell line producing mostly S RNAs and possibly HBx in the case of the Hep3B, confirming the specificity of the technique. (C) 5'RACE performed on cell fractions and supernatants of HepAD38 (Positive Control), Huh7-WT18 and Huh7-3D29. Right panel: 3.5 kb RNA species detection in the supernatant of Huh7-3D29 cells. MW, molecular weight. Note that all RNAs are smaller by 500bp due to the HBV specific reverse primer localization.

Funding

This work was supported by a public grant overseen by the French National Research Agency (ANR) as part of the second "Investissements d'Avenir" program (reference: ANR-17-RHUS-0003).

Conflict of interest statement

None declared.

References

Baylis, S.A., Heath, A.B., Chudy, M., Pisani, G., Klotz, A., Kerby, S., and Gerlich, W. (2008). An international collaborative study to establish the 2nd World Health Organization International Standard for hepatitis B virus DNA nucleic acid amplification technology-based assays. Vox Sanguinis 94, 358–362.

Butler, E.K., Gersch, J., mcnamara, A., Luk, K.-C., Holzmayer, V., de Medina, M., Schiff, E., Kuhns, M., and Cloherty, G.A. (2018). Hepatitis B Virus Serum DNA andrna Levels in Nucleos(t)ide Analog-Treated or Untreated Patients During Chronic and Acute Infection. Hepatology 68, 2106–2117.

Charre, C., Levrero, M., Zoulim, F., and Scholtès, C. (2019). Non-invasive biomarkers for chronic hepatitis B virus infection management. Antiviral Res. *169*, 104553.

Carey, I., Gersch, J., Wang, B., Moigboi, C., Kuhns, M., Cloherty, G., Dusheiko, G., and Agarwal, K. (2019). Pre-genomic HBV RNA and hbcrag predict outcomes in hbeag negative chronic hepatitis B patients suppressed on nucleos(t)ide analogue therapy. Hepatology.

Clark, D.N., Flanagan, J.M., and Hu, J. (2017). Mapping of Functional Subdomains in the Terminal Protein Domain of Hepatitis B Virus Polymerase. J. Virol. 91.

Kim D, Bousquet D, Plissonnier ML, Tak H, Grand X , Goldsmith C, Berby F, Bordes I, Paturel A, Hamilton A, Heil M, Levrero M, Testoni B and Zoulim F (2022) Circulating Hepatitis B Virus (HBV)-RNA associates with extracellular vesicles in supernatant of HBV-infected hepatocytes and serum of chronically infected patients (manuscript in preparation)

Fan, R., Peng, J., Xie, Q., Tan, D., Xu, M., Niu, J., Wang, H., Ren, H., Chen, X., Wang, M., et al. (2020). Combining HBV RNA and hepatitis B core-related antigen: guidance for safely stopping nucleos(t)ide analogues in hbeag-positive chronic hepatitis B patients. J. Infect. Dis.

Fanning, G.C., Zoulim, F., Hou, J., and Bertoletti, A. (2019). Therapeutic strategies for hepatitis B virus infection: towards a cure. Nat. Rev. Drug Discov. *18*, 827–844.

Feng, H., Chen, P., Zhao, F., Nassal, M., and Hu, K. (2013). Evidence for Multiple Distinct Interactions between Hepatitis B Virus P Protein and Its Cognate RNA Encapsidation Signal during Initiation of Reverse Transcription. PLOS ONE 8, e72798.

Fryer, J.F., Heath, A.B., Wilkinson, D.E., Minor, P.D., and Collaborative Study Group (2017). A collaborative study to establish the 3rd WHO International Standard for hepatitis B virus for nucleic acid amplification techniques. Biologicals 46, 57–63.

'Global Hepatitis Report, 2017'. n.d. Accessed 21 February 2022. https://www.who.int/publications-detail-redirect/global-hepatitis-report-2017.

Giersch K, Allweiss L, Volz T, Dandri M, Lütgehetmann M. Serum HBV pgRNA as a clinical marker for cccDNA activity. J Hepatol. 2017 Feb;66(2):460-462.

Hatakeyama, T., Noguchi, C., Hiraga, N., Mori, N., Tsuge, M., Imamura, M., Takahashi, S., Kawakami, Y., Fujimoto, Y., Ochi, H., et al. (2007). Serum HBV RNA is a predictor of early emergence of the YMDD mutant in patients treated with lamivudine. Hepatology 45, 1179–1186.

Heermann, K.-H., Gerlich, W.H., Chudy, M., Schaefer, S., and Thomssen, R. (1999). Quantitative Detection of Hepatitis B Virus DNA in Two International Reference Plasma Preparations. J Clin Microbiol 37, 68–73.

Jenkins, A., Minhas, R., Morris, C., and Berry, N. (2017). Complete Genome Sequence of the WHO International Standard for Hepatitis B Virus DNA. Genome Announc. 5.

Klymiuk, M.C., Balz, N., Elashry, M.I., Heimann, M., Wenisch, S., and Arnhold, S. (2019). Exosomes isolation and identification from equine mesenchymal stem cells. BMC Veterinary Research 15, 42.

Lai, C.-L., Dienstag, J., Schiff, E., Leung, N.W.Y., Atkins, M., Hunt, C., Brown, N., Woessner, M., Boehme, R., and Condreay, L. (2003). Prevalence and clinical correlates of YMDD variants during lamivudine therapy for patients with chronic hepatitis B. Clin. Infect. Dis. 36, 687–696.

Lampertico, P., Agarwal, K., Berg, T., Buti, M., Janssen, H.L.A., Papatheodoridis, G., Zoulim, F., and Tacke, F. (2017). EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. J. Hepatol. *67*, 370–398.

Lanford, R.E., Notvall, L., Lee, H., and Beames, B. (1997). Transcomplementation of nucleotide priming and reverse transcription between independently expressed TP and RT domains of the hepatitis B virus reverse transcriptase. J. Virol. 71, 2996–3004.

Liu S, Zhou B, Valdes JD, Sun J, Guo H. Serum Hepatitis B Virus RNA: A New Potential Biomarker for Chronic Hepatitis B Virus Infection. Hepatology. 2019 Apr;69(4):1816-1827.

Liu, S., Deng, R., Zhou, B., and Sun, J. (2020). Is World Health Organization HBV DNA Standard Appropriate for Standardizing Serum HBV RNA Assay? Clin Infect Dis.

Ma, G., Lou, B., Lv, F., Zhao, D., Zhang, Z., and Chen, Y. (2020). Hbcrag and pgrna and the therapeutic effect inhbeag-positive patients receiving anti-viral therapy, baseline serum HBV-rnais a powerful predictor of response. J. Viral Hepat.

Nassal, M. (2015). HBV cccDNA: viral persistence reservoir and key obstacle for a cure of chronic hepatitis B. Gut *64*, 1972–1984.

Ono-Nita, S.K., Kato, N., Shiratori, Y., Masaki, T., Lan, K.H., Carrilho, F.J., and Omata, M. (1999). YMDD motif in hepatitis B virus DNA polymerase influences on replication and lamivudine resistance: A study by in vitro full-length viral DNA transfection. Hepatology 29, 939–945.

Organization, W.H., and Standardization, W.E.C. on B. (2016). Collaborative study to evaluate the proposed WHO 4th International Standard for Hepatitis B Virus (HBV) DNA for Nucleic Acid Amplification Technique (NAT) based assays.

Saldanha, J., Gerlich, W., Lelie, N., Dawson, P., Heermann, K., Heath, A., and WHO Collaborative Study Group (2001). An international collaborative study to establish a World Health Organization international standard for hepatitis B virus DNA nucleic acid amplification techniques. Vox Sang 80, 63–71.

Scholtès C, Hamilton TA, Plissonnier ML, Charre C, Scott B, Wang L, Berby L, French J, Testoni B, Blair. A, Subic M, Hoppler M, Lankenau A, Grubenmann A, Levrero M, Heil ML, Zoulim F. (2022) Performance of the cobas® HBV RNA Automated Investigational Assay for the Detection and Quantification of Circulating HBV RNA in Chronic HBV Patients. J Clin Virol (in press)

Stadelmayer, B., Diederichs, A., Chapus, F., Rivoire, M., Neveu, G., Alam, A., Fraisse, L., Carter, K., Testoni, B., and Zoulim, F. (2020). Full-length 5'RACE identifies all major HBV transcripts in HBV-infected hepatocytes and patient serum. J. Hepatol. *73*, 40–51.

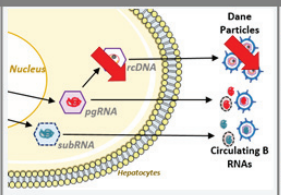
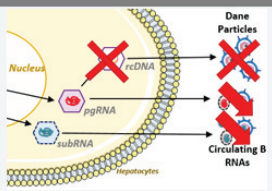
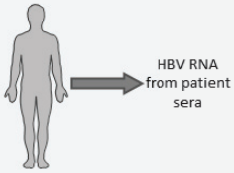
Wang, J., Shen, T., Huang, X., Kumar, G.R., Chen, X., Zeng, Z., Zhang, R., Chen, R., Li, T., Zhang, T., et al. (2016). Serum hepatitis B virus RNA is encapsidated pregenome RNA that may be associated with persistence of viral infection and rebound. J. Hepatol. 65, 700–710.

Wooddell, C.I., Yuen, M.-F., Chan, H.L.-Y., Gish, R.G., Locarnini, S.A., Chavez, D., Ferrari, C., Given, B.D., Hamilton, J., Kanner, S.B., et al. (2017). Rnai-based treatment of chronically infected patients and chimpanzees reveals that integrated hepatitis B virus DNA is a source of hbsag. Sci Transl Med 9.

Zhu, Y., Curtis, M., and Borroto-Esoda, K. (2011). The YMDD and rta194t mutations result in decreased replication capacity in wild-type HBV as well as in HBV with precore and basal core promoter mutations. Antivir. Chem. Chemother. 22, 13–22.

Zoulim, F., and Seeger, C. (1994). Reverse transcription in hepatitis B viruses is primed by a tyrosine residue of the polymerase. J. Virol. 68, 6–13.

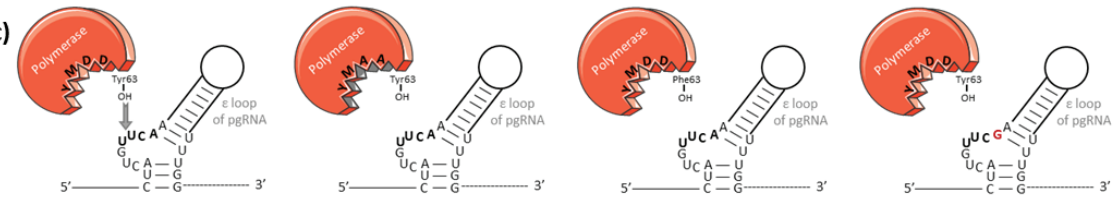
(A)

	HBV Pol ORF mutation	ATG mutation HBV Pol ORF	Synthetic RNA	Standard derived from patients sera
Scheme				
Presence of Pol protein	+	-	-	-
Encapsidation	+	-	-	?
Viral DNA	↘	-	-	+
pgRNA	+	+	+	+
Subgenomic RNAs	+/-	+/-	-	?
Production	*unlimited*	*unlimited*	*limited*	*limited*
Cost	+	+	+++	+

(B)

	Mutations	Reduction in virus DNA replication	Genotype	Encapsidation	Reference
Pol (Rnase H domain)	**R703A**	**N.I**	**D**	**+**	**Ko et al (JVI_2013)**
	D777A	**N.I**	**D**	**+**	**Ko et al (JVI_2013)**
	R781A	**-/+**	**D**	**+**	**Ko et al (JVI_2013)**
Pol (RT domain)	**rtA194T**	**40% reduction**	**D**	**N.I**	**Y Zhu et al. (SAGE journals_2011)**
	rtL180M + rtM204V	**>75% reduction**	**D**	**N.I**	**Y Zhu et al. (SAGE journals_2011)**
	rtM204I	**> 75% reduction**	**D**	**N.I**	**Y Zhu et al. (SAGE journals_2011)**
	rtL180M + rtA194T + rtM204V	**> 80% reduction**	**D**	**N.I**	**Y Zhu et al. (SAGE journals_2011)**
	rtI187V	**20% reduction**	**adr serotype**	**N.I**	**Jiyun Fan et al (JGV_2014)**
	M204I + I187V	**90% reduction**	**adr serotype**	**N.I**	**Jiyun Fan et al (JGV_2014)**
	rtL180M + rtM204V + rtI187V	**N.I**	**adr serotype**	**N.I**	**Jiyun Fan et al (JGV_2014)**
	YMAA	**severe defect in RT**	**HIV**	**N.I**	**Mulky et al (journal of virology_2004)**
Pol (TP domain)	**Y63F**	**N.I**	**D**	**+**	**Wang et al., Journal of Hepatology, 2016; Wang et al (PNAS_2013)**
HBc	**V124A**	**N.I**	**C**	**+**	**Wang et al (PNAS_2013)**
	V124M	**N.I**	**C**	**+**	**Wang et al (medical virology_2016)**
	V124G	**N.I**	**C**	**+**	**Wang et al (medical virology_2016)**
	V124E	**N.I**	**C**	**+**	**Wang et al (medical virology_2016)**
	T109N	**N.I**	**C**	**+**	**Wang et al (medical virology_2016)**
Epsilon domain of pgRNA	**A1G**	**> 90% reduction**	**N.I**	**+**	**Hui Feng et al (plos one _2013)**

(c)



WT	YMAA	Y63F	A1G
DNA accumulation: 100%	DNA accumulation: ↘	DNA accumulation: <18%	DNA accumulation: <10%
pgRNA encapsidation: 100%	pgRNA encapsidation: ≈	pgRNA encapsidation: ≈100%	pgRNA encapsidation: ≈100%
In vitro priming: 100%	In vitro priming: ↘	In vitro priming:↘	In vitro priming: 20%

Figure 1

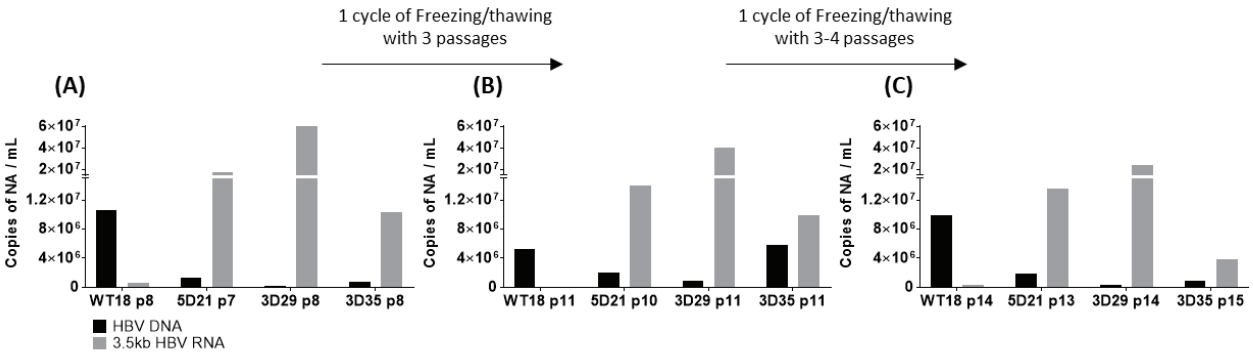
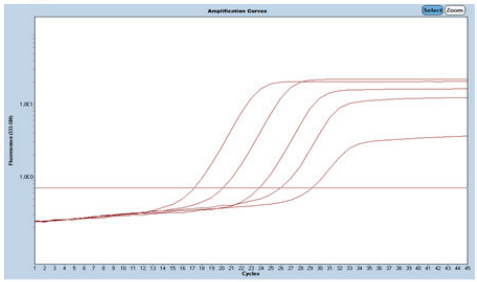

1 cycle of Freezing/thawing
with 3 passages
1 cycle of Freezing/thawing
with 3-4 passages
(A)
(B)
(C)
Copies of NA / mL
6×10^7
4×10^7
2×10^7
1.2×10^7
8×10^6
4×10^6
0
WT18 p8
5D21 p7
3D29 p8
3D35 p8
WT18 p11
5D21 p10
3D29 p11
3D35 p11
WT18 p14
5D21 p13
3D29 p14
3D35 p15
HBV DNA
3.5kb HBV RNA

Figure 2

A

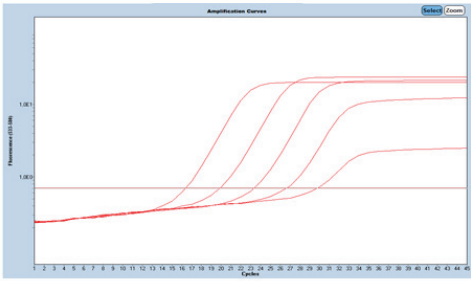
(i) Armored RNA (Roche)
10 fold dilution amplification
curves from MWF



(ii)

Ct of diluted armored RNA	Copies / mL	Ct of diluted 3D29 RNA
19,52	10^5	18,70
22,59	10^4	22,80
26,44	10^3	26,07
29,38	10^2	29,96
32,05	10^1	34,09

(iii) 3D29 HBV RNA
10 fold dilution amplification
curves from MWF



B

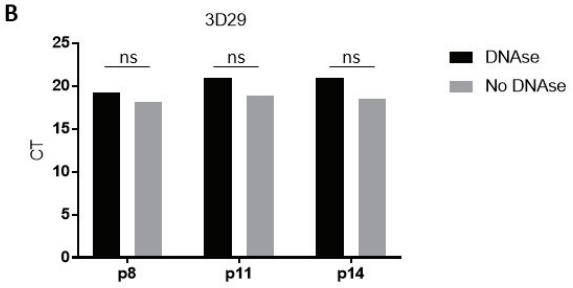


Figure 3

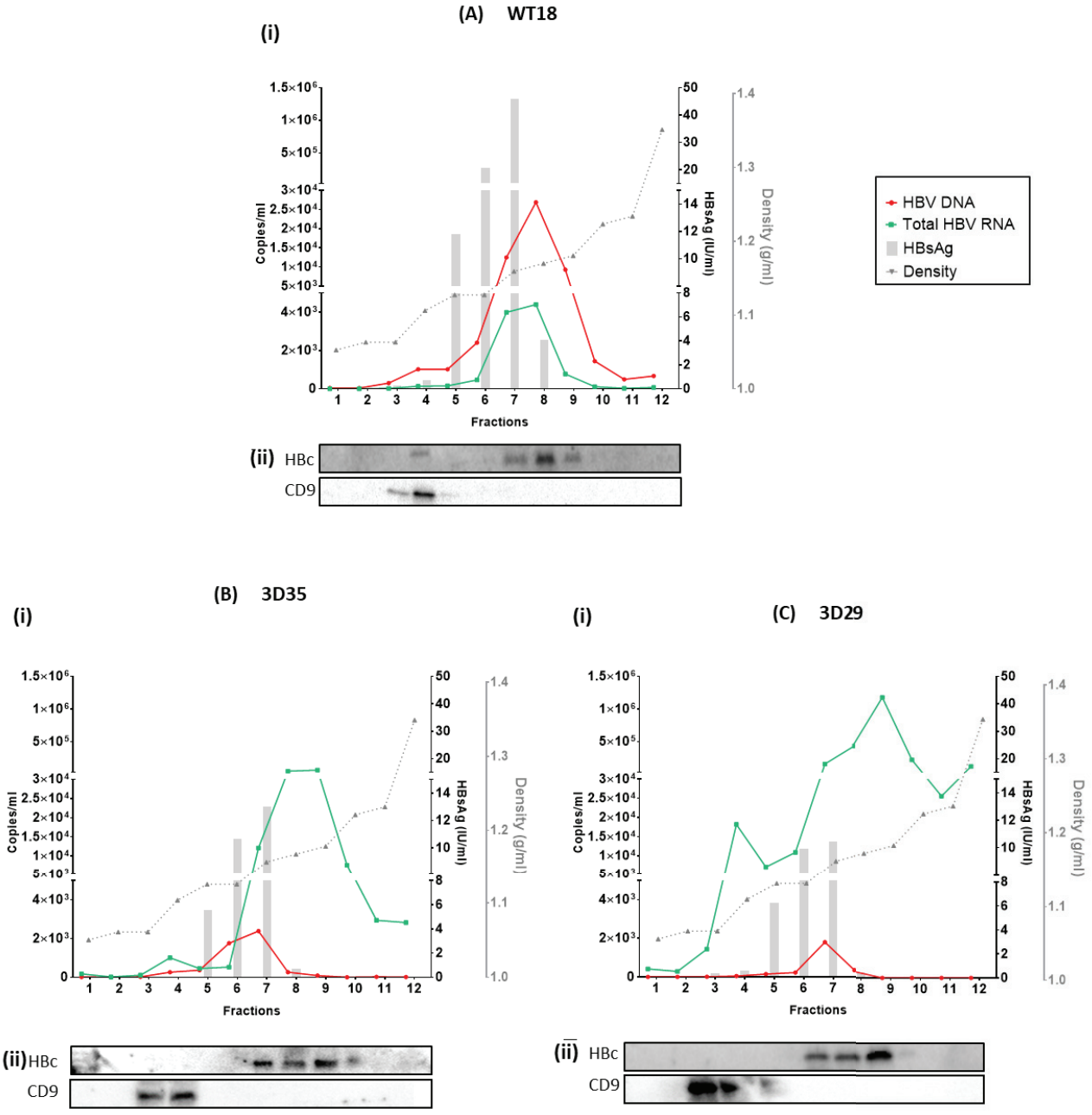

(A) WT18
(i)
Copies/ml
HBsAg (IU/ml)
Density (g/ml)
Fractions
HBV DNA
Total HBV RNA
HBsAg
Density
(ii) HBc
CD9
(B) 3D35
(i)
(ii) HBc
CD9
(C) 3D29
(i)
(ii) HBc
CD9

Figure 4

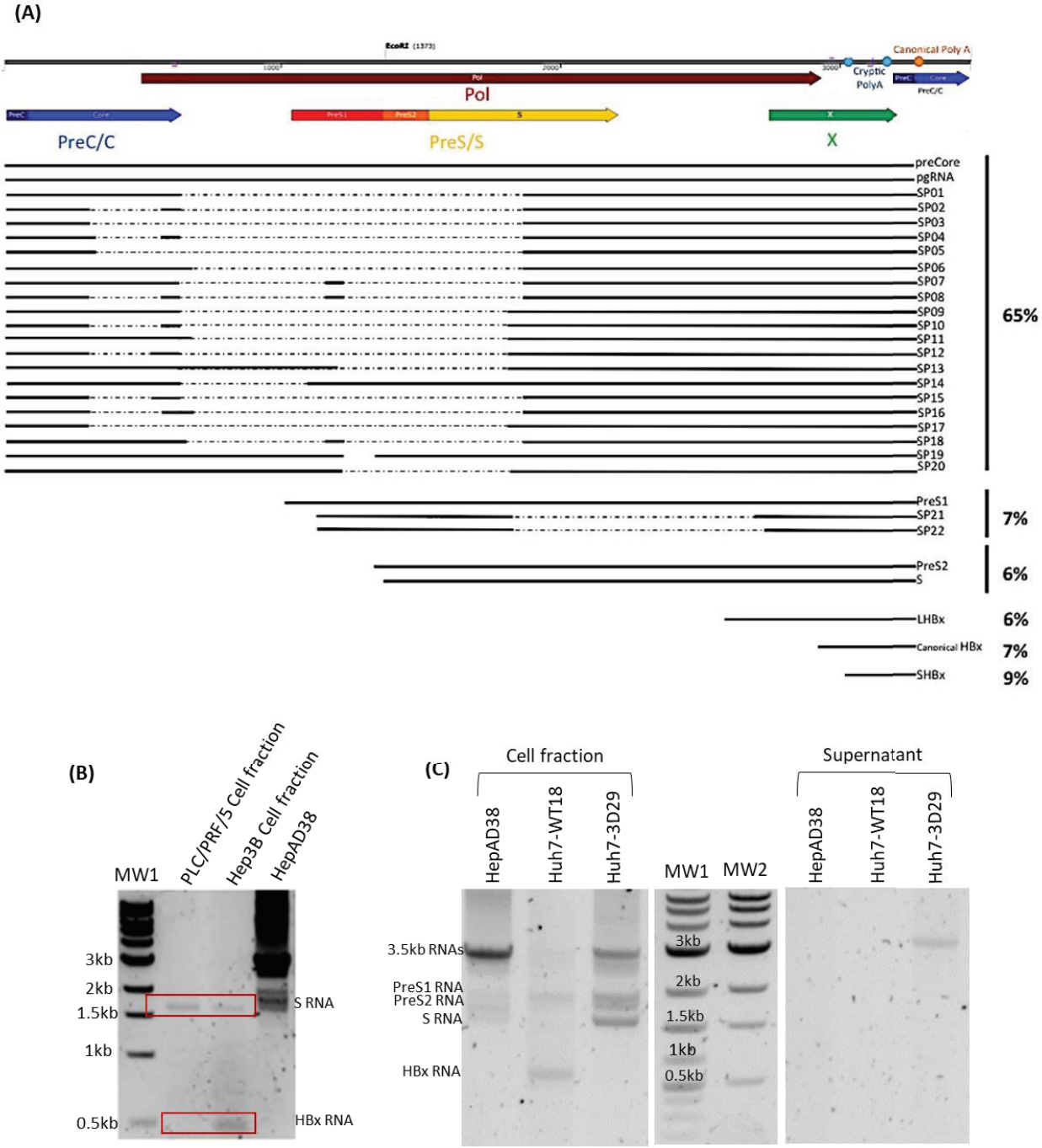
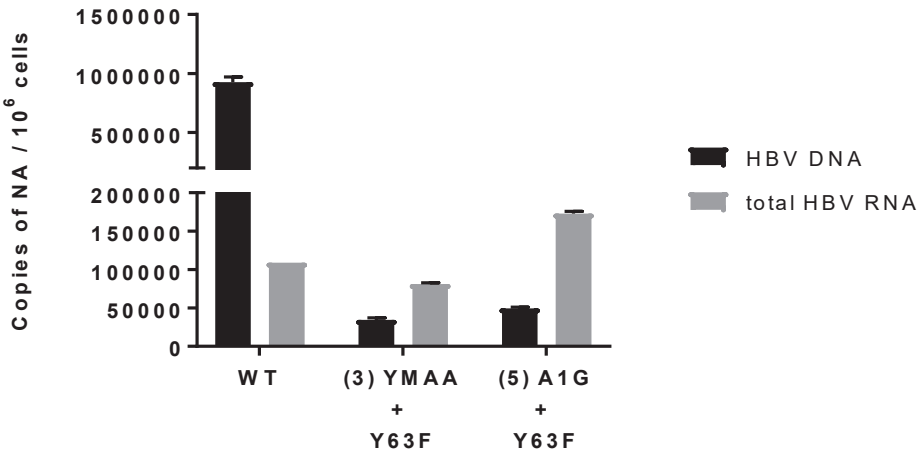
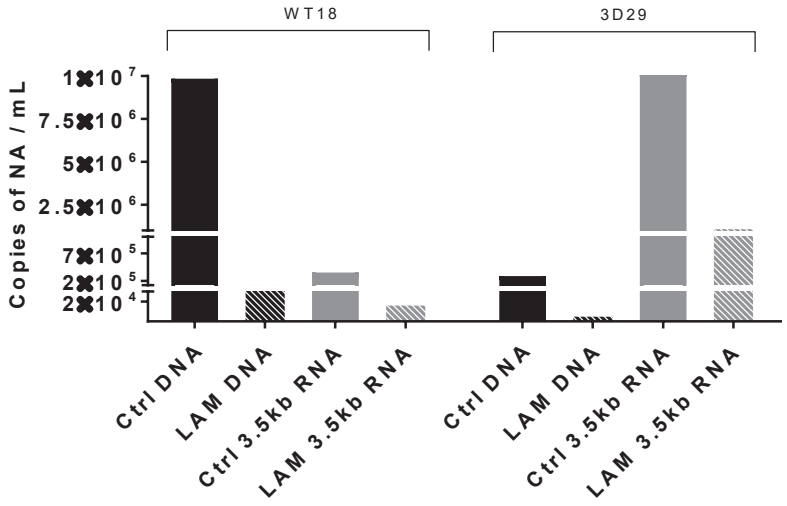

(A)
EcoRI (1373)
Canonical Poly A
Cryptic PolyA
Pol
PreC/C
PreS/S
X
preCore
pgRNA
SP01
SP02
SP03
SP04
SP05
SP06
SP07
SP08
SP09
SP10
SP11
SP12
SP13
SP14
SP15
SP16
SP17
SP18
SP19
SP20
65%
PreS1
SP21
SP22
7%
PreS2
S
6%
LHBx
6%
Canonical HBx
7%
SHBx
9%
(B)
MW1
PLC/PRF/5 Cell fraction
Hep3B Cell fraction
HepAD38
3kb
2kb
1.5kb
1kb
0.5kb
S RNA
HBx RNA
(C)
Cell fraction
Supernatant
HepAD38
Huh7-WT18
Huh7-3D29
MW1
MW2
3.5kb RNAs
PreS1 RNA
PreS2 RNA
S RNA
HBx RNA
3kb
2kb
1.5kb
1kb
0.5kb

Figure 5



Supplementary Figure 1 – Phenotypic characterization of the Huh7 polyclonal.



Supplementary Figure 2 – Study of HBV residual DNA under lamivudine treatment

Generation of a capture-high throughput DNA and RNA sequencing workflow to study HBV integration and its impact on circulating HBV RNAs in CHB patients

Generation of a capture high throughput viral DNA and RNA sequencing workflow for the study of HBV integration impact on circulating HBV RNA from CHB patients

The following people have contributed

- to design the study: Francesca GUERRIERI[1], Massimo LEVRERO[1,2,3]

- to generate and / or analyze the results: Alexia PATUREL[1], Francesca CASUSCELLI DI TOCCO[1], Silvano PIAZZA, Alfonso ESPOSITO

Affiliations

[1] Cancer Research Center of Lyon (CRCL), UMR Inserm U1052 / CNRS 5286, Lyon, France
[2] University of Lyon, University Claude Bernard Lyon 1, 69008 Lyon, France
[3] Department of Hepatology, Croix Rousse hospital, Hospices Civils de Lyon, France
[4] Department of Internal Medicine, SCIAC and the IIT Center for Life Nanoscience, Sapienza University, Rome, Italy
[5] Department of Computational Biology, International Centre for Genetic Engineering and Biotechnology (ICGEB) Padriciano, Trieste, Italy

Introduction

HBV integration into the host cell genome, an early event in the natural history of HBV infection and disease, does not contribute to viral replication but has several pathologic consequences that include a direct (insertional mutagenesis) or indirect (genetic instability) contribution to HCC development, the expression of truncated HBV proteins (PreS2/S and HBx) with new pathogenic properties, the generation of virus-host chimeric neo-epitopes that can be seen by the host immune system. HBV integration can also lead to the expression and secretion of virus-host chimeric transcripts.

Southern Blot, direct cloning and Sanger sequencing, Alu-PCR and inverted nested PCR were the initial techniques to study HBV integration. However, these techniques are relatively insensitive. Next Generation Sequencing has become the technology of reference to detect and characterize HBV DNA integrations. The rarity of integration events in non tumoral samples (~1 per 10,000 cells) and the need for relatively large amounts of starting material prompted the efforts to develop targeted sequencing protocols based on HBV capture-enrichment.

In this section we describe the development of an HBV capture-high throughput sequencing workflow for the detection of DNA and RNA virus/host junctions and chimeric virus/host RNAs.

Materials and methods

Cells

The human Hep3B and PRF/PLC/5 hepatocarcinoma cell lines were cultured at 37°C in a humidified atmosphere containing 5% CO2 in Dulbecco's modified Eagle medium supplemented with 10% FBS (Gibco) , 1% Penicillin / Streptomycin (Gibco), 1% sodium pyruvate (Gibco), 1% non-essential amino acids (NEAA, Gibco). The HepG2 derived tet-off HepAD38 cell line was a kind gift of Dr C. Seeger (Fox Chase Cancer Center, Philadelphia, USA).

Nucleic acid extraction

DNA and RNA from HepG2, Hep3B and PLC/PRF/5 hepatoma cell lines and from the non tumoral liver tissue of an HBV patient with HBV-related HCC who underwent liver resection, were extracted using the Qiagen RNeasy mini kit, as per manufacturer's instructions (Qiagen, Valencia, CA) for RNA and phenol chloroform extraction for DNA. The eluted DNA or RNA (30 µl) from each condition was stored at -80°C until further use. Quality of nucleic acids was determined by Fragment Analyzer (Agilent) or Bioanalyzer (Agilent, Santa Clara, CA, USA).

HBV Capture

SeqCap EZ Hypercap tiling probes (Roche), designed to cover the entire HBV genome of 340 HBV sequences. 1µg of nucleic acid was used to prepare libraries with the HyperPrep for RNA or HyperPlus for DNA. The quality and quantity of the amplified libraries before capture were determined using the Fragment Analyzer (Agilent, Santa Clara, CA, USA) and Qubit (Thermo Fisher, Waltham, MA, USA) respectively. For capturing, 2µg of the amplified DNA libraries were combined or not according to whether common capture or independent capture protocol was performed. For enrichment of the DNA sample library pools, the SeqCap EZ HyperCap Workflow User's Guide (Roche) was followed with some *in-house* adaptations to the manufacturer's protocol. Briefly, human Cot DNA and blocking oligos (Integrated DNA Technologies) were added to each library pool to block non-specific cross hybridization. The target regions were captured by hybridizing each pool with the SeqCap EZ probe pool overnight. The HyperCap Target Enrichment kit and Hyber Cap Bead kit were used for washing and recovery of the captured DNA. Finally, post-capture PCR amplification was performed using KAPA HiFi HotStart ReadyMix (2X) and post-LM PCR primers (5 µM), followed by DNA purification using

AMPure XP beads. Quality and quantity of the post-capture multiplexed libraries were determined by Fragment Analyzer (Agilent) or Bioanalyzer (Agilent, Santa Clara, CA, USA).

Next Generation Sequencing (MiSeq Illumina) and bioinformatic analysis

High throughput sequencing was performed on an Illumina MiSeq instrument to produce paired-end reads of 2x150 nucleotides or 2×250 nucleotides. Sequence data were aligned on HBV and human references (hg38). Paired and chimeric reads were extracted using seekSV and virus-CLIP software to identify HBV/human junctions.

Raw.fastq files were quality filtered using trim_galore (10.5281/zenodo.5127899), no reads were discarded during this phase. To detect viral insertion events, we used the SeekSV pipeline (10.1093/bioinformatics/btw591). Reads were aligned to a hybrid genome including the human genome hg38 and the HBV genome (acc.no. U95551.1). The split reads (e.g., reads mapping in part on one chromosome and another part on a viral sequences and the discordantly mapping read pairs (i.e. forward mapping on one chromosome and reverse mapping on another chromosome) were extracted and re-mapped to obtain statistics about the number of reads supporting the insertion events. The raw tabular files generated by seekSV were further filtered to remove the background noise. An insertion was considered as such only if it was supported by at least 5 reads in both its left and right junction. The number of insertion found was plotted using R bioinformatic package, the positions on the viral genome was displayed using CGView (10.1093/nar/gkn179).

Reverse transcription (RT) and droplet digital PCR (ddPCR)

For HBV RNA analysis, 4 µL of DNase-treated RNA was reverse transcribed and amplified using the SuperScriptTM IV VILOTM Master Mix (Invitrogen, Cat # 11766500). A 22 µL reaction mixture was prepared comprising 11 µL of 2X ddPCR Supermix™ for probes (no dUTP) (Bio-Rad), 1.1 µL of primers and probe mix, and 5 µL of cDNA or DNA. Nucleic acid inputs were adjusted to have acceptable rates of negative events: samples were diluted 1:5. Probes and primers include: HBV (Pa03453406_s1, Thermofischer) for total HBV RNA quantification and pgRNA (Forward primer ggagtgtggattcgcactcct, reverse primer agattgagatcttctgcgac and probe aggcaggtcccctagaagaagaactcc) for pgRNA quantification. Droplet formation was carried out using a QX100 droplet generator. Subsequent amplifications were performed in the C1000 Touch™deep-well thermal cycler (Bio-Rad) with a ramp rate of 2 °C/s and the lid heated to 105 °C, according to the Bio-Rad recommendations. First, the enzyme was activated at 95°C for 10 min followed by 40 cycles of denaturation at 94°C for 30 s and

60°C for one minute. The enzyme was deactivated at 98°C for 10 min and the reaction was kept at 4°C.

Roche Manual Workflow (MWF)

The Roche Manual Workflow (MWF) is a manual version developed for research purposes of the cobas® 6800/8800 Automated Investigational Assay (IA) for the detection and quantification of circulating HBV RNAs in chronic HBV patients (Scholtes et al., 2022). MWF uses identical chemistry and PCR primers/probe sets as the cobas® 6800/8800 automated IA but does not require high throughput and automation. Results obtained with this manual assay were highly correlated with those obtained with cobas HBV RNA (Scholtes et al., 2022). HBV RNA assay characteristics are the following: i) linear results between 10 and 10^7 copies/mL in clinical samples of several HBV genotypes; ii) precision and reproducibility with standard deviation below 0.15 $\log_{10}$ copies/mL and coefficients of variation below 5% throughout the linear range; minimal impact (<0.3 $\log_{10}$ copies/mL) of HBV DNA on HBV RNA quantification at DNA:RNA ratios of up to approximately 10^6. HBV RNA quantification was calibrated using a synthetic armored RNA (arRNA) containing 435 bp derived from the 3' end of HBV pgRNA, packaged in MS2-phage (kindly provided by Roche Diagnostics, Pleasanton, CA) arRNA was quantified by ddPCR (BioRad) using primers and a probe in the precore/core region. HBV RNA quantification with the MWF was performed according to the manufacturer's protocol.

5'RACE

5'RACE was performed as previously described in (Stadelmayer et al., 2020). Briefly, RNAs were isolated using a guanidinium thiocyanate–phenol–chloroform extraction protocol (TRI reagent (Sigma)). 5'RACE was essentially performed as described in the GeneRacer Kit manual (ThermoFisher Scientific) except Tobacco Acid Pyrophosphatase was substituted by RNA 5' Pyrophosphohydrolase (New England Biolabs) and SuperScript reverse transcriptase III by SuperScript reverse transcriptase IV (ThermoFisher Scientific). The reverse transcription reaction was performed using 3' HBV specific Gsp1 primer 5'-TTAGGCAGAGGTGAAAAAAGTTG-3'. For the 5'RACE PCR reaction Prime Star super mix DNA Polymerase (TAKARA bio), GeneRacer 5' primer and HBV specific nested primer Gsp1 5'-TTAGGCAGAGGTGAAAAAAGTTG -3' were used. 5'RACE PCR was run in a C100 Touch thermocycler (Biorad) using the following PCR program: Initial denaturation step 98°C 3 min >5x (98°C 10 s; 72°C 3 min) >5x (98°C 10 s; 70°C 3 min) >25x (98°C 10 s; 64,4 20 s; 72°C 3min) >72°C 10min.

Results and Discussion

Development of an HBV capture-high throuput sequencing workflow for the detection of DNA and RNA virus/host junctions

The strategy has been to develop a workflow that allows to detect and quantify HBV sequences integrated into the patient genome and to determine whether the integrations are transcribed. Viral sequences from both replicating HBV and integrations are in smaller proportion than the human genome, and integrated viral sequences are even less numerous than virion or episomal viral sequences. Thus, it is imperative to set up a capture-sequencing method that can be applied to both DNA and RNA and that allows sufficient enrichment of the viral sequences to detect HBV integration.

The Nimblegene-Roche SeqCap EZ HyperCap method was chosen to capture HBV-host fusion sequences due to its high capture efficiency supported by the high density tiling of 105mer biotinylated oligonucleotides and to the ability to work for both RNA and DNA capture.

To optimize the capture of HBV sequences, high density tiling of 105mer biotinylated oligonucleotides covering the entire HBV genome of >340 HBV isolates and including all HBV genotypes and subtypes, have been used. To date, only one study has used Nimblegen capture technology to study hepatitis B virus. Capture probes were based on 40 HBV reference sequences and included all 8 HBV genotypes and 4 human regions (TERT promoter, CCNE1, CCNA2 and KMT2B). The small number of HBV references included might limit the detection of HBV integrations in specific genotypes. The usefulness of the inclusion of human regions in the capture as a positive technical control remains to be discussed as it could introduce a bias towards preferential capture of certain integrations.

The workflow of the Nimblegen-Roche SeqCap EZ HyperCap method includes sample extraction, DNAse treatment, adapters ligation, libraries preparation, target capture enrichment by hybridization to the specific probes, few cycles of PCR amplification followed by paired-end sequencing on a MiSeq Illumina platform. We essentially followed the manufacturer's protocol with some *in-house* modifications detailed in the Materials and Methods section.

Different capture protocols have been tested:

- an ***independent*** HBV sequence capture where the whole protocol is performed on each single sample
- a ***common*** capture where the enrichment of the HBV sequences is done on a pool containing multiple samples. Pooling of samples is performed after the generation of the libraries
- a ***double capture*** which is performed as in the common capture, but it is followed by a second capture step.

In addition to the three capture techniques, we also compared 2 sequencing protocols to evaluate the impact of the length of the paired-ends reads and total sequencing productivity (2x250bp reads and 2x150bp reads) on the detection of virus-host junctions.

To set-up and validate the capture-sequencing workflow, we analyzed the following samples:
- HepG2 cells (HBV-integrations negative control)
- Hep3B and PLC/PRF/5 HCC cell lines that carry integrated HBV sequences (positive controls)
- the non tumoral liver tissue from an HBV patient with HBV-related HCC who underwent liver resection.

PLC/PRF/5 cell line contains multiple HBV genotype A integrations. The reconstruction of integrated HBV sequences after HBV-capture followed by short reads high throughput sequencing (NCBI reference LC533934) identified 13 virus/human junctions in 8 chromosomes (ch 3, 4, 5, 8, 11, 12, 13 and 17) (Ishii et al., 2020). A second study used a 3rd generation long-reads sequencing technique and found 84 HBV fragments inserted in 7 chromosomes (Chen et al., 2021). A third study coupled HBV-capture with 3rd generation long-reads sequencing detected HBV integrations in 8 chromosomes (ch 1, 5, 8, 11, 12, 13, 16, 17) (Ramirez et al., 2021). Collectively, 10 integrations have been found but the 3 studies that used high-thoughput sequencing coupled with different sequencing technologies, preceded or not by an oligonucleotide capture step, did not detect all the 10 integrations. HBV integrations in chromosomes 1, 8, and 3 are not transcriptionally active whereas HBV integrated sequences can express S RNAs (ch 12, 13, 16 and 17) (Chen et al., 2021), PreS1/S2/S RNAs (ch 11) and PreS2/S RNAs (Ch 11, 13).

Hep3B have two copies of the HBV genotype A DNA integrated into the host genome (Twist et al., 1981). Targeted PacBio long read DNA sequencing confirmed the presence of 2 HBV insertions linked

with 2 translocations (chromosome 16 to chromosome Y and chromosome 4 to chromosome 13) (Ramirez et al., 2021). HBV insertion on chromosome 16/Y is transcriptionally silent whereas PreS2/S RNA is transcribed from chromosome 4/13 (Ramirez et al., 2021). Due to low amount of viral integration, Hep3B cells are suitable to test different techniques and confirm the efficiency of viral enrichment to detect and sequence integrated HBV.

A total of 15 independent capture-sequencing experiments have been performed. To ensure a high specificity of the detection of virus-host chimeric reads reflecting true virus host junctions, the analysis algorithm integrated by two distinct bioinformatics softwares (SeekSV and Virus-clip) (Ho et al., 2015; Liang et al., 2017) and only the chimeric reads validated by both analysis are selected (Fig. 1A). We then evaluated the impact of the 3 different HBV capture protocol and of Illumina reads length on the detection of virus-host chimeric reads using the data from all the capture-sequencing experiments performed. As shown in Fig. 1B, the number of HBV insertions reaches a plateau at 7 insertions for PLC/PRF/5 cells with a number of reads ranging from 10^5 to 10^7. Below this threshold, the number of detected insertions decreases. Longer paired-end reads (2x250bp) allowed the detection of more HBV-host junctions (data not shown). Since a sequencing run on a MiSeq instrument generates up to 15 Gb output data and 50 M paired-end reads and that the detection window for viral insertions is 10^5 to 10^7 reads per sample, the number of reads required to efficiently detect viral integrations is well within the operating capacity of the MiSeq and does not require the use of sequencers with higher productivity. Figure 1C compares the percentages of reads aligned to HBV that captured the 3 different enrichment methods. We found that, despite the enrichment method did not impact on the number of HBV insertions detected, 90% of the PLC/PRF/5 reads enriched by *double capture* are viral reads, whereas the alignment decreased to 22% for the PLC/PRF/5 reads enriched by *common capture* and to 14.5% for the PLC/PRF/5 reads enriched by *independent capture*. However, although the *double capture* allows to recover more HBV reads, the choice of the capture protocol did not impact on the number of HBV integrations detected. Thus, within the detection window for viral insertions of 10^5 to 10^7 reads per sample, the percentage of reads aligned to HBV reads does not seem to have a direct impact on the capacity to detect human-virus junctions.

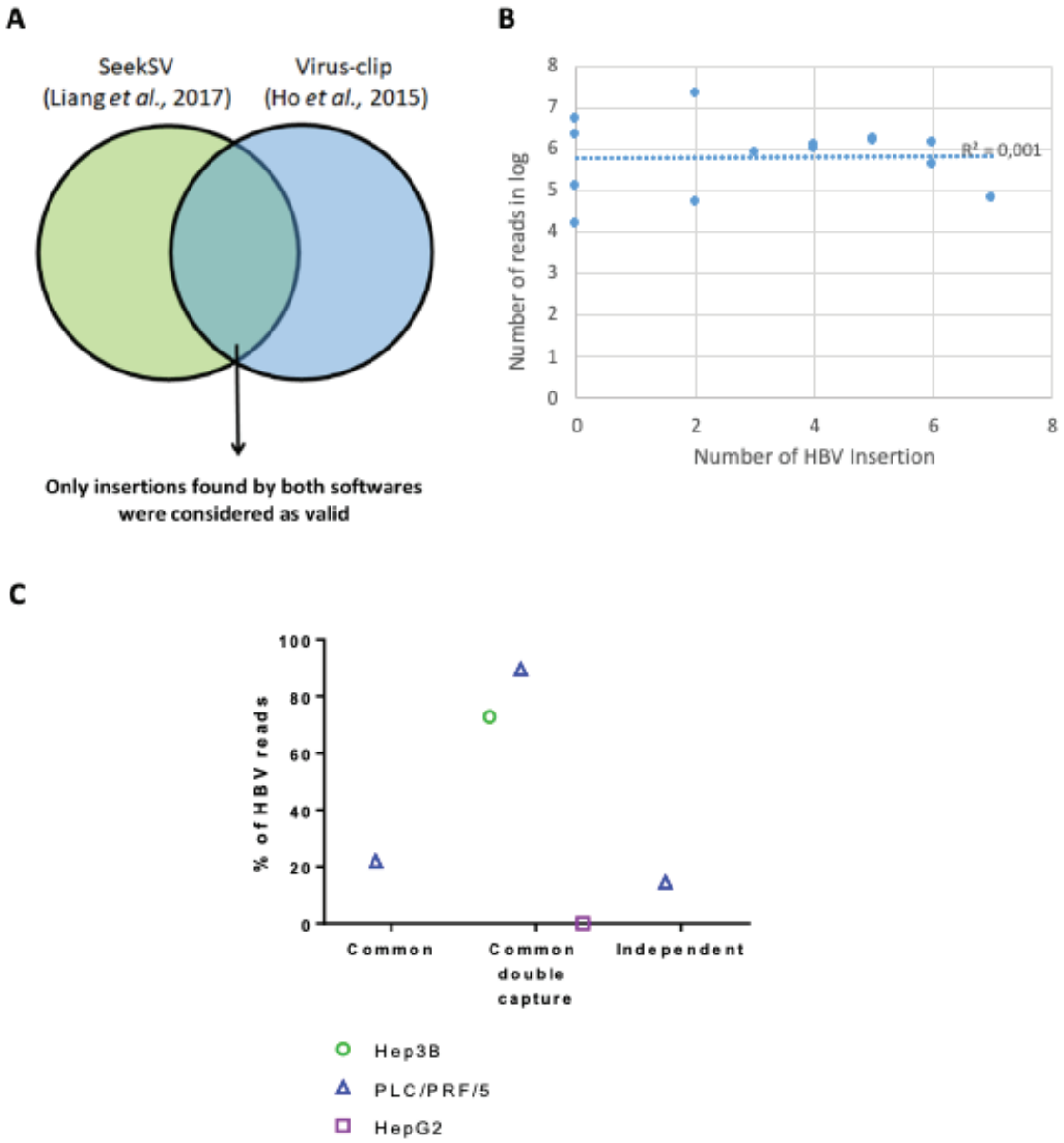


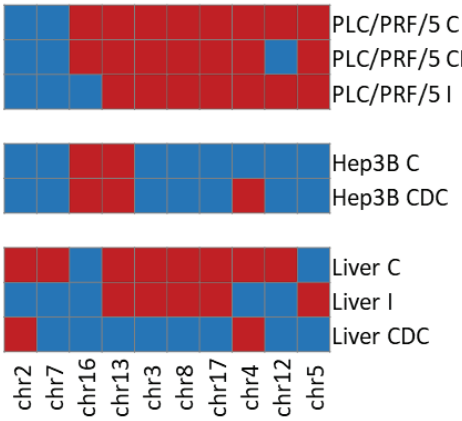
Figure 1 - Stratedies to develop a HBV capture workflow for the detection of circulating and intrahepatic HBV DNA and RNA

(A) Use of two software to build a data analysis pipeline with increased specificity for the detection chimeric virus-host reads from HBV integrations; (B) no correlation between the number of reads generated and the number of HBV insertions detected; (C) Percentage of reads aligned to HBV according to capture enrichment methods.

Figure 2A compares the number (and chromosomal location) of HBV integrations detected with the 3 capture-sequencing protocols. Between 7 and 8 HBV-host junctions have been detected in PLC/PRF/5 which is very close to what has been described in previous published data. As mentioned above, the studies that have sequenced PLC/PRF/5 cells HBV integrations using high NGS, a total of 10 HBV integrations have been identified. Using the *common capture* protocol we found 8 independent integrations, in between the 8 integrations identified by Chen et al. (2021) and the 9

integrations found by Ramirez et al. (2021) (Fig.2B). Knowing that these two studies were performed with the 3rd generation NGS long read PacBio sequencing while our results were obtained with an Illumina short read sequencing, we can validate the robustness of our method. The *double capture* and the *independent capture* protocols seem less efficient than the *common capture* as they allow the recovery of 7 HBV integrations. The difference between the different capture-sequencing protocols became more obvious when we analyzed the liver resection sample where the *double capture* identifies only 2 integrations, and the *independent capture* detects 5 integrations against the 8 integrations detected with the *common capture*.

A



B

	PLC/PRF/5	Chr 1	Chr 3	Chr 4	Chr 5	Chr 8	Chr 11	Chr 12	Chr 13	Chr 16	Chr 17
Literature results	**Ishii et al** HBV capture + Short Read Sequencing	X	✓ NA	✓ NA	✓ TERT	✓ UNC5D	✓ SAC3D1/SNX15	✓ MVK	✓ NA	X	✓ CCDC57
	Chen et al Long Read Sequencing	X	✓ NA	✓ NA	X	✓ UNC5D	X	✓ MVK	✓ N4BP2L2	✓ RP11-296	✓ CCDC57
	Ramirez et al HBV capture + Long Read Sequencing	✓	✓	X	✓	✓	✓	✓ MVK	✓	✓	✓ CCDC57
Lab results	C	X	✓	✓	✓	✓	X	✓	✓	✓	✓
	CDC	X	✓	✓	✓	✓	X	X	✓	✓	✓
	I	X	✓	✓	✓	✓	X	✓	✓	X	✓
			NA	NA	TERT	UNC5D		MVK	N4BP2L2	RP11-296	CCDC57

Figure 2 - HBV capture workflow allows robust detect ion of HBV DNA integrations
(A) Heatmap of HBV-host DNA integrations detected by the three HBV enrichment protocols: *common capture* (C), *independent capture* (I) and a common *double capture* (CDC) on PLC/PRF/5 and Hep3B cells and a liver resection of non tumor liver tissue from a patient with HBV-related HCC. (B) Comparison of the results obtained in PLC/PRF/5 cells with the 3 capture-sequencing protocols to the HBV integrations described in the literature. The host genes targeted by HBV integration are indicated when known
Chr, chromosome; C, common capture; CDC, common double capture; I, independent capture

According to the published data in PLC/PRF/5 cells, HBV was described to be integrated in 7 genes: TERT, UNC5D, SAC3D1/DNX15, MVK, N4BP2L2, RP11-296, CCDC57. In Hep3B only 1 out of the 2 integrations detected targets an annotated gene (N4BP2L2) that is also targeted by HBV in PLC/PRF/5cells (Fig2A) (Ishii et al., 2020; Chen et al., 2021; Ramirez et al., 2021).

Our results, obtained with the *common capture* HBV enrichment protocol, confirmed 6 genes targeted by HBV integration (TERT, UNC5D, MVK, N4BP2L2, RP11-296, CCDC57) in PLC/PRF/5 cells (Fig. 2B) and the N4BP2L2 gene in Hep3B cells, that is also targeted in the liver resection tissue together with the TERT and UNC5D genes (data not shown).

HBV integration in the TERT gene was found in PLC/PRF/5 in the liver resection tissue. TERT gene takes part in the G1/S transition and in formation of telomerase enzyme. Telomeres protect chromosomes from degradation. At each cell division the telomeres are reduced and when their size is critical the hepatocyte enters in apoptosis. The telomerase postpone apoptosis by adding at each cell division cycle a repeated DNA fragment at the end of the chromosome (Yuan et al., 2019).

HBV integration in the UNC5D gene was found in PLC/PRF/5 in the liver resection tissue. UNC5D codes for one of the netrin-dependent receptors. These receptors have the capacity to induce an apoptosis program, but only in the absence of their ligand. The cell that expresses these receptors is therefore dependent on the presence of the netrin ligand for survival.

The N4BP2L2 gene is targeted by HBV integration in PLC/PRF/5, Hep3B and in the liver resection tissue. This gene encodes the Nedd4 protein E3 ubiquitin ligase.

Transcriptionally activity of HBV integrations.

The alignment of the DNA libraries reads on the HBV genome shows that in the case of PLC/PRF/5 there is an almost complete mapping, e.g., the reads align on the entire genome except in the DR1 region where there is a gap (Fig. 3). This is compatible with the notion that nucleotides 1816-1832 are the ends of the viral integration and with the published results obtained with the long reads PacBio sequencing (Chen et al., 2021; Ramirez et al., 2021). In Hep3B the alignment extends from 2875 to 3163bp and from 0-2125, essentially covering the S1/S2/S and the X ORFs (Fig. 3)

RNA libraries reads align, both in PLC/PRF/5 and Hep3B, on the PreS2/S region and the 3'end domain of the Pol ORF as well as the X ORF. In PLC/PRF/5 cells show alignment to the PreC/C ORF but with very low coverage. The short reads Illumina sequencing does not allow to resolve the different viral transcripts since they are transcribed from largely overlapping ORFs. Taking into account this

limitation the result shown in Fig. 3 are compatible with those reported by Ramirez et al. (2021) using the PacBio Iso-Seq long reads sequencing.

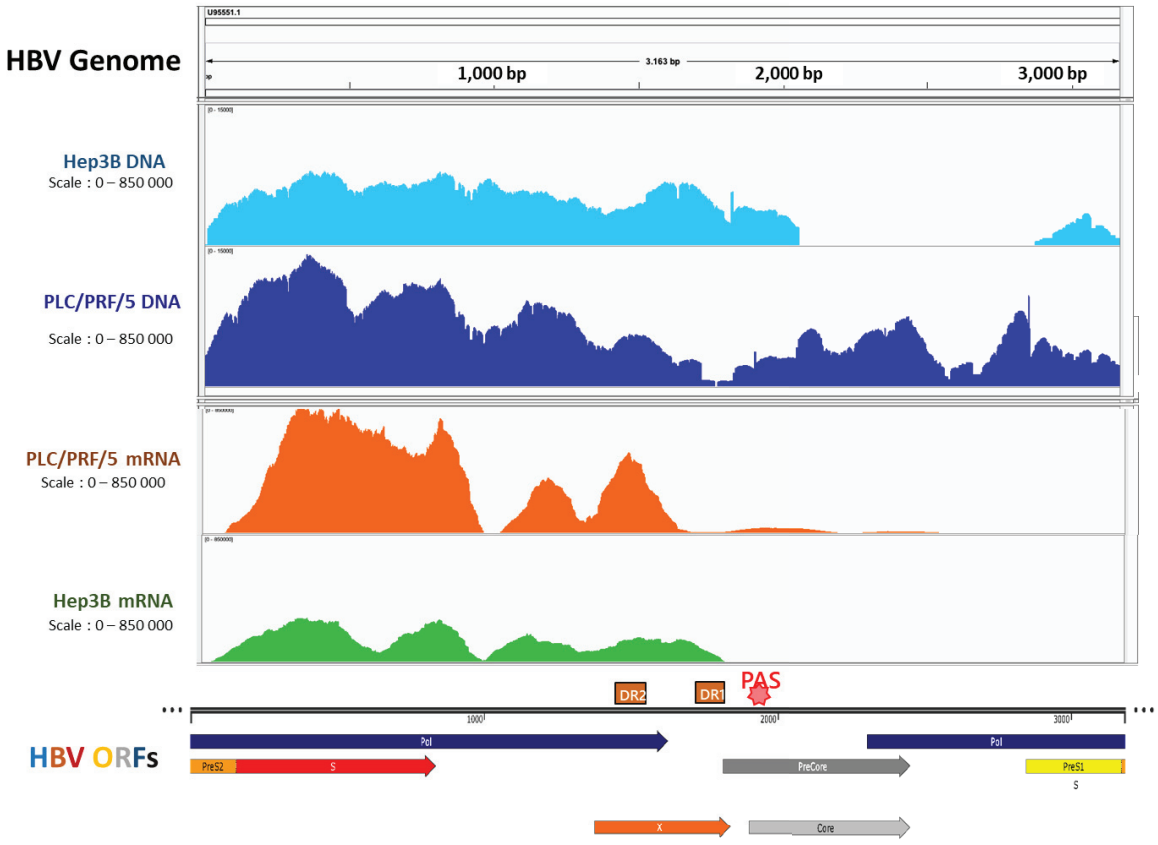


Figure 3 - HBV reads align on the viral genome
Alignment on the viral genome (GenBank accession number U95551) of PLC/PRF/5 and Hep3B HBV DNA and HBV RNA reads enriched by the *common capture* protocol.

To further investigate the HBV transcriptional profile of PLC/PRF/5 and Hep3B cells, 3.5 kb RNA species and total HBV RNAs were quantified by ddPCR with primers/probe sets specific for the Core and the X regions, respectively. Transcripts that terminate at the HBV canonical HBV poly-A site, and that are *bona fide* transcribed from the cccDNA but cannot be produced from integrated HBV sequences, were quantified by the Roche Manual Workflow (MWF) assay (Fig. 4A). ddPCR amplicons were extracted, pooled, and sequenced using the Illumina TruSeq workflow (Fig.4A). The detection of X region amplicons in the 3 cell lines was expected as both PLC/PRF/5 and Hep3B generate transcripts that contain X sequences and that are shared by all HBV RNAs. The detection of the Core amplicons in PLC/PRF/5, that were confirmed by sequencing (Fig. 4A), might be surprising since

complete virions, empty virions and core particles and HBeAg protein have not been detected in PLC/PRF/5 cells (Edman et al., 1980; Ishii et al., 2020; Malmström et al., 2012). However, two studies detected a transcript containing sequences that are unique to the pgRNA (Malmström et al., 2012; Park et al., 2008). At this stage we cannot say whether these amplicons are part of host-virus fusion transcript containing only a small HBV sequence corresponding to our amplicons. The most recent paper published using the HBV targeted PacBio technologies for long-reads high-throughput DNA and RNA sequencing could not detect any pgRNA (Ramirez et al., 2021). However, the HBV integrations in chromosomes 13 and 17 overlap the region amplified in our ddPCR (Chen et al., 2021) but no RNA data accompanied this observation.

To date, all published studies indicate that the RNAs produced by HBV integration are truncated and have lost the canonical HBV poly-A site. Consequently, the Roche MWF assay should be negative in both PLC/PRF/5 and Hep3B cells. Although this is true in Hep3B cells, a very low positive signal (0.2 copies per cell) was detected in PLC/PRF/5 cells, ~500 times lower than the signal detected on the HepAD38 positive control cells. We do not have at the moment any convincing explanation for this low reactivity, but it is evident that the Roche HBV RNA investigational assays detect preferentially, if not exclusively, HBV RNAs carrying the canonical HBV poly-A signal that are not generated, at least according to the currently available data, from HBV integrations.

Our 5'RACE results (Figure 4B) confirm the presence of PreS2/S transcripts in both PLC/PRF/5 and Hep3B cells, in agreement with the DNA sequencing data (Chen et al., 2021; Ramirez et al., 2021) and the few RNA data provided by the paper from Ramirez and coll. (2021).

HBx transcripts could not be detected so far in PLC/PRF/5 and HepG2 cells. A low number of reads that map to the HBx longer transcript described by Stadelmayer and coll (2020) have been found using PacBio long reads Iso-Seq technology (Ramirez et al., 2021). Our 5'RACE results clearly indicate the presence of a long HBx transcript in PLC/PRF/5 cells (Fig. 4B).

A

		X ORF detection by ddPCR	Core ORF detection by ddPCR	Detection of polyA RNA by Roche MWF
PLC/PRF/5	PCR quantification (in RNA copies per cell)	11	21.5	0.2
	Illumina sequencing amplicons (%)	85.1%	93.8%	8.2%
	Expected results according to literature knowledge	+	-	-
Hep3B	PCR quantification (in RNA copies per cell)	354.8	0	0
	Illumina sequencing amplicons (%)	88.3%	Not done	1.8%
	Expected results according to literature knowledge	+	-	-
HepAD38	PCR quantification (in RNA copies per cell)	1072.5	1645	96.4
	Illumina sequencing amplicons (%)	71.1%	97.2%	14.6%
	Expected results according to literature knowledge	+	+	+

B

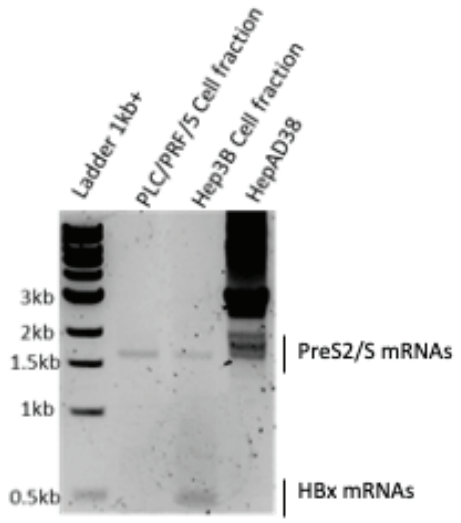


Figure 4 – PLC/PRF/5 and Hep3B transcripts characterization
(A) PCR characterization of RNAs produced by PLC/PRF/5 and Hep3B. HepAD38 was used as a positive control. 3.5 kb RNA species and total HBV RNAs were quantified by ddPCR with primers/probe sets specific for the Core and the X regions, respectively. Transcripts that terminate at the HBV canonical HBV poly-A site, and that are *bona fide* transcribed from the cccDNA but cannot be produced from integrated HBV sequences, are quantified by the Roche Manual Workflow (MWF) assay. (B) 5'RACE performed on cell fractions and of HepAD38 (Positive Control), PLC/PRF/5 and Hep3B cells. Right panel: 3.5 kb RNA species detection in the supernatant of Huh7-3D29 cells. Ladder 1kb+ = molecular weight. Note that all RNAs are smaller by 500bp due to the HBV specific reverse primer localization.

Conclusions

We set up a capture sequencing workflow for HBV DNA and RNA that consistently detects and maps HBV integrations and HBV transcripts and that is sensitive enough to detect integrations in patients' liver samples. Notably, in the non tumor liver sample from liver resection in an HBV-related HCC patient we detected HBV integration in the TERT, UNC5D and N4BP2L2 genes that code for proteins well know to contribute cell transformation and cell death regulation.

The results obtained in the model cell lines PLC/PRF/5 and Hep3B were consitent with those recently reported in the same cell line using the long reads PacBio technology, coupled or not with an HBV capture enrichment. Notably, our workflow generates workable results with a relatively low number of captured HBV reads that allows to use the common capture protocol with Illumina MiSeq instruments, that are quite widely diffused in research and reference clinical laboratories, with 6 samples multiplexing, thus allowing to limit costs.

Future work will focus on further streamligth the bioinformatics pipeline, analyze in depth the RNA-Seq results from the 15 capture-sequencing experiments and expand the analysis of human samples from chronic HBV patients and HCC patients.

Finally, we have set-up the conditions to use our capture-sequencing workflow on plasma/serum samples.

References

Chen, C.-C., Guan, G., Qi, X., Abulaiti, A., Zhang, T., Liu, J., Lu, F., and Chen, X. (2021). Pacbio Sequencing of PLC/PRF/5 Cell Line and Clearance of HBV Integration Through CRISPR/Cas-9 System. Front. Mol. Biosci. 8, 676957.

Edman, J.C., Gray, P., Valenzuela, P., Rall, L.B., and Rutter, W.J. (1980). Integration of hepatitis B virus sequences and their expression in a human hepatoma cell. Nature 286, 535–538.

Ho, D.W.H., Sze, K.M.F., and Ng, I.O.L. (2015). Virus-Clip: a fast and memory-efficient viral integration site detection tool at single-base resolution with annotation capability. Oncotarget 6, 20959–20963.

Ishii, T., Tamura, A., Shibata, T., Kuroda, K., Kanda, T., Sugiyama, M., Mizokami, M., and Moriyama, M. (2020). Analysis of HBV Genomes Integrated into the Genomes of Human Hepatoma PLC/PRF/5 Cells by HBV Sequence Capture-Based Next-Generation Sequencing. Genes 11.

Liang, Y., Qiu, K., Liao, B., Zhu, W., Huang, X., Li, L., Chen, X., and Li, K. (2017). Seeksv: an accurate tool for somatic structural variation and virus integration detection. Bioinforma. Oxf. Engl. 33, 184–191.

Malmström, S., Larsson, S.B., Hannoun, C., and Lindh, M. (2012). Hepatitis B viral DNA decline at loss of HBeAg is mainly explained by reduced cccDNA load--down-regulated transcription of PgRNA has limited impact. PloS One 7, e36349.

Park, G.-S., Kim, H.-Y., Shin, H.-S., Park, S., Shin, H.-J., and Kim, K. (2008). Modulation of hepatitis B virus replication by expression of polymerase-surface fusion protein through splicing: Implications for viral persistence. Virus Res. 136, 166–174.

Ramirez, R., van Buuren, N., Gamelin, L., Soulette, C., May, L., Han, D., Yu, M., Choy, R., Cheng, G., Bhardwaj, N., et al. (2021). Targeted Long-Read Sequencing Reveals Comprehensive Architecture, Burden, and Transcriptional Signatures from Hepatitis B Virus-Associated Integrations and Translocations in Hepatocellular Carcinoma Cell Lines. J. Virol. 95, e0029921.

Scholtès C, Hamilton TA, Plissonnier ML, Charre C, Scott B, Wang L, Berby L, French J, Testoni B, Blair. A, Subic M, Hoppler M, Lankenau A, Grubenmann A, Levrero M, Heil ML, Zoulim F. (2022) Performance of the cobas® HBV RNA Automated Investigational Assay for the Detection and Quantification of Circulating HBV RNA in Chronic HBV Patients. J Clin Virol (in press)

Stadelmayer, B., Diederichs, A., Chapus, F., Rivoire, M., Neveu, G., Alam, A., Fraisse, L., Carter, K., Testoni, B., and Zoulim, F. (2020). Full-length 5'RACE identifies all major HBV transcripts in HBV-infected hepatocytes and patient serum. J. Hepatol. *73*, 40–51.

Twist, E.M., Clark, H.F., Aden, D.P., Knowles, B.B., and Plotkin, S.A. (1981). Integration pattern of hepatitis B virus DNA sequences in human hepatoma cell lines. J. Virol. 37, 239–243.

Yuan, X., Larsson, C., and Xu, D. (2019). Mechanisms underlying the activation of TERT transcription and telomerase activity in human cancer: old actors and new players. Oncogene 38, 6172–6183.

Characterization of circulating Hepatitis B virus RNAs *in vitro* and in chronic hepatitis B patients

Circulating Hepatitis B Virus (HBV)-RNA associates with extracellular vesicles in supernatant of HBV-infected hepatocytes and serum of chronically infected patients

Doohyun Kim[1*], Delphine Bousquet[1,2*], Marie-Laure Plissonnier[1], Hyoseon Tak[1], Xavier Grand[1], Chloé Goldsmith[3], Françoise Berby[1], Isabelle Bordes[1], Alexia Paturel[1,2], Aaron Hamilton[4], Marintha Heil[4], Massimo Levrero[1,2,5], Barbara Testoni[1§] and Fabien Zoulim[1,2,6§]

[1] INSERM U1052, CNRS UMR-5286, Cancer Research Center of Lyon (CRCL), Lyon, 69008, France; [2] University of Lyon, UMR_S1052, CRCL, 69008 Lyon, France; [3] Faculty of Health, University of Canberra, ACT, Australia; [4] Roche Molecular Diagnostics, Pleasanton CA; [5] Department of Internal Medicine - DMISM and the IIT Center for Life Nanoscience (CLNS), Sapienza University, Rome, Italy; [6] Department of Hepatology, Croix Rousse hospital, Hospices Civils de Lyon, France

*These authors contributed equally to the work

Corresponding Author:

Prof. Fabien Zoulim, INSERM U1052- CRCL, 151 cours Albert Thomas, 69424 Lyon - France
Mail: fabien.zoulim@inserm.fr; Phone: (+33) 4 72 68 19 70, Fax r: (+33) 4 72 68 19 71

Abstract

Chronic hepatitis B (CHB) infection is a major health problem with 300 million of chronic carrier worldwide and is the major cause of liver failure, cirrhosis, and hepatocellular carcinoma (HCC). Due to the persistence of hepatitis B virus (HBV) covalently closed circular DNA (cccDNA), current available drugs do not allow to achieve a complete cure. New strategies towards a cure for HBV are warranted as well as serum biomarker better reflecting intrahepatic cccDNA activity. Increasing evidence suggests that circulating HBV-RNA (CirB-RNA) may serve as a serum biomarker for HBV infection, treatment, and prognosis. However, the nature of CirB-RNA and the particles in which they are contained *in vitro* and *in vivo* remains a matter of debate.

Supernatants from HBV-infected HepG2-NTCP cells treated or not with lamivudine, and sera from 9 untreated [4 HBeAg(+) and 5 HBeAg(-)] and 1 HBeAg(+) ETV-treated chronic hepatitis B (CHB) patients were subjected to Iodixanol/Sucrose ultracentrifugation. Each fraction was analyzed for HBV DNA/RNA by specific qPCR and droplet digital (dd)PCR. Viral and extracellular vesicles (EVs)-associated proteins were detected by ELISA and Western Blotting. 5' RACE PCR followed by ONT MinION sequencing was used to identify CirB-RNA species. Longitudinal serum samples before and at two time points after NUC therapy initiation were obtained from two additional patients [HBeAg(+) TDF-treated CHB and HBeAg(-) ETV-treated CHB].

CirB-RNA was mainly detected in core-associated virion-like particles, in 2 log10 less amount than HBV DNA. CirB-RNA was the predominant species in lighter density fractions (1.17-1.18 g/ml) deprived of viral proteins, both in cell supernatant and in serum. The enrichment for EVs in these fractions was confirmed by detection of CD9 and CD81 by Western Blotting, immunoprecipitation assay, Nanoparticle tracking analysis and Transmission Electron Microscopy. Distribution of CirB-RNA did not differ significantly according to HBeAg status, while in a patient with low HBsAg level, CirB-RNA was mainly detected in the EVs-enriched fractions. Lastly, CirB-RNA profiling by 5' RACE and ONT MinION sequencing identified different proportions of pgRNA-derived transcripts according to HBeAg status and HBsAg level.

Our results indicate that EVs-enriched compartment also contributes to the circulation of HBV-RNAs. Moreover, different HBV-RNA transcripts in addition to pgRNA can be detected *in vivo*. Altogether, these data could significantly contribute to the characterization of cirB-RNAs as new viral biomarker.

Key Words: Hepatitis B Virus, Extracellular Vesicles, Exosomes, Viral RNAs, Biomarker

Introduction

Despite the availability of a prophylactic vaccine and antiviral treatment, Hepatitis B virus (HBV) remains the causative agent of chronic hepatitis B (CHB), which affects more than 300 million people worldwide (‘Global Hepatitis Report, 2017). Current treatments such as nucleo(s)tides analogues (NAs) do not enable to achieve a complete cure due to the persistence of covalently closed circular DNA (cccDNA) in the nucleus of infected hepatocytes, which serves as viral genomic reservoir and matrix for transcription of viral RNAs (Carey et al. 2019). Direct intrahepatic quantification of cccDNA molecules and HBV RNA represents the most accurate way to estimate HBV replicative potential *in vivo*. However, this is hampered by the invasive nature of liver biopsies and by technical issues limiting the reproducibility and sensitivity of cccDNA and RNA detection. HBV core-related antigen (HBcrAg) and circulating HBV RNA (cirB-RNA) have been recently proposed as non-invasive surrogate markers for intrahepatic cccDNA amount and transcriptional activity (Charre et al. 2019). In particular, cirB-RNA seems a promising diagnostic biomarker, especially for monitoring NA therapy (Jansen et al. 2016). Studies suggest a wide heterogeneity in the nature and sequence of cirB-RNA according to CHB phase or antiviral treatment. Moreover, the lack of standardized methods for isolation and detection of CirB-RNAs prevents their complete characterization (Otsuka and Koike 2020). CirB-RNA are full length pgRNA (Stadelmayer et al. 2020; Hacker et al. 2004; Jansen et al. 2016; Lam et al. 2017; Prakash et al. 2018; Wang et al. 2016; Bai et al. 2018) or splicing variants derived from pgRNA (Lam et al. 2017) or HBx (Kouwaki et al. 2016). The nature of CirB-RNAs and the particles in which they are transported in the circulation remain largely unsolved.

Beside DNA containing virions, many other viral particles such as subviral particles, empty virions and naked capsids are secreted by HBV replicating cells (Liu and Hu 2019). Subviral HBsAg particles are secreted in vast excess over virions in sera of CHB patients. Empty virions share the same structure as DNA containing virions but are devoid of nucleic acids. Naked capsids (NCs), which exit cells via a different route from virions are either empty or filled with viral RNA and immature viral DNA. It has been suggested that CirB-RNAs are contained in capsids which might also be enveloped but other vesicle types such as extracellular vesicles (EVs) are not excluded (Kouwaki et al. 2016).

There is growing evidence that HBV maturation and egress depend on intraluminal vesicles. Interestingly, MVB is not only involved in virion secretion, but is also known to be a major site in generating extracellular vesicles (EVs) (Théry, Zitvogel, and Amigorena 2002). Among EVs, exosome consists of 40- to 100-nm diameter small size vesicles. Recent studies have shown that EVs are involved in multiple biological processes necessary for the cell physiology and in response to

pathogens (Alenquer and Amorim 2015). Notably, recent studies have shown that EVs play an important role in infectious disease by mediating the transfer of pathogen-derived antigens and virulence factors. EVs derived from virus-infected cells contain viral proteins and viral RNA; the contents of these vesicles are enclosed in a cellular membrane, thus enabling viruses to evade host immune response to herpes simplex virus, human immunodeficiency virus, Epstein–Barr virus, and cytomegalovirus infection (Bukong et al. 2014; Baglio et al. 2016). Furthermore, EVs from hepatitis A virus or hepatitis C virus–infected hepatocytes permit the respective viruses to invade and replicate within host hepatocytes (Longatti 2015).

In this study, we investigated the nature and the physical entities that transport circulating HBV-RNAs *in vitro* in HBV infected HepG2-NTCP cells and *in vivo* in the sera of CHB patients in the presence or not of NA treatment. Viral and non viral EVs were separated by a specifically set-up Iodixanol/sucrose density gradient approach and virological characterization was performed by droplet digital PCR (ddPCR) quantification of HBV DNA/RNA and viral proteins detection in each recovered fraction. As a key experiment to analyze all CirB-RNAs, 5' race analysis was performed using *in vitro* and *in vivo* samples, followed by single molecule sequencing by Oxford Nanopore Technologies (ONT) MinION.

Materials and methods

Cell lines, viral inocula and infection conditions

For infection purpose, HBV particles were concentrated from the supernatant of HepAD38 (HBV genotype D) cells by filtering and PEG precipitation as described previously (2137609. The HepAD38 cell line was a kind gift of Dr C. Seeger (Fox Chase Cancer Center, Philadelphia, USA) (9257747). The HepG2-NTCP cell line was a kind gift of Dr S. Urban (Heidelberg University, Germany) (24361467). HepG2-NTCP cells were infected at a multiplicity of infection of 250 vge/cell in culture media supplemented with 4% PEG (polyethylene glycol 8000). For LAM treated condition, 10 µM lamivudine (Sigma Aldrich) was added to cell cultures 3- and 5-days post infection.

Serum samples

Serum samples were collected after written informed consent from the patients and the protocol was approved by the Lyon institutional Ethic Committees (Comité de Protection des Personnes (CPP) Sud-Est IV # 11/040 and successive amendments). Patient informations are in Table 1 and Table. 2.

EV isolation

100mL to 1liter of culture supernatants were collected and centrifuged at 1500 xg for 15 min at room temperature. To thoroughly remove cellular debris, the supernatants were filtered through a 0.22µm filter (Merck Millipore, KGaA, Darmstadt, Germany). For EV concentration, filtered supernatant was purified using Amicon® Pro Purification System with 100kDa filter. Next, for EV preparation, the concentrated supernatants were ultracentrifuged at 110,000 ×g for 2h at 4˚C. The pellets were then washed with 8 mL of PBS to remove contaminant proteins and ultracentrifuged again at 110,000 ×g for 2h at 4˚C.

Density Gradient

EV pellets were resuspended in 2.5 mL of 10% iodixanol solution. 10%, 20%, 30% and 40% iodixanol solutions were prepared by mixing Optiprep™ (Axis Shield) with buffer containing 0.25 M sucrose, 10 mM Tris at pH 8.0, and 1 mM EDTA, with a final pH at 7.4. Resuspended EVs were layered on the top of the gradient and then subjected to ultracentrifugation in a SW41-Ti Rotor tube (Beckman) for 6 hours at 4˚C at 110.000 ×g. Twelve fractions of 1mL were recovered and analyzed separately.

Immuno-precipitation (IP) of EVs

Anti-CD9 and anti CD-81 antibody coupled with Dynabeads M-280 (Thermo Fisher Scientific Inc.) was added to matched set specimens in cell supernatant and serum that were diluted with EGTA, followed by incubation on a rotator at 4°C for 18 h. The beads were washed three times with PBS and stored at 4°C until further analysis.

Nucleic Acid Extraction

Viral nucleic acids were either extracted from 200µl of whole cell supernatant or gradient fractions using QiAmp MinElute Virus spin kit (Qiagen, Germany). RNA was digested with DNase 1 (RQ1 RNase-Free DNase (Promega) and retro-transcribed into cDNA using SuperScript IV Vilo (SuperScript™ IV VILO™ Master Mix) according to manufacturer's instructions (Invitrogen, Carlsbad, USA). To ensure that DNA was completely digested and that RNA was exclusively measured, a control without reverse transcriptase enzyme was performed in each experiment.

Quantification of total HBV-DNA, total HBV-RNA and 3.5 kb RNA

qPCR quantification was performed using Applied QuantStudio 7 machine (BioSystem). Precore/pgRNA (3.5 kb RNAs) were detected by the primer pair GGAGTGTGGATTCGCACTCCT (A) and AGATTGAGATCTTCTGCGAC (B) together with the Taqman hybridization probe ([6FAM]-AGGCAGGTCCCCTAGAAGAAGAACTCC-[BHQ1]). Total DNA and RNAs were measured using primers and Taqman probe Pa03453406_s1 (ThermoFisher Scientific). ddPCR was used for cells samples and low viremia patient's serum. Quantification was performed using ddPCR machine (QX200; Bio-Rad) as described in (Salerno et al., 2020). The sequences of primers and probes were the same used in qPCR. For samples with HBV DNA or RNA levels higher than 10^6 copies, the appropriate dilutions of DNA or cDNA template were performed.

ELISA for viral antigens

ELISA tests for HBeAg and HBsAg detection in cell supernatants were performed according to the manufacturer's protocol using the CLIA kits from Autobio Diagnostic.

Transmission Electron Microscopy (TEM)

EVs were fixed with 4% of paraformaldehyde (PFA) and was adsorbed on 200 Mesh nickel grids coated with formar-C. After washing once in filtrated ultra-pure water, suspensions were colored with 2% phosphotunstic acid for 2 min and examined using a JEM Jeol 1400 transmission electron microscope (Tokyo, Japan) equipped with a Orius 600 camera (USA). Particle sizes were determined with the Digital Micrograph software.

Western Blot

Fractions from density gradient were mixed with Laemmli buffer and heat at 95°C for 5min. Proteins were migrated in 4-20% mini-PROTEAN@ TGX stain-FreeTM Precast Gel (Bio-Rab Laboratories) and transferred onto a nitrocellulose membrane (Bio-Rab Laboratories). Membranes were blocked 1 hour with 5% milk or BSA (Sigma) in TBS (1 x Tris Buffer Saline (Sigma)) and stained with primary antibodies in blocking buffer overnight at 4°C and stained with HRP-conjugated secondary antibodies (1/50000) for 1 hour at room temperature. The detection was done using Clarity or Clarity Max Western ECL and the ChemiDoc XRS system (Biorad).

NanoTracking Analysis (NTA)

NanoSight NS300 (Malvern Panalytical, Orsay, France) was used to quantify particles in suspension. The number of particles and their movement were recorded for 3 × 30 s (camera level = 14, 25°C). The data obtained are presented as the average and standard deviation of the three video recordings for each replicate.

5’RACE analysis

5’RACE 5’RACE was performed as previously described in (Stadelmayer et al., 2020). Briefly, RNAs were isolated using a guanidinium thiocyanate–phenol–chloroform extraction protocol (TRI reagent (Sigma)). 5'RACE was essentially performed as described in the GeneRacer Kit manual (ThermoFisher Scientific) except Tobacco Acid Pyrophosphatase was substituted by RNA 5' Pyrophosphohydrolase (New England Biolabs) and SuperScript reverse transcriptase III by SuperScript reverse transcriptase IV (ThermoFisher Scientific). The reverse transcription reaction was performed using 3' HBV specific Gsp1 primer 5'-TTAGGCAGAGGTGAAAAAAGTTG-3'. For the 5'RACE PCR reaction Prime Star super mix DNA Polymerase (TAKARA bio), GeneRacer 5' primer and HBV specific nested primer Gsp1 5'-TTAGGCAGAGGTGAAAAAAGTTG -3' were used. 5'RACE PCR was run in a C100 Touch thermocycler

(Biorad) using the following PCR program: Initial denaturation step 98°C 3 min >5x (98°C 10 s; 72°C 3 min) >5x (98°C 10 s; 70°C 3 min) >25x (98°C 10 s; 64,4 20 s; 72°C 3min) >72°C 10min.

Oxford Nanopore Technology (ONT) MinION DNA sequencing of 5'RACE products

Products of 5'RACE PCR amplification performed with Platinum™ Taq DNA Polymerase High Fidelity (Invitrogen) were prepared for ONT MinION Sequencing using SQK-PBK004 kit, according to the manufacturer's instructions. Analysis was performed similarly to (Pertea et al., 2016) combining HISAT, StringTie and gtf compare tools. Sequences were aligned to a reference HBV genome, corresponding to full-length genome starting at forward PCR primer (position 1814).

Results

Distribution of extracellular HBV-RNA in gradient-separated supernatant of HBV-infected HepG2-NTCP cells

Extracellular Vesicles (EVs) secreted from HBV infected HepG2-NTCP cells encompass exosomes, subviral particles (SVPs), naked capsids (NCs), virions and virion-like particles (VLPs). We set-up a density gradient approach to separate each EVs subtype (Fig.1A). Cell supernatants and patient's sera were concentrated and loaded on a 10% iodixanol solution to be ultracentrifugated to separate the different particles according to their density. The particles floating remain on the surface in the 10% iodixanol solution while the denser ones will penetrate the gradient to the bottom in the 40% iodixanol solution. Finally, twelve fractions were recovered and analyzed for the presence of viral proteins such as HBs (Fig. 1B) and HBc (Fig. 1C), but also the exosome marker flotillin-1 (Flot-1) (Fig. 1C), viral DNA (Fig. 1D) and viral RNA (Fig. 1E).

HBsAg was detected in fractions 4-9 of the density gradient, with one major peak in fraction 5, and HBc was detected from fractions 7 to 11. The viral nucleic acids were found in almost all the fractions except in fractions 1, 2 and 12. A major peak is found for HBV DNA in fraction 7 while for HBV RNA there are three main peaks at fractions 4, 7 and 9. It is assumed that the detection of both HBsAg and HBc represents the detection of virion-like particles, while the detection of HBc only represents the naked capsids. HBsAg, HBc and HBV DNA were detected mainly in fractions 7 indicative of the presence of virions. The presence of both HBc and HBV DNA without HBsAg, indicative of naked capsids containing HBV DNA, was observed in fractions 9-11 with a major peak at fraction 9. Likewise, HBV RNA was quantified with peaks at fraction 7 indicative of virion-like particle containing HBV RNA and at fraction 9 indicative of naked capsids containing HBV RNA. Furthermore, total HBV RNAs levels were higher than 3.5 kb suggesting that 3.5 kb RNA would not be the only RNA specie that is secreted upon infection.

The analysis of exosomes content was performed by the detection of the exosomal membrane protein CD9 (Klymiuk et al., 2019) and Flot1, that are tetraspenin molecules markers of the EVs surface. CD9 and Flot1 were detected in fractions 2-6. Viral nucleic acids quantification in these fractions suggests the presence HBV DNA and RNA in exosomes. Noteworthy, HBV DNA was found in higher quantities than HBV RNA (10^6 vs 10^3) in exosomes fractions.

Lamivudine induced a dramatical decrease in HBV DNA, while cccDNA did not differ, and total HBV RNAs and 3.5 kb RNAs quantities increased with lamivudine treatment (Fig. S1A-C). In addition, there was no difference in HBsAg and HBeAg in whole supernatants, HBV DNA decreased, and 3.5 kb RNAs

increased by the lamivudine treatment condition (Fig. S1D-G). No significant change in exosomes markers or viral protein levels and the repartition of particle was maintained under lamivudine treatment. These results indicate that CirB-RNAs are mainly 3.5 kb RNAs but that other species are secreted and that the secretion of those nucleic acids are done via EV, VLP and NC in HBV infected-HepG2-NTCP cell.

Distribution of extracellular HBV-RNA in gradient-separated sera of chronically HBV-infected patients

Two types of serum were used for this study: one from a patient HBeAg (+) chronic infection with 8 log_{10} viral load and high HBsAg expression, while the second serum is from a patient HBeAg (-) chronic hepatitis with 5 log_{10} viral load and low HBsAg expression (Table 1).

The same experimental conditions were used for sera samples of patients than to the previously described for cell supernatant. Thus, after concentration, the samples were loaded on the iodixanol gradient. Finally, 10 fractions were obtained after Iodixanol/sucrose density ultracentrifugation for the high and low viral load patients' sera.

For the HBeAg(+) patient, and similarly to the cell supernatant, HBsAg was detected by two methods (western blot and ELISA) in fractions 4-8 of the density gradient, with one major peak in fraction 5, and HBc was detected from fractions 6 to 10 (Fig. 2 A and B). The viral nucleic acids were found in almost all the fractions. A major peak is found for HBV DNA in fraction 7 (Fig. 2C) while for HBV RNA there are, in this case, two main peaks at fractions 3 and 7. The co-detection of HBsAg, HBc and HBV DNA in the same fractions was found mainly in fractions 7, indicative of the presence of virions. The presence of both HBc and HBV DNA without HBsAg, indicative of naked capsids containing HBV DNA, was observed in fractions 8-10 with a major peak at fraction 9. Likewise, HBV RNA was quantified with a peak at fraction 7 (Fig. 2D) indicative of virion-like particle containing HBV RNA. In addition, in fractions 8-10 HBV-RNA was detected in the presence of only HBc viral proteins indicative of naked capsids containing HBV RNA. Like in the cell supernatants, total HBV RNAs levels were higher than 3.5 kb suggesting that 3.5 kb RNA would not be the only RNA specie that is secreted in patients.

The same conclusions can be drawn from the serum analysis of three other patients with the same viral parameters (data not shown). These results suggest that CirB-RNAs are mostly 3.5 kb RNAs secreted through EV, VLP and NC in patients with high viral load.

The analysis profile of viral parameters and EVs markers in the gradient fractions of an HBeAg(-) chronic hepatitis with low HBsAg patient serum differs from the HBeAg(+) patients. Although viral

DNA is still detected in fractions 6-10 (Fig. 2G), it must be noted that the amounts of HBV DNA are lower than in the serum, with $3x10^4$ copies/mL at the peak in fraction 7. HBsAg western blotting and HBV core were not detected likely due to their low expression level and the low HBV replicative activity (Fig. 2F). It is therefore impossible to locate virion secretion in the serum of this patient due to the sensitivity of the technique.

The distribution of RNA secreted in the serum of this HBeAg(-) chronic hepatitis patient with low HBsAg (Fig. 2H) is very different from the one found in the HBeAg(+) patient who presented a high viral load. Three peaks are found, one of which is one log_{10} higher than the other two and spread over fractions 2-4, where we also detected CD9 and CD81 markers, identifying EVs. This indicates that the majority of HBV RNAs are secreted into the serum by EVs and that a smaller proportion is secreted by the other NC and VLP viral particles. This is in contrast to what was shown for the HBeAg(+) patient where the greatest amount of secreted RNA was quantified in NC and VLP.

These results suggest that CirB-RNAs are secreted through EVs, VLPs and NCs in high viral load patients. However, in HBeAg negative and low HBs expression conditions of CHB patient the CirB-RNA secretion may shift from VLPs to EVs.

Distribution of extracellular HBV-RNA in gradient-separated serum of longitudinal serum sample derived from a Tenofovir-treated patient

An analysis of viral parameters and EVs was performed over several months on a patient who had at his first visit a 8 log_{10} viral load with high HBsAg expression. Following his first visit, this patient was treated with tenofovir. At the first visit as well as at the 6- and 10-months visits, serum was collected to analyze viral parameters and EV markers (Fig. 3A).

After starting HBV treatment, HBsAg and HBV DNA decreased over time (Fig. 3B and C). HBV DNA secreted in the serum decreased from 7 Log_{10} to 3 Log_{10} and further to 1 Log_{10} copies/mL (Fig. 2D), as expected with such treatment. Regarding HBV surface proteins, the amount also decreased from $2.5x10^4$, to $1.5x10^2$ and then to $1x10^2$ copies/mL. However, the HBV DNA and HBc as well as the HBsAg distributions did not change during the 10 months of follow-up. Importantly, the distribution of EVs remains identical to the different cases described above and constant during the 10-month follow-up of this patient. EVs markers and the viral parameters, whose distribution does not change during the follow-up, allowed to establish the distribution profile of the different particles: the EVs are found in fractions 2-5, the VLPs are detected in fractions 6-7 and the NCs are in fractions 8-10.

Concerning total CirB-RNAs (Fig. 3C), the levels were reduced from 6 Log_{10} to 2 Log_{10} and finally to 1 Log_{10} copies/mL as expected with long term TDF treatment. At the first visit, two peaks of circulating RNA are detectable: the first one overlapping with the EVs, and the second one more important in terms of quantity, spreading on the same fractions as the VLPs and NCs. The profile is balanced on the second visit: we still find two peaks of circulating RNAs at the level of EVs and VLPs/NCs but the difference in terms of quantity between the two peaks is less. In the third visit, there is an inversion of the profile: CirB-RNA peak at the fractions 2-5 is higher compared to the peak found at the VLPs/NCs portion ($5x10^1$ vs $2x10^1$). At the first visit, there are more total RNAs than 3.5 kb RNAs, which indicates that the secreted RNAs are mostly 3.5 kb but not exclusively and that other viral RNA species are also secreted. At the 2 follow up time points the amounts of total RNA and 3.5 kb RNA are almost identical.

Altogether theses results suggest that CirB-RNAs are secreted through EVs, VLPs and NCs in the serum of patients. The TDF treatment results in a decrease of CirB-RNA and the CirB-RNA secretion shifts from VLPs/NCs to EVs along the treatment.

Secreted HBV RNA is contained in exosomes-like particles (sEVs)

To get a better insight on exosome repartition across the gradient, we looked for a method that could allow their quantification in each fraction. Nanoparticle Tracking Analysis (NTA) is commonly used to determine EVs concentration and diameter. To further characterize EVs distribution across the fractions, we analyzed each gradient fraction by NTA. In fig. 4A, we confirmed the previous results observed in western blot with the Flot-1 exosomal marker. Exosomes distribute mainly in the light fractions of the gradient with a peak in F3 and F4 containing approximatively 30 and 25% of the total exosomes in the gradient. Notably Lamivudine treatment didn't affect particles concentration, suggesting that exosomes secretion is not enhanced by Lamivudine treatment.

TEM experiments (Fig.4B) confirmed the presence of both exosomes and SVPs in the F4 fraction of the density ultra-centrifugation. Exosomes looked as *cup shaped* structures of ~100nm in diameter containing a double membrane. SVPs comprised spheres of 22nm diameter and filaments of 18nm width and up to 400nm long. As simple negative staining could not inform us about the presence of HBs at the surface of the exosomes, we performed immunogold labelling with HBs antibody and CD9 antibody, a marker for exosome membrane. HBs and CD9 were not co-localizing in our experimental conditions suggesting that F4 is composed by two separated entities: exosomes and SVPs. Furthermore, across the gradient, we observed that neither DNA nor RNAs levels were significantly

affected by MNase digestion in F4 suggesting that HBV RNAs are mostly contained inside the exosomes in the supernatant of infected hepatocytes (Fig. S2).

Although, exosome markers were detected in HepG2-NTCP cells and in the light fractions of patients' sera, immunoprecifitation (IP) assays were performed to further confirm whether CirB-RNAs were contained in exosomes. 100 mL of HepG2-NTCP supernatant were concentrated by a ultracentrifugation using sucrose cushion, and 500 µl of patient serum were concentrated using a sucrose cushion, and then IP was performed with CD9 and CD81 antibodies (Fig. 4C). For the exosome input, a collection of light fractions (F2-5) mainly containing exosomes was used as input. IgG was used as a control antibody, and exosome-specific IP was performed using CD9 and CD81. Both total HBV RNAs and 3.5 kb RNA were detected in the case of CD9 and CD81 IPs in HepG2-NTCP cells and the patient serum (Fig. 4D-G).Taken together, this set of results suggests that sEVs derived from HBV-infected cells or CHB patient's serum indeed carry viral RNAs.

Various HBV-RNAs species are secreted in the supernatant of infected hepatocytes and in the sera of CHB patients

With the aim of identifying the HBV-RNAs species present in the different fractions of the gradient we used the 5'RACE method (Stadelmayer et al., 2020). In the case of the supernatants of HBV-infected HepG2-NTCP cells, selected fractions representing exosomes+SVPs (F4), Dane particles (F7) and NCs (F9) were analyzed separately. 3.5 kb RNA species were found in the three fractions (Fig. 5A). In addition, spliced variants were also present, mostly in F7 corresponding to virion-enriched fraction. A similar profile was observed in the serum from a patient with high viremia. Different HBx transcripts were also identified in all the fractions. The migration profile was slightly different in serum from ETV treated patient where 3.5 kb RNA was the predominant RNA specie in all the fractions. PCR-products from 5'RACE were then sequenced by long-read Oxford Nanopore Technology (ONT). MinION sequencing identified high levels of pgRNA in all fractions both in HepG2-NTCP and sera of CHB patients. As expected, pgRNA was the main RNA specie reaching up to 30% of the total reads in each fraction. The more abundant pgRNA derived spliced variants were SP1, SP3, SP9 and SP11. Slight variations in the quantity of the spliced variants were observed across the fractions *in vitro* and *in vivo*. Canonical and short HBx mRNAs were found in all fractions with a predominance in the virions associated fraction (F7). Canonical HBx mRNA was not found in ETV treated patient's serum. Altogether, these results support that HBV-RNAs in the supernatant of infected hepatocytes comprise other species besides pgRNA.

Discussion

Our results indicate that CirB-RNAs are mainly 3.5 kb RNAs and more precisely pgRNA but that other species are also secreted. Secretion is supported by EV, VLP and NC in HBV infected-HepG2-NTCP cell and patient's sera. Untreated patients with low HBV viral replication and patients treated with nucleos(t)ide analogues showed a CirB-RNA secretion shifts from VLPs/NCs to EVs.

Several points merit to be underlined. First, serum derived exosomes are a mixture of exosomes from different cell population and organs, while exosomes derived from hepatocytes like HepG2-NTCP are only hepatocytes derived exosomes (Willms et al., 2018). For this reason, the analysis of the compartments used for HBV RNA secretion in both cell line and sera from patient is quite novel. Second, it has been shown that EVs secreted from HBV infected cell can be secreted and can include HBV rcDNA that are infectious for surrounding cells (Sanada et al., 2017). Thus, it can be hypothesized that the production of viral RNAs secreted in EVs in treated patients or those with low viral parameters, may contribute to the maintenance of viral infection through the infection of new hepatocytes. Notably NUCs inhibit HBV DNA replication but do not suppress cccDNA transcriptional activity that continuously generates HBV RNAs (Fanning et al., 2019). Our results further support the rationale for CirB-RNAs as a novel viral biomarker for indirect measurement of cccDNA transcriptional activity.

The mechanism leading to the secretion of HBV-RNAs are still poorly characterized. RNA binding proteins (RBPs) regulate RNA processing, nucleocytoplasmic transport, maturation, intra-compartmental localization and turnover (Glisovic et al., 2008; König et al., 2012). As a result, RBPs could serve as key players in HBV RNAs secretion, by making complexes with HBV RNAs and transporting them into exosomes during their biosynthesis. Currently, there is no direct evidence for this RBPs-mediated transport mechanisms. Heterogeneous nuclear ribonucleoproteins (hnRNPs), a large family of 20 major and minor proteins in humans with multiple RNA-binding capacities (Dreyfuss et al., 1993) are obvious candidates.

Legends to Figures

Figure 1. Secreted HBV RNA is associated to small extracellular vescicles in supernatant of HBV-infected HepG2^{hNTCP} cells (A) Scheme of 10-40% Iodixanol/STE density gradient and theoretical small extracellular vesicles (Evs) distribution according to their buoyant density. Briefly, HepG2^{hNTCP} cells were infected at MOI 250 and supernatant was collected 8 dpi. Infected cells were treated (3TC) or not (N.T) with lamivudine (3TC) at 10µM for 5 days. (B) HBsAg ELISA was measured in each fraction. (C) Exosome marker, Flotillin-1 (Flot-1) and HBV core protein (HBc) were detected by western blot in each fraction. (D) Total HBV-DNA, (E) total HBV-RNA (tHBV-RNA) and 3.5 kb RNA were quantified by ddPCR in each fraction.

Figure 2. Analysis of viral parameters and EVs markers in gradient fractions of serum samples derived from patients in different CHB phases.
Gradient fractions were obtained according to the protocol detailed in Method section. a) HBsAg ELISA. b) Representative Western Blot analysis using anti-extracellular vesicle (EVs) markers (anti-CD9 and anti-CD81), anti-HBsAg and anti-HBV core antibody. c) HBV DNA quantification by qPCR. d) HBV RNA quantification by RT-qPCR. HBx ORF common region was used to quantify both HBV DNA and total HBV RNA. pre-Core/pgRNA 5' unique region was used to quantify 3.5 kb HBV RNA (Testoni et al. JHep 2019). CHB phases are defined according to 2017 EASL CPGs.

Figure 3. Analysis of viral parameters and EVs markers in gradient fractions of longitudinal serum samples derived from a Tenofovir-treated patient.
Gradient fractions were obteined according to the protocol detailed in Method section. a) HBsAg ELISA. b) Representative Western Blot analysis using anti-extracellular vesicle (EVs) markers (anti-CD9 and anti-CD81), anti-HBsAg and anti-HBV core antibody. c) HBV RNA quantification by RT-qPCR. HBx ORF common region was used to quantify both HBV DNA and total HBV RNA. pre-Core/pgRNA 5' unique region was used to quantify 3.5 kb HBV RNA (Testoni et al. JHep 2019). d) HBV DNA quatification by qPCR.

Figure 4. Characterization of sEVs fraction. Supernatant from HBV-infected HepG2^{hNTCP} cells was separated through density gradient. (A) Particles were quantified, and size was assessed in 3TC- treated and non-treated samples by Nanoparticle Tracking Analysis. (B) Negative staining of fraction 4 in electron microscopy (PTA 3%). (C) Schematic protocol of sEVs-Immunoprecipitation (IP) (D-E) sEVS-IP in HBV-infected HegG2^{hNTCP} supernatant (F-G) sEVS-IP in Patient serum. Graphs show the mean ± SD of 3 independent experiments.

Figure 5. Characterization of RNA in each gradient fraction by HBV 5' RACE. (A) HepG2^{hNTCP} were infected at MOI 500 and RNA was extracted using Trizol reagent from concentrated whole supernatant. Fraction F4, F7 and F9 of density gradient were subjected to HBV-5'RACE. (B-D) RNA from patients' serum was extracted with Trizol reagent and subjected to HBV-5'RACE

Supplementary Figure 1: Evaluation of Lamivudine treatment (3TC) effect on intracellular and extracellular viral parameters. HepG2^{hNTCP} cells were infected with HBV at MOI 250 and treated or not with 3TC (10µM) 3- and 5-days post infection (dpi). Cell pellets were collected 7dpi and nucleic acids were extracted. Total HBV DNA (A), cccDNA (B), 3.5 kb-RNA (C) were quantified by qPCR with HBV specific primers. HBV-DNA and cccDNA were normalized over β -Globin and HBV-RNA was normalized over GUSB expression. Graphs show the mean $\pm$ SD of 3 independent experiments. HBsAg (D) and HBeAg (E) were measured by ELISA. Total HBV-DNA (F) and 3.5 kb HBV-RNA (G) were respectively quantified by ddPCR. Graphs show the mean $\pm$ SD of 7 independent experiments.

Supplementary Figures 2. MNase protection assay of sEVs fraction. Supernatant from HepAD38 was concentrated and separated through density gradient. MNase digestion step was added and reaction was stopped by trizol addition. (A) DNA Plasmid was added to monitor MNase digestion. (B) HBV DNA and (C) RNA were quantified by ddPCR.

References

Alenquer, M., and Amorim, M.J. (2015). Exosome Biogenesis, Regulation, and Function in Viral Infection. Viruses 7, 5066–5083.

Baglio, S.R., van Eijndhoven, M.A.J., Koppers-Lalic, D., Berenguer, J., Lougheed, S.M., Gibbs, S., Léveillé, N., Rinkel, R.N.P.M., Hopmans, E.S., Swaminathan, S., et al. (2016). Sensing of latent EBV infection through exosomal transfer of 5'pppRNA. Proc. Natl. Acad. Sci. U. S. A. 113, E587-596.

Bai, L., Zhang, X., Kozlowski, M., Li, W., Wu, M., Liu, J., Chen, L., Zhang, J., Huang, Y., and Yuan, Z. (2018). Extracellular Hepatitis B Virus RNAs Are Heterogeneous in Length and Circulate as Capsid-Antibody Complexes in Addition to Virions in Chronic Hepatitis B Patients. J. Virol. 92.

Bukong, T.N., Momen-Heravi, F., Kodys, K., Bala, S., and Szabo, G. (2014). Exosomes from hepatitis C infected patients transmit HCV infection and contain replication competent viral RNA in complex with Ago2-miR122-HSP90. PLoS Pathog. 10, e1004424.

Carey, I., Gersch, J., Wang, B., Moigboi, C., Kuhns, M., Cloherty, G., Dusheiko, G., and Agarwal, K. (2019). Pre-genomic HBV RNA and HBcrAg predict outcomes in HBeAg negative chronic hepatitis B patients suppressed on nucleos(t)ide analogue therapy. Hepatol. Baltim. Md.

Charre, C., Levrero, M., Zoulim, F., and Scholtès, C. (2019). Non-invasive biomarkers for chronic hepatitis B virus infection management. Antiviral Res. 169, 104553.

Dreyfuss, G., Matunis, M.J., Piñol-Roma, S., and Burd, C.G. (1993). hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochem. 62, 289–321.

Fanning, G.C., Zoulim, F., Hou, J., and Bertoletti, A. (2019). Therapeutic strategies for hepatitis B virus infection: towards a cure. Nat. Rev. Drug Discov. 18, 827–844.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett. 582, 1977–1986.

Hacker, H.J., Zhang, W., Tokus, M., Bock, T., and Schröder, C.H. (2004). Patterns of circulating hepatitis B virus serum nucleic acids during lamivudine therapy. Ann. N. Y. Acad. Sci. 1022, 271–281.

Jansen, L., Kootstra, N.A., van Dort, K.A., Takkenberg, R.B., Reesink, H.W., and Zaaijer, H.L. (2016). Hepatitis B Virus Pregenomic RNA Is Present in Virions in Plasma and Is Associated With a Response to Pegylated Interferon Alfa-2a and Nucleos(t)ide Analogues. J. Infect. Dis. 213, 224–232.

Klymiuk, M.C., Balz, N., Elashry, M.I., Heimann, M., Wenisch, S., and Arnhold, S. (2019). Exosomes isolation and identification from equine mesenchymal stem cells. BMC Vet. Res. 15, 42.

König, J., Zarnack, K., Luscombe, N.M., and Ule, J. (2012). Protein-RNA interactions: new genomic technologies and perspectives. Nat. Rev. Genet. 13, 77–83.

Kouwaki, T., Fukushima, Y., Daito, T., Sanada, T., Yamamoto, N., Mifsud, E.J., Leong, C.R., Tsukiyama-Kohara, K., Kohara, M., Matsumoto, M., et al. (2016). Extracellular Vesicles Including Exosomes Regulate Innate Immune Responses to Hepatitis B Virus Infection. Front. Immunol. 7, 335.

Lam, A.M., Ren, S., Espiritu, C., Kelly, M., Lau, V., Zheng, L., Hartman, G.D., Flores, O.A., and Klumpp, K. (2017). Hepatitis B Virus Capsid Assembly Modulators, but Not Nucleoside Analogs, Inhibit the Production of Extracellular Pregenomic RNA and Spliced RNA Variants. Antimicrob. Agents Chemother. 61.

Liu, K., and Hu, J. (2019). Secretion of empty or complete hepatitis B virions: envelopment of empty

capsids versus mature nucleocapsids. Future Virol. 14, 95–105.

Longatti, A. (2015). The Dual Role of Exosomes in Hepatitis A and C Virus Transmission and Viral Immune Activation. Viruses 7, 6707–6715.

Otsuka, M., and Koike, K. (2020). Should Level of HBV RNA be Used to Determine When Patients Should Stop Treatment With Nucleos(t)ide Analogues. Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc. 18, 551–552.

Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat. Protoc. 11, 1650–1667.

Prakash, K., Rydell, G.E., Larsson, S.B., Andersson, M., Norkrans, G., Norder, H., and Lindh, M. (2018). High serum levels of pregenomic RNA reflect frequently failing reverse transcription in hepatitis B virus particles. Virol. J. 15, 86.

Salerno, D., Chiodo, L., Alfano, V., Floriot, O., Cottone, G., Paturel, A., Pallocca, M., Plissonnier, M.-L., Jeddari, S., Belloni, L., et al. (2020). Hepatitis B protein HBx binds the DLEU2 lncRNA to sustain cccDNA and host cancer-related gene transcription. Gut 69, 2016–2024.

Sanada, T., Hirata, Y., Naito, Y., Yamamoto, N., Kikkawa, Y., Ishida, Y., Yamasaki, C., Tateno, C., Ochiya, T., and Kohara, M. (2017). Transmission of HBV DNA Mediated by Ceramide-Triggered Extracellular Vesicles. Cell. Mol. Gastroenterol. Hepatol. 3, 272–283.

Stadelmayer, B., Diederichs, A., Chapus, F., Rivoire, M., Neveu, G., Alam, A., Fraisse, L., Carter, K., Testoni, B., and Zoulim, F. (2020). Full-length 5'RACE identifies all major HBV transcripts in HBV-infected hepatocytes and patient serum. J. Hepatol. 73, 40–51.

Théry, C., Zitvogel, L., and Amigorena, S. (2002). Exosomes: composition, biogenesis and function. Nat. Rev. Immunol. 2, 569–579.

Wang, J., Shen, T., Huang, X., Kumar, G.R., Chen, X., Zeng, Z., Zhang, R., Chen, R., Li, T., Zhang, T., et al. (2016). Serum hepatitis B virus RNA is encapsidated pregenome RNA that may be associated with persistence of viral infection and rebound. J. Hepatol. 65, 700–710.

Willms, E., Cabañas, C., Mäger, I., Wood, M.J.A., and Vader, P. (2018). Extracellular Vesicle Heterogeneity: Subpopulations, Isolation Techniques, and Diverse Functions in Cancer Progression. Front. Immunol. 9, 738.

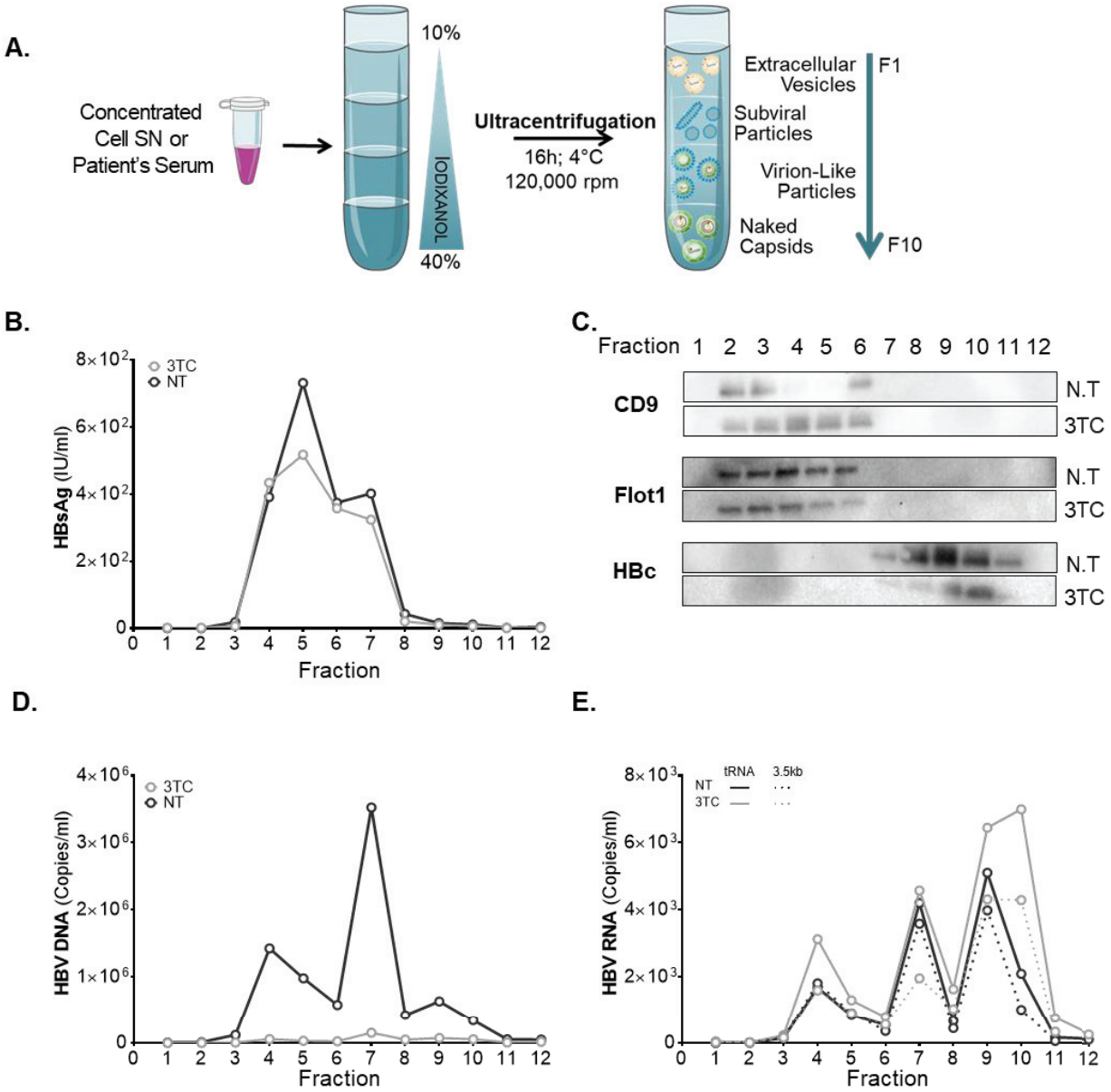


Figure 1 - Secreted HBV RNA is associated to small extracellular vescicles in supernatant of HBV-infected HepG2^{hNTCP} cells (A) Scheme of 10-40% Iodixanol/STE density gradient and theorical small extracellular vescicles (Evs) distribution according to their buoyant density. Briefly, HepG2^{hNTCP} cells were infected at MOI 250 and supernatant was collected 8 dpi. Infected cells were treated (3TC) or not (N.T) with lamivudine (3TC) at 10µM for 5 days. (B) HBsAg ELISA was measured in each fraction. (C) Exosome marker, Flotillin-1 (Flot-1) and HBV core protein (HBc) were detected by western blot in each fraction. (D) Total HBV-DNA, (E) total HBV-RNA (tHBV-RNA) and 3.5 kb RNA were quantified by ddPCR in each fraction.

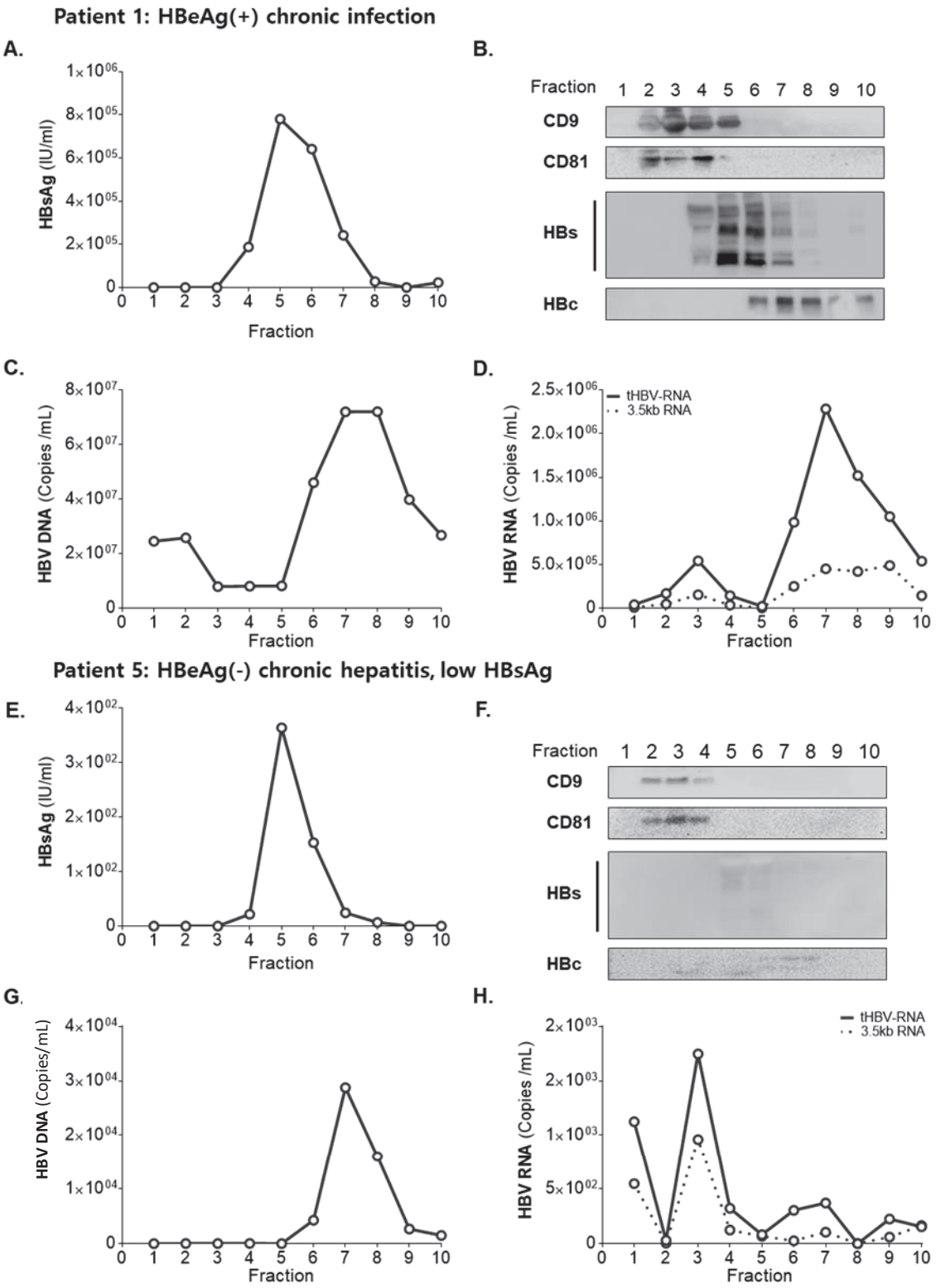


Figure 2 - Analysis of viral parameters and EVs markers in gradient fractions of serum samples derived from patients in different CHB phases.

Gradient fractions were obteined according to the protocol detailed in Method section. a) HBsAg ELISA. b) Representative Western Blot analysis using anti-extracellular vesicle (EVs) markers (anti-CD9 and anti-CD81), anti-HBsAg and anti-HBV core antibody. c) HBV DNA quantification by qPCR. d) HBV RNA quantification by RT-qPCR. HBx ORF common region was used to quantify both HBV DNA and total HBV RNA. pre-Core/pgRNA 5' unique region was used to quantify 3.5 kb HBV RNA (Testoni et al. JHep 2019). CHB phases are defined according to 2017 EASL CPGs.

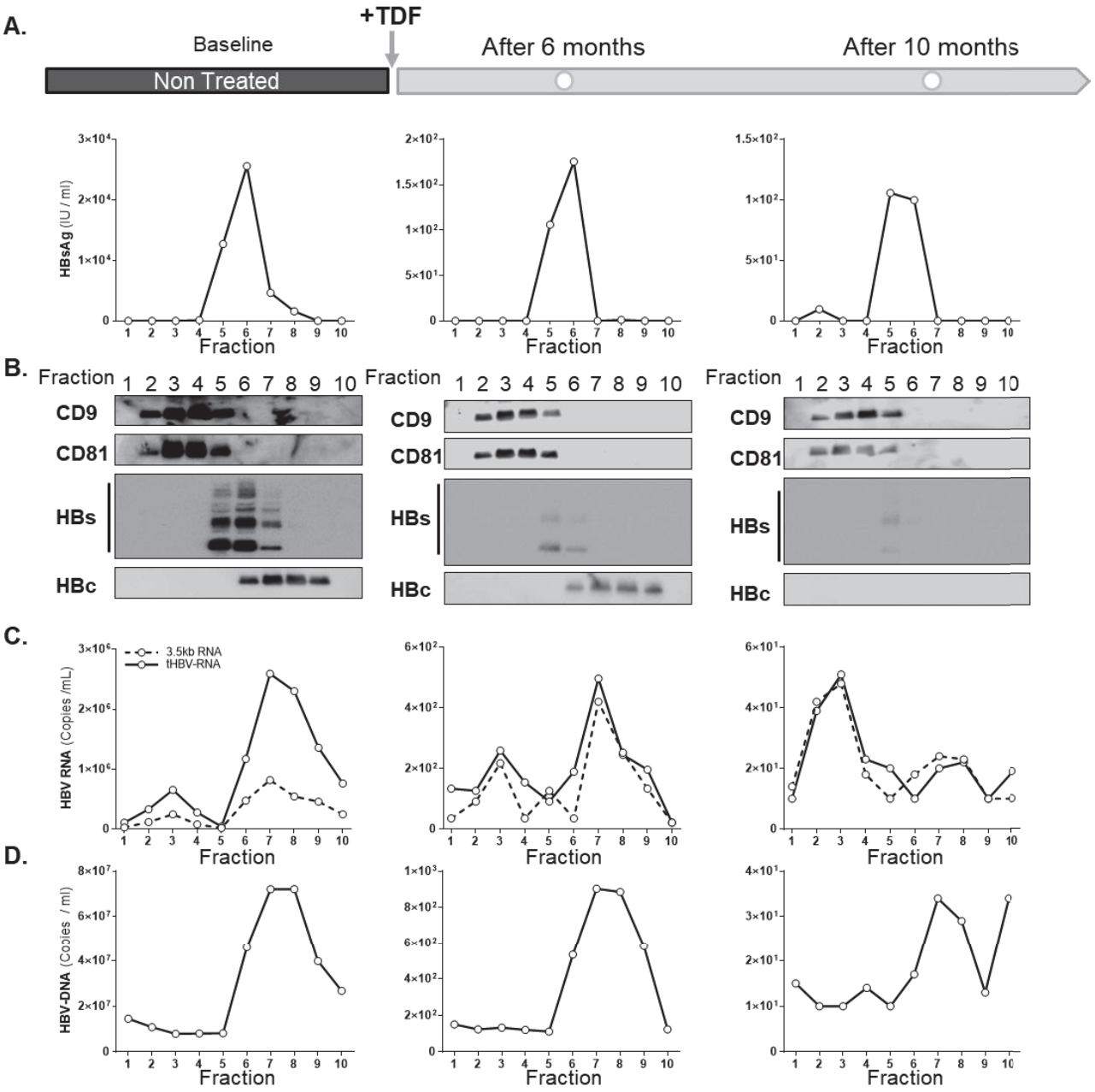


Figure 3. Analysis of viral parameters and EVs markers in gradient fractions of longitudinal serum samples derived from a Tenofovir-treated patient.

Gradient fractions were obteined according to the protocol detailed in Method section. a) HBsAg ELISA. b) Representative Western Blot analysis using anti-extracellular vesicle (EVs) markers (anti-CD9 and anti-CD81), anti-HBsAg and anti-HBV core antibody. c) HBV RNA quantification by RT-qPCR. HBx ORF common region was used to quantify both HBV DNA and total HBV RNA. pre-Core/pgRNA 5' unique region was used to quantify 3.5 kb HBV RNA (Testoni et al. JHep 2019). d) HBV DNA quatification by qPCR

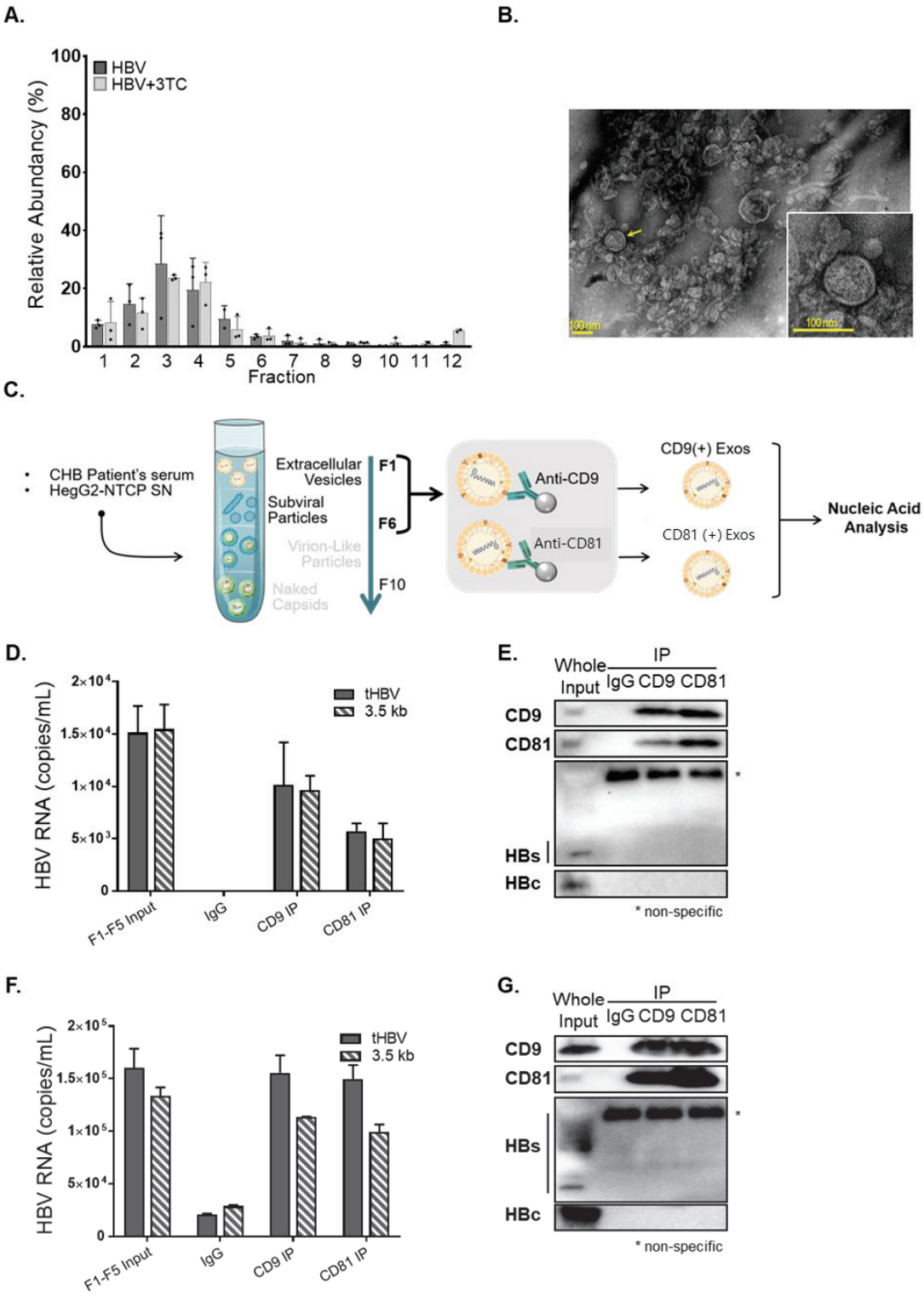


Figure 4. Characterization of sEVs fraction. Supernatant from HBV-infected HepG2^{hNTCP} cells was separated through density gradient. (A) Particles were quantified and size was assessed in 3TC- treated and non-treated samples by Nanoparticle Tracking Analysis. (B) Negative staining of fraction 4 in electron microscopy (PTA 3%). (C) Schematic protocol of sEVs-Immunoprecipitation (IP) (D-E) sEVS-IP in HBV-infected HegG2^{hNTCP} supernatant (F-G) sEVS-IP in Patient serum. Graphs show the mean ± SD of 3 independent experiments

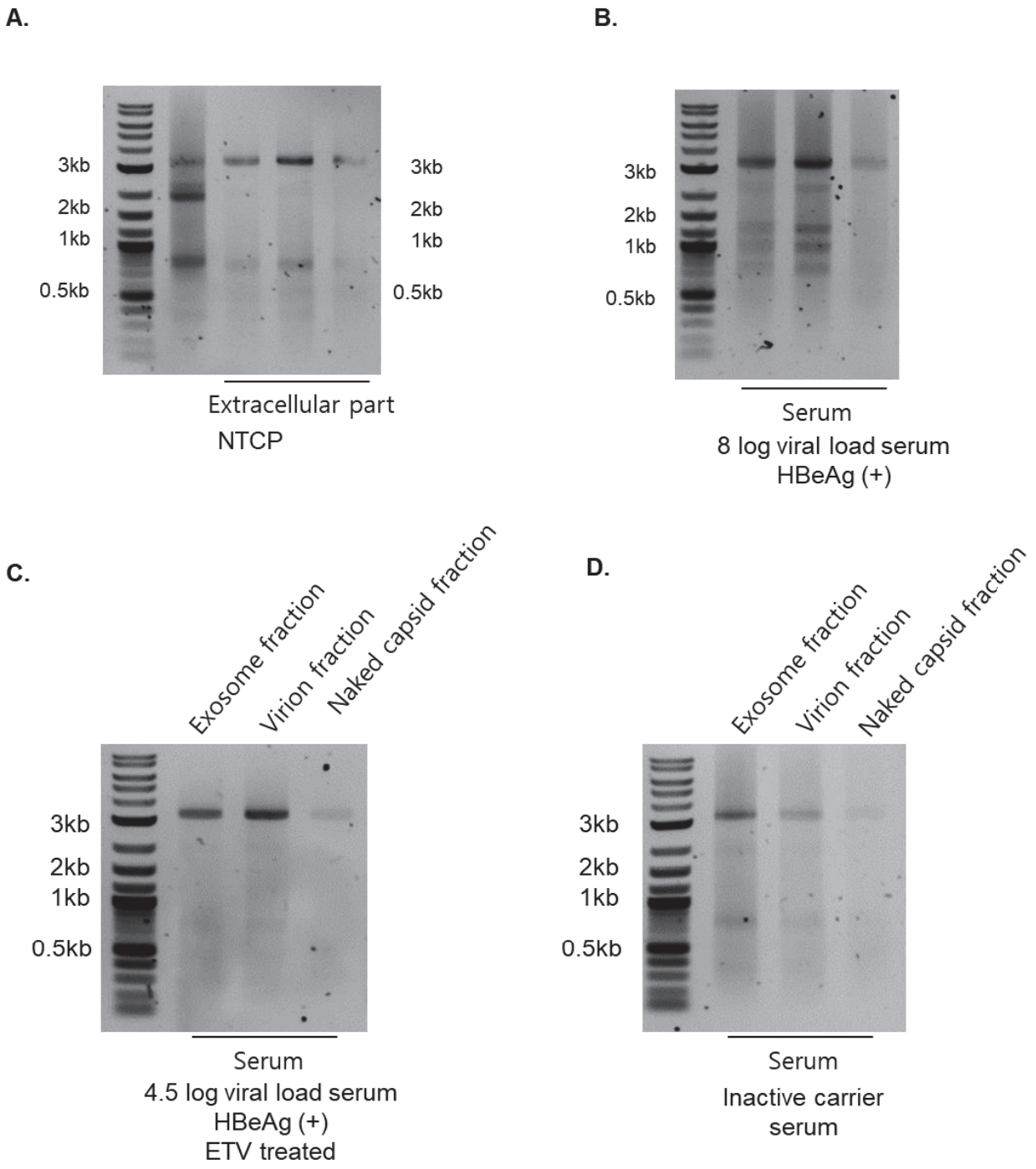
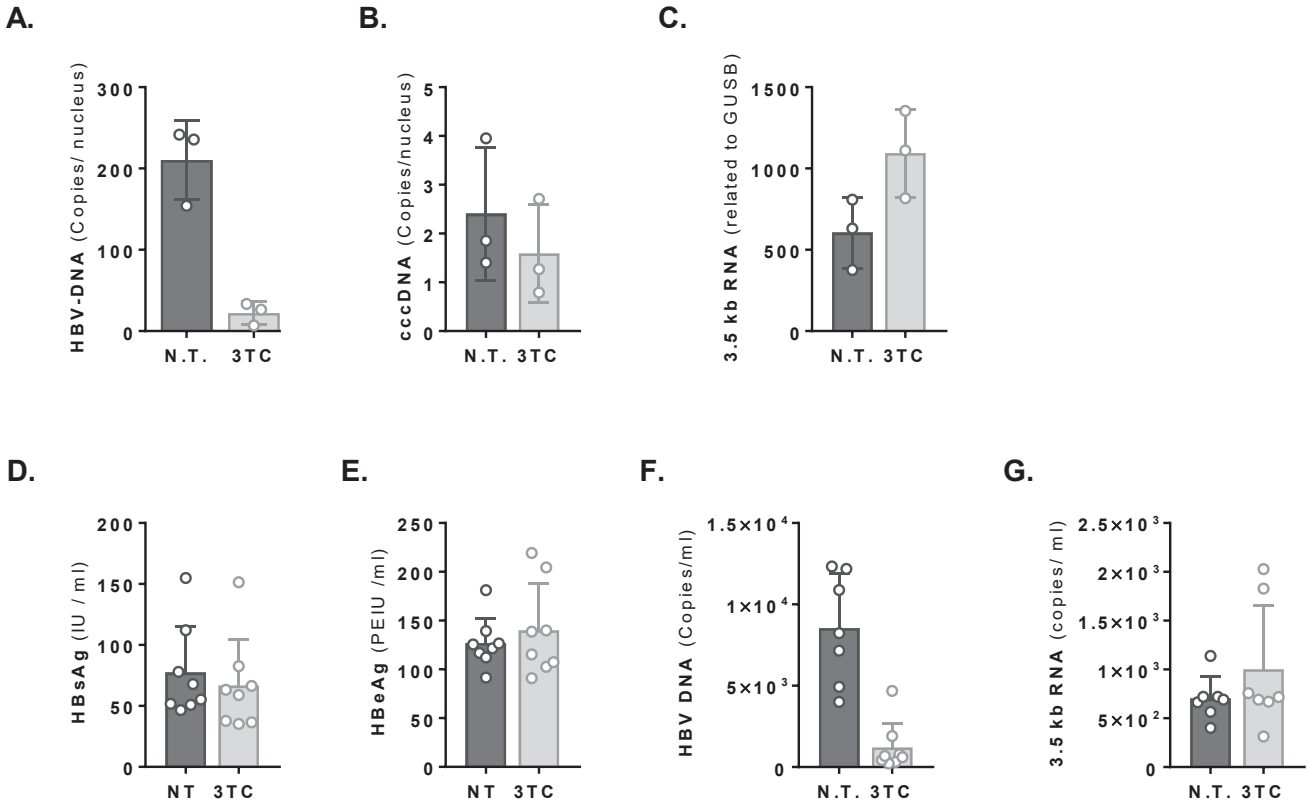
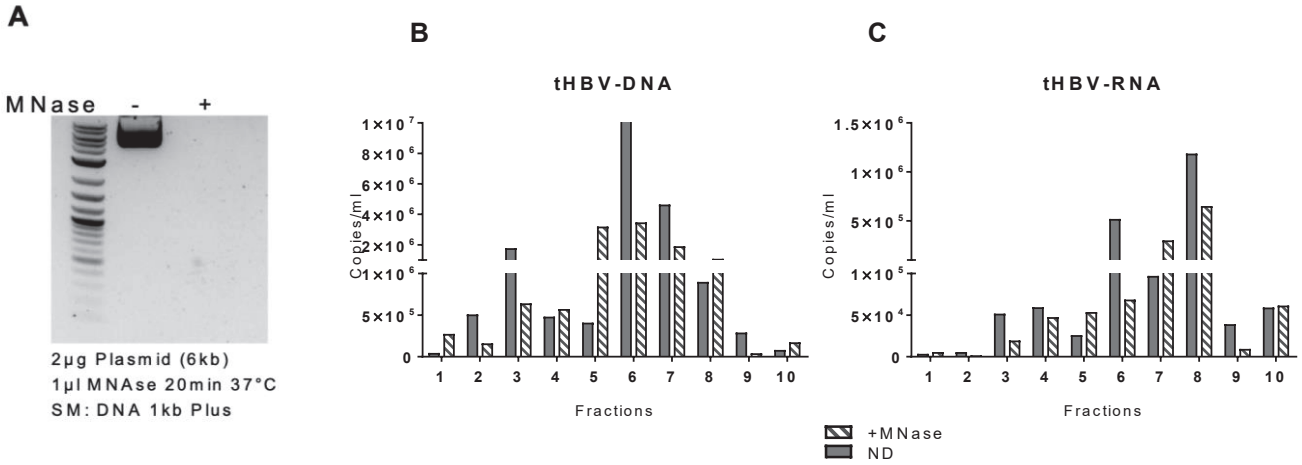


Figure 5. Characterization of RNA in each gradient fraction by HBV 5' RACE. (A) HepG2^{hNTCP} were infected at MOI 500 and RNA was extracted using Trizol reagent from concentrated whole supernatant. Fraction F4, F7 and F9 of density gradient were subjected to HBV-5'RACE. (B-D) RNA from patients' serum was extracted with Trizol reagent and subjected to HBV-5'RACE



Supplementary Figure 1: Evaluation of Lamivudine treatment (3TC) effect on intracellular and extracellular viral parameters. HepG2^{hNTCP} cells were infected with HBV at MOI 250 and treated or not with 3TC (10µM) 3- and 5-days post infection (dpi). Cell pellets were collected 7dpi and nucleic acids were extracted. Total HBV DNA (A), cccDNA (B), 3.5 kb-RNA (C) were quantified by qPCR with HBV specific primers. HBV-DNA and cccDNA were normalized over β-Globin and HBV-RNA was normalized over GUSB expression. Graphs show the mean ± SD of 3 independent experiments. HBsAg (D) and HBeAg (E) were measured by ELISA. Total HBV-DNA (F) and 3.5 kb HBV-RNA (G) were respectively quantified by ddPCR. Graphs show the mean ± SD of 7 independent experiments.



Supplementary Figures 2. MNase protection assay of sEVs fraction. Supernatant from HepAD38 was concentrated and separated through density gradient. MNase digestion step was added and reaction was stopped by trizol addition. (A) DNA Plasmid was added to monitor MNase digestion. (B) HBV DNA and (C) RNA were quantified by ddPCR.

Table. 1

Patient's No.	DNA viral load (log IU/ml)	HBe status	HBsAg quantity (UI/ML)	ALT	Anti-HBV Treatment	Geno -type
P1	8.9	P(+)	51000	27	NO	B
P2	8	P(+)	12000	42	NO	ND
P3	8.5	P(+)	81000	36	NO	D
P4	8.7	P(+)	98000	30	NO	E
P5	5.6	N(-)	570	158	NO	D
P6	4.1	N(-)	9100	30	NO	D
P7	4.5	N(-)	7200	40	NO	D
P8	2.94	N(-)	633.2	27	NO	D
P9	2.13	N(-)	160	30	NO	B
H1	-	-	-	17	-	-
H2	-	-	-	32	-	-
P10	4.5	P(+)	16000	123	Entecavir	D

Table. 2

Patient's No.	DNA viral load (log IU/ml)	HBe status	HBsAg quantity (UI/ML)	ALT	Anti-HBV Treatment	Geno-type	Follow-up date
P11 (F 1/3)	8.3	P(+)	37000	201	NO	C	04/07/2019
P11 (F 2/3)	3.1	P(+)	270	27	Tenofovir	C	09/01/2020
P11 (F 3/3)	Lo Q	P(+)/N(-)	240	24	Tenofovir	C	07/05/2020
P12 (F 1/3)	5.7	N(-)	6300	43	NO	D	19/03/2019
P12 (F 2/3)	2.4	N(-)	ND/P(+)	43	Entecavir	D	19/07/2019
P12 (F 3/3)	Lo Q	N(-)	4300	41	Entecavir	D	14/11/2019

F1-3/3: Number of Follow up

Lo Q: DNA detectable but below Low quantity

P(+)/N(-): Hospitals positive, laboratory negative

ND: not tested

CONCLUSIONS AND DISCUSSION

The development of new circulating biomarkers able to faithfully reflect the size and the transcriptional activity of the intrahepatic HBV cccDNA is an important objective in the contexts of *HBV cure* efforts and could greatly contribute to accompany the early clinical development of the many NMEs that aim to silence or decrease the pool of cccDNA to achieve a *functional* or *complete cure* with finite treatment duration.

Circulating RNA from HBV have been shown to be a good non-invasive biomarker of cccDNA activity and persistence and useful to evaluate the virological activity of HBV NMEs as well as to monitor chronic hepatitis B patients treated with the NUCs (e.g., decision to stop treatment) or IFNs (e.g., identification of *responders*).

My PhD Thesis project was part of the larger CirB-RNA project, a PIA-RHU project funded in 2017 by the French National Agency (ANR) and conducted in Pr. Zoulim's and Pr. Levrero's teams, in collaboration with Roche diagnostic (Pleasanton, California) as industrial partner.

The general aims of the CirB-RNA project are to characterize circulating HBV RNAs, validate their clinical relevance and develop the related diagnostic test. The specific objectives of my PhD thesis work, as part of the CirB-RNA effort, have been centered on : a) the generation and characterization of a molecular standard that could be used with all PCR-based HBV RNAs assays; b) to develop a

robust HBV DNA and RNA capture-sequencing workflow to study the potential impact of HBV integration on circulating HBV RNA assays and more broadly to evaluate the HBV integration burden at the DNA and RNA level in CHB patients; c) the characterization of the exact nature of secreted/circulating RNAs and their compartmentalization in *in vitro* HBV infection models and in CHB patients.

The current WHO HBV DNA standard has been used for the quantification of serum HBV RNA in some of the early reports. Its use has been challenged because the WHO HBV DNA is derived from a clinical sample, which implies the presence of unknown quantities of HBV RNA, and the lack of evidence regarding the equivalence of an HBV RNA unit to the international HBV DNA unit. Therefore, a viral RNA standard could be a good candidate as an international reference standard. Aiming to develop an ***RNA molecular standard*** to be used to calibrate PCR-based circulating HBV RNA quantification assays, we have generated several clonal cell lines transfected with HBV mutated genomes that secrete large quantities of viral RNAs. We selected the Huh7-3D29 clonal cell line, that carried a double mutation in the catalytic site (YMAA mutation) and the TP-domain (Y63F) of the polymerase, and displayed the desired RNA secretory phenotype (e.g., inversion of the secreted HBV DNA/RNA ratio). The RNA secretory phenotype of Huh7-3D29 cells, with an exceedingly high RNA/DNA ratio in cell supernatants, was stable over time. The residual HBV DNA secretion is negligible considering that the HBV RNA cobas® 6800/8800 automated investigational assay and the related Roche Manual Workflow (MWR) assay used in our experiments tolerate DNA:RNA ratios up to approximately 10^6 without losing specificity and linearity. It is important to underline that the automated HBV RNA cobas investigational assay and its manual version use the same chemistry and have the same performing characteristics (HBV genotypes inclusivity, linearity, lower limit of quantification and lower limit of detection). Notably, the standard curves obtained from 10-fold dilution of Huh7-3D29 supernatants quantified with the Roche MWF display the same profile of those obtained with a synthetic HBV armored RNA. The observation that HBV RNA quantification of Huh7-3D29 supernatants with the Roche MWF showed no significant difference with and without DNase treatment, supports the use of the Huh7-3D29 derived HBV RNA standard with both automated and manual assays. Using iodixanol/sucrose gradients to characterize the HBV RNA secreted by Huh7-3D29 cells, we found that the majority of secreted HBV RNAs are encapsidated and are found in naked capsid and in virions-like particles, but HBV RNAs are also present in secreted Extracellular Vesicles. Single molecule long reads Nanopore sequencing indicates that the majority of the Huh7-

3D29 secreted HBV RNAs are full length or spliced 3.5 kb species and terminate at the canonical HBV polyA site. All these results support the potential use of Huh7-3D29 supernatant as a standard for all HBV RNA investigational assays and *in house* PCR-based HBV RNA assays that could be adopted as an international standard.

The requirement for a WHO international standard is a lyophilized preparation that has proven long-term stability, and which is available in adequate quantities. The Huh7-3D29 cell line constitutes a potentially unlimited source of HBV RNA standard. Indeed, HBV RNA quantities produced from one 175cm^2 flask of Huh7-3D29 clonal cell line in 9 days provides enough material for 1,300 standard curves using 20 µL of 10^6 copies/mL preparations. In the case of the HBV Nucleic Acids WHO international standard it was produced 3 times over time with 2000 aliquots of 0.5 mL and a concentration of 1 x 10^6 IU/mL (Baylis et al., 2008; Fryer et al., 2017; Saldanha et al., 2001). Since Huh7-3D29 HBV RNA standard production can be further scaled up, the requirements in terms of quantity can be easily met. Next steps in the development of the Huh7-3D29 HBV RNA standard into an internationally established standard or a WHO endorsed standard Huh7-3D29 material will be the validation of its performance by several laboratories and the analysis of the results by the expert committee on biological standardization of the WHO.

The predictive nature of circulating RNAs in relevant clinical settings is more and more established but the impact of HBV sequences integrated into the human genome as a potential source of circulating HBV RNAs is not well defined. To study the ***impact of HBV integration on circulating HBV RNA assays***, we developed a robust HBV DNA and RNA capture-sequencing workflow that uses a multiplexed short reads NGS approach for the detection and quantification of DNA host-virus junctions and RNA virus-host fusion transcripts. Our workflow consistently detects and maps HBV integrations and HBV transcripts and is sensitive enough to detect integrations in patients' liver samples. Notably, in the non tumor liver sample from liver resection in an HBV-related HCC patient we detected HBV integration in the TERT, UNC5D and N4BP2L2 genes that code for proteins well know to contribute cell transformation and cell death regulation.

The results obtained in the model cell lines PLC/PRF/5 and Hep3B were consistent with those recently reported in the same cell line using the long reads PacBio technology, coupled or not with an HBV capture enrichment. Notably, our workflow generates workable results with a relatively low number of captured HBV reads that allows to use the *common capture* protocol with Illumina MiSeq

instruments, that are quite widely diffused in research and reference clinical laboratories. The possibility of multiplexing up to 6 allows to limit costs.

Next steps will include: a) a further streamligthing of the bioinformatics pipeline to optimize the balance between sensitivity and specificity of the hybrid reads calling; b) the analysis *in depth* of the RNA-Seq results from the 15 capture-sequencing experiments and the comparison of the results obtained in the model cell lines PLC/PRF/5 and Hep3B with the more recent published NGS results; c) the study of more human samples from chronic HBV patients and HCC patients; d) the adaptation of our capture-sequencing workflow to the analysis of plasma/serum samples.

The use of circulating HBV RNA as a biomarker requires a deep knowledge on the exact nature of circulating RNAs (3.5 kb species, subgenomic viral RNAs), their compartmentalization (enveloped viral particles, naked capsids, subviral HBs particles, extracellular vesicles) and their distribution in the different compartments in the different stages of the disease.

A systematic approach to the characterization of circulating Hepatitis B virus RNAs *in vitro* and chronic hepatitis B patients has been conducted within the frame of the CirB-RNA project. Multiple approaches have been used, including density gradients, ddPCR, 5' RACE PCR and single molecule long read sequencing using the Nanopore technology. CirB-RNAs are mainly 3.5 kb RNAs and more precisely pgRNA, but that other RNA species are also secreted. Secretion is supported by extracellular vesicles (EVs), enveloped virion-like particles (VLP) and naked capsids (NC) in HBV infected-HepG2-NTCP cell and patient's sera. Notably, untreated patients with low HBV viral replication and patients treated with nucleos(t)ide analogues showed a CirB-RNA secretion shifts from VLPs/NCs to EVs.

It has been shown that EVs secreted from HBV infected cell can contain HBV rcDNA and potentially infectious for surrounding cells (Sanada et al., 2017). Likewise, it is possible that viral RNAs secreted in EVs in CHB patients treated withg NUCs or in CHB patients with low viremia may contribute to the maintenance of viral infection through the infection of new hepatocytes. In this respect, it is important to underline that NUCs inhibit HBV DNA replication but do not suppress cccDNA transcriptional activity that continuously generates HBV RNAs (Fanning et al., 2019).

Altogether the results further support the rationale for CirB-RNAs as a novel viral biomarker for indirect measurement of cccDNA transcriptional activity.

APPENDICES

Hepatitis B protein HBx binds the DLEU2 lncRNA to sustain cccDNA and host cancer-related gene transcription

Hepatology

OPEN ACCESS

ORIGINAL RESEARCH

Hepatitis B protein HBx binds the DLEU2 lncRNA to sustain cccDNA and host cancer-related gene transcription

Debora Salerno,[1] Letizia Chiodo ,[2] Vincenzo Alfano ,[3] Oceane Floriot,[3] Grazia Cottone,[4] Alexia Paturel,[3] Matteo Pallocca,[5] Marie-Laure Plissonnier,[3] Safaa Jeddari,[6] Laura Belloni,[6] Mirjam Zeisel,[3] Massimo Levrero ,[3,6] Francesca Guerrieri [1,3]

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2019-319637).

For numbered affiliations see end of article.

Correspondence to
Dr Francesca Guerrieri, Center for Life NanoScience@Sapienza, Istituto Italiano di Tecnologia Center for Life Nano Science, Roma 00162, Italy; fraguerrieri@gmail.com and Professor Massimo Levrero, Cancer Research Center of Lyon (CRCL), Lyon, France; massimo.levrero@inserm.fr

DS and LC are joint first authors.

Received 14 August 2019
Revised 29 January 2020
Accepted 29 January 2020
Published Online First 29 February 2020

ABSTRACT

Objective The HBV HBx regulatory protein is required for transcription from the covalently closed circular DNA (cccDNA) minichromosome and affects the epigenetic control of both viral and host cellular chromatin.

Design We explored, in relevant cellular models of HBV replication, the functional consequences of HBx interaction with DLEU2, a long non-coding RNA (lncRNA) expressed in the liver and increased in human hepatocellular carcinoma (HCC), in the regulation of host target genes and the HBV cccDNA.

Results We show that HBx binds the promoter region, enhances the transcription and induces the accumulation of DLEU2 in infected hepatocytes. We found that nuclear DLEU2 directly binds HBx and the histone methyltransferase enhancer of zeste homolog 2 (EZH2), the catalytic active subunit of the polycomb repressor complex 2 (PRC2) complex. Computational modelling and biochemical evidence suggest that HBx and EZH2 share two preferential binding sites in DLEU2 intron 1. HBx and DLEU2 co-recruitment on the cccDNA displaces EZH2 from the viral chromatin to boost transcription and viral replication. DLEU2-HBx association with target host promoters relieves EZH2 repression and leads to the transcriptional activation of a subset of EZH2/PRC2 target genes in HBV-infected cells and HBV-related HCCs.

Conclusions Our results highlight the ability of HBx to bind RNA to impact on the epigenetic control of both viral cccDNA and host genes and provide a new key to understand the role of DLEU2 and EZH2 overexpression in HBV-related HCCs and HBx contribution to hepatocytes transformation.

Significance of this study

What is already known on this subject?

- HBx is recruited to and regulates the activity of a variety of coding genes and non-coding RNA promoters.
- Long non-coding RNAs (lncRNAs) are frequently dysregulated in both HBV-related hepatocellular carcinoma (HCC) tissues and HBV-expressing/HBx-expressing cell lines.
- LncRNAs can regulate multiple cellular processes by various mechanisms including chromatin remodelling.
- A number of lncRNAs has been described to guide chromatin-modifying complexes, including enhancer of zeste homolog 2 (EZH2)/polycomb repressor complex 2 (PRC2).

What are the new findings?

- HBx binds the promoter region of the DLEU2 lncRNA, enhances DLEU2 transcription and induces the accumulation of DLEU2 RNA.
- HBx directly binds DLEU2 and functionally competes with EZH2 for partially overlapping sites on the DLEU2.
- HBx and DLEU2 modulates transcription from the HBV covalently closed circular DNA minichromosome, hence boosting HBV replication, and from host genes promoters.
- HBx cooperates with DLEU2 by activating the promoters of a subset of EZH2/PRC2 targets in HBV-replicating cells and in HBV-related HCC.

How might it impact on clinical practice in the foreseeable future?

- These observations:
  - provide a new key to understand the role of EZH2 overexpression in HBV-related HCCs and the impact of HBV and HBx in hepatocyte transformation;
  - identify the interaction between HBx and DLEU2 as a new potential target to both inhibit HBV replication and hepatocyte transformation.

INTRODUCTION

Despite the availability of a preventive vaccine and antiviral treatments that arrest disease progression and reduce liver cancer risk, hepatitis B still remains a major global public health problem with about 257 million persons living with chronic HBV infection.[1 2] HBV accounts for over 850 000 deaths/year and is responsible for >50% of the hepatocellular carcinomas (HCCs) worldwide.[3] Currently, only 10% of chronically infected people are diagnosed and only 1% are treated.[1] Hepatitis B surface antigen (HBsAg) loss is viewed as an important

Check for updates

To cite: Salerno D, Chiodo L, Alfano V, *et al*. *Gut* 2020;**69**:2016–2024.

goal of therapy,[4] but it is achieved only in a minority of patients treated with pegylated interferon alpha2a and nucleos(t)ide analogues (NUCs) and life-long treatments are recommended with NUCs.[5] HBsAg loss is not accompanied by viral eradication[4] and most patients never clear the HBV covalently closed circular DNA (cccDNA) and bear viral sequences integrated into the host DNA.[6] The cccDNA minichromosome is the matrix for transcription of all viral mRNAs and the 3.5 kb pregenomic RNA (pgRNA) that serves as template for viral replication.[7] The level of HBV replication is determined by the size of the cccDNA pool (ie, total number of cccDNA molecules) and its transcriptional productivity (ie, the amount of HBV RNAs transcribed from each cccDNA).[7] cccDNA transcription is regulated by epigenetic modifications imposed to cccDNA-bound histones[8] by viral and cellular proteins as well as inflammatory cytokines.[7] HBx plays a key role through two independent mechanisms that cooperate to establish and maintain a transcriptionally active cccDNA minichromosome: (a) the degradation of the Smc5/6 restriction factors mediated by the activation of the DDB1-Cul4 E3 ligase early after infection[9] and (b) the prevention of transcriptional repressor recruitment on the cccDNA.[10 11] In addition to viral chromatin, HBx is recruited to and regulates a variety of coding genes and non-coding RNA promoters.[12]

Long non-coding RNAs (lncRNAs) are frequently dysregulated in HBV-related HCC and HBV-expressing/HBx-expressing cells.[13] LncRNAs play a crucial role in the modulation of signalling pathways and gene expression.[14] Viral infections may induce cellular lncRNAs with antiviral effect[15] and viruses can use lncRNAs to regulate metabolic networks promoting viral survival.[16] lncRNAs have been described to remodel chromatin[14] by guiding chromatin-modifying complexes, including enhancer of zeste homolog 2 (EZH2)/polycomb repressor complex 2 (PRC2), towards their target loci.[17 18] EZH2, the major cellular H3K27 trimethyltransferase, catalyses addition of methyl groups at lysine 27 of histone H3 (H3K27me3).[19] EZH2, as a subunit of the canonical PRC2, along with EED, SUZ12 and RBBP4,[19] is primarily involved in repression of transcription. Multiple DNA-binding proteins, including JARID2[20] and ATRX,[21] contribute to PRC2 complex recruitment to target genes to silence gene expression. EZH2 is overexpressed in many cancers including HCC.[22 23] Small molecule EZH2 inhibitors that target its catalytic SET domain are in clinical trials.[24] However, EZH2 catalytic activity may not recapitulate all EZH2/PRC2 functions in cancer and EZH2/PRC2 recruitment can be uncoupled from H3K27me3 deposition/spreading.[25] EZH2 also acts as a transcriptional co-activator of oestrogen-regulated, WNT-regulated and NOTCH-regulated proliferation pathways in breast cancer.[26 27]

Here we show, in relevant cellular models of HBV replication, that HBx activates the expression and directly binds the lncRNA DLEU2 to modulate transcription from the HBV cccDNA minichromosome, hence boosting HBV replication, and from host genes promoters. HBx-mediated increase in DLEU2 and HBx recruitment onto target gene regulatory sequences increases chromatin accessibility and activates a subset of EZH2/PRC2 targets in HBV-replicating cells and in HBV-related HCCs. These observations provide a new key to understand the role of DLEU2 and EZH2 overexpression in HBV-related HCCs and the impact of HBx in hepatocyte transformation.

RESULTS

DLEU2 lncRNA is a direct target of HBx upregulated in HBV infection and HBV-related HCC

Several lncRNAs have been shown to be deregulated in HBV infection and in HBV-related HCCs.[13 28] Recently, using a genome-wide analysis of HBx chromatin recruitment in HBV-replicating cells, we identified 16 lncRNA promoters and 26 lncRNA intragenic regions specifically bound by HBx[12] (online supplementary table S1). The expression of the 34 unique lncRNAs that are directly targeted by HBx (figure 1A, online supplementary table S1) and of 15 additional lncRNAs that are known to be modulated in human HCCs (figure 1B) has been evaluated in mock-infected and HBV-infected primary human hepatocytes (PHHs) using a custom Nanostring-GE panel assay. As shown in figures 1A, 9 out of the 34 HBx-targeted lncRNAs were significantly upregulated and 4 were strongly downregulated by HBV at day 7 postinfection. Notably, 8 out of 15 HCC-modulated lncRNAs were strongly upregulated by HBV infection (figure 1B). Among the lncRNAs with the highest levels of induction in HBV-infected PHHs, DLEU2, LINC00441, LINC00478 are direct targets of HBx (ie, positive in the anti-HBx chromatin immunoprecipitation (ChIP)), whereas for HULC, HEIH, H19 and MALAT1 we do not have direct evidence of HBx promoter recruitment.[12] We focused on DLEU2, which is differentially expressed during liver maturation[29] and overexpressed in solid human cancers[30] including HCCs[28 31] (online supplementary figures S1a-b).

DLEU2 gene extends for 142 990 bp in Ch13(q14.2) and overlaps with the miR-15a/miR-16-1 cluster, the TRIM13 gene and, partially, with the DLEU1 gene (figure 1C). The *HBx peak* detected by HBx chromatin immunoprecipitation sequencing (ChIP-Seq)[12] (figure 1C, black circle) is located ~6500 bp downstream the DLEU2 promoter previously described.[32] HBx binding on the *HBx peak* region was confirmed in independent ChIP assays performed in HBV-infected PHHs (figure 1D, left panel) and was accompanied by an increase in polymerase II (Pol II) binding (figure 1D, right panel) and DLEU2 steady state RNA levels (figure 1E–F). HBx recruitment on the DLEU2 promoter and increased DLEU2 expression were confirmed in other HBV infection/replication systems (online supplementary figure S2a-c). Figure 1G shows baseline levels of DLEU2 quantified by ddPCR in PHHs, HepG2 and HepG2-NTCP cells as well as in HBV-related HCCs and their paired HBV-infected non-tumorous liver tissues. The increase in DLEU2 RNA containing intron 1 sequences (figure 1E–G, right panels) and the accumulation of DLEU2 in the nuclear fraction of HBV-infected PHHs (figure 1H) as well as in HBV-replicating (online supplementary figures S2d) and HBx-HA-transfected HepG2 cells (online supplementary figure S2e), all supports the notion that DLEU2 induction by HBV occurs at the transcriptional level. The increase of DLEU2 levels after exogenous HBx expression in HepG2 cells (online supplementary figure S2e) and the loss of DLEU2 induction in PHHs infected with the HBx-deficient HBV (x-) virus (figure 1H) confirm the key role of HBx in increasing DLEU2 levels. Finally, HBx recruitment on the *HBx peak* region has no impact on the levels of the antisense DLEU1 lncRNA (online supplementary figure S2f).

DLEU2 lncRNA physically interacts with EZH2 and HBx

LncRNAs have been shown to guide chromatin-modifying complexes towards their chromatin targets and to regulate gene expression.[14 33] Multiple lncRNAs interact in vivo with the EZH2/PRC2 complex in embryonic stem cells (ESCs).[17] In EZH2 RIP-Seq performed in colon cancer cells,[18] DLEU2 is highly enriched, together with the lncRNAs XIST, TSIX, MALAT1 and GAS5. Since HBx has been reported to modulate EZH2/PRC2 host target gene transcription,[34 35] we sought to further investigate whether DLEU2 interacts with EZH2 and HBx in PHHs and

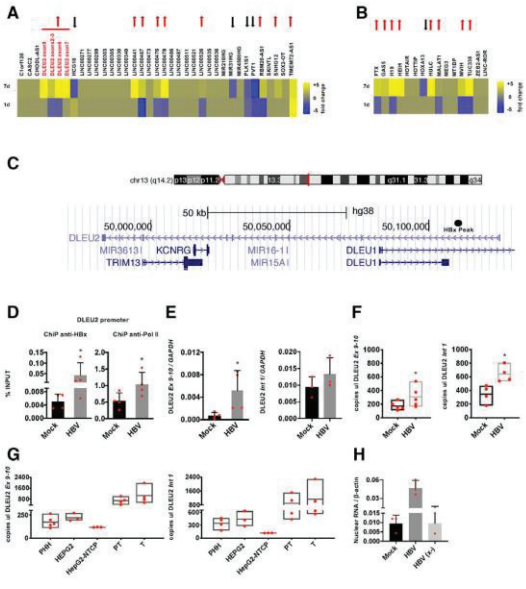


Figure 1 HBx modulates DLEU2 expression. (A) Heatmap of the expression levels in HBV-infected primary human hepatocytes (PHHs) of 34 long non-coding RNAs (lncRNAs) that are direct transcriptional targets of HBx[12] analysed by Nanostring. Blue and yellow colours in the heatmap indicate downregulated and upregulated lncRNAs, respectively. The red and black arrows identify lncRNAs upregulated ≥1.5-fold and downregulated ≤0.7-fold. LncRNA expression in HBV-infected samples is normalised to GAPDH and to mock condition. (B) Expression levels of 15 hepatocellular carcinoma (HCC)-associated lncRNAs, in HBV-infected PHHs. Nanostring analysis, data normalisation and heatmap representation as shown in figure 1A. (C) DLEU2 lncRNA genomic sequence, according to Genome Browser; the *HBx peak*, obtained by HBx chromatin immunoprecipitation sequencing (ChIP-Seq) searching, is indicated as black circle. (D) HBx and polymerase II (Pol II) occupancy on DLEU2 promoter region. Cross-linked chromatin from mock or HBV-infected PHHs (72 hours) was immunoprecipitated with anti-HBx or anti-Pol II antibodies, and then analysed by real-time quantitative PCR (qPCR) using specific DLEU2 promoter primer pairs. The detection of HBV covalently closed circular DNA (cccDNA) using specific primers in the ChIPed DNA from HBV-infected cells (cccDNA-ChIP) served as a technical positive control for the ChIP procedure. ChIP results are expressed as % of input. (E) Real-time qPCR of DLEU2 RNA (left panel: exons 9–10; right panel: intron 1) in mock-infected and HBV-infected PHHs (7 days). Results are expressed as relative values to endogenous human GAPDH mRNAs. (F) Droplet Digital PCR (ddPCR) quantification of DLEU2 RNA (left panel: exons 9–10; right panel: intron 1) from mock-infected and HBV-infected PHHs (7 days). Results are expressed as copies for 5 ng/µL of total RNA. (G) ddPCR quantification of DLEU2 RNA (left panel: exons 9–10; right panel: intron 1) from PHHs, HepG2 or HepG2-NTCP cells and from HBV-infected peritumour (PT) or tumour (T) tissue. Results are expressed as copies for 5 ng/µL of total RNA. (H) Real-time qPCR of DLEU2 RNA (intron 1) in nuclear lysates from of mock, HBV-infected and HBV (x-)-infected PHHs. Results are expressed as relative values to endogenous human β-actin mRNAs. (7 days). Data in panels (D), (E) and (F) represent means±SD from at least three independent experiments. In (D), (E) and (F), *p<0.05; Mann-Whitney U test.

HCC cell lines. As shown in figure 2A, we found that DLEU2 and EZH2 interact in PHHs. As compared with mock-infected cells, HBV-infected PHHs displayed a significant enrichment of DLEU2 in anti-EZH2 nuclear RIP (figure 2A). MALAT1, a lncRNA known to bind EZH2,[14] served as positive control (figure 2A). The increased association of EZH2 with DLEU2 in HBV-infected PHHs and not in HBx-defective HBV (x-)-infected PHHs (figure 2B) is consistent with the HBx-dependent induction of DLEU2 in HBV-infected hepatocytes. Next, we performed anti-HBx RIPs in HBV-infected PHHs (figure 2C) and HepG2 cells transfected with wild-type (wt) monomeric HBV genomes (online supplementary figure S3a). We found that HBx binds DLEU2 in HBV-infected/HBV-replicating hepatocytes (figure 2C and online supplementary figure S3a). No enrichment of MALAT1 binding to HBx was observed (figure 2C and online supplementary figure S3a). As expected, no DLEU2 can be immunoprecipitated with anti-HBx antibodies in HBx-defective HBV (x-)-infected PHHs (figure 2D). These results indicate that HBx and DLEU2 interact in the nuclei of HBV-infected cells but do not exclude that this association might be mediated by additional partners. To ascertain whether HBx and DLEU2 directly interact, we performed RNA pull-down experiments using recombinant HBx and synthetic DLEU2 species designed in intron 1 sequence. Desthiobiotinylated DLEU2 60-mer oligonucleotides were able to pull down recombinant HBx in contrast to biotinylated reverse DLEU2 control oligonucleotides (figure 2E, upper panel). The specificity of the pulled down recombinant HBx band was confirmed by immunoblot (online supplementary figure S3b) and the same antibody immunoprecipitated a 17Kd HBx band that was recognised specifically by two anti-HBx antibodies in HBV-infected HepG2-NTCP cells (online supplementary figure S3c). Altogether, these results indicate that DLEU2 binds EZH2 in liver cells and, in HBV-infected cells, the viral protein HBx.

Modelling DLEU2-EZH2-HBx interaction in silico

Our experimental data do not allow to conclude whether HBx and EZH2 compete for DLEU2 binding or whether HBx, EZH2 and DLEU2 may form a tripartite complex. We modelled via structural bioinformatics tools the binary DLEU2-EZH2 and DLEU2-HBx as well as the ternary DLEU2-HBx-EZH2 protein-RNA complexes. The *CatRapid* sequence-based method[36] and three-dimensional (3D) atomistic structures generated using the HEX and HADDOCK tools[37 38] were used to model the physical interactions between proteins and DLEU2. Methodological details and challenges related to atomistic structure prediction are discussed in online supplementary materials.

In contrast to EZH2 for which a 3D structure is known,[39] the 3D experimental structures of HBx and DLEU2 remain elusive. Using a homology modelling approach, we confirmed that HBx comprises disordered regions,[40] essential for the interaction with protein and nucleic acid partners,[41] as well as more structured domains (online supplementary figure S4). Using RNABindRplus,[42] we found that the HBx residues predicted to bind RNA mainly comprise charged amino acids as arginines and lysines that are mostly located in the disordered region corresponding to the first half of the sequence (online supplementary figure S4a). DLEU2 analysis by CatRapid[36] corroborated the interaction of intron 1 of DLEU2 with HBx and EZH2, whereas in line with the RIP experiments (figure 2C), our analysis does not favour a direct interaction between HBx and MALAT1.

Next, we modelled the secondary (online supplementary figure S5) and 3D structure (figure 3) of the DLEU2 intron 1 region encompassing the sequences that interact with HBx in the RNA pull-down experiments (figure 2C) and exploited the HEX and HADDOCK biomolecular docking methods[37 38 43] to

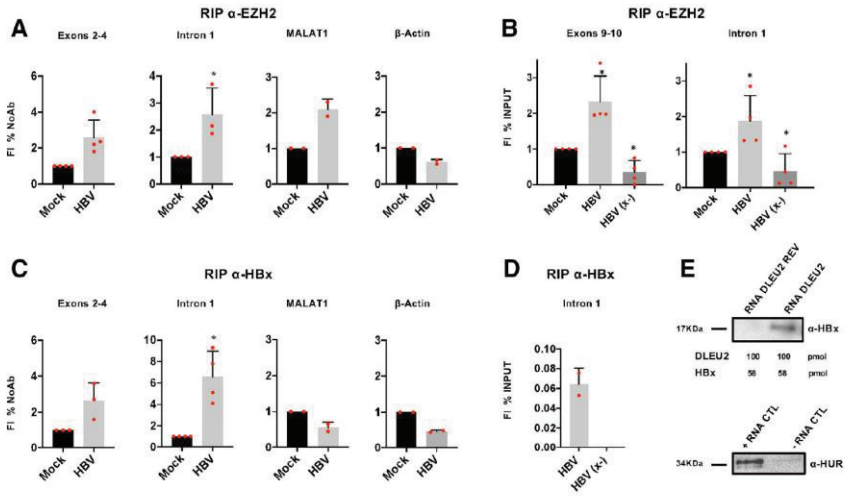


Figure 2 DLEU2 long non-coding RNA (lncRNA) interacts with enhancer of zeste homolog 2 (EZH2) and directly binds HBx. (A) RNA binding protein immunoprecipitation (RIP) using anti-EZH2 antibody. Nuclear lysates of mock-infected and HBV-infected primary human hepatocytes (PHHs) (72 hours) were subjected to immunoprecipitation with anti-EZH2 antibody or control IgG. DLEU2 exons 2–4 or intron 1, MALAT1 (positive control) and β-actin (negative control) were detected by PCR and analysed by relative densitometry. (B) RIP using anti-EZH2 antibody. Nuclear lysates of mock, HBV-infected and HBV (x-)-infected PHHs (72 hours) were subjected to immunoprecipitation with anti-EZH2 antibody or control IgG. DLEU2 exons 9–10 or intron 1 were detected by quantitative PCR (qPCR) assay. RIP results are expressed as fold induction (FI) of the % of input with respect to mock. (C) RIP using anti-HBx antibody. Nuclear lysate of mock-infected and HBV-infected PHHs (72 hours) were subjected to immunoprecipitation with anti-HBx antibody or control IgG. DLEU2 exons 2–4 or intron 1, MALAT1 and β-actin (negative controls) were detected by PCR assay and analysed by densitometry relative to the mock. RIP data in (A) and (B) represent means±SD from at least three independent experiments. (D) RIP using anti-HBx antibody. Nuclear lysate of mock, HBV-infected and HBV (x-)-infected PHHs (72 hours) were subjected to immunoprecipitation with anti-HBx antibody or control IgG. Intron 1 was detected by real-time qPCR assay. RIP results are expressed as % of input after subtraction of mock. (E) Immunoblot analysis of RNA pull-downs showing the interaction between the DLEU2 RNA (100 pmol of 3′-desthiobiotinylated RNA oligonucleotide) and recombinant HBx (1 µg), using a DLEU2 reverse sequence (DLEU2 REV) RNA as negative control (upper panel). Immunoblot analysis of positive and negative controls (provided by the manufacturer), which respectively contain RNA regions with or without binding sites for HUR, a RNA-binding proteins that regulates mRNA stability (lower panel). Data from one representative experiment out of three independent experiments are shown. In panels (A) and (B), *p<0.05; Mann-Whitney U test.

model the binary and ternary interactions of DLEU2 with EZH2 and HBx. We found that two loci on intron 1 display a high probability to be occupied by HBx (figure 3A). Interestingly, HBx preferentially binds to internal loop motifs predicted in the secondary structure of DLEU2 intron 1 (online supplementary figure S5). Using the same biomolecular docking calculations, we confirmed the well-known interaction between HBx and DDB1 (PDB entry: 3I7H[44]) and the docking results for HBx-DDBI and HBx-DLEU2 point for similar interaction strength. We identified two loci on intron 1, close but not completely overlapping with the regions potentially targeted by HBx that display a nearly equal probability to be occupied by EZH2 (figure 3B). When EZH2 is already docked to DLEU2 intron 1, HBx is still predicted to bind DLEU2 intron 1 but it is docked either to the non-occupied EZH2 site (figure 3C, panels 5 and 8) or to a docking site that is partially overlapping with the one occupied by EZH2 (figure 3C, panels 6 and 7). Notably, in both cases the presence of EZH2 slightly alters HBx docking (compare HBx docking in figure 3C, panels 5 and 7, with HBx docking in figure 3A, panel 2; HBx docking in figure 3C, panels 6 and 8, with HBx docking in figure 3A, panel 1). Similar results were obtained when HBx is docked first on DLEU2 intron 1 (compare EZH2 docking in figure 3B vs figure 3C).

Our in silico modelling is compatible with the formation of a ternary DLEU2-EZH2-HBx complex and, depending on the relative abundance of EZH2 and HBx, with a competition for partially overlapping or close by sites that may impact on EZH2/PRC2 complex functions. To verify this hypothesis, we performed DLEU2 RNA pull-down experiments with different stoichiometric ratios of HBx and EZH2.

As shown in figure 3D, both EZH2 (lane 3) and HBx (lanes 1–2) can individually bind directly DLEU2 and HBx binding is dose-dependent (lanes 1–2). At a higher EZH2:HBx ratio (2:1), EZH2 binds DLEU2 and prevents its interaction with HBx (lane 4). Conversely, at higher HBx:EZH2 ratios (6:1), HBx binding to DLEU2 precludes EZH2 binding (lane 5). In the presence of equimolar amounts of EZH2 and HBx, we observe a strong cooperative effect on HBx binding (figure 3E, lanes 1 and 3; online supplementary figure S3d), whereas EZH2 binding does not change (figure 3E, lanes 2–3). Thus, EZH2 seems to favour HBx binding to DLEU2, whereas the binding of EZH2 does not appear to be influenced by HBx. These findings can be explained if EZH2 facilitates HBx binding to DLEU2 sites that are not bound when HBx is alone. Notably, when the amount of HBx increases, EZH2 binding to DLEU2 decreases (compare figure 3E, lane 4 (HBx:EZH2 ratio 2:1) with lane 3 (equimolar amounts of HBx and EZH2) or lane 2 (EZH2 alone)). The binding of HBx in figure 3E, lane 4 (58 pmol) is higher than in lane 1 (29 pmol) but lower than in lane 3 (HBx 29 pmol and EZH2 29 pmol). The lower binding of HBx to DLEU2 observed in lane 4 as compared with lane 3 is likely the result of the displacement of EZH2 and the consequent loss of its ability to facilitate HBx binding on alternative binding sites. Altogether, the in silico modelling and the results of the RNA pull-down

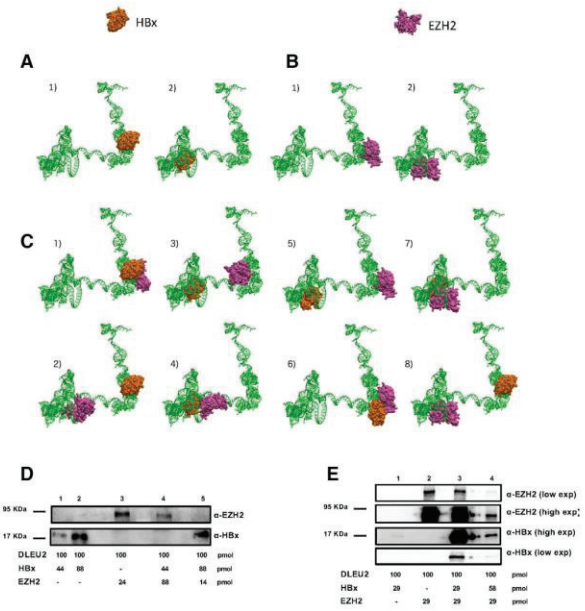


Figure 3 HBx and enhancer of zeste homolog 2 (EZH2) directly bind DLEU2 and compete for partially overlapping sites. (A) Configurations obtained from docking of HBx (surface representation, orange) on three-dimensional (3D) structure of DLEU2 intron 1 (green). Panels 1–2 present the two most probable loci for HBx adsorption on DLEU2 intron 1. (B) Configurations obtained from docking of EZH2 (surface representation, mauve) on 3D structure of DLEU2 intron 1 (green). Panels 1–2 present the two most probable loci for EZH2 adsorption on DLEU2 intron 1. (C) Panels 1–2: two most probable configurations obtained from docking of EZH2 on the most probable HBx-DLEU2 intron 1 complex (panel (A)—1). Panels 3–4: two most probable configurations obtained from docking of EZH2 on the second most probable HBx-DLEU2 intron 1 complex (panel (A)—2). Panels 5–6: two most probable configurations obtained from docking of HBx on the most probable EZH2-DLEU2 intron 1 complex (panel (B)—1). Panels 7–8: two most probable configurations obtained from docking of HBx on the second most probable EZH2-DLEU2 intron 1 complex (panel (B)—2). (D) Immunoblot analysis of RNA pull-downs showing the interaction between DLEU2 RNA, HBx and EZH2. HBx (0.75 µg=44 pmol or 1.5 µg=88 pmol) or EZH2 (1.2 µg=14 pmol, 2 µg=24 pmol or 7.5 µg=88 pmol) recombinant proteins alone or in combination (EZH2:HBx ratios of 2:1 or 1:6) were added to a fixed amount of DLEU2 RNA (100 pmol). (E) Immunoblot analysis of RNA pull-downs showing the interaction between DLEU2 RNA (100 pmol), HBx and EZH2. Increasing concentrations of HBx (0.5 µg=29 pmol, 1 µg = 58 pmol) were added at a fixed concentration of EZH2 (2.5 µg=29 pmol) recombinant proteins corresponding to EZH2:HBx ratios of 1:1 and 1:2, respectively. Data from one representative experiment out of three independent experiments are shown.

experiments reveal a highly dynamic interaction between EZH2 and HBx for DLEU2 binding that leads, depending on their relative abundance, to a functional competition between HBx and EZH2 for partially overlapping sites on DLEU2.

HBx-DLEU2 interaction regulates cccDNA transcription and HBV replication

HBx is required for the transcripion of HBV mRNAs from the cccDNA, including the HBV pgRNA that serves as template for reverse transcription and HBV replication.[9 10 45] Since DLEU2 interacts with HBx and EZH2, we next sought to assess whether DLEU2, and its interaction with HBx, contributes to cccDNA transcription. We first performed cccDNA chromatin isolation by RNA purification (ChIRP) experiments in HBV-replicating cells, using three independent antisense probe sets that specifically recover DLEU2 intronic sequences, to ascertain the binding of DLEU2 to the cccDNA (figure 4A). Notably, in HBV HBx mt-transfected HepG2 cells DLEU2 expression is reduced (figure 4B) and DLEU2 binding to the cccDNA is abrogated (figure 4C). DLEU2 depletion by specific locked nucleic acid (LNA) longRNA Gapmers (online supplementary figure S6a) resulted in a reduction of both HBV pgRNA levels (figure 4D) and total HBV DNA (data not shown) in HBV-infected PHHs, whereas cccDNA levels are not affected (figure 4E). According to these results, DLEU2 interaction with HBx and its recruitment on the viral minichromosome affect cccDNA transcription and viral replication but do not interfere with the establishment of the cccDNA pool. Similar results were obtained in both HBV-transfected HepG2 cells and HBV-infected HepG2-NTCP cells (online supplementary figure S6b). Since EZH2 is recruited onto the cccDNA in HBV-infected HepG2-NTCP cells,[46] we sought to investigate whether the ability of DLEU2 to interact with EZH2 (figure 2A) could be relevant in the regulation of cccDNA transcription. Although in the absence of HBx, the K9 methyltransferase SETDB1[11] and the arginine protein methyltransferase PRMT5[47] seem to play a predominant role in cccDNA transcriptional repression, EZH2-mediated K27 methylation of cccDNA-bound H3 also increases[11] and EZH2 silencing results in increased of HBsAg production.[47] Notably, in HBV HBx mt-transfected HepG2 cells the recruitment of EZH2 onto the cccDNA is significantly increased (figure 4F). These results and the reduction of cccDNA-bound H4 histone acetylation in the presence of DLEU2-targeting Gapmers (figure 4G) implicate DLEU2 in the epigenetic control of cccDNA transcription and in HBV replication.

HBx and DLEU2 coregulation of host genes

In view of the ability of HBx to complex with DLEU2 and EZH2, which are both overexpressed in tumour liver tissues of patients with HBV-related HCC (figure 5A), we sought to determine whether DLEU2 and EZH2 might regulate a shared set of genes and to evaluate the involvement of HBx in this process. To this aim, we extracted the genes whose expression is positively correlated with DLEU2 or EZH2 in 76 HBV-related HCCs from the transcriptomic TCGA_LIHC dataset and intersected them with the genes identified as HBx direct targets by Chip-Seq.[12] As shown in figures 5B, 261 genes were coregulated with DLEU2 in HBV-related HCCs (online supplementary table S2), with a significant enrichment in pathways associated with *Gene expression* and *cell cycle*, *DNA damage response* and *mRNA splicing* (online supplementary table S3). The intersection of DLEU2 coregulated genes and HBx ChIP data identified as HBx direct targets 33 genes (figure 5B) that enrich the same pathways with a significant propensity for splicing (online supplementary table S4). In the same patients, among the 952 genes that were coregulated with EZH2, 102 genes are direct HBx targets and 63 genes were coregulated with DLEU2. Notably, the 63 genes coregulated by both EZH2 and DLEU2 and the 102 HBx target genes that are coregulated with EZH2 enrich several pathways associated with *DNA replication*, *cell cycle* and *mitosis* (online supplementary tables S5-S6). Six genes, namely TRIM13, CCNB2, DNMT1, PRC1, POLE2 and ZBTB34 (online supplementary table S7), are

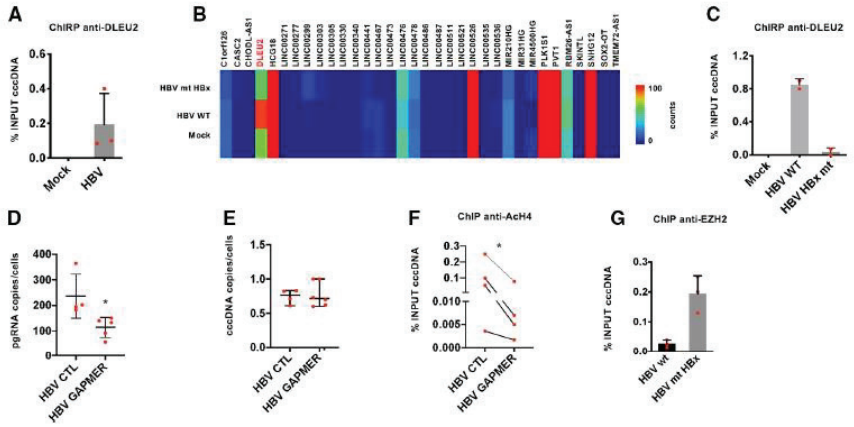


Figure 4 DLEU2 controls HBV replication. (A) DLEU2-qChIRP (chromatin isolation by RNA purification) for covalently closed circular DNA (cccDNA) using a DLEU2-specific antisense biotinylated DNA probe (100 pmol) in mock-replicating and HBV-replicating HepG2 cells (48 hours). Affinity-purified material was analysed by real-time quantitative PCR (qPCR) with DLEU2 RNA-specific primer pairs and the results are expressed as % of input. (B) Heatmap of the expression levels of 34 long non-coding RNAs (lncRNAs) that are direct transcriptional targets of HBx analysed by Nanostring in HepG2 cells replicating wild-type HBV (HBV wt) or HBV with mutant HBx (HBV mt HBx) (48 hours). Colours in the heatmap indicate log2 counts normalised to GAPDH. (C) DLEU2-qChIRP for cccDNA using a DLEU2-specific antisense biotinylated DNA probe (100 pmol) in HBV or HBV HBx mt replicating HepG2 cells (48 hours). Affinity-purified material was analysed as in 4A. Results are expressed as % of input. (D) ddPCR quantification of HBV pregenomic RNA (pgRNA) (copies/cell) from HBV-infected primary human hepatocytes (PHHs) (4 days) transfected with scrambled Gapmer (CTL) or DLEU2-specific Gapmer pools. Results are expressed as copies for 10 ng of total RNA after normalisation to endogenous human β-globin. (E) ddPCR quantification of HBV cccDNA (copies/cell) from HBV-infected PHHs (4 days) transfected with scrambled Gapmer (CTL) or DLEU2-specific Gapmer pools. Results are expressed as copies for 10 ng of total RNA after normalisation to endogenous human β-globin. (F) cccDNA-bound histone H4 acetylation. Cross-linked chromatin from HBV-replicating HepG2 cells transfected with scrambled Gapmer (CTL) or DLEU2-specific Gapmer pools (48 hours) was subjected to immunoprecipitation with anti-AcH4 antibody or IgG controls, and then analysed by real-time qPCR using cccDNA-specific primer pairs. Chromatin immunoprecipitation (ChIP) results are expressed as % of input. (G) Enhancer of zeste homolog 2 (EZH2) occupancy on cccDNA in HBV wt-replicating or HBV mt HBx-replicating HepG2 cells (48 hours). ChIP results are expressed as % of input. Data in panels (A), (B), (C), (D) and (F) represent means±SD from at least three independent experiments. In (D), and (F), *p<0.05; Mann-Whitney U test.

coregulated by HBx, DLEU2 and EZH2 (figure 5B). Notably, they are all upregulated in tumour liver tissues of patients with HBV-related HCC (figure 5C) and, with the exception of TRIM13, already reported to act as tumour suppressors in breast cancer cell lines and to be downregulated in aggressive breast cancer,[48 49] associated with a reduced overall survival (online supplementary figure S7).

LncRNAs can regulate the expression of both neighbouring genes in cis and distantly located genes in trans.[14] To gain mechanistic insights on how DLEU2 and HBx cooperate in the transcriptional regulation of host genes, and on the role of EZH2 in this process, we focused on TRIM13 and CCNB2, as examples of genes coregulated by EZH2 and DLEU2 and targeted by DLEU2 in cis and in trans, respectively. TRIM13 and CCNB2 expression is increased in HBV-infected PHHs (figure 6A) and other cell-based HBV infection/replication models (online supplementary figure S8a). The recruitment of HBx on the TRIM13 and CCNB2 promoters (figure 6B) in HBV-infected PHHs is accompanied by an increase in promoter-bound Pol II (figure 6B) and H4 acetylation (online supplementary figure S8b). Next, we investigated DLEU2 involvement in the regulation of TRIM13 and CCNB2 genes. ChIRP experiments showed that DLEU2 is recruited onto the TRIM13 and CCNB2 promoter regions in HBV-negative cells (figure 6C) and, consistent with the increased DLEU2 expression in HBV-infected cells (figure 1E–F and online supplementary figure S2b), its binding is further enriched in HBV-replicating cells (figure 6C). Selective degradation of DLEU2 by Gapmers (online supplementary figure S6a) resulted in decreased TRIM13 and CCNB2 expression in mock-infected and HBV-infected PHHs (figure 6D) and HepG2-NTCP cells (online supplementary figure S8c) as well as in wt HBV-replicating HepG2 cells (online supplementary figure S8d). Notably, TRIM13 transcription is not induced in HBx mt HBV-replicating HepG2 cells and is not affected by DLEU2-specific Gapmers (online supplementary figure S8d, left panel). In HBV-replicating HepG2 cells co-transfected with HBV monomers and DLEU2-specific Gapmers, where the degradation of DLEU2 is more efficient (online supplementary figure S6a), we also showed a sharp decrease of H4 acetylation at the TRIM13 promoter (online supplementary figure S8e). Collectively, these results indicate that DLEU2 and HBx cooperate to increase TRIM13 and CCNB2 expression. Next, we investigated whether the increased recruitment of DLEU2 on the TRIM2 and CCN2 promoters (figure 6C) and their transcriptional activation is accompanied by changes in EZH2/PRC2 occupancy. As shown in figure 6E, both EZH2 and SUZ12 bind TRIM13 and CCNB2 promoters in uninfected cells and their chromatin occupancy on both promoters remains stable or even increases in HBV-replicating HepG2-NTCP cells (figure 6E) and is accompanied by an increase of Pol II occupancy and H4 acetylation (figure 6F). The protein levels of PRC2 complex components in uninfected and HBV-infected PHHs and HepG2-NTCP cells are shown in online supplementary figure S9. Finally, in HBV-replicating HepG2 cells both EZH2 chromatin occupancy and the repressive mark H3me3K27 decrease at the TRIM13 promoter, whereas in the absence of HBx both EZH2 binding and H3me3K27 increase (online supplementary figure S8f).

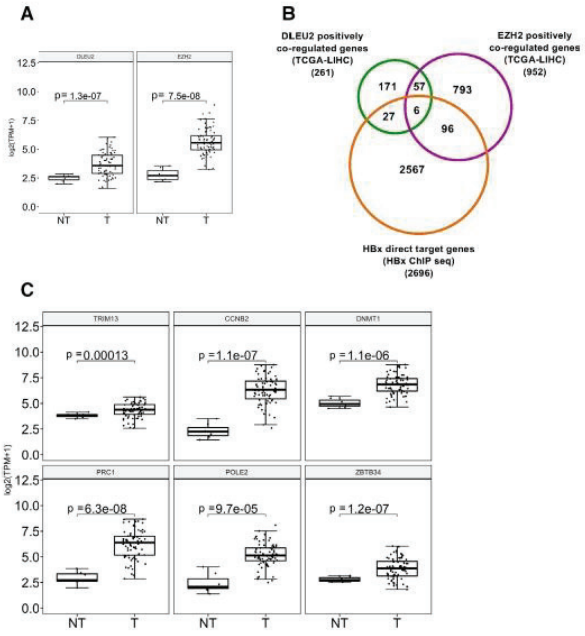


Figure 5 DLEU2 and enhancer of zeste homolog 2 (EZH2) coregulated HBx direct target gene expression in human HBV-related hepatocellular carcinoma (HCC). (A) DLEU2 and EZH2 expression data from TCGA in non-tumour (NT) and in tumour human liver tissues (T) from 76 HBV-related patients with HCC. (B) Venn diagram showing the intersection of genes whose expression is positively correlated with DLEU2 or EZH2 in 76 HBV-related HCCs from the transcriptomic TCGA-LIHC dataset accessed via the R2 platform (https://r2.amc.nl) with the genes identified as HBx direct targets by chromatin immunoprecipitation sequencing (ChIP-Seq). (C) TRIM13, CCNB2, DNMT1, PRC1, POLE2 and ZBTB34 mRNA levels in NT and T human liver tissues from HBV-related patients with HCC (transcriptomic data from TCGA, HBV-related HCCs=76).

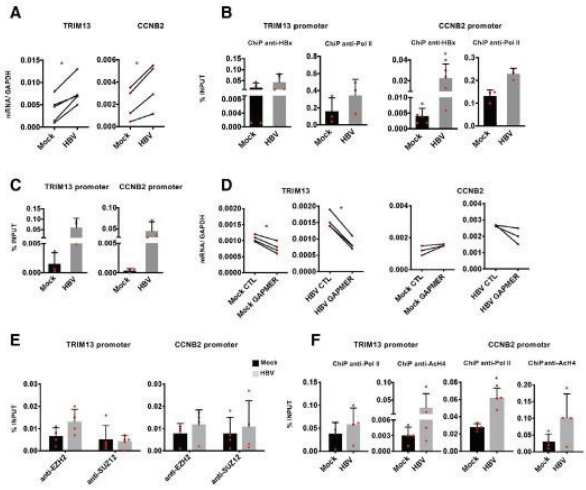


Figure 6 DLEU2 and HBx cooperate to activate TRIM13 and CCNB2 transcription. (A) Real-time quantitative PCR (qPCR) of TRIM13 and CCNB2 mRNA in mock-infected and HBV-infected primary human hepatocytes (PHHs) (7 days). Results are expressed as relative values to endogenous human GAPDH mRNAs. (B) HBx occupancy and modulation of polymerase II (Pol II) binding on TRIM13 and CCNB2 promoters in mock-infected and HBV-infected PHHs (72 hours) by chromatin immunoprecipitation (ChIP) assay. The detection of HBV covalently closed circular DNA (cccDNA) using specific primers in the ChIPed DNA from HBV-infected cells (cccDNA-ChIP) served as a technical positive control for the ChIP procedure. ChIP results are expressed as % of input. (C) DLEU2-ChIRP for TRIM13 promoter or CCNB2 promoter using DLEU2-specific antisense DNA probe in mock-replicating and HBV-replicating HepG2 cells (48 hours). (D) Real-time qPCR of TRIM13 and CCNB2 cDNA in mock-infected or HBV-infected PHHs transfected (4 days) with scrambled Gapmer (CTL) or with DLEU2 Gapmer (GAPMER). Results are expressed as relative values to endogenous human GAPDH mRNAs. (E) Enhancer of zeste homolog 2 (EZH2) and SUZ12 occupancy on TRIM13 and CCNB2 promoter regions by ChIP assay in mock-infected and HBV-infected HepG2-NTCP cells (7 days). Anti-CTCF ChIPs were included as a technical positive control. Results are expressed as % of input. (F) Modulation of H4 acetylation and Pol II binding on the TRIM13 and CCNB2 promoters in HBV-infected HepG2-NTCP cells (7 days). Results are expressed as % of input. Data in panels (A) to (F) represent means±SD from at least three independent experiments. In (A), (B), (D) and (F), *p<0.05, Mann-Whitney U test.

Altogether, these results suggest that HBx and DLEU2 cooperate to activate transcription from a subset of cellular genes, including genes that are negatively regulated by EZH2/PRC2 in non-infected cells and activated in HBV-infected cells as well as HBV-related HCCs.

CONCLUSIONS

HBx, whose primary role in the HBV life cycle is to bind the nuclear cccDNA minichromosome in order to initiate and maintain viral transcription and drive HBV replication,[10 11 45] also binds a large number of target sequences on the cellular genome.[12] HBx interacts with multiple transcription factors, coregulators and chromatin modifying complexes[12 50] in order to modulate the transcription of host genes. Here, we show that HBx binds the promoter region of DLEU2, enhances DLEU2 transcription and induces the accumulation of DLEU2 RNA species in infected hepatocytes. We showed by RIP experiments that DLEU2 binds both EZH2, the catalytic subunit of the PRC2 chromatin complex, in liver cells and HBx in HBV-infected cells. These results highlight a novel capacity of HBx to bind ncRNAs and prompted us to investigate the role of these interactions in the transcriptional regulation of cccDNA and HBx host target genes. cccDNA ChIRP experiments showed that DLEU2 binds to the cccDNA and functional experiments with DLEU2-specific Gapmers supported the role of DLEU2 recruitment in the regulation of cccDNA transcription. HBx interaction with DLEU2 on the cccDNA may represent a novel therapeutic target to silence cccDNA transcription and foster a functional cure of HBV infection.[51]

To gain mechanistic insights on how DLEU2 and HBx cooperate in the transcriptional regulation of host genes, and the role of the EZH2/PRC2 complex in this process, we focused on TRIM13 and CCNB2 as examples of genes targeted by DLEU2 in cis and in trans, respectively. Notably, TRIM13 and CCNB2 are part of a subset of six genes coregulated with DLEU2 or EZH2 in HBV-related patients with HCC and identified as direct HBx targets in ChIP-Seq experiments.[12] The contribution of PRC2-mediated deposition of the H3K27me3 repressive chromatin mark to the epigenetic silencing during embryonic development and differentiation[19] and the impact of its dysregulation in cancer[22] are well established, although

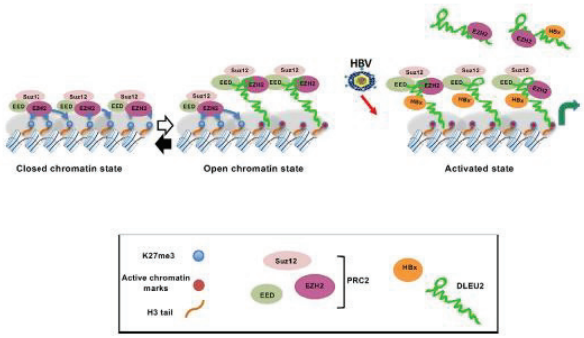


Figure 7 A model for DLEU2-PRC2/EZH2-HBx interaction. Left panel: PRC2-mediated epigenetic silencing of target genes; middle panel: DLEU2 recruitment displaces PRC2/EZH2 from close contact with target chromatin, resulting in a reduction of the methyl groups at lysine 27 of histone H3 (H3K27me3) repressive mark; right panel: in the presence of HBx (HBV-infected hepatocytes and HBV-related HCCs expressing HBx), the increase in DLEU2 levels and DLEU2 interaction with PRC2/EZH2 lead to a progressive increase of active chromatin marks associated with active transcription. HBx is part of the DLEU2-PRC2/EZH2 complex and recruited at the chromatin level together with transcription factors and co-activators to contribute to transcriptional activation (not shown in the cartoon). EZH2, enhancer of zeste homolog 2; PRC2, polycomb repressor complex 2.

EZH2-PRC2-independent and PRC2-dependent transcriptional activation have been described.[26 27 35 52 53] PRC2 complex binding to coding and non-coding RNAs is also well documented,[17 18] but there is no unified view on how PRC2 interacts with chromatin-bound RNAs and the functional consequences of this interaction. Current models, which are not necessarily mutually exclusive, of how PRC2 binding to chromatin-bound RNAs impact on transcription include: a) PRC2 recruitment to its target sites by lncRNAs[17 18]; b) RNAs serving as decoys to prevent PRC2 binding to active genes or to evict PRC2[14 54–56]; c) RNA binding to an allosteric regulatory site on EZH2 to inhibit the enzymatic activity of PRC2.[56] Notably, PRC2 binding to RNA and association with chromatin appear to be antagonistic.[54 55 57] By combining ChIRP and ChIP experiments, we showed that DLEU2 and HBx cooperate to increase TRIM13 and CNNB2 expression without changing EZH2 and SUZ12 occupancy on their promoters. Thus, HBx by increasing DLEU2 levels and by binding to DLEU2 interferes with EZH2/PRC2 repressive functions and activates transcription. Altogether, our results are consistent with a model in which HBx interaction with DLEU2 either evicts EZH2, a modality prevalent on cccDNA, or displaces the PRC2/EZH2 complex from close contact with chromatin in the case of host genes[54–57] (figure 7).

The activation of DLEU2 by HBx and the impact of its direct interaction with DLEU2 on host target genes may have important consequences on viral pathogenesis and HCC development. It is worth underlining that the genes positively coregulated with DLEU2 or EZH2 in HBV-related HCCs are enriched for pathways associated to *cell cycle*, *DNA damage response* and *mRNA splicing*. The intersection of the transcriptomic data with HBx ChIP-data[12] identified six genes coregulated by HBx, EZH2 and DLEU2, all overexpressed in HBV-related HCCs and, with one exception, associated with a reduced overall survival. Collectively, our data and these observations suggest that DLEU2 and HBx may coregulate a subset of target genes in HBV-infected cells and HCC. The ability of HBx to bind DLEU2 impacts on the epigenetic control of host genes and provide a new key to understand the role of HBx, and the overexpression of DLEU2 and EZH2 in HBV-related HCCs.

EXPERIMENTAL PROCEDURES

Materials and methods are described in the online supplementary information. Primer pairs and probes used throughout the manuscript are listed in online supplementary table 8.

Author affiliations
[1]Center for Life NanoScience@Sapienza, Istituto Italiano di Tecnologia, Rome, Italy
[2]Department of Engineering, Campus Bio-Medico University, Rome, Italy
[3]Cancer Research Center of Lyon (CRCL), UMR Inserm U1052 / CNRS 5286, Lyon, France
[4]Department of Physics and Chemistry - Emilio Segre', University of Palermo, Palermo, Italy
[5]SAFU Unit, IRCCS Regina Elena National Cancer Institute, Rome, Italy
[6]Department of Internal Medicine (DMISM), Sapienza University, Rome, Italy

Acknowledgements The authors would like to thank Professor J Faivre, Dr M Ballarino and Professor F Zoulim for fruitful discussions. The authors would like to thank M. Rivoire for providing liver samples. For experimental assistance, the authors would like to thank Dr C. Caron de Fromentel, Dr L Calvo (Nanostring), Dr A D'Alfonso (RNA extraction).

Contributors DS and LC share co-first authorship. FG and ML conceived and designed the study. FG supervised the study. DS, VA, OF, AP, M-LP, MZ, SJ, LB and FG performed experiments and analysed data. LC and GC conducted computational analysis. M-LP performed bioinformatic analysis. LC, ML and FG wrote the manuscript.

Funding This work was supported by grants from Agence Nationale pour la Recherche sur le SIDA et les hepatites virales (ANRS) to ML (n° ECTZ8323; n. ECTZ27696; n. ECTZ66014); from the IDEX Package PALSE (Programme Avenir Lyon-St Etienne); from the Agence Nationale de la Recherche (ANR@TRACTION) to ML; from the EU project 667273 HEP-CAR to ML.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement The transcriptomic data used to generate the Venn diagram in Figure 5B ('genes whose expression is positively correlated with DLEU2 or EZH2') can be accessed on the TGCA-LIHC dataset and extracted using the R2 platform (https://r2.amc.nl) with the function 'Find correlated Genes with a Single Gene'. The ChIP-Seq dataset used in Figure 5B ('genes identified as HBx direct genes'; listed in online supplementary table 8) can be accessed at http://www.ebi.ac.uk/ena/data/view/PRJEB13227. All other raw data, materials and reagents are available on request from the corresponding authors (FG and ML).

ORCID iDs
Letizia Chiodo http://orcid.org/0000-0002-8278-7075
Vincenzo Alfano http://orcid.org/0000-0001-8594-9837
Massimo Levrero http://orcid.org/0000-0002-4978-0875
Francesca Guerrieri http://orcid.org/0000-0002-2021-4394

REFERENCES

1 World Health Organization. Global hepatitis report 2017, 2017. Available: http://www.who.int/iris/handle/10665/255016
2 Schweitzer A, Horn J, Mikolajczyk RT, *et al*. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. *Lancet* 2015;386:1546–55.
3 Stanaway JD, Flaxman AD, Naghavi M, *et al*. The global burden of viral hepatitis from 1990 to 2013: findings from the global burden of disease study 2013. *Lancet* 2016;388:1081–8.
4 Lok AS, Zoulim F, Dusheiko G, *et al*. Hepatitis B cure: from discovery to regulatory approval. *Hepatology* 2017;66:1296–313.
5 European Association for the Study of the Liver. EASL 2017 clinical practice guidelines on the management of hepatitis B virus infection. *J Hepatol* 2017;67:370–98.

6 Mason WS, Gill US, Litwin S, *et al*. Hbv DNA integration and clonal hepatocyte expansion in chronic hepatitis B patients considered immune tolerant. *Gastroenterology* 2016;151:986–98.
7 Testoni B, Levrero M, Zoulim F. Challenges to a cure for HBV infection. *Semin Liver Dis* 2017;37:231–42.
8 Pollicino T, Belloni L, Raffa G, *et al*. Hepatitis B virus replication is regulated by the acetylation status of hepatitis B virus cccDNA-bound H3 and H4 histones. *Gastroenterology* 2006;130:823–37.
9 Decorsière A, Mueller H, van Breugel PC, *et al*. Hepatitis B virus X protein identifies the SMC5/6 complex as a host restriction factor. *Nature* 2016;531:386–9.
10 Belloni L, Pollicino T, De Nicola F, *et al*. Nuclear HBx binds the HBV minichromosome and modifies the epigenetic regulation of cccDNA function. *Proc Natl Acad Sci U S A* 2009;106:19975–9.
11 Rivière L, Gerossier L, Ducroux A, *et al*. Hbx relieves chromatin-mediated transcriptional repression of hepatitis B viral cccDNA involving SETDB1 histone methyltransferase. *J Hepatol* 2015;63:1093–102.
12 Guerrieri F, Belloni L, D'Andrea D, *et al*. Genome-Wide identification of direct HBx genomic targets. *BMC Genomics* 2017;18:184.
13 Moyo B, Nicholson SA, Arbuthnot PB. The role of long non-coding RNAs in hepatitis B virus-related hepatocellular carcinoma. *Virus Res* 2016;212:103–13.
14 Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell* 2011;43:904–14.
15 Liu W, Ding C. Roles of lncRNAs in viral infections. *Front Cell Infect Microbiol* 2017;7:205.
16 Wang P, Xu J, Wang Y, *et al*. An interferon-independent lncRNA promotes viral replication by modulating cellular metabolism. *Science* 2017;358:1051–5.
17 Zhao J, Ohsumi TK, Kung JT, *et al*. Genome-Wide identification of Polycomb-Associated RNAs by RIP-seq. *Mol Cell* 2010;40:939–53.
18 Guil S, Soler M, Portela A, *et al*. Intronic RNAs mediate EZH2 regulation of epigenetic targets. *Nat Struct Mol Biol* 2012;19:664–70.
19 Margueron R, Reinberg D. The polycomb complex PRC2 and its mark in life. *Nature* 2011;469:343–9.
20 Li G, Margueron R, Ku M, *et al*. Jarid2 and PRC2, partners in regulating gene expression. *Genes Dev* 2010;24:368–80.
21 Sarma K, Cifuentes-Rojas C, Ergun A, *et al*. Atrx directs binding of PRC2 to Xist RNA and polycomb targets. *Cell* 2014;159:869–83.
22 Wassef M, Margueron R. The multiple facets of PRC2 alterations in cancers. *J Mol Biol* 2017;429:1978–93.
23 Cai M-Y, Tong Z-T, Zheng F, *et al*. Ezh2 protein: a promising immunomarker for the detection of hepatocellular carcinomas in liver needle biopsies. *Gut* 2011;60:967–76.
24 Kim W, Bird GH, Neff T, *et al*. Targeted disruption of the EZH2-EED complex inhibits EZH2-dependent cancer. *Nat Chem Biol* 2013;9:643–50.
25 Lavarone E, Barbieri CM, Pasini D. Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity. *Nat Commun* 2019;10:1679.
26 Gonzalez ME, Moore HM, Li X, *et al*. Ezh2 expands breast stem cells through activation of Notch1 signaling. *Proc Natl Acad Sci U S A* 2014;111:3098–103.
27 Li J, Xi Y, Li W, *et al*. Trim28 interacts with EZH2 and SWI/SNF to activate genes that promote mammosphere formation. *Oncogene* 2017;36:2991–3001.
28 Yang Y, Chen L, Gu J, *et al*. Recurrently deregulated lncRNAs in hepatocellular carcinoma. *Nat Commun* 2017;8:14421.
29 Peng L, Paulson A, Li H, *et al*. Developmental programming of long non-coding RNAs during postnatal liver maturation in mice. *PLoS One* 2014;9:e114917.
30 Giulietti M, Righetti A, Principato G, *et al*. Lncrna co-expression network analysis reveals novel biomarkers for pancreatic cancer. *Carcinogenesis* 2018;39:1016–25.
31 Villa E, Critelli R, Lei B, *et al*. Neoangiogenesis-related genes are hallmarks of fast-growing hepatocellular carcinomas and worst survival. results from a prospective study. *Gut* 2016;65:861–9.
32 Lerner M, Harada M, Lovén J, *et al*. DLEU2, frequently deleted in malignancy, functions as a critical host gene of the cell cycle inhibitory microRNAs miR-15a and miR-16-1. *Exp Cell Res* 2009;315:2941–52.
33 Gupta RA, Shah N, Wang KC, *et al*. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 2010;464:1071–6.
34 Hu J-J, Song W, Zhang S-D, *et al*. HBx-upregulated lncRNA UCA1 promotes cell growth and tumorigenesis by recruiting EZH2 and repressing p27Kip1/CDK2 signaling. *Sci Rep* 2016;6:23521.
35 Fan H, Zhang H, Pascuzzi PE, *et al*. Hepatitis B virus X protein induces EpCAM expression via active DNA demethylation directed by RelA in complex with EZH2 and TET2. *Oncogene* 2016;35:715–26.
36 Cirillo D, Blanco M, Armaos A, *et al*. Quantitative predictions of protein interactions with long noncoding RNAs. *Nat Methods* 2017;14:5–6.
37 Ritchie DW, Kozakov D, Vajda S. Accelerating and focusing protein-protein docking correlations using multi-dimensional rotational FFT generating functions. *Bioinformatics* 2008;24:1865–73.
38 van Zundert GCP, Rodrigues JPGLM, Trellet M, *et al*. The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J Mol Biol* 2016;428:720–5.
39 Antonysamy S, Condon B, Druzina Z, *et al*. Structural context of disease-associated mutations and putative mechanism of autoinhibition revealed by X-ray crystallographic analysis of the EZH2-SET domain. *PLoS One* 2013;8:e84147.
40 Minor MM, Slagle BL. Hepatitis B virus HBx protein interactions with the ubiquitin proteasome system. *Viruses* 2014;6:4683–702.
41 Mészáros B, Tompa P, Simon I, *et al*. Molecular principles of the interactions of disordered proteins. *J Mol Biol* 2007;372:549–61.
42 Walia RR, Xue LC, Wilkins K, *et al*. RNABindRPlus: a predictor that combines machine learning and sequence homology-based methods to improve the reliability of predicted RNA-binding residues in proteins. *PLoS One* 2014;9:e97725.
43 Wass MN, Fuentes G, Pons C, *et al*. Towards the prediction of protein interaction partners using physical docking. *Mol Syst Biol* 2011;7:469.
44 Li T, Robert EI, van Breugel PC, *et al*. A promiscuous alpha-helical motif anchors viral hijackers and substrate receptors to the Cul4-Ddb1 ubiquitin ligase machinery. *Nat Struct Mol Biol* 2010;17:105–11.
45 Lucifora J, Arzberger S, Durantel D, *et al*. Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. *J Hepatol* 2011;55:996–1003.
46 Zhang H, Xing Z, Mani SKK, *et al*. Rna helicase DEAD box protein 5 regulates polycomb repressive complex 2/Hox transcript antisense intergenic RNA function in hepatitis B virus infection and hepatocarcinogenesis. *Hepatology* 2016;64:1033–48.
47 Zhang W, Chen J, Wu M, *et al*. Prmt5 restricts hepatitis B virus replication through epigenetic repression of covalently closed circular DNA transcription and interference with pregenomic RNA encapsidation. *Hepatology* 2017;66:398–415.
48 Tomar D, Singh R. TRIM13 regulates ubiquitination and turnover of NEMO to suppress TNF induced NF-κB activation. *Cell Signal* 2014;26:2606–13.
49 Chen W-xian, Cheng L, Xu L-yun, *et al*. Bioinformatics analysis of prognostic value of TRIM13 gene in breast cancer. *Biosci Rep* 2019;39.
50 Slagle BL, Bouchard MJ. Hepatitis B virus X and regulation of viral gene expression. *Cold Spring Harb Perspect Med* 2016;6:a021402.
51 Lok AS, Zoulim F, Dusheiko G, *et al*. Hepatitis B cure: from discovery to regulatory approval. *J Hepatol* 2017;67:847–61.
52 Studach LL, Menne S, Cairo S, *et al*. Subset of Suz12/PRC2 target genes is activated during hepatitis B virus replication and liver carcinogenesis associated with HBV X protein. *Hepatology* 2012;56:1240–51.
53 Xu J, Shao Z, Li D, *et al*. Developmental control of polycomb subunit composition by GATA factors mediates a switch to non-canonical functions. *Mol Cell* 2015;57:304–16.
54 Beltran M, Yates CM, Skalska L, *et al*. The interaction of PRC2 with RNA or chromatin is mutually antagonistic. *Genome Res* 2016;26:896–907.
55 Davidovich C, Wang X, Cifuentes-Rojas C, *et al*. Toward a consensus on the binding specificity and promiscuity of PRC2 for RNA. *Mol Cell* 2015;57:552–8.
56 Zhang Q, McKenzie NJ, Warneford-Thomson R, *et al*. Rna exploits an exposed regulatory site to inhibit the enzymatic activity of PRC2. *Nat Struct Mol Biol* 2019;26:237–47.
57 Wang X, Goodrich KJ, Gooding AR, *et al*. Targeting of polycomb repressive complex 2 to RNA by short repeats of consecutive guanines. *Mol Cell* 2017;65:1056–67.

SUPPLEMENTAL INFORMATION

Supplemental Information includes:

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

Salerno D, *et al. Gut* 2020;0:1–9. doi: 10.1136/gutjnl-2019-319637

MATERIALS AND METHODS

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

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

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

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

Patient samples

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Statistical Analysis

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

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

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

MODELING AND DOCKING

(a) RNA structure modelling

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

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

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

(b) HBx structure modelling

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

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

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

(c) Sequence-based interaction methods

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

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

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

(d) Docking predictions

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

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

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

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

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

e) Protein-protein and protein-lncRNAs control docking

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

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

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

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

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

Legends to Supplementary Figures

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

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

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

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

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

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

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

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

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

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

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

Figure S1

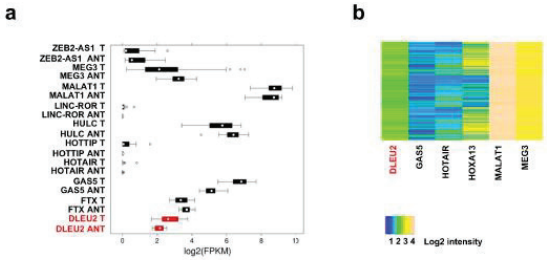

a
ZEB2-AS1 T
ZEB2-AS1 ANT
MEG3 T
MEG3 ANT
MALAT1 T
MALAT1 ANT
LINC-ROR T
LINC-ROR ANT
HULC T
HULC ANT
HOTTIP T
HOTTIP ANT
HOTAIR T
HOTAIR ANT
GAS5 T
GAS5 ANT
FTX T
FTX ANT
DLEU2 T
DLEU2 ANT
0
2
4
6
8
10
log2(FPKM)
b
DLEU2
GAS5
HOTAIR
HOXA13
MALAT1
MEG3
1 2 3 4 Log2 intensity

Figure S2

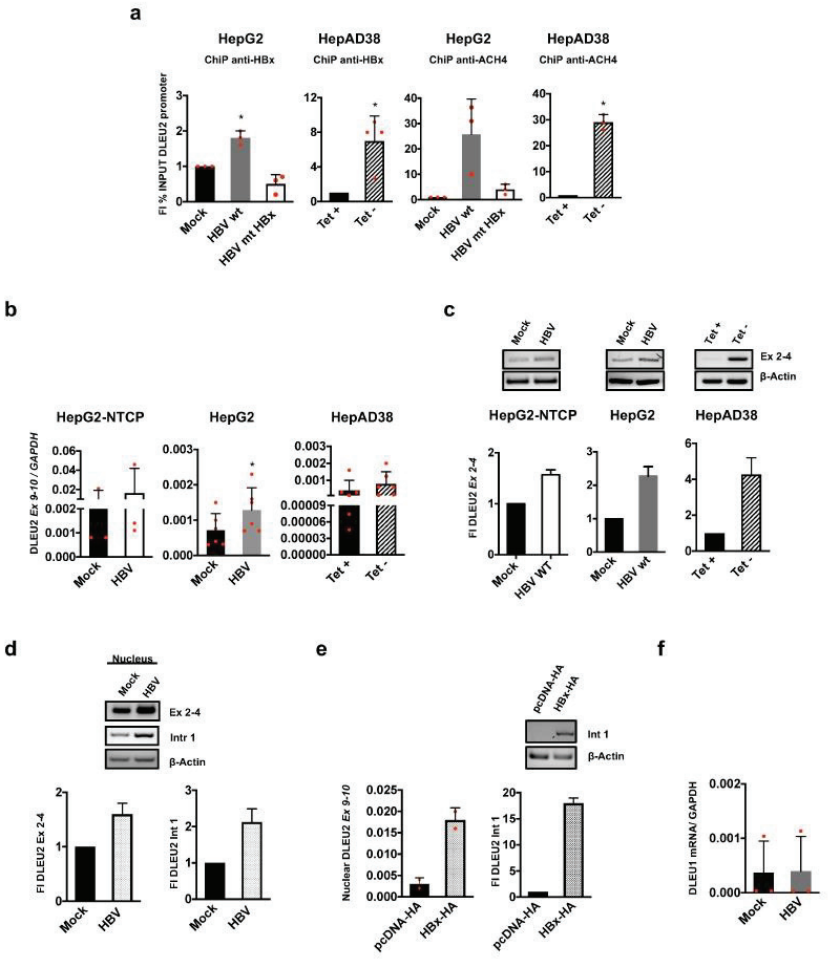

a
HepG2
ChiP anti-HBx
HepAD38
ChiP anti-HBx
HepG2
ChiP anti-ACH4
HepAD38
ChiP anti-ACH4
FI % INPUT DLEU2 promoter
Mock
HBV wt
HBV mt HBx
Tet +
Tet -
b
HepG2-NTCP
HepG2
HepAD38
DLEU2 Ex 9-10 / GAPDH
Mock
HBV
c
Ex 2-4
β-Actin
HepG2-NTCP
HepG2
HepAD38
FI DLEU2 Ex 2-4
HBV WT
HBV wt
d
Nucleus
Ex 2-4
Intr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

Figure S3

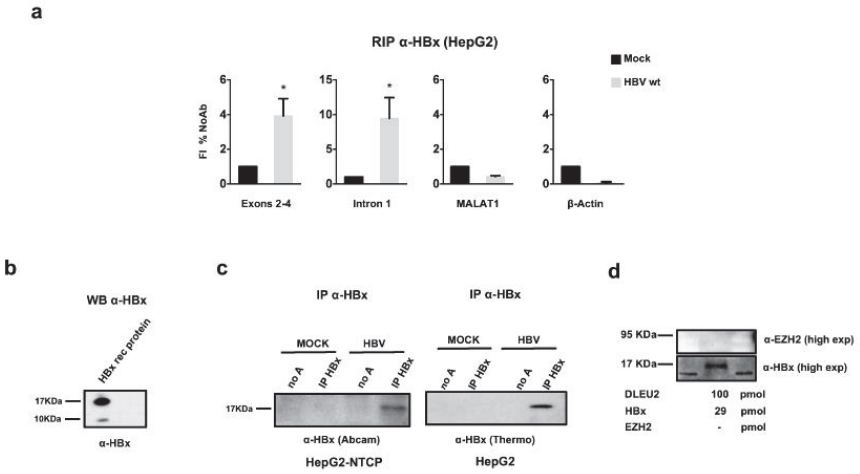

a
RIP α-HBx (HepG2)
Mock
HBV wt
Fl % NoAb
Exons 2-4
Intron 1
MALAT1
β-Actin
b
WB α-HBx
HBx rec protein
17KDa
10KDa
α-HBx
c
IP α-HBx
IP α-HBx
MOCK
HBV
MOCK
HBV
no A
IP HBx
no A
IP HBx
no A
IP HBx
no A
IP HBx
17KDa
α-HBx (Abcam)
α-HBx (Thermo)
HepG2-NTCP
HepG2
d
95 KDa
17 KDa
α-EZH2 (high exp)
α-HBx (high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

Figure S4

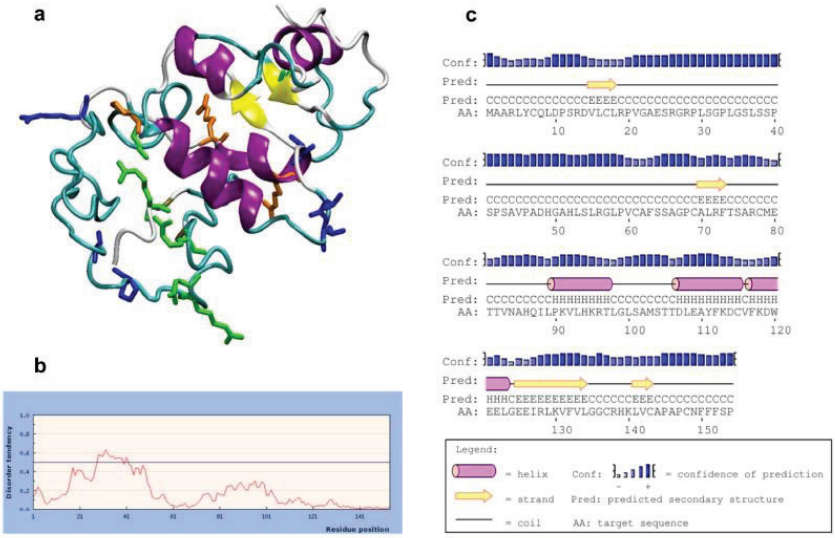
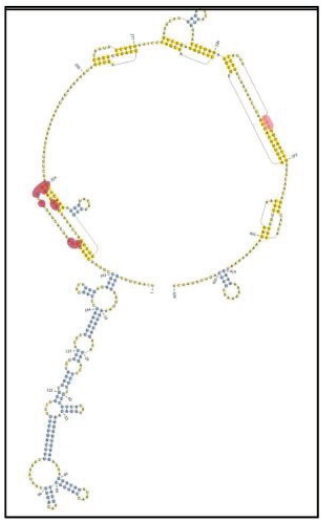

a
b
c
Conf:
Pred:
Pred: CCCCCCCCCCCCCCCEEEECCCCCCCCCCCCCCCCCCCC
AA: MAARLYCQLDPSRDVLCLRPVGAESRGRPLSGPLGSLSSP
10 20 30 40
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCEEEECCCCCCC
AA: SPSAVPADHGAHLSLRGLPVCAFSSAGPCALRFTSARCME
50 60 70 80
Pred: CCCCCCCCCHHHHHHHHCCCCCCCCCHHHHHHHHHCHHH
AA: TTVNAHQILPKVLHKRTLGLSAMSTTDLEAYFKDCVFKDW
90 100 110 120
Pred: HHHCEEEEEEEEECCCCCCEEECCCCCCCCCCC
AA: EELGEEIRLKVFVLGGCRHKLVCAPAPCNFFFSP
130 140 150
Legend:
= helix
= strand
= coil
Conf:
= confidence of prediction
Pred: predicted secondary structure
AA: target sequence
Residue position

Figure S5



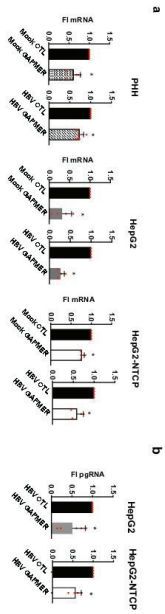
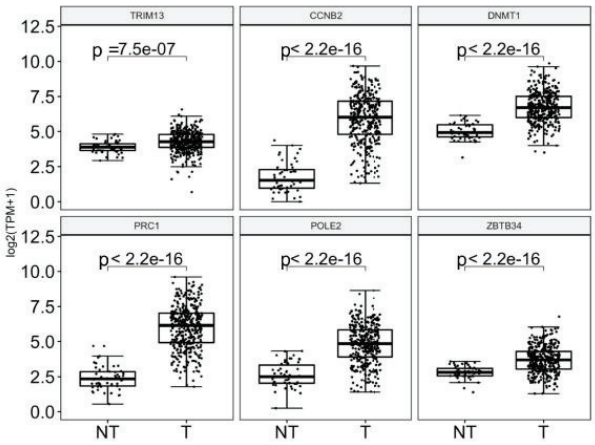

a
Fl mRNA
Mock CTL
Mock ΔAP/MER
HBV CTL
HBV ΔAP/MER
PHH
HepG2
HepG2-NTCP
b
Fl pgRNA
HepG2
HepG2-NTCP

Figure S6

Figure S7

a


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

b

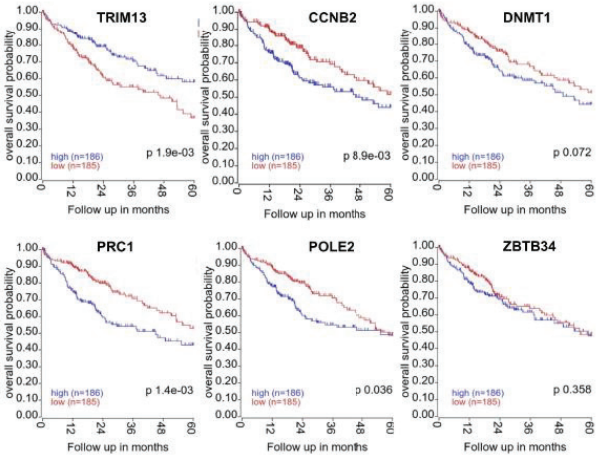

TRIM13
CCNB2
DNMT1
PRC1
POLE2
ZBTB34
overall survival probability
Follow up in months
high (n=186)
low (n=185)
p 1.9e-03
p 8.9e-03
p 0.072
p 1.4e-03
p 0.036
p 0.358

Figure S8

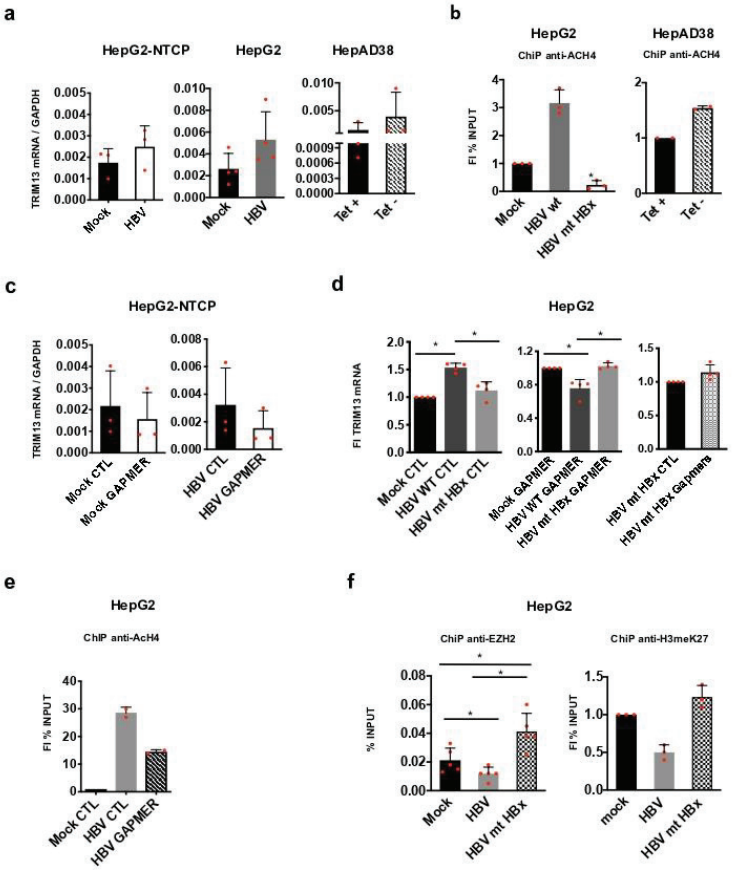

a
HepG2-NTCP
HepG2
HepAD38
TRIM13 mRNA / GAPDH
Mock
HBV
Tet +
Tet -
b
HepG2
HepAD38
ChiP anti-ACH4
FI % INPUT
Mock
HBV wt
HBV mt HBx
c
HepG2-NTCP
Mock CTL
Mock GAPMER
HBV CTL
HBV GAPMER
d
HepG2
FI TRIM13 mRNA
Mock CTL
HBV WT CTL
HBV mt HBx CTL
Mock GAPMER
HBV WT GAPMER
HBV mt HBx GAPMER
HBV mt HBx CTL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

Figure S9

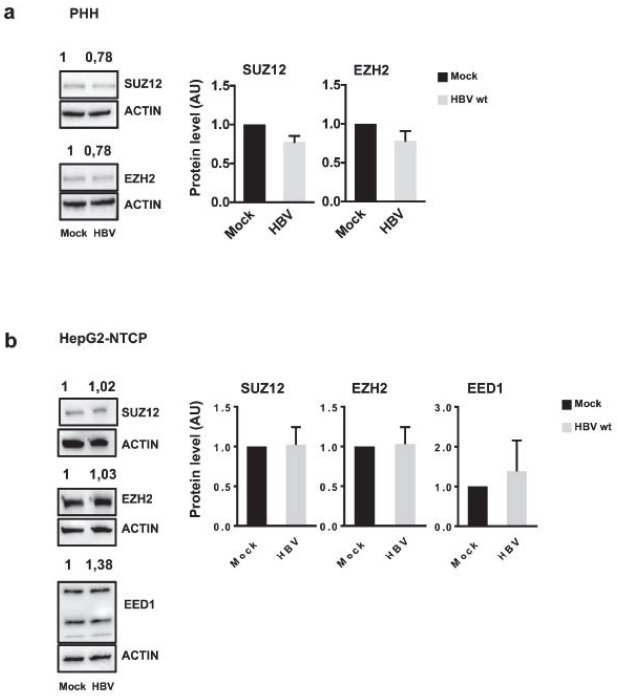

a
PHH
1 0,78
SUZ12
ACTIN
1 0,78
EZH2
ACTIN
Mock HBV
Protein level (AU)
SUZ12
EZH2
Mock
HBV wt
Mock
HBV
b
HepG2-NTCP
1 1,02
SUZ12
ACTIN
1 1,03
EZH2
ACTIN
1 1,38
EED1
ACTIN
Mock HBV
EED1

Table S1

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

Table S2

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

NAME	P-value	Adjusted p-value	Description
DNA Replication	0.0000525900	0.0028	WikiPathways 2016
Cell Cycle	0.0003112	0.006536	WikiPathways 2016
G1 to S cell cycle control	0.000274914	0.006536	WikiPathways 2016
Cell Cycle, Mitotic	0.000000000928	0.00000001096	Reactome 2016
Mitotic G2-G2/M phases_Homo sapiens_R-HSA-453274	0.00002466	0.000582	Reactome 2016
mitotic cell cycle phase transition (GO:0044772)	0.000000011	0.00000703	GO Biological Process 2018
G1/S transition of mitotic cell cycle (GO:0000082)	0.00001362	0.0010135	GO Biological Process 2018
mitotic recombination (GO:0006312)	0.0001044	0.0041421	GO Biological Process 2018

Table S7

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

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

TABLE S8

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

Supplemental Information References

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

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

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

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

PARP inhibitors and radiation potentiate liver cell death *in vitro*. Do hepatocellular carcinomas have an achilles' heel?

Clinics and Research in Hepatology and Gastroenterology 45 (2021) 101553

Available online at ScienceDirect www.sciencedirect.com

Elsevier Masson France EM|consulte www.em-consulte.com/en

ORIGINAL ARTICLE

PARP inhibitors and radiation potentiate liver cell death *in vitro*. Do hepatocellular carcinomas have an achilles' heel?

Laetitia Gerossier[a,1], Anaëlle Dubois[a,1], Alexia Paturel[a], Nadim Fares[a,2], Damien Cohen[a], Phillippe Merle[a,b], Joel Lachuer[a,c], Anne Wierinckx[a,c], Pierre Saintigny[a,d], Brigitte Bancel[e], Janick Selves[f], Anne Schnitzler[g], Bérengère Ouine[h], Aurélie Cartier[h], Leanne de Koning[h], Vincent Puard[h], Ivan Bieche[g], Hector Hernandez-Vargas[a], Janet Hall[a], Isabelle Chemin[a,*]

[a] *Univ Lyon, Université Claude Bernard Lyon 1, INSERM, CNRS, Centre Léon Bérard, Centre De Recherche En Cancérologie De Lyon, Lyon, 69008, France*
[b] *Department of Hepatology, Groupement Hospitalier Nord, Hospices Civils De Lyon, Lyon, 69000 France*
[c] *ProfileXpert, SFR-Est, CNRS UMR-S3453, INSERM US7, Lyon Cedex 08, F-69373, France*
[d] *Department of Medical Oncology, Centre Léon Bérard, Lyon, France*
[e] *Service d'Anatomopathologie, Groupement Hospitalier Est, Hospices Civils De Lyon, Lyon, 69000, France*
[f] *Anatomie Et Cytologie Pathologiques Pôle IUC Oncopole CHU Institut Universitaire Du Cancer De Toulouse - Oncopole, Toulouse, F- 31059, France*
[g] *Department of Genetics, Institut Curie, PSL Research University, Paris, F-75005, France*
[h] *Department of Translational Research, Institut Curie, PSL Research University, Paris, F-75005, France*

KEYWORDS
Liver cancer;
Hepatitis B virus (HBV) X protein;

Abstract
Background: A promising avenue for cancer treatment is exacerbating the deregulation of the DNA repair machinery that would normally protect the genome. To address the applicability of poly(ADP-ribose) polymerase (PARP) inhibitors (PARPi) combined with radiotherapy for the treatment of hepatocellular carcinoma (HCC) two approaches were used: firstly, the *in vitro*

Abbreviations: HCC, Hepatocellular carcinoma; RT, radiotherapy; HBV, hepatitis B virus; SBRT, sterotactic body RT; PARP, poly(ADP-ribose) polymerases; PARPi, PARP inhibitors; IR, ionising radiation; SMC5/6, structural maintenance of chromosomes complex 5 and 6; HR, homologous recombination; DSBs, double strand breaks; PARG, poly(ADP-ribose) glycohydrolase; gammaH2AX, serine 139 phosphorylated Histone H2AX; PT, peritumor; T, tumor; HepaRG TR, HepaRG cells containing a plasmid with the tetracycline promotor; HepaRG TRX, HepaRG cells expressing the HBx protein under a tetracycline inducible promotor; HepaRG TRX sh-scramble, cells expressing an in-efficient shRNA construct in a lentiviral GFP vector; HepaRG TRX sh-SMC6, cells expressing a 29mer shRNA construct targeting the SMC6 transcript; HU, hydroxyurea; DAPI, 4,6-diamidino-2-phenylindole; AU, arbitrary units; D_{37}, mean lethal dose; RPPA, Reverse Phase Protein Array; FPKM, Fragments Per Kilobases of exon per Million mapped reads; DDR, DNA damage response.

* Corresponding author.
E-mail address: Isabelle.chemin@inserm.fr (I. Chemin).
[1] These authors contributed equally to this work.
[2] Present address: Department of Digestive Oncology Rangueil Universitary Hospital, Hopitaux de Toulouse, Toulouse France.

https://doi.org/10.1016/j.clinre.2020.09.014

gammaH2AX;
DNA damage;
SMC5/6;
Talazoparib;
Veliparib.

sensitivity to the PARPi Veliparib and Talazoparib +/- radiation exposure was determined in liver cell lines and the impact of the HBV X protein (HBx) that deregulates cellular DNA damage repair *via* SMC5/6 degradation was investigated. Secondly, PARP expression profiles and DNA damage levels using the surrogate marker gammaH2AX were assessed in a panel of control liver *vs* HCC tissues.
Methods: Cell cytotoxicity was measured by clonogenic survival or relative cell growth and the DNA damage response using immunological-based techniques in Hep3B, PLC/PRF/5, HepG2- and HepaRG-derived models. Transcriptome changes due to HBx expression *vs* SMC6 loss were assessed by RNA sequencing in HepaRG-derived models. PARP and PARG transcripts (qPCR) and PARP1, H2AX and gammaH2AX protein levels (RPPA) were compared in control liver *vs* HBV-, HCV-, alcohol- and non-alcoholic steatohepatitis-associated HCC (tumor/peritumor) tissues.
Results: PARPi cytotoxicity was significantly enhanced when combined with X-rays (2 Gy) with Talazoparib having a greater impact than Veliparib in most *in vitro* models. HBx expression significantly lowered survival, probably driven by SMC5/6 loss based on the transcriptome analysis and higher DNA damage levels. PARP1 and PARP2 transcript levels were significantly higher in tumor than peritumor and control tissues. The HBV/HCV/alcohol-associated tumor tissues studied had reduced H2AX but higher gammaH2AX protein levels compared to peritumor and control tissues providing evidence of increased DNA damage during liver disease progression.
Conclusions: These proof-of-concept experiments support PARPi alone or combined with radiotherapy for HCC treatment, particularly for HBV-associated tumors, that warrant further investigation.

Introduction

Over 90 % of primary liver cancers are hepatocellular carcinomas (HCC) and an often too late diagnosis and few treatment options can explain their poor prognosis. Treatment with systemic chemotherapy are ineffective in HCC and many of the main oncogenic HCC drivers, such as the TERT promoter, TP53 and CTNNB1 [1], have not proven to be targetable. In contrast to many other cancer sites, ionising external beam radiotherapy (RT) has not been widely used in HCC management outside the palliative setting [2]. However, with recent technological advances in radiation delivery, hypofractionated sterotactic body RT (SBRT) used as a monotherapy or in combination with other liver-directed therapies (reviewed in [2,3]) or with radiosensitising drugs such as poly(ADP-ribose) polymerases (PARP) inhibitors (PARPi) [4] could provide promising local control and support their use for HCC.

In order to explore the potential of PARPi for HCC treatment, we have expanded an *in vitro* pilot study [5] showing that concomitant Veliparib and ionising radiation (IR) treatment lowered liver cell survival to assess a second PARPi Talazoparib and the impact of the expression of hepatitis B virus (HBV) encoded X (HBx) protein on this response. HBx degrades the cell's structural maintenance of chromosomes complex 5 and 6 (SMC5/6) to allow the completion of the HBV lifecycle [6]. The SMC5/6 complex participates in homologous recombination (HR), one of two major DNA double strand break (DSB) repair pathways and promotes sister-chromatid HR during replication. This complex also contributes to stalled replication forks' integrity and replication recombination intermediates' removal (see [7] for review), all cellular processes that can modulate the response to PARPi [8]. In parallel to this *in vitro* approach, the expression profiles of key PARP genes in a panel of control liver and HCC peritumor (PT) and tumor (T) tissues was determined. Indeed, PARP expression is a prerequisite for the use of PARPi and to investigate whether targeting PARP activity is an option for HCC therapy the transcript profiles of three PARP genes together with that of poly(ADP-ribose) glycohydrolase (PARG) that degrades the polymers and the protein levels of the major PARP protein PARP1 were also evaluated. In addition, as the cellular consequences of PARP inhibition can depend on cell proliferation rates [8] the transcript and protein expression profiles of the proliferation marker MKI67 were determined together with the protein levels of the surrogate DNA damage marker gammaH2AX (also known as serine 139 phosphorylated Histone H2AX) [9] in order to assess DNA damage levels during liver disease progression.

Materials and methods

Cell lines

The human hepatoma cell line HepG2 (ATCC HB-8065) and its derivatives HepG2 2.2.15 carrying the HBV wildtype genome and HepG2 K6 carrying a mutant form of HBV with no HBx protein [10,11], Non-differentiated HepaRG cells expressing the tetracycline inducible promotor (HepaRG TR) or the HBx protein under the tetracycline inducible promotor (HepaRG TRX) [11], HepaRG TRX sh-scramble expressing an in-efficient shRNA construct in a lentiviral GFP vector (Origene) and HepaRG TRX sh-SMC6 cells expressing a 29mer shRNA construct targeting the SMC6 transcript (Origene), Hep3B (ATCC HB-8064) and PLC/PRF/5 (ATCC CRL-8024) cells were grown in a humidified 37 °C incubator under 5% CO_2. All cell lines were regularly checked for mycoplasma contamination.

Cytotoxicity measurements

Cell survival was assessed using the colony formation assay for all cell models except the non-differentiated HepaRG TRX-based lines. Cells were seeded (1000—2000/25 cm^2 flask) and left to adhere overnight before treatment. The PARPi Talazoparib (DMSO, 10 u M) or Veliparib (PBS, 4 mM) (Selleckchem, USA) were added directly to cell culture medium to achieve the final concentrations required. To assess the impact of PARP inhibition combined with IR, PARPi were present in the culture medium for 24 h with the radiation delivered 1 h after their addition. Cells were irradiated using a 6-MeV γ-ray clinical irradiator (SL 15 Phillips) at the Anti-Cancer Centre Léon-Bérard (Lyon, France) with a dose rate of 6 Gy min^{-1} to give the required dose.

To assess hydroxurea's (HU) impact on cell survival, HU (Sigma, 0.2 M in PBS) was added directly to culture medium (final concentration 2 mM) for a period of 24 h.

After drug removal and replacement of culture medium, colonies were grown under standard culture conditions for 10 days, then fixed in methanol and stained with 0.05 % Coomassie brilliant blue in 3:1:6 ethanol/acetic acid/water. Colonies with more than 50 cells were counted and the surviving fraction (the colony count relative to mock-treated cells) determined.

Cell growth rates in non-differentiated HepaRG TRX-based lines were assessed by plating cells at a fixed density and allowing them to grow for the required time after treatment before counting viable cell numbers.

Western blot

Cells (2.10^6) were lysed in RIPA buffer (50 mM Tris pH7.4, 250 mM NaCl, 0.1 % SDS, 2 mM DTT, 0.5 % NP40) containing a protease inhibitor cocktail (Halt™) (30 min, 4 °C) or in 2% SDS in water containing Halt™ and sonicated for 10×30 s and centrifuged (15 min, 4 °C, 13,000 *g*). Protein concentrations were estimated using the BCA Protein Assay (Pierce™). Samples were diluted in Laemmli 5X, boiled and 20 μg protein extract loaded per lane. Proteins were separated on 8% (SMC6 detection) or 12 % acrylamide gels and after transfer to nitrocellulose membranes were incubated with primary antibodies overnight at 4 °C (1:1000 anti-gammaH2AX (Abgent ab18311), 1:1000 anti-H2AX (Cell Signalling 2595), 1:500 anti-Smc6 (Abgent AT3956a), 1:1000 pADPR (Abcam, clone ab14459)) except for anti-β-actin (1:10000 anti-β-actin (Sigma-Aldrich A1978) 1 h, room temperature. Membranes were then incubated (1 h, room temperature) with Horseradish-peroxidase-conjugated secondary antibodies (1:10000). Detection used ECL (BioRad), ChemiDoc and Image Lab software (Bio-Rad).

Flow cytometry

To assess the impact of HU treatment on cell cycle progression HepG2, HepG2 2.2.15, HepG2 K6 and HepaRG TRX (non-differentiated, induced or not) (2×10^6 cells/25 cm^2 flask) were seeded and left to attach overnight before treating with 2 mM HU for varying times. Cells were then processed using the BD Cycletest™ Plus DNA reagent kit (Becton Dickinson) containing propidium iodide according to the manufacturer's protocol. Flow cytometric analysis was performed using a FACSCalibur cytofluorometer (Becton-Dickinson Biosciences). At least 10^4 events were recorded and data analysis was done with CellQuest Pro software (Becton-Dickinson Biosciences).

Immunofluorescence protein detection

The presence of gammaH2AX foci in irradiated cells was assessed by immunofluorescence. Cells were plated on coverslips (after induction in the case of the non-differentiated HepaRG TRX cells) and allowed to adhere overnight before irradiation as described above. Cells were returned to the incubator to allow DNA repair to proceed and then fixed (15 min, room temperature, 4% formaldehyde), washed twice in PBS and permeabilized in ice cold lysis solution (0.5 % Triton, 10.5 % Sucrose, 0.06 % $MgCl_2$, 20 mM Hepes pH7, 50 mM NaCl, 3 min) before incubation (40 min, 37 °C) with anti-gammaH2AX (Abgent ab18311) (1:1000) in blocking buffer (PBS-BSA 3%) and incubation with an IgG secondary antibody (1:200), Alexa 488 (Fluor®) (20 min, 37 °C). After washing, coverslips were mounted with medium containing 4,6-diamidino-2-phenylindole (DAPI; Duolink®). Immunofluorescence images were acquired using an upright microscope Nikon NIE. At least 200 cells were analysed per experimental condition. The number of foci was quantified in arbitrary units (AU) by a macro using Image J software developed by J. Jacquemetton (CRCL, Lyon).

Expression shSMC6/Shcontrol for RNA seq analysis

Non-differentiated HepaRG TR, TRX, TRX + shscramble (control) and TRX + shSMC6 (construct C) cells (1×10^6) were seeded in 6-well plates and left to attach before tetracycline induction (1 μg/mL) or not, for 4 days. SMC6 depletion was validated by western blot as previously described [6] and RNA extraction performed using RNeasy™following manufacturer instructions. Three independent experiments were carried out.

RNA samples were sequenced using Illumina's NextSeq following standard procedures. Resulting base called sequences were demultiplexed with Bcl2fastq v2.17.1.14 using 0 mismatch, followed by trimming with cutadapt v1.9.1 and quality check with fastQC. Reads were mapped against hg19 genome (GRCH37, Feb. 2009) using TopHat v2.1.0 with default parameters (bowtie 2.2.9), and signals were normalised using cufflinks v2.1.1 which uses the FPKM (Fragments Per Kilobases of exon per Million mapped reads) method.

After preprocessing, FPKM values, corresponding isoform annotations, and phenotype data, were imported into R for further analysis with R/Bioconductor packages. Boxplots of normalised signals, hierarchical clustering, and multidimensional scaling were used for data inspection. Based on this, one outlier on each experimental group was removed at this step.

To define differentially expressed genes, we modelled experimental conditions as categorical variables in a linear regression using an empirical Bayesian approach [12] Pairwise comparisons with a False Discovery Rate (FDR)-adjusted P value below 0.05 and a fold-change of at least two, were

considered statistically significant. Overlaps between gene sets were performed with the VennDiagram packages.

Patients and liver samples

180 HCC patients treated by surgical resection, included between 1996 and 2014 from three French clinical centres (Lyon Croix Rousse and Toulouse Hospitals and the French National Biological Resource Centres) were retrospectively studied (Agreement DC-2008—235). RNA samples were available from matched T and PT (at least 2 cm distant from the T) tissues for 150 patients and for the other 30 only T RNA samples were analysed. Protein extracts were available from 93 tissues (46 T, 47 PT) of the following aetiologies; 29 alcohol (13 T, 16 PT), 32 HBV (17 T, 15 PT) and 32 HCV (16 T,16 PT).

RNA samples from 11 and protein extracts from 8 surgically resected normal liver tissues surrounding colorectal adenocarcinoma metastasis (Centre Léon Bérard Biological Resource Centre, ministerial agreements #AC-2013—1871 and DC-2013—1870) were used as control samples respectively. Signed informed written consents in compliance with the requirements of the local ethical committee were obtained from patients before surgery.

RNA extraction and cDNA preparation

Total RNA was extracted from frozen tissues with Extract-all (Eurobio) according to the manufacturer's instructions. One μg was treated with DNAse-I (Roche) and then retro-transcribed (RT) using the SuperScript III First-Strand Synthesis System (Thermofischer), according to the manufacturer's instructions.

Quantitative real-time PCR

Primers for *GUS* used an endogenous control were designed using Oligo 5 (National Biosciences, Plymouth, MN, USA) and their specificity checked by searches in dbEST and nr databases. Primers for *PARP1*, *PARP2*, *PARP3*, *PARG* and *MKI67* and the PCR reaction conditions are described elsewhere [13]. Quantitative values for the different transcripts were obtained from the cell cycle number (Ct value) at which the fluorescent signal became exponential. Results, were expressed as target gene expression relative to the *GUS* gene.

Protein extraction and reverse phase protein array (RPPA) analysis

Protein extracts were prepared from frozen liver sections as previously described [14] and stored at -80° until further processing from 93 tissues from HCC patients (46 T, 47 PT) of the following aetiologies; 29 alcohol (13 T, 16 PT), 32 HBV (17 T, 15 PT) and 32 HCV (16 T,16 PT). 8 surgically resected normal liver tissues served as controls.

Protein concentration was determined (Pierce BCA reducing agent compatible kit, ref 23,252) and extracts printed onto nitrocellulose covered slides (Supernova, Grace Biolabs) using a dedicated arrayer (2470 arrayer, Aushon Biosystems) in five serial dilutions (1500 to 93.75 μg/mL) and three replicates per dilution. Arrays were labeled with 4 specific antibodies (H2AX 1:2000 (CST 2595); Phospho-Histone H2AX (ser139) 1:1000 (Abcam ab2893); PARP (uncleaved) 1:1000 (Abcam ab32378); Ki67 1:250 (DAKO M7240) as described in [15]. Read-out was done using IRDye 800CW (LiCOR) on an Innoscan 710-AL scanner (Innopsys). For staining of total protein, arrays were incubated 30 min in Super G blocking buffer (Grace Biolabs), rinsed in water, incubated 5 min in 0.000005 % Fast green FCF (Sigma) and rinsed again in water. Raw data were normalised using Normacurve [16].

Statistics

The distributions of the gene mRNA levels were characterised by their median values and ranges using GraphPad Prism software (version 7.05.237). For the unsupervised hierarchical clustering of the RPPA data, Pearson was used as a distance measure and Ward as the clustering algorithm. Paired *t*-tests, Mann Whitney test, Kruskal Wallis test, Fisher's exact, Wilcoxon matched-pairs signed rank test and Spearman rank correlation test were used as appropriate. Differences were considered significant at confidence levels greater than 95 % ($P \leq 0.05$)

Other materials and methods

Additional information regarding cell culture conditions, inducible lentiviral vectors and online data for HCC livers, are provided as Supplementary Materials and Methods.

Results

In vitro assessment of sensitivity to the PARP inhibition alone and in combination with IR

Our previous *in vitro* studies showed variability in PARPi sensitivity with a short (2 h) exposure to Veliparib (10μM) reducing clonogenic survival in HepG2 but not PLC/PRF/5 or Hep3B liver cells [5]. However, in both HepG2 and the more radiation resistant PLC/PRF/5 cells, PARP inhibition lowered IR-induced cell survival based on the mean lethal dose (D_{37}) [5]. We demonstrate here that a longer Veliparib (10μM) exposure (24 h) significantly decreased cell survival in all three cell models (Fig. 1). In addition, exposure to a second PARPi, Talazoparib, chosen because it is a more potent PARPi than Veliparib in terms of induced cytotoxicity [17] at a dose (50 nM/24 h) that inhibited PARP activity as assessed by polyADPribose levels after doxorubicin treatment (Supplementary Fig. 1B), significantly lowered cell survival compared to untreated cells. Under these experimental conditions a differential impact between the two PARPi on cell survival was only seen in HepG2 cells (54.2 ± 12.7 % Veliparib *vs* 33 ± 7.9 % Talazoparib, $P \leq 0.001$) (Fig. 1A).

When combined with a fixed 2 Gy X-ray dose Veliparib (10μM/24 h) exposure significantly attenuated the radiation sensitivity of HepG2 (Fig. 1A) and Hep3B (Fig. 1C) but not PLC/PRF/5 cells (Fig. 1B). In contrast, the combined treatment of Talazoparib (50 nM/24 h) and X-rays (2 Gy) significantly lowered cell survival in all three models compared

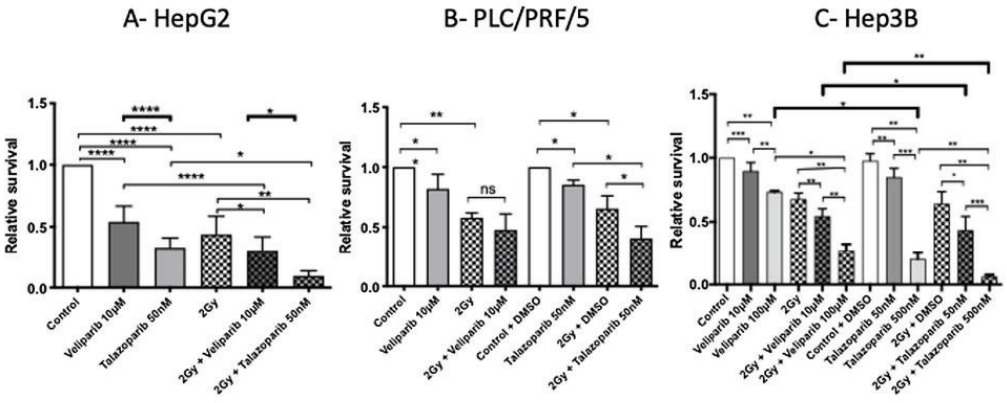


Fig. 1 Clonogenic cell survival after PARP inhibition and/or IR exposure. Cell survival assessed by colony formation in (A) HepG2, (B) PLC/PRF/5 and (C) Hep3B cells after exposure to the PARP inhibitors Veliparib or Talazoparib (24 h) and/or IR. (2 Gy). Data represents the mean +/- SD of ≥ 3 independent experiments carried out in triplicate. *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001, ns = non-significant. (Mann Witney U-test).

to radiation treatment alone, with the lowest survival seen in HepG2 cells (9.5 ± 4.2 %) (Fig. 1A). As Hep3B cells are reported to be both HR and NHEJ repair competent (see for instance [18]), the impact of 10-fold higher inhibitor doses was also investigated. These higher inhibitor doses significantly reduced cell survival with Talazoparib alone and in combination with radiation having a significantly greater impact than either Veliparib +/- radiation exposure (Fig. 1C).

Impact of HBx viral protein on cell survival after IR and PARP inhibition

As HBx expression leads to the degradation of the cell's SMC5/6 complex that is involved in processes that could modulate the impact of PARPi we next examined its impact on cell survival using HepG2-derived cell lines differing in HBx expression: HepG2 2.2.15 cells express the whole HBV genome [10] and HepG2 K6 cells carry a mutant HBV that encodes all HBV proteins except HBx [11]. In the absence of an efficient antibody against HBx, its expression was routinely deduced from SMC6 degradation in cellular protein extracts using western blotting (Supplementary Figure 1Ai).

Under the experimental conditions tested, the clonogenic survival after treatment with either Veliparib (10μM/24 h) +/- X-rays (2 Gy) was significantly lower in the presence of HBx (Fig. 2A). A second non-tumoral cell model (HepaRG TRX) was also investigated in which HBx is transiently induced for 5 days thus reducing longer-term selection effects or adaption to HBx expression. In the HepaRG TRX model HBx expression had no effect on relative cell numbers after Veliparib exposure (24 h/10μM) but cell growth was significantly reduced after Talazoparib (24 h/50 nM) (Fig. 2B) or IR (4 Gy X-rays) exposure. The combination of Talazoparib and IR significantly attenuated relative cell growth compared to radiation exposure alone. These response profiles probably reflect both the higher efficiency of PARP1 trapping by Talazoparib and the involvement of the SMC5/6 complex in the removal of such complexes in replicating cells.

DNA damage levels were next assessed in the HBx-expressing cells after X-ray exposure (4 Gy) using the formation and persistence of gamma-H2AX foci as a surrogate DNA DSB marker. A rapid increase in gammaH2AX foci/cell was seen after radiation exposure with a subsequent decline in foci numbers in the HepG2 parental and HepG2 K6 cells starting 30 min/1 h post-irradiation. In contrast, in HepG2 2.2.15 foci levels remained at the high levels seen initially after treatment over the first 2 h post-irradiation after which a decline in numbers was noted. However, foci numbers remained substantially higher than in HepG2 and HepG2 K6 cells at later time points studied suggestive of altered DNA repair capacity (Fig. 2C). Similarly, in the HepaRG model a delayed repair of DNA damage was observed in the HBx-expressing cells with higher numbers of foci persisting until 4 h post-irradiation after which a decline in numbers was observed (Fig. 2D) in support of the persistence of higher levels of DNA damage in the presence of HBx and associated with SMC5/6 complex loss.

HBx viral protein affects survival under replicative stress

As the SMC5/6 complex is also involved in a late step of cellular DNA replication, the sensitivity of HBx-expressing cells to HU induced cell killing was next assessed. This potent inhibitor of ribonucleotide reductase depletes the cellular dNTP pool resulting in stalled replication forks that, after a prolonged HU treatment (24 h/2 mM), collapse into DSBs resulting in S-phase cell cycle checkpoint activation. Under these exposure conditions HBx-expressing HepG2 2.2.15 cells had a significantly lower survival than either the parental or HepG2 K6 cells (Fig. 3Ai). A significant reduction in relative cell growth was also seen in HBx-expressing HepaRG cells (Fig. 3Bi).

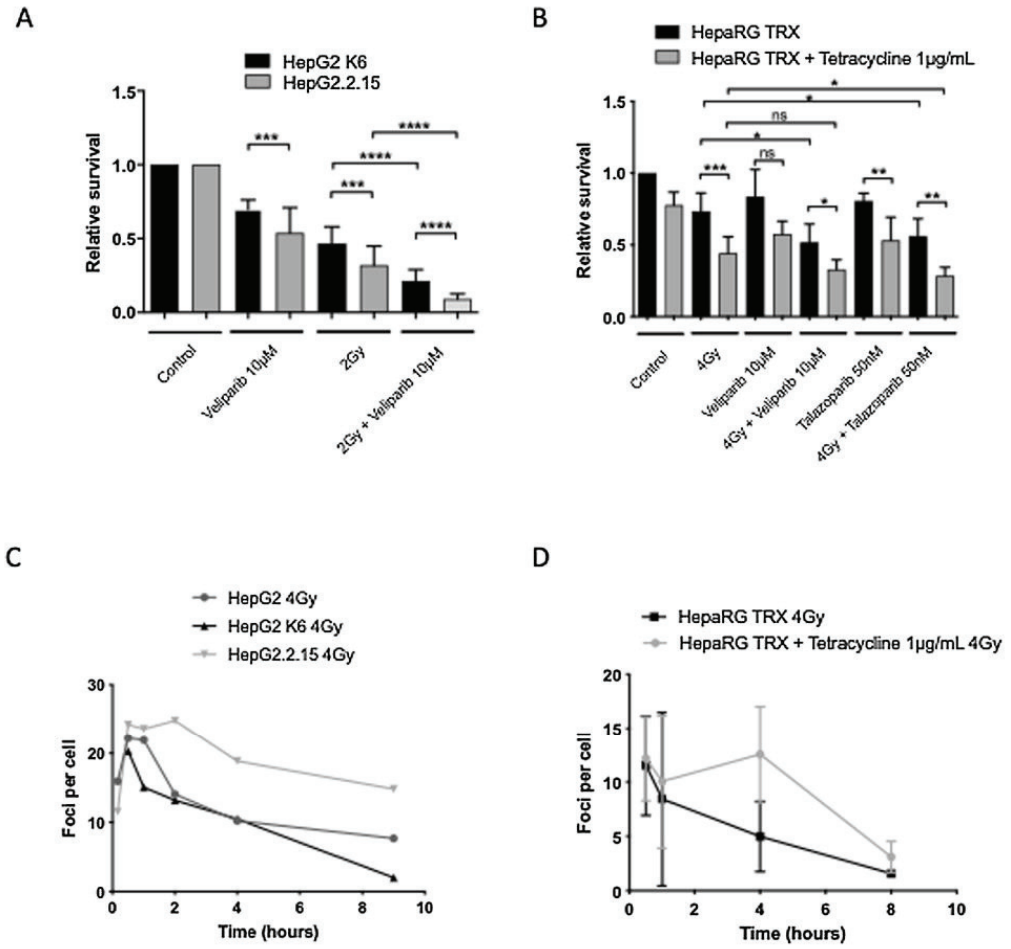


Fig. 2 Impact of HBx expression on cell survival and DNA damage levels. (A) Mean +/- SD survival in HepG2 2.2.15 (expressing HBx) or HepG2 K6 (no HBx) after Veliparib (24 h) +/- IR (2 Gy). (B) Mean +/- SD cell survival of HepaRG-TRX cells expressing HBx (+ tetracycline) or not after Veliparib or Talazoparib +/- IR. *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001, ns = non-significant (Mann Witney U-test, data from ≥ 3 independent experiments). (C) GammaH2AX foci numbers after IR (4 Gy) in HepG2-derived cells (one representative experiment, ≥ 200 cells/condition). (D) Mean +/-SD GammaH2AX foci numbers in HepaRG-TRX cells +/- HBx after IR (4 Gy) (3 independent experiments, ≥ 200 cells/condition).

The presence of blocked replication forks and the generation of DSBs caused by HU treatment will activate the intra-S cell cycle checkpoint to block the exit from S-phase of cells allowing time for DNA repair to occur. Whilst no HBx-dependent difference in the proportion of cells in S-phase were seen after HU treatment in either cell model (Fig. 3Aii/Bii), HBx-expressing cells had higher levels of gammaH2AX24 h after HU treatment indicative of higher persisting levels of DNA DSBs (Fig. 3Aiii/Biii). These profiles would suggest that the sensitivity to the cell killing effects of HU associated with HBx expression was due to an impairment in the processing of DNA strand breaks and not a defect in cell cycle checkpoint activation.

Is the HBx-associated cell sensitivity due to the degradation of SMC6 or other transcriptional changes?

As a first approach to address whether the sensitivity to IR, HU and PARPi in HBx-expressing cells is solely due to SMC5/6 complex degradation or whether HBx modulates other DNA Damage response (DDR) components that in-turn impacts on sensitivity to these agents we compared the transcriptome of proliferating HepaRG tetracycline-induced HBx cells *vs* cells depleted in SMC6 using an sh-targeting approach. Most differentially expressed genes were unique to each of these

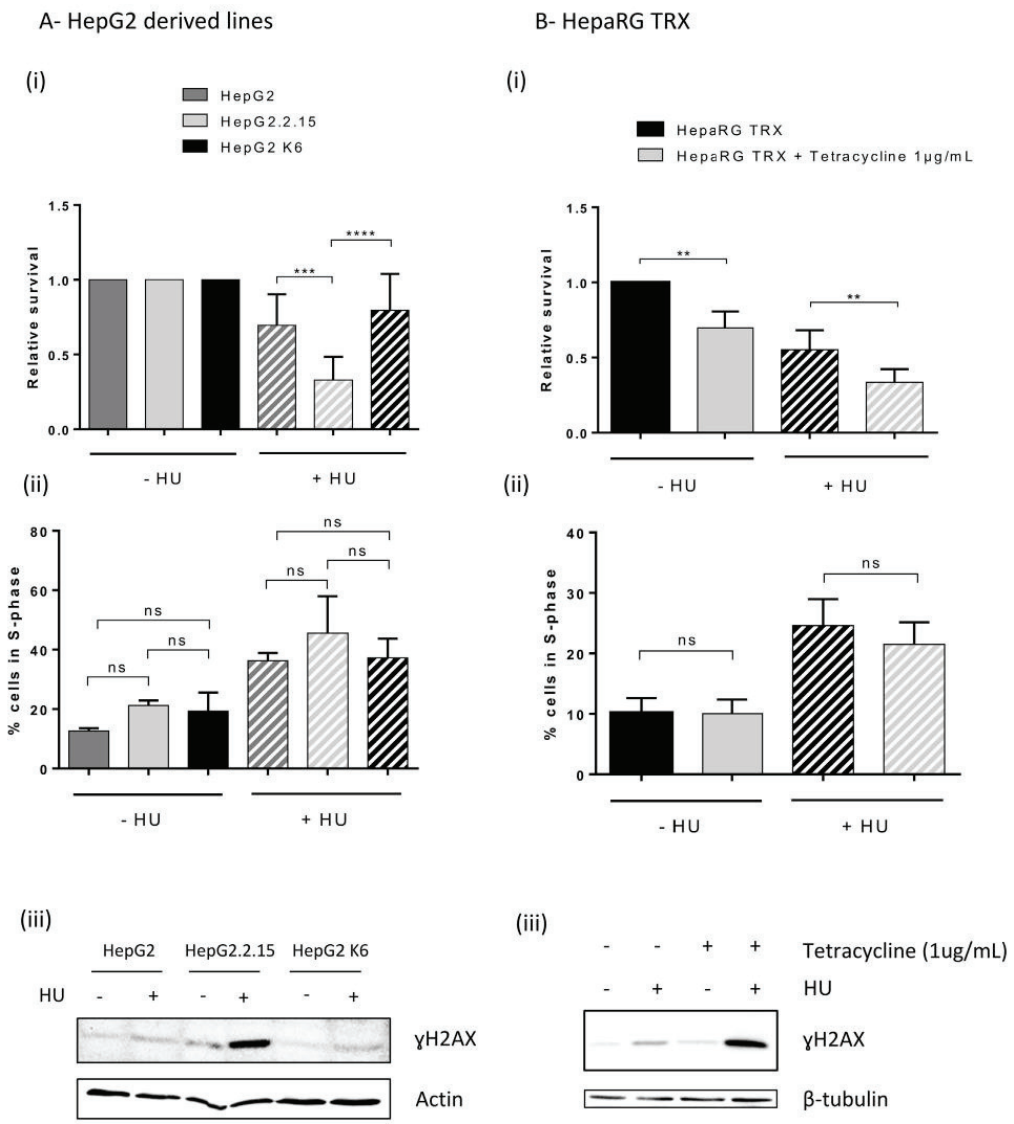


Fig. 3 Impact of HU on cell survival and cell cycle progression in HepG2 and HepaRG TRX-derived cells. (Ai and Bi) cell survival (mean +/- SD, ≥ 3 independent experiments) after HU (2 mM/24 h) exposure. (Aii and Bii) % of cells in the S phase of the cell cycle 24 h after treatment with HU (2 mM) (mean +/- SD, ≥ 3 independent experiments). (Aiii and Biii) Level of gammaH2AX in cell protein extracts after exposure to HU (2 mM/24 h). One representative experiment shown. **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001, ns = non-significant (Mann–Whitney U-test.).

two conditions with only 24 total overlapping genes (Supplementary Fig. 2). Of note no significant changes in the transcript levels of the PARP transcripts (PARP1, PARP2 and PARP3) or PARG were noted when either HBx was induced or SMC6 depleted.

Next, we focused on those genes commonly up- or down-regulated by HBx induction and SMC6 depletion, with the assumption that these genes are surrogate markers of SMC6 loss under HBx expression. Ten genes displayed such behaviour (Table 1 and Supplementary Table 1). Three were significantly down-regulated (PARPBP (also known as PARI), PBLD and LRRCC1) of which PARPBP has a potential DDR role and thus could contribute to the increased sensitivity seen in HBx-expressing cells to the panel of DNA damaging agents tested. Seven genes showed an up-regulation (SRPK2, YAF2, PRR19, HSPB3, CA12, SP3 and SRPRA).

Table 1 Genes co-regulated by HBx expression and SMC6 loss[a].

Gene	Transcript	logFC Hbx[b]	adj.P.Value[c]	logFC shSMC6[b]	adj.P.Value[c]
PBLD	uc001jns.1	−10.03	9.15E-03	−8.78	1.59E-02
PARPBP	uc001tjk.3	−8.23	4.19E-04	−8.86	3.14E-04
LRRCC1	uc010lzz.2	−2.76	4.79E-02	−8.50	1.35E-04
HSPB3	uc003jph.2	5.84	5.89E-03	6.58	3.33E-03
SRPRA	uc010sbm.2	6.02	4.32E-02	8.75	5.80E-03
SP3	uc002uig.3	6.31	3.35E-02	8.09	8.33E-03
CA12	uc002ame.3	6.99	8.15E-03	7.59	5.37E-03
YAF2	uc010sko.2	7.89	7.68E-05	7.69	1.08E-04
PRR19	uc002oth.1	8.05	7.43E-04	6.83	1.92E-03
SRPK2	uc003vct.4	8.49	2.71E-05	10.10	2.62E-05

[a] Genes that are co-regulated by HBX expression and SMC6 loss are shown in bold in the Venn diagram (Supplementary Fig. 2).
[b] Log fold change in transcript expression in presence of HBx or loss of SMC6 (shSMC6), - represents a reduction in expression.
[c] P values are false discovery rate corrected.

In addition to these overlapping genes, 27 genes were down- and 33 up-regulated when HBx was expressed and 89 genes were down- and 31 up-regulated when SMC6 was depleted (Supplementary Fig. 2, Supplementary Tables 2 and 3). KCNV1 and ZNF showed alternative splicing when HBx was expressed. DDR genes did not appear to be over-represented in these panels. For instance, no changes in BRCA1 or RAD51 transcript levels that can influence PARP inhibitor sensitivity [17] were noted. However, BABAM2 and RAD52 expression, both encoding genes involved in DSB repair, were increased in the presence of HBx. This might be a compensation mechanism for the loss of SMC6 and clearly needs further investigation.

Profile of PARP gene expression and DNA damage levels in HCC tissues

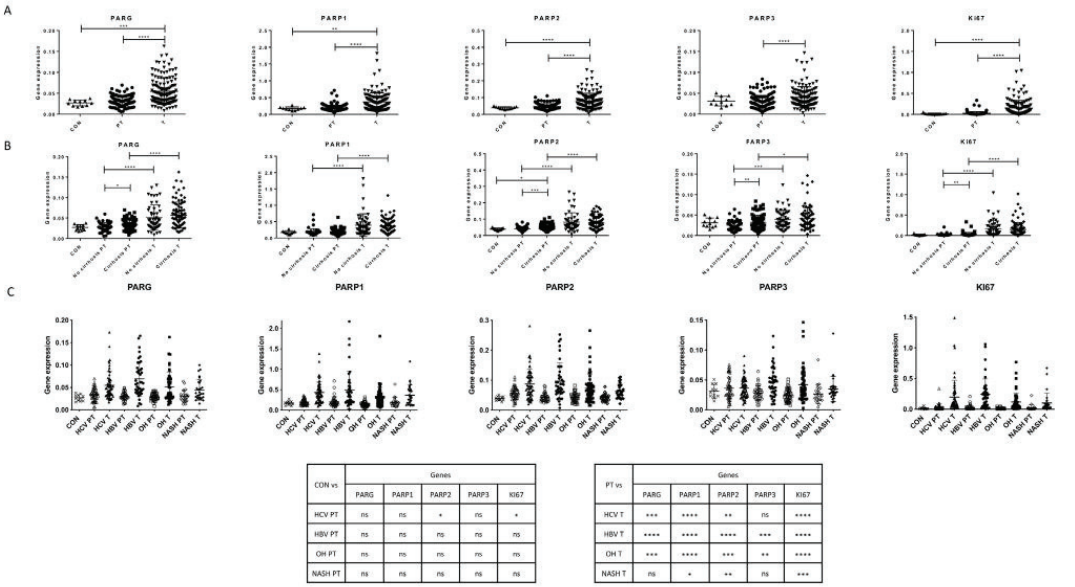
To determine the expression profile of PARP genes in HCC tissues, we assessed PARP1, PARP2 and PARP3 transcript levels together with those of PARG and MKI67 in PT and matched T tissues from a cohort of individuals with HCV-, HBV-, NASH- and alcohol-associated HCCs with or without cirrhosis (cohort characteristics are detailed in Supplementary Table 4) compared to a panel of control liver tissues. A significant increase in all the transcript studied was observed in T compared to PT tissues (Fig. 4A). Cirrhosis in PT tissues was associated with significantly higher transcript levels of PARG, PARP2, PARP3 and MKI67 compared to non-cirrhotic PT tissues but only for PARP2 were levels significantly higher than control tissues (Fig. 4B). Significantly higher levels of all the transcripts were found in T tissues compared to PT or control tissues irrespective of cirrhosis status.

In terms of aetiology-specific transcript profiles, only PT tissues from HCV-associated HCCs had significantly higher levels of PARP2 and MKI67. This difference may in part reflect that HCV-associated PT tissues had the highest proportion of tissues with cirrhosis. For all 4 aetiologies PARP1, PARP2 and MKI67 transcript levels were significantly higher in T than corresponding PT tissues, with increases in PARG seen in T compared to PT tissues from HCV-, HBV- and alcohol-associated HCCs and for PARP3 in HBV- and alcohol-associated HCCs (Fig. 4C).

Protein extracts available from a subset of HCV, HBV and alcohol-associated HCC tissues were used to assess the protein expression levels of PARP1, MKI67, gammaH2AX and H2AX. RPPA validated antibodies were not available for the other genes investigated. First, we assessed if PARP1 or MKI67 mRNA and protein levels were correlated. Weak but significant correlations for PARP1 ($r = 0.31$, $P = 0.0387$), and MKI67 ($r = 0.45$, $P = 0.0005$) mRNA/protein levels (Supplementary Fig. 3A) were found in T tissues but not in control or PT tissues (data not presented). Next, we assessed the variation in protein levels during disease progression. A non-supervised hierarchical clustering of the results (Supplementary Fig. 3B) suggests the presence of at least two major subgroups of samples, cluster A being enriched in T samples, while cluster B contains more PT and controls (Chi-square test $P = 0.0003$), suggesting that T and PT samples differ in their expression levels for the measured proteins. Yet, despite higher transcript levels in T tissues compared to control and PT tissues, no significant increases in protein levels were found for PARP1 and MKI67 ($P = 0.382$, $P = 0.078$ respectively) (Fig. 5A). However, lower H2AX levels were associated with disease progression from control to PT to T tissues, whereas gammaH2AX levels increased, resulting in a significant increase in the ratio of gamma-H2AX/H2AX (Fig. 5B). Gamma-H2AX is an integrated marker of DNA damage and repair levels and an extremely sensitive organismal stress indicator [9]. Increasing gamma-H2AX levels during HCC tumorigenesis are thus indicative of increased stress and DNA damage and might contribute to the vulnerability to PARP inhibition.

Discussion

This study has shown that *in vitro* HCC models are sensitive to the cell killing effects of PARPi when used as a single agent and that PARPi significantly attenuated radiation sensitivity in all the liver models tested. Our data is also in agreement with the literature that Talazoparib is a more potent PARPi than Veliparib in terms of induced cytotoxicity [17]. Whilst the cytotoxic activities of PARPi were initially related to their ability to block DNA single strand break repair, it is now recognized that additional cytotoxic mecha-



CON vs	Genes				
	PARG	PARP1	PARP2	PARP3	KI67
HCV PT	ns	ns	*	ns	*
HBV PT	ns	ns	ns	ns	ns
OH PT	ns	ns	ns	ns	ns
NASH PT	ns	ns	ns	ns	ns

PT vs	Genes				
	PARG	PARP1	PARP2	PARP3	KI67
HCV T	***	****	**	ns	****
HBV T	****	****	****	***	****
OH T	***	****	***	**	****
NASH T	ns	*	**	ns	***

Fig. 4 Variation in gene expression in Control and HCC tissues. (A) Comparison of normalised transcript levels in 11 control (CON) liver tissues, 148 PT and 140 T tissues. (B) Comparison of normalised transcript levels in 11 control (CON) liver tissues and PT tissues (93 cirrhotic, 55 non-cirrhotic) and T tissues (88 cirrhotic, 52 non-cirrhotic). (C) Comparison of normalised transcript levels between 11 CON liver tissues and HCV-associated (47 PT, 56 T), HBV-associated HCC (39 PT, 41 T), OH-associated HCC (42 PT, 54 T), and NASH-associated HCC (23 PT, 29 T) tissues. Error bars represent median and interquartile range *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$, ns = non-significant (Mann-Whitney U-test for (A), Kruskal Wallis test for (B)).

nisms exist including the generation of PARP-DNA complexes which block replication [17]. Indeed, stalled replication forks are a major source of genome instability in proliferating cells and strategies that exacerbate replication stress are a promising avenue to improve anticancer therapies (see [19]). The presence of the HBx protein increased this cytotoxicity which can in part be explained by the increased replicative stress concomitant to the HBx-induced loss of the SMC5/6 complex. Of the genes modulated in HBx-induced and SMC5/6 depleted cells, whilst no changes in the PARP transcripts studied were noted, the reduction in PARPBP transcript levels may contribute to these increases. Indeed, PARPBP depletion has been shown to enhance replication stress and DNA-damage accumulation [20] which is concordant with the *in vitro* cellular phenotype of the HBx-expressing cells. It has been recently reported that, based on The Cancer Genome Atlas (TCGA) and Human Protein Atlas databases, PARPBP was significantly upregulated in HCC tissues compared with normal liver tissues ($P \leq 0.05$) and correlated with worse overall survival and recurrence-free survival [21]. Of the other genes down-regulated in HBx-expressing/SMC6 depleted cells PBLD was reported down-regulated in HCC tissues [22], [23], suggesting that it's expression can be modulated by HBx-independent mechanisms, with this deletion found to be an independent poor prognosis predictor in HCC patients [23].

In addition to transcriptional modulation, it cannot be excluded that the effect of HBx on *in vitro* cell survival and DNA damage levels may be related to protein/protein interactions with a partner implicated in DDR pathways as seen for the HBx-induced loss of the SMC5/6 complex. Whilst a comprehensive proteomics evaluation of the impact of HBx expression was outside the scope of this present study, altered DDR protein expression have been reported in HBx-expressing models that potentially could impact on cell survival. For instance, in HBx-transgenic mice decreased expression of Mre11 and Rad51 expression has been noted [24]. HBx can also bind the tumor suppressor protein p53 that can modulate both DNA strand break repair pathways [25] and nucleotide excision repair [26].

Gene expression analysis carries the potential to dissect the molecular heterogeneity of cancers and can distinguish on this basis sub-classes for certain cancer types, including HCC [27]. Unfortunately, this increasing knowledge has not yet resulted in biomarker discovery. Thus, improved clinical care and identifying the most relevant characteristics of HCC and personalised medicine remains a major challenge. A comprehensive review of the expression profiles of 450 human DDR genes in different cancers [28] highlighted the therapeutic opportunities for targeting DDR, however, for HCC data was only available for 18 T samples. Subsequent bioinformatics analyses of a panel of 59 DNA repair genes in the HCC data from the TCGA to define a co-regulated gene cluster would suggest that DNA repair stratification could be useful for predicting prognosis and designing clinical trials for targeted therapy [29]. A recent analysis of the response to 31 anticancer agents in a large panel of liver cancer cell lines whose protein, RNA and mutational signatures

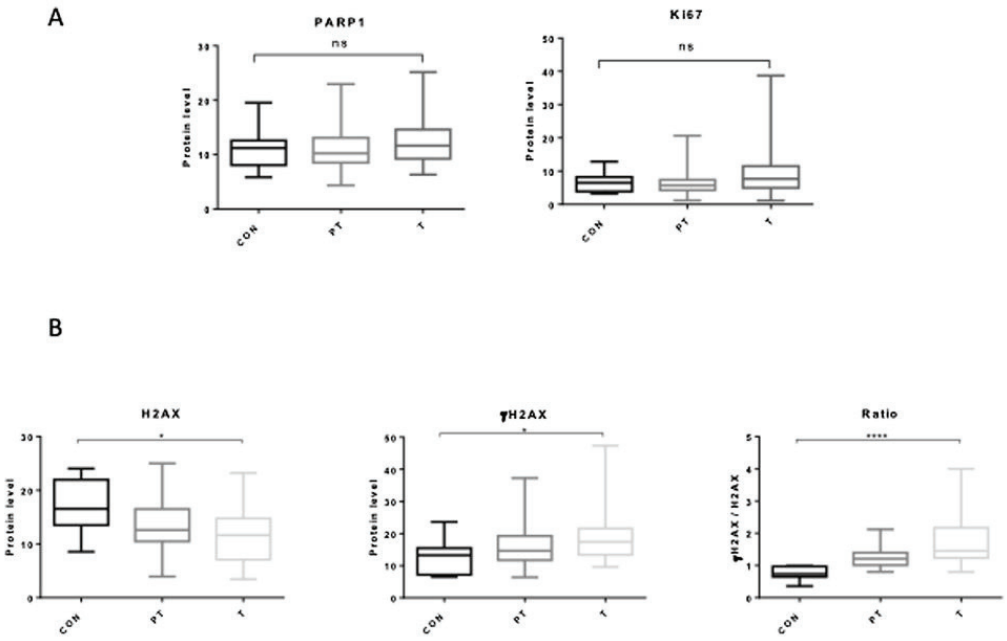


Fig. 5 Comparisons of protein levels in control, PT and T HCC tissues. (A) Comparison of normalised PARP1 and Ki67 protein levels and (B) normalised H2AX and gammaH2AX protein levels and their ratio in comparison in control (CON) (8), PT (47) and T tissues (46) from HCV, HBV and NASH-associated tumors. *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$, ns = non-significant (Kruskal Wallis test).

resembled that of the aggressive proliferation class of HCC, identified genetic alterations and gene expression patterns associated with response to these agents [30]. This information might be used to select patients for clinical trials but did not extend to DNA repair inhibitors such as PARPi that are proving promising for other cancer types. In addition, based on the impact of HBx observed here it would be of interest to determines its expression profile in HBV-associated PT and T tissues to determine whether it can impact on therapeutic responses.

There is some limited information on PARP family gene and protein expression profiles in HCC. In agreement with our data, increased PARP1 mRNA expression was reported based on the online TCGA HCC dataset [31] (and Supplementary Fig. 4) and in Chinese populations [31,32] but not found in a Polish population [33]. In our study, no correlation was found between mRNA and protein levels and no increase in PARP1 protein levels was observed in T tissues relative to control and PT tissues. This is in contrast to the data of Li et al. [31] and Xu et al. [34] who reported higher PARP1 protein levels in HCC T tissues. In agreement with our data PARP2 and PARG mRNA expression was elevated in T samples compared to the matched non-T samples in the on-line TCGA HCC dataset, however, no increases were seen for PARP3. The reasons for these discrepancies remain to be established but may be due to technical issues relating to antibody choice and tissues used for comparative purposes and reflect differences in the compositions of tissue panels and the impact of etiology on expression profiles. The elevated gammaH2AX seen in both this study and Evert et al. [35] are indicative of a higher DNA damage burden and/or replication stress that could be interpreted as an imbalance in DSB repair in HCC tissues and has been associated with cancer progression in many cancer types (see for example [36]). Intriguingly this increase is accompanied here by a reduction in H2AX levels during progression from control to T tissue. These changes parallel observations by Grusso and colleagues in triple negative breast cancer patients where chronic oxidative stress promotes H2AX degradation [37]. HCC is considered an inflammation-associated cancer, thus it is tempting to speculate that oxidative stress may be the underlying cause of the reduction in H2AX levels associated with liver disease progression. Grusso and colleagues also noted that a ROS-mediated H2AX decrease was observed after chemotherapy and was an indicator of the therapeutic efficiency and survival in Triple Negative Breast Cancer patients [37]. Whether combining PARP inhibitors that would block repair, with radiotherapy that would not only generate oxidative stress and thus further reduce H2AX levels, but also cytotoxic DNA damage, would enhance HCC tumor cytotoxicity and thus be beneficial in a therapeutic setting or lead to increased genomic instability and its consequences would need to be investigated in clinical trials.

Conclusion

In conclusion, the *in vitro* potentiation of cell death by PARPi alone or in combination with radiation exposure, taken together with the observations of elevated DNA damage levels in HCC T tissues, may represent a vulnerability that can be exploited for therapeutic benefit. Indeed, whilst RT has not been widely used for HCC treatment, it is a choice that needs further clinical evaluation [2] [3]. In addition to their direct impact on the DDR response, PARPi may have additional therapeutic benefits. For instance, the combination of IR and PARP inhibition can improve the vasculature leading to the reoxygenation of hypoxic T tissues, thus bypassing hypoxia-induced radioresistance (see [5] and references therein). Additionally, PARP inhibition can trigger the STING-dependent immune response and enhance the therapeutic efficacy of immune checkpoint blockade independent of BRCAness [38]. As the synergy between RT and the immune response is well documented the question arises if combining PARP inhibition, radiation, and immunotherapy could be a possible treatment strategy (see review [39]) and in particular for HCC where there is a growing body of evidence for each individually as therapeutic options.

Funding

This work was supported by funding from La Ligue Nationale Contre le Cancer, the ANRS (grant number 12357 to D.C.) and La Ligue Contre le Cancer Rhone Alpes.

Conflict of interest statement

None declared.

Acknowledgments

The authors would like to thank Hien Luong Nguyen, Magali Jacquet, Loïc Peyrot and Lydie Lefrancois who contributed to early stages of this project, David Cox, Agnes Tissier and Virginie Petrilli (CRCL, Lyon) for helpful discussions, J. Jacquemetton (CRCL, Lyon) for the development of macros for counting foci and Adeline Granzotto (CRCL, Lyon) for organising the irradiation of cells. The authors thank Pr. M. Rivoire (Centre Léon Bérard, Lyon) for providing human liver tissues.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.clinre.2020.09.014.

References

[1] Calderaro J, Couchy G, Imbeaud S, Amaddeo G, Letouzé E, Blanc J-F, et al. Histological subtypes of hepatocellular carcinoma are related to gene mutations and molecular tumour classification. J Hepatol 2017;67(4):727–38.
[2] Ohri N, Dawson LA, Krishnan S, Seong J, Cheng JC, Sarin SK, et al. Radiotherapy for hepatocellular carcinoma: new indications and directions for future study. JNCI: Journal of the National Cancer Institute 2016;108(9):djw133.
[3] Chino F, Stephens SJ, Choi SS, Marin D, Kim CY, Morse MA, et al. The role of external beam radiotherapy in the treatment of hepatocellular cancer. Cancer 2018;124(17):3476–89.
[4] Guillot C, Hall J, Herceg Z, Merle P, Chemin I. Update on hepatocellular carcinoma breakthroughs: poly(ADP-ribose) polymerase inhibitors as a promising therapeutic strategy. Clin Res Hepatol Gastroenterol 2014;38(2):137–42.
[5] Guillot C, Favaudon V, Herceg Z, Sagne C, Sauvaigo S, Merle P, et al. PARP inhibition and the radiosensitizing effects of the PARP inhibitor ABT-888 in in vitrohepatocellular carcinoma models. BMC Cancer 2014;14(1):603.
[6] Decorsière A, Mueller H, van Breugel PC, Abdul F, Gerossier L, et al. Hepatitis B virus X protein identifies the Smc5/6 complex as a host restriction factor. Nature 2016;531:386–9.
[7] Aragón L. The Smc5/6 complex: new and old functions of the enigmatic long-distance relative. Annu Rev Genet 2018;52(1):89–107.
[8] Mégnin-Chanet F, Bollet MA, Hall J. Targeting poly(ADP-ribose) polymerase activity for cancer therapy. Cell Mol Life Sci 2010;67(21):3649–62.
[9] Bonner WM, Redon CE, Dickey JS, Nakamura AJ, Sedelnikova OA, Solier S, et al. γH2AX and cancer. Nat Rev Cancer 2008;8(12):957–67.
[10] Sells MA, Zelent AZ, Shvartsman M, Acs G. Replicative intermediates of hepatitis B virus in HepG2 cells that produce infectious virions. J Virol 1988;62(8):2836–44.
[11] Lucifora J, Arzberger S, Durantel D, Belloni L, Strubin M, Levrero M, et al. Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. J Hepatol 2011;55(5):996–1003.
[12] Smyth Gordon K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004;3:1.
[13] Bieche I, Pennaneach V, Driouch K, Vacher S, Zaremba T, Susini A, et al. Variations in the mRNA expression of poly(ADP-ribose) polymerases, poly(ADP-ribose) glycohydrolase and ADP-ribosylhydrolase 3 in breast tumors and impact on clinical outcome. Int J Cancer 2013;133(12):2791–800.
[14] Bonnin M, Fares N, Testoni B, Estornes Y, Weber K, Vanbervliet B, et al. Toll-like receptor 3 downregulation is an escape mechanism from apoptosis during hepatocarcinogenesis. J Hepatol 2019;71(4):763–72.
[15] Meseure D, Vacher S, Lallemand F, Alsibai KD, Hatem R, Chemlali W, et al. Prognostic value of a newly identified MALAT1 alternatively spliced transcript in breast cancer. Br J Cancer 2016;114(12):1395–404.
[16] Troncale S, Barbet A, Coulibaly L, Henry E, He B, Barillot E, et al. NormaCurve: A SuperCurve-Based Method That Simultaneously Quantifies and Normalizes Reverse Phase Protein Array Data. PLoS One 2012;7(6):e38686.
[17] Murai J, Huang S-Y, Das BB, Renaud A, Zhang Y, Doroshow JH, et al. Trapping of PARP1 and PARP2 by clinical PARP inhibitors. Cancer Res 2012;72(21):5588–99.
[18] Zhang L, Zhang F, Zhang W, Chen L, Gao N, Men Y, et al. Harmine suppresses homologous recombination repair and inhibits proliferation of hepatoma cells. Cancer Biol Ther 2015;16(11):1585–92.
[19] Liao H, Ji F, Helleday T, Ying S. Mechanisms for stalled replication fork stabilization: new targets for synthetic lethality strategies in cancer treatments. EMBO Rep 2018;19(9):e46263.
[20] Nicolae CM, O'Connor MJ, Schleicher EM, Song C, Gowda R, Robertson G, et al. PARI (PARPBP) suppresses replication stress-induced myeloid differentiation in leukemia cells. Oncogene 2019;38(27):5530–40.

[21] Yu B, Ding Y, Liao X, Wang C, Wang B, Chen X. Overexpression of PARPBP correlates with tumor progression and poor prognosis in hepatocellular carcinoma. Dig Dis Sci 2019;64(10):2878—92.
[22] Long J, Lang Z, Wang H, Wang T, Wang B, Liu S. Glutamine synthetase as an early marker for hepatocellular carcinoma based on proteomic analysis of resected small hepatocellular carcinomas. HBPD INT 2010;9(3):296—305.
[23] Li A, Yan Q, Zhao X, Zhong J, Yang H, Feng Z, et al. Decreased expression of PBLD correlates with poor prognosis and functions as a tumor suppressor in human hepatocellular carcinoma. Oncotarget 2015;7(1):524—37.
[24] Ahodantin J, Bou-Nader M, Cordier C, Mégret J, Soussan P, et al. Hepatitis B virus X protein promotes DNA damage propagation through disruption of liver polyploidization and enhances hepatocellular carcinoma initiation. Oncogene 2019;38:2645—57.
[25] Menon V, Povirk L. Involvement of p53 in the repair of DNA double strand breaks: multifaceted Roles of p53 in homologous recombination repair (HRR) and non-homologous end joining (NHEJ). Subell Biohem 2014;85:321—36.
[26] Becker SA, Lee TH, Butel JS, Slagle BL. Hepatitis B virus X protein interferes with cellular DNA repair. J Virol 1998;72(1):266—72.
[27] Calderaro J, Ziol M, Paradis V, Zucman-Rossi J. Molecular and histological correlations in liver cancer. J Hepatol 2019;71(3):616—30.
[28] Pearl LH, Schierz AC, Ward SE, Al-Lazikani B, Pearl FMG. Therapeutic opportunities within the DNA damage response. Nat Rev Cancer 2015;15:166—80.
[29] Lin Z, Xu S-H, Wang H-Q, Cai Y-J, Ying L, Song M, et al. Prognostic value of DNA repair based stratification of hepatocellular carcinoma. Sci Rep 2016;6(1):25999.
[30] Caruso S, Calatayud A-L, Pilet J, La Bella T, Rekik S, Imbeaud S, et al. Analysis of liver Cancer cell lines identifies agents with likely efficacy against hepatocellular carcinoma and markers of response. Gastroenterology 2019;157(3):760—76.
[31] Li J, Dou D, Li P, Luo W, Lv W, Zhang C, et al. PARP-1 serves as a novel molecular marker for hepatocellular carcinoma in a Southern Chinese Zhuang population. Tumor Biol 2017;39(7).
[32] Qi H, Lu Y, Lv J, Wu H, Lu J, Zhang C, et al. The long noncoding RNA lncPARP1 contributes to progression of hepatocellular carcinoma through up-regulation of PARP1. Biosci Rep 2018;38(3).
[33] Krupa R, Czarny P, Wigner P, Wozny J, Jablkowski M, Kordek R, et al. The Relationship Between Single-Nucleotide Polymorphisms, the Expression of DNA Damage Response Genes, and Hepatocellular Carcinoma in a Polish Population. DNA Cell Biol 2017;36(8):693—708.
[34] Xu X, Liu Z, Wang J, Xie H, Li J, Cao J, et al. Global proteomic profiling in multistep hepatocarcinogenesis and identification of PARP1 as a novel molecular marker in hepatocellular carcinoma. Oncotarget 2016;7(12):13730—41.
[35] Evert M, Frau M, Tomasi ML, Latte G, Simile MM, Seddaiu MA, et al. Deregulation of DNA-dependent protein kinase catalytic subunit contributes to human hepatocarcinogenesis development and has a putative prognostic value. Br J Cancer 2013;109:2654—64.
[36] Bartkova J, Hořejší Z, Koed K, Krämer A, Tort F, Zieger K, et al. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature 2005;434(7035):864—70.
[37] Gruosso T, Mieulet V, Cardon M, Bourachot B, Kieffer Y, Devun F, et al. Chronic oxidative stress promotes H2AX protein degradation and enhances chemosensitivity in breast cancer patients. EMBO Mol Med 2016;8(5):527—49.
[38] Shen J, Zhao W, Ju Z, Wang L, Peng Y, Labrie M, et al. PARPi Triggers the STING-Dependent Immune Response and Enhances the Therapeutic Efficacy of Immune Checkpoint Blockade Independent of BRCAness. Cancer Res 2019;79(2):311—9.
[39] Césaire M, Thariat J, Candéias MS, Stefan D, Saintigny Y, Chevalier F. Combining PARP inhibition, radiation, and immunotherapy: a possible strategy to improve the treatment of Cancer? Int J Mol Sci 2018;19(12):3793.

PARP inhibitors and radiation potentiate liver cell death *in vitro.* Do hepatocellular carcinomas have an Achilles' heel?

Supplementary Figure legends and Materials and Methods

Laetitia Gerossier[1†], Anaëlle Dubois[1†], Alexia Paturel[1], Nadim Fares[1**], Damien Cohen[1], Phillippe Merle[1,2], Joel Lachuer[1,3], Anne Wierinckx[1,3], Pierre Saintigny[1,4], Brigitte Bancel[5], Janick Selves[6], Anne Schnitzler[7], Bérengère Ouine[8], Aurélie Cartier[8], Leanne de Koning[8], Vincent Puard[8], Ivan Bieche[7], Hector Hernandez-Vargas[1], Janet Hall[1], Isabelle Chemin[1*].

1. Univ Lyon, Université Claude Bernard Lyon 1, INSERM, CNRS, Centre Léon Bérard, Centre de Recherche en Cancérologie de Lyon, Lyon, 69008, France.
2. Department of Hepatology, Groupement Hospitalier Nord, Hospices Civils de Lyon, Lyon, 69000 France
3. ProfileXpert, SFR-Est, CNRS UMR-S3453, INSERM US7, Lyon Cedex 08, F-69373, France
4. Department of Medical Oncology, Centre Léon Bérard, Lyon, France
5. Service d'Anatomopathologie, Groupement Hospitalier Est, Hospices Civils de Lyon, Lyon, 69000, France
6. Anatomie et cytologie pathologiques Pôle IUC Oncopole CHU Institut Universitaire du Cancer de Toulouse - Oncopole, Toulouse, F- 31059, France
7. Department of Genetics, Institut Curie, PSL Research University, Paris, F-75005, France
8. Department of Translational Research, Institut Curie, PSL Research University, Paris, F-75005, France

* To whom correspondence should be addressed. Tel: +33 472681973; Fax: +33 472681971 Email: Isabelle.chemin@inserm.fr
** present address: Department of Digestive Oncology Rangueil Universitary Hospital, Hopitaux de Toulouse, Toulouse France.
† These authors contributed equally to this work.

SUPPLEMENTARY FIGURE LEGENDS

Supp. Fig. 1. Characterisation of cell lines used. **(Ai)** Expression levels of SMC6 in HepG2, HepG2 2.2.15 (expressing HBx), HepG2 K6, PLC/PRF/5 and Hep3B assessed by western blotting. Cell extracts were prepared as described in materials and methods and 30 µg of protein extract loaded onto an 8% acrylamide gel, migrated and transferred to nitrocellulose membranes, the BSA-blocked membranes were probed with primary antibodies overnight/4°, followed by secondary antibodies (1h/room temperature) and proteins revealed. B-actin was used a gel loading control and SMC1 as an internal control. **(Aii)** SMC6 expression levels in the HepaRG cells transduced with the shSMC6 targeting constructs (A to D) (reference Origene TL301502A, B, C and D) or the shscramble (TR30021, lot 0116) (left panel) and in HepaRG TRX cells induced with tetracycline (+tet) or non-induced (-tet) (right panel) determined by western blotting using the protocol outlined above. shSMC6 construct C was used for all experiments exploring the phenotype of SMC6 depleted cells and for RNA sequencing. **(B)** Activity of PARP inhibitors as determined by western blot analysis of the presence of poly(ADP-ribose) of Hep3B cells treated with the DNA damaging agent Doxorubicin. Protein extracts were prepared from cells pretreated +/- PARP inhibitor then 1h later with Doxorubicin (5µg/ml) for 1h. Cells were collected and lysed as described in supplementary materials and methods and extracts separated on a 4-15% acrylamide gel, transferred to nitrocellulose and probed with anti-pADPR antibody.

Supp. Fig. 2. Venn diagram showing the overlaps between pairwise comparison of genes significantly up- or down-regulated in HepaRG cells expressing the HBx gene or in which SMC6 was depleted using an sh-approach. The list of genes that are co-ordinately regulated (in bold) are given in Table 1 (and their functions in Supp. Table 1) and the other genes listed in Supp. Tables 2 and 3. † 2 genes show alternative splicing.

Supp. Fig. 3A. Comparison of mRNA and protein levels in 45 HCC tumor tissues of HCV, HBV and OH aetiology (16 tissues for HBV-, 16 HCV- and 13 OH-associated HCC T tissues). Open triangles HCV-, filled triangles HBV-, filled circles Alcohol-associated T tissues. **3B** Heat map of protein levels in panel of HCC PT and T tissues and liver controls assessed using RPPA. 1= PT or control tissues 2= T tissues. A non-supervised hierarchical clustering analysis found that Cluster A is enriched for T tissues and Cluster B for control and PT tissues (Chi square test P=0.0003).

Supp. Fig. 4. The expression of PARG and the PARPs' transcripts in TCGA data base**.** Comparison of transcript levels in 50 pairs of normal and HCC T tissues from the TCGA data set (Note: 23/50 are the same etiologies as our study population: 9 OH, 7, HBV, 5 HCV and 2 NASH, of the remaining samples 19 had no primary history, 5 other histories and for 3 no information was available).

SUPPLEMENTARY MATERIALS AND METHODS

Cell lines and growth conditions

The human hepatoma cell line HepG2 (ATCC HB-8065) and its derivatives HepG2 2.2.15 carrying the HBV wildtype genome and HepG2 K6 carrying a mutant form of HBV with no HBx protein were grown in Dulbecco's modified eagle medium containing 10% foetal bovine serum (FBS), 1% of GlutaMax, Penicillin, Streptomycin, sodium pyruvate, non-essential amino acids (NEAA) and 200µg/ml G418 for HepG2 2.2.15 and K6 (1,2).

HepaRG cells expressing the HBx protein under a tetracycline inducible promotor (HepaRG TRX)(2) were grown in Williams' E medium containing 10% foetal clone II serum (HyClone™), 1% of GlutaMax, Penicillin and Streptomycin, 0.1% Hydrocortisone (Upjohn), 0.025% Insulin (Sigma), 20µg/mL blasticidin, 200µg/ml Zéocin. TRX expression was induced with 1µg/ml tetracycline for 4/5 days.

HepaRG TRX sh-scramble expressing an in-efficient shRNA construct in a lentiviral GFP vector (Origene) and HepaRG TRX sh-Smc6 cells expressing a 29mer shRNA construct targeting the Smc6 transcript (Origene) were grown in HepaRG TRX medium.

Hep3B (ATCC HB-8064) and PLC/PRF/5 (ATCC CRL-8024) cells were grown in EMEM medium (Invitrogen) with heat inactivated 10% FBS and 1% each of Glutamax, Penicillin, Streptomycin, NEAA and sodium pyruvate.

Cells were routinely passaged when at 80% - 90% confluency by trypsinisation and replating at 1:10 to 1:5 densities.

Plasmid and Lentivirus preparation and Transfection of ShSMC6 and Shcontrols.

100µl TOP10 Chemically Competent E. coli was mixed with 100 ng of plasmid DNA (Origene Sh-constructs) and homogenates were incubated 30 minutes on ice. A thermic choc was performed of 42°C for 45 sec followed by 5min on ice and then bacteria were incubated 1h at 37°C. After cell plating, they were incubated overnight at 37°C. Plasmids were purified with the endotoxin-free plasmid DNA purification kit from Macherey Nagel.

293T cells were seeded at a density of about 2.4×10^6 cells per $75cm^2$ flask and were, 24h after, transfected overnight with plasmid DNA using the clonetech Lentivirus kit: 9µg of plasmid DNA, 9 µg gag pol vector, 2.4 µg env vector, and 62.5 µL of $CaCl_2$, 417.1µl of water and 500µl of HBS solution. The solution was incubated 15 min at room temperature and then added to the cell medium. 24h later the medium was changed and 48h and 72h after transfection, supernatants were harvested and filtrated (22µm), and centrifuged over an Amicon® Column at 3500rpm for 20 min at 4°C and stored at -80° until further use.

HepaRG cells were transduced with the Lentivirus constructs (shSMC6, shScramble) in 12-well plates in a total volume of 800µls of complete medium containing 0.5µl of polybrene and 500µl of lentivirus, overnight at 37°. After 16 hours the media was replaced and cells were amplified and selected with 1µg/mL puromycin. Cells were then passaged continuously in medium containing puromycin.

Assessment of PARP inhibitor activity

Cells (10^6 cells/25 cm^2 flask) were plated and left to attach overnight. After refreshing cell culture medium, the PARP inhibitor under investigation was added 1 hour before treating cells with the DNA damaging agent doxorubicin (5µg/ml). Cells were collected and protein extracts prepared using RIPA buffer (50mM Tris pH7.4, 250mM NaCl, 0.1% SDS, 2mM DTT, 0.5% NP40) containing a protease inhibitor cocktail, Halt™ and benzonase. 15 ugs of proteins were migrated on precast Criterion TGX 4–15% gels (BioRad). After migration of the proteins and their transfer to nitrocellulose membranes, the membranes were blocked with 5% nonfat milk for 1h at room temperature and then incubated with primary antibodies directed against pADPR (1/1000 dilution, clone ab14459, Abcam) overnight at 4°C (antibodies both diluted in 5% nonfat milk). After washing the membranes were incubated with secondary antibodies coupled to horseradish peroxidase (Dakocytomalion, diluted in 1/1000 in 5% nonfat milk) for 1h at room temperature. Membranes were probed with an enhanced by chemiluminescence reagent.

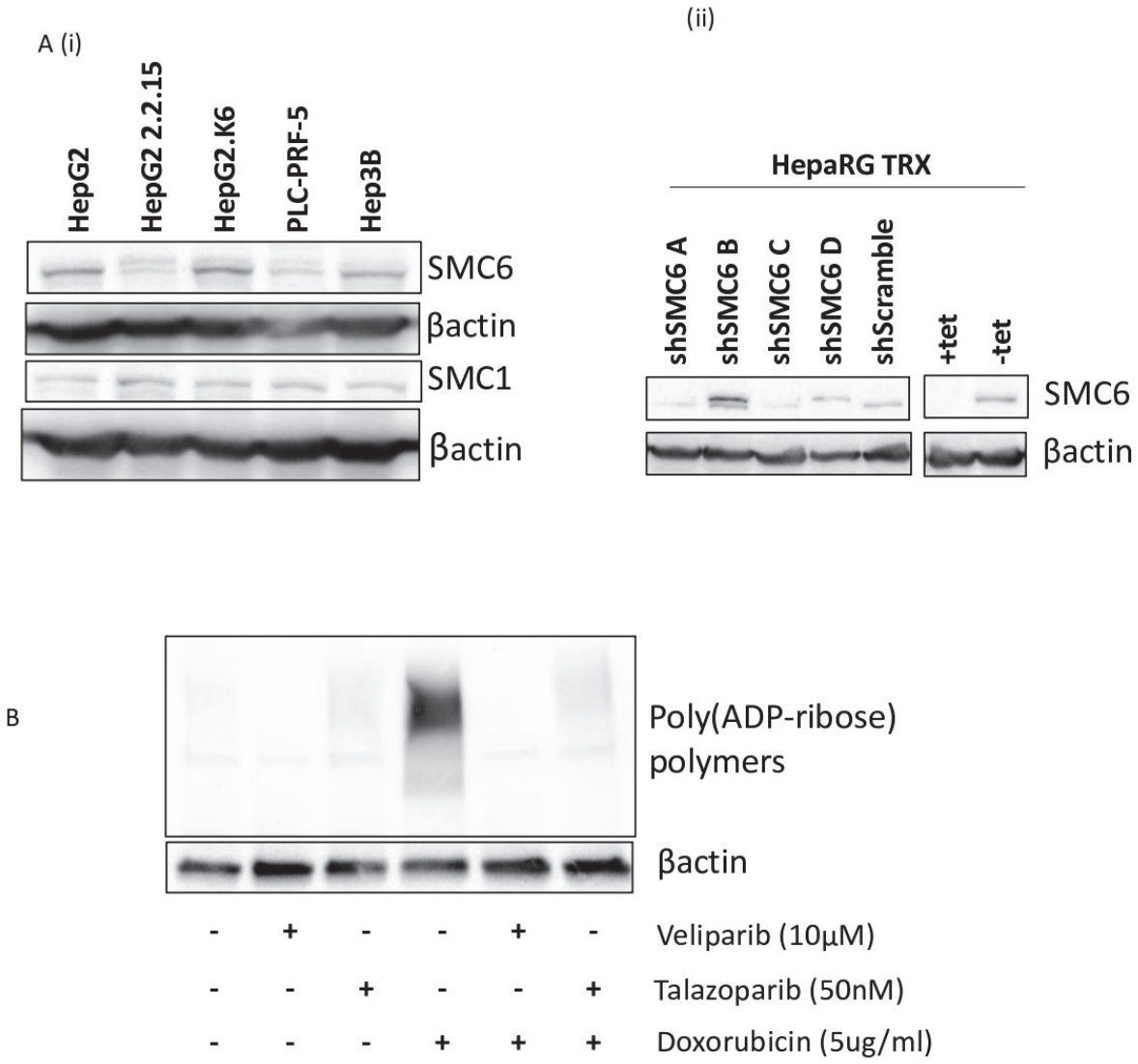
Online mRNA expression data for HCC livers

mRNA expression for the 50 resected HCCs with paired non-tumour tissue were obtained from The Cancer Genome Atlas (TCGA). Normalized gene-read counts generated from RNA sequencing (RNAseq version 2, level 3) were downloaded using the TCGA2STAT Rpackage (3) and log_2 transformed. Of note, prior to

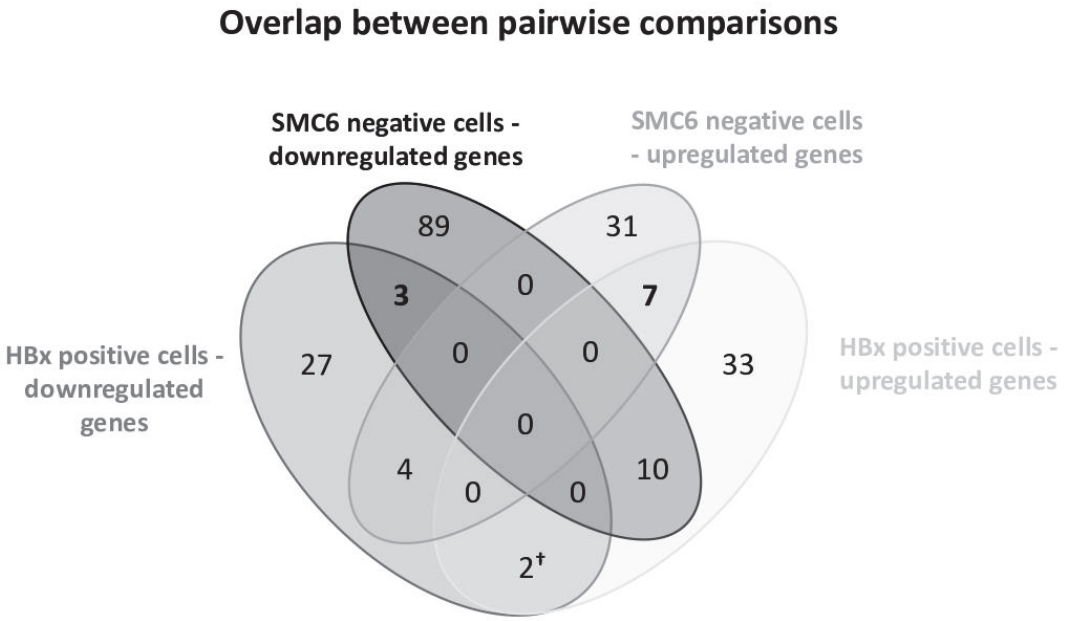
downloading data, normalization was conducted using MapSlice to do the alignment and RSEM to perform the quantification (4). Clinical data were retrieved from the cBioPortal data base (5).

1. Sells, M.A*., et al.* (1988) Replicative intermediates of hepatitis B virus in HepG2 cells that produce infectious virions. *Journal of Virology*, **62**, 2836.
2. Lucifora, J*., et al.* (2011) Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. *Journal of Hepatology*, **55**, 996-1003.
3. Wan, Y.-W*., et al.* (2015) TCGA2STAT: simple TCGA data access for integrated statistical analysis in R. *Bioinformatics*, **32**, 952-954.
4. Wang, K*., et al.* (2010) MapSplice: Accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Research*, **38**, e178-e178.
5. Cerami, E*., et al.* (2012) The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery*, **2**, 401.

Supplemental Figure 1

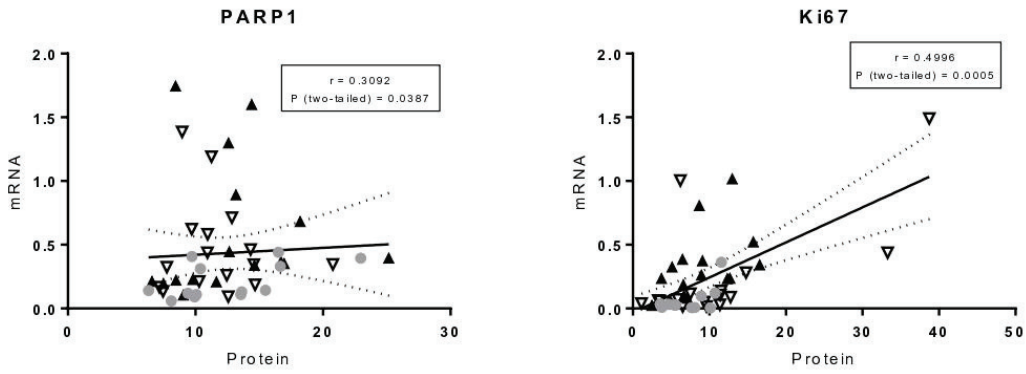


Supplemental Figure 2

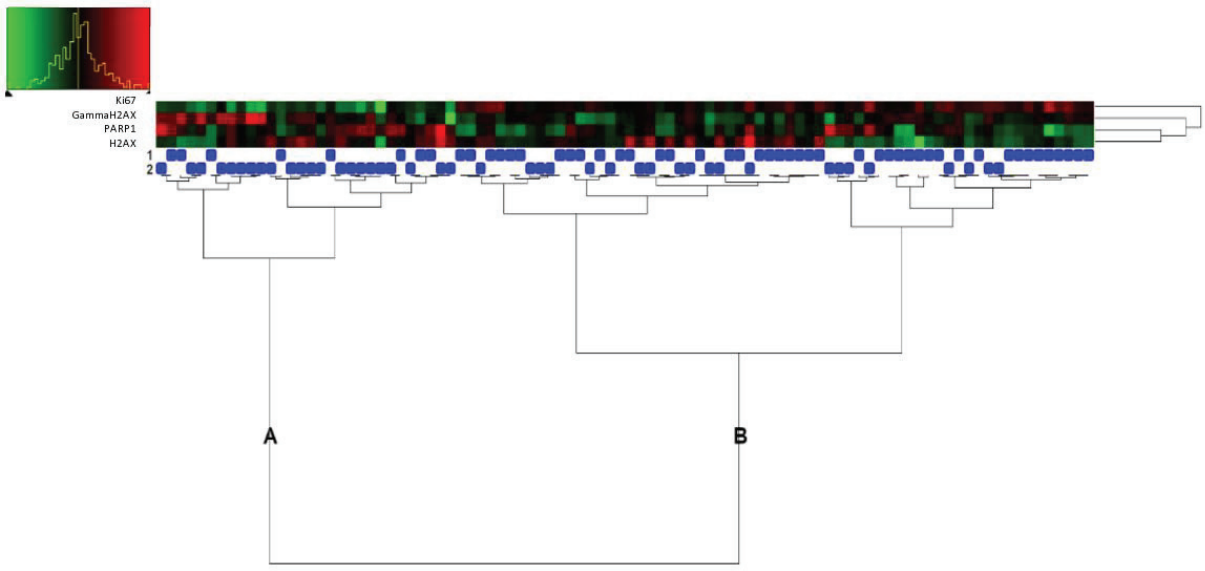


Supplemental Figure 3

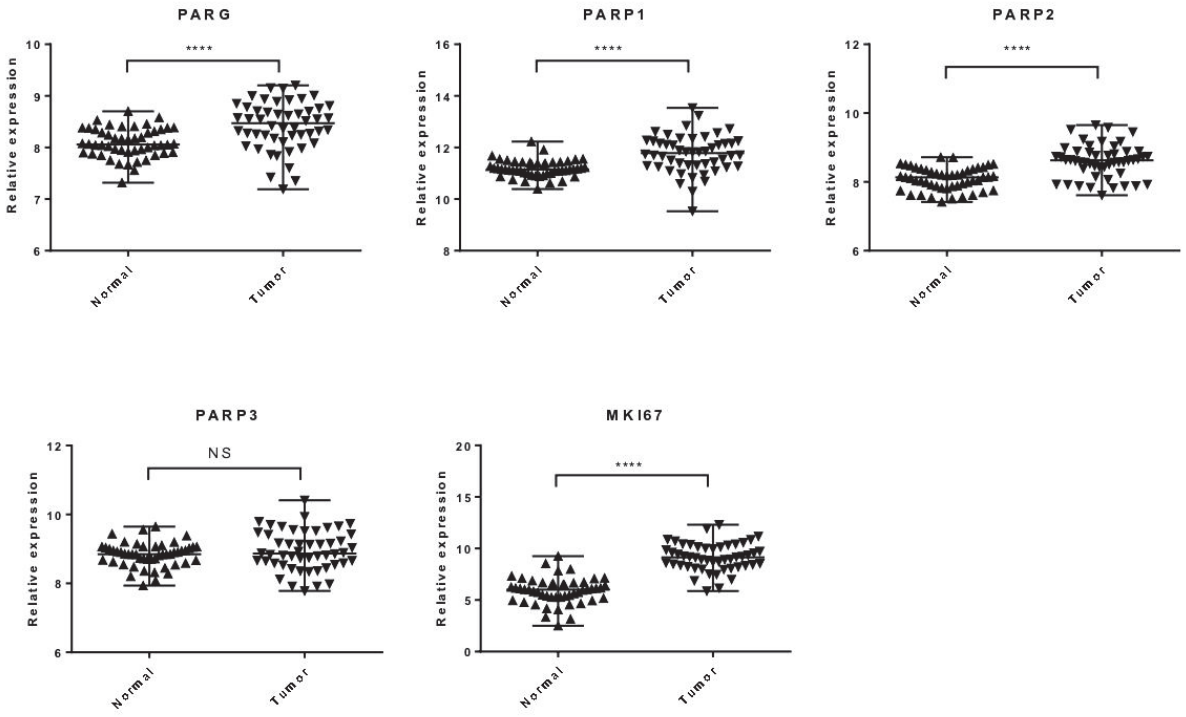
A



B



Supplemental Figure 4



Supplementary Table 1: Description of protein functions of genes regulated by both the expression of HBX and the loss of SMC6				
		Gene Card description†	UniProtKB/Swiss-Prot Summary ‡	GO molecular function
1	PBLD	PBLD (Phenazine Biosynthesis Like Protein Domain Containing) is a Protein Coding gene. Gene Ontology (GO) annotations related to this gene include *isomerase activity* .	Isomerase activity	Isomerase activity
2	PARPBP	PARPBP (PARP1 Binding Protein) is a Protein Coding gene. (Also known as PCNA interacting partner, PARI)	Required to suppress inappropriate homologous recombination, thereby playing a central role DNA repair and in the maintenance of genomic stability. Antagonizes homologous recombination by interfering with the formation of the RAD51-DNA homologous recombination structure. Binds single-strand DNA and poly(A) homopolymers. Positively regulate the poly(ADP-ribosyl)ation activity of PARP1; however such function may be indirect.	DNA binding,
3	LRRCC1	LRRCC1 (Leucine Rich Repeat And Coiled-Coil Centrosomal Protein 1) is a Protein Coding gene. Diseases associated with LRRCC1 include Meckel Syndrome, Type 1. An important paralog of this gene is DNAAF1. This gene encodes a centrosomal protein that maintains the structural integrity of the centrosome and plays a key role in mitotic spindle formation. The encoded protein contains an N-terminal leucine-rich repeat domain and a C-terminal coiled-coil domain. It associates with the centrosome throughout the cell cycle and accumulates on the mitotic centrosome.	Required for the organization of the mitotic spindle. Maintains the structural integrity of centrosomes during mitosis.	cell cycle, cell division
4	HSPB3	HSPB3 (Heat Shock Protein Family B (Small) Member 3) is a Protein Coding gene. Diseases associated with HSPB3 include Neuronopathy, Distal Hereditary Motor, Type Iic and Distal Hereditary Motor Neuronopathy Type 2. An important paralog of this gene is HSPB1. This gene encodes a muscle-specific small heat shock protein. A mutation in this gene is the cause of autosomal dominant distal hereditary motor neuropathy type 2C.	Inhibitor of actin polymerization.	response to unfolded protein
5	SRPRA	The gene encodes a subunit of the endoplasmic reticulum signal recognition particle receptor that, in conjunction with the signal recognition particle, is involved in the targeting and translocation of signal sequence tagged secretory and membrane proteins across the endoplasmic reticulum. Alternative splicing results in multiple transcript variants.	Component of the SRP (signal recognition particle) receptor. Ensures, in conjunction with the signal recognition particle, the correct targeting of the nascent secretory proteins to the endoplasmic reticulum membrane system.	GTPase activity, GTP binding, RNA binding, signal recognition particle binding,
6	SP3	This gene belongs to a family of Sp1 related genes that encode transcription factors that regulate transcription by binding to consensus GC- and GT-box regulatory elements in target genes. This protein contains a zinc finger DNA-binding domain and several transactivation domains, and has been reported to function as a bifunctional transcription factor that either stimulates or represses the transcription of numerous genes. Transcript variants encoding different isoforms have been described for this gene, and one has been reported to initiate translation from a non-AUG (AUA) start codon. Additional isoforms, resulting from the use of alternate downstream translation initiation sites, have also been noted.	Transcriptional factor that can act as an activator or repressor depending on isoform and/or post-translational modifications. Binds to GT and GC boxes promoter elements. Competes with SP1 for the GC-box promoters. Weak activator of transcription but can activate a number of genes involved in different processes such as cell-cycle regulation, hormone-induction and house-keeping.	Chromatin binding, DNA-binding transcription factor activity, RNA polymerase II-specific DNA-binding transcription repressor activity, RNA polymerase II-specific,metal ion binding proximal promoter sequence-specific DNA binding RNA polymerase II cis-regulatory region sequence-specific DNA binding, RNA polymerase II regulatory region sequence-specific DNA binding
7	CA12	Carbonic anhydrases (CAs) are a large family of zinc metalloenzymes that catalyze the reversible hydration of carbon dioxide. They participate in a variety of biological processes, including respiration, calcification, acid-base balance, bone resorption, and the formation of aqueous humor, cerebrospinal fluid, saliva, and gastric acid. This gene product is a type I membrane protein that is highly expressed in normal tissues, such as kidney, colon and pancreas, and has been found to be overexpressed in 10% of clear cell renal carcinomas. Three transcript variants encoding different isoforms have been identified for this gene.	Reversible hydration of carbon dioxide.	carbonate dehydratase activity, zinc ion binding
8	YAF2	This gene encodes a zinc finger containing protein that functions in the regulation of transcription. This protein was identified as an interacting partner of transcriptional repressor protein Yy1, and also interacts with other transcriptional regulators, including Myc and Polycomb. This protein can promote proteolysis of Yy1. Multiple alternatively spliced transcript variants have been found.	Binds to MYC and inhibits MYC-mediated transactivation. Also binds to MYCN and enhances MYCN-dependent transcriptional activation. Increases calpain 2-mediated proteolysis of YY1 in vitro. Component of the E2F6.com-1 complex, a repressive complex that methylates Lys-9 of histone H3, suggesting that it is involved in chromatin-remodeling.	transcription coactivator activity, transcription corepressor activity, metal ion binding,
9	PRR19	PRR19 (Proline Rich 19) is a Protein Coding gene	no information	no information
10	SRPK2	SRPK2 (SRSF Protein Kinase 2) is a Protein Coding gene. Diseases associated with SRPK2 include Brain Stem Infarction and Internuclear Ophthalmoplegia. Among its related pathways are mRNA Splicing - Major Pathway. Gene Ontology (GO) annotations related to this gene include *protein kinase activity* . An important paralog of this gene is SRPK1.	Serine/arginine-rich protein-specific kinase which specifically phosphorylates its substrates at serine residues located in regions rich in arginine/serine dipeptides, known as RS domains and is involved in the phosphorylation of SR splicing factors and the regulation of splicing. Promotes neuronal apoptosis by up-regulating cyclin-D1 (CCND1) expression. This is done by the phosphorylation of SRSF2, leading to the suppression of p53/TP53 phosphorylation thereby relieving the repressive effect of p53/TP53 on cyclin-D1 (CCND1) expression. Phosphorylates ACIN1, and redistributes it from the nuclear speckles to the nucleoplasm, resulting in cyclin A1 but not cyclin A2 up-regulation. Plays an essential role in spliceosomal B complex formation via the phosphorylation of DDX23/PRP28. Can mediate hepatitis B virus (HBV) core protein phosphorylation. Plays a negative role in the regulation of HBV replication through a mechanism not involving the phosphorylation of the core protein but by reducing the packaging efficiency of the pregenomic RNA (pgRNA) without affecting the formation of the viral core particles.	14-3-3 binding protein, ATP binding, magnesium ion binding, protein serine/theonine kinase activity, RNA binding,
		†www.genecards.org Stelzer G, Rosen R, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Iny Stein T, Nudel R, Lieder I, Mazor Y, Kaplan S, Dahary D, Warshawsky D, Guan - Golan Y, Kohn A, Rappaport N, Safran M, and Lancet D. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analysis , Current Protocols in Bioinformatics(2016), 54:1.30.1 - 1.30.33.doi: 10.1002 / cpbi.5.	‡www.uniprot.org The UniProt Consortium, UniPRot: a worldwide hub of protein knowledge, Nucleic Acids Res;46:D506-515 (2019)	

Supplementary Table 2 : Genes independently modulated by HBx expression and SMC6 loss and those showing alternative splicing in the presence of HBx

Lists 1-4 are genes independently regulated under only one experimental condition, Lists 5 and 6 are overlapping genes that show regulation in opposite directions by HBx/SMC6

† two genes (ZNF573 and KCNV1) show differences in isoform expression depending on HBx expression

	List 1: Down-regulated only in HBx expressing cells (29 genes)†	List2: Up-regulated only in HBx expressing cells (35 genes)†	List3: Down-regulated only in SMC6 depleted cells (89 genes)	List 4: Up-regulated only in SMC6 depleted cells (31 genes)	List 5: down-regulated by HBx and up-regulated in SMC6 depleted cells (4 genes)	List 6: up-regulated by HBx and down-regulated in SMC6 depleted cells (10 genes)
1	PTPN12	MIR877	MIR4517	RALGAPA2	AMDHD1	RANGRF
2	CTNND1	FDFT1	YY1AP1	GOSR1	ATP7B	PRXL2A
3	PDE4DIP	MSH5-SAPCD1	RSL1D1	ATF2	PEBP4	LARP4
4	ZKSCAN3	SNORD88B	DMTN	RARS2	MIPOL1	FRMD6
5	SNORA80B	MIR657	CNPPD1	ARHGEF1		UBL7
6	SHMT2	ANKS3	ELP2	TLK2		TMEM72
7	TMEM182	C2orf74	NCOA1	ASAH1		SCN8A
8	SBNO1	ZNF573†	OSTC	SHKBP1		MLLT10
9	KLHL20	OLFML3	EHD1	MGLL		RFC3
10	MIR100HG	DNAJB14	PTPRG-AS1	DMPK		PMS1
11	FMO1	IL22RA1	SPP1	SFSWAP		
12	AP2B1	MEF2B	FKBP14	PRMT2		
13	ZFP1	NLRP1	TCEANC	ZNF445		
14	C11orf54	RNF17	MVK	STAT6		
15	CDH18	PEX13	HNRNPA1L2	TPRKB		
16	MAP1S	ELAC2	PLIN3	TNIP3		
17	GDPD5	PATZ1	STXBP6	TSC22D3		
18	ZMYND8	COPE	CNTRL	FGFR1OP2		
19	IL7	PBRM1	SAA1	DDX58		
20	DEDD2	STYXL1	CRTC1	LINC00926		
21	DAG1	OGA	SEC16A	YTHDC1		
22	SERPINB5	CD27	SNX17	CRYAB		
23	CD63	WDR27	CAMK1D	GFOD1		
24	TFDP2	HSF2	DLGAP3	RAD54B		
25	SYNJ2	NKX2-1	HIST2H2AB	RALGPS1		
26	ZNF573†	CEP55	PSMC3IP	TUT7		
27	SOCS2	MIP	PDSS1	NRN1		
28	KCNV1†	BABAM2	SNORA53	RFC4		
29	GRIN2C	FOXO1	SUCO	YTHDF3		
30		SLC2A8	GLT8D1	RARG		
31		RAD52	DSCR9	HAT1		
32		KCNV1†	R3HDM2			
33		WDR37	HIST1H3C			
34		MAST2	ANKRD17			
35		CDC25B	PUDP			
36			ITGA10			
37			EDN2			
38			LVRN			
39			BRWD3			
40			SEC24D			
41			ARHGEF10L			
42			ZNF302			
43			SNORD17			
44			PFKL			
45			WDR47			
46			TCF4			
47			SHISA5			
48			LONRF3			
49			PGBD2			
50			PQBP1			
51			BTN3A1			
52			PDZD9			
53			CBFA2T2			
54			CCDC88A			
55			TIA1			
56			ZNF192P1			
57			MTMR6			
58			SMARCD3			
59			BID			
60			UBR2			
61			PPP1R2			
62			ATG16L2			
63			DBF4			
64			ATG13			
65			NEMF			
66			PSEN1			
67			AFG3L1P			
68			DLG5			
69			PHRF1			
70			TMEM86B			
71			LOC101059948			
72			DES			
73			DAAM2			
74			RABEP1			
75			CBX2			
76			ZCWPW2			
77			GFM2			
78			APTX			
79			LOC728392			
80			ASH2L			
81			KCTD20			
82			MYO16-AS1			
83			CRACR2B			
84			SYBU			
85			ENPP2			
86			TUBB3			
87			XRRA1			
88			LOC150776			
89			SVEP1			

Supplementary Table 3 : Transcripts modulated by HBx expression or SMC6 depletion (see Supplementary Figure 2 for Venn diagram)

Note: this list excludes the 10 genes that are coordinately modified by HBx expression and SMC6 depletion that are given in Table 1.

1. Transcripts up-regulated in SMC6 negative cells (35 genes in total, 4† are also down-regulated by HBx)

Transcript	logFC	adj.P.Val	seqnames	start	end	strand	Gene	
uc010zsg.2	9,293457	0,000135	chr20	20370272	20693266	-	RALGAPA2	1
uc002hfe.3	11,60334	0,000163	chr17	28804426	28853832	+	GOSR1	2
uc003pmf.3	10,54067	0,000229	chr6	88224096	88299735	-	RARS2	3
uc002ujs.4	9,658505	0,000399	chr2	175936978	176032934	-	ATF2	4
uc002osa.3	8,722362	0,000559	chr19	42388446	42411604	+	ARHGEF1	5
uc010ddp.3	6,975864	0,000888	chr17	60556386	60692841	+	TLK2	6
uc003wym.2	10,87618	0,001021	chr8	17913925	17941879	-	ASAH1	7
uc009zth.2	8,974111	0,001523	chr12	96337071	96362370	+	AMDHD1†	8
uc002ooe.3	10,55826	0,001662	chr19	41083137	41097305	+	SHKBP1	9
uc003ejv.4	8,137723	0,002051	chr3	127407905	127455200	-	MGLL	10
uc010xxt.1	9,032666	0,002271	chr19	46272976	46285815	-	DMPK	11
uc010tbn.2	8,056787	0,002628	chr12	132195632	132284283	+	SFSWAP	12
uc010tgt.1	7,054243	0,00493	chr13	52506806	52585630	-	ATP7B†	13
uc011aga.2	8,68828	0,00566	chr21	48055507	48085155	+	PRMT2	14
uc011azw.1	9,288239	0,00566	chr3	44481262	44561226	-	ZNF445	15
uc010srd.2	10,94135	0,00804	chr12	57489187	57505196	-	STAT6	16
uc003xcn.1	6,045085	0,008493	chr8	22570765	22785421	-	PEBP4†	17
uc002sjl.2	11,42273	0,008888	chr2	73956957	73961718	-	TPRKB	18
uc010ing.3	6,856491	0,008925	chr4	122052564	122085495	-	TNIP3	19
uc004enj.3	8,872419	0,008925	chrX	106956452	107019202	-	TSC22D3	20
uc001rhm.3	8,686853	0,008925	chr12	27091305	27119581	+	FGFR1OP2	21
uc010mjk.1	8,839982	0,009836	chr9	32466288	32526322	-	DDX58	22
uc010ugs.2	7,347045	0,009836	chr15	57592563	57599967	+	LINC00926	23
uc003hdy.3	9,571454	0,009836	chr4	69176105	69215824	-	YTHDC1	24
uc010rwp.1	5,867504	0,011098	chr11	111780486	111782473	-	CRYAB	25
uc001wud.3	9,262472	0,011713	chr14	37667118	38020464	+	MIPOL1†	26
uc003nau.3	7,67961	0,018571	chr6	13469502	13472295	-	GFOD1	27
uc003ygn.3	7,501066	0,021947	chr8	95439940	95487343	-	RAD54B	28
uc011mab.2	6,697432	0,030118	chr9	129677053	129946603	+	RALGPS1	29
uc004aov.3	9,621615	0,03361	chr9	88952242	88969402	-	TUT7	30
uc003mwu.3	7,12955	0,039397	chr6	5998233	6007633	-	NRN1	31
uc011bsc.2	8,730155	0,040908	chr3	186507682	186524484	-	RFC4	32
uc031tbi.1	9,415266	0,04468	chr8	64081562	64125346	+	YTHDF3	33
uc010sob.2	4,884862	0,048419	chr12	53604350	53614197	-	RARG	34
uc010fqi.2	5,793996	0,049198	chr2	172778935	172845768	+	HAT1	35

2. Transcripts down-regulated in SMC6 negative cells (99 genes in total, 10‡ are also up-regulated by HBx)											
	logFC	adj.P.Val	seqnames	start	end	strand	Gene				
uc021tfu.1	-13,4724	1,22E-05	chr16	28969904	28969982	+	MIR4517	1			
uc010pgk.2	-7,83913	2,04E-05	chr1	155638418	155658447	-	YY1AP1	2			
uc010buv.1	-11,4693	2,04E-05	chr16	11928055	11945442	-	RSL1D1	3			
uc022asr.1	-8,6468	2,98E-05	chr8	21912328	21940036	+	DMTN	4			
uc002vjv.3	-8,02494	3,54E-05	chr2	220036619	220042732	-	CNPPD1	5			
uc002kzn.2	-8,16818	3,54E-05	chr18	33709837	33754688	+	ELP2	6			
uc002rfi.3	-9,58975	3,54E-05	chr2	24787179	24993570	+	NCOA1	7			
uc002gkv.3	-9,14304	3,54E-05	chr17	8191969	8193409	+	RANGRF‡	8			
uc031sgu.1	-8,1062	3,78E-05	chr4	109571741	109588978	+	OSTC	9			
uc001obu.1	-8,95472	4,81E-05	chr11	64620208	64646191	-	EHD1	10			
uc010hno.4	-8,17208	7,35E-05	chr3	62247494	62304622	-	PTPRG-AS1	11			
uc011cde.2	-6,41621	9,03E-05	chr4	88896802	88904563	+	SPP1	12			
uc031sww.1	-9,38839	9,31E-05	chr7	30050199	30066417	-	FKBP14	13			
uc001kcd.4	-8,52	0,000111	chr10	82168242	82192753	+	PRXL2A‡	14			
uc004cvk.2	-9,23191	0,000131	chrX	13679020	13682247	+	TCEANC	15			
uc001tpc.4	-7,76518	0,000131	chr12	110012623	110035071	+	MVK	16			
uc001vgx.1	-9,6857	0,000167	chr13	53191605	53217919	+	HNRNPA1L2	17			
uc001rwn.3	-10,6171	0,000172	chr12	50794592	50855272	+	LARP4‡	18			
uc002mbl.3	-8,4096	0,000232	chr19	4838346	4867780	-	PLIN3	19			
uc001wpv.3	-9,385	0,000759	chr14	25281306	25519171	-	STXBP6	20			
uc004blb.1	-8,65234	0,000951	chr9	123914321	123939886	+	CNTRL	21			
uc001wzb.4	-6,62055	0,001021	chr14	51955839	52197444	+	FRMD6‡	22			
uc021qem.1	-10,9215	0,001162	chr11	18287808	18291523	+	SAA1	23			
uc002axx.1	-7,23201	0,001168	chr15	74738318	74753510	-	UBL7‡	24			
uc010ebv.3	-6,87535	0,001168	chr19	18794425	18893143	+	CRTC1	25			
uc010nbn.3	-7,40528	0,001495	chr9	139334548	139378211	-	SEC16A	26			
uc010ylq.2	-11,0028	0,001523	chr2	27594120	27600400	+	SNX17	27			
uc001ilo.3	-6,08371	0,001551	chr10	12391583	12871733	+	CAMK1D	28			
uc001jbn.2	-9,69189	0,001799	chr10	45406764	45430642	+	TMEM72‡	29			
uc001byc.3	-5,43952	0,002033	chr1	35331037	35370984	-	DLGAP3	30			
uc001ete.3	-7,14498	0,002259	chr1	149859019	149859466	-	HIST2H2AB	31			
uc010wgn.2	-8,13694	0,002628	chr17	40724328	40729849	-	PSMC3IP	32			
uc010qdf.2	-8,17564	0,002793	chr10	27023464	27035726	+	PDSS1	33			
uc001tfu.1	-10,7732	0,002863	chr12	98993413	98993662	+	SNORA53	34			
uc010pmo.3	-7,5379	0,003009	chr1	172502317	172580973	+	SUCO	35			
uc003dfl.3	-9,28789	0,003065	chr3	52728504	52740048	-	GLT8D1	36			
uc010gnk.3	-5,67064	0,003244	chr21	38580804	38594037	+	DSCR9	37			
uc001snr.2	-8,51106	0,00333	chr12	57647548	57690816	-	R3HDM2	38			
uc001rza.1	-8,34035	0,003519	chr12	52167973	52180610	+	SCN8A‡	39			
uc003nfv.3	-8,4794	0,003679	chr6	26045639	26046097	+	HIST1H3C	40			
uc003hgr.3	-10,4361	0,004083	chr4	73940502	74124502	-	ANKRD17	41			
uc011mho.1	-7,24441	0,004083	chrX	6975627	7066231	-	PUDP	42			
uc001eoa.3	-9,74307	0,004207	chr1	145524990	145543868	+	ITGA10	43			
uc001cgx.3	-9,02151	0,004352	chr1	41944446	41950344	-	EDN2	44			
uc003krr.3	-8,21521	0,004422	chr5	115331074	115363299	+	LVRN	45			
uc004edt.3	-7,42902	0,0058	chrX	79924987	80065233	-	BRWD3	46			
uc003icj.4	-9,45717	0,006131	chr4	119643978	119757326	-	SEC24D	47			
uc009vpe.1	-7,52402	0,00616	chr1	17866330	18016673	+	ARHGEF10L	48			
uc010xrz.1	-8,00237	0,006946	chr19	35168567	35175076	+	ZNF302	49			
uc002wqf.1	-9,93265	0,007407	chr20	17943353	17943589	-	SNORD17	50			

uc001iqx.1	-8,14506	0,00804	chr10	21823574	21884369	+	MLLT10‡	51			
uc011afd.1	-9,40479	0,008076	chr21	45719925	45733026	+	PFKL	52			
uc001dwj.3	-6,3351	0,00833	chr1	109512838	109584850	-	WDR47	53			
uc010xdy.1	-6,75792	0,00833	chr18	52889562	53257045	-	TCF4	54			
uc003ctp.2	-11,168	0,008613	chr3	48509197	48541682	-	SHISA5	55			
uc004eqy.3	-7,9842	0,008613	chrX	118108713	118151949	+	LONRF3	56			
uc001ifg.3	-7,88344	0,008925	chr1	249200442	249213345	+	PGBD2	57			
uc010nii.3	-8,92069	0,008925	chrX	48755775	48760422	+	PQBP1	58			
uc011dkj.2	-7,60955	0,009073	chr6	26402465	26415444	+	BTN3A1	59			
uc001uva.3	-7,56741	0,009154	chr13	34392206	34540695	+	RFC3‡	60			
uc021ter.1	-6,96301	0,010342	chr16	21995186	22012431	-	PDZD9	61			
uc021wbz.1	-7,89804	0,013622	chr20	32150171	32211773	+	CBFA2T2	62			
uc010fbw.2	-9,08613	0,013769	chr2	55514978	55536390	-	CCDC88A	63			
uc002sgm.3	-10,2856	0,01452	chr2	70443804	70475779	-	TIA1	64			
uc021yrq.2	-4,91827	0,015455	chr6	28129539	28137373	+	ZNF192P1	65			
uc010frz.3	-8,11955	0,015756	chr2	190656516	190742355	+	PMS1‡	66			
uc021rhi.1	-11,5123	0,015855	chr13	25831874	25834825	-	MTMR6	67			
uc003wjt.3	-8,96905	0,016838	chr7	150936059	150973820	-	SMARCD3	68			
uc002znc.2	-7,43578	0,018571	chr22	18216906	18256808	-	BID	69			
uc011dut.2	-6,10418	0,019972	chr6	42640975	42661243	+	UBR2	70			
uc003fuq.4	-5,11385	0,019972	chr3	195255532	195270224	-	PPP1R2	71			
uc009ytj.2	-6,81362	0,020699	chr11	72525451	72554717	+	ATG16L2	72			
uc011khf.1	-7,46586	0,02086	chr7	87505544	87538856	+	DBF4	73			
uc001ndb.3	-11,7765	0,021741	chr11	46638826	46697568	+	ATG13	74			
uc010anq.1	-8,46193	0,021947	chr14	50286814	50319539	-	NEMF	75			
uc010ark.3	-7,96829	0,021947	chr14	73603525	73690399	+	PSEN1	76			
uc002fqc.1	-7,17458	0,022379	chr16	90056740	90063028	+	AFG3L1P	77			
uc001jzj.3	-9,82496	0,022379	chr10	79550549	79613295	-	DLG5	78			
uc010qwe.2	-7,13316	0,023677	chr11	576483	612222	+	PHRF1	79			
uc002qjt.3	-2,98664	0,025303	chr19	55738002	55740141	-	TMEM86B	80			
uc031rlr.1	-7,56673	0,027376	chr19	49521505	49522741	-	LOC101059948	81			
uc002vll.3	-6,24831	0,027765	chr2	220283099	220291461	+	DES	82			
uc003ooy.4	-6,64107	0,028713	chr6	39816342	39828941	+	DAAM2	83			
uc010cld.1	-6,12368	0,028713	chr17	5185558	5281988	+	RABEP1	84			
uc002jxb.2	-6,81303	0,029266	chr17	77751977	77756331	+	CBX2	85			
uc003cei.3	-7,4266	0,030118	chr3	28431989	28566632	+	ZCWPW2	86			
uc010izk.1	-8,52009	0,032057	chr5	74017031	74063042	-	GFM2	87			
uc003zrn.3	-8,38983	0,032057	chr9	32972604	32997346	-	APTX	88			
uc010vtc.2	-6,77196	0,03331	chr17	5402747	5404319	-	LOC728392	89			
uc003xku.5	-8,20609	0,03331	chr8	37963311	37997598	+	ASH2L	90			
uc011dtm.2	-9,29269	0,03331	chr6	36410544	36458315	+	KCTD20	91			
uc031qng.1	-8,23398	0,034339	chr13	109816253	109819651	-	MYO16-AS1	92			
uc001lrw.2	-6,88915	0,036409	chr11	829297	831991	+	CRACR2B	93			
uc010mcu.3	-8,52664	0,038002	chr8	110586405	110704020	-	SYBU	94			
uc003yos.2	-7,00239	0,039637	chr8	120569317	120651106	-	ENPP2	95			
uc002fpf.2	-7,74983	0,039743	chr16	89985259	90002505	+	TUBB3	96			
uc001ovt.2	-9,99569	0,040959	chr11	74570203	74618318	-	XRRA1	97			
uc010zaz.1	-7,73792	0,048237	chr2	132250386	132273977	+	LOC150776	98			
uc010mty.3	-4,47111	0,048323	chr9	113127529	113191614	-	SVEP1	99			

3. Transcripts up-regulated in HBx-positive cells (45 genes in total, 10‡ are also up-regulated by HBx, 2§ show alternative transcripts)										
Transcript	logFC	adj.P.Val	seqnames	start	end	strand	Gene			
uc021yud.1	14,86966	1,32E-06	chr6	30552109	30552194	+	MIR877	1		
uc011kxf.2	9,615398	1,06E-05	chr8	11660190	11696818	+	FDFT1	2		
uc002gkv.3	9,086181	2,71E-05	chr17	8191969	8193409	+	RANGRF‡	3		
uc011dof.1	6,423624	2,71E-05	chr6	31712323	31730945	+	MSH5-SAPCD1	4		
uc010eoc.1	13,5376	2,71E-05	chr19	51302286	51302382	-	SNORD88B	5		
uc021uev.1	13,46995	2,71E-05	chr17	79099076	79099173	-	MIR657	6		
uc002cxk.3	8,250383	0,00013	chr16	4746511	4784378	-	ANKS3	7		
uc001kcd.4	8,086939	0,00013	chr10	82168242	82192753	+	PRXL2A‡	8		
uc010ypm.1	9,2244	0,000231	chr2	61389629	61391964	+	C2orf74	9		
uc001rwn.3	10,16545	0,000231	chr12	50794592	50855272	+	LARP4‡	10		
uc002ohe.3	5,007591	0,000246	chr19	38229203	38270230	-	ZNF573§	11		
uc001eet.1	6,982471	0,000285	chr1	114522030	114524876	+	OLFML3	12		
uc003hvk.4	9,76497	0,000319	chr4	100817407	100851778	-	DNAJB14	13		
uc009vrb.2	9,016784	0,000397	chr1	24446261	24469775	-	IL22RA1	14		
uc001wzb.4	7,666956	0,000397	chr14	51955839	52197444	+	FRMD6‡	15		
uc002axx.1	8,158774	0,000538	chr15	74738318	74753510	-	UBL7‡	16		
uc001jbn.2	11,37115	0,00072	chr10	45406764	45430642	+	TMEM72‡	17		
uc002nll.2	9,057388	0,000743	chr19	19256376	19281098	-	MEF2B	18		
uc002gci.3	8,153205	0,001015	chr17	5417438	5487832	-	NLRP1	19		
uc010tde.2	8,775456	0,001059	chr13	25338301	25454058	+	RNF17	20		
uc010fcj.2	8,67826	0,001331	chr2	61244697	61246216	+	PEX13	21		
uc010vvq.2	9,798082	0,001665	chr17	12894929	12921381	-	ELAC2	22		
uc003akq.3	9,487684	0,001702	chr22	31721790	31742249	-	PATZ1	23		
uc002nkm.3	10,65803	0,001702	chr19	19010323	19030199	-	COPE	24		
uc003dew.2	9,369123	0,002042	chr3	52579368	52713739	-	PBRM1	25		
uc003uen.1	10,62883	0,002582	chr7	75629671	75659842	-	STYXL1	26		
uc009xws.2	8,064098	0,002866	chr10	103544200	103578222	-	OGA	27		
uc001qod.3	6,11774	0,003339	chr12	6554051	6560884	+	CD27	28		
uc003qwv.2	6,842936	0,003418	chr6	169857303	170060912	-	WDR27	29		
uc003pyu.2	8,769949	0,004725	chr6	122720696	122754264	+	HSF2	30		
uc001rza.1	7,781249	0,004803	chr12	52167973	52180610	+	SCN8A‡	31		
uc001iqx.1	8,40185	0,006851	chr10	21823574	21884369	+	MLLT10‡	32		
uc001wtu.3	6,218687	0,009152	chr14	36985604	36989430	-	NKX2-1	33		
uc001kiq.4	9,545457	0,009452	chr10	95256389	95288849	+	CEP55	34		
uc010frz.3	9,143916	0,009452	chr2	190656516	190742355	+	PMS1‡	35		
uc001uva.3	7,598046	0,010056	chr13	34392206	34540695	+	RFC3‡	36		
uc031qhz.1	6,114278	0,01142	chr12	56846423	56862950	-	MIP	37		
uc002rlt.3	10,56554	0,013827	chr2	28113482	28561767	+	BABAM2	38		
uc010acc.1	10,90203	0,019275	chr13	41048131	41185264	-	FOXO1	39		
uc010mxj.4	8,972576	0,024878	chr9	130159417	130170170	+	SLC2A8	40		
uc001qiv.1	7,131021	0,025782	chr12	1023060	1058831	-	RAD52	41		
uc003ynr.4	7,558168	0,026746	chr8	110979233	110987421	-	KCNV1§	42		
uc001igf.1	8,831542	0,030006	chr10	1102776	1178237	+	WDR37	43		
uc009vya.3	5,972324	0,035653	chr1	46379265	46489944	+	MAST2	44		
uc002wjl.3	9,655381	0,041098	chr20	3776386	3786761	+	CDC25B	45		

4. Transcripts down-regulated in HBx-positive cells (33 genes in total, 4† are also up-regulated by SMC6 loss, 2§ show alternative transcripts)											
Transcript	logFC	adj.P.Val	seqnames	start	end	strand	Gene				
uc010ldr.1	-11,8904	1,32E-06	chr7	77200395	77214898	+	PTPN12	1			
uc001nlv.4	-8,88968	8,46E-06	chr11	57529234	57586652	+	CTNND1	2			
uc021oui.1	-8,51745	8,70E-06	chr1	144951761	144995082	-	PDE4DIP	3			
uc010jrc.3	-7,69411	1,23E-05	chr6	28317691	28336954	+	ZKSCAN3	4			
uc021vdu.1	-12,3919	1,92E-05	chr2	10586840	10586975	-	SNORA80B	5			
uc001sng.2	-8,88379	2,13E-05	chr12	57623356	57628718	+	SHMT2	6			
uc002tcd.4	-9,51176	2,71E-05	chr2	103353394	103434138	+	TMEM182	7			
uc001uev.2	-9,594	2,71E-05	chr12	123805246	123849388	-	SBNO1	8			
uc010pmr.2	-8,84649	6,09E-05	chr1	173684080	173755840	+	KLHL20	9			
uc001pya.3	-9,48916	0,000221	chr11	121959811	121971159	-	MIR100HG	10			
uc001ghl.3	-8,33043	0,000246	chr1	171217663	171255113	+	FMO1	11			
uc002hjt.3	-9,84142	0,000246	chr17	33914534	34053436	+	AP2B1	12			
uc002fdo.3	-10,3159	0,000465	chr16	75182421	75206132	+	ZFP1	13			
uc010tgt.1	-9,54078	0,001003	chr13	52506806	52585630	-	ATP7B†	14			
uc009ywi.3	-7,56906	0,00108	chr11	93474793	93496247	+	C11orf54	15			
uc021xwu.1	-8,45586	0,001226	chr5	19473155	19886381	-	CDH18	16			
uc009zth.2	-8,93059	0,001331	chr12	96337071	96362370	+	AMDHD1†	17			
uc010eaz.2	-10,425	0,001462	chr19	17830303	17838981	+	MAP1S	18			
uc001owo.4	-9,4244	0,001702	chr11	75145685	75201796	-	GDPD5	19			
uc002xsy.1	-10,3068	0,002034	chr20	45838381	45976670	-	ZMYND8	20			
uc022awh.1	-8,30133	0,002866	chr8	79645007	79717758	-	IL7	21			
uc002osw.2	-9,3715	0,003331	chr19	42702745	42721944	-	DEDD2	22			
uc021wyb.1	-7,52857	0,006168	chr3	49507565	49573051	+	DAG1	23			
uc002liz.4	-3,97241	0,009924	chr18	61144144	61172318	+	SERPINB5	24			
uc031qhv.1	-11,025	0,011638	chr12	56119227	56122434	-	CD63	25			
uc003xcn.1	-5,71758	0,011905	chr8	22570765	22785421	-	PEBP4†	26			
uc003euk.4	-8,50966	0,013889	chr3	141663270	141719229	-	TFDP2	27			
uc003qqz.2	-9,09632	0,016474	chr6	158480005	158520207	+	SYNJ2	28			
uc002ohd.3	-8,5849	0,017355	chr19	38229203	38270230	-	ZNF573§	29			
uc001tcy.2	-9,46345	0,019436	chr12	93965457	93970521	+	SOCS2	30			
uc010mcw.3	-7,74381	0,026038	chr8	110979233	110988076	-	KCNV1§	31			
uc001wud.3	-8,14827	0,026726	chr14	37667118	38020464	+	MIPOL1†	32			
uc002jlt.1	-4,65004	0,036189	chr17	72838168	72856007	-	GRIN2C	33			

Supplementary Table 4: Characteristics of HCC cohort							
HCC Etiology	**HCV**	**HBV**	**Alcohol**	**NASH**	***p-values***	**Test**	**Notes**
Number (%)	56 (31)	41 (23)	54 (30)	29 (16)			
Female (F), n (%)	15 (27)	5 (12)	2 (4)	2 (7)	0.0027	Chi-square	
Male (M), n (%)	41 (73)	36 (88)	52 (96)	27 (93)			
Median Age F, y (25th - 75th percentile)	60 (43 - 78)	51 (37 - 76)	61 (59 - 63)	61	0.3639	Kruskal Wallis	
Median Age M, y (25th - 75th percentile)	59 (34 - 76)	58 (34 - 76)	65 (52 - 83)	67.5 (52 - 83)	0.0002	Kruskal Wallis	
Paired PT-T samples, n (%)	45 (83)	39 (95)	41 (76)	23 (79)	0.0892	Chi-square	
Independant T samples, n (%)	9 (17)	2 (5)	13 (24)	6 (21)			
No cirrhosis, n (%)	8 (14)	17 (42.5)	17 (33)	18 (64)	<0.0001	Chi-square	
Cirrhosis, n (%)	48 (86)	23 (57.5)	35 (67)	10 36)			
Well-differentiated tumour, n (%)	23 (44)	16 (39)	24 (47)	10 (40)	0.0057	Chi-square	Differentiation (WHO), (data for 169/180 tumc
Moderately-differentiated tumour, n (%)	20 (38)	11 (27)	25 (49)	13 (52)			
Poorly-differentiated tumour, n (%)	9 (17)	14 (34)	2 (4)	2 (8)			
Median AFP, ng/ml (range)	25 (2 -1000000)	36.5 (2 – 100000(	10 (2-1000000)	6 (2.3 – 3086)	0.0004	Kruskal Wallis	
Number with high AFP, >200 ng/ml (%)	14 (27)	17 (44)	8 (17)	2 (7)	0.0033	Chi-square	based on 166/180 tumours
Number with low AFP, <200 ng/ml (%)	38 (73)	22 (56)	39 (83)	26 (93)			
Female (F) Tumor number of 1 (%)	5	13	1	0	0.0024	Chi-square	172/180 samples
F Tumor number of 2 (%)	0	0	1	2			
F Tumor number >2 (%)	0	2	0	0			
Male (M) Tumor number of 1 (%)	29	31	40	22	0.9567	Chi-square	
M Tumor number of 2 (%)	3	5	5	2			
M Tumor number >2 (%)	3	4	3	1			
Tumor number of 1 (%)	34	44	41	22	0.8552	Chi-square	(data for 175/180 tumours)
Tumor number of 2 (%)	3	5	6	4			
Tumor number >2 (%)	3	6	3	1			
Median Tumor size, mm (range)	45 (28 - 100)	30 (22 - 53)	53 (35 - 73)	45 (28 - 85)	0.0093	Kruskal Wallis	
Median overall survival, months (25th to 75th	24 (17 - 63)	44 (16 - 61)	29 (19 - 66)	31 (17 - 49)	0.7317	Kruskal Wallis	

Smc5/6 Antagonism by HBx Is an Evolutionarily Conserved Function of Hepatitis B Virus Infection in Mammals

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

VIRUS-CELL INTERACTIONS

Smc5/6 Antagonism by HBx Is an Evolutionarily Conserved Function of Hepatitis B Virus Infection in Mammals

Fabien Abdul,[b] Fabien Filleton,[a] Laetitia Gerossier,[c] Alexia Paturel,[c] Janet Hall,[c] Michel Strubin,[b] Lucie Etienne[a]

[a]CIRI—International Center for Infectiology Research, Inserm, U1111, Université Claude Bernard Lyon 1, CNRS, UMR5308, Ecole Normale Supérieure de Lyon, Université Lyon, Lyon, France

[b]Department of Microbiology and Molecular Medicine, University Medical Centre (CMU)/University of Geneva, Geneva, Switzerland

[c]CRCL-UMR Inserm 1052-CNRS 5286, Lyon, France

ABSTRACT Chronic infection with hepatitis B virus (HBV) is a major cause of liver disease and cancer in humans. HBVs (family *Hepadnaviridae*) have been associated with mammals for millions of years. Recently, the Smc5/6 complex, known for its essential housekeeping functions in genome maintenance, was identified as an antiviral restriction factor of human HBV. The virus has, however, evolved to counteract this defense mechanism by degrading the complex via its regulatory HBx protein. Whether the antiviral activity of the Smc5/6 complex against hepadnaviruses is an important and evolutionarily conserved function is unknown. In this study, we used an evolutionary and functional approach to address this question. We first performed phylogenetic and positive selection analyses of the Smc5/6 complex subunits and found that they have been conserved in primates and mammals. Yet, Smc6 showed marks of adaptive evolution, potentially reminiscent of a virus-host "arms race." We then functionally tested the HBx proteins from six divergent hepadnaviruses naturally infecting primates, rodents, and bats. We demonstrate that despite little sequence homology, these HBx proteins efficiently degraded mammalian Smc5/6 complexes, independently of the host species and of the sites under positive selection. Importantly, all HBx proteins also rescued the replication of an HBx-deficient HBV in primary human hepatocytes. These findings point to an evolutionarily conserved requirement for Smc5/6 inactivation by HBx, showing that Smc5/6 antiviral activity has been an important defense mechanism against hepadnaviruses in mammals. It will be interesting to investigate whether Smc5/6 may further be a restriction factor of other, yet-unidentified viruses that may have driven some of its adaptation.

IMPORTANCE Infection with hepatitis B virus (HBV) led to 887,000 human deaths in 2015. HBV has been coevolving with mammals for millions of years. Recently, the Smc5/6 complex, which has essential housekeeping functions, was identified as a restriction factor of human HBV antagonized by the regulatory HBx protein. Here we address whether the antiviral activity of Smc5/6 is an important evolutionarily conserved function. We found that all six subunits of Smc5/6 have been conserved in primates, with only Smc6 showing signatures of an "evolutionary arms race." Using evolution-guided functional analyses that included infections of primary human hepatocytes, we demonstrated that HBx proteins from very divergent mammalian HBVs could all efficiently antagonize Smc5/6, independently of the host species and sites under positive selection. These findings show that Smc5/6 antiviral activity against HBV is an important function in mammals. They also raise the intriguing possibility that Smc5/6 may restrict other, yet-unidentified viruses.

KEYWORDS hepatitis B virus, HBx, restriction factor, antagonism, Smc5/6 complex, virus-host interaction, evolution of virus and host genes, positive selection

Received 8 May 2018 **Accepted** 9 May 2018

Accepted manuscript posted online 30 May 2018

Citation Abdul F, Filleton F, Gerossier L, Paturel A, Hall J, Strubin M, Etienne L. 2018. Smc5/6 antagonism by HBx is an evolutionarily conserved function of hepatitis B virus infection in mammals. J Virol 92:e00769-18. https://doi.org/10.1128/JVI.00769-18.

Editor J.-H. James Ou, University of Southern California

Address correspondence to Lucie Etienne, lucie.etienne@ens-lyon.fr.

F.A. and F.F. contributed equally to this article.

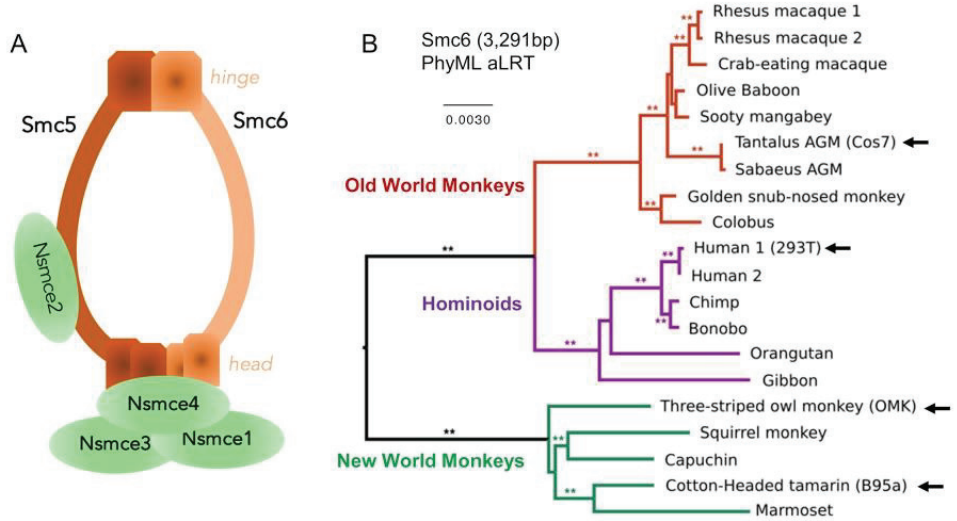
Hepatitis B virus (HBV) infects more than 250 million people worldwide and is a leading cause of chronic hepatitis and liver cancer in humans (http://www.who.int/news-room/fact-sheets/detail/hepatitis-b). HBV is a member of the *Hepadnaviridae* family of DNA viruses, which have coevolved with their host species for millions of years (1–3). Today, hepadnaviruses are found to naturally infect, in a species-specific manner, mammals, birds, fish, and amphibians. In mammals, HBVs are present in rodents, bats, and several primates, including humans, chimpanzees, gibbons, orangutans, and New World wooly monkeys. Mammalian hepadnaviruses (orthohepadnaviruses) all contain a gene encoding a small regulatory protein, the X protein or HBx, that is thought to have arisen *de novo* in the orthohepadnavirus lineage (1). HBx has long been known to play a central role in HBV replication and pathogenesis (4–6) and has recently been shown to have a key role in promoting HBV transcription by antagonizing the restriction function of the infected cell's Structural Maintenance of Chromosome (SMC) Smc5/6 complex (7, 8). However, whether this property has been conserved among the HBx-containing hepadnaviruses is unknown.

The Smc5/6 complex is, together with cohesin and condensin, one of the three SMC complexes found in eukaryotes (9, 10). As for the other SMC complexes, the core of the Smc5/6 complex is formed by a heterodimer of two SMC proteins, Smc5 and Smc6 (11), which associate with four additional subunits known as non-SMC elements (Nsmce1 to -4) (Fig. 1A). These SMC complexes all have essential housekeeping functions, playing fundamental roles in chromosome replication, segregation, and repair (reviewed in reference 10). Condensin controls chromosome condensation during mitosis, and cohesin maintains cohesion between the newly replicated sister chromatids. The role of the Smc5/6 complex is less well understood. It has reported functions in DNA replication and repair, but its exact mode of action remains elusive (12–16).

In addition to its essential cellular activities, a novel function of the human Smc5/6 complex as an HBV restriction factor has recently been uncovered: in the absence of HBx, the Smc5/6 complex binds to the HBV episomal DNA genome and inhibits viral transcription (7, 8, 17). Human HBx antagonizes this effect by hijacking the host DDB1-containing E3 ubiquitin ligase complex to target the Smc5/6 complex for ubiquitin-mediated degradation, thereby enabling productive HBV gene expression (7).

Most genes encoding antiviral restriction factors have been engaged in an "evolutionary arms race" with the viruses they inhibit (18, 19). Indeed, during long-term coevolution, pathogenic viruses and their hosts are constantly under the selective pressure of the other for survival. As a result, host restriction factors evolve rapidly and display signatures of positive (diversifying) selection. These signatures can be identified by analyzing the codon sequences of orthologous genes from a large number of related species. At virus-host interaction sites, one can witness adaptive changes, including frequent amino acid changes (where a higher nonsynonymous substitution rate [*dN*] than the synonymous rate [*dS*] is indicative of positive selection) and insertion/deletions (indels) or splicing variants as ways to modify the virus-host interface and to escape from viral antagonists (19–24).

To assess whether the antiviral function of the Smc5/6 complex has been evolutionarily important, we performed phylogenetic and evolutionary analyses of orthohepadnaviruses and host proteins in combination with functional assays. We found that all six subunits of the Smc5/6 complex have been highly conserved in primate evolution, with only Smc6 showing signatures of an evolutionary arms race. Because orthohepadnaviruses diverged millions of years ago and their HBx proteins have very little sequence homology, we then investigated the Smc5/6-antagonistic capacity of HBx proteins from six divergent orthohepadnaviruses from primates, rodents, and bats. We found that all orthohepadnavirus HBx proteins are efficient at counteracting the Smc5/6 complex, independently of the host species or amino acid variations at sites under adaptive evolution. This Smc5/6 antagonism is a strict requirement for the establishment of mammalian HBV infection, showing that the Smc5/6 complex has been an important immune defense against hepadnaviruses in mammals. Although it



C

	Smc5/6 complex						Cohesin (Smc genes)		Condensin (Smc genes)	
	Nsmce1	Nsmce2	Nsmce3	Nsmce4a	Smc5	Smc6	Smc1	Smc3	Smc2	Smc4
BUSTED (p value)	1	0.98	1	1	1	0.02*	1	1	0.81	1
PARRIS (p value)	0.48	1	1	0.81	1	0.19	1	1	1	1
PAML Codeml M7 vs M8 (p value)	1	NA	NA	NA	NA	0.48	NA	NA	NA	1
Bio++ M7 vs M8 (p value)	1	NA	NA	NA	NA	0.05*	NA	NA	NA	0.94

FIG 1 Smc6 is the least conserved subunit of the Smc5/6 complex in primates. (A) Architecture of the Smc5/6 complex. The complex is made of two core subunits (Smc5 and Smc6) and four non-SMC elements (Nsmce1 to Nsmce4). (B) Phylogenetic analysis of primate *Smc6* genes. Sequences were aligned with MUSCLE and phylogeny was performed with PhyML and an HKY+I+G model with an approximate likelihood ratio test (aLRT) as statistical support (**, aLRT > 0.8). Newly sequenced genes (arrow) are indicated. The newly sequenced *Smc6* gene from *Chlorocebus pygerythrus* (vervet African green monkey [AGM] Vero cells) is not represented because the nucleotide sequence is identical to the retrieved *Chlorocebus sabaeus* (Sabaeus AGM) sequence of *Smc6*. Alignments and phylogenies of the 10 analyzed SMC genes are available in supplemental data set 1 at https://figshare.com/articles/DatasetS1_Host_gene_alignments_used_in_the_study_fasta_format_and_phylogenetic_analyses_newick_format_Nsmce1-4_Smc1-6/6194813. (C) Positive selection analysis of the indicated genes during primate evolution. Shown are the *P* values obtained using four different methods (BUSTED, PARRIS, PAML Codeml, and Bio++; see Materials and Methods). The *P* values of the maximum-likelihood tests indicate whether the model that allows positive selection better fits the data (*, statistically significant). NA, results are not available because convergence was not obtained for these genes and/or analyses (see Materials and Methods).

would need further investigations, our findings also raise the intriguing possibility that the Smc5/6 complex may restrict other, yet-unidentified pathogenic viruses.

(This article was submitted to an online preprint archive [25].)

RESULTS

Overall evolutionary conservation of the Smc5/6 complex in primates. To trace the evolutionary history of the Smc5/6 complex, we compared the sequences of its six subunits in primates (Fig. 1). As a comparison, we also analyzed the evolutionary history of the other primate SMC genes. These include the *Smc1* and *Smc3* genes, which encode the core cohesin subunits, and the *Smc2* and *Smc4* genes, encoding the condensin core subunits. The sequences of these genes were retrieved from publicly available data sets (Table 1; see also supplemental data set 1 at https://figshare.com/articles/DatasetS1_Host_gene_alignments_used_in_the_study_fasta_format_and_phylogenetic_analyses_newick_format_Nsmce1-4_Smc1-6/6194813). To perform more

TABLE 1 Primate species and sequences that we included in the evolutionary analyses of the SMC complexes[a]

Primate group	Primate	No. of ID										
		Smc5/6 complex						Cohesin (Smc genes)			Condensin (Smc genes)	
		Nsmce1	Nsmce2	Nsmce3	Nsmce4a	Smc5	Smc6	Smc1	Smc3		Smc2	Smc4
Hominoids	Human	1	1	1	1	1	2*	1	1		1	1
	Chimp (panTro)	1	1	1	1	2	1	1	2		1	1
	Bonobo (panPan)	1	1	1		1	1	1	1		1	1
	Gorilla (gorGor)	1	1	1		1	1		2		1	1
	Orangutan (ponAbe)	1	1	1	1		1	1	1			
	Gibbon (nomLeu)	1	1	1				1			1	1
Old World monkeys	Olive baboon (papAnu)	1	1	1	1	1	1		1		1	1
	Sooty mangabey (cerAty)	1	1	1		1	1	1			1	1
	Drill (manLeu)	1	1	1		1		1	1			1
	Sabaeus AGM (chlSab)	1	1	1	1	1	1	1	1		1	1
	Tantalus AGM (chlTan)						1*					
	Rhesus macaque (macMul)	2	2	2	1	1	2	1	1		2	1
	Pig-tailed macaque (macNem)	1		1	1	1		1			1	1
	Crab-eating macaque (macFas)	1	1	1	1	1	1	1	1		1	1
	Colobus (colAng)	1	1	1	1	1	1	1	1		1	1
	Golden snub-nosed monkey (rhiRox)	1	1	1	1	1	1	1	1		1	1
	Black snub-nosed monkey (rhiBie)	1	1	1	1			1	1			1
New World monkeys	Marmoset (calJac)		1	1	1	1	1	1	1		1	1
	Cotton-headed tamarin (sagOed)						1*					
	Three-striped owl monkey (aotTri)						1*					
	Nancy Ma's owl monkey (aotNan)	1	1		1	1		1	1		1	1
	Squirrel monkey (saiBol)	1	1	1		1	1	1			1	
	Capuchin (cebCap)	1	1	1	1	1	1	1	1			1
		20 ID/19 sp	20 sp	20 ID/19 sp	14 sp	18 ID/17 sp	20 ID/18 sp	18 sp	18 ID/16 sp		17 ID/16 sp	18 sp

[a]The asterisks indicate the sequences newly obtained in this study. The total number of primate species (sp) and individuals (ID) included in each gene analysis is indicated at the bottom. The abbreviated species names are found in parentheses (i.e., three first letters of the genus followed by the three first letters of the species).

robust phylogenetic and selection analyses, we obtained additional primate species sequences using reverse transcription-PCR (RT-PCR) approaches (Table 1 and Fig. 1B; see also Materials and Methods). Overall, we included up to 20 simian primate species in our positive selection analyses to span 40 million years of evolution (26, 27). We found that the synteny of the genes was conserved during primate evolution, although some subunits had duplicated pseudogenes in a few primate species (see supplemental Fig. 1 at https://figshare.com/articles/Figure_S1_Synteny_conservation_of_Smc5_6_complex_genes_during_primate_evolution_/6194867). Among the core SMC proteins, the most conserved are the cohesin Smc1 and Smc3 subunits, which share 100% and 99.9% pairwise identity at the amino acid level in a simian primate alignment, respectively (see supplemental data set 1 at https://figshare.com/articles/DatasetS1_Host_gene_alignments_used_in_the_study_fasta_format_and_phylogenetic_analyses_newick_format_Nsmce1-4_Smc1-6/6194813). Smc6 was the least conserved SMC protein, with 97.4% pairwise identity in simian primates. Using the Genetic Algorithm for Recombination Detection (GARD) on the complete set of genes (28), evidence of recombination (GARD, $P < 0.05$) was found only for *Nsmce3*, and therefore subsequent phylogenetic and selection analyses were performed on both the whole *Nsmce3* gene and the two identified *Nsmce3* gene fragments (from bp 1 to 246 and bp 247 to 912). Phylogenetic analysis of the 10 genes showed that the gene trees derived from nucleotide alignments were largely in accordance with the accepted species tree from Perelman and colleagues (27) (Fig. 1B; see also supplemental data set 1 at https://figshare.com/articles/DatasetS1_Host_gene_alignments_used_in_the_study_fasta_format_and_phylogenetic_analyses_newick_format_Nsmce1-4_Smc1-6/6194813).

To assess whether the Smc5/6 complex and the *Smc1 to Smc4* genes have experienced diversifying selection during primate evolution, we performed four types of positive selection analyses. First, we used the BUSTED method, which tests whether a gene has experienced positive selection on at least one site or one branch during evolution (29). Of the six Smc5/6 complex subunits and the four *SMC* genes from the cohesin and condensin complexes, only one, *Smc6*, showed gene-wide evidence of episodic positive selection (BUSTED, $P < 0.05$ [Fig. 1C]). These findings were confirmed using the PARRIS method, which also detects evidence of positive selection using a codon alignment (30), although the level of significance was not reached for the *Smc6* gene ($P = 0.19$ [Fig. 1C]). Third, we ran the Codeml program from the PAML package (31) on our codon alignments to compare two models: one that allows positive selection at certain sites (M8, alternative hypothesis) and one that disallows positive selection (M7, null hypothesis). We then performed a likelihood ratio test (LRT) to examine which of the two models better fits our data. Overall, there was no convergence for 7 of the 10 genes because they were too conserved (see Materials and Methods for details), and no evidence of significant positive selection was found for the remaining genes (Fig. 1C). Finally, we used the Bio++ package from Guéguen et al. (32), which has two main advantages over PAML, to similarly test for evidence of positive selection (M7 versus M8 as implemented in Bio++). First, the DNA substitution models use nonstationary matrices, which allow nucleotide composition to change over time and therefore improve the fitting to real data (33). Second, the codon frequency better fits biological assumptions than is the case with PAML (32). Using Bio++, we obtained higher values of likelihood, and for *Smc6*, a model allowing positive selection was favored over a model that disallows positive selection (Bio++ M7 versus M8, $P = 0.05$ [Fig. 1C]). Overall, these studies show that the *SMC* genes for cohesin and condensin, as well as the six Smc5/6 complex subunits, have been highly conserved during primate evolution. This is similar to what has been described for the global evolution of SMC proteins in eukaryotes (9) and is consistent with their essential cellular housekeeping functions. However, it is in contrast to most other known antiviral restriction factors that have strongly evolved under positive selection in primates (18, 21). The only exception is the *Smc6* gene, for which two positive selection analyses found evidence of adaptive evolution.

A

Method		Codons under positive selection						
MEME	(p<0.1)		279	426	489	628		930
FUBAR	(PP>0.8 for β>α)	173		426			697	930
Bio++ M8	(PP>0.9)						697	930
PAML M8	(PP>0.9)							930
Intersect	Total	173	279	426*	489	628	697*	930*

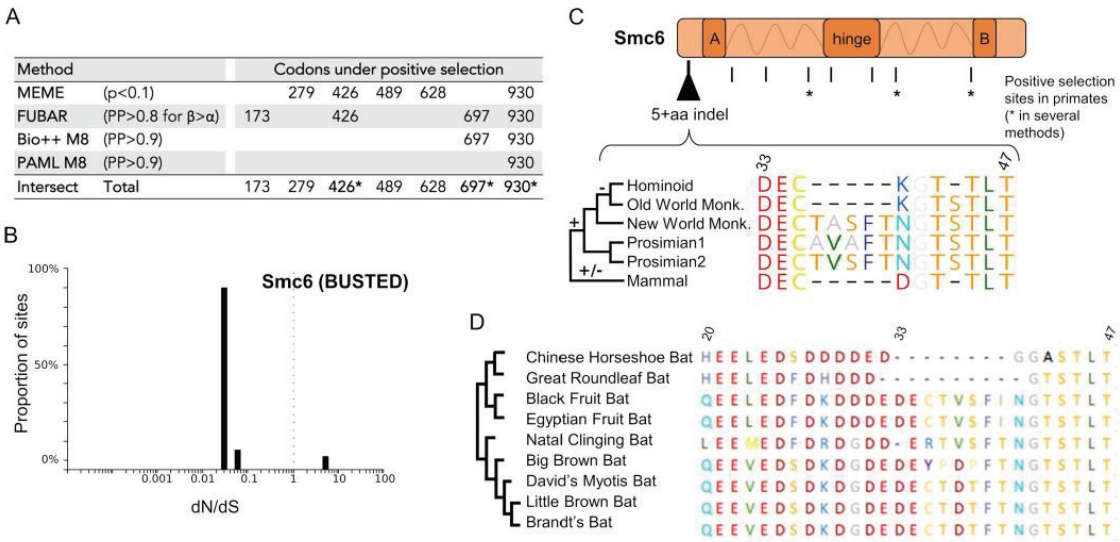


FIG 2 Evidence of episodic site-specific positive selection in Smc6 during primate evolution, as well as genetic plasticity in other mammals. (A) Specific sites in Smc6 are under positive selection. Codon alignments were analyzed using four different positive selection tests: MEME, which detects site-specific episodic positive selection; FUBAR, similar in a Bayesian framework; and Bio++ and PAML Codeml (M8), which detect site-specific positive selection (see Materials and Methods). The table shows the codon sites showing significant positive selection (i.e., that passed the widely accepted *P* value or posterior probability [PP] cutoffs for each method). The statistical thresholds used in each test are shown in the table. Codons identified as being under positive selection in at least two of the four tests are indicated by an asterisk. (B) Graphic depicting the proportion of sites in Smc6 at a given *dN*/*dS* ratio, calculated with BUSTED. A very low number of Smc6 sites are under positive selection. A *dN*/*dS* ratio of <1 indicates negative selection, a *dN*/*dS* ratio of 1 indicates neutrality, and a *dN*/*dS* ratio of >1 indicates positive selection. (C) Marks of genetic conflicts in Smc6. Sites under positive selection in primates as well as the plasticity of the N-terminal region of Smc6 in mammals are shown. Amino acid alignment was performed with MUSCLE, and residue color coding is from RasMol. A and B correspond to the globular domain that contain a Walker A and Walker B motif, respectively. Dashes indicate gaps. One nonprimate mammal species was arbitrarily chosen for the schematic representation. The mammal sequence illustrates only one possibility of the natural interspecies sequence variations that have been important in this region (see supplemental data set 3 at https://figshare.com/articles/Dataset_S3_Phylogenetic_analyses_of_Smc5_6_in_mammals_fasta_and_newick_format_/6194840 and supplemental Table 2 at https://figshare.com/articles/Table_S2_Species_used_for_the_phylogeny_of_the_mammalian_Smc6_and_for_the_experiments_/6194846). (D) Plasticity of the N-terminal region of Smc6 in bats. Truncated amino acid alignment (region from aa 20 to 47) of Smc6 sequences from bats. On the left is a cladogram of the bat Smc6. The amino acid alignment was performed with MUSCLE, and residue color coding is from RasMol (in Geneious [Biomatters]). Dashes indicate gaps.

Evidence of site-specific adaptive evolution in primate Smc6. It is formally possible that most of the Smc6 protein has been highly conserved due to its housekeeping function, while just a few sites have been engaged in a virus-host interaction. Indeed, previous studies have shown that essential cellular proteins that are usurped by viruses for their replication have evolved under a strong purifying selection background, with only the virus-host interaction sites showing rapid evolution (34–38). To determine if this was the case for *Smc6*, we characterized in more detail its evolutionary history during primate evolution. Using the BranchSite-REL algorithm from HYPHY (39), we found that positive selection has occurred on a proportion of branches of the primate *Smc6* phylogeny (episodic positive selection; BS-REL, $P < 0.05$). To look more specifically for site-specific positive selection, we used MEME and FUBAR, which are among the most accepted methods from HYPHY, as well as the posterior probabilities at codon sites in PAML and Bio++ (M8 model) (Fig. 2A). We confirmed that most codons in Smc6 have been extremely conserved, with over 90% having a ratio of nonsynonymous to synonymous nucleotide (*dN*/*dS*) substitutions lower than 1 (Fig. 2B). However, a few sites were identified as having evolved under significant positive selection by one or several methods, consistent with site-specific positive selection in the primate *Smc6* gene (Fig. 2).

In addition to single codon substitutions, other forms of genetic changes may also be adaptive in a virus-host arms race (21, 36, 40). In particular, recombination, gene deletion or duplication, and insertions and deletions (indels) can also be advantageous

for the host. These "genetic innovations" would be missed in typical methods screening for positive selection. We found some indels in genes encoding several Smc5/6 complex subunits (see supplemental data set 1 at https://figshare.com/articles/DatasetS1_Host_gene_alignments_used_in_the_study_fasta_format_and_phylogenetic_analyses_newick_format_Nsmce1-4_Smc1-6/6194813). In particular, we found a 5-amino-acid (aa) indel in the N-terminal region of Smc6 (Fig. 2C). The five New World monkey species, including the two for which sequences were determined in this study, display a TASFT motif at this position. However, this 5-aa motif was not present in any of the retrieved *Catarrhini* species sequences ($n = 13$) (Fig. 2C, Hominoid and Old World Monk.).

To decipher if this 5-aa indel was specific to simian primates, we extended our analysis to mammals. We found a significant plasticity within this region with both indels and amino acid changes during mammalian evolution. For example, the prosimians' proteins contain a 5-aa motif, but with amino acid differences in the TASFT motif (TVSFT and AVAFT, respectively [Fig. 2C]). Another remarkable example of this N-terminal plasticity was found in bats (*Chiroptera*). Most bat species contain a 5-aa stretch (residues 36 to 40) but the sequence differs significantly (TVSFI, TVSFT, PDPFT, and TDTFT), while some bats carry an 8- to 10-amino-acid deletion within this region (Chinese horseshoe bat [*Rhinolofus sinicus*] and great roundleaf bat [*Hipposideros armiger*]) (Fig. 2C and D). Therefore, although most of the Smc6 protein sequence has been very conserved in primates and more generally in mammals, a few sites have been under significant positive selection and show substantial genetic plasticity. These signatures of genetic changes in an essential protein could be reminiscent of an evolutionary arms race with pathogenic viruses.

HBx and the woodchuck WHx counterpart promote degradation of diverse mammalian Smc5/6 complexes, independently of variations at genetic innovation sites. We next used evolution-guided functional analyses to examine whether the marks of adaptive evolution and the identified interspecies variability of the Smc5/6 complex have functional consequences for the ability of HBx to promote its degradation. To test this, we put together a panel of mammalian cells encoding divergent Smc5/6 complexes and, importantly, with Smc6 orthologues with variations at positive selection sites and indels (Fig. 3A). This panel included cells derived from various primate species, including humans and the tantalus African green monkey (*Chlorocebus tantalus*), vervet African green monkey (*Chlorocebus pygerythrus*), and the New World owl monkey (*Aotus trivirgatus*), as well as ferret (*Mustela putorius furo*, carnivore) and mouse (*Mus musculus*, rodent) cells. Using this panel, as well as different human cell lines, we tested the capacity of the human HBx and the woodchuck WHx counterpart to promote degradation of the heterologous Smc5/6 complex in these cells. Consistent with previous studies (7, 8), we found that HBx transduced in various human cell lines, including HepG2 (hepatocyte carcinoma), 293T (kidney epithelial), and HeLa (epithelial adenocarcinoma), triggers similar decreases in endogenous Smc6 and Nsmce4A protein levels (Fig. 3B). Similar results were obtained with the U2OS (human osteosarcoma) and A549 (human lung adenocarcinoma) cell lines (data not shown). This indicates that this HBx activity is neither cell type specific nor affected by human polymorphism at position 697 (NCBI dbSNP reference rs1065381).

Then cells from the six different mammalian species were transduced with a lentivector encoding green fluorescent protein (GFP) alone, GFP-HBx, or GFP-WHx. Five days later, we found that HBx and WHx expression had specifically triggered the degradation of the Smc5/6 complex in all species tested (Fig. 3C). This was true even for host species like the owl monkey, which harbors a 5-aa insertion in the N-terminal region of Smc6 (Fig. 2C and 3) and for Old World monkeys, which, in contrast to New World monkeys and hominoids, are not natural hosts for hepadnaviruses (Fig. 3C). Taken together, these results show that the antagonism of endogenous Smc5/6 by the viral HBx and WHx is independent of the cell type and of the mammalian host species. Overall, amino acid differences at sites under adaptive evolution in Smc6 did not significantly impact HBx-mediated degradation of the complex (Fig. 3), suggesting that HBV has not driven Smc6 adaptation in primates.

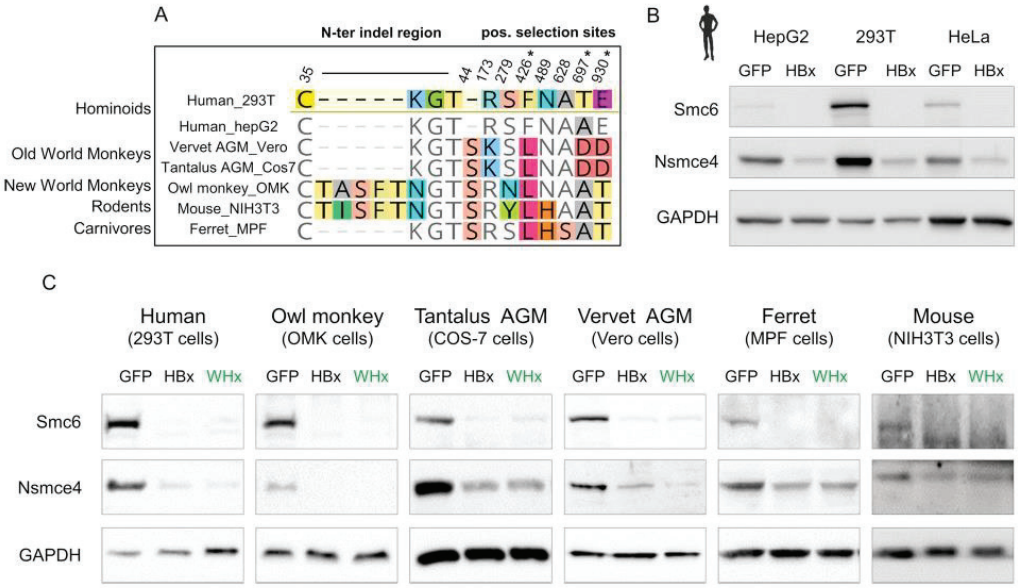


FIG 3 HBx and WHx can degrade the Smc5/6 complex in cells from diverse mammalian species. (A) Amino acid differences at the sites of genetic conflict in Smc6 between the host species tested for panels B and C. Note that all statistically significant marks of a potential evolutionary arms race identified in Fig. 2 are represented. Asterisks indicate the sites that were found under positive selection by at least two methods. (B) HBx from human HBV degrades the human Smc5/6 complex, independently of the cell type and the human polymorphism at position 697 (NCBI dbSNP reference rs1065381). Protein expression of endogenous Smc6 and Nsmce4A (two essential subunits of the Smc5/6 complex) from three human cell lines (HepG2, 293T, and HeLa cells) that were previously transduced with a lentivector expressing GFP only (GFP) or the HBx protein from human HBV fused to GFP (HBx) is shown. GAPDH serves as a loading control. (C) The human (HBx) and woodchuck (WHx) HBV X proteins promote degradation of the Smc5/6 complex in primate, rodent, and carnivore cells ($n = 6$ species). Cells were transduced with a lentivector encoding GFP alone, GFP-HBx, or GFP-WHx. The mouse Smc6 could only be detected using a different anti-Smc6 antibody (see Materials and Methods; the NIH 3T3 blots are from two SDS-PAGE loaded with the same cell lysates).

Mammalian hepadnavirus HBx proteins show a conserved ability to counteract the restriction activity of the human Smc5/6 complex. It is unknown whether the capacity to counteract the Smc5/6 complex is an important and conserved function of mammalian hepadnaviruses that diverged millions of years ago. To span the entire orthohepadnavirus evolutionary history, we examined, in addition to the human HBx and woodchuck WHx, the newly cloned HBx proteins from a hepatitis B virus that infects the New World wooly monkey (WMHBx) and from three viruses infecting distantly related bat species: *Hipposideros* cf. *ruber* (roundleaf bat), *Rhinolophus alcyone* (horseshoe bat), and *Uroderma bilobatum* (tent-making bat) (RBHBx, HBHBx, and TBHBx, respectively) (Fig. 4A). As shown in Fig. 4B, these HBx proteins have highly divergent amino acid sequences, with some regions sharing essentially no homology. Despite this low sequence identity and long-term divergence, all HBx proteins showed comparable abilities to trigger degradation of the human Smc6 and Nsmce4A proteins, when transduced in either HepG2 or 293T cells (Fig. 5A and B).

Previous work has shown that the Smc5/6 complex binds episomal DNA templates to block transcription and that inactivation of the complex leads to an increased episomal gene expression (7). Accordingly, a similar increase in expression of a transiently transfected episomal luciferase reporter construct was observed in all orthohepadnavirus HBx protein-expressing cells (Fig. 5C). As an additional control, we constructed six vectors each expressing one of the six proteins in the absence of the fused GFP and found that the X proteins in their native form could all increase the expression of the luciferase reporter (data not shown).

Because the human viral HBx protein degrades the Smc5/6 complex by recruiting DDB1, we tested whether all X proteins were capable of interacting with human

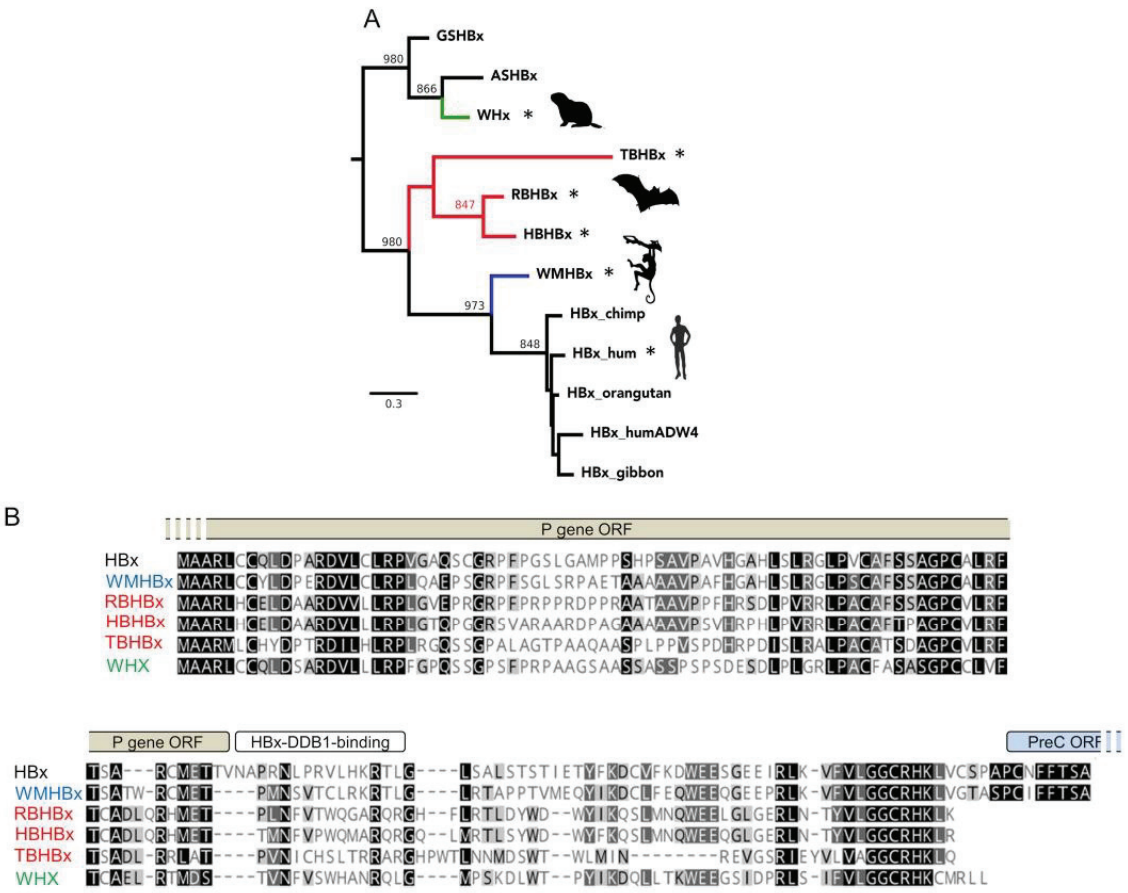


FIG 4 Evolutionary analyses of divergent mammalian HBV X proteins. (A) Phylogenetic analysis of the X proteins from hepadnaviruses that naturally infect mammals. The viral X proteins tested in our *in vitro* functional assays (Fig. 5 to 7) are indicated by an asterisk. Phylogenetic analysis of orthohepadnaviral X proteins was performed using a 161-amino-acid alignment obtained with MUSCLE (see supplemental data set 2 at https://figshare.com/articles/DatasetS2_Orthohepadnaviral_HBx_amino_acid_alignment_interleaved_phylip_format_/6194825) and the tree was built with PhyML and a JTT+I+G model with 1,000 bootstrap replicates. Bootstrap values (>800/1,000) are indicated at the nodes. The tree was rooted for representation purposes according to the work of Drexler et al. (52) (but the outgroup of orthohepadnavirus is still under debate [2]). The scale bar indicates the number of amino acid substitutions per site. We analyzed the X proteins from HBVs from the ground squirrel (GSHBV), arctic squirrel (ASHBV), and woodchuck (WHV), three bat viruses naturally infecting *Hipposideros* cf. *ruber* (roundleaf bat), *Rhinolophus alcyone* (horseshoe bat), and *Uroderma bilobatum* (tent-making bat), respectively (RBHBV, HBHBV, and TBHBV), wooly monkey HBV (WMHBV), human HBV, and HBVs from other indicated hominoids. (B) Amino acid alignment of the viral X proteins used for Fig. 5 to 7. The black-to-white gradient depicts high-to-low sequence identity (Geneious). The open reading frames (ORFs) overlapping with HBx are shown, as well as the DDB1-binding region in the human viral HBx protein (72).

DDB1. Using coimmunoprecipitation (co-IP) experiments, we found that despite significant primary amino acid differences (Fig. 4B), the six mammalian hepadnaviral X proteins interact with human DDB1 protein (Fig. 5D). This suggests that the hepadnaviral X proteins antagonize Smc5/6 via a conserved molecular mechanism.

We finally tested whether this conserved capacity to degrade the Smc5/6 complex was restricted to only the human complex and found that the X proteins from human, woodchuck, monkey, and bat hepadnaviruses could also efficiently degrade the Smc5/6 complex in mouse and owl monkey cells (Fig. 6).

Thus, the ability to degrade and to counteract the restriction activity of the Smc5/6 complex is conserved among mammalian hepadnavirus X proteins and in different species.

Divergent mammalian HBx proteins efficiently rescue replication of a human HBx-deficient hepatitis B virus in primary human hepatocytes. We finally tested

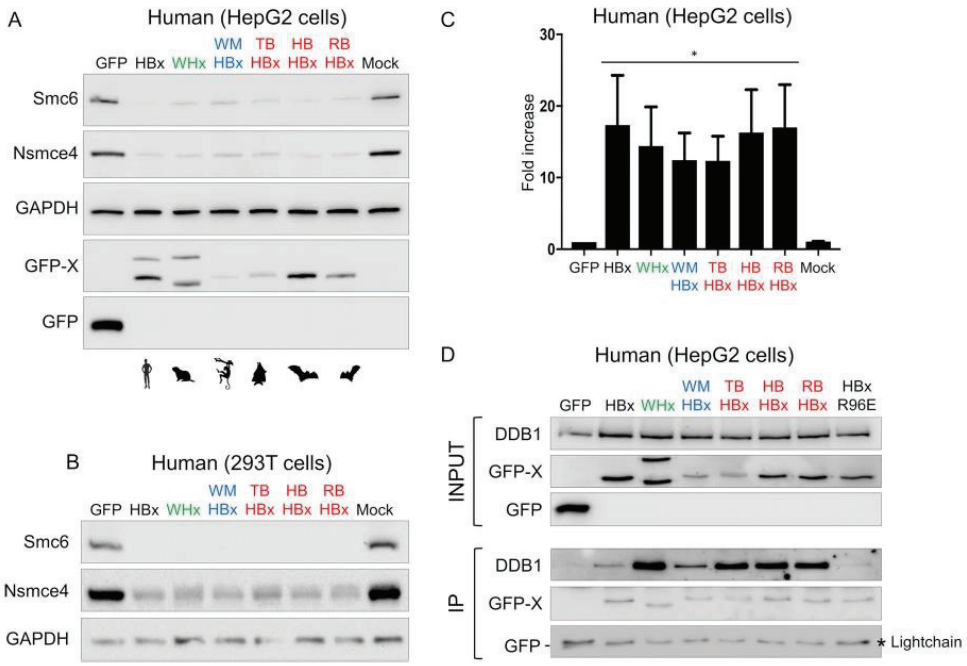


FIG 5 Highly divergent mammalian HBV X proteins show a conserved property of recruiting human DDB1 and antagonizing human Smc5/6 restriction. (A and B) Degradation of the human Smc5/6 complex by mammalian hepadnavirus X proteins. Human hepatoma HepG2 cells (A) and 293T cells (B) were transduced with a lentivector expressing only GFP (control) or the GFP-fused X protein from diverse hepadnaviruses (Fig. 4) or a mock control. Western blot analysis of the endogenous Smc6 and Nsmce4A was performed (see Materials and Methods). GAPDH served as a loading control. (C) Effect of mammalian X proteins on transiently transfected reporter gene activity. HepG2 cells were transfected with a luciferase reporter construct and the next day transduced with lentiviral vectors expressing the indicated proteins as described above. At days 5 to 7, the luciferase activity was measured; the fold increase of relative light units (RLU) versus the GFP control condition (set at 1) is shown. The means from three independent experiments are shown, along with SDs. *, *P* value = 0.1. *P* values correspond to the Wilcoxon Mann-Whitney test against the null hypothesis of no difference in the luciferase activity between the GFP control and GFP-X conditions. Of note, the same six X proteins unfused to GFP (i.e., in their native forms) also retained this activity (data not shown). (D) Interaction with human DDB1 protein was conserved for all hepadnaviral X proteins tested. The presence of DDB1 and GFP-fused protein (IP) was assessed by Western blotting. The viral X proteins could all interact with human DDB1, except for the DDB1 binding-deficient HBx mutant (R96E) that was used as a control. Note that GFP migrates to a position near the immunoglobulin light chain.

whether the HBx proteins from nonhuman orthohepadnaviruses would substitute for human HBx in an HBV replication assay. Primary human hepatocytes (PHHs) were transduced with lentiviruses expressing GFP alone or fused to one of the six orthohepadnavirus HBx proteins and 4 days later infected with the wild-type HBV or an HBx-deficient HBV mutant (HBVΔX) (Fig. 7A). As shown previously, human HBx provided in *trans* fully rescued the replication defect of HBVΔX, as measured by HBe and HBs

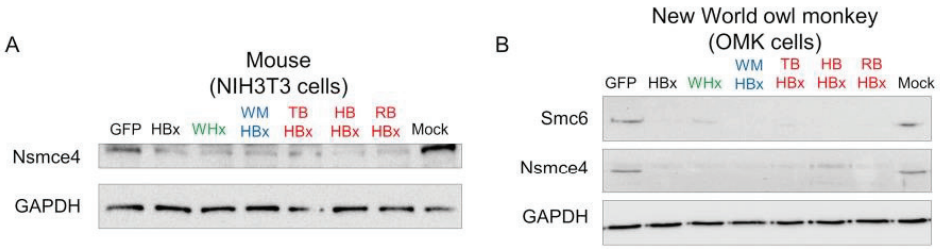


FIG 6 Conserved capacity of hepadnavirus X proteins to degrade the Smc5/6 complex in mammalian cells from mouse and New World monkey. The same experiments as for Fig. 5 were performed with mouse NIH 3T3 cells (A) and OMK owl monkey cells (B).

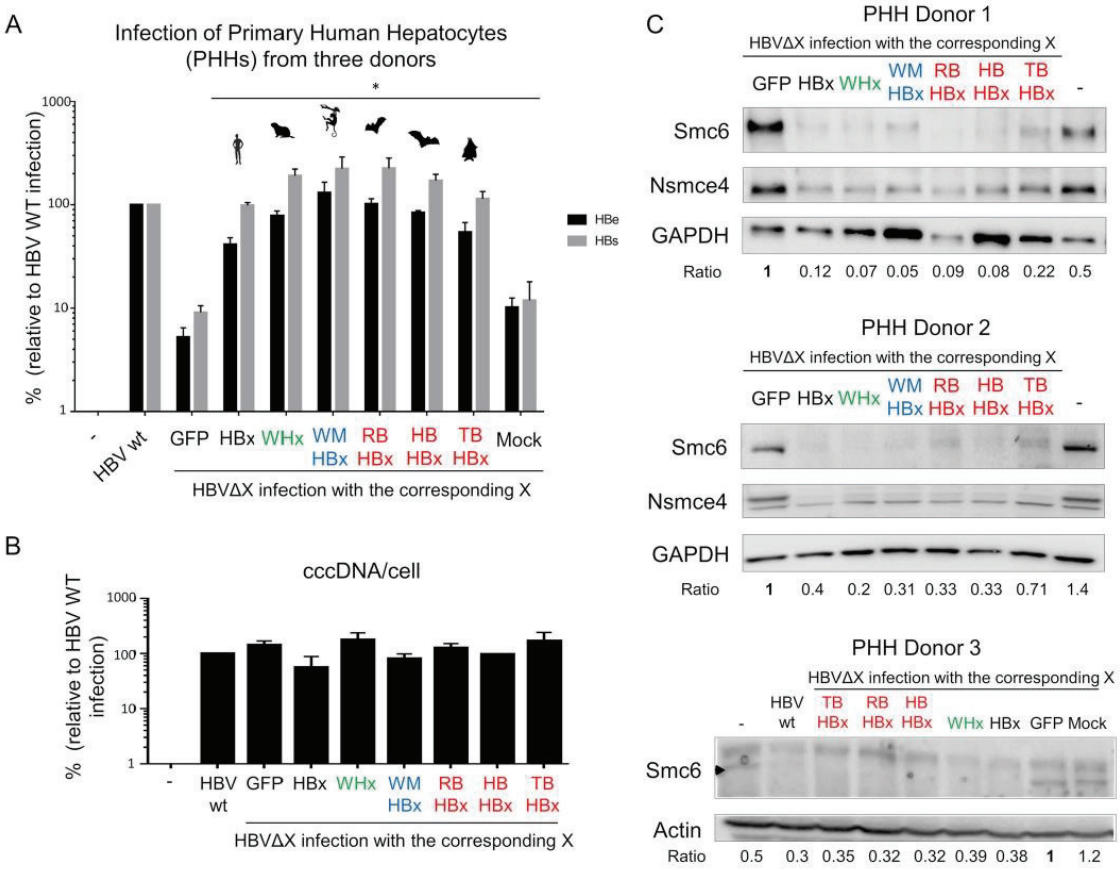


FIG 7 The X proteins from six orthohepadnaviruses can all fully rescue the replication defect of an HBx-deficient HBV in primary human hepatocytes (PHHs). (A) PHHs were mock transduced or transduced with GFP or the indicated X proteins and infected with wild-type HBV or an HBx-deficient HBV (HBVΔX). HBe and HBs antigen secretion was quantified 7 days later by ELISA. Antigen concentrations are relative to wild-type HBV, which was set to 100. Data are means ± SEMs from independent experiments performed with three different PHH donors. *, P value = 0.1. P values correspond to the Wilcoxon Mann-Whitney test against the null hypothesis of no difference in the PHH infection between HBVΔX complemented with GFP alone and the GFP-X proteins. (B) The HBV cccDNA levels were measured at day 7 postinfection by real-time PCR. Values are expressed relative to beta-globin mRNA levels to normalize to cell number. The results are the means ± SEMs for the levels seen in PHHs from two donors. (C) Smc6 degradation in PHHs expressing the X proteins from different hepadnaviral lineages. Protein extracts were prepared from the cells listed above (three different PHH donors), and Smc6 protein levels were analyzed by Western blotting. Actin or GAPDH served as a loading control. "Ratio" shows the relative protein expression level of Smc6 over the actin or GAPDH controls, normalized to the GFP condition (GFP, 1).

antigen secretion (7). Strikingly, the HBx proteins from the wooly monkey (WMHBx), woodchuck (WHx), and three bat (TBHBx, HBHBx, and RBHBx) viruses were all capable of restoring HBVΔX replication to levels comparable to those of wild-type HBV (Fig. 7A). This occurred in the absence of changes in viral covalently closed circular DNA (cccDNA) levels (Fig. 7B) and was accompanied by a concomitant decrease in Smc6 protein levels (Fig. 7C). These results provide functional evidence that orthohepadnavirus HBx proteins have a conserved capacity to antagonize the antiviral function of Smc5/6 and that this occurs in a species-independent fashion.

DISCUSSION

In this study, we found that the HBx proteins from mammalian hepadnaviruses, which diverged millions of years ago and have very little sequence homology, all have the capacity to degrade the Smc5/6 complex and to counteract its antiviral activity in a species-independent manner. The antiviral function of the Smc5/6 complex against

hepadnaviruses has therefore been an important immune defense mechanism in mammals.

We traced the evolutionary history of the genes encoding the six components of the Smc5/6 complex, as well as the genes for the Smc1 to -4 core subunits of cohesin and condensin. We show that Smc1 to -4 and the Smc5/6 complex have been highly conserved in primates. The only exception is the *Smc6* gene, which shows some signatures of adaptive evolution in primates and mammals. These include several sites under positive selection and insertion/deletion events during mammalian evolution, which could be reminiscent of a virus-host evolutionary arms race. To examine if HBV, which is currently the only virus reported to be restricted by the Smc5/6 complex, has contributed to shape the evolution of the complex, we tested several orthohepadnavirus HBx proteins for the ability to antagonize the Smc5/6 complex across species. We found that all the HBx proteins tested, including those encoded by distantly related bat HBVs, were equally efficient in antagonizing Smc5/6 in human and other mammalian cells. This conserved property suggests that the interaction between the Smc5/6 complex and HBx is independent of the sites under adaptive evolution that we identified in this study. Thus, HBx does not seem to have driven the evolution of Smc5/6 in primates and, more broadly, in mammals.

Overall, we did not identify strong signatures of positive selection in the Smc5/6 complex as typically found in other known restriction factors (18, 21, 40, 41). This is surprising given that the antagonistic relationship between HBx and Smc5/6 has been conserved in orthohepadnaviruses (this study) and has likely been played out over tens of millions of years (1). There are at least three possible explanations for this apparent inconsistency. First, Smc5/6 has an evolutionarily highly conserved architecture, and it performs fundamental functions in cellular genome maintenance and thus has likely evolved under strong purifying selection. This is in contrast to most other described restriction factors, such as APOBEC3G and TRIM5, that are dedicated to their antiviral intrinsic immune functions (41, 42). However, our findings show similarities both at the evolutionary and functional levels with the serine incorporator proteins SERINC3 and SERINC5. These cellular proteins were recently found to act as restriction factors against lentiviruses (43–46). Despite their antiviral function and their antagonism by the lentiviral Nef protein, SERINC3 and SERINC5 also show little evidence for positive selection (45). Like Smc5/6, the SERINC proteins have important cellular functions and are not part of the interferon signaling pathway. The Smc5/6 complex and the SERINC proteins may therefore fall into a separate category of restriction factors that have dual antiviral and essential cellular functions and therefore have limited evolutionary opportunities. In an extreme scenario, host proteins might not be able to "escape" from the viral antagonist, and it is likely that viruses have adapted to target such constrained host proteins (i.e., "viral strategy") (37, 47). The thorough evolutionary analysis of such constrained restriction factors would benefit from novel bioinformatic methodologies specifically designed to identify adaptive evolution that operates on a background of strong purifying selection and that may further involve genetic innovations other than site-specific positive selection (e.g., indels and recombination).

A second possible explanation for the lack of strong positive selection in the Smc5/6 complex is that orthohepadnaviruses may not have been strong drivers of mammalian genome adaptation. Indeed, acute HBV infection is mostly asymptomatic, and when evolving into chronic infection the symptoms and associated morbidity appear late in life, after the reproductive age. Arguing against this interpretation are the findings of Enard and colleagues, who provided evidence that most cellular genes reported in the literature to encode HBV-interacting proteins show a strong excess of adaptation (37). This, however, remains to be functionally demonstrated because, to our knowledge, no study to date has reported evidence for a direct evolutionary arms race between an HBV protein and a cellular factor.

A third possibility is that HBx interacts with the Smc5/6 complex indirectly, through a yet-unknown intermediate cellular protein. Although we cannot formally exclude this

possibility, none of our studies so far suggests that this is the case. In particular, no common cellular protein in addition to the DDB1 subunit of the E3 ligase and the six Smc5/6 subunits was recovered from HBx-expressing cells in pulldown experiments with either HBx or Smc5/6, as would be expected for an adaptor protein bridging HBx to the complex (F. Abdul and M. Strubin, unpublished data). We can further exclude any evolutionary arms race between DDB1 and HBV, because DDB1 is under strong purifying selection (data not shown) and our assays with heterologous species cells allow us to robustly exclude any species specificity between HBx and a hypothetical endogenous intermediate factor. Nevertheless, the exact HBx-Smc5/6 interface remains unknown. Evolutionary and mutagenesis analyses of HBx did not allow us to identify the viral determinants of the interaction with Smc5/6 (data not shown). The analysis of structural/docking models (see references 48 and 49 for examples) might contribute to solve this virus-host interface.

It is really remarkable that all the orthohepadnavirus HBx proteins that we tested antagonize the Smc5/6 complex and share the ability to fully substitute for human HBx in an HBV infection assay using primary human hepatocytes, especially given their low sequence identity (<38% identity for some pairs) and their long divergence time. These findings suggest that the Smc5/6 antiviral restriction activity is conserved and has been essential among mammals, and they provide additional evidence that antagonizing the Smc5/6 complex is a major function of HBx during HBV infection. Our findings further imply that in contrast to what has been documented for other restriction factors (53), Smc5/6 does not act as a species barrier to the potential zoonotic transmission of bat or primate HBVs to humans (50–52, 54).

HBVs are not restricted to mammals, as they are also found in birds, fishes, and amphibians (2, 55, 56). Because the latter viruses appear to lack an *HBx* gene (1, 2, 55), it would be interesting to explore if they are also restricted by Smc5/6 and, if so, what strategy these divergent hepadnaviruses use to circumvent the antiviral restriction.

Finally, because mammalian *Smc6* has some evidence of adaptive evolution independently of HBV pressure, it raises the possibility that the *Smc6* gene has been engaged in an evolutionary arms race with other pathogenic viruses. For example, the evolutionary adaptive changes identified in the lentiviral restriction factor MxB have been driven not by lentiviruses but likely by other pathogens (45, 57). In addition, restriction factors with a broad antiviral spectrum, such as MxA or PKR, have been evolutionarily driven by several pathogens (58, 59). It will be interesting to determine whether other DNA viruses are restricted by the Smc5/6 complex and, if so, whether they have contributed to shape the evolution of mammalian Smc5/6 complex.

MATERIALS AND METHODS

Cell lines and culture. Cells were maintained in Dulbecco modified Eagle medium (DMEM) supplemented with 10% fetal calf serum. Human cell lines used in the study were human embryonic kidney 293T cells and HeLa cells (gifts from Andrea Cimarelli, CIRI Lyon), as well as the human hepatoma cell line HepG2 (ATCC HB-8065). Old World monkey cells used were tantalus African green monkey (AGM) (*Chlorocebus tantalus*) COS-7 cells (a gift from Branka Horvat, CIRI Lyon) and vervet AGM (*Chlorocebus pygerythrus*) Vero cells (a gift from Andrea Cimarelli, CIRI Lyon). New World monkey cells were owl monkey (*Aotus trivirgatus*) kidney OMK cells (CelluloNet Lyon) and cotton-headed tamarin (*Saguinus oedipus*) B95a cells (a gift from Branka Horvat). We also used ferret (*Mustela putorius furo*, from Branka Horvat) and mouse cells (NIH 3T3, a gift from Theophile Olmann's lab, CIRI Lyon). The species identity of all the cell lines used in this study was confirmed by amplification and sequencing of cytochrome *b* and/or beta-actin (data not shown).

Expression plasmids. The lentivirus vectors pWPT expressing GFP, GFP-HBx, and GFP-WHx have been previously described (60–62). The X coding regions (synthesized by Genewiz) from hepadnaviruses infecting the New World wooly monkey (*Lagothrix*) (WMHBx) and three distant bat species, including the roundleaf bat (*Hipposideros* cf. *ruber*), the horseshoe bat (*Rhinolophus Alcyone*), and the tent-making bat (*Uroderma bilobatum*) (RBHBx, HBHBx, TBHBx, respectively) (52), were expressed from the same vector. The X insert was ligated to the pWPT-GFP backbone between the PstI and NotI sites following the T4 DNA (New England BioLabs [NEB]; M0202) ligation protocol. The plasmid was then transformed following the high-efficiency transformation protocol using NEB 10-beta Competent *Escherichia coli* (NEB; C3019). All the X-expressing plasmids were checked by Sanger sequencing using eGFP-F primer, 5′-CAT GGT CCT GCT GGA GTT CGT G-3′, and pWPT-R, 5′-GTC AGC AAA CAC AGT GCA CAC CA-3′. The pWPT-ΔGFP-X plasmids encoding the native form of the X proteins were obtained after digestion of the pWPT-GFP-X vectors using MluI and NotI and amplification of each X using Mlu1-pWPT-X-F primer (5′-GCT TAC GCG

TTC TGC AGT CGA CGA ATT CAC CAT G-3′) and pWPT-R primer (5′-GTC AGC AAA CAC AGT GCA CAC CA-3′). Each X insert was ligated to the pWPT-ΔGFP linearized backbone following the T4 DNA (NEB-M0202) ligation protocol.

Transfection and transduction. For recombinant-lentivirus production, plasmids were transfected in 293T cells by the calcium phosphate method. Briefly, 4.5×10^6 cells were plated in a 10-cm dish and transfected 12 h later with 15 μg of the X protein-expressing lentiviral vectors, 10 μg of packaging plasmid (psPAX2, gift from Didier Trono [Addgene plasmid 12260]), and 5 μg of envelope (pMD2G, gift from Didier Trono [Addgene plasmid 12259]). The medium was changed 6 h posttransfection. After 48 h, the viral supernatants were collected and stored at −80°C. The viral supernatants were titrated by transducing 293T or HeLa cells and measuring the GFP expression by fluorescence-activated cell sorting (FACS) 5 days later. The different mammalian cell lines were transduced by plating 0.1×10^6 to 0.5×10^6 cells in 12-well plates and adding 4 to 6 h later appropriate volumes of viral supernatants to obtain at least 65% GFP-positive cells. Because of the strong postentry block against the lentiviral vector in OMK cells due to the owl monkey TRIMcyp (63), cyclosporine was added (final concentration, 2.5 μM) to increase transduction efficiencies in these cells. Five days postransduction, cells were collected for FACS and Western blot analysis. For AGM COS-7 cells, transduction efficiencies were low and GFP-positive cells were isolated using FACS AsriaII sorting.

Luciferase reporter assay. Transfection of plasmid DNA in HepG2 cells was performed using X-tremeGENE HP DNA transfection reagent (Roche) following the manufacturer's instructions. For luciferase reporter gene assay, cells were typically seeded at a density of about 6×10^5 per 30-mm-diameter well (1×10^5 cells/cm^2) and reverse transfected with 30 ng of reporter plasmid DNA pCMV-Gluc (New England BioLabs) and 2 μg of empty EBS-PL vector. The next day, cells were transduced with lentiviral vectors encoding GFP or the diverse GFP-tagged viral X proteins. Five to 7 days later, GLuc activity was measured by adding 5 μl of the cell supernatant sample to 50 μl of room temperature assay buffer (100 mM NaCl, 35 mM EDTA, 0.1% Tween 20, 300 mM sodium ascorbate, 0.8 μM coelenterazine in 1× phosphate-buffered saline [PBS]) and immediately measuring luminescence in a luminometer (Glomax; Promega).

Western blotting. Western blot analyses were performed as previously described (7). Cells were disrupted in 2% SDS and briefly sonicated. The protein concentration was estimated and normalized using the bicinchoninic acid (BCA) protein assay (Novagen). The membranes were probed with 1:5,000 mouse monoclonal anti-GFP antibody (Roche; 11814460001) to detect the GFP-tagged X proteins, 1:1,000 mouse monoclonal antibodies against Smc6 (Abgent; AT3956a), 1:500 rabbit polyclonal antibodies against Smc6 (a gift from A. R. Lehmann) (NIH 3T3 [Fig. 3C]) (64), 1:1,000 rabbit polyclonal anti-Nse4 (Abgent; AP9909A), 1:10,000 mouse monoclonal anti-glyceraldehyde-3-phosphate dehydrogenase (anti-GAPDH; Sigma-Aldrich; G8795), and 1:500 goat polyclonal anti-DDB1 (Everest Biotech) antibodies. Horseradish peroxidase-conjugated sheep anti-rabbit and anti-mouse IgG (Amersham Biosciences; 1:5,000) were used as secondary antibodies. Detection was performed by ECL (Pierce).

Co-IP assay. HepG2 cells were harvested and washed once with 1× PBS. The cells were resuspended in 200 μl of coimmunoprecipitation (co-IP)–lysis buffer (50 mM HEPES-KOH [pH 7.5], 150 mM NaCl, 5 mM KCl, 5 mM $MgCl_2$, 50 μM $ZnCl_2$, 0.1% IGEPAL CA-630 (nonionic nondenaturing detergent; Sigma-Aldrich) and protease inhibitor cocktail from Sigma). The lysates were clarified by centrifugation at 14,000 × *g* for 20 min at 4°C. The supernatants were collected and 20 μl was set aside as input samples. The rest of the supernatants were mixed in 800 μl of co-IP–lysis buffer with a 25-μl packed-bead volume of protein A-Sepharose CL-4B (GE Healthcare) coupled to 3 μl of anti-GFP antibody (Roche; 11814460001). After 2 h of incubation at 4°C under constant rotation, the beads were sedimented by brief centrifugation (1 min at 300 × *g*) and the supernatant was discarded. The beads were washed 3 times with 1 ml of co-IP–lysis buffer. Bound protein-protein complexes were released from the beads by addition of 20 μl of 2× Laemmli SDS sample buffer. The inputs and eluted proteins were resolved in a 10% SDS-PAGE gel and detected by Western blotting.

PHH isolation, HBV infection, and ELISA. PHHs were isolated from human liver tissue from three donors as previously described (65). They were infected with PCR normalized HBV or HBVΔX at a multiplicity of infection (MOI) of 500 viral genome equivalents per cell (6). An enzyme-linked immunosorbent assay (ELISA; Autobio Diagnostics) was used to determine the amount of HBe and HBs antigens in culture media from infected or mock-infected PHHs. The efficiency of infection was controlled by quantitative PCR analyses specific for the viral cccDNA using beta-globin levels to normalize cell numbers. DNA from infected and transduced cells was extracted using the Epicentre kit. For the cccDNA analysis, T5 exonuclease (TSE44111K; Epicentre) was added and the DNA further incubated at 37°C for 30 min and 70°C for 30 min. Specific TaqMan probes and primers were then used to assess the cccDNA content as described in reference 66. Transduction levels were estimated by the GFP expression in transduced cells by FACS analyses (BD FACSCalibur).

***De novo* sequencing of Smc5/6 genes.** Total RNA was extracted from 10^7 cells following the manufacturer's instructions (NucleoSpin RNA Blood; Macheray-Nagel; 740200). Reverse transcription was performed using SuperScript III reverse transcriptase (Thermo Fisher; 18080) with random hexamers and oligo(dT). Single-round PCR of overlapping fragments was performed using Q5 high-fidelity DNA polymerase (NEB; M0491) following the manufacturer's instructions. Sequences of the *Smc6* gene from owl monkey (OMK), cotton-headed monkey (B95a), tantalus and vervet AGM (COS-7 and Vero), human (293T and HepG2), and ferret cells were obtained. Primers used for the amplification and the sequencing are available in supplemental Table 1 at https://figshare.com/articles/Table_S1_Primers_used_to_amplify_and_sequence_endogenous_Smc6_from_different_mammalian_cell_lines_/6194843.

Host phylogenetic analyses. Sequences of primate or mammalian orthologous genes were retrieved from publicly available databases using UCSC Blat and NCBI BLASTN with the human sequence as the query. After including sequences obtained from *de novo* sequencing in this study, the sequences of the orthologues were codon aligned using Muscle (67), with minor adjustments (alignments are in supplemental data set 1 at https://figshare.com/articles/DatasetS1_Host_gene_alignments_used_in_the_study_fasta_format_and_phylogenetic_analyses_newick_format_Nsmce1-4_Smc1-6/6194813). The names of the sequences have been made uniform: the first three letters of the host genus name followed by the first three letters of the species name (e.g., papAnu for *Papio anubis* or macMul for *Macaca mulata*). When sequences were retrieved using Blat on the primate full-genome assembly, the name of the genome assembly was used (e.g., panTro4 for *Pan troglodytes* or hg38 for human). The synteny of each locus of interest was analyzed in UCSC using Blat and the Genome Browser. When necessary, the genomic sequences were further retrieved and aligned to a reference gene to determine the pseudogenes and the gene orders.

We used GARD from HYPHY to perform the recombination analyses with a P value cutoff of <0.05 (28, 68). PhyML was used for the phylogenetic reconstructions with an HKY+G+I model and an approximate likelihood ratio test (aLRT) or 1,000 bootstrap replicates for branch support (69).

Positive selection analyses. Maximum-likelihood tests to assay for positive selection were performed using three platforms: HYPHY from Kosakovsky Pond and colleagues, PAML from Yang, and Bio++ from Guéguen and colleagues (31, 32, 68). In HYPHY, we used PARRIS to detect if a subset of sites in the alignment has evolved under positive selection (30). We further used the more recent BUSTED method, which detects gene-wide evidence of positive selection within a codon alignment (29). In PAML, we used Codeml with the corresponding gene tree inferred with PhyML as input. Parameters were checked using an M0 model. The gene-coding sequence alignments were fit to models that disallow (M7) or allow (M8) positive selection. For several genes, we did not get convergence certainly because genes were too conserved and the sum of dS across the tree was too low. We therefore exclude the Codeml and Bio++ analyses of seven genes, for which the p or the q parameters of the beta distribution was extreme (<0.05 or >99 [31]) (Fig. 1C, NA). The likelihood of models was compared using a chi-square test to derive P values.

For the *Smc6* gene, for which we found evidence of positive selection in some of the previously described methods, we further analyzed (i) which lineage(s) during primate evolution has been subjected to positive selection (i.e., when did the gene experience rapid evolution?) and (ii) which sites have been under positive selection (i.e., where has the gene evolved more rapidly than expected?). For branch-specific analyses, we used BS-REL from HYPHY, which identifies if certain lineages have undergone positive selection (39). To detect episodic site-specific positively selected sites, we used MEME from HYPHY (70). We also ran FUBAR in Datamonkey, which uses Bayesian inference to detect positive and negative selection at individual sites (71). In Bio++ and Codeml, we used the Bayesian posterior probability and the BEB analysis from the respective M8 model to identify codons with dN/dS ratios of >1 (sites with posterior probability of >0.90 are presented here in Fig. 2A).

Virus phylogenetics. The nucleotide sequences of the *X* gene from orthohepadnaviruses were retrieved using BLASTN with the human HBx as the query. Only one sequence per orthohepadnaviral lineage was retained for this analysis. We performed the amino acid alignment with Muscle (total length, 161 aa), and we used PhyML for the phylogenetic reconstructions with a JTT+G+I model and 1,000 bootstrap replicates for branch support (69).

Ethics statement. PHHs were prepared from adult surgical liver resections provided by Michel Rivoire's, Jean-Yves Mabrut's, and Guillaume Passot's departments. Approval from the local and national ethics committees (French Ministry of Research and Education numbers AC-2013-1871, DC-2013-1870, and DC-2008-235) and written informed consent from patients were obtained.

Accession number(s). All the new *Smc6* sequences are available at GenBank under accession numbers MF624755 to MF624761.

ACKNOWLEDGMENTS

We very sincerely thank Léa Picard and Laurent Guéguen for the Bio++ analyses and input on evolutionary analyses, Véronique Barateau for technical support with the flow cytometry analyses, Andrea Cimarelli, Olivier Hantz, and Isabelle Chemin for scientific discussions, and Loic Peyrot, Anaelle Dubois, and Francoise Berby for technical assistance. We thank Maud Michelet, Jennifer Molle, Loic Peyrot, Anaelle Dubois, Océane Floriot, Laura Dimer, Marie-Laure Plissonnier, and Julie Lucifora (CRCL) for the isolation of primary human hepatocytes and Michel Rivoire's, Jean-Yves Mabrut's, and Guillaume Passot's staff for providing liver resections. We thank A. R. Lehmann for providing anti-Smc6 antibody. We thank Michael Emerman, Andrea Cimarelli, and Oliver Fregoso for their comments on the manuscript. We thank all the contributors of publicly available genome sequences.

This work was supported by an amfAR Mathilde Krim Phase II Fellowship (109140-58-RKHF to L.E.), a Fondation pour la Recherche Médicale (FRM) Projet Innovant grant (ING20160435028 to L.E.), a FINOVI "recently settled scientist" grant (to L.E.), an ANR LabEx ECOFECT grant (ANR-11-LABX-0048 of Université de Lyon, within the program

Investissements d'Avenir [ANR-11-IDEX-0007] operated by the French National Research Agency, GrASP, to L.E.), an ANRS grant (ECTZ19143 to L.E.), a fellowship from La Ligue Contre le Cancer (to L.G.), a grant from the Swiss National Science Foundation (310030-149626 to M.S.), and by the Canton of Geneva.

F.A., L.G., M.S., and L.E. conceptualized the study. J.H., M.S., and L.E. acquired funding. F.F., F.A., L.G., and L.E. designed the methodology. F.F. and L.E. performed the phylogenetic and evolutionary analyses. F.F., F.A., L.G., A.P., and L.E. performed experiments. All the authors analyzed the data. L.E. was the project administrator. F.A., J.H., M.S., and L.E. supervised the study. F.F. and L.E. wrote the original draft of the manuscript, and all the authors reviewed and edited the manuscript.

REFERENCES

1. Suh A, Weber CC, Kehlmaier C, Braun EL, Green RE, Fritz U, Ray DA, Ellegren H. 2014. Early mesozoic coexistence of amniotes and hepadnaviridae. PLoS Genet 10:e1004559. https://doi.org/10.1371/journal.pgen.1004559.
2. Dill JA, Camus AC, Leary JH, Di Giallonardo F, Holmes EC, Ng TF. 2016. Distinct viral lineages from fish and amphibians reveal the complex evolutionary history of hepadnaviruses. J Virol 90:7920–7933. https://doi.org/10.1128/JVI.00832-16.
3. Lauber C, Seitz S, Mattei S, Suh A, Beck J, Herstein J, Borold J, Salzburger W, Kaderali L, Briggs JAG, Bartenschlager R. 2017. Deciphering the origin and evolution of hepatitis B viruses by means of a family of non-enveloped fish viruses. Cell Host Microbe 22:387–399.e386. https://doi.org/10.1016/j.chom.2017.07.019.
4. Feitelson MA, Bonamassa B, Arzumanyan A. 2014. The roles of hepatitis B virus-encoded X protein in virus replication and the pathogenesis of chronic liver disease. Expert Opin Ther Targets 18:293–306. https://doi.org/10.1517/14728222.2014.867947.
5. Benhenda S, Cougot D, Buendia MA, Neuveut C. 2009. Hepatitis B virus X protein molecular functions and its role in virus life cycle and pathogenesis. Adv Cancer Res 103:75–109. https://doi.org/10.1016/S0065-230X(09)03004-8.
6. Lucifora J, Arzberger S, Durantel D, Belloni L, Strubin M, Levrero M, Zoulim F, Hantz O, Protzer U. 2011. Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. J Hepatol 55:996–1003. https://doi.org/10.1016/j.jhep.2011.02.015.
7. Decorsière A, Mueller H, van Breugel PC, Abdul F, Gerossier L, Beran RK, Livingston CM, Niu C, Fletcher SP, Hantz O, Strubin M. 2016. Hepatitis B virus X protein identifies the Smc5/6 complex as a host restriction factor. Nature 531:386–389. https://doi.org/10.1038/nature17170.
8. Murphy CM, Xu Y, Li F, Nio K, Reszka-Blanco N, Li X, Wu Y, Yu Y, Xiong Y, Su L. 2016. Hepatitis B virus X protein promotes degradation of SMC5/6 to enhance HBV replication. Cell Rep 16:2846–2854. https://doi.org/10.1016/j.celrep.2016.08.026.
9. Cobbe N, Heck MM. 2004. The evolution of SMC proteins: phylogenetic analysis and structural implications. Mol Biol Evol 21:332–347. https://doi.org/10.1093/molbev/msh023.
10. Uhlmann F. 2016. SMC complexes: from DNA to chromosomes. Nat Rev Mol Cell Biol 17:399–412. https://doi.org/10.1038/nrm.2016.30.
11. Gligoris T, Lowe J. 2016. Structural insights into ring formation of cohesin and related smc complexes. Trends Cell Biol 26:680–693. https://doi.org/10.1016/j.tcb.2016.04.002.
12. Watanabe Y. 2005. The importance of being Smc5/6. Nat Cell Biol 7:329–331. https://doi.org/10.1038/ncb0405-329.
13. Menolfi D, Delamarre A, Lengronne A, Pasero P, Branzei D. 2015. Essential roles of the Smc5/6 complex in replication through natural pausing sites and endogenous DNA damage tolerance. Mol Cell 60:835–846. https://doi.org/10.1016/j.molcel.2015.10.023.
14. Kegel A, Sjogren C. 2010. The Smc5/6 complex: more than repair? Cold Spring Harbor Symp Quant Biol 75:179–187. https://doi.org/10.1101/sqb.2010.75.047.
15. Fernandez-Capetillo O. 2016. The (elusive) role of the SMC5/6 complex. Cell Cycle 15:775–776. https://doi.org/10.1080/15384101.2015.1137713.
16. Hwang G, Sun F, O'Brien M, Eppig JJ, Handel MA, Jordan PW. 2017. SMC5/6 is required for the formation of segregation-competent bivalent chromosomes during meiosis I in mouse oocytes. Development 144:1648–1660. https://doi.org/10.1242/dev.145607.
17. Niu C, Livingston CM, Li L, Beran RK, Daffis S, Ramakrishnan D, Burdette D, Peiser L, Salas E, Ramos H, Yu M, Cheng G, Strubin M, Delaney WE, IV, Fletcher SP. 2017. The Smc5/6 complex restricts HBV when localized to ND10 without inducing an innate immune response and is counteracted by the HBV X protein shortly after infection. PLoS One 12:e0169648. https://doi.org/10.1371/journal.pone.0169648.
18. Duggal NK, Emerman M. 2012. Evolutionary conflicts between viruses and restriction factors shape immunity. Nat Rev Immunol 12:687–695. https://doi.org/10.1038/nri3295.
19. Johnson WE. 2013. Rapid adversarial co-evolution of viruses and cellular restriction factors. Curr Top Microbiol Immunol 371:123–151.
20. Sawyer SL, Wu LI, Emerman M, Malik HS. 2005. Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. Proc Natl Acad Sci U S A 102:2832–2837. https://doi.org/10.1073/pnas.0409853102.
21. Daugherty MD, Malik HS. 2012. Rules of engagement: molecular insights from host-virus arms races. Annu Rev Genet 46:677–700. https://doi.org/10.1146/annurev-genet-110711-155522.
22. Lim ES, Malik HS, Emerman M. 2010. Ancient adaptive evolution of tetherin shaped the functions of Vpu and Nef in human immunodeficiency virus and primate lentiviruses. J Virol 84:7124–7134. https://doi.org/10.1128/JVI.00468-10.
23. Jia B, Serra-Moreno R, Neidermyer W, Rahmberg A, Mackey J, Fofana IB, Johnson WE, Westmoreland S, Evans DT. 2009. Species-specific activity of SIV Nef and HIV-1 Vpu in overcoming restriction by tetherin/BST2. PLoS Pathog 5:e1000429. https://doi.org/10.1371/journal.ppat.1000429.
24. Sauter D, Schindler M, Specht A, Landford WN, Munch J, Kim KA, Votteler J, Schubert U, Bibollet-Ruche F, Keele BF, Takehisa J, Ogando Y, Ochsenbauer C, Kappes JC, Ayouba A, Peeters M, Learn GH, Shaw G, Sharp PM, Bieniasz P, Hahn BH, Hatziioannou T, Kirchhoff F. 2009. Tetherin-driven adaptation of Vpu and Nef function and the evolution of pandemic and nonpandemic HIV-1 strains. Cell Host Microbe 6:409–421. https://doi.org/10.1016/j.chom.2009.10.004.
25. Filleton F, Abdul F, Gerossier L, Paturel A, Hall J, Strubin M, Etienne L. 1 December 2017. Smc5/6-antagonism by HBx is an evolutionary-conserved function of hepatitis B virus infection in mammals. bioRxiv https://doi.org/10.1101/202671.
26. McBee RM, Rozmiarek SA, Meyerson NR, Rowley PA, Sawyer SL. 2015. The effect of species representation on the detection of positive selection in primate gene data sets. Mol Biol Evol 32:1091–1096. https://doi.org/10.1093/molbev/msu399.
27. Perelman P, Johnson WE, Roos C, Seuanez HN, Horvath JE, Moreira MA, Kessing B, Pontius J, Roelke M, Rumpler Y, Schneider MP, Silva A, O'Brien SJ, Pecon-Slattery J. 2011. A molecular phylogeny of living primates. PLoS Genet 7:e1001342. https://doi.org/10.1371/journal.pgen.1001342.
28. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Mol Biol Evol 23:1891–1901. https://doi.org/10.1093/molbev/msl051.
29. Murrell B, Weaver S, Smith MD, Wertheim JO, Murrell S, Aylward A, Eren K, Pollner T, Martin DP, Smith DM, Scheffler K, Kosakovsky Pond SL. 2015. Gene-wide identification of episodic selection. Mol Biol Evol 32:1365–1371. https://doi.org/10.1093/molbev/msv035.
30. Scheffler K, Martin DP, Seoighe C. 2006. Robust inference of positive selection from recombining coding sequences. Bioinformatics 22:2493–2499. https://doi.org/10.1093/bioinformatics/btl427.
31. Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591. https://doi.org/10.1093/molbev/msm088.

32. Guéguen L, Gaillard S, Boussau B, Gouy M, Groussin M, Rochette NC, Bigot T, Fournier D, Pouyet F, Cahais V, Bernard A, Scornavacca C, Nabholz B, Haudry A, Dachary L, Galtier N, Belkhir K, Dutheil JY. 2013. Bio++: efficient extensible libraries and tools for computational molecular evolution. Mol Biol Evol 30:1745–1750. https://doi.org/10.1093/molbev/mst097.
33. Gueguen L, Duret L. 6 April 2017. Unbiased estimate of synonymous and non-synonymous substitution rates with non-stationary base composition. bioRxiv https://doi.org/10.1101/124925.
34. Demogines A, Abraham J, Choe H, Farzan M, Sawyer SL. 2013. Dual host-virus arms races shape an essential housekeeping protein. PLoS Biol 11:e1001571. https://doi.org/10.1371/journal.pbio.1001571.
35. Ng M, Ndungo E, Kaczmarek ME, Herbert AS, Binger T, Kuehne AI, Jangra RK, Hawkins JA, Gifford RJ, Biswas R, Demogines A, James RM, Yu M, Brummelkamp TR, Drosten C, Wang LF, Kuhn JH, Muller MA, Dye JM, Sawyer SL, Chandran K. 2015. Filovirus receptor NPC1 contributes to species-specific patterns of ebolavirus susceptibility in bats. Elife 4:e11785. https://doi.org/10.7554/eLife.11785.
36. Lou DI, Kim ET, Meyerson NR, Pancholi NJ, Mohni KN, Enard D, Petrov DA, Weller SK, Weitzman MD, Sawyer SL. 2016. An intrinsically disordered region of the DNA repair protein Nbs1 is a species-specific barrier to herpes simplex virus 1 in primates. Cell Host Microbe 20:178–188. https://doi.org/10.1016/j.chom.2016.07.003.
37. Enard D, Cai L, Gwennap C, Petrov DA. 2016. Viruses are a dominant driver of protein adaptation in mammals. Elife 5:e12469. https://doi.org/10.7554/eLife.12469.
38. Yang Z, Wong WS, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol 22:1107–1118. https://doi.org/10.1093/molbev/msi097.
39. Kosakovsky Pond SL, Murrell B, Fourment M, Frost SD, Delport W, Scheffler K. 2011. A random effects branch-site model for detecting episodic diversifying selection. Mol Biol Evol 28:3033–3043. https://doi.org/10.1093/molbev/msr125.
40. Etienne L. 2015. Virus-host evolution and positive selection, p 1–13. *In* (ed), Encyclopedia of AIDS. https://doi.org/10.1007/978-1-4614-9610-6_373-1.
41. Kluge SF, Sauter D, Kirchhoff F. 2015. SnapShot: antiviral restriction factors. Cell 163:774–774.e771. https://doi.org/10.1016/j.cell.2015.10.019.
42. Malim MH, Bieniasz PD. 2012. HIV restriction factors and mechanisms of evasion. Cold Spring Harb Perspect Med 2:a006940. https://doi.org/10.1101/cshperspect.a006940.
43. Usami Y, Wu Y, Gottlinger HG. 2015. SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. Nature 526:218–223. https://doi.org/10.1038/nature15400.
44. Rosa A, Chande A, Ziglio S, De Sanctis V, Bertorelli R, Goh SL, McCauley SM, Nowosielska A, Antonarakis SE, Luban J, Santoni FA, Pizzato M. 2015. HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation. Nature 526:212–217. https://doi.org/10.1038/nature15399.
45. Murrell B, Vollbrecht T, Guatelli J, Wertheim JO. 2016. The evolutionary histories of antiretroviral proteins SERINC3 and SERINC5 do not support an evolutionary arms race in primates. J Virol 90:8085–8089. https://doi.org/10.1128/JVI.00972-16.
46. Heigele A, Kmiec D, Regensburger K, Langer S, Peiffer L, Sturzel CM, Sauter D, Peeters M, Pizzato M, Learn GH, Hahn BH, Kirchhoff F. 2016. The potency of Nef-mediated SERINC5 antagonism correlates with the prevalence of primate lentiviruses in the wild. Cell Host Microbe 20:381–391. https://doi.org/10.1016/j.chom.2016.08.004.
47. Jäger S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, Clarke SC, Shales M, Mercenne G, Pache L, Li K, Hernandez H, Jang GM, Roth SL, Akiva E, Marlett J, Stephens M, D'Orso I, Fernandes J, Fahey M, Mahon C, O'Donoghue AJ, Todorovic A, Morris JH, Maltby DA, Alber T, Cagney G, Bushman FD, Young JA, Chanda SK, Sundquist WI, Kortemme T, Hernandez RD, Craik CS, Burlingame A, Sali A, Frankel AD, Krogan NJ. 2011. Global landscape of HIV-human protein complexes. Nature 481:365–370. https://doi.org/10.1038/nature10719.
48. Richards C, Albin JS, Demir O, Shaban NM, Luengas EM, Land AM, Anderson BD, Holten JR, Anderson JS, Harki DA, Amaro RE, Harris RS. 2015. The binding interface between human APOBEC3F and HIV-1 Vif elucidated by genetic and computational approaches. Cell Rep 13:1781–1788. https://doi.org/10.1016/j.celrep.2015.10.067.
49. Letko M, Booiman T, Kootstra N, Simon V, Ooms M. 2015. Identification of the HIV-1 Vif and human APOBEC3G protein interface. Cell Rep 13:1789–1799. https://doi.org/10.1016/j.celrep.2015.10.068.
50. Lempp FA, Wiedtke E, Qu B, Roques P, Chemin I, Vondran FWR, Le Grand R, Grimm D, Urban S. 2017. Sodium taurocholate cotransporting polypeptide is the limiting host factor of hepatitis B virus infection in macaque and pig hepatocytes. Hepatology 66:703–716. https://doi.org/10.1002/hep.29112.
51. Chouteau P, Le Seyec J, Cannie I, Nassal M, Guguen-Guillouzo C, Gripon P. 2001. A short N-proximal region in the large envelope protein harbors a determinant that contributes to the species specificity of human hepatitis B virus. J Virol 75:11565–11572. https://doi.org/10.1128/JVI.75.23.11565-11572.2001.
52. Drexler JF, Geipel A, Konig A, Corman VM, van Riel D, Leijten LM, Bremer CM, Rasche A, Cottontail VM, Maganga GD, Schlegel M, Muller MA, Adam A, Klose SM, Carneiro AJ, Stocker A, Franke CR, Gloza-Rausch F, Geyer J, Annan A, Adu-Sarkodie Y, Oppong S, Binger T, Vallo P, Tschapka M, Ulrich RG, Gerlich WH, Leroy E, Kuiken T, Glebe D, Drosten C. 2013. Bats carry pathogenic hepadnaviruses antigenically related to hepatitis B virus and capable of infecting human hepatocytes. Proc Natl Acad Sci U S A 110:16151–16156. https://doi.org/10.1073/pnas.1308049110.
53. Sawyer SL, Elde NC. 2012. A cross-species view on viruses. Curr Opin Virol 2:561–568. https://doi.org/10.1016/j.coviro.2012.07.003.
54. Yan H, Zhong G, Xu G, He W, Jing Z, Gao Z, Huang Y, Qi Y, Peng B, Wang H, Fu L, Song M, Chen P, Gao W, Ren B, Sun Y, Cai T, Feng X, Sui J, Li W. 2012. Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. Elife 1:e00049. https://doi.org/10.7554/eLife.00049.
55. Hahn CM, Iwanowicz LR, Cornman RS, Conway CM, Winton JR, Blazer VS. 2015. Characterization of a novel hepadnavirus in the white sucker (Catostomus commersonii) from the Great Lakes region of the United States. J Virol 89:11801–11811. https://doi.org/10.1128/JVI.01278-15.
56. Geoghegan JL, Duchene S, Holmes EC. 2017. Comparative analysis estimates the relative frequencies of co-divergence and cross-species transmission within viral families. PLoS Pathog 13:e1006215. https://doi.org/10.1371/journal.ppat.1006215.
57. Mitchell PS, Young JM, Emerman M, Malik HS. 2015. Evolutionary analyses suggest a function of MxB immunity proteins beyond lentivirus restriction. PLoS Pathog 11:e1005304. https://doi.org/10.1371/journal.ppat.1005304.
58. Mitchell PS, Emerman M, Malik HS. 2013. An evolutionary perspective on the broad antiviral specificity of MxA. Curr Opin Microbiol 16:493–499. https://doi.org/10.1016/j.mib.2013.04.005.
59. Elde NC, Child SJ, Geballe AP, Malik HS. 2009. Protein kinase R reveals an evolutionary model for defeating viral mimicry. Nature 457:485–489. https://doi.org/10.1038/nature07529.
60. van Breugel PC, Robert EI, Mueller H, Decorsiere A, Zoulim F, Hantz O, Strubin M. 2012. Hepatitis B virus X protein stimulates gene expression selectively from extrachromosomal DNA templates. Hepatology 56:2116–2124. https://doi.org/10.1002/hep.25928.
61. Lin-Marq N, Bontron S, Leupin O, Strubin M. 2001. Hepatitis B virus X protein interferes with cell viability through interaction with the p127-kDa UV-damaged DNA-binding protein. Virology 287:266–274. https://doi.org/10.1006/viro.2001.1036.
62. Leupin O, Bontron S, Schaeffer C, Strubin M. 2005. Hepatitis B virus X protein stimulates viral genome replication via a DDB1-dependent pathway distinct from that leading to cell death. J Virol 79:4238–4245. https://doi.org/10.1128/JVI.79.7.4238-4245.2005.
63. Sayah DM, Sokolskaja E, Berthoux L, Luban J. 2004. Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. Nature 430:569–573. https://doi.org/10.1038/nature02777.
64. Taylor EM, Moghraby JS, Lees JH, Smit B, Moens PB, Lehmann AR. 2001. Characterization of a novel human SMC heterodimer homologous to the Schizosaccharomyces pombe Rad18/Spr18 complex. Mol Biol Cell 12:1583–1594. https://doi.org/10.1091/mbc.12.6.1583.
65. Gripon P, Diot C, Guguen-Guillouzo C. 1993. Reproducible high level infection of cultured adult human hepatocytes by hepatitis B virus: effect of polyethylene glycol on adsorption and penetration. Virology 192:534–540.
66. Allweiss L, Volz T, Giersch K, Kah J, Raffa G, Petersen J, Lohse AW, Beninati C, Pollicino T, Urban S, Lutgehetmann M, Dandri M. 2018. Proliferation of primary human hepatocytes and prevention of hepatitis B virus reinfection efficiently deplete nuclear cccDNA in vivo. Gut 67:542–552. https://doi.org/10.1136/gutjnl-2016-312162.
67. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.
68. Pond SL, Frost SD, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21:676–679. https://doi.org/10.1093/bioinformatics/bti079.

69. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321. https://doi.org/10.1093/sysbio/syq010.
70. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. 2012. Detecting individual sites subject to episodic diversifying selection. PLoS Genet 8:e1002764. https://doi.org/10.1371/journal.pgen.1002764.
71. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K. 2013. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol Biol Evol 30:1196–1205. https://doi.org/10.1093/molbev/mst030.
72. Li T, Robert EI, van Breugel PC, Strubin M, Zheng N. 2010. A promiscuous alpha-helical motif anchors viral hijackers and substrate receptors to the CUL4-DDB1 ubiquitin ligase machinery. Nat Struct Mol Biol 17:105–111. https://doi.org/10.1038/nsmb.1719.

REFERENCES

Acs, G., Sells, M.A., Purcell, R.H., Price, P., Engle, R., Shapiro, M., and Popper, H. (1987). Hepatitis B virus produced by transfected Hep G2 cells causes hepatitis in chimpanzees. Proc. Natl. Acad. Sci. *84*, 4641–4644. https://doi.org/10.1073/pnas.84.13.4641.

Aden, D.P., Fogel, A., Plotkin, S., Damjanov, I., and Knowles, B.B. (1979). Controlled synthesis of HBsAg in a differentiated human liver carcinoma-derived cell line. Nature *282*, 615–616. https://doi.org/10.1038/282615a0.

Alarcon, V., Hernández, S., Rubio, L., Alvarez, F., Flores, Y., Varas-Godoy, M., De Ferrari, G.V., Kann, M., Villanueva, R.A., and Loyola, A. (2016). The enzymes LSD1 and Set1A cooperate with the viral protein HBx to establish an active hepatitis B viral chromatin state. Sci. Rep. *6*, 25901. https://doi.org/10.1038/srep25901.

Alexander, J., Bey, E., Whitcutt, J.M., and Gear, J.H. (1976a). Adaptation of cells derived from human malignant tumours to growth in vitro. S. Afr. J. Med. Sci. *41*, 89–98. .

Alexander, J.J., Bey, E.M., Geddes, E.W., and Lecatsas, G. (1976b). Establishment of a continuously growing cell line from primary carcinoma of the liver. South Afr. Med. J. Suid-Afr. Tydskr. Vir Geneeskd. *50*, 2124–2128. .

Alfaiate, D., Dény, P., and Durantel, D. (2015). Hepatitis delta virus: From biological and medical aspects to current and investigational therapeutic options. Antiviral Res. *122*, 112–129. https://doi.org/10.1016/j.antiviral.2015.08.009.

Ali, A., Abdel-Hafiz, H., Suhail, M., Al-Mars, A., Zakaria, M.K., Fatima, K., Ahmad, S., Azhar, E., Chaudhary, A., and Qadri, I. (2014). Hepatitis B virus, HBx mutants and their role in hepatocellular carcinoma. World J. Gastroenterol. *20*, 10238–10248. https://doi.org/10.3748/wjg.v20.i30.10238.

Alon, D., Stein, G.Y., Hadas-Golan, V., Tau, L., Brosh, T., and Turner, D. (2017). Immunogenicity of Sci-B-Vac (a Third-Generation Hepatitis B Vaccine) in HIV-Positive Adults. Isr. Med. Assoc. J. IMAJ *19*, 143–146. .

Alonso, S., Guerra, A.-R., Carreira, L., Ferrer, J.-Á., Gutiérrez, M.-L., and Fernandez-Rodriguez, C.M. (2017). Upcoming pharmacological developments in chronic hepatitis B: can we glimpse a cure on the horizon? BMC Gastroenterol. *17*, 168. https://doi.org/10.1186/s12876-017-0726-2.

Alter, M.J. (2006). Epidemiology of viral hepatitis and HIV co-infection. J. Hepatol. *44*, S6-9. https://doi.org/10.1016/j.jhep.2005.11.004.

Altinel, K., Hashimoto, K., Wei, Y., Neuveut, C., Gupta, I., Suzuki, A.M., Dos Santos, A., Moreau, P., Xia, T., Kojima, S., et al. (2016). Single-Nucleotide Resolution Mapping of Hepatitis B Virus Promoters in Infected Human Livers and Hepatocellular Carcinoma. J. Virol. *90*, 10811–10822. https://doi.org/10.1128/JVI.01625-16.

Alvarado-Mora, M.V., and Pinho, J.R.R. (2013). Distribution of HBV genotypes in Latin America. Antivir. Ther. *18*, 459–465. https://doi.org/10.3851/IMP2599.

Astbury, S., Costa Nunes Soares, M.M., Peprah, E., King, B., Jardim, A.C.G., Shimizu, J.F., Jalal, P., Saeed, C.H., Sabeer, F.T., Irving, W.L., et al. (2020). Nanopore sequencing from extraction-free direct PCR of dried serum spots for portable hepatitis B virus drug-resistance typing. J. Clin. Virol. *129*, 104483. https://doi.org/10.1016/j.jcv.2020.104483.

Bai, L., Zhang, X., Kozlowski, M., Li, W., Wu, M., Liu, J., Chen, L., Zhang, J., Huang, Y., and Yuan, Z. (2018). Extracellular Hepatitis B Virus RNAs Are Heterogeneous in Length and Circulate as Capsid-Antibody Complexes in Addition to Virions in Chronic Hepatitis B Patients. J. Virol. *92*. https://doi.org/10.1128/JVI.00798-18.

Barril, G., and Teruel, J.L. (2008). [Vaccination scheme in advanced chronic kidney disease]. Nefrol. Publicacion Of. Soc. Espanola Nefrol. *28 Suppl 3*, 95–99. .

Bartenschlager, R., Junker-Niepmann, M., and Schaller, H. (1990). The P gene product of hepatitis B virus is required as a structural component for genomic RNA encapsidation. J. Virol. *64*, 5324–5332. .

Basu, A.S. (2017). Digital Assays Part I: Partitioning Statistics and Digital PCR. SLAS Technol. Transl. Life Sci. Innov. *22*, 369–386. https://doi.org/10.1177/2472630317705680.

Baumert, T.F., Thimme, R., and von Weizsäcker, F. (2007). Pathogenesis of hepatitis B virus infection. World J. Gastroenterol. *13*, 82–90. https://doi.org/10.3748/wjg.v13.i1.82.

Bayard, Q., Meunier, L., Peneau, C., Renault, V., Shinde, J., Nault, J.-C., Mami, I., Couchy, G., Amaddeo, G., Tubacher, E., et al. (2018). Cyclin A2/E1 activation defines a hepatocellular carcinoma subclass with a rearrangement signature of replication stress. Nat. Commun. *9*, 5235. https://doi.org/10.1038/s41467-018-07552-9.

Baylis, S.A., Heath, A.B., Chudy, M., Pisani, G., Klotz, A., Kerby, S., and Gerlich, W. (2008). An international collaborative study to establish the 2nd World Health Organization International Standard for hepatitis B virus DNA nucleic acid amplification technology-based assays. Vox Sang. *94*, 358–362. https://doi.org/10.1111/j.1423-0410.2008.01023.x.

Bayliss, J., Lim, L., Thompson, A.J.V., Desmond, P., Angus, P., Locarnini, S., and Revill, P.A. (2013). Hepatitis B virus splicing is enhanced prior to development of hepatocellular carcinoma. J. Hepatol. *59*, 1022–1028. https://doi.org/10.1016/j.jhep.2013.06.018.

Belloni, L., Pollicino, T., De Nicola, F., Guerrieri, F., Raffa, G., Fanciulli, M., Raimondo, G., and Levrero, M. (2009). Nuclear HBx binds the HBV minichromosome and modifies the epigenetic regulation of cccDNA function. Proc. Natl. Acad. Sci. U. S. A. *106*, 19975–19979. https://doi.org/10.1073/pnas.0908365106.

Belloni, L., Allweiss, L., Guerrieri, F., Pediconi, N., Volz, T., Pollicino, T., Petersen, J., Raimondo, G., Dandri, M., and Levrero, M. (2012). IFN-α inhibits HBV transcription and replication in cell culture and in humanized mice by targeting the epigenetic regulation of the nuclear cccDNA minichromosome. J. Clin. Invest. *122*, 529–537. https://doi.org/10.1172/JCI58847.

Bertoletti, A., and Ferrari, C. (2016). Adaptive immunity in HBV infection. J. Hepatol. *64*, S71–S83. https://doi.org/10.1016/j.jhep.2016.01.026.

Bill, C.A., and Summers, J. (2004). Genomic DNA double-strand breaks are targets for hepadnaviral DNA integration. Proc. Natl. Acad. Sci. U. S. A. *101*, 11135–11140. https://doi.org/10.1073/pnas.0403925101.

Blachier, M., Leleu, H., Peck-Radosavljevic, M., Valla, D.-C., and Roudot-Thoraval, F. (2013). The burden of liver disease in Europe: a review of available epidemiological data. J. Hepatol. *58*, 593–608. https://doi.org/10.1016/j.jhep.2012.12.005.

Blumberg, B.S. (1977). Australia antigen and the biology of hepatitis B. Science *197*, 17–25. https://doi.org/10.1126/science.325649.

Blumberg, B.S. (1997). Hepatitis B virus, the vaccine, and the control of primary cancer of the liver. Proc. Natl. Acad. Sci. U. S. A. *94*, 7121–7125. https://doi.org/10.1073/pnas.94.14.7121.

Blumberg, B.S., Alter, H.J., and Visnich, S. (1965). A “NEW” ANTIGEN IN LEUKEMIA SERA. JAMA *191*, 541–546. https://doi.org/10.1001/jama.1965.03080070025007.

Bock, C.T., Schwinn, S., Locarnini, S., Fyfe, J., Manns, M.P., Trautwein, C., and Zentgraf, H. (2001). Structural organization of the hepatitis B virus minichromosome. J. Mol. Biol. *307*, 183–196. https://doi.org/10.1006/jmbi.2000.4481.

van Bömmel, F., Bartens, A., Mysickova, A., Hofmann, J., Krüger, D.H., Berg, T., and Edelmann, A. (2015). Serum hepatitis B virus RNA levels as an early predictor of hepatitis B envelope antigen seroconversion during treatment with polymerase inhibitors. Hepatol. Baltim. Md *61*, 66–76. https://doi.org/10.1002/hep.27381.

Boni, C., Vecchi, A., Rossi, M., Laccabue, D., Giuberti, T., Alfieri, A., Lampertico, P., Grossi, G., Facchetti, F., Brunetto, M.R., et al. (2018). TLR7 Agonist Increases Responses of Hepatitis B Virus–Specific T Cells and Natural Killer Cells in Patients With Chronic Hepatitis B Treated With Nucleos(T)Ide Analogues. Gastroenterology *154*, 1764-1777.e7. https://doi.org/10.1053/j.gastro.2018.01.030.

Bonilla Guerrero, R., and Roberts, L.R. (2005). The role of hepatitis B virus integrations in the pathogenesis of human hepatocellular carcinoma. J. Hepatol. *42*, 760–777. https://doi.org/10.1016/j.jhep.2005.02.005.

Böttcher, B., Vogel, M., Ploss, M., and Nassal, M. (2006). High plasticity of the hepatitis B virus capsid revealed by conformational stress. J. Mol. Biol. *356*, 812–822. https://doi.org/10.1016/j.jmb.2005.11.053.

Bottero, J., Boyd, A., Gozlan, J., Lemoine, M., Carrat, F., Collignon, A., Boo, N., Dhotte, P., Varsat, B., Muller, G., et al. (2013). Performance of rapid tests for detection of HBsAg and anti-HBsAb in a large cohort, France. J. Hepatol. *58*, 473–478. https://doi.org/10.1016/j.jhep.2012.11.016.

Bousali, M., Papatheodoridis, G., Paraskevis, D., and Karamitros, T. (2021). Hepatitis B Virus DNA Integration, Chronic Infections and Hepatocellular Carcinoma. Microorganisms *9*, 1787. https://doi.org/10.3390/microorganisms9081787.

Brechot, C., Pourcel, C., Louise, A., Rain, B., and Tiollais, P. (1980). Presence of integrated hepatitis B virus DNA sequences in cellular DNA of human hepatocellular carcinoma. Nature *286*, 533–535. https://doi.org/10.1038/286533a0.

Bruss, V. (2007). Hepatitis B virus morphogenesis. World J. Gastroenterol. *13*, 65–73. https://doi.org/10.3748/wjg.v13.i1.65.

Burns, G.S., and Thompson, A.J. (2014). Viral Hepatitis B: Clinical and Epidemiological Characteristics. Cold Spring Harb. Perspect. Med. *4*. https://doi.org/10.1101/cshperspect.a024935.

Butler, E.K., Gersch, J., McNamara, A., Luk, K.-C., Holzmayer, V., de Medina, M., Schiff, E., Kuhns, M., and Cloherty, G.A. (2018). Hepatitis B Virus Serum DNA andRNA Levels in Nucleos(t)ide Analog-Treated or Untreated Patients During Chronic and Acute Infection. Hepatol. Baltim. Md *68*, 2106–2117. https://doi.org/10.1002/hep.30082.

Cai, Y.N., Zhou, Q., Kong, Y.Y., Li, M., Viollet, B., Xie, Y.H., and Wang, Y. (2003). LRH-1/hB1F and HNF1 synergistically up-regulate hepatitis B virus gene transcription and DNA replication. Cell Res. *13*, 451–458. https://doi.org/10.1038/sj.cr.7290187.

Candotti, D., and Allain, J.-P. (2017). Biological and clinical significance of hepatitis B virus RNA splicing: an update. Ann. Blood *2*. .

Cao, F., Jones, S., Li, W., Cheng, X., Hu, Y., Hu, J., and Tavis, J.E. (2014). Sequences in the terminal protein and reverse transcriptase domains of the hepatitis B virus polymerase contribute to RNA binding and encapsidation. J. Viral Hepat. *21*, 882–893. https://doi.org/10.1111/jvh.12225.

Carey, I., Gersch, J., Wang, B., Moigboi, C., Kuhns, M., Cloherty, G., Dusheiko, G., and Agarwal, K. (2019). Pre-genomic HBV RNA and HBcrAg predict outcomes in HBeAg negative chronic hepatitis B patients suppressed on nucleos(t)ide analogue therapy. Hepatol. Baltim. Md https://doi.org/10.1002/hep.31026.

Carman, W.F., Thursz, M., Hadziyannis, S., McIntyre, G., Colman, K., Gioustoz, A., Fattovich, G., Alberti, A., and Thomas, H.C. (1995). Hepatitis B e antigen negative chronic active hepatitis: hepatitis B virus core mutations occur predominantly in known antigenic determinants. J. Viral Hepat. *2*, 77–84. https://doi.org/10.1111/j.1365-2893.1995.tb00010.x.

Caviglia, G.P., Abate, M.L., Tandoi, F., Ciancio, A., Amoroso, A., Salizzoni, M., Saracco, G.M., Rizzetto, M., Romagnoli, R., and Smedile, A. (2018). Quantitation of HBV cccDNA in anti-HBc-positive liver donors by droplet digital PCR: A new tool to detect occult infection. J. Hepatol. *69*, 301–307. https://doi.org/10.1016/j.jhep.2018.03.021.

Centers for Disease Control and Prevention (CDC) (2000). Update: expanded availability of thimerosal preservative-free hepatitis B vaccine. MMWR Morb. Mortal. Wkly. Rep. *49*, 642, 651. .

Chai, N., Chang, H.E., Nicolas, E., Han, Z., Jarnik, M., and Taylor, J. (2008). Properties of subviral particles of hepatitis B virus. J. Virol. *82*, 7812–7817. https://doi.org/10.1128/JVI.00561-08.

Chain, B.M., and Myers, R. (2005). Variability and conservation in hepatitis B virus core protein. BMC Microbiol. *5*, 33. https://doi.org/10.1186/1471-2180-5-33.

Chakraborty, P.R., Ruiz-Opazo, N., Shouval, D., and Shafritz, D.A. (1980). Identification of integrated hepatitis B virus DNA and expression of viral RNA in an HBsAg-producing human hepatocellular carcinoma cell line. Nature *286*, 531–533. https://doi.org/10.1038/286531a0.

Chami, M., Gozuacik, D., Saigo, K., Capiod, T., Falson, P., Lecoeur, H., Urashima, T., Beckmann, J., Gougeon, M.L., Claret, M., et al. (2000). Hepatitis B virus-related insertional mutagenesis implicates SERCA1 gene in the control of apoptosis. Oncogene *19*, 2877–2886. https://doi.org/10.1038/sj.onc.1203605.

Chan, D.W., and Ng, I.O. (2006). Knock-down of hepatitis B virus X protein reduces the tumorigenicity of hepatocellular carcinoma cells. J. Pathol. *208*, 372–380. https://doi.org/10.1002/path.1901.

Chan, H.L.-Y., Wong, G.L.-H., Tse, C.-H., Chim, A.M.-L., Yiu, K.K.-L., Chan, H.-Y., Sung, J.J.-Y., and Wong, V.W.-S. (2009). Hepatitis B virus genotype C is associated with more severe liver fibrosis than genotype B. Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc. *7*, 1361–1366. https://doi.org/10.1016/j.cgh.2009.08.004.

Chang, M.S., and Nguyen, M.H. (2017). Epidemiology of hepatitis B and the role of vaccination. Best Pract. Res. Clin. Gastroenterol. *31*, 239–247. https://doi.org/10.1016/j.bpg.2017.05.008.

Chang, J., Guo, F., Zhao, X., and Guo, J.-T. (2014). Therapeutic strategies for a functional cure of chronic hepatitis B virus infection. Acta Pharm. Sin. B *4*, 248–257. https://doi.org/10.1016/j.apsb.2014.05.002.

Charre, C., Levrero, M., Zoulim, F., and Scholtès, C. (2019). Non-invasive biomarkers for chronic hepatitis B virus infection management. Antiviral Res. *169*, 104553. https://doi.org/10.1016/j.antiviral.2019.104553.

Chauhan, R., Churchill, N.D., Mulrooney-Cousins, P.M., and Michalak, T.I. (2017). Initial sites of hepadnavirus integration into host genome in human hepatocytes and in the woodchuck model of hepatitis B-associated hepatocellular carcinoma. Oncogenesis *6*, e317–e317. https://doi.org/10.1038/oncsis.2017.22.

Chemin, I., and Trépo, C. (2005). Clinical impact of occult HBV infections. J. Clin. Virol. *34*, S15–S21. https://doi.org/10.1016/S1386-6532(05)80005-8.

Chen, A., and Brown, C. (2012). Distinct families of cis-acting RNA replication elements epsilon from hepatitis B viruses. RNA Biol. *9*, 130–136. https://doi.org/10.4161/rna.18649.

Chen, C.-C., Guan, G., Qi, X., Abulaiti, A., Zhang, T., Liu, J., Lu, F., and Chen, X. (2021). Pacbio Sequencing of PLC/PRF/5 Cell Line and Clearance of HBV Integration Through CRISPR/Cas-9 System. Front. Mol. Biosci. *8*, 676957. https://doi.org/10.3389/fmolb.2021.676957.

Chen, E.-Q., Wang, M.-L., Tao, Y.-C., Wu, D.-B., Liao, J., He, M., and Tang, H. (2019). Serum HBcrAg is better than HBV RNA and HBsAg in reflecting intrahepatic covalently closed circular DNA. J. Viral Hepat. *26*, 586–595. https://doi.org/10.1111/jvh.13061.

Chen, I.H., Huang, C.J., and Ting, L.P. (1995). Overlapping initiator and TATA box functions in the basal core promoter of hepatitis B virus. J. Virol. *69*, 3647–3657. .

Chen, J., Wu, M., Wang, F., Zhang, W., Wang, W., Zhang, X., Zhang, J., Liu, Y., Liu, Y., Feng, Y., et al. (2015). Hepatitis B virus spliced variants are associated with an impaired response to interferon therapy. Sci. Rep. *5*, 16459. https://doi.org/10.1038/srep16459.

Chen, P., Gan, Y., Han, N., Fang, W., Li, J., Zhao, F., Hu, K., and Rayner, S. (2013). Computational evolutionary analysis of the overlapped surface (S) and polymerase (P) region in hepatitis B virus indicates the spacer domain in P is crucial for survival. PloS One *8*, e60098. https://doi.org/10.1371/journal.pone.0060098.

Chen, W.-N., Chen, J.-Y., Jiao, B.-Y., Lin, W.-S., Wu, Y.-L., Liu, L.-L., and Lin, X. (2012). Interaction of the hepatitis B spliced protein with cathepsin B promotes hepatoma cell migration and invasion. J. Virol. *86*, 13533–13541. https://doi.org/10.1128/JVI.02095-12.

Ching, R.H.H., Sze, K.M.F., Lau, E.Y.T., Chiu, Y.-T., Lee, J.M.F., Ng, I.O.L., and Lee, T.K.W. (2017). C-terminal truncated hepatitis B virus X protein regulates tumorigenicity, self-renewal and drug resistance via STAT3/Nanog signaling pathway. Oncotarget *8*, 23507–23516. https://doi.org/10.18632/oncotarget.15183.

Chiu, Y.-T., Wong, J.K.L., Choi, S.-W., Sze, K.M.F., Ho, D.W.H., Chan, L.-K., Lee, J.M.F., Man, K., Cherny, S., Yang, W.-L., et al. (2016). Novel pre-mRNA splicing of intronically integrated HBV generates oncogenic chimera in hepatocellular carcinoma. J. Hepatol. *64*, 1256–1264. https://doi.org/10.1016/j.jhep.2016.02.005.

Chong, C.K., Cheng, C.Y.S., Tsoi, S.Y.J., Huang, F.-Y., Liu, F., Seto, W.-K., Lai, C.-L., Yuen, M.-F., and Wong, D.K.-H. (2017). Role of hepatitis B core protein in HBV transcription and recruitment of histone acetyltransferases to cccDNA minichromosome. Antiviral Res. *144*, 1–7. https://doi.org/10.1016/j.antiviral.2017.05.003.

Chou, C.K., Wang, L.H., Lin, H.M., and Chi, C.W. (1992). Glucocorticoid stimulates hepatitis B viral gene expression in cultured human hepatoma cells. Hepatol. Baltim. Md *16*, 13–18. https://doi.org/10.1002/hep.1840160104.

Chu, C.-M., and Liaw, Y.-F. (2010). Hepatitis B surface antigen seroclearance during chronic HBV infection. Antivir. Ther. *15*, 133–143. https://doi.org/10.3851/IMP1497.

Chung, R.T., King, W.C., Ghany, M.G., Lisker-Melman, M., Hinerman, A.S., Khalili, M., Sulkowski, M., Jain, M.K., Choi, E.-Y.K., Nalesnik, M.A., et al. (2021). A Prospective Cohort Study of Novel Markers of Hepatitis B Virus Replication in Human Immunodeficiency Virus Coinfection. Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc. S1542-3565(21)01362-8. https://doi.org/10.1016/j.cgh.2021.12.038.

Clark, D.N., Flanagan, J.M., and Hu, J. (2017). Mapping of Functional Subdomains in the Terminal Protein Domain of Hepatitis B Virus Polymerase. J. Virol. *91*. https://doi.org/10.1128/JVI.01785-16.

Coffin, C.S., Zhou, K., and Terrault, N.A. (2019). New and Old Biomarkers for Diagnosis and Management of Chronic Hepatitis B Virus Infection. Gastroenterology *156*, 355-368.e3. https://doi.org/10.1053/j.gastro.2018.11.037.

Cohen, D., Ghosh, S., Shimakawa, Y., Ramou, N., Garcia, P.S., Dubois, A., Guillot, C., Kakwata-Nkor Deluce, N., Tilloy, V., Durand, G., et al. (2020). Hepatitis B virus preS2Δ38-55 variants: A newly identified risk factor for hepatocellular carcinoma. JHEP Rep. Innov. Hepatol. *2*, 100144. https://doi.org/10.1016/j.jhepr.2020.100144.

Cooksley, W.G.E., Piratvisuth, T., Lee, S.-D., Mahachai, V., Chao, Y.-C., Tanwandee, T., Chutaputti, A., Chang, W.Y., Zahm, F.E., and Pluck, N. (2003). Peginterferon alpha-2a (40 kDa): an advance in the treatment of hepatitis B e antigen-positive chronic hepatitis B. J. Viral Hepat. *10*, 298–305. https://doi.org/10.1046/j.1365-2893.2003.00450.x.

Cooper, A., and Shaul, Y. (2006). Clathrin-mediated Endocytosis and Lysosomal Cleavage of Hepatitis B Virus Capsid-like Core Particles *. J. Biol. Chem. *281*, 16563–16569. https://doi.org/10.1074/jbc.M601418200.

Coppola, N., Onorato, L., Minichini, C., Di Caprio, G., Starace, M., Sagnelli, C., and Sagnelli, E. (2015). Clinical significance of hepatitis B surface antigen mutants. World J. Hepatol. *7*, 2729–2739. https://doi.org/10.4254/wjh.v7.i27.2729.

Cornberg, M., Lok, A.S.-F., Terrault, N.A., Zoulim, F., and 2019 EASL-AASLD HBV Treatment Endpoints Conference Faculty (2020). Guidance for design and endpoints of clinical trials in chronic hepatitis B - Report from the 2019 EASL-AASLD HBV Treatment Endpoints Conference‡. J. Hepatol. *72*, 539–557. https://doi.org/10.1016/j.jhep.2019.11.003.

Croagh, C.M., Desmond, P.V., and Bell, S.J. (2015). Genotypes and viral variants in chronic hepatitis B: A review of epidemiology and clinical relevance. World J. Hepatol. *7*, 289–303. https://doi.org/10.4254/wjh.v7.i3.289.

Crump, N.T., and Milne, T.A. (2019). Why are so many MLL lysine methyltransferases required for normal mammalian development? Cell. Mol. Life Sci. *76*, 2885–2898. https://doi.org/10.1007/s00018-019-03143-z.

Cui, X., McAllister, R., Boregowda, R., Sohn, J.A., Cortes Ledesma, F., Caldecott, K.W., Seeger, C., and Hu, J. (2015). Does Tyrosyl DNA Phosphodiesterase-2 Play a Role in Hepatitis B Virus Genome Repair? PloS One *10*, e0128401. https://doi.org/10.1371/journal.pone.0128401.

Dandri, M., Burda, M.R., Bürkle, A., Zuckerman, D.M., Will, H., Rogler, C.E., Greten, H., and Petersen, J. (2002). Increase in de novo HBV DNA integrations in response to oxidative DNA damage or inhibition of poly(ADP-ribosyl)ation. Hepatology *35*, 217–223. https://doi.org/10.1053/jhep.2002.30203.

Dane, D.S., Cameron, C.H., and Briggs, M. (1970). Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. Lancet Lond. Engl. *1*, 695–698. https://doi.org/10.1016/s0140-6736(70)90926-8.

De Clercq, E. (1999). Perspectives for the treatment of hepatitis B virus infections. Int. J. Antimicrob. Agents *12*, 81–95. https://doi.org/10.1016/S0924-8579(99)00060-6.

Decorsière, A., Mueller, H., van Breugel, P.C., Abdul, F., Gerossier, L., Beran, R.K., Livingston, C.M., Niu, C., Fletcher, S.P., Hantz, O., et al. (2016). Hepatitis B virus X protein identifies the Smc5/6 complex as a host restriction factor. Nature *531*, 386–389. https://doi.org/10.1038/nature17170.

Dejean, A., Brechot, C., Tiollais, P., and Wain-Hobson, S. (1983). Characterization of integrated hepatitis B viral DNA cloned from a human hepatoma and the hepatoma-derived cell line PLC/PRF/5. Proc. Natl. Acad. Sci. U. S. A. *80*, 2505–2509. .

Delaney, W.E., Yang, H., Westland, C.E., Das, K., Arnold, E., Gibbs, C.S., Miller, M.D., and Xiong, S. (2003). The Hepatitis B Virus Polymerase Mutation rtV173L Is Selected during Lamivudine Therapy and Enhances Viral Replication In Vitro. J. Virol. *77*, 11833–11841. https://doi.org/10.1128/JVI.77.21.11833-11841.2003.

Désiré, N., Ngo, Y., Franetich, J.-F., Dembele, L., Mazier, D., Vaillant, J.-C., Poynard, T., and Thibault, V. (2015). Definition of an HBsAg to DNA international unit conversion factor by enrichment of circulating hepatitis B virus forms. J. Viral Hepat. *22*, 718–726. https://doi.org/10.1111/jvh.12387.

Dong, H., Zhang, L., Qian, Z., Zhu, X., Zhu, G., Chen, Y., Xie, X., Ye, Q., Zang, J., Ren, Z., et al. (2015). Identification of HBV-MLL4 Integration and Its Molecular Basis in Chinese Hepatocellular Carcinoma. PLOS ONE *10*, e0123175. https://doi.org/10.1371/journal.pone.0123175.

Dougherty, A.M., Guo, H., Westby, G., Liu, Y., Simsek, E., Guo, J.-T., Mehta, A., Norton, P., Gu, B., Block, T., et al. (2007). A substituted tetrahydro-tetrazolo-pyrimidine is a specific and novel inhibitor of hepatitis B virus surface antigen secretion. Antimicrob. Agents Chemother. *51*, 4427–4437. https://doi.org/10.1128/AAC.00541-07.

Dreyfuss, G., Matunis, M.J., Piñol-Roma, S., and Burd, C.G. (1993). hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochem. *62*, 289–321. https://doi.org/10.1146/annurev.bi.62.070193.001445.

Dunsford, H.A., Sell, S., and Chisari, F.V. (1990). Hepatocarcinogenesis due to chronic liver cell injury in hepatitis B virus transgenic mice. Cancer Res. *50*, 3400–3407. .

Edman, J.C., Gray, P., Valenzuela, P., Rall, L.B., and Rutter, W.J. (1980). Integration of hepatitis B virus sequences and their expression in a human hepatoma cell. Nature *286*, 535–538. https://doi.org/10.1038/286535a0.

Enomoto, H., Aizawa, N., Nishikawa, H., Ikeda, N., Sakai, Y., Takata, R., Hasegawa, K., Nakano, C., Nishimura, T., Yoh, K., et al. (2016). Relationship Between Hepatic Steatosis and the Elevation of Aminotransferases in HBV-Infected Patients With HBe-Antigen Negativity and a Low Viral Load. Medicine (Baltimore) *95*, e3565. https://doi.org/10.1097/MD.0000000000003565.

Erhardt, A., Blondin, D., Hauck, K., Sagir, A., Kohnle, T., Heintges, T., and Häussinger, D. (2005). Response to interferon alfa is hepatitis B virus genotype dependent: genotype A is more sensitive to interferon than genotype D. Gut *54*, 1009–1013. https://doi.org/10.1136/gut.2004.060327.

European Association for the Study of the Liver (2012). EASL Clinical Practice Guidelines: Management of chronic hepatitis B virus infection. J. Hepatol. *57*, 167–185. https://doi.org/10.1016/j.jhep.2012.02.010.

Fan, R., Peng, J., Xie, Q., Tan, D., Xu, M., Niu, J., Wang, H., Ren, H., Chen, X., Wang, M., et al. (2020). Combining HBV RNA and hepatitis B core-related antigen: guidance for safely stopping nucleos(t)ide analogues in HBeAg-positive chronic hepatitis B patients. J. Infect. Dis. https://doi.org/10.1093/infdis/jiaa136.

Fanning, G.C., Zoulim, F., Hou, J., and Bertoletti, A. (2019). Therapeutic strategies for hepatitis B virus infection: towards a cure. Nat. Rev. Drug Discov. *18*, 827–844. https://doi.org/10.1038/s41573-019-0037-0.

Farag, M.S., van Campenhout, M.J.H., Pfefferkorn, M., Fischer, J., Deichsel, D., Boonstra, A., van Vuuren, A.J., Ferenci, P., Feld, J.J., Berg, T., et al. (2020). Hepatitis B virus RNA as Early Predictor for Response to PEGylated Interferon Alfa in HBeAg Negative Chronic Hepatitis B. Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. https://doi.org/10.1093/cid/ciaa013.

Feng, H., Chen, P., Zhao, F., Nassal, M., and Hu, K. (2013). Evidence for Multiple Distinct Interactions between Hepatitis B Virus P Protein and Its Cognate RNA Encapsidation Signal during Initiation of Reverse Transcription. PLOS ONE *8*, e72798. https://doi.org/10.1371/journal.pone.0072798.

Ferrari, C., Missale, G., Boni, C., and Urbani, S. (2003). Immunopathogenesis of hepatitis B. J. Hepatol. *39 Suppl 1*, S36-42. https://doi.org/10.1016/s0168-8278(03)00137-5.

Fletcher, G.J., Eapen, C.E., and Abraham, P. (2019). Hepatitis B genotyping: The utility for the clinicians. Indian J. Gastroenterol. https://doi.org/10.1007/s12664-019-00995-y.

Fowler, M.J., Greenfield, C., Chu, C.M., Karayiannis, P., Dunk, A., Lok, A.S., Lai, C.L., Yeoh, E.K., Monjardino, J.P., and Wankya, B.M. (1986). Integration of HBV-DNA may not be a prerequisite for the maintenance of the state of malignant transformation. An analysis of 110 liver biopsies. J. Hepatol. *2*, 218–229. https://doi.org/10.1016/s0168-8278(86)80080-0.

Freitas, N., Cunha, C., Menne, S., and Gudima, S.O. (2014). Envelope proteins derived from naturally integrated hepatitis B virus DNA support assembly and release of infectious hepatitis delta virus particles. J. Virol. *88*, 5742–5754. https://doi.org/10.1128/JVI.00430-14.

Fryer, J.F., Heath, A.B., Wilkinson, D.E., Minor, P.D., and Collaborative Study Group (2017). A collaborative study to establish the 3rd WHO International Standard for hepatitis B virus for nucleic acid amplification techniques. Biol. J. Int. Assoc. Biol. Stand. *46*, 57–63. https://doi.org/10.1016/j.biologicals.2016.12.003.

Fukai, K., Takada, S., Yokosuka, O., Saisho, H., Omata, M., and Koike, K. (1997). Characterization of a specific region in the hepatitis B virus enhancer I for the efficient expression of X gene in the hepatic cell. Virology *236*, 279–287. https://doi.org/10.1006/viro.1997.8750.

Gallucci, L., and Kann, M. (2017). Nuclear Import of Hepatitis B Virus Capsids and Genome. Viruses *9*, 21. https://doi.org/10.3390/v9010021.

Gane, E.J., Kim, H.J., Visvanathan, K., Kim, Y.J., Nguyen, A.-H., Wallin, J.J., Chen, D.Y., McDonald, C., Arora, P., Tan, S.K., et al. (2021). Safety, Pharmacokinetics, and Pharmacodynamics of the Oral TLR8 Agonist Selgantolimod in Chronic Hepatitis B. Hepatol. Baltim. Md *74*, 1737–1749. https://doi.org/10.1002/hep.31795.

Ganem, D., and Prince, A.M. (2004). Hepatitis B virus infection--natural history and clinical consequences. N. Engl. J. Med. *350*, 1118–1129. https://doi.org/10.1056/NEJMra031087.

Geier, D.A., Kern, J.K., Hooker, B.S., King, P.G., Sykes, L.K., and Geier, M.R. (2014). Thimerosal-Containing Hepatitis B Vaccination and the Risk for Diagnosed Specific Delays in Development in the United States: A Case-Control Study in the Vaccine Safety Datalink. North Am. J. Med. Sci. *6*, 519–531. https://doi.org/10.4103/1947-2714.143284.

Gentile, I., and Borgia, G. (2014). Vertical transmission of hepatitis B virus: challenges and solutions. Int. J. Womens Health *6*, 605–611. https://doi.org/10.2147/IJWH.S51138.

Gerin, J.L., Ford, E.C., and Purcell, R.H. (1975). Biochemical characterization of Australia antigen. Evidence for defective particles of hepatitis B virus. Am. J. Pathol. *81*, 651–668. .

Ghany, M.G., King, W.C., Lisker-Melman, M., Lok, A.S.F., Terrault, N., Janssen, H.L.A., Khalili, M., Chung, R.T., Lee, W.M., Lau, D.T.Y., et al. (2021). Comparison of HBV RNA and Hepatitis B Core Related Antigen With Conventional HBV Markers Among Untreated Adults With Chronic Hepatitis B in North America. Hepatol. Baltim. Md *74*, 2395–2409. https://doi.org/10.1002/hep.32018.

Glebe, D., and Urban, S. (2007). Viral and cellular determinants involved in hepadnaviral entry. World J. Gastroenterol. *13*, 22–38. https://doi.org/10.3748/wjg.v13.i1.22.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett. *582*, 1977–1986. https://doi.org/10.1016/j.febslet.2008.03.004.

Goldsmith, C., Cohen, D., Dubois, A., Martinez, M.G., Petitjean, K., Corlu, A., Testoni, B., Hernandez-Vargas, H., and Chemin, I. (2021). Cas9-targeted nanopore sequencing reveals epigenetic heterogeneity after de novo assembly of native full-length hepatitis B virus genomes. Microb. Genomics *7*. https://doi.org/10.1099/mgen.0.000507.

Golsaz Shirazi, F., Mohammadi, H., Amiri, M.M., Singethan, K., Xia, Y., Bayat, A.A., Bahadori, M., Rabbani, H., Jeddi-Tehrani, M., Protzer, U., et al. (2014). Monoclonal antibodies to various epitopes of hepatitis B surface antigen inhibit hepatitis B virus infection. J. Gastroenterol. Hepatol. *29*, 1083–1091. https://doi.org/10.1111/jgh.12483.

Grandjacques, C., Pradat, P., Stuyver, L., Chevallier, M., Chevallier, P., Pichoud, C., Maisonnas, M., Trépo, C., and Zoulim, F. (2000). Rapid detection of genotypes and mutations in the pre-core promoter and the pre-core region of hepatitis B virus genome: correlation with viral persistence and disease severity. J. Hepatol. *33*, 430–439. https://doi.org/10.1016/s0168-8278(00)80279-2.

Grinnell, F. (1984). Fibronectin and wound healing. J. Cell. Biochem. *26*, 107–116. https://doi.org/10.1002/jcb.240260206.

Grover, D., Majumder, P.P., B Rao, C., Brahmachari, S.K., and Mukerji, M. (2003). Nonrandom distribution of alu elements in genes of various functional categories: insight from analysis of human chromosomes 21 and 22. Mol. Biol. Evol. *20*, 1420–1424. https://doi.org/10.1093/molbev/msg153.

Guerrieri, F., Belloni, L., D'Andrea, D., Pediconi, N., Le Pera, L., Testoni, B., Scisciani, C., Floriot, O., Zoulim, F., Tramontano, A., et al. (2017). Genome-wide identification of direct HBx genomic targets. BMC Genomics *18*, 184. https://doi.org/10.1186/s12864-017-3561-5.

Günther, S., Sommer, G., Iwanska, A., and Will, H. (1997). Heterogeneity and Common Features of Defective Hepatitis B Virus Genomes Derived from Spliced Pregenomic RNA. Virology *238*, 363–371. https://doi.org/10.1006/viro.1997.8863.

Günther, S., Sommer, G., Von Breunig, F., Iwanska, A., Kalinina, T., Sterneck, M., and Will, H. (1998). Amplification of Full-Length Hepatitis B Virus Genomes from Samples from Patients with Low Levels of Viremia: Frequency and Functional Consequences of PCR-Introduced Mutations. J. Clin. Microbiol. *36*, 531–538. https://doi.org/10.1128/JCM.36.2.531-538.1998.

Guo, H., Xu, C., Zhou, T., Block, T.M., and Guo, J.-T. (2012). Characterization of the host factors required for hepadnavirus covalently closed circular (ccc) DNA formation. PloS One *7*, e43270. https://doi.org/10.1371/journal.pone.0043270.

Guo, X., Chen, P., Hou, X., Xu, W., Wang, D., Wang, T., Zhang, L., Zheng, G., Gao, Z., He, C.-Y., et al. (2016). The recombined cccDNA produced using minicircle technology mimicked HBV genome in structure and function closely. Sci. Rep. *6*, 25552. https://doi.org/10.1038/srep25552.

Guo, Y., Lu, H., Xu, L., Idris, N.F.B., Li, Y., Hu, J., Huang, A., and TU, Z. (2019). The response of hepatitis B virus genotype to interferon is associated with a mutation in the interferon-stimulated response element. Medicine (Baltimore) *98*, e18442. https://doi.org/10.1097/MD.0000000000018442.

Ha, S.J., Chang, J., Song, M.K., Suh, Y.S., Jin, H.T., Lee, C.H., Nam, G.H., Choi, G., Choi, K.Y., Lee, S.H., et al. (2002). Engineering N-glycosylation mutations in IL-12 enhances sustained cytotoxic T lymphocyte responses for DNA immunization. Nat. Biotechnol. *20*, 381–386. https://doi.org/10.1038/nbt0402-381.

Hadziyannis, E., and Laras, A. (2018). Viral Biomarkers in Chronic HBeAg Negative HBV Infection. Genes *9*. https://doi.org/10.3390/genes9100469.

Hadziyannis, S.J., and Vassilopoulos, D. (2001). Hepatitis B e antigen–negative chronic hepatitis B. Hepatology *34*, 617–624. https://doi.org/10.1053/jhep.2001.27834.

Hao, N., Shen, W., Du, R., Jiang, S., Zhu, J., Chen, Y., Huang, C., Shi, Y., Xiang, R., and Luo, Y. (2020). Phosphodiesterase 3A Represents a Therapeutic Target that Drives Stem Cell–like Property and Metastasis in Breast Cancer. Mol. Cancer Ther. *19*, 868–881. https://doi.org/10.1158/1535-7163.MCT-18-1233.

Hatakeyama, T., Noguchi, C., Hiraga, N., Mori, N., Tsuge, M., Imamura, M., Takahashi, S., Kawakami, Y., Fujimoto, Y., Ochi, H., et al. (2007). Serum HBV RNA is a predictor of early emergence of the YMDD mutant in patients treated with lamivudine. Hepatol. Baltim. Md *45*, 1179–1186. https://doi.org/10.1002/hep.21581.

He, X., Wang, F., Huang, B., Chen, P., and Zhong, L. Detection and analysis of resistance mutations of hepatitis B virus. 10. .

He, Z., Zhu, J., Mo, J., Zhao, H., and Chen, Q. (2020). HBV DNA integrates into upregulated ZBTB20 in patients with hepatocellular carcinoma. Mol. Med. Rep. https://doi.org/10.3892/mmr.2020.11074.

Heather, J.M., and Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. Genomics *107*, 1–8. https://doi.org/10.1016/j.ygeno.2015.11.003.

Heermann, K.H., Goldmann, U., Schwartz, W., Seyffarth, T., Baumgarten, H., and Gerlich, W.H. (1984). Large surface proteins of hepatitis B virus containing the pre-s sequence. J. Virol. *52*, 396–402. .

Heermann, K.-H., Gerlich, W.H., Chudy, M., Schaefer, S., and Thomssen, R. (1999). Quantitative Detection of Hepatitis B Virus DNA in Two International Reference Plasma Preparations. J. Clin. Microbiol. *37*, 68–73. .

Heise, T., Sommer, G., Reumann, K., Meyer, I., Will, H., and Schaal, H. (2006). The hepatitis B virus PRE contains a splicing regulatory element. Nucleic Acids Res. *34*, 353–363. https://doi.org/10.1093/nar/gkj440.

Hernán, M.A., Jick, S.S., Olek, M.J., and Jick, H. (2004). Recombinant hepatitis B vaccine and the risk of multiple sclerosis: a prospective study. Neurology *63*, 838–842. https://doi.org/10.1212/01.wnl.0000138433.61870.82.

Herrscher, C., Pastor, F., Burlaud-Gaillard, J., Dumans, A., Seigneuret, F., Moreau, A., Patient, R., Eymieux, S., de Rocquigny, H., Hourioux, C., et al. (2020a). Hepatitis B virus entry into HepG2-NTCP cells requires clathrin-mediated endocytosis. Cell. Microbiol. e13205. https://doi.org/10.1111/cmi.13205.

Herrscher, C., Roingeard, P., and Blanchard, E. (2020b). Hepatitis B Virus Entry into Cells. Cells *9*, 1486. https://doi.org/10.3390/cells9061486.

Hilger, C., Velhagen, I., Zentgraf, H., and Schröder, C.H. (1991). Diversity of hepatitis B virus X gene-related transcripts in hepatocellular carcinoma: a novel polyadenylation site on viral DNA. J. Virol. https://doi.org/10.1128/jvi.65.8.4284-4291.1991.

Hirsch, R.C., Lavine, J.E., Chang, L.J., Varmus, H.E., and Ganem, D. (1990). Polymerase gene products of hepatitis B viruses are required for genomic RNA packaging as wel as for reverse transcription. Nature *344*, 552–555. https://doi.org/10.1038/344552a0.

Hirzel, C., Pfister, S., Gorgievski-Hrisoho, M., Wandeler, G., and Zuercher, S. (2015). Performance of HBsAg point-of-care tests for detection of diagnostic escape-variants in clinical samples. J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol. *69*, 33–35. https://doi.org/10.1016/j.jcv.2015.05.024.

Ho, D.W.H., Sze, K.M.F., and Ng, I.O.L. (2015). Virus-Clip: a fast and memory-efficient viral integration site detection tool at single-base resolution with annotation capability. Oncotarget *6*, 20959–20963. https://doi.org/10.18632/oncotarget.4187.

Hoshino, T., and Inagaki, F. (2012). Molecular quantification of environmental DNA using microfluidics and digital PCR. Syst. Appl. Microbiol. *35*, 390–395. https://doi.org/10.1016/j.syapm.2012.06.006.

Hsieh, Y.-H., Su, I.-J., Wang, H.-C., Chang, W.-W., Lei, H.-Y., Lai, M.-D., Chang, W.-T., and Huang, W. (2004). Pre-S mutant surface antigens in chronic hepatitis B virus infection induce oxidative stress and DNA damage. Carcinogenesis *25*, 2023–2032. https://doi.org/10.1093/carcin/bgh207.

Hsu, I.C., Tokiwa, T., Bennett, W., Metcalf, R.A., Welsh, J.A., Sun, T., and Harris, C.C. (1993). p53 gene mutation and integrated hepatitis B viral DNA sequences in human liver cancer cell lines. Carcinogenesis *14*, 987–992. https://doi.org/10.1093/carcin/14.5.987.

Hu, J., and Liu, K. (2017). Complete and Incomplete Hepatitis B Virus Particles: Formation, Function, and Application. Viruses *9*. https://doi.org/10.3390/v9030056.

Hu, J., and Seeger, C. (1996a). Hsp90 is required for the activity of a hepatitis B virus reverse transcriptase. Proc. Natl. Acad. Sci. U. S. A. *93*, 1060–1064. https://doi.org/10.1073/pnas.93.3.1060.

Hu, J., and Seeger, C. (1996b). Expression and characterization of hepadnavirus reverse transcriptases. Methods Enzymol. *275*, 195–208. https://doi.org/10.1016/s0076-6879(96)75013-9.

Hu, J., Toft, D., Anselmo, D., and Wang, X. (2002). In vitro reconstitution of functional hepadnavirus reverse transcriptase with cellular chaperone proteins. J. Virol. *76*, 269–279. https://doi.org/10.1128/jvi.76.1.269-279.2002.

Huang, C., Xie, M.-H., Liu, W., Yang, B., Yang, F., Huang, J., Huang, J., Wu, Q., Fu, X.-D., and Zhang, Y. (2011). A structured RNA in hepatitis B virus post-transcriptional regulatory element represses alternative splicing in a

sequence-independent and position-dependent manner. FEBS J. *278*, 1533–1546. https://doi.org/10.1111/j.1742-4658.2011.08077.x.

Huang, Y.-W., Chayama, K., Kao, J.-H., and Yang, S.-S. (2015). Detectability and clinical significance of serum hepatitis B virus ribonucleic acid. Hepatobiliary Surg. Nutr. *4*, 197–202. https://doi.org/10.3978/j.issn.2304-3881.2014.11.08.

Hui, C.-K., Leung, N., Yuen, S.-T., Zhang, H.-Y., Leung, K.-W., Lu, L., Cheung, S.K.F., Wong, W.-M., and Lau, G.K. (2007). Natural history and disease progression in Chinese chronic hepatitis B patients in immune-tolerant phase. Hepatology *46*, 395–401. https://doi.org/10.1002/hep.21724.

Iannacone, M., and Guidotti, L.G. (2021). Immunobiology and pathogenesis of hepatitis B virus infection. Nat. Rev. Immunol. 1–14. https://doi.org/10.1038/s41577-021-00549-4.

Inoue, T., and Tanaka, Y. (2019). The Role of Hepatitis B Core-Related Antigen. Genes *10*. https://doi.org/10.3390/genes10050357.

Invernizzi, F., Viganò, M., Grossi, G., and Lampertico, P. (2016). The prognosis and management of inactive HBV carriers. Liver Int. *36*, 100–104. https://doi.org/10.1111/liv.13006.

Ishii, T., Tamura, A., Shibata, T., Kuroda, K., Kanda, T., Sugiyama, M., Mizokami, M., and Moriyama, M. (2020). Analysis of HBV Genomes Integrated into the Genomes of Human Hepatoma PLC/PRF/5 Cells by HBV Sequence Capture-Based Next-Generation Sequencing. Genes *11*. https://doi.org/10.3390/genes11060661.

Iwamoto, M., Saso, W., Sugiyama, R., Ishii, K., Ohki, M., Nagamori, S., Suzuki, R., Aizaki, H., Ryo, A., Yun, J.-H., et al. (2019). Epidermal growth factor receptor is a host-entry cofactor triggering hepatitis B virus internalization. Proc. Natl. Acad. Sci. *116*, 8487–8492. https://doi.org/10.1073/pnas.1811064116.

Jain, M., Koren, S., Miga, K.H., Quick, J., Rand, A.C., Sasani, T.A., Tyson, J.R., Beggs, A.D., Dilthey, A.T., Fiddes, I.T., et al. (2018). Nanopore sequencing and assembly of a human genome with ultra-long reads. Nat. Biotechnol. *36*, 338–345. https://doi.org/10.1038/nbt.4060.

Jansen, L., Kootstra, N.A., van Dort, K.A., Takkenberg, R.B., Reesink, H.W., and Zaaijer, H.L. (2016). Hepatitis B Virus Pregenomic RNA Is Present in Virions in Plasma and Is Associated With a Response to Pegylated Interferon Alfa-2a and Nucleos(t)ide Analogues. J. Infect. Dis. *213*, 224–232. https://doi.org/10.1093/infdis/jiv397.

Janssen, H.L.A., van Zonneveld, M., Senturk, H., Zeuzem, S., Akarca, U.S., Cakaloglu, Y., Simon, C., So, T.M.K., Gerken, G., de Man, R.A., et al. (2005). Pegylated interferon alfa-2b alone or in combination with lamivudine for HBeAg-positive chronic hepatitis B: a randomised trial. Lancet Lond. Engl. *365*, 123–129. https://doi.org/10.1016/S0140-6736(05)17701-0.

Janssen, H.L.A., Brunetto, M.R., Kim, Y.J., Ferrari, C., Massetto, B., Nguyen, A.-H., Joshi, A., Woo, J., Lau, A.H., Gaggar, A., et al. (2018). Safety, efficacy and pharmacodynamics of vesatolimod (GS-9620) in virally suppressed patients with chronic hepatitis B. J. Hepatol. *68*, 431–440. https://doi.org/10.1016/j.jhep.2017.10.027.

Jefferies, M., Rauff, B., Rashid, H., Lam, T., and Rafiq, S. (2018). Update on global epidemiology of viral hepatitis and preventive strategies. World J. Clin. Cases *6*, 589–599. https://doi.org/10.12998/wjcc.v6.i13.589.

Jegal, M.-E., Jung, S.-Y., Han, Y.-S., and Kim, Y.-J. (2019). C-terminal truncated HBx reduces doxorubicin cytotoxicity via ABCB1 upregulation in Huh-7 hepatocellular carcinoma cells. BMB Rep. *52*, 330–335. .

Jeng, W.-J., and Lok, A.S. (2021). Should Treatment Indications for Chronic Hepatitis B Be Expanded? Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc. *19*, 2006–2014. https://doi.org/10.1016/j.cgh.2020.04.091.

Jenkins, A., Minhas, R., Morris, C., and Berry, N. (2017). Complete Genome Sequence of the WHO International Standard for Hepatitis B Virus DNA. Genome Announc. *5*. https://doi.org/10.1128/genomeA.01576-16.

Jeong, J.H., Kwak, D.S., Rho, H.M., and Jung, G. (1996). The catalytic properties of human hepatitis B virus polymerase. Biochem. Biophys. Res. Commun. *223*, 264–271. https://doi.org/10.1006/bbrc.1996.0882.

Jia, L., Wang, X.W., and Harris, C.C. (1999). Hepatitis B virus X protein inhibits nucleotide excision repair. Int. J. Cancer *80*, 875–879. https://doi.org/10.1002/(sici)1097-0215(19990315)80:6<875::aid-ijc13>3.0.co;2-z.

Jiang, B., and Hildt, E. (2020). Intracellular Trafficking of HBV Particles. Cells *9*, 2023. https://doi.org/10.3390/cells9092023.

Jiang, B., Himmelsbach, K., Ren, H., Boller, K., and Hildt, E. (2015). Subviral Hepatitis B Virus Filaments, like Infectious Viral Particles, Are Released via Multivesicular Bodies. J. Virol. *90*, 3330–3341. https://doi.org/10.1128/JVI.03109-15.

Jiang, S., Yang, Z., Li, W., Li, X., Wang, Y., Zhang, J., Xu, C., Chen, P.-J., Hou, J., McCrae, M.A., et al. (2012). Re-evaluation of the Carcinogenic Significance of Hepatitis B Virus Integration in Hepatocarcinogenesis. PLoS ONE *7*, e40363. https://doi.org/10.1371/journal.pone.0040363.

Jourdain, G., Ngo-Giang-Huong, N., and Khamduang, W. (2019). Current progress in the prevention of mother-to-child transmission of hepatitis B and resulting clinical and programmatic implications. Infect. Drug Resist. *12*, 977–987. https://doi.org/10.2147/IDR.S171695.

Julithe, R., Abou-Jaoudé, G., and Sureau, C. (2014). Modification of the hepatitis B virus envelope protein glycosylation pattern interferes with secretion of viral particles, infectivity, and susceptibility to neutralizing antibodies. J. Virol. *88*, 9049–9059. https://doi.org/10.1128/JVI.01161-14.

Kao, J.-H. (2008). Diagnosis of hepatitis B virus infection through serological and virological markers. Expert Rev. Gastroenterol. Hepatol. *2*, 553–562. https://doi.org/10.1586/17474124.2.4.553.

Kappus, M.R., and Sterling, R.K. (2013). Extrahepatic manifestations of acute hepatitis B virus infection. Gastroenterol. Hepatol. *9*, 123–126. .

Karayiannis, P., Novick, D.M., Lok, A.S., Fowler, M.J., Monjardino, J., and Thomas, H.C. (1985). Hepatitis B virus DNA in saliva, urine, and seminal fluid of carriers of hepatitis B e antigen. Br Med J Clin Res Ed *290*, 1853–1855. https://doi.org/10.1136/bmj.290.6485.1853.

Kasai, F., Hirayama, N., Ozawa, M., Satoh, M., and Kohara, A. (2018). HuH-7 reference genome profile: complex karyotype composed of massive loss of heterozygosity. Hum. Cell *31*, 261–267. https://doi.org/10.1007/s13577-018-0212-3.

Kay, A., and Zoulim, F. (2007). Hepatitis B virus genetic variability and evolution. Virus Res. *127*, 164–176. https://doi.org/10.1016/j.virusres.2007.02.021.

Kim, S., Wang, H., and Ryu, W.-S. (2010). Incorporation of eukaryotic translation initiation factor eIF4E into viral nucleocapsids via interaction with hepatitis B virus polymerase. J. Virol. *84*, 52–58. https://doi.org/10.1128/JVI.01232-09.

Kim, Y., Yoon, J.W., Xiao, X., Dean, N.M., Monia, B.P., and Marcusson, E.G. (2007). Selective Down-Regulation of Glioma-Associated Oncogene 2 Inhibits the Proliferation of Hepatocellular Carcinoma Cells. Cancer Res. *67*, 3583–3593. https://doi.org/10.1158/0008-5472.CAN-06-3040.

Kimbi, G.C., Kramvis, A., and Kew, M.C. (2005). Integration of hepatitis B virus DNA into chromosomal DNA during acute hepatitis B. World J. Gastroenterol. *11*, 6416–6421. https://doi.org/10.3748/wjg.v11.i41.6416.

Kimura, T., Rokuhara, A., Sakamoto, Y., Yagi, S., Tanaka, E., Kiyosawa, K., and Maki, N. (2002). Sensitive enzyme immunoassay for hepatitis B virus core-related antigens and their correlation to virus load. J. Clin. Microbiol. *40*, 439–445. https://doi.org/10.1128/jcm.40.2.439-445.2002.

Kiselinova, M., Pasternak, A.O., De Spiegelaere, W., Vogelaers, D., Berkhout, B., and Vandekerckhove, L. (2014). Comparison of Droplet Digital PCR and Seminested Real-Time PCR for Quantification of Cell-Associated HIV-1 RNA. PLoS ONE *9*, e85999. https://doi.org/10.1371/journal.pone.0085999.

Kitamura, K., Que, L., Shimadu, M., Koura, M., Ishihara, Y., Wakae, K., Nakamura, T., Watashi, K., Wakita, T., and Muramatsu, M. (2018). Flap endonuclease 1 is involved in cccDNA formation in the hepatitis B virus. PLoS Pathog. *14*, e1007124. https://doi.org/10.1371/journal.ppat.1007124.

Klymiuk, M.C., Balz, N., Elashry, M.I., Heimann, M., Wenisch, S., and Arnhold, S. (2019). Exosomes isolation and identification from equine mesenchymal stem cells. BMC Vet. Res. *15*, 42. https://doi.org/10.1186/s12917-019-1789-9.

Knaus, T., and Nassal, M. (1993). The encapsidation signal on the hepatitis B virus RNA pregenome forms a stem-loop structure that is critical for its function. Nucleic Acids Res. *21*, 3967–3975. https://doi.org/10.1093/nar/21.17.3967.

Knowles, B.B., Howe, C.C., and Aden, D.P. (1980). Human hepatocellular carcinoma cell lines secrete the major plasma proteins and hepatitis B surface antigen. Science *209*, 497–499. https://doi.org/10.1126/science.6248960.

Ko, C., Bester, R., Zhou, X., Xu, Z., Blossey, C., Sacherl, J., Vondran, F.W.R., Gao, L., and Protzer, U. (2019). A New Role for Capsid Assembly Modulators To Target Mature Hepatitis B Virus Capsids and Prevent Virus Infection. Antimicrob. Agents Chemother. *64*. https://doi.org/10.1128/AAC.01440-19.

Köck, J., Theilmann, L., Galle, P., and Schlicht, H.J. (1996). Hepatitis B virus nucleic acids associated with human peripheral blood mononuclear cells do not originate from replicating virus. Hepatol. Baltim. Md *23*, 405–413. https://doi.org/10.1002/hep.510230303.

Köck, J., Rösler, C., Zhang, J.-J., Blum, H.E., Nassal, M., and Thoma, C. (2010). Generation of Covalently Closed Circular DNA of Hepatitis B Viruses via Intracellular Recycling Is Regulated in a Virus Specific Manner. PLOS Pathog. *6*, e1001082. https://doi.org/10.1371/journal.ppat.1001082.

König, J., Zarnack, K., Luscombe, N.M., and Ule, J. (2012). Protein-RNA interactions: new genomic technologies and perspectives. Nat. Rev. Genet. *13*, 77–83. https://doi.org/10.1038/nrg3141.

Königer, C., Wingert, I., Marsmann, M., Rösler, C., Beck, J., and Nassal, M. (2014). Involvement of the host DNA-repair enzyme TDP2 in formation of the covalently closed circular DNA persistence reservoir of hepatitis B viruses. Proc. Natl. Acad. Sci. U. S. A. *111*, E4244-4253. https://doi.org/10.1073/pnas.1409986111.

Kono, N., and Arakawa, K. (2019). Nanopore sequencing: Review of potential applications in functional genomics. Dev. Growth Differ. *61*, 316–326. https://doi.org/10.1111/dgd.12608.

Korolowicz, K., Czerwinski, S., Iyer, R., Skell, J., Yang, J., Tucker, R., and Menne, S. ANTIVIRAL EFFICACY AND INDUCTION OF HOST IMMUNE RESPONSES WITH SB 9200, AN ORAL PRODRUG OF THE DINUCLEOTIDE SB 9000, IN THE WOODCHUCK MODEL OF CHRONIC HEPATITIS B VIRUS (HBV) INFECTION. 1. .

Kouwaki, T., Fukushima, Y., Daito, T., Sanada, T., Yamamoto, N., Mifsud, E.J., Leong, C.R., Tsukiyama-Kohara, K., Kohara, M., Matsumoto, M., et al. (2016). Extracellular Vesicles Including Exosomes Regulate Innate Immune Responses to Hepatitis B Virus Infection. Front. Immunol. *7*, 335. https://doi.org/10.3389/fimmu.2016.00335.

Kowdley, K.V., Wang, C.C., Welch, S., Roberts, H., and Brosgart, C.L. (2012). Prevalence of chronic hepatitis B among foreign-born persons living in the United States by country of origin. Hepatol. Baltim. Md *56*, 422–433. https://doi.org/10.1002/hep.24804.

Kremsdorf, D., Lekbaby, B., Bablon, P., Sotty, J., Augustin, J., Schnuriger, A., Pol, J., and Soussan, P. (2021). Alternative splicing of viral transcripts: the dark side of HBV. Gut https://doi.org/10.1136/gutjnl-2021-324554.

Kumar, V., Jayasuryan, N., and Kumar, R. (1996). A truncated mutant (residues 58-140) of the hepatitis B virus X protein retains transactivation function. Proc. Natl. Acad. Sci. U. S. A. *93*, 5647–5652. https://doi.org/10.1073/pnas.93.11.5647.

Ladner, S.K., Otto, M.J., Barker, C.S., Zaifert, K., Wang, G.H., Guo, J.T., Seeger, C., and King, R.W. (1997). Inducible expression of human hepatitis B virus (HBV) in stably transfected hepatoblastoma cells: a novel system for screening potential inhibitors of HBV replication. Antimicrob. Agents Chemother. *41*, 1715–1720. .

Lai, C.-L., Dienstag, J., Schiff, E., Leung, N.W.Y., Atkins, M., Hunt, C., Brown, N., Woessner, M., Boehme, R., and Condreay, L. (2003). Prevalence and clinical correlates of YMDD variants during lamivudine therapy for patients with chronic hepatitis B. Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. *36*, 687–696. https://doi.org/10.1086/368083.

Lai, M.-W., Liang, K.-H., Lin, W.-R., Huang, Y.-H., Huang, S.-F., Chen, T.-C., and Yeh, C.-T. (2016). Hepatocarcinogenesis in transgenic mice carrying hepatitis B virus pre-S/S gene with the sW172* mutation. Oncogenesis *5*, e273. https://doi.org/10.1038/oncsis.2016.77.

Lam, A.M., Ren, S., Espiritu, C., Kelly, M., Lau, V., Zheng, L., Hartman, G.D., Flores, O.A., and Klumpp, K. (2017a). Hepatitis B Virus Capsid Assembly Modulators, but Not Nucleoside Analogs, Inhibit the Production of Extracellular Pregenomic RNA and Spliced RNA Variants. Antimicrob. Agents Chemother. *61*. https://doi.org/10.1128/AAC.00680-17.

Lam, Y.-F., Seto, W.-K., Wong, D., Cheung, K.-S., Fung, J., Mak, L.-Y., Yuen, J., Chong, C.-K., Lai, C.-L., and Yuen, M.-F. (2017b). Seven-Year Treatment Outcome of Entecavir in a Real-World Cohort: Effects on Clinical Parameters, HBsAg and HBcrAg Levels. Clin. Transl. Gastroenterol. *8*, e125. https://doi.org/10.1038/ctg.2017.51.

Lambert, C., Döring, T., and Prange, R. (2007). Hepatitis B virus maturation is sensitive to functional inhibition of ESCRT-III, Vps4, and gamma 2-adaptin. J. Virol. *81*, 9050–9060. https://doi.org/10.1128/JVI.00479-07.

Lampertico, P., Agarwal, K., Berg, T., Buti, M., Janssen, H.L.A., Papatheodoridis, G., Zoulim, F., and Tacke, F. (2017). EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. J. Hepatol. *67*, 370–398. https://doi.org/10.1016/j.jhep.2017.03.021.

Landers, T.A., Greenberg, H.B., and Robinson, W.S. (1977). Structure of hepatitis B Dane particle DNA and nature of the endogenous DNA polymerase reaction. J. Virol. *23*, 368–376. .

Lanford, R.E., Notvall, L., Lee, H., and Beames, B. (1997). Transcomplementation of nucleotide priming and reverse transcription between independently expressed TP and RT domains of the hepatitis B virus reverse transcriptase. J. Virol. *71*, 2996–3004. .

Larsson, S.B., Tripodi, G., Raimondo, G., Saitta, C., Norkrans, G., Pollicino, T., and Lindh, M. (2018). Integration of hepatitis B virus DNA in chronically infected patients assessed by Alu-PCR. J. Med. Virol. *90*, 1568–1575. https://doi.org/10.1002/jmv.25227.

Lau, C.-C., Sun, T., Ching, A.K.K., He, M., Li, J.-W., Wong, A.M., Co, N.N., Chan, A.W.H., Li, P.-S., Lung, R.W.M., et al. (2014). Viral-human chimeric transcript predisposes risk to liver cancer development and progression. Cancer Cell *25*, 335–349. https://doi.org/10.1016/j.ccr.2014.01.030.

Lavanchy, D. (2004). Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. J. Viral Hepat. *11*, 97–107. https://doi.org/10.1046/j.1365-2893.2003.00487.x.

Laver, T., Harrison, J., O'Neill, P.A., Moore, K., Farbos, A., Paszkiewicz, K., and Studholme, D.J. (2015). Assessing the performance of the Oxford Nanopore Technologies MinION. Biomol. Detect. Quantif. *3*, 1–8. https://doi.org/10.1016/j.bdq.2015.02.001.

Lee, G.H., Wasser, S., and Lim, S.G. (2008). Hepatitis B pregenomic RNA splicing—The products, the regulatory mechanisms and its biological significance. Virus Res. *136*, 1–7. https://doi.org/10.1016/j.virusres.2008.05.007.

Lee, J., Shin, M.-K., Lee, H.-J., Yoon, G., and Ryu, W.-S. (2004). Three Novel cis-Acting Elements Required for Efficient Plus-Strand DNA Synthesis of the Hepatitis B Virus Genome. J. Virol. *78*, 7455–7464. https://doi.org/10.1128/JVI.78.14.7455-7464.2004.

Lee, J.-H., Ku, J.-L., Park, Y.J., Lee, K.-U., Kim, W.-H., and Park, J.-G. (1999). Establishment and characterization of four human hepatocellular carcinoma cell lines containing hepatitis B virus DNA. World J. Gastroenterol. *5*, 289–295. https://doi.org/10.3748/wjg.v5.i4.289.

Lee, W.Y., Bachtiar, M., Choo, C.C.S., and Lee, C.G. (2019). Comprehensive review of Hepatitis B Virus-associated hepatocellular carcinoma research through text mining and big data analytics. Biol. Rev. *94*, 353–367. https://doi.org/10.1111/brv.12457.

Lefkowitz, E.J., Dempsey, D.M., Hendrickson, R.C., Orton, R.J., Siddell, S.G., and Smith, D.B. (2018). Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). Nucleic Acids Res. *46*, D708–D717. https://doi.org/10.1093/nar/gkx932.

Levrero, M., Pollicino, T., Petersen, J., Belloni, L., Raimondo, G., and Dandri, M. (2009). Control of cccDNA function in hepatitis B virus infection. J. Hepatol. *51*, 581–592. https://doi.org/10.1016/j.jhep.2009.05.022.

Li, A., Wu, J., Zhai, A., Qian, J., Wang, X., Qaria, M.A., Zhang, Q., Li, Y., Fang, Y., Kao, W., et al. (2019). HBV triggers APOBEC2 expression through miR-122 regulation and affects the proliferation of liver cancer cells. Int. J. Oncol. *55*, 1137–1148. https://doi.org/10.3892/ijo.2019.4870.

Li, B., Wang, Y., Shen, F., Wu, M., Li, Y., Fang, Z., Ye, J., Wang, L., Gao, L., Yuan, Z., et al. (2018a). Identification of Retinoic Acid Receptor Agonists as Potent Hepatitis B Virus Inhibitors via a Drug Repurposing Screen. Antimicrob. Agents Chemother. *62*. https://doi.org/10.1128/AAC.00465-18.

Li, C.H., Xu, F., Chow, S., Feng, L., Yin, D., Ng, T.B., and Chen, Y. (2014). Hepatitis B virus X protein promotes hepatocellular carcinoma transformation through interleukin-6 activation of microRNA-21 expression. Eur. J. Cancer Oxf. Engl. 1990 *50*, 2560–2569. https://doi.org/10.1016/j.ejca.2014.07.008.

Li, H., Bai, R., Zhao, Z., Tao, L., Ma, M., Ji, Z., Jian, M., Ding, Z., Dai, X., Bao, F., et al. (2018b). Application of droplet digital PCR to detect the pathogens of infectious diseases. Biosci. Rep. https://doi.org/10.1042/BSR20181170.

Li, H.-C., Huang, E.-Y., Su, P.-Y., Wu, S.-Y., Yang, C.-C., Lin, Y.-S., Chang, W.-C., and Shih, C. (2010). Nuclear Export and Import of Human Hepatitis B Virus Capsid Protein and Particles. PLoS Pathog. *6*. https://doi.org/10.1371/journal.ppat.1001162.

Li, J., Buckwold, V.E., Hon, M.W., and Ou, J.H. (1999). Mechanism of suppression of hepatitis B virus precore RNA transcription by a frequent double mutation. J. Virol. *73*, 1239–1244. .

Li, J., Bao, M., Ge, J., Ren, S., Zhou, T., Qi, F., Pu, X., and Dou, J. (2017). Research progress of therapeutic vaccines for treating chronic hepatitis B. Hum. Vaccines Immunother. *13*, 986–997. https://doi.org/10.1080/21645515.2016.1276125.

Li, W., Li, M., Liao, D., Lu, X., Gu, X., Zhang, Q., Zhang, Z., and Li, H. (2016). Carboxyl-terminal truncated HBx contributes to invasion and metastasis via deregulating metastasis suppressors in hepatocellular carcinoma. Oncotarget *7*, 55110–55127. https://doi.org/10.18632/oncotarget.10399.

Liang, T.J. (2009). Hepatitis B: the virus and disease. Hepatol. Baltim. Md *49*, S13-21. https://doi.org/10.1002/hep.22881.

Liang, H.-W., Wang, N., Wang, Y., Wang, F., Fu, Z., Yan, X., Zhu, H., Diao, W., Ding, Y., Chen, X., et al. (2016). Hepatitis B virus-human chimeric transcript HBx-LINE1 promotes hepatic injury via sequestering cellular microRNA-122. J. Hepatol. *64*, 278–291. https://doi.org/10.1016/j.jhep.2015.09.013.

Liang, Y., Qiu, K., Liao, B., Zhu, W., Huang, X., Li, L., Chen, X., and Li, K. (2017). Seeksv: an accurate tool for somatic structural variation and virus integration detection. Bioinforma. Oxf. Engl. *33*, 184–191. https://doi.org/10.1093/bioinformatics/btw591.

Liao, H., Liu, Y., Li, X., Wang, J., Chen, X., Zou, J., Li, Q., Liu, L., Wang, J., Huang, B., et al. (2019). Monitoring of serum HBV RNA, HBcrAg, HBsAg and anti-HBc levels in patients during long-term nucleoside/nucleotide analogue therapy. Antivir. Ther. *24*, 105–115. https://doi.org/10.3851/IMP3280.

Liaw, Y.-F. (2009). HBeAg seroconversion as an important end point in the treatment of chronic hepatitis B. Hepatol. Int. *3*, 425–433. https://doi.org/10.1007/s12072-009-9140-3.

Lim, Y.-S. (2017). Management of Antiviral Resistance in Chronic Hepatitis B. Gut Liver *11*, 189–195. https://doi.org/10.5009/gnl15562.

Limothai, U., Chuaypen, N., Poovorawan, K., Chotiyaputta, W., Tanwandee, T., Poovorawan, Y., and Tangkijvanich, P. (2019). Baseline and kinetics of serum hepatitis B virus RNA predict response to pegylated interferon-based therapy in patients with hepatitis B e antigen-negative chronic hepatitis B. J. Viral Hepat. *26*, 1481–1488. https://doi.org/10.1111/jvh.13195.

Limothai, U., Chuaypen, N., Poovorawan, K., Chotiyaputta, W., Tanwandee, T., Poovorawan, Y., and Tangkijvanich, P. (2020). Reverse transcriptase droplet digital PCR vs reverse transcriptase quantitative real-time PCR for serum HBV RNA quantification. J. Med. Virol. https://doi.org/10.1002/jmv.25792.

Lin, N., Ye, A., Lin, J., Liu, C., Huang, J., Fu, Y., Wu, S., Xu, S., Wang, L., and Ou, Q. (2020). Diagnostic Value of Detection of Pregenomic RNA in Sera of Hepatitis B Virus-Infected Patients with Different Clinical Outcomes. J. Clin. Microbiol. *58*. https://doi.org/10.1128/JCM.01275-19.

Liu, L. (2020). Clinical features of hepatocellular carcinoma with hepatitis B virus among patients on Nucleos(t) ide analog therapy. Infect. Agent. Cancer *15*, 8. https://doi.org/10.1186/s13027-020-0277-y.

Liu, C.-J., Chen, P.-J., Lai, M.-Y., Lin, F.-Y., Wang, T., Kao, J.-H., and Chen, D.-S. (2006a). Viral factors correlate with hepatitis B e antigen seroconverson in patients with chronic hepatitis B. Liver Int. *26*, 949–955. https://doi.org/10.1111/j.1478-3231.2006.01319.x.

Liu, C.-J., Lai, M.-Y., Chao, Y.-C., Liao, L.-Y., Yang, S.-S., Hsiao, T.-J., Hsieh, T.-Y., Lin, C.-L., Hu, J.-T., Chen, C.-L., et al. (2006b). Interferon alpha-2b with and without ribavirin in the treatment of hepatitis B e antigen-positive chronic hepatitis B: a randomized study. Hepatol. Baltim. Md *43*, 742–749. https://doi.org/10.1002/hep.21100.

Liu, F., Campagna, M., Qi, Y., Zhao, X., Guo, F., Xu, C., Li, S., Li, W., Block, T.M., Chang, J., et al. (2013). Alpha-interferon suppresses hepadnavirus transcription by altering epigenetic modification of cccDNA minichromosomes. PLoS Pathog. *9*, e1003613. https://doi.org/10.1371/journal.ppat.1003613.

Liu, J., Yang, H.-I., Lee, M.-H., Jen, C.-L., Batrla-Utermann, R., Lu, S.-N., Wang, L.-Y., You, S.-L., and Chen, C.-J. (2016). Serum Levels of Hepatitis B Surface Antigen and DNA Can Predict Inactive Carriers With Low Risk of Disease Progression. Hepatol. Baltim. Md *64*, 381–389. https://doi.org/10.1002/hep.28552.

Liu, S., Zhang, H., Gu, C., Yin, J., He, Y., Xie, J., and Cao, G. (2009). Associations Between Hepatitis B Virus Mutations and the Risk of Hepatocellular Carcinoma: A Meta-Analysis. JNCI J. Natl. Cancer Inst. *101*, 1066–1082. https://doi.org/10.1093/jnci/djp180.

Liu, S., Deng, R., Zhou, B., and Sun, J. (2020). Is World Health Organization HBV DNA Standard Appropriate for Standardizing Serum HBV RNA Assay? Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. https://doi.org/10.1093/cid/ciaa1503.

Liu, Y., Zhong, Y., Zou, Z., Xu, Z., Li, B., Ren, X., Bai, S., Wang, L., Li, X., Dai, J., et al. (2010). Features and clinical implications of hepatitis B virus genotypes and mutations in basal core promoter/precore region in 507 Chinese patients with acute and chronic hepatitis B. J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol. *47*, 243–247. https://doi.org/10.1016/j.jcv.2009.12.013.

Liu, Y., Jiang, M., Xue, J., Yan, H., and Liang, X. (2019). Serum HBV RNA quantification: useful for monitoring natural history of chronic hepatitis B infection. BMC Gastroenterol. *19*, 53. https://doi.org/10.1186/s12876-019-0966-4.

Lo, K.J., Tsai, Y.T., Lee, S.D., Wu, T.C., Wang, J.Y., Chen, G.H., Yeh, C.L., Chiang, B.N., Yeh, S.H., and Goudeau, A. (1985). Immunoprophylaxis of infection with hepatitis B virus in infants born to hepatitis B surface antigen-positive carrier mothers. J. Infect. Dis. *152*, 817–822. https://doi.org/10.1093/infdis/152.4.817.

Lockhart, A.C., Bauer, T.M., Aggarwal, C., Lee, C.B., Harvey, R.D., Cohen, R.B., Sedarati, F., Nip, T.K., Faessel, H., Dash, A.B., et al. (2019). Phase Ib study of pevonedistat, a NEDD8-activating enzyme inhibitor, in combination with docetaxel, carboplatin and paclitaxel, or gemcitabine, in patients with advanced solid tumors. Invest. New Drugs *37*, 87–97. https://doi.org/10.1007/s10637-018-0610-0.

Loeb, D.D., Hirsch, R.C., and Ganem, D. (1991). Sequence-independent RNA cleavages generate the primers for plus strand DNA synthesis in hepatitis B viruses: implications for other reverse transcribing elements. EMBO J. *10*, 3533–3540. .

Lohmann, V., Körner, F., Koch, J., Herian, U., Theilmann, L., and Bartenschlager, R. (1999). Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science *285*, 110–113. https://doi.org/10.1126/science.285.5424.110.

Lok, A.S.F. (2002). Chronic Hepatitis B. N. Engl. J. Med. *346*, 1682–1683. https://doi.org/10.1056/NEJM200205303462202.

Lok, A.S., Zoulim, F., Dusheiko, G., and Ghany, M.G. (2017). Hepatitis B cure: From discovery to regulatory approval. Hepatol. Baltim. Md *66*, 1296–1313. https://doi.org/10.1002/hep.29323.

Lu, H., Giordano, F., and Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. Genomics Proteomics Bioinformatics *14*, 265–279. https://doi.org/10.1016/j.gpb.2016.05.004.

Lu, J.G., Nguyen, L., Samadzadeh, S., Masouminia, M., Mendoza, A., Sweeney, O., Tillman, B., Afifyan, N., Morgan, T., French, B.A., et al. (2018). Expression of proteins upregulated in hepatocellular carcinoma in patients with alcoholic hepatitis (AH) compared to non-alcoholic steatohepatitis (NASH): An immunohistochemical analysis of candidate proteins. Exp. Mol. Pathol. *104*, 125–129. https://doi.org/10.1016/j.yexmp.2018.02.001.

Lucifora, J., Arzberger, S., Durantel, D., Belloni, L., Strubin, M., Levrero, M., Zoulim, F., Hantz, O., and Protzer, U. (2011). Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. J. Hepatol. *55*, 996–1003. https://doi.org/10.1016/j.jhep.2011.02.015.

Lucifora, J., Xia, Y., Reisinger, F., Zhang, K., Stadler, D., Cheng, X., Sprinzl, M.F., Koppensteiner, H., Makowska, Z., Volz, T., et al. (2014). Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA. Science *343*, 1221–1228. https://doi.org/10.1126/science.1243462.

Lucifora, J., Pastor, F., Charles, É., Pons, C., Auclair, H., Fusil, F., Rivoire, M., Cosset, F.-L., Durantel, D., and Salvetti, A. (2021). Evidence for long-term association of virion-delivered HBV core protein with cccDNA independently of viral protein production. JHEP Rep. *3*, 100330. https://doi.org/10.1016/j.jhepr.2021.100330.

Luo, H., Zhang, X.-X., Cao, L.-H., Tan, N., Kang, Q., Xi, H.-L., Yu, M., and Xu, X.-Y. (2019). Serum hepatitis B virus RNA is a predictor of HBeAg seroconversion and virological response with entecavir treatment in chronic hepatitis B patients. World J. Gastroenterol. *25*, 719–728. https://doi.org/10.3748/wjg.v25.i6.719.

Luo, H., Tan, N., Kang, Q., Pan, J., Chen, H., Xi, H., Yu, M., and Xu, X. (2020). Hepatitis B virus pregenomic RNA status can reveal the long-term prognoses of chronic hepatitis B patients treated with nucleos(t)ide analogues. J. Viral Hepat. *27*, 323–328. https://doi.org/10.1111/jvh.13227.

Luo, L., Chen, S., Gong, Q., Luo, N., Lei, Y., Guo, J., and He, S. (2013). Hepatitis B virus X protein modulates remodelling of minichromosomes related to hepatitis B virus replication in HepG2 cells. Int. J. Mol. Med. *31*, 197–204. https://doi.org/10.3892/ijmm.2012.1165.

Lupberger, J., Schaedler, S., Peiran, A., and Hildt, E. (2013). Identification and characterization of a novel bipartite nuclear localization signal in the hepatitis B virus polymerase. World J. Gastroenterol. *19*, 8000–8010. https://doi.org/10.3748/wjg.v19.i44.8000.

Ma, G., Lou, B., Lv, F., Zhao, D., Zhang, Z., and Chen, Y. (2020). HBcrAgandpgRNAand the therapeutic effect inHBeAg-positive patients receiving anti-viral therapy, baseline serum HBV-RNAis a powerful predictor of response. J. Viral Hepat. https://doi.org/10.1111/jvh.13299.

Ma, H., Yang, R.-F., Li, X.-H., Jin, Q., and Wei, L. (2016). HBcrAg Identifies Patients Failing to Achieve HBeAg Seroconversion Treated with Pegylated Interferon Alfa-2b. Chin. Med. J. (Engl.) *129*, 2212–2219. https://doi.org/10.4103/0366-6999.189904.

Ma, S., Chan, K.W., Lee, T.K.-W., Tang, K.H., Wo, J.Y.-H., Zheng, B.-J., and Guan, X.-Y. (2008). Aldehyde Dehydrogenase Discriminates the CD133 Liver Cancer Stem Cell Populations. Mol. Cancer Res. *6*, 1146–1153.

.

Maasoumy, B., Wiegand, S.B., Jaroszewicz, J., Bremer, B., Lehmann, P., Deterding, K., Taranta, A., Manns, M.P., Wedemeyer, H., Glebe, D., et al. (2015). Hepatitis B core-related antigen (HBcrAg) levels in the natural history of hepatitis B virus infection in a large European cohort predominantly infected with genotypes A and D. Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis. *21*, 606.e1-10. https://doi.org/10.1016/j.cmi.2015.02.010.

MacNab, G.M., Alexander, J.J., Lecatsas, G., Bey, E.M., and Urbanowicz, J.M. (1976). Hepatitis B surface antigen produced by a human hepatoma cell line. Br. J. Cancer *34*, 509–515. https://doi.org/10.1038/bjc.1976.205.

Macovei, A., Petrareanu, C., Lazar, C., Florian, P., and Branza-Nichita, N. (2013). Regulation of Hepatitis B Virus Infection by Rab5, Rab7, and the Endolysosomal Compartment. J. Virol. *87*, 6415–6427. https://doi.org/10.1128/JVI.00393-13.

Macrae, D.R., Bruss, V., and Ganem, D. (1991). Myristylation of a duck hepatitis B virus envelope protein is essential for infectivity but not for virus assembly. Virology *181*, 359–363. https://doi.org/10.1016/0042-6822(91)90503-4.

Malmström, S., Larsson, S.B., Hannoun, C., and Lindh, M. (2012). Hepatitis B viral DNA decline at loss of HBeAg is mainly explained by reduced cccDNA load--down-regulated transcription of PgRNA has limited impact. PloS One *7*, e36349. https://doi.org/10.1371/journal.pone.0036349.

Manzoor, S., Saalim, M., Imran, M., Resham, S., and Ashraf, J. (2015). Hepatitis B virus therapy: What's the future holding for us? World J. Gastroenterol. *21*, 12558–12575. https://doi.org/10.3748/wjg.v21.i44.12558.

Marcellin, P. (2016). Viral hepatitis: towards the eradication of HCV and a cure for HBV. Liver Int. Off. J. Int. Assoc. Study Liver *36 Suppl 1*, 5–6. https://doi.org/10.1111/liv.13035.

Marion, P.L., Salazar, F.H., Alexander, J.J., and Robinson, W.S. (1980). State of hepatitis B viral DNA in a human hepatoma cell line. J. Virol. *33*, 795–806. .

Martinez, M.G., Boyd, A., Combe, E., Testoni, B., and Zoulim, F. (2021). Covalently closed circular DNA: The ultimate therapeutic target for curing HBV infections. J. Hepatol. *75*, 706–717. https://doi.org/10.1016/j.jhep.2021.05.013.

Mason, A., Wick, M., White, H., and Perrillo, R. (1993). Hepatitis B virus replication in diverse cell types during chronic hepatitis B virus infection. Hepatol. Baltim. Md *18*, 781–789. https://doi.org/10.1002/hep.1840180406.

Mason, W.S., Low, H.-C., Xu, C., Aldrich, C.E., Scougall, C.A., Grosse, A., Clouston, A., Chavez, D., Litwin, S., Peri, S., et al. (2009). Detection of clonally expanded hepatocytes in chimpanzees with chronic hepatitis B virus infection. J. Virol. *83*, 8396–8408. https://doi.org/10.1128/JVI.00700-09.

Mason, W.S., Gill, U.S., Litwin, S., Zhou, Y., Peri, S., Pop, O., Hong, M.L.W., Naik, S., Quaglia, A., Bertoletti, A., et al. (2016). HBV DNA Integration and Clonal Hepatocyte Expansion in Chronic Hepatitis B Patients Considered Immune Tolerant. Gastroenterology *151*, 986-998.e4. https://doi.org/10.1053/j.gastro.2016.07.012.

Mason, W.S., Jilbert, A.R., and Litwin, S. (2021). Hepatitis B Virus DNA Integration and Clonal Expansion of Hepatocytes in the Chronically Infected Liver. Viruses *13*. https://doi.org/10.3390/v13020210.

Mast, E.E., Weinbaum, C.M., Fiore, A.E., Alter, M.J., Bell, B.P., Finelli, L., Rodewald, L.E., Douglas, J.M., Janssen, R.S., Ward, J.W., et al. (2006). A comprehensive immunization strategy to eliminate transmission of hepatitis B virus infection in the United States: recommendations of the Advisory Committee on Immunization Practices (ACIP) Part II: immunization of adults. MMWR Recomm. Rep. Morb. Mortal. Wkly. Rep. Recomm. Rep. *55*, 1–33; quiz CE1-4. .

Maxam, A.M., and Gilbert, W. (1977). A new method for sequencing DNA. Proc. Natl. Acad. Sci. *74*, 560–564. https://doi.org/10.1073/pnas.74.2.560.

McNaughton, A.L., Roberts, H.E., Bonsall, D., de Cesare, M., Mokaya, J., Lumley, S.F., Golubchik, T., Piazza, P., Martin, J.B., de Lara, C., et al. (2019). Illumina and Nanopore methods for whole genome sequencing of hepatitis B virus (HBV). Sci. Rep. *9*, 7081. https://doi.org/10.1038/s41598-019-43524-9.

Messageot, F., Salhi, S., Eon, P., and Rossignol, J.-M. (2003). Proteolytic processing of the hepatitis B virus e antigen precursor. Cleavage at two furin consensus sequences. J. Biol. Chem. *278*, 891–895. https://doi.org/10.1074/jbc.M207634200.

Michler, T., Kosinska, A.D., Festag, J., Bunse, T., Su, J., Ringelhan, M., Imhof, H., Grimm, D., Steiger, K., Mogler, C., et al. (2020). Knockdown of Virus Antigen Expression Increases Therapeutic Vaccine Efficacy in High-Titer Hepatitis B Virus Carrier Mice. Gastroenterology *158*, 1762-1775.e9. https://doi.org/10.1053/j.gastro.2020.01.032.

Milich, D., and Liang, T.J. (2003). Exploring the biological basis of hepatitis B e antigen in hepatitis B virus infection. Hepatol. Baltim. Md *38*, 1075–1086. https://doi.org/10.1053/jhep.2003.50453.

Mladenov, E., Magin, S., Soni, A., and Iliakis, G. (2016). DNA double-strand-break repair in higher eukaryotes and its role in genomic instability and cancer: Cell cycle and proliferation-dependent regulation. Semin. Cancer Biol. *37–38*, 51–64. https://doi.org/10.1016/j.semcancer.2016.03.003.

Mohebbi, A., Mohammadi, S., and Memarian, A. (2016). Prediction of HBF-0259 interactions with hepatitis B Virus receptors and surface antigen secretory factors. Virusdisease *27*, 234–241. https://doi.org/10.1007/s13337-016-0333-9.

Monjardino, J., and Crawford, E. (1979). Polypeptide profile of HBSAg excreted by a human hepatoma cell line. Virology *96*, 652–655. https://doi.org/10.1016/0042-6822(79)90123-5.

Moolla, N., Kew, M., and Arbuthnot, P. (2002). Regulatory elements of hepatitis B virus transcription. J. Viral Hepat. *9*, 323–331. https://doi.org/10.1046/j.1365-2893.2002.00381.x.

Moraleda, G., Saputelli, J., Aldrich, C.E., Averett, D., Condreay, L., and Mason, W.S. (1997). Lack of effect of antiviral therapy in nondividing hepatocyte cultures on the closed circular DNA of woodchuck hepatitis virus. J. Virol. *71*, 9392–9399. .

Murakami, Y., Saigo, K., Takashima, H., Minami, M., Okanoue, T., Bréchot, C., and Paterlini-Bréchot, P. (2005). Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas. Gut *54*, 1162–1168. https://doi.org/10.1136/gut.2004.054452.

Murphy, C.M., Xu, Y., Li, F., Nio, K., Reszka-Blanco, N., Li, X., Wu, Y., Yu, Y., Xiong, Y., and Su, L. (2016). Hepatitis B Virus X protein promotes degradation of SMC5/6 to enhance HBV replication. Cell Rep. *16*, 2846–2854. https://doi.org/10.1016/j.celrep.2016.08.026.

Mutz, P., Metz, P., Lempp, F.A., Bender, S., Qu, B., Schöneweis, K., Seitz, S., Tu, T., Restuccia, A., Frankish, J., et al. (2018). HBV Bypasses the Innate Immune Response and Does Not Protect HCV From Antiviral Activity of Interferon. Gastroenterology *154*, 1791-1804.e22. https://doi.org/10.1053/j.gastro.2018.01.044.

Nakabayashi, H., Taketa, K., Miyano, K., Yamane, T., and Sato, J. (1982). Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. Cancer Res. *42*, 3858–3863. .

Nakabayashi, H., Taketa, K., Yamane, T., Miyazaki, M., Miyano, K., and Sato, J. (1984). Phenotypical stability of a human hepatoma cell line, HuH-7, in long-term culture with chemically defined medium. Gan *75*, 151–158. .

Nakabayashi, H., Taketa, K., Yamane, T., Oda, M., and Sato, J. (1985). Hormonal control of alpha-fetoprotein secretion in human hepatoma cell lines proliferating in chemically defined medium. Cancer Res. *45*, 6379–6383. .

Nassal, M. (1992). The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. J. Virol. *66*, 4107–4116. .

Nassal, M. (2015). HBV cccDNA: viral persistence reservoir and key obstacle for a cure of chronic hepatitis B. Gut *64*, 1972–1984. https://doi.org/10.1136/gutjnl-2015-309809.

Nassal, M., and Schaller, H. (1993). Hepatitis B virus replication. Trends Microbiol. *1*, 221–228. https://doi.org/10.1016/0966-842x(93)90136-f.

Nassal, M., Rieger, A., and Steinau, O. (1992). Topological analysis of the hepatitis B virus core particle by cysteine-cysteine cross-linking. J. Mol. Biol. *225*, 1013–1025. https://doi.org/10.1016/0022-2836(92)90101-o.

Nguyen, D.H., and Hu, J. (2008). Reverse transcriptase- and RNA packaging signal-dependent incorporation of APOBEC3G into hepatitis B virus nucleocapsids. J. Virol. *82*, 6852–6861. https://doi.org/10.1128/JVI.00465-08.

Ning, X., Nguyen, D., Mentzer, L., Adams, C., Lee, H., Ashley, R., Hafenstein, S., and Hu, J. (2011). Secretion of genome-free hepatitis B virus--single strand blocking model for virion morphogenesis of para-retrovirus. PLoS Pathog. *7*, e1002255. https://doi.org/10.1371/journal.ppat.1002255.

Niu, C., Livingston, C.M., Li, L., Beran, R.K., Daffis, S., Ramakrishnan, D., Burdette, D., Peiser, L., Salas, E., Ramos, H., et al. (2017). The Smc5/6 Complex Restricts HBV when Localized to ND10 without Inducing an Innate Immune Response and Is Counteracted by the HBV X Protein Shortly after Infection. PLOS ONE *12*, e0169648. https://doi.org/10.1371/journal.pone.0169648.

Olinger, C.M., Jutavijittum, P., Hübschen, J.M., Yousukh, A., Samountry, B., Thammavong, T., Toriyama, K., and Muller, C.P. (2008). Possible new hepatitis B virus genotype, southeast Asia. Emerg. Infect. Dis. *14*, 1777–1780. https://doi.org/10.3201/eid1411.080437.

Ono-Nita, S.K., Kato, N., Shiratori, Y., Masaki, T., Lan, K.H., Carrilho, F.J., and Omata, M. (1999). YMDD motif in hepatitis B virus DNA polymerase influences on replication and lamivudine resistance: A study by in vitro full-length viral DNA transfection. Hepatol. Baltim. Md *29*, 939–945. https://doi.org/10.1002/hep.510290340.

Organization, W.H., and Standardization, W.E.C. on B. (2016). Collaborative study to evaluate the proposed WHO 4th International Standard for Hepatitis B Virus (HBV) DNA for Nucleic Acid Amplification Technique (NAT) based assays.

Orito, E., Mizokami, M., Ina, Y., Moriyama, E.N., Kameshima, N., Yamamoto, M., and Gojobori, T. (1989). Host-independent evolution and a genetic classification of the hepadnavirus family based on nucleotide sequences. Proc. Natl. Acad. Sci. U. S. A. *86*, 7059–7062. .

Otsuka, M., and Koike, K. (2020). Should Level of HBV RNA be Used to Determine When Patients Should Stop Treatment With Nucleos(t)ide Analogues. Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc. *18*, 551–552. https://doi.org/10.1016/j.cgh.2019.08.044.

Ou, J., and Rutter, W.J. (1985). Hybrid hepatitis B virus-host transcripts in a human hepatoma cell. Proc. Natl. Acad. Sci. U. S. A. *82*, 83–87. .

Palumbo, G.A., Scisciani, C., Pediconi, N., Lupacchini, L., Alfalate, D., Guerrieri, F., Calvo, L., Salerno, D., Cocco, S.D., Levrero, M., et al. (2015). IL6 Inhibits HBV Transcription by Targeting the Epigenetic Control of the Nuclear cccDNA Minichromosome. PLOS ONE *10*, e0142599. https://doi.org/10.1371/journal.pone.0142599.

Papatheodoridis, G.V., Manolakopoulos, S., Liaw, Y.-F., and Lok, A. (2012). Follow-up and indications for liver biopsy in HBeAg-negative chronic hepatitis B virus infection with persistently normal ALT: A systematic review. J. Hepatol. *57*, 196–202. https://doi.org/10.1016/j.jhep.2011.11.030.

Park, G.-S., Kim, H.-Y., Shin, H.-S., Park, S., Shin, H.-J., and Kim, K. (2008). Modulation of hepatitis B virus replication by expression of polymerase-surface fusion protein through splicing: Implications for viral persistence. Virus Res. *136*, 166–174. https://doi.org/10.1016/j.virusres.2008.05.005.

Park, S.G., Kim, Y., Park, E., Ryu, H.M., and Jung, G. (2003). Fidelity of hepatitis B virus polymerase. Eur. J. Biochem. *270*, 2929–2936. https://doi.org/10.1046/j.1432-1033.2003.03650.x.

Patel, N.H., Joshi, S.S., Lau, K.C.K., Castillo, E., and Coffin, C.S. (2019). Analysis of serum hepatitis B virus RNA levels in a multiethnic cohort of pregnant chronic hepatitis B carriers. J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol. *111*, 42–47. https://doi.org/10.1016/j.jcv.2019.01.002.

Patient, R., Hourioux, C., Sizaret, P.-Y., Trassard, S., Sureau, C., and Roingeard, P. (2007). Hepatitis B virus subviral envelope particle morphogenesis and intracellular trafficking. J. Virol. *81*, 3842–3851. https://doi.org/10.1128/JVI.02741-06.

Patient, R., Hourioux, C., and Roingeard, P. (2009). Morphogenesis of hepatitis B virus and its subviral envelope particles. Cell. Microbiol. *11*, 1561–1570. https://doi.org/10.1111/j.1462-5822.2009.01363.x.

Péneau, C., Imbeaud, S., La Bella, T., Hirsch, T.Z., Caruso, S., Calderaro, J., Paradis, V., Blanc, J.-F., Letouzé, E., Nault, J.-C., et al. (2021). Hepatitis B virus integrations promote local and distant oncogenic driver alterations in hepatocellular carcinoma. Gut https://doi.org/10.1136/gutjnl-2020-323153.

Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat. Protoc. *11*, 1650–1667. https://doi.org/10.1038/nprot.2016.095.

Pinto, M.R., Bey, E., and Bernstein, R. (1985). The PLC/PRF/5 human hepatoma cell line. I. Reevaluation of the karyotype. Cancer Genet. Cytogenet. *18*, 11–18. https://doi.org/10.1016/0165-4608(85)90033-0.

Polaris Observatory Collaborators (2018). Global prevalence, treatment, and prevention of hepatitis B virus infection in 2016: a modelling study. Lancet Gastroenterol. Hepatol. *3*, 383–403. https://doi.org/10.1016/S2468-1253(18)30056-6.

Pollicino, T., and Caminiti, G. (2021). HBV-Integration Studies in the Clinic: Role in the Natural History of Infection. Viruses *13*, 368. https://doi.org/10.3390/v13030368.

Pollicino, T., Belloni, L., Raffa, G., Pediconi, N., Squadrito, G., Raimondo, G., and Levrero, M. (2006). Hepatitis B virus replication is regulated by the acetylation status of hepatitis B virus cccDNA-bound H3 and H4 histones. Gastroenterology *130*, 823–837. https://doi.org/10.1053/j.gastro.2006.01.001.

Pollicino, T., Saitta, C., and Raimondo, G. (2011). Hepatocellular carcinoma: the point of view of the hepatitis B virus. Carcinogenesis *32*, 1122–1132. https://doi.org/10.1093/carcin/bgr108.

Pollicino, T., Cacciola, I., Saffioti, F., and Raimondo, G. (2014). Hepatitis B virus PreS/S gene variants: pathobiology and clinical implications. J. Hepatol. *61*, 408–417. https://doi.org/10.1016/j.jhep.2014.04.041.

Puisieux, A., Galvin, K., Troalen, F., Bressac, B., Marcais, C., Galun, E., Ponchel, F., Yakicier, C., Ji, J., and Ozturk, M. (1993). Retinoblastoma and p53 tumor suppressor genes in human hepatoma cell lines. FASEB J. *7*, 1407–1413. https://doi.org/10.1096/fasebj.7.14.8224613.

Qi, Y., Gao, Z., Xu, G., Peng, B., Liu, C., Yan, H., Yao, Q., Sun, G., Liu, Y., Tang, D., et al. (2016). DNA Polymerase κ Is a Key Cellular Factor for the Formation of Covalently Closed Circular DNA of Hepatitis B Virus. PLoS Pathog. *12*. https://doi.org/10.1371/journal.ppat.1005893.

Qiu, G.-H., Xie, X., Xu, F., Shi, X., Wang, Y., and Deng, L. (2015). Distinctive pharmacological differences between liver cancer cell lines HepG2 and Hep3B. Cytotechnology *67*, 1–12. https://doi.org/10.1007/s10616-014-9761-9.

Quarleri, J. (2014). Core promoter: a critical region where the hepatitis B virus makes decisions. World J. Gastroenterol. *20*, 425–435. https://doi.org/10.3748/wjg.v20.i2.425.

Rabe, B., Glebe, D., and Kann, M. (2006). Lipid-mediated introduction of hepatitis B virus capsids into nonsusceptible cells allows highly efficient replication and facilitates the study of early infection events. J. Virol. *80*, 5465–5473. https://doi.org/10.1128/JVI.02303-05.

Radziwill, G., Tucker, W., and Schaller, H. (1990). Mutational analysis of the hepatitis B virus P gene product: domain structure and RNase H activity. J. Virol. *64*, 613–620. .

Raimondo, G., Locarnini, S., Pollicino, T., Levrero, M., Zoulim, F., Lok, A.S., Allain, J.-P., Berg, T., Bertoletti, A., Brunetto, M.R., et al. (2019). Update of the statements on biology and clinical impact of occult hepatitis B virus infection. J. Hepatol. *71*, 397–408. https://doi.org/10.1016/j.jhep.2019.03.034.

Ramirez, R., van Buuren, N., Gamelin, L., Soulette, C., May, L., Han, D., Yu, M., Choy, R., Cheng, G., Bhardwaj, N., et al. (2021). Targeted Long-Read Sequencing Reveals Comprehensive Architecture, Burden, and Transcriptional Signatures from Hepatitis B Virus-Associated Integrations and Translocations in Hepatocellular Carcinoma Cell Lines. J. Virol. *95*, e0029921. https://doi.org/10.1128/JVI.00299-21.

Raney, A.K., Easton, A.J., and McLachlan, A. (1994). Characterization of the minimal elements of the hepatitis B virus large surface antigen promoter. J. Gen. Virol. *75 (Pt 10)*, 2671–2679. https://doi.org/10.1099/0022-1317-75-10-2671.

Rasche, A., Souza, B.F. de C.D., and Drexler, J.F. (2016). Bat hepadnaviruses and the origins of primate hepatitis B viruses. Curr. Opin. Virol. *16*, 86–94. https://doi.org/10.1016/j.coviro.2016.01.015.

Rausch, T., Fritz, M.H.-Y., Untergasser, A., and Benes, V. (2020). Tracy: basecalling, alignment, assembly and deconvolution of sanger chromatogram trace files. BMC Genomics *21*, 230. https://doi.org/10.1186/s12864-020-6635-8.

Ravi, R.K., Walton, K., and Khosroheidari, M. (2018). MiSeq: A Next Generation Sequencing Platform for Genomic Analysis. Methods Mol. Biol. Clifton NJ *1706*, 223–232. https://doi.org/10.1007/978-1-4939-7471-9_12.

Rendi-Wagner, P., Kundi, M., Stemberger, H., Wiedermann, G., Holzmann, H., Hofer, M., Wiesinger, K., and Kollaritsch, H. (2001). Antibody-response to three recombinant hepatitis B vaccines: comparative evaluation of multicenter travel-clinic based experience. Vaccine *19*, 2055–2060. https://doi.org/10.1016/S0264-410X(00)00410-2.

Rieger, A., and Nassal, M. (1996). Specific hepatitis B virus minus-strand DNA synthesis requires only the 5' encapsidation signal and the 3'-proximal direct repeat DR1. J. Virol. *70*, 585–589. https://doi.org/10.1128/JVI.70.1.585-589.1996.

Ringelhan, M., and Protzer, U. (2015). Oncogenic potential of hepatitis B virus encoded proteins. Curr. Opin. Virol. *14*, 109–115. https://doi.org/10.1016/j.coviro.2015.08.015.

Rivière, L., Gerossier, L., Ducroux, A., Dion, S., Deng, Q., Michel, M.-L., Buendia, M.-A., Hantz, O., and Neuveut, C. (2015). HBx relieves chromatin-mediated transcriptional repression of hepatitis B viral cccDNA involving SETDB1 histone methyltransferase. J. Hepatol. *63*, 1093–1102. https://doi.org/10.1016/j.jhep.2015.06.023.

Rokuhara, A., Matsumoto, A., Tanaka, E., Umemura, T., Yoshizawa, K., Kimura, T., Maki, N., and Kiyosawa, K. (2006). Hepatitis B virus RNA is measurable in serum and can be a new marker for monitoring lamivudine therapy. J. Gastroenterol. *41*, 785–790. https://doi.org/10.1007/s00535-006-1856-4.

Rosmorduc, O., Petit, M.A., Pol, S., Capel, F., Bortolotti, F., Berthelot, P., Brechot, C., and Kremsdorf, D. (1995). In vivo and in vitro expression of defective hepatitis B virus particles generated by spliced hepatitis B virus RNA. Hepatol. Baltim. Md *22*, 10–19. .

Rossignol, J.-F., and Bréchot, C. (2019). A Pilot Clinical Trial of Nitazoxanide in the Treatment of Chronic Hepatitis B. Hepatol. Commun. *3*, 744–747. https://doi.org/10.1002/hep4.1339.

Ruan, P., Dai, X., Sun, J., He, C., Huang, C., Zhou, R., and Chemin, I. (2020). Integration of hepatitis B virus DNA into p21-activated kinase 3 (PAK3) gene in HepG2.2.15 cells. Virus Genes *56*, 168–173. https://doi.org/10.1007/s11262-019-01725-4.

Rutsaert, S., Bosman, K., Trypsteen, W., Nijhuis, M., and Vandekerckhove, L. (2018). Digital PCR as a tool to measure HIV persistence. Retrovirology *15*, 16. https://doi.org/10.1186/s12977-018-0399-0.

Ryabova, L.A., Pooggin, M.M., and Hohn, T. (2002). Viral strategies of translation initiation: ribosomal shunt and reinitiation. Prog. Nucleic Acid Res. Mol. Biol. *72*, 1–39. https://doi.org/10.1016/s0079-6603(02)72066-7.

Saldanha, J., Gerlich, W., Lelie, N., Dawson, P., Heermann, K., Heath, A., and WHO Collaborative Study Group (2001). An international collaborative study to establish a World Health Organization international standard for hepatitis B virus DNA nucleic acid amplification techniques. Vox Sang. *80*, 63–71. https://doi.org/10.1046/j.1423-0410.2001.00003.x.

Salerno, D., Chiodo, L., Alfano, V., Floriot, O., Cottone, G., Paturel, A., Pallocca, M., Plissonnier, M.-L., Jeddari, S., Belloni, L., et al. (2020). Hepatitis B protein HBx binds the DLEU2 lncRNA to sustain cccDNA and host cancer-related gene transcription. Gut *69*, 2016–2024. https://doi.org/10.1136/gutjnl-2019-319637.

Sanada, T., Hirata, Y., Naito, Y., Yamamoto, N., Kikkawa, Y., Ishida, Y., Yamasaki, C., Tateno, C., Ochiya, T., and Kohara, M. (2017). Transmission of HBV DNA Mediated by Ceramide-Triggered Extracellular Vesicles. Cell. Mol. Gastroenterol. Hepatol. *3*, 272–283. https://doi.org/10.1016/j.jcmgh.2016.10.003.

Sanger, F., and Coulson, A.R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. *94*, 441–448. https://doi.org/10.1016/0022-2836(75)90213-2.

Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, C.A., Hutchison, C.A., Slocombe, P.M., and Smith, M. (1977a). Nucleotide sequence of bacteriophage phi X174 DNA. Nature *265*, 687–695. https://doi.org/10.1038/265687a0.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977b). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. *74*, 5463–5467. .

Sauvage, V., Boizeau, L., Candotti, D., Vandenbogaert, M., Servant-Delmas, A., Caro, V., and Laperche, S. (2018). Early MinION™ nanopore single-molecule sequencing technology enables the characterization of hepatitis B virus genetic complexity in clinical samples. PLOS ONE *13*, e0194366. https://doi.org/10.1371/journal.pone.0194366.

Schaefer, S. (2007). Hepatitis B virus taxonomy and hepatitis B virus genotypes. World J. Gastroenterol. *13*, 14–21. https://doi.org/10.3748/wjg.v13.i1.14.

Schilling, R., Ijaz, S., Davidoff, M., Lee, J.Y., Locarnini, S., Williams, R., and Naoumov, N.V. (2003). Endocytosis of Hepatitis B Immune Globulin into Hepatocytes Inhibits the Secretion of Hepatitis B Virus Surface Antigen and Virions. J. Virol. *77*, 8882–8892. https://doi.org/10.1128/JVI.77.16.8882-8892.2003.

Schinazi, R.F., Ehteshami, M., Bassit, L., and Asselah, T. (2018). Towards HBV curative therapies. Liver Int. Off. J. Int. Assoc. Study Liver *38 Suppl 1*, 102–114. https://doi.org/10.1111/liv.13656.

Schlüter, V., Meyer, M., Hofschneider, P.H., Koshy, R., and Caselmann, W.H. (1994). Integrated hepatitis B virus X and 3' truncated preS/S sequences derived from human hepatomas encode functionally active transactivators. Oncogene *9*, 3335–3344. .

Schmitt, S., Glebe, D., Tolle, T.K., Lochnit, G., Linder, D., Geyer, R., and Gerlich, W.H. (2004). Structure of pre-S2 N- and O-linked glycans in surface proteins from different genotypes of hepatitis B virus. J. Gen. Virol. *85*, 2045–2053. https://doi.org/10.1099/vir.0.79932-0.

Schulze, A., Gripon, P., and Urban, S. (2007). Hepatitis B virus infection initiates with a large surface protein-dependent binding to heparan sulfate proteoglycans. Hepatol. Baltim. Md *46*, 1759–1768. https://doi.org/10.1002/hep.21896.

Schweitzer, A., Horn, J., Mikolajczyk, R.T., Krause, G., and Ott, J.J. (2015). Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. Lancet Lond. Engl. *386*, 1546–1555. https://doi.org/10.1016/S0140-6736(15)61412-X.

Scott, R.M., Snitbhan, R., Bancroft, W.H., Alter, H.J., and Tingpalapong, M. (1980). Experimental transmission of hepatitis B virus by semen and saliva. J. Infect. Dis. *142*, 67–71. https://doi.org/10.1093/infdis/142.1.67.

Scotto, J., Hadchouel, M., Hery, C., Alvarez, F., Yvart, J., Tiollais, P., Bernard, O., and Brechot, C. (1983). Hepatitis B virus DNA in children's liver diseases: detection by blot hybridisation in liver and serum. Gut *24*, 618–624. .

Sedlak, R.H., Kuypers, J., and Jerome, K.R. (2014). A multiplexed droplet digital PCR assay performs better than qPCR on inhibition prone samples. Diagn. Microbiol. Infect. Dis. *80*, 285–286. https://doi.org/10.1016/j.diagmicrobio.2014.09.004.

Seeger, C., and Mason, W.S. (2000). Hepatitis B virus biology. Microbiol. Mol. Biol. Rev. MMBR *64*, 51–68. https://doi.org/10.1128/mmbr.64.1.51-68.2000.

Seeger, C., and Sohn, J.A. (2016). Complete Spectrum of CRISPR/Cas9-induced Mutations on HBV cccDNA. Mol. Ther. *24*, 1258–1266. https://doi.org/10.1038/mt.2016.94.

Sekiba, K., Otsuka, M., Ohno, M., Yamagami, M., Kishikawa, T., Suzuki, T., Ishibashi, R., Seimiya, T., Tanaka, E., and Koike, K. (2019a). Inhibition of HBV Transcription From cccDNA With Nitazoxanide by Targeting the HBx-DDB1 Interaction. Cell. Mol. Gastroenterol. Hepatol. *7*, 297–312. https://doi.org/10.1016/j.jcmgh.2018.10.010.

Sekiba, K., Otsuka, M., Ohno, M., Yamagami, M., Kishikawa, T., Seimiya, T., Suzuki, T., Tanaka, E., Ishibashi, R., Funato, K., et al. (2019b). Pevonedistat, a Neuronal Precursor Cell-Expressed Developmentally Down-Regulated Protein 8-Activating Enzyme Inhibitor, Is a Potent Inhibitor of Hepatitis B Virus. Hepatol. Baltim. Md *69*, 1903–1915. https://doi.org/10.1002/hep.30491.

Sells, M.A., Zelent, A.Z., Shvartsman, M., and Acs, G. (1988). Replicative intermediates of hepatitis B virus in HepG2 cells that produce infectious virions. J. Virol. *62*, 2836–2844. https://doi.org/10.1128/JVI.62.8.2836-2844.1988.

Seto, W.-K., Wong, D.K.-H., Fung, J., Huang, F.-Y., Liu, K.S.-H., Lai, C.-L., and Yuen, M.-F. (2014). Linearized hepatitis B surface antigen and hepatitis B core-related antigen in the natural history of chronic hepatitis B. Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis. *20*, 1173–1180. https://doi.org/10.1111/1469-0691.12739.

Shafritz, D.A., Shouval, D., Sherman, H.I., Hadziyannis, S.J., and Kew, M.C. (1981). Integration of hepatitis B virus DNA into the genome of liver cells in chronic liver disease and hepatocellular carcinoma. Studies in percutaneous liver biopsies and post-mortem tissue specimens. N. Engl. J. Med. *305*, 1067–1073. https://doi.org/10.1056/NEJM198110293051807.

Shamay, M., Agami, R., and Shaul, Y. (2001). HBV integrants of hepatocellular carcinoma cell lines contain an active enhancer. Oncogene *20*, 6811–6819. https://doi.org/10.1038/sj.onc.1204879.

Shaul, Y., Ziemer, M., Garcia, P.D., Crawford, R., Hsu, H., Valenzuela, P., and Rutter, W.J. (1984). Cloning and analysis of integrated hepatitis virus sequences from a human hepatoma cell line. J. Virol. *51*, 776–787. .

Shaul, Y., Rutter, W.J., and Laub, O. (1985). A human hepatitis B viral enhancer element. EMBO J. *4*, 427–430. .

Sheu, S.Y., and Lo, S.J. (1992). Preferential ribosomal scanning is involved in the differential synthesis of the hepatitis B viral surface antigens from subgenomic transcripts. Virology *188*, 353–357. https://doi.org/10.1016/0042-6822(92)90764-g.

Sirma, H., Giannini, C., Poussin, K., Paterlini, P., Kremsdorf, D., and Bréchot, C. (1999). Hepatitis B virus X mutants, present in hepatocellular carcinoma tissue abrogate both the antiproliferative and transactivation effects of HBx. Oncogene *18*, 4848–4859. https://doi.org/10.1038/sj.onc.1202867.

Sitterlin, D., Lee, T.H., Prigent, S., Tiollais, P., Butel, J.S., and Transy, C. (1997). Interaction of the UV-damaged DNA-binding protein with hepatitis B virus X protein is conserved among mammalian hepadnaviruses and restricted to transactivation-proficient X-insertion mutants. J. Virol. *71*, 6194–6199. .

Slagle, B.L., and Bouchard, M.J. (2016). Hepatitis B Virus X and Regulation of Viral Gene Expression. Cold Spring Harb. Perspect. Med. *6*, a021402. https://doi.org/10.1101/cshperspect.a021402.

Smits, N., Rasmussen, J., Bodea, G.O., Amarilla, A.A., Gerdes, P., Sanchez-Luque, F.J., Ajjikuttira, P., Modhiran, N., Liang, B., Faivre, J., et al. (2021). No evidence of human genome integration of SARS-CoV-2 found by long-read DNA sequencing. Cell Rep. *36*, 109530. https://doi.org/10.1016/j.celrep.2021.109530.

Sommese, L., Sabia, C., Paolillo, R., Parente, D., Capuano, M., Iannone, C., Cavalca, F., Schiano, C., Vasco, M., Pascale, M.R.D., et al. (2014). Screening tests for hepatitis B virus, hepatitis C virus, and human immunodeficiency virus in blood donors: Evaluation of two chemiluminescent immunoassay systems. Scand. J. Infect. Dis. *46*, 660–664. https://doi.org/10.3109/00365548.2014.926564.

Song, J.E., and Kim, D.Y. (2016). Diagnosis of hepatitis B. Ann. Transl. Med. *4*, 338. https://doi.org/10.21037/atm.2016.09.11.

Soussan, P., Garreau, F., Zylberberg, H., Ferray, C., Brechot, C., and Kremsdorf, D. (2000). In vivo expression of a new hepatitis B virus protein encoded by a spliced RNA. J. Clin. Invest. *105*, 55–60. https://doi.org/10.1172/JCI8098.

Soussan, P., Pol, J., Garreau, F., Schneider, V., Le Pendeven, C., Nalpas, B., Lacombe, K., Bonnard, P., Pol, S., and Kremsdorf, D. (2008). Expression of defective hepatitis B virus particles derived from singly spliced RNA is related to liver disease. J. Infect. Dis. *198*, 218–225. https://doi.org/10.1086/589623.

Spellman, M., and Martin, J.T. (2011). 751 TREATMENT OF CHRONIC HEPATITIS B INFECTION WITH DV-601, A THERAPEUTIC VACCINE. J. Hepatol. *54*, S302. https://doi.org/10.1016/S0168-8278(11)60753-8.

Stadelmayer, B., Diederichs, A., Chapus, F., Rivoire, M., Neveu, G., Alam, A., Fraisse, L., Carter, K., Testoni, B., and Zoulim, F. (2020). Full-length 5'RACE identifies all major HBV transcripts in HBV-infected hepatocytes and patient serum. J. Hepatol. *73*, 40–51. https://doi.org/10.1016/j.jhep.2020.01.028.

Staprans, S., Loeb, D.D., and Ganem, D. (1991). Mutations affecting hepadnavirus plus-strand DNA synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA. J. Virol. *65*, 1255–1262.
.

Steven, A.C., Conway, J.F., Cheng, N., Watts, N.R., Belnap, D.M., Harris, A., Stahl, S.J., and Wingfield, P.T. (2005). Structure, assembly, and antigenicity of hepatitis B virus capsid proteins. Adv. Virus Res. *64*, 125–164. https://doi.org/10.1016/S0065-3527(05)64005-5.

Su, T.-S., Lin, L.-H., Chou, C.-K., Chang, C., Ting, L.-P., Hu, C., and Han, S.-H. (1986). Hepatitis B virus transcripts in a human hepatoma cell line, hep 3B. Biochem. Biophys. Res. Commun. *138*, 131–138. https://doi.org/10.1016/0006-291X(86)90256-1.

Su, T.S., Hwang, W.L., and Yauk, Y.K. (1998). Characterization of hepatitis B virus integrant that results in chromosomal rearrangement. DNA Cell Biol. *17*, 415–425. https://doi.org/10.1089/dna.1998.17.415.

Summers, J., and Mason, W.S. (1982). Replication of the genome of a hepatitis B--like virus by reverse transcription of an RNA intermediate. Cell *29*, 403–415. https://doi.org/10.1016/0092-8674(82)90157-x.

Summers, J., O'Connell, A., and Millman, I. (1975). Genome of hepatitis B virus: restriction enzyme cleavage and structure of DNA extracted from Dane particles. Proc. Natl. Acad. Sci. U. S. A. *72*, 4597–4601. https://doi.org/10.1073/pnas.72.11.4597.

Sun, D., and Nassal, M. (2006). Stable HepG2- and Huh7-based human hepatoma cell lines for efficient regulated expression of infectious hepatitis B virus. J. Hepatol. *45*, 636–645. https://doi.org/10.1016/j.jhep.2006.05.019.

Sun, C.T., Lo, W.Y., Wang, I.H., Lo, Y.H., Shiou, S.R., Lai, C.K., and Ting, L.P. (2001). Transcription repression of human hepatitis B virus genes by negative regulatory element-binding protein/SON. J. Biol. Chem. *276*, 24059–24067. https://doi.org/10.1074/jbc.M101330200.

Sung, W.-K., Zheng, H., Li, S., Chen, R., Liu, X., Li, Y., Lee, N.P., Lee, W.H., Ariyaratne, P.N., Tennakoon, C., et al. (2012). Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. Nat. Genet. *44*, 765–769. https://doi.org/10.1038/ng.2295.

Sureau, C., and Salisse, J. (2013). A conformational heparan sulfate binding site essential to infectivity overlaps with the conserved hepatitis B virus a-determinant. Hepatol. Baltim. Md *57*, 985–994. https://doi.org/10.1002/hep.26125.

Sureau, C., Romet-Lemonne, J.L., Mullins, J.I., and Essex, M. (1986). Production of hepatitis B virus by a differentiated human hepatoma cell line after transfection with cloned circular HBV DNA. Cell *47*, 37–47. https://doi.org/10.1016/0092-8674(86)90364-8.

Sureau, C., Eichberg, J.W., Hubbard, G.B., Romet-Lemonne, J.L., and Essex, M. (1988). A molecularly cloned hepatitis B virus produced in vitro is infectious in a chimpanzee. J. Virol. *62*, 3064–3067. https://doi.org/10.1128/JVI.62.8.3064-3067.1988.

Sykes, P.J., Neoh, S.H., Brisco, M.J., Hughes, E., Condon, J., and Morley, A.A. (1992). Quantitation of targets for PCR by use of limiting dilution. BioTechniques *13*, 444–449. .

Sze, K.M.F., Chu, G.K.Y., Lee, J.M.F., and Ng, I.O.L. (2013). C-terminal truncated hepatitis B virus x protein is associated with metastasis and enhances invasiveness by C-Jun/matrix metalloproteinase protein 10 activation in hepatocellular carcinoma. Hepatol. Baltim. Md *57*, 131–139. https://doi.org/10.1002/hep.25979.

Sze, K.M.-F., Ho, D.W.-H., Chiu, Y.-T., Tsui, Y.-M., Chan, L.-K., Lee, J.M.-F., Chok, K.S.-H., Chan, A.C.-Y., Tang, C.-N., Tang, V.W.-L., et al. (2020). HBV-TERT Promoter Integration Harnesses Host ELF4 Resulting in TERT Gene Transcription in Hepatocellular Carcinoma. Hepatol. Baltim. Md https://doi.org/10.1002/hep.31231.

Tatematsu, K., Tanaka, Y., Kurbanov, F., Sugauchi, F., Mano, S., Maeshiro, T., Nakayoshi, T., Wakuta, M., Miyakawa, Y., and Mizokami, M. (2009). A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. J. Virol. *83*, 10538–10547. https://doi.org/10.1128/JVI.00462-09.

Testoni, B., Lebossé, F., Scholtes, C., Berby, F., Miaglia, C., Subic, M., Loglio, A., Facchetti, F., Lampertico, P., Levrero, M., et al. (2019). Serum hepatitis B core-related antigen (HBcrAg) correlates with covalently closed circular DNA transcriptional activity in chronic hepatitis B patients. J. Hepatol. *70*, 615–625. https://doi.org/10.1016/j.jhep.2018.11.030.

Tillmann, H.L. (2007). Antiviral therapy and resistance with hepatitis B virus infection. World J. Gastroenterol. *13*, 125–140. https://doi.org/10.3748/wjg.v13.i1.125.

Tipu, H.N., and Shabbir, A. (2015). Evolution of DNA sequencing. J. Coll. Physicians Surg.--Pak. JCPSP *25*, 210–215. https://doi.org/03.2015/JCPSP.210215.

Tognoni, A., Cattaneo, R., Serfling, E., and Schaffner, W. (1985). A novel expression selection approach allows precise mapping of the hepatitis B virus enhancer. Nucleic Acids Res. *13*, 7457–7472. https://doi.org/10.1093/nar/13.20.7457.

Toh, S.T., Jin, Y., Liu, L., Wang, J., Babrzadeh, F., Gharizadeh, B., Ronaghi, M., Toh, H.C., Chow, P.K.-H., Chung, A.Y.-F., et al. (2013). Deep sequencing of the hepatitis B virus in hepatocellular carcinoma patients reveals enriched integration events, structural alterations and sequence variations. Carcinogenesis *34*, 787–798. https://doi.org/10.1093/carcin/bgs406.

Tolle, T.K., Glebe, D., Linder, M., Linder, D., Schmitt, S., Geyer, R., and Gerlich, W.H. (1998). Structure and glycosylation patterns of surface proteins from woodchuck hepatitis virus. J. Virol. *72*, 9978–9985. .

Tong, S., and Revill, P. (2016). Overview of viral replication and genetic variability. J. Hepatol. *64*, S4–S16. https://doi.org/10.1016/j.jhep.2016.01.027.

Tropberger, P., Mercier, A., Robinson, M., Zhong, W., Ganem, D.E., and Holdorf, M. (2015). Mapping of histone modifications in episomal HBV cccDNA uncovers an unusual chromatin organization amenable to epigenetic manipulation. Proc. Natl. Acad. Sci. U. S. A. *112*, E5715-5724. https://doi.org/10.1073/pnas.1518090112.

Trung, N.T., Hai, L.T., Giang, D.P., Hoan, P.Q., Binh, M.T., Hoan, N.X., Toan, N.L., Meyer, C.G., Velavan, T.P., Bang, M.H., et al. (2019). No expression of HBV-human chimeric fusion transcript (HBx-LINE1) among Vietnamese patients with HBV-associated hepatocellular carcinoma. Ann. Hepatol. *18*, 404–405. https://doi.org/10.1016/j.aohep.2019.02.002.

Tseng, T.-C., Liu, C.-J., Yang, H.-C., Su, T.-H., Wang, C.-C., Chen, C.-L., Kuo, S.F.-T., Liu, C.-H., Chen, P.-J., Chen, D.-S., et al. (2012). Determinants of spontaneous surface antigen loss in hepatitis B e antigen-negative patients with a low viral load. Hepatol. Baltim. Md *55*, 68–76. https://doi.org/10.1002/hep.24615.

Tu, T., and Jilbert, A.R. (2017). Detection of Hepatocyte Clones Containing Integrated Hepatitis B Virus DNA Using Inverse Nested PCR. Methods Mol. Biol. Clifton NJ *1540*, 97–118. https://doi.org/10.1007/978-1-4939-6700-1_9.

Tu, T., Budzinska, M.A., Shackel, N.A., and Urban, S. (2017). HBV DNA Integration: Molecular Mechanisms and Clinical Implications. Viruses *9*. https://doi.org/10.3390/v9040075.

Tu, T., Budzinska, M.A., Vondran, F.W.R., Shackel, N.A., and Urban, S. (2018). Hepatitis B Virus DNA Integration Occurs Early in the Viral Life Cycle in an In Vitro Infection Model via Sodium Taurocholate Cotransporting Polypeptide-Dependent Uptake of Enveloped Virus Particles. J. Virol. *92*. https://doi.org/10.1128/JVI.02007-17.

Tu, T., Zehnder, B., Levy, M., Micali, G., Tran, L., Dabere, O., Main, N., Shackel, N., and Urban, S. (2019). Hepatitis B virus (HBV) DNA integration is not driven by viral proteins. In Zeitschrift Für Gastroenterologie, (Georg Thieme Verlag KG), p. P5.46.

Tu, T., Zhang, H., and Urban, S. (2021a). Hepatitis B Virus DNA Integration: In Vitro Models for Investigating Viral Pathogenesis and Persistence. Viruses *13*, 180. https://doi.org/10.3390/v13020180.

Tu, T., Zehnder, B., Qu, B., and Urban, S. (2021b). De novo synthesis of hepatitis B virus nucleocapsids is dispensable for the maintenance and transcriptional regulation of cccDNA. JHEP Rep. *3*, 100195. https://doi.org/10.1016/j.jhepr.2020.100195.

Tur-Kaspa, R., Burk, R.D., Shaul, Y., and Shafritz, D.A. (1986). Hepatitis B virus DNA contains a glucocorticoid-responsive element. Proc. Natl. Acad. Sci. U. S. A. *83*, 1627–1631. https://doi.org/10.1073/pnas.83.6.1627.

Twist, E.M., Clark, H.F., Aden, D.P., Knowles, B.B., and Plotkin, S.A. (1981). Integration pattern of hepatitis B virus DNA sequences in human hepatoma cell lines. J. Virol. *37*, 239–243. .

Valaydon, Z.S., and Locarnini, S.A. (2017). The virological aspects of hepatitis B. Best Pract. Res. Clin. Gastroenterol. *31*, 257–264. https://doi.org/10.1016/j.bpg.2017.04.013.

Van Der Meeren, O., Kuriyakose, S., Kolhe, D., and Hardt, K. (2012). Immunogenicity of Infanrix™ hexa administered at 3, 5 and 11 months of age. Vaccine *30*, 2710–2714. https://doi.org/10.1016/j.vaccine.2012.02.024.

Velkov, S., Ott, J.J., Protzer, U., and Michler, T. (2018). The Global Hepatitis B Virus Genotype Distribution Approximated from Available Genotyping Data. Genes *9*. https://doi.org/10.3390/genes9100495.

Verrier, E.R., Colpitts, C.C., Bach, C., Heydmann, L., Weiss, A., Renaud, M., Durand, S.C., Habersetzer, F., Durantel, D., Abou-Jaoudé, G., et al. (2016). A targeted functional RNA interference screen uncovers glypican 5 as an entry factor for hepatitis B and D viruses. Hepatol. Baltim. Md *63*, 35–48. https://doi.org/10.1002/hep.28013.

de Waal, L., Lewis, T.A., Rees, M.G., Tsherniak, A., Wu, X., Choi, P.S., Gechijian, L., Hartigan, C., Faloon, P.W., Hickey, M.J., et al. (2016). Identification of cancer cytotoxic modulators of PDE3A by predictive chemogenomics. Nat. Chem. Biol. *12*, 102–108. https://doi.org/10.1038/nchembio.1984.

Wai, C.T., Chu, C.-J., Hussain, M., and Lok, A.S.F. (2002). HBV genotype B is associated with better response to interferon therapy in HBeAg(+) chronic hepatitis than genotype C. Hepatol. Baltim. Md *36*, 1425–1430. https://doi.org/10.1053/jhep.2002.37139.

Walsh, R., and Locarnini, S. (2012). Hepatitis B precore protein: pathogenic potential and therapeutic promise. Yonsei Med. J. *53*, 875–885. https://doi.org/10.3349/ymj.2012.53.5.875.

Wang, G.H., and Seeger, C. (1992). The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. Cell *71*, 663–670. https://doi.org/10.1016/0092-8674(92)90599-8.

Wang, H., Kim, S., and Ryu, W.-S. (2009). DDX3 DEAD-Box RNA helicase inhibits hepatitis B virus reverse transcription by incorporation into nucleocapsids. J. Virol. *83*, 5815–5824. https://doi.org/10.1128/JVI.00011-09.

Wang, H.-Y., Chien, M.-H., Huang, H.-P., Chang, H.-C., Wu, C.-C., Chen, P.-J., Chang, M.-H., and Chen, D.-S. (2010). Distinct hepatitis B virus dynamics in the immunotolerant and early immunoclearance phases. J. Virol. *84*, 3454–3463. https://doi.org/10.1128/JVI.02164-09.

Wang, J., Shen, T., Huang, X., Kumar, G.R., Chen, X., Zeng, Z., Zhang, R., Chen, R., Li, T., Zhang, T., et al. (2016). Serum hepatitis B virus RNA is encapsidated pregenome RNA that may be associated with persistence of viral infection and rebound. J. Hepatol. *65*, 700–710. https://doi.org/10.1016/j.jhep.2016.05.029.

Wang, J., Yu, Y., Li, G., Shen, C., Li, J., Chen, S., Zhang, X., Zhu, M., Zheng, J., Song, Z., et al. (2018). Natural history of serum HBV-RNA in chronic HBV infection. J. Viral Hepat. *25*, 1038–1047. https://doi.org/10.1111/jvh.12908.

Wang, L., Cao, X., Wang, Z., Gao, Y., Deng, J., Liu, X., and Zhuang, H. (2019). Correlation of HBcrAg with Intrahepatic Hepatitis B Virus Total DNA and Covalently Closed Circular DNA in HBeAg-Positive Chronic Hepatitis B Patients. J. Clin. Microbiol. *57*. https://doi.org/10.1128/JCM.01303-18.

Wang, Y., Jiang, L., Ji, X., Yang, B., Zhang, Y., and Fu, X.-D. (2013). Hepatitis B Viral RNA Directly Mediates Down-regulation of the Tumor Suppressor MicroRNA miR-15a/miR-16-1 in Hepatocytes. J. Biol. Chem. *288*, 18484–18493. https://doi.org/10.1074/jbc.M113.458158.

Watanabe, T., Sorensen, E.M., Naito, A., Schott, M., Kim, S., and Ahlquist, P. (2007). Involvement of host cellular multivesicular body functions in hepatitis B virus budding. Proc. Natl. Acad. Sci. U. S. A. *104*, 10205–10210. https://doi.org/10.1073/pnas.0704000104.

Wei, L., and Ploss, A. (2021). Mechanism of Hepatitis B Virus cccDNA Formation. Viruses *13*, 1463. https://doi.org/10.3390/v13081463.

Werle-Lapostolle, B., Bowden, S., Locarnini, S., Wursthorn, K., Petersen, J., Lau, G., Trepo, C., Marcellin, P., Goodman, Z., Delaney, W.E., et al. (2004). Persistence of cccDNA during the natural history of chronic hepatitis B and decline during adefovir dipivoxil therapy. Gastroenterology *126*, 1750–1758. https://doi.org/10.1053/j.gastro.2004.03.018.

Will, H., Cattaneo, R., Koch, H.-G., Darai, G., Schaller, H., Schellekens, H., van Eerd, P.M.C.A., and Deinhardt, F. (1982). Cloned HBV DNA causes hepatitis in chimpanzees. Nature *299*, 740–742. https://doi.org/10.1038/299740a0.

Will, H., Reiser, W., Weimer, T., Pfaff, E., Büscher, M., Sprengel, R., Cattaneo, R., and Schaller, H. (1987). Replication strategy of human hepatitis B virus. J. Virol. *61*, 904–911. .

Willms, E., Cabañas, C., Mäger, I., Wood, M.J.A., and Vader, P. (2018). Extracellular Vesicle Heterogeneity: Subpopulations, Isolation Techniques, and Diverse Functions in Cancer Progression. Front. Immunol. *9*, 738. https://doi.org/10.3389/fimmu.2018.00738.

Wommack, K.E., Bhavsar, J., and Ravel, J. (2008). Metagenomics: read length matters. Appl. Environ. Microbiol. *74*, 1453–1463. https://doi.org/10.1128/AEM.02181-07.

Wooddell, C.I., Yuen, M.-F., Chan, H.L.-Y., Gish, R.G., Locarnini, S.A., Chavez, D., Ferrari, C., Given, B.D., Hamilton, J., Kanner, S.B., et al. (2017). RNAi-based treatment of chronically infected patients and chimpanzees reveals that integrated hepatitis B virus DNA is a source of HBsAg. Sci. Transl. Med. *9*. https://doi.org/10.1126/scitranslmed.aan0241.

Wu, G., Liu, B., Zhang, Y., Li, J., Arzumanyan, A., Clayton, M.M., Schinazi, R.F., Wang, Z., Goldmann, S., Ren, Q., et al. (2013). Preclinical characterization of GLS4, an inhibitor of hepatitis B virus core particle assembly. Antimicrob. Agents Chemother. *57*, 5344–5354. https://doi.org/10.1128/AAC.01091-13.

Wu, H.-C., Tsai, H.-W., Teng, C.-F., Hsieh, W.-C., Lin, Y.-J., Wang, L.H.-C., Yuan, Q., and Su, I.-J. (2014). Ground-glass hepatocytes co-expressing hepatitis B virus X protein and surface antigens exhibit enhanced oncogenic effects and tumorigenesis. Hum. Pathol. *45*, 1294–1301. https://doi.org/10.1016/j.humpath.2013.10.039.

Wu, Z.-J., Zhu, Y., Huang, D.-R., and Wang, Z.-Q. (2010). Constructing the HBV-human protein interaction network to understand the relationship between HBV and hepatocellular carcinoma. J. Exp. Clin. Cancer Res. CR *29*, 146. https://doi.org/10.1186/1756-9966-29-146.

Xia, Y., Stadler, D., Lucifora, J., Reisinger, F., Webb, D., Hösel, M., Michler, T., Wisskirchen, K., Cheng, X., Zhang, K., et al. (2016). Interferon-γ and Tumor Necrosis Factor-α Produced by T Cells Reduce the HBV Persistence Form, cccDNA, Without Cytolysis. Gastroenterology *150*, 194–205. https://doi.org/10.1053/j.gastro.2015.09.026.

Xu, Y.-B., Yang, L., Wang, G.-F., Tong, X.-K., Wang, Y.-J., Yu, Y., Jing, J.-F., Feng, C.-L., He, P.-L., Lu, W., et al. (2014). Benzimidazole derivative, BM601, a novel inhibitor of hepatitis B virus and HBsAg secretion. Antiviral Res. *107*, 6–15. https://doi.org/10.1016/j.antiviral.2014.04.002.

Yaginuma, K., and Koike, K. (1989). Identification of a promoter region for 3.6-kilobase mRNA of hepatitis B virus and specific cellular binding protein. J. Virol. *63*, 2914–2920. .

Yan, H., and Li, W. (2015). Sodium taurocholate cotransporting polypeptide acts as a receptor for hepatitis B and D virus. Dig. Dis. Basel Switz. *33*, 388–396. https://doi.org/10.1159/000371692.

Yan, Q., Lan, Y.-H., Huang, Y.-X., Fan, R.-S., Liu, L., Song, S.-P., and Li, Y.-G. (2016). Hepatitis B virus replication is upregulated in proliferated peripheral blood lymphocytes. Mol. Med. Rep. *13*, 3581–3587. https://doi.org/10.3892/mmr.2016.4973.

Yang, W., and Summers, J. (1995). Illegitimate replication of linear hepadnavirus DNA through nonhomologous recombination. J. Virol. *69*, 4029–4036. .

Yang, W., and Summers, J. (1999). Integration of hepadnavirus DNA in infected liver: evidence for a linear precursor. J. Virol. *73*, 9710–9717. .

Yang, S.-H., Lee, C.-G., Park, S.-H., Im, S.-J., Kim, Y.-M., Son, J.-M., Wang, J.-S., Yoon, S.-K., Song, M.-K., Ambrozaitis, A., et al. (2006). Correlation of antiviral T-cell responses with suppression of viral rebound in

chronic hepatitis B carriers: a proof-of-concept study. Gene Ther. *13*, 1110–1117. https://doi.org/10.1038/sj.gt.3302751.

Yang, W., Mason, W.S., and Summers, J. (1996). Covalently closed circular viral DNA formed from two types of linear DNA in woodchuck hepatitis virus-infected liver. J. Virol. *70*, 4567–4575. https://doi.org/10.1128/JVI.70.7.4567-4575.1996.

Yao, Y., Yang, B., Chen, Y., Wang, H., Hu, X., Zhou, Y., Gao, X., Lu, M., Niu, J., Wen, Z., et al. (2019). RNA-Binding Motif Protein 24 (RBM24) Is Involved in Pregenomic RNA Packaging by Mediating Interaction between Hepatitis B Virus Polymerase and the Epsilon Element. J. Virol. *93*, e02161-18. https://doi.org/10.1128/JVI.02161-18.

Yee, J.K. (1989). A liver-specific enhancer in the core promoter region of human hepatitis B virus. Science *246*, 658–661. https://doi.org/10.1126/science.2554495.

You, J., Zhao, Q., Fan, X., and Wang, J. (2019). SOX5 promotes cell invasion and metastasis via activation of Twist-mediated epithelial–mesenchymal transition in gastric cancer. OncoTargets Ther. *12*, 2465–2476. https://doi.org/10.2147/OTT.S197087.

Yu, X., and Mertz, J.E. (1996). Promoters for synthesis of the pre-C and pregenomic mRNAs of human hepatitis B virus are genetically distinct and differentially regulated. J. Virol. *70*, 8719–8726. .

Yu, Y., Schneider, W.M., Michailidis, E., Acevedo, A., Ni, Y., Ambrose, P., Zou, C., Kabbani, M., Quirk, C., Jahan, C., et al. (2019). An RNA-based system to study hepatitis B virus replication and select drug-resistance mutations. BioRxiv 787630. https://doi.org/10.1101/787630.

Yuan, X., Larsson, C., and Xu, D. (2019). Mechanisms underlying the activation of TERT transcription and telomerase activity in human cancer: old actors and new players. Oncogene *38*, 6172–6183. https://doi.org/10.1038/s41388-019-0872-9.

Yuen, M.-F., Schiefke, I., Yoon, J.-H., Ahn, S.H., Heo, J., Kim, J.H., Chan, H.L.Y., Yoon, K.T., Klinker, H., Manns, M., et al. (2019). RNA Interference Therapy with ARC-520 Results in Prolonged HBsAg Response in Patients with Chronic Hepatitis B Infection. Hepatol. Baltim. Md https://doi.org/10.1002/hep.31008.

Yuen, M.-F., Zhou, X., Gane, E., Schwabe, C., Tanwandee, T., Feng, S., Jin, Y., Triyatni, M., Lemenuel-Diot, A., Cosson, V., et al. (2021). Safety, pharmacokinetics, and antiviral activity of RO7049389, a core protein allosteric modulator, in patients with chronic hepatitis B virus infection: a multicentre, randomised, placebo-controlled, phase 1 trial. Lancet Gastroenterol. Hepatol. *6*, 723–732. https://doi.org/10.1016/S2468-1253(21)00176-X.

Yuh, C.H., Chang, Y.L., and Ting, L.P. (1992). Transcriptional regulation of precore and pregenomic RNAs of hepatitis B virus. J. Virol. *66*, 4073–4084. .

Zehender, G., Ebranati, E., Gabanelli, E., Sorrentino, C., Lo Presti, A., Tanzi, E., Ciccozzi, M., and Galli, M. (2014). Enigmatic origin of hepatitis B virus: An ancient travelling companion or a recent encounter? World J. Gastroenterol. WJG *20*, 7622–7634. https://doi.org/10.3748/wjg.v20.i24.7622.

Zhang, C., Wu, M., Zhang, L., Shang, L.-R., Fang, J.-H., and Zhuang, S.-M. (2016). Fibrotic microenvironment promotes the metastatic seeding of tumor cells via activating the fibronectin 1/secreted phosphoprotein 1-integrin signaling. Oncotarget *7*, 45702–45714. https://doi.org/10.18632/oncotarget.10157.

Zhang, D., Cao, L., Li, Y., Lu, H., Yang, X., and Xue, P. (2013). Expression of glioma-associated oncogene 2 (Gli 2) is correlated with poor prognosis in patients with hepatocellular carcinoma undergoing hepatectomy. World J. Surg. Oncol. *11*, 25. https://doi.org/10.1186/1477-7819-11-25.

Zhang, M., Li, G., Shang, J., Pan, C., Zhang, M., Yin, Z., Xie, Q., Peng, Y., Mao, Q., Xiao, X., et al. (2020). Rapidly decreased HBV RNA predicts responses of pegylated interferons in HBeAg-positive patients: a longitudinal cohort study. Hepatol. Int. *14*, 212–224. https://doi.org/10.1007/s12072-020-10015-3.

Zhang, W., Ji, Z., Wang, L., Xiao, D., and Yan, Y. (2015). A meta-analysis of HBsAg-positive rate among general Chinese populations aged 1–-59 years. Infect. Dis. *47*, 878–888. https://doi.org/10.3109/23744235.2015.1064541.

Zhang, W., Chen, J., Wu, M., Zhang, X., Zhang, M., Yue, L., Li, Y., Liu, J., Li, B., Shen, F., et al. (2017). PRMT5 restricts hepatitis B virus replication through epigenetic repression of covalently closed circular DNA transcription and interference with pregenomic RNA encapsidation. Hepatology *66*, 398–415. https://doi.org/10.1002/hep.29133.

Zhang, Y.-Y., Zhang, B.-H., Theele, D., Litwin, S., Toll, E., and Summers, J. (2003). Single-cell analysis of covalently closed circular DNA copy numbers in a hepadnavirus-infected liver. Proc. Natl. Acad. Sci. *100*, 12372–12377. https://doi.org/10.1073/pnas.2033898100.

Zhao, L.-H., Liu, X., Yan, H.-X., Li, W.-Y., Zeng, X., Yang, Y., Zhao, J., Liu, S.-P., Zhuang, X.-H., Lin, C., et al. (2016a). Genomic and oncogenic preference of HBV integration in hepatocellular carcinoma. Nat. Commun. *7*, 12992. https://doi.org/10.1038/ncomms12992.

Zhao, P., Yang, X., Qi, S., Liu, H., Jiang, H., Hoppmann, S., Cao, Q., Chua, M.-S., So, S.K., and Cheng, Z. (2013). Molecular Imaging of Hepatocellular Carcinoma Xenografts with Epidermal Growth Factor Receptor Targeted Affibody Probes. BioMed Res. Int. *2013*, 759057. https://doi.org/10.1155/2013/759057.

Zhao, X.-L., Yang, J.-R., Lin, S.-Z., Ma, H., Guo, F., Yang, R.-F., Zhang, H.-H., Han, J.-C., Wei, L., and Pan, X.-B. (2016b). Serum viral duplex-linear DNA proportion increases with the progression of liver disease in patients infected with HBV. Gut *65*, 502–511. https://doi.org/10.1136/gutjnl-2014-308989.

Zhong, S., Chan, J.Y., Yeo, W., Tam, J.S., and Johnson, P.J. (2000). Frequent integration of precore/core mutants of hepatitis B virus in human hepatocellular carcinoma tissues. J. Viral Hepat. *7*, 115–123. https://doi.org/10.1046/j.1365-2893.2000.00209.x.

Zhou, D.X., and Yen, T.S. (1991). The hepatitis B virus S promoter comprises A CCAAT motif and two initiation regions. J. Biol. Chem. *266*, 23416–23421. .

Zhu, Y., Curtis, M., and Borroto-Esoda, K. (2011). The YMDD and rtA194T mutations result in decreased replication capacity in wild-type HBV as well as in HBV with precore and basal core promoter mutations. Antivir. Chem. Chemother. *22*, 13–22. https://doi.org/10.3851/IMP1791.

Ziemer, M., Garcia, P., Shaul, Y., and Rutter, W.J. (1985). Sequence of hepatitis B virus DNA incorporated into the genome of a human hepatoma cell line. J. Virol. *53*, 885–892. .

Zlotnick, A., Venkatakrishnan, B., Tan, Z., Lewellyn, E., Turner, W., and Francis, S. (2015). Core protein: A pleiotropic keystone in the HBV lifecycle. Antiviral Res. *121*, 82–93. https://doi.org/10.1016/j.antiviral.2015.06.020.

Zoulim, F., and Seeger, C. (1994). Reverse transcription in hepatitis B viruses is primed by a tyrosine residue of the polymerase. J. Virol. *68*, 6–13. .

Zoulim, F., Carosi, G., Greenbloom, S., Mazur, W., Nguyen, T., Jeffers, L., Brunetto, M., Yu, S., and Llamoso, C. (2015). Quantification of HBsAg in nucleos(t)ide-naïve patients treated for chronic hepatitis B with entecavir with or without tenofovir in the BE-LOW study. J. Hepatol. *62*, 56–63. https://doi.org/10.1016/j.jhep.2014.08.031.

Zoulim, F., Levrero, M., and Testoni, B. (2018). Hepatitis B. Nat. Rev. Gastroenterol. Hepatol. https://doi.org/10.1038/s41575-018-0060-3.

German Collection of Microorganisms and Cell Cultures GmbH: Details.

ACKNOWLEDGMENT

Premièrement, je souhaiterais remercier les membres du jury : les rapporteurs, Prof. **Maura DANDRI** et prof. **Patrick SOUSSAN** d'avoir relu et corrigé mon manuscrit, les examinateurs, Dr. **Marintha HEIL** et Prof. **Teresa POLLICINO** d'avoir évalué mon travail de thèse. Un grand merci également au Prof. **Teresa POLLICINO** pour ses conseils lors des comités de suivi de thèse. Je remercie également le Prof. **Fabien ZOULIM** de m'avoir accueilli au sein du laboratoire et du projet RHU, mais également pour son aide et son soutien tout au long de la thèse et des comités de suivi de thèse.

Je tiens à remercier tout particulièrement le Prof. **Massimo LEVRERO** d'avoir été mon directeur de thèse. Merci de m'avoir permis de réaliser ce doctorat et de m'avoir intégrée au sein de l'équipe. Merci de m'avoir transmis vos connaissances, de m'avoir accordé du temps et de m'avoir fait confiance. Les mots sont difficiles à trouver. Alors, mille mercis pour cette expérience !

Mme CHEMIN, je tiens à vous remercier d'avoir accepté ma candidature de stage et de m'avoir soutenue tout au long de ce projet. Je souhaite exprimer toute ma gratitude à **Janet HALL** de m'avoir encadrée dans le début de ma formation de scientifique (stages de M1 et M2). Cette expérience m'a appris tellement de choses et m'a permis de vouloir continuer dans la recherche. Merci pour toute l'aide et tout votre soutien ! Un grand merci également à **Laetitia GEROSSIER** pour la transmission de ses compétences. Je n'aurai pas pu rêver mieux pour apprendre les bases techniques du scientifique. Merci de m'avoir transmis ta méthodologie, ta minutie et ton ingéniosité.

Un grand merci également à **Mirjam ZEISEL** pour toutes les discussions scientifiques et personnelles. Merci pour la relecture du manuscrit. Merci également de ta gentillesse. Cela a été un vrai plaisir de passer ces 4 années de thèses dans la même équipe.

Merci à tous les membres du projet RHU : **Marie-Laure PLISSONNIER** et **Barbara TESTONI** pour leurs précieux conseils, pour les réunions techniques RHU. Vous avez tellement apporté à ce projet ! Merci à **Bernadette VAZ** d'avoir si bien géré le projet RHU, d'avoir été toujours à l'écoute. Merci à **Françoise BERBY** et **Isabelle BORDES** pour votre aide technique dans ce projet et pour toutes les questions relatives au fonctionnement du laboratoire ! Merci aussi à **Christophe COMBET** toujours disponible pour discuter Science mais aussi pour répondre à nos problèmes d'informatiques. Merci aussi à **Delphine**, **Xavier**, **Doohyun** et **Hyoseon** sans qui cette expérience n'aurait pas été la même.

Je souhaite remercier nos collaborateurs de **Roche diagnostics** à Pleasanton, des **Hospices Civiles** de Lyon et de l'Université Claude Bernard dont notamment **Marintha HEIL**, **Ed MARINS**, **Janine FRENCH**, **Carrie LYNN** et **Laurie LAUGIER**.

Un grand merci également à **l'Agence National de Recherche** (ANR) pour le financement de cette thèse et du projet RHU sans qui tout ce travail serait impossible.

Je tiens à remercier tout particulièrement **Christophe SORDILLON** et **Marie-Agnès VITTOZ** pour leur travail, si indispensable au laboratoire mais également pour leur sourire au quotidien.

Je souhaite également présenter mes remerciements à **Francesca GUERRIERI** pour son aide sur le séquençage. Tu as toujours été disponible pour répondre à nos questions. Merci **Claude CARON** de vos précieuses questions pendant les labmeetings, toujours très constructives.

Merci à tous les **annexiens** avec qui j'ai pu partager un repas, une soirée ou boire un thé. Merci aussi à mes collègues de bureaux : **Chloé**, **Anuj**, **Caroline**, **Genia**, **Guada**, **Francesca** et **Vincenzo**, toujours prêts à manger une douceur, à discuter scientifiquement ou personnellement. Vous aurez réussi à égayer mes journées !

Un grand merci à **Francesca CASUSCELLI DI TOCCO** d'avoir été le meilleur binôme mais aussi une grande amie. Merci pour tous ces moments passés ensemble. Je suis très heureuse d'avoir fait ta connaissance ! Merci aussi à **Jenny** pour tous ce que tu fais pour le labo mais aussi et surtout pour ces repas, ces goûters et tous ces moments partagés. Merci à **Océane** d'avoir été là dans les débuts de cette thèse, de m'avoir montré le fonctionnement de l'équipe et aussi merci pour ces bons moments. Merci à **Anaëlle**, **Sarah** et **Léa** pour nos

discussions, conseils et rigolades… Merci aussi à tous **mes amis non-scientifiques** qui m'ont permis de m'évader le temps d'un week-end, d'un verre…

Un grand merci à ma famille, ma raison de vivre !

Merci **Papa** et **Maman** d'avoir fait de moi qui je suis aujourd'hui. Vous avez tellement fait pour moi. J'espère un jour donner à mes enfants tout l'amour, le soutien et l'attention que vous m'avez apportés. Finalement, cette thèse c'est aussi votre travail !

Merci à toi, **David**, pour tout. Tu es mon pillier et le meilleur des frères. Merci aussi à mes **grands-parents** pour tout l'amour qu'ils m'ont donné et donnent (Avó, Tia, este trabalho é também para vós). Merci aussi à mes oncles, tantes et cousins. Merci enfin à **Jean-Baptiste** de partager ma vie, de me soutenir (je sais que ce n'était pas toujours facile).

Je vous aime !

Enfin merci à toutes les personnes qui ont contribué, de près ou de loin, à cette thèse. Tant de monde à remercier et je m'excuse par avance si j'en oublie quelques-uns.