Virus de l'Hépatite B et transcription cellulaire : impact de la protéine HBx et de ses interactions avec les ARNs non-codants

Océane Floriot

N°d'ordre NNT : 2018LYSE1319

THESE de DOCTORAT DE L'UNIVERSITE DE LYON

opérée au sein de

l'Université Claude Bernard Lyon 1

Ecole Doctorale BMIC 340

Biologie moléculaire Intégrative et Cellulaire

Spécialité de doctorat : Virologie

Discipline : Infectiologie

Soutenue publiquement le 18/12/2018, par :

Océane Floriot

Virus de l'Hépatite B et transcription cellulaire : Impact de la protéine HBx et de ses interactions avec les ARNs non-codants

---

Devant le jury composé de :

Mr. **Pr. Zoulim Fabien**, PU-PH, CRCL HCL, Lyon	Président
Mme Dr. **Hibner Urszula** DR INSERM, IGMM, CNRS-UMR5535, Montpellier	Rapporteure
Mr. Dr. **Sureau Camille**, DR CNRS, INST, INSERM U1134, Paris	Rapporteur
Mme Dr. **Strick-Marchand Hélène**, CR Pasteur, Institut Pasteur, Paris	Examinatrice
Mr. **Pr. Levrero Massimo**, PU-PH, CRCL HCL, Lyon	Directeur de thèse

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. le Professeur Frédéric FLEURY**
Président du Conseil Académique	M. le Professeur Hamda BEN HADID
Vice-président du Conseil d'Administration	M. le Professeur Didier REVEL
Vice-président du Conseil Formation et Vie Universitaire	M. le Professeur Philippe CHEVALIER
Vice-président de la Commission Recherche	M. Fabrice VALLÉE
Directrice Générale des Services	Mme Dominique MARCHAND

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur G.RODE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. X. PERROT
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. THEVENARD
Département Chimie Biochimie	Directeur : Mme C. FELIX
Département GEP	Directeur : M. Hassan HAMMOURI
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur G. TOMANOV
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : M. le Professeur J-C PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Y.VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. le Professeur E.PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

Résumé

Le virus de l'hépatite B (VHB) reste un problème de santé majeur dans le monde malgré la disponibilité du vaccin. Le VHB n'est pas éradiqué par les thérapies actuelles et 240 millions de personnes infectées chroniquement restent à risque de développer une cirrhose du foie et un carcinome hépatocellulaire (CHC).

Le VHB est un petit virus hépatotrope doté d'un génome à ADN double brin partiel (ADNrc). Après infection l'ADNrc est converti en ADN épisomal (ADNccc) qui est ensuite organisé en minichromosome viral, qui est le modèle pour la transcription et qui initie la réplication. La protéine de l'hépatite B x (HBx) est recrutée sur l'ADNccc pour initier et maintenir la transcription de l'ADN ccc. HBx cible aussi directement des gènes cellulaires impliqué dans le développement du CHC.

Nous avons utilisé une approche ChIP-Seq pour identifier toutes les cibles génomiques de HBx dans les cellules qui répliquent le VHB. Les cibles HBx sont à la fois des gènes codant les protéines et des ARNnc (75 miARN et 34 lncRNA). Nous avons montré que HBx réprimait un sous-ensemble de miARNs qui réguleraient négativement la réplication virale (ex : miR-24) et des miARNs impliqués dans le développement du CHC (ex : miR-21). Parmi les lncARNs ciblés pour HBx, nous avons étudié DLEU2, qui est fortement surexprimé dans l'infection par le VHB et le CHC. Nous avons en outre montré que DLEU2 lie à la fois HBx et l'histone méthyltransférase Ezh2, la sous-unité catalytique du complexe répressif PRC2. L'interaction avec DLEU2 et HBx relie les fonctions Ezh2/PRC2 conduisant à l'activation constitutive d'un sous-ensemble de gènes cibles d'Ezh2 qui sont normalement conservés dans un état réprimé. Nous avons également montré que l'interaction de HBx avec DLEU2 se produisait sur le minichromosome de l'ADNccc où elle stimulait la transcription/réplication du virus. Enfin, nous avons caractérisé par ATAC-Seq les changements d'accessibilité de la chromatine imposés par HBV dans les hépatocytes humains primaires.

Mots clefs : Virus de l'Hépatite B, HBx, long non-coding RNA.

Abstract

Hepatitis B virus (HBV) remains a major health problem worldwide despite the availability of the vaccine. No cure is available for the 240 million peoples chronically infected with HBV that are at risk to develop liver cirrhosis and hepatocellular carcinoma (HCC). Viral suppression, achieved by long term treatment with nucleotides analogues (NUCs), impacts on liver fibrosis and prevents liver decompensation but HCC risk is not reduced in the first 5 years of treatment.

HBV is a small hepatotropic virus with a partially double strand DNA (rcDNA) genome. After hepatocyte infection the rcDNA is converted into the cccDNA episome that is then organized into a viral minichromosome that is the template for all viral transcripts and initiates replication. The hepatitis B x protein (HBx) is recruited on the cccDNA and is required to launch and maintain cccDNA transcription. HBx has also been shown to directly target cellular genes and this has been related to HCC development.

We used a ChIP-Seq approach to determine the full repertoire of HBx genomic targets in HBV replicating cells. HBx targets include both protein coding genes and ncRNA (75 miRNAs and 34 lncRNAs). We showed that HBx represses a subset of miRNAs that would negatively regulate viral replication (i.e. miR-24) and miRNAs involved in HCC development (i.e. miR-21). Among the HBx targeted lncRNAs we focused DLEU2, which is strongly upregulated in HBV infection and HCC. We further showed that DLEU2 binds both HBx the Ezh2 histone methyltransferase, the catalytic subunit of the repressive PRC2 complex. The interaction with DLEU2 and HBx re-wires Ezh2/PRC2 functions leading to the constitutive activation of a subset of Ezh2 target genes that are normally kept in a repressed state. We also showed that HBx interaction with DLEU2 occurs on the cccDNA minichromosome where it boosts HBV transcription/replication. Finally, we characterized by ATAC-Seq HBV imposed changes of chromatin accessibility in primary human hepatocytes.

Keywords: Hepatitis B virus, HBx, long non-coding RNA.

Table des matières

Liste des Figures

Liste des tableaux

Liste des abréviations

5hmC	5-hydroymethylcytosine
5hmU	5-hydroxymethyluracile
5mC	5-methylcytosine
aa	Acide aminé
ADN	Acide désoxyribonucléique
ADNccc	ADN circulaire cols de manière covalente
ADNrc	ARN relâché circulaire
AFB1	Aflatoxine-B1
ALAT	Transaminase alanine aminotransférases
ARN	Acide ribonucléique
ARNm	ARN messager
ARNpg	ARN prégénomique
ASAT	Aminotransférase aspartique
BCP	Basal core promoteur
BER	Base excision repair
CHB	Chronic hepatitis B (Hépatite B chronique)
CIMP	CpG Island Methylator Phenotype
CpG	Cytosine phosphate Guanine
CTCF	CCCTC-bonding factors
CTD	C-terminal domain
Cter	Carboxy terminal
DHBV	Duck hepatits B virus
DNMT	DNA methylatransferase
DR	Direct repeat
EZH2	Enhancer of zeste homologue 2
GNAT	GCN5-related N-acetyltransferase
HAT	Histones acetyl-transferase
HBcAg	Antigène HBc
HBeAg	Antigène HBe
HBsAg	Antigène HBs
HBV	Hepatitis B Virus
HCC	Carcinome hépatocellulaire (hepatocarcinome cellulaire)
HCV	Hepatitis C virus
HDAC	Histones deacetylases
HNF	Hepatocytes nuclear factors
IFN	Interféron
Kpb	Kilo paire de bases
kDa	Kilodalton
L-HBsAg	Protéine large de l'antigène HBs
lncRNA	Long non coding RNA
LSD	Lysines specific histones demethylases
MBP	Methyl-binding protein
M-HBsAg	Protéine medium de l'antigène HBs
miRNA	Micro ARN

ncRNA	Non coding ARN
NDR	Nucleosome depleted region
NES	Séquence d'exportation nucléaire
NGS	Next generation sequencing
NLS	Séquence de localisation dans le noyau
NPC	Nuclear pore complex
Nt	Nucléotides
NTCP	Sodium taurocholate cotransporting activator
NTD	N-terminal domain
Nter	Amino terminal
NUC	Nucleot(s)ides analogues
ORF	Cadre ouvert de lecture (open reading frame)
p300/CREB	P300 CREB binding protein
pb	Paire de bases
PHH	Hépatocytes primaire humain (primary humain hepatocytes)
piRNA	Piwi-interacting RNA
PKMT	Protein lysine methyltransferases
PRC2	Polycomb repressive complexe 2
PRE	Post-transcriptionnal regulatory element
PRMT	Protein arginine methyltransferase
PTM	Post-traductional modification
ROS	Reactiv oxygen species
RT	Rétrotranscriptase
SAM	S-adénosyl méthionine
SCID	Severe Combined ImmunoDeficiency
S-HBsAg	Protéine small de l'antigène HBs
siRNA	Short-interfering RNA
spARN	ARN épissés
SUMO	Small ubiquitin-like modifier
TET	Ten eleven translocase
TP	Terminal protein
TSS	Transcription start site
uPA	Urokinase-type plasminogen activator
WHBV	Woodchuck hepatitis B virus

Partie 1 : Introduction

Chapitre 1 Le virus de l'hépatite B

I. La structure et les protéines du virus de l'hépatite B

1. Virions et particules subvirales

Les particules virales du virus de l'hépatite B (HBV) sont nommées « particules de Dane » en référence au nom du chercheur qui a été le premier à l'identifier visuellement par microscopie électronique (Dane et al., 1970).

Les particules de Dane ont une taille de 42nm de diamètre (Marion and Robinson, 1983). Ce virus fait partie des virus enveloppés, il possède donc une enveloppe qui contient la capside renfermant le génome viral sous forme d'ADN. L'enveloppe est composée de lipides appartenant à la cellule hôte ainsi que de trois protéines de surface du virus nommées L-, M- et S-HBsAg (Valaydon and Locarnini, 2017). La capside fait 30-32nm de diamètre et possède une forme icosahédrique. Le détail de la composition de l'enveloppe ainsi que de la capside sera décrit plus loin.

Il existe aussi des particules subvirales (Figure 1). Ces particules sont vides, elles ne possèdent ni capsides ni génome. On trouve principalement deux types de particules subvirales : les particules sphériques et les particules en forme de filament (Hu and Liu, 2017). Le type de particules formées est déterminé par le ratio entre les différentes protéines HBsAg du virus (Patient et al., 2007).

Les particules en forme de filaments contiennent la grande forme de la protéine HBsAg que l'on nomme L-HBsAg (L pour « large ») ainsi que la petite forme de HBsAg, appelée S-HBsAg (S pour « small ») (Patient et al., 2007). Elles font 22nm de diamètre et peuvent être de longueur variable (Valaydon and Locarnini, 2017). Les particules sphériques que l'on retrouve chez les patients (Blumberg et al., 1967) font 20-25nm de diamètre et ne contiennent pas L-HBsAg. Ces deux types de particules subvirales ne sont pas infectieuses mais possèdent des propriétés immunogènes et ont été utilisées pour le développement des premiers vaccins non recombinants.

La présence de virions contenant de l'ARN a également été démontrée dans le sérum des patients (Rokuhara et al., 2006; Wang et al., 2016). La fonction dans le cycle viral de ces particules n'est pas encore très bien comprise.

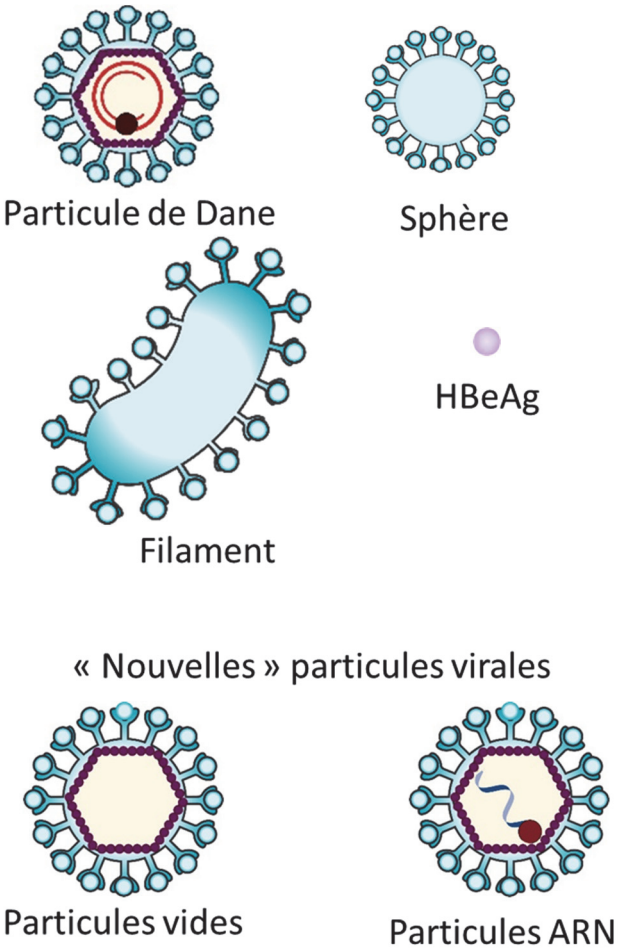


Figure 1 Particules virale et sub-virales du virus de l'Hépatite B (adapté de (Revill et al., 2016; Urban et al., 2014))

2. Structure et organisation du génome

Le HBV possède un génome ADN de 3,2kbp, qui est partiellement double brin et que l'on nomme ADN relaxé circulaire (ADNrc ou rcDNA) (Landers et al., 1977; Summers et al., 1975). Le brin complet est de polarité négative alors que le brin incomplet est de polarité positive. Les deux brins sont complémentaires au niveau de l'extrémité 5' et le brin positif possède un « gap » au niveau de l'extrémité 3' (Valaydon and Locarnini, 2017). Leur complémentarité est possible grâce au chevauchement des deux brins sur une longueur de 200 nucléotides. On retrouve aussi au niveau de ce chevauchement deux séquences répétées (« direct repeat » DR1 et DR2). A l'extrémité 5' du brin (-) est attachée de manière covalente la polymérase virale et à l'extrémité 5' du brin (+) est présente une amorce ARN coiffée d'une longueur de 18 nucléotides (Hu and Seeger, 2015).

Le génome d'HBV possède 4 cadres de lecture ouverts (« open reading frame » - ORF)) se chevauchant qui sont l'ORF PréC/C , l'ORF X, l'ORF P et l'ORF PréS1/PréS2/S. Ces 4 ORFs vont être transcrites à partir de 4 promoteurs distincts Core, X PréS1 et PréS2/S en ARNm afin de

donner les 7 protéines virales. Le génome viral possède aussi deux séquences amplificatrices « enhancer » (Enh I et Enh II) identifiés fonctionnellement (Figure 2).

L'ORF PréC/C est transcrit en ARN prégénomique (ARNpg) et en ARNm PréCore/core de 3,5kb chacun. L'ARNm PréCore/core code pour la protéine de capside Core (HBcAg) la protéine PréCore. L'ARNpg est synthétisé de façon plus abondante dans la cellule et il est également le seul transcrit nécessaire pour déclencher la réplication du génome. Etant donné que l'extrémité 5' du gène P chevauche l'extrémité 3' du gène core, la protéine P est traduite à partir de l'ARNpg lorsque la polymérase II ignore le site de polyadénylation contenue dans l'ORF lors de son premier passage (Tong and Revill, 2016).

L'ORF PréS1/PréS2/S code pour les trois protéines de surfaces L-HBsAg (PréS1/PréS2/S), M-HBsAg (M pour « medium», PréS2/S) et S-HBsAg. Ces protéines sont traduites à partir de deux ARNm. L-HBsAg est traduite à partir d'un ARNm de 2,4kb et les deux autres protéines à partir d'un ARNm de 2,1kb (Valaydon and Locarnini, 2017).

L'ORF X code pour la petite protéine X transactivatrice (HBx), son ARNm possède une taille de 0,7kb. Le promoteur de cette ORF est régulé par les enhancers. L'enhancer I chevauche le promoteur de X et l'enhancer II est localisé en amont du promoteur Core (Hu, 2016).

L'ORF P couvre 70% du génome et code pour la polymérase virale qui possède une activité de reverse transcriptase (RT), de polymérase ainsi que de RNase H.

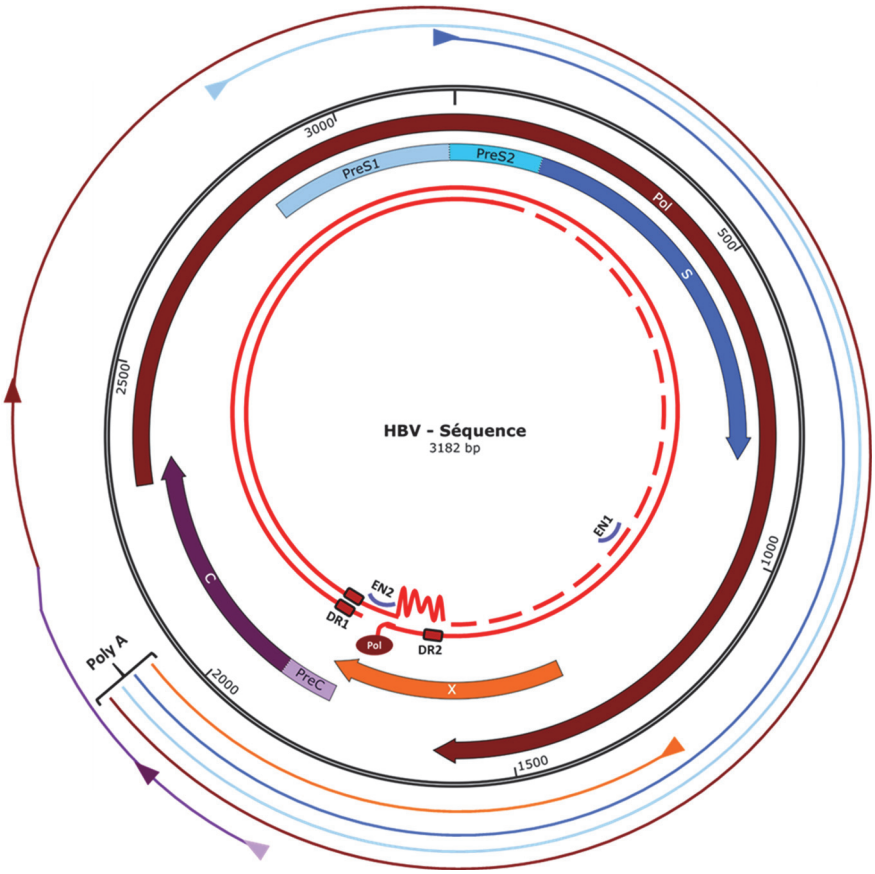


Figure 2 Structure du génome et ses transcrits (adapté de (Datta et al., 2012))

3. Les protéines virales

a. Les protéines d'enveloppe

La protéine S-HBsAg est composée de 226 acides aminées (aa) (Figure 3). Elle comporte quatre domaines hydrophobes transmembranaires (TM 1-4) qui forment des hélices membranaires (Berting et al., 1995). La boucle située entre les domaines TM2 et TM3, elle est aussi la boucle immuno-dominante et est exposée à la surface de l'enveloppe (Valaydon and Locarnini, 2017) (Figure 4).

La protéine M-HBsAg est très similaire à la petite protéine car elle ne se différencie de S-HBsAg que par la présence du domaine hydrophobique PréS2 de 55 aa à son extrémité N-terminale (Figure 3). M-HBsAg ne semble pas avoir d'influence sur la morphologie de la particule virale car elle est présente dans les mêmes proportions dans les différents types de particules (Heermann et al., 1984; Ni et al., 2010).

La plus grande protéine d'enveloppe (L-HBsAg) contient en plus du domaine S et PréS2 le domaine PréS1 de 75 aa (Figure 3). Sa longueur est variable selon les différents génotypes (voir chapitre 2.I.2). Dans sa partie N-terminale, elle possède un site de myristoylation qui est nécessaire à l'ancrage de la protéine dans la membrane du réticulum endoplasmique (Bruss, 1997; Ponsel and Bruss, 2003). La protéine L-HBsAg possède deux typologies. Lors de la conformation i-preS (pour « internal preS »), les domaines PréS1, PréS2 et le domaine TM 1 se situent du côté cytosolique. Pour la conformation e-preS (pour « external preS) les domaines PréS1 et PréS2 sont du côté externe et le domaine TM 1 est transmembranaire (Bruss et al., 1994; Prange and Streeck, 1995) (Figure 4). La grande protéine joue un rôle important dans l'enveloppement de la capside lors de la formation de particules virales et elle est essentielle pour l'infectivité virale.

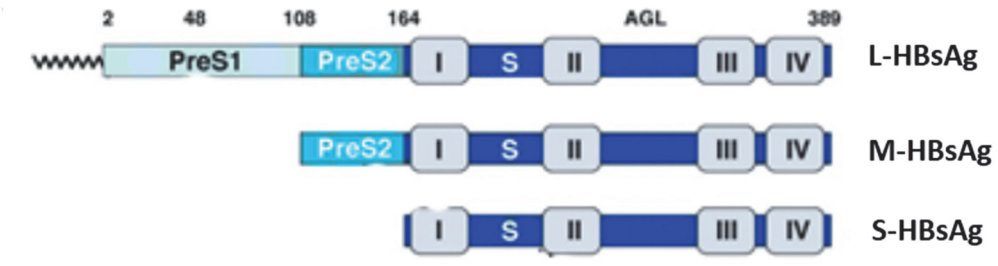


Figure 3 Structure des protéines d'enveloppe (adapté de (Urban et al., 2014))

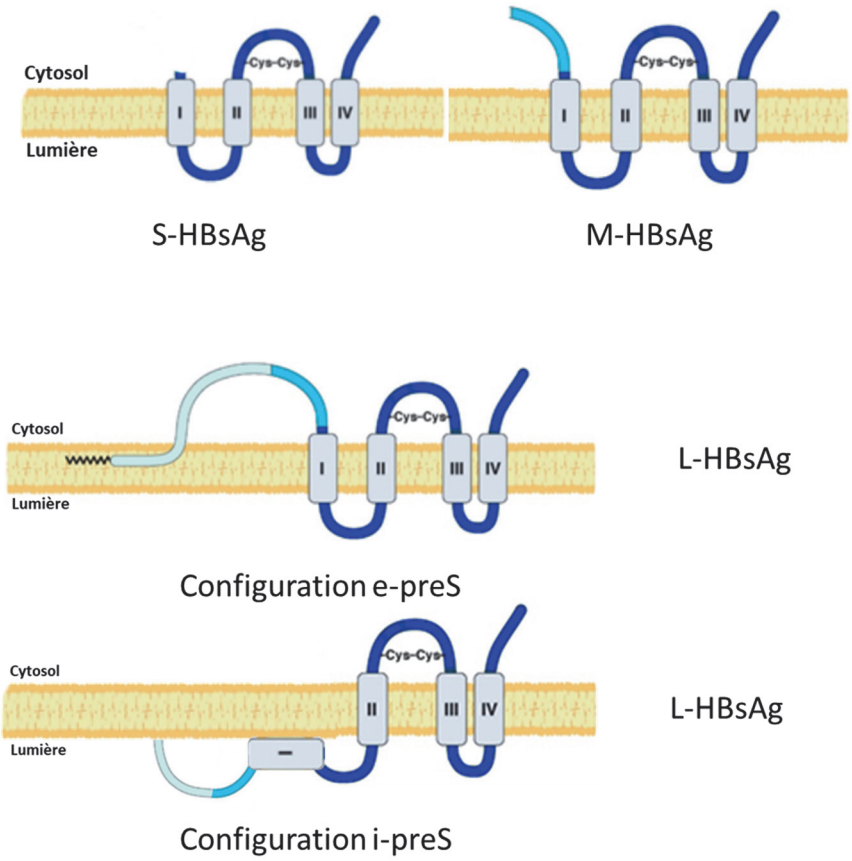


Figure 4 Conformation des différentes protéines d'enveloppe (adapté de(Prange, 2012; Urban et al., 2014))

b. La protéine Core et PréCore

La protéine PréCore

Il y a tout d'abord production d'une protéine PréCore « précurseur » (p25) qui sera directement transportée dans le réticulum endoplasmique grâce à une séquence signal de sécrétion de 29 aa située dans sa partie N-terminale (« N-terminal domain » (NTD)) (Figure 5). Elle sera tronquée de 19 aa en sa partie N-terminale afin de donner la protéine PréCore « intracellulaire » (p22). Puis elle subira une autre coupure du domaine C-terminal (« C-terminal domain » (CTD)) qui produira l'antigène HBe (HBeAg) d'une taille de 149 aa (18kDa) (Milich and Liang, 2003; Walsh and Locarnini, 2012). Une population d'HBeAg de taille hétérogène sera ensuite sécrétée. L'HBeAg est très conservée entre les différents génotypes, mais n'est pas indispensable pour la réplication virale. Par contre cette protéine semble jouer un rôle important pour l'établissement de l'infection persistante en régulant les réponses immunes de l'hôte (Hu, 2016; Milich and Liang, 2003). La positivité de l'HBeAg permet d'identifier les patients présentant une infection aiguë et les premières phases des infections chroniques (voir chapitre sur l'histoire naturelle de l'infection par HBV).

La protéine Core

La protéine Core (HBcAg), est la protéine formant la capside du virus. Sa longueur est de 183 aa (soit 21kDa). On distingue dans cette protéine principalement deux domaines, le domaine NTD et le domaine CTD.

Le NTD est composé par les 140 premiers résidus, il permet à lui seul l'assemblage des capsides. Les capsides sont composées de 180 ou 240 monomères qui vont former une structure icosaédrique que l'on nomme T=3 ou T=4. La forme majoritaire dans les virions est la structure T=4 (Birnbaum and Nassal, 1990; Wynne et al., 1999).

Le CTD ou « domaine protéamine » possède un taux élevé de résidus arginine. Ceci lui permet d'interagir avec les acides nucléiques, notamment avec l'ARNpg (Nassal, 1992; Zlotnick et al., 1997). Il contient une séquence de localisation nucléaire (« nuclear localization sequence » (NLS)) et une séquence d'exportation nucléaire (« nuclear export sequence » (NES)) (Eckhardt et al., 1991; Yeh et al., 1990). Le CTD comporte plusieurs sites de phosphorylation, dont 3 sites majeurs qui sont présents sur des résidus sérines (Lan et al., 1999). La phosphorylation de ce domaine va jouer un rôle crucial lors de la réplication du virus que nous verrons par la suite.

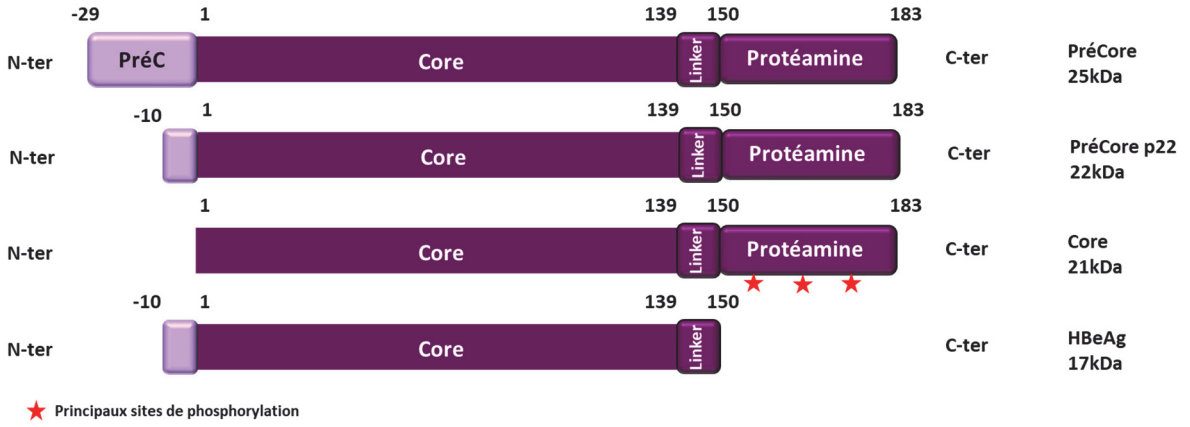


Figure 5 Structure des protéine PréCore, Core et HBeAg (adapté de (Steven et al., 2005))

c. La polymérase virale

La polymérase virale est la plus grande des protéines d'HBV, elle comporte 832 aa. On distingue 4 domaines différents : le domaine « terminal protein» TP, le « spacer », le domaine de reverse transcriptase et le domaine RNase H (Hu, 2016) (Figure 6).

Le domaine TP se lie de façon covalente au brin (-) à l'extrémité 5' du génome d'HBV. Ce domaine est aussi nécessaire pour la liaison à la boucle ε présente dans l'ARNpg, l'encapsidation de l'ARNpg et la synthèse des protéines (Bartenschlager and Schaller, 1988; Bartenschlager et al., 1990; Wang et al., 1994).

Le domaine « spacer » est la région la moins conservée. Sa fonction exacte n'est pas connue mais trois résidus cystéine localisés dans le domaine C-terminal et un résidu dans le domaine N-terminal du RT sont nécessaires pour l'encapsidation de l'ARNpg (Jones and Hu, 2013).

Le domaine RT est subdivisé en 7 sous domaines notés de A à G. Dans le sous-domaine C on trouve un motif tyrosine-méthionine-aspartate-aspartate qui est très conservé. Ce motif est essentiel pour l'activité d'ADN polymérase ADN dépendante (Jones et al., 2012). D'autre part le domaine RT possède dans le sous domaine E une activité de reverse transcriptase, , qui est importante pour la réplication et plus particulièrement dans l'interaction entre la boucle ε et la polymérase, ainsi que dans la polymérisation de l'ADN (Jones and Hu, 2013).

Le domaine RNase H se situant dans la partie C-terminale de la protéine permet la dégradation de l'ARNpg pendant la synthèse du brin (-) d'ADN (Hu, 2016).

Figure 6 Structure de la Polymérase virale (adapté de (Tong and Revill, 2016))

d. La protéine X

La protéine HBx (154 aa) est la plus petite protéine du virus (17kDa) (Benhenda et al., 2009; Slagle and Bouchard, 2016). La structure cristallographique de cette protéine n'a pas encore été déterminée mais différents domaines ont été prédits par des études de modélisation 3D (Figure 7). Cette protéine est conservée chez tous les *Hepadnavirus* mammifères mais pas chez les *Hepadnavirus* aviaires. Elle peut être localisée dans le cytoplasme et dans le noyau. HBx est nécessaire à la réplication virale et semble contribuer de manière importante à la pathogénèse virale (Benhenda et al., 2009; Lucifora et al., 2011). HBx se lie à la protéine « DNA damage-binding protein 1 » (DDB1) et recrute l'ubiquitine ligase CRL4. Le complexe CRL4-DDB1 ubiquitine les protéines Smc5/6 du complexe de maintenance structural des chromosomes ce qui va induire sa dégradation par le protéasome et l'activation de l'ADN circulaire clos de manière covalente (ADNccc). En l'absence de HBx les complexes contenant Smc5/6 inhibent la transcription de l'ADNccc et la production de l'ARNpg (Decorsière et al., 2016). HBx peut induire d'autres mécanismes afin de favoriser la réplication d'HBV comme la répression de la « DNA methyltransferase » (DNMT) DNMT3A qui va induire l'expression du miR-101(Wei et al., 2013) et donc l'induction de l'endocytose et l'autophagie, qui sont nécessaires à la réplication virale HBx jouent aussi un rôle dans la régulation de l'expression des gènes de l'hôte soit en se liant directement à des régions promotrices des gènes cellulaires (voir première étude), soit en activant plusieurs voies de signalisation cellulaires et processus cellulaires, notamment les voie de signalisation des dommages à l'ADN, du calcium, de l'apoptose, et du cycle cellulaire (Benhenda et al., 2009; Slagle and Bouchard, 2016; Wei et al., 2010).

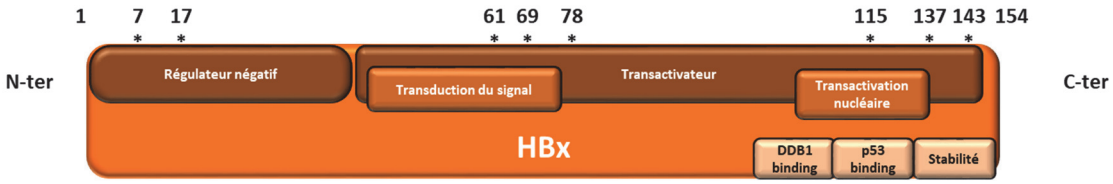


Figure 7 Structure de la protéine HBx (adapté de (Slagle and Bouchard, 2016))

II. Le cycle de réplication du virus de l'hépatite B

1. Entrée et migration au noyau

a. L'entrée du virus

La première interaction entre le virus et la cellule hôte se fait *via* l'enveloppe du virus. Des études ont montré que les régions S et PréS1 mais pas PréS2 jouaient un rôle crucial lors de l'infection (Bremer et al., 2011; Le Seyec et al., 1998; Ni et al., 2010; Salisse and Sureau, 2009). Dans la région S certaines boucles antigéniques, qui sont localisées entre les régions transmembranaires II et III, sont essentielles pour l'infection virale (Barrera et al., 2005; Bremer et al., 2011; Engelke et al., 2006; Glebe and Urban, 2007).

Le virus s'attache à la surface des hépatocytes *via* les héparanes sulfate protéoglycanes (HSPG) (Leistner et al., 2008; Schulze et al., 2007; Sureau and Salisse, 2013). Un criblage par ARN interférants a permis de déterminer le rôle du glypican 5 (GPC5) dans ce processus (Verrier et al., 2016). Le sodium taurocholate cotransporting polypeptide (NTCP) (Yan et al., 2012) un membre de la famille des gènes *SLC10A*, est le récepteur d'entrée du HBV. C'est un transporteur des acides biliaires sodium dépendant et son expression est exclusivement hépatique. Il est localisé au niveau de la membrane basolatérale des hépatocytes (Li, 2015; Watashi and Wakita, 2015) (Figure 8). L'expression spécifique d'espèce et de tissu du NTCP explique la spécificité d'infection des hépatocytes humains par HBV (Ni et al., 2014).

Le virus est internalisé par un mécanisme d'endocytose pas encore entièrement caractérisé. Suite à l'internalisation, la capside virale sera libérée dans le cytoplasme.

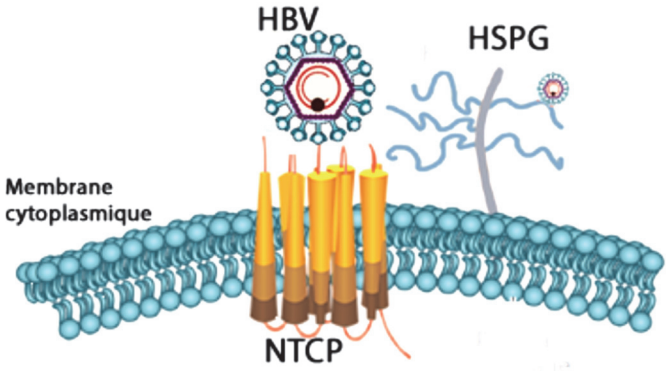


Figure 8 Attachement du virus aux récepteurs d'entrée (adapté de (Li, 2015; Revill et al., 2016))

b. Traffic intracellulaire et entrée dans le noyau

Une fois libérée de son enveloppe, la capside qui contient l'ADNrc est transportée au noyau. Les étapes de transport au noyau ne sont pas encore bien déterminées en raison de l'absence de modèles expérimentaux permettant d'étudier ce processus. La protéine Core possède dans son domaine CTD une séquence NLS permettant à la capside d'être dirigée vers le noyau de la cellule (Hu, 2016). Deux modèles de transport et de la libération du génome sont actuellement discutés. Le premier semblerait être médié par la polymérase qui est attachée à l'ADNrc (Kann et al., 1997). Dans ce cas les capsides seraient disloquées dans le cytoplasme au niveau des complexes de pores nucléaires (« nuclear pore complexe », NPC). Le génome serait transloqué dans le noyau via les importines α/β liés à la Pol. D'autre part les dimères de Core seraient transportés via l'importin-β. Dans le deuxième modèle les capsides ne seraient déstabilisées que dans le panier nucléaire (« nuclear basket ») après l'interaction de Core avec Nup-153, ce qui libèrerait le génome viral dans le noyau (Gallucci and Kann, 2017).

2. De la conversion en ADNccc à la synthèse des protéines

a. La conversion de l'ADNrc en ADN circulaire clos covalent

Dans le noyau l'ADNrc est converti en ADNccc. Le processus de chromatinisation de l'ADNccc donne lieu à un minichromosome qui est la matrice transcriptionnelle d'HBV (Figure 9). Cette étape est très rapide. En effet, l'ADNccc peut être détecté seulement quelques heures après infection dans les cellules (Hu and Seeger, 2015) (observations non publiées Testoni/Zoulim et Floriot/Levrero).

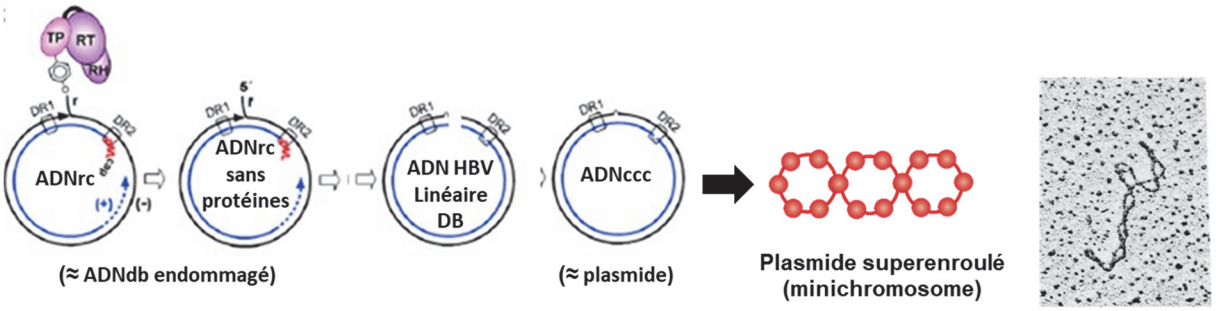


Figure 9 Conversion de l'ADNrc en ADNccc et finalement en minichromosome (adaptée de (Königer et al., 2014))

Rappelons que le brin (-) possède en son extrémité 5' une liaison avec la polymérase virale. Le brin (+) n'est pas complet à son extrémité 3' et possède un oligomère d'ARN en son extrémité 5' (Gómez-Moreno and Garaigorta, 2017). Le détachement de la polymérase virale de l'ADNrc conduit à un ADNrc sans protéines (« protein-free rcDNA », PF-rcDNA). L'oligomère d'ARN ainsi que les séquences redondantes du brin (-) sont ensuite éliminés puis le brin (+) est réparé avant que les deux brins ne soient ligués (Nassal, 2015).

La présence de la polymérase virale au niveau des capsides entrant dans le noyau après infection *de novo* a initialement suggéré que la polymérase virale pourrait être impliquée dans êe processus de réparation de l'ADNrc. Cependant il a été montré que la polymérase virale n'est pas nécessaire pour l'accumulation de l'ADNccc lors de l'infection *de novo*. L'hypothèse la plus acceptée est que la formation de l'ADNccc dépend de facteurs cellulaires normalement impliqués dans la détection et la réparation de l'ADN endommagé. Différentes études ont ainsi suggéré un rôle pour TDP2, FEN1, PRPF31, DNA Pol κ dans ce processus (Kinoshita et al., 2017; Kitamura et al., 2018; Königer et al., 2014; Qi et al., 2016).

Les mécanismes exacts de la prochaine étape de conversion de l'ADNcc en minichromosome viral ne sont pas encore bien démontrés mais il a récemment été montré que la protéine chaperonne Hira pourrait jouer un rôle important dans ce processus (Locatelli et al., 2018).

b. La transcription et traduction à partir de l'ADNccc

L'ADNccc formé et associé aux histones est utilisé comme matrice pour la transcription des ARN viraux grâce à l'ARN polymérase II de l'hôte. La transcription virale met en jeu des facteurs de transcription cellulaires ubiquitaires et des facteurs de transcription spécifiques du foie comme les facteurs nucléaires hépatocytaires (« hepatocytes nuclear factors » - HNFs) (Quasdorff and Protzer, 2010). Comme mentionné précédemment, la protéine virale HBx joue un rôle important dans la régulation de la transcription de l'ADNccc. En plus d'induire la destruction des facteurs de restriction Smc5/6, elle est recrutée sur le minichromosome viral et régule à son tour le recrutement de plusieurs co-activateurs et co-répresseurs de transcription (Belloni et al., 2009; Decorsière et al., 2016; Lucifora et al., 2011; Rivière et al., 2015; Zhang et al., 2017).

Les différents ARN viraux ont des sites d'initiation distincts, ils sont polyadénylés en leur extrémité 3' et possèdent une coiffe (« cap ») en leur extrémité 5' (Schreiner and Nassal, 2017). Les ARNs sont transportés dans le cytoplasme grâce à une séquence spécifique le « post-transcriptional regulatory element » (PRE) qui est codée par des séquences ADN qui chevauchent l'enhancer I (Hu, 2016). Plusieurs ARN épissés (spARN) ont été décrits, mais leur fonction et leur rôle dans le cycle viral ne sont pas encore complètement compris, à l'exception des spARN qui codent pour la protéine SP1 (Figure 10) (Bayliss et al., 2013).

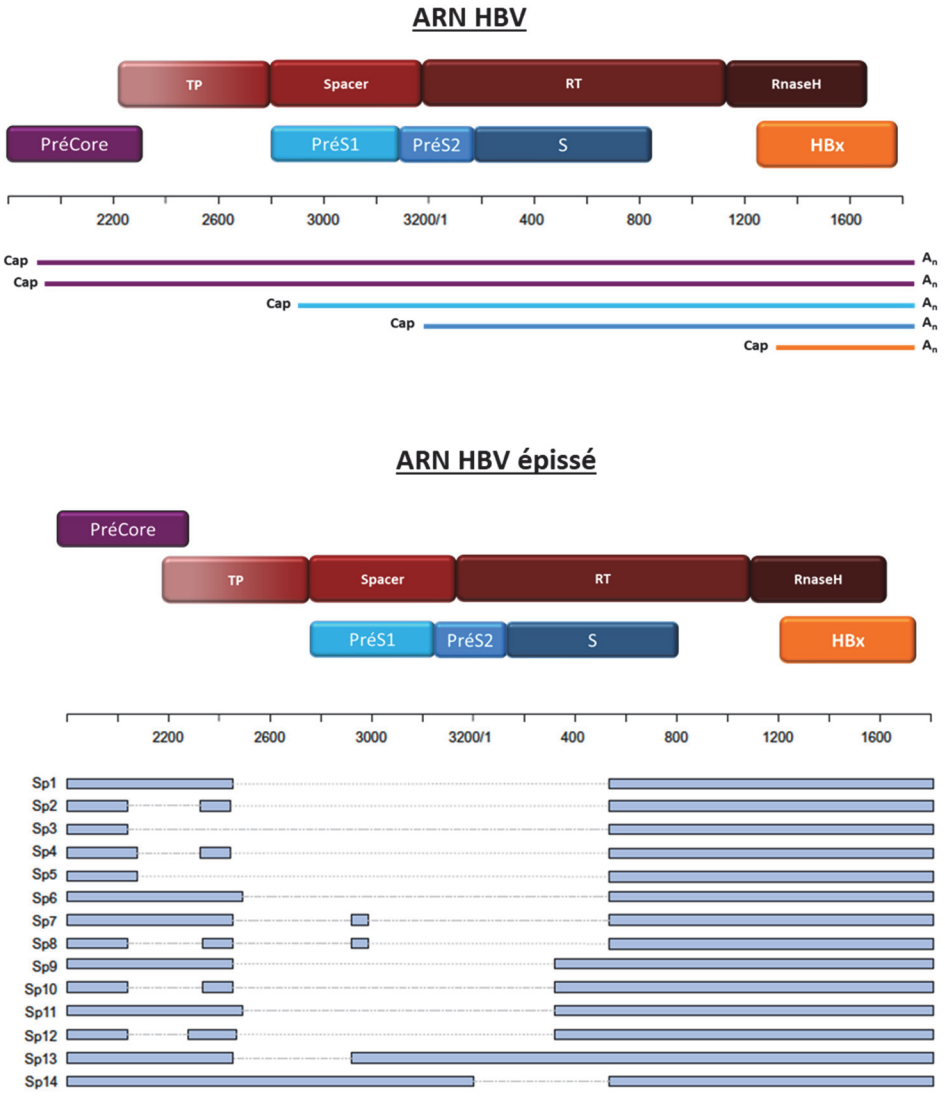


Figure 10 Schéma représentant les différents ARN HBV ainsi que des ARN épissés (adapté de (Bayliss et al., 2013))

3. Régulation épigénétique de l'ADNccc

L'organisation de l'ADNccc en tant que minichromosome viral le soumet aux mêmes modifications et régulations épigénétiques que les chromosomes de l'hôte (Figure 11). Des études biochimiques ont montré la présence sur l'ADNccc des histones H3 et H2B en forte proportion en comparaison des histones H4, H2A et H1 (Bock et al., 2001). L'adaptation de la technique d'immunoprécipitation de la chromatine (ChIP) à l'étude de l'ADNccc a permis de confirmer la présence des histones H3 et H4 et des protéines virales HBx et HBc sur le minichromosome viral et, au fil du temps, d'identifier le recrutement de plusieurs modificateurs de la chromatine de l'hôte, ainsi que de caractériser de nombreuses modifications post-traductionnelles des histones liées à l'ADNccc (Belloni et al., 2009; Hong et al., 2017; Levrero et al., 2009)

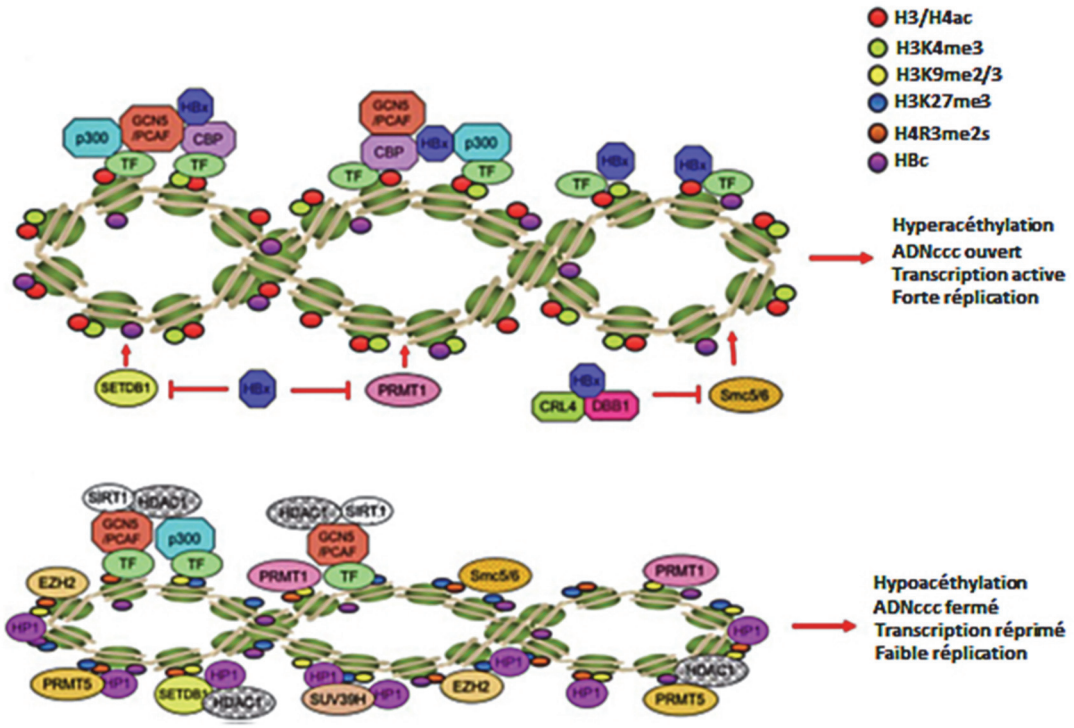


Figure 11 L'ADNccc et ses modifications épigénétiques (adapté de (Hong et al., 2017))

a. La méthylation de l'ADNccc

L'identification de 3 îlots CpG (« Cytosine phosphate Guanine ») putatifs dans la séquence de l'ADNccc (voir le chapitre 3.I.1. sur la méthylation de l'ADN) a suggéré que la méthylation de l'ADNccc pourrait contribuer au contrôle des fonctions de l'ADNccc et à la réplication de l'HBV. Le premier îlot CpG chevauche le site d'initiation du gène S (îlot I), le deuxième se situe sur la région incluant l'enhancer I et le promoteur du gène X et Core (îlot II) et le troisième se situe sur la région contenant le promoteur de préS1 et le codon stop du gène de la polymérase

(Kaur et al., 2010; Vivekanandan et al., 2008). Zhang *et al.* ont montré que, l'îlot I est peut méthylé, que la méthylation de l'îlot II est corrélée à une faible virémie et que la méthylation de l'îlot III est associée à un faible taux d'HBsAg dans le sérum de patients chroniquement infectés (Zhang et al., 2013). L'hyperméthylation de l'îlot II proche du promoteur Core peut être responsable d'une réduction de la transcription de l'ARN 3.5kb/PréC/C et donc de l'expression de HBeAg chez les patients chroniquement infectés (Guo et al., 2009). En raison de la grande stabilité de l'ADNccc il peut être transmis aux cellules filles lors de la division cellulaire tout en gardant son profil de méthylation, comme c'est le cas pour l'ADN génomique des cellules hôte. Au fil du temps et au cours de plusieurs cycles de division cellulaire, l'ADNccc peut acquérir d'autres méthylations expliquant ainsi la fréquence des méthylations de l'ADNccc chez les patients chroniquement infectés (Kim et al., 2011). L'hyperméthylation de l'îlot II proche du promoteur « core » peut inhiber la transcription de PréC/C et donc l'expression de l'HBeAg chez les patients chroniquement infectés (Guo et al., 2009).

On ne sait pas quelle DNMT est impliquée dans la méthylation du génome de HBV. Cependant, des études ont montré que l'expression de DNMT1, DNMT2 et DNMT3 est augmentée en présence de HBV mais la raison de cette induction qui est probablement responsable de l'hyperméthylation de l'ADNccc et donc d'une réduction de la réplication virale n'est pas bien comprise (Vivekanandan et al., 2009).

b. Les modifications post-traductionnelles des histones liées à l'ADNccc

L'impact des modifications post-traductionnelles (« post-translational modification » – PTMs) des histones liées à l'ADNccc sur la réplication du virus a été étudié par plusieurs groupes de recherche en utilisant d'abord la technique de l'ADNccc ChIP et plus récemment des techniques de ChIP-Seq adaptées à l'ADNccc (voir le chapitre 3.I.2.). En termes généraux, une hypoacétylation des histones liées à l'ADNccc se traduit par une chromatine ADNccc compactée avec une réduction de la transcription et de la réplication virale (Figure 11). Dans ce cas, l'ADNccc est lié à différentes classes d'histones déacétylases et a plusieurs méthyltransférases des histones à activité répressive (Figure 11). Au contraire, l'hyperacétylation des histones liées à l'ADNccc, associée au recrutement de plusieurs acétyltransférases, ouvre la chromatine de l'ADNccc en permettant des niveaux élevés de transcription et de réplication du HBV (Figure 11). Ce modèle a été confirmé *in vivo* à l'aide de biopsies du foie de patients atteints d'hépatite B chronique (Pollicino et al., 2006). Plus récemment, le groupe de Tropberger, en appliquant une approche cccDNA ChIP-Seq dans différentes lignées cellulaires ainsi que dans des hépatocytes primaires humains (PHH) infectées a confirmé que la réplication du HBV est associée à la présence d'un taux élevé de PTMs que l'on associe à une transcription active (H3K4me3, H3K27ac et H3K122ac) et très peu de PTMs répressives (Tropberger et al., 2015). La méthyltransférases « Enhancer of zeste homologue 2 » (EZH2), qui fait partie du complexe répressif « polycomb repressive complexe 2 » (PRC2) et qui est responsable de la marque répressive K27me3 sur l'histone H3, ne se trouve pas sur l'ADNccc lorsque le virus se réplique mais son inhibition conduit à

l'augmentation de la transcription de l'ADNccc et de la production des antigènes HBeAg et HBsAg (Zhang et al., 2017). Par ailleurs, la méthyltransférases « protein arginine methyltransferase 5 » (PRMT5) induit d'une part la méthylation de l'arginine en position 3 de l'histone 4 liée à l'ADNccc pour réprimer sa transcription et d'autre part interagit avec HBc et le complexe de remodelage de la chromatine Brg1 hSWI/SNF induisant l'inhibition de la liaison entre l'ARN Pol II et l'ADNccc.

Le recrutement d'HBx sur l'ADNccc est corrélé à la transcription de l'ADNccc et en l'absence de HBx sur l'ADNccc il y a une hypoacétylation des histones liées à l'ADNccc (Belloni et al., 2009) et le dépôt des marques répressives H3K9me2 et H3K9me3 par la méthyltransférases « SET domain bifurcated 1 » (SETDB1) (Rivière et al., 2015). HBx est nécessaire pour maintenir une configuration ouverte de la chromatine de l'ADNccc médiée par le recrutement des acétyltransférases PCAF/GCN5, p300 et CBP. L'activation de la transcription de l'ADNccc *via* HBx passe par le recrutement de PCAF/GCN5, p300, et l'acétyltransférase CBP. En même temps HBx inhibe le complexe protein phosphatase1/ « histone deacetylase 1 » (HDAC1) (Rivière et al., 2015).

La régulation épigénétique de l'activité transcriptionnelle de l'ADNccc peut être modulée par la cytokine inflammatoire IL6, fortement exprimée dans le foie des patients atteints d'hépatite B chronique. Le traitement des cellules qui répliquent le HBV par l'IL-6 réduit l'acétylation des histones liées à l'ADNccc et sa transcription sans affecter sa chromatinisation. Du point de vue mécanistique, l'IL-6 inhibe la liaison des facteurs de transcription HNF1α et HNF4α à l'ADNccc et relocalise STAT3, le facteur de transcription qui médie les activités de l'IL6, de l'ADNccc aux gènes cible cellulaires de l'IL-6 (Palumbo et al., 2015).

Remarquablement, le traitement des cellules qui répliquent le HBV avec des activateurs des acétyltransférases P300/CBP ou des inhibiteurs des histones déacétylases augmente l'acétylation des histones H4 liées à l'ADNccc et potentialise à la fois la transcription de l'ADNccc et la réplication du virus (Pollicino et al., 2006; Tropberger et al., 2015).

En outre, Belloni *et al.* ont démontré *in vitro* et *in vivo* qu'un traitement à l'IFN-α va inhiber la transcription de l'ADNccc via un mécanisme épigénétique impliquant le complexe répressif PRC2. Les histones liées à l'ADNccc sont hypoacéthylées et les protéines EZH2 et YY1, membres du complexe PRC2, ainsi que les histone déacétylases HDAC1 et hSirt sont recrutées sur l'ADNccc (Belloni et al., 2012).

4. Encapsidation de l'ARNpg et maturation des capsides

a. L'encapsidation de l'ARNpg et la rétrotranscription

In vivo l'assemblage des nucléocapsides d'HBV est initié par la formation d'un complexe entre l'ARNpg et la polymérase virale, un pré-requis pour l'encapsidation de l'ARNpg. La polymérase virale se lie à la structure tige boucle nommée *epsilon (*ε) qui se situe à l'extrémité 5' de

l'ARNpg (Bartenschlager et al., 1990; Hirsch et al., 1990; Junker-Niepmann et al., 1990). La structure ε va induire des modifications structurales entre la polymérase et l'ARNpg. L'encapsidation nécessite des facteurs de l'hôte comme les chaperonnes (hsp90) qui vont permettre l'initiation de l'interaction entre la polymérase et ε (Seeger and Mason, 2015).

Une fois traduite, la protéine Core forme des dimères. La dimérisation est obtenue par formation de 4 faisceaux d'hélices avec deux hélices de chaque moitié de dimères (Zlotnick et al., 2015). On retrouve deux formes de capsides dans le cytoplasme des cellules. La première forme, nommée T=3, d'une taille de 30 nm contient 90 dimères, la deuxième forme, T=4, qui est la plus répandue, contient 120 dimères pour une taille de 34 nm.

Des motifs spécifiques sur l'ARNpg médient son interaction avec le domaine C-ter de la protéine Core. Cette interaction pourrait jouer un rôle dans l'organisation de l'ARNpg pendant l'assemblage de la nucléocapside (Patel et al., 2017).

Lors de la première étape de la réplication du génome la boucle ε va servir de matrice pour la création d'une amorce ADN de 3-4 nucléotides (nt) qui va être liée de manière covalente à la polymérase virale via un résidu tyrosine dans son domaine TP (Figure 12) (Bartenschlager and Schaller, 1988; Wang and Seeger, 1992; Zoulim and Seeger, 1994). Ce brin d'ADN va ensuite être transféré à la séquence complémentaire présente sur l'extrémité 3' de l'ARNpg (« first primer shift ») (Datta et al., 2012). La reverse transcription a lieu, le brin (-) est synthétisé de 5' en 3' et simultanément l'ARNpg est dégradé par la polymérase virale *via* son activité RNase H (Summers and Mason, 1982). Dans le même temps à l'extrémité 5' coiffé de l'ARNpg quelques nucléotides ne vont pas être dégradés. Ils vont servir d'amorce pour la synthèse du brin (+) en relocalisant les paires de bases complémentaires à la région 5' DR2 du brin(-) d'ADN (« second primer shift ») (Hu and Seeger, 2015). Un transfert de brin s'effectue lorsque que le brin (-) servant de matrice est terminé (échange de matrice). Ceci est possible grâce aux séquences redondantes présentes dans le brin (-). La circularisation du génome se fait par l'appariement des séquences complémentaires sur l'ADN naissant du brin (+) et la partie redondante de l'extrémité 3' du brin (-). Ceci permet aussi au brin (+) d'avoir accès au reste du brin (-) pour s'en servir de matrice pour l'élongation (Will et al., 1987). A la fin du processus l'ADNrc sera formé.

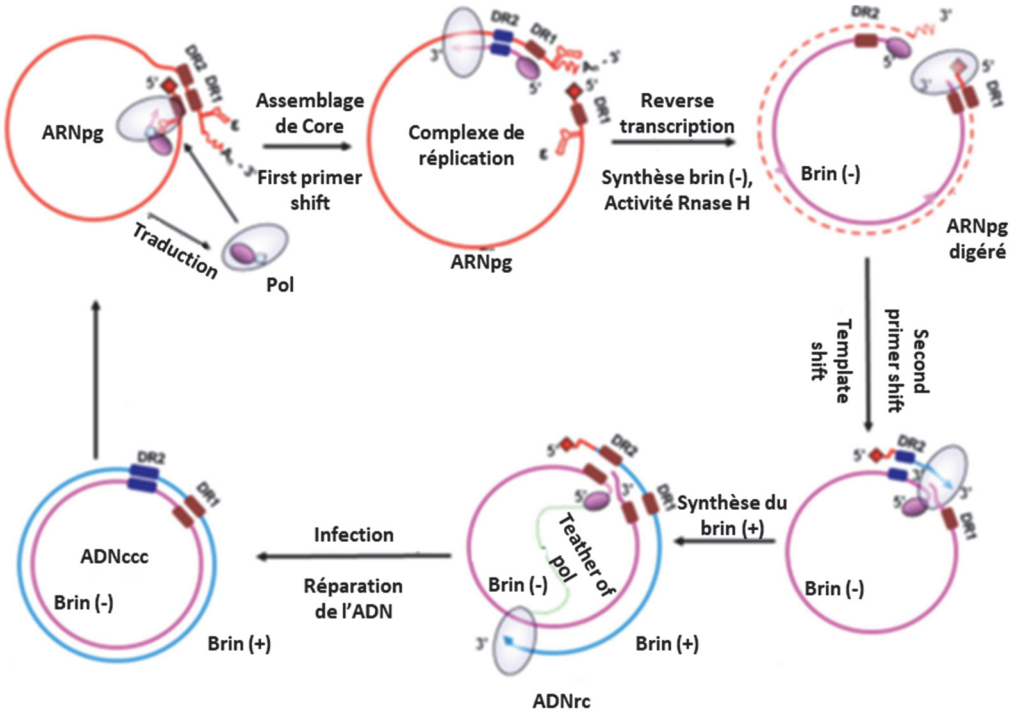


Figure 12 Schéma de la rétrotranscription de l'ARNpg en ADNrc puis en ADNccc (adapté de Datta et al., 2012))

b. La maturation des capsides

La maturation des capsides se fait en parallèle des différentes étapes de la réplication virale: reverse transcription et formation du brin ADN (-) (capsides immatures); synthèse du brin (+) et formation de l'ADNrc (capsides matures) (Figure 13). L'état de phosphorylation du CTD de Core influence l'organisation de l'ARNpg dans les nucléocapsides (Wang et al., 2012) (Figure 13). Le CTD est fortement phosphorylé dans les nucléocapsides immatures, ce qui est requis pour faciliter la synthèse du brin (-), mais par la suite il sera sous forme non phosphorylée dans les nucléocapsides matures ce qui est nécessaire à la synthèse du brin (+) et pour stabiliser les nucléocapsides quand l'ADNrc est synthétisé (Basagoudanavar et al., 2007; Liu et al., 2015a; Ning et al., 2017; Perlman et al., 2005). Des mutations de résidus spécifiques à la surface de la capside, appelé « matrix binding domain » (MBD), retardent l'incorporation de capsides immatures dans les enveloppe naissantes et la formation des virions tout en permettant la synthèse de l'ADNrc (Pairan and Bruss, 2009; Ponsel and Bruss, 2003).

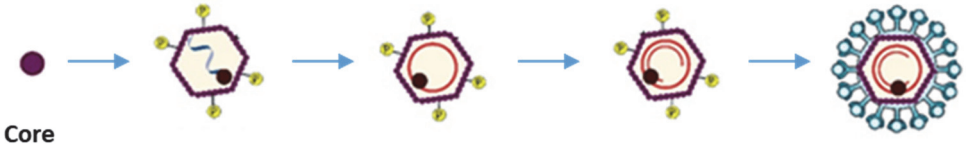


Figure 13 Etape de maturation des capsides ainsi que des états de phosphorylation (adapté de (Ning et al., 2017; Revill et al., 2016))

Une partie des nucléocapsides matures est à nouveau transportée vers le noyau afin d'amplifier le réservoir (« pool ») d'ADNccc. L'efficacité de ce phénomène de recyclage dépend fortement du modèle cellulaire et des génotypes étudiés (Hu, 2016).

5. Enveloppement et sécrétion

Les trois protéines d'enveloppe possèdent des domaines transmembranaires qui vont les aider à se loger dans le réticulum endoplasmique (Selzer and Zlotnick, 2015) (Figure 14). Quand L-HBsAg est dans sa conformation i-preS1, il y a une région conservée dans une boucle du domaine S qui est importante pour l'interaction avec les capsides et donc l'enveloppement des virions (Gripon et al., 1995; Lepère-Douard et al., 2009). La sécrétion des particules virales semble être médiée par les corps multivésiculaires (« multivesicular bodies » (MBV)).

Les virions complets contenant l'ADNrc ne représentent que 1% des particules sécrétées. L'utilisation de mutants a démontré que des virions sont produits en l'absence complète d'encapsidation de l'ARNpg ou de synthèse d'ADN (Hu and Liu, 2017). Les nucléocapsides vides sont aussi moins stables que celles contenant l'ADNrc (Cui et al., 2013).

Par ailleurs des particules contenant de l'ARN viral ont été mis en évidence dans le sérum des patients (Rokuhara et al., 2006). Il semblerait que ces particules virales contiennent de l'ARNpg qui peut être partiellement rétrotranscrit (Wang et al., 2016). Ces ARN viraux circulants pourraient être utilisés comme biomarqueurs chez les patients (van Campenhout et al., 2018).

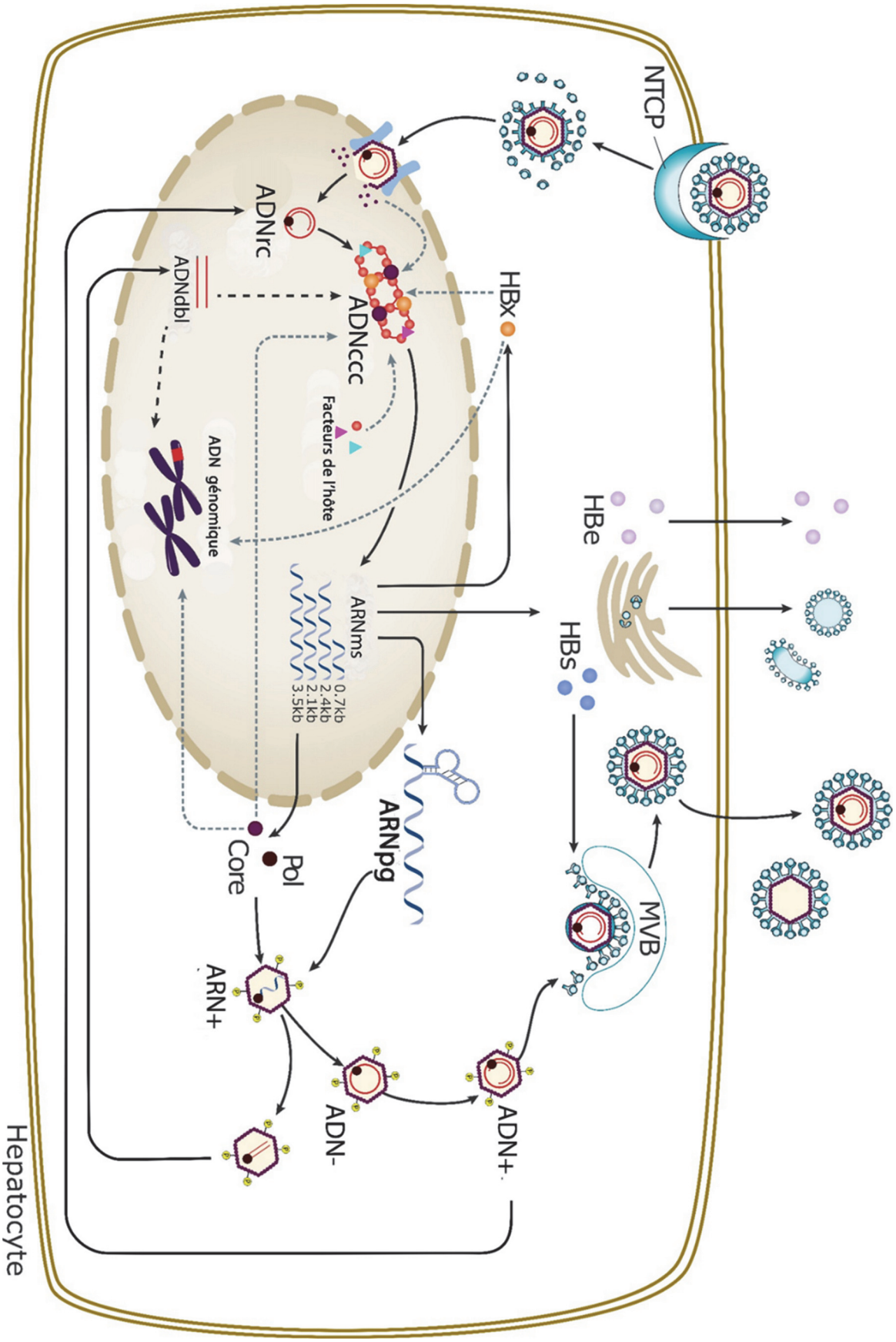


Figure 14 Cycle de réplication du virus dans les hépatocytes (adatpé de (Revill et al., 2016))

III. Modèles d'études de l'infection par le virus de l'hépatite B

HBV infecte uniquement les hépatocytes humains, et pendant longtemps l'étude de l'interaction entre le virus et la cellule hôte a été limitée par les modèles disponibles. La mise au point de nouveau modèles d'étude a permis depuis quelques années de mieux caractériser le cycle viral.

1. Les lignées cellulaires d'hépatomes

Les lignées d'hépatomes ne sont pas, en raison du faible niveau d'expression du récepteur NTCP, susceptibles à l'infection par HBV. Cependant la transfection du génome d'HBV permet l'étude de la réplication et de la sécrétion du virus. La transfection transitoire des cellules HepG2 ou HuH7 avec 1,1 ou 1,3 fois le génome d'HBV permet d'étudier la réplication du virus et sa morphogenèse. Des lignées stables clonales et produisant des particules virales infectieuses ont été aussi obtenus par transfection de > 1,0 génome HBV.

Les cellules HepG2.2.15 contiennent un dimère du génome du HBV intégré au génome de l'hôte et elles produisent de manière continue des particules virales (Sells et al., 1987; Sureau et al., 1986). La production de particules virales dans les cellules HepAD38 est quant à elle inductible par la tétracycline. Ces lignées permettent la production de particules d'HBV en quantité pour des expériences *in vitro* ou *in vivo* (Acs et al., 1987; Ladner et al., 1997).

Les cellules HepG2 H1/3Dx possèdent une intégration d'1,3 fois le génome d'HBV, avec une mutation dans la séquence de la protéine HBx, ce qui induit un codon stop. Il n'y a donc pas de production de la protéine HBx (Lucifora et al., 2011). Ceci permet l'étude de la réplication en absence de la protéine HBx.

Suite à la découverte du récepteur NTCP, plusieurs lignées HepG2 et HuH7 sur-exprimant NTCP et susceptibles à l'infection par HBV ont été créées, qui permettent d'obtenir des taux très élevés d'infection (Figure 15).

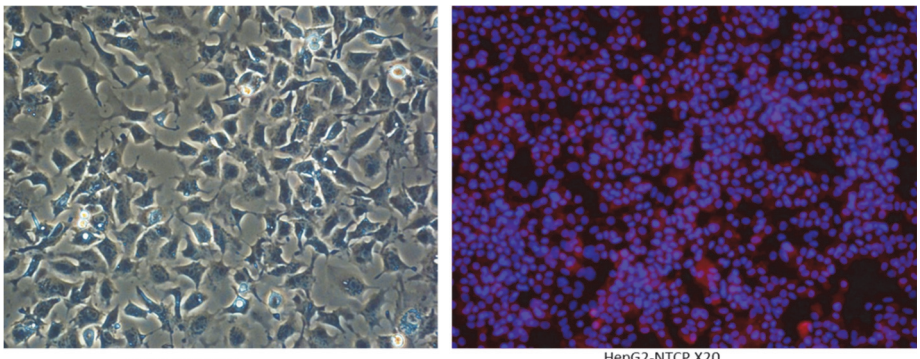


Figure 15 Photographie de cellules HepG2-NTCP en lumière blanche et après 6 jours d'infection (en rouge HBcAg)

2. La lignée cellulaire HepaRG

Les cellules HepaRG constituent une lignée de progéniteurs bipotents. Cette lignée est immortalisée mais non transformée. En culture il faut utiliser un protocole de différenciation de ces cellules afin d'obtenir des hépatocytes et des cholangiocytes (Gripon et al., 2002) Figure 16). Une fois différenciées, les cellules HepaRG deviennent susceptibles à l'infection par HBV (Hantz et al., 2009). Le taux de cellules infectées est faible (5 à 10%), ce qui ne permet pas d'étudier les effets de l'infection virale sur la régulation de l'expression des gènes de l'hôte. Cependant, il s'agit d'un bon modèle pour étudier l'immunité innée car les cellules HepRG expriment un grand nombre de récepteurs de l'immunité innée (Luangsay et al., 2015).

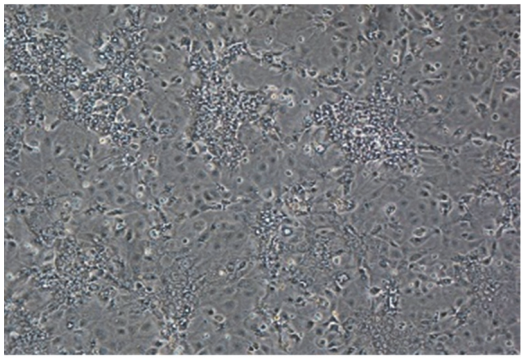


Figure 16 Photographie de cellules HepaRG différenciées

3. Les hépatocytes primaires humains (PHH)

Les hépatocytes primaires humains (PHH) restent à l'heure actuelle le meilleur modèle d'étude pour l'infection HBV (Figure 17). Ces cellules sont isolées directement à partir de résection de foie de patients atteint de cancer gastrique ou colorectal présentant des métastases au niveau du foie (LeCluyse and Alexandre, 2010). Les PHH sont susceptibles à l'infection par le virus (Gripon et al., 1988). Cependant après leur ensemencement, ces cellules peuvent se dédifférencier au cours du temps (Gripon et al., 1993; Schulze-Bergkamen et al., 2003). L'accès aux résections de foie et la qualité des PHH cryoconservés limitent l'utilisation de ce modèle.

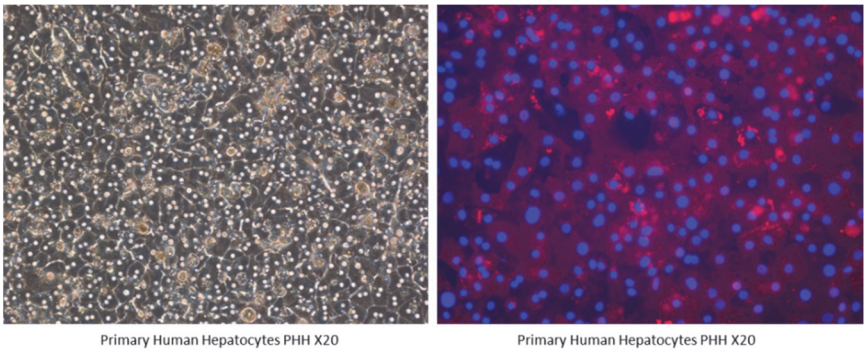


Figure 17 Photographie de PHH en lumière blanche te 12 jours après infection (en rouge HBcAg)

4. Les modèles animaux

A cause du tropisme restreint d'HBV, le chimpanzé est l'unique modèle naturel d'infection pour HBV (Maynard et al., 1972). C'est l'animal qui mime le mieux et le plus précisément l'infection naturelle par HBV chez l'homme. Mais depuis 2013 pour des raisons éthiques le National Institute of Health a progressivement limité et finalement interdit les expériences sur ces animaux.

D'autres modèles d'animaux infectés par des virus proches du HBV humain permettent d'étudier certains aspects de l'infection.

La marmotte américaine infectée par le « Woodchuck B virus » (WHB) a permis de découvrir les facteurs impliqués dans l'établissement et la persistance de l'infection et ainsi que l'hépato carcinogénèse (Dandri et al., 1996; Zoulim et al., 1994). En effet toutes les marmottes infectées développent un carcinome hépatocellulaire (HCC) mais les mécanismes moléculaires sous-jacents sont différents de ceux observés chez l'Homme (Tennant et al., 2004). Les analogues de nucléosides actuellement utilisés comme traitement antiviral ont été testés en premier dans ce modèle. De nombreuses études sur les traitements immunothérapeutiques ont aussi été effectués chez la marmotte américaine. Ce modèle a servi également à l'étude de l'infection occulte HBV qui correspond à la présence d'ADN viral dans le sérum et/ou dans le foie lorsque l'HBsAg est indétectable (Dandri et al., 2016).

Le canard de Pékin est infecté par le « Duck Hepatitis B Virus » (DHBV) (Mason et al., 1982). Le modèle DHBV a été largement utilisé pour des études *in vivo* et *in vitro* visant à comprendre la réplication des *hepadnavirus* et les mécanismes de biogenèse de l'ADNccc. Il est encore utilisé lors de tests de traitements antiviraux. Rappelons que le DHBV n'a que 40% d'homologie avec le HBV humain et que l'on ne peut étudier le développement de la maladie chronique car ce modèle ne présente pas d'atteintes du foie ou de développement de HCC (Cova et al., 1993).

Par ailleurs des modèles de souris ont été développés pour disposer de petits animaux pour l'étude de HBV.

Des souris transgéniques exprimant une partie ou le génome entier de HBV ont été générées. Ces souris transgéniques ont été utilisées afin de caractériser le rôle des protéines du virus dans l'oncogenèse (Chisari et al., 1985; Kim et al., 1991). Les souris transgéniques contenant >1,0 génome du HBV produisent des virions qui sont semblables à ceux produits chez l'Homme et ont permis d'étudier le rôle de la réponse immunitaire dans la pathogenèse virale (Guidotti et al., 1995). Etant donné que les hépatocytes de souris ne sont pas susceptibles à l'HBV, l'entrée du virus et la formation de l'ADNccc ne peuvent pas être étudiées dans ce modèle.

La transduction du génome du HBV dans le foie de souris à l'aide de vecteur « adeno associated virus » (AAV) amène à la production de particules virales et à la réplication sous le contrôle d'un promoteur HBV endogène. De plus l'ADNccc a pu être détecté dans le foie de ces souris ce qui permettra l'étude de la formation de l'ADNccc *in vivo* (Lucifora et al., 2017). Ce modèle possède toutefois des limitations car le système immunitaire compétant élimine les vecteurs au cours du temps (Dandri et al., 2016).

Enfin, des modèles de souris à foie humanisé ont été développés pour étudier l'infection HBV. Afin que la greffe d'hépatocytes soit efficace et permette la reconstruction du foie humanisé il faut : 1) qu'il y ait une destruction des hépatocytes de la souris afin de créer de la place pour le repeuplement du foie à partir des hépatocytes humains et 2) il faut inhiber les réponse immunitaires de la souris pour ne pas qu'elles détruisent la xénogreffe (Allweiss and Dandri, 2016).

Les souris uPA/SCID expriment « l'urokinase plasmogene activator » (uPA) responsable de la destruction des hépatocytes de souris et possèdent une « immunodéficience combinée sévère » (SCID) permettant la xenogreffe. Les souris FRG quant à elles sont déficientes pour l'enzyme hépatique « fumarylacetoacetate hydrolase » (FAH) - ce qui conduit à la mort des hépatocytes murins - ainsi que pour la protéine RAG2 et pour la chaine gamma du récepteur à l'IL-2 ,ce qui conduit à l'immunodéficience (Dandri and Lütgehetmann, 2014). Les deux permettent l'étude de l'infection mais pas des interactions avec le système immunitaire.

Un modèle de souris BALB/c *Rag2-/- IL2rg*$^{-/-}$NOD.sirpa uPA doublement humanisé comportant un foie et un système immunitaire humain suite à la co-transplantation de cellules souches hématopoïétiques (HSC) et d'hépatocytes humains a été généré (Strick-Marchand et al., 2015). Ce modèle permet d'étudier les réponses du système immunitaire dirigées contre le virus et de tester des stratégies immuno-thérapeutiques (Figure 18) (Dusséaux et al., 2017; Strick-Marchand et al., 2015).

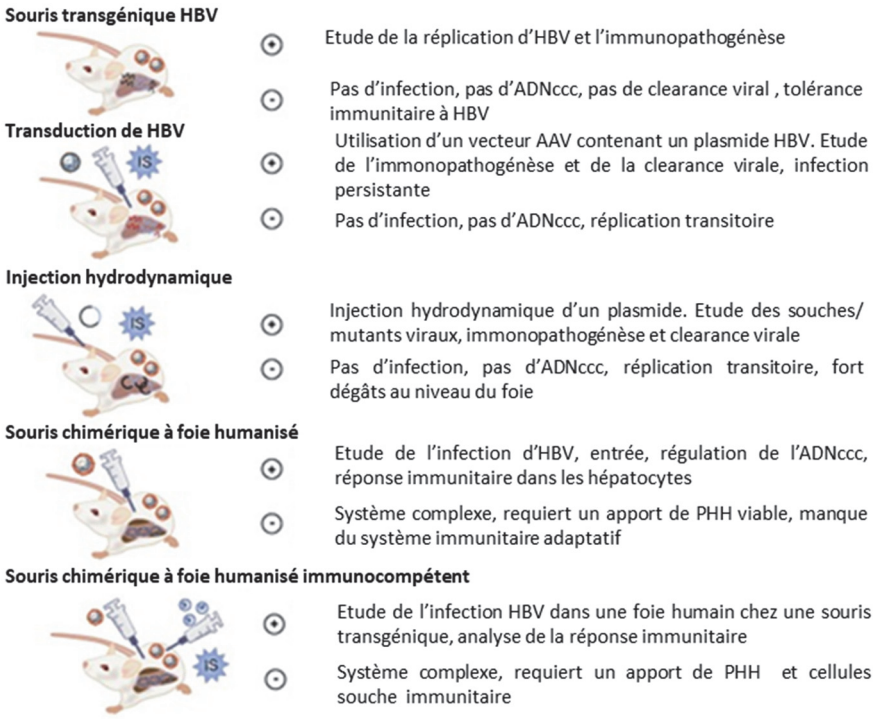


Figure 18 Différences entre les différents modèles de souris existant (adapté de(Allweiss and Dandri, 2016))

Chapitre 2 L'infection par le virus de l'hépatite B

I. De la découverte du virus de l'hépatite B au vaccin

1. De la découverte au Prix Nobel

Tout a commencé grâce aux travaux qu'effectuait Baruch Blumberg sur la variabilité de certaines protéines du sang dans des populations très diverses dans le monde. Son étude était basée sur du sang de patients polytransfusés, car celui-ci comporte une forte quantité d'anticorps dirigés contre les protéines étrangères au système immunitaire du patient (Allison and Blumberg, 1961). La détection d'une bande d'immunoprécipitation lorsque le sérum d'un patient atteint de leucémie et polytransfusé a été testée avec le sérum d'un Aborigène australien a conduit à la découverte de l'antigène « Australia ». Cet antigène était absent chez les individus sains, mais très présent chez des patients atteints de leucémie aiguë et associé à une atteinte hépatique (Blumberg et al., 1967). Cet antigène est utilisé pour le dépistage de l'hépatite B dans le sang lors de transfusions sanguines. Ensuite une étude au microscope électronique permis de visualiser le virus sous la forme de particules nommées par la suite « particules de Dane » du nom du scientifique qui fit cette découverte (Figure 19) (Dane et al., 1970). Blumberg reçu en 1976 le Prix Nobel de Médecine pour ses découvertes sur HBV.

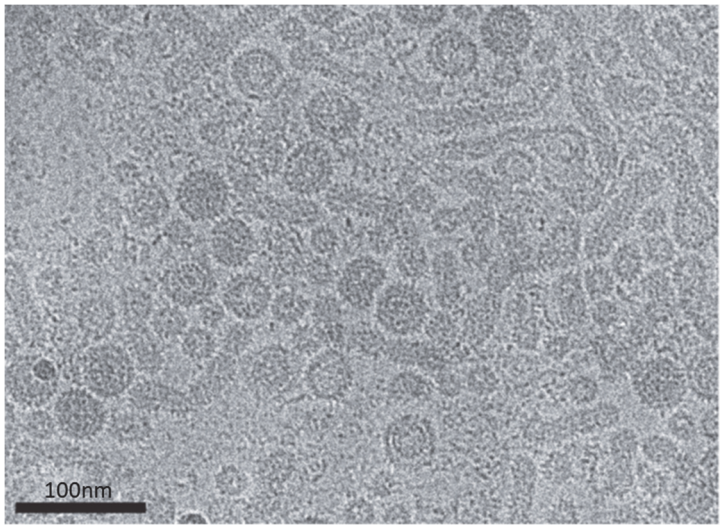


Figure 19 Cryoelectron micrograph (cryo-EM) de virions isolé à partir de patients infecté chroniquement, on distingue les particules de Dane et les particules subvirales (Selzer and Zlotnick, 2015)

2. Classification/Génotype

Tous les virus qui s'apparentent au HBV sont classés dans la famille des *Hepadnaviridae* (Figure 20). Cette famille regroupe les virus à ADN de petite taille (3-3,3kb). Il s'agit de virus

hépatotropes, qui possèdent la même structure et un cycle de réplication comportant une étape de rétrotranscription.

Cette famille comporte deux genres que sont les *orthohepadnavirus* et les *avihepadnavirus* (Fauquet and Mayo, 2005).

Les *avihepadnvirus* regroupent les *hepadnavirus* aviaires. Ce genre est principalement composé du « Heron Hepatitis B Virus » (HHBV) et du « Duck Hepatitis B Virus » (DHBV). Les *orthohepadnavirus* quant à eux infectent les mammifères. On retrouve dans ce genre HBV, mais aussi les virus apparentés infectant les primates supérieurs (chimpanzé) ou encore des rongeurs comme la marmotte d'Amérique infectée par le « Woodchuck Hepatitis Virus » (WHV) (Seeger et al., 2013).

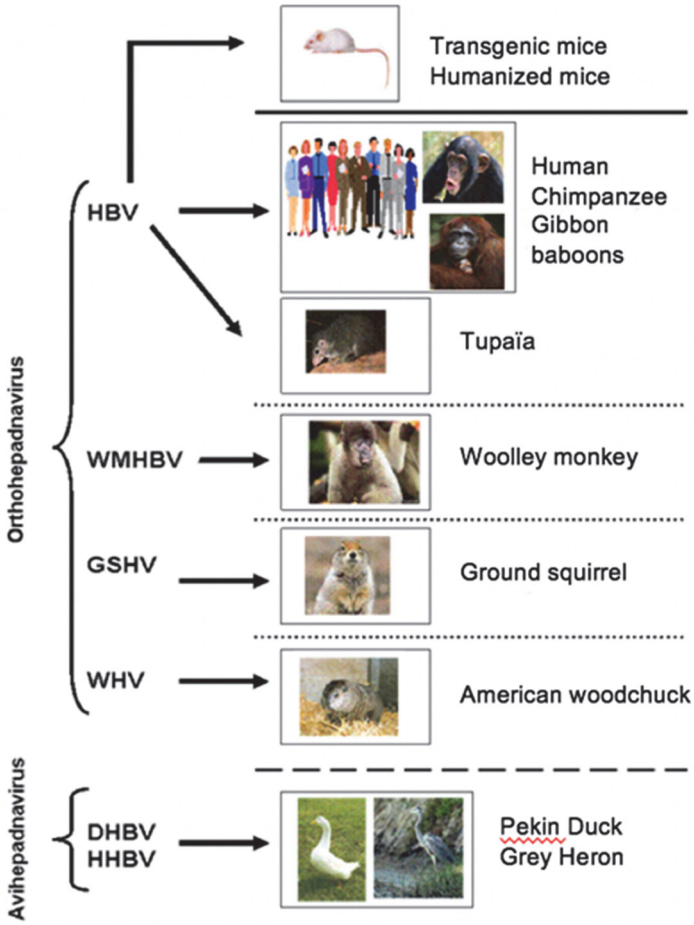


Figure 20 Arbre phylogénique des différents genre d'Hepadnavirus (adapté de (Seeger et al., 2013))

En ce qui concerne HBV, plusieurs isolats ont été retrouvés dans le monde. Ils sont classés en 8 génotypes principaux (A à H) et deux génotypes mineurs (I et J). Leur appartenance à un génotype est définie selon la divergence génétique entre les isolats. Les génotypes peuvent avoir des sous génotypes. Les génotypes A à D sont les mieux caractérisés. La répartition géographique des différents génotypes est représentée dans la Figure 21 (Kramvis, 2014). Les patients infectés par les génotypes A, B et F ont plus fréquemment une séroconversion spontanée de l'HBeAg tandis que le génotype D développe très fréquemment le mutant

HBeAg (-). L'infection avec une souche virale de génotype E est associée à une charge virale élevée et une plus haute fréquence d'HBeAg positifs. Les génotypes A1, C, B2-B4, F1 et D sont associés à un risque plus élevé de développer une cirrhose et l'HCC (Kramvis, 2014). Le génotype G est fréquemment trouvé dans les co-infections du HBV avec le virus de l'immunodéficience humaine (HIV) avec une incidence accrue de fibrose avancée (Tong and Revill, 2016). Le génotype I est issu d'un isolat vietnamien mais possède peu de divergence génétique avec le génotype C. Le génotype J a été isolé à partir d'un patient Japonais et certaines régions sont très proches d'un virus infectant l'orang-outan. Ces deux génotypes sont considérés « putatifs » du fait de leur manque de diversité génétique (Kato et al., 2016).

Figure 21 Cartographie de la répartition des différents génotypes d'HBV dans le monde (adapté de (Petit and Trepo, 2014))

3. Transmission

Le HBV est transmis par voie périnatale (de la mère à l'enfant) ou de manière horizontale via une exposition à du sang infecté.

Le risque de transmission par la mère est de 90% en Asie, car les mères possèdent souvent un fort titre d'HBV dans le sang. (Seeger et al., 2013). D'autres voies peuvent être aussi source de contamination, par exposition percutanée, à travers les muqueuses, lors des écoulements menstruels ou via des sécrétions vaginales et séminales. Le risque de transmission lors de transfusions sanguines a beaucoup diminué depuis la mise en place de tests de dépistage sérologique dans les années 1970 puis de tests d'amplification des acides nucléiques (« nucleic acid testing » (NAT)). Le virus peut aussi se transmettre lors d'injection de drogues, de tatouages, d'acuponcture si le matériel est réutilisé et contaminé. Une transmission

nosocomiale peut aussi se produire lors d'hémodialyse ou de transplantation (Seeger et al., 2013; OMS).

4. Les vaccins

L'induction de la mémoire immunitaire dirigée contre les antigènes HBs et la génération d'anticorps anti-HBs est cruciale pour la protection contre HBV. Les antigènes pre-S semblent importants pour l'induction des lymphocytes T-helper afin de produire des anticorps dirigés contre le virus (Shouval, 2003).

La démonstration dans les années 70 que les patients qui ont guéri d'une infection aiguë par HBV possèdent dans leur sérum des anticorps reconnaissant l'HBsAg du virus - ceci induisant une immunité protectrice contre HBV - a ouvert la voie au vaccin contre le HBV. Deux vaccins de première génération contenant des antigènes HBs purifiés à partir du sérum de patients ont été développés en France et aux États-Unis (Maupas et al., 1981; Szmuness et al., 1980).

La deuxième génération a été produite par les techniques de recombinaison de l'ADN chez la levure. Les deux principaux vaccins sont l'Engerix-B (SmithKline Biologicals, Belgique) et le RECOMBIVAX HB-Vax II (Merck&Co, USA). Ces deux vaccins contiennent de S-HBsAg non glycosylé qui est sécrété par les levures lors de la production et contiennent l'alun comme adjuvant (Stephenne, 1990). Il y a actuellement 140 pays où la vaccination contre HBV est présente surtout chez les enfants et les personnes à risque.

Cependant il y a des personnes chez qui la vaccination n'est pas assez efficace (« non répondeurs ») avec des taux d'anticorps anti-HBs inférieurs à 10 UI/ml et qui restent à risque d'infection.

II. Histoire naturelle de l'infection par le virus de l'hépatite B

1. L'infection aiguë

L'infection aiguë est asymptomatique dans 60 à 80% des cas. L'HBsAg va apparaitre dans le sang au cours des 10 semaines suivant l'infection suivi par l'apparition des anticorps anti-HBc et la positivité de l'ADN viral (HBV DNA) (Figure 22). Pendant cette période, il y a une augmentation du taux des « transaminases alanine aminotransférase » (ALAT) et des « aminotransférases aspartiques » (ASAT) (Dény and Zoulim, 2010). Dans les cas symptomatiques, il y a apparition d'une jaunisse avec une forte fatigue associée, le foie à la palpation peut être sensible et plus gros. La résolution de l'infection est accompagnée par la

diminution de la charge virale, une séroconversion en anti-HBe et la normalisation des ALAT (Seeger et al., 2013). Dans 0,1% des cas une forme fulminante d'hépatite peut survenir.

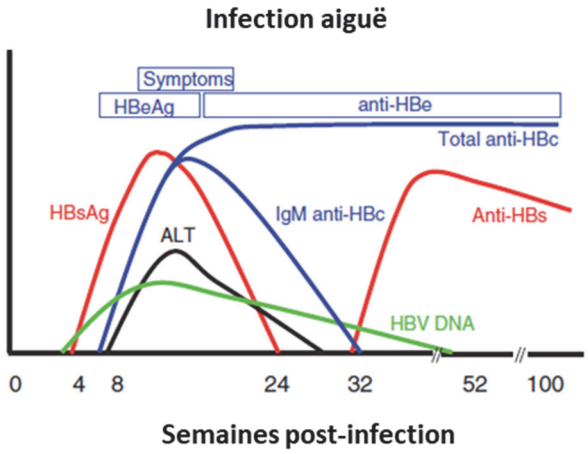


Figure 22 Profils sérologique lors de l'infection aiguë par HBV (adaptée de (Seeger et al., 2013))

Les différents biomarqueurs sérologiques du HBV au cours des infections aigues, infections résolues et chez les sujets vaccinés sont résumés dans le tableau ci-dessous (Tableau 1).

2. L'infection chronique

L'infection chronique est définie par la persistance de l'HBsAg dans le sérum au-delà de 6 mois après la contamination (Figure 23). Les patients chroniquement infectés sont positifs pour les anticorps dirigés contre la protéine core (HBc) mais négatifs pour les immunoglobulines anti-HBc de classe M (IgM). On estime que plus de 250 millions de personnes à l'échelle mondiale sont infectées chroniquement par HBV. 90% des nouveaux-nés infectés pendant la période périnatale vont développer une infection chronique, contre 30% des enfants infectés entre 1 an et 5 ans. Le risque de chronicité chez l'adulte immunocompétent est estimé à 5%, mais varie selon la population et les modalités de transmission de l'infection.

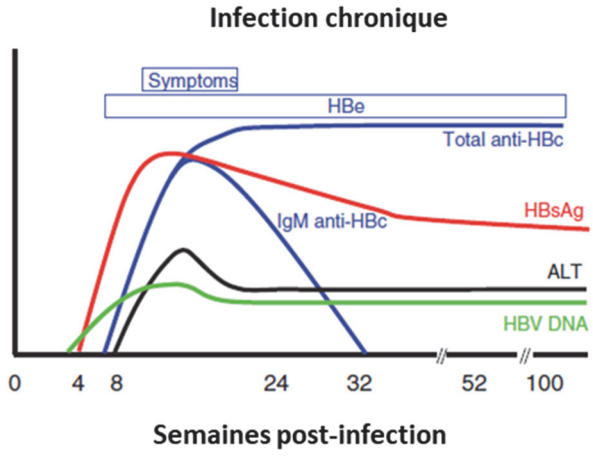


Figure 23 Profils sérologiques lors de l'infection chronique par HBV (adapté de (Seeger et al., 2013))

L'infection chronique est caractérisée par différentes phases (Tableau 1) (Lampertico et al., 2017):

- Phase 1 : infection HBV chronique HBeAg positive (« immuno-tolérance ») ;
- Phase 2 : hépatite B chronique HBeAg positive (« phase immunoréactive ») ;
- Phase 3 : infection HBV chronique HBeAg négative (« porteurs inactifs ») ;
- Phase 4: hépatite B chronique HBeAg négative;
- Phase 5: phase HBsAg négative (« infection HBV occulte »)

	Infection aigue	HBeAg positif		HBeAg négatif		Infection résolue	Personnes vaccinée
		Infection chronique	Hépatite chronique	Infection chronique	Hépatite chronique		
HBsAg	Positifs	Haut	Haut Intermédiaire	Faible	Intermédiaire	Négatifs anti-HBs > 10 UI/ml	Négatifs anti-HBs > 10 UI/ml
HBeAg	Positifs	Positif	Positif	Négatif	Négatif	Anti-HBc positifs	
ADN HBV	Positif	$>10^7$ UI/ml	10^4-10^7 UI/ml	<2000 UI/ml	>2000 UI/ml	Négatifs	
ALT	Normau/ élevé	Normal	Elevé	Normal	elevé	Normaux	
Maladie du foie		Nul/minimal	Modéré sévère	Nul	Modéré sévère		
Ancienne terminologie		Immune tolérant	Immune réactive HBeAg positif	Patient inactif	HBeAg négatifs hépatite chronique		

Tableau 1 Les différentes phases d'infection d'HBV (adapté de (Zoulim et al., 2018))

III. Les cibles thérapeutiques

1. Le traitement de l'infection chronique

Le but principal de la thérapie antivirale est de supprimer la réplication du virus et l'inflammation hépatique, ainsi que de limiter la progression de la maladie hépatique. Bien que le traitement antiviral diminue le risque de HCC, il ne l'élimine pas entièrement. Dans la grande majorité des cas le traitement antiviral ne permet pas l'élimination du virus, qui persiste dans les hépatocytes sous sa forme de minichromosome (ADNccc).

L'indication au traitement, à la fois dans les cas HBeAg positifs et HBeAg négatifs, est basée principalement sur trois critères : le taux d'ADN viral, le taux d'ALT dans le sérum et la sévérité des dommages du foie (Zoulim and Durantel, 2015).

Indépendamment du statut HBeAg , il y a deux stratégies de traitement : un traitement de courte durée avec de l'interféron α pégelé (PEG-IFN-α) ou un traitement de longue durée avec

un analogue de nucléos(t)ides (NUC) (Lampertico et al., 2017; Trépo et al., 2014; Zoulim and Durantel, 2015) (Figure 24).

L'avantage de l'IFN-α est qu'il n'induit pas de résistance et qu'il possède une activité antivirale et immunomodulatrice. Cependant il nécessite une administration parentérale par injection sous-cutanée et peut avoir des effets indésirables comme des nausées, dépressions et il est contrindiqué chez des patients avec une cirrhose décompensée (Trépo et al., 2014). Le but du traitement au PEG-IFN-α est de réduire la charge de l'ADN viral dans le sang en dessous de 2000 UI/ml (Lampertico et al., 2017).

Les NUCs inhibent l'activité polymérase ainsi que la reverse transcription de l'ARN pré-génomique en ADN. En plus de diminuer l'inflammation du foie, le traitement antiviral de longue durée peut ralentir la progression de la cirrhose. Différents NUCs sont disponibles. L'Entécavir et le Ténofovir induisent très peu de résistance et sont désormais les traitements de choix. La Lamivudine est peu coûteuse mais induit très fréquemment une résistance dans le cas de monothérapie longue. L'Adéfovir est associé à un taux de résistance intermédiaire et présente une toxicité rénale plus élevée (Zoulim and Durantel, 2015). Le but du traitement avec les NUCs est d'aboutir à une charge virale négative. Une mauvaise réponse au traitement est définie par une diminution de moins d'un log10 du taux d'ADN viral dans le sérum après trois mois de traitement (Lampertico et al., 2017).

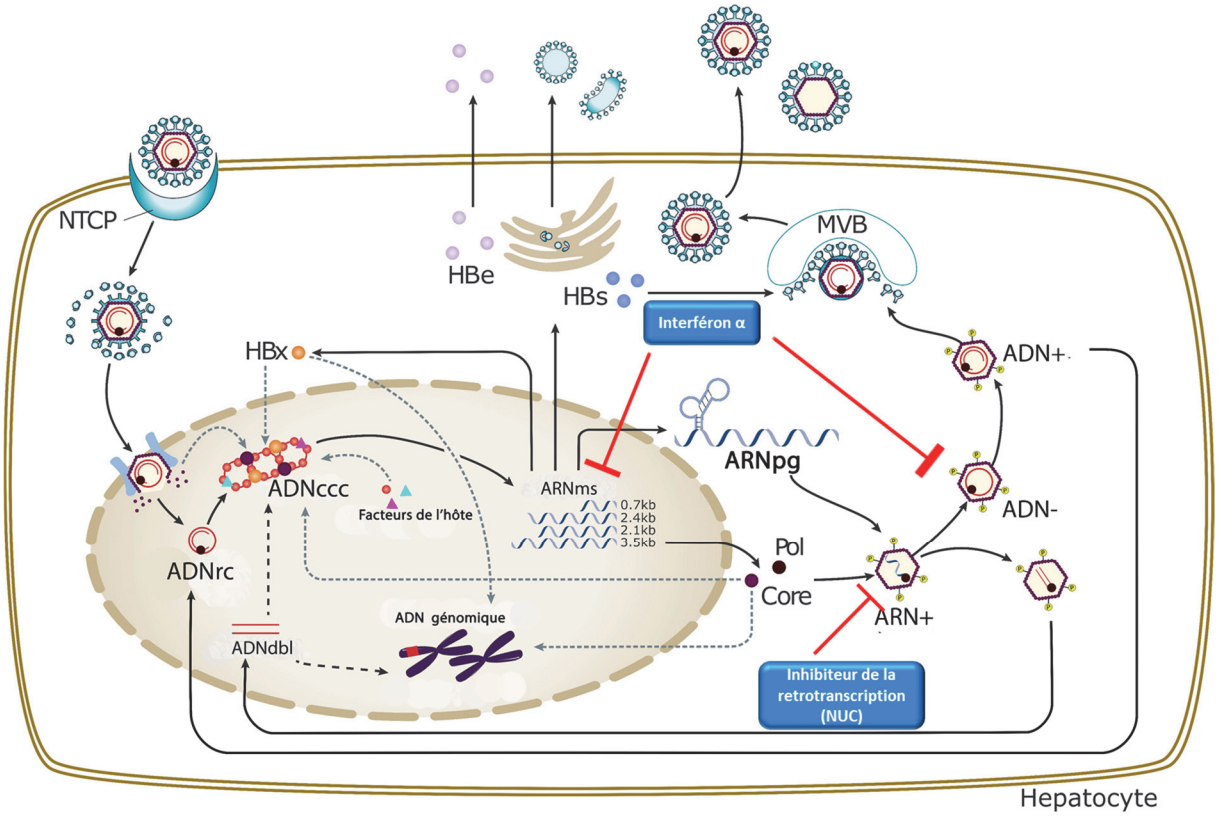


Figure 24 Traitements utilisés lors de l'infection chronique (adapté de (Lok et al., 2017))

2. Les nouvelles cibles thérapeutiques

Les grands progrès dans la compréhension des interactions HBV-hôte et des mécanismes moléculaires sous-tendant le cycle viral ont permis de développer de nouvelles stratégies thérapeutiques telles que les inhibiteurs d'entrée, les inhibiteurs de l'assemblage des capsides, les ARN interférant, des stratégies basées sur l'édition de gènes et des stratégies d'immunomodulation (Schinazi et al., 2018)). Ce domaine est en continuelle évolution et un aperçu actualisé des différents antiviraux en cours de développement peut être consulté sur le site http://www.hepb.org/treatment-and-management/drug-watch/ (Figure 25).

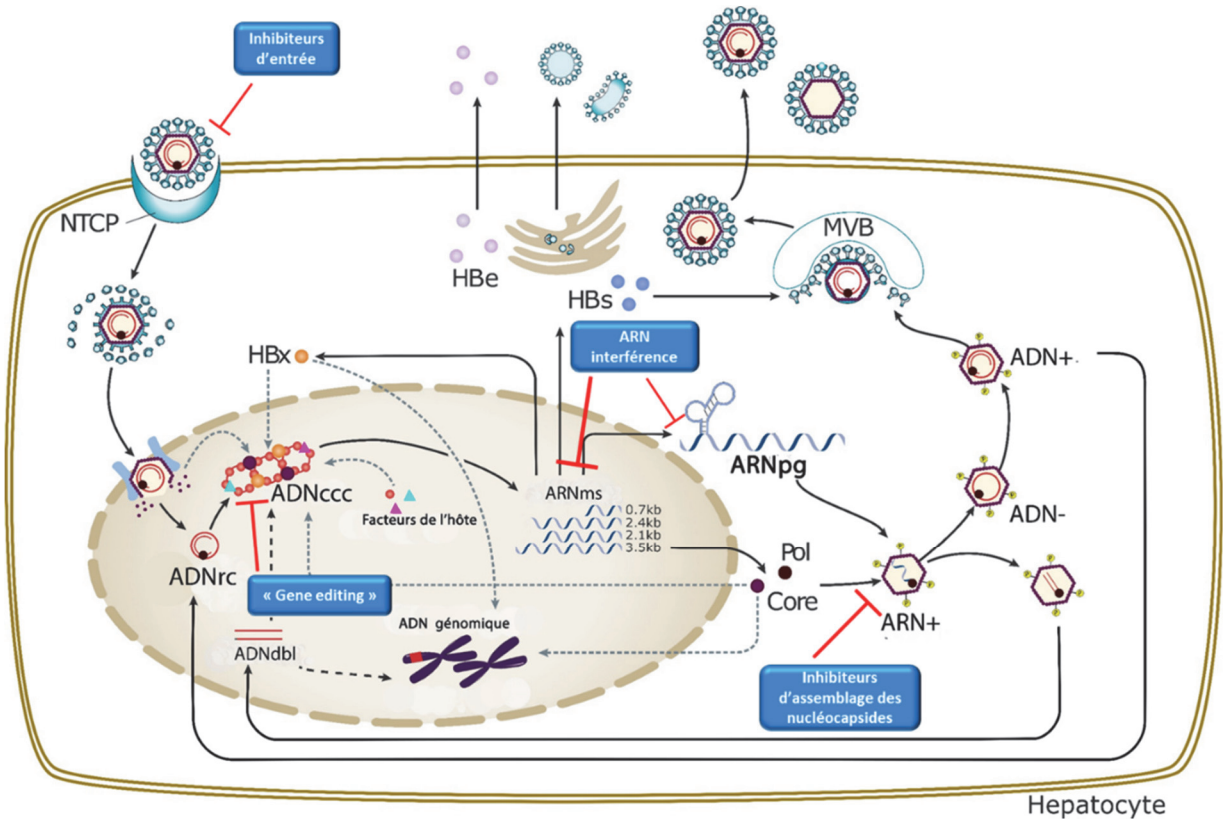


Figure 25 Schéma représentant les nouvelles cibles thérapeutiques (adapté de (Lok et al., 2017)

Chapitre 3 Virus de l'hépatite B, épigénétique et carcinome hépatocellulaire

I. Modifications épigénétiques

L'épigénétique est définie comme l'étude des changements dans l'activité des gènes, n'impliquant pas de modification de l'ADN et pouvant être transmis lors de la division cellulaire. On distingue différentes modifications épigénétiques qui sont les méthylations de l'ADN et les modifications des histones. Au vu de leur rôle important dans l'expression des gènes les ARN non-codant (ncRNA) peuvent également être considérés comme des acteurs épigénétiques (Figure 26).

L'objet de l'épigénétique concerne l'indexation du génome en fonction de la probabilité d'utilisation du matériel génétique pour la transcription et l'expression des gènes. La nécessité d'un tel système est apparue avec l'évolution des eucaryotes, qui doivent contenir une quantité accrue de matériel génétique dans l'espace confiné du noyau (Figure 26). La condensation et l'indexation du génome sont des fonctions de la chromatine, le complexe nucléoprotéique formé par l'ADN, les histones, qui sont les protéines architecturales de la chromatine, et de protéines non-histones.

La chromatine joue un rôle essentiel pas seulement dans la transcription mais aussi dans nombreux autres processus physiologiques tels que la réplication et la réparation de l'ADN.

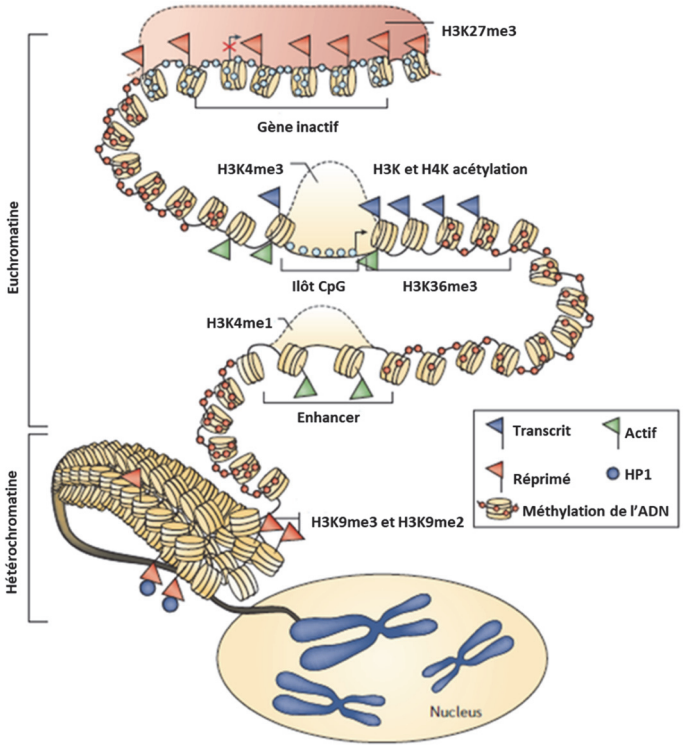


Figure 26 Représentation des modifications épigénétiques dans des cellules humaines normales (adapté de((Baylin and Jones, 2011))

1. Méthylation de l'ADN

La méthylation de l'ADN est la seule modification épigénétique qui affecte directement la molécule d'ADN et joue un rôle essentiel dans divers processus biologiques tel que le développement, l'inactivation du chromosome X et l'empreinte génomique parentale.

Les méthyltransférases de l'ADN (DNMTs) sont capables de modifier l'ADN en transférant un groupe méthyl (-CH3) du donneur de méthyl S-adénosyl-méthionine (SAM) au carbone 5 des résidus cytosines de l'ADN. Cette modification chimique de l'ADN n'influence pas la séquence d'acide nucléique. Cette réaction survient essentiellement sur les résidus cytosine de di-nucléotides palindromiques 5'-CG-3' appelés dinucléotides CpG (Jones, 2012).

a. Les différents types de méthylation de l'ADN

Il existe principalement deux types de méthylation, la méthylation *de novo* et la méthylation de maintenance.

- La méthylation *de novo* est catalysée par DNMT3A et DNMT3B et consiste en l'addition de groupement méthyl sur les cytosines des dinucléotides CpG localisés sur les deux brins d'ADN. DNMT3L, qui ne possède pas d'activité enzymatique, joue le rôle de cofacteur. La méthylation *de novo* joue aussi un rôle important lors de la gamétogenèse et de l'embryogenèse (Robertson and Jones, 2000). Dans les cellules germinales, un mécanisme particulier impliquant des petits ARN nommé piwi-interacting RNAs (piRNAs) a été décrit. Les piRNA semblent aussi recruter les DNMT3A et DNMT3B pour induire la méthylation *de novo* au niveau des transposons (Aravin and Bourc'his, 2008).

- La méthylation de maintenance implique l'ajout d'un groupement méthyl sur la cytosine d'un dinucléotide CpG localisé sur le brin néo-synthétisé de l'ADN en regard d'un dinucléotide CpG méthylé sur le brin parental (Jones, 2012). L'enzyme majeure lors de la méthylation de maintenance est DNMT1 qui possède une affinité pour l'ADN hémiméthylé et est présente au niveau de fourches de réplication de l'ADN (Liang and Weisenberger, 2017).

b. Les sites de méthylation de l'ADN

Les dinucléotides CpG sont enrichis dans les « îlots CpG » (CGI), des régions de petite taille (0,2-2 kilo bases) qui ont une teneur en GC supérieure à 50% et un rapport de CpG sur GpC d'au moins 0,6 et que l'on trouve dans 60 % des promoteurs chez l'Homme. Ils sont principalement présents au niveau des sites d'initiation de la transcription (« transcription strat site » (TSS)) mais ils peuvent aussi se trouver au niveau des régions intergéniques ou intragéniques.

La méthylation de l'ADN est une marque répressive et la méthylation d'un îlot CpG est généralement incompatible avec l'activité transcriptionnelle (Liang and Weisenberger, 2017).

Alors que la méthylation de l'ADN en proximité de TSS bloque l'initiation de la transcription tandis que les méthylations d'ADN dans les régions intragéniques peuvent stimuler l'élongation de la transcription avec un possible impact sur l'épissage des ARN. Les exons sont d'habitude plus méthylés que les introns. La forte méthylation au niveau des liaisons exon-introns suggère encore un rôle dans l'épissage. Il peut aussi y avoir des méthylations au niveau des enhancers mais leur conséquences sont moins connues (Jones, 2012).

La méthylation de l'ADN crée des sites de liaison spécifiquement reconnus par les « methyl-binding proteins » (MBPs). Ces protéines recrutent des complexes de remodelage de la chromatine et des corépresseurs transcriptionnels qui concourent à l'établissement de l'hétérochromatine (Fuks, 2005).

La liaison des protéines insulatrices, comme le CCCTC-binding factor (CTCF), aux CGI organisent l'ADN en boucles d'hétérochromatine transcriptionnellement répressives qui connectent des enhancers distaux à des promoteurs proximaux (Baylin and Jones, 2011).

La position des nucléosomes (voir plus loin le chapitre 3.I.2.) est dynamique et change pendant la réplication cellulaire de même que dans l'expression des gènes. Les promoteurs de gènes actifs, particulièrement ceux avec des régions riches en CpG et méthylation ADN faible ou absente, possèdent des régions déplétées en nucléosomes (NDRs) en amont de leur TSS. Les nucléosomes proches de ces NDR sont marqués par la modification histonique H3K4me3, ont une forte acétylation des lysines et abritent le variant H2A.Z qui déstabilise les nucléosomes pour faciliter la transcription (Baylin and Jones, 2011).

c. La déméthylation de l'ADN

Il y a deux mécanismes de déméthylation de l'ADN : la voie active et la voie passive.

La déméthylation passive de l'ADN survient lors d'un manque de maintien de la méthylation de l'ADN qui peut être due à la perte d'expression ou d'activité de DNMT1. Ceci va conduire lors de division cellulaire à l'obtention d'un ADN hypométhylé et donc à une possible activation des gènes de la région en question (Jones, 2012).

La déméthylation active de l'ADN est un processus comprenant plusieurs étapes. Il peut faire intervenir les enzymes de la famille des « Ten Eleven Translocase » (TET) qui comprend TET1, TET2 et TET3 (Liang and Weisenberger, 2017). Leur fonction est de convertir la 5-methylcytosine (5mC) en 5-hydroxymethylcytosine (5hmC) en utilisant comme cofacteur l'acide ascorbique (vitamine C). Par la suite les 5hmC peuvent être déaminées par les protéines AID/APOBEC ce qui induit l'apparition de 5-hydroxyméthyluracile (5hmU) et donc un mésappariement U/G. Les 5mC peuvent aussi être prises en charge par AID/APOBEC afin d'être désaminées ce qui conduit à la mise en place de mésappariement T/G. Les deux types de mésappariement sont pris en charge par le système de « Base Excision Repair » BER (Jones et al., 2012).

d. Le rôle de la méthylation de l'ADN dans les cancers

Une méthylation appropriée de l'ADN est essentielle au développement et au bon fonctionnement des cellules. Par conséquent, toute anomalie dans ce processus peut entraîner diverses maladies et notamment le cancer. Les cellules tumorales sont caractérisées par un méthylome différent de celui des cellules normales et on peut observer à la fois des événements d'hypométhylation et d'hyperméthylation dans le cancer. En général, une diminution globale de la teneur en CpG méthylée est observée. Ce phénomène contribue à l'instabilité génomique et, moins fréquemment, à l'activation d'oncogènes réprimés dans les cellules normales. D'autre part, l'hyperméthylation de CGI dans les promoteurs de gènes spécifiques s'est révélée être une caractéristique essentielle de nombreuses cellules cancéreuses. Un nombre croissant de gènes, qui agissent principalement comme suppresseurs de tumeurs dans les tissus normaux, a été rapporté comme étant inactivés par un mécanisme de méthylation de l'ADN pendant la tumorigenèse.

Dans une cellule normale, l'hétérochromatine péricentromérique est fortement méthylée. Les séquences satellites, les séquences génomiques répétitives (telles que les éléments LINE, SINE, IAP et Alu) sont silencieuses, garantissant ainsi l'intégrité et la stabilité génomiques. Dans une grande variété de tumeurs, cependant, ce mécanisme est perturbé et il y a une perte de méthylation de l'ADN de régions normalement inactivées. En conséquence, les risques de recombinaison mitotique non souhaitée sont plus grands. Les éléments transposables sont ensuite réactivés et peuvent s'intégrer à des sites aléatoires du génome, ce qui entraîne une mutagenèse et une instabilité génomique. Cela peut arriver dans le développement de la tumeur ou être associé aux stades ultérieurs de la progression tumorale. Par exemple, la perte de méthylation de l'ADN de Sat2 (satellite juxtacentromérique 2) et Satα (satellite centromérique α) dans le cancer du sein est liée au développement précoce de la tumeur, tandis que dans le cancer de l'ovaire, elle contribue à la progression tumorale et a été proposée comme marqueur pronostic (Costa et al., 2006; Widschwendter et al., 2004a, 2004b).

L'hypométhylation de l'ADN de gènes individuels est plutôt rare et la majorité des promoteurs affectés par la perte de méthylation de l'ADN appartient à des gènes spécifiques d'un tissu. Par exemple, dans le mélanome et le cancer colorectal, les antigènes MAGE associés au mélanome sont réactivés par hypométhylation de l'ADN de la région promotrice (Kim et al., 2006; Weber et al., 1994). La perte de méthylation de l'ADN dans le promoteur du gène CDH3 entraîne une surexpression de la P-cadhérine dans les carcinomes colorectaux et du sein. En conséquence, une polarité cellulaire diminuée favorise l'invasion, la motilité et la migration des cellules, associées à une survie pire des patients (Milich and Liang, 2003; Paredes et al., 2005; Ribeiro et al., 2010). L'hypométhylation de l'ADN pourrait être aussi impliquée dans la résistance aux médicaments anticancéreux. Par exemple, le promoteur du gène de la multirésistance aux médicaments 1 (MDR1) est hypométhylé dans les cellules cancéreuses du sein (Baker et al., 2005; Chekhun et al., 2006; Sharma et al., 2010).

Aucune enzyme de déméthylation active n'a été découverte qui pourrait expliquer l'hypométylation de l'ADN dans le cancer. En outre, il n'existe aucune preuve claire d'une réduction de l'activité des DNMT dans le cancer mais plutôt une surexpression de ces enzymes

a été rapportée dans plusieurs cancers (Ahluwalia et al., 2001; Lin et al., 2007; Roll et al., 2008). Le mécanisme de l'hypométhylation dans le cancer n'est pas complètement compris. L'hypothèse la plus créditée relie l'hypométhylation à l'expression de variants inactifs de DNMT3B qui pourrait titrer les partenaires de liaison à DNMT3B ou faire concurrence à la forme active pour se lier aux séquences d'ADN cibles. L'implication des variants inactifs de la DNMT a été démontrée dans la leucémie et le carcinome hépatocellulaire HCC (Saito et al., 2002).

Malgré l'hypométhylation globale, certains gènes sont inactivés par l'hyperméthylation des CGI dans leurs régions régulatrices, qui ne sont pas méthylées dans les correspondants tissus non malins. Ce type de répression épigénétique a été mis en évidence pour la première fois dans les études sur des patients atteints de rétinoblastome, dans lesquels l'hyperméthylation du promoteur du gène suppresseur de tumeur du rétinoblastome (RB1) a été découverte (Greger et al., 1989). Depuis lors, un grand nombre de gènes suppresseurs de tumeurs inhibés par l'hyperméthylation de l'ADN dans les tissus tumoraux a été identifié et la répression associée à la méthylation de l'ADN joue un rôle crucial dans la tumorigenèse et constitue une caractéristique de tous les cancers chez l'Homme (Esteller, 2008) (Tableau 2).

Les gènes qui acquièrent une hyperméthylation dans les régions régulatrices sont impliqués dans diverses voies cellulaires importantes. Deux gènes clé liés au cycle cellulaire, les inhibiteurs de la kinase dépendante des cyclines p16INK4a (CDKN2A) et p15INK4a (CDKN2B), qui contrôlent la progression de G1, sont inhibés par la méthylation de l'ADN dans différents types de cancer. De nombreux gènes associés aux processus de réparation de l'ADN sont également hyperméthylés dans les tissus tumoraux. C'est le cas de la MGMT (O-6-méthylguanine-DNMT) qui, dans les tissus normaux, neutralise l'effet négatif de l'alkylation de l'ADN et s'est révélée être réduite au silence dans de nombreux types de carcinomes (Esteller et al., 1999; Weller et al., 2010). Une réparation défectueuse des mésappariements de l'ADN est observée dans une portion significative des cancers gastriques et est associée à une hyperméthylation de MLH1 (Fleisher et al., 2001). Le profil de méthylation de BRCA1, qui est impliqué dans la réparation de l'ADN des coupures double-brin, de la transcription et de la recombinaison, a également été bien étudié dans les cancers du sein et de l'ovaire (Catteau and Morris, 2002). La répression par hyperméthylation de l'ADN de gènes impliqués dans l'adhésion cellulaire CDH1 (E-cadhérine) et CDH13 (H-cadhérine) peut entraîner une invasion et/ou une métastase, et donc une progression tumorale (Katoh, 2005; Kim et al., 1999). Certains des gènes hyperméthylés dans le cancer sont liés à la survie des cellules cancéreuses car ils ont des fonctions pro-apoptotiques, par exemple, la « protéine kinase 1 associée à la mort » DAPK1 (Gordian et al., 2009; Michie et al., 2010). Des facteurs de transcription (par exemple, GATA4 et GATA5) ou des gènes impliqués dans différentes voies de signalisation (par exemple, ESR1, RARB2 et CRBP1) sont aussi souvent hyperméthylés dans le cancer (Akiyama et al., 2003; Esteller, 2003; Widschwendter et al., 2004a, 2004b).

En effet, on dénombre beaucoup plus de modifications du statut de méthylation de l'ADN dans les cancers que de mutations somatiques. Des « CpG Island Methylator Phenotype » (CIMP) ont été caractérisés dans des cancers du cerveau, du sein, de l'endomètre et de l'estomac.

Il convient aussi de mentionner que s'il existe un sous-ensemble de gènes hyperméthylés dans de nombreux types de tumeurs (par exemple, p16INK4a), dans l'ensemble, chaque type de cancer peut être défini par un «méthylome» spécifique corrélé avec l'invasion de la tumeur et le résultat clinique (Liang and Weisenberger, 2017).

Gènes	Fonctions	Cancer du sein	Cancer du poumon	Cancer de la prostate	Leucémie/ Lymphome	Cancer du colon
APC	Antagoniste de la voie de signalisation Wnt qui est impliqué dans la migration cellulaire et l'adhésion	X	X	X	–	X
BMAL1	Composant principal de l'horloge circadienne	–	–	–	X	–
BRCA1	Réparation de l'ADN double brin, Tanscription	X	–	–	–	–
CDH1	E-cadherin, adhésion cellulaire	X	X	X	X	X
CDH13	H-cadherin, adhésion cellulaire	X	X	–	X	X
CDKN2A (p16INK4a)	Inhibiteur de CDK4, control de la progression du cycle cellulaire G1	X	X	X	Xa	X
CDKN2B (p15INK4b)	Inhibiteur de CDK4 et CDK6, control de la progression du cycle cellulaire G1	–	–	–	Xb	–
p14ARF	Control de la progression du cycle cellulaire G1, stabilisant de la protéine suppresseur de tumeur p53	–	–	–	–	X
COX2	Cyclooxygénase, biosynthèse des prostaglandines	–	–	–	–	X
CRBP1	Transport du rétinol nécessaire à la croissance ou à la différenciation des tissus épithéliaux	–	–	–	X	X
DAPK1	Médiateur positif de la mort cellulaire programmée induite par l'interféron γ	–	X	–	X	X
ESR1	Récepteur d'œstrogène, régulation de l'expression génique	X	X	X	X	–
GATA4	Famille GATA de facteurs de transcription à doigts de zinc	–	–	–	–	X
GATA5	Famille GATA de facteurs de transcription à doigts de zinc	–	–	–	–	X
GSTP1	Métabolisme, détoxification et élimination de composés étrangers génotoxiques	–	X	X	–	–
HIC1	Facteur de transcription	X	X	–	X	X
IGFBP3	Protéine de liaison au facteur de croissance analogue à l'insuline	–	X	–	–	–
MGMT	Réparation de l'ADN	–	X	–	Xa	X
MLH1	Réparation des mésappariements d'ADN, signalisation des dommages à l'ADN	–	–	–	–	X
NORE1A	Homologue de l'effecteur Ras	–	X	–	–	–
PYCARD	TMS1 / ASC, voies de signalisation apoptotiques	X	X	–	–	–
RARB2	Le récepteur de l'acide rétinoïque limite la croissance de nombreux types de cellules en régulant l'expression des gènes	–	X	–	X	X
RASSF1A	Inhibe l'accumulation de cycline D1, impliquée dans l'arrêt du cycle cellulaire à la transition de phase G1 / S et la réparation de l'ADN	X	X	X	X	X
TLE1	Homologue de Groucho	–	–	–	X	–
TP73	Famille de facteurs de transcription p53, réponse apoptotique aux dommages de l'ADN	–	–	–	X	–

X, a signalé une hyperméthylation fréquente . [a]une méthylation de l'ADN uniquement dans les lymphomes. [b] Méthylation de l'ADN uniquement dans la leucémie.

Tableau 2 Exemples de répression génique par hyperméthylation des îles CpG dans le promoteur dans les types de cancer les plus fréquents (adapté de (Kulis and Esteller, 2010))

2. Modifications post-traductionnelles des histones

Le nucléosome est l'unité de base de la chromatine. Il est composé d'un octamère d'histones (H3, H4, H2A, H2B) autour duquel s'enroulent 147 paires de bases d'ADN. Le domaine central des histones est très conservé mais leur partie N-terminale dépourvue de structure secondaire est variable. Les modifications intervenant sur les extrémités N-terminales (ou « queues ») des histones H3 et H4 peuvent être associées à l'euchromatine (régions ouvertes, transcriptionnellement actives) ou à l'hétérochromatine (régions fermées, transcriptionnellement réprimées) (Margueron and Reinberg, 2010).

Les PTM sur les résidus acides aminés de queue d'histones sont en mesure d'être établies simultanément sur plusieurs sites. Ces modifications permettent d'établir la structure et la fonction de la chromatine sur une région donnée et elles constituent le « code histone » (Jenuwein and Allis, 2001).

a. Acétylation des histones

L'acétylation des histones est un mécanisme dynamique contrôlant l'activation transcriptionnelle. Deux familles d'enzymes modulent l'acétylation des histones : les histones acétyl-transférases (HAT) et les histones déacétylases (HDAC) (Tableau 3).

Il existe deux types de HAT, les types A et B. Les deux utilisent comme cofacteur l'acétyl-CoA et catalysent le transfert d'un groupe acétyle de celui-ci vers un groupement ε-aminé de la chaine latérale de résidus lysines spécifiques. Ils neutralisent ainsi la charge positive des résidus lysines ce qui affaiblit l'interaction électrostatique entre l'ADN et les histones cibles relaxant la structure de la chromatine qui devient accessible (Bannister and Kouzarides, 2011).

Les HAT de type B sont plutôt cytoplasmiques et ils acétylent seulement les histones nouvellement formées et libres avant leur transport vers le noyau (Xhemalce et al., 2006). Les HAT de type A ont une localisation nucléaire et peuvent être divisés en 3 sous-groupes en fonction de leur séquence d'acides aminés : «GCN5-related N-acétyltransférases » (GNAT) ; « MOZ, Ybf2/Sas3, Sas2 et Tip60 » (MYST), « p300 CREB binding protein » (p300/CREB). Les HAT de type A sont souvent associés à des complexes multi-protéiques (Bannister and Kouzarides, 2011).

Histone	Site modifié	Enzyme	Fonction
H1	Ser27		Activation transcriptionnelle décondensation de la chromatine
H2A	Lys5	Tip60, p300/CBP	Activation transcriptionnelle
H2B	Lys5	p300, ATF2	Activation transcriptionnelle
	Lys12	p300/CBP, ATF2	Activation transcriptionnelle
	Lys15	p300/CBP, ATF2	Activation transcriptionnelle
	Lys20	p300	Activation transcriptionnelle
H3	Lys9	Gcn5, SRC-1	Activation transcriptionnelle, déposition d'histone
	Lys14	Gcn5, PCAF, Esal, Tip60, SRC-1, Elp3, Hpa2, hTFIIIC90, TAF1, Sas2, Sas3, p300	Activation transcriptionnelle, réparation de l'ADN, transcription ARN pol II &III, déposition d'histone sur l'euchromatine
	Lys18	Gcn5, p300/CBP	Activation transcriptionnelle, réparation de l'ADN ; réplication
	Lys23	Gcn5, SAs3, p300/CBP	Aéposition d'histone, activation transcriptionnelle, réparation de l'ADN
	Lys27	Gcn5	Activation transcriptionnelle
H4	Lys5	Hat1, Esal, Tip60, ATF2, Hpa2, p300	Déposition d'histone, activation transcriptionnelle, réparation de l'ADN
	Lys8	Gcn5, PCAF, Esal, Tip60, ATF2, Elp3, p300	Dctivation transcriptionnelle, réparation de l'ADN
	Lys12	Hat1, Esal, Tip60, Hpa2, p300	Déposition d'histone, activation transcriptionnelle, réparation de l'ADN, repression des télomères
	Lys16	Gcn5, Esal, Tip60, ATF2, Sas2	Activation transcriptionnelle, réparation de l'ADN

Tableau 3 Liste non exhaustive des acétylations d'histones leurs enzymes et la fonction induite.

Les HDAC catalysent la déacétylation des lysines et restaurent leur charge positive. Ceci stabilise l'architecture de la chromatine. Les HDAC sont principalement des répresseurs transcriptionnels. La même HDAC peut déacétyler plusieurs sites sur une histone (Xhemalce et al., 2006). Il existe 4 classes de HDAC. Les classes I, II et IV partagent le même mécanisme catalytique qui nécessite un ion zinc mais pas de cofacteur. Les classes I et II comprennent des enzymes proches de celles de la levure scRdp3, scHda1. La classe IV est composée par une seule HDAC, HDAC11. La classe III, nommée *sirtuins*, emploie un mécanisme catalytique qui est NAD+ dépendant (DesJarlais and Tummino, 2016).

b. Méthylation des histones

La méthylation des histones intervient sur les résidus lysine ou arginine des histones, et elle n'altère pas la charge globale des histones.

La méthylation des lysines se fait par des « protein lysine methyltransferases » (PKMT) dont le domaine SET (acronyme de [Su(var)3-9, Enhancer-of-zeste and Trithorax]) va catalyser le transfert d'un groupement méthyl du cofacteur SAM vers un groupe ε-aminé de la chaine latérale des résidus lysines (Bannister and Kouzarides, 2011). Les lysines peuvent être mono - di- ou tri-méthylées. Les méthylations associées aux gènes actifs de l'euchromatine sont H3K4, H3K36, H3K79 et celles associées aux régions de l'hétérochromatine sont H3K9, H3K27 et H4K20 (Tableau 4). Les marques actives H3K4me2/3 sont enrichies sur les TSS des gènes actifs

et H3K36me3 peut être enrichie sur le corps de gène du même gène (Xhemalce et al., 2006). Les méthylations de lysine peuvent être inversées par des méthylases de lysines dépendantes des flavines, les « lysines specific histones demethylases » (LSD). Elles ont pour cofacteur le FAD et n'interviennent que sur des lysines mono- ou di-méthylées (Ng et al., 2009).

D'autre part la méthylation des arginines est catalysée par les PRMT. Leur domaine SET va catalyser le transfert d'un groupement méthyl du cofacteur SAM vers un groupe ω-guanidine de résidus arginine (Bannister and Kouzarides, 2011). Les méthylations des sites H3R17, H3R36 sont associés à une activation des gènes et celles des sites H3R8 et H4R3 sont associés à une répression des gènes (Xhemalce et al., 2006). La famille de protéines qui vont être capables d'inverser la méthylation des arginines sont les protéines Jumonji qui sont des déméthylases dépendantes du 2-oxoglutarate (Ng et al., 2009).

Histone	Site modifié	Enzyme	Fonction
H1	Lys26	Ezh2	Répression transcriptionnelle
H3	Lys4	Set7/9, MLL, ALL-1,	Euchromatine permissive
	Arg8	PRMT5	Répression transcriptionnelle
	Lys9	Suv39h,Clr4, G9a, SETDB1	Répression transcriptionnelle (tri-Me), répression transcriptionnelle, méthylation de l'ADN, activation transcriptionnelle
	Arg17	CARM1	Activation transcriptionnelle
	Lys27	Ezh2, G9a	Répression transcriptionnelle, inactivation X (tri-Me)
	Lys36	Set2	Activation transcriptionnelle
	Lys79	Dot1	Activation transcriptionnelle
H4	Arg3	PRMT1	Activation transcriptionnelle
	Arg3	PRMT5	Activation transcriptionnelle
	Lys20	PR-Set7, Suv4-20h	Répression transcriptionnelle (mono-Me), (tri-Me), activation transcriptionnelle
	Lys59		Répression transcriptionnelle

Tableau 4 Liste non exhaustive des méthylations d'histones avec leurs enzymes et leurs fonctions.

c. Autres modifications post-traductionnelles des histones

Il existe d'autres modifications traductionnelles au niveau de histones comme la phosphorylation, l'ubiquitinylation et la sumoylation.

La phosphorylation des histones se fait sur les sérines, les thréonines et les tyrosines, principalement sur la partie N-terminale des histones. Ce mécanisme est régulé par des kinases et des phosphatases. La phosphorylation induit une augmentation de charges négatives ce qui influence la structure de la chromatine (Oki et al., 2007).

L'ubiquitine est un polypeptide de 76 acides aminés. Trois enzymes (E1 activation, E2 conjugaison et E3 ligation) catalysent l'addition d'ubiquitine aux lysines des histones. La spécificité de l'ubiquitination est très majoritairement due à l'E3 ligase qui détermine la protéine cible (Xhemalce et al., 2006). La mono-ubiquitination des histones H2A et H2B a été clairement impliquée dans la régulation de la transcription : H2Aub est plus fréquemment

corrélée à la répression génique, alors que H2Bub est principalement associée à l'activation de la transcription.

Les enzymes E1, E2 et E3 catalysent également le transfert de la petite molécule « small ubiquitin-like modifier » (SUMO) sur certains résidus de lysines des queues des histones. La sumoylation est associée à une répression transcriptionnelle (Bannister and Kouzarides, 2011) (Tableau 5).

Modification de la Chromatine	Résidus modifiés	Fonction régulée
Phosphorylation	S-ph, T-ph	Transcription, Réparation, Condensation
Ubiquitylation	K-ub	Transcription, Réparation
Sumoylation	K-su	Transcription

Tableau 5 Modification des histones et leurs conséquences fonctionnelles adaptée de (Kouzarides, 2007)

d. Modification des histones et cancer

Les profils d'acétylation et de méthylation des histones sont modifiés dans le cancer. Ceci peut être induit par deux mécanismes principaux: une altération globale des modifications des histones qui va affecter l'intégrité du génome ou une altération du programme d'expression des gènes spécifiques, comportant une mauvaise régulation d'oncogènes ou de gènes suppresseurs de tumeurs (Bannister and Kouzarides, 2011) (Figure 27).

Dans les cancers il y a une réduction globale de la monoacétylation H4K16. Cette perte d'acétylation est médiée par les différentes HDAC qui sont retrouvées surexprimées dans plusieurs types de tumeurs. Les cellules cancéreuses ont généralement une perte intégrale de la marque activatrice H3K4me3 et de la marque répressive H4K20me3. En revanche elles gagnent souvent les marques répressives H3K9me2 et H3K27me3 (Sharma et al., 2010).

Dans la leucémie myéloïde aiguë, l'oncoprotéine de fusion « mixed lineage leukemia » (MLL) induit des motifs aberrants de méthylation de H3K9 et H3K4 qui aboutit à l'altération de l'expression des gènes cibles de MLL (Portela and Esteller, 2010).

Egalement dans les tumeurs hématopoïétiques, la Janus kinase 2 (JAK2), une tyrosine kinase qui joue un rôle essentiel pour l'hématopoïèse normale, est impliquée dans l'altération de l'hétérochromatine. JAK2 peut phosphoryler la position Y41 de l'histone H3 dans les cellules hématopoïétiques et cette phosphorylation prévient la liaison entre cette région de H3 et le répresseur HP1α. Cette modification est impliquée dans la régulation de *lmo2,* un oncogène hématopoïétique clé (Xhemalce et al., 2006).

Dans beaucoup de cancers il y a une dérégulation des méthyl-transférases et en particulier de EZH2, qui est une sous unité du complexe de répression transcriptionnel PRC2. EZH2 est responsable de la marque répressive H3K27me3. EZH2 induit la prolifération et la transformation néoplasique, et est surexprimé dans plusieurs cancers, suggérant une fonction oncogène. EZH2 a été retrouvé aussi inactivé dans certaines tumeurs myéloides malignes suggérant que EZH2 puisse jouer un rôle de suppresseur de tumeur dans des contextes spécifiques. Ceci semble indiquer que le taux de H3K27me3 doit être régulé très

minutieusement pour maintenir l'équilibre homéostasique et une perte de cet équilibre induit le développement de cancers (Bannister and Kouzarides, 2011).

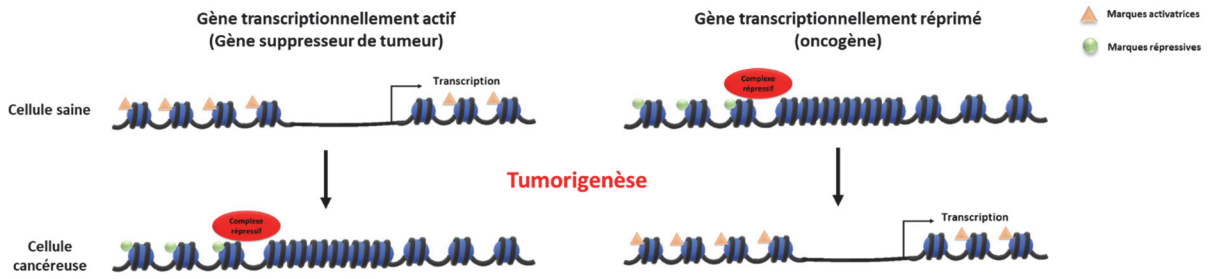


Figure 27 Modifications épigénétiques lors de la tumorigenèse.

3. Les ARN non codants et le cancer

Seulement environ 1% du génome code pour des protéines, la majorité des transcrits étant considérée comme des ARN non codant (ncRNA). Les ncRNA peuvent être divisés en deux sous-groupes : les ncRNAs de petites tailles (< 200 nucléotides (nt)) qui comprennent les « short-interfering RNA (siRNA) », les « micro-RNA (miRNA) » et les « PIWI-interacting RNA (piRNA) », et d'autre part les « long non-coding RNA (lncRNA) » (> 200 nt). Depuis plusieurs années il a été établi que ces ncRNAs jouent un rôle important dans les régulations épigénétiques, car ils peuvent moduler l'expression des gènes et la structure des chromosomes. Dans cette thèse, nous nous intéresserons en particulier aux miRNAs et aux lncRNAs.

a. Les micro-RNA (miRNA) et le cancer

Les micro-RNA sont de petits ncRNA d'une taille d'environ 22 nucléotides. Ils dérivent d'un transcrit primaire (pri-miRNA) contenant une/des structure(s) en tige boucle. Leur transcription fait intervenir l'ARN polymérase II. Dans le noyau, le pri-miRNA est clivé en pré-miARN d'environ 70 nt par Drosha associée à son co-facteur DGCR8. Le pre-miRNA sera exporté vers le cytoplasme où il sera reconnu par Dicer associée à son co-facteur TRBP 27. Le clivage par Dicer à 22 nt de l'extrémité de la tige du pre-miRNA génère un duplex de miRNA contenant un brin 5p et un brin 3p. L'un de ces brins sera incorporé dans un complexe RISC (« RNA induced silencing complex ») comrenant une protéine Argonaute (Ago). Le miRNA cible des ARNm par complémentarité de séquence grâce à sa séquence « seed » ou graine (nt 2 à 7 ou 8 du miRNA chargé dans le complexe) et permettent d'induire leur répression transcriptionnelle, leur déadénylation ou leur dégradation (Krol et al., 2010; Mattick and Makunin, 2005) (Figure 28).

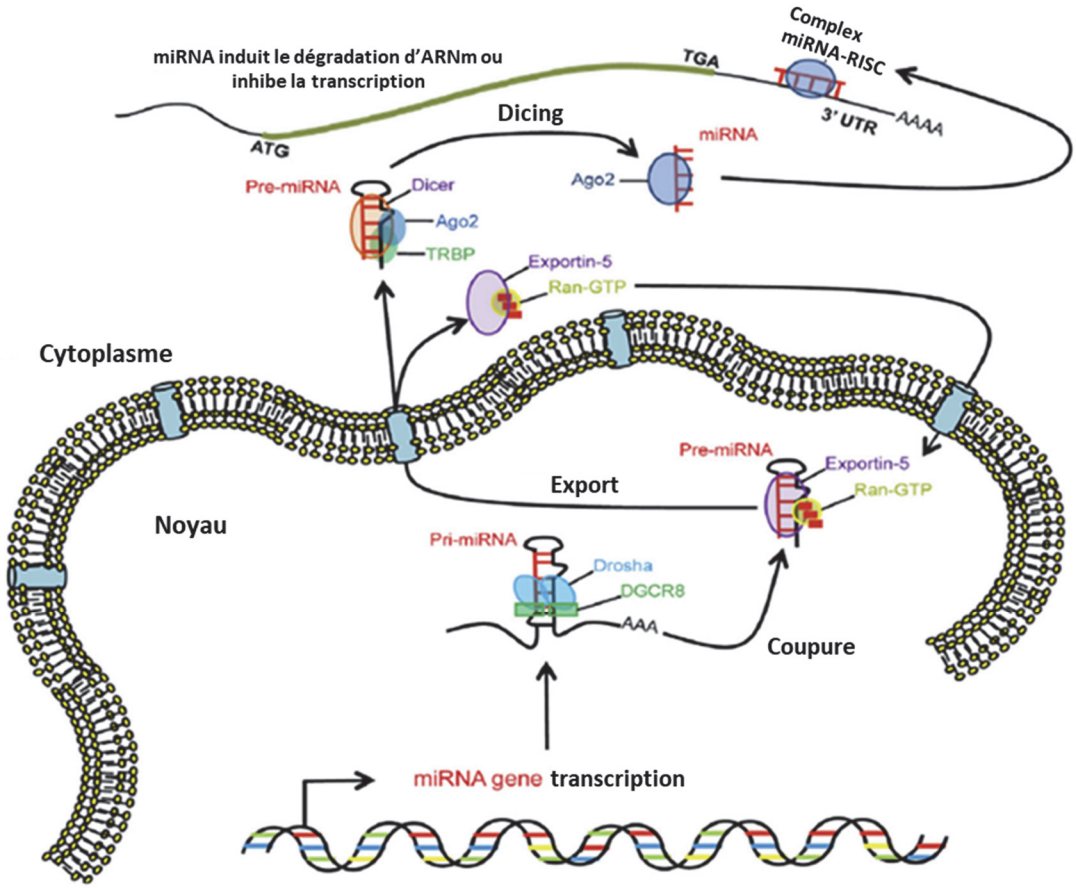


Figure 28 Schéma de la biosynthèse des miRNA (adapté de (Acunzo et al., 2015))

L'expression des miRNAs est régulée au niveau des différents types de cellules et tissus (Mattick and Makunin, 2005). Les miRNAs peuvent induire le remodelage de la chromatine en régulant les modifications histoniques. Ils peuvent aussi affecter la méthylation de l'ADN en régulant les méthylases de l'ADN.

Les miRNAs jouent des rôles fondamentaux dans la carcinogenèse et le développement tumoral (Calin et al., 2002). Il est possible de distinguer un tissu sain d'un tissu cancéreux en analysant le profil d'expression des miRNAs.

Les miRNA sont connus pour réguler l'expression des composants du complexe Polycomb. Des miRNAs dont l'expression est réprimée par des oncogènes ou d'autres acteurs de la tumorigenèse semblent cibler EZH2 induisant sa surexpression (Varambally et al., 2008). Ainsi il a été montré que miR-214 diminue l'expression d'EZH2 et induit la différentiation des muscles squelettiques (Peschansky and Wahlestedt, 2014).

Il existe des mutations qui diminuent l'efficacité du processus de maturation des miRNA, affectant le taux de miRNA matures dans la cellule. Par exemple, une mutation dans l'exportin 5 (gène XPO5) qui va séquestrer le pré-miRNA dans le noyau et empêcher la maturation des miRNAs a été observée dans des cancers avec des instabilités des microsatellites (Melo et al., 2010). Par ailleurs, la diminution de l'expression de Dicer est associée à un mauvais pronostic

du cancer ovarien et contribue à la résistance aux traitements anticancéreux (Acunzo et al., 2015; Hayes et al., 2014; Iorio and Croce, 2012).

Les nouvelles technologies comme les microarray et le séquençage de nouvelle génération (« next generation sequencing » (NGS)), permettent de déterminer le profil d'expression de miRNAs dans des tumeurs solides et de les classer selon ces profils (Acunzo et al., 2015). .

Les miRNAs peuvent être relargués par les tissus et être détectés dans les fluides corporels. Ainsi ils peuvent servir de biomarqueurs afin d'améliorer le diagnostic de certains cancers, de stratifier des patients, d'évaluer le pronostic et à déterminer l'efficacité des traitements. (Hayes et al., 2014; Iorio and Croce, 2012).

b. Les longs non-coding RNA (lncRNA) et le cancer

Les lncRNAs sont des transcrits possédant une taille supérieure à 200 nt qui ne codent pas pour des protéines. Ils possèdent une structure secondaire relativement stable, ce qui leur permet de prendre part dans la régulation de différents processus cellulaires. On peut les classer selon qu'ils sont dans une région intergénique (lincRNA) ou intronique, mais aussi en fonction de l'orientation du brin d'ADN sens ou antisens (Wang et al., 2017a).

Les modes d'action les mieux décrits des lncRNAs sont les modes d'action signal (« signal »), leurre (« decoy »), guide (« guide ») et échaffaudage (« scaffold ») (Figure 29) :

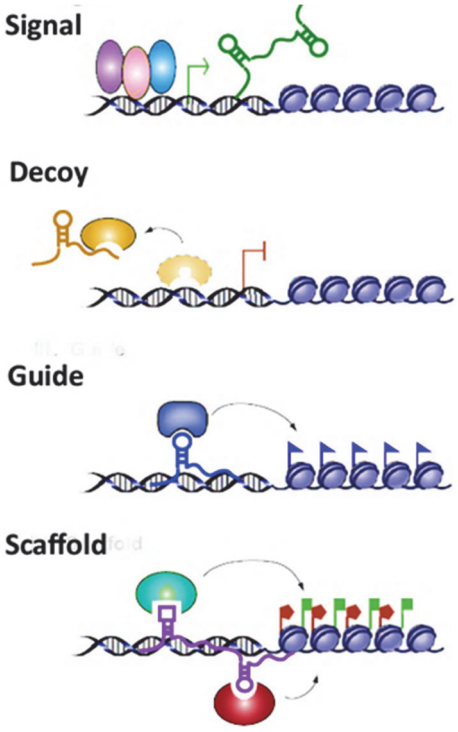


Figure 29 Mécanisme d'action des lncRNAs (adaptée de (Wang and Chang, 2011))

Signal :

La majorité des lncRNAs sont transcrits par la polymérase ARN II. Les lncRNAs ont une expression spécifique selon les types cellulaires et répondant à des stimuli, ce qui suggère que leur expression est sujette à un mécanisme de régulation transcriptionnelle. La transcription des lncRNAs se fait de manière précise dans le temps, c'est pourquoi ils peuvent servir de signal pour répondre à divers stimuli ou à un signal de développement. Des exemples de lncRNA signal sont le lncRNAs Xist, qui intervient pendant l'inactivation du chromosome X, ou HOTAIR, qui est un transcrit antisens de Hox qui peut induire la répression de gènes dans la région de HOXD (Wang and Chang, 2011).

Leurre ou « decoy » :

Ce type de lncRNA se lie à une protéine cible pour en réguler l'activité. Ces lncRNAs modulent l'activité de transcription en séquestrant des facteurs de transcription, des protéines catalytiques, des sous-unités de complexes modulateurs de la chromatine ainsi que des miRNAs. Le lncRNA PANDA, par exemple, lie le facteur de transcription NF-YA pour prévenir l'apoptose médié par p53 (Rinn and Chang, 2012).

Guide :

Les lncRNAs guide interagissent avec des ribonucleoprotéines pour les diriger vers des gènes cibles spécifiques. Ces lncRNAs peuvent agir en *cis* ou en *trans.* Au vu des rôles de ces lncRNAs dans la régulation transcriptionnelle, cela induit des modifications dans la structure de la chromatine qui va avoir des conséquences à distance de la cible. Les lncRNAs qui agissent en *cis* lient généralement des complexes polymérase ARN ou de petits ARNs régulateurs au site actif de transcription. En ce qui concerne le mécanisme en *trans,* il est facilité par des duplexes ADN ou des triplexes ARN-ADN-ADN (Moyo et al., 2016; Wang and Chang, 2011).

Echaffaudage ou « scaffold » :

Ces lncRNAs jouent un rôle structural en favorisant l'assemblage de complexes tels que les complexes de ribonucleoprotéines. Selon la nature des complexes, ils pourront activer ou réprimer la transcription. Ainsi HOTAIR est capable de lier un complexe formé par PRC2 et LSD1/CoREST/REST, qui peut réprimer l'expression de gènes via différents mécanismes.

Par ailleurs, il a été montré que certains lncRNAs peuvent coder pour de courts peptides qui peuvent présenter des activités biologiques (Ingolia et al., 2011; Ji et al., 2015).

Les lncRNAs peuvent avoir une fonction de suppresseur de tumeur ou une fonction oncogénique. Ils ont été associés aux caractéristiques des cancers suivantes : l'autosuffisance en signaux de croissance, l'insensibilité aux signaux inhibiteurs de la croissance, la capacité à éviter l'apoptose, la capacité de se répliquer indéfiniment, l'induction de l'angiogenèse et la capacité à former des métastases (Hanahan and Weinberg, 2011) (Figure 30).

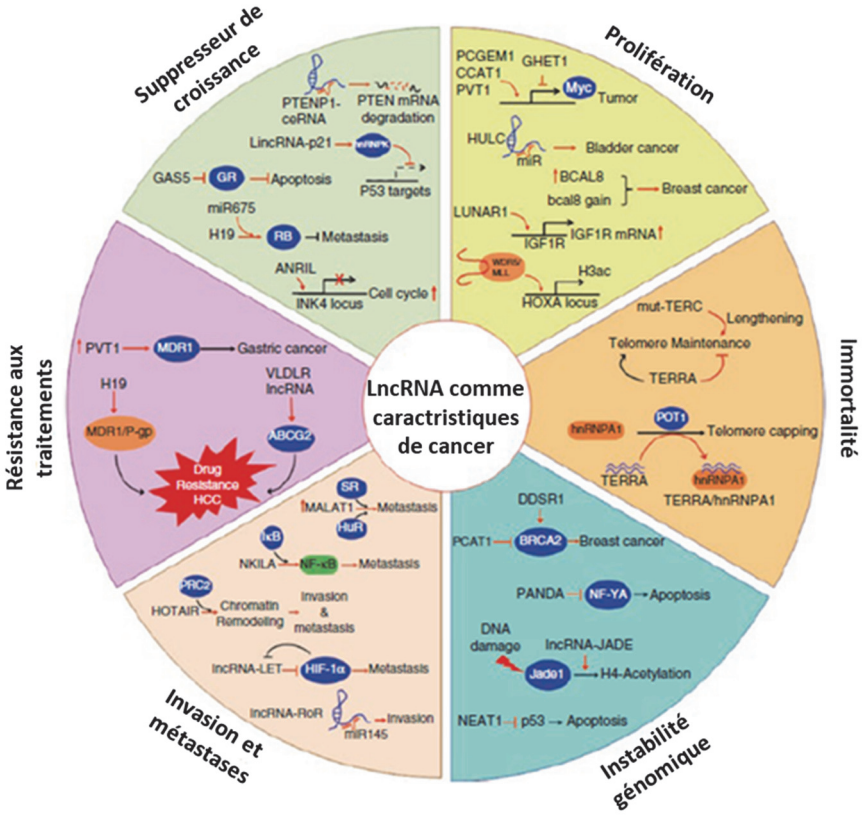


Figure 30 Caractéristiques du cancer et lncRNA (adapté de (Renganathan and Felley-Bosco, 2017))

II. Carcinome hépatocellulaire et épigénétique

1. Inflammation, cirrhose et carcinome hépatocellulaire

a. La fibrose et la cirrhose

L'histoire naturelle de l'infection chronique par HBV est caractérisée par des phases répétées d'inflammation qui s'accompagnent de mort et de régénération cellulaire. Le tissu hépatique peut initialement être réparé avec une restauration complète de l'architecture tissulaire mais si la destruction de cellules hépatiques et l'inflammation demeurent chroniques il y a une progressive accumulation des constituants de la matrice extracellulaire, se traduisant par une cicatrisation excessive ou fibrose. La fibrose est un processus évolutif qui peut être quantifié et classé en 4 stades allant du stade 0 (foie non fibrotique) au stade 4 (fibrose avancée ou cirrhose) (Bedossa, 2015). La cirrhose est le résultat final commun à toutes les maladies inflammatoires chroniques du foie, indépendamment de leur étiologie. La cirrhose est considérée comme une condition pré-néoplasique car le développement du HCC est plus fréquent dans le contexte d'un foie cirrhotique.

Dans les phases d'infection HBV chronique HBeAg positive (« immuno-tolérance ») et d'infection HBV chronique HBeAg négative (« porteurs inactifs ») l'inflammation est minime, alors qu'au cours des phases d'hépatite B chronique HBeAg positive et négative l'inflammation et l'endommagement tissulaires sont importants (Inoue and Tanaka, 2016). Au stade de cirrhose le tissu fibreux occupe environ 20 à 40% du foie avec la perte de l'architecture lobulaire (Bedossa, 2015).

Environ 20% des patients chroniquement infectés par HBV développent une cirrhose au cours du temps. Le développement de la cirrhose peut être favorisé par différents facteurs incluant l'âge, le sexe, le statut sérologique HBeAg, la charge virale élevée et le génotype d'HBV (Chu and Liaw, 2016).

b. Le carcinome hépatocellulaire (HCC)

Les cancers du foie représentent un problème de santé publique majeur avec 850 000 nouveaux cas par an. Ils représentent la deuxième cause de mortalité par cancer en raison des options thérapeutiques limitées (Llovet et al., 2016).

Les principaux facteurs étiologiques associés au développement du HCC sont l'infection chronique par les virus hépatotropes HBV et HCV, l'exposition aux aflatoxines, la consommation chronique d'alcool et les stéato-hépatites métaboliques.

L'infection chronique par le HBV et le « hepatitis C virus » (HCV) ainsi que des troubles métaboliques conduisent à une inflammation chronique, des dommages à l'ADN et des modifications épigénétiques et génétiques qui perturbent des voies oncogéniques.

- Le HBV est responsable d'environ 50% des HCC dans le monde. Les patients chroniquement infectés par le HBV présentent un risque de développer un HCC 10 à 25 fois supérieur aux individus non infectés. La suppression de la réplication du HBV par des NUCs ralentit la progression de la maladie hépatique mais pas le développement de HCC.

  Comparées aux HCCs associés à d'autres facteurs de risque, les tumeurs liées à HBV présentent des taux supérieurs d'anomalies chromosomiques, de mutations inhibitrices de p53 et de surexpression de gènes de progéniteurs hépatiques (Levrero and Zucman-Rossi, 2016).

  L'inflammation chronique et la cirrhose contribuent au potentiel oncogénique d'HBV, mais plusieurs observations suggèrent que le virus joue un rôle direct dans le développement du HCC : a) entre 25 à 30% des HCC liés à l'infection HBV se développent dans un foie non cirrhotique ; b) il existe une corrélation directe entre le niveau de la charge virale HBV et le risque de développement du HCC ; c) certaines mutations et génotypes viraux présentent un risque accru de HCC.

Les mécanismes par lesquels le HBV peut directement contribuer au processus oncogénique sont : a) l'intégration du génome viral dans le génome de la cellule hôte ; et b) les propriétés oncogéniques des protéines virales.

Des études de séquençage de nouvelle génération de HCCs dus au HBV (Fujimoto et al., 2012; Sung et al., 2012) ont montré 1) des insertions de HBV dans des exons ou régions régulatrices de TERT, MLL4 (mixed-lineage leukemia protein 4), CCNE1 (Cyclin 1), SENP5 (Sentrin-specific protease 5), ROCK1 (Rho-associated coiled-coil containing protein kinase), et 2) des mutations dans des gènes codant pour des protéines régulatrices de la chromatine telles que ARID1A, ARID1B, ARID2, MLL et MLL3 dans 50% des cas.

Les principales conséquences de l'intégration virale sont l'insertion directe dans des gènes importants pour le processus oncogénétique (« mutagenèse par insertion »), l'induction d'une instabilité génétique et la génération de protéines virales (HBx et PreS/S) tronquées présentant un potentiel oncogène accru.

La protéine HBx, qui est nécessaire à la réplication du virus, semble aussi jouer plusieurs rôles dans la promotion de la transformation des hépatocytes. HBx contribue à l'instabilité génomique caractéristique du HCC, et encore plus prononcée dans le cas des HCC liés au HBV, en induisant des altérations au niveau des centrosomes et en augmentant les erreurs de ségrégation des chromosomes. HBx peut aussi altérer plusieurs voies de signalisation, les mécanismes de réparation de l'ADN et l'activité de p53. Enfin, HBx interagit avec plusieurs facteurs nucléaires et modificateurs de la chromatine, influençant ainsi le contrôle épigénétique de la transcription de plusieurs gènes de l'hôte (Farazi and DePinho, 2006; Xu et al., 2017). Ces mécanismes seront détaillés dans le chapitre dédié.

L'expression prolongée de HBx et/ou des formes tronquées/mutées de PreS2/S dans les modèles de souris sensibilisent les hépatocytes à d'autres facteurs cancérigènes. Par ailleurs, plusieurs mutations et délétions dans la région PreS2, qui induisent le stress du réticulum endoplasmique et l'accumulation d'espèces réactives de l'oxygène (ROS), ont été associées à un risque accru de HCC chez les patients infectés par différents génotypes dans différentes régions du monde (Europe méridionale, Chine, Afrique sub-saharienne). En effet, le potentiel oncogénique du HBV est maintenu dans tous les génotypes, mais le développement de HCC est plus fréquent chez les patients infectés par les génotypes C et F. Dans le cas du génotype C il semble induire plus d'intégration dans le génome de l'hôte, de cassures double brin dans l'ADN, d'instabilité chromosomique, de dommage à l'ADN et d'accumulation de ROS, tous facteurs contribuant à un plus grand potentiel oncogénique (Xu et al., 2017). Enfin, la capacité de HBV à induire ou à favoriser le développement du HCC est préservée aussi dans le cas de l'infection HBV occulte (persistance de l'ADN viral intégré et/ou de l'ADNccc chez des patients HBsAg négatifs), soutenant ainsi le rôle direct du virus dans la pathogenèse moléculaire du HCC (Ringelhan et al., 2015).

- Le HCV est un virus non cytopathique, possédant un génome ARN (+) membre de la famille des *Flaviviridae* (Bandiera et al., 2016). Le HCV contribue au développement de HCC indirectement par l'induction de l'inflammation hépatique chronique et directement via

ses protéines (Core, NS5A, NS5B…) (Akkari et al., 2012; Simonin et al., 2013). Plus de 95% des HCC liés à HCV se développent dans le contexte d'une cirrhose.

- L'Alfatoxine-B1 (AFB1) est une mycotoxine produite par les champignons de la famille des *Aspergillus* répandue dans certaines régions d'Asie et dans l'Afrique Sub-Saharienne. Ils se développent souvent dans la nourriture comme le maïs et les cacahuètes conservés dans des conditions humides et chaudes. Le Centre International de Recherche en Cancérologie (CIRC) a qualifié cette mycotoxine comme un puissant carcinogène de classe A. Après ingestion l'AFB1 est métabolisée dans la cellule, lie l'ADN et provoque des mutations spécifiques. La mutation du codon 249 dans le gène suppresseur de tumeur p53 est caractéristique des HCC liés à l'exposition à l'AFB1. Dans le sud de la Chine et dans certaines régions d'Afrique Subsaharienne, la présence simultanée d'une exposition à l'AFB1 et de l'infection par le HBV multiplie par 10 le risque de HCC par rapport au HBV seul.

- La consommation chronique d'alcool est un cofacteur important pour le développement du HCC chez les patients atteints d'autres maladies chroniques du foie et la cirrhose alcoolique est en soi une cause importante de HCC.

- La stéatose non alcoolique (non-alcoholic fatty liver - NAFL) est définie par une accumulation de lipides dans les hépatocytes liée à des altérations du métabolisme des lipides et du glucose (syndrome métabolique). L'augmentation du nombre de personnes en surpoids ou obèses a considérablement augmenté la prévalence de cette maladie au niveau mondial dans les dernières années. Entre 10 à 20% des patients avec un NAFL vont développer une forme plus grave de maladie hépatique progressive avec une inflammation chronique et des altérations cellulaires caractéristiques, appelé stéato-hépatite non-alcoolique (non-alcoholic steatohepatitis - NASH) qui évolue en cirrhose et éventuellement en HCC (Castel and Mathurin, 2011; Desbois-Mouthon, 2011). Notamment, de 20 à 40% des HCC chez les patients avec une NASH peuvent se développer dans un foie non cirrhotique. Les éléments pathogéniques à l'origine de la cancérogenèse chez les patients NASH sont l'inflammation chronique avec des taux élevés de cytokines inflammatoires dans le foie (IL6, TNFα, LTB) et une production élevée des ROS qui causent un stress oxydatif et l'accumulation de dommages à l'ADN.

c. Les traitements

Bien que le risque de développer un HCC peut-être réduit par le traitement de la cause sous-jacente (élimination virale ou abstinence alcoolique), il n'existe à ce jour aucune stratégie pour prévenir le développement de cancer chez les patients avec une fibrose avancée ou une cirrhose. Un diagnostic précoce permet d'améliorer la prise en charge mais la détection de petites tumeurs du foie par échographie n'est pas aisée et il n'existe aucun biomarqueur sérique robuste qui pourrait être utilisé dans les programmes de dépistage. Les options thérapeutiques pour le HCC sont limitées et actuellement elles ne sont pas adaptées aux

différentes étiologies. Les patients éligibles pour un traitement curatif peuvent bénéficier d'une résection chirurgicale, une ablation par radiofréquence ou une transplantation hépatique (Galle et al., 2018). Cependant les rechutes sont fréquentes. Chez les patients avec un HCC avancé, qui ne sont pas éligibles aux traitements curatifs, l'inhibiteur de protéine kinase Sorafenib, qui cible les récepteurs au VEGF et la kinase c-Raf, permet d'augmenter de 3 mois la survie médiane de ces patients (Ganten et al., 2017). L'inhibiteur multi-kinase regorafenib, utilisé comme traitement de deuxième intention chez les patients avec une progression tumorale après Sorafenib, a montré une augmentation de la médiane de survie de 10 mois (Finn et al., 2018). Malgré ces avancées et l'espoir suscité par les nouvelles stratégies immuno-modulatrices basées sur les inhibiteurs de checkpoints des cellules T (anti-PD1 et anti-PDL1), il y a un besoin médical urgent de découvrir de nouveaux médicaments efficaces pour le HCC et de définir de nouvelles stratégies de personnalisation du traitement.

2. Epigénétique et carcinome hépatocellulaire

Les modifications épigénétiques tels que la méthylation de l'ADN, les modifications post-traductionnelles d'histones et des modulations de l'expression génique par ncRNA surviennent tôt au cours du développement du HCC.

La modulation du profil de méthylation dans le génome au niveau des promoteurs représente une modification épigénétique importante dans les cancers humains. Plusieurs études ont décrit des méthylations aberrantes dans les HCC (Levrero and Zucman-Rossi, 2016; Singh et al., 2018; Villanueva et al., 2015). Une liste non exhaustive est présentée dans le Tableau 6.

Le rôle que jouent les modifications d'histones dans les maladies du foie et la pathogenèse du HCC est moins bien connu. Cependant il a été montré que plusieurs HDACs sont surexprimées dans les HCCs (Lachenmayer et al., 2012; Wu et al., 2010) et un traitement combinant des inhibiteurs de HDAC et le sorafenib permet d'augmenter l'efficacité du traitement et la survie dans des modèles d'HCC in vivo (Lachenmayer et al., 2012) et chez les patients (Bitzer et al., 2016).

L'activation de EZH2 et du complexe PRC2, due à la surexpression d'EZH2 plutôt qu'à des mutations activatrices, est fréquemment observée dans les cancers de la prostate, du sein et dans les HCCs (Simon and Lange, 2008). En tant que sous-unité catalytique de PRC2, EZH2 catalyse la triméthylation de H3K27 (H3K27me3) afin de réprimer l'expression de gènes qui régulent le développement et la différentiation cellulaire. Au cours du cancer, l'activation de EZH2-PRC2 réprime l'expression de gènes suppresseurs de tumeur mais EZH2 peut également activer la transcription indépendamment de PRC2.

Plusieurs études ont également montré que 20% de toutes les tumeurs, y compris des HCCs dus à l'abus d'alcool et aux virus des hépatites, comportent des mutations dans des protéines qui régulent l'organisation et les fonctions de la chromatine (*ARID1A, ARID1B, ARID2, MLL,* et *MLL3*) (Guichard et al., 2012; Li et al., 2011; Sung et al., 2012).

Chromosome	Gène
1q	RASSF5
3p	RASSF1
3p	ROSSF1A
3p	CTNNB1
3p	TGF-1Br11
5q	APC
8p	DLC-1
8p	LPTS
8p	CSMD1
9p	CDKN2A
9p	CDKN2B
10q	NKX6-2
10q	SFRP5
11p	KAI1
11p	IGF-2
11p	TSLC1
16q	CDH1
16p	AXIN1
16p	SOCS-1
22q	NEFH

Tableau 6 Liste de gènes présentant une méthylation aberrante dans les HCCs (adapté de (Singh et al., 2018; Villanueva et al., 2015))

3. HBx, épigénétique et HCC

HBx est connue pour induire de nombreuses modifications dans des voies de signalisation potentiellement impliquées dans la cancérogenèse : a) HBx intervient dans la réponse aux dommages à l'ADN en réduisant la capacité de réparation; b) HBx conduit au stress oxydatif et à des dommages dans les mitochondries ; c) HBx interfère avec les voies de signalisation activées par plusieurs cytokines ; d) HBx possède des propriétés pro-apoptotiques et anti-apoptotiques (Figure 31) (Guerrieri et al., 2016; Levrero and Zucman-Rossi, 2016; Liu et al., 2016; Slagle and Bouchard, 2016).

HBx possède la capacité de modifier la transcription de gènes cellulaires en modulant les modifications épigénétiques via l'interaction avec de nombreux facteurs de transcriptions, et d'enzymes de modification de la chromatine. HBx peut interférer avec différentes HDAC, les DNMTs, le complexe Polycomb dont les fonctions sont d'induire des modifications épigénétiques et d'entrainer des modifications de l'expression des gènes - et ainsi engager les cellules sur la voie de la tumorigenèse (Levrero and Zucman-Rossi, 2016).

HBx peut former un complexe avec CBP, p300 et la « methyl-CpG binding domain protein 2 » (MBD2). Ce complexe est recruté sur les promoteurs P3 et P4 de « l'insulin-like growth factor-2 » (IGF-2), ce qui va induire des hypométhylations et la transcription des deux promoteurs et augmenter l'expression de l'IGF-2 dans le HCC (Liu et al., 2015b). L'hyperméthylation du promoteur de CD82, un gène suppresseur de tumeur qui prévient des métastases, médiée par HBx inhibe son expression (Yu et al., 2014). Le recrutement du complexe mSin3A/HDAC par HBx sur le promoteur du gène E-cadherin (CDH1) va induire une déacétylation et donc l'inhibition de l'expression de CDH1 (Arzumanyan et al., 2012). On observe une

hypométhylation de multiples régions GpC chez des souris HBx transgéniques avant le développement de HCC. La plupart de ces régions sont impliquées dans la tumorigenèse.

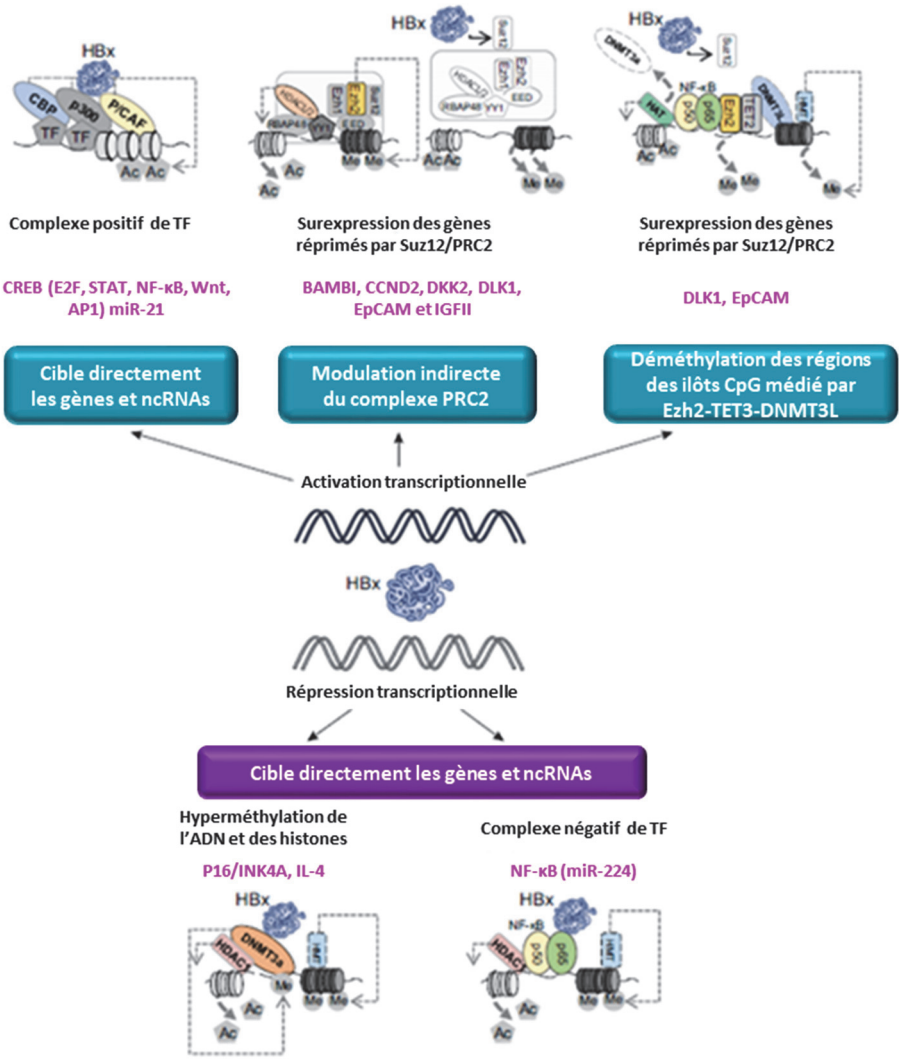


Figure 31 HBx et modifications épigénétiques(adapté de (Levrero and Zucman-Rossi, 2016))

4. Rôle des ncRNA pendant l'infection et l'HCC

a. miRNA et l'infection HBV

Le rôle des miRNAs dans les infections virales est encore une source de controverse. De manière générale les miRNAs possèdent une activité antivirale, mais de plus en plus d'études démontrent qu'ils peuvent être détournés par le virus afin d'établir un environnement propice à sa réplication et à la persistance virale. En effet les miRNAs vont jouer un grand rôle dans la régulation des gènes de l'hôte ou en ciblant directement les transcrits viraux.

Lors de l'infection par HBV il y a une forte interaction entre plusieurs grandes familles de miRNAs et le virus :

- Le miR-122 est fortement exprimé dans le foie et y représente environ 70% de la population totale des miRNAs. Beaucoup d'études ont cherché à déterminer quelle est l'interaction entre le miR-122 et HBV. Les première études dans des cellules transfectées par HBV et en présence d'une surexpression du miR-122 montrent une forte diminution des taux d'HBeAg et HBsAg (Lamontagne et al., 2015). La régulation d'HBV par le miR-122 pourrait être médiée par la cycline G1 qui est une cible directe de miR-122. La cycline G1 interagit avec p53, inhibant l'interaction de p53 sur les enhancers I et II d'HBV induisant une répression de la transcription d'HBV (Fornari et al., 2009; Gramantieri et al., 2007). D'autres études ont montré dans des cellules HepG2.2.15 ou dans des cellules HepG2 transfectées que le miR-122 peut supprimer l'expression des protéines et la réplication d'HBV (Chen et al., 2011; Fan et al., 2011; Peng et al., 2014).

- La famille des miR-125 est l'une des plus étudiées car son homologue chez *C. elegans, lin-4*, est le premier miRNA a avoir été décrit. En comparant l'expression de miR-125 dans des cellules HepG2.2.15 et des cellules HepG2 il a été montré qu'il était surexprimé en présence de HBV. Ceci a été confirmé dans chez des patients infectés par HBV (Lamontagne et al., 2015). Cette surexpression pourrait jouer un rôle en faveur d'HBV pour réguler sa réplication en contrôlant le taux de miRNA ayant un rôle anti-viral.

- La famille de miR-15 possède 6 membres (miR-15a, miR-15b, miR-16-1, miR-16-2, miR-195, miR-497) et une activité suppresseur de tumeur. Elle est très exprimée dans différents cancers incluant le HCC. Les cibles de cette famille sont principalement des régulateurs du cycle cellulaire et de l'apoptose. L'expression des membres de cette famille est réduite dans des cellules transfectées HBV, dans des cellules HepG2.2.15, dans des foies infectés par HBV et dans des lignées exprimant HBx (Wang et al., 2013; Wu et al., 2011). En effet HBx est capable d'induire c-Myc qui va à son tour induire la répression de miR-15a/miR-16. Une analyse avec des techniques de « pull-down » a démontré que le miR-15a peut se lier à l'ARN d'HBx et il a été émis l'hypothèse que HBx peut séquestrer des miRNAs afin de diminuer leurs taux dans la cellule (Wang et al., 2013).

D'autres familles de miRNAs sont impliquées dans la régulation de la réplication comme la famille des let-7, miR-199a-3p, leur effet sur la réplication d'HBV sont présentés brièvement dans le Tableau 7 ci-après (Deng and Lu, 2016; Hayes and Chayama, 2016; Lamontagne et al., 2015):

miRNA	Effet sur la réplication d'HBV
miR-1	Augmentée
miR-15b	Augmentée
miR-18a	Augmentée
miR-372/373	Augmentée
miR-449a	Augmentée
miR-34c	Diminuée
miR-122	Diminuée
miR-130a	Diminuée
miR-141	Diminuée
miR-155	Diminuée
miR-370	Diminuée
miR-939	Diminuée

Tableau 7 Liste de quelques miRNA et leur effet sur la réplication d'HBV (adapté de ((Deng and Lu, 2016))

Du fait de la modification de l'expression de beaucoup de miRNAs lors de l'infection par HBV, on retrouve- également leur expression modifiée dans le HCC induit par HBV. Ces modifications d'expression de miRNAs vont notamment permettre de définir de nouveaux biomarkers qui permettront l'identification des HCC liés à HBV (Hayes and Chayama, 2016).

b. lncRNA dans le HCC induit par HBV

Depuis quelques années de nombreuses études ont démontré le rôle des lncRNAs dans plusieurs cancers et notamment le HCC. Plusieurs lncRNA sont associés au HCC et certains lncRNAs pourraient être spécifiques de l'infection HBV (Figure 32). On retrouve 83 lncRNA derégulés associés au HCC dont 31 sont aussi associé à HBV et 52 à HCV (Pan et al., 2013). Aucun des transcrits n'est commun aux deux virus et ceux qui sont liés à HBV sont associés à la transcription, à la transduction du signal et aux métastases (Iizuka et al., 2002).

- Le lncRNA « highly up-regulated in liver cancer » (HULC) est présent en plus grande concentration dans le tissu cancéreux que dans le tissu adjacent. HULC semble agir comme une « éponge » endogène afin de réduire la fonction du miR-372. Ce miRNA est impliqué dans la régulation du cycle cellulaire, l'apoptose, l'invasion, la prolifération dans différents cancers (Wang et al., 2010). On retrouve une surexpression de HULC dans des cellules produisant HBV (Du et al., 2012). La détection de HULC dans le sérum de patients atteints de HCC suggère une utilisation possible comme biomarqueur.
- Le gène H19, qui se trouve en aval du gène IGF-2, code pour le lncRNA H19 qui est normalement exprimé uniquement pendant le développement embryonnaire. Normalement l'allèle maternel est actif et l'autre allèle inactif grâce à une empreinte épigénétique. H19 possède une activité suppresseur de tumeur (Moyo et al., 2016). Lorsqu'il y a une activation bi-allélique, H19 joue un rôle dans les mécanismes épigénétiques impliqué dans le HCC (Kim and Lee, 1997). La surexpression de H19 est

régulée par Myc. Cette surexpression est retrouvée dans les HCCs induits par HBV et elle est associée aux métastases, l'invasion et le développement tumoral (Iizuka et al., 2002).

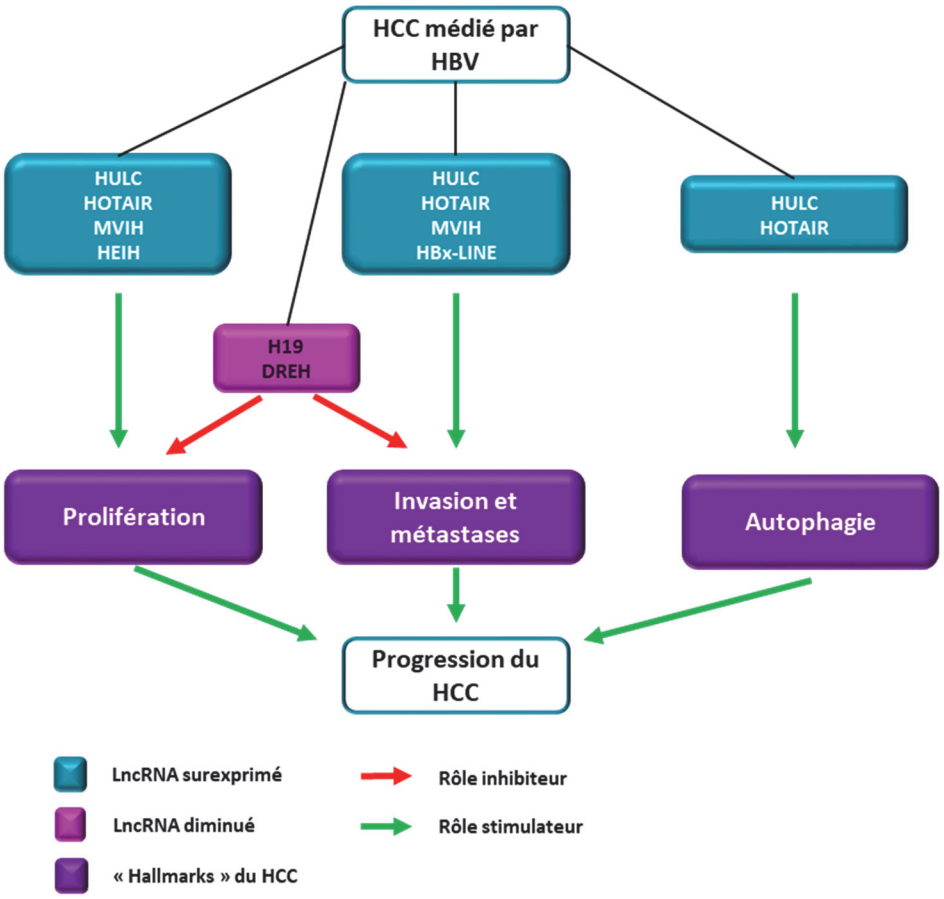


Figure 32 Exemple de régulation de lncRNA dans le HCC médié par HBV (adapté de (Moyo et al., 2016; Qiu et al., 2017))

- HOTAIR est un lncRNA transcrit à partir du locus *HOXC*. HOTAIR interagit avec plusieurs protéines de modification de la chromatine. En 5' il lie le complexe PRC2 et en 3' il interagit avec un complexe comprenant la « lysine-specific demethylase 1 » (LSD1), le «co-repressor element-1 silencing transcription factor » (Co-REST) et HDAC1. HOTAIR possède un mécanisme d'échauffaudage (« scaffold ») et de guide (Moyo et al., 2016). L'interaction de HOTAIR avec la «polo-like kinase 1 » (Plk1), SUZ12 (un membre de PRC2) et la «zinc finger protein 198 » (ZNF198) a été impliquée dans la transformation des hépatocytes causée par HBx (Studach et al., 2012).

D’autres lncRNA sont impliqués dans le HCC lié à HBV et quelques exemples sont décrits dans le Tableau 8 ci-dessous :

Nom du lncRNA	Description du gène	Dérégulation	Fonction biologique dans HCC médié par HBV
DREH		Répression	Inhibe le développement du HCC et des métastases en interagissant avec la vimentine
HBx-LINE	Fusion de HBx avec le LINE1 human lors de l’intégration d’HBV	Surexpression	Induit le développement du HCC en activant la voie Wnt/β-catenine
HEIH	lncRNA associé à EZH2	Surexpression	Induit le développement du HCC en interagissant avec EZH2 et en suprimant p15, p16, p21 et p57
MVIH	Localisé dans un intron de la protéine ribosomal s24	Surexpression	Interagit avec PGK-1, inhibe sa sécrétion et augmente l’angiogenèse

Tableau 8 Exemple de lncRNA impliqué dans le HCC médié par HBV (adapté de (Moyo et al., 2016; Qiu et al., 2017))

Partie 2 : Etudes de la thèse

Matériels et Méthodes

Cette partie comporte une description « graphique » et « succincte » des méthodes utilisées au cours de cette thèse basée sur des schémas. Les détails des matériels et méthodes sont décrits dans chaque publication.

1. Production de virus et infection

Afin de préparer les stocks d'inoculum d'HBV nous utilisons la lignée HepAD38, une lignée stable dérivée des cellules HepG2 qui contient un transgène HBV (génotype D souche ayw) et produit le virus (Figure 33)

Figure 33 Production du virus

2. Préparation des hépatocytes primaires

Des pièces hépatiques sont utilisées pour la préparation des hépatocytes primaires. Elles sont issues d'une résection hépatique chez des patients présentant des métastases du foie suite à un cancer colorectal. (LeCluyse and Alexandre, 2010) (Figure 34).

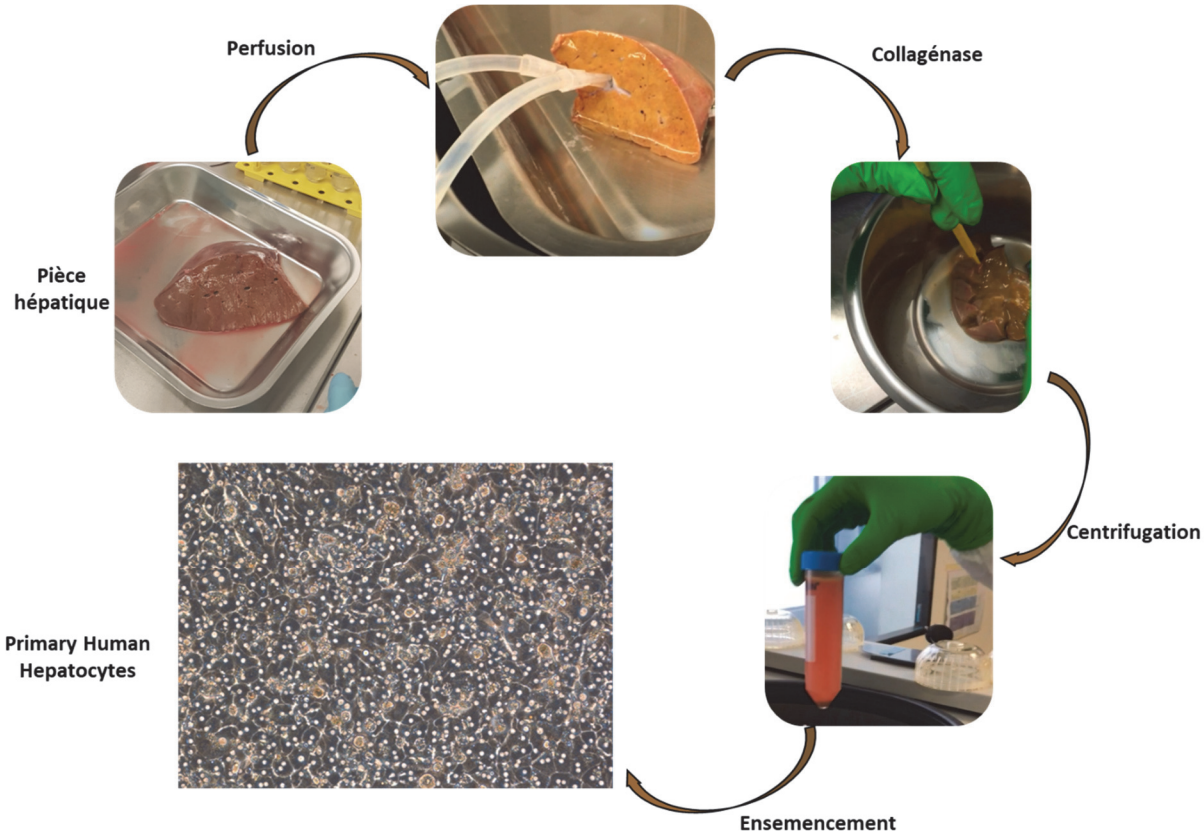


Figure 34 Préparation des hépatocytes primaires

3. Chromatin immunoprecipitation (ChIP)

L'immuniprécipitation de la chromatine (ChIP) est utilisée afin de déterminer le recrutement de protéines d'intérêt sur différentes régions du génome (Figure 35).

Après infection des cellules, celles-ci sont *« cross-linked »* (fixation des protéines sur la chromatine) puis la chromatine est fragmentée par sonication. Des anticorps spécifiques pour la protéine d'intérêt ou dirigés contre des modifications post-traductionnelles spécifiques des histones sont utilisés pour immunoprécipiter la chromatine. Après une étape de « reverse-crosslink » on obtient de l'ADN qui peut être utilisé : a) en qPCR, afin de vérifier soit le recrutement de la protéine d'intérêt sur un gène déterminé ou les modifications post-traductionnelles d'intérêt sur les histones dans la région amplifiée ; b) ou pour le séquençage à haut débit afin de déterminer toutes les régions dans un génome où la protéine est recrutée où la modification post-traductionnelle d'intérêt des histones est présente.

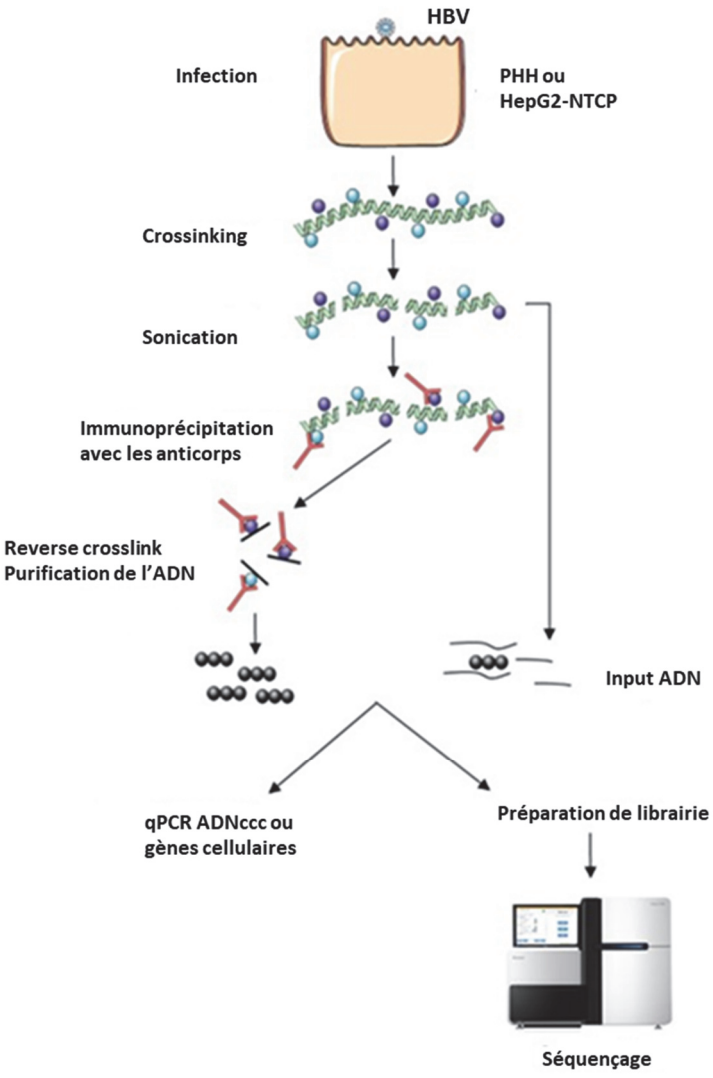


Figure 35 Représentation schématique des techniques de chromatin immunoprecipitation (ChIP) et ChIP-Seq

4. RNA immunoprecitpitation (RIP)

L'immunoprecipitation de l'ARN (RIP) est utilisée afin de déterminer le recrutement de protéines d'intérêt sur l'ARN (Figure 36).

Après infection des cellules, celles-ci sont « cross-linked » (fixation des protéines sur les ARN) puis la chromatine est fragmentée par sonication et des anticorps spécifiques de la protéine d'intérêt sont utilisés pour immunoprécipiter l'ARN. Après un reverse-crosslink on obtient l'ARN immuno-précipité qui peut être utilisé : a) en RTqPCR, afin de vérifier la liaison de la protéine d'intérêt sur un ARN déterminé ; b) en séquençage à haut débit afin de déterminer tous les ARN où la protéine est recrutée.

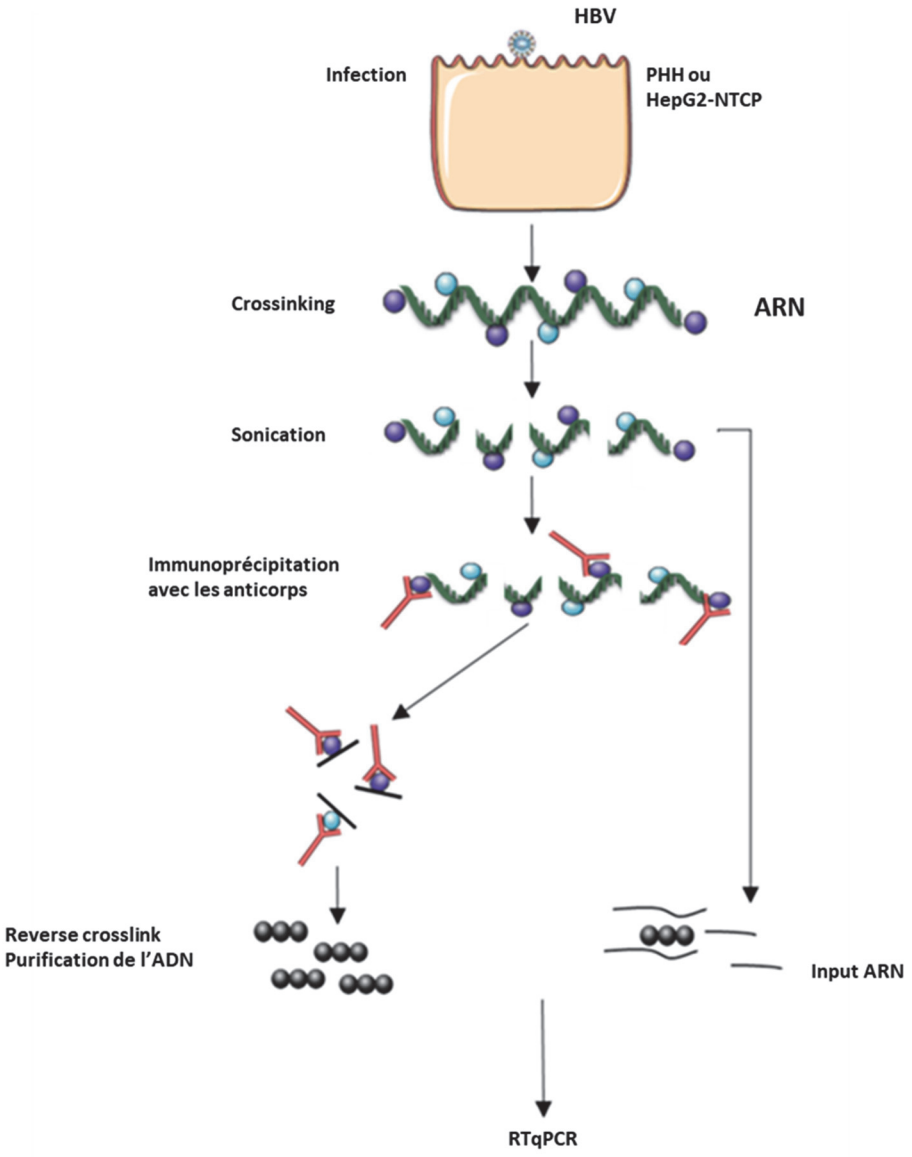


Figure 36 Représentation schématique de la technique d'RNA immunoprecipitation

5. Chromatin isolation by RNA immunoprecipitation (ChIRP)

« Isolation de la chromatine par immunoprécipitation des ARN »

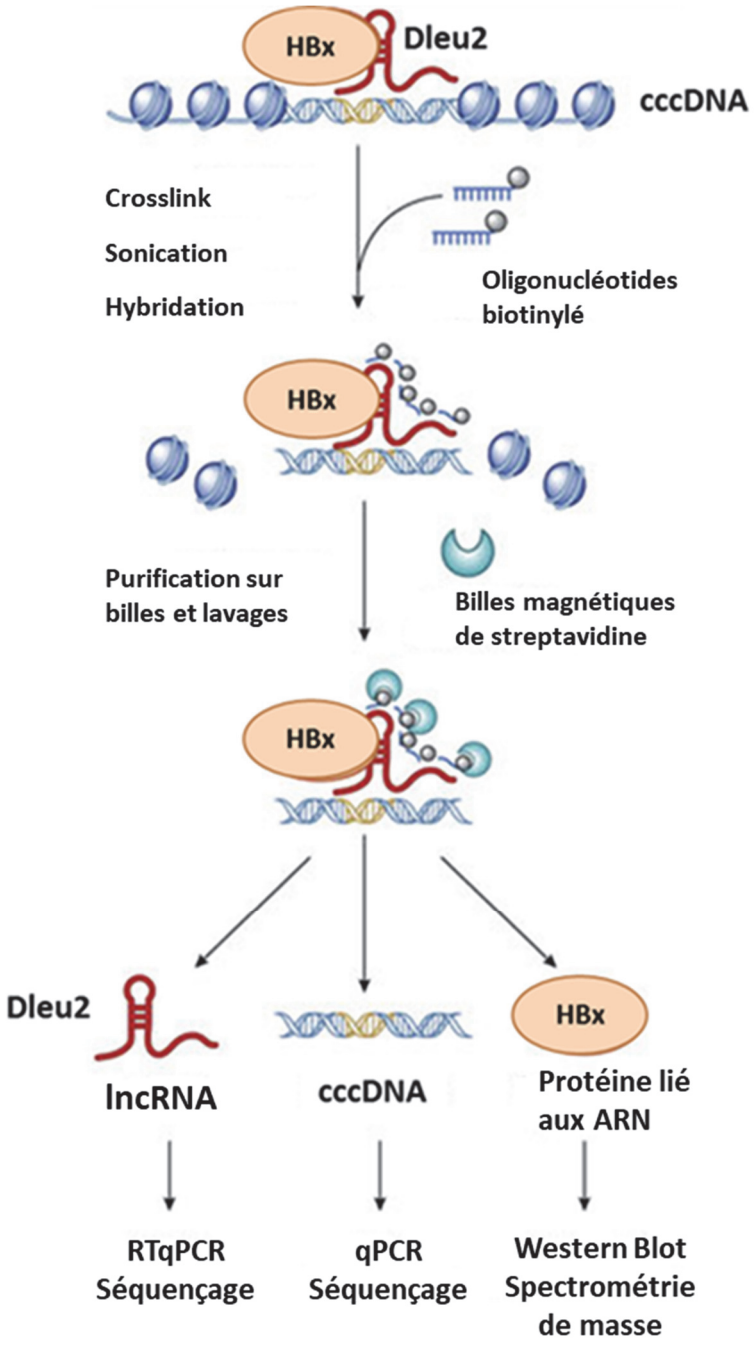


Figure 37 Schéma représentant la technique de Chromatine isolation by RNA immunoprecipitation

6. Assay for Transposase-Accessible Chromatin (ATAC-Seq)

La méthode d'«Assay for Transposase-Accessible Chromatin » (analyse de la chromatine accessible à la transposase) permet grâce au séquençage à haut débit de déterminer les régions accessibles de la chromatine. En comparaison des autres méthodes d'accessibilité de la chromatine l'ATAC-seq utilise un très faible nombre de cellules (50000 cellules) et aucun agent chimique qui pourrait induire des biais dans l'expérience. Cette technique est basée sur la capacité de la transposase 5 (Tn5) à couper l'ADN entre les nucléosomes de manière très régulière. Ensuite les fragments sont amplifiés et une librairie de séquençage est préparée afin de pouvoir séquencer directement les fragments (Buenrostro et al., 2013) (Figure 38 Figure 39).

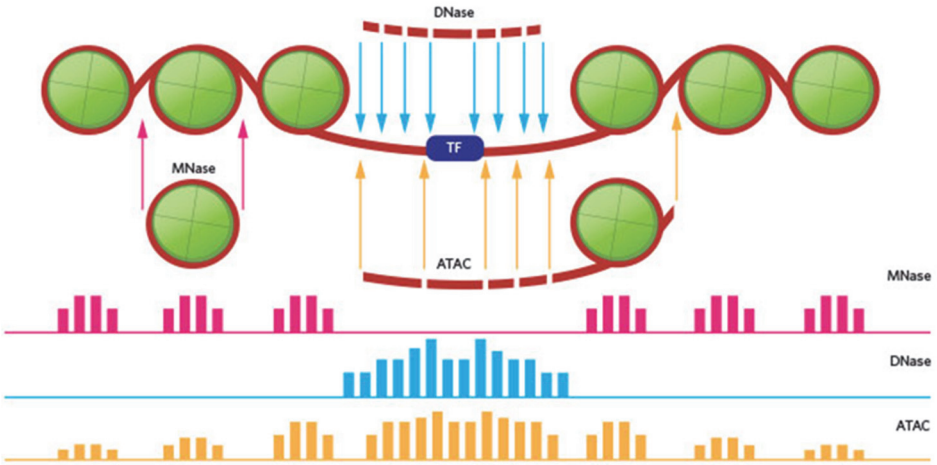


Figure 38 Comparaison schématique des différentes méthodes d'accessibilité de la chromatine (https://www.biostars.org/p/209592/)

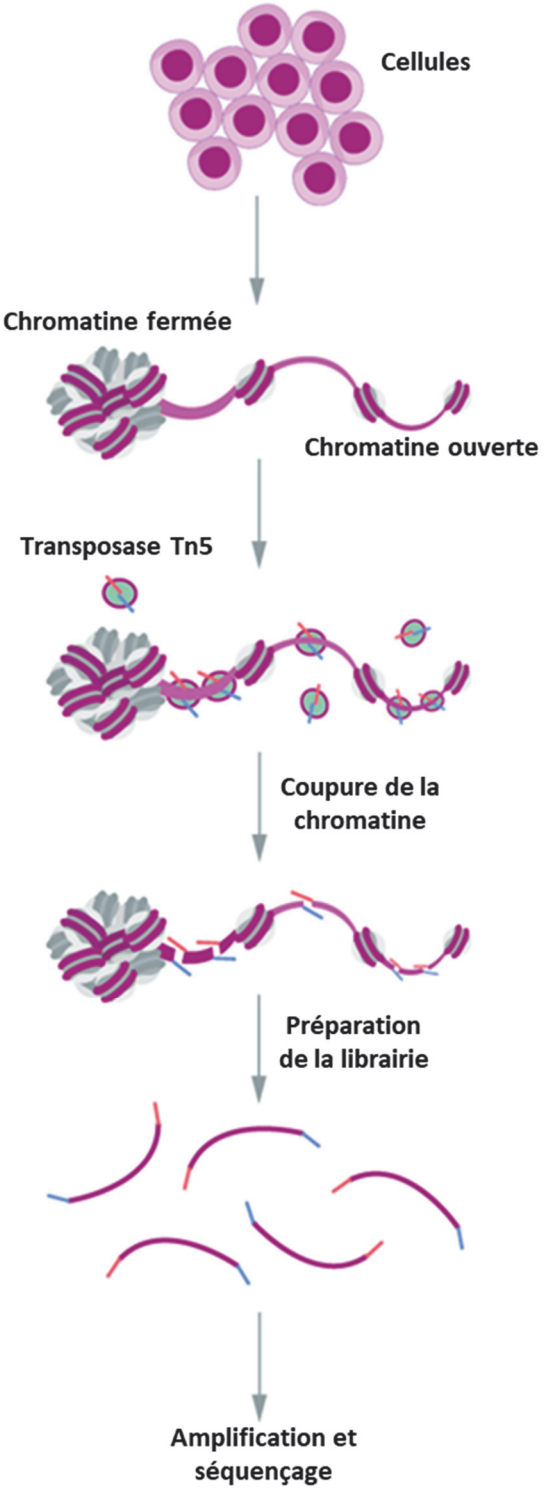


Figure 39 Schéma représentant les différentes étapes de l'ATAC-seq (adapté de (Buenrostro et al., 2013))

7. RNA-Seq

La méthode de séquençage de l'ARN (RNA-seq) permet, à partir des ARN totaux ou des ARN messagers (enrichissement avec des billes oligo-dT), et après reverse transcription en ADNc (conservation du sens des transcrits) de préparer des librairies pour le séquençage de l'ensemble des ARNs présents dans les différentes conditions choisies (Figure 40).

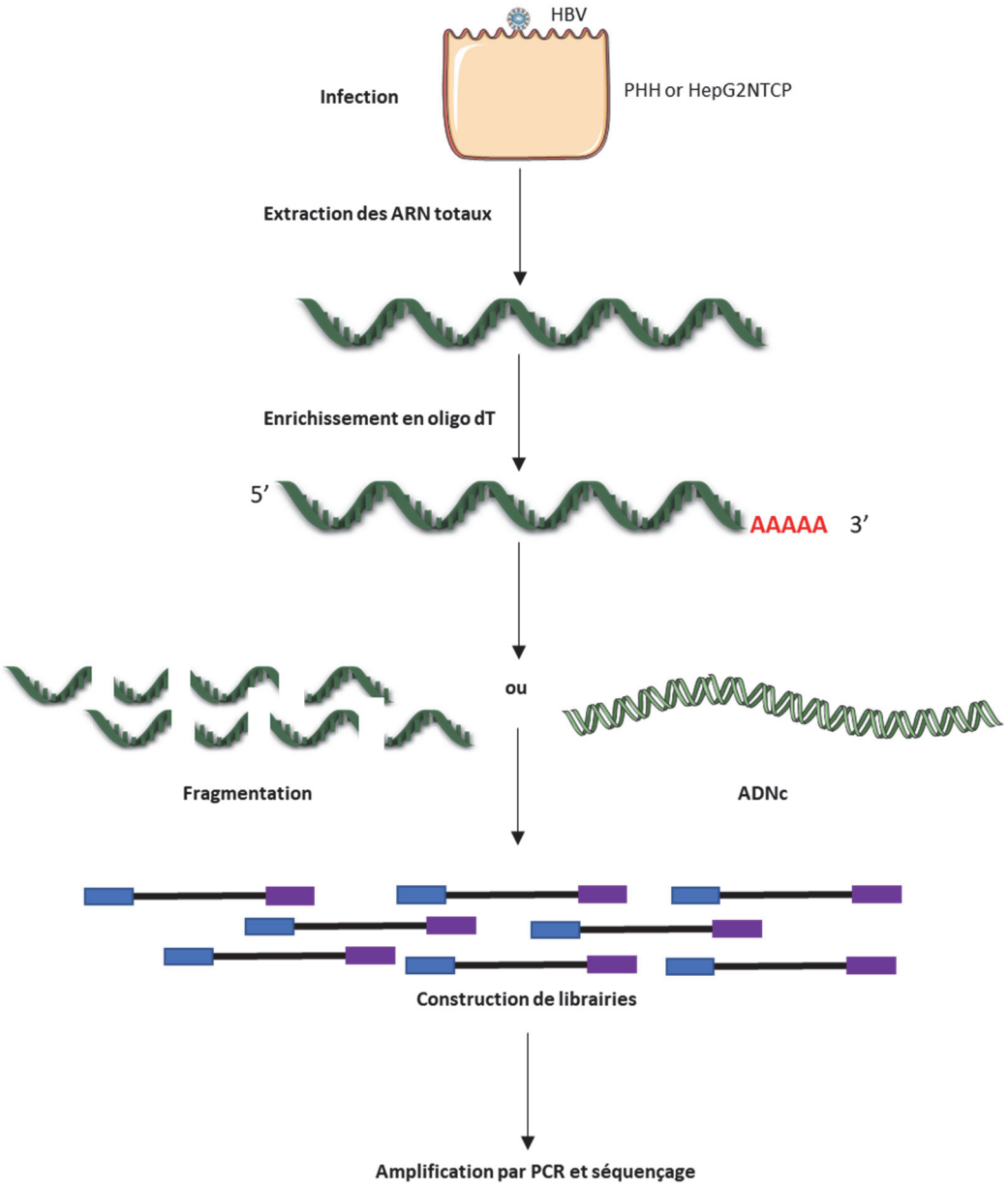


Figure 40 Schéma représentant les différentes étapes du RNA-seq

8. Résumé des techniques génomiques utilisées

Technique	Méthode	Cible/Identification
ChIP	Immunoprécipitation	ADN cible lié à une protéine
RIP	Immunoprécipitation	ARN cible lié à une protéine
ChIRP	Immunoprécipitation	ADN cible lié à un ARN via une protéine
ATAC-seq	Tagmentation	Régions accessibles du génome
RNA-seq	Enrichissement oligo dT	Transcriptome

Tableau 9 Résumé des différentes techniques utilisées.

Objectifs des études

L'objectif de ma thèse était de contribuer à élucider l'interaction entre le HBV et les cellules infectées dans la pathogenèse virale, la progression de la maladie et le développement du HCC. En particulier, j'ai abordé deux points spécifiques : a) l'impact de HBV sur l'accessibilité de la chromatine et le contrôle épigénétique de la transcription des cellules de l'hôte ; b) le rôle de la protéine HBx dans la transcription des gènes cellulaires et son interaction avec la machinerie épigénétique.

Plusieurs études ont montré qu'il y avait relativement peu de modifications de la transcription et surtout peu d'induction de la réponse immunitaire innée suite à une infection aiguë par le HBV dans des modèles animaux (Wieland et al., 2004) ou dans des cellules infectées *in vitro* (Niu et al., 2017; Suslov et al., 2018). Ces résultats ont conduit au concept, largement accepté, selon lequel le HBV est un virus « stealth » ou « furtif » et qu'il n'induit pas ou très peu de modifications dans la cellule hôte. Cependant, d'autres études indiquent clairement que le virus a un impact sur la transcription des cellules hôtes (Ancey et al., 2015 ; Lamontagne et al., 2016), qu'il est détecté par les mécanismes de l'immunité innée (Lebossé et al., 2017; Luangsay et al., 2015; Lucifora et al., 2018) et, plus important, que l'infection virale a un impact profond sur les gènes de l'immunité innée dans le foie des patients chroniquement infectés au cours des différentes phases de la maladie (Lebossé et al., 2017; Luangsay et al., 2015; Lucifora et al., 2018).

La protéine HBx a toujours été indiquée comme l'acteur principal dans la modulation de la transcription des gènes cellulaires, soit par interaction directe avec la machinerie transcriptionnelle, soit en activant des voies de signalisation contrôlant des facteurs de transcription et des programmes de transcription spécifiques. La découverte de sa capacité a lier le minichromosome viral (Belloni et al., 2009; Lucifora et al., 2011) et de son rôle essentiel dans l'initiation (Decorsière et al., 2016) et le maintien (Belloni et al., 2009; Rivière et al., 2015) de la transcription de l'ADNccc ont confirmé son importance dans le cycle viral et ont permis d'identifier HBx en tant que cible thérapeutique. Malgré le grand nombre de publications, le rôle, l'étendue et les mécanismes précis par lesquels HBx affecte la transcription des gènes de l'hôte dans le contexte d'une infection virale et *in vivo* sont encore discutés.

Les résultats obtenus lors de ma thèse sont présentés dans 3 études :

a) dans la première étude, nous avons utilisé une approche d'immunoprécipitation de la chromatine suivie d'un séquençage à haut débit (ChIP-Seq) pour identifier toutes les cibles génomiques de HBx (gènes codant les protéines et ncRNA) dans les cellules qui répliquent le HBV. Nous avons aussi caractérisé la fonction d'un sous-groupe de miRNA réprimés par HBx ;
b) dans la deuxième étude, parmi les lncRNAs ciblés par HBx, nous nous sommes concentrés sur DLEU2, qui est fortement surexprimé dans l'infection par le HBV et dans le HCC, et nous avons étudié l'interaction entre DLEU2, HBx et l'histone méthyltransférase EZH2, la sous-unité catalytique du complexe répressif PRC2.
c) enfin, dans la troisième étude nous avons caractérisé par ATAC-Seq (assay for transposase-accessible chromatin with highthroughput sequencing) les changements d'accessibilité de la chromatine imposés par HBV dans les hépatocytes primaires humains infectés.

Etude 1

Genome-wide identification of direct HBx genomic targets.

Guerrieri F, Belloni L, D'Andrea D, Pediconi N, Le Pera L, Testoni B, Scisciani C, Floriot O, Zoulim F, Tramontano A, Levrero M.

BMC Genomics. 2017 Feb 17;18(1):184. doi: 10.1186/s12864-017-3561-5.

Des fonctions multiples de la protéine HBx ont été décrites. Son rôle principal est de promouvoir la transcription de l'ADNccc. HBx se lie et active le complexe ubiquitine ligase CRL4-DDB1, conduisant à la dégradation par le protéasome du facteur de restriction Smc5/6 et l'activation de la transcription de l'ADNccc. En l'absence de HBx les complexes contenant Smc5/6 inhibent la transcription de l'ADNccc et la production de l'ARNpg (Decorsière et al., 2016). La présence de HBx est également nécessaire pour maintenir la transcription de l'ADNccc en empêchant la liaison des répresseurs transcriptionnels, tels que les désacétylases et les méthyltransférases des histones, et sa répression épigénétique (Belloni et al., 2009; Rivière et al., 2015). HBx joue aussi un rôle dans la régulation de l'expression des gènes de l'hôte soit en se liant directement à des régions promotrices des gènes cellulaires soit en activant plusieurs voies de signalisation cellulaires et facteurs de transcription (Levrero and Zucman-Rossi, 2016; Slagle and Bouchard, 2016).

Nous avons utilisé une approche d'immunoprécipitation de la chromatine suivie d'un séquençage à haut débit (ChIP-Seq) pour identifier toutes les cibles génomiques de HBx dans les cellules qui répliquent leHBV. Les cibles directes de HBx sont à la fois des gènes codant des protéines et des ncRNA (75 miRNA et 34 lncRNA).

L'analyse fonctionnelle d'environ 4 000 gènes codant pour les protéines et de 75 miRNA potentiellement régulés au niveau transcriptionnel par HBx a révélé un enrichissement en gènes/ncRNA impliqués dans le métabolisme cellulaire, la dynamique de la chromatine et le cancer, mais également en gènes/ncRNA ? qui modulent la réplication de HBV.

Nous avons identifié de nouveaux mécanismes par lesquels HBx stimule la réplication de HBV, médiée par l'activation transcriptionnelle directe de gènes et d'ARNm qui potentialisent l'endocytose (famille RAB) et l'autophagie (ATGs, beclin-1, miR33a) et la répression transcriptionnelle d'un groupe des miRNA (miR138, miR224, miR576, miR596) qui ciblent directement l'ARNpg de HBV et inhibent la réplication de HBV.

Guerrieri et al. BMC Genomics (2017) 18:184
DOI 10.1186/s12864-017-3561-5

BMC Genomics

RESEARCH ARTICLE Open Access

Genome-wide identification of direct HBx genomic targets

Francesca Guerrieri[1], Laura Belloni[1], Daniel D'Andrea[2], Natalia Pediconi[3], Loredana Le Pera[1,2], Barbara Testoni[4], Cecilia Scisciani[5], Oceane Floriot[4], Fabien Zoulim[4], Anna Tramontano[1,2,6] and Massimo Levrero[1,4,5,7*]

Abstract

Background: The Hepatitis B Virus (HBV) HBx regulatory protein is required for HBV replication and involved in HBV-related carcinogenesis. HBx interacts with chromatin modifying enzymes and transcription factors to modulate histone post-translational modifications and to regulate viral cccDNA transcription and cellular gene expression. Aiming to identify genes and non-coding RNAs (ncRNAs) directly targeted by HBx, we performed a chromatin immunoprecipitation sequencing (ChIP-Seq) to analyse HBV recruitment on host cell chromatin in cells replicating HBV.

Results: ChIP-Seq high throughput sequencing of HBx-bound fragments was used to obtain a high-resolution, unbiased, mapping of HBx binding sites across the genome in HBV replicating cells. Protein-coding genes and ncRNAs involved in cell metabolism, chromatin dynamics and cancer were enriched among HBx targets together with genes/ncRNAs known to modulate HBV replication. The direct transcriptional activation of genes/miRNAs that potentiate endocytosis (Ras-related in brain (RAB) GTPase family) and autophagy (autophagy related (ATG) genes, beclin-1, miR-33a) and the transcriptional repression of microRNAs (miR-138, miR-224, miR-576, miR-596) that directly target the HBV pgRNA and would inhibit HBV replication, contribute to HBx-mediated increase of HBV replication.

Conclusions: Our ChIP-Seq analysis of HBx genome wide chromatin recruitment defined the repertoire of genes and ncRNAs directly targeted by HBx and led to the identification of new mechanisms by which HBx positively regulates cccDNA transcription and HBV replication.

Keywords: Hepatitis B virus, HBx, Epigenetics, miRNAs, ChIP-Seq

Background

Despite the availability of an effective prophylactic vaccine and potent antiviral therapies hepatitis B virus (HBV) is still a major health problem. Over 240 million people chronic hepatitis B virus (HBV) carriers worldwide remain at risk of developing hepatocellular carcinoma (HCC) [1, 2]. The persistence of viral replication in the liver and high serum HBV-DNA levels correlate with disease severity, progression of liver fibrosis and HCC development in the clinical setting [2].

The HBV regulatory protein HBx is both required for HBV replication [3] and implicated in HBV-related oncogenesis [4]. The mechanisms underlying the pleiotropic activities of HBx have been only partially elucidated. HBx regulates transcription both directly, at the chromatin level, and indirectly, by affecting intracellular signaling pathways that modulate the activity of multiple transcription factors [4]. HBx has been shown to target the epigenetic control of cellular genes expression by interacting with chromatin modifying enzymes [5–7]. HBx is also recruited to the cccDNA in HBV-infected cells [3] and is required for the transcription of all viral RNAs from the cccDNA minichromosome in the nucleus [3, 8]. HBx is thought to regulate cccDNA transcription by two main mechanisms: a) the degradation of the Smc5/6 restriction factor mediated by the DDB1 – Cul4 E3 ligase complex [9, 10] and b) the prevention of cccDNA epigenetic silencing by the histone deacetilase HDAC1 [8], the protein arginine N-methyltransferase 1 (PRMT1) [11], the Tudor-domain protein Spindlin-1 [12] and the histone methyl-transferase SETDB1 [13, 14].

* Correspondence: massimo.levrero@uniroma1.it; massimo.levrero@inserm.fr
[1]Center for Life NanoScience@Sapienza, Istituto Italiano di Tecnologia, Viale Regina Elena 291, Rome 00161, Italy
[4]INSERM U1052, Cancer Research Center of Lyon (CRCL), 151 cours Albert Thomas, Lyon 69424, France
Full list of author information is available at the end of the article

Additional HBx activities that boost HBV replication are DNA Methyltransferase 3 Alpha (DNMT3A) downregulation mediated by miR-101 induction [15], the elevation of cytosolic calcium levels [16] and the induction of autophagy [17].

Results

ChIP-Seq analysis of genome wide HBx recruitment

To obtain a high-resolution, unbiased, mapping of HBx binding sites across the genome we sequenced HBx-bound fragments by ChIP-Seq in a cccDNA-driven HBV replication system [18]. Four independent anti-HBx paired ChIPs were carried out on formaldehyde-crosslinked chromatin from mock and HBV-transfected HepG2 cells. On average, ~62% of the reads could be aligned to the hg19 version of the human genome (Additional file 1: Figure S1a); ~1% of the total number of reads aligned to the HBV genome (ayw strain, NCBI Reference Sequence: NC_003977.2) and ~1% of reads falls on the mitochondrial genome. All reads with multiple alignments and mismatches with the reference genomes were excluded (~6.3% with at least one unknown base, ~11% with multiple alignment and ~13% with 1 or 2 mismatches). ~4000 statistically significant peaks/run (*P*-value < 10^{-4}) were detected. About 12% of the HBx peaks were located within 10kb from the transcription start site (TSS) of known genes, 44.6% were located in intergenic regions and 39.5% within introns (Additional file 1: Figure S1b-S1c).

Functional analysis of the genes potentially regulated by HBx

We first performed a functional enrichment analysis of the whole genome repertoire of the genes potentially regulated by HBx in HBV replicating HepG2 cells. Figure 1a shows the categories enriched, respectively, among the *KEGG* and *Reactome* Biological Pathways; the *Gene Ontology* (GO) Biological Processes and the *InterPro families*, a database that integrates diverse information about protein

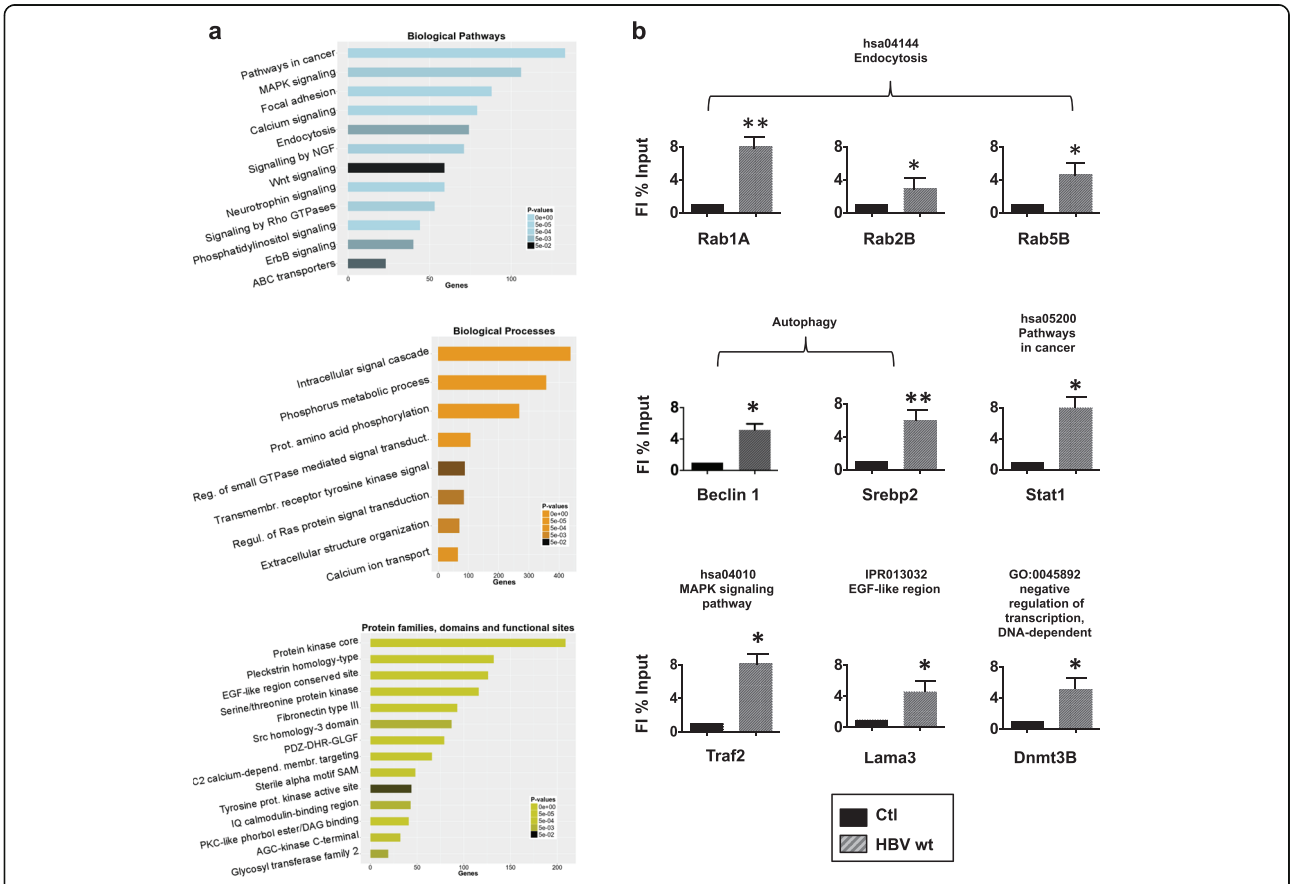


Fig. 1 Functional analysis of protein-coding genes potentially regulated by HBx. **a** Biological Pathways (KEGG, REACTOME), Biological Processes (GO) and Protein families, domains and important sites (InterPro) over-represented (Benjamini-corrected *p*-value < 5×10^{-2}) from the analysis of ~5500 protein-coding genes with HBx-binding sites are shown. In each plot, the categories are in descending order with respect to the associated gene number, with the corresponding color scale of p-values shown on the *right*. **b** Validation of HBx occupancy on selected gene promoters in independent ChIP experiments. ChIP results are expressed as Fold Induction (FI) of the % of Input. Histograms show the mean from three independent experiments; *bars* indicate SD. P-values: * 0,02 ≤ *P* < 0,05; ** 0,01 ≤ *P* < 0.02; *** 0,005 ≤ *P* < 0.01. Gene specific promoter primer pairs are detailed in Additional file 1: Table S6

families, domains and functional sites. Overall, the results show a significant enrichment in genes involved in cell metabolism, chromatin dynamics and cancer, but also in biological pathways that have been associated with the control of HBV replication [i.e. Ras/Src [19], calcium transport, endocytosis]. It is important to underline that, since the HBx ChIP-Seq dataset was generated in HBV replicating HepG2 cells, which are derived from a primary liver tumor, and that an enrichment in genes belonging to the metabolism, chromatin dynamics and cancer pathways is often found in many immortalized and transformed cell lines, a number of the HBx genomic binding sites we found might reflect a bias for transcriptionally active chromatin regions in the "tumor cell" chromatin environment.

Independent anti-HBx ChIP experiments confirmed HBx recruitment, with a 3 to 8-fold enrichment relative to the control IgG ChIPed samples, to the promoter regions of selected genes belonging to different functional categories [Hsa04144: Rab1A, Rab2B, Rab5B; GO:0006357: SREBP2, Beclin1; Has04010: Traf2; Hsa05200: Stat1; IPR013032: Lama3; GO:0045892: Dnmt3B] in HBV replicating HepG2 cells (Fig. 1b) and in HBV-infected NTCP-HepG2 cells (Additional file 1: Figure S2a) and primary human hepatocytes from 2 different donors (Additional file 1: Figure S2b).

Validated PCR primers for genomic target sites that were negative for anti-HBx ChIP-Seq peak assignment showed no enrichment in the HBx immunoprecipitated chromatin, thus confirming the lack of off-target HBx recruitment and the specificity of the anti-HBx ChIP (Additional file 1: Figure S3).

HBx directly enhances endocytosis

The enrichment of genes involved in endocytosis among HBx targets (Fig. 1a) represents a potential link between its role in viral replication [20, 21] and its contribution to liver cancer development [22]. Our ChIP-Seq analysis showed that HBx is recruited and potentially regulates 24 members of the Ras-related in brain (RAB) GTPase family of genes, including several RAB GTPases implicated in endocytosis (Additional file 1: Table S1). HBx binding to the RAB1A, RAB2B and RAB5B promoter regions (see Fig. 1b) was accompanied by an increase in promoter-bound histone H4 acetylation (Fig. 2a) and RAB1A, RAB2B and RAB5B expression (Fig. 2b) in HBV-wt replicating cells but not in HBx-mt replicating cells. The up-regulation of RAB1A, RAB2B and RAB5B expression was confirmed in HBV-infected NTCP-HepG2 cells (Fig. 2c) and human primary hepatocytes (PHH) (Fig. 2d). The specificity of HBx binding to the RAB1A, RAB2B and RAB5B promoter regions was further confirmed in NTCP-HepG2 cells infected with HBV-wt and HBx-mt virus (Additional file 1: Figure S4). Notably, HBV-wt HepG2 replicating cells showed a robust uptake of transferrin, with intense perinuclear accumulation, whereas transferrin distribution was more diffuse in mock HepG2 cells and in HBx-mt HepG2 replicating cells due to a diminished rate of transferrin internalization (Fig. 3).

HBx is also recruited to and potentially regulates several genes involved in autophagy, including Beclin-1, SREBP2 and several members of the ATG (AuTophaGy related) gene family (Additional file 1: Table S1). HBV has been show to induce ER stress and cellular autophagy, which is known to enhance viral replication and virions assembly [17]. Beclin-1 is a major effector of autophagy in mammalian cells and has been shown to be upregulated in HBV-related HCCs, HBx-overexpressing cells and 2.2.15 HBV replicating cells [23]. The SREBP-2 transcription factor, known to be a major regulator of cholesterol metabolism, has also been involved in the direct activation of several genes involved in autophagy [24]. HBx recruitment to the Beclin-1 and SREBP2/miR-33a promoter regions was confirmed in independent ChIP experiments performed in HBV-wt replicating cells (Fig. 1b). HBV-wt replicating cells, but not HBx-mt replicating cells, displayed higher levels of histone H4 acetylation on the Beclin-1 and SREBP2 promoters as compared to HBV mock cells (Fig. 2a) and increased expression of Beclin-1 and SREBP2 (Fig. 2b). These results were confirmed in HBV-infected NTCP-HepG2 cells (Fig. 2c) and HBV-infected human primary hepatocytes (PHH) (Fig. 2d). Altogether, these results indicate that HBx targets endocytosis and autophagy gene expression in HBV replicating cells.

ChIP-Seq identifies ncRNAs directly targeted by HBx

Analyses of the distribution of all the anti-HBx ChIP-Seq peaks revealed the occurrence of putative HBx binding sites in 16 lncRNA promoters and 32 lncRNA intragenic regions, in 44 snoRNA, 3 snRNA and 75 miRNA promoter regions (Additional file 1: Figure S1d). 39 out of the 75 HBx targeted miRNAs are classified as intragenic [25] and 15 of them display HBx peaks in the promoter region of their target genes. Independent anti-HBx ChIP experiments in HBV replicating HepG2 cells (Fig. 4a), HBV-infected NTCP-HepG2 cells (Additional file 1: Figure S5a) and PHHs (Additional file 1: Figure S5b) confirmed HBx binding to selected miRNA promoters (>4-fold enrichment relative to the control IgG ChIPed samples in 11 out of 13). Primers designed to overlap an HBx peak detected with a higher *P*-value (*P* = 1E-3) showed no enrichment in anti-HBx ChIPed DNA (Additional file 1: Figure S6).

The 75 miRNAs targeted by HBx include several miRNAs that have been implicated in the regulation of specific liver functions ($n = 9$), in HBV replication ($n = 6$), in hepatocarcinogenesis ($n = 12$), cancer ($n = 30$) and others the function of which is not yet known ($n = 40$) (Additional file 1: Table S2). The analysis of the selected miRNA promoter sequences bound by HBx using the Genomatix MatInspector

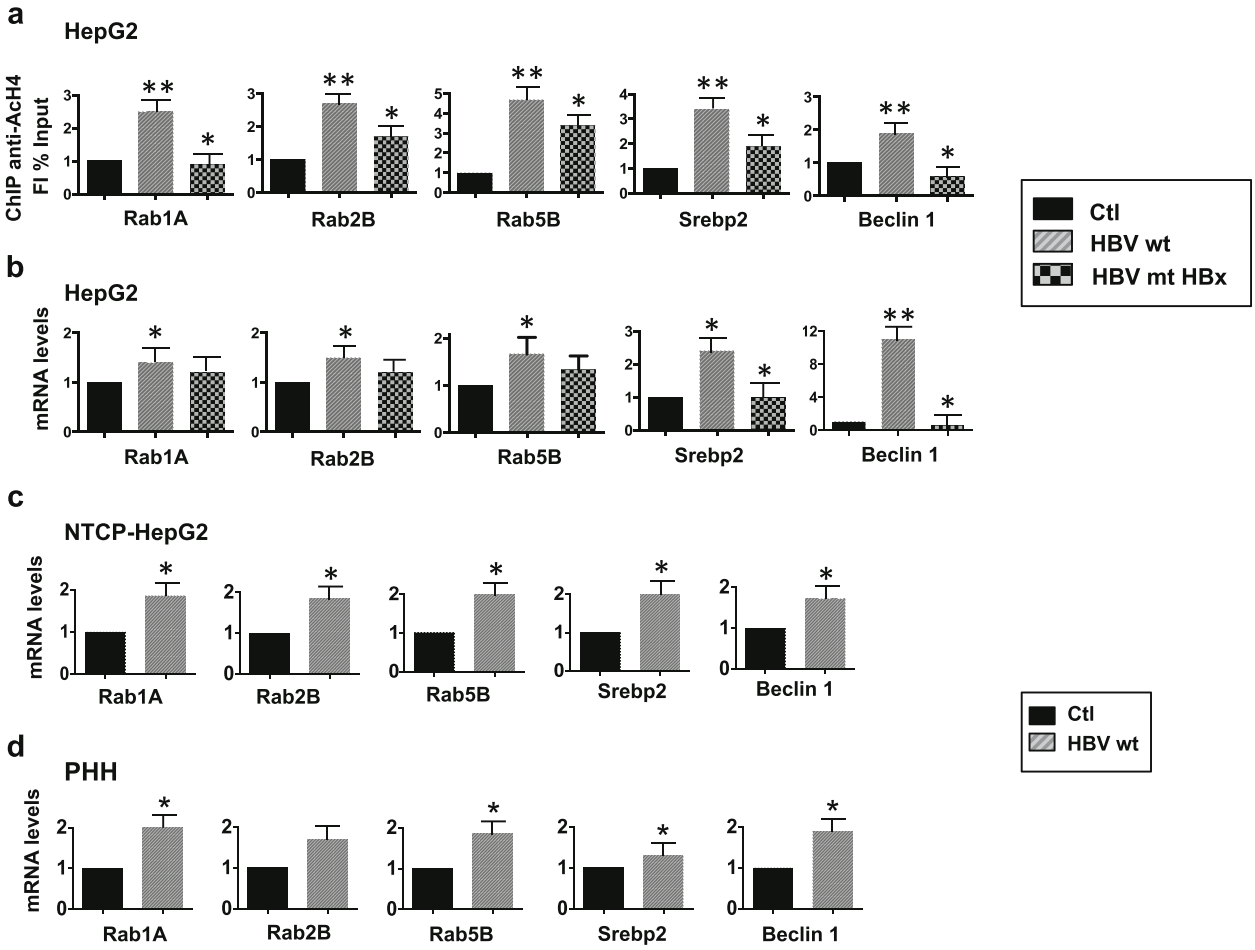


Fig. 2 HBx ChIP-seq identifies direct target genes involved in endocytosis. **a** H4 acetylation on Rab1A, Rab2B, Rab5B, Srebp2 and Beclin 1 genes promoters assessed in independent ChIP experiments. Anti-AcH4 ChIPed chromatin from mock, HBV-wt and HBx-mt HBV replicating HepG2 cells and NTCP-HepG2 HBV-infected cells. Results are expressed as in Fig. 1b. **b**–**d** qPCR quantification of Rab1A, Rab2B Rab5B, Srebp2 and Beclin-1 mRNA levels in mock, HBV-wt and HBx-mt HepG2 cells replicating HBV (**b**) HBV-infected NTCP-HepG2 (**c**) and PHHs (**d**). Results are expressed as fold induction relative to the mock-transfected/infected controls (Ctl) after normalization to endogenous human β-actin mRNAs. Data represent means ± SD from at least three independent experiments performed in duplicate. P-values: * $0{,}02 \leq P < 0{,}05$; ** $0{,}01 \leq P < 0.02$; *** $0{,}005 \leq P < 0.01$

resource showed the presence of binding sites for multiple transcription factors that have been reported to interact with HBx, including NFAT, CREB, NFkB, SREBP1, STAT, E2F and SMAD [4], a finding consistent with a mechanism of piggy-backing for HBx mediated by multiple transcription factors (Additional file 1: Table S3).

Next, we analyzed the impact of HBx chromatin recruitment on the expression of HBx targeted miRNAs. Figure 4b shows the differential expression, analyzed by real-time qPCR, of the 15 miRNAs encoded from the 13 promoters used for the independent validation of HBx chromatin recruitment in Fig. 4a. 8 miRNAs [miR-1913, miR-129-3p miR-129-5p, miR-26b; miR-551b-3p, miR-551b-5p, miR-3648, miR-3687] are consistently upregulated (>1.3 fold) and 5 miRNAs [miR-138, miR-224, miR-302e, miR-576-3p, miR-596] are downregulated (<0.7 fold) in HBV-wt HepG2 replicating cells. 4 miRNAs [miR-21, miR-26a-5p, miR-452, miR-552] are apparently not modulated or only slightly modulated by HBV replication in this assay (Fig. 4b). Notably, concordant results were obtained for 9/13 miRNAs included in the miRNAs Taqman microarrays profiles from HBV-infected (12 dpi) PHHs (Fig. 4c and Additional file 1: Figure S7). miR-21 and miR-452, that were only slightly up-regulated 48 h after transfection in HBV-replicating HepG2 cells, were significantly up-regulated in PHH 12 days after infection. HBV up-regulation of miR-21 expression is consistent with previous reports [26]. miR-552 was dowregulated in HBV-wt replicating HepG2 cells at 48 h withouth reaching <0.7 fold threshold but its levels were significantly reduced in HBV-infected PHH. miR-596 was dowregulated 48 h after transfection in HBV-wt replicating HepG2 cells and significantly up-regulated in HBV-infected PHH.

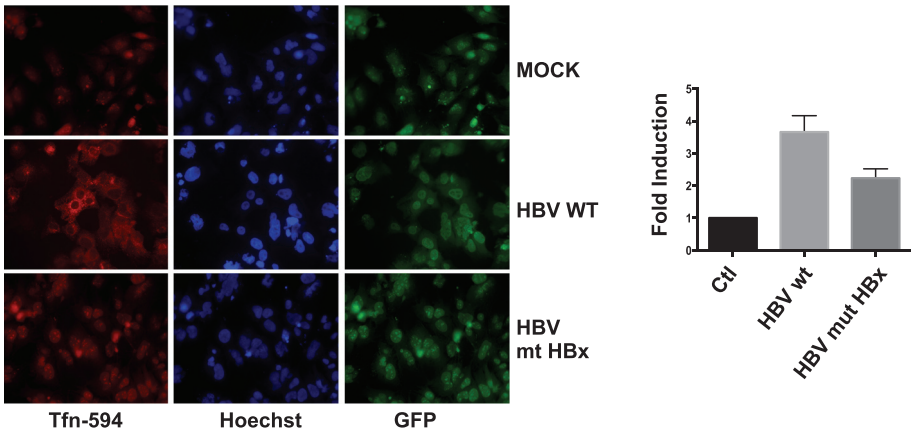


Fig. 3 HBx directly enhances endocytosis. Endocytosis of transferrin is increased in HBV-wt replicating HepG2 cells. Representative images of Alexa Fluor 594-conjugated human transferrin uptake (*red*, *left panels*), nuclear staining with Hoechst (*blue*, *middle panels*) and nuclear GFP positive transfected cells (*green*, *right panels*) in mock, HBV-wt and HBx-mt HBV transduced HepG2 cells. Histograms represent the mean and SD of Alexa 488-conjugated transferrin signal quantified by ImageJ software in 100 cells from each of 3 independent experiments

HBx recruitment impacts on H4 histone acetylation of neighboring chromatin promoters. As shown in Fig. 4d, H4 acetylation was decreased on the promoter regions of miR-138-2, miR-224, miR-302e, miR-576-3p and miR-596, all repressed by HBx. Conversely, HBx recruitment was accompanied by an increased H4 acetylation at the miR-26b promoter (Fig. 4d), whose expression is activated in HBV-wt replicating HepG2 cells and HBV-infected PHH (Fig. 4b and Additional file 1: Figure S5). These results confirm the HBx ability to influence transcription by directly modulating the epigenetic status of target promoters and provide a mechanism for miRNA repression by HBx.

A subset of HBx targeted miRNAs represses HBV replication

A number of miRNAs directly or indirectly promote or repress HBV replication (see ref. [27] and Additional file 1 : Table S4). Thus, miR-122 abrogates p53-mediated inhibition of HBV replication by targeting cyclin G1 and its interaction with p53 [28]. miR-372/373 and miR-501 promote HBV gene expression by targeting the transcription factor nuclear factor I/B [29] and HBXIP [30], respectively. miR-1 increases HBV transcription and replication by targeting HDAC4, which in turn represses HBV transcription [31]. miR-15b potentiates HBV replication by targeting HNF1a and relieving its repressive activity on HBV Enhancer 1 [32]. Conversely, miR-141 suppresses HBV expression and replication in HepG2 cells by targeting PPARa [33] and mir-130a inhibits hepatitis B virus replication by targeting PGC1α and PPARγ [34]. miR-199a-3p and miR-210 [35], miR-15a/miR-16-1 [36], the miR-17-92 cluster [37] and miR-1231 [38] have been shown to target HBV mRNAs directly and to inhibit HBV replication. None of these miRNAs showed HBx recruitment in our ChIP-Seq analysis, neither to their putative promoter sequence nor to the promoter regions of their host genes, with the exception of miR-15a and miR16.1 that are embedded into the DLeu2 lncRNA gene, the promoter of which harbors an HBx binding peak.

We investigated whether the miRNAs repressed by HBx binding to their regulatory regions might directly target HBV transcripts, and in particular the HBV pgRNA. HBx repression could relieve the miRNA-directed down-regulation of HBV replication and unveal new miRNA–dependent auto-regulatory loops in HBV replicating cells. *In silico* analysis revealed the presence of several putative seed sequences on the HBV genome specific for HBx-regulated miRNAs, which are also conserved across HBV genotypes (Additional file 1: Table S5). As shown in Fig. 5a, pre-miR-138, pre-miR-224 and pre-miR-596 overexpression reduces HBV pgRNA levels, whereas pre-miR-302e does not. Similarly, pre-mir-26a2, that is not modulated by HBV in our systems and we use as a control, did not affect HBV pgRNA levels (Fig. 5a). Notably, co-transfection of HBV-wt together with pre-miR-138-2, pre-miR-224, and pre-miR-596 resulted in a significant reduction of HBV replication, measured as cytoplasmic core particles associated rcHBV-DNA (Fig. 5b). Altogether, these results suggest that HBx repression of miR-138, miR-224 and miR-596 expression relieves the negative effects of these miRNAs on HBV replication. On the other hand, mir-302e likely downregulates HBV regulation indirectly, by targeting one or more genes involved in the regulation of HBV replication.

In order to better characterize the interplay between HBx, HBx-targeted miRNAs and HBV replication, we selected miR-224, whose expression is repressed by HBV replication (ref [39] bis and Fig. 4b and c) and increased in HCC patients [39–42] where its expression is driven

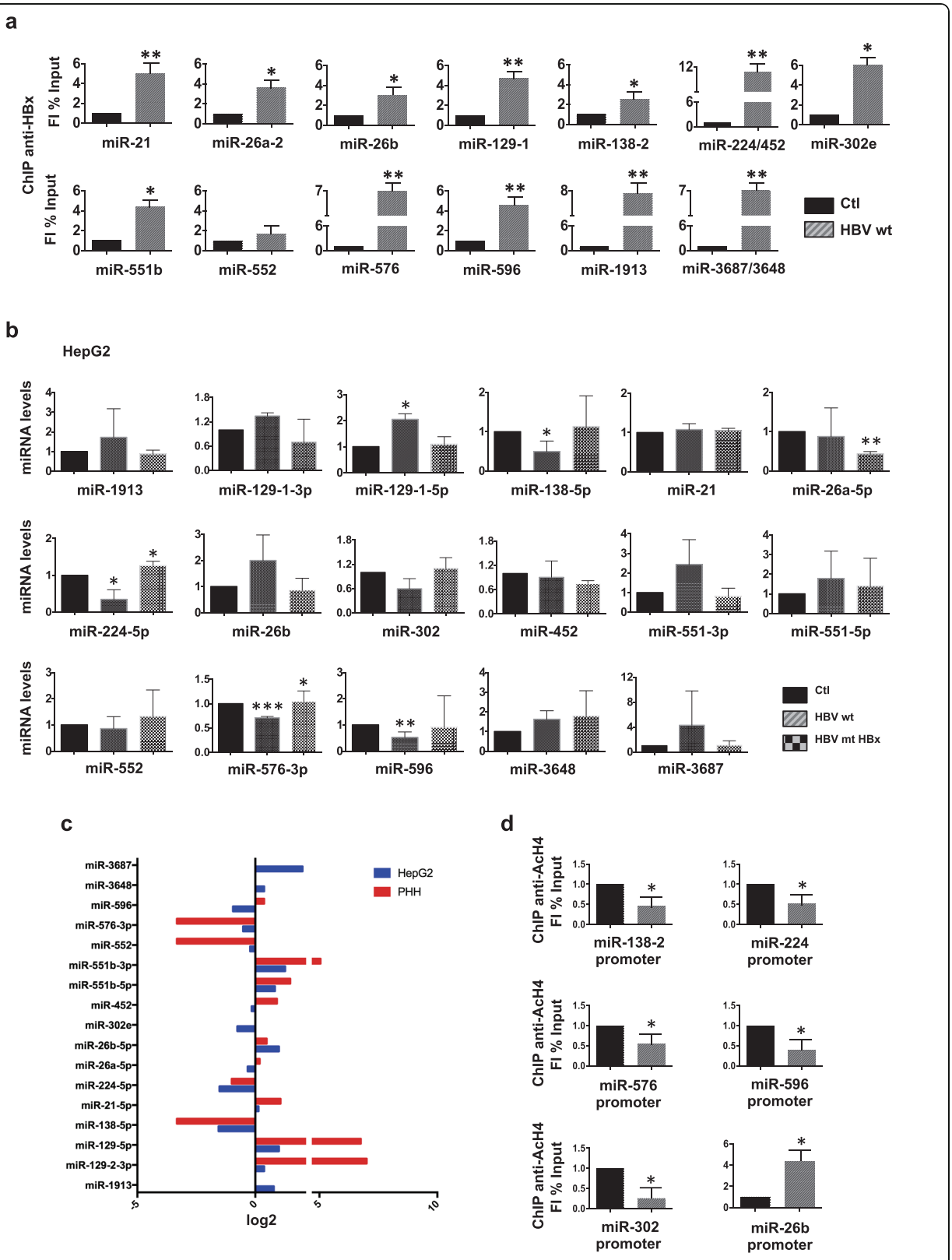


Fig. 4 (See legend on next page.)

(See figure on previous page.)
Fig. 4 ChIP-seq miRNA peak validation and HBx targeted miRNAs expression. **a** HBx occupancy on the putative promoter regions (–5000 + 1000 nt relative to position +1 of each miRNA in MirBase v18) of selected miRNAs (Additional file 1: Figure S1d) was validated in independent ChIP experiments. Cross-linked chromatin from mock or HBV-wt replicating HepG2 cells was immunoprecipitated with a specific anti-HBx antibody or relevant IgG controls, and then analyzed by real-time qPCR using specific primer pairs (see Additional file 1: Table S6). ChIP results are expressed as Fold Induction (FI) of the % of Input and the histograms show the mean from three independent experiments; *bars* indicate SD. **b**–**c** miRNA profiles were analyzed by real-time qPCR, and normalized with respect to RNU38, in mock, HBV-wt and HBx-mt replicating HepG2 cells (**b**) and by Taqman PCR-arrays in HBV infected (12 dpi) PHHs (**c**). Data represent means ± SD from at least three independent experiments performed in duplicate. **d** HBx recruitment impacts on H4 histone acetylation of neighboring chromatin promoters. Anti-AcH4 ChIPs were performed and analysed as in *a*). P-values: * $0{,}02 \leq P < 0{,}05$; ** $0{,}01 \leq P < 0.02$; *** $0{,}005 \leq P < 0.01$

by the TNF/LT-NFkB signaling [42]. Alignment of ChIP-Seq HBx occupancy on the intronic miR-224 regulatory region revealed the presence of a NFkB/p65 consensus motif, which is known to be a direct transcriptional regulator of miR-224 expression and up-regulation in human hepatocellular carcinoma [42]. HBx recruitment is accompanied by a strong co-recruitment of p65 and the DNMT3A DNA methyltransferase (Fig. 5c) in the same region of the miR-224 promoter, that is accompanied by an hypoacetylation of Histone 4 lysines (Fig. 4d). We also show that exogenously expressed HBx repressed miR-224 promoter in a transient luciferase reporter assay (Fig. 5d). Altogether these results strongly suggest the recruitment of HBx as part of a transcriptionally inactive NFkB/p65 complex on the miR-224 promoter, which is responsible for the silencing of miR-224 transcription.

Next, to confirm that the inhibition of HBV replication by miR-224 is indeed the result of a direct targeting of HBV 3.5Kb/pgRNA, the 5 putative seed sequences identified by the RNAhybrid software in the HBV genome (Fig. 6a) were cloned at the 3′UTR of the Renilla Luciferase gene in the pRL-TK vector (Fig. 6b). Co-transfection of the different Renilla Luciferase constructs together with pre-miR-224 showed a significative inhibition of the Luciferase activity for constructs containing seed 1 and seed 4, thus confirming that HBV is a direct target of miR-224 which inhibits viral replication by directly binding HBV genome on seed 1 and seed 4 (Fig. 6b).

Discussion

Our ChIP-Seq genome wide analysis of HBx chromatin recruitment in HBV replicating cells provides a repertoire of genes and ncRNAs directly targeted by HBx and led us to propose new mechanisms by which HBx potentiates HBV replication. As already mentioned, a possible limitation of our study is that our HBx ChIP-Seq data set was generated in HBV replicating HepG2 cells, which are transformed cells derived from primary liver tumor, wild type for p53 and mutated for β-catenin. Indeed, some of the "enriched genes" are associated with pathways, such as metabolism, chromatin dynamics and cancer, that are often found to be deregulated in transcriptomic analysis of immortalized and tumor cell lines. It cannot be excluded that a number of HBx genomic binding sites identified in our study might reflect a bias for transcriptionally active chromatin regions present in HepG2 cells. It is conceivable that the same regions might not be transcriptionally active, and therefore potentially not be targeted by HBx, in normal hepatocytes (i.e. a human liver recently infected by HBV or primary human hepatocytes infected in vitro by HBV). However, strong clinical evidence links HBV replication with the development of HBV-related HCCs and HBx expression has been widely implicated in HCC development and progression. Thus, a repertoire of HBx genomic binding sites identified in a tumor derived hepatocytic cell line, such as the HepG2 cells, might be very relevant for HBV-induced hepatocarcinogenesis. Most HBx targets validated in our study have been confirmed in primary hepatocytes from different donors and, notably, HBx is recruited very early post-infection on the regulatory sequence of miR-21, a well established *onco-miR*, also in HBV-infected primary human hepatocytes.

We found that HBx activates several genes and miRNAs that potentiate endocytosis and autophagy to favor HBV replication, and represses miRNAs (miR-224, miR-138 and miR-596) that potentially target the HBV pgRNA and would inhibit HBV replication (Fig. 7). Thus, HBx binding to the cccDNA increases pgRNA transcription and HBV replication and, at the same time, HBx protects the pgRNA from the negative effects of miRNAs miR-224, miR-138 and miR-596 by targeting their promoters and inhibiting their expression. To better characterize the interplay between HBx, HBx targeted miRNAs and HBV replication we focused on miR-224. The down-regulation of miR-224 expression in HBV replicating cells (ref [35] and Fig. 4b and c) and HBV infected primary hepatocytes (Additional file 1: Figure S5) is in apparent contradiction with the reported up-regulation of miR-224 in elevated in HCC [40–42] including HBV-related HCCs [42]. Notably, Amaddeo et al. have reported that HBx inactivating mutations, including missense/stop and frameshift mutations as well as deletions, are selected in HBV-related HCCs that carry a lower number of viral copies as compared to

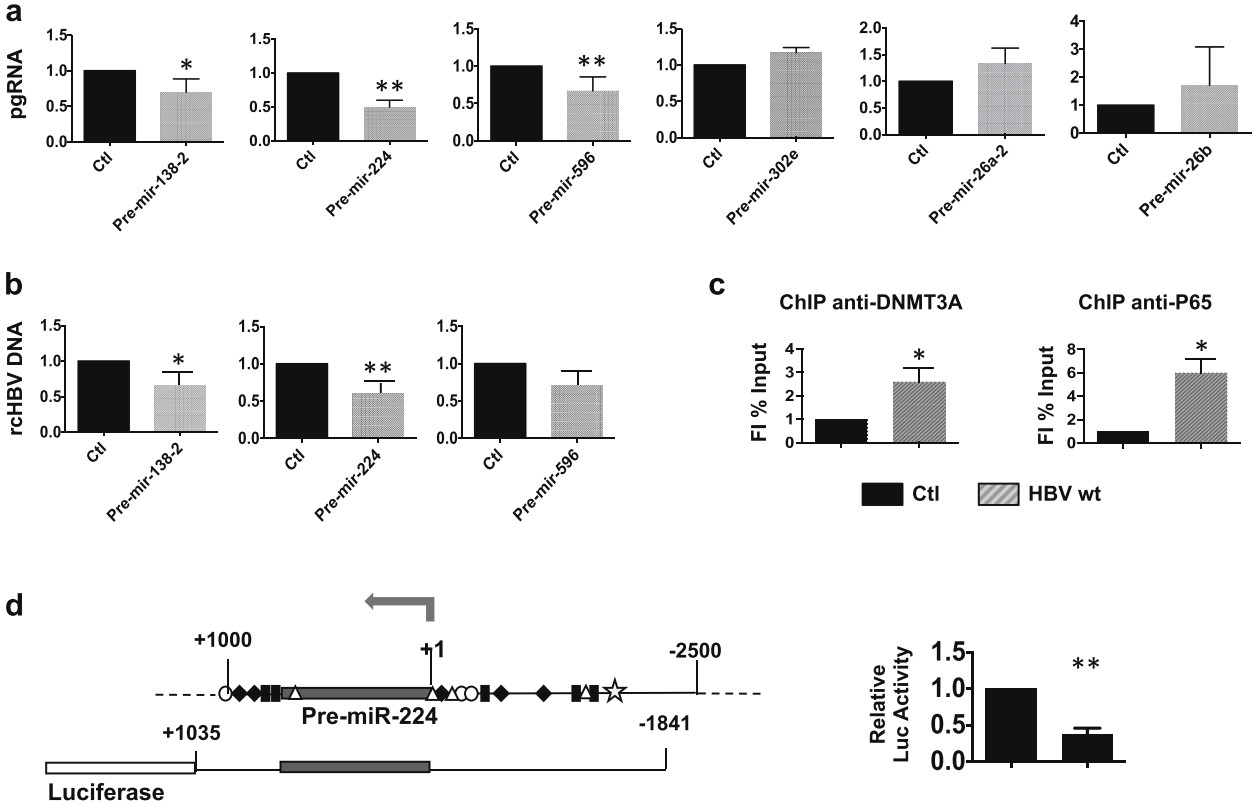


Fig. 5 Impact of HBx-targeted miRNAs on 3.5Kb/pgRNA transcription and HBV replication. **a** HepG2 cells are co-transfected with linear wild type HBV genomes (HBV) and the indicated pre-hsa-miRNAs. HBV 3.5Kb/pgRNA levels were determined by real-time qPCR using specific primers. Results were expressed as fold induction relative to mock controls (Ctl) after normalization towards endogenous human β-actin mRNAs. **b** rcHBV-DNA levels associated to cytoplasmatic core particles were determined by real-time qPCR after normalization to β-globin. Histograms represent the mean from 3 independent experiments; *bars* indicate SDs. P values were determined using the Student's *t* test. *$P < 0.05$. **c** DNMT3A and p65 recruitment on miR224 promoter. ChIPed chromatin from mock and HBV-wt replicating HepG2 cells. Results are expressed as in Fig. 4a. **d** *Left panel*. Putative regulatory region (−2500/+1000 from TSS) and schematic representation of the luciferase construct containing the miR-224 promoter region. The sequence was analyzed by motif software (http://www.genome.jp/tools/motif/) using a cut off score of 85%. *Black diamonds* represent p53 binding sites; white circles represent NFkB consensus; white triangles represent AP1 consensus; *black rectangles* represent TCF/LEF consensus and grey star represents HNF-3β consensus. *Right panel*. HepG2 cells were transfected with 250 ng of the −1841/+1035 miR-224 luciferase construct together with 250 ng of HBx expression vector. Luciferase activity was assayed 30h after transfection and expressed as fold induction over the control. Histograms represent the mean of 3 independent experiments each performed in duplicate; *bars* indicate S.D. P-values: * $0{,}02 \leq P < 0{,}05$; ** $0{,}01 \leq P < 0.02$; *** $0{,}005 \leq P < 0.01$

the non-tumor tissues [43]. Although it would be challenging to obtain a formal demonstration in vivo, it is tempting to speculate that, in the early phases of HBV infection, HBx binding at or near the p65/NFkB sites in the miR-224 promoter leads to the repression of miR-224 expression, that would be detrimental for viral replication, whereas HBx inactivation leaves the miR-224 promoter free to respond to TNF/LT-NFkB signaling in HCCs and highly displastic nodules.

Conclusions

The analysis of HBx genome wide chromatin recruitment provides a repertoire of genes and ncRNAs directly targeted by HBx and led to the identification of new mechanisms by which HBx positively regulates cccDNA transcription and HBV replication.

Methods

Cell cultures

HepG2 (ATCC HB-8065), HepG2.2.15 (a HepG2 clone transfected with a plasmid containing two head-to-tail dimers of the HBV genome, kindly provided by H. Will, MPI, Martinsried, Germany) and NTCP-HepG2 (a HepG2 clone transfected with a plasmid expressing the HBV receptor Na^+-taurocholate cotransporting polypeptide (NTCP), kindly provided by S Urban, Heidelberg University, Germany) cells were cultured in supplemented Dulbecco's modified Eagle's medium (DMEM)

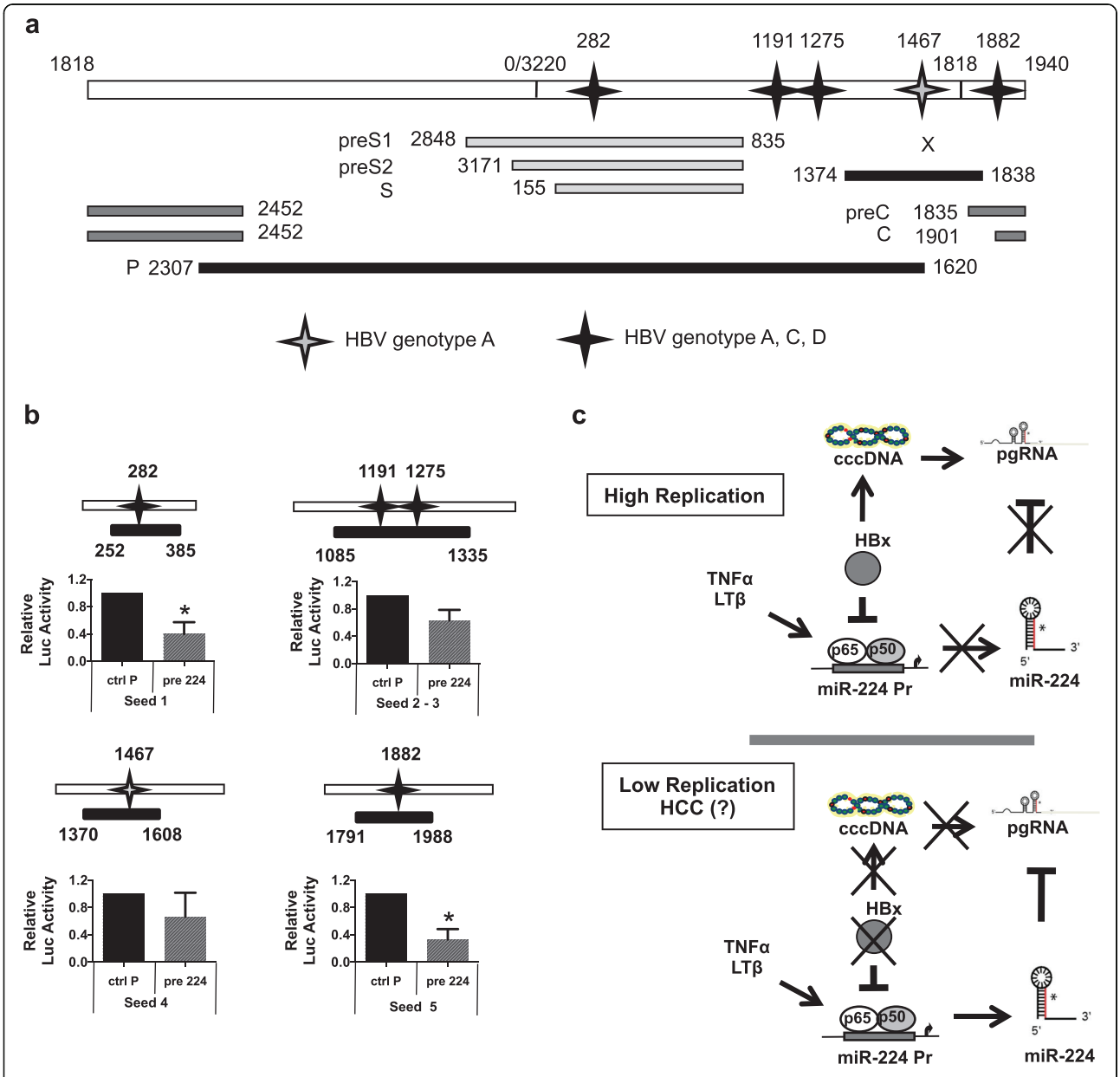


Fig. 6 Direct targeting of HBV 3.5Kb/pgRNA by miR-224. **a** In silico analysis of the HBV genome, performed with the RNAhybrid tool on the Bielefeld University Bioinformatics Server, revealed the presence of several putative miR-224 seed sequences that are conserved across HBV genotypes. **b** *Upper panels*. Schematic representation of the putative seed sequences of HBV genome cloned on the 3′ UTR of the Renilla luciferase gene in the pRL-TK vector (Promega). *Lower panels*. HepG2 cells were transfected with the above indicated luciferase constructs together with 30nM of pre-miR-224 or its relative control. Luciferase activity was assayed 30h after transfection and expressed as fold induction over the control. All histograms represent the mean of three indipendent experiments each performed in duplicate; *bars* indicate S.D. P-values: * $0{,}02 \leq P < 0{,}05$; ** $0{,}01 \leq P < 0.02$; *** $0{,}005 \leq P < 0.01$. Grey stars, HBV genotype A; *black stars*, genotype A, C, D. **c** Proposed model for a differential modulation of miR-224 expression by HBx and inflammatory cytokines in ealrly phases of HBV infection and in HCC

as described [44] and maintained in a 5% CO2 humidified incubator at 37 C.

Transient transfection of full-length HBV DNA genomes

The pCR.HBV.A.EcoRI and the pCR.HBXmt.A.EcoRI plasmids were restricted by EcoRI-PvuI (New England Biolabs) to generate monomeric linear full-length WT (genotype A, adw) and HBx-mt (carrying a termination site at the codon 8 of the X ORF) HBV genomes [44]. Linear HBV monomers were transfected into HepG2 human hepatoma cells using the Mirrus Bio transIT-LT1 (Mir 2300A). Briefly, 2–3 × 10^6 or 1–2 × 10^7 HepG2 cells are seeded in 100-mm or 150-mm-diameter Petri dishes, respectively, transfected after 24 h with 1 to 2 μg of digested

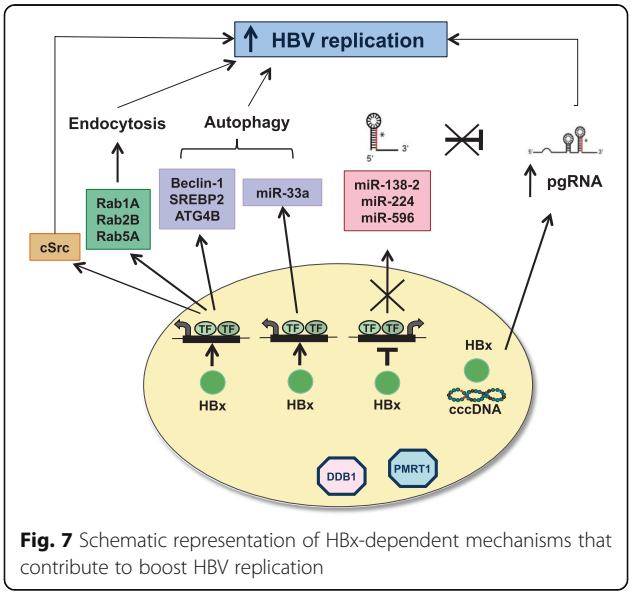


Fig. 7 Schematic representation of HBx-dependent mechanisms that contribute to boost HBV replication

HBV DNA and harvested 48 h post-transfection. A GFP expression vector (500 ng) was included to assess transfection efficiency (range 28%–32%). The HBV region spanning the predicted ends of the linear dsDNA was amplified and sequenced to exclude the generation of circular HBV DNA molecules carrying sequence modifications at the recombination site. Southern blot analysis has shown that the HBV cccDNA species in the nuclear extracts from HepG2 transfected cells co-migrate with the cccDNA isolated from an HBV-infected liver used as a positive control and are converted into a linear DNA by XhoI digestion [18].

Primary culture of human hepatocytes and HBV infection

Primary human hepatocytes (PHH) (provided by Prof. M. Rivoire) 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. PHH were prepared and cultured as described elsewhere [45]. Briefly, the cells are plated overnight in collagen-coated dishes (BD Biosciences) at 2×10^5 cells/cm^2 in William's medium (Life Technologies) supplemented with 10% FetalClone II (Thermo Scientific), 1% penicillin/streptomicine and 1% glutamine (Life Technologies), as well as 5 μg/ml Insulin and 5×10^{-7} M hydrocortisone hemisuccinate (Sigma Aldrich). After 24 h PHH are extensively washed in serum-free medium and kept in serum-free medium for one more day to counter select the growth of contaminating fibroblast and endothelial cells, and infected 48 hours after plating with HBV virus (i.e. inoculum) produced in HepG2.2.15 cells [46].

ChIP assays

Cells are resuspended in 0,2–0,4 ml of ChIP lysis buffer (50 mM Tris HCL, pH 8, 0.5% NP40, 1 mM EDTA, and 100 mM NaCL), incubated 10 min at 4°C and the lysate centrifuged at 10,000 g for 2 min. The nuclei are fixed in 1% formaldehyde for 30 min at 4°C, extracted with a 20 mM Tris, pH 8, 3 mM MgCl2, 20 mM KCl buffer containing protease inhibitors, pelleted by microcentrifugation and lysed in SDS lysis buffer (1% sodium dodecyl sulfate, 10 mM EDTA, 50 mM Tris-chloride, pH 8.1) containing protease inhibitors. Chromatin is sonicated using a Bioruptor Sonicator (Diagenode Inc) to generate 300- to 1000-bp DNA fragments (5 pulses of 45 s at 80% power) for ChIP and 200- to 300-bp DNA fragments (total time of 20 min, 30 s ON, 30 s OFF) for ChIP-Seq. After microcentrifugation, the supernatant is diluted 1:10 in dilution buffer (0.01% sodium dodecyl sulfate, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-chloride, pH 8.1, 167 mM NaCl, containing protease inhibitors). 100 μl of antibody-coupled magnetic beads (Invitrogen Dynalbeads) are added to each 1 ml chromatin preparation and incubated on a rotator for 14–16 h at 4°C. The antibodies used are anti-HBx (MA1-081; mouse monoclonal IgG1), anti-AcH4 (06–866, Upstate; rabbit polyclonal IgG recognizing histone H4 acetylated at lysines 6, 9, 13, and 17), anti-DNMT3A (sc-20703; Santa Cruz Biotechnology Inc.), anti-p65 (sc-372; Santa Cruz Biotechnology Inc.). Nonspecific immunoglobulins (Santa Cruz Biotechnology Inc.) are used as negative controls. After the reverse cross-linking, immunoprecipitated chromatin was purified by phenol/chloroform (1:1) extraction and ethanol precipitation. ChIPed chromatin is analyzed by real-time PCR amplification using either primers (NCC1 and CCCAS) and probes (FL and Red) specific for the HBV cccDNA [3] or specific primers for genes and miRNAs promoters and the SYBR Green DNA Master mix (Applied Biosystems, Inc., Foster City, US-CA) (Additional file 1: Table S6). For ChIP-Seq chromatin is prepared from ~2×10^7 cells / experimental point and processed according to the Illumina ChIp-Seq libraries generation protocol (IP-102-1001) and sequenced on Illumina GAII/GAIIx platforms. The eluate was quantified by using a Qubit (Invitrogen) fluorometer. The quality of each anti-HBx ChIP processed for ChIP-Seq was prealably assessed by controlling HBx binding to the HBV cccDNA [8] and to the MT1F gene promoter [47]. DNA fragments recovered from reverse cross-linked chromatin are repaired, ligated to adapters, size selected and PCR-amplified to generate NGS libraries following the Illumina DNA Library Construction Kit.

ChIP-Seq data analysis

Reads (36 bp) were mapped to the UCSC human genome hg19 using Bowtie (version 0.12.9) [48]. Uniquely aligned reads with no mismatches were retained for

subsequent analyses. Peak-calling was performed using model-based analysis for ChIP-seq (MACS, version 1.4.2) [49], following the protocol described in Feng et al. [50]. Peak-calling was performed against a reference input sample from the same HepG2 cell line. Home-made Perl scripts (www.biocomputing.it; "Tools" section; Scripts related to Guerrieri et al.) were used to locate HBx candidate binding sites along the genome, taking into account three categories: promoters (10 kb upstream or 1 kb downstream the transcriptional start site (TSS) of genes), intragenic (from 1 kb after TSS up to the transcript end site (TES)) and intergenic. For microRNAs, in particular, the promoter HBx binding sites were located within the range of (−5000; +1000) bps relative to the TSS. All potential target protein-coding genes were annotated according to UCSC, RefSeq and GenCode; for non-coding genes (snoRNAs, snRNAs and lincRNAs) we referred to the Ensembl Annotation and for microRNAs to the miRBase database v18 [51].

Functional enrichment analyses of candidate HBx target genes

All protein-coding genes with putative HBx binding sites (promoters or intragenic) were analysed by DAVID [52] and FIDEA [53] web servers using KEGG [54] and Reactome pathways [55, 56], the InterPro families [57] and the Gene Ontology (GO) Biological Processes [58]. Only biological terms with Benjamini-corrected p-value < 0.05 were considered enriched.

Plasmids and luciferase assay

The miR-224 promoter region −1841/+1035 (relative to the pre-miR sequence was amplified by PCR from HepG2 genomic DNA with the following primers: luc-miR-224_FOR 5′-AATCCTGTGCACCTCATCCTCTGT-3′; luc-miR-224_REV 5′-GACGAGCGGAG AAGGTTCTT-3′. The amplified region was cloned into the pGL3 Basic Vector (Promega, Madison, WI, USA) using the T4 DNA ligase (Promega, Madison, WI, USA) and the construct was confirmed by sequencing. Putative miR-224 seed sequences in the HBV genome (genotipe A; ADW) were amplified by PCR with specific primers (Additional file 1: Table S6), cloned in the 3′ UTR of the Renilla luciferase gene in the pRL-TK vector (Promega) and confirmed by sequencing. For luciferase assay, cells are transfected with the indicated luciferase reporter constructs and expression vectors using the Lipofectamine Plus reagent (Invitrogen Inc., Carlsbad, CA, USA). Luciferase activity in cell lysates is measured using the Luciferase Assay System (Promega, Madison, WI, USA) 30 h after transfection and the results normalized for protein levels (Bradford, Bio-Rad Laboratories, Hercules, CA, USA). Pre-miRNA-224 and anti-miRNA-224 (Ambion #AM17100 and #AM17000) are transfected with the Lipofectamine plus reagent (Invitrogen Inc., Carlsbad, CA, USA) at a final concentration of 30 nmol/l and cells were analyzed 30 h after transfection.

HBV pgRNA and cellular mRNA analysis

Total RNA was extracted using the TRIzol reagent (Invitrogen). RNA samples are treated with RQ1 RNase-Free DNase (Promega) for 30 min at 37°C and RNA quality and quantity monitored by ethidium bromide staining and UV absorbance. For pgRNA analysis, 2 μg of DNase-treated RNA are reverse transcribed (ThermoScript RT-PCR System, Invitrogen) and quantified by real-time PCR analysis (Light Cycler; Roche Diagnostics) using the following pgRNA-specific primers and probes: forward primer, 5′-GCCTTAGAGTCTCCTGAGCA-3′, reverse primer, 5′-GAGGGAGTTCTTCTTCTAGG-3′, FRET hybridization probes, 5′–AGTGTGGATTCGCACTCCTCCAGC-FL-3′, and Red640-5′ATAGACCACCAAATGCCCCTATCTTAT CAAC-3′. RNA samples are normalized using the hGAPDH housekeeping gene Light Cycler Set (Roche Diagnostics). miRNAs and mRNAs levels in cultured cells were assayed, after reverse transcription (ThermoScript, Invitrogen, Inc., Carlsbad, US-CA) using miRNA specific stem-loop primers or random primers, by quantitative qPCR using either a specific TaqMan FAM-probe (Applied Biosystems, Inc.) with the specific primers. Results were normalized with respetct to RNU38 or β-actin and expressed as described above.

Taqman® array microRNA cards

Total RNA is extracted from PHH at different time points after HBv infection with TRIzol® reagent (Life Technologies) according to manufacturer's instructions. The Megaplex™ RT primer human pool A + B v3.0 (Life Technologies) was used for miRNA-specific reverse transcription. RNA quality and quantity were monitored by ethidium-bromide staining and by UV absorbance. The Megaplex™ RT product was then loaded on Taqman® microRNA cards A + B and run performed with default thermal cycling conditions (Taqman® array user bulletin – Life Technologies). Data were collected through SDS 2.3 software and analyzed through RQ Manager 1.2 program (Life Technologies).

Core-particles associated HBV DNA purification and quantitation

Cells are washed with ice-cold PBS, lysed in 50 mmol Tris–HCl, pH 7.4, 1 mmol EDTA, and 1% NP-40 (lysis buffer A) and nuclei pelleted by centrifugation for 1 min at 10,000 g. The supernatants are adjusted to 100 mmol MgCl2, treated with 100 mg/ml of DNase I for 30 min at 37°C and then with 0.5 mg/ml proteinase K and 1% SDS for 2 h at 50°C. Nucleic acids are purified by phenol- chloroform (1:1) extraction and ethanol precipitation adding glycogen. HBV DNA is quantified by real-time PCR in a Light Cycler

instrument (Roche) using the following primers and probes: forward, 5′-CTCGTGGTGGACTTCTCTC-3′, and reverse 5′-CAGCAGGATGAAGAGGAA-3′. We also used specific FRET hybridization probes: 5′-CACTCACCAACCTC CTGTCCTCCAA-FL-3′, Red640, 5′-TGTCCTGGTTAT CGCTGGATGT GTCT-3′. Amplifications were performed as follows: 95°C for 5 min, followed by 45 cycles at 95°C for 10 s, 58°C for 10 s, and 72°C for 20 s.

Immunofluoresence and transferrin-uptake assay

HepG2 cell lines were grown overnight on glass coverslips in 6-well plates. The cells were serum-starved at 37 °C for 1 h in starving medium (DMEM 0%), then incubated at 37 °C for 15–20 min in starving medium containing 25 μg/ml of biotinylated holo-transferrin (Sigma T-3915). Cells were fixed in 4% paraformaldehyde at room temperature for 20 min, washed thrice with PBS, permeabilised, blocked with TBS + 0.2% Saponin + 10% serum anti rabbit for 15 min, and then incubated at 37 °C for 1 h with 1:500 dilution of Streptavidin from Molecular Probes conjugated to the fluorophor Alexa 594 (Invitrogen). After three washes with TBS + 0.02% Saponin + 1% serum, the coverslips were mounted onto glass slides. Images of Transferrin, DAPI and GFP-fluorescence were acquired using a *Zeiss 510 Meta* confocal microscope (all images were obtained with the same exposure setting). Alexa 594-conjugated transferrin signal intensity was quantified for 10 transfected cells in each set, and the mean intensity calculated and normalized against that of the nontransfected cells (Mock) with the software ImageJ.

Statistics

P values were determined using the 2-tailed Student's *t* test. Wilcoxon rank-sum test was used for nonparametric pairwise comparisons. $P < 0.05$ was considered significant.

Additional file

Additional file 1: Figure S1. HBx ChIP-Seq experiments analysis. **Figure S2.** HBx occupancy on cellular promoters. **Figure S3.** HBx occupancy on the control RFP2 promoter. **Figure S4.** HBx occupancy on cellular promoters in NTCP-HepG2 cells infected with wild-type and mt-HBx HBV. **Figure S5.** HBx occupancy on miRNAs promoters in NTCP-HepG2 cells and PHHs infected with HBV-wt virus. **Figure S6.** ChIP validation of high P value HBx peak. **Figure S7.** HBx targeted miRNAs expression in HBV-infected Primary Human Hepatocytes. **Table S1.** Rab family and ATG family. **Table S2.** HBx targeted miRNAs functions (PubMed). **Table S3.** TF binding sites overlapping HBx peaks (Genomatix MatInspector). **Table S4.** miRNAs that activate or repress HBV replication HBV replication. **Table S5.** miRNA seeds predicted in HBV pgRNA sequence and conserved across HBV genotypes. **Table S6.** List of primers used thoughout the manuscript. (PDF 647 kb)

Abbreviations

ATG: Autophagy related genes; ChIP-Seq: Chromatin immunoprecipitation sequencing; DDB1: DNA damage-binding protein 1; DNMT3A: DNA Methyltransferase 3 Alpha; HBV: Hepatitis B Virus; HCC: Hepatocellular carcinoma; lncRNAs: long non-coding RNAs; miRNAs: microRNAs; ncRNAs: non-coding RNAs; NTCP: Na + −taurocholate cotransporting polypeptide; pgRNA: pregenomic RNA; PHH: Human primary hepatocytes; PRMT1: Protein arginine N-methyltransferase; RAB: Ras-related in brain GTPase family of genes; TSS: Transcription start site

Acknowledgements

None

Funding

This work was supported by grants from: the Italian Ministry of University and Research (MIUR-FIRB), the Italian Ministry of Health (Ricerca Finalizzata: RF 2010–2317822), the CARIPLO Foundation, the University of Lyon-St Etienne (PALSE PROGRAM), the Agence National de la Recherche (ANR@TRACTION), the Center for Life NanoSciences of the Italian Institute of Technology (CLNS-IIT) to ML; ANRS to ML and FZ; KAUST [KUKI1-012-43] and Epigenomics Flagship Project – EPIGEN to AT; DevWeCan French Laboratories of Excellence Network (Labex, Grant #ANR-10-LABX-61) to FZ; the Gilead Sciences Research Scholars Program in Liver Diseases to LB. FG, LL and LB are recipients of research contracts from IIT.

Availability of data and materials

The ChIP-seq dataset supporting the conclusions of this article is available on the European Nucleotide Archive (ENA) portal and can be found using the following direct URL: http://www.ebi.ac.uk/ena/data/view/PRJEB13227 .

Authors' contributions

FG, ML and AT conceived and designed the project. Experiments were conducted by FG (NGS, functional validations, virology), LB, CS, OF and NP (virology) and BT (virology and microarrays). DD and LL performed bioinformatics analysis. FZ and AT contributed to data interpretation and presentation. FG and ML wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent to publish

Not applicable

Ethics approval and consent to participate

Local and national ethics committees (French Ministry of Research and Education numbers AC-2013-1871 and DC-2013-1870) approval and informed consent from patients were obtained.

Author details

[1]Center for Life NanoScience@Sapienza, Istituto Italiano di Tecnologia, Viale Regina Elena 291, Rome 00161, Italy. [2]Biocomputing Lab, Department of Physics, Sapienza University, Rome, Italy. [3]Department of Molecular Medicine, Sapienza University, Viale Regina Elena 291, Rome 00161, Italy. [4]INSERM U1052, Cancer Research Center of Lyon (CRCL), 151 cours Albert Thomas, Lyon 69424, France. [5]Department of Internal Medicine - DMISM, Sapienza University, Viale del Policlinico 155, 00161 Rome, Italy. [6]Istituto Pasteur Fondazione Cenci Bolognetti, Viale Regina Elena 291, Rome 00161, Italy. [7]Cancer Research Center of Lyon (CRCL) - INSERM U1052, 151 cours Albert Thomas, 69424 Lyon Cedex 03, France.

Received: 7 April 2016 Accepted: 7 February 2017
Published online: 17 February 2017

References

1. Ott JJ, Stevens GA, Groeger J, Wiersma ST. Global epidemiology of hepatitis B virus infection: new estimates of age-specific hbsag seroprevalence and endemicity. Vaccine. 2012;30:2212–9. 2.
2. European Association for the Study of the Liver (EASL) clinical practice guidelines: Management of chronic hepatitis B virus infection. J Hepatol. 2012;57:167–85.
3. 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;55:996–1003.
4. Guerrieri F, Belloni L, Pediconi N, Levrero M. Molecular mechanisms of HBV-associated hepatocarcinogenesis. Semin Liver Dis. 2013;33:147–56.

5. Levrero M, Pollicino T, Petersen J, Belloni L, Raimondo G, Dandri M. Control of cccDNA function in hepatitis B virus infection. J Hepatol. 2009;51:581–92.
6. Cougot D, Wu Y, Cairo S, Caramel J, Caramel J, Renard CA, Lévy L, et al. The hepatitis B virus X protein functionally interacts with CREB-binding protein/p300 in the regulation of CREB-mediated transcription. J Biol Chem. 2007;282:4277–87.
7. Cougot D, Allemand E, Rivière L, Benhenda S, Duroure K, Levillayer F, et al. Inhibition of PP1 phosphatase activity by HBx: a mechanism for the activation of hepatitis B virus transcription. Sci Signal. 2012;5:205–11.
8. Belloni L, Pollicino T, De Nicola F, Guerrieri F, Raffa G, Fanciulli M, 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.
9. Decorsière A, Mueller H, van Breugel PC, Abdul F, Gerossier L, Beran RK, et al. Hepatitis B virus X protein identifies the Smc5/6 complex as a host restriction factor. Nature. 2016;531:386–9.
10. Murphy CM, Xu Y, Li F, Nio K, Reszka-Blanco N, Li X, et al. Hepatitis B virus X protein promotes degradation of SMC5/6 to enhance HBV replication. Cell Rep. 2016;16:2846–54.
11. Benhenda Ducroux A, Rivière L, Sobhian B, Ward MD, Dion S, et al. Methyltransferase PRMT1 is a binding partner of HBx and a negative regulator of hepatitis B virus transcription. J Virol. 2013;87:4360–71.
12. Ducroux A, Benhenda S, Rivière L, Semmes OJ, Benkirane M, Neuveut C. The Tudor domain protein Spindlin1 is involved in intrinsic antiviral defense against incoming hepatitis B Virus and herpes simplex virus type 1. PLoS Pathog. 2014;10:e1004343.
13. Rivière L, Gerossier L, Ducroux A, Dion S, Deng Q, Michel ML, et al. HBx relieves chromatin-mediated transcriptional repression of hepatitis B viral cccDNA involving SETDB1 histone methyltransferase. J Hepatol. 2015;63:1093–102.
14. Alarcon V, Hernández S, Rubio L, Alvarez F, Flores Y, Varas-Godoy M, et al. The enzymes LSD1 and Set1A cooperate with the viral protein HBx to establish an active hepatitis B viral chromatin state. Sci Rep. 2016;6:25901.
15. Wei X, Xiang T, Ren G, Tan C, Liu R, Xu X, et al. miR-101 is down-regulated by the hepatitis B virus x protein and induces aberrant DNA methylation by targeting DNA methyltransferase 3A. Cell Signal. 2013;25:439–46.
16. Yang B, Bouchard MJ. The hepatitis B virus X protein elevates cytosolic calcium signals by modulating mitochondrial calcium uptake. J Virol. 2012;86:313–27.
17. Sir D, Tian Y, Chen WL, Ann DK, Ann DK, Yen TS, Ou JH. The early autophagic pathway is activated by hepatitis B virus and required for viral DNA replication. Proc Natl Acad Sci U S A. 2010;107:4383–8.
18. Pollicino T, Belloni L, Raffa G, Pediconi N, Squadrito G, Raimondo 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.
19. Klein NP, Schneider RJ. Activation of Src family kinases by hepatitis B virus HBx protein and coupled signaling to Ras. Mol Cell Biol. 1997;17:6427–36.
20. Huang HC, Chen CC, Chang WC, Tao MH, Huang C. Entry of hepatitis B virus into immortalized human primary hepatocytes by clathrin-dependent endocytosis. J Virol. 2012;86:9443–53.
21. Macovei A, Petrareanu C, Lazar C, Florian P, Branza-Nichita N. Regulation of hepatitis B virus infection by Rab5, Rab7, and the endolysosomal compartment. J Virol. 2013;87:6415–27.
22. Mellman I, Yarden Y. Endocytosis and cancer. Cold Spring Harb Perspect Biol. 2013;5:a016949.
23. Tang H, Da L, Mao Y, Li Y, Li D, Xu Z, et al. Hepatitis B virus X protein sensitizes cells to starvation-induced autophagy via up-regulation of beclin 1 expression. Hepatology. 2009;49:60–71.
24. Seo YK, Jeon TI, Chong HK, Biesinger J, Xie X, Osborne TF. Genome-wide localization of SREBP-2 in hepatic chromatin predicts a role in autophagy. Cell Metab. 2011;13:367–75.
25. He C, Li Z, Chen P, Huang H, Hurst LD, Chen J. Young intragenic miRNAs are less coexpressed with host genes than old ones: implications of miRNA-host gene coevolution. Nucleic Acids Res. 2012;40:4002–12.
26. Qiu X, Dong S, Qiao F, Lu S, Song Y, Lao Y, et al. HBx-mediated miR-21 upregulation represses tumor-suppressor function of PDCD4 in hepatocellular carcinoma. Oncogene. 2013;32:3296–305.
27. Zhang X, Hou J, Lu M. Regulation of hepatitis B virus replication by epigenetic mechanisms and microRNAs. Front Genet. 2013;4:202.
28. Wang S, Qiu L, Yan X, Jin W, Wang Y, Chen L, et al. Loss of microRNA 122 expression in patients with hepatitis B enhances hepatitis B virus replication through cyclin G(1) -modulated P53 activity. Hepatology. 2012;55:730–41.
29. Guo H, Liu H, Mitchelson K, Rao H, Luo M, Xie L, et al. MicroRNAs-372/373 promote the expression of hepatitis B virus through the targeting of nuclear factor I/B. Hepatology. 2011;54:808–19.
30. Jin J, Tang S, Xia L, Du R, Xie H, Song J, et al. MicroRNA-501 promotes HBV replication by targeting HBXIP. Biochem Biophys Res Commun. 2013;430:1228–33.
31. Zhang X, Zhang E, Ma Z, Pei R, Jiang M, Schlaak JF, et al. Modulation of hepatitis B virus replication and hepatocyte differentiation by MicroRNA-1. Hepatology. 2011;53:1476–85.
32. Dai X, Zhang W, Zhang H, Sun S, Yu H, Guo Y, et al. Modulation of HBV replication by microRNA-15b through targeting hepatocyte nuclear factor 1α. Nucleic Acids Res. 2014;42:6578–90.
33. Hu W, Wang X, Ding X, Li Y, Zhang X, Xie P, et al. MicroRNA-141 represses HBV replication by targeting PPARa. PLoS One. 2012;7:e34165.
34. Huang JY, Chou SF, Lee JW, Chen HL, Chen CM, Tao MH, et al. MicroRNA-130a can inhibit hepatitis B virus replication via targeting PGC1α and PPARγ. RNA. 2015;21:385–400.
35. Zhang GL, Li YX, Zheng SQ, Liu M, Li X, Tang H. Suppression of hepatitis B virus replication by microRNA-199a-3p and microRNA-210. Antiviral Res. 2010;88:169–75.
36. Wang Y, Jiang L, Ji X, Yang B, Zhang Y, Fu XD. Hepatitis B viral RNA directly mediates down-regulation of the tumor suppressor microRNA miR-15a/miR-16-1 in hepatocytes. J Biol Chem. 2013;288:18484–9.
37. Jung YJ, Kim JW, Park SJ, Min BY, Jang ES, Kim NY, et al. c-Myc-mediated overexpression of miR-17-92 suppresses replication of hepatitis B virus in human hepatoma cells. J Med Virol. 2013;85(6):969–78.
38. Kohno T, Tsuge M, Murakami E, Hiraga N, Abe H, Miki D, et al. Human microRNA hsa-miR-1231 suppresses hepatitis B virus replication by targeting core mRNA. J Viral Hepat. 2014;21:e89–97.
39. Zhang ZZ, Liu X, Wang DQ, Teng MK, Niu LW, Huang AL, et al. Hepatitis B virus and hepatocellular carcinoma at the miRNA level. World J Gastroenterol. 2011;17:3353–8.
40. Li W, Xie L, He X, Li J, Tu K, Wei L. Diagnostic and prognostic implications of microRNAs in human hepatocellular carcinoma. Int J Cancer. 2008;123:1616–22.
41. Gao P, Wong CC, Tung EK, Lee JM, Wong CM, Ng IO. Deregulation of microRNA expression occurs early and accumulates in early stages of HBV-associated multistep hepato-carcinogenesis. J Hepatol. 2011;54:1177–84.
42. Scisciani C, Vossio S, Guerrieri F, Schinzari V, De Iaco R, Cervello M, et al. Transcriptional regulation of miR-224 upregulated in human HCCs by NFκB inflammatory pathways. J Hepatol. 2012;56:855–61.
43. Amaddeo G, Cao Q, Ladeiro Y, Imbeaud S, Nault JC, Jaoui D, et al. Integration of tumour and viral genomic characterisations in HBV-related hepatocellular carcinomas. Gut. 2015;64:820–9.
44. Belloni L, Allweiss L, Guerrieri F, Pediconi N, Volz T, Pollicino T, et al. 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. 2012;122:529–37.
45. Lecluyse EL, Alexandre E. Isolation and culture of primary hepatocytes from resected human liver tissue. Methods Mol Biol. 2010;640:57–82.
46. Hantz O, Parent R, Durantel D, Gripon P, Guguen-Guillouzo C, Zoulim F. Persistence of the hepatitis B virus covalently closed circular DNA in HepaRG human hepatocyte-like cells. J Gen Virol. 2009;90:127–35.
47. Zheng DL, Zhang L, Cheng N, Xu X, Deng Q, Teng XM, et al. Epigenetic modification induced by hepatitis B virus X protein via interaction with de novo DNA methyltransferase DNMT3A. J Hepatol. 2009;50:377–87.
48. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.
49. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9:R137.
50. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. Nat Protoc. 2012;7:1728–40.
51. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. Nucleic Acids Res. 2008;36:D154–8.
52. Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4:44–57.
53. D'Andrea D, Grassi L, Mazzapioda M, et al. FIDEA: a server for the functional interpretation of differential expression analysis. Nucleic Acids Res. 2013;41: W84-8.
54. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30.

55. Croft D. Building models using Reactome pathways as templates. Methods Mol Biol. 2013;1021:273–83.
56. Milacic M, Haw R, Rothfels K, Wu G, Croft D, Hermjakob H, et al. Annotating cancer variants and anti-cancer therapeutics in reactome. Cancers (Basel). 2012;4:1180–211.
57. Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, et al. InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res. 2011;40:D306–12.
58. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25:25–9.

Supplementary Files

Genome-wide identification of direct HBx targets that control HBV replication

Francesca Guerrieri [1, 2], Laura Belloni [1, 2], Daniel D'Andrea [3], Natalia Pediconi [4], Loredana Le Pera[1, 3], Barbara Testoni [4], Cecilia Scisciani [2], Oceane Floriot [5], Fabien Zoulim [5], Anna Tramontano [1,3, 6], Massimo Levrero [1, 2, 5]

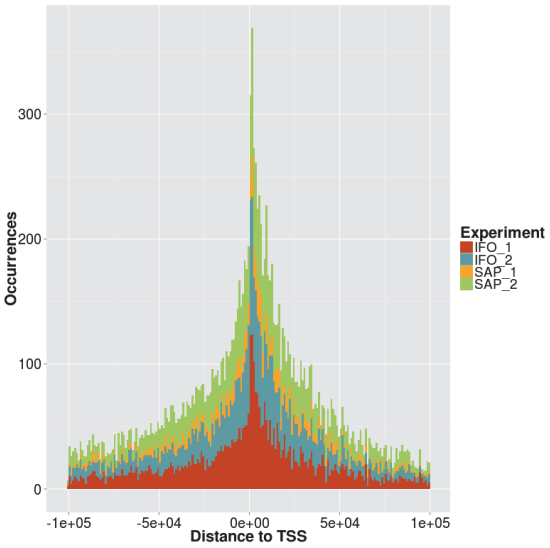
[1] *Center for Life NanoScience@Sapienza, Istituto Italiano di Tecnologia, Rome, Italy*
[2] *Dept of Internal Medicine - DMISM, Sapienza University, Rome Italy*
[3] *Biocomputing Lab, Department of Physics, Sapienza University of Rome, Italy*
[4] *Dept of Molecular Medicine, Sapienza University, Rome, Italy*
[5] *INSERM U1052, Cancer Research Center of Lyon (CRCL), Lyon, 69008, France*
[6] *Istituto Pasteur Fondazione Cenci Bolognetti, Rome, Italy*

Tabel of contents:

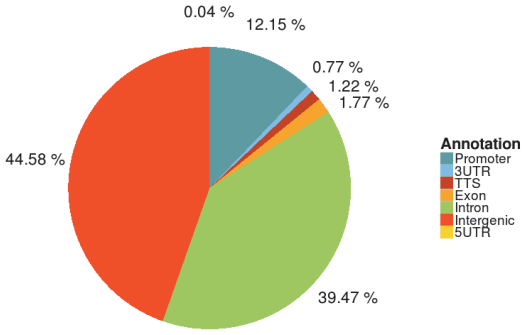
a

Exp	Name	#Good unique reads	#Uniquely mapped reads	%Uniquely mapped reads
SAP	Input	33,167,085	25,687,082	77
SAP	HBx_wt_1	29,109,646	21,261,284	73
SAP	HBx_wt_2	13,090,991	7,994,287	61
IFO	Input	17,516,410	12,157,995	69
IFO	HBx_wt_1	17,148,341	7,996,242	47
IFO	HBx_wt_3	15,629,897	7,141,652	46

b



c



d

Hsa-miR	fdr	Hsa-miR	fdr	Hsa-miR	fdr	Hsa-miR	fdr
Hsa-miR-224	**0**	Hsa-miR-3139	0.24	Hsa-miR-4698	0.76	Hsa-miR-3617	1.75
Hsa-miR-452	**0**	Hsa-miR-548p	0.24	Hsa-miR-640	0.79	Hsa-miR-944	1.80
Hsa-miR-21	**0**	Hsa-miR-302e	0.25	Hsa-miR-4740	0.80	**Hsa-miR-26a-2**	**2.14**
Hsa-miR-3648	**0**	Hsa-miR-663a	0.29	Hsa-miR-4662b	0.80	Hsa-miR-4488	2.15
Hsa-miR-576	**0.14**	Hsa-miR-138-2	0.30	Hsa-miR-5698	0.82	Hsa-miR-3197	2.26
Hsa-miR-1913	**0.14**	**Hsa-miR-551b**	**0.36**	Hsa-miR-4681	0.82	Hsa-miR-4770	2.42
Hsa-miR-26b	**0.15**	Hsa-miR-5095	0.47	Hsa-miR-584	0-84	Hsa-miR-543	2.76
Hsa-miR-5588	0.15	Hsa-miR-4682	0.51	Hsa-miR-4648	0.87	Hsa-miR-495	2.76
Hsa-miR-3691	0.15	Hsa-miR-1539	0.55	Hsa-miR-3650	0.88	Hsa-miR-4664	3.54
Hsa-miR-4703	0.16	Hsa-miR-4429	0.59	Hsa-miR-133a-2	0.89	Hsa-miR-4317	3.73
Hsa-miR-5193	0.16	Hsa-miR-663b	0.59	Hsa-miR-3914-1	0.95	Hsa-miR-4321	3.87
Hsa-miR-943	0.16	Hsa-miR-663a	0.63	Hsa-miR-3170	0.95	**Hsa-miR-3648**	**11.1**
Hsa-miR-129-1	**0.16**	Hsa-miR-4666a	0.67	Hsa-miR-548a-2	1.04	**Hsa-miR-3687**	**11.1**
Hsa-miR-552	**0.17**	**Hsa-miR-3648**	**0.69**	Hsa-miR-3622b	1.18	Hsa-miR-4448	37.5
Hsa-miR-3657	0.17	**Hsa-miR-3687**	**0.69**	Hsa-miR-4286	1.22	Hsa-miR-3916	37.6
Hsa-miR-1973	0.17	Hsa-miR-639	0.69	Hsa-miR-555	1.23	Hsa-miR-4442	37.9
Hsa-miR-4276	0.17	**Hsa-miR-596**	**0.70**	Hsa-miR-4429	1.24	Hsa-miR-4446	38.0
Hsa-miR-3909	0.18	Hsa-miR-4501	0.71	Hsa-miR-4425	1.25	Hsa-miR-4438	43.1
Hsa-miR-4476	0.18	Hsa-miR-378g	0.73	Hsa-miR-4309	1.29		
Hsa-miR-1244-2	0.20	Hsa-miR-3667	0.73	Hsa-miR-639	1.73		
Hsa-miR-626	0.22	Hsa-miR-4710	0.74	Hsa-miR-5006	1.74		

Figure S1. HBx ChIP-Seq experiments analysis. Libraries were sequenced using the Illumina Genome Analyzer II (two IFO libraries) or the Genome Analyzer IIx (two SAP libraries). Total input DNA from each experiment served as control sample. *a)* Number of reads obtained in the different experiments; *b)* Histogram showing the distribution of peaks with respect to the Transcription Start Site (TSS); *c)* Distribution of HBx peaks in the different genomic regions. *d)* list of the 75 HBx targeted miRNA promoters sorted according to their False Discovery Rate (FDR).

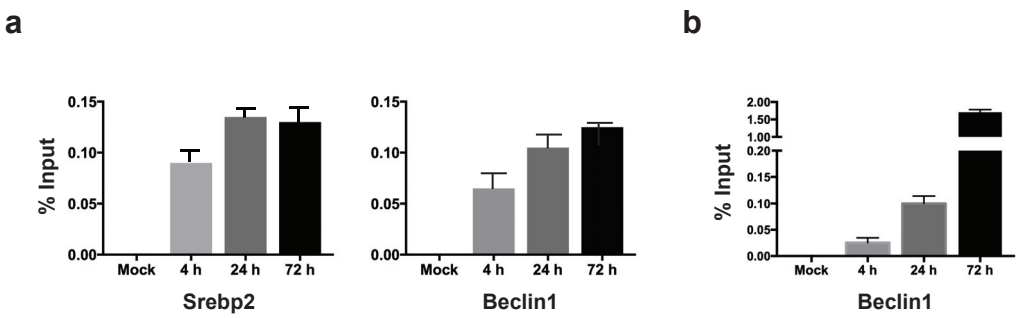


Figure S2. HBx occupancy on cellular promoters in: a) NTCP-HepG2 cells and b) PHHs infected with HBV-wt virus. Anti-HBx ChIPed chromatin from NTCP-HepG2 cells (a) and PHHs (b) infected with HBV-wt virus is analyzed with primers specific for the SREBP2 and Beclin-1 promoters. ChIP results are expressed as % of Input. Histograms show the mean from two independent experiments; bars indicate SD.

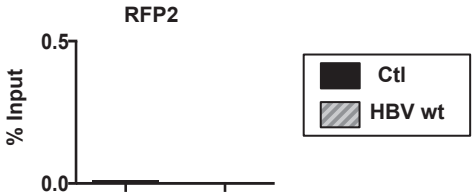


Figure S3. HBx occupancy on the control RFP2 promoter. HBx occupancy on the RFP2 promoter showed no HBx enrichment in independent ChIP experiments. Results are expressed as in Figure S2.

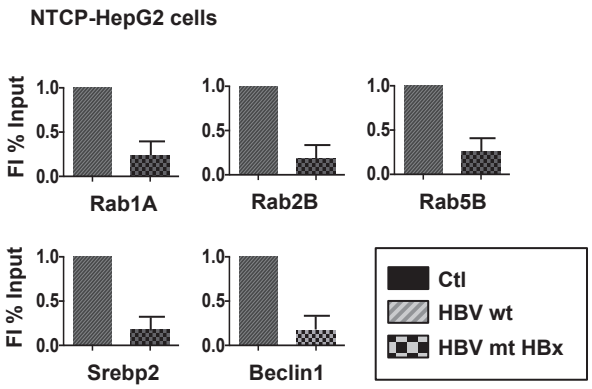


Figure S4. HBx occupancy on cellular promoters in NTCP-HepG2 cells infected with wild-type and mt-HBx HBV. Anti-HBx ChIPed chromatin from NTCP-HepG2 cells infected with HBV-wt and HBV-mt-HBx virus is analyzed with primers specific for the RAB1A, RAB2B, RAB5B, Beclin-1 and SREBP2 promoters. ChIP results are expressed as Fold Induction (FI) of the % of Input. Histograms show the mean from three independent experiments; bars indicate SD.

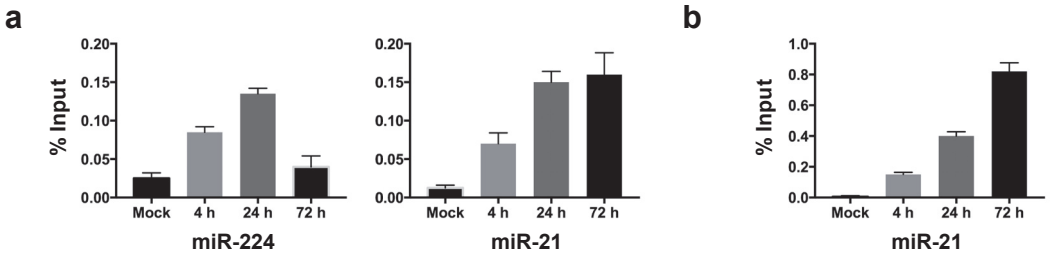


Figure S5. HBx occupancy on miRNAs promoters in: a) NTCP-HepG2 cells and b) PHHs infected with HBV-wt virus. Anti-HBx ChIPed chromatin from NTCP-HepG2 cells (a) and PHHs (b) infected with HBV-wt virus is analyzed with primers specific for the miR-224 and miR-21 promoters. Results are expressed as in Figure S2.

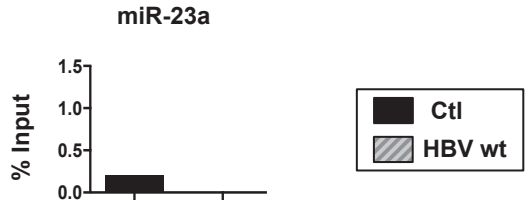


Figure S6. ChIP validation of high P value HBx peak. HBx occupancy on miR-23a (peak P value =1E-3) showed no enrichment in anti-HBx in an independent ChIP validation experiment. Results are expressed as in Figure S2.

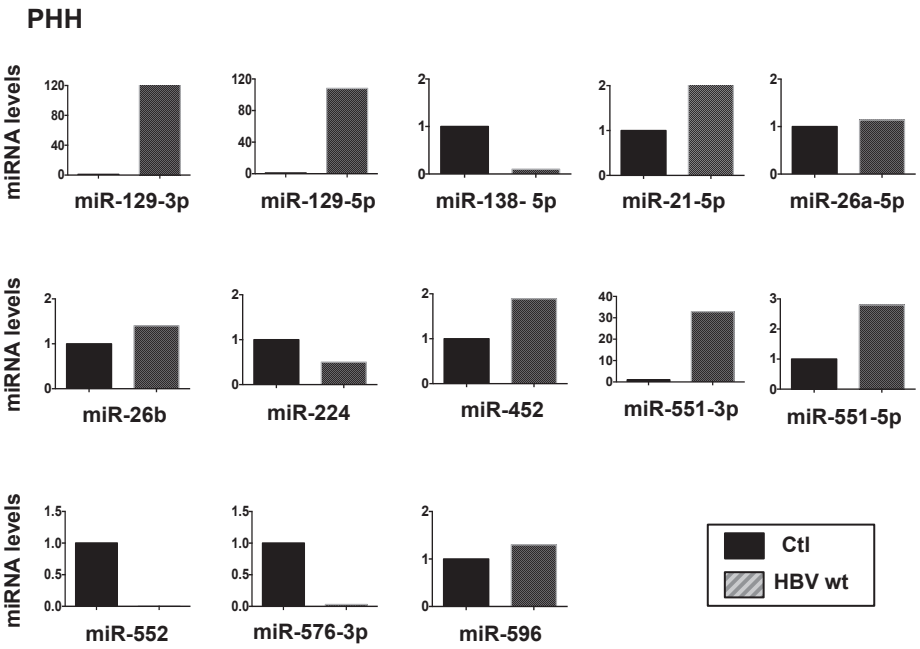


Figure S7. HBx targeted miRNAs expression in HBV-infected Primary Human Hepatocytes. miRNAs expression was analysed by Taqman PCR-arrays in HBV infected (12 dpi) Primary Human hepatocytes (PHH).

Table S1

Rab family & ATG family table

Genes	Location
RAB1A, RAB2B, RAB5A, RAB6C, RAB8A, RAB12, RAB27A, RAB28, RAB31, RAB38, RAB11FIP2, RAB11FIP3	Promoter
RAB3B, RAB3C, RAB3D, RAB4A, RAB5B, RAB8BA, RAB8B, RAB10, RAB18, RAB27A, RAB28, RAB31P, RAB37, RAB32, RAB11FIP3, RAB11FIP5	In gene
ATG4B	Promoter
ATG3, ATG4B, ATG5, ATG7, ATG9A, ATG10, ATG16L1, ATG9B	In gene

Table S2

HBx targeted miRNAs functions (PubMed)

Gene	Liver	HCC	HBV	Cancer	NA	Other
hsa-mir-1244-2						X
hsa-mir-129-1	X	X	X	X		
hsa-mir-133a-2		X		X		
hsa-mir-138-2		X	X	X		
hsa-mir-1539				X		
hsa-mir-1913					X	
hsa-mir-1973				X		
hsa-mir-21	X	X	X	X		
hsa-mir-224	X	X	X	X		
hsa-mir-26a-2	X	X	X	X		
hsa-mir-26b	X	X		X		
hsa-mir-302e					X	
hsa-mir-3139					X	
hsa-mir-3170				X		
hsa-mir-3197						X
hsa-mir-3617						X
hsa-mir-3622b					X	
hsa-mir-3648						
hsa-mir-3648						
hsa-mir-3648				X		
hsa-mir-3650					X	
hsa-mir-3657					X	
hsa-mir-3667					X	
hsa-mir-3687					X	
hsa-mir-3691					X	
hsa-mir-378g				X		
hsa-mir-3909					X	
hsa-mir-3914-1					X	
hsa-mir-3916					X	
hsa-mir-4276					X	
hsa-mir-4286				X		
hsa-mir-4309					X	
hsa-mir-4317				X		
hsa-mir-4321					X	
hsa-mir-4425					X	
hsa-mir-4429						X
hsa-mir-4438					X	
hsa-mir-4442					X	
hsa-mir-4446				X		

Gene	Liver	HCC	HBV	Cancer	NA	Other
hsa-mir-4448	X			X		
hsa-mir-4476					X	
hsa-mir-4488					X	
hsa-mir-4501				X		
hsa-mir-452	X	X		X		
hsa-mir-4648					X	
hsa-mir-4662b				X		
hsa-mir-4664					X	
hsa-mir-4666a					X	
hsa-mir-4681					X	
hsa-mir-4682					X	
hsa-mir-4698					X	
hsa-mir-4703					X	
hsa-mir-4710					X	
hsa-mir-4740					X	
hsa-mir-4770					X	
hsa-mir-495		X		X		
hsa-mir-5006					X	
hsa-mir-5095					X	
hsa-mir-5193					X	
hsa-mir-543	X			X		
hsa-mir-548a-2						X
hsa-mir-548p					X	
hsa-mir-551b		X	X			
hsa-mir-552				X		
hsa-mir-555					X	
hsa-mir-5588					X	
hsa-mir-5698					X	
hsa-mir-576				X		
hsa-mir-584	X			X		
hsa-mir-596		X		X		
hsa-mir-626						X
hsa-mir-639				X		
hsa-mir-640				X		
hsa-mir-663a				X		
hsa-mir-663b				X		
hsa-mir-943		X				
hsa-mir-944				X		

Table S3
TF binding sites overlapping HBx peaks
(Genomatix MatInspector)

	NFAT	CREB	NFkB	p53	STAT	E2F	SMAD
miR-224							
miR-129-1							
miR-138-2							
miR-302e							
miR-596							
miR-26a-2							
miR-26b							
miR-551b							
miR-552							
miR-1913							
miR-576							
miR-3648							
miR-21							

Table S4

miRNAs that activate or repress HBV replication

	HBV replication	Target / mechanisms	References
miR-1	up	HDAC4	23
miR-15a miR-16-1	down	Viral RNAs	28
miR-15b	up	HNF1a	24
miR-17-92 cluster	down	Viral RNAs	29
miR-122	down	Cyclin G1-p53	20
miR-130a	down	PGC1a-PPARg	26
miR-141	down	PPARa	25
miR-372/373	up	nuclear factor I/B	21
miR-199a-3p miR-210	down	Viral RNAs	27
miR-1231	down	Viral RNAs	30
miR-501	up	HBXIP	22

Table S5

miRNA seeds predicted in HBV pgRNA sequence and conserved across HBV genotypes (RNA-hybrid a)

miRNA	Seeed 5p	Seed 3p
21	-22,2 kcal/mol	-21,4 kcal/mol
26a-2	-17,8 kcal/mol	-19,6 kcal/mol
26b	-21,5 kcal/mol	-25,4 kcal/mol
129-1	-24,5 kcal/mol	-27,9 kcal/mol
138-2	-31,2 kcal/mol	-22,3 kmol/cal
224	-27 kcal/mol	-25,8 kcal/mol
302e	-22,6 kcal/mol	/
452	-25,6 kcal/mol	-19,7 kcal/mol
551b	-26,2 kcal/mol	-24,8 kmol/cal
552	-25 kcal/mol	/
576	-22,2 kcal/mol	-23,3 kcal/mol
596	-31,3 kcal/mol	/
1913	-35 kcal/mol	/
3648	-29,8 kcal/mol	/
3687	-24,6 kcal/mol	/

[a] KrugerJ & Rehmsmeier M, NAR 2006, 24, W451-W454

Table S6

List of primers used thoughout the manuscript

ChIP primers	Forward	Reverse
SREBP2	5′- CCCTCTCCAGCCTCAGACAATGA-3′	5′-GGCACATTCACCTGCTTGCTAACT-3 ′
RAB1A	5′- AACTTGCACTACTGCCTCCTGTCA-3	5′-TTAGGCCACCAGGGATTGTTCACT-3 ′
RAB2B	5′- GCCACTGTGTCTGGCCCATAATTT-3′	5′-CCCTTGATCCTTCTTCCTGTTCCT-3 ′
RAB5B	5′- CTTGTTGGCTGGACATGGT-3 ′	5′-GCTATGGATGAGAGAATGGAGG-3 ′
BECLIN1	5- CAGTACATCCTGAGCCAGAAC-3 ′	5-GGGTGCATGAAGAGGTATAGTC -3 ′
TRAF2	5- CCCAGGTTCAAGTGATTTTCCT-3 ′	5′-CTGGGCATGGTGGCAGATGC-3 ′
STAT1	5- GAGGCAGCCATTCGGGAATCTA-3 ′	5′-CTCACCTCGCAACACCTGATCC-3 ′
LAMA3	5- CACCATGCCTGGCTAAGT-3 ′	5′-TGATATCAAAAGCGCAAGCA-3 ′
DNMT3B	5- CCATAAAGTGACCTTCTGG-3 ′	5′-CTGCTGCTCCTCAATACACCAA-3 ′
hsa-mir-21	5′-CCGGCATAGGTCCATCTCTGCA 3′	5′-TCAGCCCGAAACTTCCCAGAG-3′
hsa-mir-26a-2	5′-CACGTGGCCAGCACTGTTAAATGA-3′	5′-GGTTACTACTGATTTCCCTGCCTG -3′
hsa-mir-26b	5′-AAAGAAGTGCACGGGCAACGG-3′	5′-GCGCTTTCCTCTCCCTCTGTGAA -3′
hsa- mir-129-1	5′-TCTGGGCCTCACTTATGCCAC-3′	5′-ATGCTGGGAGAGCTGTGTGACCTT -3′
hsa-mir-138-2	5′- ACCCTGGTACAAGCCAGAGGAAAT-3′	5′-ACAAGCTGAGGCGGGAAGACATAA-3′
hsa-mir-224	5′-CAGAGCAACCCTGGAAACACTCTT -3′	5′-AGCAGAACAAGATGCCTCTCACGA-3′
hsa-mir-302e	5′-GAAGAGACAACACAGAAGCTCAGCAG-3′	5′-CATGTTAGGTCATTGGGAAGTGAGGG-3′
hsa-mir-551B	5′-CAGAAACTGATGGGACTGATCTGC -3′	5′- TGACGAGATATAGTGGCTTGCCTG-3′
hsa-mir-552	5′-GTGATGGACACACACACACACACA-3′	5′-ACAGCTCACCAGTCTTACCCAGAT -3′
hsa-mir-576	5′- TTGGGTCAAGAGTCAGAAGTTT -3′	5′- TGGCTTCTACTTGTCCTTTCC -3′
hsa-mir-596	5′-GGGAGGAAGGAAGGAAAGAAA -3′	5′-TCTGTGAAAGCTGGATGCTAA -3′
hsa-mir-1913	5′-TTGAAAGTGTTCTGCGCTATCC-3′	5′-ACCTGGGCTGAAGGAGATTTA -3′
hsa-mir-3648	5′-GGCCGTGCCTGAGGTTTCT -3′	5′-AACGGACGTGAAGCCGGTGA -3′

mRNA primers	Forward	Reverse	Probe
SREBP2	5′-GCCCTGGAAGTGACAGAGAG -3′	5′-TGCTTTCCCAGGGAGTGA -3′	#21, cat.no. 04686942001
RAB1A	5′-GGGAAAACAATCAAGCTTCAAA -3	5′-CTGGAGGTGATTGTTCGAAAT -3′	#72, cat. no. 04688953001
RAB2B	5′-ACAAATCAAACTGCAAATCTGG -3′	5′-AACGGGTGATAGAACGGAAG -3′	#61, cat. no. 04688597001
RAB5B	5′-GCCAAGACAGCTATGAACGTG -3 ′	5′-GCCAAGACAGCTATGAACGTG -3′	#72, cat.no. 04688953001
BECLIN1	5'-GCAGCTGGATAAGCTGAAGA-3'	5'-CGACCCAGCCTGAAGTTATT-3'	

miRNAs primers			
miRNAs	***Probe***	***miRNAs***	***Probe***
hsa-mir-21	ID 000397	hsa-mir-452	ID 002329
hsa-mir-26a	ID 000405	hsa-mir-551B -3p	ID 001535
hsa-mir-26b	ID 000407	hsa-mir-551B -5p	ID 002346
hsa-mir-33a	ID 002135	hsa-mir-552	ID 001520
hsa- mir-129-5p	ID 000590	hsa-mir-576-3p	ID 002351
hsa- mir-129-3p	ID 002298	hsa-mir-596	ID 001550
hsa-mir-138	ID 002284	hsa-mir-1913	ID 121113
hsa-mir-224	ID 002099	hsa-mir-3648	ID 464401
hsa-mir-302e	ID 002802	hsa-mir-3687	ID 464645
RNU38B	ID 001004		

miRNAs PreMir	
hsa-mir-26a-5p PM10249	hsa-mir-302e PM13467
hsa-mir-138-5p PM11727	hsa-mir-596 PM11608
hsa-mir-224-5p PM12571	

Primers specific for SEED sequences	Forward	Reverse
Seed1	5'-ACCCTGTGACGAACATGGAGAACA-3'	5'-ACACATCCAGCGATAACCAGGACA-3'
Seed2	5'-ACAGGCTTTCACTTTCTCGCCAAC-3'	5'-AGTTCCGATGAGCTTTGCTCCAGA-3'
Seed3	5'-TTTCCATGGCTGCTAGGCTGTACT-3'	5'-AACGTGCAGAGGTGAAGCGAAGT-3'
Seed4	5'-ATAAATTGGTCTGCGCACCAGCAC-3'	5'-ACGGAAGGAAAGAAGTCAGAAGGC-3'

Etude 2

Hepatitis B protein HBx binds the DLEU2 lncRNA to sustain viral and host gene transcription.

Debora Salerno [*], Letizia Chiodo [*], Oceane Floriot [*], Grazia Cottone, Matteo Pallocca, Safaa Jeddari , Massimo Levrero [#], Francesca Guerrieri [#]

Soumis à Nature Microbiology

Plusieurs lncRNA (HULC, HEIH, LET, MVIH, H19, HOTAIR, PCAT-29, EVI1, UCA1) ont été montré être dérégulés dans l'infection par le HBV et dans les HCC associés au HBV (Esposti et al., 2016; Hu et al., 2016; Moyo et al., 2016; Yang et al., 2017; Zhang et al., 2016).

Dans l'étude 1, nous avons identifié 16 promoteurs de lncRNA et 34 régions intragéniques de lncRNA liés spécifiquement par HBx (Guerrieri et al., 2017). Nous avons évalué dans des hépatocytes primaires humains (PHH) infectés par le HBV l'expression des 34 lncRNA uniques qui sont directement ciblés par HBx et de 15 lncRNA supplémentaires modulés dans les HCC humains. Neuf des 34 lncRNA ciblés par HBx sont induits de manière significative (DLEU2, LINC00441, LINC00467, LINC00476, LINC00478, LINC00526, RBM26-AS1, SNHG12, TMEM72-AS1) et 4 sont réprimés (HCG18, MIR31HG, PLK1, PK) 7 jours après l'infection. Par ailleurs 7 sur 15 des lncRNA modulés dans les HCC (FTX, GAS5, H19, HEIH, HULC, MVIH, TUC338) sont fortement induit par le HBV. Parmi les lncRNA présentant les niveaux d'induction les plus élevés dans les PHH infectés par le HBV, DLEU2, LINC00441, LINC00478 sont des cibles directes de HBx, alors que nous n'avons pas de preuve directe du recrutement de HBx sur les promoteurs de HULC, HEIH et H19 (Guerrieri et al., 2017).

Nous nous sommes concentrés sur DLEU2, qui est fortement activé par le HBV, et qui est exprimé lors de la maturation du foie (Peng et al., 2014) et surexprimé dans les HCC humains (Villa et al., 2016; Yang et al., 2017). Le gène DLEU2 s'étend sur 2768 pb et chevauche le cluster miR-15a / miR-16-1, le gène TRIM13 et, partiellement, le gène DLEU1. Le recrutement de HBx s'accompagne d'une augmentation de la liaison de la polymérase II, de l'acétylation des histones H4 et des ARN DLEU2.

DLEU2 est hautement enrichi, avec quelques autres lncRNA, dans les RIP-Seq (immunoprécipitation d'ARN suivie d'un séquençage à haut débit) EZH2 réalisés dans des cellules cancéreuses du côlon (Guil et al., 2012). Nous avons pu confirmer par des expériences de RIP que DLEU2 se lie à l'histone méthyltransférase EZH2, la sous-unité catalytique du complexe répressif PRC2, dans les cellules du foie et à la protéine virale HBx dans les cellules infectées par leHBV. Ces résultats mettent en évidence une nouvelle propriété de HBx de se lier aux ncRNA. Nous avons aussi pu montrer que l'interaction entre DLEU2 et HBx module les fonctions du complexe EZH2/PRC2 conduisant à l'activation constitutive d'un sous-ensemble de gènes cibles d'EZH2 qui sont normalement conservés dans un état réprimé et sont surexprimé dans les HCC.

Nous avons également montré, avec des expériences d'isolement de la chromatine ADNccc par purification d'ARN (ChIRP), que l'interaction de HBx avec DLEU2 se produisait sur le minichromosome de l'ADNccc où elle stimulait la transcription/réplication du virus.

Hepatitis B protein HBx binds the DLEU2 lncRNA to sustain viral and host gene transcription

Debora Salerno [1 #], Letizia Chiodo [2 #], Oceane Floriot [3#], Grazia Cottone [4], Matteo Pallocca [5], Safaa Jeddari [6], Massimo Levrero [1,3,6 *], Francesca Guerrieri [1 *]

[1] Center for Life Nano Science, Istituto Italiano di Tecnologia, Rome, Italy

[2] Unit of Nonlinear Physics and Mathematical Modeling, Department of Engineering, Università Campus Bio-Medico, Roma, Italy

[3] Cancer Research Center of Lyon (CRCL), INSERM U1052, Lyon, 69008, France

[4] Department of Physics and Chemistry, University of Palermo, Palermo, Italy

[5] UOSD SAFU, Regina Elena National Cancer Institute, Rome

[6] Department of Internal Medicine - DMISM, Sapienza University, Rome, Italy

[#] DS and LC are joint first authors

* ML and FG are co-corresponding authors

Abstract

The Hepatitis B virus (HBV) HBx regulatory protein is required for transcription from the cccDNA minichromosome and affects the epigenetic control of both viral and host cellular chromatin. Here we show, by combining computational tools and functional experiments in relevant cellular models of HBV replication, that HBx binds DLEU2, a lncRNA expressed in the liver and activated in human hepatocellular carcinoma (HCC). HBx and the histone metyltransferase EZH2, the catalytic active subunit of the PRC2 complex, share the same preferential binding sites on DLEU2. DLEU2 and HBx co-recruitment on the cccDNA displaces EZH2 from the viral chromatin to boost transcription and viral replication. DLEU2 association to the EZH2/PRC2 target promoters TRIM13 and CCNB2 relieves EZH2 repression. In HBV replicating cells, the concomitant increase in DLEU2 levels and HBx recruitment further increase transcription from a subset of active EZH2/PRC2 target genes. Collectively, our results highlight the ability of HBx to bind RNA to impact on the epigenetic control of host genes and provide a key to understand the role of DLEU2 and EZH2 overexpression in HBV-related HCCs and HBx contribution.

Introduction

Despite availability of a preventive vaccine and anti-viral treatments that arrest disease progression and reduce liver cancer risk, hepatitis B still remains a major public health problem worldwide with about 257 million persons living with HBV chronic infection, that makes it the most common chronic viral infection[1,2]. HBV (hepatitis B, HBV-related cirhhosis and HBV-related liver cancer) accounts for over 850,000 deaths/year[3] and is responsible for >50% of the hepatocellular carcinomas (HCCs) worldwide[3]. Currently, only 10% of chronically infected people have been diagnosed worldwide and as few as 1% are adequately treated[1]. Hepatitis B surface antigen (HBsAg) loss is viewed as an important goal of therapy[4] but it is achieved only in a minority of patients treated with pegylated interferon alpha2a (PegIFN) and nucleos(t)ide analogues (NUCs) and life-long treatments are recommended[5]. HBsAg loss is not accompanied by viral eradication[4] and most patients never clear the HBV cccDNA and bear viral sequences integrated into the host DNA[6]. cccDNA minichromosome is the matrix for transcription of all viral mRNAs and the 3.5Kb pregenomic RNA, that serves as template for viral replication within the capsids in the cytoplasm[7]. The level of HBV replication is determined by the size of the cccDNA pool (ie total number of cccDNA molecules present) and its transcriptional productivity (ie the amount of HBV RNAs transcribed from each cccDNA)[7]. cccDNA transcription is regulated by the epigenetic modifications imposed to cccDNA-bound histones[8] by viral and cellular proteins as well as by inflammatory cytokines, such as IL6, IL1b and IFNα[7]. Low serum HBV-DNA levels and low HBV replication are associated with cccDNA bound histone deacetylation and, conversely, high serum HBV DNA levels are associated with cccDNA hyper-acetylated histones in the liver of HBeAg negative chronic HBV patients[8]. HBx regulates cccDNA transcription by 2 mechanisms: 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 repressors recruitment and cccDNA silencing[10,11]. In addition to viral chromatin, HBx is recruited to and regulates the activity of a variety of coding genes and ncRNAs promoters[12]. Long non coding RNAs (lncRNAs) are frequently dysregulated in both HBV-related HCC tissues and HBV-/HBx-expressing cell lines[13]. The biological effects of lncRNAs are less understood compared to small ncRNAs (microRNAs), but they have been

recently recognized to play a crucial role in the modulation of signaling pathways and regulation of gene expression[14]. A number of lncRNAs have been described to guide chromatin-modifying complexes, including EZH2/PRC2, toward their chromatin target loci[14–17]. In this paper, we show, by combining computational tools and functional experiments in relevant cellular models of HBV replication, that HBx binds the lncRNA DLEU2 to modulate transcription from the HBV cccDNA minichromosome, hence boosting HBV replication and from host genes promoters. HBx recruitment to DLEU2-EZH2/PCR2 regulatory complexes freezes the promoters of a subset of EZH2/PRC2 targets into a constitutionally active state. 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 hepatocytes 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 (i.e. HULC, HEIH, LET, MVIH, H19, HOTAIR, PCAT-29, EVI1, UCA1)[13,18–21]. Recently, using a genome wide analysis of HBx chromatin recruitment in HBV-replicating cells, we identified 16 lncRNA promoters and 34 lncRNA intragenic regions specifically bound by HBx[12]. The expression of the 34 unique lncRNAs that are directly targeted by HBx (Figure 1a) and of 15 additional lncRNAs that are known to be modulated in human HCCs (Figure 1b) has been evaluated in mock and HBV-infected human primary hepatocytes (PHHs) using a custom Nanostring-GE panel assay. As shown in Figure 1a, 9 out of the 34 HBx-targeted lncRNAs were significantly upregulated (DLEU2, LINC00441, LINC00467, LINC00476, LINC00478, LINC00526, RBM26-AS1, SNHG12, TMEM72-AS1) and 4 were strongly down-regulated (HCG18, MIR31HG, PLK1S1, PVT1) by HBV at day 7 post infection. Notably, 7 out of 15 HCC-modulated lncRNA were strongly upregulated (FTX, GAS5, H19, HEIH, HULC, MVIH, TUC338) by HBV infection (Figure 1b). Among the lncRNAs with the highest levels of induction in HBV infected PHHs, DLEU2, LINC00441, LINC00478 were HBx direct targets (i.e positive in the anti-HBx ChIP) whereas for HULC, HEIH and H19 we do not have direct evidence of HBx promoter recruitment[12]. We focused on DLEU2, which is differentially expressed during liver maturation[22] and overexpressed in human HCCs[23,24] (and Figures S1a and S1b).

DLEU2 gene extends for 2768 bp and overlaps with the miR-15a/miR-16-1 cluster, TRIM13 and, partially, with DLEU1 genes (Figure 1c). The *HBx peak* detected by HBx ChIP seq[12] (Figure 1c, black circle) is located ~6600 bp upstream the DLEU2 promoter previously described in Lerner et al.[25]. HBx binding on *HBx peak* region was confirmed in independent ChIP assays performed in HBV-infected PHH (Figure 1d, left panel), HepG2 cells transfected with HBV monomers[10] and HepAD38 induced cell lines (Figure S2a, upper panels). HBx recruitment is accompanied by an increase in Polymerase II binding (Figure 1d, right panel), H4 acetylation (Figure S2a, lower panels) and DLEU2 steady state RNA levels (Figure 1e and Figure S2b). The modulation of DLEU2 expression was further confirmed in all the HBV infection/replication systems (HBV-

infected PHHs and NTCP-HepG2 cells; HBV transfected HepG2 cells and HepAD38 induced cells) using RT-PCR primers specific for exons 2-4 and intron 2 (Figure 1f and Figure S2c). The increase in DLEU2 RNA containing intron 2 sequences in the nuclei of HBV replicating cells (Figure 1g) supports the notion that DLEU2 induction by HBx and HBV occurs at the transcriptional level. Notably, HBx recruitment on the *HBx peak* region has no impact on the levels of the antisense DLEU1 lncRNA (Figure 1c and data not shown).

DLEU2 lncRNA physically interacts with EZH2 and HBx

A number of lncRNAs have been described to guide chromatin-modifying complexes toward their chromatin loci of origin and regulate gene expression[14,15]. Thus, the lncRNA HOTAIR induces a genome-wide reprogramming of PRC2 targets to promote cancer metastasis in breast and colon cancer[26,27]. Several lncRNAs interact *in vivo* with the EZH2/PRC2 complex in embrionic stem cells (ESCs)[16] and DLEU2 is highly enriched, together with the XIST, TSIX, MALAT1 and GAS5 lncRNAs, in EZH2 RIP-Seq performed in colon cancer cells[17]. Since HBx has been reported to bind EZH2/PRC2 in order to modulate transcription of specific host genes[21,28] we sought to further investigate the interaction between DLEU2, EZH2 and HBx in hepatocytes and HCC cell lines. As shown in Figure 2a, we found that DLEU2 and EZH2 interact *in vivo* in PHHs. As compared to mock-infected cells, HBV-infected PHHs displayed a significant enrichment of DLEU2 lncRNA (intron 2 and exons 2-4) in anti-EZH2 RNA immunoprecipitations (Figure 2a, left and middle panels). MALAT1, a lncRNA known to bind EZH2, served as positive control (Figure 2a, right panels). Next, we performed anti-HBx RIPs in HBV-infected PHHs and HepG2 cells transfected with wild type (wt) monomeric HBV genomes. We found that HBx binds DLEU2 (intron 2 and exons 2-4) in HBV infected / replicating hepatocytes (Figure 2b and 2c, left and middle panels). No enrichment of MALAT1 binding to HBx was observed (Figure 2b and 2c, right panels). Altogether these results indicate that the DLEU2 lncRNA binds EZH2 in liver cells and, in HBV infected cells, the viral protein HBx.

Modeling the DLEU2-EZH2-HBx interaction in silico.

The RIP results suggest, but do not formally prove, the presence of a tripartite complex containing DLEU2, EZH2 and HBx in HBV infected / replicating cells. To gain insights on the formation of the complex and the interactions between the three partners we modeled DLEU2-EZH2 and DLEU2-HBx protein complexes and the ternary DLEU2, HBx and EZH2 protein-RNA complex using bioinformatics tools.

The three-dimensional atomistic structure is an essential ingredient to model protein-RNA physical interactions. The EZH2 structure has been established[29], but the structure prediction for partially disordered proteins (i.e. HBx) and lncRNAs (i.e. DLEU2) is a challenging task. Methodological details and issues are discussed at lenght in the Supplementary Materials Section.

We used DLEU2 exons 1-10 and intron 2 sequences to model the corresponding secondary (Figure S3) and tertiary (Figure 3a and Figure S4) structures[30–32]. We also modelled the atomistic structure of HBx obtaining a partially disordered stucture, as expected from predictions based on the sequence analysis, and in agreement with literature[33-34]. Actually, the presence of intrinsically disordered regions, in combination with structured domains, is essential for HBx interaction behaviour.

We first investigated the interaction between DLEU2 derived RNA structures and HBx using modeling tools based both on primary sequences[35,36] and on 3D atomistic models, in order to verify if any interaction exists, and if it is driven by sequence, structure, or peptides/nucleotides length, all previously suggested mechanisms for RNAs-protein interactions[37]. The RNA-binding residues of HBx, predicted from the protein sequence, are reported in Table S1 (line 1) and highlighted in Figure S5a. These residues, mainly charged amino acids as arginines and lysines, are mostly located in the first middle part of the sequence, i.e. in the disordered region. Methods based on primary sequences matching predicted exons 6, 1-3 and 4 as the most interacting ones, with scores in decreasing order. HBx interaction with the intron 2 is predicted within the same score range.

To gain more insights into HBx - DLEU2 interactions we also performed docking calculations exploiting the 3D atomistic models of both the HBx protein and the DLEU2 RNA sequences. The

docking protocols are fully described in Supplemental Materials, in the Section “Docking predictions”. We found that HBx, thanks to its partially disordered nature, interacts with multiple DLEU2 species. The 3D structures of the HBx-DLEU2 exons complexes provided by docking calculations are shown in Figure S4. Although it is debated in the literature[38] if and to which extent docking scores may be used to rank different interacting couples, we use here with some caution the HADDOCK scores[39,40] as a relative indicator of the strength of the interaction. We found higher propensity of HBx for exons 6, 2, 8, 3 (in decreasing order) than for other exons. The score value for HBx-intron 2 interaction is in the same range as the above highly interacting exons. The 3D structures of HBx-intron 2 complexes are shown in Figure 3a, panels 1,2. Notably, two loci on intron 2 show a nearly equal probability to be occupied by HBx.

Additional information on the interaction between DLEU2 species and HBx came from the docking results by inspecting the protein-RNA interface according to a distance-based criterion. In Table S1 (lines 2-11, and 12-13), are listed the HBx residues that, based upon docking analysis, are predicted to preferentially interact with DLEU2 RNA exons and intron 2, respectively. According to these data and those from the sequence-based interaction affinities calculations, HBx binds RNAs preferentially via its more disordered part. HBx also shows preferential binding to double strands and internal loops motifs of ncRNAs (helical regions, stems, highlighted on the secondary structure representation in Figure S3), rather than complex 3D structures.

To further assess the quality of our docking protocols, we exploited some experimental information on other HBx complexes (see Supplemental Material, “Protein-protein and protein-lncRNAs control docking“). In particular, for the strongly interacting DDB1-HBx complex (PDB entry: 3I7H[33]), the complex structure provided by docking is very similar to the experimental one. Notably, the HADDOCK score is in the same range or slightly higher than the most interacting DLEU2 species.

Next we used the DLEU2 intron 2 - EZH2 binary complex as a reference to evaluate the possible interfering activity of HBx. We docked the DLEU2 intron 2 with EZH2 (see Figure 3a, panels 3, 4), and we identified two loci on intron 2, 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 3b (panels 5-8). HBx can occupy either one or the other of the two most probable

EZH2 loci on the intron 2 (see Figure 3b, panels 5 and 8) or even the same locus on intron 2 (see Figure 3b, panels 6 and 7). The opposite mechanism was also investigated, i.e. we used as docking targets for EZH2 the two binary HBx-intron 2 complexes shown in Figure 3a, obtaining the ternary complexes shown in Figure 3b, panels 1-4. Also in this case, we could find configurations in which the two proteins (HBx and EZH2) bind the intron 2 of DLEU2 in the same region (Figure 3b, panels 1 and 4) or in which EZH2 occupies one or the other HBx loci on the intron 2 (Figure 3b, panels 2 and 3). We verified that, on exon 6 (the most prone to protein interaction among considered DLEU2 exons), HBx and EZH2 share the same preferential binding site. If the RNA locus is occupied by one of the two proteins, the other one arranges to bind close by.

Our *in silico* model does not 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. Despite these limitations, our results suggest that, independently from the initial triggering event, HBx and EZH2 likely share the same binding site on DLEU2 and, as a consequence, HBx, binding to DLEU2 may interfere with EZH2/PRC2 complex functions and transcriptional regulation.

HBx-DLEU2 interaction regulates cccDNA transcription and HBV replication

RIP experiments and *in silico* modeling both indicate that HBx and DLEU2 interact. HBx is required for the transcripion of HBV mRNAs from the nuclear cccDNA viral minichromosome, including the HBV pgRNA that serves as template for reverse transcription and HBV replication[9,10,41]. To assess whether DLEU2, and its interaction with HBx, contribuite to cccDNA transcription, we first performed cccDNA chromatin isolation by RNA purification (ChIRP) experiments, using three independent antisense probe sets that specifically recover DLEU2 intronic sequences, in HBV replicating cells to confirm the binding of DLEU2 to the viral cccDNA minichromosome (Figure 4a). DLEU2 depletion by specific LNA longRNA Gapmers (Figure 4b) resulted in a reduction of cccDNA-bound H4 histones acetylation (Figure 4c), indicating that DLEU2 contributes to the epigenetic regulation of the HBV mini-chromosome. The impact of DLEU2 on HBx-driven cccDNA

transcription and HBV replication was confirmed by the decrease in HBV pgRNA levels (Figure 4d) and total HBV DNA (data not shown) in the presence of Gapmers targeting DLEU2 in both HBV infected PHHs and HepG2-NTCP cells as well as in HepG2 cells transfected with HBV linear monomers. HBx regulates cccDNA transcription by 2 mechanisms: 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 repressors recruitment and cccDNA silencing[10,11]. EZH2 is recruited onto the cccDNA in HepG2-NTCP infected cells[42]. Although in the absence of HBx, the K9 methylltransferase SETDB1[11] and the arginine protein methyltransferase PRMT5[43] seem to play a predominant role in cccDNA transcriptional repression, K27 methylation of cccDNA-bound H3 also increases[11] and EZH2 silencing results in an increased HBsAg production[43]. Notably, in HepG2 cells transfected with HBV monomers carrying a mutant HBx, where DLEU2 expression is reduced (Figure S2a, lower panels and Figure 4e), the recruitment of EZH2 onto the cccDNA is significantly increased (Figure 4f). Taken together these results implicate DLEU2 in the epigenetic control of cccDNA transcription and in HBV replication.

HBx / DLEU2 / EZH2 coregulation of TRIM13 expression

LncRNAs can regulate the expression of both neighbouring genes in *cis* or distantly located genes in *trans*[14]. As shown in Figure 1c, DLEU2 partially overlaps with the TRIM13 gene, that encodes for an ubiquitin E3 ligase implicated in ER-stress induced autophagy[44] and in the negative regulation of MDA5-mediated activation of type I interferon production in response to viral infections[45]. TRIM13 expression is upregulated in HBV-infected PHH (Figure 5a) and HepG2-NTCP cells, in HepG2 cells transfected with HBV monomers and in HepAD38 inducible cell lines (Figure S6a). HBx is recruited on the TRIM13 promoter (Figure 5b, left panel) in HBV infected PHHs. Coherent with the modulation of TRIM13 expression, the recruitment of HBx is accompanied by an increase in promoter-bound Polymerase II (Figure 5b, right panel) and H4 acetylation (Figure S6b). These results indicate that HBx regulates both DLEU2 and TRIM13 expression and we investigated whether DLEU2 was also involved in TRIM13 regulation. The reduction of DLEU2 RNA levels by selective degradation using Gapmers (see Figure 4b) resulted in a comparable and

statistically significant decrease in TRIM13 expression both in mock infected and HBV-infected PHHs (Figure 5c), in HBV-infected HepG2-NTCP and in HBV replicating HepG2 cells (Figure S6c). Notably, in HBV-replicating HepG2 cells co-trasfected with HBV monomers and DLEU2 specific LNA longRNA Gapmers, where the degradation of DLEU2 RNA levels is more efficient (Figure 4b, right panels), we observed a sharp decrease in H4 acetylation at the TRIM13 promoter (Figure 5d). Taken together these results indicate that DLEU2 and HBx cooperate to increase TRIM13 expression. Interestingly, DLEU2 RNA levels correlate positively with TRIM13 expression in a cohort of 373 hepatocellular carcinoma (HCC) patients (TCGA, https:// cancergenome.nih.gov/) (r=0,21, P=0,000048) (Figure 5e, left panel). A significant correlation is maintained in HBV-related HCCs (r=0,37, P=0,001) but not in HCV- or alcohol-related HCCs (Figure 5e, middle and right panels), confirming that DLEU2 and TRIM13 expression are strongly and positively correlated with HBV infection *in vivo*.

To better understand the mechanism underlying the cooperation between DLEU2 and HBx in the transcriptional regulation of host genes and the role of the EZH2/PRC2 complex, we further investigated TRIM13 promoter regulation. First, to determine whether DLEU2 is associated with the TRIM13 promoter in HBV negative and HBV positive cells, we performed ChIRP experiments. DLEU2 is recruited onto the TRIM13 promoter region in HBV negative cells (Figure 6a) and, consistent with the induction of DLEU2 expression (Figure 1e-f and Figure S2b, middle panel), its binding is enriched in HBV replicating cells (Figure 6a). Next, we showed that EZH2 binds the TRIM13 promoter in HBV negative HepG2 cells (Figure 6b, left panel). In HBV replicating HepG2 cells both EZH2 chromatin occupancy and the repressive mark H3 trimethyl Lys27 (H3me3K27) at the TRIM13 promoter decrease (Figure 6b, right panel). In the absence of HBx, both EZH2 binding and H3me3K27 increase at the TRIM13 promoter (Figure 6b). Finally, TRIM13 upregulation in cells exposed to the DZNep EZH2 inhibitor (Figure 6c) confirmed that TRIM13 is a EZH2/PRC2 target.

In addition to TRIM13, 260 additional genes were co-regulated with DLEU2 RNA in HCC patients (TCGA data, https://cancergenome.nih.gov/) (Table S2), with a significant enrichment in pathways associated to *Gene expression* and *Cell Cycle* (*Reactome* Biological Pathways) or in *Base excision repair* and *DNA replication* (*KEGG*)(data not shown), and 70 out of them were identified

as direct HBx target in ChIP-Seq experiments[12] (Table S2). We could validate by anti-HBx ChIP and anti-DLEU2 ChIRP (Figure 6d and 6e) the CCNB2 gene as a direct target of the HBx-DLEU2-EZH2 complex.

Taken together these results would suggest that the functional interaction between HBx, DLEU2 and EZH2/PRC2 results in the transcriptional activation of a subset of EZH2/PRC2 target genes in HBV infected cells and HBV-related HCCs.

Conclusions

The essential role of the PRC2 complex and the deposition of the H3K27me3 repressive chromatin mark in the promotion of epigenetic silencing during embryonic development and differentiation[46] and the impact of its disregulation in cancer development[47] are well established. PRC2 complex binding to coding and ncRNAs is also well documented[16,17] but there is no unified view on how PRC2 interacts with chromatin bound RNAs. Current models, that 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]; b) RNAs serving as decoys to prevent PRC2 binding to active genes or to evict PRC2[48]; c) PRC2 scanning of all nascent RNAs to locate possible sites of action[49]. Notably, PRC2 binding to RNA and chromatin association appear to be antagonistic[48]. Both motif specific and apparently promiscous binding have been described[49–51]. More recently, the identification of folded guanine quadruplexes (G4) structures as high affinity motifs for PRC2 binding and the widespread presence of G-tract motifs in many RNAs has provided an unifying model for PRC2/RNA interactions[52].

Here we show that HBx binds the promoter region of the DLEU2 lncRNA, enhances DLEU2 transcription and induces the accumulation of DLEU2 RNA species in the nucleus. DLEU2 is highly enriched, together with few other lncRNAs, in EZH2 RIP-Seq performed in colon cancer cells[17]. We could confirm by RIP experiments that DLEU2 binds EZH2 in liver cells and, in HBV infected cells, the viral protein HBx. These results highlight a novel capacity of the viral protein to bind ncRNAs. HBx is required to initiate and maintain transcripion from the nuclear cccDNA viral minichromosome[10,11,41] in order to drive HBV replication. cccDNA chromatin isolation by RNA

purification (ChIRP) experiments showed that DLEU2 binds to the cccDNA and functional experiments with DLEU2 specific LNA longRNA Gapmers supported the role of DLEU2 recruitment in the regulation of cccDNA transcription. Interestingly, multiple putative G4 structures can be identified on the HBV cccDNA and on the HBV pgRNA (data not shown) and a G-quadruplex motif in the envelope gene promoter has been shown to regulate HBV transcription[53]. Altogether, our results and these observations underline the potential for strategies interfering with G-quadrupex and DLEU2 functional interaction on the cccDNA as potential new targets for anti-HBV therapeutic approaches.

HBx has been reported to bind a large number of target sequences on the cellular genome[12] and to interacts with EZH2/PRC2 in order to modulate transcription of specific host genes[21,28]. Our *in silico* modeling of DLEU2-EZH2-HBx interaction suggested that HBx and EZH2 share the same binding sites on DLEU2. As a consequence, HBx, by binding to DLEU2, may interfere with EZH2/PRC2 complex functions and transcriptional regulation. LncRNAs are known to regulate the expression of both neighbouring genes in *cis* or distantly located genes in *trans*[14]. DLEU2 partially overlaps with the TRIM13 gene and, by combining ChIRP and ChIP experiments, we showed that both DLEU2 and EZH2 bind the TRIM13 promoter. Upon HBV infection, the increase in TRIM13 expression is accompanied by HBx recruitment (ChIP) and an increase in DLEU2 occupancy (ChIRP). Whereas DLEU2 association with EZH2 (RIP) increases, reflecting the HBx dependent activation of DLEU2 expression, the recruitment of EZH2 on the TRIM13 promoter (ChIP) is decreased. This observation likely reflects the known competition between PRC2 binding to RNA (as detected by RIP) vs PRC2 association with chromatin (as detected by ChIP). Notably, the increase of both EZH2 binding to DLEU2 (RIP) and the incresaed occupancy of DLEU2 on the TRIM13 promoter (ChIRP) is consistent with a model in which DLEU2 interaction with PRC2/EZH2 displaces the PRC2/EZH2 complex from close contact with chromatin[48] but still remains in the proximity of the original targeted chromatin due to the persitence of the RNA binding (Figure 6g).

Interestingly, DLEU2 and TRIM13 expression correlate positively in HBV-related HCC patients (TCGA, https:// cancergenome.nih.gov/). In addition to TRIM13, 260 additional genes were co-regulated with DLEU2 RNA in HCC HBV-related patients (TCGA, https:// cancergenome.nih.gov/)

and 70 out of them were identified as direct HBx target in ChIP-Seq experiments[12] suggesting that the interaction between DLEU2 and HBx might co-regulate a subset of EZH2/PRC2 targets in HBV replicating cells and HCC. DLEU2 interaction with the PRC2 complex inhibits EZH2 repressive activity[14] (Figure 6g) and the concomitant increase in DLEU2 and HBx recruitment in HBV replicating cells and in HBV-related HCCs would further increase transcription and freeze the promoters of a subset of EZH2/PRC2 targets in an active state (Figure 6g).

Methods

Cell cultures. HepG2, HepG2 derived clones HepG2.2.15 (HepAD38) and NTCP-HepG2 cells were cultured in supplemented Dulbecco's modified Eagle's medium (DMEM) as described[12] and maintained in a 5% CO2 humidified incubator at 37 °C.

Primary culture of human hepatocytes and HBV infection. Primary human hepatocytes (PHH) (provided by Prof. M. Rivoire, Centre Leon Berard de Cancerologie, 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. PHH were prepared and cultured as described elsewhere[54]. Briefly, the cells are plated overnight in collagen-coated dishes (BD Biosciences) at 2×10^5 cells/cm^2 in William's medium (Life Technologies) supplemented with 10% FetalClone II (Thermo Scientific), 1% penicillin/streptomicine and 1% glutamine (Life Technologies), as well as 5 µg/ml Insulin and 5×10-7 M hydrocortisone hemisuccinate (Sigma Aldrich). After 24 h PHH are extensively washed in serum-free medium and kept in serum-free medium for one more day to counter select the growth of contaminating fibroblast and endothelial cells, and infected 48 hours after plating with HBV virus (i.e. inoculum) produced in HepG2.2.15 cells[55].

Transient transfection of full-length HBV DNA genomes. Monomeric linear full-length WT and HBX mt HBV genomes were released from the pCR.HBV.A.EcoRI and the pCR.HBXmt.A.EcoRI

plasmids using EcoRI-PvuI (New England Biolabs). Linear HBV monomers were transfected into HepG2 human hepatoma cells using the Mirrus Bio trans IT-LT1 reagent (Mirrus, Cat#MIR2305) as previously described[12]. For inhibition of DLEU2 experiments in HepG2 cells, mixed LNA/DNA oligonucleotides (GAPMERS) were generated (Exiqon, Cat#300600) against human DLEU2 ncRNA. The DNA based antisense oligonucleotides against DLEU2 RNA mediates the RNaseH catalysed degradation of DLEU2 RNA as shown for other nuclear retained RNAs. Gapmers and relative control were transfected for 48 hours at a final concentration of 40nM into HepG2 human hepatoma cells using the Lipofectamine Plus reagent (Invitrogen Cat#11514015 and Cat#18324020).

***ChIP assays*.** 48 hours after transfection with linear HBV monomers and/or Gapmers*,* HepG2 cells were resuspended in 1-2 ml of ChIP lysis buffer (50 mM Tris HCL, pH 8, 0.5% NP40, 1 mM EDTA, and 100 mM NaCL) and incubated 10 minutes at 4°C. The lysate was centrifuged at 10,000 g for 2 minutes to pellet the nuclei. The supernatant was removed, and the nuclei fixed in 1% formaldehyde for 30 minutes at 4°C. Isolated cross-linked nuclei were extracted with a 20 mM Tris, pH 8, 3 mM MgCl2, 20 mM KCl buffer containing protease inhibitors, pelleted by microcentrifugation, and lysed by incubation in SDS lysis buffer (1% sodium dodecyl sulfate, 10 mM EDTA, 50 mM Tris-chloride, pH 8.1) containing protease inhibitors. For ChIP the resulting chromatin solution was sonicated for 5 pulses of 45 seconds at 80% power to generate 300- to 1000-bp DNA fragments using a Bioruptor Sonicator (Diagenode Inc). 100 µl of Dynabeads Protein G (Invitrogen, Cat#10003D) were added to each 1 ml chromatin preparation and incubated on a rotator for 14–16 hours at 4°C. The antibodies used are anti-HBx (ThermoFisher, Cat#MA1-081), anti-AcH4 (Millipore, Cat#06-866), anti-Pol II (Diagenode, Cat#C15200004), anti-EZH2 (Cell Signalling, Cat#5246), anti-Histone H3 (tri methyl K27) (Abcam, Cat#6002). 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 is analyzed by real-time PCR amplification using either primers (NCC1 and CCCAS) and probes (FL and Red) specific for the

HBV cccDNA or specific primers for genes promoters (Table S3) and the Lightcycler 480 Probes Master (Roche).

HBV pgRNA and cellular mRNA analysis. Total RNA was extracted from HepG2 cells 48 hours after transfection with the linear monomeric HBV DNA or HBV infected PHH using the TRIzol reagent (Invitrogen) as recommended by the manufacturer. Nuclear and cytoplasmatic extraction was performed by PARIS KIT (Ambion, Cat#AM1921). The RNA samples were treated with RQ1 RNase-Free DNase (Promega, Cat#M6101) for 30 minutes at 37°C and stored until used. RNA quality and quantity were monitored by ethidium bromide staining and UV absorbance. For pgRNA analysis, 2 ug of DNase-treated RNA was reverse transcribed and amplified by 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 indicates in Tables S3. The human Beta Actin housekeeping gene Light Cycler Set (Roche Diagnostics) was used to normalize the RNA samples. mRNAs levels in cultured cells 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 either a specific TaqMan FAM-probe (Applied Biosystems, Inc.) with the specific primers (Table S3). Results were normalized with respect to human Beta Actin and expressed as described above.

RIP (RNA immunoprecipitation) 48 hours after transfection with linear HBV monomers HepG2 or HBV infected PHH nuclear pellets were fixed in 1% formaldehyde for 30 min at room temperature, 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 RNAase inhibitor SUPERASin, Protease inhibitors) and splitted into two fractions of 500 ml each (mock and IP). Then we performed a mechanically shearing using a dounce homogenizer with 15–20 strokes. After centrifugation at 13,000 rpm for 10 min, the supernatant was immunoprecipitated with 10ug of anti-HBx or anti-EZH2 and incubated overnight at 4°C with gently rotation. Then we added protein G beads (40 ul) and incubated for 1 hr at 4°C

with gentle rotation. Pellet beads were centrifugated at 2500 rpm for 30 s, removed the supernatant and resunspended in 500ul of RIP buffer (three total RIP washes) to wash off the unbound material. The washes were followed by one wash in PBS. Finally, isolate coprecipitated RNAs were extracted by TRIZOL and eluted with nuclease-free water. Reverse transcription of DNAse treated RNA was performed with ThermoScript (Invitrogen, Inc., Carlsbad, US-CA) according to manufacturer's instructions. cDNA was used to amplify the DLEU2 regions and MALAT1.

ChIRP (**Chromatin Isolation by RNA Purification).** Antisense DNA probes were designed against *DLEU2* isoforms using an online designer at https://www.biosearchtech.com/support/tools/design-software/chirp-probe-designer. 48 hours after transfection with linear HBV monomers, HepG2 nuclear pellets were Crosslinked in presence of 1% glutaraldehyde in PBS for 10 minutes at room temperature by shaking and then quenched with 0.125 M glycine for 5 minutes. The pellets were washed in PBS, freezed in liquid nitrogen and stored at -80°C. Pellets were resuspend in 1 ml freshly 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 sec ON and 45 sec OFF, cycle time -15 min each, until DNA was in the size range of 100–500 bp. From each condition 10 ul were taken and preserved for DNA/RNA INPUT each. 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 RNAase inhibitor were added to the buffer prior to use). DLEU2 probes (100 pmol) were added to chromatin, mixed by rotation at 37°C overnight. 100µl of washed Streptavidin magnetic C1 beads (Invitrogen) were added per 100 pmol of probes, and the whole reaction was mixed for another 30 min at 37°C. Beads/biotin-probes/RNA/chromatin complex were collected on magnets, washed 5 times with wash buffer (2X SSC, 0.5% SDS and PMSF added just before use). 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 RNAase inhibitor were added

to the buffer prior to use) 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 S3).

NanoString nCounter lncRNA gene expression assay and data analysis. NanoString nCounter assays were performed using 100ng of purified RNA following manufacturer's instructions (NanoString Technologies). The long non conding RNA probe set was selected from the custom probe set. Sample preparation and Hybridization reactions wes performed according to manufacturer's instructions (NanoString Technologies). All hybridization reactions were incubated at 65°C for a minimum of 16 hrs. 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 added for data normalization.

***In silico* modeling. a) Structure**. Exons FASTA sequences of DLEU2, as given by GenomeBrowser (https://genome.ucsc.edu/cgi-bin/hgGateway), have been used to obtain the secondary structure of the exons and intron 2, via DotKnot[30,31]. The FASTA sequence and the secondary structure of the considered lncRNAs have been used to obtain the tertiary structure via the RNAcomposer tool[32]. Secondary structure predictions for HBx, based on aminoacid sequence, have been performed with IUPred[56] (http://iupred.enzim.hu/), PONDR[57] (http://www.pondr.com/) and DISOPRED2[58] (http://bioinf.cs.ucl.ac.uk/web_servers/disopred/disopred_overview/). The HBx structure has been modelled by I-TASSER[59–61] (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[35] (http://pridb.gdcb.iastate.edu/RPISeq/) and catRAPID[62] (http://s.tartaglialab.com/page/catrapid_group). RNA-binding residues are predicted via RNABindRPlus[36] (http://ailab1.ist.psu.edu/RNABindRPlus/), based on the aminoacid sequence. HBx protein and RNAs models were docked by using HEX[63] (http://hex.loria.fr/) and NPDock[64–66] (http://genesilico.pl/NPDock) for the preliminary screening and determination of the interface residues, and by using HADDOCK[39,40] for a more refined, interacting-residues driven docking.

Quantification and statistical analysis. Data 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 averages ± SD. P values were determined using Unpaired t-test. *p≤0.05 **p≤0.005, ***p ≤ 0.0005, ****p ≤ 0.00005 were considered significant. Statistical details can be found directly in the figures or in the corresponding figure legends.

REFERENCES

1. World Health Organization (WHO). Global Hepatitis Report 2017. Geneva: World Health Organization; 2017. Licence: CC BY-NC-SA 3.0 IGO
2. Schweitzer, A., Horn, J., Mikolajczyk, R. T., Krause, G. & Ott, J. J. Estimations of worldwide prevalence of chronic hepatitis B virus infection: A systematic review of data published between 1965 and 2013. *Lancet* **386,** 1546–1555 (2015).
3. Stanaway, J. D. *et al.* The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of Disease Study 2013. *Lancet* **388,** 1081–1088 (2016).
4. Lok, A. S., Zoulim, F., Dusheiko, G. & Ghany, M. G. Hepatitis B cure: From discovery to regulatory approval. *Hepatology* **66,** 1296–1313 (2017).
5. European Association for the Study of the Liver (EASL). EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. *J. Hepatol.* **67,** 370–398 (2017).
6. Mason, W. S. *et al.* HBV DNA Integration and Clonal Hepatocyte Expansion in Chronic Hepatitis B Patients Considered Immune Tolerant. *Gastroenterology* **151,** 986–998.e4 (2016).
7. Testoni, B., Levrero, M. & Zoulim, F. Challenges to a Cure for HBV Infection. *Semin. Liver Dis.* **37,** 231–242 (2017).
8. Pollicino, T. *et al.* Hepatitis B virus replication is regulated by the acetylation status of hepatitis B virus cccDNA-bound H3 and H4 histones. *Gastroenterology* **130,** 823–837 (2006).
9. Decorsière, A. *et al.* Hepatitis B virus X protein identifies the Smc5/6 complex as a host restriction factor. *Nature* **531,** 386–389 (2016).
10. Belloni, L. *et al.* Nuclear HBx binds the HBV minichromosome and modifies the epigenetic regulation of cccDNA function. *Proc. Natl. Acad. Sci.* **106,** 19975–19979 (2009).
11. Rivière, L. *et al.* HBx relieves chromatin-mediated transcriptional repression of hepatitis B viral cccDNA involving SETDB1 histone methyltransferase. *J. Hepatol.* **63,** 1093–1102 (2015).

12. Guerrieri, F. *et al.* Genome-wide identification of direct HBx genomic targets. *BMC Genomics* **18,** (2017).
13. Yang, X. *et al.* The emergence of long non-coding RNAs in the tumorigenesis of hepatocellular carcinoma. *Cancer Letters* **360,** 119–124 (2015).
14. Wang, K. C. & Chang, H. Y. Molecular Mechanisms of Long Noncoding RNAs. *Molecular Cell* **43,** 904–914 (2011).
15. Mercer, T. R., Dinger, M. E. & Mattick, J. S. Long non-coding RNAs: Insights into functions. *Nature Reviews Genetics* **10,** 155–159 (2009).
16. Zhao, J. *et al.* Genome-wide Identification of Polycomb-Associated RNAs by RIP-seq. *Mol. Cell* **40,** 939–953 (2010).
17. Guil, S. *et al.* Intronic RNAs mediate EZH2 regulation of epigenetic targets. *Nat. Struct. Mol. Biol.* **19,** 664–670 (2012).
18. Moyo, B., Nicholson, S. A. & Arbuthnot, P. B. The role of long non-coding RNAs in hepatitis B virus-related hepatocellular carcinoma. *Virus Res.* **212,** 103–113 (2015).
19. Zhang, Q. *et al.* Analysis of long noncoding RNA expression in hepatocellular carcinoma of different viral etiology. *J. Transl. Med.* **14,** (2016).
20. Degli Esposti, D. *et al.* Identification of novel long non-coding RNAs deregulated in hepatocellular carcinoma using RNA-sequencing. *Oncotarget* **7,** 31862–31877 (2016).
21. Hu, J. J. *et al.* HBx-upregulated lncRNA UCA1 promotes cell growth and tumorigenesis by recruiting EZH2 and repressing p27Kip1/CDK2 signaling. *Sci. Rep.* **6,** (2016).
22. Peng, L. *et al.* Developmental programming of long non-coding RNAs during postnatal liver maturation in mice. *PLoS One* **9,** (2014).
23. Yang, Y. *et al.* Recurrently deregulated lncRNAs in hepatocellular carcinoma. *Nat. Commun.* **8,** (2017).
24. Villa, E. *et al.* Neoangiogenesis-related genes are hallmarks of fast-growing hepatocellular carcinomas and worst survival. results from a prospective study. *Gut* **65,** 861–869 (2016).
25. Lerner, M. *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.* **315,** 2941–

2952 (2009).

26. Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464,** 1071–1076 (2010).

27. Kogo, R. *et al.* Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res.* **71,** 6320–6326 (2011).

28. Fan, H., Zhang, H., Pascuzzi, P. E. & Andrisani, O. Hepatitis B virus X protein induces EpCAM expression via active DNA demethylation directed by RelA in complex with EZH2 and TET2. *Oncogene* **35,** 715–726 (2016).

29. Antonysamy, S. *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* **8,** (2013).

30. Sperschneider, J. & Datta, A. Dot knot: Pseudoknot prediction using the probability dot plot under a refined energy model. *Nucleic Acids Res.* **38,** (2010).

31. Sperschneider, J., Datta, A. & Wise, M. J. Heuristic RNA pseudoknot prediction including intramolecular kissing hairpins. *RNA* **17,** 27–38 (2011).

32. Popenda, M. *et al.* Automated 3D structure composition for large RNAs. *Nucleic Acids Res.* **40,** (2012).

33. Minor, M. M. & Slagle, B. L. Hepatitis B virus HBx protein interactions with the ubiquitin proteasome system. *Viruses* **6,** 4683–4702 (2014).

34. Sidhu, K., Kumar, S., Reddy, V. S. & Kumar, V. Mass spectrometric determination of disulfide bonds in the biologically active recombinant HBx protein of hepatitis B virus. *Biochemistry* **53,** 4685–4695 (2014).

35. Muppirala, U. K., Honavar, V. G. & Dobbs, D. Predicting RNA-Protein Interactions Using Only Sequence Information. *BMC Bioinformatics* **12,** (2011).

36. Walia, R. R. *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* **9,** (2014).

37. Somarowthu, S. *et al.* HOTAIR Forms an Intricate and Modular Secondary Structure. *Mol. Cell* **58,** 353–361 (2015).

38. Wass, M. N., Fuentes, G., Pons, C., Pazos, F. & Valencia, A. Towards the prediction of protein interaction partners using physical docking. *Mol. Syst. Biol.* **7,** (2011).

39. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125,** 1731–1737 (2003).

40. Van Zundert, G. C. P. *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **428,** 720–725 (2016).

41. Lucifora, J. *et al.* Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. *J. Hepatol.* **55,** 996–1003 (2011).

42. Zhang, H. *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* **64,** 1033–1048 (2016).

43. Zhang, W. *et al.* 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 (2017).

44. Tomar, D., Singh, R., Singh, A. K., Pandya, C. D. & Singh, R. TRIM13 regulates ER stress induced autophagy and clonogenic ability of the cells. *Biochim. Biophys. Acta - Mol. Cell Res.* **1823,** 316–326 (2012).

45. Narayan, K. *et al.* TRIM13 Is a Negative Regulator of MDA5-Mediated Type I Interferon Production. *J. Virol.* **88,** 10748–10757 (2014).

46. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469,** 343–349 (2011).

47. Bracken, A. P. & Helin, K. Polycomb group proteins: Navigators of lineage pathways led astray in cancer. *Nature Reviews Cancer* **9,** 773–784 (2009).

48. Beltran, M. *et al.* The interaction of PRC2 with RNA or chromatin s mutually antagonistic. *Genome Res.* **26,** 896–907 (2016).

49. Davidovich, C., Zheng, L., Goodrich, K. J. & Cech, T. R. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat. Struct. & Mol. Biol.* **20,** 1250–1257 (2013).

50. Kaneko, S., Son, J., Bonasio, R., Shen, S. S. & Reinberg, D. Nascent RNA interaction keeps PRC2 activity poised and in check. *Genes Dev.* **28,** 1983–1988 (2014).

51. Davidovich, C. *et al.* Toward a consensus on the binding specificity and promiscuity of PRC2 for RNA. *Mol. Cell* **57,** 552–559 (2015).

52. Wang, X. *et al.* Targeting of Polycomb Repressive Complex 2 to RNA by Short Repeats of Consecutive Guanines. *Mol. Cell* **65,** 1056–1067.e5 (2017).

53. Biswas, B., Kandpal, M. & Vivekanandan, P. A G-quadruplex motif in an envelope gene promoter regulates transcription and virion secretion in HBV genotype B. *Nucleic Acids Res.* **45,** 11268–11280 (2017).

54. Lecluyse, E. L. & Alexandre, E. Isolation and culture of primary hepatocytes from resected human liver tissue. *Methods Mol. Biol.* **640,** 57–82 (2010).

55. Hantz, O. *et al.* Persistence of the hepatitis B virus covalently closed circular DNA in HepaRG human hepatocyte-like cells. *J. Gen. Virol.* **90,** 127–135 (2009).

56. Dosztányi, Z., Csizmok, V., Tompa, P. & Simon, I. IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21,** 3433–3434 (2005).

57. Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K. & Uversky, V. N. PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta - Proteins Proteomics* **1804,** 996–1010 (2010).

58. Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F. & Jones, D. T. The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **20,** 2138–2139 (2004).

59. Yang, J. *et al.* The I-TASSER suite: Protein structure and function prediction. *Nature Methods* **12,** 7–8 (2014).

60. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5,** 725–738 (2010).

61. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9,**

(2008).

62. Bellucci, M., Agostini, F., Masin, M. & Tartaglia, G. G. Predicting protein associations with long noncoding RNAs. *Nature Methods* **8,** 444–445 (2011).

63. Ritchie, D. W., Kozakov, D. & Vajda, S. Accelerating and focusing protein-protein docking correlations using multi-dimensional rotational FFT generating functions. *Bioinformatics* **24,** 1865–1873 (2008).

64. Puton, T., Kozlowski, L., Tuszynska, I., Rother, K. & Bujnicki, J. M. Computational methods for prediction of protein-RNA interactions. *J. Struct. Biol.* **179,** 261–268 (2012).

65. Tuszynska, I. *et al.* Computational modeling of protein-RNA complex structures. *Methods* **65,** 310–319 (2014).

66. Tuszynska, I., Magnus, M., Jonak, K., Dawson, W. & Bujnicki, J. M. NPDock: A web server for protein-nucleic acid docking. *Nucleic Acids Res.* **43,** W425–W430 (2015).

Author Contributions

DS and LC share co-first authorship. F.G. supervised the study. D.S., O.F., S.J., and F.G. performed the experiments and analyzed the data. L.C. and G.C. conducted computational analysis. M.P. performed bioinformatic analysis. F.G and M.L. conceived and designed the study. L.C., M.L. and F.G. wrote the manuscript.

Declaration of Interests

The authors disclose no conflicts.

ACKNOWLEDGMENTS

We warmly thank Dr. M. Ballarino and Prof. Zoulim for fruitful discussions. For experimental assistance, we thank Dr. L. Calvo (Nanostring), Dr. A. D'Alfonso (RNA extraction), Dr. L. Belloni and Dr. L. Lupacchini (cell lines).

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 Agence National de la Recherche (ANR@TRACTION) to ML; from the EU project 667273 HEP-CAR to ML.

Figure 1

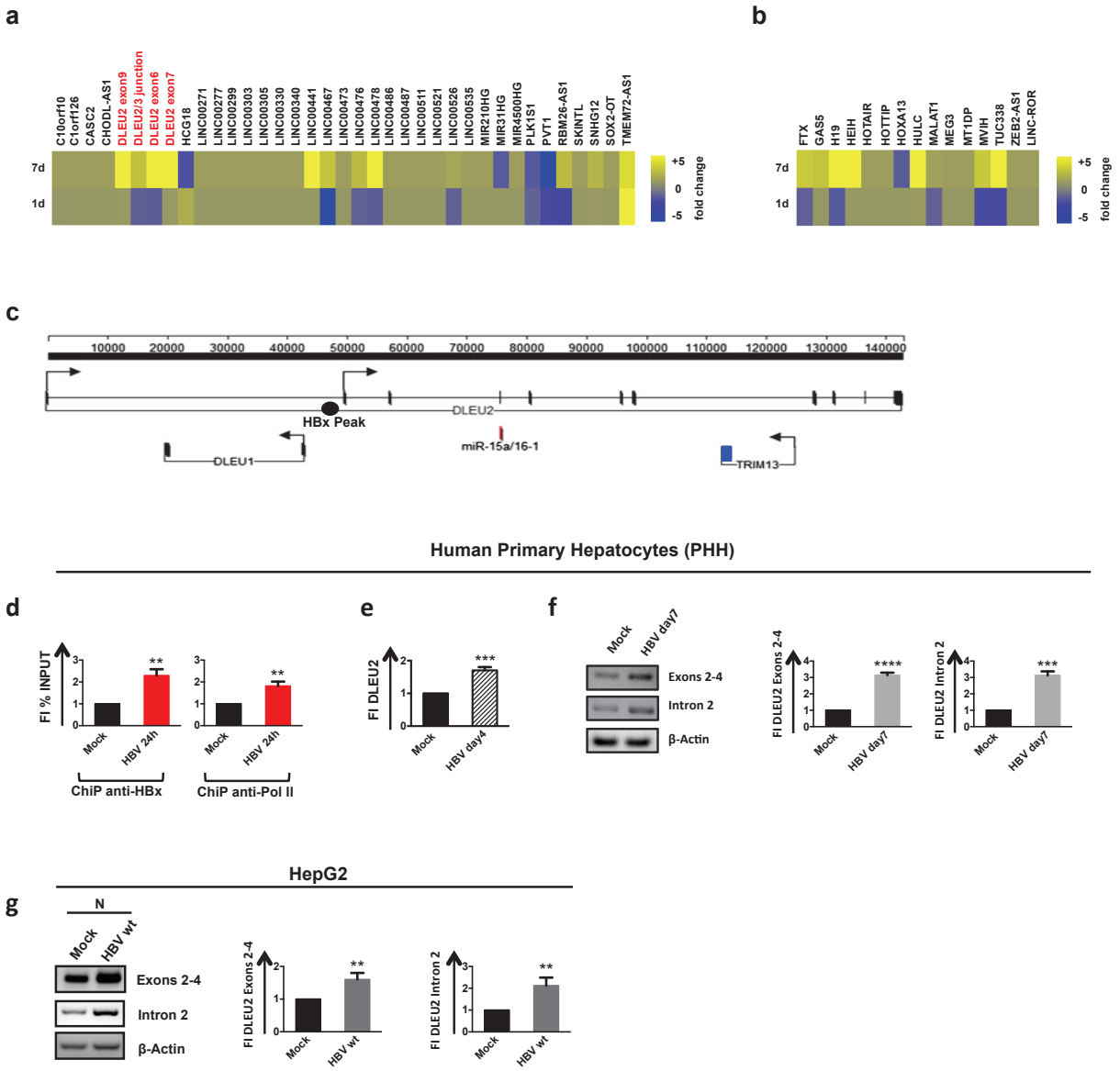

a
C10orf10
C1orf126
CASC2
CHODL-AS1
DLEU2 exon9
DLEU2/3 junction
DLEU2 exon6
DLEU2 exon7
HCG18
LINC00271
LINC00277
LINC00299
LINC00303
LINC00305
LINC00330
LINC00340
LINC00441
LINC00467
LINC00473
LINC00476
LINC00478
LINC00486
LINC00487
LINC00511
LINC00521
LINC00526
LINC00535
MIR210HG
MIR31HG
MIR4500HG
PLK1S1
PVT1
RBM26-AS1
SKINTL
SNHG12
SOX2-OT
TMEM72-AS1
7d
1d
+5
0
-5
fold change
b
FTX
GAS5
H19
HEIH
HOTAIR
HOTTIP
HOXA13
HULC
MALAT1
MEG3
MT1DP
MVIH
TUC338
ZEB2-AS1
LINC-ROR
7d
1d
+5
0
-5
fold change
c
10000
20000
30000
40000
50000
60000
70000
80000
90000
100000
110000
120000
130000
140000
HBx Peak
DLEU2
DLEU1
miR-15a/16-1
TRIM13
Human Primary Hepatocytes (PHH)
d
FI % INPUT
**
**
Mock
HBV 24h
Mock
HBV 24h
ChiP anti-HBx
ChiP anti-Pol II
e
FI DLEU2
***
Mock
HBV day4
f
Mock
HBV day7
Exons 2-4
Intron 2
β-Actin
FI DLEU2 Exons 2-4
****
Mock
HBV day7
FI DLEU2 Intron 2
***
Mock
HBV day7
HepG2
g
N
Mock
HBV wt
Exons 2-4
Intron 2
β-Actin
FI DLEU2 Exons 2-4
**
Mock
HBV wt
FI DLEU2 Intron 2
**
Mock
HBV wt

Figure 2

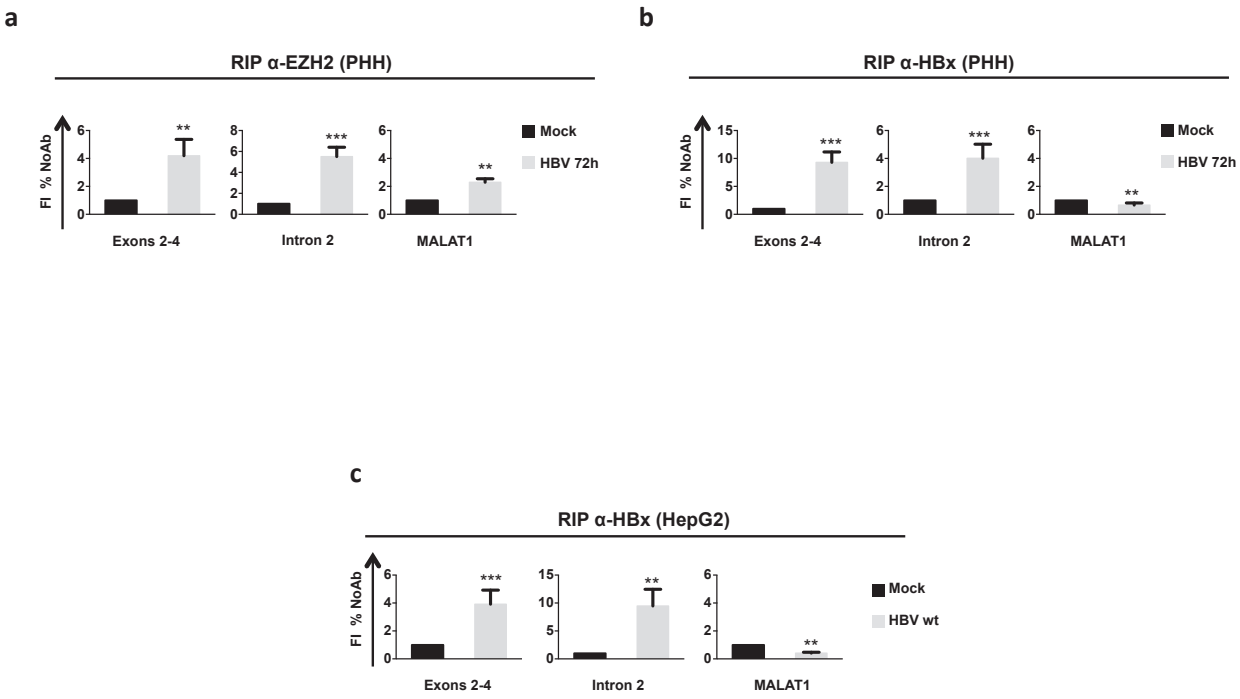

a
RIP α-EZH2 (PHH)
FI % NoAb
Exons 2-4
Intron 2
MALAT1
Mock
HBV 72h
b
RIP α-HBx (PHH)
FI % NoAb
Exons 2-4
Intron 2
MALAT1
Mock
HBV 72h
c
RIP α-HBx (HepG2)
FI % NoAb
Exons 2-4
Intron 2
MALAT1
Mock
HBV wt

Figure 3

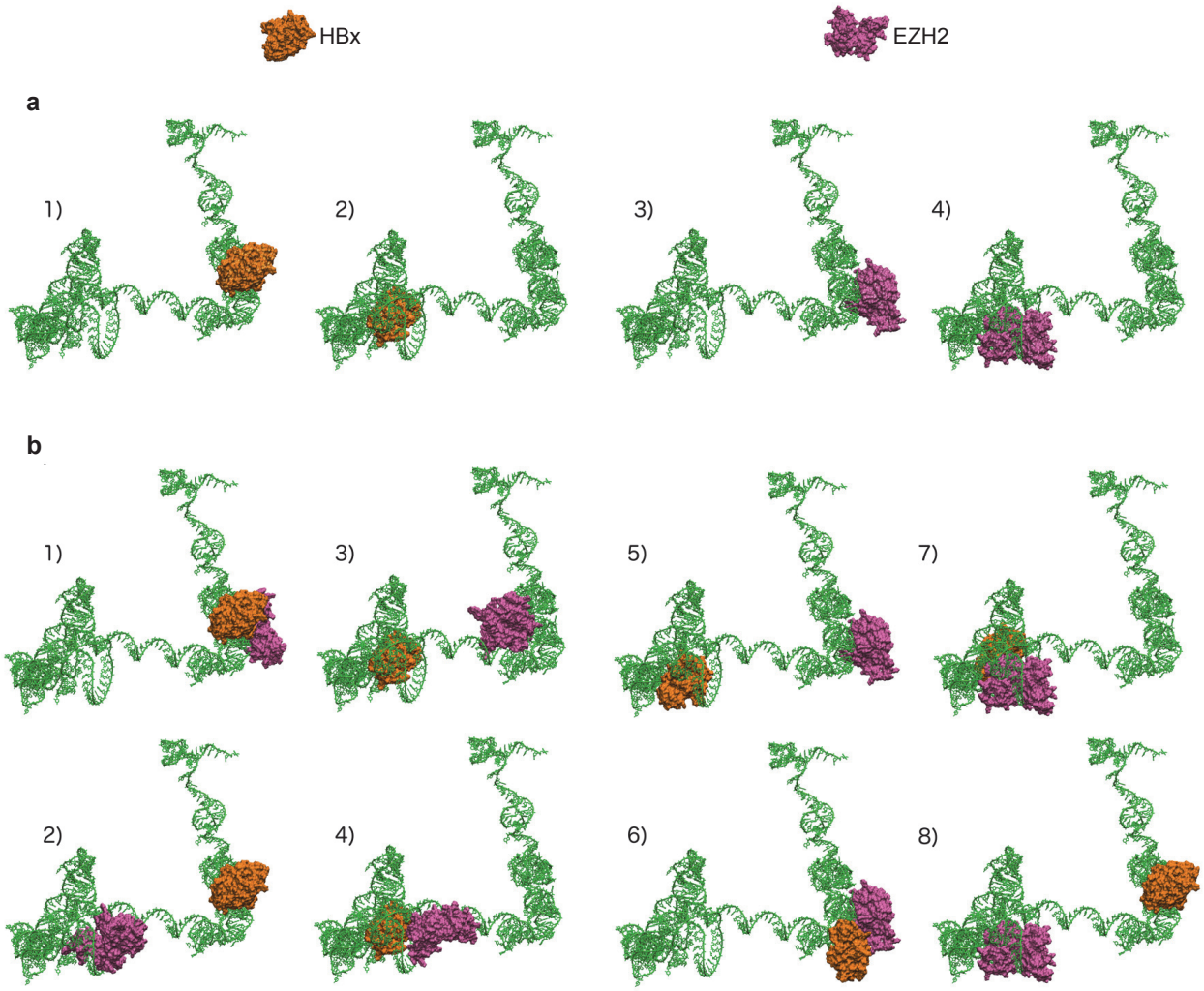


Figure 4

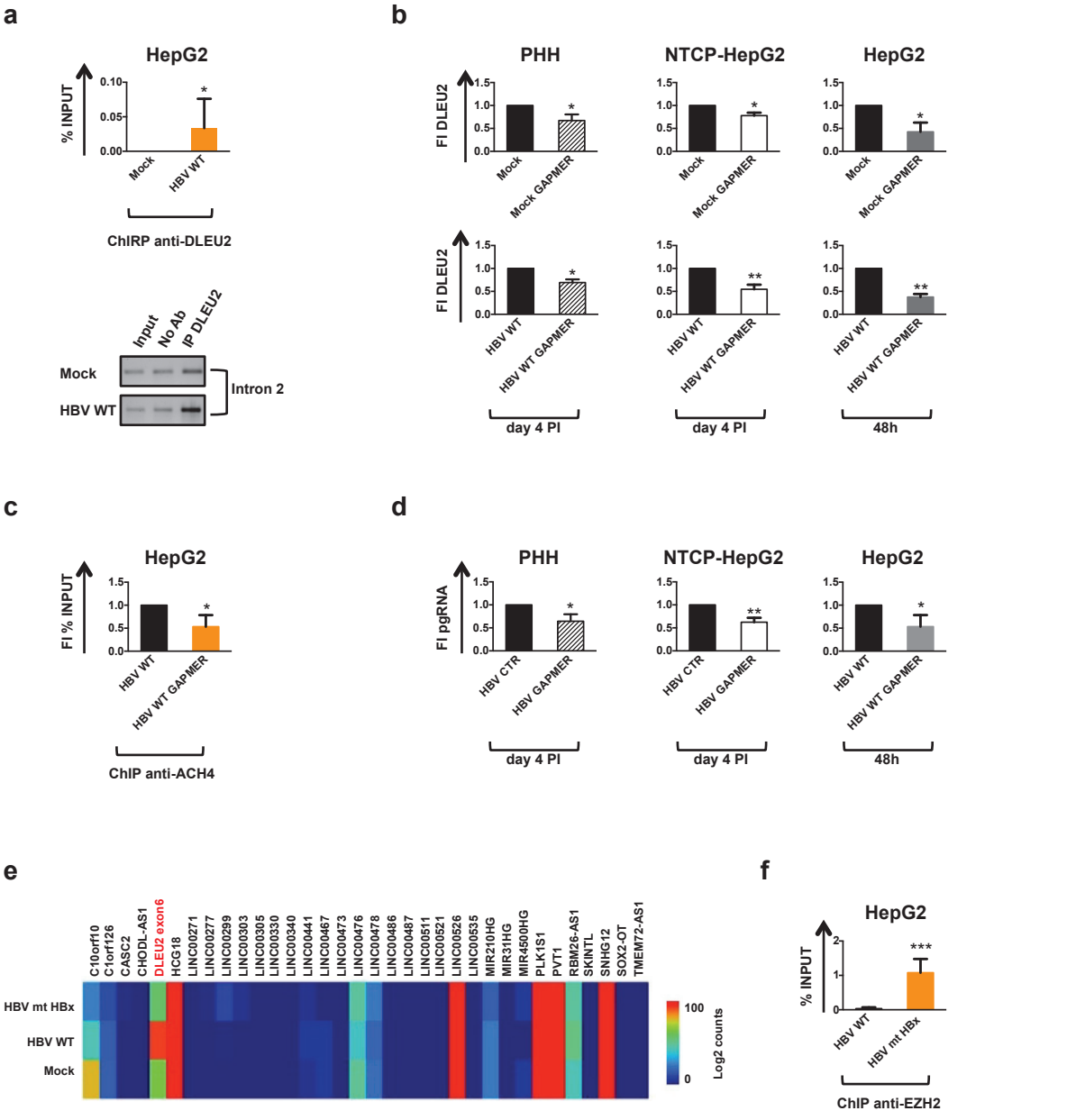
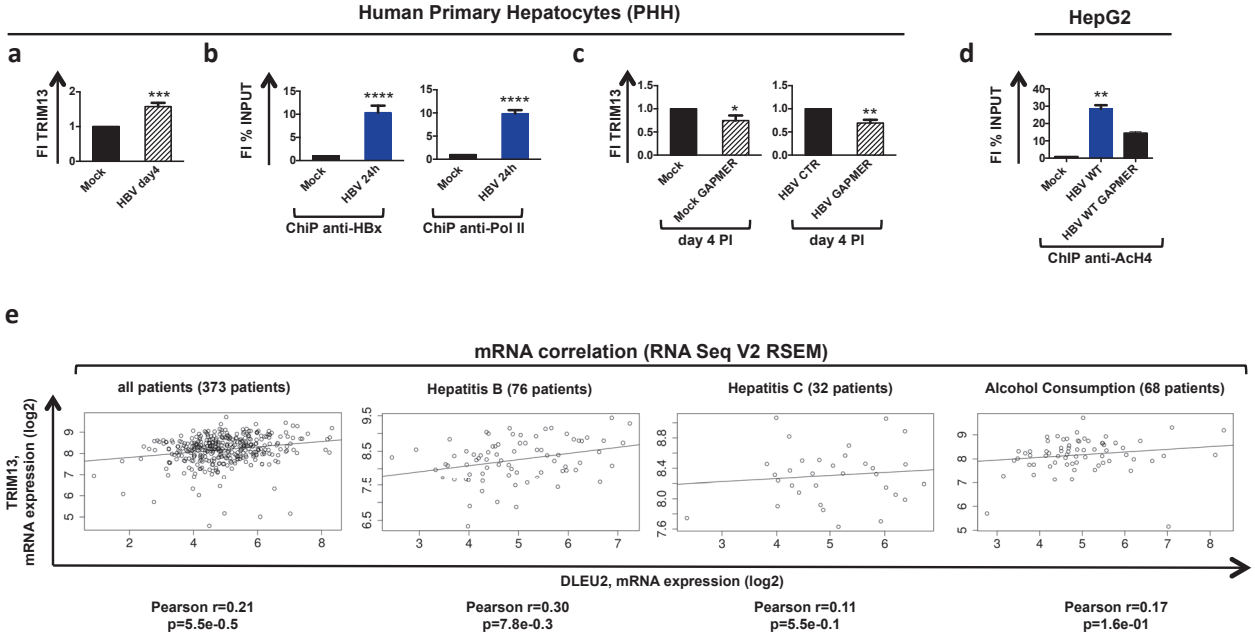
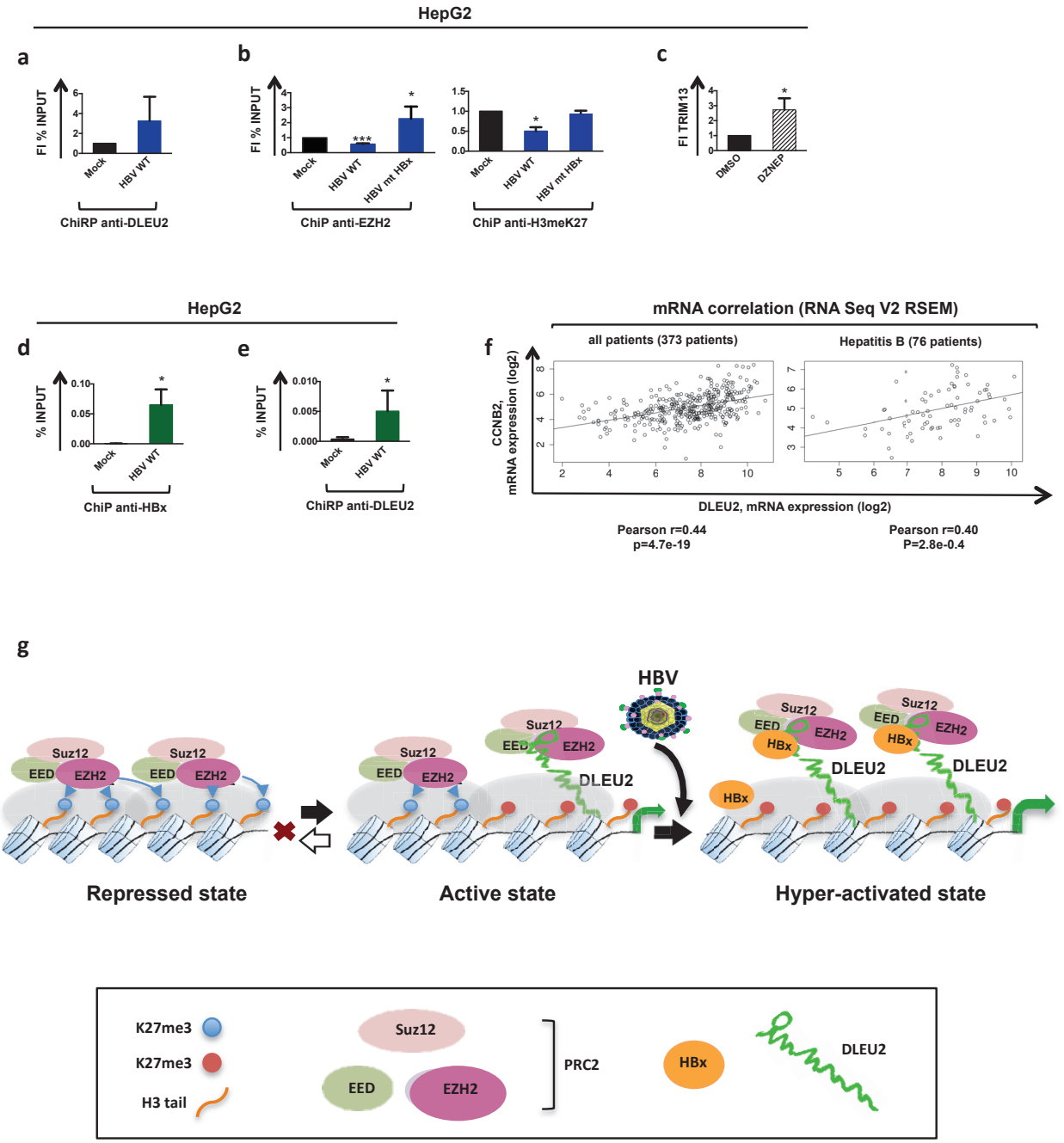

a
HepG2
% INPUT
0.10
0.05
0.00
Mock
HBV WT
ChIRP anti-DLEU2
Input
No Ab
IP DLEU2
Mock
HBV WT
Intron 2
b
PHH
NTCP-HepG2
HepG2
FI DLEU2
Mock
Mock GAPMER
HBV WT
HBV WT GAPMER
day 4 PI
48h
c
HepG2
FI % INPUT
HBV WT
HBV WT GAPMER
ChIP anti-ACH4
d
PHH
NTCP-HepG2
HepG2
FI pgRNA
HBV CTR
HBV GAPMER
e
C10orf10
C1orf126
CASC2
CHODL-AS1
DLEU2 exon6
HCG18
LINC00271
LINC00277
LINC00299
LINC00303
LINC00305
LINC00330
LINC00340
LINC00441
LINC00467
LINC00473
LINC00476
LINC00478
LINC00486
LINC00487
LINC00511
LINC00521
LINC00526
LINC00535
MIR210HG
MIR31HG
MIR4500HG
PLK1S1
PVT1
RBM26-AS1
SKINTL
SNHG12
SOX2-OT
TMEM72-AS1
HBV mt HBx
HBV WT
Mock
100
0
Log2 counts
f
HepG2
% INPUT
***
HBV WT
HBV mt HBx
ChIP anti-EZH2

Figure 5


Human Primary Hepatocytes (PHH)
HepG2
a
b
c
d
FI TRIM13
FI % INPUT
Mock
HBV day4
HBV 24h
ChiP anti-HBx
ChiP anti-Pol II
Mock GAPMER
HBV CTR
HBV GAPMER
day 4 PI
HBV WT
HBV WT GAPMER
ChIP anti-AcH4

e
mRNA correlation (RNA Seq V2 RSEM)
all patients (373 patients)
Hepatitis B (76 patients)
Hepatitis C (32 patients)
Alcohol Consumption (68 patients)
TRIM13, mRNA expression (log2)
DLEU2, mRNA expression (log2)
Pearson r=0.21
p=5.5e-0.5
Pearson r=0.30
p=7.8e-0.3
Pearson r=0.11
p=5.5e-0.1
Pearson r=0.17
p=1.6e-01

Figure 6


HepG2
a
FI % INPUT
Mock
HBV WT
ChiRP anti-DLEU2
b
FI % INPUT
Mock
HBV WT
HBV mt HBx
ChiP anti-EZH2
ChiP anti-H3meK27
c
FI TRIM13
DMSO
DZNEP
HepG2
d
% INPUT
Mock
HBV WT
ChiP anti-HBx
e
% INPUT
Mock
HBV WT
ChiRP anti-DLEU2
mRNA correlation (RNA Seq V2 RSEM)
f
all patients (373 patients)
Hepatitis B (76 patients)
CCNB2, mRNA expression (log2)
DLEU2, mRNA expression (log2)
Pearson r=0.44
p=4.7e-19
Pearson r=0.40
P=2.8e-0.4
g
HBV
Suz12
EED
EZH2
DLEU2
HBx
Repressed state
Active state
Hyper-activated state
K27me3
K27me3
H3 tail
Suz12
EED
EZH2
PRC2
HBx
DLEU2

Legends to Figures

Figure 1. HBx modulates DLEU2 expression. (a) Heatmap of 34 lncRNAs in Human Primary Hepatocytes (PHH), potentially deregulated following HBV infection, by Nanostring Digital Counter. The blue and yellow colors in the heatmap indicate down- and up-regulated lncRNAs respectively. Normalized on GAPDH and respect to mock condition. (b) Heatmap of 15 lncRNAs known in literature to be modulated in HCC identified from NanoString analysis PHH. Normalized on GAPDH and respect to mock condition. (c) DLEU2 lncRNA genomic sequence, according with Genome Browser; it is also indicated HBx Peak (black circle) obtained by HBx ChIP Seq searching[12]. (d) HBx and Pol II occupancy on DLEU2 promoter region was validated in three independent ChIP experiments. Cross-linked chromatin from PHH not infected or infected, respectively mock and HBV 24h, was immunoprecipitated with a specific anti-HBx and anti-Pol II antibody or relevant IgG controls, and then analyzed by real-time qPCR using specific DLEU2 promoter primer pairs. ChIP results are expressed as Fold Induction (FI) of the % of Input respect to mock and the histograms show the mean from three independent experiments; bars indicate SD. (e) Real Time qPCR of DLEU2 mRNA in mock and 4 days HBV-infected PHH. Results are expressed as FI relative to the mock after normalization to endogenous human β-ACTIN mRNA. Data represent means ± SD from at least three independent experiments performed in duplicate. (f) PCR, and relative densitometry of the exons 2-4, intron2 of DLEU2 and β-ACTIN in mock and 7 days HBV-infected PHH. Data from 1 representative experiment out of 3 are shown. (g) PCR, and relative densitometry, of the DLEU2 exons 2-4, intron 2 and β-ACTIN in mock and HBV WT, nuclear fraction (N), of HepG2 cell lines. Data from 1 representative experiment out of 3 are shown.

Figure 2. DLEU2 lncRNA interacts with PRC2 complex. (**a**) RNA binding protein immunoprecipitation (RIP). The lysate of mock and 72 hours HBV-infected PHH were applied to immunoprecipitation with antibodies against IgG or EZH2. DLEU2 exons 2-4, intron 2 and MALAT1 lncRNA were detected by PCR assay and analyzed by relative densitometry. Data from 1 representative experiment out of 3 are shown. (b) RIP. The lysate of mock and 72 hours HBV-

infected PHH were applied to immunoprecipitation with antibodies against IgG or HBx. DLEU2 exons 2-4, intron 2 and MALAT1 were detected by PCR assay and analyzed by relative densitometry. Data from 1 representative experiment out of 3 are shown. (**c**) HepG2 HBV WT-replicating cell lysates were subjected to RIP using IgG control or Anti-HBx antibodies. DLEU2 Exons 2-4, intron2 and MALAT1 were detected by PCR assay and analyzed by relative densitometry.

Figure 3. HBx and EZH2 bind DLEU2 RNA (a) Docked configurations of HBx (surface representation, orange) and EZH2 (surface representation, mauve) on tertiary structure of intron 2 (green). 1 and 2 panels present the two most probable loci for HBx adsorption on intron 2. 3 and 4 panels present the two most probable loci for EZH2 adsorption on intron 2. (b) 1-4 panels: Docked configurations of EZH2 on the complex HBx-intron 2. EZH2 is docked on the two most probable conformations of the HBx-intron 2 complex. 5-8 panels: Docked configurations of HBx on the complex EZH2-intron 2. HBx is docked on the two most probable conformations of the EZH2-intron 2 complex.

Figure 4. DLEU2 controls HBV replication (a) DLEU2-qChIRP for cccDNA using DLEU2-specific antisense DNA probe in mock and HBV WT HepG2 cell lines. Affinity-purified material was analyzed by real time PCR with DLEU2 RNA specific primers (lower panel). (b) Real Time PCR of DLEU2 in mock and HBV WT +/- Gapmer pools in PHH, NTCP and HepG2 replicating cells. Results are expressed as FI relative to the endogenous human β-Actin mRNAs. Data represent means ± SD from at least three independent experiments performed in duplicate. (c) H4 histone acetylation on cccDNA was validated in three independent ChIP experiments. Cross-linked chromatin from HBV WT and HBV WT with Gapmer pools HepG2 replicating cells was immunoprecipitated with a specific anti-AcH4 antibody or relevant IgG controls, and then analyzed by real-time qPCR using specific primer pairs. ChIP results are expressed as FI of the % of Input respect to HBV WT and the histograms show the mean from three independent experiments; bars indicate SD. (d) HBV pgRNA levels in 4 days HBV-infected PHH and in HepG2 and NTCP

replicating cells with Gapmer pools by real time PCR. Results are normalized to endogenous human β-Actin mRNAs. Data represent means ± SD from at least three independent experiments performed in duplicate (e) Heatmap of 34 lncRNAs in HepG2 cell lines deregulated following HBV WT or HBV mutant HBx (HBV mt HBx) transfection, by Nanostring Digital Counter. The colors in the Heatmap indicate the log2 counts normalized on GAPDH. (f) Assessment of EZH2 on cccDNA in HepG2 (HBV WT and HBV mutant HBx) -replicating cells transfected or not with specific Gapmer for DLEU2 RNA species. ChIP results are expressed as % of Input.

Figure 5. **DLEU2 and HBx are required for TRIM13 expression.** (a) Real Time PCR of TRIM13 mRNA in mock and 4 days HBV-infected PHH. Results are expressed as Fold Induction relative to the mock after normalization to endogenous human β-Actin mRNAs. Data represent means ± SD from at least three independent experiments performed in duplicate. (b) HBx occupancy and Polimerase II binding on TRIM13 promoter region by ChIP assay and analyzed as in 1d. (c) Real Time PCR of TRIM13 cDNA in mock or HBV WT infected PHH +/- Gapmer pools. Results are expressed as FI relative to the mock after normalization to endogenous human β-Actin mRNAs. Data represent means ± SD from at least three independent experiments performed in duplicate. (d) Assessment of H4 acetylation on the TRIM13 promoter in HBV-replicating HepG2 cells (HBV WT) transfected or not with specific Gapmers for DLEU2 RNA species. ChIP results are expressed as FI of the % of Input respect to mock and analyzed as in 4c. (e) Pearson correlation (r) of DLEU2 and TRIM13 transcript expression values among 373 HCC Patients (TCGA data).

Figure 6. HBx-DLEU2-EZH2 complex disables the direct epigenetic silencing to specific regions. (a) DLEU2-qChIRP for TRIM13 promoter using DLEU2-specific antisense DNA probe in mock and HBV WT HepG2 cell lines. (b) EZH2 occupancy and H3meK27 mark on TRIM13 promoter regions by ChIP assay in HBV-replicating HepG2 cells (HBV WT/ HBV mt HBx). ChIP results are expressed as FI of the % of Input respect to mock and analyzed as in 4f. (c) Real Time PCR of TRIM13 cDNA in DZNeP treated HepG2 cells respect to DMSO. Results are expressed as a FI relative to the mock after normalization to endogenous human β-Actin mRNAs. Data represent

means ± SD from at least three independent experiments performed in duplicate. (d) HBx occupancy on CCNB2 promoter region by ChIP assay and analyzed as in Figure 1d. (e) DLEU2-qChIRP for CCNB2 promoter using DLEU2-specific antisense DNA probe in mock and HBV WT HepG2 cell lines. (f) Pearson correlation (r) of DLEU2 and CCNB2 transcript expression values among 373 HCC Patients (TCGA data). (g) A model for DLEU2-PRC2/EZH2-HBx interaction on cell host genes. *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 the H3K27me3 repressive mark and transcriptional activation; *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 leads to a progresive loss of the repressive H3K27me3 mark and chromatin bound EZH2 that is associated with active constitutive transcription. HBx is both recruited at the chromatin levels to contribute to transcriptional activation and it is part of the DLEU2-PRC2/EZH2 complex.

SUPPLEMENTAL INFORMATION

Supplemental Information includes: Supplementary Material: Modeling and Docking section; six Supplementary figures; three Supplementary tables;

Supplementary Materials

Modeling and Docking

a) RNA structure modelling.

Experimental structures of lncRNAs are scarce[1] and methods to predict *de novo* RNA/ lncRNA structures are subject of intense and challenging investigation[2-3], at the edge of the RNA research.

The DLEU2 species (exons 1-10 and intron 2) were modeled starting from the FASTA sequences (https://genome.ucsc.edu/cgi-bin/hgGateway). We first obtained the secondary structures (Figure S3) via DotKnot[4-5], a hybrid heuristic algorithm, based on sequence matching and free energy minimization. The algorithm is on overall faster than methods based on free energy minimization only, that can become unaffordable for long sequences[6]. It performs a preliminary detection of pseudoknots, used in a second step during the secondary structure prediction. Pseudoknot structures (Figure S3) can become important for long RNAs, in particular in relationship with transcription and regulatory activities[7].

The FASTA sequence and the secondary structure have been then used to obtain the tertiary structure of the considered lncRNAs (Figure S4: exons 1-10; Figure 3: intron 2) via the RNAcomposer tool[8], a quite recent method, based on a machine translation principle. RNAcomposer has been tested for high resolution large RNA structure prediction (in reliability assessment of prediction methods for tertiary structure in case of lncRNAs, see Miao et al., 2015[3], starting from a user-defined secondary structure. It allows to model up to 500 residues, therefore it is suited for the large exons/intron here considered.

b) HBx structure modelling.

HBx is a small protein (154 aa), with subdomains devoted to different functions[9]. It is a partially disordered protein (hybrid type Intrinsically Unstructured Protein, IUP), as it contains some tens % 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 aa sequence lead to the formation of four disulfide bonds, which role is to stabilize and to control the protein function[10]. In this work, the HBx structure has been modelled by I-TASSER[11-12-13], 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 S5a); 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 S5b) based on the primary aa sequence, as given by IUPred[14], and PONDR[15], 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 S5c) from DISOPRED2[16], while our model is different from a previous 3D model[17], not including constraints on cysteines.

c) Sequence based interaction methods.

We estimated the interaction probabilities for the HBx-DLEU2 species with various tools based on primary sequences analysis. The method RPISeq[18] indicates that HBx preferentially interacts with the exons 1-4, 6. RPISeq results are compared, and are in reasonable agreement, with the ones from catRAPID[19], a method providing local information on residues more prone to interaction with nucleotides. However, prediction from catRAPID may be not fully reliable for DLEU2 species because the tool does not include complex RNA structures as pseudoknots.

HBx protein sequence has been analyzed with RNABindRPlus[20], which predicts RNA-binding residues based on the protein primary sequence. Residues with higher interacting propensity for RNA are listed in Table S1 (line 1), and 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.

d) Docking predictions.

HBx protein and RNAs models were docked by using three different tools, two for the preliminary screening and determination of the interface residues (HEX[21]; NPDock[22-23-24]), and one for a more refined docking (HADDOCK[25-26]). 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 use preliminary docking results from HEX and NPDock 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 (exon/intron 2 nucleotides) with at least one atom within 5 Å of exon/intron 2 nucleotides (protein residues) in the complexes produced by HEX/NPDock. These aminoacids (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.

As for the exons, in most cases HEX and NPDock provide the same adsorption site for HBx; in the cases of exons 2, 3, 6, 8, for which the HEX and NPDock predictions do not coincide, we used HADDOCK to refine the docking and find the most stable conformation. In three cases, the HEX configurations are found to be more stable than the ones produced by NPDock. In a fourth case a new configuration is found by HADDOCK, different from both the HEX and NPDock ones. The 3D structures of the HBx-exon final complexes provided by HADDOCK are shown in Figure S4.

Given this preliminary study, we defined, for the following dockings, a protocol based on a pre-screening of configurations by HEX, to be refined and confirmed by HADDOCK.

Docking of HBx on intron 2 has been at first carried out with HEX, by including also the electrostatic contributions in the scoring function, beside the volumetric terms. Interestingly two loci have been found on the intron 2 with nearly equal probability to be occupied by HBx (see Figure 3a, panels 1-2). The same docking protocol has been applied to investigate the interaction of the PRC2 protein EZH2 with intron 2. Also in this case, HEX calculations provide two loci on the intron 2 with nearly equal probability to be occupied by EZH2 (Figure 3a, panels 3-4). The most stable configurations of HBx (EZH2)-intron2 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.

While docking is scarcely used to predict interaction partners, it has been proven, with a high-throughput docking experiment, that different complexes have different overall signals in their docking scores[27], and that standard docking algorithms scores are able to distinguish real interactors from non-interacting cases. We use the score rankings from HEX and HADDOCK to evaluate the propensity of HBx to bind DLEU species. The ranking provided by HEX, even when based only on the volumetric part of the interaction[27], gives a good match with the ranking of *interaction probabilities* as given by sequence-based methods as RPISeq and with the HADDOCK final scores ranking.

Finally, HEX calculations have been carried out to investigate the ternary EHZ2-intron 2-HBX complexes. The most stable configurations of the proteins-intron 2 complexes are confirmed by HADDOCK. Results are shown in Figure 3b, panels 1-8.

e) Interface analysis

Based on docking results, information on the interaction between RNA species and HBx has been provided by investigating the protein-RNA interface at molecular level. The interaction interface is here defined by considering protein residues (nucleotides) with at least one atom within 5 Å of RNA nucleotides (protein residues). We report in table S1 results obtained by HADDOCK, compared to the sequence-based prediction obtained by RNABindRPlus. We observe a preferential adsorption of HBx on stems and double simple helices, while hairpins and pseudoknots are mostly not involved (see also Figure S3).

f) Protein-protein and protein-lncRNAs control docking

As our protocol includes prediction of structure for large lncRNAs, and docking of such structures with an intrinsically disordered protein, whose structure has also been modeled, we assessed the quality of both our HBx model and docking protocol by using some experimental information on other HBx complexes and RNA structures.

X-rays data on DDB1-HBx short stretch complex[28] (PDB entry: 3I7H) show that HBx interacts with

DBB1 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-intron 2 and in the HBx-exon 6 (the most computationally interacting exon) cases. 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 van Hemert et al., 2011[17].

We also relied on some control dockings performed using other RNAs, looking for a negative control case, by choosing a lncRNA (of known experimental structure) that in principle has no biological interaction with HBx. In particular, docking HBx to a portion of MALAT1, its triple helix (4PLX.pdb), returns an interaction score lower by the 25% with respect to HBx-intron 2 and HBx-exon 6 cases. This is a first case of negative control for HBx-RNA interactions.

On the other hand, given the strong interaction of HBx and DBB1, 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 that neither DBB1 nor the HBX-DBB1 complex significantly interact with the lncRNA, providing a second case of negative control for protein-lncRNA interaction.

In other words, our docking procedure is able to discriminate among real or just supposed protein-RNA interactors.

SUPPLEMENTARY REFERENCES

1. Liu F, Somarowthu S, Pyle AM. Visualizing the secondary and tertiary architectural domains of lncRNA RepA. *Nature Chemical Biology* **13**, 282–289 (2017).
2. Yan K, Arfat Y, Li D, Zhao F, Chen Z, Yin C, Sun Y, Hu L, Yang T, Qian A. Structure Prediction: New Insights into Decrypting Long Noncoding RNAs. *Int. J. Mol. Sci*. **17**, 132 (2016).

3. Miao Z, Adamiak RW, Blanchet MF, Boniecki M, Bujnicki JM, Chen SJ, Cheng C, Chojnowski G, Chou FC, Cordero P et al. RNA-Puzzles Round II: assessment of RNA structure prediction programs applied to three large RNA structures. *RNA* **21(6)**:1066-84 (2015).
4. Sperschneider J, Datta A. DotKnot: pseudoknot prediction using the probability dot plot under a refined energy model. *Nucleic Acids Research*. (2010).
5. Sperschneider J, Datta A, Wise MJ. Heuristic RNA pseudoknot prediction including intramolecular kissing hairpins. *RNA* **17(1)**: 27-38. (2011).
6. Sperschneider J, Datta A. KnotSeeker: heuristic pseudoknot detection in long RNA sequences. *RNA* **14** :630-640. (2008).
7. Brierley I, Pennell S, Gilbert RJC. Viral RNA pseudoknots: versatile motifs in gene expression and replication. *Nature Reviews Microbiology* **5**, 598–610 (2007).
8. Popenda M, Szachniuk M, Antczak M, Purzycka KJ, Lukasiak P, Bartol N, Blazewicz J, Adamiak RW. Automated 3D structure composition for large RNAs. *Nucleic Acids Research* **40**, No. 14 e112 (2012).
9. Lee S-H, Cha E-J, Lim J-E, Kwon S-H, Kim D-H, Cho H, Han K-H. Structural Characterization of an Intrinsically Unfolded Mini-HBX Protein from Hepatitis B Virus. *Mol. Cells* **34**, 165-169 (2012).
10. Sidhu K, Kumar S, Reddy VS, Kumar V. Mass spectrometric determination of disulfide bonds in the biologically active recombinant HBx protein of hepatitis B virus. *Biochemistry* **22** ;53(28):4685-95 (2014).
11. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nature Methods* **12** (1), 7-8 (2015).
12. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, **5** (4), 725 (2010).
13. Zhang Y. I-TASSER server for protein 3D structure prediction, *BMC Bioinformatics* **9**:40 (2008).

14. Dosztányi Z, Csizmok V, Tompa P, Simon I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics Applications Note* **21** no. 16, pages 3433–3434, (2005).
15. Xue B, Dunbrack RL, Williams RW, Dunkera AK, Uversky VN. PONDR-FIT: A Meta-Predictor of Intrinsically Disordered Amino Acids, *Biochim Biophys Acta* **1804** (4): 996–1010 (2010).
16. Ward JJ, McGuffin LJ, Bryson K, Buxton BF, Jones DT. The DISOPRED server for the prediction of protein disorder, *Bioinformatics Applications Note* **20**, 2138–2139 (2004).
17. van Hemert FJ, van de Klundert MAA, Lukashov VV, Kootstra NA, Berkhout B, Zaaijer HL. Protein X of Hepatitis B Virus: Origin and Structure Similarity with the Central Domain of DNA Glycosylase. *PLoS ONE* **6** (8): e23392 (2011).
18. Muppirala UK, Honavar VG, Dobbs D. Predicting RNA-Protein Interactions Using Only Sequence Information. *BMC Bioinformatics* **12** :489 (2011).
19. Bellucci M, Agostini F, Masin M, Tartaglia GG. Predicting protein associations with long noncoding RNAs. *Nature Methods* **8**(6), 445 (2011).
20. Walia RR, Xue LC, Wilkins K, El-Manzalawy Y, Dobbs D, Honavar V. 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*, **9**(5) e9772 (2014).
21. Ritchie DW, Kozakov D, Vajda S. Accelerating and focusing protein–protein docking correlations using multi-dimensional rotational FFT generating functions. *Bioinformatics Original Paper* **24** no. 17, pages 1865–1873 (2008).
22. Puton T, Kozlowski L, Tuszynska I, Rother K, Bujnicki JM. Computational methods for prediction of protein-RNA interactions. *J Struct Biol*. **179** (3):261-8 (2012).
23. Tuszynska I, Matelska D, Magnus M, Chojnowski G, Kasprzak JM, Kozlowski LP, Dunin-Horkawicz S, Bujnicki JM. Computational modeling of protein-RNA complex structures. *Methods* **65**(3):310-9 (2014).

24. Tuszynskal, Magnus M, Jonak K, Dawson W, Bujnicki JM. NPDock: a web server for protein–nucleic acid docking. *Nucleic Acids Research* **43**, Web Server issue W425–W430 (2015).

25. Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **125**, 1731-1737(2003).

26. van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ, Bonvin AMJJ. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol* **428**, 720–725 (2016).

27. Wass MN, Fuentes G, Pons C, Pazos F, Valencia A. Towards the prediction of protein interaction partners using physical docking. *Molecular Systems Biology* **7**:469, 3 (2011).

28. Minor MM, Slagle BL. Hepatitis B virus HBx protein interactions with the ubiquitin proteasome system. *Viruses* **6**, 4683-4702 (2014).

29. Yang Y, Chen L, Gu J, Zhang H, Yuan J, Lian Q, Lv G, Wang S, Wu Y, Yang YT et al. Recurrently deregulated lncRNAs in hepatocellular carcinoma. *Nat Commun*. **8**:14421. (2017).

Figure S1

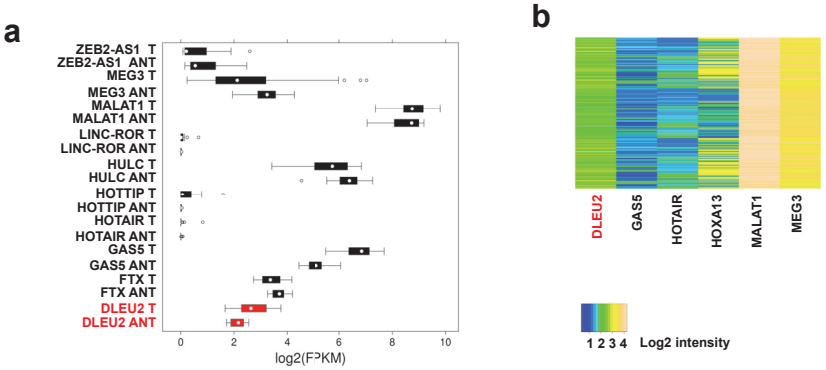


Figure S2

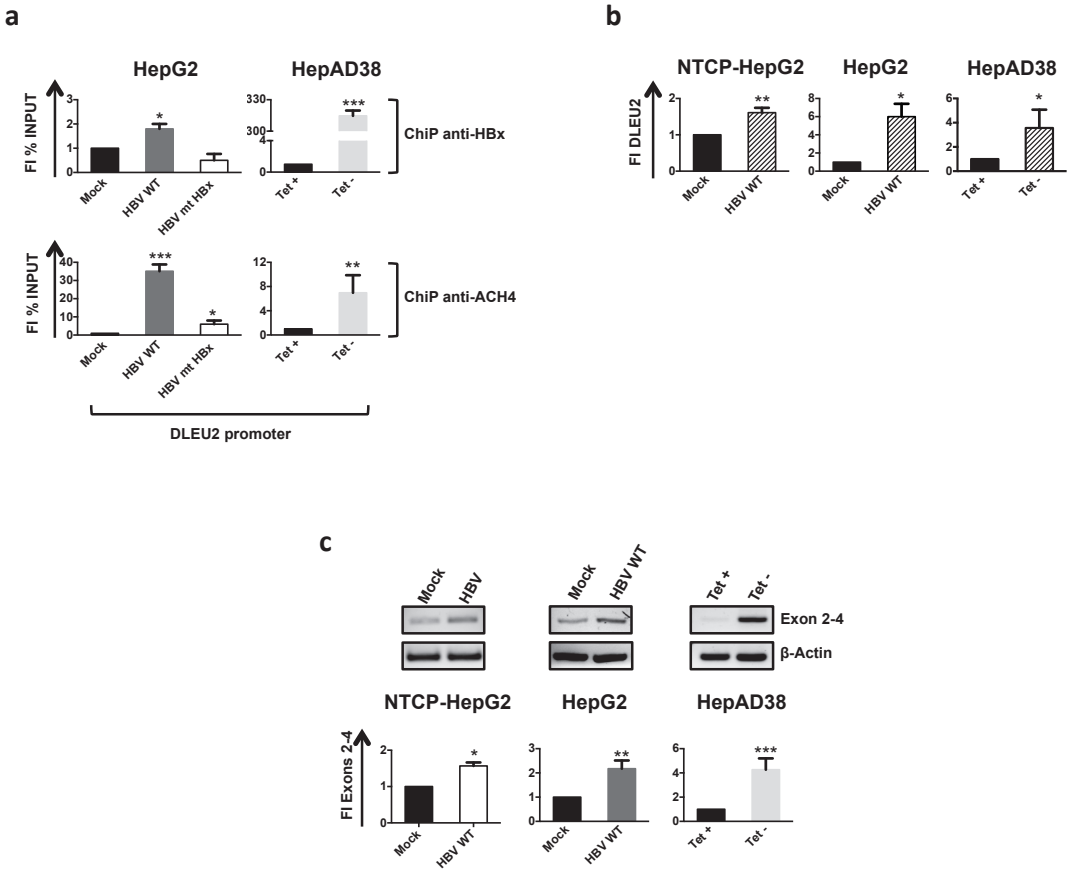

a
HepG2
HepAD38
Fl % INPUT
ChiP anti-HBx
Mock
HBV WT
HBV mt HBx
Tet +
Tet -
ChiP anti-ACH4
DLEU2 promoter
b
NTCP-HepG2
HepG2
HepAD38
Fl DLEU2
c
Mock
HBV
HBV WT
Tet +
Tet -
Exon 2-4
β-Actin
NTCP-HepG2
HepG2
HepAD38
Fl Exons 2-4

Figure S3

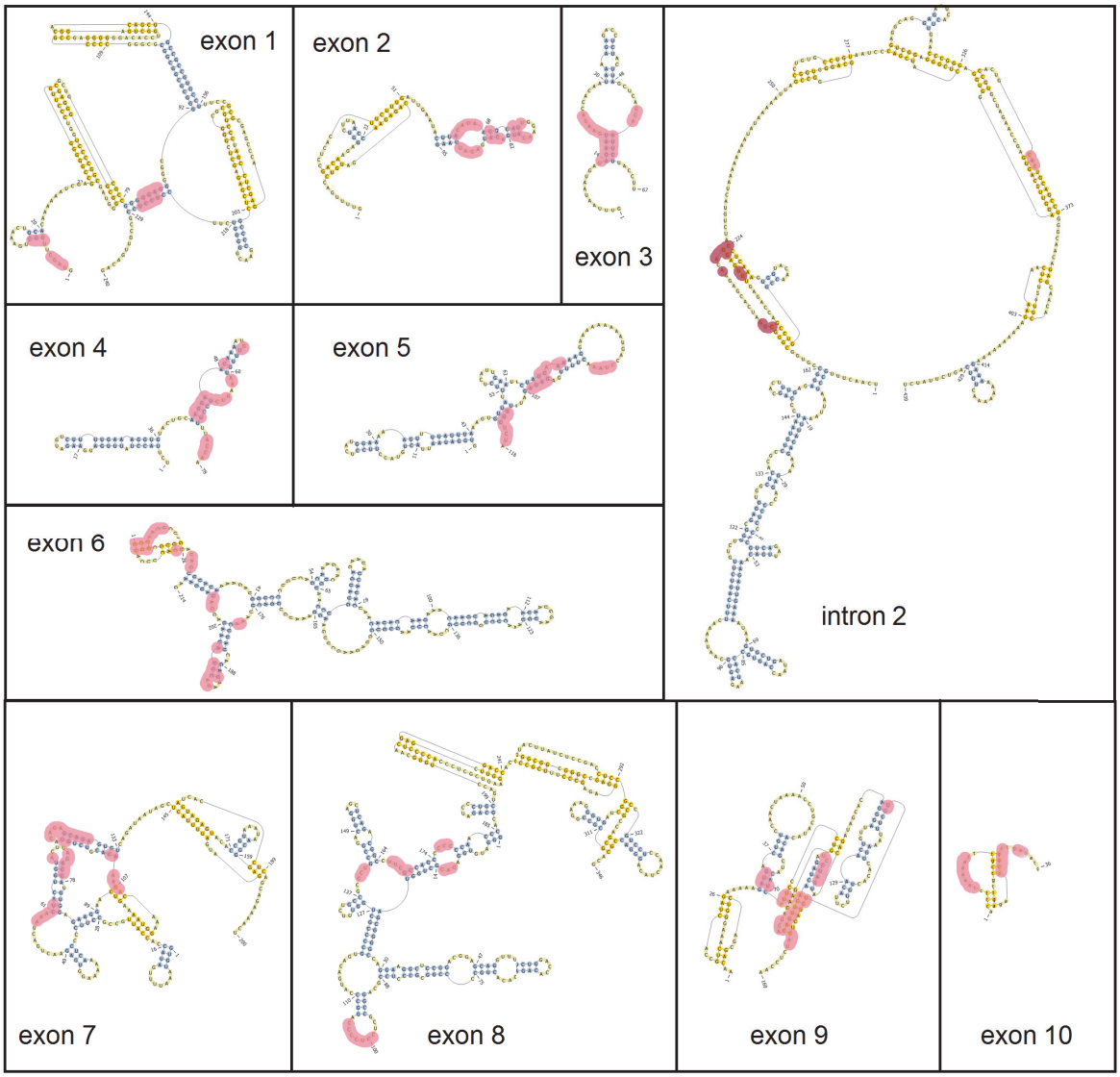

exon 1
exon 2
exon 3
exon 4
exon 5
exon 6
intron 2
exon 7
exon 8
exon 9
exon 10

Figure S4

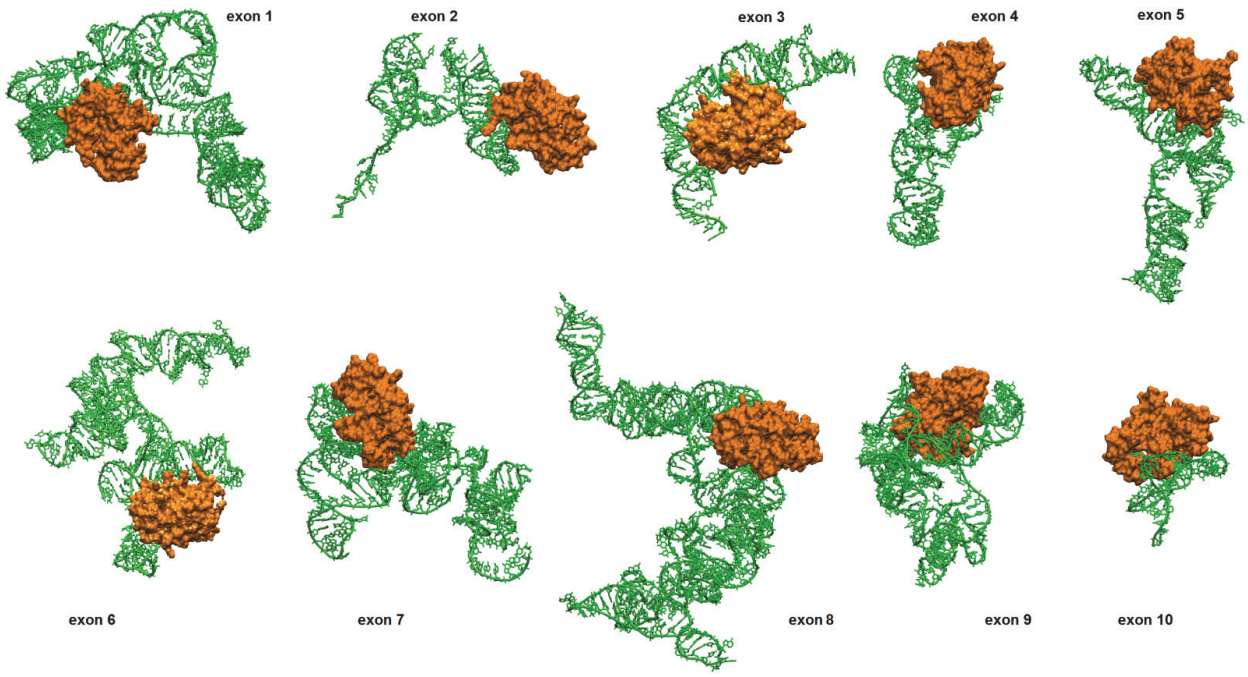

exon 1
exon 2
exon 3
exon 4
exon 5
exon 6
exon 7
exon 8
exon 9
exon 10

Figure S5

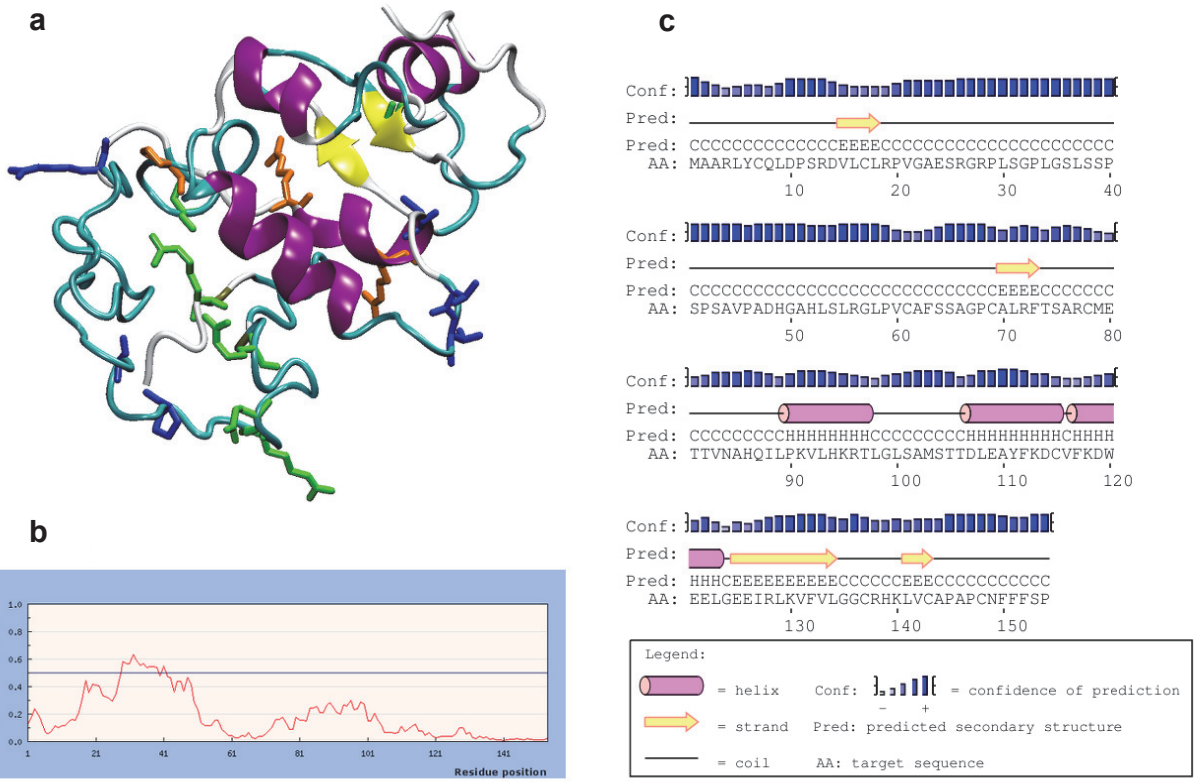

a
b
c
Disorder tendency
Residue position
Conf:
Pred:
Pred: CCCCCCCCCCCCCCCEEEECCCCCCCCCCCCCCCCCCCCC
AA: MAARLYCQLDPSRDVLCLRPVGAESRGRPLSGPLGSLSSP
10 20 30 40
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCEEEECCCCCCCC
AA: SPSAVPADHGAHLSLRGLPVCAFSSAGPCALRFTSARCME
50 60 70 80
Pred: CCCCCCCCCCHHHHHHHHCCCCCCCCCCHHHHHHHHHCHHHH
AA: TTVNAHQILPKVLHKRTLGLSAMSTTDLEAYFKDCVFKDW
90 100 110 120
Pred: HHHCEEEEEEEEEECCCCCCCEEECCCCCCCCCCCC
AA: EELGEEIRLKVFVLGGCRHKLVCAPAPCNFFFSP
130 140 150
Legend:
= helix
= strand
= coil
Conf:
= confidence of prediction
Pred: predicted secondary structure
AA: target sequence

Figure S6

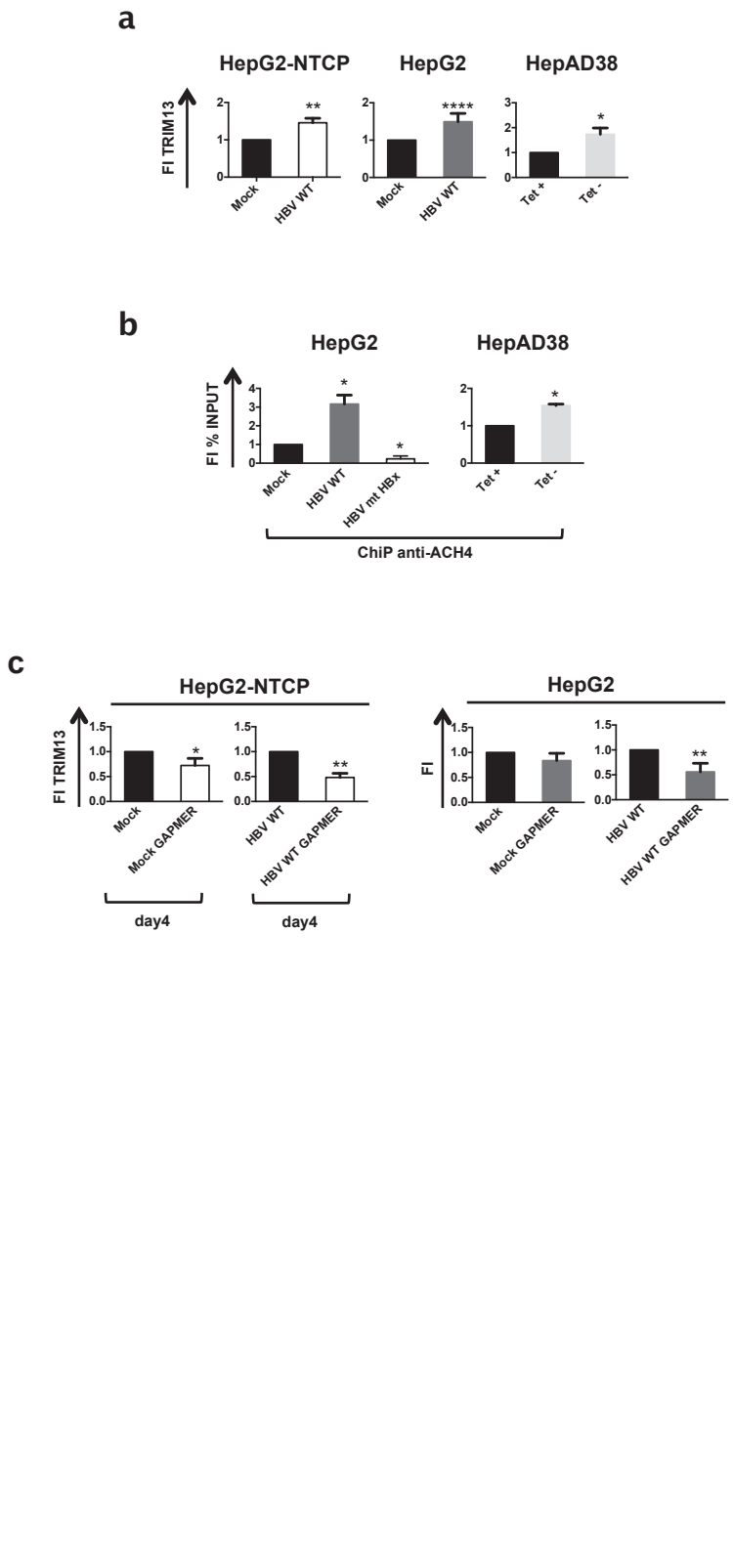

a
HepG2-NTCP
HepG2
HepAD38
FI TRIM13
Mock
HBV WT
Tet +
Tet -
b
HepG2
HepAD38
FI % INPUT
Mock
HBV WT
HBV mt HBx
ChiP anti-ACH4
c
HepG2-NTCP
HepG2
FI TRIM13
FI
Mock
Mock GAPMER
HBV WT
HBV WT GAPMER
day4

Table S1

HBx residues mostly interacting with RNA

T	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
1	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
2	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
3	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
4	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
5	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
6	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
7	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
8	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
9	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
10	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
l2-a	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
l2-b	M	A	A	R	L	Y	C	Q	L	D	P	S	R	D	V	L	C	L	R	P	V	G	A	E	S	R	G	R	P	L	S	G	P	L	G	S	L	S	S	P
	1									10										20									30									40

T	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
1	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
2	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
3	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
4	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
5	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
6	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
7	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
8	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
9	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
10	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
l2-a	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
l2-b	S	P	S	A	V	P	A	D	H	G	A	H	L	S	L	R	G	L	P	V	C	A	F	S	S	A	G	P	C	A	L	R	T	F	S	A	R	C	M	E
										50										60									70									80

T	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
1	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
2	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
3	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
4	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
5	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
6	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
7	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
8	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
9	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
10	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
l2-a	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
l2-b	T	T	V	N	A	H	Q	I	L	P	K	V	L	H	K	R	T	L	G	L	S	A	M	S	T	T	D	L	E	A	Y	F	K	D	C	V	F	K	D	W
										90										100									110									120

T	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
1	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
2	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
3	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
4	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
5	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
6	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
7	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
8	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
9	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
10	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
l2-a	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
l2-b	E	E	L	G	E	E	I	R	L	K	V	F	V	L	G	G	C	R	H	K	L	V	C	A	P	A	P	C	N	F	F	F	S	P
										130										140									150									

Table S3

Primers and probes sequences

ChIP primers	Forward (5'-3')	Reverse (5'-3')
DLEU2 promoter	GCTGATAACCAGTGCCACTAA	CCTCTCAAAGTGCTGGGATTA
TRIM13 promoter	ACCCAAACTTCCTCAACTGG	GGAATGGCTCCTCCAGAATTTA
PCR primers	**Forward (5'-3')**	**Reverse (5'-3')**
DLEU2 exon 2-4	ACCTGTAGCAGAGAACCAATT	TTCCTTGCAGTACACCTTTCA
DLEU2 intron2	GCCAAACATGCTAACCAACTC	ATGCCACTACATCCTGCTTATC
28S	GGCGAAGCCAGAGGAAACT	GACGACCGATTTGCACGTC
MALAT1	AGGCGTTGTGCGTAGAGGA	GGATTTTTACCAACCACTCGC
β-Actin	GCACTCTTCCAGCCTTCC	AGGTCTTTGCGGATGTCC
BECLIN	GCAGCTGGATAAGCTGAAGA	CGACCCAGCCTGAAGTTATT
Real Time probes	**Probe**	
DLEU2	Hs00863925_m1 (Applied)	
EZH2	Hs00544833_m1 (Applied)	
TRIM13	Hs00328634_s1 (Applied)	
β-Actin	cat.no. 05532957001 (Roche)	
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	

Legends to Supplementary Figures and Tables

Figure S1. DLEU2 expression in hepatocellular carcinoma (a) lncRNAs expression from RNA-seq data ([29]), from 60 patients tumor and adjacent normal tissues. T= Tumor. ANT= Adjacent Normal Tissues. **(b)** Expression of selected lncRNAs on HCC (81 patients) microarray dataset (Villa et al. 2016).

Figure S2. HBx modulates DLEU2 expression in HepG2, NTCP and HepAD38 cell lines. (a) HBx occupancy (upper panel) and H4 histone acetylation (lower panel) on the promoter region of DLEU2. Cross-linked chromatin from mock, HBV wt or HBV HBx mutant replicating HepG2 and AD38 cells was immuno-precipitated with a specific anti-HBx and anti-AcH4 antibody or relevant IgG controls, and then analyzed by real-time qPCR using specific primer pairs. ChIP results are expressed as Fold Induction of the % of Input respect to Mock and the histograms show the mean from three independent experiments; bars indicate SD. **(b)** Real Time PCR of DLEU2 mRNA in Mock and HBV WT in HepG2, HepAD38 and NTCP cells. Results are expressed as Fold Induction relative to the mock after normalization to endogenous human β-Actin mRNAs. Data represent means ± SD from at least three independent experiments performed in duplicate. **(c)** PCR, and relative densitometry of 2-4 exon and β-Actin in Mock and HBV WT in HepG2, AD38 and NTCP cells. Data from 1 representative experiment out of 3 are shown. For all experiments, unpaired t-test was performed: *p≤0.05 **p≤0.005, ***p ≤ 0.0005.

Figure S3. DLEU2 RNA secondary structure. Secondary structure of exons 1-10 and intron 2 of DLEU2, predicted with DotKnot and plotted with Pseudowiever. Pseudoknots are highlighted in bright yellow. Parts interacting with HBx (from docking results) are highlighted as light red zones. The two adsorption loci on intron2 are highlighted as light red and bright red zones. Exon 11 has not been modelled, because its sequence lenght is beyond the current limits of prediction tools. Intron 2 has also a size lenght beyond current computational limits, but, based on chip seq HBx

data, we could select as interacting the central region of the sequence, basis 280-718 (439 nucleotides).

Figure S4. Hbx binds RNA. Docked configurations of HBx (surface representation, orange) on the tertiary structure of exons 1-10 (green). The 3D exons structures are built with RNAcomposer, by using FASTA sequence and secondary structure. All considered exons are below the 500 residues size limit of RNAcomposer.

Figure S5. HBx protein. (a) HBx tertiary structure, with color highlighted secondary structure, modelled from I-Tasser, by using default parameters values for the homology modeling. Constraints on four disulfide bonds ([28, 10]) 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 ([17]). Residues involved in interaction with RNA are depicted in the same color code used in Table S1, according to RNABindRPlus results. The 3D model is in good agreement with both the **(b)** IUPRED results on the disordered regions, and the **(c)** DISOPRED results on the secondary structure.

Figure S6. TRIM13 expression in HBV-replicating hepatic cell lines. (a) Real Time PCR of TRIM13 mRNA in Mock and HBV WT in HepG2-NTCP, HepG2 and HepAD38 cell lines. Results are expressed as fold induction relative to the mock after normalization to endogenous human β-Actin mRNAs. Data represent means ± SD from at least three independent experiments performed in duplicate. For all experiments, unpaired t-test was performed: *p≤0.05 **p≤0.005, ***p≤0.0005, ****p≤0.00005. **(b)** H4 acetylation on TRIM13 promoter regions cross-linked chromatin from mock, HBV wt or HBV HBx mutant replicating HepG2 and AD38 cells was immunoprecipitated with anti-AcH4 antibody or relevant IgG controls, and then analyzed by real-time qPCR using specific primer pairs. ChIP results are expressed as Fold Induction of the % of Input respect to Mock and analyzed as in 4c. **(c)** Real Time PCR of TRIM13 in HBV wt +/- Gapmer in HepG2-NTCP and HepG2 replicating cells. Results are expressed as fold induction relative to the Mock after

normalization to endogenous human β-Actin mRNAs. Data represent means ± SD from at least three independent experiments performed in duplicate. For all experiments, unpaired t-test was performed: *p≤0.05 **p≤0.005.

Etude 3

Identification of chromatin-accessible domains on host genome and HBV genome in HBV infected primary human hepatocytes.

Océane Floriot, Giuseppe Pascucci, Matteo Pallocca, Francesca De Nicola, Michel Rivoire, Marie-Laure Plissonnier, Mirjam Zeisel, Massimo Levrero, Francesca Guerrieri

En préparation

L'ampleur de l'impact de HBV sur le transcriptome cellulaire reste controversée. Plusieurs études ont montré peu de modifications de la transcription suite à une infection aiguë par le HBV dans des modèles animaux (Wieland et al., 2004) ou dans des hépatocytes primaires infectées *in vitro* (Niu et al., 2017; Suslov et al., 2018). En revanche, d'autres études montrent clairement que le virus a un impact sur la transcription des cellules hôtes (Ancey et al., 2015; Lamontagne et al., 2016).

L'activation de la transcription dans les eucaryotes a été liée à la perturbation de l'organisation des nucléosomes au niveau des promoteurs, des *enhancers*, des *silencers* et des *insulators* en raison de la liaison des facteurs de transcription. Ainsi, les séquences régulatrices ont tendance à coïncider avec les sites génomiques ouverts ou accessibles de la chromatine remodelée. Des tests actuels d'accessibilité de la chromatine sont utilisés pour séparer le génome par des moyens enzymatiques ou chimiques et pour isoler les régions accessibles ou protégées. L'ADN isolé est ensuite quantifié à l'aide d'une plate-forme de séquençage de nouvelle génération. Ces tests se sont révélés utiles pour identifier les événements de remodelage de la chromatine et les changements épigénétiques qui se traduisent par une expression génique différentielle, une prolifération cellulaire et le développement de certaines maladies.

La technique d'ATAC-Seq (assay for transposase-accessible chromatin suivie d'un séquençage à haut débit) est la méthode la plus courante d'étude de la chromatine accessible, basée sur la capacité de la transposase Tn5 à fragmenter l'ADN et à s'intégrer *in vivo* dans des régions régulatrices actives. L'ATAC-Seq analyse la chromatine accessible ouverte, le positionnement des nucléosomes et les empreintes de TF au niveau du génome entier. L'ATAC-seq n'implique aucune modification chimique, est simple et rapide avec un protocole en deux étapes, est très sensible et nécessite donc un petit nombre de cellules (500 à 50 000 cellules).

Nous avons utilisé l'ATAC-seq pour détecter les changements précoces d'accessibilité de la chromatine dans des hépatocytes primaires humains infecté par le HBV et pour les corréler avec des données transcriptomiques d'ARN-Seq.

Le nombre de sites génomiques qui modifient fortement leur accessibilité à la chromatine augmente avec le temps d'infection (278 à 2 heures, 350 à 24 heures et 1095 at 72 heures), avec une prévalence de régions plus ouvertes (potentiellement actives) que de régions fermées (potentiellement réprimés) et environ 30% dencRNA dans toutes les conditions. Les régions accessibles de la chromatine sont enrichies pour les motifs de liaison des facteurs de transcription SMAD, STAT3, TP53 et de Suz12, un partenaire de EZH2 dans le complexe PRC2.

L'analyse fonctionnelle des gènes présents dans les régions avec des changements d’accessibilité révéle un enrichissement en gènes / ncRNA impliqués dans le cancer du foie, la transcription, le métabolisme du calcium et la voie de signalisation du TGF-b.

Manuscript in preparation

Identification of chromatin-accessible domains on host genome and HBV genome in HBV infected primary human hepatocytes

Océane Floriot[1#], Giuseppe Pascucci[2#], Matteo Pallocca[3], Francesca De Nicola[4], Michel Rivoire[5], Marie-Laure Plissonnier[1], Mirjam B. Zeisel[1], Massimo Levrero[1,2*], Francesca Guerrieri[1,3*]

[1] Cancer Research Center of Lyon (CRCL), INSERM U1052, Lyon, 69008, France

[2] Center for Life Nano Science, Istituto Italiano di Tecnologia, Rome, Italy

[3] UOSD SAFU, Regina Elena National Cancer Institute, Rome

[4] Oncogenomic and Epigenetic Unit, Regina Elena National Cancer Institute, Rome

[5] INSERM U1032, Centre de lutte contre le cancer Léon Bérard (CLB), Lyon, 69008, France

OF and GP are co-first authors

* ML and FG are co-corresponding authors

Authors Contribution

OF and GP share co-first authorship. FG supervised the study. OF, FDN, MLP, MBZ and FG performed the experiments and analyzed the data. GP and MP performed bioinformatic analysis. FG and ML conceived and designed the study. FG and ML wrote the manuscript.

Declaration of Interests

The authors disclose no conflicts.

Acknowledgments

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 Agence National de la Recherche (ANR@TRACTION) to ML; from the EU project 667273 HEP-CAR to ML.

Abstract

Hepatitis B virus (HBV) remains a major health problem worldwide despite the availability of an efficient vaccine, with 250 million peoples chronically infected and at risk to develop liver cirrhosis and hepatocellular carcinoma (HCC). A complex host-virus interaction is responsible for both the dysfunction of HBV specific B/T cells and the persistence of the cccDNA viral mini-chromosome in the infected hepatocytes, the two key challenges for HBV cure. The extent of HBV impact on the cellular transcriptome remains controversial and led to the commonly accepted concept of HBV being a "stealth" virus that induces no or very few changes in the host genome. We used ATAC-seq to detect early changes in chromatin accessibility in HBV-infected primary human hepatocytes (PHHs). We show that after HBV infection an increasing number of genomic sites change their chromatin accessibility over time (278 at 2 hours, 350 at 24 hours and 1095 at 72 hours), with a prevalence of more open potentially transcriptionally active regions and about 30% of ncRNA. RNA-seq analysis performed at the same times post infection showed an increasing number of differentially expressed genes (DEGs). The accessible chromatin regions are enriched for the binding motifs of the transcription factors SMADs, STAT3, TP53 and Suz12, a partner of Ezh2 in the PRC2 complex. Functional analysis of genes neighboring the regions with accessibility changes revealed enrichment in genes / ncRNAs involved in liver cancer, transcription, calcium metabolism and TGF-β signaling. Altogether these results show that HBV infection impacts on host cells chromatin landscape and transcription.

Introduction

Approximately 250 million individuals have chronic hepatitis B (CHB) and more than 780,000 people die each year due to HBV-associated liver diseases, such as cirrhosis and hepatocellular carcinoma (HCC) (WHO Global Hepatitis Report 2017; Schweitzer et al., 2015, Glogal Burden of Diseases, 2016). Multiple nucleos(t)ide analogs as well as pegylated interferon-α are approved for the treatment of CHB, but since these therapies rarely lead to cure (Zoulim et al., 2016; EASL CPG 2017) there is an urgent need to develop novel antiviral therapies. The key challenges for HBV cure are to overcome the dysfunction of HBV specific B/T cells and to target the pool of cccDNA, the viral mini-chromosome responsible for viral persistence in the nucleus of infected hepatocytes (Testoni et al., 2017). Several new antivirals and immuno-modulatory compounds have reached preclinical and/or early clinical evaluation with the aim of silencing cccDNA and/or reducing the size of the cccDNA pool in order to achieve *functional cure* with finite treatment duration (Testoni et al., 2017; Lok et al., 2017), the ultimate goal being the elimination or silencing of cccDNA.

Lack or low sensing of HBV by the innate immune system and HBV ability to inhibit signaling through pattern recognition receptor (PRRs) pathways likely contribute to deficient immune responses in CHB (Maini & Gehring, 2016). The little induction of the innate immune response following acute HBV infection in chimpanzees (Wieland et al., 2004), in chimeric mice carrying a humanized liver (Giersch et al, 2015), in woodchucks acutely infected with the woodchuck hepatitis virus (WHV) (Fletcher et al., 2013) as well as in HBV infected primary human hepatocytes *in vitro* (Niu et al., 2017) and in fresh liver biopsies *ex vivo* (Suslov et al., 2018) has led to the widely accepted concept that HBV is a "stealth" virus that induces no or very few changes in the host cell (Wieland and Chisari, 2005; Durantel and Zoulim, 2009). However, other studies have shown that HBV is detected by innate immunity sensors (Luangsay et al., 2015; Sato et al., 2015; Yoneda et al., 2016; Lucifora et al., 2018) and it has an impact on the transcription of host cells (Ancey et al., 2015, Lamontagne et al., 2016, Hösel et al., 2017). Moreover, HBV has a profound impact on the genes of innate immunity in the liver of chronically infected patients during the different phases of the disease (Lebossé et al., 2017) and DNA methylation profiles are altered in chronically infected patients (Jin et al., 2018). In these studies, however, the patients were in a chronic situation with many years of intrahepatic inflammation and ongoing viral replication and it is therefore difficult or impossible to identify the direct contribution of the virus. Thus, the extent of the impact of HBV on the cellular transcriptome remains controversial.

Transcriptional activation in eukaryotic cells has been tightly linked with disruption of nucleosome organization at promoters, enhancers, silencers, insulators and locus control

regions due to transcription factor binding. Regulatory DNA tends to coincide with open or accessible genomic sites of remodeled chromatin. Current genomic chromatin accessibility assays, that include DNase-seq, FAIRE-seq and ATAC-seq, are used to fragment the genome by enzymatic or chemical means and isolate either the accessible or protected locations (Tsompana and Buck, 2014). The isolated DNA is then quantified using a next-generation sequencing platform. These assays have been applied to identify the epigenetic changes responsible for differential gene expression, cell proliferation, functional diversification and disease development (De Dieleveult et al., 2016; Dechassa et al., 2018; Jonhson et al., 2018; Wang et al., 2018). ATAC-seq (assay for transposase-accessible chromatin followed by high throughput sequencing) is the most current method for probing open chromatin, based on the ability of hyperactive Tn5 transposase (Adey at al., 2010) to fragment DNA and integrate into active regulatory regions *in vivo* (Adey et al., 2010; Buenrostro et al., 2013). Differently from other chromatin accessibility techniques, ATAC-seq does not involve any chemical modification allowing to study simultaneously and at high resolution multiple aspects of chromatin architecture and remodeling, including the detection of open chromatin, nucleosome positioning and TF fingerprints on the whole genome using a small number (500 to 50,000) of cells.

Here we applied ATAC-seq to detect early changes in chromatin accessibility in HBV-infected PHHs. We showed that the number of genomic sites with altered chromatin accessibility increases with the time of infection, reaching >1000 at day 3 post-infection, with a prevalence of open, potentially transcriptionally active regions at all time points. Motif discovery and pathway analysis showed an increase in SMADs, STAT3, TP53 and Suz12 putative binding sites and an enrichment in genes / ncRNAs involved in liver cancer, transcription, calcium metabolism and the TGF-β signaling at the chromatin-accessible regions.

Material and methods

Primary culture of human hepatocytes. 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 Le Cluyse & Alexandre (2010) with minor modifications. Liver samples are 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 are seeded at 2,5x10^5 cells per cm² on collagen type IV (Corning, #354236) pretreated plates 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). PHH are then extensively washed in serum-free medium, kept in serum-free medium for 24 hours 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 hours before infection and then incubated for 16 h with an HBV inoculum produced in HepAD38 cells at a multiplicity of infection of 250 vge (virus genome equivalents) per cell, in presence of 4% PEG-8000 (Polyethylene glycol, Sigma–Aldrich, #1546605).

HBV viral inoculum preparation and infection. The HBV inoculum used in this study was prepared from HepAD38 cells as previously described (Ladner et al., 1997). 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. HepAD38 cells were cultured in HYPER*Flask*® (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 at 3-4 months, clarified first 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 hour and the pellets resuspended in Opti-MEM (Invitrogen,

#31985070) to achieve a 50 to 100 fold concentration. After DNA extraction (QIAmp Ultrasens Virus kit, Qiagen) the HBV inoculum was tittered by qPCR using serial dilutions of an HBV plasmid to build a standard curve. Primers are detailed in Supplementary Figure 1. All viral preparations were tested for the presence of endotoxins (Lonza Verviers, Belgium).

Total DNA/RNA extraction and reverse transcription. Total cellular DNA was extract using the MasterPure™ Complete DNA/RNA Purification Kit (Epicentre, #MC85200) according to manufacturer's recommendations. Total cellular RNA was extracted with the Extract-All reagent (Eurobio, #GEXEXT040U) following the manufacturer's protocol. Extracted RNA was treated for 1 h at 37 °C with RNAse-free DNase I (Invitrogen, #AM2222), which was subsequently inactivated at 70 °C for 20 min. RNA reverse transcription was performed using the Superscript® III RT kit (Life Technologies, #11752250), according to manufacturer's instructions.

Quantitative polymerase chain reaction (qPCR). Quantitative PCR was performed using SYBR-Green mix (Power Up Master Mix, Life Technologies, #A25742) with 10 nanograms of cDNA or total cellular DNA for the quantification of total HBV RNAs and HBV DNA, respectively. Following cycling conditions were used: a denaturation step for 20 s at 95 °C; and then 40 cycles of 3 s at 95 °C, 30 s at 60 °C. HBV DNA or messenger RNA (mRNA) relative amount was normalized over the expression of housekeeping gene *RPLP0* for DNA and *GUS* for mRNA (primers are detailed in Supplementary Table 1).
Before cccDNA quantification, 500 ng of total DNA were digested for 6 h, at 37 °C, with a Plasmid-safe DNase (Epicentre, #E3150K) to get rid of contaminating rcDNA (relaxed circular HBV DNA) species and heat inactivated for 30 min at 80°C, before being used in a FRET-based qPCR (Werle-Lapostolle et al., 2004), with cccDNA and *β-globin* primers and probes as detailed in the Supplementary Table 1.

Enzyme-linked immunoassay (ELISA) for secreted HBV antigens. HBeAg and HBsAg were quantified in cell culture in supernatants from HBV infected PHHs using commercially available ELISA kits (Autobio, China, #CL0313-2 and #CL0311-2), according to the manufacturer's protocol.

Immunofluorescence for HBc detection. Cells were fixed with 4% formaldehyde (Sigma-Aldrich, #F8775) in PBS 1X (Eurobio, #CS1PBS01KBP), permeabilized with 0.1% Triton X-100 (Sigma-Aldrich, #T8787) in PBS 1X and then treated with 1,5% of BSA (Euromedex, #04-100-812-E) in PBS 1X. Cells were then incubated for 1 hour at room temperature with an anti-HBc mouse monoclonal antibody (Abcam ab18684; 1:1000 dilution). After washing, cells were

incubated with a goat anti-mouse Alexa Fluor 555-conjugated secondary antibody (Cell Signaling, #4409; 1:1000 dilution). Nuclei were counterstained using DAPI (Cell Signaling, #4083; 1:5000 dilution).

Chromatin Tagmentation and Sequencing. 50 000 HBV or mock infected cells harvested 2, 24 and 72 hours post-infection were used for the transposase reaction, which was carried out essentially as described in Buenrostro et al. (2015). PHHs are trypsinized (Trypsine 0,5% EDTA; Invitrogen, #25300054) at 37°C and the reaction was stopped by adding cell cultured medium on ice. After a wash in cold PBS 1X (Eurobio, #CS1PBS01KBP) cells were centrifugated at 500g for 5 minutes at 4°C, the pellets resuspend the in 50 µl of cold nuclei lysis buffer (10mM Tris-HCL [pH7,4], 10mM NaCl, 3mM MgCl2, 0,1% NP-40). After 30 minutes, the cells were centrifuged at 500g, for 5 minutes at 4°C, the supernatants removed and pelleted nuclei used for the tagmentation protocol (Buenrostro et al., 2015). Briefly, pellets were resuspended in the transposase reaction mix (25 µL 2xTD buffer, 2.5 µL transposase (Nextera Illumina) and 22.5 µL nuclease-free water) for 40 min at 37 °C and then was purified on Qiagen MinElute columns. Libraries fragments were amplified using 1× NEB next PCR master mix and 1,25 µM of custom Nextera PCR primers (Buenrostro et al., 2015) using the following PCR conditions: 72 °C for 5 min; 98 °C for 30 s; and thermocycling at 98 °C for 10 s, 63 °C for 30 s and 72 °C for 1 min. Library amplification were monitored by qPCR on 1/10 of the reaction after 5 cycles in order to avoid saturation and the risk of GC and size bias. Libraries were typically amplified for a total of 10–15 cycles followed by size-selection using AMPure beads and library quality control by Bioanalyzer (Agilent). Each ATAC-sequencing library was pair-end sequenced (150 cycles) to an average of 40 million reads per sample using a NextSeq V2 high output sequencing kit on an NextSeq 500 Illumina platform.

RNA sequencing. RNA are extract with the RNAeasy mini kit (Qiagen, #74104) and treated with the RNAse Free DNAse set (Qiagen, #79254) and controlled for quality on a Bioanalyzer (Agilent). mRNA library were prepared using the TruSeq® Stranded RNA HT (w/Ribo-Zero™) kit to increases the depth of sequencing. Each library was pair-end sequenced to an average of 50 – 60 million reads per sample using 75x2 cycles on a NextSeq 500 Illumina platform.

***ChIP assays*.** PHH were harvested at the indicated time post-infection, resuspended in 1-2 ml of ChIP lysis buffer (50 mM Tris HCL, pH 8, 0.5% NP40, 1 mM EDTA, and 100 mM NaCL) and incubated 10 minutes at 4°C. The lysate was centrifuged at 10,000 g for 2 minutes to pellet the nuclei. The supernatant was removed, and the nuclei fixed in 1% formaldehyde for 30 minutes at 4°C. Isolated cross-linked nuclei were extracted with a 20 mM Tris, pH 8, 3 mM MgCl2, 20 mM KCl buffer containing protease inhibitors, pelleted by microcentrifugation, and lysed by

incubation in SDS lysis buffer (1% sodium dodecyl sulfate, 10 mM EDTA, 50 mM Tris-chloride, pH 8.1) containing protease inhibitors. For ChIP the resulting chromatin solution was sonicated for 5 pulses of 45 seconds at 80% power to generate 300- to 1000-bp DNA fragments using a Bioruptor Sonicator (Diagenode Inc). 100 µl of Dynabeads Protein G (Invitrogen, Cat#10003D) were added to each 1 ml chromatin preparation and incubated on a rotator for 14–16 hours at 4°C. with an anti-HBx antibody (ThermoFisher, #MA1-081). 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 is analyzed by real-time PCR amplification using specific primers for genes promoters and the Lightcycler 480 Probes Master (Roche).

ATAC-seq data processing. Reads (25-30 million per sample) were trimmed using ngsqctoolkit (2.3.3) and aligned to the GRCh38/hg38 assembly of the human genome using bowtie2 (2.2.3). At the same time, trimmed reads were also aligned on HBV genome using bowtie2 (2.2.3). Additional analysis, visualization, and quality control were done using various in-house R scripts and published Bioconductor packages (ATACSeqQC). Peak calling was performed with MACS2 and to annotate peaks with promoter/distal labels, and nearest gene, we used the Homer package, with the command “annotatePeaks.pl”. The chromatin accessibility of each region in each sample was quantified using HTSeq-count R package, counting the number of reads from the filtered BAM file that overlapped each region. Data normalization and significant analysis was performed by pair-wise comparing the 3 categories using DESeq2 with p value < 0.05 and log2 fold change > 1.8. The ATAC-seq analysis workflow is shown in Supplementary Figure 1.
The ENRICHR gene set enrichment analysis suite (Kuleshov et al., 2016) was used for motif discovery and transcription factor target enrichment at the chromatin-accessible regions (ChEA and ENCODE databases) as well as to analyze the pathway enrichment (KEGG) and the biological function of the genes associated with the accessible chromatin regions (GO_MF, GO_BP).

RNA-seq analysis. Sequencing quality control was performed using Sequence Analysis Viewer (SAV) and FastQC reports. Base Phred quality score ≥30, checking the sequencing quality has been used to remove bases with low qualities values. RNA-seq reads were processed with the RAP cloud pipeline (D’Antonio et al., 2015), a Tuxedo-suite based workflow that employs TopHat (Kim et al., 2013) for read mapping and Cuffdiff-Cufflinks (Trapnell et al., 2012; Trapnell et al., 2013) for gene-transcript expression quantification and differential

expression analysis. FPKM (Fragments Per Kilobase mapped per Million) values matrixes were downloaded and processed with custom R/Python scripting. Only genes with FDR P-values ≤ 0.05 and a fold change in absolute value ≥1.5 will be considered as differentially expressed. Functional enrichment analysis of differentially expressed genes has been investigated using the ENRICHR gene set enrichment analysis suite (Gene ontology, KEEG, Reactome).

Results and Discussion

ATAC-seq chromatin accessibility landscape in primary human hepatocytes (PHHs). We used the assay of transposase-accessible chromatin sequencing (ATAC-seq) to establish the genome-wide chromatin accessibility landscape in HBV-infected PHHs. PHHs were infected at MOI 250 and infection efficiency was controlled by HBeAg and HBsAg secretion (Figure S2A), total intracellular HBV DNA, total HBV RNA, cccDNA copies per cell (Figure S2B) and HBcAg immunostaining (Figure S2C). ATAC-seq data were from four replicate samples from HBV-infected (HBV) or not (mock) hepatocytes. We obtained an average of 71.7% mappability and >40 million qualified fragments per sample. In non-infected and HBV-infected PHHs samples 28600 and 31500 high-confidence open chromatin regions (peaks) were identified, respectively. The analysis of ATAC-seq fragment sizes generated Tn5 transposase fragmentation revealed DNA fragments that corresponded to nucleosome depleted, mono-nucleosome, and di-nucleosome patterns (Figure S3A), which is consistent with previous findings (Buenrostro et al., 2013). The estimated ATAC-seq libraries complexity (Figure S3B) and the distribution of filtered reads mapped to each chromosome (Figure S3C) were evaluated to ensure the quality and the representativeness of the ATAC-seq libraries.

Increased chromatin accessibility in HBV infected hepatocytes. Next, we compared chromatin accessibility in non-infected and HBV-infected PHHs at different times post-infection. The hierarchical bi-clustering analysis of differentially accessible regions clearly separated the samples into two groups: one from not infected cells (Mock) and the other from HBV infected (HBV) cells (Figure 1A), indicating a substantial shift of chromatin accessibility in HBV infected cells. The Volcano Plot analysis shows how the separations between the HBV-infected PHHs and controls takes place at statistically significant P value (< 0,05), with a Log2 fold change of 1,8 (Figure 1B). Genes that are highly dysregulated are farther to the left and right sides, while highly significant changes appear higher on the plot. Green dots represent dysregulated genes that meet both the P value and the fold change criteria for significantly more open (on the right) and closed (on the left) chromatin after HBV infection. In red are the genes that pass the P value filter but have a Log2 fold change below the chosen threshold whereas the yellow dots define genes that display a high Log2 fold change but do not reach a P value < 0.05.

The genomic sites displaying altered chromatin accessibility varies over time, with a relatively low overlap at the different time points, and their number increases with the time of infection (278 at 2 hours; 350 at 24 hours and 1095 at 72 hours) (Figure 2A), possibly reflecting the kinetics of HBV replication and cccDNA transcription after infection in PHHs (Niu et al., 2017

and Figure 8). Indeed, 2 hours and 24 hours post infection the HBc protein from the incoming viral capsids is likely present and bound to the cccDNA (Belloni et al., 2009; Riviere et al., 2015) whereas the only HBV RNA species detected is the HBx transcript and the 3.5 kb RNA species (pgRNA and preC RNA) are detected after 24 hours (Niu et al., 2017). It is tempting to speculate that the 3 waves of chromatin accessibility changes we observed might be related to, respectively, the early sensing of the viral infection, the early activity of the HBx protein in the absence of full HBV genome expression and the complete establishment of HBV infection. However, additional experiments, including PHH infections with mutant HBVs and HBx ChIP experiments performed at early times post-infection, will be needed to support this hypothesis. At all time points post infection, open potentially transcriptionally active genes were more represented (Figure 2B) with about 2/3 of protein-coding genes and 25 to 31 % of non-coding RNAs (Figure 2C). The complete list of unique genes showing an increased or decreased chromatin accessibility in HBV-infected PHHs is presented in Supplementary Table 2. Supplementary Tables 3, 4 and 5 contain the complete annotated list of the genes displaying an altered chromatin accessibility (gene name, P value, log2 fold change, chromosome location and open closed region position in the genome, distance to TSS, gene description and gene type – i.e. protein coding, ncRNA, pseudo-gene) at 2 hours, 24 hours and 72 hours post-infection.

Accessible chromatin regions are highly enriched for SMADs and STAT3 binding motifs. Figure 3A shows that the accessible chromatin regions were enriched around the TSSs at all time points post-infection, which is consistent with the fact that chromatin around the TSS's throughout the genome is more accessible than the surrounding genomic regions.
We used the ENRICHR gene set enrichment analysis suite (Kuleshov et al., 2016) and the ChEA and ENCODE databases for motif discovery and transcription factor target enrichment at the altered chromatin-accessible regions. The analysis performed combining all replicates with no time distinctions identified the binding sites for the SMADs and the STAT3 proteins as the most highly enriched motifs (Figure 3B). Notably, the SMAD family of transcription factors mediates TGF-β signaling that is involved both in liver fibrosis and HCC development (Levrero and Zucman Rossi, 2016) and is downregulated in HBV infected PHHs (Xia et al., 2018). Moreover, Interleukin-6-dependent and -independent activation of STAT3 has been detected in HBV-replicating hepatocytes in cell culture and in vivo and prevention of STAT3 activation reduced HBV replication and viral genes expression (Hösel et al., 2017). Other significantly enriched motifs (P value <0.05) were the p53 responsive elements (p53RE), bound by the TP53 tumor suppressor and its TP73 and TP63 paralogs, and Suz12, a component of the PRC2 polycomb repressive complex. The same motif analysis performed at the different times post-infection (Figure 3C) confirmed the enrichment of binding sites for SMADs, STAT3 and

Suz12 as well as the p53REs and identified additional less enriched potential binding sites for SMC3 (structural maintenance of chromosomes 3), a component of the Cohesin protein complex that is a major organizer of interphase chromatin and sister chromatin cohesion, and TCF3, a partner of β-catenin. SMC3/Cohesin cooperates with the insulator binding protein CTCF to facilitate the developmental silencing of imprinted genes (Chatzinikolaou et al., 2017) and drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres (Remeseiro et al., 2012).

Pathway enrichment analysis of genes displaying changes in chromatin accessibility after HBV infection. To model the impact of HBV infection on cell functions and viral replication we analyzed the pathway enrichment (KEGG) and the biological function of the genes associated with the open chromatin regions (GO_MF, GO_BP) using the ENRICHR gene set enrichment analysis suite. Overall, the regions with increased chromatin accessibility were enriched in genes involved in liver cancer development and/or progression (Figure 4A, Jensen DISEASES dataset). The pathway enrichment analysis performed at the different times post-infection showed an enrichment of genes involved in: a) histones demethylation (KEGG) and calcium signaling (GO) at 2 hours (Figure 4B); b) downregulation and upregulation of different subsets of genes controlled by TGF-β signaling (KEGG) at 24 hours (Figure 4C); c) TP53 and DNA damage responses (GO) at 24 hours (Figure 4C); cAMP/CREB regulated genes (KEGG) and cell adhesion (GO) at 72 hours (Figure 4D). Conversely, chromatin regions with decreased accessibility were enriched in genes involved in a) fat metabolism (GO) at 2 hours (Figure 5A); b) histone acetylation and phosphoproteins phosphatases (GO) at 24 hours (Figure 5B) and c) retinoic acid receptor (RAR) signaling (KEGG) at 72 hours (Figure 5C). Altogether these results indicate that the changes in chromatin accessibility imposed by HBV infection target a number of pathways involved in HBV replication and, although our analysis is restricted to early events after hepatocytes infection, also pathways that may potentially favor the establishment of a premalignant phenotype in infected hepatocytes. Thus, calcium signaling is important for HBV replication (Yang and Bouchard, 2012). The activation of histone demethylases may play an important role in counteracting cccDNA silencing by histone methyltransferases such as SETDB1 and SET1A (Riviere et al., 2015; Alarcon et al., 2016). The modulation of different subsets of TGF-β target genes can play a role both in viral replication and hepatocytes transformation (Xia et al. 2018; Levrero and Zucman-Rossi, 2016). Finally, the downregulation of RAR signaling would promote viral replication (Li et al., 2018).
The RNA-seq analysis of non-infected and HBV-infected PHHs identified 563, 1259 and 636 differentially expressed genes (DEGs) 2 hours, 24 hours and 72 hours post-infection, respectively. Interestingly, the functional enrichment analysis of these DEGs, investigated using the ENRICHR suite and the Gene ontology, KEEG and REACTOME datasets, confirmed

the modulation, in HBV infected PHHs, of genes involved in cell cycle regulation (2 hours post-infection, REACTOME) (Figure 6A); calcium signaling (72 hours post-infection, REACTOME) (Figure 6C); PKA, cAMP and CREB signaling (72 hours post-infection, REACTOME and KEGG) (Figure 6C) as well as genes regulated by the TP53 and the E2F1 transcription factors (2 hours post infection, REACTOME) (Figure 6A).

The integration of the ATAC-seq and RNA-seq analysis did not allow to identify a univocal correlation between the changes in chromatin accessibility and gene expression at the different times post-infection. Temporal or cell type specific dissociations between chromatin accessibility and transcription have been reported (Ackermann et al., 2016). Our analysis at present is limited to correlate gene expression with changes in chromatin accessibility at or near an identified promoter region extended to neighboring upstream and downstream genes (Supplementary Tables 3, 4 and 5). It been shown that distal noncoding regions can exhibit dynamic changes in chromatin accessibility behaving as functional distal enhancers (Daugherty et al., 2017) and our analysis would have not included most of these distal regions. The correlation might also be missed due to different temporal dynamics in chromatin accessibility and transcriptional changes. However, this does not seem to be the explanation in our case since the results were only modestly improved when we repeated the ATAC-seq and RNA-seq integration analysis comparing each ATAC-seq time point with the next RNA-seq time points (i.e. ATAC-seq 2 hours vs RNA-seq 24 hours or 72 hours; ATAC-seq 24 hours vs RNA-seq 72 hours). Altogether, our results concur with the growing understanding that gene activation depends on multiple regulatory regions, many of which are located far from the gene locus itself and with the increasing awareness regarding the existence of multiple intermediate *chromatin states* between the accessible/transcribed and non-accessible/non-transcribed extremes.

We have recently identified, using a ChIP-seq approach in HBV replicating cells, the whole genome repertoire of genes that are directly targeted by the viral protein HBx (Guerrieri et al., 2017). Although the primary role of HBx in HBV life cycle is to bind the cccDNA viral mini-chromosome in order to promote and maintain cccDNA transcription and viral replication (Pollicino et al, 2006; Belloni et al., 2009; Lucifora et al., 2011; Riviere et al., 2015; Decorsière et al., 2016), HBx has been also shown to activate the transcription of a number of cellular genes by interacting with host epigenetic regulators (reviewed in Levrero and Zucman-Rossi, 2016). Thus, we investigated whether the activation of selected host genes transcription by HBx was associated with, or preceded by, changes in chromatin accessibility at their promoter regions. In Figure 7 we compared the ATAC-seq profiles from PHHs non-infected and HBV-infected (24 hours) with the profile of H3 Lysine 27 acetylation (H3K27Ac), a reliable histone mark of transcriptionally active genes, from the ENCODE dataset for the TRIM13 (panel A), SREBP2 (panel B) and DNMT3b (panel C) genes. TRIM13 encodes for a ubiquitin E3 ligase

implicated in ER-stress induced autophagy (Tomar et al. 2012) and in the negative regulation of MDA5-mediated activation of type I interferon production in response to viral infections (Narayan et al. 2014). The SREBP2 transcription factor, known to be a major regulator of cholesterol metabolism, has also been involved in the direct activation of several genes involved in autophagy (Seo et al., 2011). The DNMT3B (DNA Methyltransferase 3 Beta) is required for genome-wide de novo methylation and is essential for the establishment of DNA methylation patterns during development. TRIM13, SREBP2 and DNMT3b are direct target genes of HBx (Guerrieri et al. 2017) and we confirmed by anti-HBx ChIP that the recruitment of HBx on the 3 promoter regions in HBV-infected PHHs is accompanied by an increase in RNA levels (Figure 7A-C).

ATAC-seq analysis of cccDNA chromatin dynamics. The alignment of ATAC-seq and RNA-seq reads to the HBV genome allowed us to generate ATAC-seq profiles of the HBV minichromosome at the different times post-infection and to correlate the evolution of cccDNA chromatin accessibility with cccDNA transcription, the recruitment of Pol2 on the cccDNA and cccDNA-bound histones post-translational modifications (Figure 8). cccDNA episome formation and its chromatinization occur both very early after HBV infection in PHHs. Indeed, cccDNA can be detected by qPCR and the cccDNA minichromosome immunoprecipitated by anti-H3 cccDNA ChIP as early as 2 hours post-infection (Locatelli et al., 2017 and data not show). As shown in Figure 8C, cccDNA chromatin accessibility increases over time after infection with a well-defined profile 72 hours post-infection, when cccDNA is fully transcribed and pgRNA and preC 3.5 kb RNA species are detected in the RNA-seq (Figure 8D). The ATAC-seq profile at 72 hours post-infection shows 2 major regions of high chromatin accessibility (CA1 and CA2). The CA1 region overlap a peak of Pol2 recruitment and nucleosomes with highly acetylated H3K27 that are reflected in the accessibility profile of CA1. Altogether these results demonstrate the feasibility of studying the epigenetics of the cccDNA mini-chromosome and the dynamics of cccDNA chromatin accessibility by ATAC-seq.

Conclusions. We used ATAC-seq to detect early changes in host and viral cccDNA chromatin accessibility in HBV-infected PHHs. Our results show a progressive increase of genomic sites with altered chromatin accessibility (>1000 at day 3 post-infection) with a prevalence of more open/potentially transcriptionally active vs closed/transcriptionally inactive regions. Chromatin accessibility changes affected both proximal extended promoter regions and distal intragenic and intergenic regions. About 30% of the impacted genes were ncRNA genes. RNA-seq analysis performed at the same time post infection showed an increasing number of differentially expressed genes (DEGs). Motif discovery and pathway analysis of accessible chromatin regions and their neighboring genes showed an increase in the binding motifs for

SMADs, STAT3, TP53 and Suz12, a partner of Ezh2 in the PRC2 polycomb repressive complex and an enrichment in genes/ncRNAs involved in viral replication and the establishment of a pre-neoplastic. Altogether our results challenge the commonly accepted concept of HBV being a "stealth" virus that induces no or very few changes in the host genome.

References

Ackermann A, Wang Z, Schug J, Naji A, Kaestner K. (2016) Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. *Mol Metab* 5(3): 233-244.

Adey A, Morrison HG, Asan XX. et al (2010) Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol* 11(12): R119.

Ancey PB, Testoni B, Gruffaz M et al. (2015) Genomic responses to hepatitis B virus (HBV) infection in primary human hepatocytes. *Oncotarget* 6: 44877-44891.

Alarcon V, Hernández S, Rubio L et al. (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.

Belloni L, Pollicino T, De Nicola F et al. (2009) Nuclear HBx binds the HBV minichromosome and modifies the epigenetic regulation of cccDNA function. *Proc Natl Acad Sci USA* 106(47): 19975-9.

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10(12): 1213-8.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Prot Mol Biol* 109: 21-29.

Chatzinikolaou G, Apostolou Z, Aid-Pavlidis T et al. (2017) ERCC1-XPF cooperates with CTCF and cohesin to facilitate the developmental silencing of imprinted genes. *Nat Cell Biol* 19(5): 421-432.

Daugherty AC, Yeo RW, Buenrostro JD, Greenleaf WJ, Kundaje A, Brunet A (2017) Chromatin accessibility dynamics reveal novel functional enhancers in C. elegans. *Genome Res* 12: 2096-2107.

D'Antonio M, D'Onorio De Meo P, Pallocca M et al. (2015) RAP: RNA-seq Analysis Pipeline, a new cloud-based NGS web application. *BMC Genomics* 16 (S6): S3.

Dechassa ML, Tryndyak V, de Conti A, Xiao W, Beland FA, Pogribny IP. (2018) Identification of chromatin-accessible domains in non-alcoholic steatohepatitis-derived hepatocellular carcinoma. *Mol Carcinogenesis* 57: 978–987.

Decorsière A, Mueller H, van Breugel PC et al. (2016) Hepatitis B virus X protein identifies the Smc5/6 complex as a host restriction factor. *Nature* 531(7594): 386-9.

de Dieuleveult M, Yen K, Hmitou I et al. (2016) Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. *Nature* 530(7588): 113-6.

Durantel D and Zoulim F (2009) Innate response to hepatitis B virus infection: observations challenging the concept of a stealth virus. *Hepatology* 50: 1692-1695.

EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. (2017) *J Hepatol* 67(2): 370-398.

Fletcher SP, Chin DJ, Cheng DT, et al. (2013) Identification of an intrahepatic transcriptional signature associated with self-limiting infection in the woodchuck model of hepatitis B. *Hepatology* 57: 13-22.

Giersch K, Allweiss L, Volz T et al. (2015) Hepatitis Delta co-infection in humanized mice leads to pronounced induction of innate immune responses in comparison to HBV mono-infection. *J Hepatol* 63: 346-353.

GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016 388(10053): 1459-1544.

Global Hepatitis Report 2017. Geneva: World Health Organization; 2017. Licence: CC BY-NC-SA 3.0 IGO.

Guerrieri F, Belloni L, D'Andrea D et al. (2017). Genome-wide identification of direct HBx genomic targets. *BMC Genomics* 18(1): 184

Hösel M, Quasdorff M, Ringelhan M et al. (2017) Hepatitis B Virus activates signal transducer and activator of transcription 3 (STAT3) supporting hepatocyte survival and virus replication. *Cell Mol Gastroenterol Hepatol* 4(3): 339-363.

Jin J, Xu H, Wu R, Niu J and Li S. (2018) Aberrant DNA methylation profile of hepatitis B virus infection *J Med Virol* Aug 17 [Epub ahead of print]

Johnson JL, Georgakilas G, Petrovic J, Kurachi M, Cai S, Harly C et al. (2018) Lineage-determining transcription factor TCF-1 initiates the epigenetic identity of T cells. *Immunity* 48(2): 243-257.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14: R36.

Kuleshov MV, Jones MR, Rouillard AD et al. (2016) ENRICHR: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 44(W1): W90-7.

Ladner SK, Otto MJ, Barker CS et al. (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.

Lamontagne J, Mell JC, and Bouchard MJ. (2016) Transcriptome-wide analysis of Hepatitis B Virus-mediated changes to normal hepatocyte gene expression. *PLoS Pathog* 12: e1005438.

Lebossé F, Testoni B, Fresquet J et al. (2017). Intrahepatic innate immune response pathways are downregulated in untreated chronic hepatitis B. *J. Hepatol* 66: 897-909.

Le Cluyse EL and Alexandre E. (2010) Isolation and culture of primary hepatocytes from resected human liver tissue. In *Hepatocytes* (Humana Press), pp. 57–82.

Levrero M and Zucman-Rossi J. (2016) Mechanisms of HBV-induced hepatocellular carcinoma. *J Hepatol* 64 (S1): S84-S101.

Li B, Wang Y, Shen F et al. (2018) Identification of Retinoic Acid Receptor agonists as potent Hepatitis B Virus inhibitors via a drug repurposing screen. *Antimicrob Agents Chemother* pii: AAC.00465-18.

Locatelli M, Fresquet, J, Quivy JP et al. (2017) The histone chaperone HIRA is crucial for the early establishment of hepatitis B virus mini-chromosome. *Hepatology*, 66: S1, 783A.

Lok AS, Zoulim F, Dusheiko G, Ghany MG. Hepatitis B Cure: from discovery to regulatory approval. (2017) *Hepatology* 66(4): 1296-1313.

Luangsay S, Ait-Goughoulte M, Michelet M et al. (2015). Expression and functionality of Toll- and RIG-like receptors in HepaRG cells. *J. Hepatol* 63: 1077-1085.

Lucifora J, Arzberger S, Durantel D et al. (2011) Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. *J Hepatol* 55(5): 996-1003.

Lucifora J, Bonnin M, Aillot L et al. (2018). Direct antiviral properties of TLR ligands against HBV replication in immune-competent hepatocytes. *Sci Rep* 8: 5390.

Maini MK, Gehring AJ. (2016) The role of innate immunity in the immunopathology and treatment of HBV infection. *J Hepatol* 64: S60-S70.

Narayan K, Waggoner L, Pham ST et al. (2014). TRIM13 is a negative regulator of MDA5 mediated type I interferon production. *J Virol* 88(18): 10748-57.

Niu C, Livingston CM, Li L 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.

Pollicino T, Belloni L, Raffa G et al. (2006) Hepatitis B virus replication is regulated by the acetylation status of hepatitis B virus cccDNA-bound H3 and H4 histones. *Gastroenterology* 130(3): 823-37.

Remeseiro S, Cuadrado A, Carretero M et al. (2012) Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres. *EMBO J* 31(9): 2076-89.

Rivière L, Gerossier L, Ducroux A et al. (2015) HBx relieves chromatin-mediated transcriptional repression of hepatitis B viral cccDNA involving SETDB1 histone methyltransferase. *J Hepatol* 63(5): 1093-102.

Sato S, Li K, Kameyama T et al. (2015) The RNA sensor RIG-I dually functions as an innate sensor and direct antiviral factor for hepatitis B virus. *Immunity* 42: 123±132.

Schweitzer A, Horn J, Mikolajczyk RT, Krause G, Ott JJ. (2015) Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published

between 1965 and 2013. *Lancet* 386(10003):1546-55.

Seo YK, Jeon TI, Chong HK, Biesinger J, Xie X, Osborne TF. (2011) Genome-wide localization of SREBP-2 in hepatic chromatin predicts a role in autophagy. *Cell Metab* 13: 367-375.

Suslov A, Boldanova T, Wang X, Wieland S and Heim MH. (2018) Hepatitis B virus does not interfere with innate immune responses in the human liver. *Gastroenterology* 154: 1778–1790.

Testoni B, Levrero M, Zoulim F. Challenges to a cure for HBV Infection. *Semin Liver Dis* 2017, 37(3): 231-242.

Tomar D, Singh R, Singh AK, Pandya CD, Singh R. (2012) TRIM13 regulates ER stress induced autophagy and clonogenic ability of the cells. B*iochim Biophys Acta* 1823(2): 316-26.

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* 7: 562-78.

Trapnell C, Hendrickson D,Sauvageau S, Goff L, Rinn JL, Pachter L. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology* 31: 46-53.

Tropberger P, Mercier A, Robinson M et al. (2015) Mapping of histone modifications in episomal HBV cccDNA uncovers an unusual chromatin organization amenable to epigenetic manipulation. *Proc Natl Acad Sci USA* 112(42): E5715-24.

Tsompana M and Buck MJ. (2014) Chromatin accessibility: a window into the genome. *Epigenetics & Chromatin* 7:33.

Wang J, Zibetti C, Shang P et al. (2018) ATAC-seq analysis reveals a widespread decrease of chromatin accessibility in age-related macular degeneration. *Nat Commun* 9(1): 1364.

Werle-Lapostolle B, Bowden S, Locarnini S et al. (2004) Persistence of cccDNA during the natural history of chronic hepatitis B and decline during adefovir dipivoxil therapy. *Gastroenterology* 126: 1750–1758.

Wieland S, Thimme R, Purcell RH, Chisari FV. (2004) Genomic analysis of the host response to hepatitis B virus infection. *Proc Natl Acad Sci USA* 101: 6669-6674.

Wieland SF and Chisari FV. (2005) Stealth and cunning: hepatitis B and hepatitis C viruses. *J Virol* 79: 9369±9380.

Xia Y, Cheng X, Li Y, Valdez K, Chen W, Liang TJ. (2018) Hepatitis B Virus deregulates the cell cycle to promote viral replication and a premalignant phenotype. *J Virol* 12: 92(19).

Yang B and Bouchard MJ. (2012) The hepatitis virus X protein elevates cytosolic calcium signals by modulating mitochondrial calcium uptake. *J Virol* 86: 313-22

Yoneda M, Hyun J, Jakubski S, et al. (2016) Hepatitis B Virus and DNA Stimulation Trigger a Rapid Innate Immune Response through NF-kappaB. *J Immunol* 197: 630±643.

Zoulim F, Lebossé F, Levrero M. (2016) Current treatments for chronic hepatitis B virus infections. *Curr Opin Virol* 18: 109-16.

Figure 1

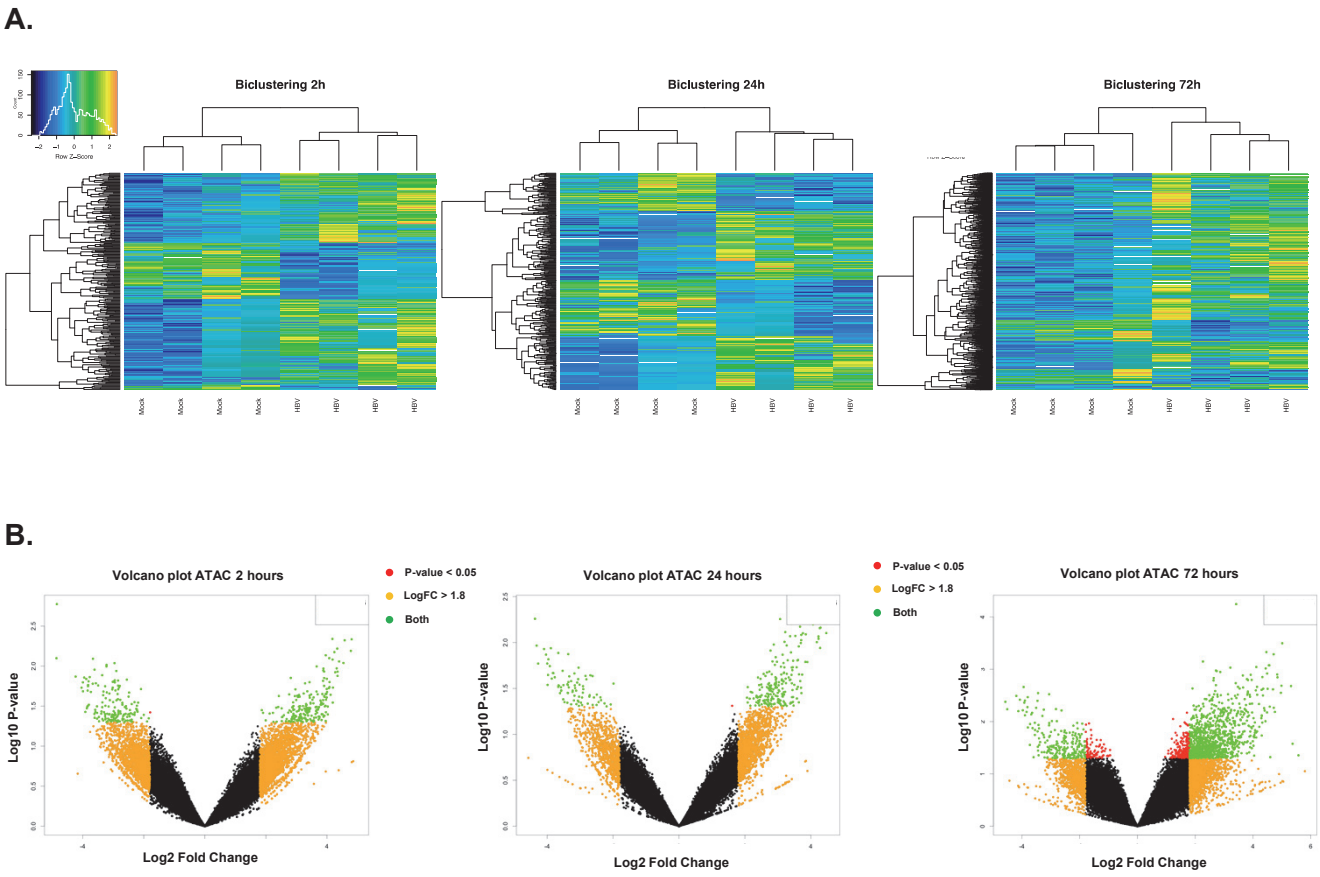
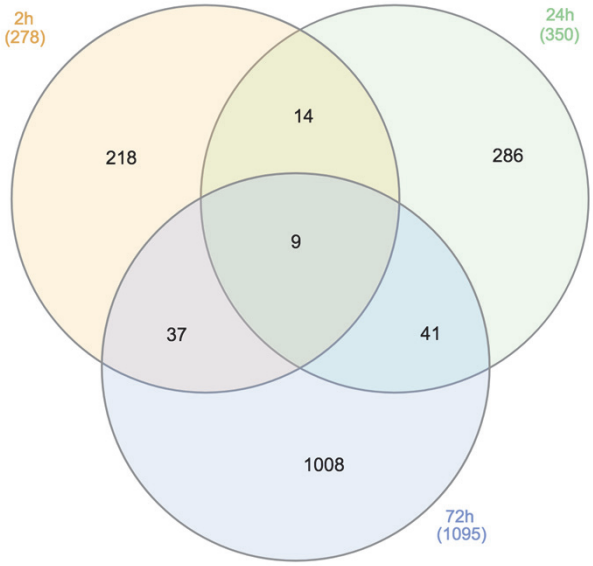


Figure 1. ATAC-Seq analysis of chromatin accessibility in HBV-infected PHHs. **A.** Heat map of bi-clustering analysis of differentially accessible regions in non-infected (Mock) and HBV-infected (HBV) PHHs. Color change from blue to yellow/red reflects the entity of accessibility change from closed (blue) to open (red). Samples include 4 HBV-infected and 4 non-infected PHH samples. **B.** Volcano plot analysis of genes showing changes in chromatin accessibility in mock and HBV infected PHHs. x-axis = log2 fold change; y-axis = log10 P value. Statistically significant threshold are set at P value < 0,05 with a Log2 fold change of 1,8. Genes that are highly dysregulated are farther to the left and right sides, while highly significant changes appear higher on the plot.

Figure 2

A.



B.

	Open genes	Closed genes
2 hours post-infection	199	79
24 hours of post-infection	200	150
72 hours of post-infection	904	191

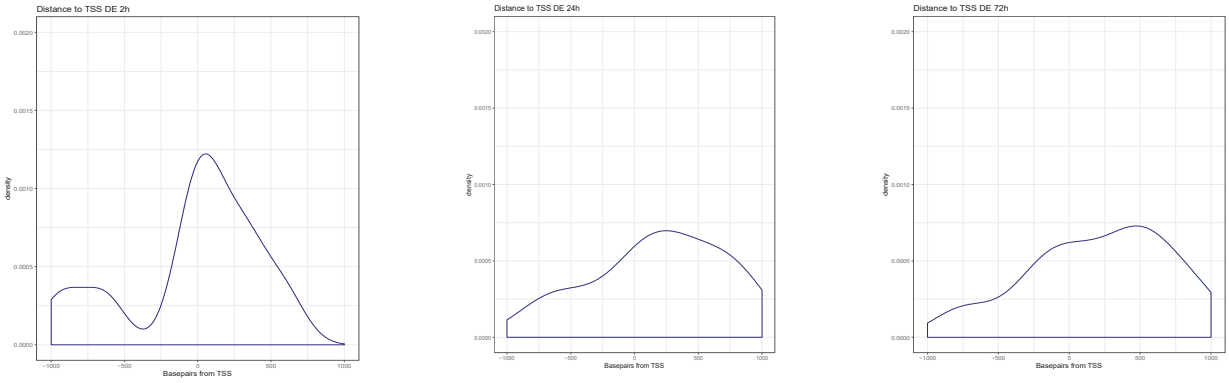
C.

	% ncRNA	% mRNA	% scRNA	% snRNA
2 hours post-infection	25,8	70,32	0	0,35
24 hours post-infection	28,46	67,48	0	0,81
72 hours post-infection	30,74	63,77	0,16	0,57

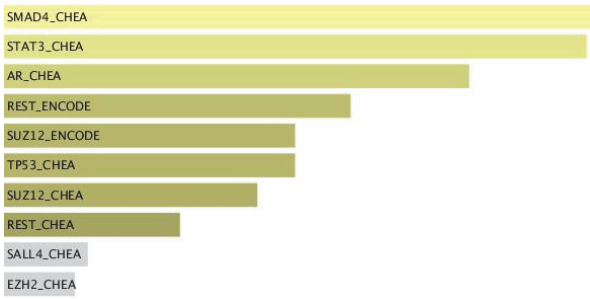
Figure 2. Changes of chromatin accessibility in HBV infected PHHs. A. Venn diagram of genes with altered chromatin accessibility at 2 hours (orange), 24 hours (green) or 72 hours (light blue) post-infection. **B.** ATAC-Seq open and closed genes in HBV infected PHHs 2 hours / 24 hours / 72 hours post-infection. **C**. Percentage of protein coding RNAs and ncRNA genes annotated genes.

Figure 3

A.



B.



Index	Name	P-value	Adjusted p-value	Z-score	Combined score
1	SMAD4_CHEA	2.850e-11	2.964e-9	-1.75	42.39
2	AR_CHEA	2.080e-7	0.00001081	-1.60	24.57
3	SUZ12_CHEA	0.00004145	0.001437	-1.47	14.85
4	STAT3_CHEA	0.0001919	0.004990	-1.67	14.26
5	SUZ12_ENCODE	0.001761	0.03662	-1.49	9.48
6	TP53_CHEA	0.01619	0.2806	-1.49	6.16
7	REST_ENCODE	0.03833	0.4642	-1.53	4.98
8	REST_CHEA	0.03561	0.4642	-1.43	4.76
9	NFE2L2_CHEA	0.04017	0.4642	-1.35	4.34
10	SALL4_CHEA	0.1256	1.000	-1.37	2.85

C.

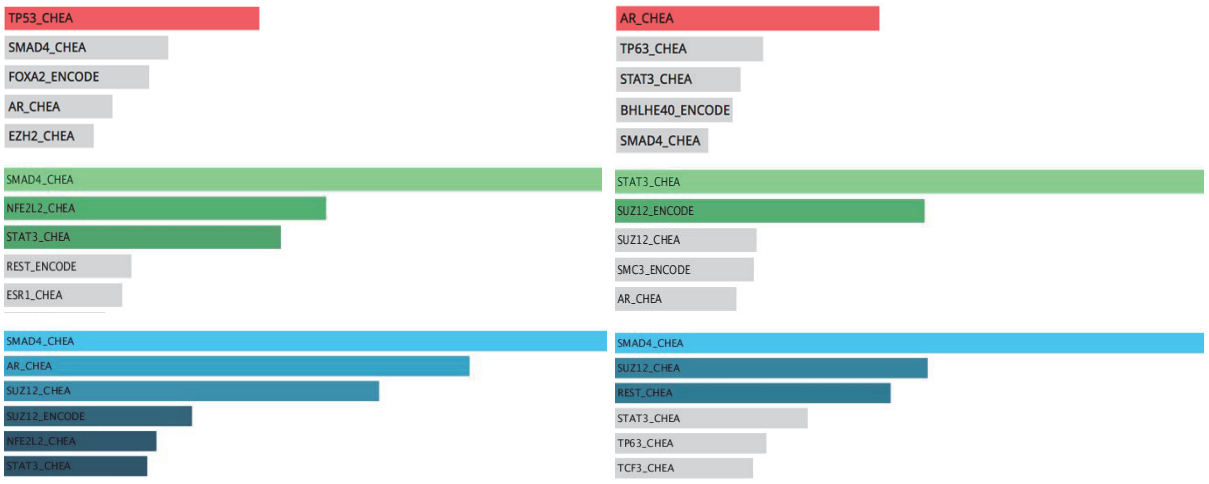


Figure 3. Motif discovery and transcription factor target enrichment analysis. A. Chromatin accessibility around the TSS. x axis = distance from the TSS; y axis = ATAC-seq reads density. **B-C.** ENRICHR gene set enrichment analysis based upon the ChEA and ENCODE databases for motif discovery and transcription factor target enrichment at the altered chromatin-accessible regions performed in all replicates without time distinction (B, yellow) or at 2 hours (red), 24 hours (green) and 72 hours (blue) post-infection. In C. the left panes refer to differentially open genomic sites / genes whereas the right panels refer to differentially closed genomic sites / genes.

Figure 4

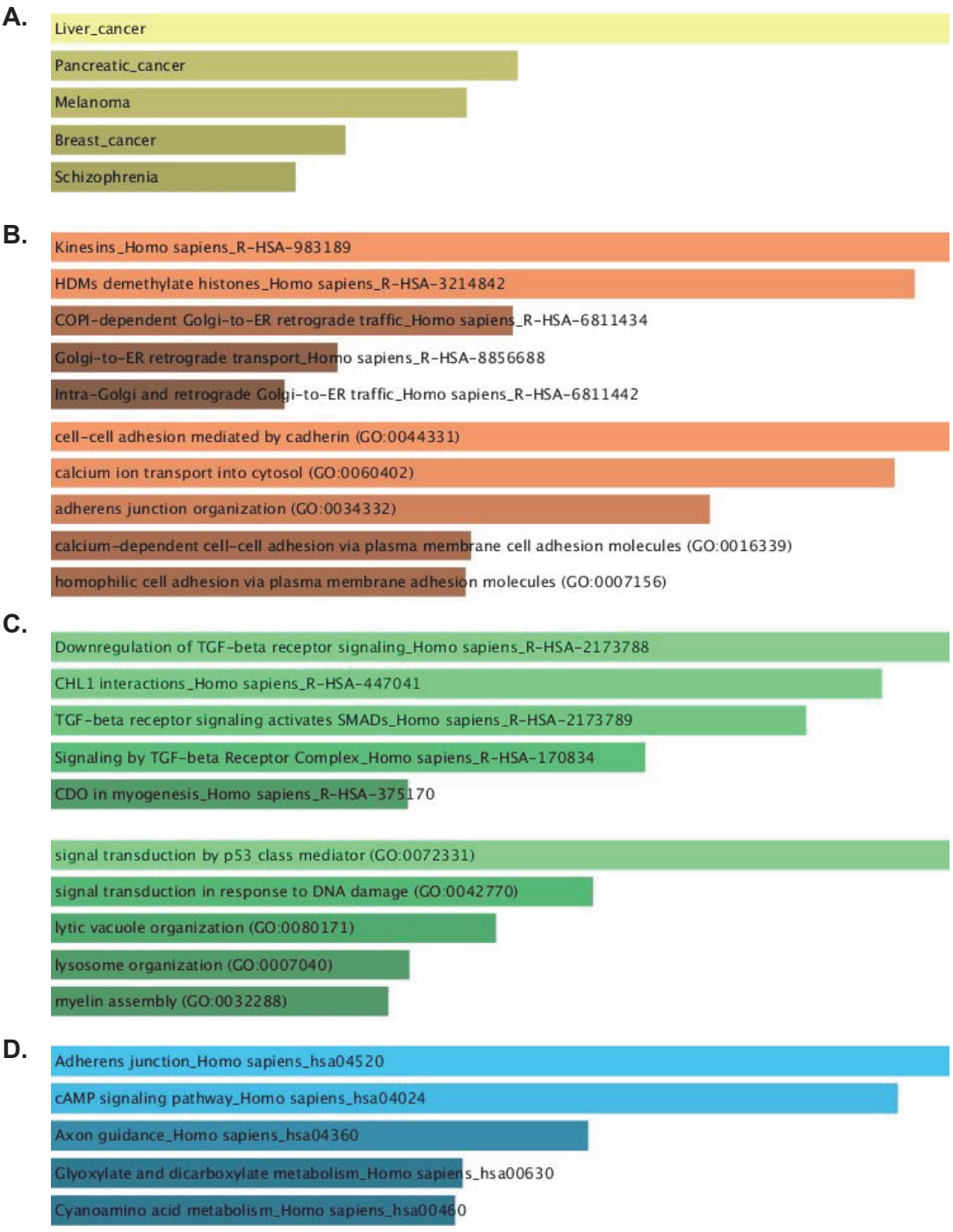


Figure 4. Pathway enrichment analysis of genes associated with open chromatin regions after HBV infection. Pathway enrichment (KEGG) and the genes biological function (GO_MF, GO_BP) were analyzed using the ENRICHR gene set enrichment analysis tool. **A.** Jensen DISEASES analysis **B. C. D.** KEGG and Gene Ontology analysis of the pathways and biological process associated with accessible chromatin regions in at 2 hours (B, orange), 24 hours (C, green) and 72 hours (D, blue) post-infection.

Figure 5

A.

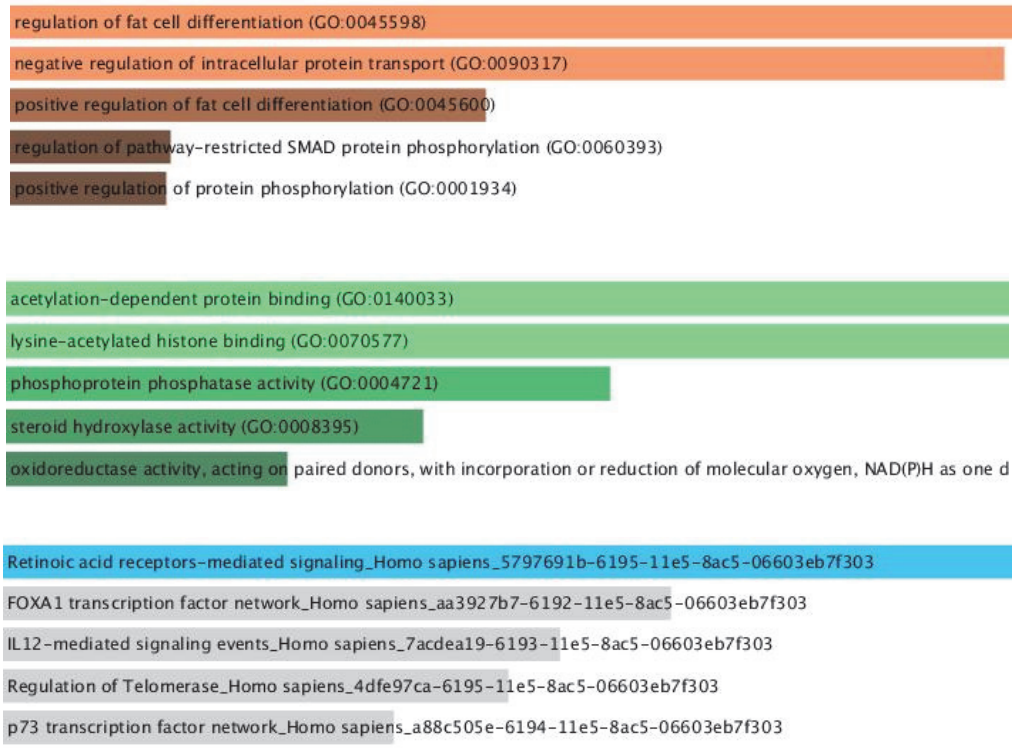


Figure 5. Pathway enrichment analysis of genes associated with closed chromatin regions after HBV infection. Pathway enrichment (KEGG) and the genes biological function (GO_MF, GO_BP) were analyzed using the ENRICHR gene set enrichment analysis tool. **A.** Jensen DISEASES analysis **B. C. D.** KEGG and Gene Ontology analysis of the pathways and biological process associated with genes displaying decreased chromatin accessibility at 2 hours (B, orange), 24 hours (C, green) and 72 hours (D, blue) post-infection.

Figure 6

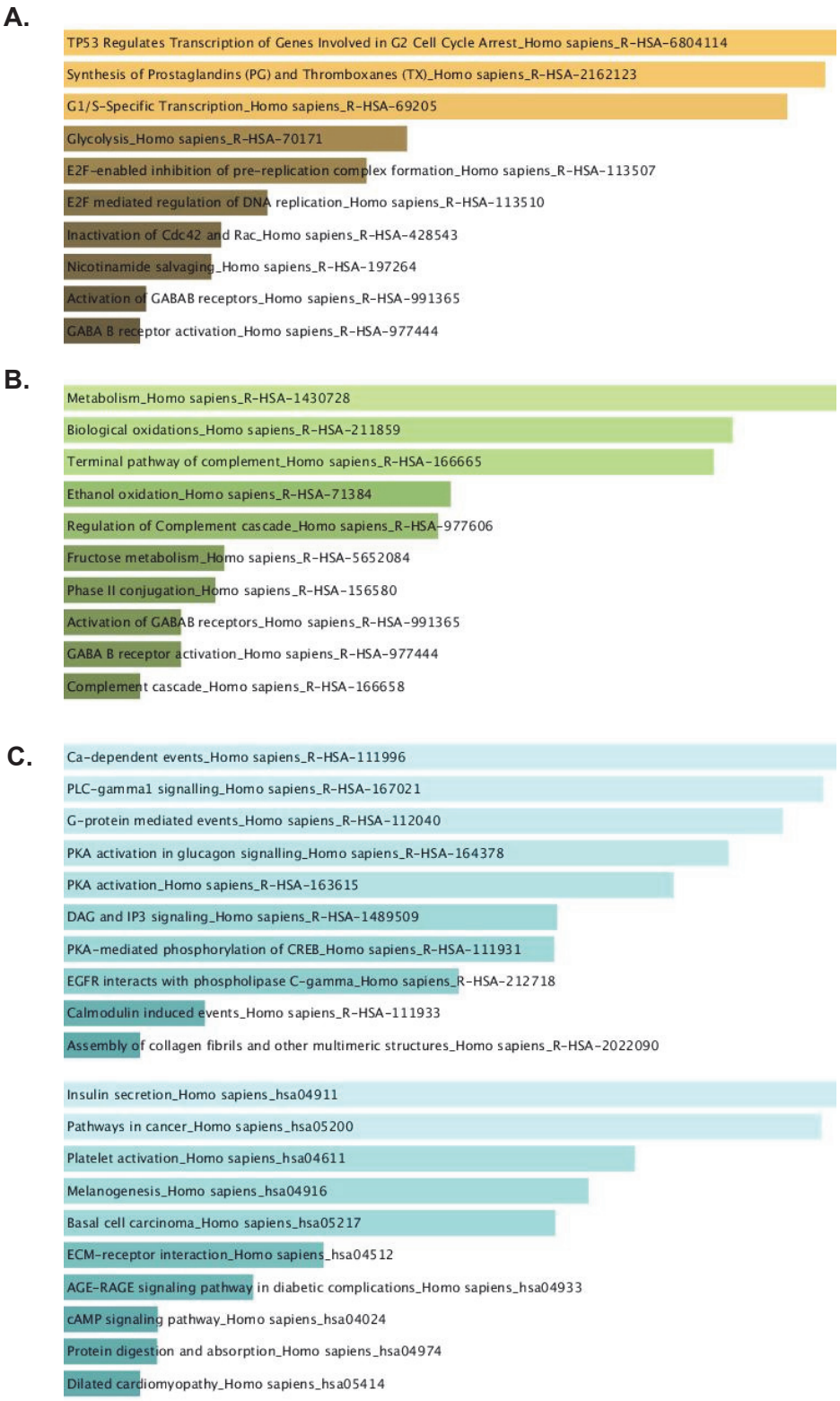


Figure 6. Pathway enrichment analysis and gene ontology of RNA-seq differentially expressed genes (DEGs) after HBV infection. Functional enrichment analysis of DEGs using the ENRICHR suite and the Gene ontology, KEEG, REACTOME datasets. **A. B. C.** Pathways and biological process associated with DEGs genes at 2 hours (A, orange, REACTOME), 24 hours (B, green, REACTOME) and 72 hours (D, blue; upper panel = REACTOME; lower panel = KEGG) post-infection.

Figure 7

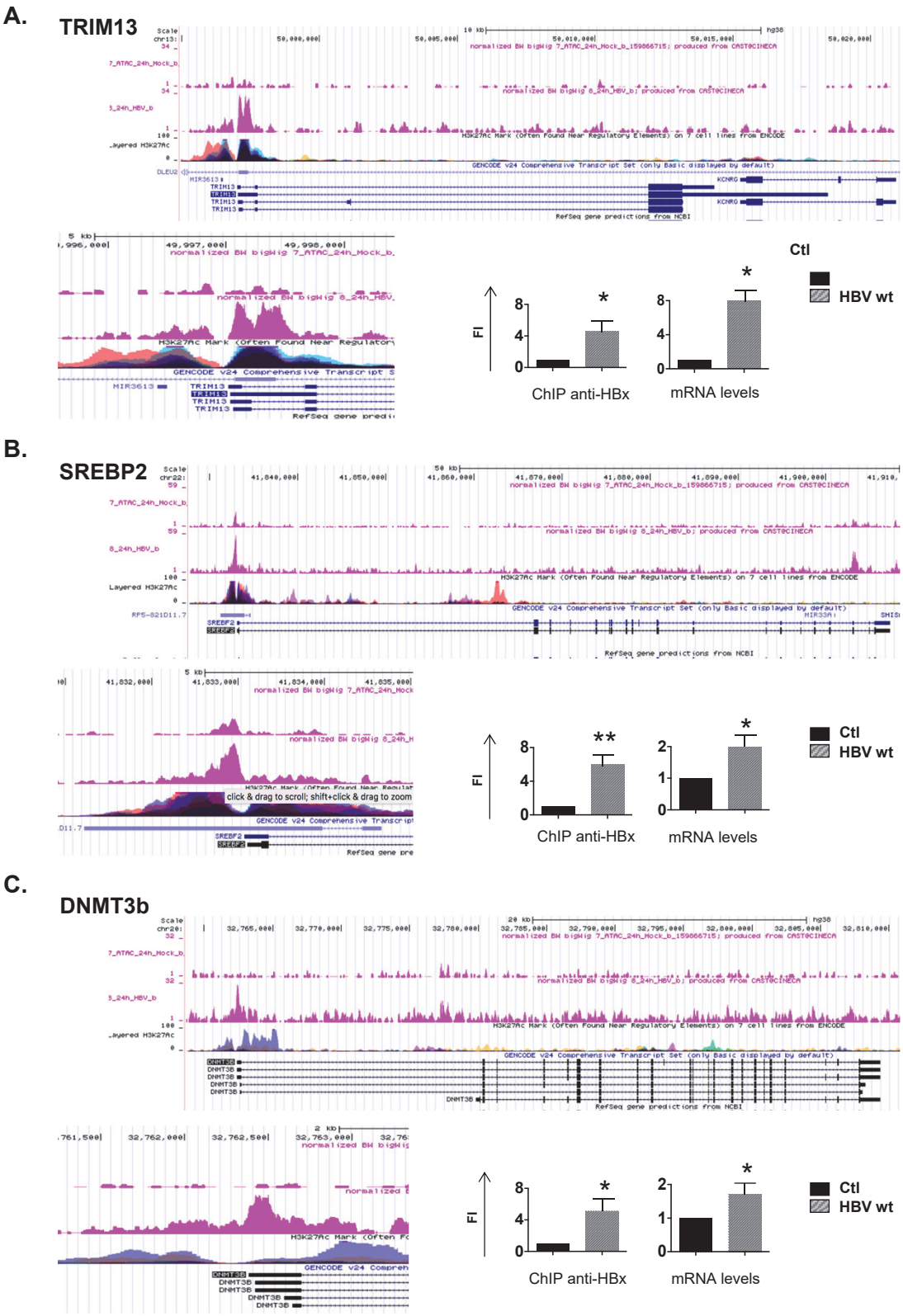


Figure 7. ATAC-seq chromatin accessibility, H3 Lysine 27 acetylation, HBx recruitment and transcription at the TRIM13, SREBF2 and DNMT3B genes. ATAC-Seq data peaks (pink) for the TRIM13 (**A**), SREBP2 (**B**) and DNMT3b (**C**) in non-infected (mock) and HBV-infected PHHs were integrated with H3K27Ac ChIP-Seq profiles from the ENCODE datasets using the UCSC Genome Browser. *Lower left* panels in A, B and C show the same promoter regions at higher magnification. *Lower right* panels show HBx recruitment on the 3 promoter regions analyzed by ChIP and the 3 genes mRNA levels analyzed by qPCR in HBV-infected and non-infected (mock) PHHs. ChIP results are expressed as Fold Induction of the % of Input respect to mock Real Time results are expressed as Fold Induction relative to the mock after normalization to endogenous human β-Actin mRNAs. Data represent means ± SD from at least three independent experiments performed in duplicate. * = $p < 0.05$; ** = $p<0.01$.

Figure 8

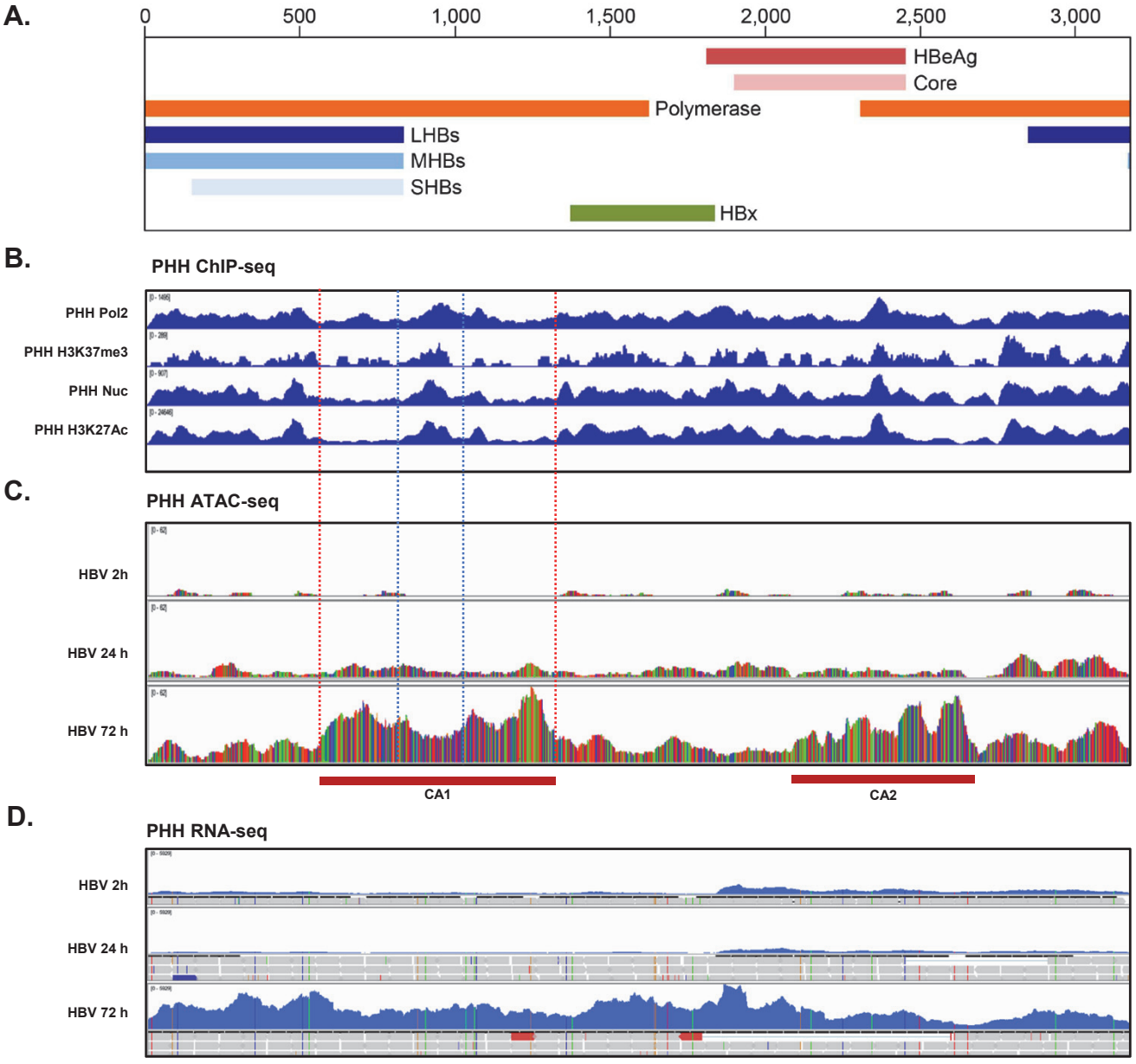
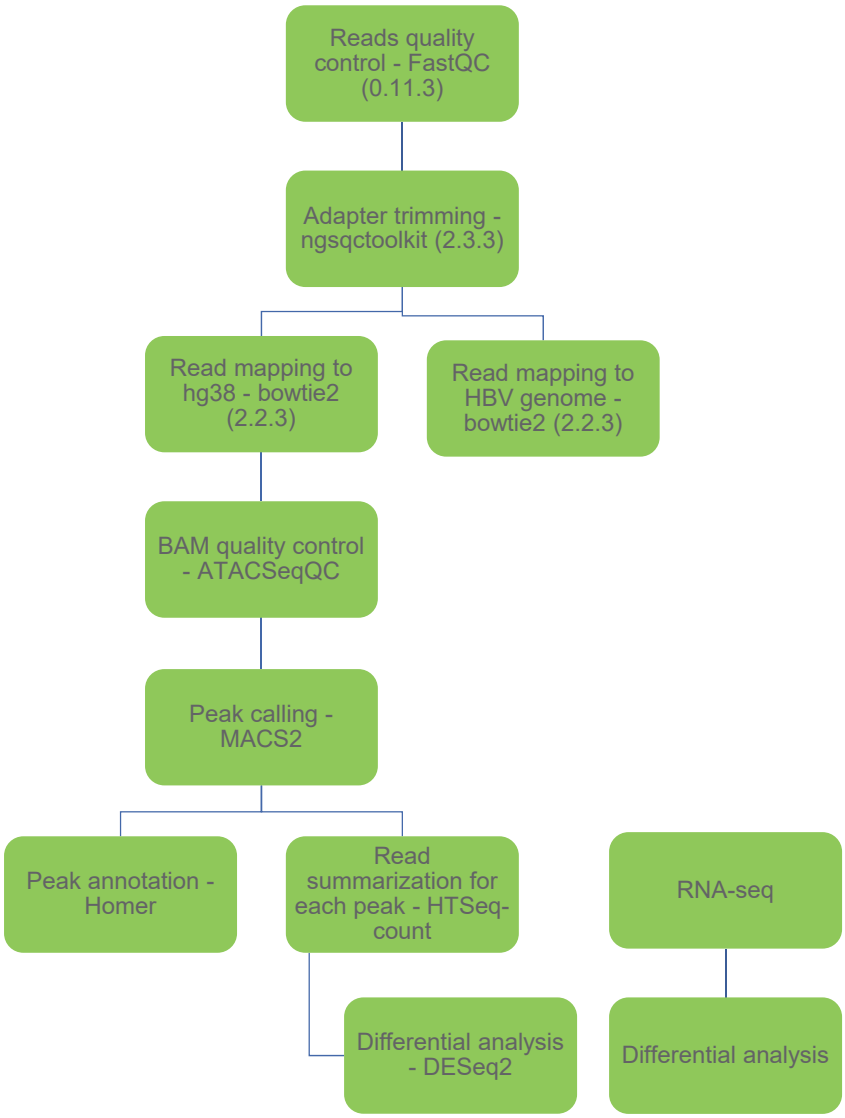


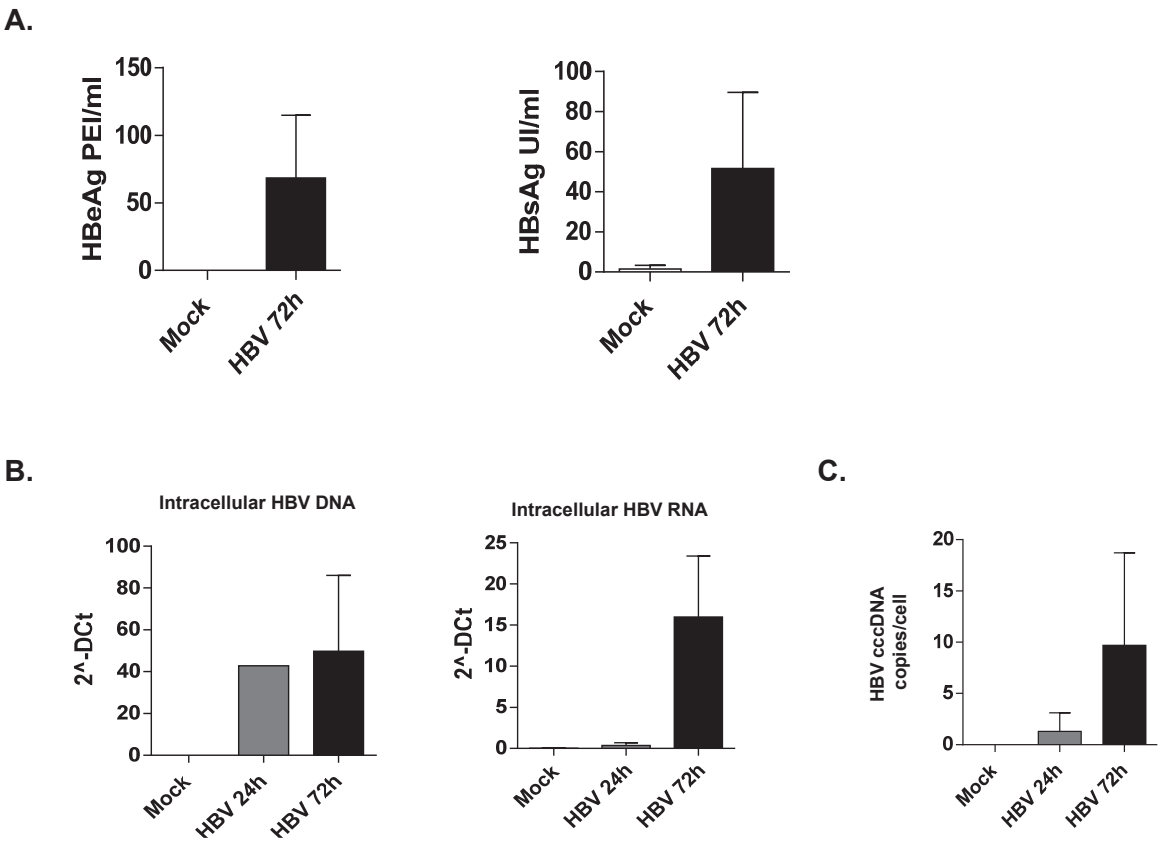
Figure 8. ATAC-Seq analysis of the cccDNA minichromosome in HBV-infected hepatocytes. ATAC-sec and RNA-sec read are aligned to the HBV genotype D genome (NCBI Reference: NC_003977.2) (**A**) to generate ATAC-seq (**C**) and RNA-seq (**D**) profiles in HBV-infected PHHs 2 hours, 24 hours and 72 hours post-infection. In **B.** are shown Pol 2, H3K27me3 and H3K27Ac cccDNA ChIP-seq profiles in HBV infected PHHs from Troberger et al. (2015). Nuc = nucleosome mapping on the cccDNA from Tropberger et al. (2015).

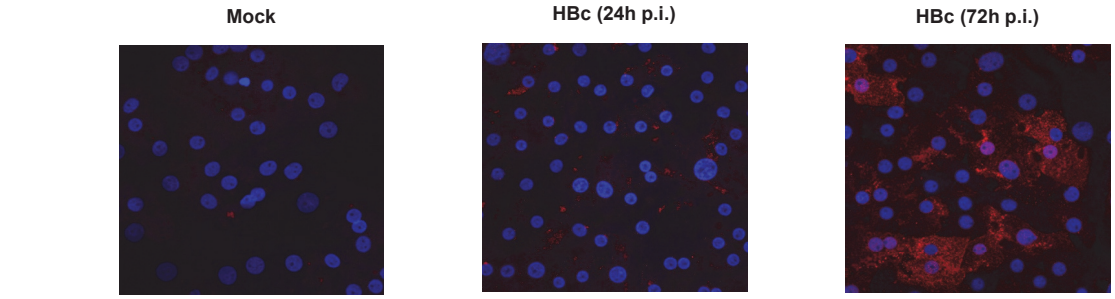
Figure S1



Supplementary Figure 1. ATAC-seq analysis workflow

Figure S2

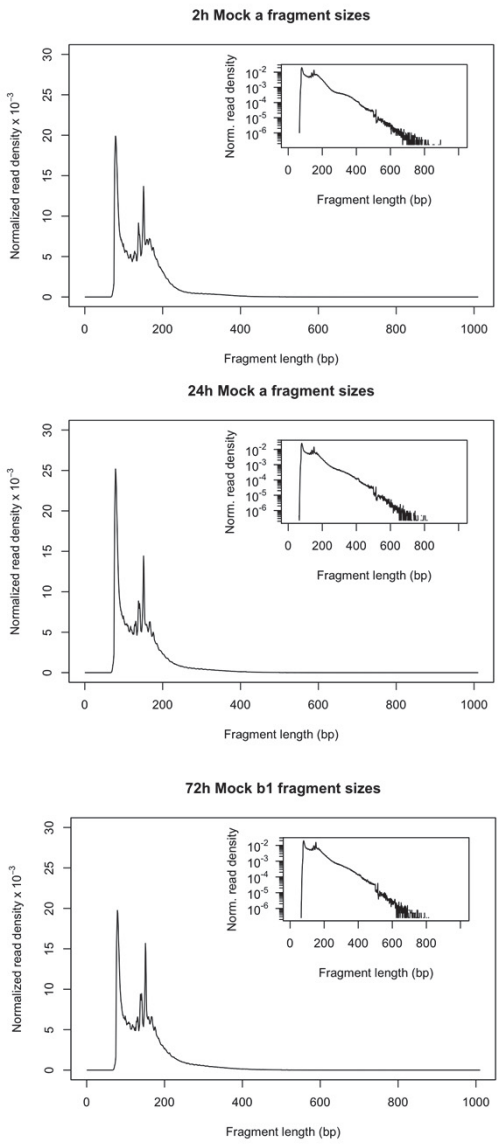




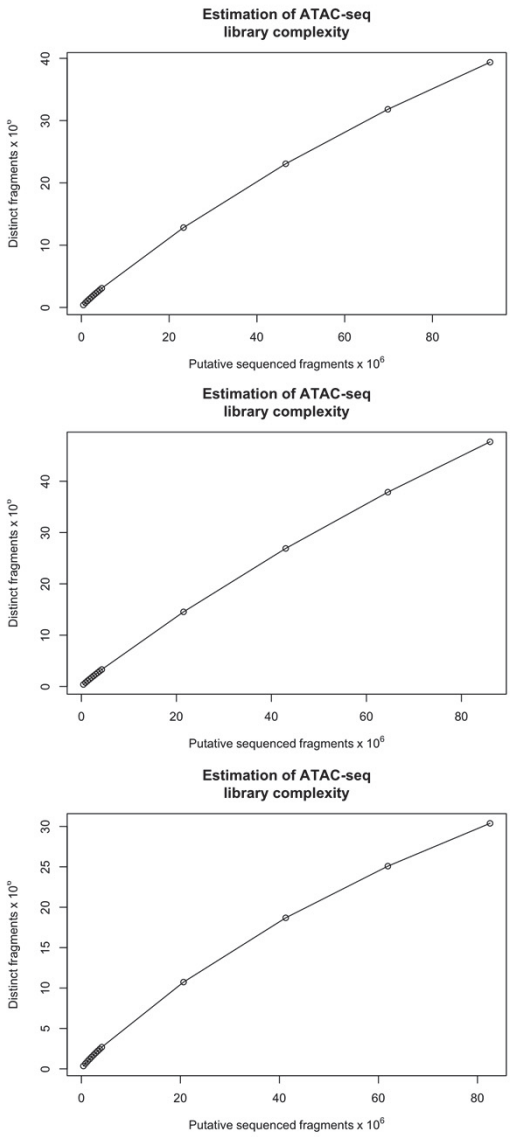
Supplementary Figure 2. HBV infection in PHHs. PHHs are infected at MOI 250 with a HepAD38-derived viral inoculum. Viral parameters are controled 24 hours and 72 hours post-infection. **A.** ELISA assay for secreted HBeAg (*left panel;* PEI/ml) and HBsAg (*right panel*; UI/ml). **B.** Total intracellular HBV DNA (*left panel*) and total intracellular HBV RNAs (*right panel*) were quantifued by qPCR (primers are detailed in Suplementary Table 1). Results are expressed as 2^-DCt relative to endogenous house keeping genes. **C.** HBV cccDNA levels were quantified by qPCR using cccDNA specific probes and primers (Supplementary Table 1). Results are expressed as copies / cell using HBV and β-Globin standards. **D.** HBcAg expression was assessed by immunoflurescence with an anti-HBc mouse monoclonal antibody (Abcam ab18684; 1:1000 dilution).

Figure S3

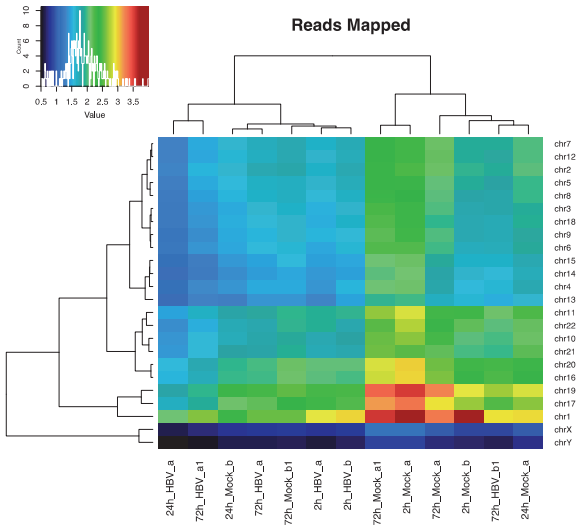
A.



B.



C.



Supplementary Figure 3. A. ATAC-seq fragment sizes generated from PHH nuclei after Tn5 transposition. **B.** ATAC-seq libraries complexity. x axis = putative sequenced fragments x 10^6; y axis = number of distinct fragments x 10^6. **C.** Distribution of filtered reads mapped to each chromosome.

Supplementary Table 1

Primers and probes for used for HBV DNA, HBV RNA and cccDNA quantification.

HBV	primers	Forward	GCTGACGCAACCCCCACT
		Reverse	AGGAGTTCCGCAGTATGG
RPLPO	primers	Forward	CACCATTGAAATCCTGAGTGATGT
		Reverse	TGACCAGCCCAAAGGAGAAG
GUS	primers	Forward	CGTGGTTGGAGAGCTCATTTGGAA
		Reverse	ATTCCCCAGCACTCTCGTCGGT
cccDNA	primers	Forward	CTC CCC GTC TGT GCC TTC T
		Reverse	GCC CCA AAG CCA CCC AAG
	probe	Forward	GTT CAC GGT GGT CTC CAT GCA ACG T
		Reverse	AGG TGA AGC GAA GTG CAC ACG GAC C
βGlobine	primers	Forward	ACA CAA CTG TGT TCA CTA GC
		Reverse	CAA CTT CAT CCA CGT TCA CC
	probe	Forward	CAA ACA GAC ACC ATG GTG CAC CTG ACT CCT GAG GA
		Reverse	AAG TCT GCC GTT ACT GCC CTG TGG GGC AA

Supplementary Table 2

List of unique genes showing an increased (UP) or decreased (DOWN) chromatin accessibility in HBV infected human primary hepatocytes

2 h post-infection		24 h post-infection		72 h post-infection	
Chromatin accessibility UP	*Chromatin accessibility DOWN*	*Chromatin accessibility UP*	*Chromatin accessibility DOWN*	*Chromatin accessibility UP*	*Chromatin accessibility DOWN*
MIR7977	C1QL2	PPP1R2P3	MIR3922	BCL2L14	MIR6076
LURAP1L	TAS2R1	CINP	DMRT2	MEOX2-AS1	TUBB4A
GRID1-AS1	LINC01254	HACE1	SSH2	OPA1-AS1	CDH6
MIR548BB	MIR924	DSG1-AS1	PTPN4	C14orf180	FAT3
MIR3170	AACS	TMEM167A	HLA-DQB1	ITGB6	GRIK4
GALNT15	EBF1	LINC00603	SNORA105A	NDST4	LRRC32
SEMA5A-AS1	LOC100294145	VRK2	EYA4	RNF219-AS1	LINGO2
MADCAM1	YTHDC2	TBC1D23	OAZ1	ADAMTSL1	LINC00494
LINC01309	MIR1202	DRD3	DNAH5	ROCK1	LINC01949
MIR4637	NBEAL1	APOLD1	DEPDC1	EIF3E	WWOX
HEATR9	LINC00238	CTD-2194D22.4	ZADH2	ELFN1	GZMH
MIR3152	MTCL1	TBC1D5	PTPRS	LOC93463	C3orf79
ARMC7	SAMD4A	LINC01127	TARS	ANKRD33	COPB1
LINC01029	BARD1	LOC105378098	STMND1	SPTBN1	RFX2
LINC00538	PMEPA1	TENM2	DST	DAZL	SELL
LOC102724957	IZUMO3	ITGA8	LOC105372028	LINC02026	METTL4
METTL4	CSMD1	ADAMTSL1	VASH1	LOC101929468	LINC01250
DLGAP1	ZDHHC21	FAM133CP	UBE2QL1	CNTNAP5	SEPT10
GPR25	PPP4R3A	MIR548AD	SLC6A3	NAALADL2-AS1	DPYD-AS2
LINC02111	SMU1	CDH18	SPRED2	SEPT7P2	LINC01121
HGC6.3	ATP6V1C1	LOC101927168	PLXNA2	FAM120C	ISY1-RAB43
MIR6083	SIAH3	OPCML	JARID2	LINC01608	PRG1
SMCO1	ACVR2A	PPM1A	LINC01239	SNX14	MIR3936
FUNDC2P2	WDFY2	CSMD3	KIF21A	COL6A4P2	PLBD1
CCR7	IPMK	E2F7	LRFN5	DMD	GDNF-AS1
NIPSNAP3B	MIR4454	GPR151	LOC101927697	LOC645166	OR4E1
LOC101928591	PLA2G5	MIR548AD	EDN1	ZNF385B	MYOM3
SUN1	GNG4	MGC32805	LSP1P3	MDGA2	LINC00967
CEP76	ITGA5	C7orf72	PPP2R2A	RUNX2	RCE1
PLIN3	MIR4324	APOL5	BRE-AS1	LOC644919	PROSER2
RXFP2	POTEA	AHCYL2	ACSM6	HYI	ELAVL2
CD28	ANKRD55	LINC01242	SCN7A	SMAD7	PPM1M
TLE6	AKAIN1	PIN1P1	LINC01714	LINC01258	KCNQ1
PEG13	NSFP1	WDR89	TRIO	PTPRD-AS2	NRIP1
AA06	MAP2	ARPC5L	WFDC3	DEPDC1	FCRL4
UTY	LINC00601	LINC00906	TEK	LCORL	LOC101929243
SF3B1	LRRC37A2	LINC01568	ROCK1	LOC646241	FBXO22
PCAT1	MAGEB16	OR5T2	TTC39C	ERCC6L	REEP6
ONECUT3	FNIP2	SEPT10	MIR4477A	LOC284395	MIR548AU
JPH1	FAM46A	CHL1	CD200R1	MIR2115	CLDN5
COQ10A	PTPRS	MSR1	PHEX-AS1	LRFN5	TICAM1
POM121L12	FGGY	FAM46A	MYOM1	LRRTM4	PACRGL
CDH18	GPR21	EGFR-AS1	LINGO4	LINC01913	PTPRD
LOC101928161	TOX3	ARHGAP29	MLANA	DAPK1	LRRTM3
NPSR1	TRIO	MT1M	RORA	LAMA4	PPP1R1C
ARL8B	SLC9A2	MEAT6	NDUFV2	FLJ37505	DCD
CX3CR1	LOC102723692	SNX5	DLEC1	UBBP4	PEX14
LOC101929122	TMEM45B	SMAD7	BTBD2	LCAL1	UHRF1
RGPD1	SLC7A11-AS1	LINC01848	SCHLAP1	LOC102546299	SLC1A4
NECTIN1	MEAT6	TINAG	DMRT1	RTKN2	PSMD6-AS2
TMC3	B3GALT1	BBOX1	RDM1	LINC01697	MIR1296
DNAH10	TMEM261	PRDM15	TRIP13	ST8SIA4	ZBTB16
ARID5B	CCDC71L	MS4A13	TRIO	DTL	RARB
LOC100131532	KIAA0040	VWC2L	LINC00367	HULC	REREP3
BAIAP2L1	LINC00999	FAM134B	SLC25A51	TMEM261	ANAPC10
LINC00906	LINC01060	CLUL1	OR2F1	EMB	EVL
VAX1	LOC101926915	LINC01395	TAF1L	TNR	TPD52L3
ACBD3	CA12	KCNK18	LINC01107	LOC100132078	PI4KAP1
CCBE1	LINC02052	MAPK13	LOC101928622	LINC01648	PLOD2
KIF16B	LINC01126	KCMF1	UBAP2	GAS6-AS1	ROPN1L
CSF3R	MIR3973	LOC105369486	SLC8A2	C3orf58	MIR1587
TTC39A	RAB4A	DICER1	LEPROT	CLECL1	NNMT
BTBD3	LINC01601	IL1RL1	SLC29A4	ACSBG1	ADAT3
OTUB1	RORA-AS2	MIR4473	SORT1	BAP1	TYRP1

2 h post-infection		24 h post-infection		72 h post-infection	
Chromatin accessibility UP	*Chromatin accessibility DOWN*	*Chromatin accessibility UP*	*Chromatin accessibility DOWN*	*Chromatin accessibility UP*	*Chromatin accessibility DOWN*
SNORA4	LINC00376	RPS29	LINC01121	PRR18	ABL1
DMRT2	NIT2	C19orf47	MIR4643	LSAMP-AS1	TSSK1B
ANKK1	LYPD1	LOC102724691	CELF5	LOC101927272	SYNDIG1
NRG1	PP2D1	GBX1	GTF2IP1	PACRG-AS3	MLIP
ATP2C1	RERE	LINC00261	SYAP1	LOC105372038	MOV10L1
XIRP2	SEPT9	CPNE8	MARCH6	SNORA107	FCGBP
KCTD3	JMJD1C	ABCG2	SF3A3	SSBP3	POM121L2
KIF26B	UBE2D1	LINC01519	LOC101929563	LOC105369739	TMEM261
MAP7	ZNF521	MAMDC2	LOC100507194	NCS1	MIR4306
GNE	MS4A10	LINC01123	MIR1267	PABPC4L	GHR
PLXNA2	ADCYAP1	STAMBPL1	CBX4	FBXL5	SELENOP
TNRC18	C5orf22	CTSO	DRD2	SUSD6	AMACR
MIR4777	LINC01602	GRM5-AS1	TRIO	TRAPPC11	LOC100506858
STAC3	NWD1	TMEM119	CYP46A1	UBE2U	LINC01781
TEK	EPAS1	MVB12B	CCDC140	ACER2	MBOAT4
EPS15L1		LOC101927588	NUDT12	VSTM2A	SVIL-AS1
DDX1		C7orf71	SH3YL1	FRMPD2	CTNND2
LINC00563		LINC01898	NNT-AS1	ZSCAN20	NDFIP1
GLDC		LOC101928535	RASSF5	LINC01468	PRMT8
IZUMO3		RPL13AP17	LYPD6B	SGIP1	FAM84B
CDH11		LINC01611	LINC00593	CSNK1G2-AS1	ABHD6
MISP		B3GALNT1	LOC105374972	FUT10	ETV6
LOC101928052		POU4F2	RORA	CYP2C18	SIK2
FBXO7		COL8A1	GLIS3	LINC01910	SIK2
SMYD3		SHROOM4	IMPG1	C4orf33	PSG7
ZSCAN30		MSC-AS1	KCNQ5-AS1	PRKCZ	PLEKHA7
MCF2L2		RAB18	IL15	PDE3A	PCSK2
LOC401134		LOC101928851	CXADR	UBBP4	SULF2
BRCAT107		LOC101927907	TP53TG3	FLJ26245	LINC00535
TMTC1		NHSL1	USP17L6P	LOC105378146	STMN4
SALL2		LRIG3	ANK3	LOC100506858	LOC105374704
FAM63B		LINC00906	SMARCA2	FNDC3A	ATP1A1
ATP2C2		CAT	FAM183BP	NANOS1	LINC02072
DNAJC25-GNG10		NRG1-IT1	CTNND2	ACAP2	MIR6788
RAB11FIP2		MIR592	PCCB	POLR2M	LINC-PINT
TMX3		MIR4438	CBX4	RSPRY1	SEMA6A-AS1
OR4E1		CCDC26	SLC41A3	ACO1	SLC2A2
DNAH7		BMPR1B	WASF3	LOC105371046	LINC01288
TRMT10C		MYF6	NTS	PPFIA2	LINC01600
HDHD2		DHFR2	EDN1	LRRTM3	ENO4
CELF2		ICA1L	LOC105377962	PCNX1	LINC01696
NCOA2		TTC6	NTRK3-AS1	ACTR3BP5	STBD1
IFT43		BASP1	ECT2L	FAT2	PSPC1
LOC645166		MRPS30	PLA2G4B	PDE4B	ARHGAP18
CAPZA3		DLGAP1	GIPC3	LOC105369860	RIC1
ACACA		SHB	LOC101928226	MIR2054	MIR548AC
HNRNPKP3		PKHD1	AKT3	MYT1L	KIF26B
HNRNPA3		MIR8066	LINC00603	CELSR2	ASCC3
RGS2		RAI14	LOC101929698	BANF2	LOC105376194
CRBN		ABCG2	ADGRL4	KCNA5	STUM
FREM2		ACTR6	LINC01256	PKIB	IL1R1
MIR1237		LOC105370605	LOC101927752	LOC441025	VGLL2
CHD1		LOC102723661	LINC01361	LINC01790	IMPAD1
C19orf25		LOC101927911	MIR187	GRID2	MTO1
EDN2		LINC01020	F5	LINGO1-AS2	RNF185
LINC01923		SEMA6A	PDE1C	ALG10	LINC01929
OSCP1		CSMD3	LEPROT	ROBO2	IMPA1
GNS		MAN1A1	MGC16025	HOMER1	DPP6
SLC1A5		LINC00613	PGR	UBE2NL	RND3
NCKAP5		LOC729506	CNTFR	KIAA1324L	FAM46A
IQGAP2		SH3BP4	ARHGEF10L	PPT1	OR56A1
MIR548AG1		LINC00458	LOC100130264	RPS6KC1	NFIB
BMPER		GNA11	GRHPR	VAPA	GOLGA2P2Y
UGT2A3		GMPR	PRSS48	POU4F3	ZNF665
FST		MYO6	LINC00550	MIR1260B	LOC645166
LINC01007		MAPRE2	ALG1L2	RAB3C	ZCCHC6
CTNND2		DOCK3	EDN1	SLCO4C1	LOC101927849
NPR3		PXDN	MIR3134	LINC00348	KCNE4
CD2AP		KIF16B	CDH18	TMEM120B	SLC17A6
LINGO2		LOC105369509	PEX5L	LINC01235	DLGAP1-AS4
LINC01654		PTGS2	SRD5A2	CARMIL1	LINC00200
MLLT10P1		LINC01192	CCDC71L	ATF3	SNORA87
LINC00836		LINC01243	FGF10	TUNAR	MIR3152
SLC25A51		ZNF8	CYP4F12	WDR18	NPAS2
NPR3		TBX18	DLGAP1-AS2	LINC01376	PSAT1
RUSC2		UCHL5	PIP	LRRIQ1	LINC00907

2 h post-infection		24 h post-infection		72 h post-infection	
Chromatin accessibility UP	*Chromatin accessibility DOWN*	*Chromatin accessibility UP*	*Chromatin accessibility DOWN*	*Chromatin accessibility UP*	*Chromatin accessibility DOWN*
KIF2A		ZEB1	LINC01091	KCNA4	LHPP
MIR4445		GPC1	RAB11A	DLGAP1-AS5	LRFN2
TM4SF1		MIR8054	MIR6776	NEPRO	LOC105373338
LINC01793		LINGO2	FAM105A	LINC00330	BMP3
MN1		ZNF727	LOC101928820	CNOT9	LOC100506422
FUT6		NRG1-IT3	TMCC1	ATOH1	TGFBR2
TFEC		LOC105374194	SMAD5	LINC01242	LINC01526
KCNAB1		LOC285766	YTHDC2	PREP	MIS18BP1
LOC389705		CNTN6	ANK3	POTEA	PTBP2
PFKP		LINC01243	NELL1	TRY2P	PTPRD
CDH26		TMEM179	LOC101927915	API5	SPEF2
MIR548AG2		ST8SIA6	BRD3	MIR4666B	KCNMA1
OSTF1		HACE1	COBL	GRIK2	LINC01243
KIF4B		ATXN2	SNORA33	ZMAT4	NRSN1
CMSS1		IMMP2L	CNTFR-AS1	FGF14	TIPARP
CCNYL2		RER1	CYP3A7	ZNF716	C1orf198
EDIL3		LINC01813	MKRN9P	ALG10	TSPAN12
GRIN2D		LINC00442	OCA2	PDE5A	GLUL
LOC100506858		MIR7848	MIR4454	RIMS2	LINC02140
MIR924		LOC102724874	OPRK1	RHOBTB2	RRN3P2
ATL2		ROR2		LINC00710	LOC105370369
CD226		KCNN2		FSHB	MIR3691
DNAH5		RAB27B		LINC00243	RBM47
ZBTB20-AS3		LOC100507377		CXCL1	FASLG
C4orf26		TRPS1		SLC16A7	LOC101928618
LINC01029		MYOF		FUT1	A4GALT
NLGN4Y		LOC105377143		CLIP4	OR2S2
RNLS		HSP90B2P		GRIA2	SELENOP
BEND5		FAM8A1		LRRTM3	SLC1A3
CCNY		PMEPA1		MIR1913	FAM135B
LINC01182		ROR2		REG3A	MYO3A
MIPOL1		TCF4		BMPR1B	LOC101927901
FAM170A		GBE1		PHLDB2	BDKRB1
LOC101927078		PNLIP		PEG10	S100A7
TRPM2		SETBP1		MIR548AA1	STN1
TMCC1		APOOP5		PSAT1	MIR3977
KIAA1217		SLC9A4		CCNI2	LOC439933
FAM201A		MIR887		HS3ST1	SMG6
LINC01299		PSMB7		CAMK2D	RBM43
MIR5582		ZNF385D-AS2		URB2	LOC100287072
LOC105379514		SNORD15B		LOC105377267	LINC00865
RNF180		CLVS2		NRXN1	TBX5
TUBB6		SP140L		LOC100996724	MIR4743
LOC101927450		KLHL3		C9orf72	SHD
LINC01243		AGPS		LOC101929284	FLNB-AS1
NEGR1		MIR4782		SNORA98	FGF20
LOC103352541		CPS1		LOC101927188	KCNAB2
STPG2-AS1		DMRT2		MIR4431	LOC100505797
PIN4P1		PATE1		RHOA	ARSJ
TUSC5		IRAIN		RBM19	KCNS3
PTPRD-AS2		MIR6073		PRKDC	SAXO1
RANBP3L		LOC91450		L3MBTL3	NAV2
LINC02027		TMEM106B		VRK2	LOC102723778
MTRNR2L6		ATRX		HEY2	MIR3942
SLC4A7		GJB2		OXGR1	FAM181B
LOC101929563		ANO4		LMO7DN-IT1	TNPO3
NGF		UBAP1L		ZNF544	TAPT1
NBEA		CHAC2		TPTE	MTCL1
GLI3		DLEU7-AS1		SUCLG2	FAM47E-STBD1
CEP192		LNCPRESS2		MRPL46	SLC1A1
KDM6A		LRRC17		SPTSSB	PTPRD
STPG2-AS1		SLC35G2		ANKRD20A4	LOC440311
LINC01242		BPY2		FAT4	WASHC2A
SULT1C2P1		INTU		FAM230C	RIC1
		CIITA		LINC00882	CDH8
		HACE1		OR4F5	MIR7641-2
		SRP14		LINC00163	ASPA
		TFEB		LOC100130331	LSM8
		CNTNAP2		POMZP3	EXOSC3
				CD99	DKFZP586I1420
				LINC01207	
				GMPS	
				C15orf32	
				PICK1	
				MIR3974	
				DAP	

2 h post-infection		24 h post-infection		72 h post-infection	
Chromatin accessibility UP	*Chromatin accessibility DOWN*	*Chromatin accessibility UP*	*Chromatin accessibility DOWN*	*Chromatin accessibility UP*	*Chromatin accessibility DOWN*
				LOC102467226 ANKRD36C SPOCK1	

RAB37, RAD51AP2, GPT2, TMEM184C, GRIFIN, CCNT2, LINC00290, LINC01103, LINC00898, TESMIN, PTX3, TPO, LOC100287632, PTCHD4, LOC101927243, MIR4459, LOC153910, LOC101927588, SERPINA7, PLCL2, LINC01039, LOC105369187, ACAT1, LOC105378146, TYR, KIF5C, LOC105376430, NOBOX, SMG1P2, SSBP2, SPATA4, CNTNAP5, TENM4, SREBF2, IRX1, MIR548BA, MSL3, LNX1, SNX16, GATSL2, LINC01435, LINC01085, EFHB, MEIS2, ZFYVE28, LOC643711, ASZ1, DIO3OS, EMB,SCG5, BCKDHB, IQCA1, USP22, KCNN3, LRRC4C, LOC101927557, INTU, AP1S3, C3orf67-AS1, COBL, PCDHA4, LOC105375972, MIR4465, KCNJ3, NEB, SMARCA2, VIPR2, CD180, ATP5G3, GIMD1, EDN2, ZNF733P, ACTG1P4, DCDC2C, MME, ROBO2, TNFAIP8L3, CDH9, RAP2A, TRAM1L1, ANGPT1, LOC644919, NIPBL-AS1, IPMK, LOC101928137, BNC2, ROCK1P1, SCML4, RAB35, LINC00366, RAB28, CKMT2-AS1, ARIH2OS, ANO3, EMBP1, LINC01411, LINC02015, C19orf35, KLF12, LOC105378269, SNAR-H, SCUBE3, SNORA107, LINC02114, TFPI, STXBP6, FRG1CP, GBA3, RGS18, SREK1, DMRTA1, GGT7, PLCE1-AS2, CCDC167, MYO16-AS1, ATOH1, ZNF733P, KL, MPO, CTTNBP2NL, KDM4A, LINC00680, ZNF407, LINC00578, FAM13A-AS1, LRP1B, LOC400940, TRIM42, FAM27B, LINC01246, FXYD6, NPTXR, KLF9, LOC100507468, LINC01760, AKAP7, DNAJC6, HCG25, LOC105369911, MAF, MIR4439, LINC00867, UNC5C, GALNT2, MTHFD2P1, DACH1, BBX, LINC00371, CD226, MIR4268, RNF7, FAM133CP, AKAIN1, MROH2B, TFG, REG3A, LOC101928944, ADGRL1, DEPDC1, MBL2, MIR3976HG, DNAH11, TRAF6, LOC100287944, MIR5007, PDE10A, TRIQK, NOX4, FAM46C, LINC01793, MTRNR2L3, LOC401134, BMP3, LINC00316, EXT1, PRICKLE2, TMEM106B, MAPKAP1, MIR4426, NBPF20, CDRT15P1, MIR378H, KLF5, MIS18BP1, ANXA5, LINC01592, GOT2, CWH43, DOCK3, RGS21, MIR548AB, LOC102723778, CALB2, LINC01608, PTPRD-AS1, LRRTM4, PRDM13, LINC01087, ZNHIT6, ARHGAP15, YWHAEP7, CACNG5, LINC01605, LINC00680, GOLGA6C, PLEKHM2, P2RY1, STXBP6, TOMM70, YIPF7, LOC105369785, SLC44A1, LOC101927686, NKAIN2, LINC01625, KIF13B, LINC02035, MVK, FEZF2, FAM71C, LOC101927815, ZFHX4-AS1, ETNK1, LOC105378183, AASDH, MCTP2, SLC9A9-AS1, MTRNR2L6, ATP8A2, MIR6124, PRG4, AXIN2, LOC101928944, PRKCH, ZFPM2-AS1, CBLN1, BMPER, LINC00347, TRHDE-AS1, RNU2-1, PDGFC, RAG2, LOC105376365, ELF2, LOC100996349, UBE2E3, XPNPEP2, MYO1B, NR2F1-AS1, CCSER1, MEAT6, AREG, CDH10, LINC01118, FRG1CP, GATA5, TBC1D4, FTMT, LOC105374704, KCNK12, LINC01959, SEMA6D, YTHDC2, SPATA4, SEC31B, LOC101929621, LINC00392, DSCAM-IT1, GNPDA2, MBOAT1, LOC105369891, LOC105370369, NPVF, EGFLAM-AS4, AQP4, LINC00972, LINC00907, CDH8, LINC01676, LOC102467214, MUT, TCF7L1, LINC00693, RPTOR, FXYD4, ARX, LINC01192, CD48, MIR1267, EMB, LRP1B, LINC01309, GAREM2, LOC101927078, MGAT4C, SNTG1, MIR4774, LOC340017, CXXC4, LINC01947, LVCAT5, DCN, TRIB2, KCNH7, BTC, CABLES1, ZBBX, LMX1A, MIR8084, GHSR, LOC101928433, LOC101928052, RNU105B, APBA2, EGFLAM, LOC105370586, ZNF536, TMEM45A, PDE4DIP, HCN1, SNORA98, POTEA, MOXD1, PTPRD-AS2, BNC2, LINC01361, LOC257396, MIR4650-1, LINC00456, DISC1FP1, RBFOX1, EDIL3, LOC101928372, LOC101927849, MIR4735, DLC1, IL15, DUSP6, TRIO, DBET, CCR4, LINC00636, CDH6, PRR16, GLRA1, COL11A1, LOC439933, MDGA2, LINC01602, PLXNC1, MBL2, LOC100287010, LINC02032, BLOC1S5, LOC100129138, SRGAP2C, HMCN1, RNLS, CCT5, C22orf24, GABBR2, QKI, C10orf107, SMAD5, ACTR3BP5, LINC01242, LINC00358, LINC01297, TEX26, LOC653712, PDPR, RAMP3, LOC440040, NLGN1-AS1, OR4F17, FRG2EP, FBXO15, GRIN2B, OR2T33, ADGRA2, MIR4445, AFF3, PRR20B, LRRC37A, ARHGAP32, PDGFC, LINC01600, RAD18, DPP10-AS3, BCL2L11, DUSP26, DIO2, LINC00879, PKHD1, KDM4C, RAMP3, KCNH8, SMIM23, NEU4, LINC02103, TRIB2, PLAC4, HCN1, LOC100192426, SND1, LINC01052, ZNF733P, SH2D7, LYST, LHFP, TPK1, PCDH15, ACO1, LOC101928052, LINC01592, LINC00587, FRG1CP, MTAP, LOC101928052, TMEM266, MIR499A, SERINC1, RABGAP1L, RP1L1, RIIAD1, PACRG, ABCA8, LOC102723439, LIMS1, MIR1263, SAMD9, MCC, MIR8066, LOC101927100, LINC00924, NAB1, HIP1, CIB4, HSPA12A, OCSTAMP, RNF4, LOC339593, MTUS2, MMP16, TMEM261, SIK2, FGF14, CTSO, MIR1973, HNRNPKP3, CXCL12, SUCNR1, LINC02085, RTKN2, PLD5, MIR3170, LOC102725080 IL15, LINC00434, ARG1, GIMD1, COL6A3, MIR7157, VPS4B, MIR302F, C1QTNF3-, AMACR, ANO5, TAGAP, FRG1CP, ZNF717, LNCPRESS2, HTR1B, CYP2C19, LINC00293, SCG2, SYN2, DIAPH3-AS1, MIR2054, MTSS1, PTN, SORCS3, PRKAG2, CAPS2, MAGED4B, TDO2, KIF17, GNPDA2, SCN3B, SOX1, NFASC, C5orf47, DRD1, LOC101926942, GPR176, LOC100131496, ETS1, CSMD3, LOC101927915, LINC01252, GALNT1, MIR7641-2, LINC01207, LOC100133920, CTNNA3, AHCTF1P1, MIR147A, MIR4643, UCK2, FAM170A, RANBP3L, LOC102467222, EMBP1, MIR4641, CCNY, LOC101927269, ZAN, ZNF716, GLYATL1, LAMC1, CACNA1A, OR7E156P, ACTR3BP5, EPHA7, MAT2B, IQCJ, METRNL, NPY5R, LOC644285, TMCO5A, ARHGAP42, KIF26B, LINC01847, LOC100507534, LOC101929470, LOC441666, LOC389705, ANKRD20A5P, SYT10, CYP7B1, LILRA5, PLOD2, LNPEP, LINC00273, LINC02120, LOC101928510, RUFY1, LINC00922, WNT16, PON1, TACR3, RNGTT, CYP4Z1, SLC16A7, CST9L, LOC100133920, FGF8, LINC02126, FAM95C, DIRC3, LOC101927050, NXPH2, LOC100507377, ROBO2, GFRA1, TMED4, LINC00423, LINC00593, SLC8A1-AS1, FRAS1, NOVA1, BOLA3, C8orf87, STN1, XIAP, NT5C1B, LINC01372, TYRP1, LUCAT1, UGT8, XIRP2, MIR4459, DCLK2, GPC6-AS2, CTNNA2, HDDC2, LINC01821, DNAH6, WDR18, NMBR, NXPH1, C11orf49, LOC90768, LOC101927691, DPYS, LOC100506675, LINC02120, TMEM71, MIR4799, LINC01324, USP12-AS1, IL1RAP, GBE1, PLCL2, AASDH, LINC02059, CACNA1I, MET, CNTN6, DHFRP3, MIR4465, TBC1D14, NBPF8, MAPRE1, SP4, INPP4B, HTR2A, CXCL8, LINC00609, CNTNAP5, LINC00491, LINC00363, ATP5G3, XKR9, LOC644919, CCT6B, BMPER, LOC643542, LINC00273, RGS17, RAI14, LOC644669, OR6P1, NPY, FAM193A, LOC105369509, ARMC2-AS1, MIR6888, LOC101929721, MRC1, MTRNR2L1, LINC02032, LINC01309, LOC101929411, LSM6, MIR592, LINC01372, XIRP2, PCDH7, C5orf52, SSX2IP, NRG1-IT3, LOC102467222, WSCD2, LINC01582, VSX2, BCOR, LOC10013033,1 VWCE, CNTLN, BBS7, NFU1, MEI4, PI4K2B, FAM19A2, UBASH3A, LOC105374704, PPARGC1A, LARP1B, MIR5007, OR8S1, LOC101927069, EPAS1, LOC101927087, DIAPH3-AS1, UBE2E1, ZNF638, SNORA70B, LINC01098, MAP3K7, CHRNA9, STPG2, TRIML1, RRM1, MRPS25, LINC01950, GNB5, TET2-AS1, ASIC5, FAM230C, VAV1, ATMIN, LOC101927822, LINC01209, PRKAR1A, LINC01781, PAK5, OSTF1, FRG1CP, CTNND2, TTC29, LINC01474, GBE1, LOC102723886, VCY1B, FYN, SLC9A2, ARHGAP26- AS1, MIR4500, SLC39A12-AS1, ADAMTS16, GRID1-AS1, OLFM3, CNR1, BTNL8, DCDC5, LVCAT1, MIR4473, ZNF716, CDH12, LINC00997, LOC105372028, LOC100506422, CDH19, MIR4454, ALG10B, C2orf76, LOC101928944, GORASP2, DACH1, MANEA-AS1, OR4D11, SMA4, ESR1, OR8K1, LOC729506, ARSB, RFC3, CHRNB3, RAI2, LINC01393, LINC00316, POU4F3, IPO11-LRRC70, MIR135A2, OR10W1, WWOX, STOML3, SLC39A9, GFRA4, SH3GL2, NLGN1-AS1, CERS6, SOX5, TCOF1, NEDD1, DENND1A, BRCAT107, VPS13D, PLCE1-AS2, C7orf71, SLC8A1, LINC00290, LINC00613, RPL38, MIR4454, RIMS2, LINC01242, TENM4, LINC00879, BMP7-AS1, LINC00564, NEIL3, ALG10B, C8orf34-AS1, CDS1, LOC101929116, LINC01378, MIR5702, OR7E37P, KCNC2, MIR548AU, EIF5A2, LINC01184, WDFY3-AS2, RPS7P5, RERGL, LOC101929550, TRIM49B, CCDC144NL- AS1, ATOH1, TLE1

Supplementary Table 3
List of genes with altered chromatin accessibility 2 hrs post-infection

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
MIR7977	0.00381731335508109	4.01	chr3	176417098	176417364	-97872	NR_107017	microRNA 7977	ncRNA
LURAP1L	0.00458997926056706	4.14	chr9	12749907	12750080	-25020	NM_203403	leucine rich adaptor protein 1 like	protein-coding
GRID1-AS1	0.00484159754509172	4.42	chr10	85632270	85632468	54638	NR_038986	GRID1 antisense RNA 1	ncRNA
MIR548BB	0.00538768001644301	4.52	chr3	60927059	60927418	-309368	NR_128708	microRNA 548bb	ncRNA
C1QL2	0.00551560415907369	-4.37	chr2	119174459	119174706	-15687	NM_182528	complement C1q like 2	protein-coding
MIR3170	0.00555495132467479	3.07	chr13	98293366	98293969	85143	NR_036129	microRNA 3170	ncRNA
GALNT15	0.00561284682633775	3.59	chr3	16090360	16090550	-83867	NM_001319051	polypeptide N-acetylgalactosaminyltransferase 15	protein-coding
SEMA5A-AS1	0.00583779393888807	3.8	chr5	9348536	9348772	-162679	NR_109935	SEMA5A antisense RNA 1	ncRNA
MADCAM1	0.00625907116609458	3.42	chr19	497912	498114	1523	NM_130760	mucosal vascular addressin cell adhesion molecule 1	protein-coding
LINC01309	0.00633566897912118	4.07	chr13	104260303	104260532	835217	NR_126374	long intergenic non-protein coding RNA 1309	ncRNA
MIR4637	0.00672262786626071	3.67	chr5	14788132	14788331	37781	NR_039780	microRNA 4637	ncRNA
HEATR9	0.00689622329168422	4.29	chr17	35847833	35848013	20968	NR_135630	HEAT repeat containing 9	protein-coding
MIR3152	0.00696477009930733	4.27	chr9	18672621	18672831	99420	NR_036107	microRNA 3152	ncRNA
ARMC7	0.00770047879474616	3.22	chr17	75115856	75116167	6059	NM_001304273	armadillo repeat containing 7	protein-coding
LINC01029	0.00788854837756934	4.47	chr18	77948255	77948416	45392	NR_104127	long intergenic non-protein coding RNA 1029	ncRNA
LINC00538	0.00792470016319152	3.76	chr1	213572902	213573087	-351755	NR_046189	long intergenic non-protein coding RNA 538	ncRNA
LOC102724957	0.00813588640536662	3.85	chr11	15559859	15560015	-83948	NM_001293172	uncharacterized LOC102724957	protein-coding
METTL4	0.00825869458009818	3.71	chr18	2171760	2171950	399648	NM_022840	methyltransferase like 4	protein-coding
DLGAP1	0.00848351359369875	3.53	chr18	4607123	4607354	-151972	NM_004746	DLG associated protein 1	protein-coding
GPR25	0.00853928705231771	3.37	chr1	200850659	200850839	-22206	NM_005298	G protein-coupled receptor 25	protein-coding
LINC02111	0.00857999983235076	4.23	chr5	17391787	17391976	-4571	NR_033975	long intergenic non-protein coding RNA 2111	ncRNA
HGC6.3	0.00910757621204706	4.23	chr6	167941942	167942104	34916	NM_001129895	uncharacterized LOC100128124	protein-coding
MIR6083	0.00942181323365116	3.71	chr3	124263467	124263685	-110756	NR_106731	microRNA 6083	ncRNA
SMCO1	0.00983457364442926	3.2	chr3	196537544	196537769	-22290	NM_001077657	single-pass membrane protein with coiled-coil domains 1	protein-coding
FUNDC2P2	0.0099020642054713	4.18	chr2	84240238	84240411	-50358	NR_003663	FUN14 domain containing 2 pseudogene 2	pseudo
CCR7	0.0101263427235449	3.7	chr17	40533224	40533393	27086	NM_001301718	C-C motif chemokine receptor 7	protein-coding
NIPSNAP3B	0.0105324239299598	2.88	chr9	104835231	104835400	71574	NR_130759	nipsnap homolog 3B	protein-coding
LOC101928591	0.0107676013306919	3.27	chr11	41902522	41902719	182838	NR_135065	uncharacterized LOC101928591	ncRNA
TAS2R1	0.0108004033358542	-4.32	chr5	9741678	9741835	-111405	NM_019599	taste 2 receptor member 1	protein-coding
SUN1	0.0110226874168709	3.63	chr7	844346	844566	11951	NM_001130965	Sad1 and UNC84 domain containing 1	protein-coding
CEP76	0.0114608986970246	4.28	chr18	12686756	12687030	15884	NM_024899	centrosomal protein 76	protein-coding
LINC01254	0.0115303720132601	-3.89	chr18	10413077	10413265	1199	NR_110775	long intergenic non-protein coding RNA 1254	ncRNA
MIR924	0.011768444306112	-4.09	chr18	39111096	39111458	510898	NR_030628	microRNA 924	ncRNA
PLIN3	0.0122382295710575	2.78	chr19	4861289	4861534	6357	NM_001164189	perilipin 3	protein-coding
RXFP2	0.0128058476663498	3.4	chr13	31711379	31711601	-28052	NM_001166058	relaxin/insulin like family peptide receptor 2	protein-coding
AACS	0.012994649672535	-3.66	chr12	125057741	125057961	-7516	NM_001319839	acetoacetyl-CoA synthetase	protein-coding
EBF1	0.0133910769079382	-3.36	chr5	158769199	158769380	330473	NM_001324109	early B-cell factor 1	protein-coding
LOC100294145	0.0135403765102475	-4.03	chr6	32916595	32916800	22521	NR_037178	uncharacterized LOC100294145	ncRNA
CD28	0.0135733884962174	3.14	chr2	203693486	203693644	-12910	NM_006139	CD28 molecule	protein-coding
TLE6	0.0144519992188119	2.93	chr19	2969093	2969333	-8320	NR_138612	transducin like enhancer of split 6	protein-coding
YTHDC2	0.0152282845664534	-3.96	chr5	113835424	113835612	321856	NM_001345975	YTH domain containing 2	protein-coding
PEG13	0.0155046635713508	3.61	chr8	140169401	140169578	-68954	NR_144431	paternally expressed 13	ncRNA
AA06	0.0156020082361353	3.3	chr17	33627373	33627571	-93712	NR_037584	uncharacterized LOC100506677	pseudo
UTY	0.015650771247608	3.59	chrY	13552972	13553139	-72385	NM_001258257	ubiquitously transcribed tetratricopeptide repeat containing, Y-linked	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
SF3B1	0.0157938939667043	3.46	chr2	197407219	197407669	27649	NM_001005526	splicing factor 3b subunit 1	protein-coding
PCAT1	0.016062986791325	2.84	chr8	126851525	126851703	-161540	NR_045262	prostate cancer associated transcript 1 (non-protein coding)	ncRNA
ONECUT3	0.0162183039850832	2.97	chr19	1733091	1733351	-20442	NM_001080488	one cut homeobox 3	protein-coding
JPH1	0.0162640673322914	3.83	chr8	74265250	74265507	55983	NM_001317830	junctophilin 1	protein-coding
COQ10A	0.0163095472265002	3.89	chr12	56278201	56278383	11009	NM_001099337	coenzyme Q10A	protein-coding
POM121L12	0.016703377375279	3.17	chr7	53386177	53386332	350598	NM_182595	POM121 transmembrane nucleoporin like 12	protein-coding
MIR1202	0.0169323590774415	-4	chr6	155910193	155910387	-36507	NR_031606	microRNA 1202	ncRNA
NBEAL1	0.0169630520140878	-4.27	chr2	203087541	203087710	72746	NM_001114132	neurobeachin like 1	protein-coding
CDH18	0.0175241370630679	2.91	chr5	20899130	20899293	-323338	NM_001291956	cadherin 18	protein-coding
LINC00238	0.0180788983093802	-3.92	chr14	65979808	65980072	-506431	NR_110307	long intergenic non-protein coding RNA 238	ncRNA
LOC101928161	0.0183432177799607	3.09	chr2	133267149	133267409	1084	NR_110294	uncharacterized LOC101928161	ncRNA
NPSR1	0.0185919402321231	3.57	chr7	34690047	34690466	31971	NM_207172	neuropeptide S receptor 1	protein-coding
ARL8B	0.0186048029748406	3.46	chr3	5129145	5129344	6999	NM_018184	ADP ribosylation factor like GTPase 8B	protein-coding
CX3CR1	0.018633047596395	3.2	chr3	39250780	39251026	29133	NM_001337	C-X3-C motif chemokine receptor 1	protein-coding
LOC101929122	0.0186898310853942	3.42	chr6	159179715	159179891	-9796	NR_121668	uncharacterized LOC101929122	ncRNA
MTCL1	0.0187763491674134	-3.05	chr18	8747338	8747756	30176	NM_015210	microtubule crosslinking factor 1	protein-coding
RGPD1	0.0199860775392017	3.44	chr2	87979082	87979530	6484	NM_001024457	RANBP2-like and GRIP domain containing 1	protein-coding
SAMD4A	0.0203273908285343	-3.53	chr14	54727945	54728161	-26735	NM_001161577	sterile alpha motif domain containing 4A	protein-coding
NECTIN1	0.0207393918102737	2.74	chr11	119713042	119713365	15522	NM_203286	nectin cell adhesion molecule 1	protein-coding
TMC3	0.0211899455492221	2.67	chr15	81373886	81374186	41	NM_001080532	transmembrane channel like 3	protein-coding
DNAH10	0.0212210227934924	3.19	chr12	123739229	123739388	-23187	NM_207437	dynein axonemal heavy chain 10	protein-coding
BARD1	0.0214866661760563	-3.83	chr2	214695925	214696089	113704	NM_001282543	BRCA1 associated RING domain 1	protein-coding
PMEPA1	0.0215508263186894	-3.81	chr20	57753466	57753624	-42009	NM_199171	prostate transmembrane protein, androgen induced 1	protein-coding
ARID5B	0.0215735653314915	3.18	chr10	61951253	61951839	50292	NM_032199	AT-rich interaction domain 5B	protein-coding
IZUMO3	0.0217135423653089	-3.17	chr9	25004212	25004409	-458634	NM_001271706	IZUMO family member 3	protein-coding
LOC100131532	0.02190895107796	3.27	chr6	170001164	170001323	38620	NR_027434	uncharacterized LOC100131532	ncRNA
BAIAP2L1	0.0219747935205575	2.57	chr7	98470328	98470693	-69395	NM_018842	BAI1 associated protein 2 like 1	protein-coding
LINC00906	0.0221634752632242	2.82	chr19	28901156	28901536	-63785	NR_027318	long intergenic non-protein coding RNA 906	ncRNA
VAX1	0.0222024319519768	3.64	chr10	117140556	117140726	-2340	NM_199131	ventral anterior homeobox 1	protein-coding
ACBD3	0.0222860444711419	3.1	chr1	226168633	226169223	17794	NM_022735	acyl-CoA binding domain containing 3	protein-coding
CCBE1	0.022342514659669	3.74	chr18	59548146	59548297	149191	NM_133459	collagen and calcium binding EGF domains 1	protein-coding
KIF16B	0.0225048458543974	3.41	chr20	16533606	16533804	39729	NM_001199865	kinesin family member 16B	protein-coding
CSMD1	0.0225512072869785	-3.97	chr8	5128053	5128245	-133343	NM_033225	CUB and Sushi multiple domains 1	protein-coding
ZDHHC21	0.022738231276169	-2.69	chr9	14652894	14653067	40502	NM_178566	zinc finger DHHC-type containing 21	protein-coding
PPP4R3A	0.0231081967953466	-2.84	chr14	91581484	91581713	-71118	NM_001284280	protein phosphatase 4 regulatory subunit 3A	protein-coding
CSF3R	0.0231393528543887	2.62	chr1	36630394	36630582	-147174	NM_000760	colony stimulating factor 3 receptor	protein-coding
SMU1	0.0231914570931774	-2.62	chr9	33101303	33101482	-24676	NM_018225	DNA replication regulator and spliceosomal factor	protein-coding
ATP6V1C1	0.0233522712102892	-3.33	chr8	103006168	103006417	-14728	NM_001695	ATPase H+ transporting V1 subunit C1	protein-coding
TTC39A	0.0238515719984039	2.93	chr1	51314236	51314448	7926	NM_001297665	tetratricopeptide repeat domain 39A	protein-coding
SIAH3	0.0238988781623651	-3.77	chr13	45836099	45836491	15416	NM_198849	siah E3 ubiquitin protein ligase family member 3	protein-coding
BTBD3	0.0239986459773624	3.33	chr20	12048769	12048939	130989	NM_001282551	BTB domain containing 3	protein-coding
OTUB1	0.0240298242812867	2.4	chr11	64036463	64036699	50728	NR_003089	OTU deubiquitinase, ubiquitin aldehyde binding 1	protein-coding
SNORA4	0.0245460473274974	2.95	chr3	186791453	186791625	3926	NR_002588	small nucleolar RNA, H/ACA box 4	snoRNA
ACVR2A	0.0247274384100102	-3.86	chr2	147564700	147564870	-279732	NM_001278579	activin A receptor type 2A	protein-coding
DMRT2	0.0247560942356952	3.24	chr9	1472733	1472897	422195	NM_181872	doublesex and mab-3 related transcription factor 2	protein-coding
ANKK1	0.0248496243226195	2.47	chr11	113387427	113388058	-49	NM_178510	ankyrin repeat and kinase domain containing 1	protein-coding
WDFY2	0.0250583831717412	-3.06	chr13	51684021	51684790	100057	NM_052950	WD repeat and FYVE domain containing 2	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
IPMK	0.0251470595883928	-2.81	chr10	57992705	57992930	275117	NM_152230	inositol polyphosphate multikinase	protein-coding
NRG1	0.0251998616984254	3.12	chr8	32966652	32966811	244978	NM_001322197	neuregulin 1	protein-coding
ATP2C1	0.0253260825039996	3.38	chr3	130905244	130905702	10883	NM_001001486	ATPase secretory pathway Ca2+ transporting 1	protein-coding
XIRP2	0.025399389761448	2.94	chr2	166988656	166988881	100281	NM_001199143	xin actin binding repeat containing 2	protein-coding
KCTD3	0.0255996215870951	3.36	chr1	215548131	215548323	-19152	NM_016121	potassium channel tetramerization domain containing 3	protein-coding
KIF26B	0.0258521682653248	3.12	chr1	245235067	245235239	80168	NM_018012	kinesin family member 26B	protein-coding
MAP7	0.0263020882962981	3.68	chr6	136582783	136583146	-32145	NM_001198616	microtubule associated protein 7	protein-coding
GNE	0.0263416581121416	3.17	chr9	36240379	36240606	18007	NM_001190383	glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase	protein-coding
MIR4454	0.0263588736479066	-3.27	chr1	229131589	229131787	43113	NR_039659	microRNA 4454	ncRNA
PLXNA2	0.0266116988025156	3.17	chr1	208310788	208310966	-66557	NM_025179	plexin A2	protein-coding
TNRC18	0.0269369272447968	2.95	chr7	5362440	5362629	61012	NM_001080495	trinucleotide repeat containing 18	protein-coding
PLA2G5	0.0269689639099663	-3.35	chr1	20089510	20089671	19382	NM_000929	phospholipase A2 group V	protein-coding
MIR4777	0.0270738415276069	3.03	chr2	231307286	231307465	-55333	NR_039937	microRNA 4777	ncRNA
GNG4	0.027269100344063	-3.09	chr1	235604355	235604545	45543	NM_001098722	G protein subunit gamma 4	protein-coding
STAC3	0.0277019289171556	2.73	chr12	57287321	57287597	-36266	NM_145064	SH3 and cysteine rich domain 3	protein-coding
TEK	0.0277134035674124	3.25	chr9	27127172	27127358	18124	NM_001290077	TEK receptor tyrosine kinase	protein-coding
EPS15L1	0.0278016259230087	3.42	chr19	16434216	16434376	37716	NR_047665	epidermal growth factor receptor pathway substrate 15 like 1	protein-coding
DDX1	0.0278676047416308	2.91	chr2	15618904	15619102	27382	NM_004939	DEAD-box helicase 1	protein-coding
ITGA5	0.0280079404053959	-1.99	chr12	54415391	54415771	3685	NM_002205	integrin subunit alpha 5	protein-coding
LINC00563	0.028233213274236	3.13	chr13	46318853	46319011	-21088	NR_047493	long intergenic non-protein coding RNA 563	ncRNA
GLDC	0.0282891553721113	2.94	chr9	6545296	6545461	100314	NM_000170	glycine decarboxylase	protein-coding
IZUMO3	0.0284745004682989	2.71	chr9	24852266	24852971	-306942	NM_001271706	IZUMO family member 3	protein-coding
MIR4324	0.0285394160340261	-3.39	chr19	49245242	49245518	63488	NR_036209	microRNA 4324	ncRNA
CDH11	0.0286279095491569	2.9	chr16	63984764	63985317	1137097	NM_001797	cadherin 11	protein-coding
MISP	0.0286383892230219	2.5	chr19	766237	766542	15276	NR_135168	mitotic spindle positioning	protein-coding
POTEA	0.0287825125698099	-3.02	chr8	43280919	43281167	-11399	NM_001005365	POTE ankyrin domain family member A	protein-coding
LOC101928052	0.0288635442366286	3.5	chr4	161952436	161952612	-70209	NR_125888	uncharacterized LOC101928052	ncRNA
ANKRD55	0.0289363316477154	-3.99	chr5	56177423	56177602	55847	NM_024669	ankyrin repeat domain 55	protein-coding
AKAIN1	0.0289713045471146	-3.37	chr18	4830815	4831077	366310	NM_001145194	A-kinase anchor inhibitor 1	protein-coding
FBXO7	0.0290636612526028	2.43	chr22	32571683	32571985	96597	NM_001257990	F-box protein 7	protein-coding
NSFP1	0.0290977515267246	-3.51	chr17	46598577	46598782	7945	NR_033799	N-ethylmaleimide-sensitive factor pseudogene 1	pseudo
SMYD3	0.0292095374683513	3.02	chr1	246116261	246116553	301005	NM_022743	SET and MYND domain containing 3	protein-coding
MAP2	0.0292736888643935	-3.16	chr2	209334368	209334737	-89495	NM_001039538	microtubule associated protein 2	protein-coding
LINC00601	0.029315855707129	-3.66	chr10	126436109	126436268	-14309	NR_073453	long intergenic non-protein coding RNA 601	ncRNA
ZSCAN30	0.0293870588126322	2.79	chr18	35289513	35289793	592	NM_001288711	zinc finger and SCAN domain containing 30	protein-coding
LRRC37A2	0.0294493300175409	-3.65	chr17	46501692	46501888	-10920	NM_001006607	leucine rich repeat containing 37 member A2	protein-coding
MCF2L2	0.0296289485115738	2.75	chr3	183378399	183378631	49754	NM_015078	MCF.2 cell line derived transforming sequence-like 2	protein-coding
LOC401134	0.0296936929889463	2.94	chr4	65027248	65027651	-22949	NR_033976	uncharacterized LOC401134	ncRNA
MAGEB16	0.0299230951096738	-2.91	chrX	35443014	35443301	-355185	NM_001099921	MAGE family member B16	protein-coding
BRCAT107	0.0303232211974105	2.77	chr5	44536611	44536771	-26409	NR_131946	breast cancer-associated transcript 107	ncRNA
FNIP2	0.0303839336013969	-3.66	chr4	158890472	158890655	84469	NM_001323916	folliculin interacting protein 2	protein-coding
TMTC1	0.0303961613279439	3.61	chr12	29906334	29906532	-121674	NM_175861	transmembrane and tetratricopeptide repeat containing 1	protein-coding
SALL2	0.0305719192244567	2.67	chr14	21525914	21526171	366	NM_001291447	spalt like transcription factor 2	protein-coding
FAM63B	0.0305837952611253	2.08	chr15	58795860	58796680	24908	NM_001040453	family with sequence similarity 63 member B	protein-coding
ATP2C2	0.0306473930722624	2.52	chr16	84367465	84367718	-932	NM_001286527	ATPase secretory pathway Ca2+ transporting 2	protein-coding
FAM46A	0.030723179097624	-3.58	chr6	81076246	81076445	676366	NM_017633	family with sequence similarity 46 member A	protein-coding
PTPRS	0.0308268125872134	-3.34	chr19	5265273	5265439	75447	NM_130855	protein tyrosine phosphatase, receptor type S	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
DNAJC25-GNG10	0.0309466708864858	2.75	chr9	111641236	111641401	9966	NM_004125	DNAJC25-GNG10 readthrough	protein-coding
RAB11FIP2	0.0309825201710405	2.97	chr10	117960580	117960746	85940	NM_014904	RAB11 family interacting protein 2	protein-coding
TMX3	0.0310820384417511	3.7	chr18	68388012	68388202	327009	NM_019022	thioredoxin related transmembrane protein 3	protein-coding
OR4E1	0.031111265498585	2.92	chr14	22076733	22076886	-405792	NM_001317107	olfactory receptor family 4 subfamily E member 1 (gene/pseudogene)	protein-coding
FGGY	0.0313011551881316	-3.15	chr1	59317179	59317562	7292	NM_001244714	FGGY carbohydrate kinase domain containing	protein-coding
DNAH7	0.0315937371798118	2.71	chr2	196051816	196052195	16807	NM_018897	dynein axonemal heavy chain 7	protein-coding
TRMT10C	0.0316302886167508	2.8	chr3	101551999	101552181	-9746	NM_017819	tRNA methyltransferase 10C, mitochondrial RNase P subunit	protein-coding
HDHD2	0.0317092517099528	3.14	chr18	47106997	47107455	43274	NM_032124	haloacid dehalogenase like hydrolase domain containing 2	protein-coding
CELF2	0.0321282019622511	3.57	chr10	10803546	10803704	4914	NM_001326327	CUGBP, Elav-like family member 2	protein-coding
GPR21	0.0321660486495814	-2.9	chr9	123046661	123047013	13194	NM_005294	G protein-coupled receptor 21	protein-coding
NCOA2	0.032367458174871	2.86	chr8	70255902	70256117	147246	NM_001321713	nuclear receptor coactivator 2	protein-coding
IFT43	0.032416890901636	2.77	chr14	76001734	76002010	16119	NR_045664	intraflagellar transport 43	protein-coding
LOC645166	0.0327153931336301	2.45	chr1	144893598	144893798	175527	NR_027354	lymphocyte-specific protein 1 pseudogene	pseudo
CAPZA3	0.0329332276354735	3.44	chr12	18905338	18905509	167312	NM_033328	capping actin protein of muscle Z-line alpha subunit 3	protein-coding
ACACA	0.0329952518173111	3.03	chr17	37309036	37309220	-9361	NM_198837	acetyl-CoA carboxylase alpha	protein-coding
HNRNPKP3	0.0331379609405046	2.75	chr11	43052203	43052812	216862	NR_033868	heterogeneous nuclear ribonucleoprotein K pseudogene 3	pseudo
TOX3	0.0332025849906617	-3.57	chr16	52496677	52496861	50125	NM_001080430	TOX high mobility group box family member 3	protein-coding
TRIO	0.0333836754044159	-3.22	chr5	14354654	14354840	211045	NM_007118	trio Rho guanine nucleotide exchange factor	protein-coding
SLC9A2	0.0334032607840978	-3.12	chr2	102640932	102641150	21507	NM_003048	solute carrier family 9 member A2	protein-coding
HNRNPA3	0.0334909177876978	2.78	chr2	176948998	176949332	-263529	NR_138470	heterogeneous nuclear ribonucleoprotein A3	protein-coding
LOC102723692	0.0335925801728471	-3.57	chr16	16982360	16982534	-152057	NR_135179	uncharacterized LOC102723692	ncRNA
RGS2	0.0338652228781427	3.04	chr1	192790176	192790392	-18755	NM_002923	regulator of G-protein signaling 2	protein-coding
CRBN	0.0343505204042145	2.84	chr3	3318344	3318512	-138711	NM_001173482	cereblon	protein-coding
FREM2	0.0344405440078429	3.22	chr13	38730340	38730542	43405	NM_207361	FRAS1 related extracellular matrix protein 2	protein-coding
MIR1237	0.0344697557585728	2.18	chr11	64379547	64379822	11082	NR_031602	microRNA 1237	ncRNA
TMEM45B	0.0345026613458705	-3.23	chr11	129786142	129786353	-29534	NM_001331210	transmembrane protein 45B	protein-coding
CHD1	0.0347573951713755	3.18	chr5	98891529	98891693	34923	NM_001270	chromodomain helicase DNA binding protein 1	protein-coding
C19orf25	0.0350361335930624	2.22	chr19	1471023	1471197	8119	NM_152482	chromosome 19 open reading frame 25	protein-coding
SLC7A11-AS1	0.0351889499706202	-2.92	chr4	137820079	137820261	-268844	NR_038380	SLC7A11 antisense RNA 1	ncRNA
MEAT6	0.0353375635417427	-3.38	chr6	164317797	164318047	504141	NR_131926	melanoma-associated transcript 6	ncRNA
EDN2	0.0353487499444445	3	chr1	41482776	41482939	1826	NR_126098	endothelin 2	protein-coding
B3GALT1	0.0354595916157943	-3.37	chr2	167787512	167787711	-31061	NM_020981	beta-1,3-galactosyltransferase 1	protein-coding
LINC01923	0.0355542654914578	3.71	chr2	198437290	198437536	-62316	NR_110267	long intergenic non-protein coding RNA 1923	ncRNA
TMEM261	0.0355775622860996	-3.61	chr9	7473351	7473530	326366	NM_001318058	transmembrane protein 261	protein-coding
OSCP1	0.0360173362087644	2.67	chr1	36431015	36431237	19359	NM_145047	organic solute carrier partner 1	protein-coding
GNS	0.0363216988064608	3.13	chr12	64706566	64706722	52802	NM_002076	glucosamine (N-acetyl)-6-sulfatase	protein-coding
SLC1A5	0.0364112346844723	2.87	chr19	46804707	46804878	-16207	NM_005628	solute carrier family 1 member 5	protein-coding
NCKAP5	0.0365384475427193	2.71	chr2	133700002	133700168	-131625	NM_207481	NCK associated protein 5	protein-coding
IQGAP2	0.0365975052165283	2.78	chr5	76332090	76332256	-71082	NM_006633	IQ motif containing GTPase activating protein 2	protein-coding
MIR548AG1	0.0369031874820239	3.19	chr4	60680465	60680621	-242076	NR_039652	microRNA 548ag-1	ncRNA
CCDC71L	0.0369124114870842	-2.49	chr7	106621297	106621644	39718	NM_175884	coiled-coil domain containing 71-like	protein-coding
BMPER	0.03699678626822	3.16	chr7	34239415	34239580	334586	NM_133468	BMP binding endothelial regulator	protein-coding
UGT2A3	0.037429214665149	2.51	chr4	68895688	68895952	55971	NR_024010	UDP glucuronosyltransferase family 2 member A3	protein-coding
KIAA0040	0.0374317363421064	-2.71	chr1	175167470	175167632	25262	NM_014656	KIAA0040	protein-coding
LINC00999	0.0379200476612775	-3.17	chr10	38429690	38429880	1639	NR_024497	long intergenic non-protein coding RNA 999	ncRNA
FST	0.0379340098360524	2.94	chr5	53324109	53324334	-156213	NM_013409	follistatin	protein-coding
LINC01007	0.038026099990676	2.35	chr7	101598630	101598810	-29714	NR_103747	long intergenic non-protein coding RNA 1007	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
CTNND2	0.0383475108744326	2.63	chr5	11283202	11283416	305608	NM_001288715	catenin delta 2	protein-coding
NPR3	0.0384086806852077	2.47	chr5	32794285	32794507	83064	NM_000908	natriuretic peptide receptor 3	protein-coding
CD2AP	0.0384551435238963	3.14	chr6	47481912	47482090	4212	NM_012120	CD2 associated protein	protein-coding
LINGO2	0.0384601536786631	2.99	chr9	29056042	29056554	156702	NM_001258282	leucine rich repeat and Ig domain containing 2	protein-coding
LINC01654	0.0385012069084928	2.71	chr1	17938995	17939283	-126518	NR_125946	long intergenic non-protein coding RNA 1654	ncRNA
MLLT10P1	0.0387865364866868	3.15	chr20	30652152	30652319	-248773	NR_045115	myeloid/lymphoid or mixed-lineage leukemia; translocated to, 10 pseudogene 1	pseudo
LINC01060	0.0388043106677037	-2.72	chr4	188914263	188915024	459065	NR_033869	long intergenic non-protein coding RNA 1060	ncRNA
LINC00836	0.0388556552388748	2.45	chr10	25650077	25650632	-1358	NR_108067	long intergenic non-protein coding RNA 836	ncRNA
LOC101926915	0.0389028870356047	-3.03	chr6	46273886	46274232	-2265	NR_125838	uncharacterized LOC101926915	ncRNA
SLC25A51	0.0390262268089809	2.82	chr9	37879645	37879869	24596	NM_033412	solute carrier family 25 member 51	protein-coding
NPR3	0.039186345420097	2.35	chr5	32687392	32687596	-23143	NM_001204376	natriuretic peptide receptor 3	protein-coding
RUSC2	0.039209576809738	2.47	chr9	35538833	35538999	284	NM_001135999	RUN and SH3 domain containing 2	protein-coding
KIF2A	0.0392169408952471	2.38	chr5	62162446	62163008	-143435	NM_001243953	kinesin family member 2A	protein-coding
MIR4445	0.0392386034877956	3.46	chr3	109554193	109554558	-48453	NR_039647	microRNA 4445	ncRNA
TM4SF1	0.039348113782636	3.43	chr3	149417482	149417676	-39798	NM_014220	transmembrane 4 L six family member 1	protein-coding
LINC01793	0.0394676049178432	2.69	chr2	59728106	59728280	510485	NR_110219	long intergenic non-protein coding RNA 1793	ncRNA
CA12	0.0395270967544349	-2.38	chr15	63331473	63331773	50487	NM_206925	carbonic anhydrase 12	protein-coding
LINC02052	0.0396588756568454	-2.58	chr3	186514592	186515319	-21294	NR_033844	long intergenic non-protein coding RNA 2052	ncRNA
LINC01126	0.0397044664796747	-2.94	chr2	43397377	43397544	170249	NR_027251	long intergenic non-protein coding RNA 1126	ncRNA
MN1	0.0397657070419207	2.38	chr22	27697665	27697944	103694	NM_002430	MN1 proto-oncogene, transcriptional regulator	protein-coding
FUT6	0.0398520462505295	2.06	chr19	5838102	5838572	1394	NM_001040701	fucosyltransferase 6	protein-coding
TFEC	0.0406765883208978	2.26	chr7	115997085	115997352	-28905	NM_001244583	transcription factor EC	protein-coding
KCNAB1	0.0406805506406443	3.58	chr3	156286332	156286506	-4568	NM_172159	potassium voltage-gated channel subfamily A member regulatory beta subunit 1	protein-coding
LOC389705	0.0408845267409729	2.15	chr9	14993269	14993613	114	NR_003920	chromosome 4 open reading frame 27 pseudogene	pseudo
PFKP	0.0409885425346045	2.81	chr10	3046072	3046226	-21371	NM_001323068	phosphofructokinase, platelet	protein-coding
CDH26	0.041094902393441	2.86	chr20	60010975	60011133	14692	NM_021810	cadherin 26	protein-coding
MIR548AG2	0.0413223177190319	3.01	chr20	60568322	60568484	3841	NR_039653	microRNA 548ag-2	ncRNA
MIR3973	0.0415662558981225	-3.51	chr11	35999792	35999957	-10224	NR_039769	microRNA 3973	ncRNA
OSTF1	0.0418298750002473	2.71	chr9	75273337	75273763	185068	NM_012383	osteoclast stimulating factor 1	protein-coding
KIF4B	0.0418514381367556	3.04	chr5	155547553	155547711	533877	NM_001099293	kinesin family member 4B	protein-coding
CMSS1	0.0421906658995165	2.45	chr3	99795073	99795265	-22665	NM_032359	cms1 ribosomal small subunit homolog (yeast)	protein-coding
CCNYL2	0.0422142418170764	2.66	chr10	42453974	42454314	18096	NR_103829	cyclin Y-like 2 (pseudogene)	pseudo
EDIL3	0.042348380230676	3.01	chr5	84813347	84813504	-428558	NM_005711	EGF like repeats and discoidin domains 3	protein-coding
RAB4A	0.0423900834248746	-2.36	chr1	229223590	229223827	-47354	NR_073545	RAB4A, member RAS oncogene family	protein-coding
GRIN2D	0.0424655884983748	2.6	chr19	48397979	48398131	3180	NM_000836	glutamate ionotropic receptor NMDA type subunit 2D	protein-coding
LINC01601	0.0427500814200583	-3.19	chr18	44595270	44595497	63604	NR_131768	long intergenic non-protein coding RNA 1601	ncRNA
LOC100506858	0.0427588839852574	3.08	chr5	2274325	2274503	37787	NR_104616	uncharacterized LOC100506858	ncRNA
MIR924	0.0429255337875858	2.79	chr18	38680943	38681179	941114	NR_030628	microRNA 924	ncRNA
ATL2	0.0429581821320631	3.08	chr2	38422098	38422265	-43597	NM_001330462	atlastin GTPase 2	protein-coding
CD226	0.0429675493497117	3.02	chr18	69728054	69728215	219719	NM_001303618	CD226 molecule	protein-coding
RORA-AS2	0.0430010585287045	-3.07	chr15	60722508	60722720	41045	NR_120318	RORA antisense RNA 2	ncRNA
DNAH5	0.0430075754612667	2.41	chr5	13986342	13986502	-41942	NM_001369	dynein axonemal heavy chain 5	protein-coding
ZBTB20-AS3	0.0430342986271268	2.56	chr3	114874321	114874763	1509	NR_126422	ZBTB20 antisense RNA 3	ncRNA
C4orf26	0.043690236363545	3.17	chr4	75536500	75536651	-19473	NR_046429	chromosome 4 open reading frame 26	protein-coding
LINC01029	0.0440404106725579	2.7	chr18	77919331	77919504	74310	NR_104127	long intergenic non-protein coding RNA 1029	ncRNA
LINC00376	0.0441834116856357	-3.05	chr13	63446420	63446574	-118403	NR_126409	long intergenic non-protein coding RNA 376	ncRNA
NLGN4Y	0.0443431196759791	3.46	chrY	14328683	14328851	-193841	NM_001206850	neuroligin 4, Y-linked	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
RNLS	0.0443702190701183	2.43	chr10	88254216	88254510	328962	NM_001031709	renalase, FAD dependent amine oxidase	protein-coding
BEND5	0.0445320098125782	3.04	chr1	48826350	48826518	-49465	NM_001302082	BEN domain containing 5	protein-coding
NIT2	0.0445399743124099	-2.44	chr3	100330000	100330322	-4557	NM_020202	nitrilase family member 2	protein-coding
LYPD1	0.0445492513393136	-3.47	chr2	132589931	132590109	80350	NM_001321234	LY6/PLAUR domain containing 1	protein-coding
CCNY	0.0447379140242039	2.6	chr10	35409221	35409376	72424	NM_001282853	cyclin Y	protein-coding
LINC01182	0.0447621495884011	2.35	chr4	13642135	13642433	-12895	NR_121681	long intergenic non-protein coding RNA 1182	ncRNA
MIPOL1	0.0448157544138095	3.64	chr14	37277959	37278136	80134	NM_001195296	mirror-image polydactyly 1	protein-coding
FAM170A	0.0448595982072632	3.4	chr5	119662310	119662506	32849	NM_182761	family with sequence similarity 170 member A	protein-coding
LOC101927078	0.0449176244449094	3	chr5	114704624	114704785	68709	NR_130785	uncharacterized LOC101927078	ncRNA
TRPM2	0.0449720055925	2.08	chr21	44400934	44401308	-24408	NM_001320352	transient receptor potential cation channel subfamily M member 2	protein-coding
PP2D1	0.0451568414964198	-2.97	chr3	20013365	20013572	-1195	NM_001252657	protein phosphatase 2C like domain containing 1	protein-coding
TMCC1	0.0451626478443241	3.42	chr3	129802643	129802857	77809	NM_001017395	transmembrane and coiled-coil domain family 1	protein-coding
RERE	0.0452476662109823	-3.39	chr1	8417097	8417296	6491	NM_001042682	arginine-glutamic acid dipeptide repeats	protein-coding
KIAA1217	0.0453546462086651	3.2	chr10	24151467	24151647	-57234	NM_001282768	KIAA1217	protein-coding
FAM201A	0.0457882663401929	3.04	chr9	38769184	38769438	148223	NR_027294	family with sequence similarity 201 member A	ncRNA
LINC01299	0.0459193779799917	2.76	chr8	65536530	65536681	26061	NR_033893	long intergenic non-protein coding RNA 1299	ncRNA
MIR5582	0.0459359044399863	2.41	chr11	46744213	46744372	8900	NR_049846	microRNA 5582	ncRNA
LOC105379514	0.0459506042609631	2.6	chr21	9647364	9647524	-14745	NR_135513	uncharacterized LOC105379514	ncRNA
sept-09	0.0460176031479512	-2.26	chr17	77473978	77474196	-1156	NM_001113495	septin 9	protein-coding
RNF180	0.0461960827566167	3.37	chr5	64074847	64075062	-90889	NM_178532	ring finger protein 180	protein-coding
TUBB6	0.0461968805459582	2.64	chr18	12336554	12337062	28749	NM_001303528	tubulin beta 6 class V	protein-coding
JMJD1C	0.0462937182020723	-2.77	chr10	63474562	63474963	-8761	NM_032776	jumonji domain containing 1C	protein-coding
LOC101927450	0.0464104883198378	2.74	chr9	78753800	78753976	391778	NR_109771	uncharacterized LOC101927450	ncRNA
LINC01243	0.0464395518335581	2.9	chr9	31726723	31727474	-345608	NR_135134	long intergenic non-protein coding RNA 1243	ncRNA
NEGR1	0.0464584470927469	3.23	chr1	72540592	72540751	-258077	NM_173808	neuronal growth regulator 1	protein-coding
LOC103352541	0.0466638328092629	2.8	chr6	140064566	140064788	-73416	NR_121624	uncharacterized LOC103352541	ncRNA
STPG2-AS1	0.0467088966598815	3.08	chr4	97064430	97064598	-302412	NR_102713	STPG2 antisense RNA 1	ncRNA
UBE2D1	0.0468736038618061	-3.11	chr10	58322737	58322904	-12159	NM_001204880	ubiquitin conjugating enzyme E2 D1	protein-coding
PIN4P1	0.0469005329269471	2.87	chr15	44007710	44007885	-129626	NR_003571	peptidylprolyl cis/trans isomerase, NIMA-interacting 4 pseudogene 1	pseudo
TUSC5	0.0469447082073642	2.73	chr17	1278949	1279099	-639	NM_172367	tumor suppressor candidate 5	protein-coding
PTPRD-AS2	0.0470566120979931	2.51	chr9	11187682	11187856	574563	NR_110696	PTPRD antisense RNA 2 (head to head)	ncRNA
ZNF521	0.0470865768882327	-3.27	chr18	25233804	25233968	118364	NM_001308225	zinc finger protein 521	protein-coding
MS4A10	0.0472328742479918	-2.07	chr11	60775468	60775783	-9723	NM_206893	membrane spanning 4-domains A10	protein-coding
RANBP3L	0.0475981192001352	3.22	chr5	36392919	36393086	-91093	NM_001161429	RAN binding protein 3 like	protein-coding
LINC02027	0.0476675242460189	2.69	chr3	80574995	80575207	-418767	NR_132411	long intergenic non-protein coding RNA 2027	ncRNA
ADCYAP1	0.047778481331758	-2.62	chr18	881967	882228	-22846	NM_001099733	adenylate cyclase activating polypeptide 1	protein-coding
C5orf22	0.0478394769956341	-2.5	chr5	31557293	31557866	25313	NM_018356	chromosome 5 open reading frame 22	protein-coding
LINC01602	0.047926686102663	-2.94	chr8	57888333	57888556	-89914	NR_130934	long intergenic non-protein coding RNA 1602	ncRNA
MTRNR2L6	0.0479569434943269	2.73	chr7	142619358	142619531	-46855	NM_001190487	MT-RNR2-like 6	protein-coding
SLC4A7	0.0481985890955637	2.72	chr3	27563354	27563516	-79015	NM_001321103	solute carrier family 4 member 7	protein-coding
LOC101929563	0.0482293626152925	3.1	chr9	23255044	23255306	417224	NR_121602	uncharacterized LOC101929563	ncRNA
NGF	0.0484168855444761	2.44	chr1	115214253	115214445	123887	NM_002506	nerve growth factor	protein-coding
NBEA	0.0484666244393094	2.32	chr13	35526630	35526790	49961	NM_001204197	neurobeachin	protein-coding
GLI3	0.0486177000388111	3.48	chr7	42401660	42401829	-164725	NM_000168	GLI family zinc finger 3	protein-coding
CEP192	0.0487443718395803	2.16	chr18	13003626	13003932	12417	NM_032142	centrosomal protein 192	protein-coding
KDM6A	0.0487612640583013	3.17	chrX	44883242	44883434	10163	NM_001291418	lysine demethylase 6A	protein-coding
STPG2-AS1	0.0489235315556565	3.27	chr4	97086581	97086746	-280263	NR_102713	STPG2 antisense RNA 1	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
NWD1	0.0489622570481963	-3.04	chr19	16719869	16720064	-10	NM_001290355	NACHT and WD repeat domain containing 1	protein-coding
LINC01242	0.0490714052806714	2.71	chr9	30506154	30506321	-97783	NR_046204	long intergenic non-protein coding RNA 1242	ncRNA
SULT1C2P1	0.0495229900254999	2.81	chr2	108340184	108340406	18057	NR_037191	sulfotransferase family 1C member 2 pseudogene 1	pseudo
EPAS1	0.0496219619302549	-2.2	chr2	46259636	46260014	-37577	NM_001430	endothelial PAS domain protein 1	protein-coding

Supplementary Table 4
List of genes with altered chromatin accessibility 24 hrs post-infection

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
MIR3922	0.00167846109606298	-4.85	chr12	104754068	104754263	162532	NR_037487	microRNA 3922	ncRNA
PPP1R2P3	0.0018460989318751	4.97	chr5	156716235	156716399	-134221	NR_038443	protein phosphatase 1 regulatory inhibitor subunit 2 pseudogene 3	pseudo
CINP	0.00457042651016713	4.18	chr14	102380658	102380800	-17813	NM_032630	cyclin dependent kinase 2 interacting protein	protein-coding
HACE1	0.0046205531201099	4.81	chr6	103676921	103677245	1182463	NM_001321083	HECT domain and ankyrin repeat containing E3 ubiquitin protein ligase 1	protein-coding
DSG1-AS1	0.00475709102490307	4.58	chr18	31403236	31403427	23657	NR_110788	DSG1 antisense RNA 1	ncRNA
TMEM167A	0.00604098215635939	4.34	chr5	83197761	83197966	-120410	NM_174909	transmembrane protein 167A	protein-coding
LINC00603	0.00644505953605041	4.78	chr5	40132305	40132467	80095	NR_104633	long intergenic non-protein coding RNA 603	ncRNA
VRK2	0.0066774423020012	4.24	chr2	57883083	57883230	-24495	NM_001288837	vaccinia related kinase 2	protein-coding
TBC1D23	0.00731806720061929	3.96	chr3	100276876	100277229	16235	NM_018309	TBC1 domain family member 23	protein-coding
DRD3	0.00794669164420907	4.07	chr3	114178756	114178933	-68	NM_001290809	dopamine receptor D3	protein-coding
DMRT2	0.00796543334352703	-4.86	chr9	1458621	1458788	408084	NM_181872	doublesex and mab-3 related transcription factor 2	protein-coding
SSH2	0.00809195414620171	-3.67	chr17	29844398	29844628	-83081	NM_001282131	slingshot protein phosphatase 2	protein-coding
APOLD1	0.00827310486611747	4.18	chr12	12779552	12779867	-5898	NM_030817	apolipoprotein L domain containing 1	protein-coding
CTD-2194D22.4	0.00873940133323193	4.4	chr5	2001470	2001618	114213	NR_109912	uncharacterized LOC101929081	ncRNA
TBC1D5	0.00913860739246859	4.54	chr3	17413075	17413395	286785	NM_001134381	TBC1 domain family member 5	protein-coding
PTPN4	0.00916598067919051	-2.96	chr2	119731258	119731686	-28159	NM_002830	protein tyrosine phosphatase, non-receptor type 4	protein-coding
HLA-DQB1	0.00965465550119866	-3.22	chr6	32668256	32668600	-1739	NM_002123	major histocompatibility complex, class II, DQ beta 1	protein-coding
LINC01127	0.0100821511712936	3.66	chr2	101967841	101968060	-15515	NR_103791	long intergenic non-protein coding RNA 1127	ncRNA
SNORA105A	0.010092655668002	-3.22	chr5	21806895	21807355	-77329	NR_132787	small nucleolar RNA, H/ACA box 105A	snoRNA
LOC105378098	0.0106702101561261	4.32	chr6	161636023	161636181	165125	NR_134594	uncharacterized LOC105378098	ncRNA
TENM2	0.0107097476510678	3.88	chr5	167417640	167417791	132877	NM_001122679	teneurin transmembrane protein 2	protein-coding
EYA4	0.0110737944282832	-3.62	chr6	133472494	133472701	231098	NM_001301012	EYA transcriptional coactivator and phosphatase 4	protein-coding
ITGA8	0.0113204961260818	3.88	chr10	15581179	15581355	139068	NM_001291494	integrin subunit alpha 8	protein-coding
OAZ1	0.011494877649696	-2.9	chr19	2266630	2266799	-2772	NM_004152	ornithine decarboxylase antizyme 1	protein-coding
ADAMTSL1	0.0117098481940856	3.91	chr9	18190269	18190475	-283709	NM_001040272	ADAMTS like 1	protein-coding
FAM133CP	0.0117275897401877	4.35	chr10	58757663	58757982	42807	NR_027508	family with sequence similarity 133, member A pseudogene	pseudo
MIR548AD	0.0119359960678406	4.27	chr2	35370944	35371158	-100354	NR_039629	microRNA 548ad	ncRNA
DNAH5	0.012983269425378	-3.36	chr5	14025942	14026184	-81583	NM_001369	dynein axonemal heavy chain 5	protein-coding
CDH18	0.013236209074784	4.35	chr5	19973760	19973918	14405	NM_004934	cadherin 18	protein-coding
DEPDC1	0.0135066491814665	-4.23	chr1	68818749	68818899	-321603	NM_017779	DEP domain containing 1	protein-coding
ZADH2	0.0136936632028093	-3.77	chr18	74963958	74964242	241413	NM_001306093	zinc binding alcohol dehydrogenase domain containing 2	protein-coding
PTPRS	0.013888260583939	-3.95	chr19	5320714	5320902	19995	NM_130853	protein tyrosine phosphatase, receptor type S	protein-coding
LOC101927168	0.0140388474813231	2.81	chr18	6653903	6654144	75839	NR_134645	uncharacterized LOC101927168	ncRNA
OPCML	0.0140895544315469	3.91	chr11	133263318	133263662	269122	NM_001319104	opioid binding protein/cell adhesion molecule like	protein-coding
PPM1A	0.0144116504577761	4	chr14	60222915	60223198	-22696	NM_177952	protein phosphatase, Mg2+/Mn2+ dependent 1A	protein-coding
TARS	0.0144668457801469	-3.88	chr5	33195938	33196279	-244588	NM_001258437	threonyl-tRNA synthetase	protein-coding
STMND1	0.0147583160769847	-2.56	chr6	17115868	17116323	13837	NM_001190766	stathmin domain containing 1	protein-coding
DST	0.0150413298446148	-2.95	chr6	56637534	56638023	5118	NM_015548	dystonin	protein-coding
LOC105372028	0.0151419736498011	-2.84	chr18	24991105	24991567	265555	NR_134604	uncharacterized LOC105372028	ncRNA
CSMD3	0.0152612952916675	3.87	chr8	113105441	113105622	271622	NM_198124	CUB and Sushi multiple domains 3	protein-coding
VASH1	0.0154753685021713	-3.81	chr14	76599707	76599926	-162076	NM_014909	vasohibin 1	protein-coding
UBE2QL1	0.01558324127808	-3	chr5	6476277	6476446	27738	NM_001145161	ubiquitin conjugating enzyme E2 Q family like 1	protein-coding
SLC6A3	0.015706249211786	-3.02	chr5	1437433	1437782	7821	NM_001044	solute carrier family 6 member 3	protein-coding
E2F7	0.0158149640139037	4.08	chr12	77246542	77246720	-181051	NM_203394	E2F transcription factor 7	protein-coding
GPR151	0.0158777022425558	3.43	chr5	146565886	146566164	-49912	NM_194251	G protein-coupled receptor 151	protein-coding
SPRED2	0.015939308891163	-4	chr2	65631637	65631802	-199197	NM_181784	sprouty related EVH1 domain containing 2	protein-coding
PLXNA2	0.0159690435493456	-2.95	chr1	208652422	208652731	-408256	NM_025179	plexin A2	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
JARID2	0.0163646989586876	-3.57	chr6	15254226	15254454	5485	NM_001267040	jumonji and AT-rich interaction domain containing 2	protein-coding
MIR548AD	0.016383979763502	3.28	chr2	35472148	35472325	831	NR_039629	microRNA 548ad	ncRNA
LINC01239	0.0165985780824211	-3.67	chr9	23117833	23118024	471728	NR_038977	long intergenic non-protein coding RNA 1239	ncRNA
KIF21A	0.0172064527400032	-3.83	chr12	39474112	39474269	-30800	NM_001173463	kinesin family member 21A	protein-coding
LRFN5	0.0174466618915173	-3.5	chr14	42455683	42455919	848945	NM_001346173	leucine rich repeat and fibronectin type III domain containing 5	protein-coding
LOC101927697	0.0175292128206602	-2.84	chr5	158370085	158370239	39611	NR_109888	uncharacterized LOC101927697	ncRNA
MGC32805	0.0176480080852434	3.91	chr5	122500569	122500746	-21570	NR_051996	uncharacterized LOC153163	ncRNA
C7orf72	0.0176480080852434	3.91	chr7	50179797	50179966	83795	NM_001161834	chromosome 7 open reading frame 72	protein-coding
EDN1	0.0176599660411047	-2.79	chr6	12348583	12349085	58538	NM_001955	endothelin 1	protein-coding
LSP1P3	0.0177933484828549	-3.83	chr5	28787418	28787645	-139339	NR_033961	lymphocyte-specific protein 1 pseudogene 3	pseudo
PPP2R2A	0.0178503782355238	-3.27	chr8	26185535	26185679	-105884	NM_002717	protein phosphatase 2 regulatory subunit Balpha	protein-coding
APOL5	0.0182900883834954	4.08	chr22	35684526	35684961	-33129	NM_030642	apolipoprotein L5	protein-coding
BRE-AS1	0.018454182477246	-2.81	chr2	27896222	27896831	-5412	NR_028308	BRE antisense RNA 1	ncRNA
AHCYL2	0.0185346037959754	4.2	chr7	129375061	129375221	-502	NM_001130722	adenosylhomocysteinase like 2	protein-coding
ACSM6	0.0185479439116412	-3.02	chr10	95241748	95241981	47664	NM_207321	acyl-CoA synthetase medium-chain family member 6	protein-coding
SCN7A	0.0185884992840664	-4.04	chr2	166628847	166629013	-141959	NM_002976	sodium voltage-gated channel alpha subunit 7	protein-coding
LINC01242	0.018857717102377	4.06	chr9	30328593	30328754	79781	NR_046204	long intergenic non-protein coding RNA 1242	ncRNA
PIN1P1	0.0193095384731256	3.99	chr1	69991319	69991474	72074	NR_023916	peptidylprolyl cis/trans isomerase, NIMA-interacting 1 pseudogene 1	pseudo
LINC01714	0.0193432462130756	-2.09	chr1	8234124	8234498	25639	NR_125998	long intergenic non-protein coding RNA 1714	ncRNA
TRIO	0.0194711057364159	-2.84	chr5	14265610	14265827	122016	NM_007118	trio Rho guanine nucleotide exchange factor	protein-coding
WDR89	0.0194991613749515	2.97	chr14	63663447	63663614	-21607	NM_001258272	WD repeat domain 89	protein-coding
WFDC3	0.0195162174212616	-3.03	chr20	45779067	45779480	12635	NM_080614	WAP four-disulfide core domain 3	protein-coding
TEK	0.0201287509910024	-2.93	chr9	27125048	27125394	16080	NM_001290077	TEK receptor tyrosine kinase	protein-coding
ROCK1	0.0201878448746534	-3.84	chr18	19091626	19091815	2020131	NM_005406	Rho associated coiled-coil containing protein kinase 1	protein-coding
ARPC5L	0.0206208116118049	3.36	chr9	124894597	124894847	32604	NM_030978	actin related protein 2/3 complex subunit 5 like	protein-coding
TTC39C	0.0206367311585076	-3.99	chr18	24076223	24076613	-36935	NM_001292030	tetratricopeptide repeat domain 39C	protein-coding
LINC00906	0.0207583138066376	3.66	chr19	28947035	28947154	-18037	NR_027318	long intergenic non-protein coding RNA 906	ncRNA
MIR4477A	0.0207686349741845	-3.32	chr20	30106943	30107120	-226568	NR_039688	microRNA 4477a	ncRNA
CD200R1	0.0210296493190514	-3.73	chr3	112931749	112931918	43257	NM_138939	CD200 receptor 1	protein-coding
LINC01568	0.0210453600996103	4	chr16	73753197	73753493	366540	NR_038234	long intergenic non-protein coding RNA 1568	ncRNA
PHEX-AS1	0.0212899200808918	-3.21	chrX	22134905	22135303	37879	NR_046639	PHEX antisense RNA 1	ncRNA
MYOM1	0.0213999959755124	-2.48	chr18	3207611	3208059	12273	NM_003803	myomesin 1	protein-coding
OR5T2	0.0219829217193681	3.24	chr11	56205257	56205593	27760	NM_001004746	olfactory receptor family 5 subfamily T member 2	protein-coding
LINGO4	0.0220030432849005	-3.13	chr1	151809544	151809911	-4285	NM_001004432	leucine rich repeat and Ig domain containing 4	protein-coding
sept-10	0.0222493892505717	3.98	chr2	109540821	109540980	72757	NM_001321513	septin 10	protein-coding
CHL1	0.0222493892505717	3.98	chr3	179838	179988	-16683	NM_006614	cell adhesion molecule L1 like	protein-coding
MSR1	0.0223008156878476	3.85	chr8	16285856	16286066	-93170	NM_138715	macrophage scavenger receptor 1	protein-coding
FAM46A	0.0223026078708696	3.96	chr6	81257331	81257480	495306	NM_017633	family with sequence similarity 46 member A	protein-coding
MLANA	0.0223124239732133	-2.86	chr9	5879720	5880070	-11014	NM_005511	melan-A	protein-coding
EGFR-AS1	0.0224740973260902	3.69	chr7	55198300	55198500	-9451	NR_047551	EGFR antisense RNA 1	ncRNA
ARHGAP29	0.0225286124133973	2.76	chr1	94310393	94310649	-62761	NM_001328665	Rho GTPase activating protein 29	protein-coding
MT1M	0.022590623978912	3.48	chr16	56632903	56633091	375	NM_176870	metallothionein 1M	protein-coding
MEAT6	0.0227866998126211	4.12	chr6	164437035	164437197	384947	NR_131926	melanoma-associated transcript 6	ncRNA
SNX5	0.0229269204263302	3.08	chr20	17996065	17996186	-27134	NM_001282454	sorting nexin 5	protein-coding
RORA	0.0230592700261484	-3.33	chr15	60551887	60552073	40528	NM_134262	RAR related orphan receptor A	protein-coding
SMAD7	0.0233409156118489	3.43	chr18	48958039	48958240	-7428	NM_005904	SMAD family member 7	protein-coding
LINC01848	0.0234349038948551	3.03	chr5	109877325	109877874	-5583	NR_033175	long intergenic non-protein coding RNA 1848	pseudo
TINAG	0.0236283337752846	2.42	chr6	54307462	54307903	-723	NM_014464	tubulointerstitial nephritis antigen	protein-coding
BBOX1	0.0236629529424134	2.95	chr11	27081042	27081296	40207	NM_003986	gamma-butyrobetaine hydroxylase 1	protein-coding
PRDM15	0.0236744385131973	3.91	chr21	41839935	41840068	23301	NR_104260	PR/SET domain 15	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
NDUFV2	0.0240320345142988	-2.5	chr18	8982465	8982693	-120051	NM_021074	NADH:ubiquinone oxidoreductase core subunit V2	protein-coding
MS4A13	0.0240737333391207	3.11	chr11	60512486	60512712	-2814	NM_001100909	membrane spanning 4-domains A13	protein-coding
VWC2L	0.024233103384272	3.92	chr2	214225016	214225181	-186470	NM_001080500	von Willebrand factor C domain containing protein 2-like	protein-coding
FAM134B	0.0243315739052413	3.05	chr5	16757784	16757956	-140812	NM_001034850	family with sequence similarity 134 member B	protein-coding
CLUL1	0.0245866379503516	3.56	chr18	607681	607831	-8916	NM_001318522	clusterin like 1	protein-coding
LINC01395	0.0246319209468919	3.99	chr11	129681472	129681635	-64284	NR_120582	long intergenic non-protein coding RNA 1395	ncRNA
KCNK18	0.0247821238063703	3.44	chr10	117201322	117201493	3918	NM_181840	potassium two pore domain channel subfamily K member 18	protein-coding
DLEC1	0.0251412307854105	-3.25	chr3	38066904	38067084	27789	NM_007337	deleted in lung and esophageal cancer 1	protein-coding
BTBD2	0.0258694797746179	-2.85	chr19	1986292	1986574	29270	NM_017797	BTB domain containing 2	protein-coding
MAPK13	0.0262641112221604	3.97	chr6	36108988	36109191	-21395	NM_002754	mitogen-activated protein kinase 13	protein-coding
KCMF1	0.0263105243804947	3.52	chr2	84939609	84939802	-31403	NM_020122	potassium channel modulatory factor 1	protein-coding
SCHLAP1	0.0263732293237294	-2.77	chr2	180706845	180707415	15026	NR_104320	SWI/SNF complex antagonist associated with prostate cancer 1 (non-protein coding)	ncRNA
LOC105369486	0.0264782789088122	2.99	chr11	110339699	110339866	-15332	NR_135100	uncharacterized LOC105369486	ncRNA
DICER1	0.0265866463482418	3.87	chr14	95174428	95174662	-16535	NM_030621	dicer 1, ribonuclease III	protein-coding
DMRT1	0.0266192954665201	-3.11	chr9	846927	847118	5332	NM_021951	doublesex and mab-3 related transcription factor 1	protein-coding
RDM1	0.0267186684584315	-2.9	chr17	35930600	35930887	33	NR_027998	RAD52 motif containing 1	protein-coding
IL1RL1	0.0268958649366624	3.79	chr2	102302089	102302297	-9309	NM_001282408	interleukin 1 receptor like 1	protein-coding
TRIP13	0.0281800873393137	-3.01	chr5	915340	915522	22577	NM_004237	thyroid hormone receptor interactor 13	protein-coding
TRIO	0.0281833192431259	-2.3	chr5	14268273	14268612	124740	NM_007118	trio Rho guanine nucleotide exchange factor	protein-coding
LINC00367	0.028373995722897	-3.36	chr13	20993972	20994241	-44678	NR_104058	long intergenic non-protein coding RNA 367	ncRNA
MIR4473	0.0283755023308038	2.89	chr9	20125825	20125962	285345	NR_039684	microRNA 4473	ncRNA
RPS29	0.0284708108829632	3.81	chr14	48936823	48937027	649491	NM_001032	ribosomal protein S29	protein-coding
C19orf47	0.0285172218095024	4.19	chr19	40329074	40329255	19245	NR_046202	chromosome 19 open reading frame 47	protein-coding
SLC25A51	0.0285648744282328	-2.19	chr9	37964380	37964673	-60173	NM_033412	solute carrier family 25 member 51	protein-coding
LOC102724691	0.0286429534204952	3.46	chr2	105052809	105053095	-14456	NR_110603	uncharacterized LOC102724691	ncRNA
OR2F1	0.028747490145379	-2.7	chr7	143960083	143960256	242	NM_012369	olfactory receptor family 2 subfamily F member 1 (gene/pseudogene)	protein-coding
GBX1	0.0289925781443191	3.83	chr7	151148585	151148742	18885	NM_001098834	gastrulation brain homeobox 1	protein-coding
LINC00261	0.0291249699857289	3.56	chr20	22570526	22570701	8029	NR_001558	long intergenic non-protein coding RNA 261	ncRNA
CPNE8	0.0293286382176258	3.46	chr12	38883734	38883924	21789	NM_153634	copine 8	protein-coding
ABCG2	0.0296001586011524	3.88	chr4	88247445	88247611	-16206	NM_001257386	ATP binding cassette subfamily G member 2 (Junior blood group)	protein-coding
LINC01519	0.0297477206969234	3.32	chr10	84825364	84825758	373344	NR_120668	long intergenic non-protein coding RNA 1519	ncRNA
TAF1L	0.0297761908946236	-2.4	chr9	32644243	32645017	-8961	NM_153809	TATA-box binding protein associated factor 1 like	protein-coding
MAMDC2	0.030067987324907	3.8	chr9	70085461	70085627	41963	NM_001347990	MAM domain containing 2	protein-coding
LINC01123	0.03019062764657	3.82	chr2	110361837	110361956	22629	NR_046112	long intergenic non-protein coding RNA 1123	ncRNA
LINC01107	0.0303337712163921	-2.7	chr2	238650923	238651072	-95498	NR_037809	long intergenic non-protein coding RNA 1107	ncRNA
LOC101928622	0.0303798151170532	-3.8	chr4	33494333	33494535	545459	NR_125902	uncharacterized LOC101928622	ncRNA
STAMBPL1	0.0304401154973422	3.8	chr10	88897514	88897674	17407	NM_020799	STAM binding protein like 1	protein-coding
CTSO	0.0304881500625175	3.33	chr4	156378070	156378287	-424282	NM_001334	cathepsin O	protein-coding
UBAP2	0.0312385589726512	-2.23	chr9	33931733	33932349	2323	NM_001282530	ubiquitin associated protein 2	protein-coding
SLC8A2	0.031266513709182	-3.26	chr19	47456689	47456848	15282	NM_015063	solute carrier family 8 member A2	protein-coding
GRM5-AS1	0.0312979245824906	3.95	chr11	88557932	88558050	53415	NR_049724	GRM5 antisense RNA 1	ncRNA
LEPROT	0.0314064291276195	-3.83	chr1	65372549	65372745	-47801	NM_001198683	leptin receptor overlapping transcript	protein-coding
TMEM119	0.0314726013684511	2.9	chr12	108593909	108594041	4143	NM_181724	transmembrane protein 119	protein-coding
MVB12B	0.0315322019370292	3.21	chr9	126342519	126342644	15737	NM_001011703	multivesicular body subunit 12B	protein-coding
SLC29A4	0.031635046442245	-2.58	chr7	5260970	5261583	-21654	NM_001040661	solute carrier family 29 member 4	protein-coding
LOC101927588	0.0316418485776375	3.31	chr8	124299337	124299495	-52018	NR_125420	uncharacterized LOC101927588	ncRNA
C7orf71	0.0316819960059669	3.12	chr7	26686575	26686745	47575	NM_001302625	chromosome 7 open reading frame 71	protein-coding
LINC01898	0.0317724187538734	3.93	chr18	75875721	75875888	-163401	NR_134646	long intergenic non-protein coding RNA 1898	ncRNA
SORT1	0.0317868072480067	-3.53	chr1	109371359	109371617	21869	NM_001205228	sortilin 1	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
LINC01121	0.0319576523788556	-3.15	chr2	45285228	45285408	-30377	NR_033831	long intergenic non-protein coding RNA 1121	ncRNA
LOC101928535	0.0319577312497688	3.08	chr11	106306946	106307116	55624	NR_120551	uncharacterized LOC101928535	ncRNA
MIR4643	0.0319868729388467	-3.85	chr6	91168089	91168394	-353419	NR_039786	microRNA 4643	ncRNA
RPL13AP17	0.0319927594461418	4.03	chr7	78615191	78615447	268077	NR_003680	ribosomal protein L13a pseudogene 17	pseudo
LINC01611	0.032267387094948	3.66	chr6	84581910	84582082	-107067	NR_132100	long intergenic non-protein coding RNA 1611	ncRNA
B3GALNT1	0.0323833073099633	2.4	chr3	161031190	161031335	73633	NM_033169	beta-1,3-N-acetylgalactosaminyltransferase 1 (globoside blood group)	protein-coding
POU4F2	0.0324181549484366	3.74	chr4	146756068	146756232	117257	NM_004575	POU class 4 homeobox 2	protein-coding
CELF5	0.0324520154137756	-2.62	chr19	3290455	3290690	65087	NR_033342	CUGBP, Elav-like family member 5	protein-coding
COL8A1	0.0325387301188105	3.63	chr3	99287886	99288052	-350627	NM_001850	collagen type VIII alpha 1 chain	protein-coding
GTF2IP1	0.0325909663355069	-3.36	chr7	73185990	73186157	31144	NR_002206	general transcription factor IIi pseudogene 1	pseudo
SHROOM4	0.0325950969767832	3.74	chrX	50708069	50708281	105869	NR_027121	shroom family member 4	protein-coding
MSC-AS1	0.0327280916821511	2.86	chr8	71683552	71683701	-159497	NR_033652	MSC antisense RNA 1	ncRNA
SYAP1	0.0327774038166528	-2.95	chrX	16745453	16746010	26147	NM_032796	synapse associated protein 1	protein-coding
RAB18	0.0328914686685027	2.88	chr10	27483024	27483194	-21065	NM_001256410	RAB18, member RAS oncogene family	protein-coding
LOC101928851	0.0330641780399511	2.34	chr4	57661675	57661894	235912	NR_125907	uncharacterized LOC101928851	ncRNA
LOC101927907	0.0332054266145048	3.36	chr2	77028040	77028413	42261	NR_110284	uncharacterized LOC101927907	ncRNA
NHSL1	0.0333591886779436	3.44	chr6	138520485	138520709	-21155	NM_001144060	NHS like 1	protein-coding
LRIG3	0.0334160417076283	3.42	chr12	59165136	59165332	-244696	NM_153377	leucine rich repeats and immunoglobulin like domains 3	protein-coding
mars-06	0.0336338461536775	-3.87	chr5	10344549	10344751	-8989	NM_005885	membrane associated ring-CH-type finger 6	protein-coding
LINC00906	0.0336645771120947	3.45	chr19	28914458	28914609	-50598	NR_027318	long intergenic non-protein coding RNA 906	ncRNA
CAT	0.0336851841866944	2.51	chr11	34474271	34474462	35441	NM_001752	catalase	protein-coding
NRG1-IT1	0.0338778710045109	3.26	chr8	32140740	32140971	114637	NR_104156	NRG1 intronic transcript 1	ncRNA
SF3A3	0.0339549643551593	-2.57	chr1	37995851	37996177	-5904	NM_001320830	splicing factor 3a subunit 3	protein-coding
MIR592	0.0340803875948217	3.1	chr7	126899984	126900427	157979	NR_030323	microRNA 592	ncRNA
LOC101929563	0.0341913059413914	-2.25	chr9	23307934	23308889	363988	NR_121602	uncharacterized LOC101929563	ncRNA
MIR4438	0.0342138168430065	3.3	chr2	214019349	214019724	261469	NR_039640	microRNA 4438	ncRNA
CCDC26	0.0343449279627109	3.71	chr8	129708057	129708231	-27905	NR_130917	CCDC26 long non-coding RNA	ncRNA
LOC100507194	0.0343940869398448	-2.23	chr6	3176757	3177127	18825	NR_038296	uncharacterized LOC100507194	ncRNA
BMPR1B	0.0345023777109496	3.43	chr4	94769735	94769966	11873	NM_001203	bone morphogenetic protein receptor type 1B	protein-coding
MYF6	0.0345087012948412	3.1	chr12	80708734	80708948	1212	NM_002469	myogenic factor 6	protein-coding
DHFR2	0.034606587930239	3.4	chr3	94130679	94130954	-67593	NM_001195643	dihydrofolate reductase 2	protein-coding
ICA1L	0.0349031709474068	3.71	chr2	202761418	202761566	67742	NM_001288624	islet cell autoantigen 1 like	protein-coding
TTC6	0.0349348463465712	2.92	chr14	37718704	37718900	123603	NM_001310135	tetratricopeptide repeat domain 6	protein-coding
BASP1	0.0351706403364739	3.18	chr5	17077728	17077896	-139011	NM_001271606	brain abundant membrane attached signal protein 1	protein-coding
MIR1267	0.035189936731872	-3.59	chr13	107128556	107128754	402593	NR_031671	microRNA 1267	ncRNA
MRPS30	0.035205873659732	2.9	chr5	44879798	44880073	71010	NM_016640	mitochondrial ribosomal protein S30	protein-coding
DLGAP1	0.0352441180691758	3.66	chr18	4505948	4506474	-50945	NM_004746	DLG associated protein 1	protein-coding
CBX4	0.0354214619843197	-2.45	chr17	79874552	79874777	-35250	NM_003655	chromobox 4	protein-coding
SHB	0.0354659178551666	1.86	chr9	38153357	38153897	-84414	NM_003028	SH2 domain containing adaptor protein B	protein-coding
PKHD1	0.0354803515603095	3.33	chr6	51993737	51993883	93815	NM_170724	polycystic kidney and hepatic disease 1 (autosomal recessive)	protein-coding
MIR8066	0.035642668577513	2.77	chr4	101222656	101222869	18110	NR_107033	microRNA 8066	ncRNA
RAI14	0.0358179183697697	3.3	chr5	34532706	34532827	-123562	NM_001145522	retinoic acid induced 14	protein-coding
DRD2	0.0359334508962733	-2.22	chr11	113474277	113474719	781	NM_000795	dopamine receptor D2	protein-coding
ABCG2	0.0360024231505771	3.07	chr4	88093871	88094157	64845	NM_004827	ATP binding cassette subfamily G member 2 (Junior blood group)	protein-coding
TRIO	0.0360960316118183	-2.83	chr5	14334754	14334957	191153	NM_007118	trio Rho guanine nucleotide exchange factor	protein-coding
ACTR6	0.0360997801851409	3.65	chr12	100215150	100215377	14486	NR_048568	ARP6 actin-related protein 6 homolog	protein-coding
LOC105370605	0.0363692718668869	3.41	chr14	84750981	84751327	-636121	NR_135155	uncharacterized LOC105370605	ncRNA
LOC102723661	0.0365066428930189	3.64	chr1	95994664	95994965	2746	NR_125991	uncharacterized LOC102723661	ncRNA
CYP46A1	0.0365545531270209	-2.99	chr14	99667227	99667372	-17119	NM_006668	cytochrome P450 family 46 subfamily A member 1	protein-coding
CCDC140	0.0366527262573337	-2.77	chr2	221969252	221969661	-328691	NM_153038	coiled-coil domain containing 140	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
NUDT12	0.0367894482703298	-3.39	chr5	103674959	103675149	-112253	NM_001300741	nudix hydrolase 12	protein-coding
LOC101927911	0.0368820057925682	3.42	chr17	2933912	2934036	31921	NR_110818	uncharacterized LOC101927911	ncRNA
LINC01020	0.0369133873285017	3.01	chr5	4957590	4957738	-76695	NR_026994	long intergenic non-protein coding RNA 1020	ncRNA
SEMA6A	0.03700175396829	3.31	chr5	116667649	116667831	-92814	NM_001300780	semaphorin 6A	protein-coding
CSMD3	0.0370048030618373	3.39	chr8	113229828	113230044	147217	NM_198124	CUB and Sushi multiple domains 3	protein-coding
MAN1A1	0.0371045004058714	2.83	chr6	119308135	119308482	41458	NM_005907	mannosidase alpha class 1A member 1	protein-coding
SH3YL1	0.0372108504629208	-2.21	chr2	246642	246872	9583	NM_001282687	SH3 and SYLF domain containing 1	protein-coding
LINC00613	0.0372347768226311	3.39	chr4	135238377	135238557	649626	NR_103762	long intergenic non-protein coding RNA 613	ncRNA
NNT-AS1	0.0373336028604168	-3.92	chr5	43642445	43642621	-39303	NR_073113	NNT antisense RNA 1	ncRNA
RASSF5	0.0374685112922722	-2.82	chr1	206557229	206557531	216	NM_182665	Ras association domain family member 5	protein-coding
LYPD6B	0.0375660557926187	-3.59	chr2	149030981	149031156	-7399	NM_001317005	LY6/PLAUR domain containing 6B	protein-coding
LINC00593	0.0375664523340004	-3.31	chr15	69802739	69802911	-32409	NR_026764	long intergenic non-protein coding RNA 593	ncRNA
LOC729506	0.0376254991799157	3.96	chr5	8300615	8300784	156865	NR_039984	uncharacterized LOC729506	ncRNA
LOC105374972	0.0379764352425818	-2.3	chr6	22880914	22881381	-163223	NR_134614	uncharacterized LOC105374972	ncRNA
SH3BP4	0.0380730990880951	2.73	chr2	235075384	235075753	123584	NM_014521	SH3 domain binding protein 4	protein-coding
LINC00458	0.0381555637188321	3.68	chr13	54137465	54137741	-4732	NR_108062	long intergenic non-protein coding RNA 458	ncRNA
RORA	0.0382241448209988	-2.83	chr15	61038578	61038859	190585	NM_134261	RAR related orphan receptor A	protein-coding
GLIS3	0.0383227566730988	-2.95	chr9	4143807	4144071	8244	NM_152629	GLIS family zinc finger 3	protein-coding
IMPG1	0.0383375072308888	-3.53	chr6	76701828	76702060	-629266	NM_001563	interphotoreceptor matrix proteoglycan 1	protein-coding
GNA11	0.0383697225561857	1.91	chr19	3098796	3098967	4471	NM_002067	G protein subunit alpha 11	protein-coding
GMPR	0.0383895186831891	2.23	chr6	16245003	16245267	6555	NM_006877	guanosine monophosphate reductase	protein-coding
KCNQ5-AS1	0.0387505470678911	-2.42	chr6	73174773	73174937	-31341	NR_046621	KCNQ5 antisense RNA 1	ncRNA
MYO6	0.0388043490886287	3.36	chr6	75735811	75736104	-13219	NM_001300899	myosin VI	protein-coding
MAPRE2	0.039143637810436	2.12	chr18	35085696	35085947	44171	NR_046177	microtubule associated protein RP/EB family member 2	protein-coding
IL15	0.0391717254385967	-2.76	chr4	141942692	141942861	306180	NM_000585	interleukin 15	protein-coding
CXADR	0.0392674696753653	-3.39	chr21	17560326	17560585	47549	NM_001207066	coxsackie virus and adenovirus receptor	protein-coding
TP53TG3	0.0393548526802804	-3.2	chr16	33338945	33339097	-21856	NR_110910	TP53 target 3	protein-coding
DOCK3	0.0393634756522217	2.9	chr3	50930688	50930871	255538	NM_004947	dedicator of cytokinesis 3	protein-coding
PXDN	0.0394281655075347	3.24	chr2	1886627	1886980	-142256	NM_012293	peroxidasin	protein-coding
USP17L6P	0.0394607384052715	-2.59	chr4	9406015	9406383	38325	NR_027279	ubiquitin specific peptidase 17-like family member 6, pseudogene	pseudo
KIF16B	0.0395808863019469	3.14	chr20	16634873	16635051	-61528	NM_001199866	kinesin family member 16B	protein-coding
ANK3	0.0395873197854135	-3.62	chr10	60309147	60309394	80714	NM_020987	ankyrin 3	protein-coding
SMARCA2	0.0396401075542034	-2.48	chr9	2115953	2116493	-41457	NM_001289398	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	protein-coding
LOC105369509	0.0396471580542342	3.62	chr11	115135170	115135349	-198318	NR_135108	uncharacterized LOC105369509	ncRNA
FAM183BP	0.0397308933482288	-3.27	chr7	38793245	38793558	-106312	NR_028347	acyloxyacyl hydrolase (neutrophil)	pseudo
CTNND2	0.0397809517811421	-3.12	chr5	11954129	11954300	-50171	NR_109988	catenin delta 2	protein-coding
PCCB	0.0398809081641958	-2.96	chr3	136276041	136276212	25801	NM_000532	propionyl-CoA carboxylase beta subunit	protein-coding
PTGS2	0.039881829584069	3.83	chr1	186577278	186577444	103066	NM_000963	prostaglandin-endoperoxide synthase 2	protein-coding
CBX4	0.0400030738776374	-2.36	chr17	79843928	79844325	-4712	NM_003655	chromobox 4	protein-coding
LINC01192	0.0400293899024345	3.86	chr3	163089789	163089960	213427	NR_033945	long intergenic non-protein coding RNA 1192	ncRNA
SLC41A3	0.040098233259778	-2.51	chr3	126032541	126032700	24155	NM_001008487	solute carrier family 41 member 3	protein-coding
LINC01243	0.0401220712697704	4.03	chr9	31427036	31427222	-45639	NR_135134	long intergenic non-protein coding RNA 1243	ncRNA
WASF3	0.040123921444191	-2.83	chr13	26435120	26435526	-122380	NM_006646	WAS protein family member 3	protein-coding
NTS	0.0401771075254726	-3.14	chr12	85943623	85943821	69427	NM_006183	neurotensin	protein-coding
EDN1	0.0402353296291226	-2.65	chr6	12492400	12492655	202231	NM_001955	endothelin 1	protein-coding
LOC105377962	0.0402369002943943	-3.03	chr6	115732656	115732949	-88967	NR_134602	uncharacterized LOC105377962	ncRNA
NTRK3-AS1	0.0402938600237409	-2.64	chr15	88246722	88247069	-5835	NR_038229	NTRK3 antisense RNA 1	ncRNA
ECT2L	0.0404621647761807	-2.92	chr6	138816020	138816199	19998	NM_001077706	epithelial cell transforming 2 like	protein-coding
ZNF8	0.0404661763469635	3.11	chr19	58268699	58268864	-10171	NM_021089	zinc finger protein 8	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
PLA2G4B	0.0405832513331771	-2.3	chr15	41836446	41837149	-2016	NM_001114633	phospholipase A2 group IVB	protein-coding
GIPC3	0.0406044403013161	-2	chr19	3579987	3580409	-5373	NM_133261	GIPC PDZ domain containing family member 3	protein-coding
TBX18	0.0406435771437021	3.28	chr6	84876637	84876800	-112482	NM_001080508	T-box 18	protein-coding
UCHL5	0.0409111802547902	3.29	chr1	193067842	193068018	-7823	NR_037607	ubiquitin C-terminal hydrolase L5	protein-coding
LOC101928226	0.040930584940284	-2.94	chr1	248703866	248704058	6900	NR_125950	uncharacterized LOC101928226	ncRNA
ZEB1	0.0410562468608402	2.59	chr10	31353535	31354063	32178	NM_001323671	zinc finger E-box binding homeobox 1	protein-coding
AKT3	0.0410690883380926	-3.01	chr1	243819177	243819436	23976	NM_005465	AKT serine/threonine kinase 3	protein-coding
GPC1	0.0412718068697885	2.91	chr2	240320530	240320706	-115080	NM_002081	glypican 1	protein-coding
MIR8054	0.041374809662916	3.13	chr11	23942839	23943023	-523741	NR_107021	microRNA 8054	ncRNA
LINGO2	0.0414291016481568	3.38	chr9	29500965	29501140	-288052	NM_001258282	leucine rich repeat and Ig domain containing 2	protein-coding
ZNF727	0.0414954230608174	2.96	chr7	64043440	64043578	-1934	NM_001159522	zinc finger protein 727	protein-coding
NRG1-IT3	0.0416587966381213	3.49	chr8	32493007	32493216	52365	NR_047475	NRG1 intronic transcript 3	ncRNA
LOC105374194	0.0416750712887115	3.92	chr3	166444167	166444285	-454457	NR_135545	uncharacterized LOC105374194	ncRNA
LINC00603	0.0418393222052677	-3.05	chr5	40265965	40266144	213763	NR_104633	long intergenic non-protein coding RNA 603	ncRNA
LOC101929698	0.0421369407320162	-3.27	chr20	32532859	32533022	22981	NR_110619	uncharacterized LOC101929698	ncRNA
ADGRL4	0.0421930411911497	-2.68	chr1	79367711	79367925	-361008	NM_022159	adhesion G protein-coupled receptor L4	protein-coding
LOC285766	0.0423321553515464	2.88	chr6	186083	186259	19313	NR_126020	uncharacterized LOC285766	ncRNA
LINC01256	0.0424118724634957	-2.86	chr4	132648290	132648617	57364	NR_126401	long intergenic non-protein coding RNA 1256	ncRNA
LOC101927752	0.0424477273091485	-2.46	chr8	1746098	1746523	17024	NR_134302	uncharacterized LOC101927752	ncRNA
CNTN6	0.0424617767650693	2.99	chr3	1110817	1111013	18257	NM_014461	contactin 6	protein-coding
LINC01243	0.0425668677670194	3.73	chr9	31563409	31563587	-182008	NR_135134	long intergenic non-protein coding RNA 1243	ncRNA
TMEM179	0.042755405930074	3.14	chr14	104624846	104624997	-20161	NM_001286390	transmembrane protein 179	protein-coding
ST8SIA6	0.0427952187165898	3.89	chr10	17452360	17452505	1955	NM_001004470	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 6	protein-coding
HACE1	0.0428354968743661	3.76	chr6	104800572	104800694	58913	NM_001321083	HECT domain and ankyrin repeat containing E3 ubiquitin protein ligase 1	protein-coding
LINC01361	0.0429821701274491	-3	chr1	83134516	83134724	-148412	NR_110633	long intergenic non-protein coding RNA 1361	ncRNA
MIR187	0.0430813076206922	-2.54	chr18	35893865	35894378	10805	NR_029616	microRNA 187	ncRNA
F5	0.0431142782258829	-3	chr1	169541117	169541314	45316	NM_000130	coagulation factor V	protein-coding
ATXN2	0.0431251434302605	2.54	chr12	111636239	111636682	-36784	NM_002973	ataxin 2	protein-coding
PDE1C	0.0431869990111193	-3.47	chr7	31932469	31932696	138280	NM_001191056	phosphodiesterase 1C	protein-coding
LEPROT	0.0432544706773868	-3.06	chr1	65407999	65408274	-12312	NM_001198683	leptin receptor overlapping transcript	protein-coding
IMMP2L	0.0433547186009881	2.53	chr7	111532121	111532287	30087	NM_032549	inner mitochondrial membrane peptidase subunit 2	protein-coding
RER1	0.043437511193691	2.86	chr1	2401826	2402037	10156	NM_007033	retention in endoplasmic reticulum sorting receptor 1	protein-coding
LINC01813	0.0435870115790664	3.08	chr2	56108986	56109231	-18726	NR_135590	long intergenic non-protein coding RNA 1813	ncRNA
LINC00442	0.0438138845857701	3.64	chr13	19075872	19076078	67716	NR_026852	long intergenic non-protein coding RNA 442	ncRNA
MGC16025	0.0438349529958623	-2.9	chr2	239214189	239214352	-18813	NR_026664	uncharacterized LOC85009	ncRNA
MIR7848	0.0439419144813248	3.17	chr8	133000062	133000387	46357	NR_107002	microRNA 7848	ncRNA
LOC102724874	0.0441207091573621	3.58	chr8	77935888	77936072	485283	NR_125830	uncharacterized LOC102724874	ncRNA
PGR	0.0442078185822669	-3.61	chr11	101166159	101166401	-36467	NM_000926	progesterone receptor	protein-coding
ROR2	0.0442345023680908	2.94	chr9	92018055	92018375	-68053	NM_004560	receptor tyrosine kinase like orphan receptor 2	protein-coding
KCNN2	0.0443485533181295	3.13	chr5	114467907	114468129	34488	NM_170775	potassium calcium-activated channel subfamily N member 2	protein-coding
RAB27B	0.0443614462673438	3.26	chr18	54830811	54831022	2439	NM_004163	RAB27B, member RAS oncogene family	protein-coding
CNTFR	0.044539105996786	-2.09	chr9	34591071	34591265	-1028	NM_001207011	ciliary neurotrophic factor receptor	protein-coding
LOC100507377	0.044551401840864	3.75	chr12	74017309	74017494	275230	NR_038300	uncharacterized LOC100507377	ncRNA
TRPS1	0.0447157979734061	3.29	chr8	115228243	115228425	439678	NM_001282902	transcriptional repressor GATA binding 1	protein-coding
MYOF	0.0447460699320353	2.86	chr10	93289367	93289578	192845	NM_013451	myoferlin	protein-coding
LOC105377143	0.0447893794259169	3.18	chr3	66727302	66727433	10969	NR_135533	uncharacterized LOC105377143	ncRNA
ARHGEF10L	0.0448476429482719	-1.84	chr1	17559896	17560602	-20304	NM_001319837	Rho guanine nucleotide exchange factor 10 like	protein-coding
LOC100130264	0.0448552992471816	-3.25	chr20	19254177	19254376	30320	NR_024564	uncharacterized LOC100130264	ncRNA
HSP90B2P	0.0450264005938516	3.28	chr15	99302149	99302472	44785	NR_073383	heat shock protein 90 beta family member 2, pseudogene	pseudo
GRHPR	0.0451396523504419	-2.93	chr9	37290992	37291237	-131596	NM_012203	glyoxylate and hydroxypyruvate reductase	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
PRSS48	0.0452533055049706	-2.7	chr4	151340922	151341397	63986	NM_183375	protease, serine 48	protein-coding
LINC00550	0.0452889139268274	-3.17	chr13	68897426	68897614	-12195	NR_038878	long intergenic non-protein coding RNA 550	ncRNA
ALG1L2	0.0453479862700715	-3.53	chr3	130036774	130036913	-44988	NM_001136152	ALG1, chitobiosyldiphosphodolichol beta-mannosyltransferase like 2	protein-coding
EDN1	0.0454672854414144	-2.98	chr6	12419655	12420042	129552	NM_001955	endothelin 1	protein-coding
FAM8A1	0.0455553245250473	2.92	chr6	17559972	17560156	-40223	NM_016255	family with sequence similarity 8 member A1	protein-coding
PMEPA1	0.04568863956363	2.84	chr20	57734441	57734567	-22968	NM_199171	prostate transmembrane protein, androgen induced 1	protein-coding
ROR2	0.04568863956363	2.84	chr9	91888429	91888569	61663	NM_001318204	receptor tyrosine kinase like orphan receptor 2	protein-coding
MIR3134	0.045690714731102	-3.36	chr3	15687841	15688026	9438	NR_036085	microRNA 3134	ncRNA
TCF4	0.0457401001272765	3.18	chr18	55256737	55256979	45763	NM_001243236	transcription factor 4	protein-coding
GBE1	0.0457535456317659	2.68	chr3	81708893	81709064	52821	NM_000158	1,4-alpha-glucan branching enzyme 1	protein-coding
PNLIP	0.0458777601647955	3.45	chr10	116490221	116490612	-55508	NM_000936	pancreatic lipase	protein-coding
SETBP1	0.045912688780881	2.93	chr18	44803144	44803317	122332	NM_015559	SET binding protein 1	protein-coding
CDH18	0.0459206263541069	-2.66	chr5	18947178	18947503	938932	NM_001291957	cadherin 18	protein-coding
APOOP5	0.0460121197812955	3.07	chr16	59367844	59368023	387258	NR_028471	apolipoprotein O pseudogene 5	pseudo
PEX5L	0.0461634880577258	-3.26	chr3	180210281	180210457	-173316	NM_001256751	peroxisomal biogenesis factor 5 like	protein-coding
SLC9A4	0.0465242549768386	3.72	chr2	102520302	102520481	47088	NM_001011552	solute carrier family 9 member A4	protein-coding
SRD5A2	0.0465660261821028	-2.4	chr2	31567792	31568341	12905	NM_000348	steroid 5 alpha-reductase 2	protein-coding
MIR887	0.046660909407815	2.84	chr5	15951780	15951948	16682	NR_030616	microRNA 887	ncRNA
PSMB7	0.046662668433576	3.09	chr9	124419975	124420095	-4562	NM_002799	proteasome subunit beta 7	protein-coding
ZNF385D-AS2	0.0467681351719569	3.32	chr3	22485608	22485794	528562	NR_046876	ZNF385D antisense RNA 2	ncRNA
CCDC71L	0.0468279227796483	-3.12	chr7	106732430	106732616	-71335	NM_175884	coiled-coil domain containing 71-like	protein-coding
FGF10	0.0469182931455934	-3.05	chr5	44090187	44090354	298412	NM_004465	fibroblast growth factor 10	protein-coding
SNORD15B	0.047096613208981	3.82	chr11	75406419	75406673	2125	NR_000025	small nucleolar RNA, C/D box 15B	snoRNA
CLVS2	0.047096613208981	3.82	chr6	123030341	123030734	34566	NM_001010852	clavesin 2	protein-coding
SP140L	0.0471772922166335	3.36	chr2	230360607	230360756	33501	NM_001308162	SP140 nuclear body protein like	protein-coding
CYP4F12	0.0471777450588339	-2.7	chr19	15667154	15667317	-5783	NR_117085	cytochrome P450 family 4 subfamily F member 12	protein-coding
KLHL3	0.0473229934059101	3.1	chr5	137719574	137719842	1602	NM_001257194	kelch like family member 3	protein-coding
DLGAP1-AS2	0.0474392780685646	-2.41	chr18	3652869	3653105	49250	NR_119377	DLGAP1 antisense RNA 2	ncRNA
AGPS	0.0475573320736652	3	chr2	177396619	177396816	3974	NM_003659	alkylglycerone phosphate synthase	protein-coding
MIR4782	0.0475888868965432	3.57	chr2	113676367	113676533	44918	NR_039943	microRNA 4782	ncRNA
CPS1	0.0476050536783378	1.99	chr2	210549000	210549339	-7430	NM_001875	carbamoyl-phosphate synthase 1	protein-coding
PIP	0.0476619934591291	-2.85	chr7	143140789	143141001	8814	NM_002652	prolactin induced protein	protein-coding
DMRT2	0.0476842373100162	3.07	chr9	1291606	1291771	241068	NM_181872	doublesex and mab-3 related transcription factor 2	protein-coding
LINC01091	0.047791256923292	-2.46	chr4	123819509	123819754	45367	NR_027106	long intergenic non-protein coding RNA 1091	ncRNA
PATE1	0.0478142311626139	3.48	chr11	125728839	125728989	-17379	NM_138294	prostate and testis expressed 1	protein-coding
IRAIN	0.0478582287616755	3.06	chr15	98668790	98668908	-17628	NR_126453	IGF1R antisense imprinted non-protein coding RNA	ncRNA
RAB11A	0.0478793480666918	-2.53	chr15	65857067	65857272	-12290	NM_004663	RAB11A, member RAS oncogene family	protein-coding
MIR6776	0.0478880684051366	-3.33	chr17	2659635	2659818	33193	NR_106834	microRNA 6776	ncRNA
FAM105A	0.0480233187737708	-2.59	chr5	14467436	14467735	-114197	NM_019018	family with sequence similarity 105 member A	protein-coding
LOC101928820	0.0480439449367098	-2.99	chr6	85248754	85248948	141335	NR_134634	uncharacterized LOC101928820	ncRNA
MIR6073	0.0480606361957104	3.53	chr11	15849472	15849639	120066	NR_106721	microRNA 6073	ncRNA
TMCC1	0.048064254946614	-2.92	chr3	129690561	129690761	-1929	NR_033361	transmembrane and coiled-coil domain family 1	protein-coding
LOC91450	0.0481604715222496	3.02	chr15	77966744	77967018	27344	NR_026998	uncharacterized LOC91450	ncRNA
SMAD5	0.0481638786988814	-3.1	chr5	136107849	136108014	-24914	NM_001001419	SMAD family member 5	protein-coding
YTHDC2	0.0482741485209411	-2.44	chr5	113753155	113753435	239633	NM_001345975	YTH domain containing 2	protein-coding
TMEM106B	0.048380112702715	3.57	chr7	12129059	12129229	-82078	NM_018374	transmembrane protein 106B	protein-coding
ATRX	0.048380112702715	3.57	chrX	77651980	77652144	134207	NM_000489	ATRX, chromatin remodeler	protein-coding
ANK3	0.0484142417837067	-3.05	chr10	60123230	60123919	17442	NM_001149	ankyrin 3	protein-coding
GJB2	0.0484339647592279	3.96	chr13	20198723	20198973	-5873	NM_004004	gap junction protein beta 2	protein-coding
ANO4	0.0486327666555982	3.15	chr12	100866832	100866949	-34756	NM_001286616	anoctamin 4	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
UBAP1L	0.0486653086944772	2.77	chr15	65105630	65105775	657	NM_001163692	ubiquitin associated protein 1 like	protein-coding
NELL1	0.0486710571139199	-3.18	chr11	21220100	21220432	550715	NM_006157	neural EGFL like 1	protein-coding
CHAC2	0.0486739449093943	2.52	chr2	53437604	53437869	-330047	NM_001346127	ChaC cation transport regulator homolog 2	protein-coding
DLEU7-AS1	0.0486864967299535	2.74	chr13	50800778	50800943	-6995	NR_046551	DLEU7 antisense RNA 1	ncRNA
LNCPRESS2	0.0486969227282587	3.14	chr4	92189095	92189256	88127	NR_125920	lncRNA p53 regulated and ESC associated 2	ncRNA
LOC101927915	0.0488558713071197	-2.41	chr8	137492899	137493288	-79505	NR_125428	uncharacterized LOC101927915	ncRNA
BRD3	0.0488791873195468	-3.47	chr9	134080882	134081038	-12941	NM_007371	bromodomain containing 3	protein-coding
LRRC17	0.0488918259243005	2.65	chr7	102928656	102928823	15842	NM_005824	leucine rich repeat containing 17	protein-coding
SLC35G2	0.0489692192112843	3.33	chr3	136790538	136790664	-28418	NM_025246	solute carrier family 35 member G2	protein-coding
COBL	0.0490191358949223	-2.79	chr7	51543889	51544160	-227163	NM_001346444	cordon-bleu WH2 repeat protein	protein-coding
SNORA33	0.0490638990868723	-3.06	chr6	132820841	132821023	3713	NR_002436	small nucleolar RNA, H/ACA box 33	snoRNA
CNTFR-AS1	0.049107395216069	-1.89	chr9	34547944	34548218	-19931	NR_024369	CNTFR antisense RNA 1	ncRNA
BPY2	0.0492936855483905	2.84	chrY	24619547	24619850	1694	NM_004678	basic charge, Y-linked, 2	protein-coding
INTU	0.0494809639551204	2.35	chr4	127538805	127539033	-94013	NM_015693	inturned planar cell polarity protein	protein-coding
CIITA	0.0495019867076472	2.91	chr16	10888988	10889127	11859	NM_001286403	class II major histocompatibility complex transactivator	protein-coding
CYP3A7	0.0495192116627951	-2.38	chr7	99751539	99751854	-16496	NM_000765	cytochrome P450 family 3 subfamily A member 7	protein-coding
HACE1	0.0496878043551715	3.56	chr6	103541754	103541922	1317708	NM_001321083	HECT domain and ankyrin repeat containing E3 ubiquitin protein ligase 1	protein-coding
SRP14	0.0497305662407535	2.81	chr15	40003454	40003688	35631	NM_003134	signal recognition particle 14	protein-coding
TFEB	0.0497319947027034	2.25	chr6	41727566	41727802	6717	NM_001271945	transcription factor EB	protein-coding
MKRN9P	0.0498423955668643	-2.86	chr12	87799858	87800097	-15266	NR_033410	makorin ring finger protein 9, pseudogene	pseudo
OCA2	0.0498423955668643	-2.86	chr15	27813423	27813595	285803	NM_000275	OCA2 melanosomal transmembrane protein	protein-coding
MIR4454	0.0498659811354539	-2.64	chr5	12772386	12772588	104628	NR_039659	microRNA 4454	ncRNA
CNTNAP2	0.0498786764182062	3.21	chr7	145623626	145623825	-492636	NM_014141	contactin associated protein-like 2	protein-coding
OPRK1	0.0499513498520724	-3.22	chr8	53527112	53527308	-275513	NM_000912	opioid receptor kappa 1	protein-coding

Supplementary Table 5

List of genes with altered chromatin accessibility 72 hrs post-infection

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
MEOX2-AS1	0.000318131479529961	5.01	chr7	15841167	15841409	152910	NR_110094	MEOX2 antisense RNA 1 (head to head)	ncRNA
OPA1-AS1	0.000469140963215896	4.46	chr3	193717306	193717473	-90057	NR_046634	OPA1 antisense RNA 1	ncRNA
C14orf180	0.000516575499668394	4.71	chr14	104476355	104476506	-103254	NM_001008404	chromosome 14 open reading frame 180	protein-coding
ITGB6	0.000710777391317259	2.27	chr2	160221777	160222128	-21639	NM_000888	integrin subunit beta 6	protein-coding
NDST4	0.000840593391869134	4.55	chr4	115680031	115680236	-566257	NM_022569	N-deacetylase and N-sulfotransferase 4	protein-coding
RNF219-AS1	0.000840643185625352	4.57	chr13	78276405	78276568	221631	NR_047001	RNF219 antisense RNA 1	ncRNA
ADAMTSL1	0.000905172581752592	4.12	chr9	18516577	18516837	42626	NM_001040272	ADAMTS like 1	protein-coding
ROCK1	0.000930708995819819	3.22	chr18	18629444	18629603	2482328	NM_005406	Rho associated coiled-coil containing protein kinase 1	protein-coding
EIF3E	0.000962333295625573	4.44	chr8	108205936	108206154	42685	NM_001568	eukaryotic translation initiation factor 3 subunit E	protein-coding
ELFN1	0.00105780437157664	2.72	chr7	1674979	1675181	-34082	NM_001128636	extracellular leucine rich repeat and fibronectin type III domain containing 1	protein-coding
LOC93463	0.0010813536897086	2.75	chr2	236874456	236875042	181418	NR_135202	uncharacterized LOC93463	ncRNA
ANKRD33	0.00112672272130934	2.87	chr12	51874499	51874706	-13407	NM_001304460	ankyrin repeat domain 33	protein-coding
SPTBN1	0.00120409058129951	4.49	chr2	54618536	54618698	60223	NM_178313	spectrin beta, non-erythrocytic 1	protein-coding
DAZL	0.00125677464467644	4.84	chr3	16700451	16700611	-95032	NM_001351	deleted in azoospermia like	protein-coding
LINC02026	0.00127511932852604	3.92	chr3	194024135	194024311	-20564	NR_033944	long intergenic non-protein coding RNA 2026	ncRNA
LOC101929468	0.00131307263693627	3.65	chr4	104394511	104394704	-96358	NR_125926	uncharacterized LOC101929468	ncRNA
CNTNAP5	0.00140769156852111	4.33	chr2	124733919	124734167	708756	NM_130773	contactin associated protein like 5	protein-coding
NAALADL2-AS1	0.00142447596493914	4.85	chr3	175606906	175607075	169343	NR_046714	NAALADL2 antisense RNA 1	ncRNA
SEPT7P2	0.00147394869260097	2.64	chr7	45827007	45827167	-58069	NR_024271	septin 7 pseudogene 2	pseudo
FAM120C	0.0015427247155043	3.79	chrX	54162666	54162939	20456	NM_017848	family with sequence similarity 120C	protein-coding
LINC01608	0.00159577704703767	3.95	chr8	110748422	110748624	278910	NR_125416	long intergenic non-protein coding RNA 1608	ncRNA
SNX14	0.00160643802244995	3.83	chr6	85575398	85575680	18617	NM_001297614	sorting nexin 14	protein-coding
COL6A4P2	0.00163569280617407	2.63	chr3	130182067	130182612	-30481	NR_027898	collagen type VI alpha 4 pseudogene 2	pseudo
DMD	0.00184159819013682	3.4	chrX	31424168	31424463	83983	NM_004014	dystrophin	protein-coding
LOC645166	0.00201216732739505	2.72	chr1	143423561	143423784	1012	NR_027355	lymphocyte-specific protein 1 pseudogene	pseudo
ZNF385B	0.00206464621346405	4.99	chr2	179605722	179605945	-19927	NR_104234	zinc finger protein 385B	protein-coding
MDGA2	0.00209093160114678	5.32	chr14	47340887	47341048	2279	NR_103766	MAM domain containing glycosylphosphatidylinositol anchor 2	protein-coding
RUNX2	0.00216609210700747	3.48	chr6	45540434	45540594	117937	NM_001278478	runt related transcription factor 2	protein-coding
LOC644919	0.00216835244601175	4.14	chr14	40295578	40295739	-659053	NR_109758	uncharacterized LOC644919	ncRNA
MIR6076	0.00218124124405174	-3.94	chr14	49950843	49951063	-15446	NR_106724	microRNA 6076	ncRNA
HYI	0.00222811259794472	3.66	chr1	43472877	43473043	-18971	NM_001330526	hydroxypyruvate isomerase (putative)	protein-coding
SMAD7	0.00232771643982834	1.91	chr18	48913173	48913491	29475	NM_001190823	SMAD family member 7	protein-coding
LINC01258	0.00233466780420293	3.54	chr4	38397059	38397486	125908	NR_110951	long intergenic non-protein coding RNA 1258	ncRNA
PTPRD-AS2	0.00235411265799271	4.35	chr9	10904743	10904904	291617	NR_110696	PTPRD antisense RNA 2 (head to head)	ncRNA
DEPDC1	0.00236144671912197	4.51	chr1	68548483	68548633	-51337	NM_017779	DEP domain containing 1	protein-coding
LCORL	0.00242058761137759	4.69	chr4	18397147	18397317	-375357	NM_153686	ligand dependent nuclear receptor corepressor like	protein-coding
LOC646241	0.00247529618872356	4.3	chr5	17974501	17974674	167313	NR_134287	uncharacterized LOC646241	ncRNA
ERCC6L	0.00253006157876209	4.1	chrX	72246249	72246418	-7325	NM_017669	ERCC excision repair 6 like, spindle assembly checkpoint helicase	protein-coding
LOC284395	0.00260686475346161	2.69	chr19	29382313	29382466	143363	NR_040029	uncharacterized LOC284395	ncRNA
MIR2115	0.00266311270706123	3.82	chr3	48318669	48318824	-2287	NR_031749	microRNA 2115	ncRNA
LRFN5	0.00281792594178422	4.97	chr14	42803218	42803445	1196475	NM_001330106	leucine rich repeat and fibronectin type III domain containing 5	protein-coding
LRRTM4	0.00283967436716731	3.82	chr2	77270410	77270613	251922	NM_001330370	leucine rich repeat transmembrane neuronal 4	protein-coding
TUBB4A	0.00285369998296327	-3.44	chr19	6501969	6502168	516	NM_001289131	tubulin beta 4A class IVa	protein-coding
CDH6	0.00300330718432846	-3.06	chr5	30539090	30539520	-654350	NM_004932	cadherin 6	protein-coding
LINC01913	0.00300813017524543	3.14	chr2	41381766	41381963	-495691	NR_033996	long intergenic non-protein coding RNA 1913	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
DAPK1	0.0030084739966362	2.66	chr9	87343803	87344149	-153252	NM_001288731	death associated protein kinase 1	protein-coding
LAMA4	0.00303304294077919	2.77	chr6	112299388	112299984	-44971	NM_002290	laminin subunit alpha 4	protein-coding
FLJ37505	0.00305619509091283	3.47	chr12	127779211	127779378	-102323	NR_033987	uncharacterized LOC400087	ncRNA
UBBP4	0.00306743111272877	4.15	chr17	22123944	22124121	-78963	NR_144546	ubiquitin B pseudogene 4	pseudo
LCAL1	0.00307037641056822	3.8	chr6	79355921	79356131	-42642	NR_130915	lung cancer associated lncRNA 1	ncRNA
LOC102546299	0.00308130384059806	3.89	chr5	165183752	165183908	713551	NR_105065	uncharacterized LOC102546299	ncRNA
RTKN2	0.00312295839653739	2.49	chr10	62280441	62280605	-11660	NM_001282941	rhotekin 2	protein-coding
LINC01697	0.00312958002036239	2.89	chr21	28051664	28051856	3346	NR_126010	long intergenic non-protein coding RNA 1697	ncRNA
ST8SIA4	0.00316486905154704	4.34	chr5	101150185	101150380	-246997	NM_005668	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4	protein-coding
DTL	0.00321992988935371	3.38	chr1	211938908	211939195	-96502	NM_001286229	denticleless E3 ubiquitin protein ligase homolog	protein-coding
FAT3	0.00322749167071838	-4.19	chr11	92104613	92104782	-247399	NM_001008781	FAT atypical cadherin 3	protein-coding
HULC	0.00333500294268826	2.77	chr6	8982603	8982892	330538	NR_004855	hepatocellular carcinoma up-regulated long non-coding RNA	ncRNA
TMEM261	0.00335293433311314	3.9	chr9	8033144	8033294	-233413	NM_033428	transmembrane protein 261	protein-coding
GRIK4	0.00341202192049776	-3.8	chr11	120690148	120690353	125814	NM_001282473	glutamate ionotropic receptor kainate type subunit 4	protein-coding
EMB	0.00342066689821084	2.61	chr5	49704637	49705383	736390	NM_198449	embigin	protein-coding
TNR	0.00354369329682191	3.43	chr1	175544402	175544609	199111	NM_003285	tenascin R	protein-coding
LOC100132078	0.00366596557652266	3.26	chr11	111462224	111462396	-44124	NR_046085	uncharacterized LOC100132078	ncRNA
LINC01648	0.00369595329080067	2.57	chr1	30330797	30330959	-293266	NR_110790	long intergenic non-protein coding RNA 1648	ncRNA
GAS6-AS1	0.00369927565145115	2.34	chr13	113788800	113788975	-26723	NR_044995	GAS6 antisense RNA 1	ncRNA
C3orf58	0.00372149498252629	3.71	chr3	144854027	144854209	880793	NM_001134470	chromosome 3 open reading frame 58	protein-coding
CLECL1	0.00373291800601734	2.39	chr12	9739715	9740123	-6620	NM_001253750	C-type lectin like 1	protein-coding
ACSBG1	0.00377965076153212	2.59	chr15	78245663	78245866	-11057	NM_001199377	acyl-CoA synthetase bubblegum family member 1	protein-coding
LRRC32	0.00381805186038032	-4.07	chr11	76605927	76606083	63995	NM_005512	leucine rich repeat containing 32	protein-coding
BAP1	0.00386402347546189	3.15	chr3	52386853	52387002	23178	NM_004656	BRCA1 associated protein 1	protein-coding
PRR18	0.00388991633216755	2.55	chr6	166296775	166296925	11533	NM_175922	proline rich 18	protein-coding
LSAMP-AS1	0.00392637506491	2.59	chr3	116129542	116129785	-230361	NR_109998	LSAMP antisense RNA 1	ncRNA
LINGO2	0.00396994793282743	-3.9	chr9	28667646	28667887	51539	NM_152570	leucine rich repeat and Ig domain containing 2	protein-coding
LINC00494	0.00401015950814457	-3.46	chr20	48372932	48373096	13103	NR_026958	long intergenic non-protein coding RNA 494	ncRNA
LOC101927272	0.00402274872617519	2.5	chr16	50644968	50645180	-23781	NR_110908	uncharacterized LOC101927272	ncRNA
LINC01949	0.00418646831459433	-4.57	chr5	87232281	87232441	6840	NR_130914	long intergenic non-protein coding RNA 1949	ncRNA
PACRG-AS3	0.00429293840976898	2.83	chr6	163231574	163231736	-39653	NR_122120	PACRG antisense RNA 3	ncRNA
WWOX	0.00434551237250674	-3.92	chr16	78740450	78740676	641150	NM_001291997	WW domain containing oxidoreductase	protein-coding
LOC105372038	0.00444668995067161	2.93	chr18	27405221	27405417	-62663	NR_134589	uncharacterized LOC105372038	ncRNA
SNORA107	0.0045286054887073	3.97	chr13	84535350	84535561	-248199	NR_132852	small nucleolar RNA, H/ACA box 107	snoRNA
SSBP3	0.00454349519391172	3.3	chr1	54385020	54385227	21272	NM_018070	single stranded DNA binding protein 3	protein-coding
LOC105369739	0.00455193573570523	3.53	chr12	43108289	43108502	416179	NR_135028	uncharacterized LOC105369739	ncRNA
NCS1	0.00456673319171218	2.27	chr9	130219873	130220150	19418	NM_001128826	neuronal calcium sensor 1	protein-coding
PABPC4L	0.00458403955217053	3.89	chr4	135037186	135037382	-835536	NM_001114734	poly(A) binding protein cytoplasmic 4 like	protein-coding
FBXL5	0.00461350266753687	4.21	chr4	15607762	15607937	47563	NR_036464	F-box and leucine rich repeat protein 5	protein-coding
SUSD6	0.00469681835122881	2.58	chr14	69663414	69663619	51923	NM_014734	sushi domain containing 6	protein-coding
TRAPPC11	0.00482286408674584	3.76	chr4	183726826	183726977	67634	NM_199053	trafficking protein particle complex 11	protein-coding
UBE2U	0.00487462207792743	2.66	chr1	64309408	64309616	105894	NM_152489	ubiquitin conjugating enzyme E2 U (putative)	protein-coding
ACER2	0.00494087356124411	1.94	chr9	19548384	19549505	140017	NM_001010887	alkaline ceramidase 2	protein-coding
VSTM2A	0.00500564862266944	2.12	chr7	54546688	54547036	698	NR_133927	V-set and transmembrane domain containing 2A	protein-coding
FRMPD2	0.00506599913077323	2.73	chr10	48254397	48254728	-2805	NM_001318191	FERM and PDZ domain containing 2	protein-coding
ZSCAN20	0.00512328012359842	2.29	chr1	33441019	33441198	-31523	NM_145238	zinc finger and SCAN domain containing 20	protein-coding
LINC01468	0.00519524459421287	4.58	chr10	52566817	52567116	-96433	NR_120641	long intergenic non-protein coding RNA 1468	ncRNA
SGIP1	0.00522297526572635	2.04	chr1	66533580	66534251	-41	NM_032291	SH3 domain GRB2 like endophilin interacting protein 1	protein-coding
GZMH	0.00529212036188445	-3.81	chr14	24605553	24605702	4093	NM_033423	granzyme H	protein-coding
CSNK1G2-AS1	0.0054179831846019	3.17	chr19	1964765	1964926	-10296	NR_033400	CSNK1G2 antisense RNA 1	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
FUT10	0.00543414789804882	3.14	chr8	33381637	33381784	91436	NM_032664	fucosyltransferase 10	protein-coding
CYP2C18	0.00544608943834588	4.72	chr10	94657864	94658037	-25544	NM_000772	cytochrome P450 family 2 subfamily C member 18	protein-coding
LINC01910	0.00544646537758895	3.55	chr18	70511629	70511777	-127112	NR_110764	long intergenic non-protein coding RNA 1910	ncRNA
C3orf79	0.00545553111453442	-3.51	chr3	153791841	153792119	307485	NM_001101337	chromosome 3 open reading frame 79	protein-coding
C4orf33	0.00553150551063327	4.39	chr4	129193627	129193818	97595	NM_001099783	chromosome 4 open reading frame 33	protein-coding
PRKCZ	0.00555000975664328	3.01	chr1	2085411	2085563	11501	NM_001242874	protein kinase C zeta	protein-coding
PDE3A	0.00560452917991013	2.56	chr12	20269426	20269698	-99683	NM_000921	phosphodiesterase 3A	protein-coding
UBBP4	0.00561039731597824	2.97	chr17	22042627	22042800	-160282	NR_144546	ubiquitin B pseudogene 4	pseudo
FLJ26245	0.00565750446667771	2.2	chr16	36963959	36964284	1217569	NR_033985	uncharacterized LOC400533	ncRNA
LOC105378146	0.00567533004822332	2.8	chr6	168763541	168763835	200694	NR_136250	uncharacterized LOC105378146	ncRNA
LOC100506858	0.00573129587247125	3.71	chr5	2429763	2429948	-117654	NR_104616	uncharacterized LOC100506858	ncRNA
COPB1	0.00576535471315005	-2.82	chr11	14443541	14444080	56048	NM_016451	coatomer protein complex subunit beta 1	protein-coding
FNDC3A	0.00577366195648685	2.96	chr13	48999014	48999187	22491	NR_103528	fibronectin type III domain containing 3A	protein-coding
RFX2	0.00579802426259119	-4.51	chr19	6120140	6120315	-9574	NM_000635	regulatory factor X2	protein-coding
NANOS1	0.00580814106235966	3.72	chr10	118905722	118905884	-123913	NM_199461	nanos C2HC-type zinc finger 1	protein-coding
SELL	0.00590955483040786	-3.24	chr1	169719129	169719519	-7622	NM_000655	selectin L	protein-coding
ACAP2	0.00591291183240616	1.98	chr3	195478594	195479271	-35844	NM_012287	ArfGAP with coiled-coil, ankyrin repeat and PH domains 2	protein-coding
POLR2M	0.00592373466370012	2.02	chr15	57732508	57732931	26198	NM_001018102	RNA polymerase II subunit M	protein-coding
RSPRY1	0.00595556743045068	2.87	chr16	57200244	57200426	14051	NM_133368	ring finger and SPRY domain containing 1	protein-coding
ACO1	0.00595582000820861	3.24	chr9	32264115	32264309	-120391	NM_001278352	aconitase 1	protein-coding
LOC105371046	0.00598433998030801	3.03	chr16	1562924	1563084	-17523	NR_135176	uncharacterized LOC105371046	ncRNA
PPFIA2	0.00603200456008884	3.75	chr12	81641707	81641966	-43465	NM_001220477	PTPRF interacting protein alpha 2	protein-coding
LRRTM3	0.00606690287795074	2.9	chr10	66389783	66390189	-536048	NM_178011	leucine rich repeat transmembrane neuronal 3	protein-coding
PCNX1	0.00607688102324815	2.94	chr14	70964454	70964774	57209	NM_014982	pecanex homolog 1 (Drosophila)	protein-coding
METTL4	0.00607962897052454	-3.97	chr18	2223794	2224031	347591	NM_001308401	methyltransferase like 4	protein-coding
ACTR3BP5	0.00608924478416346	3.36	chr10	40461184	40461356	1764674	NR_045000	ACTR3B pseudogene 5	pseudo
LINC01250	0.00612631108785573	-2.63	chr2	3023716	3024059	102139	NR_110228	long intergenic non-protein coding RNA 1250	ncRNA
FAT2	0.00614795599648376	2.67	chr5	151551039	151551209	17820	NM_001447	FAT atypical cadherin 2	protein-coding
PDE4B	0.00614898556596658	4.4	chr1	65807075	65807237	13983	NM_002600	phosphodiesterase 4B	protein-coding
LOC105369860	0.00617421187743134	2.16	chr12	78406710	78407320	-47269	NR_135021	uncharacterized LOC105369860	ncRNA
MIR2054	0.00620713058040456	3.24	chr4	126147680	126147874	640518	NR_031746	microRNA 2054	ncRNA
MYT1L	0.00626861748281584	3.77	chr2	2535254	2535411	-204019	NM_015025	myelin transcription factor 1 like	protein-coding
sept-10	0.00631156821766923	-4.12	chr2	109480541	109480819	132977	NM_001321514	septin 10	protein-coding
CELSR2	0.00632466904770201	3.48	chr1	109242118	109242281	-7820	NM_001408	cadherin EGF LAG seven-pass G-type receptor 2	protein-coding
BANF2	0.00633031887204402	3.09	chr20	17729182	17729380	29333	NM_178477	barrier to autointegration factor 2	protein-coding
KCNA5	0.00642128096659378	3.59	chr12	5134307	5134590	90529	NM_002234	potassium voltage-gated channel subfamily A member 5	protein-coding
PKIB	0.00646113919398467	2.9	chr6	122540146	122540299	68305	NM_001270394	cAMP-dependent protein kinase inhibitor beta	protein-coding
LOC441025	0.00650682446962498	3.64	chr4	75179560	75179710	-174441	NR_033964	uncharacterized LOC441025	ncRNA
DPYD-AS2	0.00652479457955517	-3.84	chr1	97499604	97499852	-297193	NR_046591	DPYD antisense RNA 2	ncRNA
LINC01790	0.00653703580543687	4.76	chr2	195049717	195049912	319219	NR_110223	long intergenic non-protein coding RNA 1790	ncRNA
GRID2	0.00663631167125483	2.99	chr4	93002420	93002843	698232	NM_001510	glutamate ionotropic receptor delta type subunit 2	protein-coding
LINGO1-AS2	0.00664376543036377	2.61	chr15	77651834	77652009	-8130	NR_120361	LINGO1 antisense RNA 2	ncRNA
ALG10	0.00664615951373052	2.95	chr12	36016916	36017277	1994815	NM_032834	ALG10, alpha-1,2-glucosyltransferase	protein-coding
ROBO2	0.00664663829656909	3.01	chr3	76679709	76679875	-360351	NM_002942	roundabout guidance receptor 2	protein-coding
HOMER1	0.00664910795870179	3.83	chr5	79537477	79537654	-23729	NM_001277078	homer scaffolding protein 1	protein-coding
UBE2NL	0.00668116134180352	4.49	chrX	143977304	143977591	93376	NM_001012989	ubiquitin conjugating enzyme E2 N like (gene/pseudogene)	protein-coding
KIAA1324L	0.0067003809799743	3.97	chr7	87013969	87014165	26167	NM_001291991	KIAA1324 like	protein-coding
PPT1	0.00677975379910807	2.64	chr1	40070507	40070787	26823	NM_001142604	palmitoyl-protein thioesterase 1	protein-coding
LINC01121	0.00678978503424825	-3.65	chr2	45373784	45373989	-118945	NR_033831	long intergenic non-protein coding RNA 1121	ncRNA
RPS6KC1	0.00679775466957006	3.48	chr1	213455549	213455767	404425	NM_001287220	ribosomal protein S6 kinase C1	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
VAPA	0.00682176296180144	3.21	chr18	10123467	10123618	209584	NM_003574	VAMP associated protein A	protein-coding
POU4F3	0.00696737070778689	3.71	chr5	146284123	146284329	-54798	NM_002700	POU class 4 homeobox 3	protein-coding
MIR1260B	0.00698905903946083	2.88	chr11	96180262	96180468	-161073	NR_036125	microRNA 1260b	ncRNA
RAB3C	0.00701201226293203	3.15	chr5	58772141	58772296	189174	NM_138453	RAB3C, member RAS oncogene family	protein-coding
SLCO4C1	0.00705555189721742	2.51	chr5	101962529	101962689	333940	NM_180991	solute carrier organic anion transporter family member 4C1	protein-coding
LINC00348	0.00713232522999267	3.44	chr13	70623491	70623688	-391552	NR_047699	long intergenic non-protein coding RNA 348	ncRNA
TMEM120B	0.00721428866194988	2.23	chr12	121748459	121748625	35790	NM_001080825	transmembrane protein 120B	protein-coding
LINC01235	0.00725597228519589	3.57	chr9	13427756	13427920	3491	NR_033863	long intergenic non-protein coding RNA 1235	ncRNA
CARMIL1	0.00728582112041433	2.98	chr6	25327046	25327281	47735	NM_017640	capping protein regulator and myosin 1 linker 1	protein-coding
ISY1-RAB43	0.00728763494786738	-3.68	chr3	129171423	129171592	-10277	NM_001204890	ISY1-RAB43 readthrough	protein-coding
ATF3	0.00742319930061371	3.68	chr1	212589319	212589473	-19232	NM_001674	activating transcription factor 3	protein-coding
TUNAR	0.00750866743227924	3.14	chr14	95953374	95953599	76714	NR_132398	TCL1 upstream neural differentiation-associated RNA	ncRNA
WDR18	0.00751293275858178	2.44	chr19	964266	964427	-19982	NM_024100	WD repeat domain 18	protein-coding
LINC01376	0.00754486274043717	2.04	chr2	19163989	19164311	-137179	NR_135287	long intergenic non-protein coding RNA 1376	ncRNA
LRRIQ1	0.00757102945694449	2.22	chr12	85089676	85090052	53543	NM_001079910	leucine rich repeats and IQ motif containing 1	protein-coding
PRG1	0.00758880131458023	-3.89	chr19	43330122	43330338	-18826	NR_026881	p53-responsive gene 1	ncRNA
KCNA4	0.00761606526221012	4.3	chr11	30049178	30049394	-32256	NM_002233	potassium voltage-gated channel subfamily A member 4	protein-coding
DLGAP1-AS5	0.00762800393418065	3.86	chr18	4199173	4199340	-65346	NR_036489	DLGAP1 antisense RNA 5	ncRNA
NEPRO	0.00767096551845613	2.29	chr3	113072699	113073038	-53135	NR_134969	nucleolus and neural progenitor protein	protein-coding
LINC00330	0.00770192110896021	2.68	chr13	44735030	44735196	74517	NR_038433	long intergenic non-protein coding RNA 330	ncRNA
CNOT9	0.00774886521241172	3.43	chr2	218581456	218581616	12956	NM_001271634	CCR4-NOT transcription complex subunit 9	protein-coding
MIR3936	0.00778751500095185	-3.29	chr5	132363371	132363545	2141	NR_037500	microRNA 3936	ncRNA
ATOH1	0.00780025918472925	4.3	chr4	93670448	93670627	-158390	NM_005172	atonal bHLH transcription factor 1	protein-coding
LINC01242	0.00780891965288126	1.93	chr9	29816612	29816999	591649	NR_046204	long intergenic non-protein coding RNA 1242	ncRNA
PREP	0.00781774596484462	2.8	chr6	105379676	105379865	23354	NM_002726	prolyl endopeptidase	protein-coding
POTEA	0.00782135738649704	2.26	chr8	43237756	43238000	-54564	NM_001005365	POTE ankyrin domain family member A	protein-coding
TRY2P	0.00783168335384426	2.87	chr7	142440324	142440609	-168217	NR_036483	trypsinogen-like pseudogene	pseudo
API5	0.00785185688195639	2.94	chr11	43306108	43306260	-5771	NM_006595	apoptosis inhibitor 5	protein-coding
MIR4666B	0.00787526106557243	2.4	chrX	29478440	29479069	-95524	NR_049877	microRNA 4666b	ncRNA
GRIK2	0.00798251521003849	3.83	chr6	102619850	102620024	1220952	NM_021956	glutamate ionotropic receptor kainate type subunit 2	protein-coding
ZMAT4	0.00807328739445244	3.81	chr8	40738368	40738575	159353	NM_024645	zinc finger matrin-type 4	protein-coding
FGF14	0.00813517325281965	2.37	chr13	102198841	102199131	111700	NM_001321933	fibroblast growth factor 14	protein-coding
ZNF716	0.00813744870367916	2.68	chr7	58652476	58652674	1202398	NM_001159279	zinc finger protein 716	protein-coding
ALG10	0.00814735994934152	2.29	chr12	35175613	35175823	1153437	NM_032834	ALG10, alpha-1,2-glucosyltransferase	protein-coding
PDE5A	0.00825047784037554	3.12	chr4	119647478	119647723	-18774	NM_001083	phosphodiesterase 5A	protein-coding
RIMS2	0.00826194538094081	3.63	chr8	103787201	103787350	-31913	NM_001282881	regulating synaptic membrane exocytosis 2	protein-coding
RHOBTB2	0.00826770304430018	2.06	chr8	22973158	22973863	-13907	NM_001160036	Rho related BTB domain containing 2	protein-coding
LINC00710	0.00827041887632224	2.23	chr10	10924879	10925388	27030	NR_024410	long intergenic non-protein coding RNA 710	ncRNA
FSHB	0.00827555967652225	1.92	chr11	30196594	30197012	-34213	NM_001018080	follicle stimulating hormone beta subunit	protein-coding
LINC00243	0.00840293927033215	2.4	chr6	30815784	30816171	14682	NR_130726	long intergenic non-protein coding RNA 243	ncRNA
CXCL1	0.00845217340810244	3.1	chr4	73891927	73892098	22620	NM_001511	C-X-C motif chemokine ligand 1	protein-coding
SLC16A7	0.00851017939098013	3.31	chr12	59796960	59797205	107745	NR_073055	solute carrier family 16 member 7	protein-coding
FUT1	0.00851154638957112	1.96	chr19	48752546	48752939	918	NM_001329877	fucosyltransferase 1 (H blood group)	protein-coding
CLIP4	0.00855148367969949	2.56	chr2	29116297	29116461	953	NM_001287527	CAP-Gly domain containing linker protein family member 4	protein-coding
GRIA2	0.00855654484332035	3.14	chr4	157187652	157187801	-32858	NM_001083619	glutamate ionotropic receptor AMPA type subunit 2	protein-coding
LRRTM3	0.00856582630280095	2.76	chr10	66281259	66281510	-644650	NM_178011	leucine rich repeat transmembrane neuronal 3	protein-coding
MIR1913	0.00860186150978845	1.9	chr6	166514905	166515180	-5609	NR_031734	microRNA 1913	ncRNA
REG3A	0.00860283184091844	4.02	chr2	79216728	79216890	-57055	NM_138937	regenerating family member 3 alpha	protein-coding
BMPR1B	0.00861848480737424	4.29	chr4	95156392	95156604	65007	NM_001256794	bone morphogenetic protein receptor type 1B	protein-coding
PHLDB2	0.00873448231382883	2.22	chr3	111803020	111803415	-55963	NM_145753	pleckstrin homology like domain family B member 2	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
PEG10	0.00876885622624571	1.88	chr7	94697280	94697790	41210	NM_001184962	paternally expressed 10	protein-coding
MIR548AA1	0.00883630325866425	2.53	chr8	123315338	123315645	-32543	NR_037516	microRNA 548aa-1	ncRNA
PSAT1	0.00883859451425298	2.51	chr9	78410462	78410693	113502	NM_058179	phosphoserine aminotransferase 1	protein-coding
CCNI2	0.00907451199637456	1.86	chr5	132747151	132747825	43	NM_001287253	cyclin I family member 2	protein-coding
HS3ST1	0.00913798948159257	3.57	chr4	11582381	11582543	-153549	NM_005114	heparan sulfate-glucosamine 3-sulfotransferase 1	protein-coding
CAMK2D	0.00917122817676049	3.07	chr4	113511902	113512160	249896	NM_001321567	calcium/calmodulin dependent protein kinase II delta	protein-coding
URB2	0.00924145826128904	2.87	chr1	229802747	229802924	176619	NM_014777	URB2 ribosome biogenesis 2 homolog (S. cerevisiae)	protein-coding
LOC105377267	0.00928409106009806	3.19	chr4	69153093	69153284	-28912	NR_136191	uncharacterized LOC105377267	ncRNA
NRXN1	0.00933259423188358	3.57	chr2	50442057	50442224	-94384	NM_138735	neurexin 1	protein-coding
LOC100996724	0.00939199306872944	2.07	chr1	120657897	120658228	7374	NR_144517	phosphodiesterase 4D interacting protein-like	pseudo
C9orf72	0.00939617829609683	3.58	chr9	28082613	28082795	-508838	NM_145005	chromosome 9 open reading frame 72	protein-coding
LOC101929284	0.00944729863500154	3.09	chr5	8846986	8847185	7353	NR_109932	uncharacterized LOC101929284	ncRNA
SNORA98	0.00946128984038553	2.78	chr6	144989082	144989578	583675	NR_132778	small nucleolar RNA, H/ACA box 98	snoRNA
LOC101927188	0.00958403175164835	2.85	chr18	6994903	6995070	40309	NR_126040	uncharacterized LOC101927188	ncRNA
MIR4431	0.00960218522565381	3.79	chr2	52648359	52648506	54183	NR_039630	microRNA 4431	ncRNA
RHOA	0.00965430649984243	2.86	chr3	49389876	49390072	22123	NM_001313947	ras homolog family member A	protein-coding
RBM19	0.00969735539643447	2.63	chr12	114045490	114045684	-79216	NM_001146698	RNA binding motif protein 19	protein-coding
PRKDC	0.00969950391630461	3.11	chr8	47855887	47856042	104219	NM_006904	protein kinase, DNA-activated, catalytic polypeptide	protein-coding
L3MBTL3	0.00975392208189243	3.32	chr6	130037399	130037608	17859	NM_001346550	l(3)mbt-like 3 (Drosophila)	protein-coding
PLBD1	0.00975816197095886	-2.96	chr12	14619684	14620046	-52008	NM_024829	phospholipase B domain containing 1	protein-coding
GDNF-AS1	0.00981752330797934	-2.79	chr5	37892092	37892267	18754	NM_001278094	GDNF antisense RNA 1 (head to head)	protein-coding
VRK2	0.00983379569707591	2.17	chr2	57217722	57217893	-689844	NM_001288837	vaccinia related kinase 2	protein-coding
HEY2	0.00985387978812787	4.43	chr6	125737134	125737283	-12378	NM_012259	hes related family bHLH transcription factor with YRPW motif 2	protein-coding
OXGR1	0.00985569789549428	3.38	chr13	97051524	97051740	-56902	NM_001346195	oxoglutarate receptor 1	protein-coding
LMO7DN-IT1	0.00986381215595718	3.11	chr13	76189476	76189634	312669	NR_126373	LMO7DN intronic transcript 1	ncRNA
ZNF544	0.00988721163151098	2.83	chr19	58248691	58248919	19891	NM_001320781	zinc finger protein 544	protein-coding
TPTE	0.00990218139838507	2.18	chr21	10746765	10747088	225412	NM_199260	transmembrane phosphatase with tensin homology	protein-coding
SUCLG2	0.00990621887415555	2.92	chr3	67569438	67569606	85092	NM_001177599	succinate-CoA ligase GDP-forming beta subunit	protein-coding
MRPL46	0.00991020116087559	3.33	chr15	88437489	88437642	29837	NM_022163	mitochondrial ribosomal protein L46	protein-coding
SPTSSB	0.00991649200440139	2.44	chr3	161319376	161319748	52521	NM_001040100	serine palmitoyltransferase small subunit B	protein-coding
ANKRD20A4	0.00996605173523842	2.74	chr9	64316561	64316722	-52922	NM_001098805	ankyrin repeat domain 20 family member A4	protein-coding
FAT4	0.0100737349005398	2.98	chr4	125215762	125215940	-100561	NM_001291303	FAT atypical cadherin 4	protein-coding
FAM230C	0.0102952717229043	2.07	chr13	16984974	16985324	-1210148	NR_027278	family with sequence similarity 230 member C	ncRNA
LINC00882	0.0103926788565175	4.07	chr3	106965994	106966196	274543	NR_028303	long intergenic non-protein coding RNA 882	ncRNA
OR4F5	0.0104433400274461	4.32	chr1	67945	68099	-1069	NM_001005484	olfactory receptor family 4 subfamily F member 5	protein-coding
LINC00163	0.0104752151361273	2.12	chr21	44981829	44982244	12050	NR_033840	long intergenic non-protein coding RNA 163	ncRNA
LOC100130331	0.0104908594564969	2.52	chr1	237857419	237857609	-4661	NR_027247	POTE ankyrin domain family, member F pseudogene	pseudo
OR4E1	0.010539116489533	-3.55	chr14	22025150	22025353	-354234	NM_001317107	olfactory receptor family 4 subfamily E member 1 (gene/pseudogene)	protein-coding
MYOM3	0.010607589951212	-2.09	chr1	24070966	24071220	41082	NM_152372	myomesin 3	protein-coding
LINC00967	0.0106613183917577	-3.05	chr8	66212551	66212730	20526	NR_039979	long intergenic non-protein coding RNA 967	ncRNA
POMZP3	0.010683960371534	3.38	chr7	76638453	76638606	-11226	NM_012230	POM121 and ZP3 fusion	protein-coding
CD99	0.0107410833387331	3.2	chrX	2707303	2707479	16258	NM_002414	CD99 molecule	protein-coding
LINC01207	0.0107815529663775	2.35	chr4	164623356	164623539	-130684	NR_038834	long intergenic non-protein coding RNA 1207	ncRNA
GMPS	0.0108581435361011	3.35	chr3	155885872	155886033	15416	NM_003875	guanine monophosphate synthase	protein-coding
C15orf32	0.0109158948782249	3.72	chr15	92510665	92510829	39070	NM_001301106	chromosome 15 open reading frame 32	protein-coding
RCE1	0.0109284706405775	-1.91	chr11	66834379	66834715	-8865	NM_005133	Ras converting CAAX endopeptidase 1	protein-coding
PICK1	0.0109336516332475	2.04	chr22	38024730	38024924	-32428	NM_001039583	protein interacting with PRKCA 1	protein-coding
MIR3974	0.0109809511018078	2.99	chr12	17396990	17397145	-276232	NR_039770	microRNA 3974	ncRNA
DAP	0.0109839680795686	2.71	chr5	10839117	10839276	-77921	NM_001291963	death associated protein	protein-coding
LOC102467226	0.011016994598997	3.65	chr5	121285844	121286087	-36585	NR_104999	uncharacterized LOC102467226	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
ANKRD36C	0.0110767236363352	2.03	chr2	95950613	95950780	41624	NM_001310154	ankyrin repeat domain 36C	protein-coding
SPOCK1	0.0110777446450977	2.38	chr5	137529412	137529576	-30165	NM_004598	SPARC/osteonectin, cwcv and kazal like domains proteoglycan 1	protein-coding
PROSER2	0.0111342195617935	-3.72	chr10	11799417	11799600	-23890	NM_153256	proline and serine rich 2	protein-coding
RAB37	0.0111492505215335	3.01	chr17	74669494	74669711	-1515	NM_175738	RAB37, member RAS oncogene family	protein-coding
RAD51AP2	0.0113361283092885	3.25	chr2	17193917	17194090	324436	NM_001099218	RAD51 associated protein 2	protein-coding
GPT2	0.0113812372746996	2.14	chr16	46899799	46900032	14801	NM_001142466	glutamic--pyruvic transaminase 2	protein-coding
TMEM184C	0.0113952299190535	2.54	chr4	147589799	147590508	-27235	NM_018241	transmembrane protein 184C	protein-coding
GRIFIN	0.0114085092401169	1.81	chr7	2459077	2459474	17108	NM_001291784	galectin-related inter-fiber protein	protein-coding
CCNT2	0.0114266952378582	2.99	chr2	134964088	134964250	45347	NM_001320748	cyclin T2	protein-coding
LINC00290	0.0115068582606414	2.23	chr4	181512040	181512208	-352975	NR_033918	long intergenic non-protein coding RNA 290	ncRNA
LINC01103	0.0115203801482944	2.45	chr2	104549836	104550078	-39448	NR_110113	long intergenic non-protein coding RNA 1103	ncRNA
LINC00898	0.0115322127119181	2.79	chr22	47587127	47587325	44343	NR_033377	long intergenic non-protein coding RNA 898	ncRNA
TESMIN	0.0115659555668987	2.44	chr11	68748956	68749187	2449	NM_001039656	testis expressed metallothionein like protein	protein-coding
ELAVL2	0.0116010496516606	-3.43	chr9	24003836	24004056	-177881	NM_004432	ELAV like RNA binding protein 2	protein-coding
PTX3	0.0116026939501793	1.82	chr3	157437475	157438472	1182	NM_002852	pentraxin 3	protein-coding
TPO	0.0116140792902528	1.89	chr2	1344860	1345124	-68469	NM_001206745	thyroid peroxidase	protein-coding
LOC100287632	0.0116314898759006	3.49	chr6	118801661	118801827	19036	NR_037662	selenoprotein K pseudogene	pseudo
PTCHD4	0.0117174158843554	3.53	chr6	48026062	48026216	42550	NM_001013732	patched domain containing 4	protein-coding
LOC101927243	0.0117684127746113	2.39	chr7	77316695	77316844	-98429	NR_110073	uncharacterized LOC101927243	ncRNA
PPM1M	0.0118018615832026	-3.06	chr3	52255860	52256057	9788	NM_001122870	protein phosphatase, Mg2+/Mn2+ dependent 1M	protein-coding
MIR4459	0.0118467445184851	2.62	chr6	132931912	132932109	-17682	NR_039664	microRNA 4459	ncRNA
LOC153910	0.0118581779454162	2.34	chr6	142506091	142506248	131667	NR_027311	uncharacterized LOC153910	ncRNA
LOC101927588	0.0118803080703814	2.34	chr8	124228820	124229078	18449	NR_125420	uncharacterized LOC101927588	ncRNA
SERPINA7	0.0118841650619708	1.89	chrX	106007876	106008448	30565	NM_000354	serpin family A member 7	protein-coding
PLCL2	0.0119123430190506	3.38	chr3	16747492	16747943	-137242	NM_001144382	phospholipase C like 2	protein-coding
LINC01039	0.0119162583659758	2.99	chr13	99565766	99565951	-11254	NR_126390	long intergenic non-protein coding RNA 1039	ncRNA
KCNQ1	0.011969287366956	-2.22	chr11	2471785	2472050	10463	NM_181798	potassium voltage-gated channel subfamily Q member 1	protein-coding
LOC105369187	0.0119699018788164	3.02	chr12	65487870	65488120	69316	NR_135033	uncharacterized LOC105369187	ncRNA
ACAT1	0.0119838181705005	3.85	chr11	108147653	108147820	26205	NM_000019	acetyl-CoA acetyltransferase 1	protein-coding
LOC105378146	0.0120111109734921	2.59	chr6	168742903	168743129	221366	NR_136250	uncharacterized LOC105378146	ncRNA
TYR	0.0120471376523202	2.83	chr11	89328076	89328263	150297	NM_000372	tyrosinase	protein-coding
KIF5C	0.0120780020000738	2.57	chr2	148852739	148852946	-22381	NM_004522	kinesin family member 5C	protein-coding
LOC105376430	0.0121573515706288	1.89	chr10	14991476	14991824	19086	NR_134493	uncharacterized LOC105376430	ncRNA
NOBOX	0.0121837029735167	2.82	chr7	144476200	144476456	-66101	NM_001080413	NOBOX oogenesis homeobox	protein-coding
SMG1P2	0.0121986902855629	4.01	chr16	29553648	29553813	41362	NR_135305	SMG1P2, nonsense mediated mRNA decay associated PI3K related kinase pseudogene 2	pseudo
SSBP2	0.0122058851803051	3.87	chr5	81820487	81820793	-69387	NM_001256734	single stranded DNA binding protein 2	protein-coding
SPATA4	0.0122223232382856	1.98	chr4	176225960	176226184	-30401	NM_144644	spermatogenesis associated 4	protein-coding
CNTNAP5	0.0122479440699126	2.72	chr2	124002639	124002909	-22513	NM_130773	contactin associated protein like 5	protein-coding
TENM4	0.0122497862737762	2.83	chr11	80014859	80015089	-574323	NM_001098816	teneurin transmembrane protein 4	protein-coding
SREBF2	0.0123911724244821	3.02	chr22	41859480	41859697	26509	NM_004599	sterol regulatory element binding transcription factor 2	protein-coding
IRX1	0.0124120416917464	2.35	chr5	4080510	4080683	484542	NM_024337	iroquois homeobox 1	protein-coding
MIR548BA	0.0125166281235182	2.47	chr2	48983944	48984799	-75232	NR_107011	microRNA 548ba	ncRNA
MSL3	0.0125862137289724	4.22	chrX	11787771	11787932	28235	NM_006800	male-specific lethal 3 homolog (Drosophila)	protein-coding
LNX1	0.0125987310343358	2.51	chr4	53576885	53577397	14445	NM_001126328	ligand of numb-protein X 1	protein-coding
SNX16	0.0126011836062675	3.07	chr8	81996728	81997033	-154594	NM_022133	sorting nexin 16	protein-coding
GATSL2	0.0126878681656231	2.39	chr7	75007824	75007972	43080	NM_001145064	GATS protein like 2	protein-coding
LINC01435	0.0127413865651722	2.08	chr10	108426994	108427328	-357868	NR_125760	long intergenic non-protein coding RNA 1435	ncRNA
LINC01085	0.0127906662252034	2.93	chr4	14369384	14369534	257491	NR_033931	long intergenic non-protein coding RNA 1085	ncRNA
EFHB	0.0128428624626842	2.21	chr3	19899834	19900049	34273	NM_144715	EF-hand domain family member B	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
MEIS2	0.0128764212745164	3.94	chr15	37133878	37134051	-32665	NM_001220482	Meis homeobox 2	protein-coding
NRIP1	0.0129550398974566	-3.79	chr21	15102663	15102840	-37946	NM_003489	nuclear receptor interacting protein 1	protein-coding
ZFYVE28	0.0130035477887253	1.84	chr4	2320608	2321284	43931	NM_001172660	zinc finger FYVE-type containing 28	protein-coding
LOC643711	0.0131598002035339	3.3	chr12	97671649	97671821	81863	NR_077241	platelet activating factor acetylhydrolase 1b catalytic subunit 2 pseudogene	pseudo
ASZ1	0.0131715402638271	2.36	chr7	117393852	117394211	33492	NM_001301822	ankyrin repeat, SAM and basic leucine zipper domain containing 1	protein-coding
DIO3OS	0.013173446098597	2.18	chr14	101499123	101499300	56465	NR_002770	DIO3 opposite strand/antisense RNA (head to head)	ncRNA
FCRL4	0.0133287980854694	-3.63	chr1	157636749	157636911	-38750	NM_031282	Fc receptor like 4	protein-coding
EMB	0.0133419684096985	2.05	chr5	49786444	49786612	654872	NM_198449	embigin	protein-coding
SCG5	0.0133627903084562	1.91	chr15	32649393	32649628	7897	NM_003020	secretogranin V	protein-coding
BCKDHB	0.0133904541092861	3.96	chr6	80728642	80728822	622122	NR_134945	branched chain keto acid dehydrogenase E1 subunit beta	protein-coding
IQCA1	0.0134395971609088	2.33	chr2	236444645	236444953	58898	NM_001270585	IQ motif containing with AAA domain 1	protein-coding
USP22	0.013452691643882	2.09	chr17	20998042	20998225	44906	NM_015276	ubiquitin specific peptidase 22	protein-coding
LOC101929243	0.01347538732145	-3.41	chr6	34639300	34639464	-56772	NR_134625	uncharacterized LOC101929243	ncRNA
KCNN3	0.013481109575601	2.09	chr1	154751754	154751981	107995	NM_170782	potassium calcium-activated channel subfamily N member 3	protein-coding
FBXO22	0.0135426693765801	-2.07	chr15	75885399	75886172	-18074	NM_147188	F-box protein 22	protein-coding
LRRC4C	0.013550331608806	2.31	chr11	41431707	41431927	27819	NM_020929	leucine rich repeat containing 4C	protein-coding
LOC101927557	0.0136076486398652	2.28	chr17	57406794	57406947	116727	NR_110808	uncharacterized LOC101927557	ncRNA
INTU	0.0136944437115804	2.81	chr4	127163798	127163962	-469052	NM_015693	inturned planar cell polarity protein	protein-coding
AP1S3	0.0138422214184942	2.76	chr2	223889205	223889360	-51680	NM_001039569	adaptor related protein complex 1 sigma 3 subunit	protein-coding
C3orf67-AS1	0.0138545624106195	2.47	chr3	58794284	58794874	-29892	NR_110820	C3orf67 antisense RNA 1	ncRNA
COBL	0.0139231878428401	1.97	chr7	51806405	51806803	-489743	NM_001346442	cordon-bleu WH2 repeat protein	protein-coding
PCDHA4	0.0140092538541238	2.17	chr5	140808635	140808888	1687	NM_031500	protocadherin alpha 4	protein-coding
LOC105375972	0.0141104136471878	2.48	chr9	9750736	9750916	-48566	NR_135135	uncharacterized LOC105375972	ncRNA
REEP6	0.0141143791347379	-1.87	chr19	1501051	1501581	10311	NM_001329556	receptor accessory protein 6	protein-coding
MIR4465	0.0143324956204818	2.85	chr6	141260684	141260836	576946	NR_039675	microRNA 4465	ncRNA
KCNJ3	0.0143661683177186	2.08	chr2	154587791	154588025	-110673	NM_001260508	potassium voltage-gated channel subfamily J member 3	protein-coding
MIR548AU	0.0144841455826818	-3.54	chr2	107419865	107420030	-70422	NR_049853	microRNA 548au	ncRNA
CLDN5	0.0145302280732875	-2.86	chr22	19506403	19506598	18837	NM_003277	claudin 5	protein-coding
NEB	0.0145495989272936	2.71	chr2	151633218	151633393	101182	NM_001164508	nebulin	protein-coding
SMARCA2	0.0146199261272132	4.07	chr9	1605221	1605415	-409901	NM_139045	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	protein-coding
VIPR2	0.0146585081328304	2.3	chr7	159171598	159172146	-26915	NR_130758	vasoactive intestinal peptide receptor 2	protein-coding
CD180	0.0146725872079265	2.14	chr5	67102662	67103184	93866	NM_005582	CD180 molecule	protein-coding
ATP5G3	0.014719084795277	2.85	chr2	175377842	175378198	-196258	NM_001689	ATP synthase, H+ transporting, mitochondrial Fo complex subunit C3 (subunit 9)	protein-coding
TICAM1	0.01472662754809	-2.52	chr19	4818017	4818286	13591	NM_182919	toll like receptor adaptor molecule 1	protein-coding
GIMD1	0.0148034054035991	2.56	chr4	106681416	106681613	-314079	NM_001195138	GIMAP family P-loop NTPase domain containing 1	protein-coding
EDN2	0.0148112104266554	3.21	chr1	41689788	41689954	-205188	NM_001302269	endothelin 2	protein-coding
ZNF733P	0.0148589149468242	2.01	chr7	63113967	63114416	189865	NR_003952	zinc finger protein 733, pseudogene	pseudo
ACTG1P4	0.0149633749013191	1.97	chr1	103569823	103570114	564	NR_024438	actin gamma 1 pseudogene 4	pseudo
DCDC2C	0.0149730659836559	2.33	chr2	3703789	3703936	270	NM_001287444	doublecortin domain containing 2C	protein-coding
MME	0.0149859515047458	2.54	chr3	154780515	154780690	-299045	NM_000902	membrane metalloendopeptidase	protein-coding
ROBO2	0.0150699520740899	2.22	chr3	77039632	77039922	-366	NM_002942	roundabout guidance receptor 2	protein-coding
PACRGL	0.0151456164443736	-2.64	chr4	20773349	20773603	73063	NM_001330748	PARK2 coregulated like	protein-coding
TNFAIP8L3	0.0152067102520883	2.63	chr15	51157222	51157390	-52030	NM_207381	TNF alpha induced protein 8 like 3	protein-coding
CDH9	0.015268894250784	3	chr5	27106507	27106685	-68014	NM_016279	cadherin 9	protein-coding
RAP2A	0.0152712611198173	3.71	chr13	97560029	97560193	125890	NM_021033	RAP2A, member of RAS oncogene family	protein-coding
TRAM1L1	0.0152951188308345	2.66	chr4	117134440	117134670	-48975	NM_152402	translocation associated membrane protein 1-like 1	protein-coding
ANGPT1	0.0153348326185444	3.06	chr8	107356804	107356967	-20370	NM_001314051	angiopoietin 1	protein-coding
LOC644919	0.0153637479338939	3.22	chr14	40409228	40409409	-545393	NR_109758	uncharacterized LOC644919	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
NIPBL-AS1	0.0153875423175057	4.14	chr5	36806848	36807007	69767	NR_046262	NIPBL antisense RNA 1 (head to head)	ncRNA
PTPRD	0.0153945719562654	-3.46	chr9	10560203	10560417	52413	NM_002839	protein tyrosine phosphatase, receptor type D	protein-coding
IPMK	0.0154643540888268	3.91	chr10	57916969	57917142	350879	NM_152230	inositol polyphosphate multikinase	protein-coding
LOC101928137	0.0154795231671043	2.06	chr12	72740631	72740793	-418478	NR_110130	uncharacterized LOC101928137	ncRNA
BNC2	0.0154897983003303	2.83	chr9	16653510	16653666	217200	NM_017637	basonuclin 2	protein-coding
ROCK1P1	0.0155669280614306	2.24	chr18	96777	96994	-12180	NR_033770	Rho associated coiled-coil containing protein kinase 1 pseudogene 1	pseudo
SCML4	0.0156158589903548	2.83	chr6	107785351	107785511	38886	NM_198081	sex comb on midleg-like 4 (Drosophila)	protein-coding
RAB35	0.0156439881955273	2.64	chr12	120099348	120099524	17403	NM_001167606	RAB35, member RAS oncogene family	protein-coding
LINC00366	0.0156538880817587	1.81	chr13	38582389	38582915	14928	NR_046999	long intergenic non-protein coding RNA 366	ncRNA
RAB28	0.0156563791184451	2.97	chr4	13340397	13340593	143870	NM_004249	RAB28, member RAS oncogene family	protein-coding
CKMT2-AS1	0.0156872930725125	3.66	chr5	81291740	81291921	9739	NR_034122	CKMT2 antisense RNA 1	ncRNA
ARIH2OS	0.0157214074011682	2.35	chr3	48952775	48953147	-33576	NM_001123040	ariadne homolog 2 opposite strand	protein-coding
ANO3	0.0157505095773736	3.53	chr11	25573616	25573811	-615410	NM_001313726	anoctamin 3	protein-coding
LRRTM3	0.015766795002816	-2.99	chr10	67173800	67174065	161666	NR_111909	leucine rich repeat transmembrane neuronal 3	protein-coding
EMBP1	0.0158651821847064	2.14	chr1	123868403	123868569	2349374	NR_003955	embigin pseudogene 1	pseudo
LINC01411	0.0159769026411449	2.21	chr5	174327517	174327723	-8734	NR_125806	long intergenic non-protein coding RNA 1411	ncRNA
PPP1R1C	0.0159781031780589	-2.23	chr2	182050359	182050677	64694	NR_048566	protein phosphatase 1 regulatory inhibitor subunit 1C	protein-coding
LINC02015	0.015982495715654	1.89	chr3	177917647	177917873	100895	NR_110826	long intergenic non-protein coding RNA 2015	ncRNA
DCD	0.0159952903560306	-2.53	chr12	54688689	54688864	-40283	NM_053283	dermcidin	protein-coding
C19orf35	0.0160087434947696	2.42	chr19	2282330	2282479	-222	NM_198532	chromosome 19 open reading frame 35	protein-coding
KLF12	0.0160766513195272	3.29	chr13	73891717	73891913	242114	NM_007249	Kruppel like factor 12	protein-coding
LOC105378269	0.0160986535198164	2.18	chr10	42754766	42754918	-300	NR_134498	uncharacterized LOC105378269	ncRNA
SNAR-H	0.0161039193078709	3.02	chr2	77938105	77938262	16843	NR_024342	small ILF3/NF90-associated RNA H	snRNA
SCUBE3	0.0161426724371817	2.72	chr6	35216310	35216476	2331	NM_001303136	signal peptide, CUB domain and EGF like domain containing 3	protein-coding
SNORA107	0.0161543559400134	4.11	chr13	84659336	84659558	-124207	NR_132852	small nucleolar RNA, H/ACA box 107	snoRNA
LINC02114	0.0161733392165232	2.53	chr5	4523797	4523946	-249610	NR_104619	long intergenic non-protein coding RNA 2114	ncRNA
TFPI	0.0162017988338494	1.91	chr2	187658492	187659212	-104347	NM_001329240	tissue factor pathway inhibitor	protein-coding
STXBP6	0.0162233812091279	2.82	chr14	25210094	25210311	-159905	NM_001304476	syntaxin binding protein 6	protein-coding
FRG1CP	0.0162526268348871	1.98	chr20	27436652	27436935	1165872	NR_132315	FSHD region gene 1 family member C, pseudogene	pseudo
GBA3	0.0162677552661661	2.41	chr4	23237764	23238149	523867	NM_001277225	glucosylceramidase beta 3 (gene/pseudogene)	protein-coding
RGS18	0.0162752851466988	4.07	chr1	191873018	191873203	-285352	NM_130782	regulator of G-protein signaling 18	protein-coding
PEX14	0.0162817545719277	-2.34	chr1	10519069	10519595	44386	NM_004565	peroxisomal biogenesis factor 14	protein-coding
SREK1	0.0163911641521623	2.65	chr5	66320201	66320475	175503	NM_139168	splicing regulatory glutamic acid and lysine rich protein 1	protein-coding
DMRTA1	0.0164139381434701	3.35	chr9	22529332	22529526	82588	NM_022160	DMRT like family A1	protein-coding
GGT7	0.0164288567886603	3.27	chr20	34864903	34865058	7878	NM_178026	gamma-glutamyltransferase 7	protein-coding
PLCE1-AS2	0.0164336184813878	2.86	chr10	94188490	94188650	-79776	NR_120616	PLCE1 antisense RNA 2	ncRNA
CCDC167	0.0164596986920384	2.03	chr6	37491409	37491583	8428	NM_138493	coiled-coil domain containing 167	protein-coding
MYO16-AS1	0.0164886461467268	2.08	chr13	108944809	108945023	222387	NR_047700	MYO16 antisense RNA 1	ncRNA
UHRF1	0.0165083138605352	-2.46	chr19	4904634	4904793	1633	NM_001290050	ubiquitin like with PHD and ring finger domains 1	protein-coding
SLC1A4	0.0165487521941551	-3.3	chr2	64978574	64978730	-9793	NM_001193493	solute carrier family 1 member 4	protein-coding
ATOH1	0.0165701949085978	3.27	chr4	93569365	93569515	-259487	NM_005172	atonal bHLH transcription factor 1	protein-coding
ZNF733P	0.0165842465460026	2.97	chr7	62254886	62255066	1049080	NR_003952	zinc finger protein 733, pseudogene	pseudo
KL	0.016631573839323	2.07	chr13	32969613	32969876	-46689	NM_004795	klotho	protein-coding
MPO	0.0166356798261378	2.52	chr17	58300505	58300709	-19672	NM_000250	myeloperoxidase	protein-coding
CTTNBP2NL	0.0166644013040578	2.28	chr1	112329318	112329482	-66778	NM_018704	CTTNBP2 N-terminal like	protein-coding
KDM4A	0.0167397535797264	1.85	chr1	43671362	43671549	21329	NM_014663	lysine demethylase 4A	protein-coding
LINC00680	0.0167699021823096	2.97	chr6	58839238	58839463	-877904	NR_125728	long intergenic non-protein coding RNA 680	ncRNA
ZNF407	0.0168282428340268	2.34	chr18	74657592	74657751	26708	NM_017757	zinc finger protein 407	protein-coding
PSMD6-AS2	0.016847882058369	-2.14	chr3	64003731	64004177	-68	NR_038286	PSMD6 antisense RNA 2	ncRNA
LINC00578	0.0168794556282396	2.61	chr3	177516057	177516226	74220	NR_047568	long intergenic non-protein coding RNA 578	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
FAM13A-AS1	0.0169817078542833	2.97	chr4	88721189	88721354	11482	NR_002806	FAM13A antisense RNA 1	ncRNA
LRP1B	0.0170552233235751	1.83	chr2	141399654	141400210	731769	NM_018557	LDL receptor related protein 1B	protein-coding
LOC400940	0.0170787883083294	2.05	chr2	6037505	6037698	55623	NR_026833	uncharacterized LOC400940	ncRNA
TRIM42	0.0171846909532879	3.03	chr3	140699923	140700093	21984	NM_152616	tripartite motif containing 42	protein-coding
FAM27B	0.017200493666324	2.04	chr9	67761819	67762023	-35178	NR_027422	family with sequence similarity 27 member B	ncRNA
LINC01246	0.0172078215990125	3.16	chr2	6570894	6571052	62814	NR_110500	long intergenic non-protein coding RNA 1246	ncRNA
FXYD6	0.0172397950333121	1.86	chr11	117860063	117860295	16852	NM_001164837	FXYD domain containing ion transport regulator 6	protein-coding
NPTXR	0.0172576180552628	2.71	chr22	38844719	38844894	-794	NM_014293	neuronal pentraxin receptor	protein-coding
MIR1296	0.017276889628229	-1.93	chr10	63364804	63365179	8057	NR_031566	microRNA 1296	ncRNA
KLF9	0.0173456481492335	2.18	chr9	70398263	70398422	16315	NM_001206	Kruppel like factor 9	protein-coding
LOC100507468	0.017350980372392	3.23	chr7	69168706	69168874	428705	NR_108105	uncharacterized LOC100507468	ncRNA
LINC01760	0.0173525459885548	2.87	chr1	95286524	95286675	31664	NR_135588	long intergenic non-protein coding RNA 1760	ncRNA
AKAP7	0.017424709613035	3.63	chr6	131167775	131167953	32178	NM_016377	A-kinase anchoring protein 7	protein-coding
ZBTB16	0.0174341092965739	-2.03	chr11	114134102	114134296	73633	NM_001018011	zinc finger and BTB domain containing 16	protein-coding
DNAJC6	0.0176108705920257	3.14	chr1	65305578	65305770	-3861	NM_001256864	DnaJ heat shock protein family (Hsp40) member C6	protein-coding
RARB	0.0177177218386841	-2.74	chr3	25459210	25459587	1088	NM_001290217	retinoic acid receptor beta	protein-coding
HCG25	0.0178468506649152	2.73	chr6	33238603	33238823	-10823	NR_044997	HLA complex group 25 (non-protein coding)	ncRNA
LOC105369911	0.0178474367852255	1.89	chr12	93979156	93979402	-35673	NR_135017	uncharacterized LOC105369911	ncRNA
MAF	0.0178643282126535	3.24	chr16	79333056	79333229	267583	NM_001031804	MAF bZIP transcription factor	protein-coding
MIR4439	0.017917288951952	2.7	chr2	225006543	225006772	3883	NR_039641	microRNA 4439	ncRNA
REREP3	0.0179498952929988	-3.25	chr15	22307153	22307347	48636	NR_033735	arginine-glutamic acid dipeptide repeats pseudogene 3	pseudo
LINC00867	0.0179823058113519	2.8	chr10	118469996	118470170	112975	NR_108045	long intergenic non-protein coding RNA 867	ncRNA
UNC5C	0.0180047004607728	4.05	chr4	95637014	95637203	-87898	NM_003728	unc-5 netrin receptor C	protein-coding
ANAPC10	0.0180175751682002	-3.55	chr4	145039669	145039903	57596	NM_001318367	anaphase promoting complex subunit 10	protein-coding
GALNT2	0.0180278072166251	2.19	chr1	230151445	230151601	84328	NR_120373	polypeptide N-acetylgalactosaminyltransferase 2	protein-coding
MTHFD2P1	0.0180506659185481	1.98	chr3	95340601	95341717	342034	NR_077228	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase pseudogene 1	pseudo
DACH1	0.0180524452453267	2.75	chr13	71851929	71852082	15187	NM_004392	dachshund family transcription factor 1	protein-coding
EVL	0.0181083090432353	-2.25	chr14	99995744	99995976	24385	NM_001330221	Enah/Vasp-like	protein-coding
BBX	0.0182318070611486	1.83	chr3	107525735	107525900	2881	NM_001142568	BBX, HMG-box containing	protein-coding
LINC00371	0.0182648804141675	2.44	chr13	51148588	51148754	23551	NR_102432	long intergenic non-protein coding RNA 371	ncRNA
CD226	0.0182750557620528	3.08	chr18	70011414	70011613	-54337	NM_001303619	CD226 molecule	protein-coding
MIR4268	0.018350682235712	2.73	chr2	220056764	220056916	-150275	NR_036228	microRNA 4268	ncRNA
TPD52L3	0.0184647876178902	-2.16	chr9	6354082	6354289	25836	NM_001001875	tumor protein D52 like 3	protein-coding
RNF7	0.018481608980873	4.3	chr3	141750989	141751164	12867	NM_001201370	ring finger protein 7	protein-coding
FAM133CP	0.0185094413237429	3.12	chr10	58754528	58754764	39631	NR_027508	family with sequence similarity 133, member A pseudogene	pseudo
AKAIN1	0.0185400489693466	3.07	chr18	5133891	5134067	63277	NM_001145194	A-kinase anchor inhibitor 1	protein-coding
MROH2B	0.0185596344642345	2.03	chr5	41000103	41000291	71145	NM_173489	maestro heat like repeat family member 2B	protein-coding
TFG	0.0185671916883775	2.46	chr3	100718197	100718370	8741	NM_001195478	TRK-fused gene	protein-coding
REG3A	0.0185860560826103	2.8	chr2	79251283	79251453	-91614	NM_138937	regenerating family member 3 alpha	protein-coding
PI4KAP1	0.0185939726366249	-2.53	chr22	18549720	18550024	-1044	NR_003563	phosphatidylinositol 4-kinase alpha pseudogene 1	pseudo
LOC101928944	0.0186522995197796	3.57	chr11	80393980	80394172	368727	NR_120570	uncharacterized LOC101928944	ncRNA
PLOD2	0.0186758019750678	-3.03	chr3	145846101	145846360	315265	NM_182943	procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	protein-coding
ADGRL1	0.0186784800851357	2.3	chr19	14222945	14223345	-16960	NM_001008701	adhesion G protein-coupled receptor L1	protein-coding
DEPDC1	0.018705036900843	3.91	chr1	68876488	68876637	-379341	NM_017779	DEP domain containing 1	protein-coding
MBL2	0.0187684827719012	2.03	chr10	53071299	53071719	-299809	NM_000242	mannose binding lectin 2	protein-coding
ROPN1L	0.0187934114144943	-1.98	chr5	10452649	10453220	11072	NM_031916	rhophilin associated tail protein 1 like	protein-coding
MIR3976HG	0.0188570740470648	3.01	chr18	5747909	5748164	-783	NR_038839	MIR3976 host gene	ncRNA
DNAH11	0.0189896509347449	2.56	chr7	21733786	21734209	190782	NM_001277115	dynein axonemal heavy chain 11	protein-coding
TRAF6	0.0190346246337555	1.89	chr11	36502775	36503390	7231	NM_004620	TNF receptor associated factor 6	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
LOC100287944	0.0191426560230749	2.67	chr12	106862471	106862686	-87747	NR_040246	uncharacterized LOC100287944	ncRNA
MIR5007	0.0191835591748814	4.02	chr13	55483677	55483867	309318	NR_049804	microRNA 5007	ncRNA
PDE10A	0.0192150156589225	3.29	chr6	165636215	165636396	25795	NM_001130690	phosphodiesterase 10A	protein-coding
TRIQK	0.0192597515120407	2.63	chr8	92888114	92888304	77380	NM_001191036	triple QxxK/R motif containing	protein-coding
NOX4	0.0193102592282551	3.49	chr11	89619607	89619820	-30102	NM_001143837	NADPH oxidase 4	protein-coding
FAM46C	0.0193122946582259	2.12	chr1	117663587	117663838	57730	NM_017709	family with sequence similarity 46 member C	protein-coding
LINC01793	0.0193559379650981	3.88	chr2	59302841	59303020	85222	NR_110219	long intergenic non-protein coding RNA 1793	ncRNA
MTRNR2L3	0.0194838378213498	1.95	chr20	57358693	57359190	881	NM_001190472	MT-RNR2-like 3	protein-coding
LOC401134	0.019508798414092	3.24	chr4	65306097	65306261	-301679	NR_033976	uncharacterized LOC401134	ncRNA
BMP3	0.0195167173526292	2.04	chr4	81031272	81031603	472	NM_001201	bone morphogenetic protein 3	protein-coding
LINC00316	0.0195600300578456	1.82	chr21	45367410	45367573	-25501	NR_103811	long intergenic non-protein coding RNA 316	ncRNA
EXT1	0.0196313478268571	2.17	chr8	118075224	118075406	36504	NM_000127	exostosin glycosyltransferase 1	protein-coding
MIR1587	0.0196539941141185	-3.05	chrX	39936719	39937294	99445	NR_039763	microRNA 1587	ncRNA
PRICKLE2	0.0196541181354051	3.59	chr3	64215172	64215342	10198	NM_198859	prickle planar cell polarity protein 2	protein-coding
NNMT	0.0197004012949781	-2.48	chr11	114270678	114270916	-25016	NM_006169	nicotinamide N-methyltransferase	protein-coding
TMEM106B	0.0197161017656662	3.38	chr7	12031235	12031429	-179890	NM_001134232	transmembrane protein 106B	protein-coding
MAPKAP1	0.0197693985167084	3.59	chr9	125603862	125604063	103272	NM_024117	mitogen-activated protein kinase associated protein 1	protein-coding
MIR4426	0.0197823645227085	2.8	chr16	61276302	61276454	-220612	NR_039624	microRNA 4426	ncRNA
NBPF20	0.0198141112739942	2.15	chr1	145486872	145487042	-81179	NM_001278267	neuroblastoma breakpoint family member 20	protein-coding
CDRT15P1	0.0198152313685064	2.84	chr17	14011336	14011507	-13077	NR_003261	CMT1A duplicated region transcript 15 pseudogene 1	pseudo
MIR378H	0.0198269713329713	2.44	chr5	154827903	154828095	-1459	NR_039667	microRNA 378h	ncRNA
ADAT3	0.0199109203363673	-2.1	chr19	1897618	1898053	-7379	NM_138422	adenosine deaminase, tRNA specific 3	protein-coding
KLF5	0.0199137696939026	2.56	chr13	73194250	73194429	135547	NM_001730	Kruppel like factor 5	protein-coding
MIS18BP1	0.0199321009153657	2.67	chr14	45332578	45332734	-79254	NM_018353	MIS18 binding protein 1	protein-coding
ANXA5	0.0199501062339354	2.23	chr4	121718022	121718356	-21197	NM_001154	annexin A5	protein-coding
LINC01592	0.0199849573613766	2.93	chr8	68868292	68868479	235805	NR_039986	long intergenic non-protein coding RNA 1592	ncRNA
TYRP1	0.0199922310868145	-3.23	chr9	12543253	12543407	-150056	NM_000550	tyrosinase related protein 1	protein-coding
GOT2	0.0200006307870562	4.04	chr16	59063441	59063620	-329173	NM_002080	glutamic-oxaloacetic transaminase 2	protein-coding
CWH43	0.0200006307870562	4.04	chr4	49924304	49924480	937750	NM_001286791	cell wall biogenesis 43 C-terminal homolog	protein-coding
DOCK3	0.02007837791894	2.16	chr3	50921997	50922256	246885	NM_004947	dedicator of cytokinesis 3	protein-coding
RGS21	0.0201186983885245	2.93	chr1	192263285	192263668	-53516	NM_001039152	regulator of G-protein signaling 21	protein-coding
MIR548AB	0.0201295422090967	2.83	chr3	104214485	104214666	-690459	NR_039611	microRNA 548ab	ncRNA
LOC102723778	0.0201385954071159	4.11	chr4	31487140	31487298	316075	NR_125935	uncharacterized LOC102723778	ncRNA
CALB2	0.0202093897330819	1.93	chr16	71357159	71357359	-1454	NM_001740	calbindin 2	protein-coding
LINC01608	0.02023782855562	3.55	chr8	110685403	110685573	341945	NR_125416	long intergenic non-protein coding RNA 1608	ncRNA
ABL1	0.0202683923300254	-1.99	chr9	130771092	130771405	57367	NM_007313	ABL proto-oncogene 1, non-receptor tyrosine kinase	protein-coding
PTPRD-AS1	0.0203099770004546	2.37	chr9	9206590	9206766	348660	NR_121600	PTPRD antisense RNA 1	ncRNA
LRRTM4	0.0203905905402307	2.19	chr2	77457750	77458269	64424	NM_001330370	leucine rich repeat transmembrane neuronal 4	protein-coding
PRDM13	0.0204023961769846	2.54	chr6	99710403	99710737	103796	NM_021620	PR/SET domain 13	protein-coding
LINC01087	0.0204552492295214	2.57	chr2	131626900	131627048	-10051	NR_108087	long intergenic non-protein coding RNA 1087	ncRNA
ZNHIT6	0.0205109902594781	3.48	chr1	85713612	85713775	-5260	NM_017953	zinc finger HIT-type containing 6	protein-coding
ARHGAP15	0.0205203188301195	2.81	chr2	143333350	143333606	204148	NM_018460	Rho GTPase activating protein 15	protein-coding
YWHAEP7	0.0205627421504587	2.1	chr17	37889017	37889342	-4436	NR_024178	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon pseudogene 7	pseudo
CACNG5	0.0205921358631548	1.95	chr17	66836149	66836781	-40808	NM_145811	calcium voltage-gated channel auxiliary subunit gamma 5	protein-coding
LINC01605	0.0206018399109642	2.2	chr8	37501207	37501368	-7374	NR_121620	long intergenic non-protein coding RNA 1605	ncRNA
TSSK1B	0.0206152710414604	-2.55	chr5	113379102	113379256	55852	NM_032028	testis specific serine kinase 1B	protein-coding
LINC00680	0.0206162631239478	2.33	chr6	58560780	58561485	-599686	NR_125728	long intergenic non-protein coding RNA 680	ncRNA
GOLGA6C	0.0206229955989794	2.26	chr15	75246426	75246628	-12031	NM_001164404	golgin A6 family member C	protein-coding
PLEKHM2	0.0206470576757046	2.5	chr1	15673352	15673525	-10894	NM_015164	pleckstrin homology and RUN domain containing M2	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
P2RY1	0.0206820299246323	3.17	chr3	152957441	152957606	122830	NM_002563	purinergic receptor P2Y1	protein-coding
STXBP6	0.0208155209666944	2.9	chr14	25425590	25425751	-375373	NM_001304476	syntaxin binding protein 6	protein-coding
SYNDIG1	0.0208479510362528	-2.52	chr20	24419449	24419616	-49667	NM_024893	synapse differentiation inducing 1	protein-coding
TOMM70	0.0208694563397138	3.27	chr3	100425362	100425530	-24048	NM_014820	translocase of outer mitochondrial membrane 70	protein-coding
MLIP	0.0209188923774088	-3.41	chr6	54057260	54057504	38466	NM_138569	muscular LMNA interacting protein	protein-coding
YIPF7	0.0209456048127094	2.63	chr4	44659860	44660090	-8334	NM_182592	Yip1 domain family member 7	protein-coding
LOC105369785	0.0210798544802137	2.91	chr12	58105319	58105471	17503	NR_135022	uncharacterized LOC105369785	ncRNA
SLC44A1	0.0210803693349162	2.82	chr9	105159921	105160085	-84602	NM_001330731	solute carrier family 44 member 1	protein-coding
MOV10L1	0.0211787513470588	-2.22	chr22	50119474	50119821	-27408	NM_001164106	Mov10 RISC complex RNA helicase like 1	protein-coding
LOC101927686	0.0212720485516417	2.21	chr6	113354463	113354633	295527	NR_125844	uncharacterized LOC101927686	ncRNA
NKAIN2	0.0213040960054515	3.19	chr6	123812914	123813085	9153	NM_001040214	Sodium/potassium transporting ATPase interacting 2	protein-coding
LINC01625	0.0213182705870582	2.91	chr6	139590493	139590707	-116004	NR_033919	long intergenic non-protein coding RNA 1625	ncRNA
KIF13B	0.0213241261945424	2.12	chr8	29289457	29289888	-26579	NM_015254	kinesin family member 13B	protein-coding
LINC02035	0.0213509332142649	2.34	chr3	122898483	122898683	12070	NR_024618	long intergenic non-protein coding RNA 2035	ncRNA
MVK	0.0214033375125025	2.18	chr12	109604147	109604314	30535	NM_001114185	mevalonate kinase	protein-coding
FEZF2	0.0215083002372027	2.53	chr3	62591264	62591451	-217842	NM_018008	FEZ family zinc finger 2	protein-coding
FAM71C	0.0215190016083093	2.63	chr12	99540306	99540470	-107362	NM_153364	family with sequence similarity 71 member C	protein-coding
LOC101927815	0.0215620507487089	2.04	chr8	3634797	3635004	-906449	NR_125426	uncharacterized LOC101927815	ncRNA
ZFHX4-AS1	0.0216010981824185	2.21	chr8	76771113	76771353	-87958	NR_024360	ZFHX4 antisense RNA 1	ncRNA
FCGBP	0.0216358733733438	-2.09	chr19	39897925	39898154	36587	NM_003890	Fc fragment of IgG binding protein	protein-coding
ETNK1	0.0216545277222911	2.33	chr12	22729125	22729305	104073	NM_001039481	ethanolamine kinase 1	protein-coding
LOC105378183	0.0216689686206653	1.99	chr14	103254179	103254594	-62898	NR_135269	uncharacterized LOC105378183	ncRNA
AASDH	0.0217123660697768	2.21	chr4	56316034	56316223	71380	NM_001323899	aminoadipate-semialdehyde dehydrogenase	protein-coding
MCTP2	0.0217390995497213	2.83	chr15	94313685	94313900	15591	NM_018349	multiple C2 and transmembrane domain containing 2	protein-coding
SLC9A9-AS1	0.0217698276354045	2.4	chr3	143528715	143529111	186667	NR_048544	SLC9A9 antisense RNA 1	ncRNA
MTRNR2L6	0.021830459657943	3.14	chr7	142650899	142651049	-15325	NM_001190487	MT-RNR2-like 6	protein-coding
ATP8A2	0.0219234840914039	2.01	chr13	25372470	25372792	620	NM_016529	ATPase phospholipid transporting 8A2	protein-coding
MIR6124	0.0220678866698519	1.86	chr11	12197823	12198452	34454	NR_106739	microRNA 6124	ncRNA
PRG4	0.0222092702631737	3.24	chr1	186137010	186137200	-159168	NM_005807	proteoglycan 4	protein-coding
AXIN2	0.0223239522617612	2.26	chr17	65630501	65630770	-69013	NM_004655	axin 2	protein-coding
POM121L2	0.0225431331844441	-2.55	chr6	27280095	27280275	32047	NM_033482	POM121 transmembrane nucleoporin like 2	protein-coding
LOC101928944	0.0225475259115525	3.41	chr11	81293240	81293442	-530538	NR_120570	uncharacterized LOC101928944	ncRNA
PRKCH	0.0225667589996543	2.87	chr14	61367478	61367642	46117	NM_006255	protein kinase C eta	protein-coding
ZFPM2-AS1	0.0225738277736151	3.28	chr8	106001415	106001627	58982	NR_125797	ZFPM2 antisense RNA 1	ncRNA
CBLN1	0.0226907886226692	2.17	chr16	49302532	49302698	-20784	NM_004352	cerebellin 1 precursor	protein-coding
BMPER	0.0227591801614322	2.18	chr7	33678848	33679183	-225896	NM_133468	BMP binding endothelial regulator	protein-coding
LINC00347	0.0228553200933808	1.86	chr13	74814303	74814755	261686	NR_034025	long intergenic non-protein coding RNA 347	ncRNA
TRHDE-AS1	0.0228656607619212	2.44	chr12	72677564	72677769	-404157	NR_026836	TRHDE antisense RNA 1	ncRNA
RNU2-1	0.0228905032228386	1.82	chr17	43259008	43259251	-7110	NR_002716	RNA, U2 small nuclear 1	snRNA
PDGFC	0.0229206206449056	2.95	chr4	156823887	156824173	147364	NR_036641	platelet derived growth factor C	protein-coding
TMEM261	0.0230503226387189	-3.12	chr9	7936125	7936344	-136428	NM_033428	transmembrane protein 261	protein-coding
RAG2	0.0230886949346319	2.45	chr11	36931235	36931490	-333083	NM_001243785	recombination activating 2	protein-coding
LOC105376365	0.0231257881211574	2.39	chr10	3838413	3838577	4545	NR_134490	uncharacterized LOC105376365	ncRNA
ELF2	0.0231258322882546	2.39	chr4	139152979	139153325	-13676	NM_201999	E74 like ETS transcription factor 2	protein-coding
MIR4306	0.0231759483121473	-2.52	chr13	99630439	99630909	-12385	NR_036191	microRNA 4306	ncRNA
LOC100996349	0.0232046346630242	2.05	chr19	22607194	22607380	-729	NR_103470	testis expressed 264 pseudogene	pseudo
UBE2E3	0.0232398062402015	3	chr2	180903039	180903216	-77258	NM_001278555	ubiquitin conjugating enzyme E2 E3	protein-coding
XPNPEP2	0.02327396574949	1.83	chrX	129713199	129713518	-25612	NM_003399	X-prolyl aminopeptidase 2	protein-coding
GHR	0.0233439830411988	-2.01	chr5	42581828	42582133	16116	NM_001242462	growth hormone receptor	protein-coding
MYO1B	0.0233613097640666	2.15	chr2	191305000	191305353	59900	NM_001161819	myosin IB	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
NR2F1-AS1	0.0233621930998046	2.52	chr5	93249316	93249699	321836	NR_109825	NR2F1 antisense RNA 1	ncRNA
CCSER1	0.023382572650723	3.02	chr4	90655533	90655784	420627	NM_207491	coiled-coil serine rich protein 1	protein-coding
MEAT6	0.0234682052724319	2.71	chr6	164907787	164907960	-85810	NR_131926	melanoma-associated transcript 6	ncRNA
AREG	0.0234690246592163	3.46	chr4	74473378	74473541	28361	NM_001657	amphiregulin	protein-coding
CDH10	0.0236170749221	3.27	chr5	24317587	24317760	275826	NM_001317224	cadherin 10	protein-coding
LINC01118	0.023619169439112	2.22	chr2	46779835	46780029	-36736	NR_034099	long intergenic non-protein coding RNA 1118	ncRNA
FRG1CP	0.0236401779444645	3.14	chr20	28384115	28384282	218467	NR_132315	FSHD region gene 1 family member C, pseudogene	pseudo
GATA5	0.0236892316397422	1.96	chr20	62475396	62475569	488	NM_080473	GATA binding protein 5	protein-coding
SELENOP	0.0237302695463792	-2.1	chr5	42888943	42889759	-77429	NM_001085486	selenoprotein P	protein-coding
AMACR	0.023780282408183	-1.83	chr5	34001308	34001744	6589	NM_001167595	alpha-methylacyl-CoA racemase	protein-coding
TBC1D4	0.0237915407877346	3.99	chr13	75380815	75380968	101277	NM_014832	TBC1 domain family member 4	protein-coding
FTMT	0.0238641941488914	3.42	chr5	121891504	121891662	39628	NM_177478	ferritin mitochondrial	protein-coding
LOC105374704	0.0238676466309396	3.82	chr5	30289545	30289703	409064	NR_134264	uncharacterized LOC105374704	ncRNA
KCNK12	0.0238795082496254	2.23	chr2	47561253	47561478	8966	NM_022055	potassium two pore domain channel subfamily K member 12	protein-coding
LINC01959	0.0238837516325145	2.1	chr5	135625233	135625486	28576	NR_027127	long intergenic non-protein coding RNA 1959	ncRNA
LOC100506858	0.0239921393750975	-2.79	chr5	2429263	2429480	-117170	NR_104616	uncharacterized LOC100506858	ncRNA
LINC01781	0.0239922426603459	-2.48	chr1	80892462	80892788	251338	NR_125942	long intergenic non-protein coding RNA 1781	ncRNA
SEMA6D	0.02402035540821	2.34	chr15	46526224	46526452	-657868	NM_001198999	semaphorin 6D	protein-coding
YTHDC2	0.0240282658707148	2.24	chr5	113752957	113753135	239384	NM_001345976	YTH domain containing 2	protein-coding
MBOAT4	0.0240398438391498	-2.24	chr8	30137189	30137426	7377	NM_001100916	membrane bound O-acyltransferase domain containing 4	protein-coding
SVIL-AS1	0.0241660601250505	-1.83	chr10	29352886	29353176	-56503	NR_110921	SVIL antisense RNA 1	ncRNA
SPATA4	0.024205960416674	2.62	chr4	176150886	176151292	44582	NM_144644	spermatogenesis associated 4	protein-coding
SEC31B	0.0242793390409345	2.2	chr10	100498287	100498519	21435	NM_015490	SEC31 homolog B, COPII coat complex component	protein-coding
LOC101929621	0.0243158187342091	2.43	chr4	108394912	108395173	-162191	NR_134679	uncharacterized LOC101929621	ncRNA
LINC00392	0.0244153455180117	2.71	chr13	73711959	73712118	147794	NR_047009	long intergenic non-protein coding RNA 392	ncRNA
CTNND2	0.0244499025680489	-2.61	chr5	11466883	11467042	121955	NM_001288716	catenin delta 2	protein-coding
DSCAM-IT1	0.024468536190398	2.14	chr21	40653003	40653164	-22316	NR_046774	DSCAM intronic transcript 1	ncRNA
GNPDA2	0.0244801003158106	2.6	chr4	45371971	45372165	-645434	NR_073094	glucosamine-6-phosphate deaminase 2	protein-coding
MBOAT1	0.024538760785798	2.37	chr6	20183208	20183372	29174	NR_073465	membrane bound O-acyltransferase domain containing 1	protein-coding
LOC105369891	0.0245413173452995	2.77	chr12	89844465	89844689	-103116	NR_135019	uncharacterized LOC105369891	ncRNA
LOC105370369	0.0245413173452995	2.77	chr13	111589054	111589272	-6846	NR_135814	uncharacterized LOC105370369	ncRNA
NPVF	0.0245633964955473	1.87	chr7	25313130	25313759	-84958	NM_022150	neuropeptide VF precursor	protein-coding
EGFLAM-AS4	0.024655711268266	3.4	chr5	38312017	38312189	-21117	NR_046219	EGFLAM antisense RNA 4	ncRNA
NDFIP1	0.0246969777786253	-3.37	chr5	142120315	142120516	11656	NM_030571	Nedd4 family interacting protein 1	protein-coding
AQP4	0.0247285088933612	2.36	chr18	26971381	26971690	-105691	NM_001317387	aquaporin 4	protein-coding
LINC00972	0.0247738513466741	3.54	chr7	85464924	85465106	43893	NR_134240	long intergenic non-protein coding RNA 972	ncRNA
LINC00907	0.0247788787781085	3.71	chr18	42365560	42365750	178987	NR_046456	long intergenic non-protein coding RNA 907	ncRNA
CDH8	0.0248069857650987	3.79	chr16	62267630	62267795	-230877	NM_001796	cadherin 8	protein-coding
LINC01676	0.0248168865834829	3.32	chr1	105527200	105527433	91619	NR_125954	long intergenic non-protein coding RNA 1676	ncRNA
LOC102467214	0.0248608293979387	2.47	chr5	112572107	112572449	-56158	NR_104672	uncharacterized LOC102467214	ncRNA
MUT	0.0248939200117733	2.46	chr6	49194922	49195134	268300	NM_000255	methylmalonyl-CoA mutase	protein-coding
TCF7L1	0.0248994685413589	2.3	chr2	85158900	85159075	25527	NM_031283	transcription factor 7 like 1	protein-coding
LINC00693	0.0249000502103792	2.1	chr3	28575301	28575535	140	NR_038840	long intergenic non-protein coding RNA 693	ncRNA
RPTOR	0.0249294055194473	2.49	chr17	80619032	80619195	74288	NM_020761	regulatory associated protein of MTOR complex 1	protein-coding
FXYD4	0.0250325310127141	2.12	chr10	43337181	43337340	-34384	NM_001184963	FXYD domain containing ion transport regulator 4	protein-coding
ARX	0.0250385160357136	2.07	chrX	25012995	25013414	2744	NM_139058	aristaless related homeobox	protein-coding
LINC01192	0.0250387038809385	3.22	chr3	163081804	163081953	221423	NR_033945	long intergenic non-protein coding RNA 1192	ncRNA
CD48	0.0250979536779129	2.35	chr1	160696191	160696404	15554	NM_001256030	CD48 molecule	protein-coding
MIR1267	0.0251028769453223	2.72	chr13	107204205	107204366	326963	NR_031671	microRNA 1267	ncRNA
EMB	0.0251650229168757	2.98	chr5	50112873	50113041	328443	NM_198449	embigin	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
LRP1B	0.0251747763080953	3.1	chr2	141952679	141952839	178942	NM_018557	LDL receptor related protein 1B	protein-coding
PRMT8	0.0251843280982084	-3.19	chr12	3391018	3391186	9753	NM_001256536	protein arginine methyltransferase 8	protein-coding
FAM84B	0.0252493203566743	-2.27	chr8	126744231	126744453	-185876	NM_174911	family with sequence similarity 84 member B	protein-coding
LINC01309	0.0252966728815779	2.71	chr13	104315246	104315711	890278	NR_126374	long intergenic non-protein coding RNA 1309	ncRNA
GAREM2	0.0253056189598121	3.63	chr2	26103767	26103959	-69228	NM_001168241	GRB2 associated regulator of MAPK1 subtype 2	protein-coding
LOC101927078	0.0254337533478713	3.41	chr5	114683531	114683706	89795	NR_130785	uncharacterized LOC101927078	ncRNA
MGAT4C	0.0254517483851759	2.47	chr12	86565789	86565951	273034	NM_013244	MGAT4 family member C	protein-coding
SNTG1	0.0255146836856778	2.15	chr8	50039938	50040125	-124100	NM_001321778	syntrophin gamma 1	protein-coding
MIR4774	0.0256862744056141	3.13	chr2	168621823	168621986	38961	NR_039933	microRNA 4774	ncRNA
LOC340017	0.0256965125992304	1.85	chr4	157444071	157444574	-128168	NR_026992	uncharacterized LOC340017	ncRNA
CXXC4	0.0257004801857867	3.39	chr4	104583641	104583802	-88820	NM_025212	CXXC finger protein 4	protein-coding
LINC01947	0.025735960765165	3.24	chr5	166760530	166760744	165733	NR_108020	long intergenic non-protein coding RNA 1947	ncRNA
LVCAT5	0.025735960765165	3.24	chr7	119682188	119682366	225098	NR_131960	liver cancer-associated transcript 5	ncRNA
DCN	0.0257396864628811	2.67	chr12	91229005	91229653	-46205	NM_001920	decorin	protein-coding
TRIB2	0.0258075811875951	2.19	chr2	12685342	12685524	-31439	NR_027303	tribbles pseudokinase 2	protein-coding
KCNH7	0.0258480657135803	2.34	chr2	163155416	163155696	-316809	NM_033272	potassium voltage-gated channel subfamily H member 7	protein-coding
ABHD6	0.0258767029941427	-2	chr3	58246317	58246806	9059	NM_001320126	abhydrolase domain containing 6	protein-coding
BTC	0.0258848938504408	1.98	chr4	74769154	74769648	25285	NM_001729	betacellulin	protein-coding
CABLES1	0.0258976533228971	2.64	chr18	23115304	23115485	-19170	NM_001256438	Cdk5 and Abl enzyme substrate 1	protein-coding
ZBBX	0.0259079062691839	2.6	chr3	167148316	167148676	231787	NM_001199202	zinc finger B-box domain containing	protein-coding
LMX1A	0.0259228375423776	2.29	chr1	165282737	165282930	73408	NM_001174069	LIM homeobox transcription factor 1 alpha	protein-coding
MIR8084	0.0259438092377464	2.87	chr8	93003523	93003698	-26141	NR_107051	microRNA 8084	ncRNA
GHSR	0.025958739856738	2.1	chr3	172401418	172402086	46704	NM_004122	growth hormone secretagogue receptor	protein-coding
LOC101928433	0.0260001006753276	3.92	chr6	16752501	16752676	-8550	NR_136240	uncharacterized LOC101928433	ncRNA
ETV6	0.0260615001775708	-2.54	chr12	11712696	11713120	63054	NM_001987	ETS variant 6	protein-coding
LOC101928052	0.0260990156327188	3.19	chr4	161554021	161554192	-468627	NR_125888	uncharacterized LOC101928052	ncRNA
RNU105B	0.0261587143669827	2.92	chr20	8709841	8710088	-121221	NR_004386	RNA, U105B small nucleolar	snoRNA
APBA2	0.0261685191944573	2.65	chr15	28922790	28923042	1279	NM_005503	amyloid beta precursor protein binding family A member 2	protein-coding
EGFLAM	0.0261755182141455	2.19	chr5	38263924	38264193	5649	NM_001205301	EGF like, fibronectin type III and laminin G domains	protein-coding
SIK2	0.0261961146619585	-2.42	chr11	111671506	111671790	69257	NM_015191	salt inducible kinase 2	protein-coding
SIK2	0.0262309310300866	-2.13	chr11	111586492	111586898	-15696	NM_015191	salt inducible kinase 2	protein-coding
PSG7	0.026253898383251	-2.77	chr19	42952214	42952437	-15147	NM_001290042	pregnancy specific beta-1-glycoprotein 7 (gene/pseudogene)	protein-coding
LOC105370586	0.0262556106653685	2.28	chr14	78941874	78942024	-258033	NR_135159	uncharacterized LOC105370586	ncRNA
ZNF536	0.0264070033302272	1.97	chr19	30375472	30375649	3167	NM_014717	zinc finger protein 536	protein-coding
TMEM45A	0.0264596715147954	2.19	chr3	100500332	100500733	7913	NM_018004	transmembrane protein 45A	protein-coding
PDE4DIP	0.0264776998243908	2.16	chr1	148885335	148885515	-1948	NM_001195260	phosphodiesterase 4D interacting protein	protein-coding
HCN1	0.026532444413925	5.37	chr5	45371515	45371694	324514	NM_021072	hyperpolarization activated cyclic nucleotide gated potassium channel 1	protein-coding
SNORA98	0.0266089326400475	3.39	chr6	144438518	144438742	32975	NR_132778	small nucleolar RNA, H/ACA box 98	snoRNA
POTEA	0.0266221267389193	2.23	chr8	44430868	44431031	1138507	NM_001002920	POTE ankyrin domain family member A	protein-coding
MOXD1	0.0266426362195678	2.53	chr6	132384848	132385016	16602	NM_015529	monooxygenase DBH like 1	protein-coding
PTPRD-AS2	0.0266605066390579	2.18	chr9	11197519	11197670	584388	NR_110696	PTPRD antisense RNA 2 (head to head)	ncRNA
BNC2	0.0266970243252272	2.09	chr9	16876695	16876899	-6009	NM_001317939	basonuclin 2	protein-coding
PLEKHA7	0.0267236049651158	-2.24	chr11	16936622	16936978	77616	NM_001329631	pleckstrin homology domain containing A7	protein-coding
LINC01361	0.0267466976919072	3.51	chr1	82596182	82596355	388654	NR_110632	long intergenic non-protein coding RNA 1361	ncRNA
PCSK2	0.0267846716175966	-2.8	chr20	17268883	17269052	42015	NM_002594	proprotein convertase subtilisin/kexin type 2	protein-coding
LOC257396	0.0268203682761887	3.22	chr5	53253650	53253806	143886	NR_034107	uncharacterized LOC257396	ncRNA
MIR4650-1	0.026839074436338	1.89	chr7	72757827	72757997	60009	NR_039793	microRNA 4650-1	ncRNA
LINC00456	0.0270425024325591	2.68	chr13	97191770	97191941	-12233	NR_125725	long intergenic non-protein coding RNA 456	ncRNA
SULF2	0.0270593325716886	-1.96	chr20	47786718	47786898	-192	NM_001161841	sulfatase 2	protein-coding
DISC1FP1	0.0271303688415887	2.09	chr11	90363960	90364191	112843	NR_104190	DISC1 fusion partner 1 (non-protein coding)	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
RBFOX1	0.0272439694333839	3.67	chr16	6318018	6318187	298399	NM_001308117	RNA binding protein, fox-1 homolog 1	protein-coding
EDIL3	0.0272995440510437	2.4	chr5	85049159	85049378	-664401	NM_005711	EGF like repeats and discoidin domains 3	protein-coding
LOC101928372	0.0273063157383068	2.51	chr1	160933966	160934189	1612	NR_110695	uncharacterized LOC101928372	ncRNA
LOC101927849	0.0274204898216639	3.01	chr4	149423429	149423608	269223	NR_125887	uncharacterized LOC101927849	ncRNA
MIR4735	0.0274261950600984	3.32	chr1	195040846	195041011	1541553	NR_039888	microRNA 4735	ncRNA
DLC1	0.0274312048275212	3.3	chr8	13506876	13507054	7955	NM_182643	DLC1 Rho GTPase activating protein	protein-coding
IL15	0.0274673665579631	2.89	chr4	141605506	141605673	-31007	NM_172175	interleukin 15	protein-coding
DUSP6	0.0274792820444024	1.85	chr12	89205687	89206281	146875	NM_001946	dual specificity phosphatase 6	protein-coding
TRIO	0.027505388027115	2.52	chr5	14342277	14342441	198657	NM_007118	trio Rho guanine nucleotide exchange factor	protein-coding
DBET	0.0275604125308049	2.53	chr4	190117035	190117261	52646	NR_121644	D4Z4 binding element transcript (non-protein coding)	ncRNA
CCR4	0.0275617531871602	1.85	chr3	32940310	32940538	-11150	NM_005508	C-C motif chemokine receptor 4	protein-coding
LINC00636	0.0276118908248536	2.14	chr3	107938226	107938436	55126	NR_015394	long intergenic non-protein coding RNA 636	ncRNA
CDH6	0.027677359731056	2.85	chr5	30620946	30621120	-572622	NM_004932	cadherin 6	protein-coding
PRR16	0.0277138415712817	2.63	chr5	120346911	120347076	-117285	NM_001300783	proline rich 16	protein-coding
GLRA1	0.0277422441370771	2.06	chr5	151858013	151858254	66703	NM_001292000	glycine receptor alpha 1	protein-coding
COL11A1	0.0277984228233305	2.52	chr1	102826661	102826813	281759	NM_080629	collagen type XI alpha 1 chain	protein-coding
LINC00535	0.0278444698430017	-2.81	chr8	93641909	93642154	58402	NR_033858	long intergenic non-protein coding RNA 535	ncRNA
LOC439933	0.0278680249274434	3.36	chr4	35163744	35163915	-1080287	NR_122079	uncharacterized LOC439933	ncRNA
MDGA2	0.0278833810388566	3.06	chr14	47253650	47253879	89482	NR_103766	MAM domain containing glycosylphosphatidylinositol anchor 2	protein-coding
STMN4	0.0279103204023394	-3.72	chr8	27181529	27181696	76827	NM_001283053	stathmin 4	protein-coding
LOC105374704	0.0279422696898209	-3.4	chr5	29806904	29807106	-73555	NR_134264	uncharacterized LOC105374704	ncRNA
LINC01602	0.0280002138879949	2.71	chr8	57975811	57975961	-2472	NR_130934	long intergenic non-protein coding RNA 1602	ncRNA
PLXNC1	0.0281107116896142	2.6	chr12	94170909	94171209	22336	NM_005761	plexin C1	protein-coding
MBL2	0.0281580848004963	2.48	chr10	53469484	53469674	-697879	NM_000242	mannose binding lectin 2	protein-coding
LOC100287010	0.0281828470718166	2.55	chr2	104332764	104332951	75475	NR_037885	uncharacterized LOC100287010	ncRNA
LINC02032	0.0282067472832035	1.82	chr3	148315181	148315579	237221	NR_134937	long intergenic non-protein coding RNA 2032	ncRNA
BLOC1S5	0.0282398575430224	2.44	chr6	8059098	8059309	5211	NM_001199322	biogenesis of lysosomal organelles complex 1 subunit 5	protein-coding
LOC100129138	0.028304509107789	2.13	chr1	104147571	104147739	74632	NR_033990	THAP domain containing, apoptosis associated protein 3 pseudogene	pseudo
ATP1A1	0.0283356191481449	-2.03	chr1	116291058	116291369	-81960	NM_000701	ATPase Na+/K+ transporting subunit alpha 1	protein-coding
SRGAP2C	0.0284589977743318	2.14	chr1	121281803	121282166	97017	NM_001271872	SLIT-ROBO Rho GTPase activating protein 2C	protein-coding
HMCN1	0.0284686896195903	2.07	chr1	185789892	185790123	55456	NM_031935	hemicentin 1	protein-coding
RNLS	0.0285138363458193	2.45	chr10	88226389	88226549	356856	NM_001031709	renalase, FAD dependent amine oxidase	protein-coding
CCT5	0.028537130464208	2.07	chr5	10275910	10276093	25464	NM_001306156	chaperonin containing TCP1 subunit 5	protein-coding
C22orf24	0.0285487368633429	2.78	chr22	31932274	31932499	6068	NM_001302819	chromosome 22 open reading frame 24	protein-coding
GABBR2	0.0285503155623031	2.3	chr9	98494714	98494912	214384	NM_005458	gamma-aminobutyric acid type B receptor subunit 2	protein-coding
QKI	0.0285865211026286	2.77	chr6	163526366	163526526	111803	NM_001301085	QKI, KH domain containing RNA binding	protein-coding
C10orf107	0.0285910801951704	2.88	chr10	61587028	61587207	-75844	NM_173554	chromosome 10 open reading frame 107	protein-coding
SMAD5	0.028606657511394	2.49	chr5	136119835	136120049	-12903	NM_001001420	SMAD family member 5	protein-coding
ACTR3BP5	0.0286591343719727	3.46	chr10	39241333	39241539	544840	NR_045000	ACTR3B pseudogene 5	pseudo
LINC01242	0.0286780780582611	3.02	chr9	30276992	30277154	131381	NR_046204	long intergenic non-protein coding RNA 1242	ncRNA
LINC00358	0.0288022019149156	3.12	chr13	62119925	62120177	-90503	NR_046995	long intergenic non-protein coding RNA 358	ncRNA
LINC01297	0.028844073213224	1.94	chr22	15760750	15760969	14185	NR_132385	long intergenic non-protein coding RNA 1297	ncRNA
TEX26	0.0288719479938694	2.26	chr13	30990694	30991061	58180	NM_152325	testis expressed 26	protein-coding
LOC653712	0.0288816762290502	2.46	chr3	128865223	128865377	6241	NR_034179	intraflagellar transport 122 homolog (Chlamydomonas) pseudogene	pseudo
PDPR	0.0288873973867505	1.83	chr16	70108324	70108532	-5198	NM_001322119	pyruvate dehydrogenase phosphatase regulatory subunit	protein-coding
RAMP3	0.0288995279150204	2.52	chr7	45348750	45348939	191076	NM_005856	receptor activity modifying protein 3	protein-coding
LINC02072	0.0289091207153721	-2.9	chr17	51640330	51640482	303691	NR_122104	long intergenic non-protein coding RNA 2072	ncRNA
MIR6788	0.0289295852442469	-1.82	chr18	10914815	10915041	-155279	NR_106846	microRNA 6788	ncRNA
LOC440040	0.0289343296720732	4.02	chr11	49672442	49672627	114006	NR_027044	glutamate metabotropic receptor 5 pseudogene	pseudo
NLGN1-AS1	0.0289343296720732	4.02	chr3	173899282	173899478	21416	NR_046664	NLGN1 antisense RNA 1	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
OR4F17	0.0289965332161734	2.37	chr19	141052	141229	30461	NM_001005240	olfactory receptor family 4 subfamily F member 17	protein-coding
LINC-PINT	0.0290782589875161	-1.91	chr7	131088042	131089152	18433	NR_024153	long intergenic non-protein coding RNA, p53 induced transcript	ncRNA
FRG2EP	0.0291110424564679	2.29	chr20	29309494	29309663	107383	NR_037925	FSHD region gene 2 family member E, pseudogene	pseudo
SEMA6A-AS1	0.0292091499093372	-2.96	chr5	116492188	116492412	44802	NR_109879	SEMA6A antisense RNA 1	ncRNA
FBXO15	0.0292824834898017	2.09	chr18	73767326	73767726	380238	NM_152676	F-box protein 15	protein-coding
GRIN2B	0.0293670177831381	2.04	chr12	13850371	13850543	129631	NM_000834	glutamate ionotropic receptor NMDA type subunit 2B	protein-coding
OR2T33	0.0293691321261378	2.69	chr1	248272793	248272974	931	NM_001004695	olfactory receptor family 2 subfamily T member 33	protein-coding
SLC2A2	0.029399911099428	-2.73	chr3	171060149	171060310	-33250	NM_001278659	solute carrier family 2 member 2	protein-coding
ADGRA2	0.029448353771648	2.06	chr8	37801445	37801610	4644	NM_032777	adhesion G protein-coupled receptor A2	protein-coding
LINC01288	0.0296262637873351	-3.02	chr8	34437219	34437380	-346622	NR_125746	long intergenic non-protein coding RNA 1288	ncRNA
MIR4445	0.0296533342327407	2.02	chr3	109532403	109532586	-70334	NR_039647	microRNA 4445	ncRNA
AFF3	0.029724526703104	2.54	chr2	100019341	100019505	86160	NM_001025108	AF4/FMR2 family member 3	protein-coding
PRR20B	0.0297664850210574	3.44	chr13	56198632	56198797	-942204	NM_001130404	proline rich 20B	protein-coding
LRRC37A	0.0298020151151983	2.2	chr17	46277048	46277291	-17962	NM_014834	leucine rich repeat containing 37A	protein-coding
ARHGAP32	0.0298627139527618	2.12	chr11	129221967	129222344	-29957	NM_001142685	Rho GTPase activating protein 32	protein-coding
PDGFC	0.0299256988592423	2.6	chr4	156774448	156774645	196848	NR_036641	platelet derived growth factor C	protein-coding
LINC01600	0.0300120342126392	2.39	chr6	2450798	2450969	183720	NR_131168	long intergenic non-protein coding RNA 1600	ncRNA
RAD18	0.0300270861729622	2.58	chr3	8997133	8997290	-33736	NM_020165	RAD18, E3 ubiquitin protein ligase	protein-coding
DPP10-AS3	0.0300321906954234	2.51	chr2	114804591	114804777	28864	NR_132105	DPP10 antisense RNA 3	ncRNA
LINC01600	0.0300351405673204	-2.98	chr6	2564013	2564176	70509	NR_131168	long intergenic non-protein coding RNA 1600	ncRNA
BCL2L11	0.0300456203516764	2.73	chr2	111200918	111201087	80088	NM_138625	BCL2 like 11	protein-coding
DUSP26	0.0300633523947919	2.44	chr8	33722786	33722945	-122759	NM_024025	dual specificity phosphatase 26	protein-coding
DIO2	0.0301057362094346	2.77	chr14	80377109	80377285	-146143	NM_000793	iodothyronine deiodinase 2	protein-coding
LINC00879	0.0302003974163168	3.29	chr3	94692397	94692650	-245740	NR_015400	long intergenic non-protein coding RNA 879	ncRNA
ENO4	0.0302016392927216	-2.02	chr10	116840587	116840942	-8748	NM_001242699	enolase family member 4	protein-coding
PKHD1	0.0302322715139331	2.81	chr6	51962421	51962595	125117	NM_138694	polycystic kidney and hepatic disease 1 (autosomal recessive)	protein-coding
LINC01696	0.0303379290831565	-3.45	chr1	209177014	209177190	-148354	NR_135085	long intergenic non-protein coding RNA 1696	ncRNA
KDM4C	0.0303999358757586	2.16	chr9	7045933	7046820	288735	NM_001304341	lysine demethylase 4C	protein-coding
STBD1	0.0304416833900865	-2.08	chr4	76281040	76281294	-24859	NM_003943	starch binding domain 1	protein-coding
RAMP3	0.030510642073037	2.19	chr7	45312103	45312276	154421	NM_005856	receptor activity modifying protein 3	protein-coding
PSPC1	0.0305141056952452	-3.11	chr13	19726850	19727048	56070	NM_001042414	paraspeckle component 1	protein-coding
KCNH8	0.0305227355230767	2.86	chr3	18847788	18848157	-300553	NM_144633	potassium voltage-gated channel subfamily H member 8	protein-coding
SMIM23	0.0305282130911503	1.89	chr5	171623754	171623930	-162030	NM_001289970	small integral membrane protein 23	protein-coding
ARHGAP18	0.0305694899238741	-2.27	chr6	129748355	129749013	-38459	NM_033515	Rho GTPase activating protein 18	protein-coding
NEU4	0.0306451313210157	2.48	chr2	241833764	241834138	23019	NM_001167602	neuraminidase 4	protein-coding
RIC1	0.0307074566382113	-2.64	chr9	5696618	5697240	67810	NM_001135920	RIC1 homolog, RAB6A GEF complex partner 1	protein-coding
LINC02103	0.0307176677360328	2.83	chr5	28563703	28563864	-276118	NR_134265	long intergenic non-protein coding RNA 2103	ncRNA
TRIB2	0.0307293278269898	2.34	chr2	12704224	12704423	-12549	NR_027303	tribbles pseudokinase 2	protein-coding
PLAC4	0.0307742745228629	2.22	chr21	41236611	41236822	-51477	NM_182832	placenta specific 4	protein-coding
HCN1	0.030786489596072	3.9	chr5	45774568	45774748	-78540	NM_021072	hyperpolarization activated cyclic nucleotide gated potassium channel 1	protein-coding
MIR548AC	0.0308953138005431	-2.92	chr1	116555880	116556059	4142	NR_039621	microRNA 548ac	ncRNA
KIF26B	0.0309818208334409	-2.53	chr1	245492002	245492186	337109	NM_018012	kinesin family member 26B	protein-coding
LOC100192426	0.0310303572796605	1.87	chr18	8063112	8063412	303772	NR_024419	uncharacterized LOC100192426	ncRNA
SND1	0.0310380813777663	1.97	chr7	127821035	127821446	169251	NM_014390	staphylococcal nuclease and tudor domain containing 1	protein-coding
LINC01052	0.031085785999082	2.66	chr13	65828488	65828666	-37470	NR_132362	long intergenic non-protein coding RNA 1052	ncRNA
ZNF733P	0.031106857676514	2.7	chr7	62554865	62555133	749057	NR_003952	zinc finger protein 733, pseudogene	pseudo
ASCC3	0.0311465030564911	-1.82	chr6	100719811	100720600	161167	NM_022091	activating signal cointegrator 1 complex subunit 3	protein-coding
SH2D7	0.0311519820932954	1.92	chr15	78092174	78092624	-186	NM_001101404	SH2 domain containing 7	protein-coding
LYST	0.0311847147045494	2.08	chr1	235910575	235910754	-26956	NM_001301365	lysosomal trafficking regulator	protein-coding
LHFP	0.0311917376153169	3.37	chr13	39581864	39582076	21249	NM_005780	lipoma HMGIC fusion partner	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
TPK1	0.0312552770830566	2.67	chr7	145054920	145055070	-218942	NM_001042482	thiamin pyrophosphokinase 1	protein-coding
LOC105376194	0.0313276377628213	-2.41	chr9	104221043	104221256	-49080	NR_135141	uncharacterized LOC105376194	ncRNA
PCDH15	0.0313485410231034	2.59	chr10	54792963	54793147	8236	NM_001142772	protocadherin related 15	protein-coding
ACO1	0.0313899057123139	1.93	chr9	32382786	32383041	-1690	NM_001278352	aconitase 1	protein-coding
LOC101928052	0.0314374696870711	2.77	chr4	161092580	161092771	-930058	NR_125888	uncharacterized LOC101928052	ncRNA
LINC01592	0.0314485455293249	2.22	chr8	69078443	69078716	25611	NR_039986	long intergenic non-protein coding RNA 1592	ncRNA
LINC00587	0.0314645706027904	3.51	chr9	102185466	102185704	-334052	NR_103830	long intergenic non-protein coding RNA 587	ncRNA
FRG1CP	0.0314660574741702	1.9	chr20	28481390	28481691	121125	NR_132315	FSHD region gene 1 family member C, pseudogene	pseudo
MTAP	0.031501283803231	2.16	chr9	21836688	21836868	34142	NM_002451	methylthioadenosine phosphorylase	protein-coding
LOC101928052	0.0315492125465661	2.56	chr4	160960514	160960682	-1062135	NR_125888	uncharacterized LOC101928052	ncRNA
TMEM266	0.0316750394407366	2.35	chr15	76069119	76069361	9282	NM_152335	transmembrane protein 266	protein-coding
MIR499A	0.0317072827591597	2.52	chr20	34976068	34976235	-14225	NR_030223	microRNA 499a	ncRNA
STUM	0.0317152582330576	-2.71	chr1	226542199	226542378	-6512	NM_001003665	stum, mechanosensory transduction mediator homolog	protein-coding
SERINC1	0.0318073996642278	2.02	chr6	122449118	122449916	22364	NM_020755	serine incorporator 1	protein-coding
RABGAP1L	0.0318272733793813	2.26	chr1	174726350	174726702	-73371	NM_001330989	RAB GTPase activating protein 1 like	protein-coding
RP1L1	0.0318650051127522	1.89	chr8	10598652	10598852	56355	NM_178857	RP1 like 1	protein-coding
RIIAD1	0.0318802303310946	2.1	chr1	151739917	151740152	18497	NM_001144956	regulatory subunit of type II PKA R-subunit (RIIa) domain containing 1	protein-coding
PACRG	0.0318945587804617	2.54	chr6	162526970	162527129	-200083	NM_152410	PARK2 coregulated	protein-coding
ABCA8	0.0318958777028848	2.59	chr17	68809534	68809709	145771	NM_007168	ATP binding cassette subfamily A member 8	protein-coding
LOC102723439	0.0319009067077714	2.39	chr10	74608690	74608856	-79449	NR_120673	uncharacterized LOC102723439	ncRNA
LIMS1	0.0319390463979331	2.35	chr2	108570302	108570587	-17857	NM_004987	LIM zinc finger domain containing 1	protein-coding
MIR1263	0.0319577221976488	3.34	chr3	164365281	164365442	-193805	NR_031665	microRNA 1263	ncRNA
IL1R1	0.0320631643514587	-2.81	chr2	102035270	102035463	-35010	NM_001288706	interleukin 1 receptor type 1	protein-coding
SAMD9	0.0321239041833844	2.2	chr7	93017355	93017691	100500	NM_001193307	sterile alpha motif domain containing 9	protein-coding
VGLL2	0.0321753858697882	-2.85	chr6	117198895	117199056	-66583	NM_153453	vestigial like family member 2	protein-coding
MCC	0.0321892518530266	2.18	chr5	113361801	113362152	-67061	NM_002387	mutated in colorectal cancers	protein-coding
MIR8066	0.0322426261722207	3.96	chr4	101131046	101131214	109742	NR_107033	microRNA 8066	ncRNA
LOC101927100	0.0322815284921029	3.3	chr5	115586465	115586685	-15580	NR_109875	uncharacterized LOC101927100	ncRNA
LINC00924	0.0322839448658582	2.08	chr15	95494141	95494417	61186	NR_027132	long intergenic non-protein coding RNA 924	ncRNA
NAB1	0.0325276701904504	2.67	chr2	190722306	190722498	63244	NM_001321314	NGFI-A binding protein 1	protein-coding
HIP1	0.0325777670497985	2.35	chr7	75756154	75756326	-17268	NM_005338	huntingtin interacting protein 1	protein-coding
IMPAD1	0.0326085191782472	-3.06	chr8	56973176	56973376	20595	NM_017813	inositol monophosphatase domain containing 1	protein-coding
CIB4	0.0326557338778325	2.32	chr2	26635006	26635175	6278	NM_001029881	calcium and integrin binding family member 4	protein-coding
MTO1	0.0327172490586195	-2.69	chr6	73462842	73463068	1224	NM_012123	mitochondrial tRNA translation optimization 1	protein-coding
HSPA12A	0.0327320965044276	1.91	chr10	116771445	116771775	-29036	NM_025015	heat shock protein family A (Hsp70) member 12A	protein-coding
OCSTAMP	0.0328538798307548	2.22	chr20	46579045	46579271	-28584	NM_080721	osteoclast stimulatory transmembrane protein	protein-coding
RNF4	0.0328671326426199	2	chr4	2510735	2510913	41371	NM_001185009	ring finger protein 4	protein-coding
LOC339593	0.0328802831128177	2.57	chr20	11274323	11274502	-1029	NR_038972	uncharacterized LOC339593	ncRNA
MTUS2	0.0328964687601685	2.46	chr13	29294800	29294980	-133750	NM_015233	microtubule associated tumor suppressor candidate 2	protein-coding
RNF185	0.0329563289144622	-2.54	chr22	31177680	31177939	17657	NM_001135825	ring finger protein 185	protein-coding
MMP16	0.0329883185988913	2.11	chr8	88773136	88773288	-445724	NM_005941	matrix metallopeptidase 16	protein-coding
TMEM261	0.0329965825543689	2.39	chr9	8258800	8258998	-459093	NM_033428	transmembrane protein 261	protein-coding
SIK2	0.0330108235291621	2.72	chr11	111574362	111574692	-27864	NM_015191	salt inducible kinase 2	protein-coding
FGF14	0.0330154062550158	3.06	chr13	102275962	102276159	34626	NM_001321933	fibroblast growth factor 14	protein-coding
CTSO	0.0330479825621077	3.37	chr4	156113321	156113508	-159518	NM_001334	cathepsin O	protein-coding
MIR1973	0.0330780339216086	2.63	chr4	116243406	116243584	-56230	NR_031737	microRNA 1973	ncRNA
HNRNPKP3	0.0330849664366904	1.99	chr11	43157658	43157870	111605	NR_033868	heterogeneous nuclear ribonucleoprotein K pseudogene 3	pseudo
LINC01929	0.03317002317406	-2.46	chr18	55207056	55207220	-82832	NR_110743	long intergenic non-protein coding RNA 1929	ncRNA
CXCL12	0.0332478765721383	1.93	chr10	44575850	44576108	-190882	NM_001277990	C-X-C motif chemokine ligand 12	protein-coding
SUCNR1	0.0332675285574934	2.11	chr3	151845505	151845678	-28052	NM_033050	succinate receptor 1	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
LINC02085	0.0333701518306063	3.16	chr3	101928995	101929166	-11779	NR_026934	long intergenic non-protein coding RNA 2085	ncRNA
RTKN2	0.0333842891745991	2.52	chr10	62243663	62243864	25100	NM_145307	rhotekin 2	protein-coding
PLD5	0.0334799489789895	2.87	chr1	242349042	242349244	100339	NM_001195812	phospholipase D family member 5	protein-coding
MIR3170	0.0335047668352359	2.26	chr13	98201351	98201511	-7093	NR_036129	microRNA 3170	ncRNA
LOC102725080	0.0335118636990632	2.99	chr8	13727842	13728091	98276	NR_134450	uncharacterized LOC102725080	ncRNA
IMPA1	0.0335817784531888	-2.81	chr8	81668959	81669134	17308	NM_001144878	inositol monophosphatase 1	protein-coding
IL15	0.0336104327760157	2.41	chr4	141973989	141974174	337485	NM_000585	interleukin 15	protein-coding
DPP6	0.0336141653221128	-2.21	chr7	153669510	153669670	-217507	NM_001039350	dipeptidyl peptidase like 6	protein-coding
LINC00434	0.0336549697364631	2.57	chr13	60324053	60324244	-56044	NR_047022	long intergenic non-protein coding RNA 434	ncRNA
ARG1	0.0338127020444913	2.15	chr6	131473433	131473782	-99597	NM_000045	arginase 1	protein-coding
RND3	0.0338507941680356	-2.09	chr2	150175155	150175525	312355	NM_005168	Rho family GTPase 3	protein-coding
GIMD1	0.0338807251841161	1.89	chr4	106385021	106385446	-17798	NM_001195138	GIMAP family P-loop NTPase domain containing 1	protein-coding
COL6A3	0.0339081567897682	2.54	chr2	237393999	237394186	20115	NM_057165	collagen type VI alpha 3 chain	protein-coding
MIR7157	0.0339438269783867	2.94	chr2	141085724	141085966	-499160	NR_106979	microRNA 7157	ncRNA
VPS4B	0.0340185198666255	2.37	chr18	63397548	63397814	24838	NM_004869	vacuolar protein sorting 4 homolog B	protein-coding
MIR302F	0.0340644882546439	1.98	chr18	30281049	30281290	-17741	NR_031684	microRNA 302f	ncRNA
C1QTNF3-AMACR	0.0341516729309403	2.4	chr5	34099996	34100194	24433	NR_037951	C1QTNF3-AMACR readthrough (NMD candidate)	ncRNA
FAM46A	0.0341927956140755	-3.06	chr6	81294287	81294467	458334	NM_017633	family with sequence similarity 46 member A	protein-coding
OR56A1	0.0343659398305769	-1.86	chr11	6062651	6063076	-35122	NM_001001917	olfactory receptor family 56 subfamily A member 1	protein-coding
NFIB	0.0343836843341912	-2.98	chr9	14272876	14273047	41085	NM_005596	nuclear factor I B	protein-coding
ANO5	0.0343918103452709	3.34	chr11	21990443	21990616	-202647	NM_001142649	anoctamin 5	protein-coding
TAGAP	0.0344550473612269	2.42	chr6	159035967	159036139	6300	NM_001278733	T-cell activation RhoGTPase activating protein	protein-coding
GOLGA2P2Y	0.0344684744647147	-2.58	chrY	25356799	25356977	-98423	NR_001555	golgin A2 pseudogene 2, Y-linked	pseudo
FRG1CP	0.0345350292548028	1.81	chr20	28841344	28841575	-238794	NR_132315	FSHD region gene 1 family member C, pseudogene	pseudo
ZNF717	0.0345672041964786	2.06	chr3	76108078	76108395	-322653	NM_001324026	zinc finger protein 717	protein-coding
LNCPRESS2	0.0345917842356942	2.2	chr4	92182391	92182579	94817	NR_125920	lncRNA p53 regulated and ESC associated 2	ncRNA
HTR1B	0.0346442396768775	2.57	chr6	76823576	76823899	640285	NM_000863	5-hydroxytryptamine receptor 1B	protein-coding
CYP2C19	0.0346886074294003	2.36	chr10	94813062	94813232	50466	NM_000769	cytochrome P450 family 2 subfamily C member 19	protein-coding
LINC00293	0.0346986387484376	2.12	chr8	46572897	46573269	-267803	NR_027012	long intergenic non-protein coding RNA 293	ncRNA
SCG2	0.0347525317862525	2.81	chr2	223602453	223602636	-45	NM_003469	secretogranin II	protein-coding
SYN2	0.0347595123157148	2.01	chr3	11929162	11929425	-75067	NM_133625	synapsin II	protein-coding
DIAPH3-AS1	0.034761472852552	2.52	chr13	59501388	59501585	-511232	NR_109838	DIAPH3 antisense RNA 1	ncRNA
MIR2054	0.0347820681120143	2.5	chr4	125679922	125680195	172799	NR_031746	microRNA 2054	ncRNA
MTSS1	0.0348304792499528	2.12	chr8	124688088	124688332	40297	NM_014751	MTSS1, I-BAR domain containing	protein-coding
PTN	0.0348682640867885	2.13	chr7	137585800	137585971	-241895	NM_001321387	pleiotrophin	protein-coding
ZNF665	0.0349356147073514	-1.9	chr19	53192457	53193059	608	NM_024733	zinc finger protein 665	protein-coding
SORCS3	0.0349639698323761	2.39	chr10	104616183	104616371	-24824	NM_014978	sortilin related VPS10 domain containing receptor 3	protein-coding
PRKAG2	0.035005873021284	2.1	chr7	151612468	151612794	19627	NM_024429	protein kinase AMP-activated non-catalytic subunit gamma 2	protein-coding
CAPS2	0.0350795411978324	2.44	chr12	75305473	75305640	24500	NM_032606	calcyphosine 2	protein-coding
MAGED4B	0.0351384460947644	2.62	chrX	52068846	52069059	320	NM_030801	MAGE family member D4B	protein-coding
TDO2	0.0352608134904907	2.25	chr4	155909626	155909851	6045	NM_005651	tryptophan 2,3-dioxygenase	protein-coding
KIF17	0.0352643572708023	2.13	chr1	20702770	20702932	15166	NM_001122819	kinesin family member 17	protein-coding
GNPDA2	0.0353259376067387	1.95	chr4	44823480	44823824	-97018	NR_073094	glucosamine-6-phosphate deaminase 2	protein-coding
LOC645166	0.0353674501544203	-2.55	chr1	143498415	143498700	-73873	NR_027356	lymphocyte-specific protein 1 pseudogene	pseudo
ZCCHC6	0.0353787289476442	-2.55	chr9	86336149	86336332	18210	NM_001330718	zinc finger CCHC-type containing 6	protein-coding
SCN3B	0.0354121125282058	1.98	chr11	123653853	123654073	644	NM_001040151	sodium voltage-gated channel beta subunit 3	protein-coding
SOX1	0.0354555077262284	2.19	chr13	112062899	112063169	-4565	NM_005986	SRY-box 1	protein-coding
LOC101927849	0.0355204694020151	-2.77	chr4	149443055	149443282	288873	NR_125887	uncharacterized LOC101927849	ncRNA
NFASC	0.0355627440600223	3.07	chr1	204842306	204842503	13750	NM_001005388	neurofascin	protein-coding
C5orf47	0.0355776049380698	1.91	chr5	173989356	173989535	286	NM_001144954	chromosome 5 open reading frame 47	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
DRD1	0.0356926346191885	2.25	chr5	175460028	175460288	-15998	NM_000794	dopamine receptor D1	protein-coding
LOC101926942	0.0357558148105276	2.98	chr10	90509520	90509704	31193	NR_110657	uncharacterized LOC101926942	ncRNA
GPR176	0.0358228408852429	1.93	chr15	39893882	39894402	26750	NM_001271854	G protein-coupled receptor 176	protein-coding
KCNE4	0.0358262206170769	-3.44	chr2	223324223	223324398	272380	NM_080671	potassium voltage-gated channel subfamily E regulatory subunit 4	protein-coding
LOC100131496	0.0358421473286647	1.86	chr20	47255018	47255261	-63363	NR_024594	uncharacterized LOC100131496	ncRNA
ETS1	0.0359007043732651	2	chr11	128604015	128604177	-16538	NM_001143820	ETS proto-oncogene 1, transcription factor	protein-coding
SLC17A6	0.0359965356994991	-2.7	chr11	22334032	22334301	-3955	NM_020346	solute carrier family 17 member 6	protein-coding
CSMD3	0.0360176780496245	1.83	chr8	113432333	113432827	4433	NM_052900	CUB and Sushi multiple domains 3	protein-coding
LOC101927915	0.0360623547029463	1.99	chr8	137154108	137154448	259310	NR_125428	uncharacterized LOC101927915	ncRNA
LINC01252	0.0360860867797433	1.96	chr12	11500849	11501005	-47103	NR_033890	long intergenic non-protein coding RNA 1252	ncRNA
DLGAP1-AS4	0.036088446293948	-3.1	chr18	3998724	3998909	36463	NR_102696	DLGAP1 antisense RNA 4	ncRNA
GALNT1	0.0360902520071647	2.99	chr18	35755210	35755375	100723	NM_020474	polypeptide N-acetylgalactosaminyltransferase 1	protein-coding
MIR7641-2	0.0361221856930233	2.36	chr5	34771499	34771685	-30726	NR_106988	microRNA 7641-2	ncRNA
LINC01207	0.0361361082185763	2.12	chr4	164667021	164667660	-86791	NR_038834	long intergenic non-protein coding RNA 1207	ncRNA
LOC100133920	0.0361690643310237	2.56	chr2	93830119	93830448	756900	NR_024443	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like pseudogene	pseudo
CTNNA3	0.0361770615108817	2.81	chr10	67624925	67625120	40663	NM_001291133	catenin alpha 3	protein-coding
AHCTF1P1	0.0362508366647757	3.05	chr2	161506657	161506822	1163	NR_077058	AT-hook containing transcription factor 1 pseudogene 1	pseudo
MIR147A	0.0363148366673969	2.77	chr9	120326989	120327179	-82034	NR_029604	microRNA 147a	ncRNA
MIR4643	0.0363775858193175	2.23	chr6	91437300	91437537	-84242	NR_039786	microRNA 4643	ncRNA
UCK2	0.0363911880651173	2.83	chr1	165834905	165835164	7539	NM_012474	uridine-cytidine kinase 2	protein-coding
FAM170A	0.0364895557639526	1.95	chr5	119542854	119543372	-86446	NM_001163991	family with sequence similarity 170 member A	protein-coding
RANBP3L	0.0365432278415808	2.67	chr5	36272323	36272503	29489	NM_001323280	RAN binding protein 3 like	protein-coding
LOC102467222	0.0365601363459264	2.72	chr8	20765667	20765859	-187870	NR_130774	uncharacterized LOC102467222	ncRNA
EMBP1	0.0366379878967718	1.94	chr1	125162266	125162538	3643290	NR_003955	embigin pseudogene 1	pseudo
LINC00200	0.0366788918978982	-2.17	chr10	1141004	1141345	-18594	NR_015376	long intergenic non-protein coding RNA 200	ncRNA
MIR4641	0.0367065644475402	1.85	chr6	41596532	41596776	-2069	NR_039784	microRNA 4641	ncRNA
SNORA87	0.0367343681289231	-2.5	chr10	113284228	113284396	238957	NR_132769	small nucleolar RNA, H/ACA box 87	snoRNA
CCNY	0.0367364343858981	2.46	chr10	35235554	35235704	-11352	NM_181698	cyclin Y	protein-coding
LOC101927269	0.0367605785037324	1.98	chr7	80373351	80373733	897	NR_110075	uncharacterized LOC101927269	ncRNA
ZAN	0.0367836881313405	2.15	chr7	100765044	100765210	31501	NR_111918	zonadhesin (gene/pseudogene)	protein-coding
ZNF716	0.0368760635014537	2.25	chr7	58663028	58663387	1213030	NM_001159279	zinc finger protein 716	protein-coding
GLYATL1	0.0369221542370065	2.84	chr11	58977439	58977675	34308	NM_080661	glycine-N-acyltransferase like 1	protein-coding
LAMC1	0.0369334533044305	2.59	chr1	183098247	183098505	74916	NM_002293	laminin subunit gamma 1	protein-coding
CACNA1A	0.0369501612995927	2.01	chr19	13383608	13383757	122778	NM_001174080	calcium voltage-gated channel subunit alpha1 A	protein-coding
OR7E156P	0.0369646935807147	2.17	chr13	63636405	63636889	-100788	NR_002171	olfactory receptor family 7 subfamily E member 156 pseudogene	pseudo
ACTR3BP5	0.0369682020645699	2.1	chr10	39998914	39999097	1302409	NR_045000	ACTR3B pseudogene 5	pseudo
EPHA7	0.0370051767807108	2.93	chr6	93538791	93539031	-119329	NM_004440	EPH receptor A7	protein-coding
MAT2B	0.0370394678370107	2.75	chr5	163774439	163774625	269004	NM_013283	methionine adenosyltransferase 2B	protein-coding
IQCJ	0.0370401038786572	2.4	chr3	159090568	159090756	21410	NM_001197100	IQ motif containing J	protein-coding
METRNL	0.0370568516969694	2.28	chr17	83083322	83083502	3721	NM_001004431	meteorin like, glial cell differentiation regulator	protein-coding
MIR3152	0.0370720155820744	-3.07	chr9	18800442	18800629	227229	NR_036107	microRNA 3152	ncRNA
NPY5R	0.037102446631943	2.54	chr4	163396022	163396196	52007	NM_001317091	neuropeptide Y receptor Y5	protein-coding
LOC644285	0.0372375744705325	2.94	chr5	87348844	87349008	-25298	NR_130929	uncharacterized LOC644285	ncRNA
TMCO5A	0.037276985459289	2.03	chr15	37891979	37892162	-42537	NM_152453	transmembrane and coiled-coil domains 5A	protein-coding
ARHGAP42	0.0373135200902935	2.01	chr11	100470117	100470331	-217452	NM_152432	Rho GTPase activating protein 42	protein-coding
KIF26B	0.0373185197941691	3.49	chr1	245757364	245757523	602458	NM_018012	kinesin family member 26B	protein-coding
LINC01847	0.0373219742791572	2.13	chr5	159872565	159872728	-1262	NR_109891	long intergenic non-protein coding RNA 1847	ncRNA
LOC100507534	0.0373260005949633	2.01	chr16	47904542	47904753	-17517	NR_110649	uncharacterized LOC100507534	ncRNA
LOC101929470	0.0373541952741261	1.97	chr8	29804563	29804777	23834	NR_125817	uncharacterized LOC101929470	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
LOC441666	0.0373572584892878	2.53	chr10	41300950	41301155	1066993	NR_024380	zinc finger protein 91 pseudogene	pseudo
LOC389705	0.0373598679332309	3.27	chr9	14954971	14955175	-38254	NR_003920	chromosome 4 open reading frame 27 pseudogene	pseudo
ANKRD20A5P	0.0374112893356891	2.1	chr18	14225409	14225845	46530	NR_040113	ankyrin repeat domain 20 family member A5, pseudogene	pseudo
SYT10	0.0374211859102463	2.24	chr12	33439290	33439484	432	NM_198992	synaptotagmin 10	protein-coding
CYP7B1	0.0374557123414106	1.84	chr8	64772193	64772405	26492	NM_001324112	cytochrome P450 family 7 subfamily B member 1	protein-coding
LILRA5	0.0374597674046822	2.19	chr19	54316325	54316496	-3271	NM_181985	leukocyte immunoglobulin like receptor A5	protein-coding
PLOD2	0.0375601132962455	2.23	chr3	146072007	146072230	89377	NM_182943	procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	protein-coding
LNPEP	0.0375649780617276	2.56	chr5	96984733	96985025	26427	NM_175920	leucyl and cystinyl aminopeptidase	protein-coding
NPAS2	0.0375868653235714	-2.13	chr2	100715402	100715670	-104615	NM_002518	neuronal PAS domain protein 2	protein-coding
LINC00273	0.0375939345324801	2.26	chr16	34619961	34620189	-460039	NR_038368	long intergenic non-protein coding RNA 273	ncRNA
PSAT1	0.0376383423357514	-1.94	chr9	78369606	78369869	72662	NM_058179	phosphoserine aminotransferase 1	protein-coding
LINC00907	0.0376843187987857	-2.01	chr18	42392600	42392769	206016	NR_046456	long intergenic non-protein coding RNA 907	ncRNA
LINC02120	0.0376896112844809	2.52	chr5	32831713	32832045	-115564	NR_033832	long intergenic non-protein coding RNA 2120	ncRNA
LOC101928510	0.0377403499186254	2.18	chr11	36159965	36160142	-82152	NR_135064	uncharacterized LOC101928510	ncRNA
LHPP	0.0377502712035056	-2.15	chr10	124478033	124478226	16357	NM_001318332	phospholysine phosphohistidine inorganic pyrophosphate phosphatase	protein-coding
RUFY1	0.0377742247645531	1.96	chr5	179523153	179523326	-27322	NM_025158	RUN and FYVE domain containing 1	protein-coding
LINC00922	0.0377968706488789	2.34	chr16	65562816	65563011	13387	NR_027755	long intergenic non-protein coding RNA 922	ncRNA
WNT16	0.0378019982312564	2.25	chr7	121327641	121327911	-1260	NM_057168	Wnt family member 16	protein-coding
PON1	0.0378278388175169	1.9	chr7	95311544	95311949	12826	NM_000446	paraoxonase 1	protein-coding
TACR3	0.0378747569842752	1.83	chr4	103934156	103934416	-214470	NM_001059	tachykinin receptor 3	protein-coding
RNGTT	0.0379682390791618	2.42	chr6	88828086	88828239	135467	NM_001286428	RNA guanylyltransferase and 5'-phosphatase	protein-coding
CYP4Z1	0.0379944162226157	2.29	chr1	47080873	47081074	13485	NM_178134	cytochrome P450 family 4 subfamily Z member 1	protein-coding
SLC16A7	0.0380823289647928	2.85	chr12	59724280	59724470	35038	NR_073055	solute carrier family 16 member 7	protein-coding
LRFN2	0.0380936005503995	-2.18	chr6	40550422	40550596	36955	NM_020737	leucine rich repeat and fibronectin type III domain containing 2	protein-coding
CST9L	0.0381353370201198	3.01	chr20	23575719	23575899	-7060	NM_080610	cystatin 9-like	protein-coding
LOC100133920	0.0381897967593	2.02	chr2	93954901	93955054	632206	NR_024443	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like pseudogene	pseudo
FGF8	0.0382001201438895	2.13	chr10	101779735	101779919	542	NM_001206389	fibroblast growth factor 8	protein-coding
LINC02126	0.0382569767075092	2.16	chr16	65334689	65334881	-158212	NR_110918	long intergenic non-protein coding RNA 2126	ncRNA
LOC105373338	0.0382653411874088	-1.89	chrX	130676792	130677072	-125704	NR_135655	uncharacterized LOC105373338	ncRNA
FAM95C	0.038275477422055	2.05	chr9	38555126	38555343	-9862	NR_047651	family with sequence similarity 95 member C	ncRNA
DIRC3	0.0383350951466315	2.07	chr2	217756478	217756844	-68	NR_026597	disrupted in renal carcinoma 3	ncRNA
LOC101927050	0.038381487369629	2.2	chr2	90317223	90317441	-679425	NR_136329	lysine methyltransferase 2C pseudogene	pseudo
BMP3	0.0383866805193993	-3.22	chr4	80744632	80744883	-286208	NM_001201	bone morphogenetic protein 3	protein-coding
NXPH2	0.0384583419309674	2.65	chr2	138646793	138646948	133371	NM_007226	neurexophilin 2	protein-coding
LOC100507377	0.0385041718797124	2.94	chr12	74370829	74370992	-78279	NR_038300	uncharacterized LOC100507377	ncRNA
ROBO2	0.0385056297522938	3.08	chr3	76426242	76426414	-613815	NM_002942	roundabout guidance receptor 2	protein-coding
GFRA1	0.0385090989355762	3.05	chr10	116254656	116254825	17567	NM_145793	GDNF family receptor alpha 1	protein-coding
TMED4	0.0385350956826798	2.11	chr7	44588309	44588502	-6110	NM_182547	transmembrane p24 trafficking protein 4	protein-coding
LOC100506422	0.0385380299095823	-3.31	chr9	25962350	25962502	-104249	NM_001004352	putative deoxyuridine 5'-triphosphate nucleotidohydrolase-like protein FLJ16323	protein-coding
LINC00423	0.0386875551455923	2.24	chr13	32819561	32819756	91994	NR_047020	long intergenic non-protein coding RNA 423	ncRNA
LINC00593	0.0387362722036605	2.14	chr15	69853494	69853653	18339	NR_026764	long intergenic non-protein coding RNA 593	ncRNA
SLC8A1-AS1	0.0387515206724167	1.84	chr2	39960845	39961007	43292	NR_038441	SLC8A1 antisense RNA 1	ncRNA
FRAS1	0.0387642822856711	3.87	chr4	78100613	78100773	43123	NM_025074	Fraser extracellular matrix complex subunit 1	protein-coding
NOVA1	0.0387754584469196	2.35	chr14	26010150	26010563	587398	NM_006491	NOVA alternative splicing regulator 1	protein-coding
BOLA3	0.0388080362816116	1.87	chr2	74100838	74101072	46957	NM_001035505	bolA family member 3	protein-coding
C8orf87	0.0388419576744212	2.95	chr8	93348736	93348894	-181965	NM_001242668	chromosome 8 open reading frame 87	protein-coding
STN1	0.0388489170631807	1.91	chr10	103912499	103912812	5632	NM_024928	STN1, CST complex subunit	protein-coding
TGFBR2	0.0389362892868413	-2.79	chr3	30440108	30440494	-166201	NM_003242	transforming growth factor beta receptor 2	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
XIAP	0.0389604246351982	2.04	chrX	123833598	123833820	-26103	NM_001204401	X-linked inhibitor of apoptosis	protein-coding
NT5C1B	0.0389629794176342	2.96	chr2	18781690	18781965	-192247	NM_001199088	5'-nucleotidase, cytosolic IB	protein-coding
LINC01372	0.0389870336189238	1.94	chr7	67643815	67644039	307951	NR_108104	long intergenic non-protein coding RNA 1372	ncRNA
LINC01526	0.0389985299975089	-1.82	chr6	81820817	81821341	-6887	NR_126363	long intergenic non-protein coding RNA 1526	ncRNA
TYRP1	0.0390000822269068	2.37	chr9	11839988	11840178	-853303	NM_000550	tyrosinase related protein 1	protein-coding
LUCAT1	0.0390096718293034	2.01	chr5	91028770	91028951	285542	NR_103549	lung cancer associated transcript 1 (non-protein coding)	ncRNA
UGT8	0.0390141237834773	2.5	chr4	114615681	114615879	-6587	NM_003360	UDP glycosyltransferase 8	protein-coding
XIRP2	0.0390210510143754	3.22	chr2	167265617	167265831	78441	NM_001199144	xin actin binding repeat containing 2	protein-coding
MIR4459	0.0391405120783677	2.25	chr5	54114826	54115048	-39354	NR_039664	microRNA 4459	ncRNA
MIS18BP1	0.0391876280416239	-3.14	chr14	45281723	45281898	-28408	NM_018353	MIS18 binding protein 1	protein-coding
DCLK2	0.0392151195876978	3.3	chr4	150004521	150004733	-73647	NM_001040260	doublecortin like kinase 2	protein-coding
PTBP2	0.0392915052673732	-2.52	chr1	96792128	96792521	69686	NR_125356	polypyrimidine tract binding protein 2	protein-coding
GPC6-AS2	0.0394271072109638	2	chr13	93799943	93800557	58223	NR_046536	GPC6 antisense RNA 2	ncRNA
CTNNA2	0.0395028023797906	2.79	chr2	80709613	80709774	396314	NM_001282599	catenin alpha 2	protein-coding
HDDC2	0.0395131015775039	1.81	chr6	125279982	125280209	22041	NM_016063	HD domain containing 2	protein-coding
LINC01821	0.0395626790581659	2.46	chr2	193603912	193604093	-740267	NR_110222	long intergenic non-protein coding RNA 1821	ncRNA
DNAH6	0.0395979506615304	1.83	chr2	84585353	84585531	68987	NM_001370	dynein axonemal heavy chain 6	protein-coding
WDR18	0.0396230056226339	2.09	chr19	958878	959042	-25368	NM_024100	WD repeat domain 18	protein-coding
NMBR	0.0396510754614315	2.66	chr6	141563234	141563460	525452	NM_002511	neuromedin B receptor	protein-coding
NXPH1	0.0396627673123655	2.47	chr7	8365707	8365893	-68155	NM_152745	neurexophilin 1	protein-coding
C11orf49	0.0396807885769897	2.54	chr11	46978147	46978378	41573	NM_001003676	chromosome 11 open reading frame 49	protein-coding
LOC90768	0.0397035845704173	2.91	chr4	181668830	181669048	475576	NR_027107	uncharacterized LOC90768	ncRNA
LOC101927691	0.0397342488348592	2.13	chr6	819004	819159	-49417	NR_110853	uncharacterized LOC101927691	ncRNA
PTPRD	0.0397731621605636	-3.01	chr9	10232299	10232507	380320	NM_002839	protein tyrosine phosphatase, receptor type D	protein-coding
DPYS	0.0398192514903356	1.92	chr8	104417631	104417799	49334	NM_001385	dihydropyrimidinase	protein-coding
LOC100506675	0.0398537994923635	1.86	chr11	31939157	31939362	102059	NR_120526	uncharacterized LOC100506675	ncRNA
LINC02120	0.0399530353580028	2.28	chr5	33098192	33098423	150864	NR_033832	long intergenic non-protein coding RNA 2120	ncRNA
TMEM71	0.0399890280877086	3.13	chr8	132753669	132753825	6921	NM_144649	transmembrane protein 71	protein-coding
MIR4799	0.0400096665134576	2.03	chr4	148015333	148015608	232875	NR_039962	microRNA 4799	ncRNA
LINC01324	0.0400447104567176	2.02	chr3	164805991	164806338	25316	NR_126405	long intergenic non-protein coding RNA 1324	ncRNA
SPEF2	0.0400448937896093	-1.96	chr5	35683351	35683776	65676	NM_024867	sperm flagellar 2	protein-coding
USP12-AS1	0.0401473964679615	2.87	chr13	27079691	27079870	-45753	NR_046547	USP12 antisense RNA 1	ncRNA
KCNMA1	0.0401728591462592	-2.25	chr10	77601492	77601778	36184	NM_001271522	potassium calcium-activated channel subfamily M alpha 1	protein-coding
IL1RAP	0.0401979544401336	2.35	chr3	190543632	190543848	29689	NM_134470	interleukin 1 receptor accessory protein	protein-coding
GBE1	0.0402205393027501	3.6	chr3	83160715	83160910	-1399013	NM_000158	1,4-alpha-glucan branching enzyme 1	protein-coding
PLCL2	0.0402245726559737	2.45	chr3	16808910	16809098	-75955	NM_001144382	phospholipase C like 2	protein-coding
LINC01243	0.0402489608602052	-3	chr9	31304022	31304216	77371	NR_135134	long intergenic non-protein coding RNA 1243	ncRNA
NRSN1	0.0402587545121877	-2.13	chr6	23587596	23588047	-538365	NM_080723	neurensin 1	protein-coding
AASDH	0.0402982565529359	2.55	chr4	56395137	56395289	-7705	NM_001286672	aminoadipate-semialdehyde dehydrogenase	protein-coding
LINC02059	0.0404237646443829	2.3	chr5	86917182	86917444	170495	NR_105017	long intergenic non-protein coding RNA 2059	ncRNA
CACNA1I	0.0404359610139016	2.14	chr22	39557590	39557761	-13078	NM_021096	calcium voltage-gated channel subunit alpha1 I	protein-coding
MET	0.040440232036294	2.2	chr7	116665575	116665849	-6647	NM_001324402	MET proto-oncogene, receptor tyrosine kinase	protein-coding
CNTN6	0.0404713612249507	2.13	chr3	1034552	1034785	-57990	NM_001289081	contactin 6	protein-coding
DHFRP3	0.0405284562522674	3.79	chr2	83320927	83321105	464213	NR_033423	dihydrofolate reductase pseudogene 3	pseudo
MIR4465	0.040559630056024	2.24	chr6	141253076	141253523	569485	NR_039675	microRNA 4465	ncRNA
TBC1D14	0.0405740624907883	2.68	chr4	6859968	6860167	-49377	NM_001113361	TBC1 domain family member 14	protein-coding
NBPF8	0.0406098100988267	2.41	chr1	120491289	120491553	55070	NR_102404	neuroblastoma breakpoint family member 8	protein-coding
TIPARP	0.0406355446934158	-2.94	chr3	156703799	156703966	27218	NM_001184718	TCDD inducible poly(ADP-ribose) polymerase	protein-coding
C1orf198	0.040636865543612	-2.7	chr1	230844535	230844719	11409	NM_001136495	chromosome 1 open reading frame 198	protein-coding
TSPAN12	0.0406562790051395	-2.69	chr7	120845861	120846099	12143	NM_012338	tetraspanin 12	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
MAPRE1	0.0407215884948744	2.95	chr20	32855118	32855264	35298	NM_012325	microtubule associated protein RP/EB family member 1	protein-coding
SP4	0.0408077064613589	1.98	chr7	21319444	21319612	-108506	NM_001326542	Sp4 transcription factor	protein-coding
INPP4B	0.040814236019436	2.48	chr4	142983651	142983842	-137211	NM_001101669	inositol polyphosphate-4-phosphatase type II B	protein-coding
HTR2A	0.0408613961533424	2.88	chr13	47311989	47312189	-415013	NM_000621	5-hydroxytryptamine receptor 2A	protein-coding
CXCL8	0.0408725299868854	2.06	chr4	73773672	73773983	33321	NM_000584	C-X-C motif chemokine ligand 8	protein-coding
LINC00609	0.0408897604888748	2.1	chr14	36014198	36014384	-56136	NR_073454	long intergenic non-protein coding RNA 609	ncRNA
CNTNAP5	0.0408989500345662	2.3	chr2	123944437	123944642	-80748	NM_130773	contactin associated protein like 5	protein-coding
LINC00491	0.0409401635828057	1.92	chr5	102693237	102693550	-21929	NR_103756	long intergenic non-protein coding RNA 491	ncRNA
GLUL	0.0409556412047467	-2.21	chr1	182379827	182380234	11374	NM_001033056	glutamate-ammonia ligase	protein-coding
LINC00363	0.040958710566642	2.56	chr13	93088959	93089176	43726	NR_126360	long intergenic non-protein coding RNA 363	ncRNA
ATP5G3	0.0409731248604326	2.48	chr2	175282182	175282503	-100580	NM_001689	ATP synthase, H+ transporting, mitochondrial Fo complex subunit C3 (subunit 9)	protein-coding
XKR9	0.0410234175398491	2.25	chr8	70787533	70788117	118378	NM_001287260	XK related 9	protein-coding
LOC644919	0.0410396559352229	2.78	chr14	40411994	40412154	-542637	NR_109758	uncharacterized LOC644919	ncRNA
CCT6B	0.0410776944338768	1.92	chr17	34909523	34909719	51888	NM_001193529	chaperonin containing TCP1 subunit 6B	protein-coding
BMPER	0.0411061155450276	1.93	chr7	33998233	33998625	93518	NM_133468	BMP binding endothelial regulator	protein-coding
LOC643542	0.0411065708395185	3.27	chr18	67478144	67478368	-38290	NR_033921	uncharacterized LOC643542	ncRNA
LINC00273	0.0411479316075811	2.28	chr16	34618534	34618786	-458624	NR_038368	long intergenic non-protein coding RNA 273	ncRNA
LINC02140	0.0411965727614414	-1.84	chr16	54459727	54460031	-89180	NR_134923	long intergenic non-protein coding RNA 2140	ncRNA
RGS17	0.0412228170227161	2.32	chr6	153073691	153073853	57482	NM_012419	regulator of G-protein signaling 17	protein-coding
RAI14	0.0412510191888866	1.99	chr5	34695953	34696266	8550	NM_001145523	retinoic acid induced 14	protein-coding
LOC644669	0.0412910087439811	2.32	chr18	15693238	15693398	-367399	NR_027417	ankyrin repeat domain 30B pseudogene	pseudo
OR6P1	0.0412956975902522	3.26	chr1	158569222	158569385	-5699	NM_001160325	olfactory receptor family 6 subfamily P member 1	protein-coding
NPY	0.0413155204181811	2.25	chr7	24382032	24382222	97939	NM_000905	neuropeptide Y	protein-coding
FAM193A	0.0413225927660258	2.43	chr4	2670713	2670885	45367	NR_046335	family with sequence similarity 193 member A	protein-coding
RRN3P2	0.0413272695716763	-2.54	chr16	29080211	29080373	5450	NR_003369	RRN3 homolog, RNA polymerase I transcription factor pseudogene 2	pseudo
LOC105370369	0.0413785189060694	-2.38	chr13	111596374	111596548	452	NR_135814	uncharacterized LOC105370369	ncRNA
LOC105369509	0.0413923589303491	2.42	chr11	115226547	115226866	-106871	NR_135108	uncharacterized LOC105369509	ncRNA
ARMC2-AS1	0.0414300957341301	1.87	chr6	108957591	108957782	-33583	NR_104137	ARMC2 antisense RNA 1	ncRNA
MIR6888	0.0415013189367252	2.08	chr2	159169944	159170105	-16811	NR_106948	microRNA 6888	ncRNA
LOC101929721	0.0415801883937104	2.52	chr1	49353457	49353662	96148	NR_125988	uncharacterized LOC101929721	ncRNA
MRC1	0.0416606665768715	1.85	chr10	17794250	17794615	-14911	NM_002438	mannose receptor C-type 1	protein-coding
MTRNR2L1	0.0416745201081715	2.3	chr17	22940939	22941097	417907	NM_001190452	MT-RNR2-like 1	protein-coding
LINC02032	0.0417327320445452	3.24	chr3	147890152	147890456	-187855	NR_134937	long intergenic non-protein coding RNA 2032	ncRNA
LINC01309	0.0417402516366022	2.41	chr13	104085628	104085781	660504	NR_126374	long intergenic non-protein coding RNA 1309	ncRNA
LOC101929411	0.0418794280020349	2.23	chr3	102077693	102078157	-85780	NR_135549	uncharacterized LOC101929411	ncRNA
LSM6	0.0418892489486083	2.45	chr4	146260782	146261101	85258	NM_007080	LSM6 homolog, U6 small nuclear RNA and mRNA degradation associated	protein-coding
MIR592	0.0419382181241316	1.93	chr7	126657713	126658017	400319	NR_030323	microRNA 592	ncRNA
LINC01372	0.042010293073596	3.28	chr7	67343132	67343299	7239	NR_108104	long intergenic non-protein coding RNA 1372	ncRNA
XIRP2	0.0420307260085427	2.33	chr2	167006276	167006452	117877	NM_152381	xin actin binding repeat containing 2	protein-coding
MIR3691	0.0420313249585624	-2.05	chr6	5140437	5141021	7593	NR_037462	microRNA 3691	ncRNA
RBM47	0.0420771941611607	-2.25	chr4	40688201	40688358	-58413	NM_001098634	RNA binding motif protein 47	protein-coding
PCDH7	0.0420890987425695	2.91	chr4	30232299	30232458	-488030	NM_032456	protocadherin 7	protein-coding
C5orf52	0.0420929617540282	1.82	chr5	157693741	157694029	22332	NM_001145132	chromosome 5 open reading frame 52	protein-coding
FASLG	0.0420964561516753	-2.58	chr1	172692044	172692289	33158	NM_000639	Fas ligand	protein-coding
LOC101928618	0.0421301010749877	-2.83	chr7	36027682	36027837	-67551	NR_108089	uncharacterized LOC101928618	ncRNA
SSX2IP	0.0421580388343612	3.45	chr1	84737315	84737489	-46845	NM_001166294	SSX family member 2 interacting protein	protein-coding
NRG1-IT3	0.0421889652652788	2.58	chr8	32459754	32459925	19093	NR_047475	NRG1 intronic transcript 3	ncRNA
LOC102467222	0.0421902616357302	2.38	chr8	20907199	20907378	-46345	NR_130774	uncharacterized LOC102467222	ncRNA
WSCD2	0.0422055759629546	2.7	chr12	108181996	108182160	50347	NM_001304447	WSC domain containing 2	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
LINC01582	0.0422268411034033	1.9	chr15	98180283	98180475	-91626	NR_120325	long intergenic non-protein coding RNA 1582	ncRNA
VSX2	0.042265696337913	2.23	chr14	74245915	74246086	6528	NM_182894	visual system homeobox 2	protein-coding
BCOR	0.0423802825191047	2.05	chrX	40176633	40176810	608	NM_001123384	BCL6 corepressor	protein-coding
LOC100130331	0.0424387345561715	2.03	chr1	237875377	237875681	13354	NR_027247	POTE ankyrin domain family, member F pseudogene	pseudo
A4GALT	0.0425437594806357	-2.59	chr22	42721091	42721242	135	NM_001318038	alpha 1,4-galactosyltransferase	protein-coding
VWCE	0.0425458191241164	2.36	chr11	61275930	61276104	19299	NM_152718	von Willebrand factor C and EGF domains	protein-coding
CNTLN	0.0425568370276126	2.32	chr9	17240718	17240937	105836	NM_017738	centlein	protein-coding
BBS7	0.0425647966799588	3.03	chr4	121898086	121898305	-27698	NM_176824	Bardet-Biedl syndrome 7	protein-coding
NFU1	0.0425988178808096	2.59	chr2	69471277	69471451	-33736	NM_015700	NFU1 iron-sulfur cluster scaffold	protein-coding
OR2S2	0.0428454153546079	-2.26	chr9	35945597	35945963	12374	NM_019897	olfactory receptor family 2 subfamily S member 2 (gene/pseudogene)	protein-coding
MEI4	0.0429480274861495	2.42	chr6	77739479	77739663	86517	NM_001322247	meiotic double-stranded break formation protein 4	protein-coding
PI4K2B	0.0429704355052642	1.89	chr4	25247240	25247403	13290	NR_144633	phosphatidylinositol 4-kinase type 2 beta	protein-coding
SELENOP	0.0431002725853659	-2.25	chr5	42858228	42858475	-46429	NM_001085486	selenoprotein P	protein-coding
FAM19A2	0.0432012291845876	1.98	chr12	61355765	61355915	836999	NM_178539	family with sequence similarity 19 member A2, C-C motif chemokine like	protein-coding
UBASH3A	0.0432157467888628	2.34	chr21	42407472	42407651	3699	NM_001001895	ubiquitin associated and SH3 domain containing A	protein-coding
LOC105374704	0.0433257172833789	2.15	chr5	30500779	30500964	620311	NR_134264	uncharacterized LOC105374704	ncRNA
PPARGC1A	0.0433290907351505	1.81	chr4	24192170	24192503	280493	NM_001330751	PPARG coactivator 1 alpha	protein-coding
LARP1B	0.0433361518482995	2.06	chr4	128101109	128101354	39965	NM_001278604	La ribonucleoprotein domain family member 1B	protein-coding
MIR5007	0.0433449258097705	2.84	chr13	55727647	55727831	553285	NR_049804	microRNA 5007	ncRNA
OR8S1	0.0433787596624216	2.18	chr12	48542462	48542739	16968	NM_001005203	olfactory receptor family 8 subfamily S member 1	protein-coding
LOC101927069	0.0434037935864989	2.78	chr9	68929550	68929704	46214	NR_110647	uncharacterized LOC101927069	ncRNA
EPAS1	0.0434121451614331	1.97	chr2	46142725	46142944	-154568	NM_001430	endothelial PAS domain protein 1	protein-coding
LOC101927087	0.0434190767684384	1.97	chr4	124374279	124374459	184065	NR_110838	uncharacterized LOC101927087	ncRNA
DIAPH3-AS1	0.0434304443633454	2.05	chr13	59318209	59318507	-694360	NR_109838	DIAPH3 antisense RNA 1	ncRNA
UBE2E1	0.0434306938045182	2.55	chr3	23762805	23762984	-42999	NM_003341	ubiquitin conjugating enzyme E2 E1	protein-coding
ZNF638	0.0434329762106003	1.86	chr2	71299669	71299869	-31986	NM_014497	zinc finger protein 638	protein-coding
SNORA70B	0.0435381847259003	2.44	chr2	61362310	61362475	54985	NR_003707	small nucleolar RNA, H/ACA box 70B	snoRNA
LINC01098	0.0436482717906001	2.63	chr4	177667628	177667855	-61016	NR_028342	long intergenic non-protein coding RNA 1098	ncRNA
MAP3K7	0.0436705256008471	2.18	chr6	90800979	90801181	-213779	NM_145331	mitogen-activated protein kinase kinase kinase 7	protein-coding
CHRNA9	0.0436716453524297	1.85	chr4	40405731	40405903	70488	NM_017581	cholinergic receptor nicotinic alpha 9 subunit	protein-coding
STPG2	0.0437840499163665	2.02	chr4	98125719	98125997	17382	NM_174952	sperm tail PG-rich repeat containing 2	protein-coding
TRIML1	0.0439008164980713	2.27	chr4	188242263	188242451	102916	NM_178556	tripartite motif family like 1	protein-coding
SLC1A3	0.0439149012301524	-2.32	chr5	36528821	36528981	-77454	NM_004172	solute carrier family 1 member 3	protein-coding
RRM1	0.0440733993282199	2.64	chr11	3999322	3999483	-95283	NM_001033	ribonucleotide reductase catalytic subunit M1	protein-coding
FAM135B	0.0440754158949034	-2.58	chr8	138505531	138505778	-8832	NM_015912	family with sequence similarity 135 member B	protein-coding
MYO3A	0.044089679210524	-2.03	chr10	25933875	25934253	-9	NM_017433	myosin IIIA	protein-coding
LOC101927901	0.0440898739730806	-3.42	chr12	67716759	67716915	12638	NR_110063	uncharacterized LOC101927901	ncRNA
BDKRB1	0.0441142951610781	-2.05	chr14	96309245	96309535	53180	NM_000710	bradykinin receptor B1	protein-coding
S100A7	0.0441627448517441	-1.95	chr1	153456460	153456845	4009	NM_002963	S100 calcium binding protein A7	protein-coding
MRPS25	0.0441638556147345	1.83	chr3	15016965	15017163	48273	NR_135246	mitochondrial ribosomal protein S25	protein-coding
LINC01950	0.0441724374424575	5.57	chr5	106179819	106180015	831097	NR_104671	long intergenic non-protein coding RNA 1950	ncRNA
STN1	0.0442838712599075	-2.4	chr10	103908193	103908481	9950	NM_024928	STN1, CST complex subunit	protein-coding
GNB5	0.0443178005882002	1.95	chr15	52163362	52163582	16493	NM_006578	G protein subunit beta 5	protein-coding
TET2-AS1	0.0443541777715094	2.95	chr4	105411548	105411720	-58753	NR_126420	TET2 antisense RNA 1	ncRNA
ASIC5	0.0443701321650871	2.35	chr4	155830979	155831171	35198	NM_017419	acid sensing ion channel subunit family member 5	protein-coding
FAM230C	0.0443742804234425	2.36	chr13	17226788	17226962	-968422	NR_027278	family with sequence similarity 230 member C	ncRNA
MIR3977	0.0445276540186919	-1.84	chr5	82594220	82594652	-245719	NR_039773	microRNA 3977	ncRNA
VAV1	0.0445332827170118	1.9	chr19	6791237	6791466	18683	NM_001258206	vav guanine nucleotide exchange factor 1	protein-coding
ATMIN	0.044553389386005	2.61	chr16	81056154	81056384	19087	NM_001300728	ATM interactor	protein-coding
LOC101927822	0.0446844817647074	1.85	chr8	133854899	133855286	-31404	NR_125424	uncharacterized LOC101927822	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
LINC01209	0.044690568116901	2.12	chr3	176814451	176814749	2401	NR_110819	long intergenic non-protein coding RNA 1209	ncRNA
PRKAR1A	0.0446978967669491	1.93	chr17	68389419	68389571	-24128	NM_001278433	protein kinase cAMP-dependent type I regulatory subunit alpha	protein-coding
LINC01781	0.0448308682418888	4.04	chr1	80258130	80258345	-277518	NR_125940	long intergenic non-protein coding RNA 1781	ncRNA
PAK5	0.0448471093476907	2.31	chr20	9750609	9750799	88335	NM_020341	p21 (RAC1) activated kinase 5	protein-coding
OSTF1	0.0449331988987287	3.29	chr9	75474463	75474614	386056	NM_012383	osteoclast stimulating factor 1	protein-coding
FRG1CP	0.0449438654596203	2.4	chr20	27943238	27943386	659353	NR_132315	FSHD region gene 1 family member C, pseudogene	pseudo
LOC439933	0.0449901765274379	-2.18	chr4	36206067	36206479	-37843	NR_122079	uncharacterized LOC439933	ncRNA
CTNND2	0.0450612202490306	2.77	chr5	11761187	11761395	142752	NM_001288717	catenin delta 2	protein-coding
SMG6	0.0450859198775306	-2.87	chr17	2276671	2276863	-10568	NM_001256827	SMG6, nonsense mediated mRNA decay factor	protein-coding
TTC29	0.0450878587931776	2.81	chr4	147093407	147093624	-147629	NM_001300761	tetratricopeptide repeat domain 29	protein-coding
LINC01474	0.0450926529551335	1.83	chr9	72839764	72840049	34203	NR_121120	long intergenic non-protein coding RNA 1474	ncRNA
RBM43	0.0451157040155538	-1.82	chr2	151239010	151239438	22651	NM_198557	RNA binding motif protein 43	protein-coding
LOC100287072	0.0451559718527681	-2.52	chr17	20734708	20734863	22339	NR_073509	ribosomal protein S6 kinase B1 pseudogene	pseudo
GBE1	0.0453044386158926	2.47	chr3	81993685	81994134	-232110	NM_000158	1,4-alpha-glucan branching enzyme 1	protein-coding
LINC00865	0.0453080469672157	-2.81	chr10	89818306	89818470	-11105	NR_038382	long intergenic non-protein coding RNA 865	ncRNA
LOC102723886	0.0453318166278565	2.31	chr1	219419941	219420148	137280	NR_135822	uncharacterized LOC102723886	ncRNA
VCY1B	0.0453465654680436	2.2	chrY	13883079	13883271	103333	NM_181880	variable charge, Y-linked 1B	protein-coding
FYN	0.0453867430265907	1.94	chr6	111798449	111798700	-39443	NM_153048	FYN proto-oncogene, Src family tyrosine kinase	protein-coding
SLC9A2	0.045405773942462	3.3	chr2	102582130	102582318	-37310	NM_003048	solute carrier family 9 member A2	protein-coding
ARHGAP26-AS1	0.0454230254341779	2.77	chr5	142962137	142962311	-93314	NR_046680	ARHGAP26 antisense RNA 1	ncRNA
TBX5	0.0454239136980185	-2.35	chr12	114401212	114401390	4862	NM_181486	T-box 5	protein-coding
MIR4500	0.0455355305347836	3.39	chr13	87283745	87283937	334899	NR_039722	microRNA 4500	ncRNA
SLC39A12-AS1	0.0455517384998039	2.53	chr10	18070316	18070476	-59834	NR_038419	SLC39A12 antisense RNA 1	ncRNA
ADAMTS16	0.0456442471118376	3.09	chr5	5184020	5184172	43766	NR_136935	ADAM metallopeptidase with thrombospondin type 1 motif 16	protein-coding
GRID1-AS1	0.045652899159156	3.47	chr10	85677144	85677296	99489	NR_038986	GRID1 antisense RNA 1	ncRNA
OLFM3	0.0457096361922298	2.01	chr1	101996703	101997055	355	NM_001288823	olfactomedin 3	protein-coding
CNR1	0.0457128974872176	1.96	chr6	88238080	88238232	-72108	NM_001160258	cannabinoid receptor 1	protein-coding
BTNL8	0.0457152550032173	2.26	chr5	180935382	180935531	25892	NM_001159710	butyrophilin like 8	protein-coding
DCDC5	0.0457163766722335	2	chr11	30984534	30984877	7981	NM_020869	doublecortin domain containing 5	protein-coding
LVCAT1	0.0457288076912799	2.44	chr4	43510342	43510510	-506435	NR_131959	liver cancer-associated transcript 1	ncRNA
MIR4473	0.0458772722091954	2.33	chr9	20219303	20219465	191854	NR_039684	microRNA 4473	ncRNA
MIR4743	0.0459320640819482	-2.36	chr18	48710595	48710851	40123	NR_039897	microRNA 4743	ncRNA
ZNF716	0.04594272257126	2.37	chr7	59836756	59837419	2386910	NM_001159279	zinc finger protein 716	protein-coding
CDH12	0.0459687662769158	2.6	chr5	22971181	22971360	-117648	NM_004061	cadherin 12	protein-coding
LINC00997	0.0460155371610722	2.56	chr7	32768902	32769134	10732	NR_036501	long intergenic non-protein coding RNA 997	ncRNA
LOC105372028	0.0460244646886616	3.13	chr18	24915216	24915424	189539	NR_134604	uncharacterized LOC105372028	ncRNA
LOC100506422	0.0460657240621303	2.51	chr9	26135273	26135493	68708	NM_001004352	putative deoxyuridine 5'-triphosphate nucleotidohydrolase-like protein FLJ16323	protein-coding
SHD	0.0462067982788078	-1.95	chr19	4274255	4274635	-4156	NM_020209	Src homology 2 domain containing transforming protein D	protein-coding
FLNB-AS1	0.0462867190790283	-2.52	chr3	58134535	58134761	35987	NR_135534	FLNB antisense RNA 1	ncRNA
CDH19	0.0463231269277678	2.19	chr18	66310347	66310509	293710	NM_021153	cadherin 19	protein-coding
MIR4454	0.0463513869656866	2.28	chr5	13015294	13015687	347631	NR_039659	microRNA 4454	ncRNA
ALG10B	0.0463599846473432	2.74	chr12	37354654	37354833	-962012	NM_001013620	ALG10B, alpha-1,2-glucosyltransferase	protein-coding
C2orf76	0.0464940939892728	2.1	chr2	119324811	119324982	41846	NM_001322330	chromosome 2 open reading frame 76	protein-coding
LOC101928944	0.046611760420298	3.04	chr11	80943442	80943639	-180737	NR_120570	uncharacterized LOC101928944	ncRNA
GORASP2	0.0466250320766522	1.81	chr2	170916257	170916653	-11983	NM_001201428	golgi reassembly stacking protein 2	protein-coding
DACH1	0.046668735620174	2.88	chr13	71686046	71686213	181063	NM_004392	dachshund family transcription factor 1	protein-coding
MANEA-AS1	0.0466722437722995	2.85	chr6	95270052	95270233	307309	NR_104136	MANEA antisense RNA 1 (head to head)	ncRNA
FGF20	0.046703149063034	-2.09	chr8	16919637	16919940	82377	NM_019851	fibroblast growth factor 20	protein-coding
OR4D11	0.046783623881151	2.88	chr11	59502810	59502983	-680	NM_001004706	olfactory receptor family 4 subfamily D member 11	protein-coding

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
SMA4	0.0468658731146528	3.03	chr5	70177669	70177839	48041	NR_024054	glucuronidase beta pseudogene	pseudo
ESR1	0.046866039738215	3.3	chr6	151970374	151970536	161348	NM_001328100	estrogen receptor 1	protein-coding
OR8K1	0.0468868479371892	1.86	chr11	56337011	56337209	-8929	NM_001002907	olfactory receptor family 8 subfamily K member 1	protein-coding
LOC729506	0.046919920551208	2.46	chr5	8400972	8401142	56507	NR_039984	uncharacterized LOC729506	ncRNA
ARSB	0.0469521612026393	2.9	chr5	78826245	78826420	159611	NM_198709	arylsulfatase B	protein-coding
RFC3	0.0469629605508804	2.2	chr13	33828474	33828626	10481	NM_181558	replication factor C subunit 3	protein-coding
CHRNB3	0.0469708781940975	2.94	chr8	42663513	42663712	-33764	NM_001347717	cholinergic receptor nicotinic beta 3 subunit	protein-coding
RAI2	0.0469823213451576	2.59	chrX	17945061	17945260	-83823	NM_021785	retinoic acid induced 2	protein-coding
LINC01393	0.0472866233486281	2.45	chr7	115153852	115154012	74974	NR_120521	long intergenic non-protein coding RNA 1393	ncRNA
LINC00316	0.0472992994269729	2.41	chr21	45369376	45369539	-27467	NR_103811	long intergenic non-protein coding RNA 316	ncRNA
POU4F3	0.0473234180843376	2.22	chr5	146375705	146375874	36765	NM_002700	POU class 4 homeobox 3	protein-coding
IPO11-LRRC70	0.0473781693598626	2.59	chr5	62979259	62979499	400644	NR_073584	IPO11-LRRC70 readthrough	ncRNA
MIR135A2	0.0474299319401931	2.36	chr12	97556420	97556575	-7315	NR_029678	microRNA 135a-2	ncRNA
OR10W1	0.0475304910699848	2.49	chr11	58266908	58267060	1276	NM_207374	olfactory receptor family 10 subfamily W member 1	protein-coding
WWOX	0.0475542417905861	2.15	chr16	78240242	78240409	140912	NM_001291997	WW domain containing oxidoreductase	protein-coding
STOML3	0.0475964298470559	1.88	chr13	38986566	38987230	3961	NM_001144033	stomatin like 3	protein-coding
SLC39A9	0.0477150863504278	2.26	chr14	69437840	69438006	39255	NM_001330185	solute carrier family 39 member 9	protein-coding
KCNAB2	0.0477310025245774	-2.04	chr1	6052599	6052849	6610	NM_001199863	potassium voltage-gated channel subfamily A regulatory beta subunit 2	protein-coding
GFRA4	0.0477365674537113	2.09	chr20	3644042	3644280	19238	NM_022139	GDNF family receptor alpha 4	protein-coding
SH3GL2	0.04777850803309	4.6	chr9	17454976	17455165	-123884	NM_003026	SH3 domain containing GRB2 like 2, endophilin A1	protein-coding
NLGN1-AS1	0.0478153270385615	2.26	chr3	173711902	173712078	208806	NR_046664	NLGN1 antisense RNA 1	ncRNA
LOC100505797	0.0478371643640669	-2.55	chr18	72957355	72957526	88400	NR_134647	myosin heavy chain IB-like	ncRNA
ARSJ	0.0478413049950325	-1.93	chr4	114006034	114006314	-26452	NM_024590	arylsulfatase family member J	protein-coding
CERS6	0.0478563252254198	2.96	chr2	168499602	168499878	43491	NM_203463	ceramide synthase 6	protein-coding
SOX5	0.047857580567463	2.94	chr12	23415449	23415646	169065	NM_178010	SRY-box 5	protein-coding
TCOF1	0.0478596984041809	2.02	chr5	150367202	150367384	9654	NM_001135243	treacle ribosome biogenesis factor 1	protein-coding
NEDD1	0.0478689945870549	1.83	chr12	96831417	96831694	-75668	NM_001135176	neural precursor cell expressed, developmentally down-regulated 1	protein-coding
DENND1A	0.0479525349566797	2.18	chr9	123897430	123897691	32578	NM_020946	DENN domain containing 1A	protein-coding
BRCAT107	0.047971593457929	1.84	chr5	44603653	44604006	-93547	NR_131946	breast cancer-associated transcript 107	ncRNA
KCNS3	0.0479766229733552	-2.77	chr2	18013648	18013818	135076	NM_002252	potassium voltage-gated channel modifier subfamily S member 3	protein-coding
SAXO1	0.0479774019966754	-3.6	chr9	19028775	19028996	4373	NM_001287050	stabilizer of axonemal microtubules 1	protein-coding
VPS13D	0.048045661812017	2.37	chr1	12347006	12347210	117069	NM_018156	vacuolar protein sorting 13 homolog D	protein-coding
PLCE1-AS2	0.0480890783761436	2.93	chr10	94195084	94195247	-86371	NR_120616	PLCE1 antisense RNA 2	ncRNA
C7orf71	0.0480903235277504	2.15	chr7	26686624	26686815	47634	NM_001302625	chromosome 7 open reading frame 71	protein-coding
SLC8A1	0.0481265482720285	2.06	chr2	41123667	41123834	-611315	NM_001112802	solute carrier family 8 member A1	protein-coding
LINC00290	0.0481592274069648	2.8	chr4	180134004	180134188	1025053	NR_033918	long intergenic non-protein coding RNA 290	ncRNA
NAV2	0.0481633487700596	-2.38	chr11	19935191	19935367	-87277	NM_001111019	neuron navigator 2	protein-coding
LINC00613	0.0482420373592496	2.77	chr4	135832275	135832453	55729	NR_103762	long intergenic non-protein coding RNA 613	ncRNA
LOC102723778	0.0482877628785146	-3.11	chr4	31242257	31242412	71190	NR_125935	uncharacterized LOC102723778	ncRNA
MIR3942	0.0483073060081484	-2.53	chr15	35450535	35450789	-78298	NR_037507	microRNA 3942	ncRNA
RPL38	0.0483326073242739	1.98	chr17	74025092	74025365	-178428	NM_000999	ribosomal protein L38	protein-coding
MIR4454	0.0483628740016825	2.32	chr5	12797173	12797354	129404	NR_039659	microRNA 4454	ncRNA
RIMS2	0.048373891901155	3.25	chr8	104129962	104130119	-93295	NM_001282882	regulating synaptic membrane exocytosis 2	protein-coding
LINC01242	0.0483851340224077	2.98	chr9	29950559	29950748	457801	NR_046204	long intergenic non-protein coding RNA 1242	ncRNA
TENM4	0.0484318260099058	2.47	chr11	80078141	80078303	-637571	NM_001098816	teneurin transmembrane protein 4	protein-coding
FAM181B	0.0484910156423778	-3.21	chr11	82758828	82759002	-25051	NM_175885	family with sequence similarity 181 member B	protein-coding
TNPO3	0.0485110037144554	-1.92	chr7	129014667	129015125	40277	NM_001191028	transportin 3	protein-coding
LINC00879	0.0485788440793346	2.91	chr3	94929020	94929193	-9157	NR_015400	long intergenic non-protein coding RNA 879	ncRNA
TAPT1	0.0486465897909637	-2.2	chr4	16224489	16224702	1943	NM_153365	transmembrane anterior posterior transformation 1	protein-coding
BMP7-AS1	0.0486652575470709	2.48	chr20	56925530	56925711	-288076	NR_110631	BMP7 antisense RNA 1	ncRNA

Gene Name	p value	log2 Fold Change	Chr	Start	End	Distance to TSS	Nearest Promoter ID	Gene Description	Gene Type
LINC00564	0.0486840794944263	2.19	chr13	82258861	82259038	-1031966	NR_047494	long intergenic non-protein coding RNA 564	ncRNA
NEIL3	0.0486922283271262	2.59	chr4	177105815	177106042	-203909	NM_018248	nei like DNA glycosylase 3	protein-coding
ALG10B	0.048754899521384	2.07	chr12	36439903	36440091	-1876758	NM_001013620	ALG10B, alpha-1,2-glucosyltransferase	protein-coding
C8orf34-AS1	0.048774765213096	2.8	chr8	68406844	68407030	-75446	NR_038877	C8orf34 antisense RNA 1	ncRNA
MTCL1	0.0488029878169337	-2.56	chr18	8813311	8813475	96022	NM_015210	microtubule crosslinking factor 1	protein-coding
FAM47E-STBD1	0.0489310060968405	-2.11	chr4	76256907	76257125	5316	NM_001242939	FAM47E-STBD1 readthrough	protein-coding
CDS1	0.0489361546459413	2.25	chr4	84751186	84751360	168369	NM_001263	CDP-diacylglycerol synthase 1	protein-coding
LOC101929116	0.049045808283313	1.85	chr9	126082634	126082829	193184	NR_121584	uncharacterized LOC101929116	ncRNA
LINC01378	0.0490749080839803	3.09	chr4	117551301	117551467	122986	NR_125757	long intergenic non-protein coding RNA 1378	ncRNA
SLC1A1	0.0492187378591966	-2.43	chr9	4538185	4538382	47856	NM_004170	solute carrier family 1 member 1	protein-coding
MIR5702	0.0492287793888767	2.11	chr2	226581057	226581219	77655	NR_049887	microRNA 5702	ncRNA
OR7E37P	0.0493010281248102	2.36	chr13	41403031	41403214	-39442	NR_002163	olfactory receptor family 7 subfamily E member 37 pseudogene	pseudo
PTPRD	0.0493057467873341	-2.94	chr9	8750976	8751167	-17125	NM_130393	protein tyrosine phosphatase, receptor type D	protein-coding
KCNC2	0.0493069709589854	2.49	chr12	74985351	74985525	224310	NM_139137	potassium voltage-gated channel subfamily C member 2	protein-coding
LOC440311	0.049315465607258	-2.22	chr15	94670388	94670677	-184831	NR_077061	glioma tumor suppressor candidate region gene 2 pseudogene	pseudo
WASHC2A	0.0493918294972107	-3.12	chr10	50147398	50147604	79613	NM_001330102	WASH complex subunit 2A	protein-coding
MIR548AU	0.0494337603657224	1.96	chr2	107364833	107365000	-15391	NR_049853	microRNA 548au	ncRNA
EIF5A2	0.0495077953780688	2.2	chr3	170948664	170948842	-40116	NM_020390	eukaryotic translation initiation factor 5A2	protein-coding
RIC1	0.0495307064333402	-1.81	chr9	5580433	5580920	-48443	NM_020829	RIC1 homolog, RAB6A GEF complex partner 1	protein-coding
LINC01184	0.0495790375295369	2.7	chr5	127959517	127959769	123431	NR_015360	long intergenic non-protein coding RNA 1184	ncRNA
WDFY3-AS2	0.0497123530798797	1.96	chr4	85127844	85128020	161114	NR_015359	WDFY3 antisense RNA 2	ncRNA
RPS7P5	0.0497187711472539	3.23	chr1	240026900	240027093	19472	NR_036695	ribosomal protein S7 pseudogene 5	pseudo
CDH8	0.0497629460181422	-2.34	chr16	63403899	63404083	-1367156	NM_001796	cadherin 8	protein-coding
RERGL	0.0497871791703444	2.12	chr12	18141650	18141831	-51547	NM_001286201	RERG like	protein-coding
LOC101929550	0.049792537208319	2.62	chr8	35969665	35969816	-259420	NR_125819	uncharacterized LOC101929550	ncRNA
MIR7641-2	0.0498177558881081	-2.14	chr6	4313192	4313438	-114687	NR_106988	microRNA 7641-2	ncRNA
ASPA	0.0498764554473462	-2.53	chr17	3475669	3475886	-225	NM_000049	aspartoacylase	protein-coding
LSM8	0.0499057191640607	-2.87	chr7	118038370	118038539	-145578	NM_016200	LSM8 homolog, U6 small nuclear RNA associated	protein-coding
TRIM49B	0.0499066466159664	1.88	chr11	48941983	48942251	-89483	NM_001206626	tripartite motif containing 49B	protein-coding
EXOSC3	0.0499230470163535	-2.3	chr9	37775594	37776034	9278	NM_016042	exosome component 3	protein-coding
CCDC144NL-AS1	0.0499470302166307	2.06	chr17	20842311	20842484	-26036	NR_104185	CCDC144NL antisense RNA 1	ncRNA
ATOH1	0.049948183970945	1.82	chr4	93772049	93772292	-56757	NM_005172	atonal bHLH transcription factor 1	protein-coding
TLE1	0.0499515156235147	2.51	chr9	81132655	81132864	556776	NM_001303104	transducin like enhancer of split 1	protein-coding
DKFZP586I1420	0.0499558403779887	-2.82	chr7	30340563	30340719	-29409	NR_002186	uncharacterized protein DKFZp586I1420	pseudo
BCL2L14	5.67642824786759e-05	3.42	chr12	12071021	12071267	200	NM_138723	BCL2 like 14	protein-coding

Conclusions et perspectives :

Les résultats des 3 études inclus dans ce manuscrit apportent dans leur ensemble un certain nombre de connaissance nouvelles pour répondre aux 2 objectifs principaux définis au début de mon projet de thèse:

a) le rôle, l'étendue et les mécanismes par lesquels HBx affecte la transcription des gènes cellulaires dans le contexte de l'infection virale *in vitro* et *in vivo* et en particulier les interactions de HBx avec la machinerie épigénétique ;

b) l'impact de HBV sur l'accessibilité de la chromatine et le contrôle épigénétique de la transcription des cellules de l'hôte.

Les deux objectifs sont liées à la caractérisation de l'interaction entre le virus et la cellule hôte infectée et à la question toujours controversée de savoir si et dans quelle mesure le HBV est un virus « stealth » ou « furtif » qui n'induit pas ou très peu de modifications transcriptionnelles dans des cellules infectées *in vitro* (Niu et al., 2017; Suslov et al., 2018) ou à la suite d'une infection aiguë dans des modèles animaux (Wieland et al., 2004).

Comme déjà mentionné, la situation est clairement différente dans le foie des patients atteints d'hépatite B chronique où l'on observe une répression de plusieurs gènes de l'immunité inné d'entité variable dans les différentes phases de la maladie (Lebossé et al., 2017). On pourrait soutenir que, dans cette étude, les patients se trouvaient dans une situation d'infection chronique avec de nombreuses années d'inflammation intrahépatique et de réplication virale et qu'il pouvait donc être difficile voire impossible d'identifier la contribution directe du virus. Cependant, d'autres études montrent que le virus est détecté par les récepteurs de l'immunité innée (Lebossé et al., 2017; Luangsay et al., 2015; Lucifora et al., 2018) et a un impact sur la transcription des cellules hôtes (Ancey et al., 2015; Lamontagne et al., 2016). Les résultats des expériences de RNA-Seq présentés dans l'**étude 3** confortent la notion selon laquelle l'infection par le HBV module l'expression de nombreux gènes dans les hépatocytes primaires humains, avec 563 gènes dérégulés à 2 heures, 1259 gènes à 24 heures et 636 gènes à 72 heures post-infection.

Les résultats des expériences de ChIP-Seq décrits dans l'**étude 1**, qui ont permis d'identifier le répertoire des cibles génomiques de HBx dans des cellules qui répliquent le HBV, représentent encore une autre confirmation de la capacité de HBV à avoir un impact sur le génome de la cellule hôte et soulignent le rôle de la protéine HBx. Bien que le rôle principal de HBx soit d'initier et de maintenir la transcription de l'ADNccc, plusieurs articles ont décrit des gènes cellulaires ciblés par HBx. HBx est co-recruté avec l'acétyltransférase CBP et active la transcription de plusieurs gènes cellulaires régulés par le facteur de transcription CREB (Cougot et al., 2007). Le recrutement du complexe mSin3A/HDAC et HBx sur le promoteur du gène E-cadhérine (CDH1) induit une déacétylation et l'inhibition de son expression (Arzumanyan et al., 2012). De même, HBx est recruté sur les promoteurs P3 et P4 de l' « insulin-like growth factor-2 » (IGF-2) et forme un complexe avec les acétyl-transférases CBP et p300 et la « methyl-CpG binding domain protein 2 » (MBD2) pour induire l'hypo-méthylation et la transcription des deux promoteurs et augmenter l'expression de l'IGF-2 (Liu

et al., 2015b). La plupart des données sur l'activation de gènes cellulaires par HBx a été générée dans des cellules sur-exprimant HBx et non dans le contexte de réplication de HBV. Notre **étude 1** présente pour la première fois une évaluation au niveau du génome entier de toutes les cibles potentielles de HBx dans un contexte de réplication virale.

Les cibles directes de HBx que nous avons mises en évidence sont à la fois des gènes codant des protéines et des ncRNA (75 miRNA, 34 lncRNA, 44 snoRNA et 3 snRNA). L'analyse fonctionnelle d'environ 4 000 gènes codant pour des protéines et de 75 miRNA potentiellement régulés au niveau transcriptionnel par HBx a révélé un enrichissement en gènes/ncRNA impliqués dans le métabolisme cellulaire, la dynamique de la chromatine et le cancer, mais également en gènes/ncRNA qui modulent la réplication de HBV. Parmi les gènes/ncRNA potentiellement importants pour la transformation des hépatocytes, il est intéressant de mentionner l'oncogène cellulaire c-scr et miR-21, un onco-miR bien connu.

Les mécanismes les plus importants par lesquels HBx augmente la transcription de l'ADNccc et, par conséquent, la réplication virale, sont d'induire la dégradation du facteur de restriction Smac5/6 pour démarrer la transcription de l'ADNccc (Decorsière et al., 2016) et d'empêcher sa répression épigénétique par des désacétylases et méthyltransférases cellulaires (Belloni et al., 2009; Rivière et al., 2015). D'autres mécanismes utilisé par HBx afin de favoriser la transcription de l'ADNccc et la réplication deHBV ont été décrits, comme la répression de l'ADN méthyltransférase DNMT3A qui va induire l'expression du miR-101 et donc l'induction de l'endocytose et l'autophagie, qui sont nécessaires à la réplication virale (Wei et al., 2013).

La caractérisation des cibles transcriptionnelles de HBx dans l'**étude 1** nous a permis d'identifier de nouveaux mécanismes par lesquels HBx stimule la réplication de HBV. L'activation transcriptionnelle directe de gènes et de ncRNA impliqués dans l'endocytose (famille RAB) et l'autophagie (ATGs, beclin-1, miR33a) contribue à la réplication virale. D'autre part, HBx réprime la transcription d'un groupe de miRNA (miR138, miR224, miR576, miR596) qui ciblent directement l'ARNpg de HBV et inhibent la réplication de HBV. Ainsi, d'un côté HBx régule l'activité de transcription de l'ADNccc et, de l'autre, protège l'ADNccc de la dégradation.

Perspectives (étude 1). *Les modèles cellulaires utilisés dans l'étude 1 pour les expériences ChIP-Seq et les validations en termes de ChIP et d'expression d'ARNm et de ncRNA, nous ont permis d'étudier les cibles directes de HBx dans un contexte de réplication et d'infection virale. Néanmoins, ces modèles représentent, au mieux, un modèle d'infection aiguë par le HBV plutôt que de refléter un état d'infection HBV chronique, beaucoup plus pertinent dans le contexte du développement du HCC. Pour ces raisons, nous prévoyons d'effectuer d'autres expériences de ChIP-Seq chez des souris humanisées infectées par le HBV et comportant un foie humanisé (modèle de réplication chroniqueHBV) ou une double humanisation du foie et du système immunitaire (modèle de traitement de l'hépatite B chronique). Enfin, des expériences de ChIP-Seq anti-HBx seront effectuées chez les souris trangéniques HBV-1.3, dans lesquelles le HBV se réplique et qui développent à l'âge de 15-18 mois des tumeurs du foie ressemblant aux HCC humains.*

Parmi les lncRNA ciblés par HBx et identifiés dans l'**étude 1**, la caractérisation de l'interaction entre HBx et l'un de ces lncRNA, DLEU2, nous a donné l'occasion de préciser davantage

l'impact de HBx sur la machinerie épigénétique et de dévoiler une voie potentiellement importante dans le développement et/ou la progression du HCC.

Dans l'**étude 2** nous nous sommes donc concentrés sur DLEU2, qui est fortement activé par le HBV, est exprimé lors de la maturation du foie (Peng et al., 2014) et est surexprimé dans les HCC humains (Villa et al., 2016; Yang et al., 2017). Un autre lncRNA surexprimé après infection par le HBV et cible directe de HBx, linc00441, est aussi surexprimé dans les HCC et a été montré comme pouvant lier et inactiver le suppresseur de tumeur pRB1 (Tang et al., 2017).

Le recrutement de HBx sur le gène DLEU2 s'accompagne d'une augmentation de la liaison de la polymérase II, de l'acétylation des histones H4 et de l'ARN DLEU2, confirmant le rôle de HBx dans l'induction de l'expression de DLEU2.

Plusieurs lncRNA ont été décrits comme servant de guide pour les complexes modifiant la chromatine vers leurs cibles pour réguler l'expression génique (Mercer et al., 2009; Wang and Chang, 2011). Ainsi, le lncRNA HOTAIR induit une reprogrammation à l'échelle du génome des cibles du complexe PRC2 afin de favoriser les métastases du cancer du sein et du côlon (Gupta et al., 2010; Kogo et al., 2011). Plusieurs lncRNA interagissent *in vivo* avec le complexe EZH2/PRC2 dans les cellules souches embryonnaires (ESC) (Zhao et al., 2010) et DLEU2 ainsi que les xARN, TSIX, MALAT1 et GAS5 sont hautement enrichis dans les RIP-Seq EZH2 réalisés dans des cellules de cancer du côlon (EZH2) (Guil et al., 2012). Dans l'**étude 2** nous avons montré, par des expériences de RIP, que DLEU2 se lie à EZH2 dans les cellules du foie et à la protéine virale HBx dans les cellules infectées par le HBV. Ces résultats mettent en évidence une nouvelle propriété de la protéine virale de se lier aux lncRNA.

Les lncRNA peuvent réguler l'expression de gènes voisins, par rapport à leur propre gène (régulation en *cis*), ou distants (régulation en *trans*) (Wang and Chang, 2011). Le gène DLEU2 s'étend sur 2768 pb et chevauche le cluster miR-15a / miR-16-1, le gène TRIM13 et, partiellement, le gène DLEU1. L'étude de la régulation du gène TRIM13 par HBx, EZH2 et DLEU2 nous a permis de dévoiler un nouveau mécanisme par lequel HBx modifie les propriétés des complexes de modification de la chromatine EZH2/PRC2 conduisant à l'activation constitutive d'un sous-ensemble de gènes cibles d'EZH2 qui sont normalement conservés dans un état réprimé et sont surexprimés dans les HCC (Figure 41). En effet, l'expression de DLEU2 est corrélée avec l'expression de TRIM13 dans une cohorte de 373 patients atteints d'un HCC (TCGA, https: // cancergenome.nih.gov/). Cette corrélation existe dans les HCC liés à HBV mais pas dans le HCC liés à HCV ou à l'alcool, confirmant que l'expression de DLEU2 et TRIM13 est fortement et positivement corrélée à l'infection par le HBV *in vivo*. En plus de TRIM13, 260 autres gènes sont co-régulés avec l'ARN DLEU2 chez les patients atteints de HCC (données TCGA, https://cancergenome.nih.gov/) et il est important de noter que 70 d'entre eux ont été identifiés comme cibles directes de HBx dans les expériences de ChIP-Seq de notre **étude 1**.

Enfin, nous avons également montré que l'interaction de HBx avec DLEU2 et EZH2 se produisait sur le minichromosome de l'ADNccc où elle stimulait la transcription/réplication du virus (Figure 41). Il a été proposé récemment que les structures « folded guanine quadruplexes » (G4) puissent agir en tant que motifs de haute affinité pour la liaison des complexes PRC2 aux ARN (Wang et al., 2017b). La présence plutôt généralisée de motifs « G4-tract » dans de nombreux ARN a fourni un modèle unificateur pour expliquer comment les

complexes PRC2 interagissent avec plusieurs ARN. Plusieurs structures putatifs G4 ont été identifiés sur l'ADNccc et sur l'ARNpg de HBV et un motif « G4-quadruplex » dans le promoteur du gène S a été montré réguler la transcription de HBV (Biswas et al., 2017). Ces résultats ouvrent la possibilité de cibler l'interaction entre DLEU2, HBx et les structures « G4-quadruplex » sur l'ADNccc dans une perspective thérapeutique.

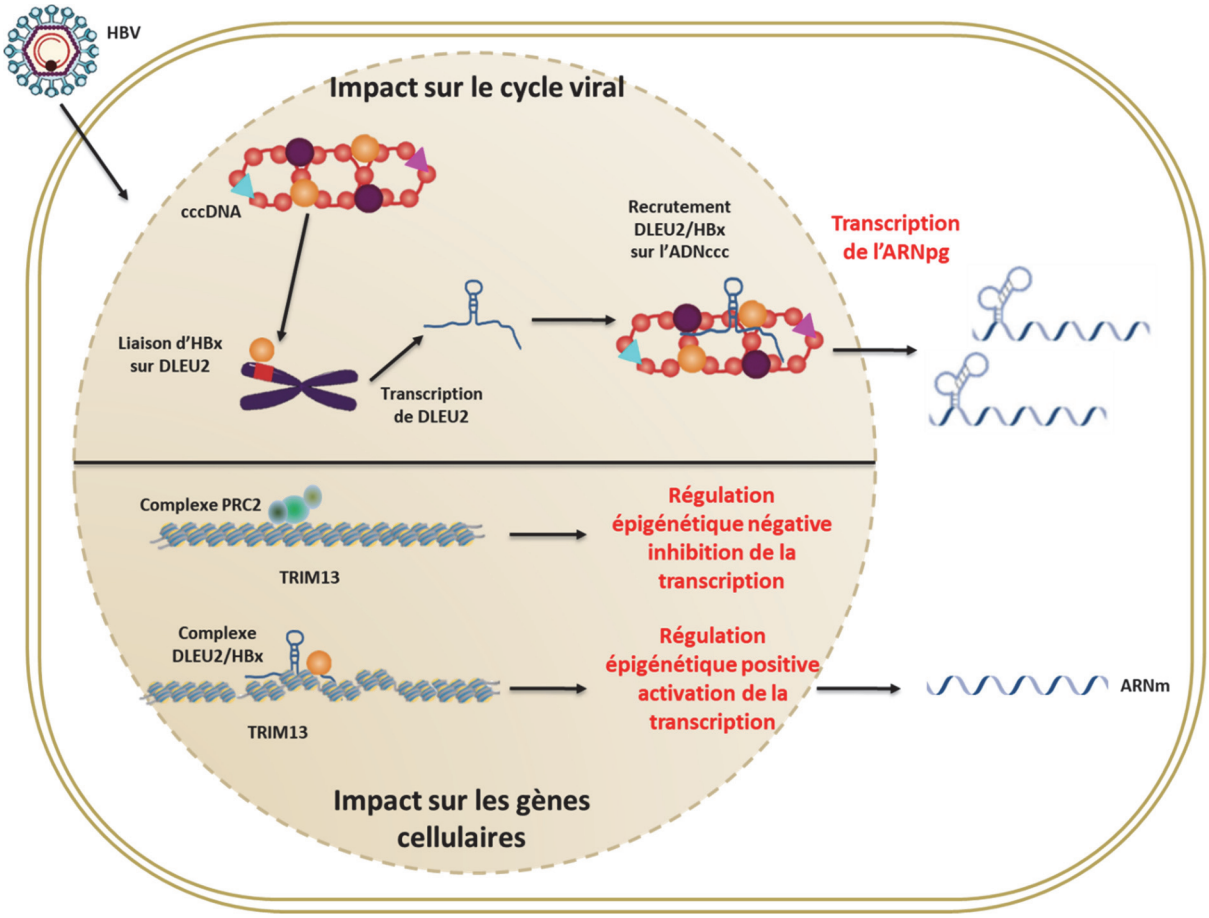


Figure 41 Schéma récapitulatif de l'impact de l'interaction entre HBx et DLEU2 sur le cycle viral ainsi que sur les gènes cellulaires.

Perspectives (étude 2). *Le rôle de l'interaction de DLEU2 avec HBx et EZH2 dans la réplication de HBV sera approfondi en générant une lignée cellulaire permissive à HBV dans laquelle l'expression de DLEU2 a été supprimée par une approche CRISPER/Cas9. D'autre part, nous poursuivrons l'identification de gènes pertinents pour le développement et la progression du HCC co-activés par HBx, DLEU2 et EZH2 en utilisant une combinaison de techniques omiques (ChIP-Seq anti-HBx et anti-EZH2, ChIRP DLEU2).*

Dans l'**étude 3**, toujours avec l'objectif d'étudier l'impact de l'infection par le HBV sur les fonctions du génome de la cellule hôte, nous avons utilisé la technique d'ATAC-Seq pour étudier les changements d'accessibilité de la chromatine imposés par HBV dans les hépatocytes primaires humains infectés. L'activation de la transcription dans les eucaryotes a été liée à la perturbation de l'organisation des nucléosomes au niveau des promoteurs, des « enhancers », des « silencers » et des « insulators » en raison de la liaison des facteurs de transcription et de l'identification des événements de remodelage de la chromatine. Ceci est

corrélé aux changements épigénétiques qui se traduisent par une expression génique différentielle, une prolifération cellulaire et le développement de certaines maladies.

La technique d'ATAC-Seq a été choisie parce qu'elle permet en même temps d'analyser la chromatine ouverte, le positionnement des nucléosomes et les empreintes de facteurs de transcription au niveau du génome entier. De plus, l'ATAC-Seq n'implique aucune modification chimique, est rapide et très sensible, nécessitant donc qu'un petit nombre de cellules (500 à 50 000 cellules).

Nos résultats montrent clairement des changements précoces d'accessibilité de la chromatine dans des hépatocytes primaires humains infectés par le HBV. En effet, le nombre de sites génomiques dont l'accessibilité de la chromatine est modifiée augmente fortement avec le temps d'infection (278 à 2 heures, 350 à 24 heures et 1095 at 72 heures), avec une prévalence de régions plus ouvertes (potentiellement actives) que de régions fermées (potentiellement réprimées) et environ 30% de ncRNA dans toutes les conditions.

Nous avons également constaté que les régions accessibles de la chromatine sont enrichies pour les motifs de liaison aux facteurs de transcription SMAD, STAT3, TP53 et a Suz12, un partenaire d'EZH2 dans le complexe PRC2. L'analyse fonctionnelle des gènes présents dans les régions avec des changements d'accessibilité de la chromatine a révélé un enrichissement en gènes/ncRNA impliqués dans le cancer du foie, la transcription, le métabolisme du calcium et la voie de signalisation du TGF-b.

Les données transcriptomiques de RNA-Seq dans les mêmes cellules ont montré que le HBV avait un impact sur l'expression de gènes appartenant à plusieurs voies, mais nous n'avons pas détecté de corrélation directe et univoque entre les changements d'accessibilité de la chromatine et la transcription. Ceci est de plus en plus rapporté dans de nombreuses études et cela est probablement dû à la difficulté de corréler le « niveau » d'accessibilité de la chromatine et la « localisation » des domaines de la chromatine plus ou moins accessibles avec l'expression génique. Une solution possible pour tenter de surmonter ce problème serait de coupler l'ATAC-Seq non seulement au RNA-Seq mais également aux profils ChIP-Seq pour des modifications post-traductionnelles des histones. Une autre limite de notre étude concerne l'absence de temps plus longs après l'infection.

Perspectives (étude 3). *Afin d'identifier la contribution de HBx aux changements très précoces (avant que la transcription complète de l'ADNccc se produise), nous prévoyons de répéter des expériences d'ATAC-Seq après l'infection d'hépatocytes primaires humains par un virus défectif pour HBx. De même, il sera important d'évaluer le « paysage » d'accessibilité de la chromatine dans un environnement plus proche de l'infection chronique tel que les souris humanisése.*

Annexes:

1. Articles publiés sur des études réalisées en master:

a. Expression and functionality of Toll- and RIG-like receptors in HepaRG cells

Expression and functionality of Toll- and RIG-like receptors in HepaRG cells

Souphalone Luangsay1,2,5,_, Malika Ait-Goughoulte1,2,5,_, Maud Michelet1,2, Océane Floriot1,2,Marc Bonnin1,2, Marion Gruffaz1,2, Michel Rivoire3,4, Simon Fletcher5,§, Hassan Javanbakht5,Julie Lucifora1,2, Fabien Zoulim1,2,6,7,⇑,_, David Durantel1,2,⇑,_

Journal of Hepatology 2015 vol. 63 j 1077–1085

Research Article

EASL JOURNAL OF HEPATOLOGY

Expression and functionality of Toll- and RIG-like receptors in HepaRG cells

Souphalone Luangsay[1,2,5,†], Malika Ait-Goughoulte[1,2,5,†], Maud Michelet[1,2], Océane Floriot[1,2], Marc Bonnin[1,2], Marion Gruffaz[1,2], Michel Rivoire[3,4], Simon Fletcher[5,§], Hassan Javanbakht[5], Julie Lucifora[1,2], Fabien Zoulim[1,2,6,7,*,‡], David Durantel[1,2,*,‡]

[1]*INSERM U1052, CNRS UMR_5286, Cancer Research Centre of Lyon (CRCL), 69008 Lyon, France;* [2]*University of Lyon, Université Claude Bernard (UCBL), UMR_S1052, 69008 Lyon, France;* [3]*Centre Léon Bérard (CLB), 69008 Lyon, France;* [4]*INSERM U1032, 69003 Lyon, France;* [5]*Pharma Research & Early Development (pRED), Roche Innovation Center Basel, F. Hoffmann-La Roche AG, 4070 Basel, Switzerland;* [6]*Hospices Civils de Lyon (HCL), 69002 Lyon, France;* [7]*Institut Universitaire de France (IUF), 75005 Paris, France*

Background & Aims: HepaRG cells are considered as the best surrogate model to primary human hepatocyte (PHH) culture to investigate host-pathogen interactions. Yet their innate immune functions remain unknown. In this study, we explored the expression and functionality of Toll-like (TLR) and retinoic acid-inducible gene-1 (RIG-I)-like receptors (RLR) in these cells.
Methods: Gene and protein expression levels of TLR-1 to 9 and RLR in HepaRG were mainly compared to PHH, by RT-qPCR, FACS, and Western blotting. Their functionality was assessed, by measuring the induction of *toll/rig-like* themselves and several target innate gene expressions, as well as the secretion of IL-6, IP-10, and type I interferon (IFN), upon agonist stimulation. Their functionality was also shown by measuring the antiviral activity of some TLR/RLR agonists against hepatitis B virus (HBV) infection.
Results: The basal gene and protein expression profile of TLR/RLR in HepaRG cells was similar to PHH. Most receptors, except for TLR-7 and 9, were expressed as proteins and functionally active as shown by the induction of some innate genes, as well as by the secretion of IL-6 and IP-10, upon agonist stimulation. The highest levels of IL-6 and IP-10 secretion were obtained by TLR-2 and TLR-3 agonist stimulation respectively. The highest preventive anti-HBV activity was obtained following TLR-2, TLR-4 or RIG-I/MDA-5 stimulations, which correlated with their high capacity to produce both cytokines.
Conclusions: Our results indicate that HepaRG cells express a similar pattern of functional TLR/RLR as compared to PHH, thus qualifying HepaRG cells as a surrogate model to study pathogen interactions within a hepatocyte innate system.

Keywords: Pathogen recognition receptor; Toll-like receptor; Hepatocyte; Hepatitis B virus; Antiviral activity.
Received 5 June 2014; received in revised form 15 June 2015; accepted 16 June 2015; available online 3 July 2015
* Corresponding authors. Address: INSERM U1052, 151 cours Albert Thomas, 69003 Lyon, France. Tel.: +33 4 72 68 19 70; fax: +33 4 72 68 19 71.
E-mail addresses: fabien.zoulim@inserm.fr (F. Zoulim), david.durantel@inserm.fr (D. Durantel).
† Contributed equally as joint first authors.
‡ Share senior authorship.
§ Present address: Gilead Sciences, Foster City, CA, USA.
Abbreviations: PHH, primary human hepatocyte; TLR, toll-like receptor; RIG-I, retinoic acid-inducible gene 1; RLR, RIG-like receptor; RT, reverse transcription; qPCR, quantitative polymerase chain reaction; FACS, fluorescence-activated cell sorting; IL-6, interleukin-6; IP-10, interferon-gamma-inducible protein-10; IFN, interferon; HBV, hepatitis B virus; MDA-5, melanoma differentiation-associated gene 5; HAV, hepatitis A virus; HEV, hepatitis E virus; HDV, hepatitis D virus; HCV, hepatitis C; WHO, world health organization; HCC, hepatocellular carcinoma; PRR, pathogen recognition receptor; CLR, C-type lectin receptor; NLR, NOD-like receptor; PAMP, pathogen-associated molecular pattern; NTCP, sodium taurocholate cotransporting polypeptide; dHepaRG, differentiated HepaRG; ATCC, american type culture collection; DMSO, dimethysulfoxide; CT, cycle threshold; LTA, lipoteichoic acid; LPS, lipopolyipopolysaccharide; HBeAg, hepatitis B «e» antigen; HBsAg, hepatitis B «s» antigen; pHepaRG, proliferative HepaRG; NF-κB, nuclear factorkappa B.

Introduction

The liver is targeted by several important human pathogens, including plasmodium falciparum or hepatotropic viruses [1]. Among them, hepatitis A (HAV) and E virus (HEV) do not usually establish persistent infections, whereas hepatitis B (HBV), D (HDV), and C virus (HCV) may evolve toward chronicity [1]. The high number of chronic carriers, around 240 million for HBV (WHO data) and 130 million for HCV worldwide, account for the rapid increase in the incidence of hepatocellular carcinoma (HCC) [2]. Pathogenesis of HBV and HCV infections, and the severity of virus induced liver injury are mainly determined by immunological-mediated events rather than the virus itself, since only limited cytopathic effects are noticed in hepatocytes following infection [3].

Infection of cells by microorganisms leads to the activation of the host inflammatory defense response through the initial sensing of pathogen mediated by innate pattern recognition receptors (PRR) [4]. PRR are mainly represented by Toll-like receptors (TLR), C-type lectin receptors (CLR), retinoic acid-inducible gene-1 (RIG-I)-like receptors (RLR), and NOD-like receptors (NLR). In addition, intracellular DNA sensors are proposed as new types of PRR [5]. To date, ten functional TLR have been identified in human (TLR-1 to 10) and detect specific pathogen-associated molecular patterns (PAMP) derived from microorganisms [6]. Upon recognition of respective PAMP, TLRs recruit distinct sets

Research Article

of adaptor molecules which initiate downstream signaling events that lead to the secretion of inflammatory cytokines, type I interferon (IFN), chemokines and antimicrobial peptides [7]. *In vitro* models have shown that an increasing number of viruses can activate TLR pathways [8]. Cells also express cytoplasmic RNA helicases such as RIG-I and melanoma-differentiation-associated gene-5 (MDA-5) that function as an alternative class of PRR through the recognition of double stranded-RNA (dsRNA) produced during viral replication.

Viruses can evade IFN response, and recent studies have increased our understanding of the underlying molecular mechanisms in relation to inhibition [9]. Following infection, viruses can induce complex intracellular events that affect many components of host signaling pathways to its own benefits [10,11]. Elucidating the strategies that are used by viruses to inhibit a host response is important for the understanding of the mechanism(s) by which viruses can establish persistent infection and to design strategies to unlock this inhibition for therapeutic purposes.

Although the expression of some PRR genes was shown in human liver tissue and PHH, the expression and functional activity of a wide spectrum of PRR on human hepatocytes remains poorly defined [12,13]. Transformed hepatic cell lines such as HepG2 and HuH7 are commonly used to study hepatotropic virus life cycle and early host/virus interaction, however, in terms of innate immunity, these cells differ from PHH in that they do not have fully functional IFN pathways [14,15] and display profound defect in PRR signaling [16]. Beside recently developed models based hepatoma cells overexpressing NTCP [17], or on the complex use of induced pluripotent cells (iPS) [18], only freshly prepared PHH and the hepatocyte-like HepaRG cells can support a complete HBV and HDV life cycle, including the early events of infection [19–21].

HepaRG cells are bipotent liver progenitor cells that differentiate into both cholangiocyte-like and hepatocyte-like cells in culture. Throughout differentiation, HepaRG cells evolve from a homogeneous dedifferentiated, depolarized epithelial phenotype showing no specific organization to a structurally well-defined and polarized monolayer closely resembling those formed in PHH culture [22–24]. This cell line represents an *in vitro* human model for hepatocytic differentiation and was shown to be functional for IFN signaling [25]. However, the functionality of TLRs/RLRs in this cell line has not been described so far.

Material and methods

Cell culture and HBV infection

HepaRG cells were cultured and differentiated in presence of DMSO (cell culture grade, Sigma) as previously described [20]. PHH were freshly prepared, as previously described and maintained in same media as dHepaRG [26]. HuH7 and HepG2 (ATCC) cells were also maintained in complete William's medium and cultured in presence of DMSO for three days before RNA extraction. The transfection of siRNA (SMARTpool™ from Darmacon/GE) into dHeppaRg or PHH cells was performed using Darmafect-1 reagent, as recommended by provider (Darmacon/GE). HBV recombinant virus (i.e. inoculum) was produced in HepG2.2.15 cells and differentiated HepaRG infected as previously described [27].

Nucleic acid extraction and RT-qPCR/qPCR

Total RNA was extracted from cells with the NucleoSpin RNA II kit according to manufacturer's instructions (Macherey-Nagel). RNA reverse transcription was performed using the Superscript III RT (Life Technologies). cDNA was then analyzed by qPCR with the "Express SYBR GreenER™ qPCR SuperMix Universal" (Invitrogen). RNA expression level was determined using the comparative cycle threshold (Ct) method, where the amount of target DNA was normalized to housekeeping genes *rplp0* and *β-actin* cDNA ($2^{-\Delta Ct}$). When cells were stimulated with ligands the amount of target DNA was normalized to housekeeping genes and to the unstimulated control cDNA ($2^{-\Delta\Delta Ct}$). Primers used for the qPCR step are presented in Table 1. Total DNA was extracted from HepaRG using the MasterPure™ Complete DNA/RNA Purification Kit (Epicentre). Quantitative PCR for intracellular HBV DNA was performed as indicated above using HBV specific primers.

Toll- and RIG-like receptor agonists and stimulation

All agonists were purchased from Invivogen and used according to provider's recommendation. Pam3CSK4, LTA, Poly(I:C)-LMW, LPS, Flagellin (FLA-BS), FSL-1, Imiquimod (R837) or CL264, R848, ssRNA40/lyoVec™, CpG ODN-2395 (class-C + CpG control), Poly(I:C)-LMW/LyoVec™ were respectively used to stimulate/engage TLR-1/2, TLR-2, TLR-3, TLR-4, TLR-5, TLR-2/6, TLR-7, TLR7/8, TLR-8, TLR-9, and RIG-I/MDA-5. The concentrations used are indicated in the figure legends.

Flow cytometry analysis

To detect TLR expression, 10^6 HepaRG or PHH cells were fixed with paraformaldehyde (2%; Sigma) and permeabilized with saponin (0.25%; Sigma). After saturation (saponin 0.25%, fraction-V albumin 3%; Sigma), cells were stained with purified anti-human TLR-2, 4, 5, 6 (sc-73181, sc-52062, sc-57461, sc-30001; Santa Cruz), anti-human TLR-7, 8, 9 (Alexis Biochemicals (ALX-210-874-C100, ALX-804-376-C100) or Abcam (ab184943, ab85859, ab45371)), or anti-human TLR-3 (mAb hTLR-3; Invivogen), followed by incubation with the appropriate secondary antibody conjugated to Alexa Fluor 488 (Invitrogen). For double staining, either anti-cytochrome P450-3A4 antibody (ab135813; Abcam) or anti-cytochrome P450-3A1 (ab22724; Abcam), which specifically stain hepatocytes and not cholangiocytes (Supplementary Fig. 1) were used in combination with anti-PRR antibodies from either mouse or rabbit origin. Cells were analyzed using a FACScalibur flow cytometer and analyzed using CellQuestPro software (BD Biosciences).

Table 1. Primer sequences used for RT-qPCR.

Gene	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')
TRL1	caggccctcttcctcgttaga	ttcctaaaggtagaagctgttcttca
TRL2	ctctcggtgtcggaat	cccgctcactgtaagaa
TLR3	tttgcaagaggaatgtttaaatct	cacctatccgttctttctgaactg
TLR4	tcttggtggaagttgaacg	gccacaccgggaataa
TLR5	gtaccctgactcgttct	ttctgcacccatgtga
TLR6	agtggacccagactcg	agttcgtaatggcacc
TLR8	tgagctgcgctaccac	cttctgccttcgggtt
TLR9	ggcccccggccttctt	caggagtggtccactgtcttga
TLR10	gcattcccaccaggtatcataaac	aaagcccacatttacgcctatc
RIGI	gctgatgaaggcattgacattg	cagcattactagtcagaaggaagca
MDA5	cccaagacacagaatgaacaaaa	cgagaccataacggataacaatgt
IFNα	gtgaggaaatacttccaaagaatcac	tctcatgatttctgctctgacaa
IFNβ	gccgcattgaccatgtatgaga	gagatcttcagtttcggaggtaac
IL28b	taagagggccaaagatgcctt	ctggtccaagacatccccc
IL29	gtgactttggtgctaggcttg	gcctcaggtcccaattccc
IL6	acccctgacccaaccacaaat	agctgcgcagaatgagatgagtt
ISG56	agccagatgtcctcacagac	cttctaccactgtttcatgc
β-actin	tggcattgccgacaggatgc	tctgctggaggtggacagcga
RPLP0	caccattgaaatcctgagtgatgt	tgaccagcccaaaggagaag

Western blot analysis

Cells were harvested in lysis buffer (140 mM NaCl, 10 mM Tris pH 7.6, 1 mM EDTA, 1% Triton X-100, 0.05% sodium dodecyl sulfate, 1× protease inhibitor cocktail 5 Roche). Clarified lysates were subjected to SDS-PAGE and Western blot transfer to nitrocellulose membranes (Millipore). Membranes were probed with anti-RIG-I, anti-MDA5, or anti-TLR2 antibodies (Alexis Biochemicals and Santa Cruz) followed by peroxidase-conjugated secondary antibodies (Life Technologies) and detection by chemiluminescence (Pierce). Anti-actin (Sigma) was used as a loading control.

Enzyme-linked immunoassay for IL-6, IP-10, and secreted HBV antigens

Supernatant from stimulated culture was harvested and assayed for IL-6 and IP-10 protein by ELISA using the human IL-6 or human IP-10 assay kit according to the manufacturer's protocol (Invitrogen). Cut-offs for these ELISA were 2.5 pg/ml. HBeAg and HBsAg were detected in the supernatant of infected HepaRG by ELISA using Autobio kits (Autobio, China). Cut-offs for these ELISA were respectively 1 NCU/ml (i.e. 1 NCU ≈ 13 ng) and 2.5 ng/ml.

Statistical analysis

Statistical analysis was performed using the non-parametric Mann-Whitney tests using the GraphPad Prism software. For all tests, a p value $\leqslant 0.05$ (*) was considered as significant.

Results

Analysis of TLR and RLR RNA expression in dHepaRG cells and PHH at RNA level

Differentiated HepaRG cells, PHH, as well as DMSO-partially re-differentiated HuH7 [28] and HepG2 [29] cells were investigated for total RNA expression levels for *TLR-2* to *TLR-10*, as well as *RIG-I* and *MDA-5*. All *TLRs* were similarly expressed in dHepaRG and PHH, whereas levels of expression were significantly lower in the two transformed hepatoma cell lines, except for *TLR-8* expression, which was slightly higher in the latters. *RIG-I* and *MDA-5* were well expressed in all cell types, but more expressed in PHH as compared to other cells. *TLR-7*, *8*, *9*, and *TLR-10* were less expressed than other sensors in both dHepaRG and PHH (Fig. 1A, B). The expression level of these PRR was further analyzed according to the HepaRG differentiation status (Fig. 1C). Differentiated HepaRG cells expressed slightly more *TLR-2, 3, 4* and *TLR-6* as compared to proliferative HepaRG (pHepaRG) whereas the expression level of the other sensors tested was not significantly changed in dHepaRG *vs.* pHepaRG.

Analysis of TLR and RLR protein production in HepaRG cells and PHH

TLR protein expression in pHepaRG and dHepaRG was further analyzed (Fig. 2A, B). FACS analysis showed that all TLR (TLR-2 to 9, with TLR-6 being less expressed) were produced in both proliferative and differentiated HepaRG cells (Fig. 2A), and a positive reactivity was also obtained for TLR-7, 8, and 9, although their expression at RNA level was very low. The expression of TLR in dHepaRG cells was found qualitatively similar, but quantitatively slightly higher as compared to PHH (two different batches) (Fig. 2B; Supplementary Fig. 2), whereas the RNA levels were similar suggesting differences in post-transcriptional or post-translational regulations. The expression of the cytosolic sensors RIG-I and MDA-5 was analyzed by Western blot in PHH and differentiated HepaRG cells. Both sensors could be detected

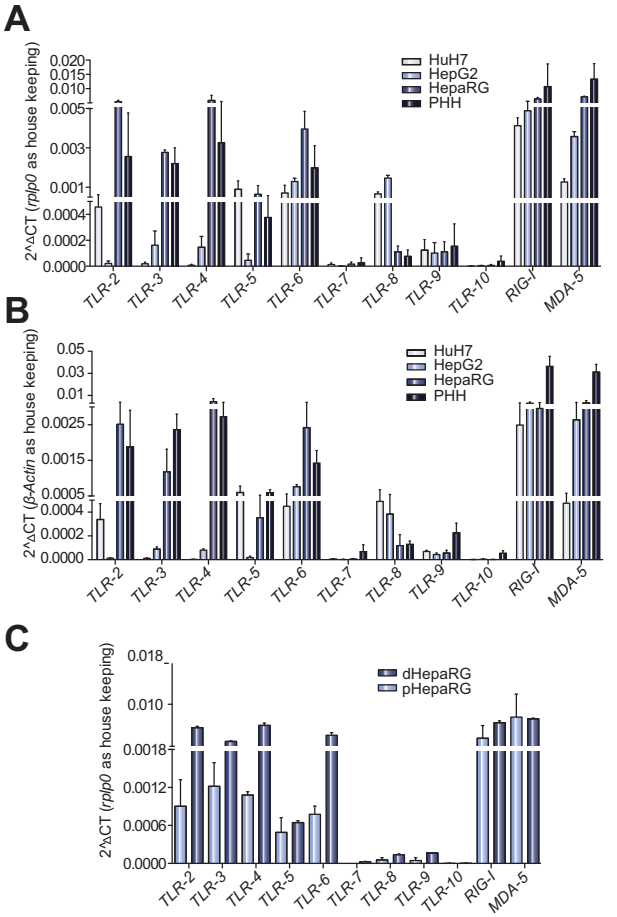


Fig. 1. mRNA expression of TLR and RLR in differentiated and proliferative HepaRG. (A and B) RT-qPCR analysis of TLR/RLR expression in dHepaRG, PHH, HepG2, and HuH7. The relative RNA expression was normalized to housekeeping genes *rplp0* (A) or *β-actin* cDNA ($2^{-\Delta Ct}$) (B). (C) Comparative RNA expression of TLR/RLR in pHepaRG and dHepaRG cells. RT-qPCR analysis of TLR/RLR expression was normalized to housekeeping genes (*rplp0* and *β-actin*) and relatively compared to pHepaRG cells. Results are given as a mean ± SEM of 3 (or 4 for PHH; n ⩾3) independent experiments (with biological triplicate) with three batches of PHH from different donors.

at basal level, but were highly expressed following cell stimulation (24 h) with IFN-α (1000 IU/ml) (Fig. 2C). Since differentiation of HepaRG cells leads to a mixture of hepatocyte- and cholangiocyte-like cells, we sought to determine whether expression of PRR would be differentially distributed in each cell type by flow cytometry using a double staining with antibodies directed against selected TLR and hepatocyte specifically-expressed P450 cytochromes. We found that the selected TLR were more expressed in hepatocyte-like cells, as compared to cholangiocyte-like cells (Fig. 2D).

Inducible cytokine RNA expression upon stimulation with TLR/RLR prototypic ligands/agonists

Differentiated HepaRG cells were stimulated for 2, 4, 8, and 16 h by TLR/RLR cognate ligands at concentrations indicated in Fig. 3. This stimulation could induce the expression of the

Research Article

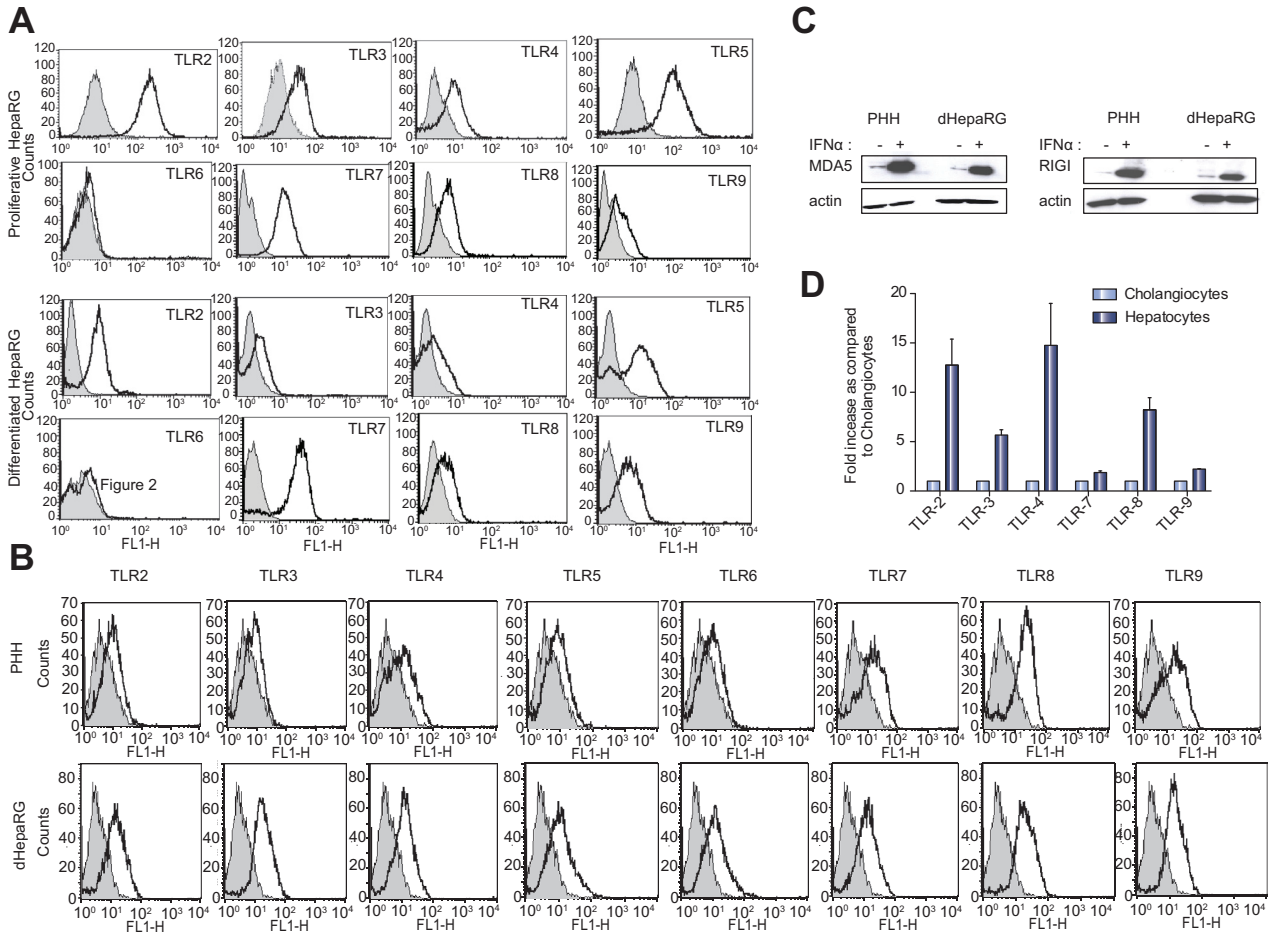


Fig. 2. Protein expression of PRR in pHepaRG and dHepaRG cells. (A and B) TLR expression analyzed by FACS. Black histograms represent staining with the indicated TLR antibody, shaded gray histograms represent the isotype control, in pHepaRG, dHepaRG and PHH. (C) Western blot analysis of RIG-I and MDA5 in dHepaRG (and PHH), with or without IFNα (1000 IU/ml) induction. (D) Expression of some TLRs in hepatocyte-like cells as compared to cholangiocyte-like cells measured by FACS, and reported on graph as fold increase of MFI (using cholangiocyte as reference, set at 1).

pro-inflammatory cytokine gene (*IL-6*), the type I and type III IFNs (*IFNα, β* and *λ*) and interferon stimulated genes, ISGs (*ISG56*). With the exception of TLR7 and 9 ligands, all cognate ligands could induce *IL-6* gene expression at the concentration tested (Fig. 3A, panel IL-6). A strong IL-6 induction was observed after 16 h of stimulation for TLR-2, 3, 4, 5, 6, RIG-I/MDA-5, whereas an induction was observed for TLR-8 after only 2 h of stimulation which returned to baseline rapidly. Because production of type I IFN, in particular IFN-β, is one of the first cellular antiviral defenses, we tested whether this gene was activated at RNA level in HepaRG cells after some TLR/RLR stimulations. Indeed *IFN-β* gene expression was found induced after stimulation by some cognate ligands (Fig. 3A, panel IFN-β). *IFN-β* gene moderately increased after stimulation with TLR-3, 7, 8, and 9 ligands, but strongly increased after stimulation with RIG-I/MDA-5 ligands. It is worth noting that the concentration of poly(I:C)-LMW/lyoVec™ (1 μg/ml) used in these experiments was shown to be slightly toxic when multiple dosing were used for antiviral studies; this led to a readjustment of the concentration in experiments presented hereafter. Stimulation of TLR-3, 7, 9 and RIG-I/MDA-5 also slightly induced IFN-α expression (Fig. 3A, panel IFN-α). Type III IFN expression has been shown to depend on the same stimulations, i.e. viral infection or dsRNA ligands, and signal transduction pathway, as those inducing type I IFN expression in epithelial cells [30]. IL-28B (IFN-λ3) expression was only induced by TLR-3 or RIG-I/MDA-5 stimulations, although the activation kinetics were different (i.e. earlier for TLR-3) (Fig. 3A, panel IL-28B). The induction of IL-29 (IFN-λ1) was unexpectedly lower than that of IL-28B (Fig. 3A, panel IL-29). The activation of all IFN transcripts was associated to the activation of prototypic ISGs such as *ISG56* induced after TLR-3 stimulation and RIG-I/MDA-5.

In PHH, low-level of gene activation was also observed with TLR-7/9 ligands (Fig. 4A), but as PHH cultures are not completely pure, we cannot exclude that residual activation was due to the presence of other liver cells.

Secretion profile of IL-6, IP-10, and type I IFN after stimulation with TLR/RLR prototypic ligands

The cell culture supernatants of HepaRG cells were first tested for the production of IL-6 and IP-10 following a 24 h TLR/RLR

JOURNAL OF HEPATOLOGY

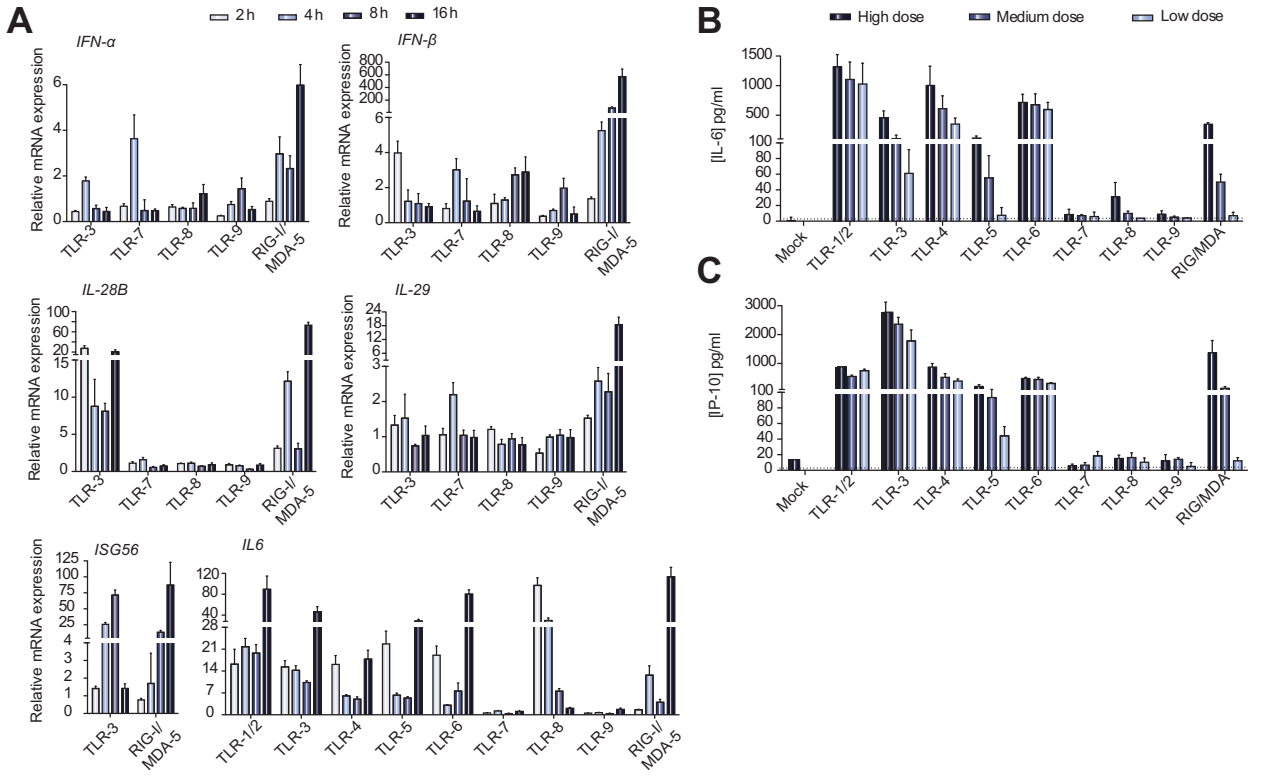


Fig. 3. Induction of *PRR*, *cytokine*, and *ISG* gene expression in dHepaRG cells & secretion of IL-6 and IP-10 after stimulation with prototypic ligands. (A) Gene expression was analyzed (n = 3) following 2 h, 4 h, 8 h or 16 h stimulation with prototypic ligands of TLR1/2 (Pam3CSK4, 500 ng/ml), TLR3 (Poly(I:C)-LMW, 10 μg/ml), TLR4 (LPS, 1 μg/ml), TLR5 (FLA-BS, 10 μg/ml), TLR6 (FSL1, 5 μg/ml), TLR7 (Imiquimod, 10 μg/ml), TLR8 (ssRNA40, 10 μg/ml), TLR9 (CpG-2395, 5 μM) RIG-I/MDA5 (Poly(I:C)-LMW-lyovec, 1 μg/ml). After stimulation RT-qPCR were performed on targeted *PRR* (self-activation by cognate ligands; *PRR* are indicated on the x-axis), *IL-6*, *IFN-β*, *IFN-α*, *IL-28B*, *Il-29*, and *ISG56* genes. The RNA level expression was normalized to housekeeping genes and relatively compared to the non-stimulated control. (B and C) The secretion of IL-6 and IP-10, after 24 h of stimulation with different doses of indicated ligands, was quantified by ELISA. The concentration used for the ligands were: TLR-1/2 (Pam3CSK4, 500-50-5 ng/ml), TLR-3 (Poly(I:C)-LMW, 10-1-0.1 μg/ml), TLR-4 (LPS, 1-0.1-0.01 μg/ml), TLR-5 (Flagellin, 10-1-0.1 μg/ml), TLR-6 (FSL1, 5-1-0.2 μg/ml), TLR-7 (Imiquimod, 10-1-0.1 μg/ml), TLR-8 (ssRNA40, 10-2-0.4 μg/ml), TLR9 (CpG, 5-1-0.2 μM) RIG-I/MDA5 (Poly(I:C)-LMH-Lyovec, 1-0.2-0.04 μg/ml). Cut-offs for IL-6 and IP-10 ELISA assays are 2.5 pg/ml; there are materialized on panels B and C by dotted lines.

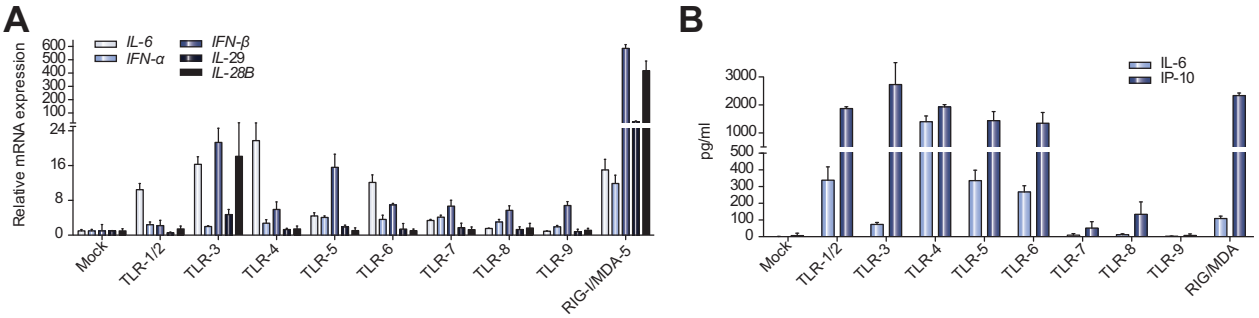


Fig. 4. Induction of IL-6 and IFNs gene expression in PHH cells & secretion of IL-6 and IP-10 after stimulation with prototypic ligands. (A) Gene expression was analyzed (n = 3) after 16 h of stimulation with prototypic ligands of TLR1/2 (Pam3CSK4, 500 ng/ml), TLR3 (Poly(I:C)-LMW, 10 μg/ml) TLR4 (LPS, 1 μg/ml), TLR5 (FLA-BS, 10 μg/ml) TLR6 (FSL1, 5 μg/ml), TLR7 (Imiquimod, 10 μg/ml) TLR8 (ssRNA40, 10 μg/ml), TLR9 (CpG-2395, 5 μM) RIG-I/MDA5 (Poly(I:C)-LMW-lyovec, 1 μg/ml). After stimulation RT-qPCR were performed on *IL-6*, *IFN-β*, *IFN-α*, *IL-28B*, and *IL-29* genes. The RNA level expression was normalized to housekeeping genes and relatively compared to the non-stimulated control. (B) Secretion of IL-6 and IP-10, after 24 h of stimulation with a single dose of indicated ligands, was quantified by ELISA. The concentration used for the ligands were: TLR-1/2 (Pam3CSK4, 500 ng/ml), TLR-3 (Poly(I:C)-LMW, 10 μg/ml), TLR-4 (LPS, 1 μg/ml), TLR-5 (Flagellin, 10 μg/ml), TLR-6 (FSL1, 5 μg/ml), TLR-7 (Imiquimod, 10 μg/ml), TLR-8 (ssRNA40, 10 μg/ml), TLR9 (CpG, 5 μM) RIG-I/MDA5 (Poly(I:C)-LMH-Lyovec, 1 μg/ml). Cut-offs for IL-6 and IP-10 ELISA assays are 2.5 pg/ml.

Research Article

stimulation with increasing doses of prototypic ligands (Fig. 3B, C). Basically, all TLR/RLR ligands tested, but not TLR-7/8/9 ligands, could induce a strong secretion of IL-6 in supernatants in a dose dependent manner. The highest IL-6 secretions (between 500 and 1000 pg/ml) were obtained with TLR-1/2, 4, and 6 ligands used at a high concentration, as expected. Double-stranded RNA ligands, engaging either TLR-3 or RIG-I/MDA-5, led to a lower, but yet significant level of IL-6 production (i.e. between 300 and 400 pg/ml). Despite a good RNA expression induction, TLR-8 ligands did not lead to a strong accumulation of IL-6 in the supernatant (only around 30 pg/ml). This could be due to fast turnover of the secreted cytokine after a very early expression of IL-6 at RNA level. In agreement with RNA expression data, TLR-7 and 9 ligands did not lead to any detectable level of IL-6. The same ligands that were able to induce IL-6 secretion could also increase the IP-10 production. However, TLR-3 and RIG-I/MDA-5 ligands led to the higher production, with a respective secretion of around 2800 and 1400 pg/ml at the highest doses of ligands, as compared to the secretion of around 800 pg/ml with TLR-1/2 or 4, and respectively 200 and 500 pg/ml with TLR-5 and 6 ligands. Similarly to what was obtained with IL-6 secretion, TLR-7, 8, and 9 did not lead to any detectable level of IP-10. Results obtained in PHH are shown in Fig. 4B. In PHH even more IP-10, but less IL-6 (with the exception of TLR-4 agonist stimulation) were produced as compared to dHepaRG.

In contrast, we could only detect biologically active type I IFN, using a functional assay, in the supernatant of HepaRG cell stimulated only with very high (and rather cytotoxic in the context of multiple dosing; data not shown) doses of RIG-I and MDA/5 ligands (Supplementary Fig. 3), thus suggesting that the IFN levels produced after stimulation by other prototypic ligands might be under the detection threshold of the assay.

Exemplary comparison of expression and functionality of TLR-2 in dHepaRG and PHH

As a first attempt to compare the expression and functionality of all TLRs/RLRs in dHepaRg and PHH, we studied TLR-2 expression and function in both cell types using two additional batches of cells. Using Pam3CSK4 to stimulate and a specific TLR-2 siRNA (siRNA against HCV was used as control) to control specificity, we have shown that: i) the basal level of expression (RNA and protein) of this sensor is rather even in both cell types (Fig. 5A, B; Supplementary Fig. 2); ii) its expression can be strongly induced by Pam3CSK4 agonist stimulation and decreased by specific siRNA (Fig. 5A–C); and iii) its specific agonist stimulation by ligands leads to a strong secretion of IL-10 (slightly stronger in PHH as compared to dHepaRG) (Fig. 5D). Thus this pathway seems to be operational in hepatocytes.

Antiviral activity of some TLR/RLR ligands in HBV-infected dHepaRG

To fully characterize the functionality of the TLR/RLR receptors that are expressed in HepaRG cells, we treated cells with TLR-1/2, 3, 4, 7, 8, 9 or RIG-I/MDA-5 ligands with non-toxic doses (as determined by 3× exposure of HepaRG cells to ligands; Fig. 6D) for 24 h before infecting them with HBV. 24 h inoculation with HBV was made in the presence of cytokine produced during the 24 h exposure to ligands. Then, infected cells were washed and further retreated twice with the same dose of ligands

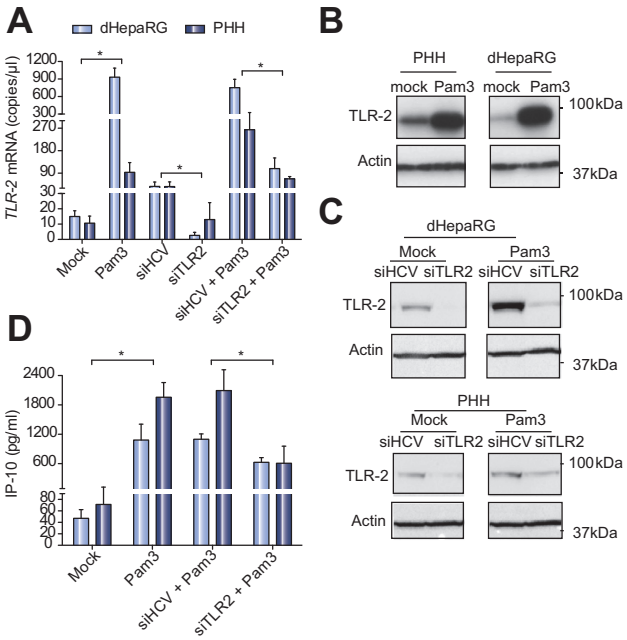


Fig. 5. Comparison of the functionality of TLR-2 in dHepaRG and PHH cells. (A) dHepaRG and PHH were either not or transfected with either siRNA against TLR2 or control siRNA (i.e. targeting HCV) for 24 h, then treated or not for 24 h with Pam3CSK4 (500 ng/ml). Then RNAs were extracted and subjected to RT-qPCR with TLR-2 specific primers. A plasmid containing *TLR-2* gene was used to calculate copies/µL. (B and C) Protein were extracted and subjected to Western blot analysis with an anti-TLR2 and an anti-actin to control loading. (D) ELISA was performed on supernatant for the detection of IP-10. Results for (A and D) are given as a mean ± SEM of two independent experiments (n = 2, but with biological quadruplicate for each experiment) and differences were considered as statistically significant to the control condition when p value was $\leqslant 0.05$ (*).

respectively at days 1 and 4 post-infection. The impact of TLR/RLR ligand treatment on the establishment of HBV infection was analyzed at day 7 post-infection by measuring HBV viral protein secretion (HBsAg/HBeAg) and the accumulation of intracellular HBV DNA by qPCR. A strong antiviral activity was obtained with TLR-1/2 and TLR-4 ligands, as well as with RIG-I/MDA-5 ligand (Fig. 6A–C), whereas weaker or no antiviral activity was observed with TLR-3, 7, and 8 ligands. Surprisingly, an antiviral activity was obtained with the TLR-9 ligand/agonist, which did not correlate with any production of IL-6 or IP-10. However, the same antiviral activity was obtained with the control CpG ligand, thus suggesting that the phenotype obtained was independent of innate functions.

Discussion

Hepatoma cell lines or transformed human hepatocytes are commonly used to study hepatotropic pathogens, however the relevance of these models is limited when studying the interplay between pathogens and hepatocyte innate immunity. Indeed pathogen sensing, IFN response and NF-κB signaling are commonly altered in these cell types compared to PHH [16]. RNA levels for TLR and a set of downstream signaling molecules are lower in these cells compared to PHH, and not modified in response to stimulation. Moreover, we previously showed that

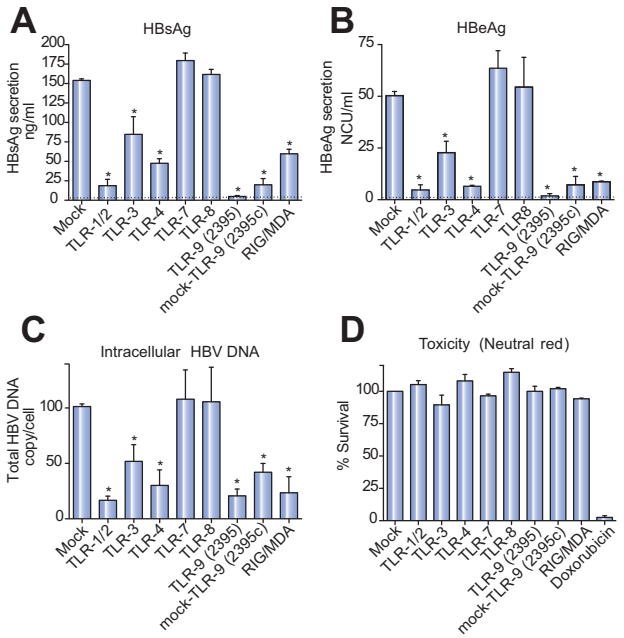


Fig. 6. Effect of stimulation of some PRR receptors by ligands on HBV infection establishment. Secreted HBsAg (A), HBeAg (B), and intracellular HBV DNA accumulation (C) were measured at day 7 post-infection by ELISA and qPCR in HBV-infected dHepaRG cells, which were pre-treated (24 h prior HBV inoculation), then twice treated post-infection (day 1 and 4 post-infection) with indicated ligands. Toxicity (D) was evaluated in the same conditions by the neutral red assay. The concentration used were: 500 ng/ml of Pam3C4 (TLR-1/2), 10 μg/ml of poly(I:C)-LMH (TLR-3), 500 ng/ml of LPS (TLR-4), 5 μg/ml of CL264 (TLR-7), 10 μg/ml of R848 (TLR-7/8), 5 μM of CpG-2395 (TLR-9), 5 μM of CpG-2395 control, 0.01 μg/ml of poly(I:C)-LMW-lyovec (RIG-I/MDA-5). Results are given as a mean ± SEM of three independent experiments (each with biological triplicate) and differences were considered as statistically significant to the control condition when p value was $\leqslant 0.05$ (*). Cut-offs for HBeAg and HBsAg ELISA assays are respectively 1 NCU/ml and 2.5 ng/ml; there are materialized on panels A and B by dotted lines.

a type I IFN response to HBV replication does not lead to an antiviral effect in transformed HepG2 cells while it suppressed viral replication in differentiated HepaRG cells [31]. Furthermore it has been observed that the patterns of cytokine secretion were also different even when a common transcription factor (NF-κB) was involved. The defects in TLR signaling observed in hepatoma cells was consistent with the hypothesis that innate immunity could play an important role in suppressing tumorigenesis [16]. These defects make them of limited interest to study innate immune response induced by a hepatotropic pathogen [14,15].

In contrast, PHH and HepaRG cells, which are mainly non-cancerous, functional for IFN signaling, and permissive to HBV replication, are more relevant models [20,21]. Their transcriptomes have been shown to be closer to each other, as compared to hepatoma cells, and closer to that of liver biopsies [32]. The relatively low replication rate of hepatotropic viruses (low rate for HBV and virtually not permissive to HCV) in HepaRG cells might result, at least in part, from a cellular antiviral response [33]. Indeed, an experimentally strong HBV expression and replication after baculoviral delivery of HBV was shown to induce a potent anti-HBV IFN response, while in the same experimental conditions this IFN response did not show any antiviral effect in transformed HepG2 cells [31]. Therefore HepaRG cells, because physiologically related to PHH [34], were expected to be a valuable tool to study host/pathogen interactions. However, detailed features of this immune response were still unclear and the present study provides the first characterization and functionality of PRRs gene expression in an HBV susceptible and non-cancerous cell line, the HepaRG cell in comparison with PHH.

In this respect, our study showed that RNA of all TLR as well as cytoplasmic helicases RIG-I and MDA5 receptors, could be detected in either proliferative or differentiated HepaRG cells, in similar amount than in PHH. While all RNA could be detected, the level of expression of *TLR-7*, *8*, *9* and *TLR-10* proved to be very low in hepatocytes. Protein expression analysis showed that all TLR were expressed in both dHepaRG and PHH, with slightly higher levels of expression in HepaRG. The detection of TLR-7, 8, and 9 at protein level contrasted with the low RNA expression, and suggests that post-transcriptional mechanisms could be involved to allow detectable protein expression. RIG-I and MDA5 were also expressed at the protein level and highly inducible after interferon stimulation. For most PRR, the stimulation by their prototypic ligands could induce their own expression by positive feed-back. These stimulations could also lead to the RNA induction of several cytokines such as IL-6, with the exception of TLR-7 and 9. However, TLR-7 and 9 ligands could somehow induce a weak *IFN-α* and *β* RNA expression implying that all these TLRs could be slightly functional, yet not leading to detectable secretion of IL-6 or IP-10. In humans, the expression of TLR-7 and TLR9- is mainly confined to plasmacytoid dendritic cells and B-lymphocytes. However, low levels of TLR7 and or TLR9 have been reported in other cell types, including hepatocytes [35] particularly in the setting of chronic inflammation [36]. In addition, the functionality of the TLR-3 and RIG-I/MDA-5 pathway was demonstrated after cell stimulation with extracellular dsRNA (poly(I:C)) and cytoplasmic dsRNA (poly(I:C)-lyovec) respectively showing an induction of IFN-α, β, and λ as well as ISGs (such as ISG56). More importantly, after stimulation with prototypic ligands targeting PRR, using different doses of ligands, IL-6 could be secreted and detected at high levels up to 1000 pg/ml with TLR-1/2, 4, and 6 stimulations, and between 100 and 500 pg/ml for TLR-3, 5, as well as RLR stimulations. A weak IL-6 expression was obtained with TLR-8 stimulation, but no secretion could be shown for TLR-9 and TLR-7 ligands. In contrast, dsRNA ligands engaging either TLR-3 or RIG-I/MDA-5 were more prone to induce secretion of IP-10 (i.e. with production of 2800 and 1400 pg/ml respectively), an IFN-related chemokine. Other functional PRR, based on their ability to induce IL-6 secretion, were also capable to induce IP-10. Interestingly, type I IFN could only be detected in cell supernatants after stimulation with high, yet rather toxic doses of RIG-I and MDA-5 ligands, thus suggesting that it is difficult to measure the production of type I IFN by hepatocytes.

An important part of the demonstration of PRR functionality in HepaRG relies on the antiviral effect observed with some ligands regarding the establishment of HBV infection in these HepaRG cells. Indeed, while it was shown that IL-6 could have an antiviral effect against HBV [37], many other cytokines may also have antiviral effect. Therefore, the engagement of PRR in HepaRG was expected to lead to conditioned media having potent antiviral activity. This was strongly obtained with TLR-1/2 and 4 agonists, as well as with RIG-I/MDA-5 ligands, but surprisingly less with a TLR-3 ligand, despite its ability to bet-

Research Article

ter induce the production of IL-6 and IP-10 as compared to a RIG-I/MDA-5 agonist. This result could only be explained by the ability of RIG-I/MDA-5 agonists to induce measurable amounts of type I IFNs compared to TLR-3 agonists. The strongest antiviral activity was obtained by Pam3CSK4 agonist stimulation of TLR-1/2, a pathway which is well functional in hepatocytes, as shown by more advanced comparison done in both dHepaRG and PHH (Fig. 5). These interesting results warrant further investigation on this pathway and its antiviral effectors.

Infections by hepatotropic pathogens can be cleared once innate immune activation and pathogen-specific T cell responses predominate over pathogen immune escape mechanisms, as it occurs in HAV infections, and in cases of resolution of acute HBV or HCV infections. Viral persistence in hepatocytes can occur by escaping both innate and adaptive immune responses, as observed in chronic HBV or HCV infections. HepaRG exhibits functional TLR and dsRNA-activated signaling pathways in contrast to cultured hepatoma cells [38]. It appears unclear why hepatoma cell lines are unresponsive to a wide range of TLR agonists, but RNA levels for TLR receptor and a set of downstream signaling molecules appear to be lower in hepatoma cell lines than in primary hepatocytes [16]. However, low responsiveness to ligands was observed even when the cognate TLR was expressed at apparently normal levels, suggesting that other downstream factors must be involved. Our data show that HepaRG cells express functional PRRs. The activation of these PRRs leads to the production of cytokines/chemokines, including cytokines bearing antiviral activity. This study opens doors to relevant and detailed analyzes of hepatocyte-like cell response to hepatotropic viruses or other hepatotropic pathogens in the absence of non-parenchymal or liver resident immune cells. Therefore, the HepaRG cell line that maintains the functional properties of PHHs provides a unique tool for the understanding of the interplay between the innate immune response and hepatotropic pathogens. This work may have potential applications to study the activation of hepatocyte innate immunity by these pathogens, the evasion to innate immunity by these pathogens, and mechanism by which innate responses could be restored.

Financial support

This work was supported by grants from ANRS (French national agency for research on AIDS and viral hepatitis; several grants from CSS4), FINOVI (Foundation for innovation in infectiology; project call n°#4), FRM (Foundation for medical research; DEQ20110421327), Hoffmann-La-Roche Ltd (Switzerland division) and by INSERM core grants. This work was also supported by the DEVweCAN LABEX (ANR-10-LABX-0061) of the "Université de Lyon", within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

Conflict of interest

DD and FZ have received research grants from Hoffmann-La-Roche; SF and HJ was/is employees at Hoffmann-La-Roche; SL and MAG were post-doctoral fellows at INSERM at the time the research was performed, but are now employees at Hoffmann-La-Roche.

Authors' contribution

SL, MAG, MM, OF, MG, MB, JL, DD performed experiments; SL, MAG, SF, HJ, FZ, DD designed experiments; SL, MAG, MM, OF, MB, JL, DD analyzed results and drawn figures; SL, MAG, FZ, DD wrote the manuscript.

Acknowledgements

The authors would like to thank Lydie Lefrançois and Judith Fresquet for the isolation of PHHs, as well as the staff from Pr Michel Rivoire's surgery room for providing us with liver resection.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhep.2015.06.022.

References

[1] Protzer U, Maini MK, Knolle PA. Living in the liver: hepatic infections. Nat Rev Immunol 2012;12:201–213.
[2] El-Serag HB. Epidemiology of viral hepatitis and hepatocellular carcinoma. Gastroenterology 2012;142:e1261.
[3] Guidotti LG, Chisari FV. Immunobiology and pathogenesis of viral hepatitis. Annu Rev Pathol 2006;1:23–61.
[4] Takeuchi O, Akira S. Pattern recognition receptors and inflammation. Cell 2010;140:805–820.
[5] Wu J, Chen ZJ. Innate immune sensing and signaling of cytosolic nucleic acids. Annu Rev Immunol 2014;32:461–488.
[6] Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. Cell 2006;124:783–801.
[7] Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat Immunol 2010;11:373–384.
[8] Bowie AG, Haga IR. The role of Toll-like receptors in the host response to viruses. Mol Immunol 2005;42:859–867.
[9] Ait-Goughoulte M, Lucifora J, Zoulim F, Durantel D. Innate antiviral immune responses to hepatitis B virus. Viruses 2010;2:1394–1410.
[10] Heim MH. Innate immunity and HCV. J Hepatol 2013;58:564–574.
[11] Bertoletti A, Ferrari C. Innate and adaptive immune responses in chronic hepatitis B virus infections: towards restoration of immune control of viral infection. Gut 2012;61:1754–1764.
[12] Nishimura M, Naito S. Tissue-specific mRNA expression profiles of human toll-like receptors and related genes. Biol Pharm Bull 2005;28:886–892.
[13] Liu S, Gallo DJ, Green AM, Williams DL, Gong X, Shapiro RA, et al. Role of toll-like receptors in changes in gene expression and NF-kappa B activation in mouse hepatocytes stimulated with lipopolysaccharide. Infect Immun 2002;70:3433–3442.
[14] Melen K, Keskinen P, Lehtonen A, Julkunen I. Interferon-induced gene expression and signaling in human hepatoma cell lines. J Hepatol 2000;33:764–772.
[15] Keskinen P, Nyqvist M, Sareneva T, Pirhonen J, Melen K, Julkunen I. Impaired antiviral response in human hepatoma cells. Virology 1999;263:364–375.
[16] Alexopoulos LG, Saez-Rodriguez J, Cosgrove BD, Lauffenburger DA, Sorger PK. Networks inferred from biochemical data reveal profound differences in toll-like receptor and inflammatory signaling between normal and transformed hepatocytes. Mol Cell Proteomics 2010;9:1849–1865.
[17] Ni Y, Lempp FA, Mehrle S, Nkongolo S, Kaufman C, Falth M, et al. Hepatitis B and D viruses exploit sodium taurocholate co-transporting polypeptide for species-specific entry into hepatocytes. Gastroenterology 2014;146:1070–1083.
[18] Shlomai A, Schwartz RE, Ramanan V, Bhatta A, de Jong YP, Bhatia SN, et al. Modeling host interactions with hepatitis B virus using primary and induced pluripotent stem cell-derived hepatocellular systems. Proc Natl Acad Sci U S A 2014;111:12193–12198.

[19] Parent R, Marion MJ, Furio L, Trepo C, Petit MA. Origin and characterization of a human bipotent liver progenitor cell line. Gastroenterology 2004;126:1147–1156.
[20] Gripon P, Rumin S, Urban S, Le Seyec J, Glaise D, Cannie I, et al. Infection of a human hepatoma cell line by hepatitis B virus. Proc Natl Acad Sci U S A 2002;99:15655–15660.
[21] Gripon P, Diot C, Theze N, Fourel I, Loreal O, Brechot C, et al. Hepatitis B virus infection of adult human hepatocytes cultured in the presence of dimethyl sulfoxide. J Virol 1988;62:4136–4143.
[22] Estornes Y, Toscano F, Virard F, Jacquemin G, Pierrot A, Vanbervliet B, et al. DsRNA induces apoptosis through an atypical death complex associating TLR3 to caspase-8. Cell Death Differ 2012;19:1482–1494.
[23] Khvalevsky E, Rivkin L, Rachmilewitz J, Galun E, Giladi H. TLR3 signaling in a hepatoma cell line is skewed towards apoptosis. J Cell Biochem 2007;100:1301–1312.
[24] Troadec MB, Glaise D, Lamirault G, Le Cunff M, Guerin E, Le Meur N, et al. Hepatocyte iron loading capacity is associated with differentiation and repression of motility in the HepaRG cell line. Genomics 2006;87:93–103.
[25] Maire M, Parent R, Morand AL, Alotte C, Trepo C, Durantel D, et al. Characterization of the double-stranded RNA responses in human liver progenitor cells. Biochem Biophys Res Commun 2008;368:556–562.
[26] Lecluyse EL, Alexandre E. Isolation and culture of primary hepatocytes from resected human liver tissue. Methods Mol Biol 2010;640:57–82.
[27] Hantz O, Parent R, Durantel D, Gripon P, Guguen-Guillouzo C, Zoulim F. Persistence of the hepatitis B virus covalently closed circular DNA in HepaRG human hepatocyte-like cells. J Gen Virol 2009;90:127–135.
[28] Sainz Jr B, Chisari FV. Production of infectious hepatitis C virus by well-differentiated, growth-arrested human hepatoma-derived cells. J Virol 2006;80:10253–10257.
[29] Jammart B, Michelet M, Pecheur EI, Parent R, Bartosch B, Zoulim F, et al. Very-low-density lipoprotein (VLDL)-producing and hepatitis C virus-replicating HepG2 cells secrete no more lipoviroparticles than VLDL-deficient Huh7.5 cells. J Virol 2013;87:5065–5080.
[30] Ank N, Iversen MB, Bartholdy C, Staeheli P, Hartmann R, Jensen UB, et al. An important role for type III interferon (IFN-lambda/IL-28) in TLR-induced antiviral activity. J Immunol 2008;180:2474–2485.
[31] Lucifora J, Durantel D, Testoni B, Hantz O, Levrero M, Zoulim F. Control of hepatitis B virus replication by innate response of HepaRG cells. Hepatology 2010;51:63–72.
[32] Hart SN, Li Y, Nakamoto K, Subileau E-A, Steen D, Zhong X-B. A comparison of whole genome gene expression profiles of HepaRG cells and HepG2 cells to primary human hepatocytes and human liver tissues. Drug Metab Dispos 2010;38:988–994.
[33] Marion MJ, Hantz O, Durantel D. The HepaRG cell line: biological properties and relevance as a tool for cell biology, drug metabolism, and virology studies. Methods Mol Biol 2010;640:261–272.
[34] Guguen-Guillouzo C, Guillouzo A. General review on in vitro hepatocyte models and their applications. Methods Mol Biol 2010;640:1–40.
[35] Lee J, Wu CC, Lee KJ, Chuang TH, Katakura K, Liu YT, et al. Activation of anti-hepatitis C virus responses via Toll-like receptor 7. Proc Natl Acad Sci U S A 2006;103:1828–1833.
[36] Takii Y, Nakamura M, Ito M, Yokoyama T, Komori A, Shimizu-Yoshida Y, et al. Enhanced expression of type I interferon and toll-like receptor-3 in primary biliary cirrhosis. Lab Invest 2005;85:908–920.
[37] Hosel M, Quasdorff M, Wiegmann K, Webb D, Zedler U, Broxtermann M, et al. Not interferon, but interleukin-6 controls early gene expression in hepatitis B virus infection. Hepatology 2009;50:1773–1782.
[38] Lanford RE, Guerra B, Lee H, Averett DR, Pfeiffer B, Chavez D, et al. Antiviral effect and virus-host interactions in response to alpha interferon, gamma interferon, poly(i)-poly(c), tumor necrosis factor alpha, and ribavirin in hepatitis C virus subgenomic replicons. J Virol 2003;77:1092–1104.

Expression and functionality of Toll- and RIG-like receptors in HepaRG cells

Souphalone Luangsay, Malika Ait-Goughoulte, Maud Michelet, Océane Floriot, Marc Bonnin, Marion Gruffaz, Michel Rivoire, Simon Fletcher, Hassan Javanbakht, Julie Lucifora, Fabien Zoulim, David Durantel

Contents

Supplementary material and methods

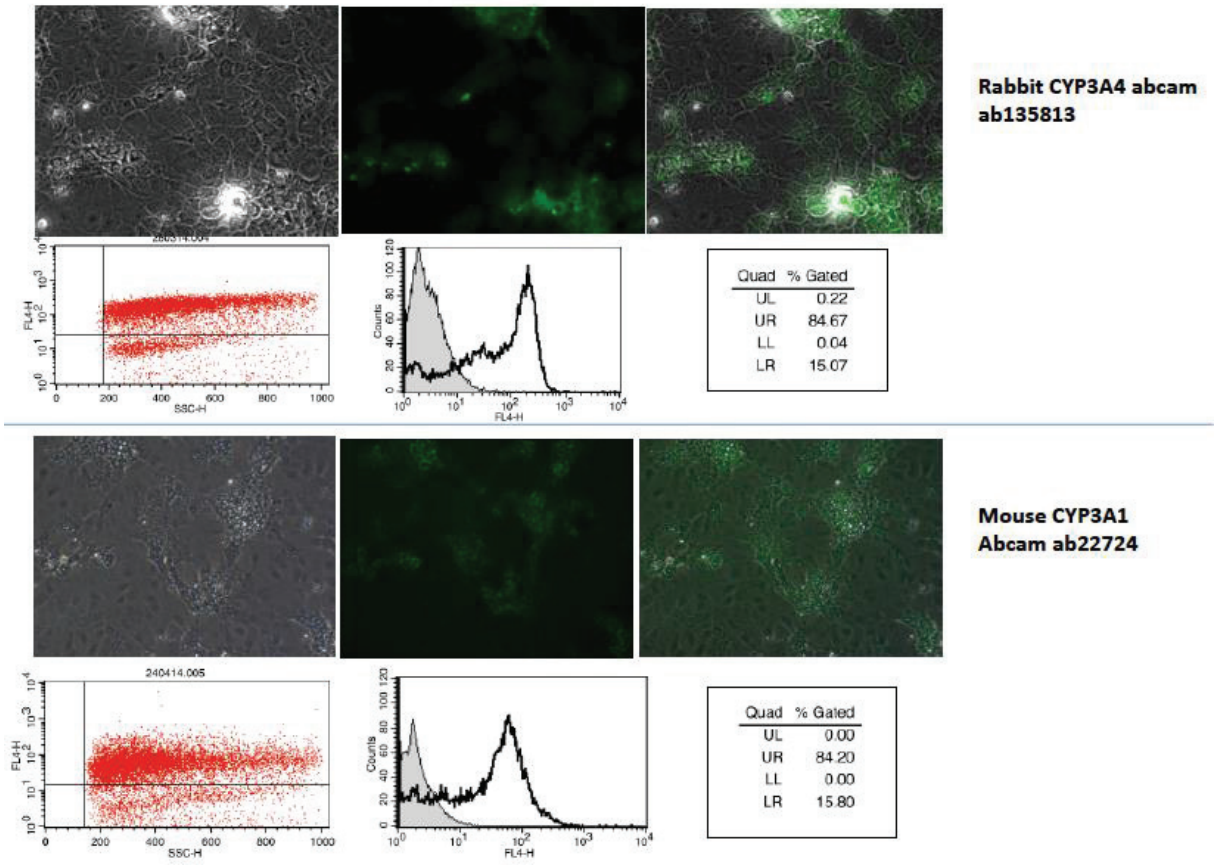
Analysis of Secreted Type I Interferon

Two millions of Huh7.5 cells were transfected with 10 µg pISRE-Luc vector (Stratagene) in a 10 cm diameter dish using Mirus TransIT-LT1 transfection reagent according to manufacturer's instructions (Mirus Bio LLC). The pISRE-Luc plasmid (plasmid containing the luciferase gene under the control of the Interferon Stimulated Response Element) expresses luciferase under type-I IFN inducible promoter. After a 16h incubation with transfection mixture, cells were trypsinized and reseeded in a 96-well plate at approximately 3.10^4 cells/well in a volume of 50 µL. Six hours later, 50 µL of conditioned HepaRG supernatants, previously stimulated by various ligands at various concentrations as indicated. After 24h at 37°C, cells were washed with phosphate-buffered saline and lysed before luciferase activity was monitored using the Renilla Luciferase Assay System (Promega).

Supplementary figures and figure legends

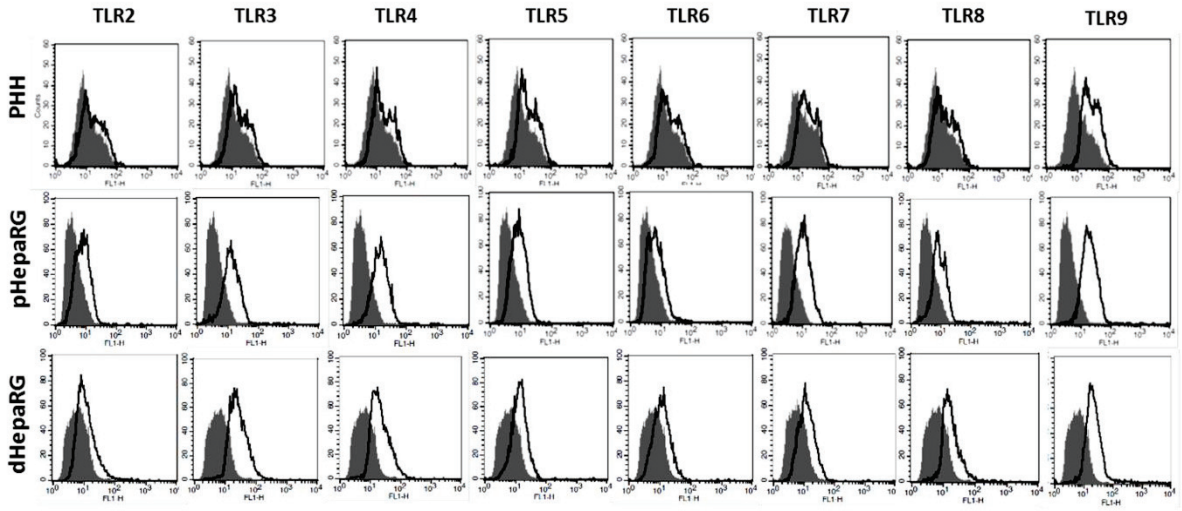
Supplementary-Fig. 1. Cyp stainings by FACS and immunofluorescence in differentiated HepaRG.

The expression of CYP3A4 (upper panel) and CYP3A1 (lower panel) was assessed by immunofluorescence and FACS in dHepaRG to evaluate the hepato-specificity and the percentage of hepatocyte-lie cells in a monolayer of dHepaRG cells.



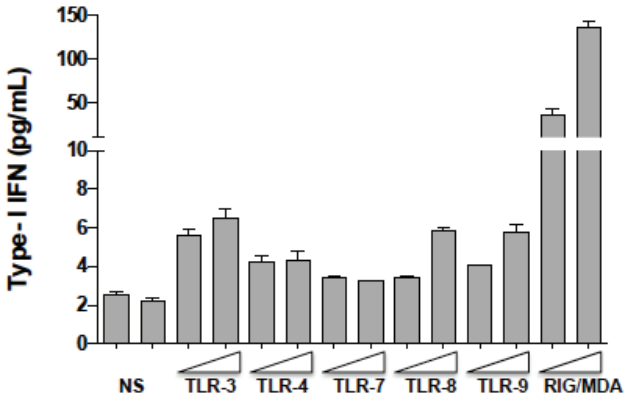
Supplementary-Fig. 2. Protein expression of PRR in pHepaRG, dHepaRG and PHH cells.

TLR expression analyzed by FACS. Black histograms represent stainings with the indicated TLR antibody, shaded gray histograms represent the isotype control, in pHepaRG, dHepaRG and PHH.



Supplementary-Fig. 3. Analysis of secreted type-I IFN upon prototypic ligand treatment.

Differentiated HepaRG cells were stimulated with increasing doses of PRR ligands (TLR-3 (Poly(I:C)-LMW, 1 and 10 µg/mL), TLR-4 (LPS, 0.1 and 1 µg/mL), TLR-7 (Imiquinod, 10-1-0.1 µg/mL), TLR-8 (ssRNA40, 10-2-0.4 µg/mL), TLR9 (CpG, 5-1-0.2 µM), and RIGI/MDA5 (Poly(I:C)-LMH-Lyovec, 0.2 and 1 µg/mL)). Twenty four hours after stimulation, supernatant were tested along with non stimulated cells (NS) for type I interferon activity following procedure described in supplementary materials and methods.



b. Direct antiviral properties of TLR ligands againts HBV replication in immune-competent hepatocytes

Direct antiviral properties of TLR ligands against HBV replication in immune-competent hepatocytes

Julie Lucifora1, Marc Bonnin1, Ludovic Aillot1, Floriane Fusil2, Sarah Maadadi1, Laura Dimier1 ,Maud Michelet1, Océane Floriot1, Anaïs Ollivier2, Michel Rivoire3, Malika Ait-Goughoulte4, Stéphane Daffis5, Simon P. Fletcher5, Anna Salvetti1, François-Loïc Cosset2, Fabien Zoulim1,6 & David Durantel1

Scientific Reports | *(2018) 8:5390* | Doi:10.1038/S41598-018-23525-W

SCIENTIFIC REPORTS

OPEN

Direct antiviral properties of TLR ligands against HBV replication in immune-competent hepatocytes

Received: 25 October 2017
Accepted: 13 March 2018
Published online: 29 March 2018

Julie Lucifora[1], Marc Bonnin[1], Ludovic Aillot[1], Floriane Fusil[2], Sarah Maadadi[1], Laura Dimier[1], Maud Michelet[1], Océane Floriot[1], Anaïs Ollivier[2], Michel Rivoire[3], Malika Ait-Goughoulte[4], Stéphane Daffis[5], Simon P. Fletcher[5], Anna Salvetti[1], François-Loïc Cosset[2], Fabien Zoulim[1,6] & David Durantel[1]

Current therapies for chronic hepatitis B virus (HBV) infections are effective at decreasing the viral load in serum, but do not lead to viral eradication. Recent studies highlighted the therapeutic or "adjuvant" potential of immune-modulators. Our aim was to explore the direct anti-HBV effect of Toll-Like-Receptors (TLR) agonists in hepatocytes. HBV-infected primary human hepatocytes (PHH) or differentiated HepaRG cells (dHepaRG) were treated with various TLR agonists. Amongst all TLR ligands tested, Pam3CSK4 (TLR1/2-ligand) and poly(I:C)-(HMW) (TLR3/MDA5-ligand) were the best at reducing all HBV parameters. No or little viral rebound was observed after treatment arrest, implying a long-lasting effect on cccDNA. We also tested Riboxxol that features improved TLR3 specificity compared to poly(I:C)-(HMW). This agonist demonstrated a potent antiviral effect in HBV-infected PHH. Whereas, poly(I:C)-(HMW) and Pam3CSK4 mainly induced the expression of classical genes from the interferon or NF-κB pathway respectively, Riboxxol had a mixed phenotype. Moreover, TLR2 and TLR3 ligands can activate hepatocytes and immune cells, as demonstrated by antiviral cytokines produced by stimulated hepatocytes and peripheral blood mononuclear cells. In conclusion, our data highlight the potential of innate immunity activation in the direct control of HBV replication in hepatocytes, and support the development of TLR-based antiviral strategies.

Hepatitis B virus (HBV) is a small DNA virus that persists within hepatocytes thanks to the establishment and maintenance of a covalently-closed-circular DNA (cccDNA). This cccDNA serve as the main template for viral RNA synthesis, including the pre-genomic RNA (pgRNA), which is subsequently converted into relaxed-circular DNA (rcDNA) by a HBV polymerase-mediated reverse-transcription step taking place inside nucleocapsids. Different viral particles and antigens circulate in the blood of infected patients, including HBe antigens (HBeAg), Dane particles (infectious particles or virions), spheres and rod (empty enveloped particles); the later three features envelope proteins at their surface and all comprise the pool of secreted HBs antigens (HBsAg). Nucleic acid-free subviral particles are produced in large excess compared to virions, and therefore are thought to play an important role in terms of immune subversion[1].

With around 250 million people chronically infected, who are at high risk to develop severe liver diseases, such as cirrhosis or hepatocellular carcinoma (HCC), HBV infection remain a major medical burden worldwide. Current therapies for chronic HBV infections, mainly relying on nucleos(t)ides analogues (e.g. entecavir, tenofovir…), are effective at suppressing viremia in blood of patients and somehow improve long-term outcome, but have only low rates of HBsAg loss, with or without associated anti-HBs seroconversion, and typically do not lead to cccDNA elimination[2]. Moreover, there are life-long treatments, as drug administration arrest almost

[1]INSERM, U1052, Cancer Research Center of Lyon (CRCL), Université de Lyon (UCBL1), CNRS UMR_5286, Centre Léon Bérard, Lyon, France. [2]CIRI – International Center for Infectiology Research, Inserm, U1111, Université Claude Bernard Lyon 1, CNRS, UMR5308, Ecole Normale Supérieure de Lyon, Univ Lyon, F-69007, Lyon, France. [3]INSERM U1032, Centre Léon Bérard (CLB), Lyon, France. [4]Roche Pharma Research and Early Development (pRED), Roche Innovation Center Basel, F. Hoffmann-La Roche, 4070, Basel, Switzerland. [5]Gilead Sciences, Inc, Foster City, CA, USA. [6]Department of Hepatology, Croix-Rousse Hospital, Hospices Civils de Lyon, Lyon, France. Julie Lucifora and Marc Bonnin contributed equally to this work. Correspondence and requests for materials should be addressed to J.L. (email: julie.lucifora@inserm.fr) or D.D. (email: david.durantel@inserm.fr)

universally leads to a rebound in viremia and liver diseases. This warrants the identification of new antiviral strategies, including immune-therapeutic components, to improve the *functional cure* rates and decrease further the risk of end-stage liver diseases.

Infection of human cells by microorganisms initially leads to the activation of the host innate immune response through a sensing mediated by pattern recognition receptors (PRR). PRR include Toll-like receptors (TLR), C-type lectin receptors (CLR), RIG-I-like receptors (RLR) and NOD-like receptors (NLR), intracellular DNA sensors and cytoplasmic RNA helicases such as RIG-I (Retinoic-acid Inducible Gene-I) and MDA-5 (melanoma differentiation-associated gene-5)[3]. Each PRR detects pathogen-associated molecular patterns (PAMP) derived from viruses, bacteria, mycobacteria, fungi or parasites and activates downstream signaling events leading to specific gene expression programs and the secretion of interferons (IFN), inflammatory cytokines/chemokines, and other antimicrobial peptides[3]. Many IFNs and pro-inflammatory cytokines have been shown to have direct anti-HBV effect in hepatocytes[4]. The use of PRR agonists, in particular TLR7-L (TLR7-Ligands), to induce endogenous IFNs and other cytokines, has been successfully applied to animal models of HBV infection. It is worth noting that the sub-cutaneous injection of Pegylated-IFN-α is currently a therapeutic option for treating CHB patients, although it is associated with many adverse effects[5]. The induction of endogenous IFN-α might therefore be of interest to increase the antiviral effect and lower side effects. It was shown both in HBV-infected chimpanzees and WHV-infected woodchucks that the orally delivered TLR7 agonist GS-9620 significantly reduced viremia and HBsAg[6–8]. Moreover, GS-9620 also impacted on cccDNA expression, and led to anti-HBs seroconversion in WHV-infected woodchucks[6–8]. In human, GS-9620 was shown to be safe[9], but its anti-HBV effect was modest at used doses (AASLD 2017). TLR9 ligands are likely to be used in the HBV field as adjuvants for HBV prophylactic vaccination[10,11], but have also been recently shown to induce a weak antiviral activity in mono-therapy in animal models of HBV infection[12,13]. However, neither TLR7 nor TLR9 are expressed in hepatocytes[14] suggesting that their antiviral effect results from the activation of non-hepatocyte cells, such as plasmacytoid dendritic cells, and the endogenous action of IFN and other cytokines. We previously showed that hepatocytes (primary human hepatocytes (PHH) and differentiated HepaRG cells (dHepaRG)) express a number of other PRR and that their agonization could inhibit HBV replication when the ligands were applied before viral infection and maintained throughout the experiment[14]. To provide further insight into the potential of TLR agonists to induce efficient and durable antiviral effect in already established HBV infections, we tested the direct anti-HBV effect of different TLR agonists in HBV-infected hepatocytes *in vitro*, and extended the testing for TLR2-L and TLR3-L *in vivo*.

Results

Antiviral effect of various TLR agonists *in vitro*. To assess the effects of TLR stimulation on established HBV infection, HBV-infected PHH and dHepaRG cells were treated with Pam3CSK4 (TLR-1/2 agonist), poly(I:C)-HMW (TLR3-agonist), LPS (TLR-4 agonist), FLA-BS (TLR-5 agonist), FSL-1 (TLR-2/6-agonist), Imiquimod (ImiQ, TLR-7 agonist), CL264 (TLR-7 agonist), ssRNA (TLR-8 agonist) or 2395 (CpG ODN, TLR-9 agonist). As expected from our previous study and the pattern of TLR expression in dHepaRG and PHH[14], cells produced significant amount of IL-6 (Figs 1A,D and S1) and IP-10 (Fig. 1A and D) mainly in response to Pam3CSK4, poly(I:C)-(HMW), LPS and FSL. Pam3CSK4, poly(I:C)-(HMW) and LPS treatments significantly decreased both total intracellular HBV DNA (Fig. 1B and E) and HBeAg secretion (Fig. 1C and F) in both HBV-infected dHepaRG cells or PHH. Treatments with FSL induced a 55–65% reduction of total intracellular HBV DNA in both HBV-infected dHepaRG and PHH (Fig. 1B), but only a significant effect on HBeAg secretion in PHH (Fig. 1C and F). As hepatocytes do not express TLR7[14], TLR7 ligands (Imiquimod; ImiQ or CL264) induced neither IL-6 nor IP-10 following stimulation of dHepaRG. However, surprisingly, ImiQ treatment decreased total intracellular HBV DNA (Fig. 1B) as well as secretion of HBeAg (Fig. 1C) in dHepaRG, although not in PHH. This phenotype in dHepaRG was nevertheless independent of an innate response, as no production of cytokines was detected (Fig. 1A and D); we therefore did not pursue investigations on this agonist *in vitro*, as the phenotype in PHH is more relevant and likely more predictive as compared to that obtained with dHepaRG cells. As Pam3CSK4 and poly(I:C)-(HMW) were the most potent inhibitors of HBV in both *in vitro* models (Fig. 1), we further focused on the characterization of their antiviral effects for the rest of the study.

Dose-response antiviral effect, in the absence of toxicity, of Pam3CSK4 and poly(I:C)-(HMW). Both Pam3CSK4 and poly(I:C)-(HMW) induced a dose-dependent reduction of total intracellular HBV DNA, as well as extracellular HBeAg and HBsAg secretions without significant toxicity (Fig. S2, panels A and B). It is worth noting that, despite an absence of cytotoxicity, a significant reduction in apolipoprotein B secretion suggests that Pam3CSK4 and poly(I:C)-(HMW) could altered hepatocyte differentiation status (Fig. S2, bottom graphs of panel A). However, this phenomenon was similar to that observed with IFN-α treatment, which is used to treat CHB patients.

Long lasting antiviral effect of Pam3CSK4 and poly(I:C)-(HMW). In contrast to the rebound observed after the arrest of nucleos(t)ides analogues[15], reduction of total intracellular HBV DNA, HBeAg and HBsAg secretions after Pam3CSK4 or poly(I:C)-(HMW) treatment of dHepaRG cells were sustained for at least 10 days after the end of treatment (Fig. 2A–C). Additionally, using qPCR or Southern blot analyses, we observed a decrease in cccDNA levels in HBV-infected dHepaRG treated with Pam3CSK4 or poly(I:C)-(HMW) (Figs 2D and S3). Importantly, this decrease of cccDNA following Pam3CSK4 or poly(I:C)-(HMW) treatment could be confirmed in two out of three experiments using three different donors of PHH (Fig. S4A).

Riboxxol: a novel TLR3 agonist with potent anti-HBV activity. As poly(I:C)-(HMW) can also activate MDA5 (Fig. 3A), we tested Riboxxol, a synthetic 50 base pair double-stranded RNA, which is a specific ligand of TLR3 (Figs 3A and S5). Similarly to poly(I:C)-(HMW), treatments of HBV-infected PHH with Riboxxol

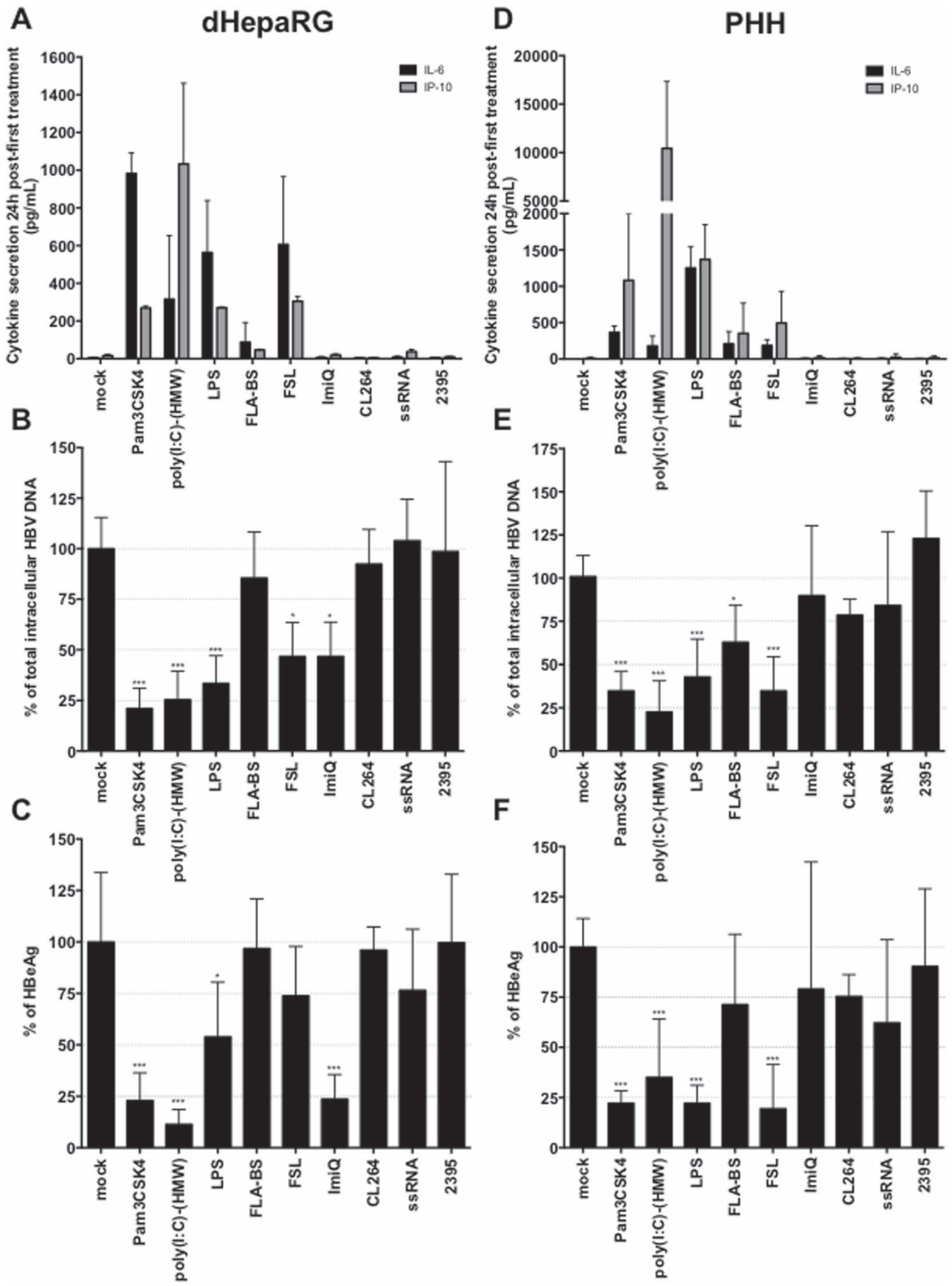


Figure 1. Antiviral activity of TLR-agonists in HBV-infected primary human hepatocytes (PHH) and differentiated HepaRG cells (dHepaRG). (**A**–**C**) dHepaRG or (**D**–**F**) PHH were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for (**A**–**C**) seven days or (D–F) 4 days, then treated twice with the indicated molecules (see method part for concentrations) for seven additional days. (**A**,**D**) Twenty-four hours after the first treatment, a fraction of cell supernatants was collected, and IL-6 and IP-10 contents were quantified by ELISA. (**B**,**E**) At the end of treatments, cells were harvested, total DNAs extracted, and total intracellular HBV DNA amounts quantified by qPCR. (**C**,**F**) Supernatants were also collected and HBeAg levels assessed by ELISA. Results are the mean +/− SD of three independent experiments (three batches of dHepaRG and three donors of PHH), each performed with three biological replicates.

dose-dependently decreased total intracellular HBV DNA, as well as extracellular HBeAg and HBsAg (Fig. 3B). Of note, a weak decrease in cccDNA levels was observed in two out of four experiments using four different donors of PHH (Fig. S4B), thus suggesting a weaker activity on this marker as compared to poly(I:C)-(HMW).

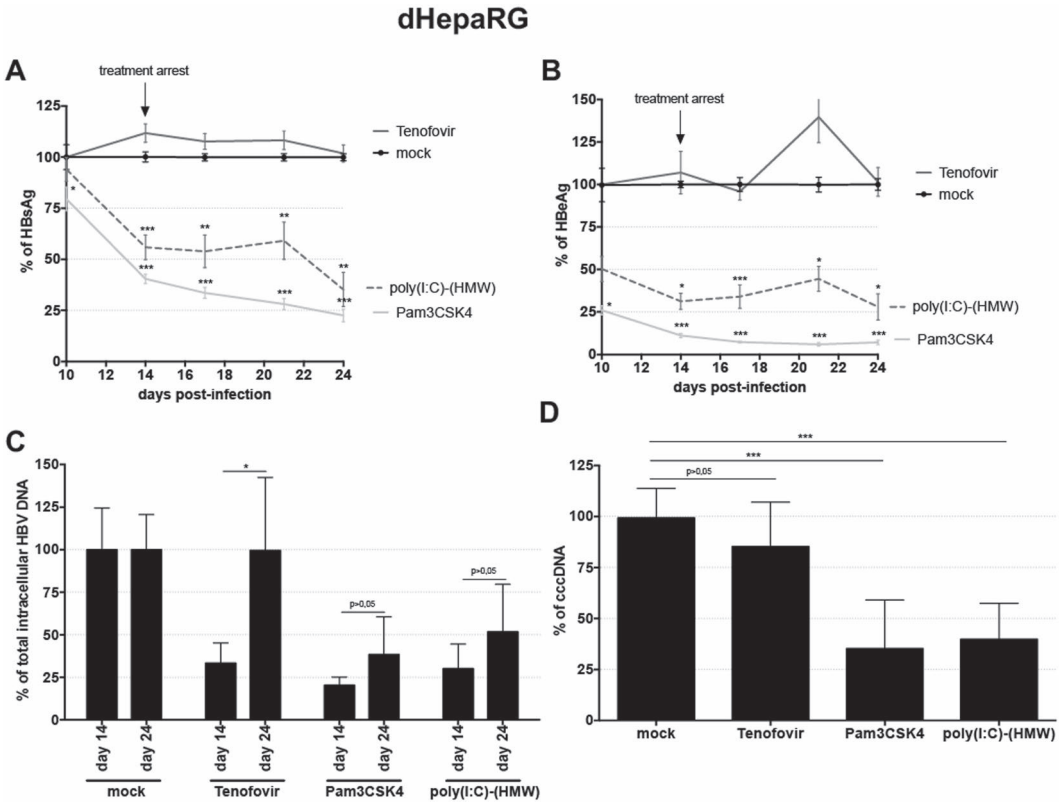


Figure 2. Inhibition of HBV by Pam3CSK4 and poly(I:C)-(HMW) is maintained after treatment arrest in dHepaRG cells. dHepaRG cells were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven days, treated twice with the indicated molecules for seven days, and further cultured without treatment for 10 additional days. Supernatant were collected at the indicated times and (**A**) HBsAg and (**B**) HBeAg levels assessed by ELISA. Cells were harvested at (**C**) day-14 or (**C**,**D**) day-24 post-infection, total DNAs extracted, and (**C**) total intracellular HBV DNAs or (**D**) cccDNA amounts quantified by qPCR. Results are the mean +/− SD of at least three independent experiments each performed with three biological replicates.

Interestingly, besides its resistance to body fluids and its relevant physiochemical characteristics, which allow *in vivo* injection[16], Riboxxol did not affect hepatocyte viability and differentiation (assessed by ApoB secretion), in contrast to Pam3CSK4 and poly(I:C)-(HMW) ligands (Figs 3B and S6).

Kinetics of inhibition and pathways involved. The antiviral effect of Pam3CSK4 was rapid, with a 50% reduction of total HBV RNAs in HBV-infected dHepaRG or PHH (measured at 24 h or 48 h post-treatment, respectively) with a single administration of the compound (Fig. 4A,D). The reduction in HBV RNAs, following a single dose of a TLR3 ligand, was less efficient and inhibition of shorter duration in HBV-infected dHepaRG (Fig. 4A,D). As expected Pam3CSK4 treatment induced a transient increase of IL-6 mRNA, but not OAS1 mRNA levels, two prototypic genes of NF-κB and IFN pathways. In contrast, poly(I:C)-(HMW) mostly induced a transient increase of OAS mRNA level (Fig. 4B,C,E,F). Interestingly, Riboxxol induced similar (or higher) OAS mRNA level relative to poly(I:C)-(HMW), but slightly higher IL-6 mRNA level (2 fold more), as compared to poly(I:C)-(HMW) (Fig. 4B,C,E,F). This correlated well with the pattern of IL-6 and IP-10 secretion (Fig. S7). Altogether, those data indicate that Riboxxol and poly(I:C)-(HMW) can activate both NF-κB and type-I IFN pathways.

The antiviral effect of TLR2 and TLR3 ligands in hepatocyte cultures is not due to type-I IFN and IL-6 production. IL-6 is the main cytokine produced by hepatocytes in response to TLR1/2 or TLR3 ligands (Fig. S1). It was previously shown that IL-6 and/or type-I IFN have antiviral activity against HBV[4,15,17–19], which was confirmed here (Fig. 5A and B). In order to investigate if secretion of IL-6 and/or type-I IFN by the stimulated hepatocytes contribute to the observed antiviral phenotypes of Pam3CSK4, poly(I:C)-(HMW) or Riboxxol, dHepaRG cells were treated with these ligands together with blocking antibodies against IL-6, IFNAR1 or both. We confirmed that the amount of anti-IL-6 antibody used was able to fully neutralize (Fig. 5A) the antiviral effect of a dose of rIL-6 hundred times higher than that secreted by dHepaRG cells (Figs 5A and S1,S7). The amount of anti-IFNAR1 antibody used was able to substantially decrease the antiviral effect of a dose of rhPEG-IFN-α at least twenty times higher than the level of type-I IFN produced by hepatocytes (Figs S1 and 5B). The anti-IFNAR1

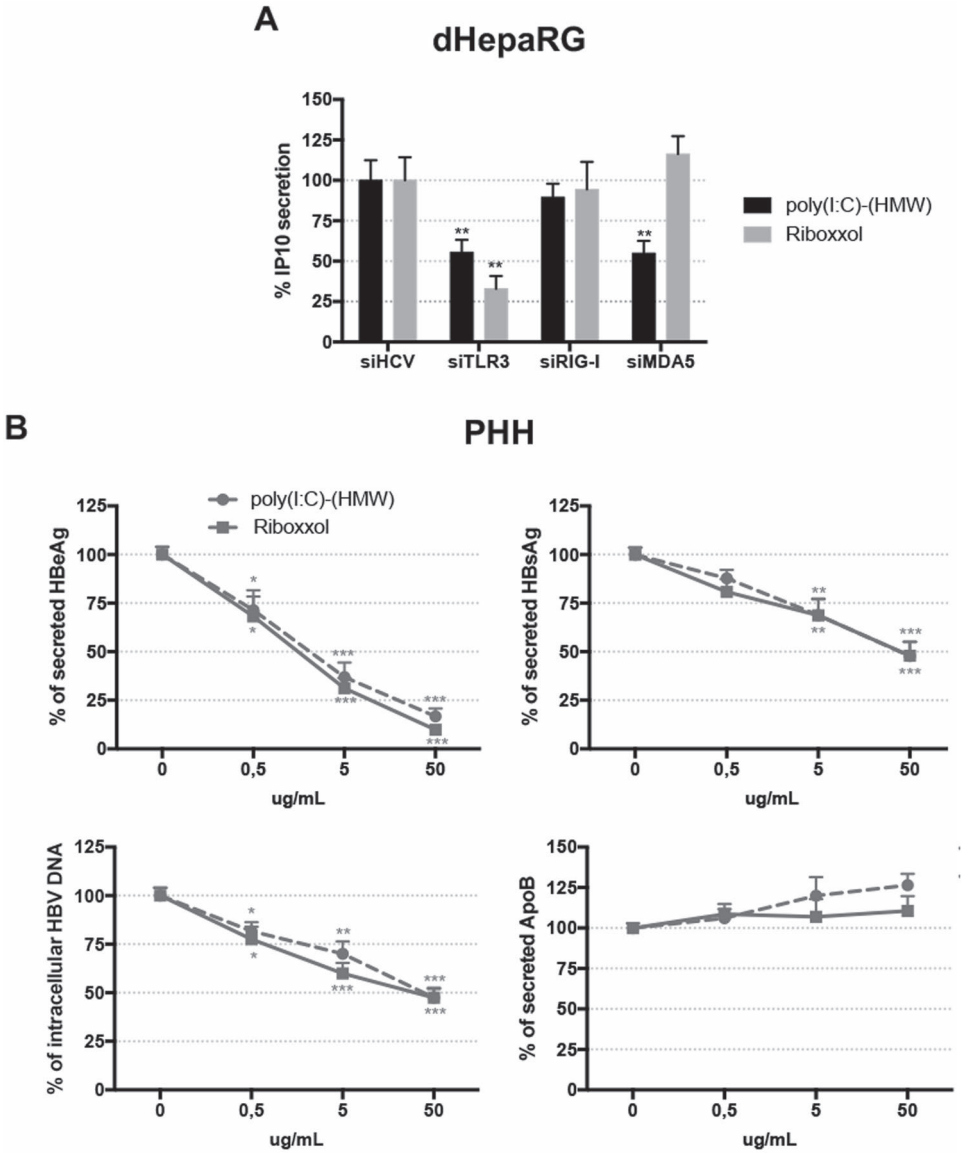


Figure 3. Inhibition of HBV by PolyIC(HMW) and Riboxol. (**A**) Uninfected dHepaRG cells were transfected with the indicated siRNA. 14 days later, cells were stimulated with 10 µg/ml of the indicated ligands. Cells were lyzed 24 h later and IP10 secretions were assessed by ELISA. Results are the mean +/− SD of three independent experiments each performed in technical duplicates. (**B**) PHH were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven days and treated twice with poly(I:C)-(HMW) or Riboxxol at the indicated concentration for seven days. At the end of the experiments, supernatants were collected, and HBeAg, HBsAg and ApoB levels assessed by ELISA. Cells were harvested, total DNAs extracted and total intracellular HBV DNA amounts were quantified by qPCR. Results are the mean +/− SD of four independent experiments (four donors of PHH) each performed with three biological replicates.

antibody did not rescue the antiviral effect of Pam3CSK4, poly(I:C)-(HMW), or Riboxxol (Fig. 5C–E). The neutralizing anti-IL-6 antibody only induced a slight rescue of HBeAg when combined with Pam3CSK4 or Riboxxol (Fig. 5C–E). These data indicate that hepatocyte-produced IL-6 or type-I IFN are not the main drivers of the antiviral effect of Pam3CSK4, poly(I:C)-(HMW) or Riboxxol.

Cytokines produced by immune cells upon TLR2 or TLR3 stimulation and antiviral effect in mice models. If IL-6 and type-I IFNs production by hepatocytes do not seem to be solely involved in the antiviral effect in our experimental conditions, their production in the liver microenvironment by other cells types are expected to be beneficial for the overall antiviral effect we can anticipate *in vivo*. To this end, we demonstrated that treatment of human peripheral blood mononuclear cells (PBMCs) with Pam3CSK4 strongly induced IL-6, as well as some IL-1β and TNF-α (Fig. 6A). Poly(I:C)-(HMW) induced less IL-6 but more IFN-γ and IFN-α whereas Riboxxol induced all these cytokines, which are known to inhibit HBV[4,15,17–20] (Fig. 6A). As expected HBeAg

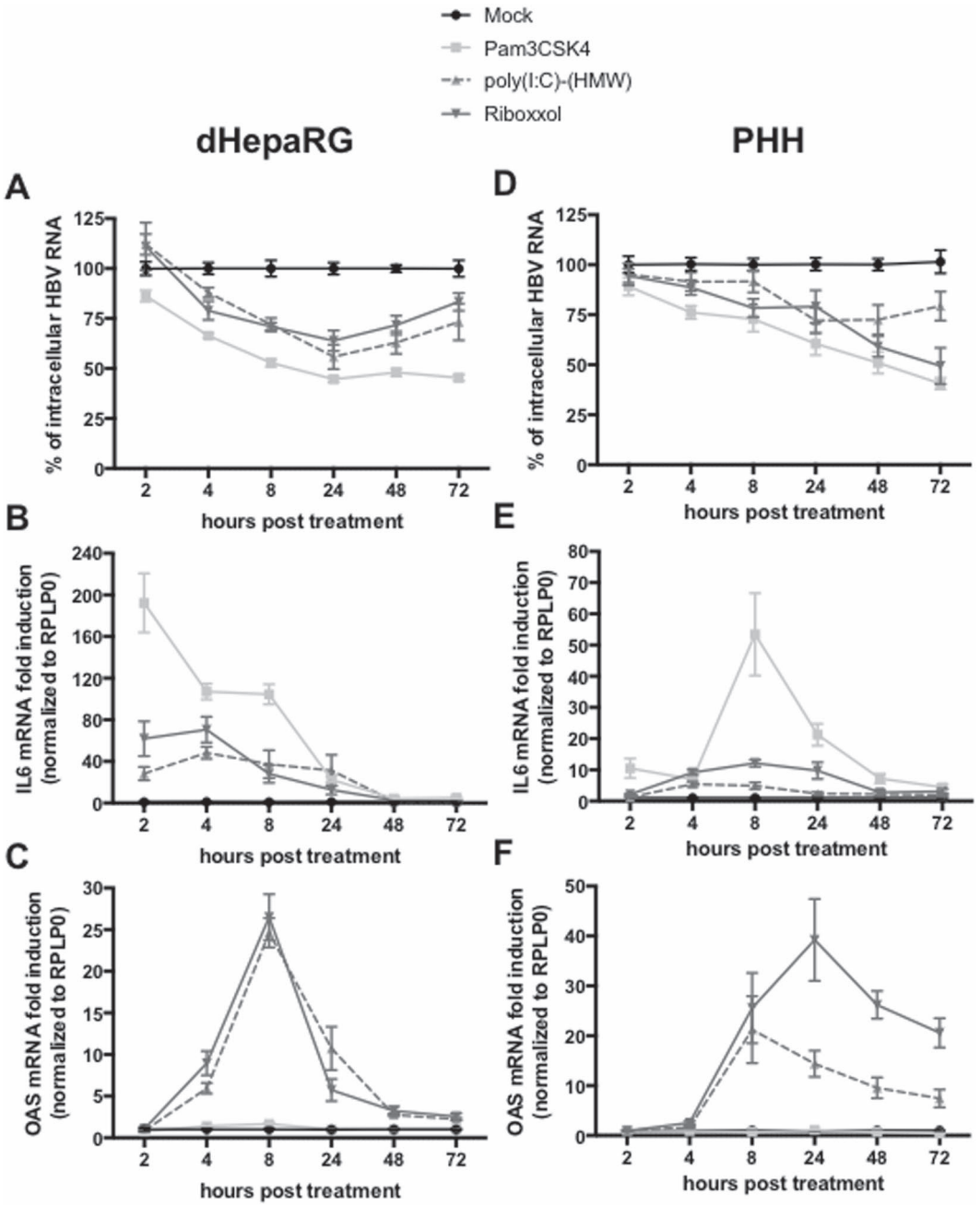


Figure 4. Riboxxol induced both type I IFN and NFkB pathways. (**A**–**C**) dHepaRG cells or (**D**–**F**) PHH were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven days and treated twice or not with Pam3CSK4 (1 ug/mL), poly(I:C)-(HMW) (20 ug/mL for dHepaRG, 5 ug/mL for PHH) or Riboxxol (50 ug/mL for dHepaRG, 5 ug/mL for PHH) for seven days. At the indicated time post-treatment, cells were harvested, (**A**,**D**) total RNAs extracted and levels of HBV RNA, (**B**,**E**) IL-6 mRNA or (**C**,**F**) OAS mRNA assessed by RT-qPCR. Results are the mean +/− SEM of three independent experiments (three batches of dHepaRG and three donors of PHH) each performed with three biological replicates.

secretion by HBV-infected dHepaRG was reduced in the presence of supernatants from these stimulated PBMCs (Fig. 6B). Of note, the amount of cytokines produced by PBMCs and their antiviral effect on HBV-infected hepatocytes may vary from one donor to another (compared Figs 6 to S8). To assess the effect of TLR2-L or TLR3-L *in vivo*, HBV-infected liver-humanized (HuHep) mice mice were intravenously (IV) injected 2 times per week during 3 weeks with Pam3CSK4 or Riboxxol. The dose escalation protocol we have chosen led to very moderate decreases of HBV parameters in this animal model (Fig. S9A) suggesting that either higher doses or a liver-specific delivery of these ligands have to be further tested.

Discussion

Toll-like receptors are important molecular mediators linking innate and adaptive immunity, and their stimulation by cognate agonists induced an antiviral response in animal models of HBV infection. Indeed, treatments of chimpanzees or woodchucks with TLR7 or TLR9 agonists led to reduction of HBV replication markers[6–8,12]. The strongest antiviral phenotype was obtained with GS-9620, a TLR-7 agonist, in the woodchuck model, with a strong effect of treatment on cccDNA and a long-lasting efficacy associated with an anti-HBsAg seroconversion. These antiviral effects probably resulted from the activation of non-parenchymal cells and subsequent production of anti-HBV cytokines/IFNs, since hepatocytes do not express TLR7 or TLR9 and cannot therefore be directly activated by these agonists[14]. However, hepatocytes express a number of other TLRs and we investigated here if their stimulation with specific agonists could directly inhibit an already established HBV infection in isolated hepatocytes (i.e., in the absence of immune cells).

We demonstrated that ligands of TLR1/2 (Pam3CSK4) and TLR3 (poly(I:C)-(HMW) or Riboxxol activated hepatocytes (PHH or dHepaRG) innate responses and efficiently decreased levels all HBV replication markers, including a strong phenotype on HBV RNAs. This suggests that those treatments are able to prevent transcription of viral RNAs from cccDNA and/or to interfere with their stability. Underlying mechanisms are likely to involve either direct negative epigenetic regulations on cccDNA or a direct IRF3-mediated (i.e., IFN-independent) ISG induction, which would be independent from the production of either IL-6 or IFN-α and their secondary effect on cells via downstream pathways[17,18], as shown by our neutralization studies. As HBV-infected dHepaRG cells are able to maintain HBV replication for months[21], we could show that Pam3CSK4 and poly(I:C)-(HMW)-mediated antiviral effect was sustained after arrest of the treatments contrary to the antiviral effect of nucleos(t)ides analogues, thus suggesting a role of epigenetic regulations. In addition, this lack of (or weak) rebound was associated to a slight, but significant, reduction of cccDNA levels measured by both qPCR and Southern blotting in dHepaRG. This suggests that Pam3CSK4 and poly(I:C)-(HMW) could trigger cellular pathways that can subsequently affect cccDNA stability, as shown previously with some cytokines or lymphotoxin-beta agonists[15,20]. It is nevertheless worth noting that the effect on cccDNA of Pam3CSK4, poly(I:C)-(HMW) or Riboxxol could not be confirmed in all PHH donors tested suggesting that (i) either dHepaRG are particularly efficient to trigger innate cellular pathways and/or (ii) the low level of replication of HBV in dHepaRG is more prone to cccDNA degradation and/or (iii) PHH donor heterogeneity (e.g. genetic background, alteration of liver function due the underlying donor disease or treatment) may affect pathways triggering cccDNA destabilization. This also emphasizes the importance of performing experiments with multiple PHH donors. In addition to the direct antiviral effect observed in infected hepatocytes, Pam3CSK4, poly(I:C)-(HMW) and Riboxxol also activated PBMCs to produce different cytokines, which could in turn inhibit HBV replication. Among them, IL-6 might be the main driver of this indirect antiviral effect against HBV since its lower induction in PMBCs from donor 2 stimulated with Riboxxol led to a lower inhibition of HBeAg secretion (Fig. S8B).

Our data are in accordance with experiments performed in HBV transgenic mice or WHV infected woodchuck that identified respectively TLR3 and TLR2 pathways as potential therapeutic targets[22–24]. However, in contrast to agonists of TLR7 (e.g., GS-9620), TLR8 (e.g., GS-9688) and RIG-I/NOD2 (e.g., SB-9200) that can be delivered orally to human, TLR3 and TLR2 ligands cannot and will likely require strategy for delivery since their IV administration with a dose escalation protocol (from 20 to 80 ug during 3 weeks) only led to very moderate decreases of HBV parameters in HBV-infected HuHep mice. In this respect, nanoparticles will be used to reduce the active dose of ligands, protect these ligands from degradation and specifically deliver specifically them to the liver, thereby preventing systemic exposure and potential adverse effects[25]. These ligands may also be used as part of a combination therapies with currently, e.g. with currently used anti-HBV drugs, as well as with immunotherapies, such as therapeutic vaccines and/or checkpoint inhibitors.

Methods

Cell culture and HBV infection. HepaRG cells were cultured, differentiated, and infected by HBV as previously described[26]. Primary human hepatocytes (PHH) were freshly prepared from human liver resection obtained from the Centre Léon Bérard (Lyon) with French ministerial authorizations (AC 2013-1871, DC 2013–1870, AFNOR NF 96 900 sept 2011) as previously described[27]. HBV inoculum was prepared from HepAD38[28] supernatants by polyethylene-glycol-MW-8000 (PEG8000, SIGMA) precipitation (8% final) as previously described[29]. Viral stocks with titer reaching at least 1×10^{10} vge/mL were tested endotoxin free. Blood from healthy donors was obtained from "Etablissement Français du Sang", Lyon France. PBMCs were isolated by gradient centrifugation using Histopaque (Sigma-Aldrich) and cultivated in RPMI medium, 10% FBS, 5% Penicillin-Streptomycin. Poly(I:C)-(HMW), Riboxxol and Pam3CSK4 were added at 5 µg/mL concentration. PBMCs were incubated for 24 h at 37 °C and 5%CO2 after TLRs stimulation.

Reagents. The following PRR agonists were purchased from InvivoGen and used at the indicated final concentration unless stated otherwise in the figure legends: Pam3CSK4 (TLR-1/2 agonist, 0.5 µg/mL), Poly(I:C)-HMW (TLR3-agonist, 5 µg/mL), LPS (TLR-4 agonist, 0.5 µg/mL), FLA-BS (TLR-5 agonist, 1 µg/mL), FSL-1 (TLR-2/6-agonist, 1 µg/mL), Imiquimod (R837, TLR-7 agonist, 5 µg/mL), CL264 (TLR-7 agonist 5 µg/mL), ssRNA40/lyoVecTM (ssRNA, TLR-8 agonist, 1 µg/mL), 2395 (CpG ODN-2395 (class-C), TLR-9 agonist, 5 µM). Riboxxol, recombinant IL6 and rhPEG-IFN-α (Roferon) were purchased respectively from Riboxx life science, R&D Systems and Roche. Neutralizing anti-IFNRA1 antibodies (IgG2a, #21385-1) and their isotype control antibodies (IgG2a, #281-010) were purchased respectively from PBL assay Science and Ancell. Neutralizing anti-IL-6 antibodies (IgG1, #501110) and their isotype control antibodies (IgG2, #400414) were purchased from Biolegend. Transfection of siRNA (SMARTpoolTM from Dharmacon/GE) were performed using Dharmafect-1 transfection reagent as recommended by the manufacturer (Darmacon/GE). siRNA against TLR3(# L-007745-00) RIG-I (# L-012511-00), MDA5 (# L-013041-00) and HCV (as siRNA control) was used at final concentration of 25 nM.

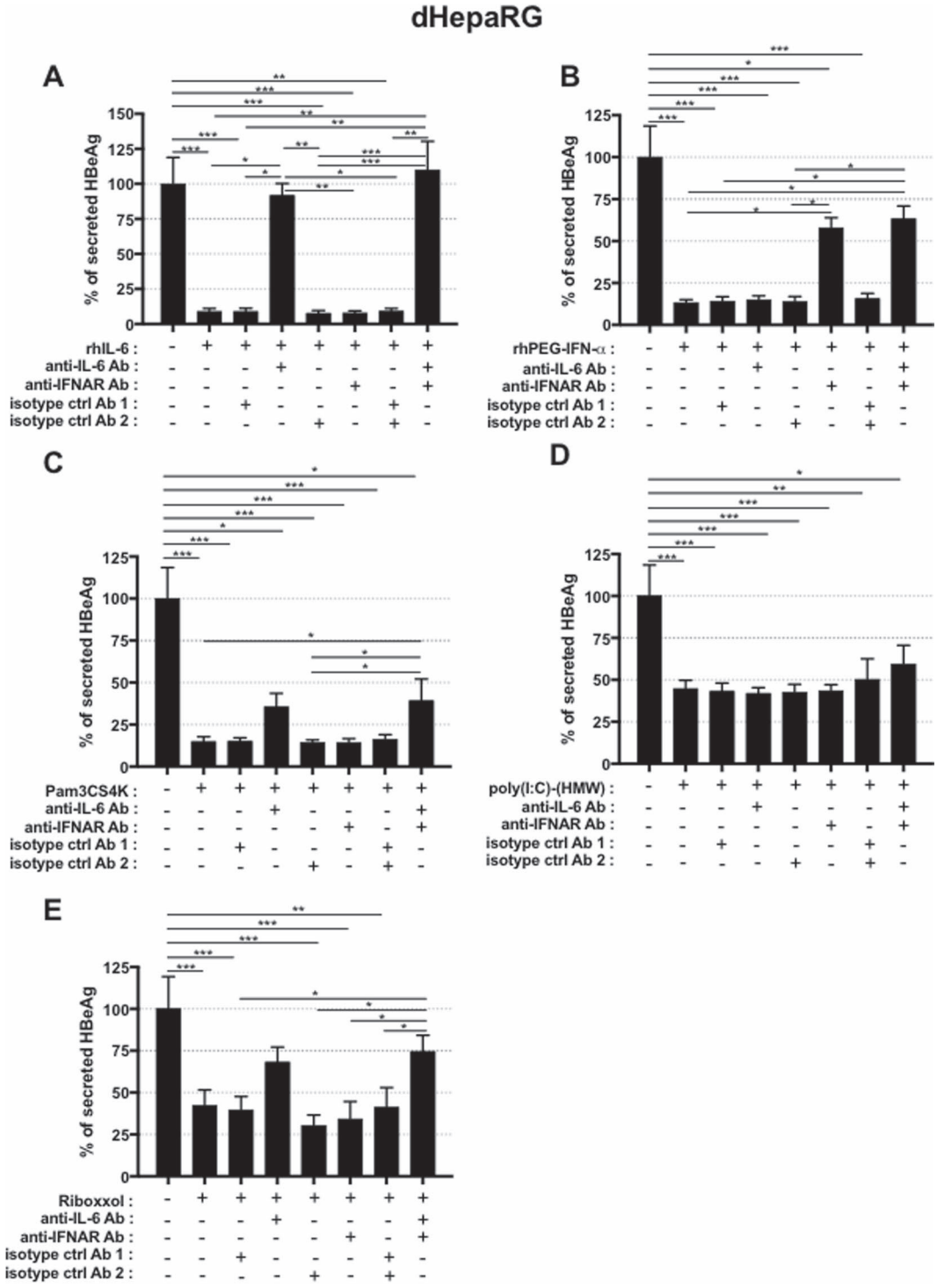


Figure 5. Inhibition of HBV by TLR-L is partially due to IL-6 secretion. dHepaRG cells were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven day and treated, or co-treated, twice as indicated with (**A** to **E**) anti-IL-6 antibodies (Ab) (2 ug/mL), (**A** to **E**) anti-IFNAR1 Ab (2 ug/mL), (**A** to **E**) isotype control (ctrl) Abs (1 for control of anti-IL6 Ab, 2 for control of anti-IFNRA1 Ab) (2 ug/mL), (**A**) rhIL-6 (100 ng/mL), (**B**) rhPEG-IFN-α (500 IU/mL), (**C**) Pam3CSK4 (1 ug/mL), (**D**) poly(I:C)-(HMW) (20 ug/mL) or (**E**) Riboxxol (50 ug/mL) for seven more days (2 treatments). At the end of the experiment, supernatants were collected and HBeAg levels assessed by ELISA. Results are the mean +/− SD of two independent experiments each performed with three biological replicates.

Western blot analysis. Cells were harvested in RIPA lysis buffer (NaCl 150 mM, Tris HCl pH = 8,0 50 mM, SDS 0,1%, NP40 1%, Na Deoxycholate 0,5%) containing protease inhibitors (Protein Cocktail Inhibitors from Sigma-Aldrich, NaF 10 mM, Na Orthovanadate 10 mM). Clarified lysates were subjected to 10% SDS-PAGE and

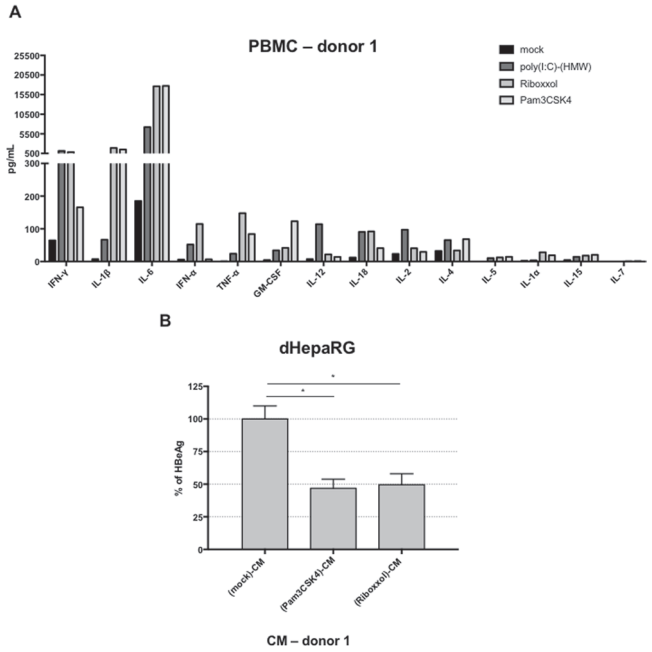


Figure 6. TLR2 and 3 ligands lead to production of inflammatory cytokines by immune blood cells. (**A**) Fresh PBMC from a healthy donor were cultivated and stimulated or not with Pam3CSK4 (5 µg/mL), poly(I:C)-(HMW) (5 µg/mL) or Riboxxol (5 µg/mL) for 24 h. Supernatant were collected and cytokines content was analyzed with Luminex Assay. (**B**) dHepaRG cells were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven days and treated or not during 10 days with the indicated conditioned media (CM) diluted 1/100. Supernatant were collected and HBeAg levels were assessed by ELISA. Results are the mean +/− SD of two independent experiments each performed with two or three biological replicates.

Western Blot transfer onto PVDF or nitrocellulose membranes using the TransTurbo Blot apparatus according to the manufacturer (Biorad). Primary antibodies are the TLR3 antibody (D10F10, #6961) RIG-I antibody (D14G6, #3743) MDA5 antibody (D74E4, #5321) from Cell Signaling and the anti-huActin (clone C4, #08691002) from MP Biomedicals. Secondary HRP antibodies anti-rabbit and anti-mouse were purchased from Sigma-Aldrich.

Nucleic acid extractions, reverse transcription and qPCR analyses. Total RNA and total DNA were respectively extracted from cells with the NucleoSpin RNA II kit and Nucleospin® 96 tissue kit according to the manufacturer's instructions (Macherey-Nagel). RNA reverse transcription was performed using the Superscript III RT (Life Technologies). Quantitative PCR for HBV were performed using HBV specific primers and normalized to PRNP housekeeping gene as previously described[15].

cccDNA detection. After total DNA isolation and T5 exonuclease digestion (NEB, Fance) digestion for 1 h, cccDNA was quantified as previously described[30]. Alternatively, cccDNA was analyzed by Southern blot after Hirt extraction as described previously[31].

Enzyme-linked Immunoassay for IL-6, IP10, HBeAg, HBsAg, and ApoB detection in cell supernatant. Human IL-6 and human IP-10 cytokines were detected in the supernatants using the DuoSet® ELISA kit according to the manufacturer (R&D Systems). HBeAg and HBsAg were detected in the supernatant of HBV-infected cells using the Autobio kit according to the manufacturer (AutoBio, China). ApoB secretion was detected in the supernatant of cells using the Total Human Apolipoprotein B assay according to the manufacturer (AlerCHEK, USA).

Measurement of cytokines by Luminex. IL-12p70, GM-CSF, IFN-α, IFN-γ, IL-1α, IL-1β, IL-1RA, IL-13, IL-15, IL-18, IL-2, IL-31, IL-4, IL-5, IL-6, IL-7, TNF-α and TNF-β were measured with a ProcartaPlex Human Th1/Th2 & Cytokine Panel 1 C (18 plex) kit (eBiosciences). Samples were analyzed using a Bio-Plex® 200 System (Bio-Rad, Hercules, CA) according to the manufacturer's instructions.

Viability/cytotoxicity assays. Neutral red uptake assay and Sulforhodamine staining to estimate cell viability/cytotoxicity were performed as previously described[4].

***In vivo* experiments.** All experiments were performed in accordance with the European Union guidelines for approval of the protocols by the local ethics committee (Authorization Agreement C2EA-15, "Comité Rhône-Alpes d'Ethique pour l'Expérimentation Animale", Lyon, France - APAFIS#1570-2015073112163780). Primary Human Hepatocytes (PHH, Corning, BD Gentest) were injected in FRG mice intrasplenically 48 h after

adeno-uPA conditioning[32]. Mice were subjected to NTBC cycling during the liver repopulation process. Mice with HSA levels >19 mg/ml, as determined using a Cobas C501 analyzer, Roche Applied Science, were infected by IP injection with 5×10^8 vge/mL of HBV. Six weeks later, mice were injected intravenously twice a week during 3 weeks with Pam3CSK4 or Riboxxol (20 ug per mouse per injection the first week, 40 ug per mouse per injection the second week and 80 ug per mouse per injection the third week). Sera were collected at different time before and after treatments. Mice were sacrificed 10,5 weeks post HBV infection and liver pieces were flash frozen in liquid nitrogen and kept at −80 °C before further processing. HBV viremia was determined by extraction of DNA from sera (supplemented with a known amount of non-HBV DNA; pUC19 plasmid from New England Biolabs) using the NucleoSpin® RNA Virus kit (Macherey Nagel) followed by qPCR. For each serum, quantity of viral genome equivalent was calculated according to a standard curve (build with dilutions of a plasmid containing the HBV sequence) and normalized to the ratio between the actual and theoretical amount of recovered pUC19 DNA. Total DNA were extracted from liver pieces and cccDNA amount was quantified after T5 digestions as described previously[33].

Statistical analysis. Statistical analyses were performed using the XLStat software and Kruskal-Wallis tests with multiple comparison respect to non-treated cells (Dunn's post-test). For all tests, a p value $\leq$ 0,05 was considered as significant. *Correspond to p value $\leq$ 0.05; **correspond to p value $\leq$ 0.01; ***correspond to p value $\leq$ 0.001.

References

1. Zeisel, M. B. *et al*. Towards an HBV cure: state-of-the-art and unresolved questions-report of the ANRS workshop on HBV cure. *Gut*. https://doi.org/10.1136/gutjnl-2014-308943 (2015).
2. Durantel, D. & Zoulim, F. New antiviral targets for innovative treatment concepts for hepatitis B virus and hepatitis delta virus. *J Hepatol* **64**, S117–131, https://doi.org/10.1016/j.jhep.2016.02.016 (2016).
3. Pandey, S., Kawai, T. & Akira, S. Microbial sensing by Toll-like receptors and intracellular nucleic acid sensors. *Cold Spring Harb Perspect Biol* **7**, a016246, https://doi.org/10.1101/cshperspect.a016246 (2015).
4. Isorce, N. *et al*. Antiviral activity of various interferons and pro-inflammatory cytokines in non-transformed cultured hepatocytes infected with hepatitis B virus. *Antiviral Res* **130**, 36–45, https://doi.org/10.1016/j.antiviral.2016.03.008 (2016).
5. Konerman, M. A. & Lok, A. S. Interferon Treatment for Hepatitis B. *Clin Liver Dis* **20**, 645–665, https://doi.org/10.1016/j.cld.2016.06.002 (2016).
6. Lanford, R. E. *et al*. GS-9620, an oral agonist of Toll-like receptor-7, induces prolonged suppression of hepatitis B virus in chronically infected chimpanzees. *Gastroenterology* **144**(1508–1517), 1517 e1501–1510, https://doi.org/10.1053/j.gastro.2013.02.003 (2013).
7. Lawitz, E. *et al*. Safety, pharmacokinetics and pharmacodynamics of the oral toll-like receptor 7 agonist GS-9620 in treatment-naive patients with chronic hepatitis C. *Antivir Ther* **20**, 699–708, https://doi.org/10.3851/IMP2845 (2015).
8. Menne, S. *et al*. Sustained efficacy and seroconversion with the Toll-like receptor 7 agonist GS-9620 in the Woodchuck model of chronic hepatitis B. *J Hepatol* **62**, 1237–1245, https://doi.org/10.1016/j.jhep.2014.12.026 (2015).
9. Gane, E. J. *et al*. The oral toll-like receptor-7 agonist GS-9620 in patients with chronic hepatitis B virus infection. *J Hepatol* **63**, 320–328, https://doi.org/10.1016/j.jhep.2015.02.037 (2015).
10. Dammermann, W., Dornbrack, J., Broker, K., Bentzien, F. & Luth, S. CpG oligonucleotides increase HBV-specific cytokine responses in whole blood and enhance cytokine release assay sensitivity. *J Virol Methods* **248**, 195–201, https://doi.org/10.1016/j.jviromet.2017.07.011 (2017).
11. Campbell, J. D. Development of the CpG Adjuvant 1018: A Case Study. *Methods Mol Biol* **1494**, 15–27, https://doi.org/10.1007/978-1-4939-6445-1_2 (2017).
12. Meng, Z. *et al*. Combination therapy including CpG oligodeoxynucleotides and entecavir induces early viral response and enhanced inhibition of viral replication in a woodchuck model of chronic hepadnaviral infection. *Antiviral Res* **125**, 14–24, https://doi.org/10.1016/j.antiviral.2015.11.001 (2016).
13. Huang, L. R. *et al*. Intrahepatic myeloid-cell aggregates enable local proliferation of CD8(+) T cells and successful immunotherapy against chronic viral liver infection. *Nature immunology* **14**, 574–583, https://doi.org/10.1038/ni.2573 (2013).
14. Luangsay, S. *et al*. Expression and functionality of Toll- and RIG-like receptors in HepaRG cells. *J Hepatol* **63**, 1077–1085, https://doi.org/10.1016/j.jhep.2015.06.022 (2015).
15. Lucifora, J. *et al*. Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA. *Science* **343**, 1221–1228, https://doi.org/10.1126/science.1243462 (2014).
16. Jelinek, I. *et al*. TLR3-specific double-stranded RNA oligonucleotide adjuvants induce dendritic cell cross-presentation, CTL responses, and antiviral protection. *J Immunol* **186**, 2422–2429, https://doi.org/10.4049/jimmunol.1002845 (2011).
17. Hosel, M. *et al*. Not interferon, but interleukin-6 controls early gene expression in hepatitis B virus infection. *Hepatology* **50**, 1773–1782 (2009).
18. Belloni, L. *et al*. Mimicking Interferon-α (IFNα) inhibitory activity on hep- atitis B virus (HBV) transcription and replication by tar- geting the epigenetic control of nuclear cccDNA minichromosome with epigenetic small molecules. *Hepatology* **56**, 369A (2012).
19. Palumbo, G. A. *et al*. 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 (2015).
20. Xia, Y. *et al*. Interferon-gamma and Tumor Necrosis Factor-alpha 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 (2016).
21. Lucifora, J. *et al*. Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. *J Hepatol* **55**, 996–1003, doi:S0168-8278(11)00197-8 [pii] https://doi.org/10.1016/j.jhep.2011.02.015 (2011).
22. Isogawa, M., Robek, M. D., Furuichi, Y. & Chisari, F. V. Toll-like receptor signaling inhibits hepatitis B virus replication *in vivo*. *J Virol* **79**, 7269–7272 (2005).
23. Wu, J. *et al*. Hepatitis B virus suppresses toll-like receptor-mediated innate immune responses in murine parenchymal and nonparenchymal liver cells. *Hepatology* **49**, 1132–1140 (2008).
24. Zhang, X. *et al*. Role of Toll-like receptor 2 in the immune response against hepadnaviral infection. *J Hepatol* **57**, 522–528, https://doi.org/10.1016/j.jhep.2012.05.004 (2012).
25. Tyler, B., Gullotti, D., Mangraviti, A., Utsuki, T. & Brem, H. Polylactic acid (PLA) controlled delivery carriers for biomedical applications. *Adv Drug Deliv Rev* **107**, 163–175, https://doi.org/10.1016/j.addr.2016.06.018 (2016).
26. Gripon, P. *et al*. Infection of a human hepatoma cell line by hepatitis B virus. *Proc Natl Acad Sci USA* **99**, 15655–15660 (2002).
27. Lecluyse, E. L. & Alexandre, E. Isolation and culture of primary hepatocytes from resected human liver tissue. *Methods Mol Biol* **640**, 57–82, https://doi.org/10.1007/978-1-60761-688-7_3 (2010).

28. Ladner, S. K. *et al*. 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 (1997).
29. Luangsay, S. *et al*. Early inhibition of hepatocyte innate responses by hepatitis B virus. *J Hepatol* **63**, 1314–1322, https://doi.org/10.1016/j.jhep.2015.07.014 (2015).
30. Werle-Lapostolle, B. *et al*. Persistence of cccDNA during the natural history of chronic hepatitis B and decline during adefovir dipivoxil therapy. *Gastroenterology* **126**, 1750–1758 (2004).
31. Lucifora, J. *et al*. Detection of the hepatitis B virus (HBV) covalently-closed-circular DNA (cccDNA) in mice transduced with a recombinant AAV-HBV vector. *Antiviral Res* **145**, 14–19, https://doi.org/10.1016/j.antiviral.2017.07.006 (2017).
32. Calattini, S. *et al*. Functional and Biochemical Characterization of Hepatitis C Virus (HCV) Particles Produced in a Humanized Liver Mouse Model. *J Biol Chem* **290**, 23173–23187, https://doi.org/10.1074/jbc.M115.662999 (2015).
33. Allweiss, L. *et al*. Proliferation of primary human hepatocytes and prevention of hepatitis B virus reinfection efficiently deplete nuclear cccDNA *in vivo*. *Gut*. https://doi.org/10.1136/gutjnl-2016-312162 (2017).

Acknowledgements

The authors would like to thank Lydie Lefrançois and Jennifer Molle for help with the isolation of primary human hepatocytes, the staff from Prof Michel Rivoire's surgery room for providing us with liver resection and Dr. Serge Lebecque and Dr. Touffic Reno for advises on Riboxxol. This work was supported by grants from ANRS (French national agency for research on AIDS and viral hepatitis; grants from CSS4), The "ANRS consortium on humanized mouse models for viral hepatitis", FINOVI (Foundation for innovation in infectiology), ARC (French Agency for Research against Cancer), Gilead Science, the European Research Council (ERC-AdG-233130-HEPCENT - to FLC) and INSERM. This work was also supported by the DEVweCAN LABEX (ANR-10-LABX-0061) and the LabEx Ecofect (ANR-11-LABX-0048) of the "Université de Lyon", within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). MB was the recipient of RPF (Roche Postdoc Fellowship) program supported by Hoffmann-La-Roche (http://www.roche.com/careers/workplaces/wp_research/postdoc_fellowship_rpf_program.htm). The authors thanks the Plateforme de Thérapie Génique in Nantes (France) for the production of the *in vivo* certified lots of adeno-uPA vector as well as J.F. Henry, N. Aguilera and J.L. Thoumas from the animal facility (PBES, Plateau de Biologie Experimental de la Souris, ENS de Lyon).

Author Contributions

Study concept and design: J.L., D.D., F.Z., S.F., S.D. Acquisition of data: J.L., M.B., L.A., F.F., S.M., O.F., L.D., M.M., S.D., A.O. Analysis and interpretation of data: J.L., M.B., M.A.G. and D.D. Drafting of the manuscript: J.L. and D.D. Statistical analyses: J.L. Material support: M.R., A.S., F.L.C.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-23525-w.

Competing Interests: S.M., L.D., O.F., M.M., L.A., A.S., F.F., F.L.C., A.O. have nothing to declare. F.Z., J.L. and D.D. received a research grant from Gilead Science to perform experiments. S.D. and S.F. are employees of Gilead Science. M.B. was the recipient of RPF (Roche Postdoc Fellowship) program supported by Hoffmann-La-Roche. MAG is an employee of Hoffman-La-Roche.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary figures

Direct antiviral properties of TLR ligands against HBV replication in immune-competent hepatocytes

Julie Lucifora[1,*,$], Marc Bonnin[1,$], Ludovic Aillot[1], Floriane Fusil[2], Sarah Maadadi[1], Laura Dimier[1], Maud Michelet[1], Océane Floriot[1], Anaïs Ollivier[2], Michel Rivoire[3], Malika Ait-Goughoulte[4], Stéphane Daffis[5], Simon P. Fletcher[5], Anna Salvetti[1], François-Loïc Cosset[2], Fabien Zoulim[1,6] and David Durantel[1,*]

[1] INSERM, U1052, Cancer Research Center of Lyon (CRCL), Université de Lyon (UCBL1), CNRS UMR_5286, Centre Léon Bérard, Lyon, France
[2] CIRI – International Center for Infectiology Research, Inserm, U1111, Université Claude Bernard Lyon 1, CNRS, UMR5308, Ecole Normale Supérieure de Lyon, Univ Lyon, F-69007, Lyon, France
[3] INSERM U1032, Centre Léon Bérard (CLB), Lyon, France
[4] Roche Pharma Research and Early Development (pRED), Roche Innovation Center Basel, F. Hoffmann-La Roche, 4070 Basel, Switzerland
[5] Gilead Sciences, Inc., Foster City, CA, USA;
[6] Department of Hepatology, Croix-Rousse Hospital, Hospices Civils de Lyon, Lyon, France

[$] equal contribution

***Correspondence:**
Dr. David Durantel and Dr. Julie Lucifora
Centre de Recherche en Cancérologie de Lyon (CRCL), UMR Inserm 1052 - CNRS 5286, 151 cours Albert Thomas, 69424 Lyon Cedex 03, France ; Phone: + 33 4 72 68 19 70 ; Fax : +33 4 72 68 19 71 ; E-mail: david.durantel@inserm.fr and julie.lucifora@inserm.fr

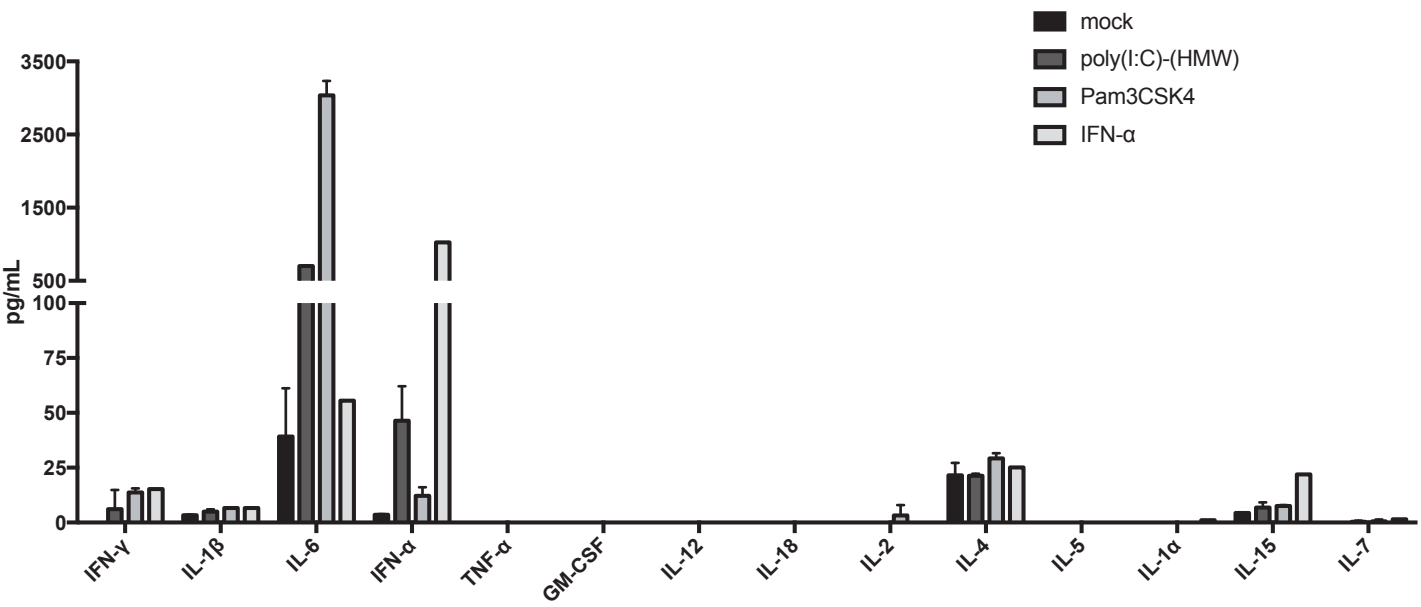


Figure S1: **Cytokines secretion by dHepaRG stimulated with poly(I:C)-(HMW) or Pam3CSK4.** dHepaRG cells were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven days, treated or not with Pam3CSK4 (0,5 ug/mL), poly(I:C)-(HMW) (5 ug/mL) or PEG-IFN-α (1000 IU/mL) for 24h. Supernatants were collected and amount of the indicated cytokines were measured by Luminex assay. Results are the mean +/- SD one experiment performed with two biological replicate.

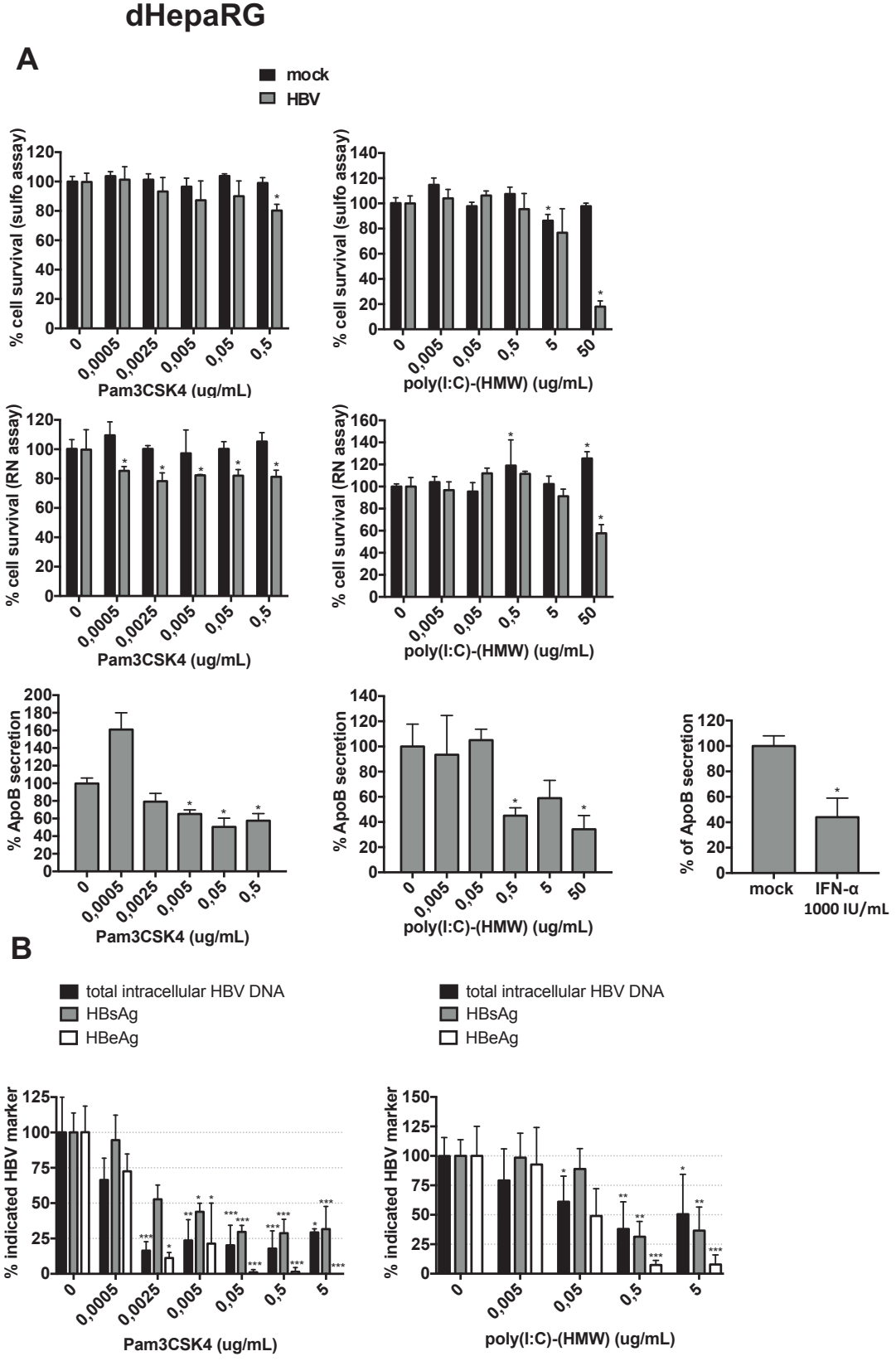


Figure S2: **Pam3CSK4 and poly(I:C)-(HMW) inhibit HBV in HBV-infected dHepaRG in a dose dependent manner without toxicity**. dHepaRG were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven days and treated with the indicated molecules at the indicated concentrations for seven more days (2 treatments). At the end of the experiments, (A) cells survival rates was assessed by sulforhodamin or neutral red assays, (B) total DNAs were extracted and total intracellular HBV DNA amounts were evaluated by qPCR analyses. Supernatant were collected and (A) ApoB, (B) HBeAg or HBsAg levels were assessed by ELISA. Results are the mean +/- standard deviation (SD) of three independent experiments each performed with three biological replicate.

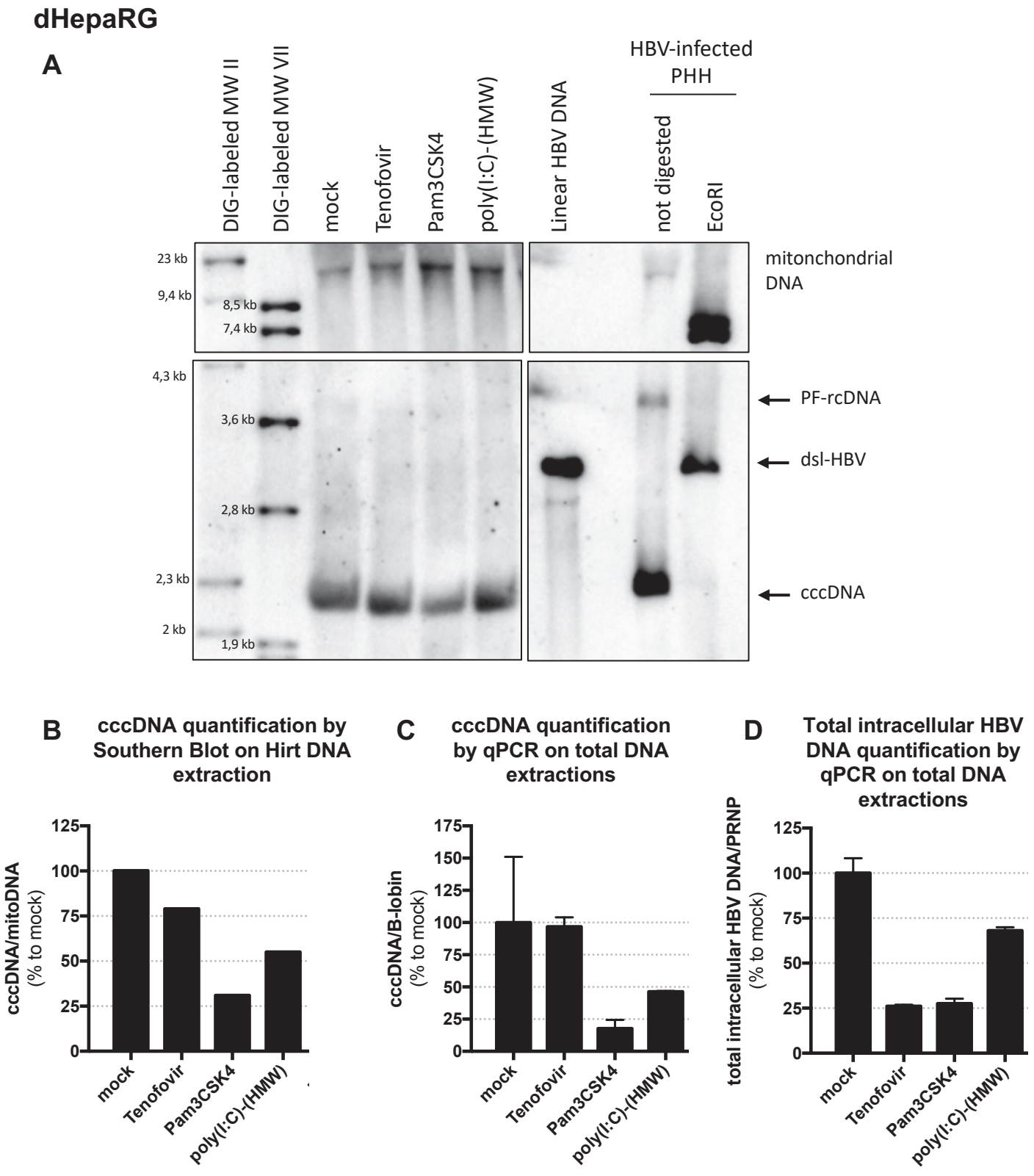


Figure S3: **Pam3CSK4 and poly(I:C)-(HMW) treatments decreased cccDNA levels in HBV-infected dHepaRG**. dHepaRG were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven days and treated with Tenofovir (10 uM), Pam3CSK4 (0,5 ug/mL) or poly(I:C)-(HMW) (0,5 ug/mL) for seven more days (2 treatments). At the end of the experiments, cells were harvested, (A, B) DNAs were extracted following a Hirt procedure or (C, D) a total DNA extraction procedure and submitted to (A) Southern blot analyses using HBV-DIG labeled probes or (C, D) qPCR analyses to detect (C) cccDNA or (D) total intracellular HBV DNAs. (B) cccDNA and mitochondrial DNA were quantified on the southern blot analyses using the « ImageLab » software from Bio-Rad.

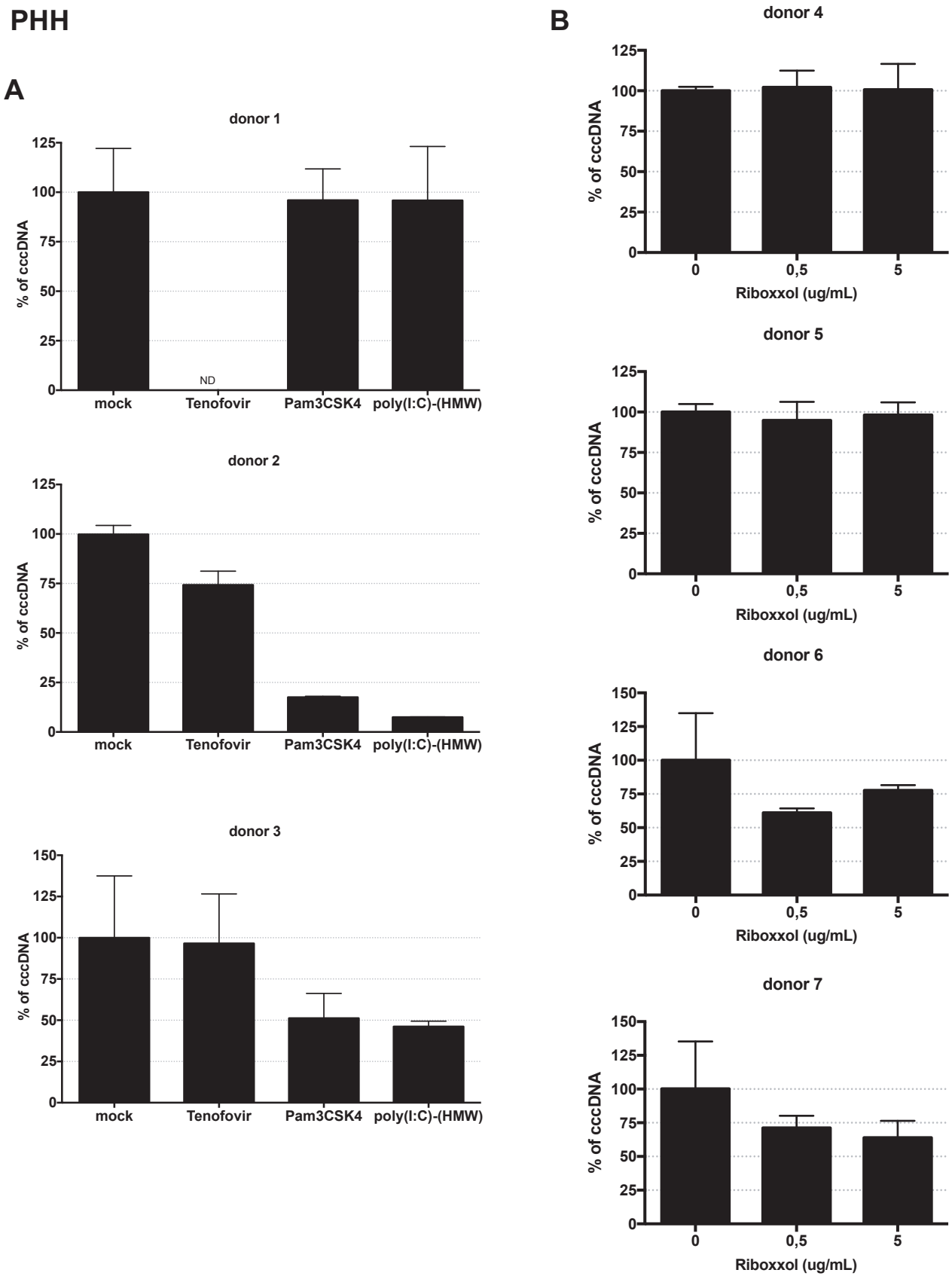


Figure S4: **Effect of Pam3CSK4, poly(I:C)-(HMW) and riboxxol treatments on cccDNA levels in HBV-infected PHH**. PHH were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for four days and treated with Tenofovir (10 uM), Pam3CSK4 (0,5 ug/mL), poly(I:C)-(HMW) (0,5 ug/mL) or Riboxxol (indicated concentrations) for seven more days (2 treatments). At the end of the experiments, cells were harvest, total DNA**s** were extracted and cccDNA amounts were evaluated by qPCR analyses. For each donor, results are the mean +/- SD one experiment performed with three biological replicate.

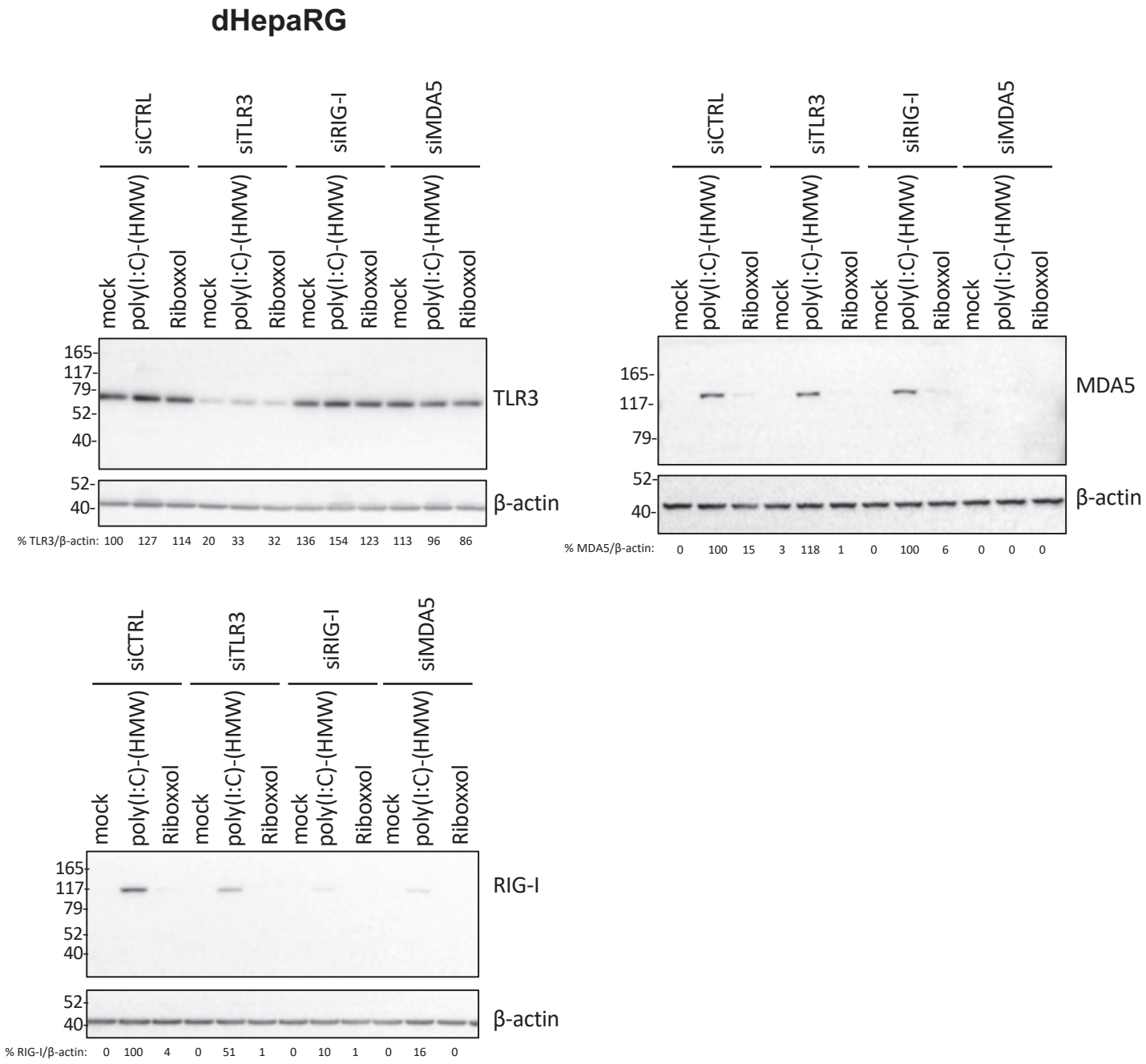


Figure S5**: TLR3, MDA5 and RIG-I protein levels after siRNA transfection.** Uninfected dHepaRG cells were transfected with the indicated siRNA. 14 days later, cells were stimulated with 10 µg/ml of the indicated ligands. Cells were lyzed 24h later and TLR3, MDA5 or RIG-I protein levels were analyzed by western blot. Blots presented here are representative of three independent experiments.

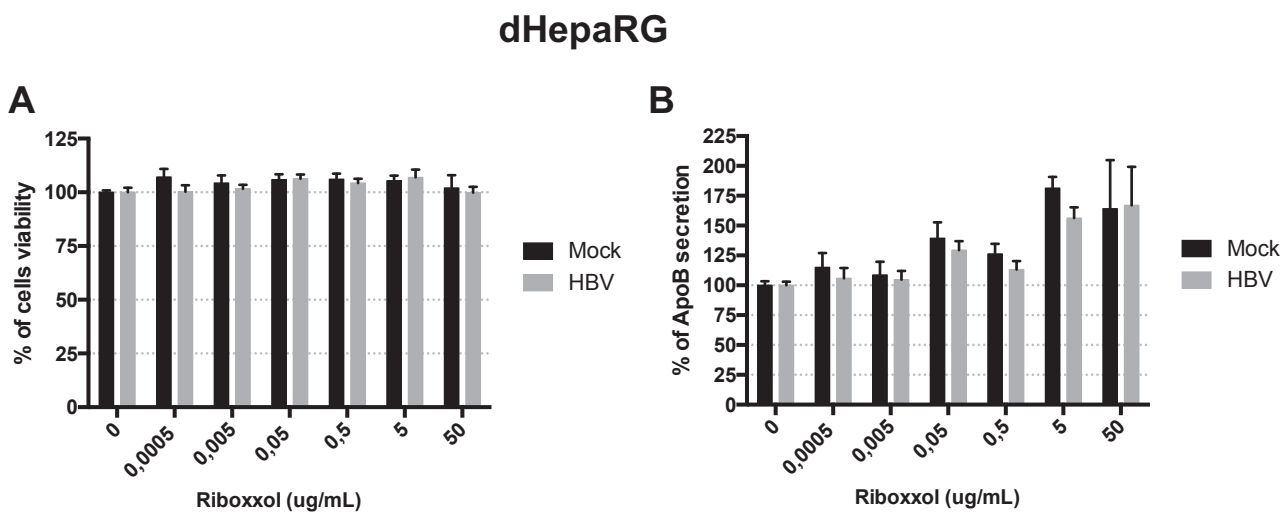


Figure S6: **Treatments with Riboxxol did not impact on dHepaRG survival and differentiation status.** dHepaRG were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven days and treated with the riboxxol at the indicated concentrations for seven more days (2 treatments). At the end of the experiments, (A) cells survival rates was assessed by sulforhodamin assay, (B) supernatant were collected and ApoB levels were assessed by ELISA. Results are the mean +/- standard deviation (SD) of three independent experiments each performed with three biological replicate.

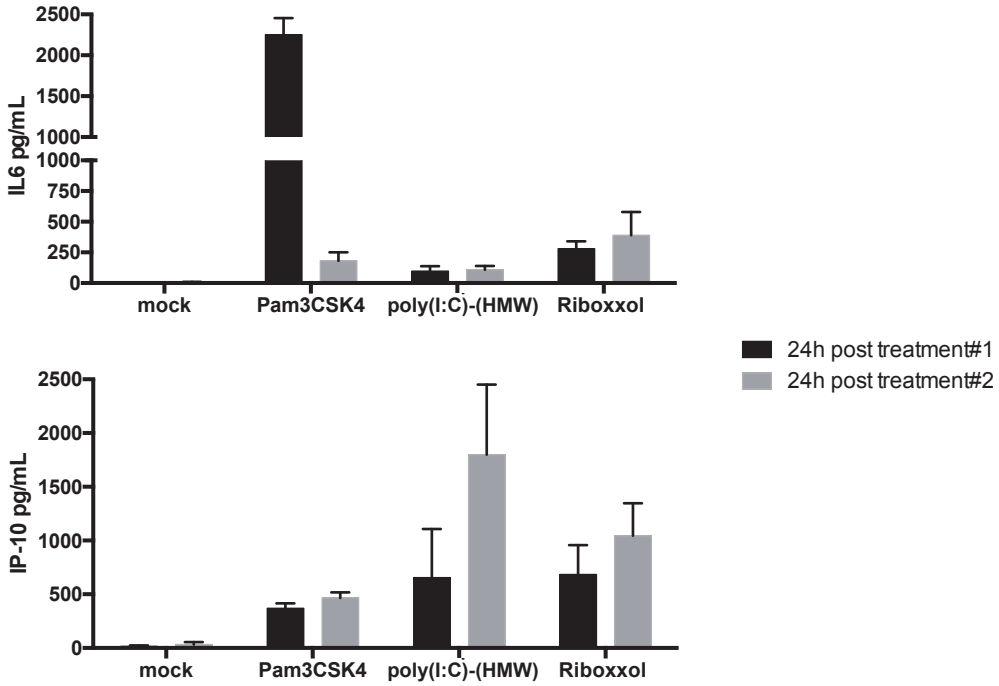


Figure S7: **Riboxxol induced secretion of both IL-6 and IP-10.** dHepaRG cells were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven days, treated or not with Pam3CSK4 (0,5 ug/mL), poly(I:C)-(HMW) (5 ug/mL) or Riboxxol (50 ug/mL) for seven days (2 treatments). Supernatants were collected and amount of IL6 and IP-10 were quantified by ELISA. Results are the mean +/- SD of three independent experiments each performed with three biological replicate.

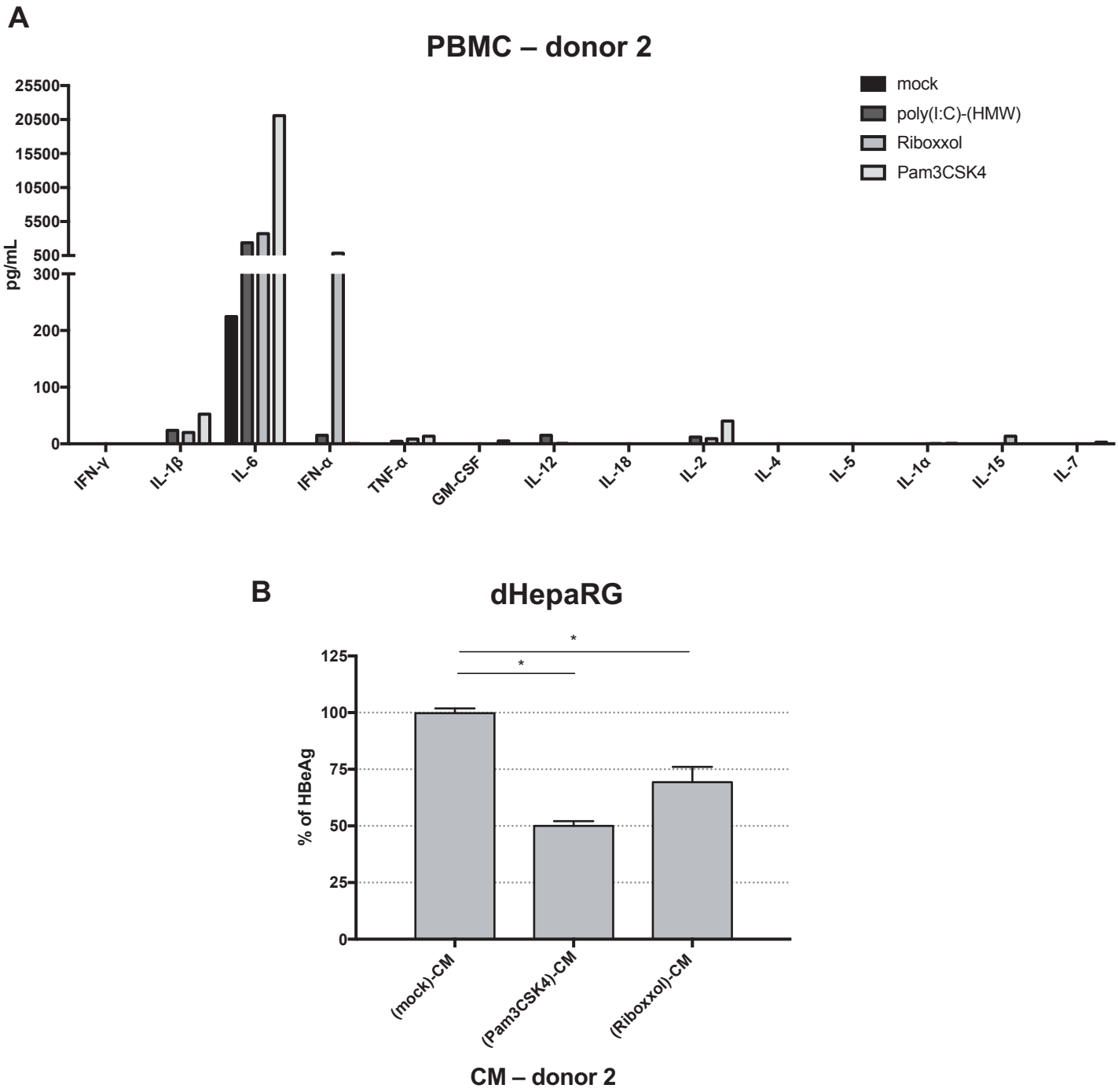


Figure S8: **TLR2 and 3 ligands lead to production of inflammatory cytokines by immune blood cells**. (A) Fresh PBMC from a healthy donor were cultivated and stimulated or not with Pam3CSK4 (5 µg/mL), poly(I:C)-(HMW) (5 µg/mL) or Riboxxol (5 µg/mL) for 24h. Supernatant were collected and cytokines content was analyzed with Luminex Assay. (B) dHepaRG cells were infected by HBV at a multiplicity of infection of 100 viral genome equivalent (vge)/cell for seven days and treated or not during 10 days with the indicated conditioned media (CM) diluted 1/100. Supernatant were collected and HBeAg levels were assessed by ELISA.

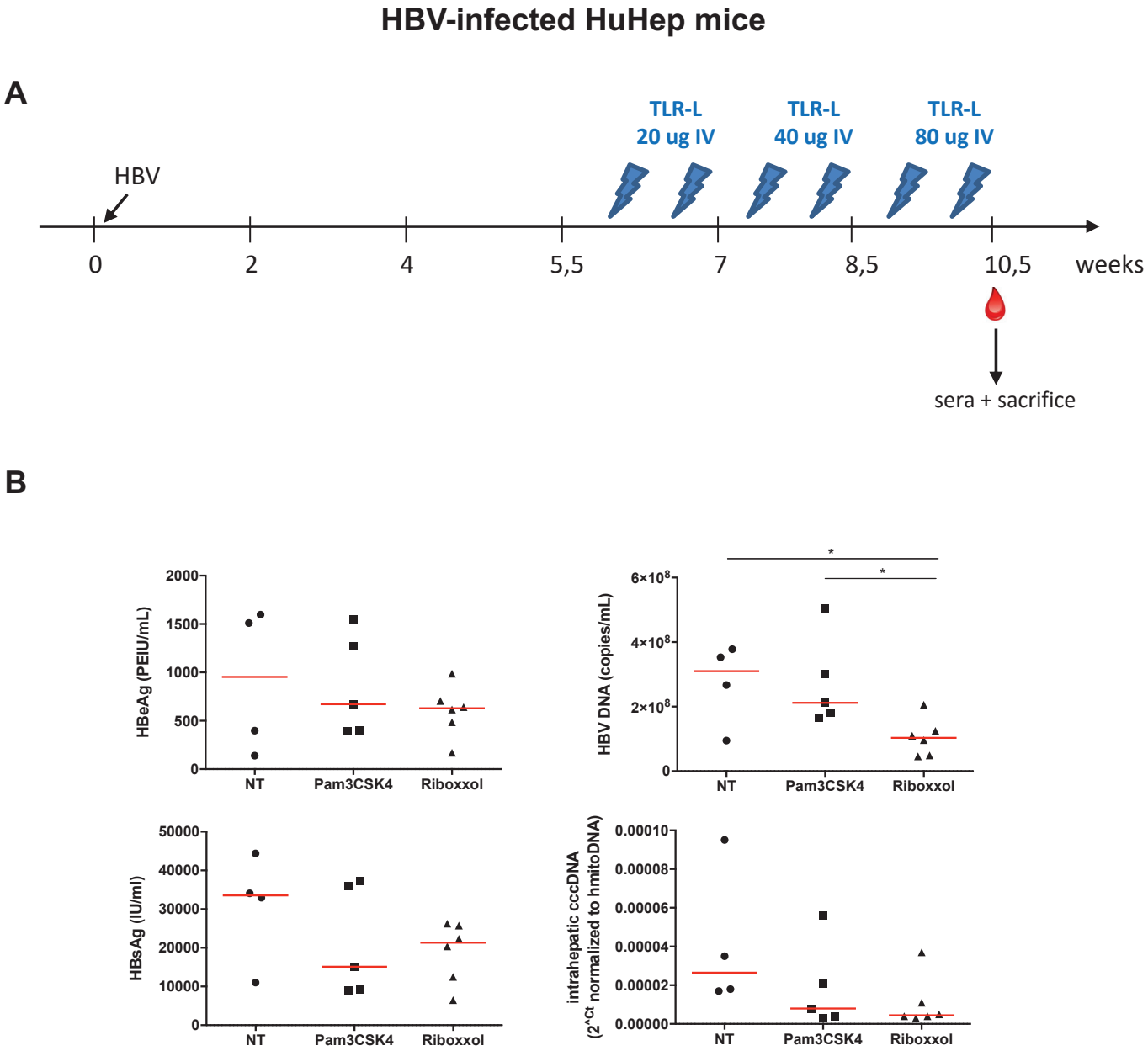


Figure S9: **Effect of Pam3CS4K and Riboxxol in HBV-infected HuHep mice.** HuHep mice were infected IP with HBV and treated IV with TLR-L (Pam3CSK4 or Riboxxol) or not (NT; no treatment) (A) according to the scheme. Mice were sacrificed after 3 weeks of treatment and (B) levels of HBV markers (HBeAg, HBSAg, viremia and intrahepatic HBV cccDNA) were assessed by ELISA, qPCR or RT-qPCR. Each dot represents one animal and red lines the median from the 4-5 animals.

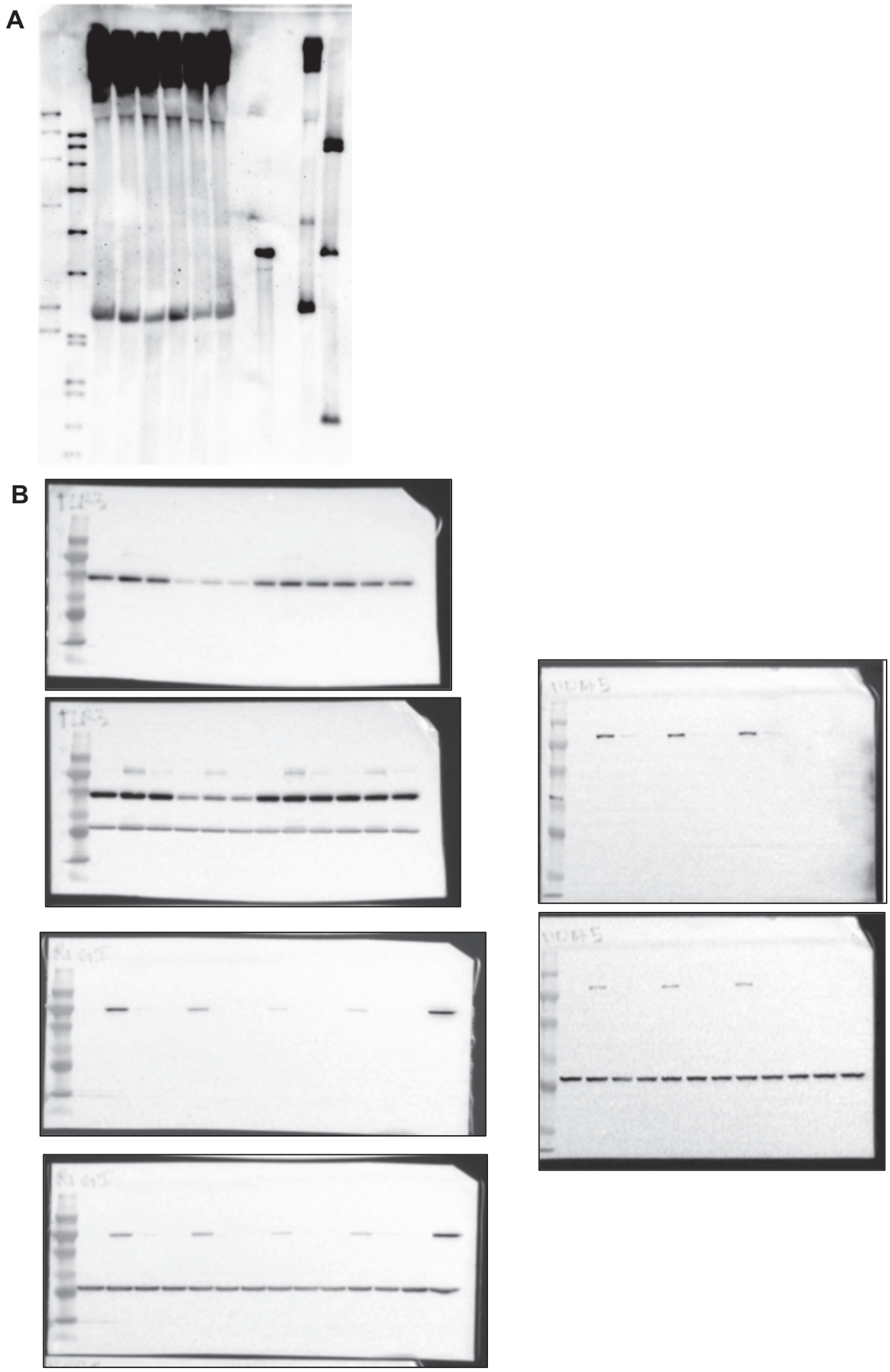


Figure S11: **Full blots** from (A) fig S3A, (B) fig S5

Remerciements

Je souhaiterais tout d'abord remercier les membres du jury qui ont accepté de lire, de juger mon travail : je remercie le **Dr Urzsula Hibner** et le **Dr Camille SUREAU** pour leurs lectures critiques du manuscrit, ainsi que le **Dr Hélène Strick-Marchand** d'être examinatrice de ma thèse.

Je tiens à remercier le **Prof. Fabien Zoulim** de m'avoir tout d'abord accueilli au sein de son laboratoire pendant mon master et par la suite d'avoir suivis l'évolution de ma thèse durant les comités et enfin de présider ma thèse.

Je tiens sincèrement à remercier le **Prof. Massimo Levrero** pour avoir été mon directeur de thèse. Vous avez fait un pari un peu fou il y a 4 ans de me prendre en thèse lors de la création de l'équipe. Merci de m'avoir fait confiance, cela n'a pas toujours été facile mais je pense que nous nous en sommes bien sorti. Merci de ne pas seulement penser à la science mais aussi de vous préoccuper de nous sur le plan personnel.

Merci énormément **Mirjam** pour toute l'aide à la fin de cette thèse que ce soit pour les manips ou pour les corrections !

Barbara je veux te remercier pour m'avoir aidé au début de ma thèse à apprendre et maîtriser ces fameux ChIP dont tu es l'experte !! mais aussi d'avoir été là pendant toutes ces années autant pour des conseils que pour partager des moments en dehors !! Et je crois que la boucle est bouclée on a commencé les manips ensembles et une des dernières était avec toi…
Marie-Laure que dire… Je suis tellement contente que tu ais rejoint l'équipe !! Tu nous apporte tellement de choses, l'aides pour les manips, les cours de confiance en soi ! Merci d'être toujours très sincère surtout en conseils mode. Reste comme tu es !!

Je tiens à remercier toutes les personnes en Italie et principalement **Francesca** pour l'aide qu'ils ont apporté au cours de ma thèse.

Merci à **Marie-Agnès** et **Christophe** pour tout le travail que vous faites au sein de ce laboratoire.

Je remercie énormément la Team PHH car sans vous rien de nos projets ne serait possible et je sais combien c'est compliqué !! Je suis ravie d'avoir fait partie de cette team. Merci en particulier à **Maud** pour ton énergie et ta bonne humeur ! Merci ma **mamie Lolo** pour ta joie et tous les thés du matin, les discussions de danse et tout le reste !

Mama, **Susu**, nous avons commencé cette aventure ensemble il y a un peu plus de 5 ans. Ca y est je clôture cette génération. Je suis ravie d'avoir partagé tous ces moments avec vous. Je ne m'étale pas je vous ai déjà laissé des petits mots dans vos thèses… Continuer à m'envoyer des cartes postales de vos aventures !!

Merci **Nathalie** et **Dulce** les grandes d'avoir été là au début et puis aussi par la suite !!

Merci à tous les annexiens, vous participez à la bonne ambiance de ce laboratoire !! merci **Fleur** d'être comme tu es et **Damien** pour être un super coloc de bureau (enfin je ne saurais décrire cette colocation…)

Papa un jour tu m'avais dit « ne fait pas médecine c'est un boulot de con !! », je t'ai presque écouté…. **Papa**, **Maman** je veux vous remercier de m'avoir toujours soutenue malgré mes hésitations sur mes choix professionnels. Merci d'être des parents formidables !! Je vous aime très fort !
Merci à toute ma famille **Yannick**, **Virginie**, **Isa**, **Anne-Laure**, **Pierre-Luc**….

Merci **Fred** pour toutes les conversations que l'on a pu avoir sur l'avenir, d'avoir été aussi patient avec moi pour m'apprendre à avoir un peu plus confiance en moi à cheval et dans la vie et qu'avec les chevaux rien n'est jamais vraiment acquit. Cela me sert aussi dans mon travail avec la science !!

Mon **club des Démons,** que dire pour vous remercier sans trop m'étendre, le challenge va être difficile…

Ma petite **Alexia**, je suis tellement heureuse que tu fasses partie de cette équipe !! tu seras une grande scientifique. Je te remercie pour ta joie et ton petit côté démon tout mignon qui nous fait craquer quand on a besoin de se remonter le moral (ou pas). Promis j'essaie de te faire un manuel pour gérer la thèse !! **Léa** la pro des travaux et des manips, tu as toujours des super idées pour faire avancer tes projets comme avec tes bébés norganoïdes. Merci de m'avoir fait confiance pour m'en occuper pendant tes vacances !! **Anaëlle** encore une ingé qui cours partout…. Merci pour tous ces moments de fous rires cet été et aussi au quotidien. Merci les filles pour ces soirées « l'Amour est dans le pré » et glace maison ! **Jennifer** bravo pour tout ce que tu fais au laboratoire et qui contribue à son bon fonctionnement. Merci de m'avoir fait découvrir ton univers de la danse et d'être toujours là pour moi et tes amies !! Pour résumer : merci énormément les filles d'être là pour tout ce que vous faites et avez fait pendant ma thèse ! Vous êtes des vraies amies formidables. Mais ça c'est parce que vous faites partie de l'élite !!

Pour **Bernard**…. Car je pense que tu aurais voulu la lire…

Bibliographie

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. U. S. A. *84*, 4641–4644.

Acunzo, M., Romano, G., Wernicke, D., and Croce, C.M. (2015). MicroRNA and cancer--a brief overview. Adv. Biol. Regul. *57*, 1–9.

Ahluwalia, A., Hurteau, J.A., Bigsby, R.M., and Nephew, K.P. (2001). DNA methylation in ovarian cancer. II. Expression of DNA methyltransferases in ovarian cancer cell lines and normal ovarian epithelial cells. Gynecol. Oncol. *82*, 299–304.

Akiyama, Y., Watkins, N., Suzuki, H., Jair, K.-W., van Engeland, M., Esteller, M., Sakai, H., Ren, C.-Y., Yuasa, Y., Herman, J.G., et al. (2003). GATA-4 and GATA-5 transcription factor genes and potential downstream antitumor target genes are epigenetically silenced in colorectal and gastric cancer. Mol. Cell. Biol. *23*, 8429–8439.

Akkari, L., Grégoire, D., Floc'h, N., Moreau, M., Hernandez, C., Simonin, Y., Rosenberg, A.R., Lassus, P., and Hibner, U. (2012). Hepatitis C viral protein NS5A induces EMT and participates in oncogenic transformation of primary hepatocyte precursors. J. Hepatol. *57*, 1021–1028.

Allison, A.C., and Blumberg, B.S. (1961). An isoprecipitation reaction distinguishing human serum-protein types. Lancet Lond. Engl. *1*, 634–637.

Allweiss, L., and Dandri, M. (2016). Experimental in vitro and in vivo models for the study of human hepatitis B virus infection. J. Hepatol. *64*, S17–S31.

Ancey, P.-B., Testoni, B., Gruffaz, M., Cros, M.-P., Durand, G., Le Calvez-Kelm, F., Durantel, D., Herceg, Z., and Hernandez-Vargas, H. (2015). Genomic responses to hepatitis B virus (HBV) infection in primary human hepatocytes. Oncotarget *6*, 44877–44891.

Aravin, A.A., and Bourc'his, D. (2008). Small RNA guides for de novo DNA methylation in mammalian germ cells. Genes Dev. *22*, 970–975.

Arzumanyan, A., Friedman, T., Kotei, E., Ng, I.O.L., Lian, Z., and Feitelson, M.A. (2012). Epigenetic repression of E-cadherin expression by hepatitis B virus x antigen in liver cancer. Oncogene *31*, 563–572.

Baker, E.K., Johnstone, R.W., Zalcberg, J.R., and El-Osta, A. (2005). Epigenetic changes to the MDR1 locus in response to chemotherapeutic drugs. Oncogene *24*, 8061–8075.

Bandiera, S., Billie Bian, C., Hoshida, Y., Baumert, T.F., and Zeisel, M.B. (2016). Chronic hepatitis C virus infection and pathogenesis of hepatocellular carcinoma. Curr. Opin. Virol. *20*, 99–105.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Barrera, A., Guerra, B., Notvall, L., and Lanford, R.E. (2005). Mapping of the hepatitis B virus pre-S1 domain involved in receptor recognition. J. Virol. *79*, 9786–9798.

Bartenschlager, R., and Schaller, H. (1988). The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. EMBO J. *7*, 4185–4192.

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.

Basagoudanavar, S.H., Perlman, D.H., and Hu, J. (2007). Regulation of hepadnavirus reverse transcription by dynamic nucleocapsid phosphorylation. J. Virol. *81*, 1641–1649.

Baylin, S.B., and Jones, P.A. (2011). A decade of exploring the cancer epigenome — biological and translational implications. Nat. Rev. Cancer *11*, 726.

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.

Bedossa, P. (2015). Fibrose hépatique : physiopathologie, diagnostic, pronostic.

Belloni, L., Pollicino, T., Nicola, F.D., 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. *106*, 19975–19979.

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.

Benhenda, S., Cougot, D., Buendia, M.-A., and Neuveut, C. (2009). Chapter 4 Hepatitis B Virus X Protein: Molecular Functions and Its Role in Virus Life Cycle and Pathogenesis. B.-A. in C. Research, ed. (Academic Press), pp. 75–109.

Berting, A., Hahnen, J., Kröger, M., and Gerlich, W.H. (1995). Computer-aided studies on the spatial structure of the small hepatitis B surface protein. Intervirology *38*, 8–15.

Birnbaum, F., and Nassal, M. (1990). Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein. J. Virol. *64*, 3319–3330.

Biswas, B., Kandpal, M., and Vivekanandan, P. (2017). A G-quadruplex motif in an envelope gene promoter regulates transcription and virion secretion in HBV genotype B. Nucleic Acids Res. *45*, 11268–11280.

Bitzer, M., Horger, M., Giannini, E.G., Ganten, T.M., Wörns, M.A., Siveke, J.T., Dollinger, M.M., Gerken, G., Scheulen, M.E., Wege, H., et al. (2016). Resminostat plus sorafenib as second-line therapy of advanced hepatocellular carcinoma - The SHELTER study. J. Hepatol. *65*, 280–288.

Blumberg, B.S., Gerstley, B.J., Hungerford, D.A., London, W.T., and Sutnick, A.I. (1967). A serum antigen (Australia antigen) in Down's syndrome, leukemia, and hepatitis. Ann. Intern. Med. *66*, 924–931.

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 minichromosome1. J. Mol. Biol. *307*, 183–196.

Bremer, C.M., Sominskaya, I., Skrastina, D., Pumpens, P., El Wahed, A.A., Beutling, U., Frank, R., Fritz, H.-J., Hunsmann, G., Gerlich, W.H., et al. (2011). N-terminal myristoylation-dependent masking of neutralizing epitopes in the preS1 attachment site of hepatitis B virus. J. Hepatol. *55*, 29–37.

Bruss, V. (1997). A short linear sequence in the pre-S domain of the large hepatitis B virus envelope protein required for virion formation. J. Virol. *71*, 9350–9357.

Bruss, V., Lu, X., Thomssen, R., and Gerlich, W.H. (1994). Post-translational alterations in transmembrane topology of the hepatitis B virus large envelope protein. EMBO J. *13*, 2273–2279.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., et al. (2002). Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc. Natl. Acad. Sci. U. S. A. *99*, 15524–15529.

van Campenhout, M.J.H., van Bömmel, F., Pfefferkorn, M., Fischer, J., Deichsel, D., Boonstra, A., van Vuuren, A.J., Berg, T., Hansen, B.E., and Janssen, H.L.A. (2018). Host and viral factors associated with serum hepatitis B virus RNA levels among patients in need for treatment. Hepatol. Baltim. Md.

Castel, H., and Mathurin, P. (2011). Stéatohépatite non alcoolique.

Catteau, A., and Morris, J.R. (2002). BRCA1 methylation: a significant role in tumour development? Semin. Cancer Biol. *12*, 359–371.

Chekhun, V.F., Kulik, G.I., Yurchenko, O.V., Tryndyak, V.P., Todor, I.N., Luniv, L.S., Tregubova, N.A., Pryzimirska, T.V., Montgomery, B., Rusetskaya, N.V., et al. (2006). Role of DNA hypomethylation in the development of the resistance to doxorubicin in human MCF-7 breast adenocarcinoma cells. Cancer Lett. *231*, 87–93.

Chen, Y., Shen, A., Rider, P.J., Yu, Y., Wu, K., Mu, Y., Hao, Q., Liu, Y., Gong, H., Zhu, Y., et al. (2011). A liver-specific microRNA binds to a highly conserved RNA sequence of hepatitis B virus and negatively regulates viral gene expression and replication. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. *25*, 4511–4521.

Chisari, F.V., Pinkert, C.A., Milich, D.R., Filippi, P., McLachlan, A., Palmiter, R.D., and Brinster, R.L. (1985). A transgenic mouse model of the chronic hepatitis B surface antigen carrier state. Science *230*, 1157–1160.

Chu, C.-M., and Liaw, Y.-F. (2016). Natural History of Hepatitis B Virus Infection. In Hepatitis B Virus in Human Diseases, (Humana Press, Cham), pp. 217–247.

Costa, F.F., Paixão, V.A., Cavalher, F.P., Ribeiro, K.B., Cunha, I.W., Rinck, J.A., O'Hare, M., Mackay, A., Soares, F.A., Brentani, R.R., et al. (2006). SATR-1 hypomethylation is a common and early event in breast cancer. Cancer Genet. Cytogenet. *165*, 135–143.

Cougot, D., Wu, Y., Cairo, S., Caramel, J., Renard, C.-A., Lévy, L., Buendia, M.A., and Neuveut, C. (2007). The Hepatitis B Virus X Protein Functionally Interacts with CREB-binding Protein/p300 in the Regulation of CREB-mediated Transcription. J. Biol. Chem. *282*, 4277–4287.

Cova, L., Duflot, A., Prave, M., and Trepo, C. (1993). Duck hepatitis B virus infection, aflatoxin B1 and liver cancer in ducks. Arch. Virol. Suppl. *8*, 81–87.

Cui, X., Ludgate, L., Ning, X., and Hu, J. (2013). Maturation-associated destabilization of hepatitis B virus nucleocapsid. J. Virol. *87*, 11494–11503.

Dandri, M., and Lütgehetmann, M. (2014). Mouse models of hepatitis B and delta virus infection. J. Immunol. Methods *410*, 39–49.

Dandri, M., Schirmacher, P., and Rogler, C.E. (1996). Woodchuck hepatitis virus X protein is present in chronically infected woodchuck liver and woodchuck hepatocellular carcinomas which are permissive for viral replication. J. Virol. *70*, 5246–5254.

Dandri, M., Volz, T., and Lütgehetmann, M. (2016). Experimental Models: Cell Culture and Animal Models. In Hepatitis B Virus in Human Diseases, (Humana Press, Cham), pp. 35–62.

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.

Datta, S., Chatterjee, S., Veer, V., and Chakravarty, R. (2012). Molecular biology of the hepatitis B virus for clinicians. J. Clin. Exp. Hepatol. *2*, 353–365.

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.

Deng, W., and Lu, M. (2016). The role of microRNAs in hepatocyte metabolism and hepatitis B virus replication. Virol. Sin. *31*, 472–479.

Dény, P., and Zoulim, F. (2010). Hepatitis B virus: From diagnosis to treatment. Pathol. Biol. *58*, 245–253.

Desbois-Mouthon, C. (2011). De la stéatohépatite non alcoolique au carcinome hépatocellulaire : mécanismes physiopathologiques. /data/revues/19572557/00050001/38/.

DesJarlais, R., and Tummino, P.J. (2016). Role of Histone-Modifying Enzymes and Their Complexes in Regulation of Chromatin Biology. Biochemistry *55*, 1584–1599.

Du, Y., Kong, G., You, X., Zhang, S., Zhang, T., Gao, Y., Ye, L., and Zhang, X. (2012). Elevation of highly up-regulated in liver cancer (HULC) by hepatitis B virus X protein promotes hepatoma cell proliferation via down-regulating p18. J. Biol. Chem. *287*, 26302–26311.

Dusséaux, M., Masse-Ranson, G., Darche, S., Ahodantin, J., Li, Y., Fiquet, O., Beaumont, E., Moreau, P., Rivière, L., Neuveut, C., et al. (2017). Viral Load Affects the Immune Response to HBV in Mice With Humanized Immune System and Liver. Gastroenterology *153*, 1647-1661.e9.

Eckhardt, S.G., Milich, D.R., and McLachlan, A. (1991). Hepatitis B virus core antigen has two nuclear localization sequences in the arginine-rich carboxyl terminus. J. Virol. *65*, 575–582.

Engelke, M., Mills, K., Seitz, S., Simon, P., Gripon, P., Schnölzer, M., and Urban, S. (2006). Characterization of a hepatitis B and hepatitis delta virus receptor binding site. Hepatol. Baltim. Md *43*, 750–760.

Esposti, D.D., Hernandez-Vargas, H., Voegele, C., Fernandez-Jimenez, N., Forey, N., Bancel, B., Le Calvez-Kelm, F., McKay, J., Merle, P., and Herceg, Z. (2016). Identification of novel long non-coding RNAs deregulated in hepatocellular carcinoma using RNA-sequencing. Oncotarget *7*, 31862–31877.

Esteller, M. (2003). Cancer epigenetics: DNA methylation and chromatin alterations in human cancer. Adv. Exp. Med. Biol. *532*, 39–49.

Esteller, M. (2008). Epigenetics in cancer. N. Engl. J. Med. *358*, 1148–1159.

Esteller, M., Hamilton, S.R., Burger, P.C., Baylin, S.B., and Herman, J.G. (1999). Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. Cancer Res. *59*, 793–797.

Fan, C.-G., Wang, C.-M., Tian, C., Wang, Y., Li, L., Sun, W.-S., Li, R.-F., and Liu, Y.-G. (2011). miR-122 inhibits viral replication and cell proliferation in hepatitis B virus-related hepatocellular carcinoma and targets NDRG3. Oncol. Rep. *26*, 1281–1286.

Farazi, P.A., and DePinho, R.A. (2006). Hepatocellular carcinoma pathogenesis: from genes to environment. Nat. Rev. Cancer *6*, 674–687.

Fauquet, C.M., and Mayo, M.A. (2005). The Viruses. In Virus Taxonomy, (San Diego: Academic Press), p. 9.

Finn, R.S., Merle, P., Granito, A., Huang, Y.-H., Bodoky, G., Pracht, M., Yokosuka, O., Rosmorduc, O., Gerolami, R., Caparello, C., et al. (2018). Outcomes of sequential treatment with sorafenib followed by regorafenib for HCC: Additional analyses from the phase III RESORCE trial. J. Hepatol. *69*, 353–358.

Fleisher, A.S., Esteller, M., Tamura, G., Rashid, A., Stine, O.C., Yin, J., Zou, T.T., Abraham, J.M., Kong, D., Nishizuka, S., et al. (2001). Hypermethylation of the hMLH1 gene promoter is associated with microsatellite instability in early human gastric neoplasia. Oncogene *20*, 329–335.

Fornari, F., Gramantieri, L., Giovannini, C., Veronese, A., Ferracin, M., Sabbioni, S., Calin, G.A., Grazi, G.L., Croce, C.M., Tavolari, S., et al. (2009). MiR-122/cyclin G1 interaction modulates p53 activity and affects doxorubicin sensitivity of human hepatocarcinoma cells. Cancer Res. *69*, 5761–5767.

Fujimoto, A., Totoki, Y., Abe, T., Boroevich, K.A., Hosoda, F., Nguyen, H.H., Aoki, M., Hosono, N., Kubo, M., Miya, F., et al. (2012). Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. Nat. Genet. *44*, 760–764.

Fuks, F. (2005). DNA methylation and histone modifications: teaming up to silence genes. Curr. Opin. Genet. Dev. *15*, 490–495.

Galle, P.R., Forner, A., Llovet, J.M., Mazzaferro, V., Piscaglia, F., Raoul, J.-L., Schirmacher, P., and Vilgrain, V. (2018). EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. J. Hepatol. *69*, 182–236.

Gallucci, L., and Kann, M. (2017). Nuclear Import of Hepatitis B Virus Capsids and Genome. Viruses *9*.

Ganten, T.M., Stauber, R.E., Schott, E., Malfertheiner, P., Buder, R., Galle, P.R., Göhler, T., Walther, M., Koschny, R., and Gerken, G. (2017). Sorafenib in Patients with Hepatocellular Carcinoma-Results of the Observational INSIGHT Study. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. *23*, 5720–5728.

Glebe, D., and Urban, S. (2007). Viral and cellular determinants involved in hepadnaviral entry. World J. Gastroenterol. *13*, 22–38.

Gómez-Moreno, A., and Garaigorta, U. (2017). Hepatitis B Virus and DNA Damage Response: Interactions and Consequences for the Infection. Viruses *9*.

Gordian, E., Ramachandran, K., and Singal, R. (2009). Methylation mediated silencing of TMS1 in breast cancer and its potential contribution to docetaxel cytotoxicity. Anticancer Res. *29*, 3207–3210.

Gramantieri, L., Ferracin, M., Fornari, F., Veronese, A., Sabbioni, S., Liu, C.-G., Calin, G.A., Giovannini, C., Ferrazzi, E., Grazi, G.L., et al. (2007). Cyclin G1 is a target of miR-122a, a microRNA frequently down-regulated in human hepatocellular carcinoma. Cancer Res. *67*, 6092–6099.

Greger, V., Passarge, E., Höpping, W., Messmer, E., and Horsthemke, B. (1989). Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. Hum. Genet. *83*, 155–158.

Gripon, P., Diot, C., Thézé, N., Fourel, I., Loreal, O., Brechot, C., and Guguen-Guillouzo, C. (1988). Hepatitis B virus infection of adult human hepatocytes cultured in the presence of dimethyl sulfoxide. J. Virol. *62*, 4136–4143.

Gripon, P., Diot, C., and 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.

Gripon, P., Le Seyec, J., Rumin, S., and Guguen-Guillouzo, C. (1995). Myristylation of the hepatitis B virus large surface protein is essential for viral infectivity. Virology *213*, 292–299.

Gripon, P., Rumin, S., Urban, S., Le Seyec, J., Glaise, D., Cannie, I., Guyomard, C., Lucas, J., Trepo, C., and Guguen-Guillouzo, C. (2002). Infection of a human hepatoma cell line by hepatitis B virus. Proc. Natl. Acad. Sci. U. S. A. *99*, 15655–15660.

Guerrieri, F., Belloni, L., Pediconi, N., and Levrero, M. (2016). Pathobiology of Hepatitis B Virus-Induced Carcinogenesis. In Hepatitis B Virus in Human Diseases, (Humana Press, Cham), pp. 95–121.

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.

Guichard, C., Amaddeo, G., Imbeaud, S., Ladeiro, Y., Pelletier, L., Maad, I.B., Calderaro, J., Bioulac-Sage, P., Letexier, M., Degos, F., et al. (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat. Genet. *44*, 694–698.

Guidotti, L.G., Matzke, B., Schaller, H., and Chisari, F.V. (1995). High-level hepatitis B virus replication in transgenic mice. J. Virol. *69*, 6158–6169.

Guil, S., Soler, M., Portela, A., Carrère, J., Fonalleras, E., Gómez, A., Villanueva, A., and Esteller, M. (2012). Intronic RNAs mediate EZH2 regulation of epigenetic targets. Nat. Struct. Mol. Biol. *19*, 664–670.

Guo, Y., Li, Y., Mu, S., Zhang, J., and Yan, Z. (2009). Evidence that methylation of hepatitis B virus covalently closed circular DNA in liver tissues of patients with chronic hepatitis B modulates HBV replication. J. Med. Virol. *81*, 1177–1183.

Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.-C., Hung, T., Argani, P., Rinn, J.L., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature *464*, 1071–1076.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Hantz, O., Parent, R., Durantel, D., Gripon, P., Guguen-Guillouzo, C., and Zoulim, F. (2009). Persistence of the hepatitis B virus covalently closed circular DNA in HepaRG human hepatocyte-like cells. J. Gen. Virol. *90*, 127–135.

Hayes, C.N., and Chayama, K. (2016). MicroRNAs as Biomarkers for Liver Disease and Hepatocellular Carcinoma. Int. J. Mol. Sci. *17*, 280.

Hayes, J., Peruzzi, P.P., and Lawler, S. (2014). MicroRNAs in cancer: biomarkers, functions and therapy. Trends Mol. Med. *20*, 460–469.

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.

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.

Hong, X., Kim, E.S., and Guo, H. (2017). Epigenetic regulation of hepatitis B virus covalently closed circular DNA: Implications for epigenetic therapy against chronic hepatitis B. Hepatol. Baltim. Md *66*, 2066–2077.

Hu, J. (2016). Hepatitis B Virus Virology and Replication. In Hepatitis B Virus in Human Diseases, (Humana Press, Cham), pp. 1–34.

Hu, J., and Liu, K. (2017). Complete and Incomplete Hepatitis B Virus Particles: Formation, Function, and Application. Viruses *9*.

Hu, J., and Seeger, C. (2015). Hepadnavirus Genome Replication and Persistence. Cold Spring Harb. Perspect. Med. *5*, a021386.

Hu, J.-J., Song, W., Zhang, S.-D., Shen, X.-H., Qiu, X.-M., Wu, H.-Z., Gong, P.-H., Lu, S., Zhao, Z.-J., He, M.-L., et al. (2016). HBx-upregulated lncRNA UCA1 promotes cell growth and tumorigenesis by recruiting EZH2 and repressing p27Kip1/CDK2 signaling. Sci. Rep. *6*, 23521.

Iizuka, N., Oka, M., Yamada-Okabe, H., Mori, N., Tamesa, T., Okada, T., Takemoto, N., Tangoku, A., Hamada, K., Nakayama, H., et al. (2002). Comparison of gene expression profiles between hepatitis B virus- and hepatitis C virus-infected hepatocellular carcinoma by oligonucleotide microarray data on the basis of a supervised learning method. Cancer Res. *62*, 3939–3944.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789–802.

Inoue, T., and Tanaka, Y. (2016). Hepatitis B virus and its sexually transmitted infection - an update. Microb. Cell Graz Austria *3*, 420–437.

Iorio, M.V., and Croce, C.M. (2012). MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. EMBO Mol. Med. *4*, 143–159.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074–1080.

Ji, Z., Song, R., Regev, A., and Struhl, K. (2015). Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. ELife *4*, e08890.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. *13*, 484.

Jones, S.A., and Hu, J. (2013). Hepatitis B virus reverse transcriptase: diverse functions as classical and emerging targets for antiviral intervention. Emerg. Microbes Infect. *2*, e56.

Jones, S.A., Boregowda, R., Spratt, T.E., and Hu, J. (2012). In vitro epsilon RNA-dependent protein priming activity of human hepatitis B virus polymerase. J. Virol. *86*, 5134–5150.

Junker-Niepmann, M., Bartenschlager, R., and Schaller, H. (1990). A short cis-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. EMBO J. *9*, 3389–3396.

Kann, M., Bischof, A., and Gerlich, W.H. (1997). In vitro model for the nuclear transport of the hepadnavirus genome. J. Virol. *71*, 1310–1316.

Kato, H., Sugiyama, M., and Mizokami, M. (2016). Hepatitis B Virus Genotypes. In Hepatitis B Virus in Human Diseases, (Humana Press, Cham), pp. 63–78.

Katoh, M. (2005). Epithelial-mesenchymal transition in gastric cancer (Review). Int. J. Oncol. *27*, 1677–1683.

Kaur, P., Paliwal, A., Durantel, D., Hainaut, P., Scoazec, J.-Y., Zoulim, F., Chemin, I., and Herceg, Z. (2010). DNA methylation of hepatitis B virus (HBV) genome associated with the development of hepatocellular carcinoma and occult HBV infection. J. Infect. Dis. *202*, 700–704.

Kim, K.S., and Lee, Y.I. (1997). Biallelic expression of the H19 and IGF2 genes in hepatocellular carcinoma. Cancer Lett. *119*, 143–148.

Kim, C.-M., Koike, K., Saito, I., Miyamura, T., and Jay, G. (1991). HBx gene of hepatitis B virus induces liver cancer in transgenic mice. Nature *351*, 317.

Kim, D.Y., Cho, M.H., Yang, H.K., Hemminki, K., Kim, J.P., Jang, J.J., and Kumar, R. (1999). Detection of methylation damage in DNA of gastric cancer tissues using 32P postlabelling assay. Jpn. J. Cancer Res. Gann *90*, 1104–1108.

Kim, J.-W., Lee, S.H., Park, Y.S., Hwang, J.-H., Jeong, S.-H., Kim, N., and Lee, D.H. (2011). Replicative activity of hepatitis B virus is negatively associated with methylation of covalently closed circular DNA in advanced hepatitis B virus infection. Intervirology *54*, 316–325.

Kim, K.-H., Choi, J.-S., Kim, I.-J., Ku, J.-L., and Park, J.-G. (2006). Promoter hypomethylation and reactivation of MAGE-A1 and MAGE-A3 genes in colorectal cancer cell lines and cancer tissues. World J. Gastroenterol. *12*, 5651–5657.

Kinoshita, W., Ogura, N., Watashi, K., and Wakita, T. (2017). Host factor PRPF31 is involved in cccDNA production in HBV-replicating cells. Biochem. Biophys. Res. Commun. *482*, 638–644.

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.

Kogo, R., Shimamura, T., Mimori, K., Kawahara, K., Imoto, S., Sudo, T., Tanaka, F., Shibata, K., Suzuki, A., Komune, S., et al. (2011). Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. Cancer Res. *71*, 6320–6326.

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.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. Cell *128*, 693–705.

Kramvis, A. (2014). Genotypes and genetic variability of hepatitis B virus. Intervirology *57*, 141–150.

Krol, J., Loedige, I., and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. Nat. Rev. Genet. *11*, 597–610.

Kulis, M., and Esteller, M. (2010). DNA methylation and cancer. Adv. Genet. *70*, 27–56.

Lachenmayer, A., Toffanin, S., Cabellos, L., Alsinet, C., Hoshida, Y., Villanueva, A., Minguez, B., Tsai, H.-W., Ward, S.C., Thung, S., et al. (2012). Combination therapy for hepatocellular carcinoma: Additive preclinical efficacy of the HDAC inhibitor panobinostat with sorafenib. J. Hepatol. *56*, 1343–1350.

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.

Lamontagne, J., Steel, L.F., and Bouchard, M.J. (2015). Hepatitis B virus and microRNAs: Complex interactions affecting hepatitis B virus replication and hepatitis B virus-associated diseases. World J. Gastroenterol. *21*, 7375–7399.

Lamontagne, J., Mell, J.C., and Bouchard, M.J. (2016). Transcriptome-Wide Analysis of Hepatitis B Virus-Mediated Changes to Normal Hepatocyte Gene Expression. PLoS Pathog. *12*, e1005438.

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.

Lan, Y.T., Li, J., Liao, W., and Ou, J. (1999). Roles of the three major phosphorylation sites of hepatitis B virus core protein in viral replication. Virology *259*, 342–348.

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.

Le Seyec, J., Chouteau, P., Cannie, I., Guguen-Guillouzo, C., and Gripon, P. (1998). Role of the pre-S2 domain of the large envelope protein in hepatitis B virus assembly and infectivity. J. Virol. *72*, 5573–5578.

Lebossé, F., Testoni, B., Fresquet, J., Facchetti, F., Galmozzi, E., Fournier, M., Hervieu, V., Berthillon, P., Berby, F., Bordes, I., et al. (2017). Intrahepatic innate immune response pathways are downregulated in untreated chronic hepatitis B. J. Hepatol. *66*, 897–909.

LeCluyse, E.L., and Alexandre, E. (2010). Isolation and Culture of Primary Hepatocytes from Resected Human Liver Tissue. In Hepatocytes, (Humana Press), pp. 57–82.

Leistner, C.M., Gruen-Bernhard, S., and Glebe, D. (2008). Role of glycosaminoglycans for binding and infection of hepatitis B virus. Cell. Microbiol. *10*, 122–133.

Lepère-Douard, C., Trotard, M., Le Seyec, J., and Gripon, P. (2009). The first transmembrane domain of the hepatitis B virus large envelope protein is crucial for infectivity. J. Virol. *83*, 11819–11829.

Levrero, M., and Zucman-Rossi, J. (2016). Mechanisms of HBV-induced hepatocellular carcinoma. J. Hepatol. *64*, S84–S101.

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.

Li, W. (2015). The hepatitis B virus receptor. Annu. Rev. Cell Dev. Biol. *31*, 125–147.

Li, M., Zhao, H., Zhang, X., Wood, L.D., Anders, R.A., Choti, M.A., Pawlik, T.M., Daniel, H.D., Kannangai, R., Offerhaus, G.J.A., et al. (2011). Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma. Nat. Genet. *43*, 828–829.

Liang, G., and Weisenberger, D.J. (2017). DNA methylation aberrancies as a guide for surveillance and treatment of human cancers. Epigenetics *12*, 416–432.

Lin, R.-K., Hsu, H.-S., Chang, J.-W., Chen, C.-Y., Chen, J.-T., and Wang, Y.-C. (2007). Alteration of DNA methyltransferases contributes to 5'CpG methylation and poor prognosis in lung cancer. Lung Cancer Amst. Neth. *55*, 205–213.

Liu, K., Ludgate, L., Yuan, Z., and Hu, J. (2015a). Regulation of multiple stages of hepadnavirus replication by the carboxyl-terminal domain of viral core protein in trans. J. Virol. *89*, 2918–2930.

Liu, S., Koh, S.S.Y., and Lee, C.G.L. (2016). Hepatitis B Virus X Protein and Hepatocarcinogenesis. Int. J. Mol. Sci. *17*.

Liu, X.Y., Tang, S.H., Wu, S.L., Luo, Y.H., Cao, M.R., Zhou, H.K., Jiang, X.W., Shu, J.C., Bie, C.Q., Huang, S.M., et al. (2015b). Epigenetic modulation of insulin-like growth factor-II overexpression by hepatitis B virus X protein in hepatocellular carcinoma. Am. J. Cancer Res. *5*, 956–978.

Llovet, J.M., Zucman-Rossi, J., Pikarsky, E., Sangro, B., Schwartz, M., Sherman, M., and Gores, G. (2016). Hepatocellular carcinoma. Nat. Rev. Dis. Primer *2*, 16018.

Locatelli, M., Fresquet, J., Maadadi, S., Quivy, J.P., Testoni, B., and Zoulim, F. (2018). PS-113 - Characterization of host and viral proteins involved in the chromatinization of the hepatitis B virus minichromosome. J. Hepatol. *68*, S61.

Lok, A.S., Zoulim, F., Dusheiko, G., and Ghany, M.G. (2017). Hepatitis B cure: From discovery to regulatory approval. J. Hepatol. *67*, 847–861.

Luangsay, S., Ait-Goughoulte, M., Michelet, M., Floriot, O., Bonnin, M., Gruffaz, M., Rivoire, M., Fletcher, S., Javanbakht, H., Lucifora, J., et al. (2015). Expression and functionality of Toll- and RIG-like receptors in HepaRG cells. J. Hepatol. *63*, 1077–1085.

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.

Lucifora, J., Salvetti, A., Marniquet, X., Mailly, L., Testoni, B., Fusil, F., Inchauspé, A., Michelet, M., Michel, M.-L., Levrero, M., et al. (2017). Detection of the hepatitis B virus (HBV) covalently-closed-circular DNA (cccDNA) in mice transduced with a recombinant AAV-HBV vector. Antiviral Res. *145*, 14–19.

Lucifora, J., Bonnin, M., Aillot, L., Fusil, F., Maadadi, S., Dimier, L., Michelet, M., Floriot, O., Ollivier, A., Rivoire, M., et al. (2018). Direct antiviral properties of TLR ligands against HBV replication in immune-competent hepatocytes. Sci. Rep. *8*, 5390.

Margueron, R., and Reinberg, D. (2010). Chromatin structure and the inheritance of epigenetic information. Nat. Rev. Genet. *11*, 285–296.

Marion, P.L., and Robinson, W.S. (1983). Hepadna viruses: hepatitis B and related viruses. Curr. Top. Microbiol. Immunol. *105*, 99–121.

Mason, W.S., Aldrich, C., Summers, J., and Taylor, J.M. (1982). Asymmetric replication of duck hepatitis B virus DNA in liver cells: Free minus-strand DNA. Proc. Natl. Acad. Sci. U. S. A. *79*, 3997–4001.

Mattick, J.S., and Makunin, I.V. (2005). Small regulatory RNAs in mammals. Hum. Mol. Genet. *14 Spec No 1*, R121-132.

Maynard, J.E., Berquist, K.R., Krushak, D.H., and Purcell, R.H. (1972). Experimental infection of chimpanzees with the virus of hepatitis B. Nature *237*, 514–515.

Melo, S.A., Moutinho, C., Ropero, S., Calin, G.A., Rossi, S., Spizzo, R., Fernandez, A.F., Davalos, V., Villanueva, A., Montoya, G., et al. (2010). A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. Cancer Cell *18*, 303–315.

Mercer, T.R., Dinger, M.E., and Mattick, J.S. (2009). Long non-coding RNAs: insights into functions. Nat. Rev. Genet. *10*, 155–159.

Michie, A.M., McCaig, A.M., Nakagawa, R., and Vukovic, M. (2010). Death-associated protein kinase (DAPK) and signal transduction: regulation in cancer. FEBS J. *277*, 74–80.

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.

Moyo, B., Nicholson, S.A., and Arbuthnot, P.B. (2016). The role of long non-coding RNAs in hepatitis B virus-related hepatocellular carcinoma. Virus Res. *212*, 103–113.

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 gutjnl-2015-309809.

Ng, S.S., Yue, W.W., Oppermann, U., and Klose, R.J. (2009). Dynamic protein methylation in chromatin biology. Cell. Mol. Life Sci. *66*, 407.

Ni, Y., Sonnabend, J., Seitz, S., and Urban, S. (2010). The pre-s2 domain of the hepatitis B virus is dispensable for infectivity but serves a spacer function for L-protein-connected virus assembly. J. Virol. *84*, 3879–3888.

Ni, Y., Lempp, F.A., Mehrle, S., Nkongolo, S., Kaufman, C., Fälth, M., Stindt, J., Königer, C., Nassal, M., Kubitz, R., et al. (2014). Hepatitis B and D viruses exploit sodium taurocholate co-transporting polypeptide for species-specific entry into hepatocytes. Gastroenterology *146*, 1070–1083.

Ning, X., Basagoudanavar, S.H., Liu, K., Luckenbaugh, L., Wei, D., Wang, C., Wei, B., Zhao, Y., Yan, T., Delaney, W., et al. (2017). Capsid Phosphorylation State and Hepadnavirus Virion Secretion. J. Virol. *91*.

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.

Oki, M., Aihara, H., and Ito, T. (2007). Role of histone phosphorylation in chromatin dynamics and its implications in diseases. Subcell. Biochem. *41*, 319–336.

Pairan, A., and Bruss, V. (2009). Functional surfaces of the hepatitis B virus capsid. J. Virol. *83*, 11616–11623.

Palumbo, G.A., Scisciani, C., Pediconi, N., Lupacchini, L., Alfaiate, D., Guerrieri, F., Calvo, L., Salerno, D., Di Cocco, S., Levrero, M., et al. (2015). Correction: IL6 Inhibits HBV Transcription by Targeting the Epigenetic Control of the Nuclear cccDNA Minichromosome. PLoS ONE *10*, e0145555.

Pan, Y., Qin, T., Feng, L., and Yu, Z. (2013). Expression profile of altered long non-coding RNAs in patients with HBV-associated hepatocellular carcinoma. J. Huazhong Univ. Sci. Technol. Med. Sci. Hua Zhong Ke Ji Xue Xue Bao Yi Xue Ying Wen Ban Huazhong Keji Daxue Xuebao Yixue Yingdewen Ban *33*, 96–101.

Paredes, J., Albergaria, A., Oliveira, J.T., Jerónimo, C., Milanezi, F., and Schmitt, F.C. (2005). P-cadherin overexpression is an indicator of clinical outcome in invasive breast carcinomas and is associated with CDH3 promoter hypomethylation. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. *11*, 5869–5877.

Patel, N., White, S.J., Thompson, R.F., Bingham, R., Weiß, E.U., Maskell, D.P., Zlotnick, A., Dykeman, E., Tuma, R., Twarock, R., et al. (2017). HBV RNA pre-genome encodes specific motifs that mediate

interactions with the viral core protein that promote nucleocapsid assembly. Nat. Microbiol. *2*, 17098.

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.

Peng, F., Xiao, X., Jiang, Y., Luo, K., Tian, Y., Peng, M., Zhang, M., Xu, Y., and Gong, G. (2014). HBx down-regulated Gld2 plays a critical role in HBV-related dysregulation of miR-122. PloS One *9*, e92998.

Perlman, D.H., Berg, E.A., O'connor, P.B., Costello, C.E., and Hu, J. (2005). Reverse transcription-associated dephosphorylation of hepadnavirus nucleocapsids. Proc. Natl. Acad. Sci. U. S. A. *102*, 9020–9025.

Peschansky, V.J., and Wahlestedt, C. (2014). Non-coding RNAs as direct and indirect modulators of epigenetic regulation. Epigenetics *9*, 3–12.

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.

Ponsel, D., and Bruss, V. (2003). Mapping of amino acid side chains on the surface of hepatitis B virus capsids required for envelopment and virion formation. J. Virol. *77*, 416–422.

Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nat. Biotechnol. *28*, 1057–1068.

Prange, R. (2012). Host factors involved in hepatitis B virus maturation, assembly, and egress. Med. Microbiol. Immunol. (Berl.) *201*, 449–461.

Prange, R., and Streeck, R.E. (1995). Novel transmembrane topology of the hepatitis B virus envelope proteins. EMBO J. *14*, 247–256.

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*, e1005893.

Qiu, L., Wang, T., Xu, X., Wu, Y., Tang, Q., and Chen, K. (2017). Long Non-Coding RNAs in Hepatitis B Virus-Related Hepatocellular Carcinoma: Regulation, Functions, and Underlying Mechanisms. Int. J. Mol. Sci. *18*, 2505.

Quasdorff, M., and Protzer, U. (2010). Control of hepatitis B virus at the level of transcription. J. Viral Hepat. *17*, 527–536.

Renganathan, A., and Felley-Bosco, E. (2017). Long Noncoding RNAs in Cancer and Therapeutic Potential. Adv. Exp. Med. Biol. *1008*, 199–222.

Revill, P., Testoni, B., Locarnini, S., and Zoulim, F. (2016). Global strategies are required to cure and eliminate HBV infection. Nat. Rev. Gastroenterol. Hepatol. *13*, 239–248.

Ribeiro, A.S., Albergaria, A., Sousa, B., Correia, A.L., Bracke, M., Seruca, R., Schmitt, F.C., and Paredes, J. (2010). Extracellular cleavage and shedding of P-cadherin: a mechanism underlying the invasive behaviour of breast cancer cells. Oncogene *29*, 392–402.

Ringelhan, M., O'Connor, T., Protzer, U., and Heikenwalder, M. (2015). The direct and indirect roles of HBV in liver cancer: prospective markers for HCC screening and potential therapeutic targets. J. Pathol. *235*, 355–367.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. *81*.

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.

Robertson, K.D., and Jones, P.A. (2000). DNA methylation: past, present and future directions. Carcinogenesis *21*, 461–467.

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.

Roll, J.D., Rivenbark, A.G., Jones, W.D., and Coleman, W.B. (2008). DNMT3b overexpression contributes to a hypermethylator phenotype in human breast cancer cell lines. Mol. Cancer *7*, 15.

Saito, Y., Kanai, Y., Sakamoto, M., Saito, H., Ishii, H., and Hirohashi, S. (2002). Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b4, associated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis. Proc. Natl. Acad. Sci. U. S. A. *99*, 10060–10065.

Salisse, J., and Sureau, C. (2009). A function essential to viral entry underlies the hepatitis B virus "a" determinant. J. Virol. *83*, 9321–9328.

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.

Schreiner, S., and Nassal, M. (2017). A Role for the Host DNA Damage Response in Hepatitis B Virus cccDNA Formation-and Beyond? Viruses *9*.

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.

Schulze-Bergkamen, H., Untergasser, A., Dax, A., Vogel, H., Büchler, P., Klar, E., Lehnert, T., Friess, H., Büchler, M.W., Kirschfink, M., et al. (2003). Primary human hepatocytes – a valuable tool for investigation of apoptosis and hepatitis B virus infection. J. Hepatol. *38*, 736–744.

Seeger, C., and Mason, W.S. (2015). Molecular biology of hepatitis B virus infection. Virology.

Seeger, C., Zoulim, F., and Mason, W.S. (2013). Hepadnaviruses. In Fields Virology, (Lippincott Williams and Wilkins), pp. 2185–2221.

Sells, M.A., Chen, M.L., and Acs, G. (1987). Production of hepatitis B virus particles in Hep G2 cells transfected with cloned hepatitis B virus DNA. Proc. Natl. Acad. Sci. U. S. A. *84*, 1005–1009.

Selzer, L., and Zlotnick, A. (2015). Assembly and Release of Hepatitis B Virus. Cold Spring Harb. Perspect. Med. *5*, a021394.

Sharma, S., Kelly, T.K., and Jones, P.A. (2010). Epigenetics in cancer. Carcinogenesis *31*, 27–36.

Shouval, D. (2003). Hepatitis B vaccines. J. Hepatol. *39 Suppl 1*, S70-76.

Simon, J.A., and Lange, C.A. (2008). Roles of the EZH2 histone methyltransferase in cancer epigenetics. Mutat. Res. Mol. Mech. Mutagen. *647*, 21–29.

Simonin, Y., Vegna, S., Akkari, L., Grégoire, D., Antoine, E., Piette, J., Floc'h, N., Lassus, P., Yu, G.-Y., Rosenberg, A.R., et al. (2013). Lymphotoxin signaling is initiated by the viral polymerase in HCV-linked tumorigenesis. PLoS Pathog. *9*, e1003234.

Singh, A.K., Kumar, R., and Pandey, A.K. (2018). Hepatocellular Carcinoma: Causes, Mechanism of Progression and Biomarkers. Curr. Chem. Genomics Transl. Med. *12*, 9–26.

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.

Stephenne, J. (1990). Development and production aspects of a recombinant yeast-derived hepatitis B vaccine. Vaccine *8 Suppl*, S69-73; discussion S79-80.

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. In Advances in Virus Research, (Academic Press), pp. 125–164.

Strick-Marchand, H., Dusséaux, M., Darche, S., Huntington, N.D., Legrand, N., Masse-Ranson, G., Corcuff, E., Ahodantin, J., Weijer, K., Spits, H., et al. (2015). A novel mouse model for stable engraftment of a human immune system and human hepatocytes. PloS One *10*, e0119820.

Studach, L.L., Menne, S., Cairo, S., Buendia, M.A., Hullinger, R.L., Lefrançois, L., Merle, P., and Andrisani, O.M. (2012). Subset of Suz12/PRC2 target genes is activated during hepatitis B virus replication and liver carcinogenesis associated with HBV X protein. Hepatology *56*, 1240–1251.

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.

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.

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.

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.

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.

Suslov, A., Boldanova, T., Wang, X., Wieland, S., and Heim, M.H. (2018). Hepatitis B Virus Does Not Interfere With Innate Immune Responses in the Human Liver. Gastroenterology *154*, 1778–1790.

Tang, J., Xie, Y., Xu, X., Yin, Y., Jiang, R., Deng, L., Tan, Z., Gangarapu, V., Tang, J., and Sun, B. (2017). Bidirectional transcription of Linc00441 and RB1 via H3K27 modification-dependent way promotes hepatocellular carcinoma. Cell Death Dis. *8*, e2675.

Tennant, B.C., Toshkov, I.A., Peek, S.F., Jacob, J.R., Menne, S., Hornbuckle, W.E., Schinazi, R.D., Korba, B.E., Cote, P.J., and Gerin, J.L. (2004). Hepatocellular carcinoma in the woodchuck model of hepatitis B virus infection. Gastroenterology *127*, S283–S293.

Tong, S., and Revill, P. (2016). Overview of hepatitis B viral replication and genetic variability. J. Hepatol. *64*, S4–S16.

Trépo, C., Chan, H.L.Y., and Lok, A. (2014). Hepatitis B virus infection. Lancet Lond. Engl. *384*, 2053–2063.

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.

Urban, S., Bartenschlager, R., Kubitz, R., and Zoulim, F. (2014). Strategies to Inhibit Entry of HBV and HDV Into Hepatocytes. Gastroenterology *147*, 48–64.

Valaydon, Z.S., and Locarnini, S.A. (2017). The virological aspects of hepatitis B. Best Pract. Res. Clin. Gastroenterol. *31*, 257–264.

Varambally, S., Cao, Q., Mani, R.-S., Shankar, S., Wang, X., Ateeq, B., Laxman, B., Cao, X., Jing, X., Ramnarayanan, K., et al. (2008). Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. Science *322*, 1695–1699.

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.

Villa, E., Critelli, R., Lei, B., Marzocchi, G., Cammà, C., Giannelli, G., Pontisso, P., Cabibbo, G., Enea, M., Colopi, S., et al. (2016). Neoangiogenesis-related genes are hallmarks of fast-growing hepatocellular carcinomas and worst survival. Results from a prospective study. Gut *65*, 861–869.

Villanueva, A., Portela, A., Sayols, S., Battiston, C., Hoshida, Y., Méndez-González, J., Imbeaud, S., Letouzé, E., Hernandez-Gea, V., Cornella, H., et al. (2015). DNA Methylation-based prognosis and epidrivers in hepatocellular carcinoma. Hepatology n/a-n/a.

Vivekanandan, P., Thomas, D., and Torbenson, M. (2008). Hepatitis B viral DNA is methylated in liver tissues. J. Viral Hepat. *15*, 103–107.

Vivekanandan, P., Thomas, D., and Torbenson, M. (2009). Methylation regulates hepatitis B viral protein expression. J. Infect. Dis. *199*, 1286–1291.

Walsh, R., and Locarnini, S. (2012). Hepatitis B precore protein: pathogenic potential and therapeutic promise. Yonsei Med. J. *53*, 875–885.

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.

Wang, K.C., and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Mol. Cell *43*, 904–914.

Wang, C., Wang, L., Ding, Y., Lu, X., Zhang, G., Yang, J., Zheng, H., Wang, H., Jiang, Y., and Xu, L. (2017a). LncRNA Structural Characteristics in Epigenetic Regulation. Int. J. Mol. Sci. *18*, 2659.

Wang, G.H., Zoulim, F., Leber, E.H., Kitson, J., and Seeger, C. (1994). Role of RNA in enzymatic activity of the reverse transcriptase of hepatitis B viruses. J. Virol. *68*, 8437–8442.

Wang, J., Liu, X., Wu, H., Ni, P., Gu, Z., Qiao, Y., Chen, N., Sun, F., and Fan, Q. (2010). CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. Nucleic Acids Res. *38*, 5366–5383.

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.

Wang, J.C.-Y., Dhason, M.S., and Zlotnick, A. (2012). Structural organization of pregenomic RNA and the carboxy-terminal domain of the capsid protein of hepatitis B virus. PLoS Pathog. *8*, e1002919.

Wang, X., Goodrich, K.J., Gooding, A.R., Naeem, H., Archer, S., Paucek, R.D., Youmans, D.T., Cech, T.R., and Davidovich, C. (2017b). Targeting of Polycomb Repressive Complex 2 to RNA by Short Repeats of Consecutive Guanines. Mol. Cell *65*, 1056-1067.e5.

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.

Watashi, K., and Wakita, T. (2015). Hepatitis B Virus and Hepatitis D Virus Entry, Species Specificity, and Tissue Tropism. Cold Spring Harb. Perspect. Med. *5*, a021378.

Weber, J., Salgaller, M., Samid, D., Johnson, B., Herlyn, M., Lassam, N., Treisman, J., and Rosenberg, S.A. (1994). Expression of the MAGE-1 tumor antigen is up-regulated by the demethylating agent 5-aza-2'-deoxycytidine. Cancer Res. *54*, 1766–1771.

Wei, X., Xiang, T., Ren, G., Tan, C., Liu, R., Xu, X., and Wu, Z. (2013). miR-101 is down-regulated by the hepatitis B virus x protein and induces aberrant DNA methylation by targeting DNA methyltransferase 3A. Cell. Signal. *25*, 439–446.

Wei, Y., Neuveut, C., Tiollais, P., and Buendia, M.-A. (2010). Molecular biology of the hepatitis B virus and role of the X gene. Pathol. Biol. *58*, 267–272.

Weller, M., Stupp, R., Reifenberger, G., Brandes, A.A., van den Bent, M.J., Wick, W., and Hegi, M.E. (2010). MGMT promoter methylation in malignant gliomas: ready for personalized medicine? Nat. Rev. Neurol. *6*, 39–51.

Widschwendter, M., Siegmund, K.D., Müller, H.M., Fiegl, H., Marth, C., Müller-Holzner, E., Jones, P.A., and Laird, P.W. (2004a). Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. Cancer Res. *64*, 3807–3813.

Widschwendter, M., Jiang, G., Woods, C., Müller, H.M., Fiegl, H., Goebel, G., Marth, C., Müller-Holzner, E., Zeimet, A.G., Laird, P.W., et al. (2004b). DNA hypomethylation and ovarian cancer biology. Cancer Res. *64*, 4472–4480.

Wieland, S., Thimme, R., Purcell, R.H., and Chisari, F.V. (2004). Genomic analysis of the host response to hepatitis B virus infection. Proc. Natl. Acad. Sci. U. S. A. *101*, 6669–6674.

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.

Wu, G., Yu, F., Xiao, Z., Xu, K., Xu, J., Tang, W., Wang, J., and Song, E. (2011). Hepatitis B virus X protein downregulates expression of the miR-16 family in malignant hepatocytes in vitro. Br. J. Cancer *105*, 146–153.

Wu, L.-M., Yang, Z., Zhou, L., Zhang, F., Xie, H.-Y., Feng, X.-W., Wu, J., and Zheng, S.-S. (2010). Identification of Histone Deacetylase 3 as a Biomarker for Tumor Recurrence Following Liver Transplantation in HBV-Associated Hepatocellular Carcinoma. PLoS ONE *5*, e14460.

Wynne, S.A., Crowther, R.A., and Leslie, A.G. (1999). The crystal structure of the human hepatitis B virus capsid. Mol. Cell *3*, 771–780.

Xhemalce, B., Dawson, M.A., and Bannister, A.J. (2006). Histone Modifications. In Reviews in Cell Biology and Molecular Medicine, (Wiley-VCH Verlag GmbH & Co. KGaA), p.

Xu, W., Yu, J., and Wong, V.W.-S. (2017). Mechanism and prediction of HCC development in HBV infection. Best Pract. Res. Clin. Gastroenterol. *31*, 291–298.

Yan, H., Zhong, G., Xu, G., He, W., Jing, Z., Gao, Z., Huang, Y., Qi, Y., Peng, B., Wang, H., et al. (2012). Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. ELife *1*, e00049.

Yang, Y., Chen, L., Gu, J., Zhang, H., Yuan, J., Lian, Q., Lv, G., Wang, S., Wu, Y., Yang, Y.-C.T., et al. (2017). Recurrently deregulated lncRNAs in hepatocellular carcinoma. Nat. Commun. *8*, 14421.

Yeh, C.T., Liaw, Y.F., and Ou, J.H. (1990). The arginine-rich domain of hepatitis B virus precore and core proteins contains a signal for nuclear transport. J. Virol. *64*, 6141–6147.

Yu, G., Bing, Y., Li, W., Xia, L., and Liu, Z. (2014). Hepatitis B virus inhibits the expression of CD82 through hypermethylation of its promoter in hepatoma cells. Mol. Med. Rep. *10*, 2580–2586.

Zhang, Q., Matsuura, K., Kleiner, D.E., Zamboni, F., Alter, H.J., and Farci, P. (2016). Analysis of long noncoding RNA expression in hepatocellular carcinoma of different viral etiology. J. Transl. Med. *14*, 328.

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. Hepatol. Baltim. Md *66*, 398–415.

Zhang, Y., Li, C., Zhang, Y., Zhu, H., Kang, Y., Liu, H., Wang, J., Qin, Y., Mao, R., Xie, Y., et al. (2013). Comparative analysis of CpG islands among HBV genotypes. PloS One *8*, e56711.

Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M., and Lee, J.T. (2010). Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol. Cell *40*, 939–953.

Zlotnick, A., Cheng, N., Stahl, S.J., Conway, J.F., Steven, A.C., and Wingfield, P.T. (1997). Localization of the C terminus of the assembly domain of hepatitis B virus capsid protein: implications for morphogenesis and organization of encapsidated RNA. Proc. Natl. Acad. Sci. U. S. A. *94*, 9556–9561.

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.

Zoulim, F., and Durantel, D. (2015). Antiviral therapies and prospects for a cure of chronic hepatitis B. Cold Spring Harb. Perspect. Med. *5*.

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., Saputelli, J., and Seeger, C. (1994). Woodchuck hepatitis virus X protein is required for viral infection in vivo. J. Virol. *68*, 2026–2030.

Zoulim, F., Levrero, M., and Testoni, B. (2018). Hepatitis B.

OMS | Hépatite B.