Factors involved in the regulation of Long-Interspersed-Nuclear-Elements (L1) retrotransposons in the context of Hepatocellular Carcinoma

PRAVEEN DHONDURAO SUDHINDAR

PhD thesis

A thesis submitted for the degree of Doctor of Philosophy in the Newcastle University Centre for Cancer, Biosciences Institute, Newcastle University

July 2021

Abstract

Hepatocellular carcinoma (HCC) generally develops on the background of a chronic liver disease following the accumulation of genetic damage and epigenetic alterations of growth regulatory genes, leading to activation of oncogenes and loss of function of tumour suppressor genes. Recent studies indicate that epigenetic aspects play an important role in the initiation of HCC. This includes dysregulation of repeat elements belonging to the Long Interspersed Nuclear Elements (LINE1 or L1) class. The L1 elements are autonomous mobile elements and upon activation contribute towards genomic instability via insertional mutagenesis. The thesis is aimed at understanding the factors leading to aberrant activation of retrotransposons and regulators of active retrotransposition in the context of HCC. All the liver cancer cell lines (Huh7, HepG2, Hep3B, PLC-PRF/5 and SK-Hep1) supported active retrotransposition in vitro irrespective of their basal L1 expression status or TP53 status. Since, active L1 retrotransposition through 'Target Primed Reverse Transcription' (TPRT) involves first DNA strand nicking by ORF2 endonuclease followed by second strand cleavage, we hypothesised that the DNA damage response pathways are involved in regulating the process. To decipher the influence of individual DNA repair pathway elements on the process of active retrotransposition, small molecule inhibitors towards ATM (KU-55933), DNA-PK (NU-7441), ATR (VE-821), CHK1 (SRA737) and PARP (Rucaparib) were utilised. Overall, inhibition of ATR (Ataxia Telangiectasia And Rad3-Related Protein), a serine/threonine kinase involved in DNA replication stress and DNA damage signalling increased retrotransposition rate in all the cell lines. In addition, an increase in active retrotransposition was observed in Huh7 cell in presence of subgenomic copy of Hepatitis C Virus (HCV, a prevalent cause of HCC and contributes towards hepatocarcinogenesis by inducing oxidative stress, DNA damage and epigenetic changes in hepatocytes). Interestingly, the rate of retrotransposition remained higher in cells compared to control cell lines even when they were treated with PSI7977 (antiviral agent) successfully eliminating the viral genome from the cells. Hence, HCV upregulated active retrotransposition even beyond viral clearance and thus can contribute towards hepatocarcinogenesis by a 'hit-and-run' mechanism. Interrogating publicly available datasets - GSE84346 (RNAseq of Chronic HCV Hepatitis (CHC) patients and controls) and RNAseq data of non-tumour liver from the Cancer Genome Atlas HCC study - confirmed upregulation of L1 transcripts in chronic hepatitis patients liver. Hence, L1s can be activated

before oncogenic transformation in CHC patients, with HCV-activated retrotransposition contributing towards genomic instability leading to HCC development. However, direct role of L1 in cellular transformation has not been demonstrated so far. Hence, to evaluate the potential of L1 insertions to initiate cellular transformation in hepatocytes, we developed a model of active L1 retrotransposition in immortalised hepatocyte cell line (HHL-5). In brief, GFP-based retrotransposition assay was set up in HHL-5 cell line and the GFP positive population was FACS sorted and the clones which emerged were screened for L1/GFP insertion and oncogenic properties. One out of 10 clones exhibited potential oncogenic transformation based on in vitro assessment however, further characterisation is needed to confirm this. In parallel, a CRISPR-Cas9 based system containing deactivated Cas9 (dCas9) and L1 promoter specific gRNA was developed with the aim to selectively isolate L1 promoters and study chromatin interactions at the loci under different physiological conditions. A 9.7 fold enrichment of L1 promoter was observed as compared to background DNA in Huh7 cells transiently transfected with dCas9+L1gRNA construct. However, in the given timeframe the system was not scaled up to carryout proteomic analysis to identify factors bound at the promoter. Overall, the study highlights potential of using HCC-related cell lines to study the influence of exogenous factors (such as HCV) and endogenous regulatory pathways (such as DNA repair pathways) affecting L1 expression/retrotransposition and have generated several tools for future investigations.

Acknowledgements

First of all, I would like to thank my primary supervisor Dr Ruchi Shukla for her invaluable supervision and support during the course of my PhD degree. I would like to thank Ruchi for the advice and encouragement provided throughout my time as a PhD student. I feel extremely lucky to have a supervisor who responded to my queries so promptly. Also, I would like to thank Prof. John Lunec for his helpful feedback in every team meeting that helped shape this project. Thank you, John, for your critical advice and for suggesting many important additions and improvements. I would also like to express my gratitude to Prof Helen Reeves for providing valuable feedback on my yearly reports. I wish to thank my lab colleagues for their help and support. I have enjoyed working with every one of them and hope to stay in touch in future. I would also like to thank Newcastle University NUORS fellowship and JGW Patterson special grant for the funding support.

Finally, I would like to thank my wonderful family and friends for their unparalleled love and support. I wish to thank my dear wife Mrs. Sindhu Padmanaban who has supported the family during much of my PhD study. My wife Sindhu has been extremely supportive of me during this entire time and has made countless sacrifices which I am so grateful of. I remember my late father Mr. Sudhindar Dhondurao on this occasion who always encouraged me in every aspect of life. I dedicate this thesis to him. I am forever indebted to my parents for providing me the opportunities and experiences that has made me who I am. I would have never been able to submit this thesis without your love and support.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

WHO	World Health Organisation
BCLC	Barcelona-Clinic Liver Cancer
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
NAFLD	Non-Alcoholic Fatty Liver Disease
NASH	Non-Alcoholic Steatohepatitis
LINE1 or L1	Long Interspersed Nuclear Elements
EMT	Epithelial Mesenchymal Transition
UPR	Unfolded Protein Response
DAA's	Direct Acting Antivirals
ETV	Entecavir
TDF	Tenofovir Disoproxil Fumarate
TAF	Tenofovir Alafenamide
DNMT	DNA methyltransferases
HDAC	Histone Deacetylases
ADV	Adefovir dipivoxil
LdT	Telbivudin
ALD	Alcoholic Liver Disease
ROS	Reactive Oxygen Species
GSEA	Gene Set Enrichment Analysis
AFP	Alpha Fetoprotein
4mC	N4-methylcytosine
6mA	N6-methyladenine
5hmC	5-hydroxymethylcytosine
5hmU	5-hydroxymethyluracil
TET	Ten Eleven Translocation
5fC	5-formyl cytosine
5caC	5-carboxy cytosine
ctDNA	Circulating tumour DNA

HMT	Histone lysine methyltransferases
HAT	Histone acetyltransferases
FDA	Food and Drug Administration
PCR2	Polycomb Repressor Complex 2
ncRNAs	non-coding RNAs
miRNAs	micro RNAs
siRNAs	Small interfering RNAs
piRNAs	PIWI interacting RNAs
lncRNAs	long non-coding RNAs
snRNAs	small nuclear RNAs
snoRNAs	small nucleolar RNAs
LTR	Long Terminal Repeats
5'-UTR	5'-untranslated region
TPRT	Target Primed Reverse Transcription
L1-MET	L1-inserted c-MET
BaP	Benzo(a)Pyrene
YY1	Ying Yang1
CLD	Chronic Liver Disease
CHC	Chronic Hepatitis C
Huh7-J17	Huh7 cells with HCV replicon
Huh7-J17+PSI7977	Post HCV clearance model
TCGA	The Cancer Genome Atlas
DSB	Double Strand Break
HRR	Homologous Recombination Repair
NHEJ	Non-Homologous End Joining
BER	Basal Excision Repair
DDR	DNA Damage Response
HHL-5-RTN	Retrotransposed lines generated using L1 GFP reporter construct
HHL-5-RTNBlast^{res}	Retrotransposed clones generated by blasticidin resistance

enChIP	Engineered DNA-binding molecule-mediated ChIP
ChIP	Chromatin Immunoprecipitation
RACE	Rapid Amplification of cDNA ends
sgRNA	Single guide RNA
PB-TET	Piggy-bac vector with tetracycline inducible promoter
PB-TET-dCas9	Piggy-bac vector with dCas9 insert
Huh7 PB-TET-dCas9	Huh7 stable cell line expressing Dox-inducible Flag-dCas9
Huh7 PB-TET-dCas9-L1gRNA	Huh7 stable cell line expressing Dox-inducible Flag-dCas9-L1gRNA
HHL-5 PB-TET-dCas9	HHL-5 stable cell line expressing Dox-inducible Flag-dCas9
HHL-5 PB-TET-dCas9-L1gRNA	HHL-5 stable cell line expressing Dox-inducible Flag-dCas9-L1gRNA

Chapter 1: Introduction

1.1 Hepatocellular Carcinoma

Cancer is a name given to a collection of diseases involving abnormal cell growth. It is a genetic disease caused by changes in 3 main types of driver genes namely the proto-oncogenes, tumour suppressor genes and DNA repair genes (*Cancer definition*). According to estimates from WHO Cancer is the second leading cause of deaths globally and was responsible for an estimated 9.6 million deaths in 2018 (Bray F. *et al.*, 2018). In the UK, accounts for approximately 28% of all deaths registered in 2016, with the highest mortality rate registered in the north east region of England (Patel V., 2017). It is estimated that over 50% of people in the UK above 65 years of age will be diagnosed with cancer at some point in their lives (Ahmad A.S. *et al.*, 2015). Liver cancer accounts for the third leading cause of cancer deaths globally (Bray F. *et al.*, 2018). There are two subtypes of liver cancer- Hepatocellular Carcinoma (HCC) which arises in the hepatocytes and Cholangiocarcinoma arising in the bile ducts (Dhondurao-sudhindar.P., 2017; Bray F. *et al.*, 2018). HCC accounts for the majority of primary liver cancers worldwide (Bray F. *et al.*, 2018). HCC development mainly happens on a background of chronic liver disease, follows the accumulation of damage induced genetic and epigenetic alterations of growth regulatory genes, leading to increase in activation of oncogenes and loss of function of tumour suppressor genes (Kanda M. *et al.*, 2015). Advances in novel methodologies like next generation sequencing, genome wide methylation and proteomic studies have identified molecular classes of HCC (Llovet J.M. *et al.*, 2003) and several abnormalities associated with HCC in early stages proposed as early detection biomarkers (Lozada M.E. *et al.*, 2015). However, detection presently is typically by imaging and thus detects cancer mostly at an advanced stage when treatment options are limited (Llovet J.M. *et al.*, 2003). Consequently the 5 year survival rate of HCC patients is still poor (Buendia M., 2015) and the molecular mechanisms driving tumorigenesis that can be targeted therapeutically remain elusive – in part due to the heterogeneity of the tumour (Fig 1.1).

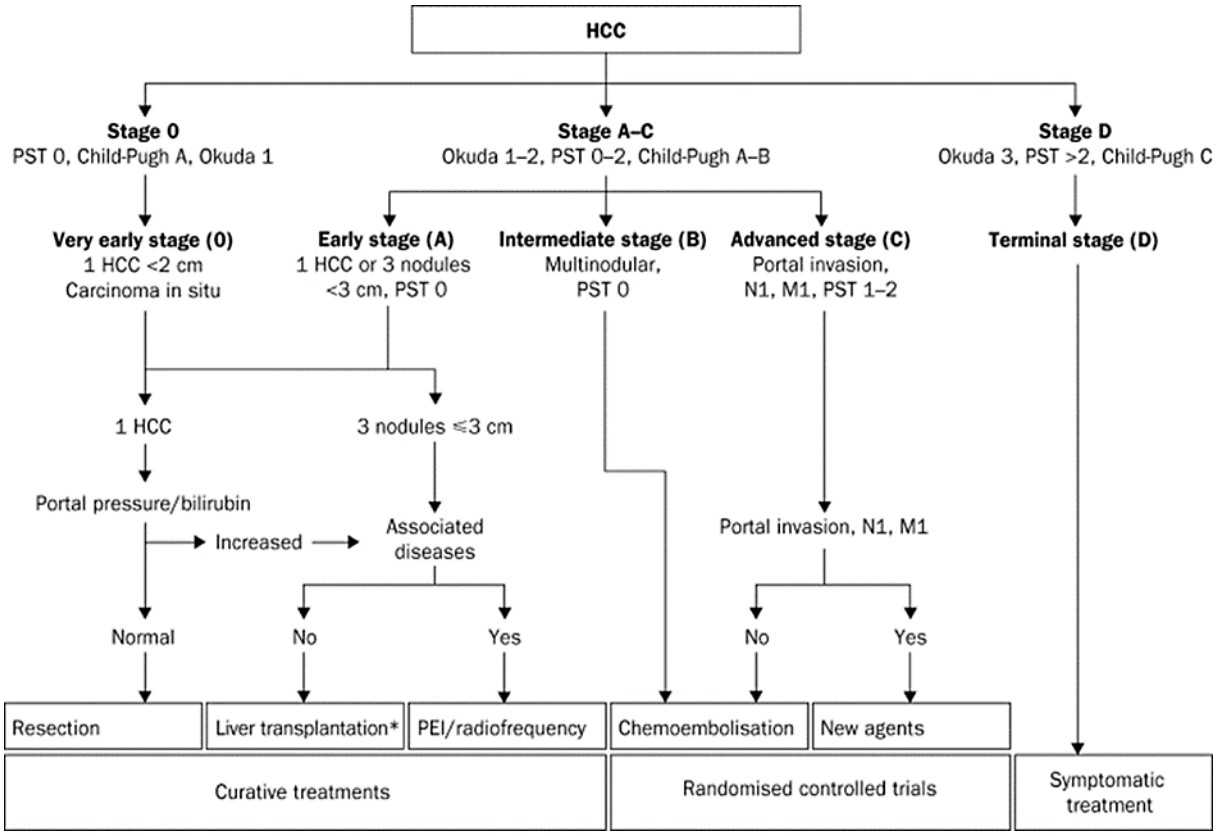

Figure 1.1 *Barcelona-Clinic Liver Cancer (BCLC) staging classification*

Figure shows BCLC staging classification and treatment options (Llovet J.M. et al., 2003)

Recent studies indicate that epigenetic aspects play a major role in initiation of HCC pathogenesis. These include dysregulation of Long Interspersed Nuclear Elements (LINE1 or L1), which can reduce the tumour suppressive capacity of somatic cells by L1 mediated retro-transposition. Studying epigenetic changes such as these may identify novel candidate biomarkers for early detection and possibly novel molecular therapeutic targets (Anestopoulos I. *et al.*, 2015; Kanda M. *et al.*, 2015; Dhondurao-sudhindar.P., 2017).

1.2 Risk factors for HCC

The major risk factors for HCC include carcinogens like aflatoxin, Hepatitis B virus (HBV), Hepatitis C virus (HCV) infections, chronic alcoholism (Perz J.F. *et al.*, 2006), as well as obesity, diabetes and the associated condition - non-alcoholic fatty liver disease (NAFLD) (El-Serag H.B., 2012; Dhondurao-sudhindar.P., 2017).

1.2.1 Hepatitis virus infection

There are 5 main hepatitis viruses, referred to as types A, B, C, D and E. In particular, types B and C lead to chronic disease in hundreds of millions of people. Chronic infection with Hepatitis B and Hepatitis C viral infections are the most common risk factors for HCC worldwide. HBV infection is mostly acquired by birth or in early childhood whereas HCV infection can occur at any age acquired mainly through contaminated blood and needles (de Martel C. *et al.*, 2015). According to the GLOBOCAN database (2012) 770,000 cases of liver cancer occurred worldwide in 2012, out of which 56% were attributable to HBV and 20% to HCV infections. The prevalence of HBV infection is high in East Asian and Sub Saharan African countries compared to Europe (Maucort-Boulch D. *et al.*, 2018). Maucort-Boulch et al developed a statistical model to study the distribution of HBV and HCV infections worldwide and identified that two out of three cases of liver cancer arises on the background of HBV infection in less developed nations compared to one in four in more developed countries whereas one in two cases of liver cancer in the developed world is attributable to HCV infections (Maucort-Boulch D. *et al.*, 2018).

1.2.1a HCV

HCV is a single stranded RNA virus in the *Flaviviridae* family, approximately 9600 nucleotides in length (Choo Q.L. *et al.*, 1989). It causes acute and chronic hepatitis in humans, if left untreated can progress to cirrhosis and hepatocellular carcinoma (Alter H.J. and Seeff L.B., 2000). The ability of HCV to cause chronic infection in most patients is partly due to its ability to evade host innate immune responses and its ability to regulate critical signalling pathways in hepatocytes (Chan S.T. and Ou J.J., 2017). The HCV genome displays remarkable genetic diversity due to a highly error prone RNA polymerase (De Francesco R. and Migliaccio G., 2005). The HCV genome, core protein, structural envelope glycoproteins E1 and E2, and non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B) are the known viral components of the virion. Collectively these proteins contribute to various aspects of HCV life cycle including the Virus entry, Fusion and uncoating, Translation, RNA replication, Virion assembly and release (Ploss A. and Dubuisson J., 2012) (Fig 1.2).

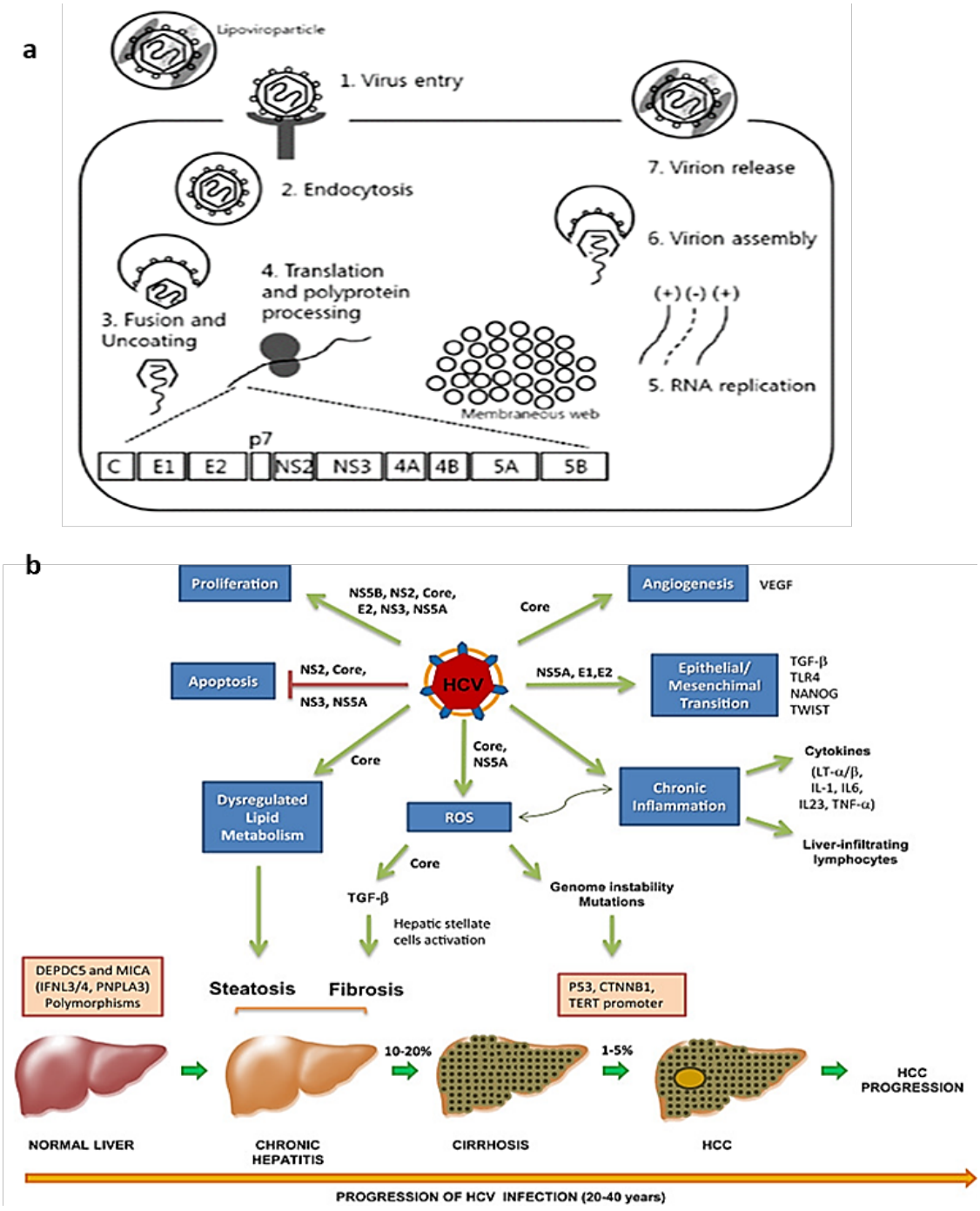

Figure 1.2 *HCV associated risk of Hepatocarcinogenesis*

Figure shows the life cycle of HCV (Kim C.W. and Chang K., 2013) (a) and the molecular mechanisms of HCV induced Hepatocarcinogenesis (Vescovo et al., 2016) (b).

HCV particles bind to the host cells via a specific interaction between HCV envelope glycoprotein and a yet unknown host cellular factor. Bound particles are then internalised by receptor mediated "**Endocytosis" (Fig 1.2a).** After the viral genome is liberated from the

nucleocapsid ("**Uncoating**") and translated at the rough endoplasmic reticulum, NS4B in conjunction with other viral or cellular factors induces the formation of "**Membranous web**" which serves as a scaffold for viral replication complex. After the genome amplification and HCV protein expression progeny virions are "**Assembled**". Newly produced viral particles then leave the host cells by constitutive secretive pathway ("**Virion release**") (Chevaliez S. and Pawlotsky J.M., 2006) (Fig 1.2a).

Chronic HCV infection can deregulate several host signalling pathways which has severe implications in cancer development (Bandiera S. *et al.*, 2016; Virzì A. *et al.*, 2018). For example, Epidermal Growth Factor Receptor (EGFR) is a key entry point for HCV into the hepatocytes and hence HCV induces the activation of EGFR signalling pathway which increases the risk of liver disease (Diao J. *et al.*, 2012; Igloi Z. *et al.*, 2015). Similarly HCV has been shown to downregulate tumour suppressors like pRb (McGivern D.R. *et al.*, 2009)and p53 (Sato Y. and Tsurumi T., 2013) thereby increasing cell survival (Fig 1.2b). NS5B protein interacts with pRb and retains it in the cytoplasm of the hepatocytes leading to its proteosomal degradation (Munakata T. *et al.*, 2005). NS5A protein binds with p53 in the cytoplasm thereby reducing its nuclear presence which inhibits apoptosis (Lan K.H. *et al.*, 2002). HCV also targets the TGFβ signalling via interaction of the HCV core protein with SMAD3 (Cheng P.L. *et al.*, 2004) thereby promoting epithelial mesenchymal transition (EMT) (Fig 1.2b), contributing to metastatic tumour development (Thiery J.P. and Sleeman J.P., 2006). It has been shown in-vitro that transient transfection of HCV core protein in Huh7 cells promoted cell proliferation and cell cycle progression mainly via upregulation of Wnt signalling, which play a crucial role in HCV induced hepatocarcinogenesis (Fukutomi T. *et al.*, 2005). HCV infection also enhances the notch and hedgehog signalling pathways which are involved in several morphological key functions like cell proliferation, differentiation, migration and survival (Pereira Tde A. *et al.*, 2010; Iwai A. *et al.*, 2011; Virzì A. *et al.*, 2018). HCV core protein mediates the development of hepatic angiogenesis by triggering the production of TGFβ2 and VEGF proteins by multiple pathways (Hassan M. *et al.*, 2009) (Fig 1.2b). The HCV viral genome replicates in the hepatocytes leading to the accumulation of viral particles in the Endoplasmic Reticulum (ER) where the virus induces stress (Dash S. *et al.*, 2016). The host cell responds by increasing the expression of Unfolded Protein Response (UPR) genes which relieves the stress response by downregulating the overall protein

synthesis (Tardif K. *et al.*, 2005; Wang M. and Kaufman R.J., 2014). In addition the HCV infection induces autophagy to enhance its replication and prolong its survival by mitigating the host apoptotic response (Dreux M. *et al.*, 2009; Dash S. *et al.*, 2016).

Various steps in HCV life cycle could serve as targets for novel therapeutics like Direct Acting Antivirals (DAAs) or Host targeting antivirals. HCV entry can be targeted by monoclonal or polyclonal neutralising antibodies (ITX5061, Ezetimibe, and Erlotinib) although their efficacy in-vivo turned out to be much lower (Schiano T.D. *et al.*, 2006; Meuleman P. *et al.*, 2011). Small molecule inhibitors targeting HCV envelope glycoproteins have been developed (EI-1, lectins, EGCG (Baldick C.J. *et al.*, 2010; Calland N. *et al.*, 2012)) but they were found to be mostly genotype specific (Matsumura T. *et al.*, 2009). HCV non-structural proteins (NS3, NS4A) can be targeted using second generation protease inhibitors- telaprevir and boceprevir both of which have been approved for treatment of HCV infection (McHutchison J.G. *et al.*, 2009; Jacobson I.M. *et al.*, 2011; Poordad F. *et al.*, 2011). Clemizole was identified as a potential inhibitor of HCV RNA replication by blocking the binding of RNA to NS4B (Einav S. *et al.*, 2008). Similarly HCV NS5A inhibitor, BMS-790052 was found to be effective in controlling the viral load both in-vitro and in-vivo (Gao M. *et al.*, 2010). These agents in combination with standard treatment of care has led to therapeutic regimen with better tolerability and improved clinical outcomes (Gao M. *et al.*, 2010).

1.2.1b HBV

HBV is a partially double stranded DNA virus, approximately 3.2kb in length in the *Hepadnaviridae* family (Li H. *et al.*, 2020). According to WHO global hepatitis report 2017, an estimated 257 million people were chronically infected with HBV and 887,000 people die each year from HBV related liver disease, approximately half of them in China ('WHO (2017) Global Hepatitis Report,' 2017). HBV infection can cause acute to chronic hepatitis and in the absence of any anti-viral treatment can progress to liver cirrhosis and HCC (Fattovich G. *et al.*, 2008). HBV particles are composed of 4 overlapping ORFs (C, P, S and X). HBc and HBe proteins are produced from ORF C, HBV DNA polymerase from ORF P, the surface proteins- small, medium and large HBs are produced from ORF S and HBx is produced from ORF X (Tsukuda S. and Watashi Koichi., 2020). Together these proteins contribute to various aspects of HBV life cycle (Fig 1.3).

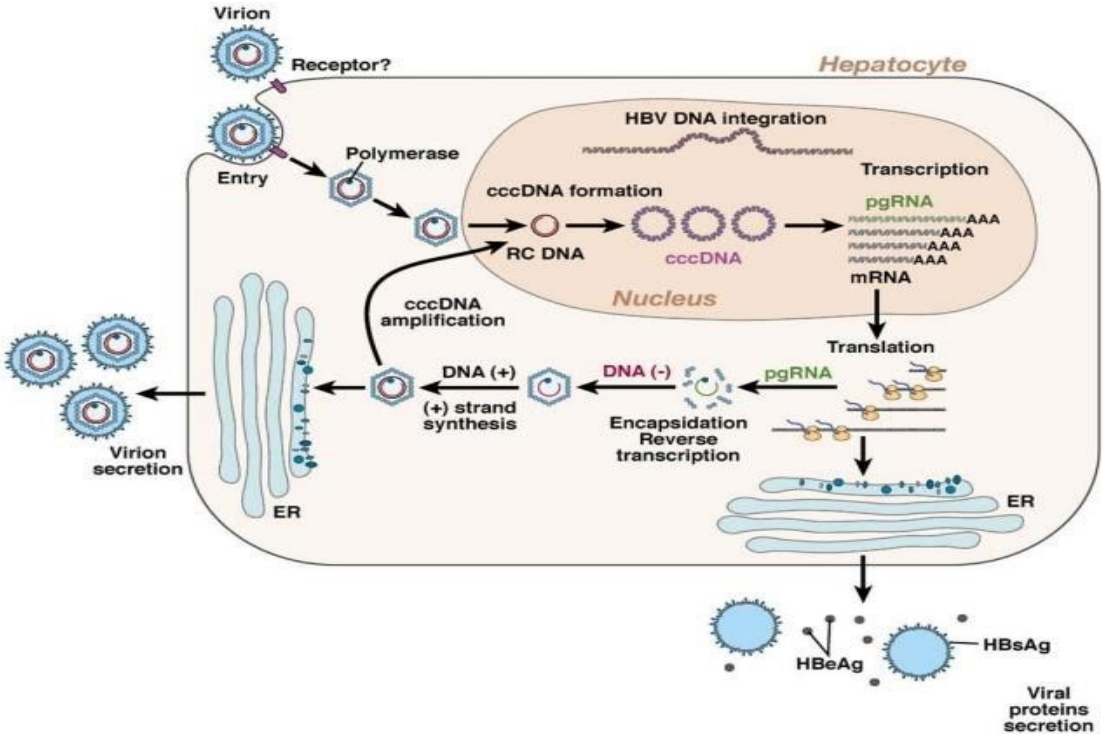

Figure 1.3 *HBV life cycle*

Figure 1.3 shows life cycle of HBV (Zoulim and Locarnini, 2009).

The life cycle of HBV involves the viral entry into host cells; rcDNA's entry into the nucleus to form cccDNA; transcription and translation of viral RNAs and proteins; viral nucleocapsid formation and assembly; reverse transcription and rcDNA synthesis; and, finally, viral packaging, maturation, and budding (Li H. et al., 2020).

Chronic infection with HBV triggers several oncogenic pathways like Jak-Stat, Ras/MAPK, PI-3K/Akt and NF-κB (Bouchard M.J. *et al.*, 2006). Prolonged expression of HBx and Long HBs surface proteins has also been shown to trigger epigenetic modifications of the host genes (Jia L. *et al.*, 2020). For example, HBx protein is shown to interact with DNA methyltransferases (DNMT3a) and Histone Deacetylases (HDAC1). HBx modulates the transcriptional activation of DNMT3a to promote regional hypermethylation of tumour suppressor genes like p16 (Zheng D.L. *et al.*, 2009). HBV integration frequently occurs in genes upregulated in tumours like TERT, MLL4, Cyclin E1, SENP5 and ROCK1 (Zhao L.H. *et al.*, 2016). HBV insertions into L1 retrotransposons are known to generate HBx-L1 chimeric transcripts which promotes β-catenin signalling activation, E-cadherin reduction and cell migration which

increases the liver injury and risk of HCC in mice (Liang H.W. *et al.*, 2016). HBx also increases the genomic instability by its interaction with p53 (Shahnazari P. *et al.*, 2014) and DDB1 (Hodgson A.J. *et al.*, 2012) involved in DNA repair.

Vaccine against HBV was successfully developed by Merck in 1984, who used yeast expressed HBsAg to target HBV (McAleer W.J. *et al.*, 1984). Since then, mass vaccination drive was carried out by the Chinese government which led to marked decline of HBV infection in china (Li H. *et al.*, 2020). IFN-α therapy (Wong D.K. *et al.*, 1993) and nucleoside analogue anti-viral drugs like Lamuvidine (Lai C.L. *et al.*, 2003), Adefovir dipivoxil (ADV) (Trepo C., 2014) and Telbivudin (LdT) (Marcellin P. *et al.*, 2008) were used as first line anti-viral therapies against HBV infection until late 1990's, but due to low efficacy and high rate of drug resistance these drugs are not used as first choice now. Conventional IFN-α therapy has been replaced by Peg-IFN due to its long-acting effect (Piratvisuth T. *et al.*, 2008). Low resistant nucleoside analogues like Entecavir (ETV), Tenofovir disoproxil fumarate (TDF) and Tenofovir alafenamide (TAF) strongly inhibits HBV replication and found to have low drug resistance, consequently used as first line therapy in chronic hepatitis B patients (Yokosuka O. *et al.*, 2010; Kitrinos K.M. *et al.*, 2014). Several stages in HBV lifecycle can also be targeted by Direct Acting Antivirals (DAA's) and indirect anti-viral drugs that modulate the host immune response. For example DAA's like Myrcludex-B, targeting the HBV entry and ARC-520 (RNAi based agent) are currently in phase II clinical trials (Blank A. *et al.*, 2016; Wooddell C.I. *et al.*, 2017). Nucleocapsid assembly inhibitors like NVR 3-778, JNJ-6379, GLS4 and ABI-H0731 are in different stages of clinical trials (Zhang H., 2018; Lam A.M. *et al.*, 2019; Ma X.L. *et al.*, 2019; Vandenbossche J. *et al.*, 2019). Indirect anti-viral drugs that reactivate the host immune response in HBV infected individuals are also under development. For example Toll like receptor agonists (TLR-7 and TLR-8) are under clinical trials (Lanford R.E. *et al.*, 2013; Grant E., 2018). Similarly anti-PD-1 antibody (Nivolumab) could be employed to block the PD-1 pathway which plays an important role in T cell exhaustion during HBV infection (Wenjin Z. *et al.*, 2012).

1.2.2 Chronic alcoholism

Alcoholism is one of the leading causes of liver disease globally with WHO estimating approximately 2.5 million deaths worldwide each year related to consumption of alcohol

(*World Health Organisation. Global status report on alcohol and health 2018.*, 2018). Alcohol consumption also increases the risk of developing HCC significantly (*World Health Organisation. Global status report on alcohol and health 2018.*, 2018). A daily consumption of 60-80 g/d of alcohol for 10 years or longer in men and 20 g/d in women may lead to development of chronic liver disease in approximately 40% of the cases (Mandayam S. *et al.*, 2004; Rocco A. *et al.*, 2014). Alcohol also synergistically interacts with other causative agents like hepatitis viral infection and risk factors like diabetes, obesity and NAFLD to promote the progression of Alcoholic Liver Disease (ALD) (Day C.P., 2000). Ethanol and its metabolites like reactive oxygen species (ROS), nitric oxide and hydroxyethyl radical can exert a direct cytotoxic effect on hepatocytes (Page A. *et al.*, 2015). These metabolites can induce hepatic inflammation via cytokines like TNFα which acts as indirect causative agents for alcohol induced hepatocellular damage (McClain C.J. *et al.*, 2004; Page A. *et al.*, 2015). Although liver is considered as the main organ affected by alcohol abuse, recent studies indicates that it is a systemic disease. Hence there is urgent need to promote preventive policy strategies to reduce the clinical and economic burden of chronic alcohol abuse (Rocco A. *et al.*, 2014).

1.2.3 Non-Alcoholic Fatty Liver Disease

Non-Alcoholic Fatty Liver Disease (NAFLD) represents a spectrum of liver diseases including simple steatosis (fat infiltration) progressing to inflammation (Non-Alcoholic Steatohepatitis, NASH) and fibrosis and ultimately leading to cirrhosis and hepatocellular carcinoma (Anstee Q.M. *et al.*, 2011). It is the most common cause of liver dysfunction in most of the developed and developing countries and is increasing owing to its close association with diabetes and obesity (Hardy T. *et al.*, 2016). NAFLD is projected to be the primary cause for liver transplantation in the next decade (Baumeister S.E. *et al.*, 2008) as the incidence of chronic viral hepatitis is decreasing and incidence of NAFLD is increasing (Holmberg S.D. *et al.*, 2013). To distinguish simple steatosis from NASH, ultrasonography is largely used in current clinical practise although it is sensitive only if >33% of the liver is steatotic (Dasarathy S. *et al.*, 2009; Schwenzer N.F. *et al.*, 2009). The incidence of lobular inflammation and NASH increases with age as the anti-inflammatory responses fail. The pathogenesis of NAFLD is incredibly complex as it involves interplay between several environmental and genetic factors. Hence, there are no reliable biomarkers and no licensed therapy exist as of now (Hardy T. *et al.*, 2016).

1.3 Molecular mechanisms of HCC development

1.3.1 Key signalling pathway alterations and common mutations in HCC

Hepatocellular Carcinoma is highly resistant to conventional chemotherapy and radiotherapy due to the fact that the tumour is heterogeneous in nature arising from the chromosomal instability (Rao C.V. *et al.*, 2017). Recent cancer genome next generation sequencing studies have revealed several signalling pathways that contribute to HCC development like oncogenic pathways (WNT/β-catenin, TGFβ and EGFR), DNA damage checkpoint repair pathways and oxidative stress response pathways (Rao C.V. *et al.*, 2017). Commonly mutated genes involved in the pathways leading to HCC development are listed in Table 1.1. For example, mutation in the TERT promoter leading to overactive telomerase is the most common gain of function mutation identified in early development of HCC contributing to approximately 60% of HCC cases (Totoki Y. *et al.*, 2014; Zucman-Rossi J. and et al., 2015). Early HCC harbouring TERT promoter mutations are at higher risk of complete malignant transformation into more advanced case of HCC (Zucman-Rossi J. and et al., 2015). TP53 is the most common tumour suppressor gene mutated in almost 30% of the HCC patients (Fujimoto A. *et al.*, 2015; Schulze K. *et al.*, 2015). Mutations in TP53 gene leads to alterations in P53 cell cycle pathway and cell cycle progression (Totoki Y. *et al.*, 2014; Schulze K. *et al.*, 2015). The oncogenic WNT/β-catenin pathway is frequently altered in HCC due to activating mutations in c-terminus of the CTNNB1/β-catenin gene (20-40%) and loss of function mutations of AXIN1 (9-13%) (Satoh S. *et al.*, 200; La Coste de A. *et al.*, 1998). Although CTNNB1 and AXIN1 genes are involved in the same WNT/β-catenin pathway they alter the WNT signalling in different ways as CTNNB1 is a positive regulator and AXIN1 is the negative regulator of WNT signalling (Rao C.V. *et al.*, 2017). These alterations in WNT signalling pathways lead to genomic instability and contributes to HCC development (Rao C.V. *et al.*, 2017). Apart from these frequently mutated genes, there are other genes with a mutation rate of less than 5% that have been described like WWP1 (Zhang X.F. and et al., 2015), ATM (Daugherity E.K. and et al., 2012), p16INK4A (Jenkins N.C. and et al., 2011), NRF2 (Sporn M.B. and Liby K.T., 2012), ARID2 (Li M. *et al.*, 2011) and FGF19 (Sawey E.T. *et al.*, 2011). Likewise, somatic mutations in individual genes involved in the TGFβ pathway remain less than 5% but when taken together, 38% of HCC cases among a cohort of 202 HCC samples from TCGA HCC dataset had mutations in one or other genes involved in the pathway (Chen J. *et al.*, 2018). Hence, alterations in genes involved in TGFβ superfamily is

common in HCC. Mutations in genes involved in the TGFβ pathway also correlated with decreased survival in HCC patients (Chen J. *et al.*, 2018). Although several studies in the last decade have been able to delineate the mutations and genomic alterations in HCC development and progression, it has not been translated into clinical practice yet for actual advantage of patients (Zucman-Rossi J. and et al., 2015).

Commonly mutated genes involved in HCC development	Percentage of samples with one or more mutations
TP53	30.1%
CTNNB1	26.0%
ALB	12.8%
PCLO	10.7%
LRP1B	9.0%
ARIDIA	7.9%
AXIN1	6.8%
PRKDC	6.6%
KMT2D	5.5%
BAP1	5.5%
RB1	5.2%
ARID2	5.2%

Table 1.1 *List of commonly mutated genes involved in HCC development*

Table shows the list of commonly mutated genes and the percentage of mutation in HCC development. Hepatocellular Carcinoma TCGA PanCancer data obtained from 366 profiled samples (Cerami E. et al., 2012)

1.3.2 HCC classification based on transcriptomic changes

Hoshida et al performed integrative transcriptomic meta-analysis in samples from 8 independent HCC patient cohorts (total of 603 patients) from different parts of the world and were able to classify them into three subclasses- S1, S2 and S3 based on tumour size, extent of cellular differentiation and alpha-fetoprotein levels in the serum (Hoshida Y. *et al.*, 2009). The Gene Set Enrichment Analysis (GSEA) of the datasets revealed that the three subclasses were associated with distinct biological processes- S1 subclass was associated with WNT signalling pathway activation, S2 subclass tumours were characterised by MYC and AKT

activation, S3 subclass tumours showed abundant expression and differential activation of p53 and p21 target genes (Hoshida Y. *et al.*, 2009). Therapeutic agents selectively targeting some of the features associated with molecular subclasses are in development like GC33 a humanised monoclonal antibody against GPC3, a marker of S2 subclass showed promising results in phase 1 clinical trials (Zhu A.X. *et al.*, 2013). LY2157299, a small molecule inhibitor of TGFβ (S1 subclass) is now in phase II clinical trial development (Giannelli G. *et al.*, 2014), Galunicertib which also targets the TGFβ pathway is in phase II trial (Giannelli G. *et al.*, 2016). Dasatinib, an Src/Ab1 small molecule kinase inhibitor (S1 subclass) was found to be more effective in-vitro(Finn R.S. *et al.*, 2013). Tivantinib (MET inhibitor) (Santoro A. *et al.*, 2013) and Ramucirumab (VEGFR2 inhibitor) (Zhu A.X. *et al.*, 2015) showed better response in AFP positive aggressive tumours (S2 subclass). Development of these molecular therapeutic agents selectively targeting some of the features of molecular subclasses of HCC, has led to an hypothesis that they can have clinical utility as a predictive indicator of drug response (Goossens N. *et al.*, 2015; Hirschfield H. *et al.*, 2018).

1.4 Key Epigenetic changes in HCC

Epigenetic mechanisms are modifications that occur in the genetic material that do not change the nucleotide sequence, but instead may cause conformational modifications in DNA, which determines how the genome is interpreted by the cell to generate a phenotype (Korkmaz A. *et al.*, 2011). There are basically three epigenetic modifications to be considered - DNA methylation, histone modification and regulation by non-coding RNA. These epigenetic mechanisms play a key role in development, human physiology and diseases including liver fibrosis and cancer (Sharma S. *et al.*, 2010; Wilson C.L. *et al.*, 2017).

1.4.1 DNA methylation

DNA methylation can be described as one of the most common DNA modifications where methyl group is attached to 5’ position of cytosine base (5mC), mainly in context of a CpG dinucleotide and is associated with suppression or silencing of gene expression (Tourancheau A. *et al.*, 2020). Other types of methylated bases are also reported like N4-methylcytosine (4mC), N6-methyladenine (6mA), 5-hydroxymethylcytosine (5hmC) and 5-hydroxymethyluracil (5hmU) (Tourancheau A. *et al.*, 2020). DNA methyltransferases (DNMT) and Ten Eleven Translocation (TET) enzymes are two of the most important components of

the DNA methylation machinery (Fig 1.4) (Robertson K.D., 2005). Disturbances in the balance between these components can contribute to disease progression(Robertson K.D., 2005; Scourzic L. *et al.*, 2015). The addition of methyl group is controlled by a family of enzymes called DNA methyltransferases (DNMTs) (Bestor T.H., 2000). DNMT1, DNMT3a and DNMT3b are the 3 most important DNMTs involved in maintenance of DNA methylation patterns (Bestor T.H., 2000). DNMT1's are classified as maintenance methyltransferases involved in the maintenance of the pre-existing DNA methylation patterns with strong preference for hemi methylated DNA's (Cui and Xu, 2018). Whereas DNMT3a and DNMT3b are involved in establishment of new DNA methylation patterns. They catalyse the methylation of new DNA elements in early embryonic development (Cui and Xu, 2018).

On the other hand, TET proteins promote removal of methyl group from DNA (Ito S. *et al.*, 2010). The TET enzymes (TET1, TET2 and TET3) oxidises 5mC to 5hmC and subsequently to 5-formyl cytosine (5fC) and 5-carboxy cytosine (5caC) (Jiang Y.Z. *et al.*, 2015). In normal cells, most CpG dinucleotides are often methylated and most unmethylated CpG's exist as CpG islands in the promoter region which are the often the transcription factor binding sites required for gene expression(Lim D.H.K. and Maher E.R., 2010). CpG islands are genomic regions with at least 200 bp, a GC percentage greater than 50%, and an observed-to-expected CpG ratio greater than 60% wherein the CpG sites (cytosine-guanine in linear sequence of bases along 5'-3' direction) occur with high frequency (Deaton A.M. and Bird A., 2011). Methylation of CpG islands can attract the methyl CpG binding proteins (MBDs) which can cause modification of histones and chromatin condensation hence resulting in gene silencing (Lim D.H.K. and Maher E.R., 2010). Cancer cells are characterized by global hypomethylation and focal hypermethylation of the tumour suppressor genes (Lim D.H.K. and Maher E.R., 2010; Jueliger S. *et al.*, 2016).

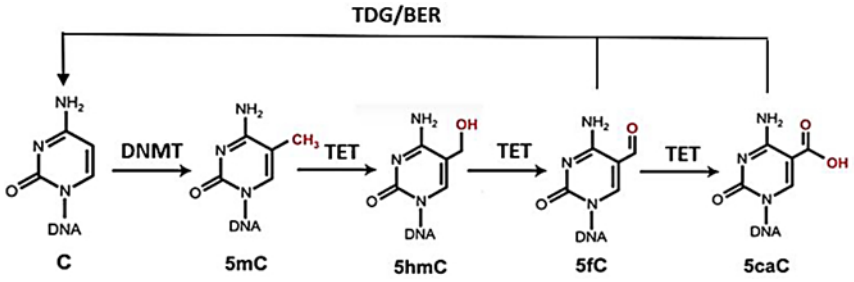

Figure 1.4 *DNA methylation and demethylation*

Figure shows the DNA methylation and demethylation process mediated by DNMTs and TET enzymes (Lan Y. and Evans T., 2019). Enzymes involved in Thymine DNA glycosylase/Base Excision Repair pathway can excise and remove the oxidised 5mC bases thereby completing the demethylation process (Lan Y. and Evans T., 2019).

Song et al performed genome wide methylation study in a total of 27 HCC samples and 20 adjacent normal liver tissues and identified differential methylation pattern in 13% of the CpG loci (Song M.A. *et al.*, 2013). In another such study, genome wide methylation analysis of 71 human HCC samples identified SMPD3 and NEFH as tumour suppressor genes in HCC silenced by DNA methylation (Revill K. *et al.*, 2013). Villenueva et al analysed methylation based prognostic signatures from HCC tumour tissue from 304 patients with surgical resection and validated 36 DNA methylation markers which accurately predicted poor survival in HCC patients (Villanueva A. *et al.*, 2015). Further integrative analysis of the differential methylation status of 646 tumour and 134 non-tumour samples of HCC patients revealed 222 candidate genes whose levels of expression were negatively regulated by promoter methylation(Zheng Y. *et al.*, 2016). These studies suggest a role for changes in DNA methylation in patients with HCC, some of which may be good candidates for biomarkers (Jueliger S. *et al.*, 2016; Zheng Y. *et al.*, 2016). Although several studies analysed DNA methylation changes in HCC, very little was known about the DNA methylation changes that govern the transition from cirrhosis to early stages of HCC. Until a recent study by Hernandez-Meza et al mapping DNA methylation changes across different stages of hepatocarcinogenesis identified novel epigenetic gatekeepers in HCC (Hernandez-Meza G. *et al.*, 2020). Some of these promoter methylations have been explored as HCC detection biomarkers in circulating tumour DNA (ctDNA) such as DBX2, THY1, TGR5, MT1M, MT1G, INK4A, VIM, FBLN1, RGS10, ST8SIA6, RUNX, and SEPT9 (Zhang P. *et al.*, 2013; Wu X. *et al.*, 2020). Methylation profiles of the HCC tumour DNA was matched with that of the plasma ctDNA in a large clinical cohort study (1,098 HCC patients and 835 normal controls) conducted by Xu et al suggesting the utility of ctDNA in prognosis of HCC (Xu R.H. *et al.*, 2017).

Studies on DNA methylation signature have also been carried out to understand drug resistance, in one such study the DNA methylation signatures of 22 tumour suppressor genes were compared between HCC patient derived xenografts resistant to Sorafenib (a tyrosine kinase inhibitor approved for treatment in advanced HCC patients) and normal liver tissue from healthy donors (Jueliger S. *et al.*, 2016). Out of the 22 tumour suppressor genes

compared in the study, 10 genes (CDKN1A, CDKN2A, DLEC1, E2F1, GSTP1, OPCML, E2F1, RASSF1, RUNX3 and SOCS1) showed striking difference in methylation status wherein the promoter regions were hypermethylated in HCC patient derived xenografts compared to healthy liver tissues (Jueliger S. *et al.*, 2016). Hypomethylating agents like Azacytidine and Decitabine (cytosine analogues) have been approved by the FDA for the treatment of Myelodysplastic syndrome (Sato T. *et al.*, 2017). Guadecitabine is currently in phase III clinical trial for treatment of Acute Myeloid Leukemia. Other cytosine analogue hypomethylating agents like Zebularine, CP-4200 and small molecule inhibitors like RG108 and Nanaomycin A are in preclinical development (Sato T. *et al.*, 2017) which can be used as therapies.

Another key role of DNA methylation in a normal cell is repression of L1 elements (Kannan M. *et al.*, 2017). The L1s belongs to a class of transposable elements, which are the repetitive sequences, can migrate within the genome and play an important role in human evolution and pathogenesis of several diseases including cancer (Anwar S.L. *et al.*, 2017). Alterations in global methylation patterns in cancer has been shown to activate L1 elements in various cancer types including HCC (Rodriguez-Martin B. *et al.*, 2020).Effects of L1 reactivation and retrotransposition in the context of HCC has been discussed in this thesis (from section 1.5).

1.4.2 Histone modification and chromatin remodelling

Nucleosomes, the building blocks of chromatin, are composed of two copies each of histones (H2A, H2B, H3 and H4) and a short segment of DNA (~145-147bp) wrapped around them (Luger K. *et al.*, 2012). Histone methylation and acetylation are two important modifications essential for gene expression regulated by Histone lysine methyltransferases (HMT's) and Histone acetyltransferases (HAT's) (Morera L. *et al.*, 2016). Histone lysine deacetylases (HDAC's) are known as epigenetic erasers - responsible for removing the epigenetic marks on the histones (Falkenberg K.J. and Johnstone R.W., 2014). Lachenmayer et al showed aberrant expression of several HDAC's in HCC, including HDAC3 and HDAC5, which correlated significantly with DNA copy number gains (Lachenmayer A. *et al.*, 2012). They also showed knockdown of HDAC's leading to anti-tumour effects in preclinical mouse models of HCC (Lachenmayer A. *et al.*, 2012). In another study, HDAC8 was suggested to drive hepatocarcinogenesis in NAFLD patients, as knockdown of HDAC8 in a mouse NAFLD - HCC model resulted in restoration of normal metabolic profile of the liver (Tian Y. *et al.*, 2015).

H3K27 methyltransferase- EZH2 is also highly expressed in some HCC and is associated with malignant transformation (Sasaki M. *et al.*, 2008). EZH2, a functional enzymatic component of polycomb repressor complex 2, is responsible for methylation of H3K27 (Mann J. *et al.*, 2010). Methylation activity of EZH2 facilitates heterochromatin formation thereby leading to gene silencing (Viré E. *et al.*, 2006). Knockdown of the EZH2 gene expression in preclinical animal models, as well as in-vitro in cell lines, has shown HCC growth inhibition (Chen Y. *et al.*, 2007). In human studies, over expression of EZH2 closely correlated with poor prognosis in HCC patients, while inhibition of EZH2 with small molecule inhibitors blocked the aggressive nature of HCC cells (Gao S.B. *et al.*, 2014).

ARID1A/B or ARID2, a subunit of SWI/SNF chromatin remodelling complex is frequently mutated in HCC (Kelso T.W.R. *et al.*, 2017). An in vitro study by Duan et al showed that overexpression of ARID2 supressed cell proliferation and migration in HCC cell lines, whereas siRNA mediated knockdown of ARID2 increased proliferation and migration capacities in vitro (Duan Y. *et al.*, 2016). A latest study by Jiang et al demonstrated that ARID2 expression is significantly decreased in metastatic HCC tissues, showing positive correlation with survival of HCC patients (Jiang H. *et al.*, 2020). To summarize, HDAC's, EZH2 and ARID2 may serve as diagnostic of HCC, as well as potentially being targets for HCC therapy.

1.4.3 Role of non-coding RNA's

The vast majority of the human genome consists of non-coding RNAs (ncRNAs). ncRNAs do not encode proteins and include micro RNAs (miRNAs), small interfering RNAs (siRNAs), PIWI interacting RNAs (piRNAs), long non-coding RNAs (lncRNAs), small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) (Amicone L. *et al.*, 2015). They are involved in transcriptional regulation (Amicone L. *et al.*, 2015). Deregulation of these ncRNAs has been linked to disease progression and cancer (Amicone L. *et al.*, 2015). lncRNAs in particular have been reported to act as tumour suppressors or oncogenes in HCC (Prensner J.R. and Chinnaiyan A.M., 2011). A recent RNA-seq study by Esposti et al, in which 23 tumour and non-tumour liver tissues were assessed, revealed 57 differentially expressed lncRNAs in HCC compared to the adjacent non-tumour tissue (Esposti D.D. *et al.*, 2016). The team also identified oncogenic roles for lncRNAs like H-19 and CRNDE, reporting their involvement in cell cycle deregulation occurring during HCC development (Esposti D.D. *et al.*, 2016).

In recent years a number of studies have reported multiple alterations and roles of mi-RNAs in HCC tumour progression, including upregulation or downregulation in HCC, as well as changes in the circulation – raising the potential of circulating mi-RNAs as biomarkers in HCC (Mao B. and Wang G., 2015). Recently, miRNA-873 has been shown to enhance HCC tumour progression by regulating cell growth and metastasis through excessive phosphorylation of the PI3K/AKT/mTOR pathway (Han G. *et al.*, 2018). miRNA-493 has been shown to act as a tumour suppressor by regulating Zinc Finger Protein-X (ZFX) expression in HCC (Ding W. *et al.*, 2018). These studies raise the possibility of miRNAs as potential HCC therapeutic candidates.

1.5 Transposable Elements- L1 Retrotransposons

Transposable Elements (TE), also called 'jumping genes', are mobile genetic elements that can move from one part of the genome to another (Cordaux R. and Batzer M.A., 2009). They constitute nearly 50% of the human genome (Lander E.S. and et al., 2001). They are classified into two types - DNA Transposons and Retrotransposons (Cordaux R. and Batzer M.A., 2009). The DNA transposons migrate as DNA from one genomic location to another and have no defined active function in human genome (Pace J. K. and Feschotte C., 2007; Cordaux R. and Batzer M.A., 2009). The Retrotransposons migrate as RNA intermediates that are reverse transcribed to DNA before integrating into the genome (Cordaux R. and Batzer M.A., 2009). The retrotransposons are further subdivided into two types based on the presence or absence of Long Terminal Repeats (LTR) as LTR Retrotransposons and Non-LTR Retrotransposons (Fig 1.5) (Cordaux R. and Batzer M.A., 2009). Long Interspersed Nuclear Elements (LINE-1 or L1), Alu and SVA elements form the group of Non-LTR retrotransposons, which collectively constitutes one third of the genome (Fig 1.5) (Lander E.S. and et al., 2001).

L1s are the only autonomous transposable elements capable of retrotransposition and integration into new locations in the genome. These can cause insertional mutations and the creation of novel transcription factor binding sequences resulting in the binding and activation of proximal promoters has been reported (Erwin J. *et al.*, 2014; Dhondurao-sudhindar.P., 2017). Insertion of the L1 elements near oncogenes or tumour suppressor genes can contribute to tumorigenesis, as was first reported by Miki et al, who described the disruption of the APC gene (which encodes a tumour suppressor) by a somatic insertion of L1 in colon cancer (Morse B. *et al.*, 1988; Miki Y. *et al.*, 1992). However, not all L1 copies are

capable of retrotransposition as the majority are truncated or have point mutations or rearrangements making them non-functional. Only about 50-100 full length retrotransposition competent copies are present in a human genome, which in turn are silenced epigenetically by DNA methylation, Small Interfering RNA (siRNA) and piwi-interacting RNA mediated mechanisms (Pezic D. *et al.*, 2014). Recent studies indicate the role of L1 elements in HCC development, however the mechanism behind which is still unclear (Shukla R. *et al.*, 2013a; Dhondurao-sudhindar.P., 2017). I will refer to these in detail in the relevant sections described below.

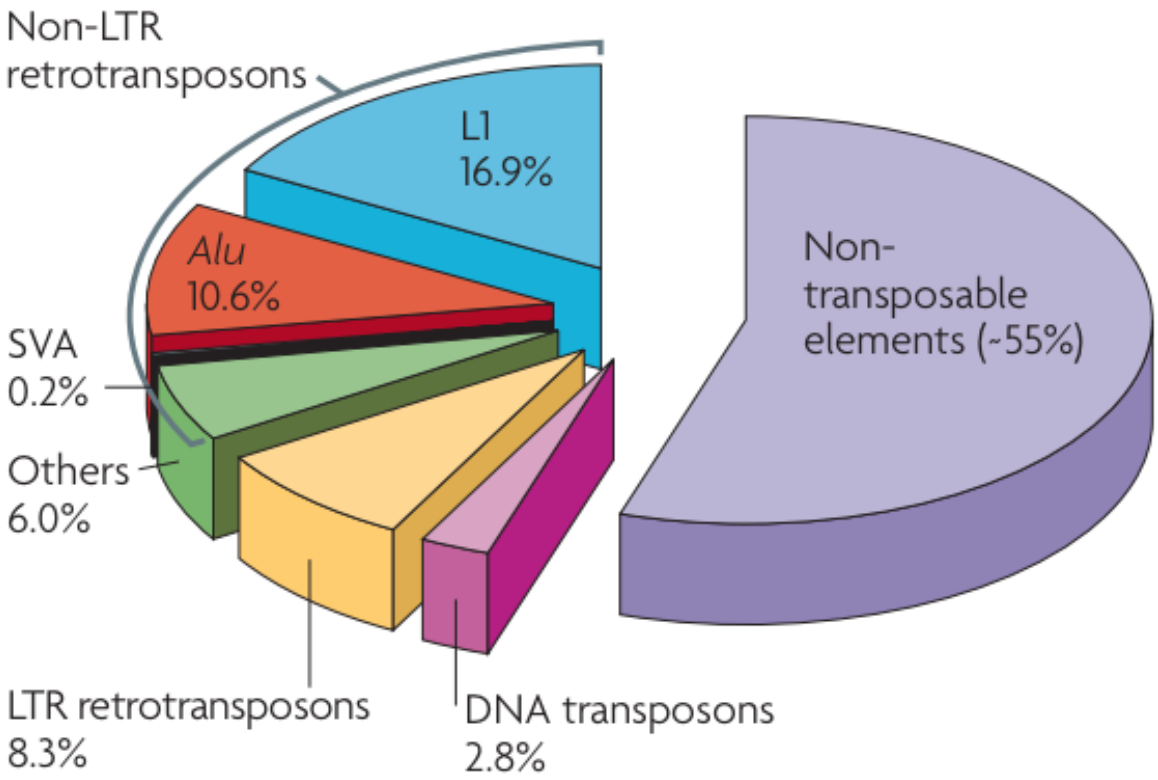

Figure 1.5 *Pie chart visualises the Transposable Element Content of Human Genome*

Figure adapted from (Cordaux R. and Batzer M.A., 2009)

1.5.1 L1 Structure and the process of Retrotransposition

A full length L1 element is 6kb in length, comprises of a 5'-untranslated region (UTR) (900bp in length) (Rahbari R. *et al.*, 2015)containing its own sense and anti-sense promotors (Speek M., 2001), three open reading frames- ORF0, ORF1 and ORF2 encoding proteins ORF1P(40KDa) and ORF2P(150KDa) respectively and a 3' UTR ending with a poly(a) tail (Xiao-Jie L. *et al.*, 2016). ORF0 is a 71 amino acid peptide located in the antisense promoter of L1 5'UTR. A recent study by Denli et al has found that ORF0 enhances L1 mobility, although other functions of ORF0 are largely unknown (Denli *et al.*, 2015). ORF1P contains a RNA recognition motif and ORF2P has 'Endonuclease' and 'Reverse Transcriptase' activities (Fig 1.6a) (Xiao-Jie L. *et al.*, 2016). There are about half a million copies of these L1 elements present in human genome but only about 80-100 are competent for retro transposition (active L1s) and the remaining 99% are inactive due to 5' truncations, internal rearrangements or deletions (Rahbari R. *et al.*, 2015). Major polymorphisms of the L1 elements occur within the 5'-UTR (Rahbari R. *et al.*, 2015). First 155bp of the L1 sense promoter within the L1-5'UTR region is involved in L1 expression (Rahbari R. *et al.*, 2015). Transcription of the L1 elements occurs in the nucleus and the L1 mRNA is transported into the cytoplasm where it is either suppressed by the epigenetic mechanisms (described in section 1.6.1) or translated into ORF1P and ORF2P (Xiao-Jie L. *et al.*, 2016). Both ORF1P and ORF2P then preferentially bind to their encoding RNA and the resultant L1 ribonucleoprotein particles migrate back into the nucleus (Wei W. *et al.*, 2001; Xiao-Jie L. *et al.*, 2016), where the L1 mRNA is reverse transcribed into cDNA mainly by a process termed as 'Target Primed Reverse Transcription' (TPRT) (Luan D.D. *et al.*, 1993; Dhondurao-sudhindar.P., 2017). During this process, the ORF2P endonuclease creates a nick in the DNA which acts as a primer for synthesis of L1 cDNA (Fig 1.6b), thus integrating into the genome and responsible for 'Insertional Mutagenesis' (Mandal P.K. *et al.*, 2013; Ariumi Y., 2016; Dhondurao-sudhindar.P., 2017). Morrish et al reported an alternate 'Endonuclease independent retrotransposition' in p53 deficient CHO cell line, which also carried mutations in genes required for NHEJ DNA damage repair pathway and dysfunctional telomeres. L1 elements rely upon pre-existing DNA lesions like sites of genomic DNA damage (single or double strand break) to initiate Target Prime Reverse Transcription in absence of an endonuclease (Morrish T.A. *et al.*, 2002). The L1 elements, which utilize this alternate

pathway for retrotransposition occurs at a relatively low frequency in the human genome (Morrish T.A. *et al.*, 2002; Morrish T.A. *et al.*, 2007).

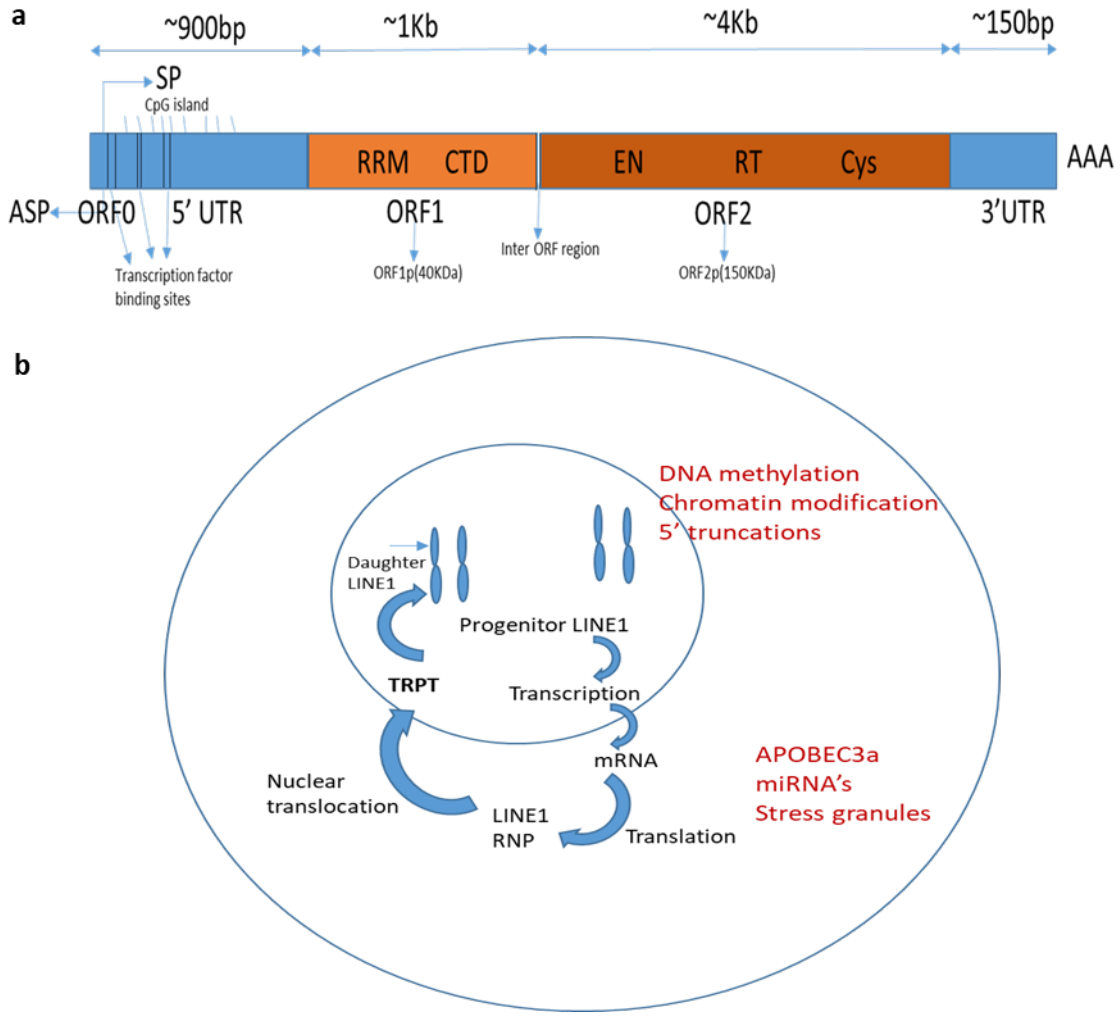

Figure 1.6 *Structure of L1 and the process of retrotransposition*

Figure shows the structure of an active L1 element (a) and the process of retrotransposition and the host defence mechanisms (b)

1.5.2 Influence of active L1s and retrotransposition on transcriptional deregulation

The majority of the L1 elements are not mobile but retain strong promoter activity, as demonstrated by the fact that while only 80-100 L1 elements are capable of retrotransposition, approximately 7000 L1 copies can initiate transcription (Khan H. *et al.*, 2006). These aberrant transcriptionally active L1 promoters are potentially as relevant to cancer pathogenesis as the L1 retrotransposition and de-novo insertion elements (Tufarelli C. and Badge R.M., 2017). L1 promoter activity is bidirectional so not only it can initiate

transcription from the sense strand, it can also drive anti-sense transcription (Tufarelli C. and Badge R.M., 2017). The transcription of the L1 element can extend into the adjacent genomic sequence beyond L1, known as the flanking sequences, giving rise to novel RNA transcripts which can either code for a novel protein or may act as regulatory RNA's (Tufarelli C. and Badge R.M., 2017). Currently there is little evidence of L1 5'UTR driven retrotransposition in normal somatic cells, although an in-vivo study using transgenic mouse models from Muotri et al showed L1 somatic retrotransposition in neuronal cells (Muotri A.R. *et al.*, 2005). The group demonstrated in-vitro that retrotransposition could alter the expression of several neuronal genes (Muotri A.R. *et al.*, 2005). In human cancer cells, both the L1 5'UTR sense and anti-sense promotor have been shown to drive the transcription of the protein coding sequences creating chimeric transcripts (Cruickshanks H.A. and Tufarelli C., 2009; Vafadar-Isfahani N. *et al.*, 2017). c-Met is a tyrosine receptor kinase over-expressed in several cancers and harbours an intronic L1 promoter. Activation of the L1 promoter within the c-Met gene leads to the formation of a truncated L1-Met chimeric transcript, whose role has been implicated in several cancers (Roman-Gomez J. *et al.*, 2005; Wolff E.M. *et al.*, 2010).

L1 mediated 3' transduction has been shown to alter the genome. When L1s retrotranspose and insert into a target sequence the 3' machinery may skip the L1 polyadenylation and instead use the second downstream polyadenylation site (Goodier J.L. *et al.*, 2000). An example for such L1 mediated transduction was first shown by Miki et al, wherein the disruption of the APC gene in colon cancer caused by somatic insertion of an L1 element, composed of the 3' portion of L1 sequence with nearly 180 base pairs of polyadenylate tract (Miki Y. *et al.*, 1992). Pan-Cancer analysis of the whole genomes from a sample size of 2954 cancer genomes and 38 histological cancer subtypes, identified 19,166 somatic retrotransposition events making it a predominant somatic structural variation in several cancer types (Rodriguez-Martin B. *et al.*, 2020). Integration of L1 transduction on CDKN2A tumour suppressor gene caused mega base size deletion, which led to loss of one copy of the gene, implicating the role of L1 mediated deletion on oncogenic potential in oesophageal tumours (Rodriguez-Martin B. *et al.*, 2020). Similarly L1 mediated rearrangements induced breakage-fusion-bridge cycles that triggered CCDN1 oncogene amplification in oesophageal adenocarcinoma (Rodriguez-Martin B. *et al.*, 2020). CCDN1 is a known oncogene, whose role is implicated in several cancers (Lü J. *et al.*, 2020; Rodriguez-Martin B. *et al.*, 2020).

Shukla et al have mapped the L1 integration sites in HBV and HCV related HCC using Retrotransposon Capture sequencing (RC-seq) and reported that germline L1 insertion in MCC gene (tumour suppressor mutated in colorectal cancer) activate oncogenic β-Catenin/Wnt Signalling pathway (Shukla R. *et al.*, 2013a). They also observed the activation of a potential liver oncogene ST18 (suppression of tumorigenicity-18) upon intronic L1 insertion (Shukla R. *et al.*, 2013a). In a follow up study, they have observed active retrotransposition in cases of alcoholic liver related disease HCC (Schauer S.N. *et al.*, 2018) however, the rate of somatic retrotransposition was found to be lower in non-viral cases (8 somatic L1 insertions in 25 individuals) than virus associated HCC (12 somatic L1 insertions in 19 donors) (Shukla R. *et al.*, 2013a; Schauer S.N. *et al.*, 2018).

1.6 Regulation of L1s

1.6.1 L1 Promoter Regulation- Role of DNA methylation and Chromatin modifications

Literature studies predominantly suggests that the epigenetic abnormality especially global hypomethylation is a leading cause of cancers, including HCC and is associated with a poor prognosis (Ehrlich M., 2002; Ateeq B. *et al.*, 2008; Gao X. *et al.*, 2014; Xiao-Jie L. *et al.*, 2016). The CpG islands of the L1 promoter region are often methylated, blocking the access of transcription factors binding to the L1 promoter and thereby preventing the initiation of transcription and expression is silenced in normal somatic cells (Inamura K. *et al.*, 2014). However, a consequence of the global hypomethylation prevailing in cancer cells is the reactivation of L1 elements and notably, global hypomethylation has been classified as a major contributor for HCC tumorigenesis (Gao X. *et al.*, 2014; Tufarelli C. and Badge R.M., 2017). The association between the activation of L1 elements due to global hypo-methylation of the genome and poor prognosis in the HCC patients still remain unclear (Gao X. *et al.*, 2014). However, these reports suggest that L1 hypomethylation has potential for use as a diagnostic and prognostic biomarker (Shigaki H. *et al.*, 2013; Xu R.H. *et al.*, 2017; Wu X. *et al.*, 2020). Zhu et al quantified the methylation status of 3 individual CpG sites within the L1 promoter and analysed L1-inserted c-MET (L1-MET) gene expression and its correlation with L1 methylation levels. They concluded that L1 hypomethylation is associated with poor prognosis in HCC patients due to activation of c-Met expression (Zhu C. *et al.*, 2014). Several tumour suppressor and oncogenes involved in Jak/Stat, Wnt/β-catenin and *Ras* pathways are differently methylated in very early stages of hepatocarcinogenesis (Calvisi D.F. *et al.*, 2007) coupled with

cellular stress including but not limited to HBV and HCV infections, chronic alcoholism, oxidative stress and other liver diseases may drive retrotransposition and L1 expression leading to hepatocarcinogenesis (Morikawa T. *et al.*, 2012). This hypothesis is supported by a literature study from Goodier et al which states that L1-ORF1p co-localizes with markers of cytoplasmic stress granules in stressed cells (Goodier J.L. *et al.*, 2007) although L1 expression varies from cell to cell depending on the stressor.

Chromatin modifications can also influence L1 regulation. For instance a study by Tenang et al showed that exposing HeLa cells to Benzo(a)Pyrene (BaP) induced early enrichment of H3K4me3 and H3K9Ac chromatin markers and also reduced the association of DNA methyltransferases to L1 promoter(Fig 3) (Teneng I. *et al.*, 2011). These DNA methyltransferases play a critical role in DNA methylation of L1 promoters and hence, if they get blocked, reactivation of L1 elements occurs within these cells (Teneng I. *et al.*, 2011). This study demonstrates a link between DNA methylation, genomic stress and chromatin modifications in regulation of the L1 elements (Teneng I. *et al.*, 2011). During the early phases of embryonic development the DNA methylation marks in the genome are erased, which keeps the retrotransposons active, however an epigenetic switch to histone-based control of the retro-elements is activated when the DNA methylation disappears (Walter M. *et al.*, 2016). A study by Walter et al showed that when H3K9 dimethylation (H3K9me2) mark disappeared during early embryogenesis, H3K9 and H3K27 trimethylation (H3K9me3 and H3K27me3) chromatin markers secured the control of keeping the L1 retrotransposons in check (Walter M. *et al.*, 2016).

1.6.2 Role of transcription factors

The L1 5' UTR contains a sense promoter whose activity is both RNA polymerase II (pol-II) and RNA polymerase III (pol-III) dependant. A study from Athanikar et al showed that the Ying Yang 1 (YY-1) binding site acts as a component of the L1 promoter, directing transcription initiation (Athanikar J.N. *et al.*, 2004). Those L1 element progeny lacking the YY1 binding site would lack an internal promoter, hence would lead to their extinction (Athanikar J.N. *et al.*, 2004; Sanchez-Luque F.J. *et al.*, 2019). Harris et al proposed a model by which p53 can limit L1 retrotransposition in somatic cells. The presence of active p53 in normal somatic cells activates the transcription of the L1 elements, creating more ORF2 mediated double strand

break and more p53 activity, thereby creating a positive feedback loop by increasing the p53 dependent DNA damage response and eventually leading to apoptosis of the cell (Harris C.R. *et al.*, 2009).Other transcription factors like RUNX3 and SOX have been shown to interact with L1 promoter as well. Tchenio et al demonstrated the role of SOX proteins in L1 promoter regulation via a reporter assay where a 10 fold increase in L1 promoter activity was observed by overexpressing SOX-11 in HEK-293 cells (Tchénio T. *et al.*, 2000). In a similar study Yang et al showed that exogenous expression of RUNX3 increased the L1 transcription and retrotransposition in HeLa cells (Yang N. *et al.*, 2003). Sun et al, 2018 developed a MapRRCon pipeline and performed a more comprehensive computational analysis of transcriptional factors binding to L1 5'UTR using EN-CODE ChIP-seq datasets and identified ~175 novel transcriptional factors binding to the L1 promoter apart from the previously known ones. Important among them are Myc and CTCF as they play major role in cancer progression (Sun X. *et al.*, 2018). They also identified Myc as the transcriptional repressor of L1 promoter in HEK293 cells, wherein knockdown of Myc by specific siRNA increased the promoter activity of L1 5'UTR in HEK293 cells (Sun X. *et al.*, 2018).

1.6.3 Post-transcriptional regulation of L1 elements

Silencing of the L1 promoter occurs in the nuclear compartment through DNA and chromatin modifications and 5' UTR truncations (Rahbari R. *et al.*, 2015). However once L1 gets activated, the vast majority of the L1 elements are regulated by the cytoplasmic machinery post-transcriptionally through small-interfering RNAs (si-RNAs) (Yang N. and Kazazian H.H., 2006) or piwi-interacting RNA mediated mechanisms (Malone C.D. *et al.*, 2009) and APOBEC proteins(Fig 2b) (Lovsin N. and Peterlin B.M., 2009). A study from Soifer et al, showed that the RNAi (RNA-interference) mechanism can regulate the expression of L1 elements (Soifer H.S. *et al.*, 2005). The RNAi mechanism is initiated by the cleavage of L1 double-stranded RNA by RNase III enzyme DICER and the resultant L1-siRNAs recognises the cognate L1mRNAs leading to the degradation of the transcript (Soifer H.S. *et al.*, 2005). In addition to this, there are small non-coding RNAs called the piwi-interacting RNAs (pi-RNAs) shown to control the L1 regulation (Sigurdsson M.I. *et al.*, 2012). Pezic et al showed that in mice one of the PIWI proteins called MIWI-2 forms a complex with pi-RNA which then recruits a histone methyltransferase (Pezic D. *et al.*, 2014). This complex leaves a mark of H3K9me3 on 5' UTR

region of a full length actively transcribed L1 element, which in turn suppresses its expression in germ cells and also in somatic cells (Pezic D. *et al.*, 2014).

Inhibition of L1 activity has also been shown to be mediated by cellular deaminases like ADAR and APOBEC family proteins (Orecchini E. *et al.*, 2018). The APOBEC proteins are a part of the innate immune system which fights against exogenous retroviruses (Schumann G.G., 2007). The role of APOBEC3B proteins in L1 regulation has been shown by Wissing et al where in shRNA mediated knockdown of APOBEC3B proteins in HeLa and hESCs cells led to increase in L1 retrotransposition by about 3-fold; however, the exact mechanism of inhibition was not demonstrated (Wissing S. *et al.*, 2011). More recently, Richardson et al demonstrated that APOBEC3A can inhibit L1 retrotransposition by deaminating transiently exposed single-stranded DNA that arises during new L1 integration events (Richardson S.R. *et al.*, 2014). Similarly, recent studies by Orecchini et al demonstrated that knockdown of ADAR1 expression in HeLa cells increased L1 retrotransposition and conversely overexpression of ADAR1 caused reduction in L1 retrotransposons activity (Orecchini E. *et al.*, 2017a; Orecchini E. *et al.*, 2017b).

1.7 Thesis aims

The purpose of this thesis was to investigate the regulators of active L1 retrotransposition and to understand the factors leading to aberrant activation of L1 elements in the context of HCC. Specific aims are as follows

- To evaluate the potential of external stimuli related to HCC (HCV infection) to reactivate retrotransposons and/or retrotransposition.
- To study the regulation of active L1 retrotransposition and expression in HCC related cell lines.
- To determine the oncogenic potential of active retrotransposition in immortalised human hepatocyte cell line (HHL-5).
- To develop a CRISPR-Cas9 based system to selectively isolate L1 promoter and study chromatin interactions at the loci under different physiological conditions.

Chapter 2: Materials and Methods

2.1 Cell culture

Human HCC related cell lines (HepG2, Huh7, Huh1, PLC-PRF/5, Hep3B, SNU182, SNU475 and SK-Hep1) were obtained from NICR cell biobank and Human Immortalized Hepatocytes;HHL-5 (Clayton R.F. *et al.*, 2005) were obtained from Prof Arvind Patel, MRC Virology Unit, Institute of Virology, Glasgow, UK. HepG2, Huh7, Huh1, PLC-PRF/5 and SK-Hep1 were maintained in Dulbecco's Modified Eagle Medium/nutrient mixture F-12 ham (SIGMA-D6421), SNU182 and SNU475 were maintained in RPMI1640 (SIGMA-R5886) supplemented with 10% Foetal Bovine Serum (FBS), 1% Glutamine (SIGMA-G7513) and 1% Penicillin-Streptomycin solutions (SIGMA-P0781). HHL5 cells were maintained in Dulbecco's Modified Eagle Medium/ high glucose (SIGMA-D5671) supplemented with 10% FBS, 1% MEM Non-Essential amino acids (GIBCO 11140050) and 1% Penicillin-streptomycin solutions. Cells were passaged and maintained on T-75 flasks and were incubated at 37°C, 5% CO2 in a humidified chamber.

2.2 Transfection efficiency analysis using Fluorescent Assorted Cell Sorting (FACS)

Cells were seeded on a 6 well plate (~$2x10^5$ cells/well) to obtain around 80% confluency next day. Keeping one well as un-transfected control other wells were transfected with EGFP plasmid using 2 different transfection reagents- TransIT-LT1 and TransIT-X2 (Mirus Bio, Cat no. 6003) and by altering the DNA to reagent ratio- 1:3 to 1:6 as per manufacturer's recommendation. 48h after transfection cells were harvested and analysed for transfection efficiency by FACS to quantify EGFP positive population. The condition which yielded the maximum transfection efficiency was selected for future transfections.

2.3 Transfection with plasmids

The cells were seeded onto 6 well plates and transfected next day at around 80-90% confluency with appropriate plasmid DNA's using TransIT-LT1 transfection reagent (Mirus Bio, Cat no. MIR 2304) in 1:3 ratio (1µg of plasmid DNA to 3µl of TransIT-LT1) and incubated under standard conditions for 48 hours. If required, selection with Puromycin (2µg/ml, Sigma P8833) was carried out for plasmid transfections having a puromycin resistance cassette for 2 to 3 days until untransfected cells were all dead.

2.4 *In vitro* Retrotransposition assay

2.4.1 EGFP as a reporter system

HCC cell lines were transfected with 1µg of 99GFPLRE3 (retrotransposition competent) or 1µg of 99GFPJM111 [retrotransposition incompetent by introduction of point mutation in ORF1, ARR conserved block of amino acid near c-terminus of ORF1 protein at residues 260-262 mutated to AAA (Moran *et al.*, 1996)] plasmid containing a GFP based retrotransposition cassette and puromycin resistance gene using TRANSIT-LT1 transfection reagent (the plasmids are a kind gift from Dr Jose Luis Garcia-Perez, MRC Human Genetics Unit, University of Edinburgh). Selection with 2µg/ml puromycin was initiated 4 days after transfection and continued for further 3-4 days. The cells were harvested and divided into two fractions. One fraction of cells was analysed by FACS to quantify GFP expression. The number of live cells that expressed GFP represented the retrotransposition efficiency (cells that had undergone a complete cycle of retrotransposition) and the other fraction of cells were pelleted for genomic DNA extraction that was further analysed quantitatively by Taqman qPCR (see section 2.7.1).

2.4.2 Blasticidin resistance as a reporter system

HCC cell lines were transfected with following plasmids: 1. A wild type L1 retrotransposition plasmid (pJJ101/L1.3) containing blasticidin-based retrotransposition cassette ; 2. A mutant LINE1 retrotransposition plasmid [pJM105/L1.3 has a missense mutation (D702Y) in RT domain of ORF2 protein (Moran *et al.*, 1996)] which served as a negative control; and 3. pCEP4-blasticidin which served as a positive control (the plasmids are a kind gift from Dr Jose Luis Garcia-Perez, MRC Human Genetics Unit, University of Edinburgh). 5 days after the transfection the cells were harvested and transferred to 10cm dishes with 4µg/ml of Blasticidin. Fresh Blasticidin media was added every 3 days and the selection continued for further 12-14 days or until the cells in the negative control were completely dead and Blasticidine resistant colonies developed in pJJ101/L1.3 transfected cell plate. To quantify for L1 retrotransposition the colonies were fixed and stained with crystal violet (see section 2.5) and quantified by colony counter.

2.5 Clonogenic Assay and crystal violet staining

200 cells were seeded per 10cm dish of respective conditions and the dishes incubated at 37°C, 5% CO2 in a humidified chamber for 2 weeks. The dishes were fixed at the end of 2

weeks with methanol enough to cover the surface for 20 minutes. Once the cells were fixed the fixative was removed and dishes were left open to air dry and rinsed with distilled water once. Then 0.4% crystal violet solution (SIGMA C0775) was added and stained for 5 minutes. At the end of staining period the plates were washed with distilled water and air dried. The colonies were observed and counted by an automated cell counter.

2.6 Cell Proliferation

2.6.1 Sulforhodamine B (SRB) Assay

Appropriate number of cells (500-1000) were seeded onto 96 well plate with 200µl media to get 10-20% confluency the next day and incubated at 37°C, 5% CO2 in a humidified chamber. The cells were fixed on respective days with 50µl of Carnoy fixative (3 parts of methanol and one part of glacial acetic acid) and stored at 4°C for at least 24 hrs. Then the plates were rinsed with deionised water and air dried. 0.4% of Sulforhodamine B (SRB) solution in an amount enough to cover the surface was added to each well and stained for 40 minutes. At the end of staining period, SRB solution was removed and the plates were washed with 1% acetic acid and air dried. Finally, 200µl of 10mM Tris base was added to each well to dissolve the stain and optical density (OD) values were measured at 540nm using a Microplate spectrophotometer.

2.6.2 Incucyte zoom

Cells were seeded onto 96 well plates to get 10-20% confluency the next day and incubated at 37°C, 5% CO2 in the IncuCyte® system. The cell growth was monitored using the IncuCyte® Live-Cell Analysis System to capture phase contrast images every four hours. 4 images per well were captured and analysed using the integrated confluence algorithm. Mean of the 4 images was taken as the % confluence of the well at a time point, 3-5 wells were set per condition.

2.7 Genomic DNA extraction, PCR and Agarose gel electrophoresis

Genomic DNA was extracted using DNeasy Blood and Tissue Kit (Qiagen-Cat#69504) as per manufacturer's protocol and DNA concentration was measured by Nanodrop. Oligonucleotide primers were designed for genomic DNA regions using NCBI primer blast. PCR was carried out using Platinum Green Hot start 2x PCR master mix (Invitrogen-Cat#13001013) on a thermocycler. Cycling parameters were as follows- 94°c for 2 minutes then 30-40 cycles

with following parameters- 94°c for 30seconds, 50-60°c (depending on the primer melting temperature) for 30seconds, 72°c for 1 minute/kb (see appendix table 1 and 2 for primer details). The PCR products were subjected to agarose gel electrophoresis and the bands were visualized by exposing to UV light and the picture taken with Biorad chemi doc imaging system. The agarose gel images were quantified using ImageJ.

2.7.1 Taqman qPCR

Quantitaive PCR was carried out using QuantStudio ™ 7 flex Real Time PCR system and 2x Taqman genotyping qPCR master mix (Thermo Fischer Scientific-Cat#4371353). Standard curves were performed by serial dilutions (gDNA for RNaseP and EGFP plasmid for GFP) and primer pair efficiency was quantified. To assay for the removal of intron from the GFP based retrotransposition assay, 200ng of genomic DNA was used as a template in a 10µl PCR reaction with forward and reverse primers (50µM), FAM dye labelled BHQ1 probes (10µM) designed to target the GFP region and VIC dye labelled TAMRA probes designed to target RNaseP region. To compare the GFP insertion rate L1-GFP copy numbers were normalized to RNaseP and calculated using the ddct relative quantification method.

2.8 Quantitative Real Time PCR

RNA was extracted using RNeasy Mini Kit (Qiagen-Cat#74104) as per manufacturer's protocol. DNase treatment was performed to avoid genomic DNA contamination. Total RNA was treated with DNase digestion reagents (Turbo DNase and 10x buffer) then incubated at 37°C for 30 minutes followed by addition of DNase inactivation reagent. 1µg of total RNA was diluted with distilled water-one sample for cDNA and other for no-RT control using Promega Cat#A3500 kit. In short, the samples were first heated at 70°C for 10 minutes and kept on ice. Then MgCL2 (4µL), Reverse Transcriptase (10µl), d-NTP mix (2µl), Recombinant RNasin Ribonuclease Inhibitor (0.5µl), AMV Reverse Transcriptase (0.5µl) and oligo-dt 15 or random hexamers primer (1µl) were added to each sample except no AMV Reverse Transcriptase in no-RT control tubes. The samples were incubated at 42°C for 15 minutes followed by 95°C for 5 minutes, then kept on ice for 5 minutes. The cDNA samples were stored at -20°C for RT-qPCR analysis. Quantitative Real Time PCR was carried out as 10µl reactions in triplicates containing-5µl platinum SYBR green (Invitrogen-Catalogue number 11744), 0.2µl forward and reverse primer mix (10µM) and 2µl C-DNA. Real time PCR amplification was completed using

QuantStudio ™ 7 flex Real Time PCR system. Data was analyzed using the QuantStudio software and the Melting curves were used to ensure no primer-dimer formation. To evaluate the primer efficiency of each primer set, standard curve was set up by diluting cDNA and the slope and R^2 were determined. Final quantifications were done by ddCt method using TBP/HPRT/GAPDH/18s as housekeeping genes (See Appendix table 1 and 2 for all the primers used in the project). Primers were designed against the coding region for each target genes using NCBI-Primer Blast.

2.9 Western Blot analysis of Total Protein Extract

The cell pellets were homogenized in SDS lysis buffer [Tris HCL (pH 6.8)-12.5ml, SDS-2g, Glycerol-10ml, dissolved in 67.5ml distilled water] by heating at 100°C for 10 mins and then sonicated (3 pulses of 10seconds each). Then the samples were spun at 13000rpm for 10 minutes at 4°c and the supernatant containing the protein lysate were transferred to newly labelled Eppendorf tubes. The protein concentration was estimated by Pierce BCA protein assay kit. 30-50μg of protein extract was run on BIORAD Mini PROTEAN 4-15% gradient SDS-PAGE gel (Cat#456-8086) with 1x Gel running buffer (Glycine-144g, 30g Tris-base, 10g SDS dissolved in 1L of distilled water). Proteins were transferred onto a nitrocellulose membrane for western blot analysis. The membrane was blocked for 1 hour in 5% milk/BSA (5% skimmed milk powder/BSA in 1x TBS tween) then incubated overnight at 4°C with indicated primary antibody in blocking solution, washed with TBS tween at room temperature (3 times 10 minutes each) followed by 1 hour incubation at room temperature with anti-mouse/rabbit HRP-conjugated secondary antibody 1:2000 (DAKO Polyclonal Goat Anti-Mouse-HRP or Polyclonal Goat Anti-Rabbit-HRP) in the blocking solution. The membrane was washed again 3 times before visualizing the antibody tagged protein bands using ECL chemiluminescence system (Amesham ECL Western Botting detecting reagent-product code-RPN2106). ECL signals were captured on x-ray films or Biorad chemi doc system. The membrane was then stripped with mild stripping solution (Composition-15g glycine, 1g SDS, 10ml Tween-20 dissolved in 1L of distilled water with pH 2.2) and re-probed with another primary antibody. Finally the membranes were probed with anti-GAPDH or anti-Tubulin to be used as loading controls (See Appendix table 3 for primary antibodies used in this study for western blotting)

Chapter 3: Regulation of L1 retrotransposition and expression in context of HCC

3.1 Introduction

One of the most common factors associated with HCC development is chronic hepatits B or C virus (HBV or HCV) infection (discussed in detail in chapter 1- section 1.2). Around 214,000 people are chronically infected with HCV in the UK while globally, HCV infection is prevalant in 2.2-3% of the world population (130-170 million people) and more than 350,000 people die of HCV-related conditions (including HCC) per year ('Global burden of disease (GBD) for hepatitis C,' 2004; Lavanchy, 2009). A major advance in the field has been the introduction of direct acting antivirals (DAAs) targeting HCV infection, as now the majority of patients can be cured of HCV (Chung R.T. and Baumert T.F., 2014). However, what is also increasingly clear is that, while this reduces HCC risk in infected individuals, the risk is not eliminated- in either the global population treated with the antivirals, or in those with HCV receiving curative treatments for HCC (Rinaldi L. *et al.*, 2020). Both de novo cancer risk and HCC recurrence risk persist at levels in order of 2-5 fold above patients previously cured of HCV with interferon containing regimes (Reig *et al.*, 2016; Baumert *et al.*, 2017; Reig *et al.*, 2017). The coexistence of chronic liver disease (CLD) makes the use of traditional cytotoxic agents risky, with no survival benefit. Moreover, the majority of patients have incurable disease at presentation, due to late detection when curative treatments (resection, ablation) cannot be applied. Hence, it is widely recognised that to have a major impact on survival, early detection and delivery of curative therapies would be highly preferable. Thus even after advent of successful DAA therapies, there remains an unmet need to understand HCC risk in HCV patients, as well as a need to develop tools for early detection and HCC treatment strategies (Ioannou, 2021). We believe that this can be achieved by understanding the molecular mechanisms that lead to hepatocarcinogenesis.

Why all those infected with HCV do not develop HCC is unknown, as are the mechanisms whereby HCC risks persists after viral infection treatment. Upon HCV infection, the virus modulates the host cells for its own survival and replication. The HCV genome is directly translated at the rough endoplasmic reticulum (ER) in a single polyprotein precursor that is eventually cleaved by cellular and viral proteases into ten mature products. These virus

encoded proteins then further participate in the process of viral replication and assembly (Dustin *et al.*, 2016). Overall, HCV replication process induces oxidative and ER stress in the cells, thus promoting hepatocarcinogenesis (Maki *et al.*, 2007; Medvedev *et al.*, 2016; Zhang *et al.*, 2019). Moreover, HCV infection is demonstrated to induce autophagy (Chu and Ou, 2021), innate immune response (Chan and Ou, 2017) and impair DNA damage repair pathways (Machida *et al.*, 2010; Ryan *et al.*, 2016; Nguyen *et al.*, 2018). Hence, HCV encoded proteins interact with various host proteins dysregulating different pathways, contributing towards hepatocarcinogenesis. Several of these changes are also epigenetic in nature (Rongrui *et al.*, 2014; Wijetunga *et al.*, 2017). Moreover, many of these epigenetic changes have been shown to persist even after HCV infection gets cleared by DAA treatment (Hamdane N. *et al.*, 2019; Perez *et al.*, 2019). Okamoto et al have demonstrated global DNA hypomethylation (measured using L1 promoter) by HBV and HCV in a humanised mouse model of hepatitis virus infection (Okamoto *et al.*, 2014). Furthermore, Shukla et al, demonstrated L1 activation in association with active retrotransposition and its impact on oncogenic signaling pathways in HBV and HCV related HCC (Shukla *et al.*, 2013). In a follow up study they have observed active retrotransposition in cases of alcoholic liver related disease HCC (Schauer *et al.*, 2018), however the rate of somatic retrotransposition was found to be lower in non-viral cases (8 somatic L1 insertions in 25 individuals) than virus associated HCC (12 somatic L1 insertions in 5 out of 26 donors) (Shukla *et al.*, 2013; Schauer *et al.*, 2018). In addition, Shukla et al have demonstrated active retrotransposition in the non-tumour liver tissue of one HBV-HCC case (Shukla *et al.*, 2013) showing that although not as frequent as in tumour tissues, active retrotransposition can occur in non-tumour tissues (at least in presence of HBV infection). It is well established in literature that L1 elements can get activated in tumours of various cancer types and are involved in somatic structural variations leading to cancer evolution as well as can play a potential role in human cancer development (Lee *et al.*, 2012; Rodriguez-Martin *et al.*, 2020).

Hence, we hypothesise that L1s may get activated as opportunistic parasites in Chronic Hepatitis C (CHC) patients, due to virus assisted epigenetic remodeling and suppression of host defense factors. Thus in this chapter, we **aim** to test this in human patients and evaluate the influence of HCV on the process of active retrotransposition using *in vitro* models. We also aim to look at the role of DNA damage response elements in HCV mediated L1

retrotransposition. Once activated, L1s can continue towards genomic instability contributing towards cancer development even beyond HCV clearance. Thus, these mechanistic understandings can give a direction to develop therapeutic or cancer preventative approaches for the CHC patients.

3.2 Materials and methods

3.2.1 Cell lines

Besides the cell lines mentioned in section 2.1 additional cell lines used in this chapter were Huh7-J17 cell line and its corresponding parental Huh7 cells. The cell lines were a kind gift from Prof. Arvind Patel, Professor of viral vaccinology, University of Glasgow. Upon arrival, the Huh7-J17 cells were confirmed for the presence of HCV replicon by luciferase assay and western blotting of HCV protein (NS5a) (see Fig 3.2a). The cell lines were maintained in RPMI-1640 (Sigma) media supplemented with 10% foetal bovine serum, 1% Glutamine and 1% Penicillin-Streptomycin for further experimentation. The Huh7-J17 cell line was maintained with puromycin (2µg/ml) in the media.

3.2.2 Patient samples

Archived diagnostic formalin-fixed paraffin embedded (FFPE) liver biopsies from patients with HCV with or without associated HCC were obtained from Newcastle Cancer Centre biobank. All patients had provided written consent for use of their tissues for research purposes. Ethical approval was obtained for the use of FFPE CHC patient biopsies (study reference: NAHPB-126) by the National Research Ethics Service (NRES) Committee North East (REC ref: 12/NE/0395) sponsored by NUTH Trust R&D (Ref: 6579).

3.2.3 Recovery of HCV infection by sofosbuvir treatment

Huh7-J17 cells were seeded in puromycin-free media and treated with Sofosbuvir (PSI-7977-HCV Ns5b polymerase inhibitor, Adooq Biosciences Cat#A11529) at 10µM dose for 12 days wherein the media was replaced with fresh drug treatment every 72 hrs. At the end of the treatment, the cells were harvested and pelleted for total RNA (sec 2.8) and protein extraction (sec 2.9).

3.2.4 L1 retrotransposition with DNA damage response inhibitors

Retrotransposition assay was performed using GFP based retrotransposition reporter constructs as stated in section 2.3.1. Four days after transfection, puromycin selection (2µg/ml) was initiated in presence of small molecule inhibitors that target various DNA damage response pathway proteins: 10µM KU-55933 (ATMi, TOCRIS, 3544), 1µM NU-7441 (DNA-PKi, Apex Bio, A8315), 1µM VE-821 (ATRi, Axon, 1893), 1µM SRA737 (CHK1i, Selleckchem, S8253) or PARP (Rucaparib, 10µM). Cells were harvested after 3-4 days of selection in presence of DNA damage response inhibitors and were further analysed by FACS and Taqman PCR as stated in section 2.7.1 to quantify retrotransposition rates.

3.2.5 DNA damage repair plasmid re-joining assays

Homologous recombination repair (HRR) and nonhomologous end joining repair (NHEJ) activity of the indicated cell lines was assessed by plasmid re-joining assays (Bradbury *et al.*, 2020). In brief, pDRGFP contains an in vivo homologous recombination substrate that is composed of two differentially mutated GFP genes oriented as direct repeats and separated by a drug selection marker that can be excised by I-SceI (Sudhindar *et al.*, 2021). pCBASce1 expresses the I-SceI endonuclease that introduces a DSB at an I-SceI site. Upon successful repair, GFP is expressed in the cells and thus is an indicator of HRR. The pimEJ5GFP plasmid is an I-SceI-based chromosomal break reporter for NHEJ. In this reporter, end joining between two distal tandem I-SceI recognition sites restores an EGFP expression cassette, caused by deletion of the intervening pgkPURO cassette; thus, GFP expression is an indicator of NHEJ (Sudhindar *et al.*, 2021).

The cells were seeded onto 24 well plates to obtain 80-90% confluency the next day. Then the cells were transfected with equimolar concentrations of DR-GFP and pCBAScel for HR activity, pEJ5-GFP and pCBAScel for NHEJ activity using TransIT-LT1 transfection reagent. pEGFP expressing plasmid was used as a transfection control. The cells were harvested 48hrs after transfection for FACS analysis. GFP positive populations were gated based on SSC-FSC scatter plots with 50,000 cells analysed per condition. Percentages of cells with active HR were calculated as 100× (% DR-GFP+ cells)/(% pEGFP+ cells) and the percentage of cells with active NHEJ were calculated as 100× (% pEJ5-GFP+ cells)/(% pEGFP+ cells).

3.2.6 Immunohistochemistry (IHC)

L1orf1p IHC was performed on a Ventana Discovery XT system using standard protocol (done by Miss. Misti McCain, IHC technician, Newcastle University Centre for Cancer). In short, antigen retrieval was performed using Discovery CC1 buffer (Roche 06414575001 (950-500)) followed by incubation with primary antibody against L1orf1p (1:2000, Mouse Monoclonal, MABC1152, Merck) followed by anti-mouse-HRP secondary antibody (Roche 05266556001 (760-150)). An expert liver-pathologist (Dr Yvonne Bury, Cellular Pathology, Royal Victoria Infirmary, Newcastle upon Tyne hospital) assessed the staining.

3.2.7 Luciferase assay

The cells were lysed with passive lysis buffer (Promega) and luciferase activity was measured using luciferase assay reagent (Promega, E1500) and Omega plate reader as per the instructions.

3.2.8 DNA damage induction by Gemcitabine treatment to assess ATR and Chk1 inhibitors efficacies

Huh7 cells were seeded in 12 well plates to attain approximately 90% confluency the next day. Cells were pre-treated or not with ATRi (VE-821, 1µM) and CHK1i (SRA737, 1µM) for 30 minutes followed by treatment with Gemcitabine (100nM) for 4 hrs. Finally, the cells were harvested for protein extraction using Phospho safe lysis buffer (Merck Millipore, cat no. 71296).

3.2.9 DNA damage induction by Doxorubicin treatment to assess ATM inhibitor efficacy

HepG2 cells were maintained in RPMI-1640 (Sigma) media supplemented with 10% foetal bovine serum, 1% Glutamine and 1% Penicillin-Streptomycin. Cells were seeded in 6 well plates to attain approximately 90% confluency the next day. Wells were pre-treated or not with ATMi (KU-55933, 10µM) for 30 minutes followed by treatment with 100nM of Doxorubicin and harvested 24 hrs after treatment for protein extraction using Phospho safe lysis buffer.

3.2.10 Bioinformatics

Human HCC RNAseq data was downloaded from the cancer genome atlas hepatocellular carcinoma (TCGA-LIHC) project and RNAseq data of CHC patients and healthy controls GSE84346 was obtained from NCBI GEO database. The reads were mapped to the human L1-

Ta sequence (5'UTR-promoter, Genbank: L19092) by BLAT alignment using an in-house algorithm to obtain L1 counts. The counts were normalised by the total number of reads in each library and expressed here as counts per million (bioinformatics analysis was performed by Dr Ruchi Shukla in collaboration with Prof Geoffrey Faulkner, Mater Research Institute, University of Queensland, Australia).

3.2.11 Statistical analysis

GraphPad Prism software (GraphPad 8.0 and 9.0) was used for statistical analysis. P values < 0.05 were considered significant. Mean ± standard errors are shown in figures where applicable. * = P < 0.05, ** = P < 0.01, *** = P < 0.001, **** = P < 0.0001. Data was analysed by one-sample t-test (for fold change), student's t-test (2 groups) or one- or two-way ANOVA with Tukey's multiple comparison correction when required (3 groups).

3.3 Results

3.3.1 HCV activates L1 expression in non-tumour tissue of patients with chronic hepatitis

To evaluate L1 activation in the liver of patients with chronic hepatitis C (CHC) prior to the development of HCC, we have analysed a publicly available RNAseq dataset of CHC patients and control (healthy) individuals (GSE84346). There was significant upregulation of L1 transcripts in the liver of CHC patients (Fig 3.1a). Likewise, interrogation of the RNAseq dataset of the cancer genome atlas hepatocellular carcinoma (TCGA-LIHC) study revealed upregulation of L1 expression in the non-tumour liver of patients with a history of viral hepatitis compared to patients with no history of any known HCC risk factors (Fig 3.1b). However, upon cancer development L1 was found to be upregulated in all the HCC cases irrespective of the underlying aetiology (Fig 3.1b). We also evaluated the presence of L1orf1 encoded protein (L1orf1p) expression immunohistochemically in HCV-infected liver biopsies from our own biobank. Again, L1orf1p expression was observed in the non-tumour tissue of some individuals (6 out of 11) years before HCC development. In two of the cases with subsequent diagnostic HCC biopsy tissue available, the earlier non-tumour L1 status matched that of the HCC (Fig 3.1c). Hence, the data demonstrates that L1s can be activated in a chronically diseased HCV infected pre-neoplastic liver and thus may contribute towards cancer development.

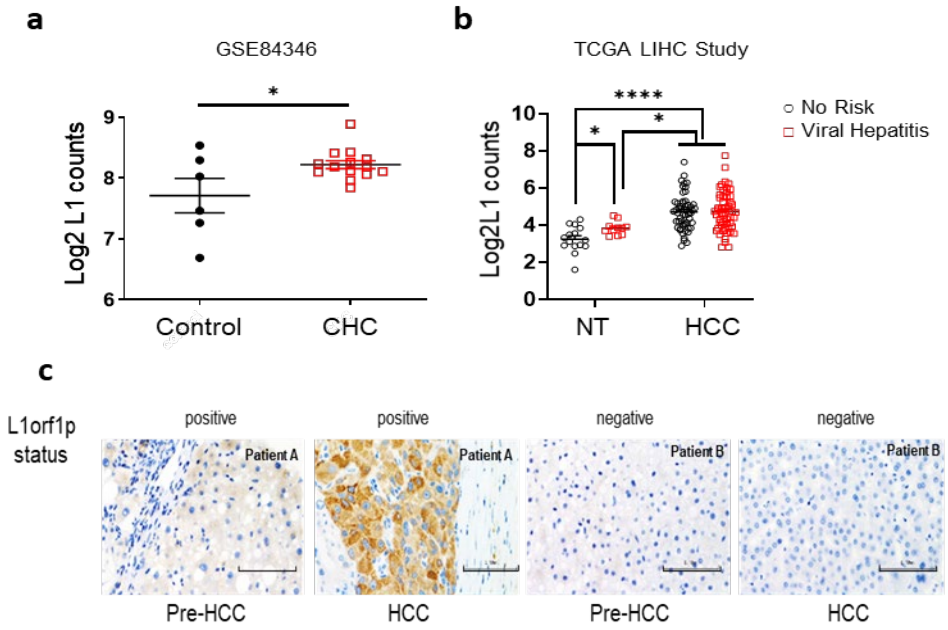

Figure 3.1 *HCV activates L1 expression in non-tumour liver*

*Graphs represent normalised L1 transcripts count in the liver of chronically infected HCV patients versus healthy volunteers control group (a), in the non-tumour liver and HCC tissue of HCC patients with indicated etiologies (b). Representative L1orf1p IHC images of CHC patients liver biopsies pre-HCC and HCC. Scale bar represents 0.1mm (c). * p<0.05, ****p<0.0001, unpaired t test (a) and two-way ANOVA (b).*

3.3.2 HCV actives L1 retrotransposition

Further, to evaluate the direct influence of HCV infection on L1 retrotransposition process, Huh7-J17 cells containing a plasmid-based HCV replication system (HCV genome without envelope proteins fused with puromycin-resistance cassette and luciferase expression cassette) was employed (Angus *et al.*, 2012; Magri *et al.*, 2016). Presence of the viral replicon was confirmed by western blotting showing expression of NS5A (an HCV encoded protein) and increased autophagy in Huh7-J17 cells as indicated by decreased level of p62 (an autophagic cargo adapter also called as Sequestosome-1) compared to parental Huh7 cells (Fig 3.2). Interestingly, upregulation of L1orf1p was observed in Huh7-J17 compared to Huh7 cells (Fig 3.2) (Sudhindar *et al.*, 2021).

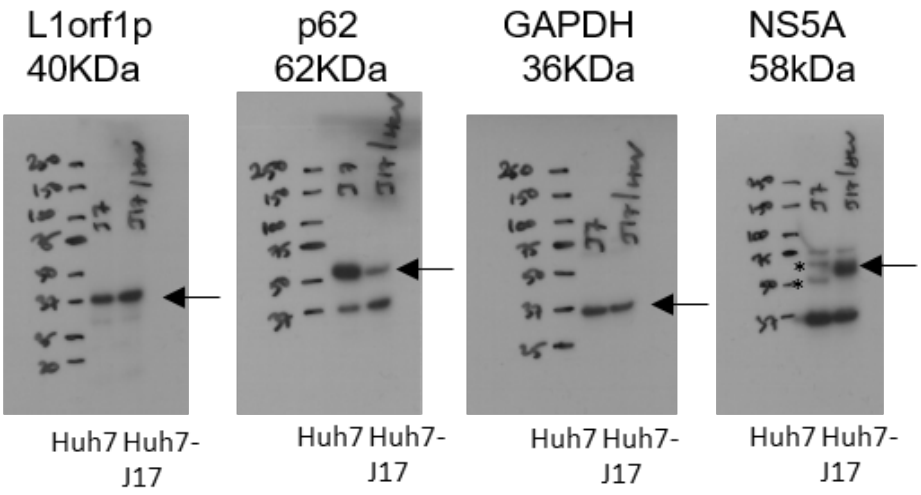

Figure 3.2 *HCV activates L1s in vitro*

*Western blot image (whole x-ray film scan) of whole cell lysates of Huh7 and Huh7-J17 cells with indicated antibodies. Expected molecular weights are indicated under each protein label. GAPDH was used as a loading control. Non-specific bands are marked with *. This experiment was repeated 3 independent times (n=3). One representative example is shown here.*

Next, we compared the retrotransposition efficiency of the Huh7-J17 cell line with the corresponding parental or naïve Huh7 cells by an *in-vitro* EGFP-based retrotransposition assay. In this system, an L1 retrotransposition reporter plasmid contains an EGFP cassette interrupted by an intron so that it can be expressed only after a successful cycle of retrotransposition leading to EGFP splicing and integration into the genome (Moran *et al.*, 1996) (Fig 3.3a). A significant increase in cells undergoing active retrotransposition was observed in Huh7-J17 cells compared to the Huh7 control cells, as evident by EGFP positive cells by FACS analysis (fold change ~3.2, Fig 3.3b-d). Insertion of L1-EGFP plasmid in the genome was further verified by genomic PCR to detect intron-less GFP and was quantified by Taqman qPCR, revealing a ~2.9 folds increase in retrotransposition efficiency in Huh7-J17 cells with HCV replicon compared to the naïve cells (Fig 3.4a and b). The influence of HCV on active retrotransposition was further independently verified by blasticidin-based retrotransposition assay, wherein attainment of blasticidin resistance acts as a marker of active retrotransposition (Moran *et al.*, 1996). Again, a significant increase in blasticidin resistant colonies were observed in Huh7-J17 compared to Huh7 indicating increased retrotransposition rate in the cells in presence of HCV (fold change 6.8 ± 1.36, n = 3, Fig 3.4c) (Sudhindar *et al.*, 2021).

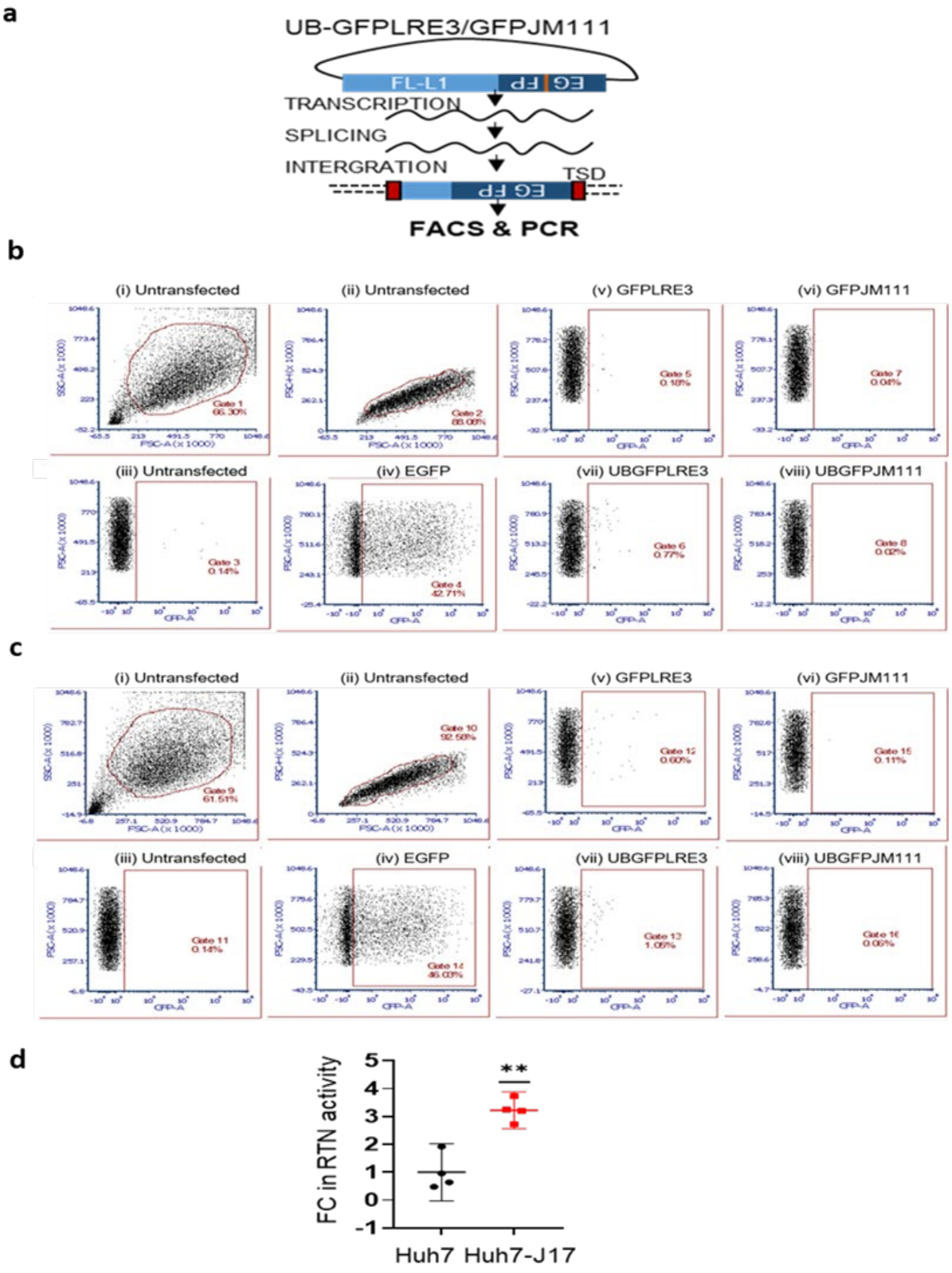

Figure 3.3 *EGFP-based retrotransposition assay gating strategy*

*Schematics representing retrotransposition assay with EGFP indicator cassette. UB denotes presence of UB promoter before the L1-retortransposition cassette. JM111 plasmid contains mutation in L1ORF1 (RR260-261AA) of LRE3 sequence rendering it to be retrotransposition incompetent and thus is used as a negative control (a). EGFP positive cells were selected by setting a gate on untransfected cells to remove the background debris signal (FSC-A SSC-A plot-i) and then selecting for single cell population (FSC-A FSC-H plot-ii). Then a gate on the single cell population was set on untransfected cells to set auto fluorescence of the cells in BL1 channel (plot-iii). Shift in fluorescence was then assessed for eGFP plasmid transfected cells as a positive control (plot-iv) and level of retrotransposition was assessed in the GFPLRE3 and UBGFPLRE3 transfected cells (plots-v and vii) and respective JM111 transfected cells (plots-vi and viii) for Huh7 (b) and Huh7-J17 (c). Results of the retrotransposition assay in Huh7 cells in the presence and absence of HCV: graph representing fold change (FC) in number of retrotransposition (RTN) positive cells observed by FACS analysis 5 days after transfection with UB-GFPLRE3 plasmid where negative gate was set using cells transfected with UB-GFPJM111 plasmid, n=4 independent repeats (d). **p<0.01 one-sample t test.*

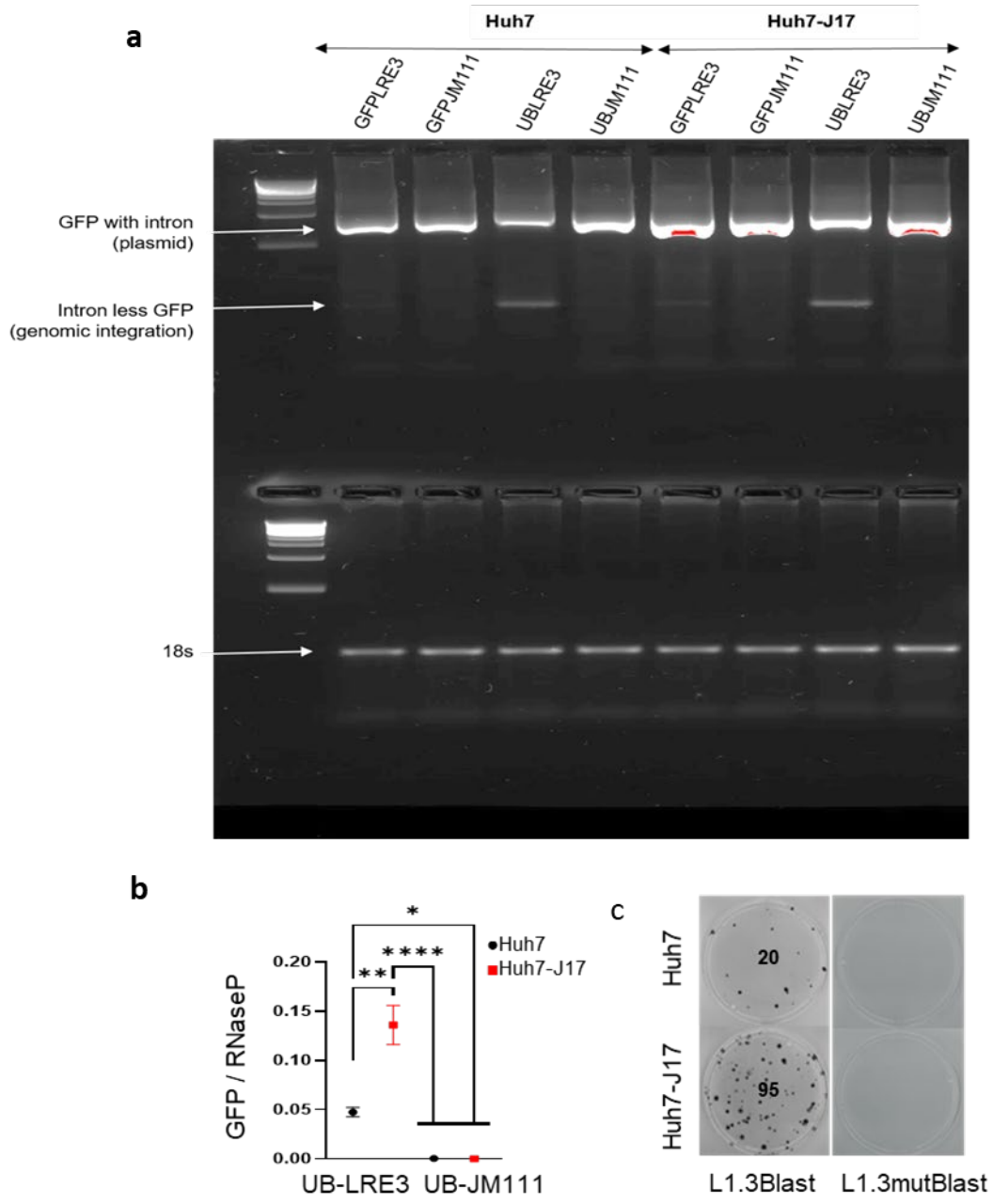

Figure 3.4 *EGFP-based retrotransposition assessment by genomic PCR followed by agarose gel electrophoresis*

*Top band in the top gel indicates GFP amplification from the LRE3/JM111 plasmids while bottom band in the top gel indicates intron less GFP and is visible only in cells transfected with LRE3 plasmids. Bottom gel represents amplification of 18S genomic region and is used as a loading control of the genomic DNA (a). Quantification of GFP insertions in the genome using RNaseP as a control, n=3 technical replicates of one representative Taqman qPCR assay (b) and Blasticidin-resistant colonies representative of active retrotransposition events visualised by crystal violet stain 3 weeks after selecting cells with Blasticidin. Image is representative of 3 independent repeats done in duplicates. Numbers represent number of colonies in the plate (c). * p<0.05, **p<0.01,****p<0.0001, one-way ANOVA with Tukey's multiple correction (b).*

3.3.3 HCV upregulates L1 retrotransposition potentially by inhibition of DNA damage repair pathways

L1 retrotransposition takes place via a process commonly known as Target Prime Reverse Transcription (TPRT) and less often by EN independent ways (described in chapter 1, section 1.5.1). Since L1 retrotransposition induces genomic instability, it is regulated by several host factors at various levels (refer to chapter 1, section 1.6) including DNA damage response pathways. There are several reports demonstrating effect of DNA repair factors on L1 retrotransposition but observations are inconsistent between studies for example, activating as well as inhibitory role of ATM on L1 retrotransposition has been reported (Gasior *et al.*, 2006; Coufal *et al.*, 2011). Recently, a systematic analysis of host factors affecting L1 mobility was carried out by whole genome siRNA screen and identified the double-stranded break (DSB) repair especially BRCA1 homologous recombination repair (HRR) pathway and Fanconi anemia (FA) factors as potent inhibitors of L1 activity in HeLa cells (Mita *et al.*, 2020).

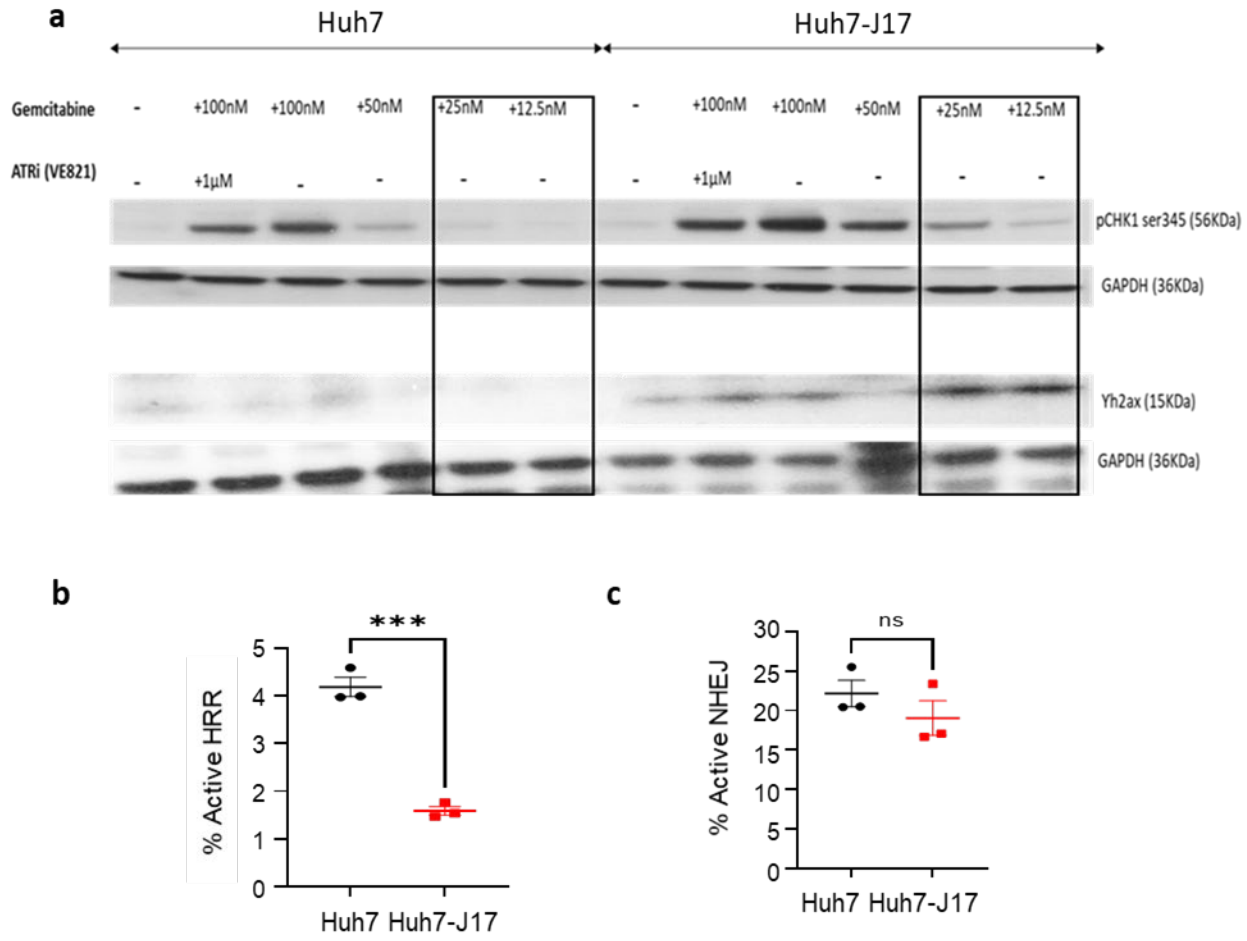

Figure 3.5 *HCV impairs DNA damage repair pathways in Huh7 cells*

*Western blot image of comparison of DNA damage in Huh7 wt vs Huh7-J17 cell lines. A representative image of two independent experiments is shown here. Cells were treated with decreasing doses of Gemcitabine from 100nM to 12.5nM. The lanes highlighted in rectangle shows clear upregulation of pCHK1 induction in Huh7-J17 cells. Whole x-ray film scan is provided in Appendix FigS1 (a). Graphs representing percentage of GFP positive cells representing active HRR (b) and NHEJ (c) repair in the indicated cell lines.*** p<0.001, ns = non-significant, unpaired t test.*

Since HCV is known to impair DNA repair mechanisms, we compared the sensitivity of Huh7 and Huh7-J17 cells in response to the potent DNA damaging agent i.e. Gemcitabine. As evident from western blot image (Fig 3.5a), more accumulation of γH2AX (a molecular marker of DNA damage and repair) was observed with Gemcitabine treatment in Huh7 cells in presence of HCV compared to the parental cells. ATR/CHK1 DNA damage response pathway is constitutively more active to repair the DNA damage induced by the HCV infection as evident by increase in pCHK1 at ser345 (Fig 3.5a compare 100nMGem with and without ATRi).

This was further confirmed by using plasmid-based reporter assays for HRR and NHEJ pathways in Huh7 and Huh7-J17 cells. The reporter assay revealed significant downregulation of the HRR pathway in Huh7-J17 cells compared to Huh7 parental cells (Fig 3.5b). While no significant influence of HCV replicon was observed on NHEJ pathway (Fig 3.5c). Hence, we speculated HRR pathway may also influence the regulation of L1 retrotransposition in Huh7 cells (Sudhindar *et al.*, 2021).

We therefore examined the effect of blocking DNA damage response enzymes on L1 retrotransposition frequencies in Huh7 cell line using small molecule inhibitors that target ATM (KU- 55933, 10µM), DNAPK (NU-7441, 1µM), ATR (VE-821, 1µM), CHK1 (SRA737, 1µM) and PARP (Rucaparib, 10µM) (doses selected based on literature search). A significant increase in L1 retrotransposition efficiency upon inhibition of ATR, CHK1 and PARP was observed (Fig 3.6a and b, effectiveness of ATM, ATR and CHK1 inhibitors is shown in Fig 3.6c-e, effectiveness of DNA-PKi and PARPi were tested by Hannah Smith, PhD student under Prof. Nicola Curtin, Newcastle University Centre for Cancer, data not shown here).

PARP is known to get activated by either SSBs or DSBs further activating the DNA damage response by the cells (Ronson G.E. *et al.*, 2018). Inhibition of PARP activity increased L1 retrotransposition in Huh7 cells indicating that PARP1 might play a role in regulation of the single strand breaks generated directly by L1-ORF2 endonuclease during TPRT through Base Excision Repair (BER) pathway (Ronson G.E. *et al.*, 2018). In addition, ATR and CHK1 are known to play an important role in the maintenance of DNA integrity in the face of DNA damaging insults principally through their involvement in the HRR as well as cell cycle checkpoints (Rundle *et al.*, 2017). Regulation of L1 retrotransposition by ATR and Chk1 in Huh7 cells proved that L1s can contribute to genomic instability through HRR pathway.

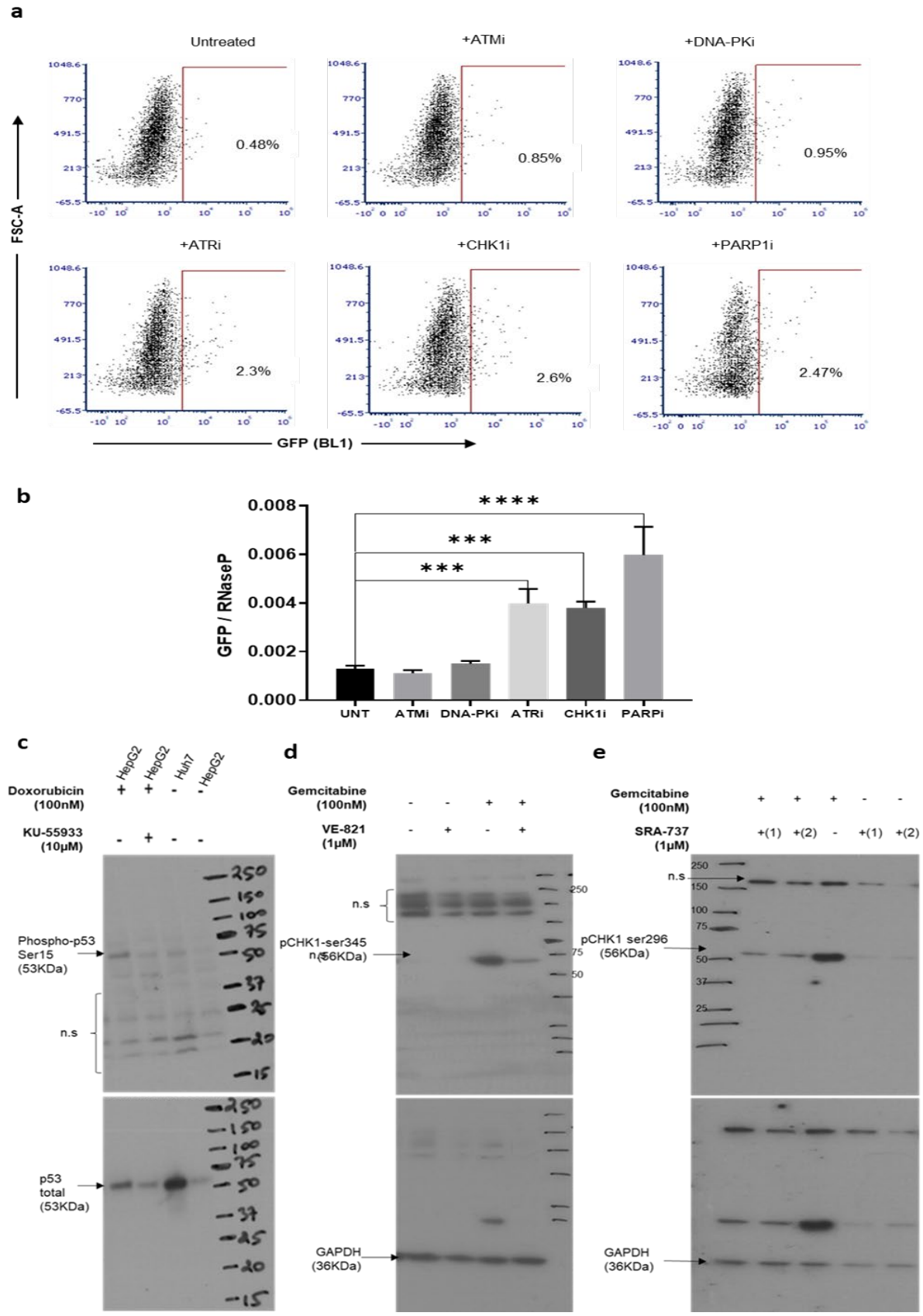

Figure 3.6 *Regulation of active retrotransposition in Huh7 cells by DDR response pathways*

*Numbers on the dotplots represent % EGFP positive cells (representative of cells undergoing active retrotransposition). Untreated control represents cells treated only with DMSO (1 in 1000 dilution in cell line media) (a), graph represents quantification of GFP insertions in the genome using RNAseP as a control, n=3 technical replicates (b). ***p<0.001, one-way ANOVA with multiple comparisons. Since Huh7 cells are mutant for p53, DNA damage was induced by Doxorubicin treatment (100nM) in HepG2 cells (wild type p53 cell line) in the presence or absence of ATMi (KU-55933) and cell lysates were prepared 24 hrs later to evaluate phosphor-p53 ser15. Bottom panel shows the membrane reprobed for total p53 after stripping. Untreated Huh7 cells lysate was used as a positive control for total p53; accumulated p53 in Huh7 cells confirmed mutant p53 status (c). DNA damage was induced by gemcitabine treatment (100nM) in the presence or absence of indicated inhibitors in Huh7 cells and cell lysates were prepared 4 hrs after the induction to evaluate levels of pCHK1 as a surrogate for the ATR-Chk1 pathway activation. Bottom panel shows the membranes reprobed for GAPDH without stripping to assess protein loading. n.s.=nonspecific bands. 1 and 2 corresponds to biological replicate lysates (d and e).*

3.3.4 Influences of L1 retrotransposition and active L1 retrotransposition continues even after viral clearance

A number of genes are reported to be dysregulated by HCV via epigenetic mechanisms and the dysregulation continues beyond viral clearance (Hamdane *et al.*, 2019; Perez *et al.*, 2019). Likewise, influence of L1-mediated somatic mutagenesis will continue in a cell beyond viral clearance. Moreover, activated L1 elements may continue to retrotranspose contributing towards genetic instability. To answer whether active viral infection is essential to the upregulated retrotransposition rate we observed *in vitro*, we treated Huh7-J17 cells with PSI7977 (Sofosbuvir, a NS5B polymerase inhibitor that inhibits virus replication). As shown in Fig 3.7a, PSI7977 treatment exhibited dose-dependent decrease in HCV replicon levels in Huh7-J17 cells as judged by overall luciferase activity of the cells. 10µM dose was selected and Huh7-J17 cell line was treated with 10µM of PSI7977 in absence of puromycin for 3 weeks to mimic DAA treatment and obtain a cell line that is clear of the virus to be used as post-HCV clearance model (Huh7-J17+PSI7977). Loss of the HCV replicon from the cells was confirmed by checking the puromycin sensitivity of the cells post PSI7977 treatment. As expected, 100% cell death was observed in Huh7-J17+PSI7977 cells upon puromycin treatment (2µg/ml for 48hrs) and no puromycin resistance gene transcript was observed by RT-qPCR indicating complete loss of the HCV replicon (Fig 3.7b). In addition, no NS5A was detected in Huh7-J17+PSI7977 whole cell lysate confirming the clearance of HCV replicon from the cells (Fig 3.7c) (Sudhindar *et al.*, 2021).

Next, EGFP- and blasticidin- based retrotransposition assays were carried out in Huh7-J17+PSI7977 using Huh7 as control. The level of active retrotransposition remained upregulated in Huh7-J17+PSI7977 cells compared to Huh7 cells (fold change 2.75 ± 0.2, n= 2, Fig 3.7d and e) thus indicating the influence of HCV on L1 retrotransposition includes a 'hit-and-run' mechanism via pathway(s) which remain dysregulated in cells even after viral clearance. However, the HRR pathway of Huh7-J17+PSI7977 cells restored to parental Huh7 cells level (Fig 3.7f) indicating that the stress induced by HCV and dysregulation of DNA damage repair pathways return to steady state level upon viral clearance (The BER pathway changes between parental Huh7, Huh7-J17 and Huh7-J17+PSI7977 cell lines could not be tested in this project due to time constraint). Hence, the underlying mechanisms regulating active retrotransposition are different in the presence of active HCV infection versus post-viral clearance and warrants further investigation (Sudhindar *et al.*, 2021).

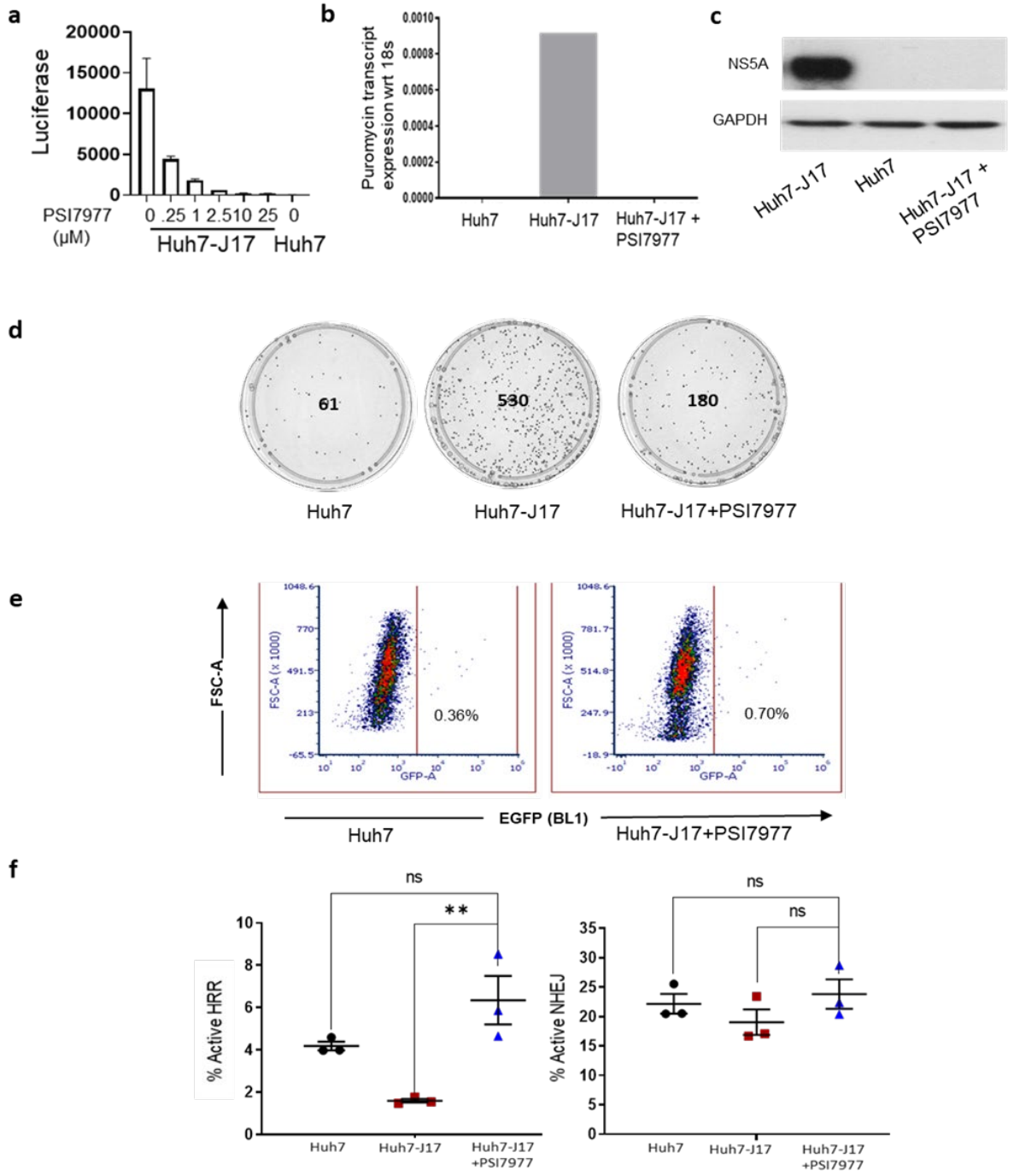

Figure 3.*7 Influence of HCV on L1 retrotranspositon continues even after viral clearance*

Luciferase expression levels of Huh7-J17 cells upon treatment with indicated doses of PSI7977 for 48hrs. Parental Huh7 cells was used as negative control. Error bars represent mean ± SE of 2 technical repeats (a). Graph shows puromycin transcript expression of Huh7, Huh7-J17 and Huh7-J17+PSI7977 cell lines measured by Real time PCR. 18s was used as the house keeping control (b). Western blot image of whole cell lysates of indicated cells to confirm clearance of HCV replicon 3 weeks after treatment with 10μM PSI7977 using anti-NS5A antibody. GAPDH was used as a loading control (c). Plates showing retrotransposition rates in indicated cell lines as assessed by Blasticidin-based system by resistant colonies. Image is representative of 2 independent repeats. Numbers represent number of colonies in the plate (d). EGFP-based

*retrotransposition assay in indicated cells analysed by FACS. Cells were transfected with UBGFPLRE3 plasmid and analysed 5 days after transfection (e). Graphs representing percentage of GFP positive cells representing active HRR and NHEJ repair in the indicated cell lines (f).** p<0.01, ns = non-significant, One way Anova.*

3.3.5 Role of DNA damage response pathway in modulating L1 retrotransposition in Liver cancer cell lines

Similar to Huh7 cells the retrotransposition frequency was identified in a panel of liver cancer cell lines (HepG2, Hep3B, PLC-PFR/5 and SK-Hep1) using EGFP-based retrotransposition assay. As shown in Fig 3.8a, all the cell lines supported active retrotransposition even if they exhibit varying levels of endogenous L1-ORF1p expression (Fig 3.8b).

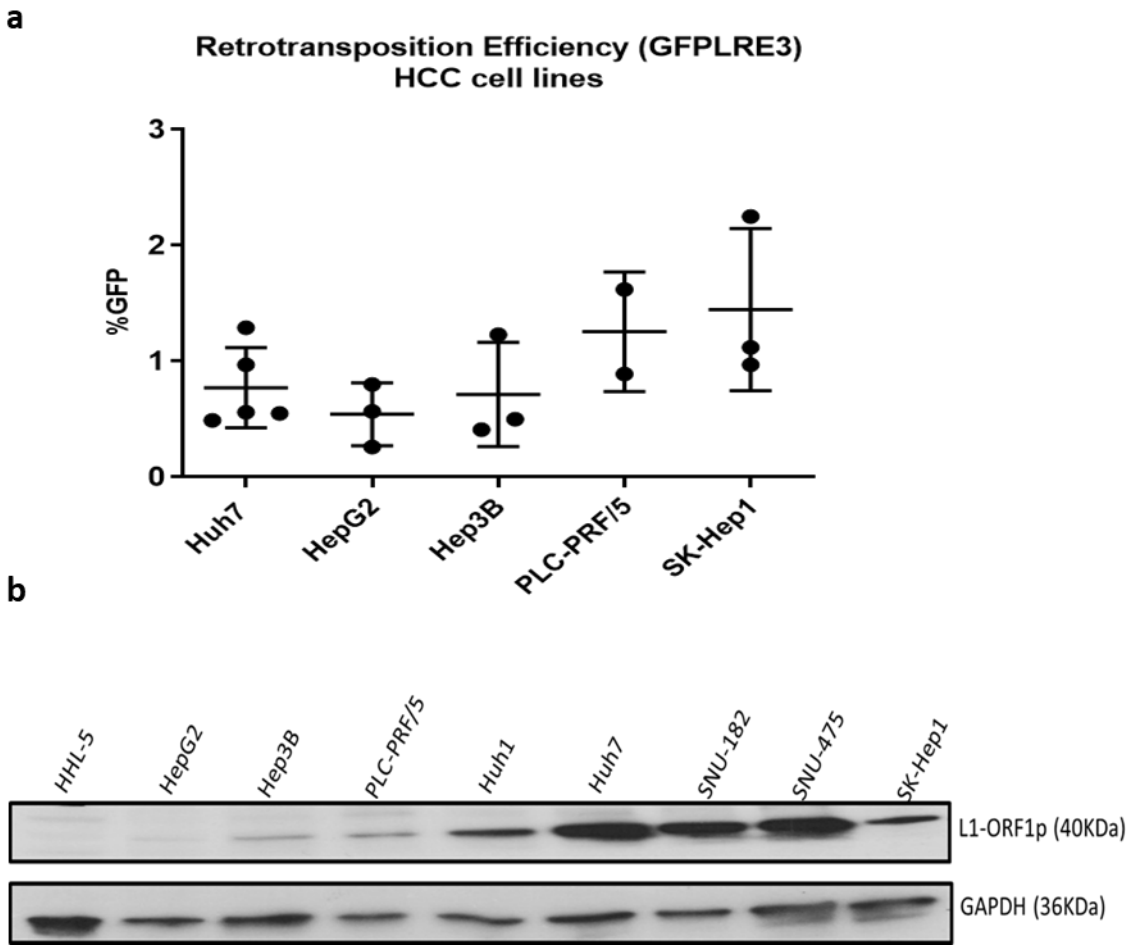

Figure 3.8 *Retrotranspositon efficiency and L1 expression in HCC cell lines*

Shown are the mean frequencies of GFP positive cells across different cell lines with standard deviation represented by error bars (a). Western blot analysis of total protein lysates from the indicated liver cancer cell lines using antibodies against L1 ORF1p to determine the endogenous ORF1p expression. GAPDH was used as the loading control (b).

Next, we examined if L1 retrotransposition is regulated by DNA damage response enzymes in other liver cancer related cell lines similar to Huh7 cells. For this purpose, PLC-PRF/5 (containing a p53 mutation at codon 249, G:C-T:A, with substitution of serine to arginine (Hsu I.C. *et al.*, 1993)) and SK-Hep1 (TP53 wild type) were chosen and the effect of blocking DNA damage response enzymes on L1 retrotransposition frequencies using small molecule inhibitors as indicated in section 3.3.3 were tested. In PLC-PRF/5 cell line, increase in retrotransposition efficiency was observed upon inhibition of DNA-PK, ATR and CHK1 (Fig 3.9a) whereas in SK-Hep1 cells increase in retrotransposition was observed only upon inhibition of DNA-PK (Fig 3.9b). Hence, L1 retrotranspostion is regulated by DNA damage response pathways in all the cell lines but different pathways are implicated in different cell lines.

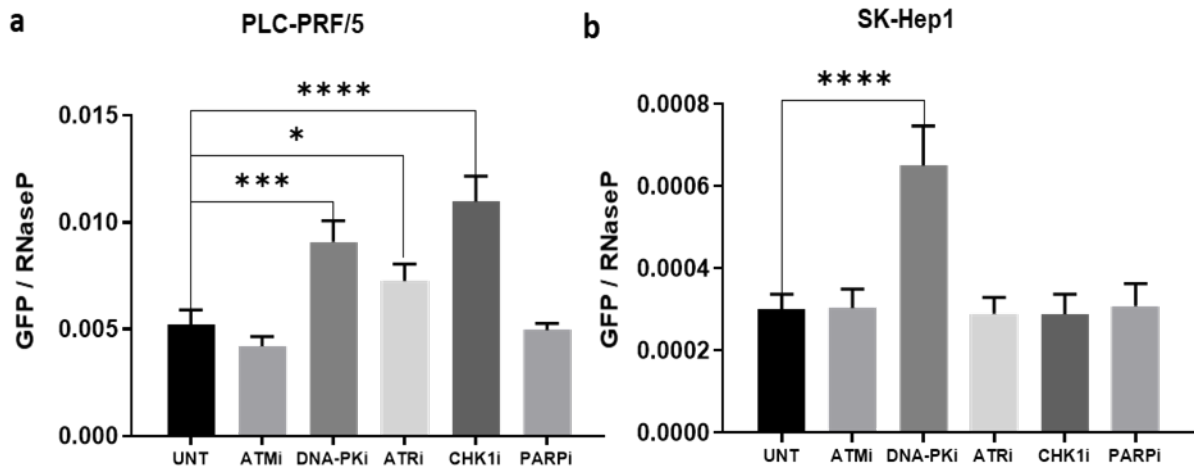

Figure 3.9 *Effect of DNA damage response inhibitors on L1 retrotransposition*

*Assayed by Taqman genomic DNA qPCR in, PLC-PRF/5 (a) and SK-Hep1 (b) cell lines. Error bars indicate mean±SEM of three technical repeats. ****p value <0.0001, ***p value <0.001, *p value <0.1. One-way Anova with Dunnets multiple comparisons test.*

To check the influence of the DNA damage response inhibitors on cell growth in order to identify which of these pathways are essential for cell cycle regulation, we performed a cell proliferation assay by SRB staining. Huh7 and PLC-PRF/5 cell lines were more sensitive to CHK1i and ATMi at Day 5 timepoint (Fig 3.10a and b) although ATR and CHK1 inhibition influenced L1 retrotransposition (Fig 3.6b and 3.9a). Likewise, SK-Hep1 cells were more

sensitive to CHK1i and ATMi as well (Fig 3.10c), although L1 retrotransposition was found to be regulated via DNA-PK mediated mechanism (Fig 3.9b). Hence, L1 retrotransposition is regulated independent of cell cycle regulation.

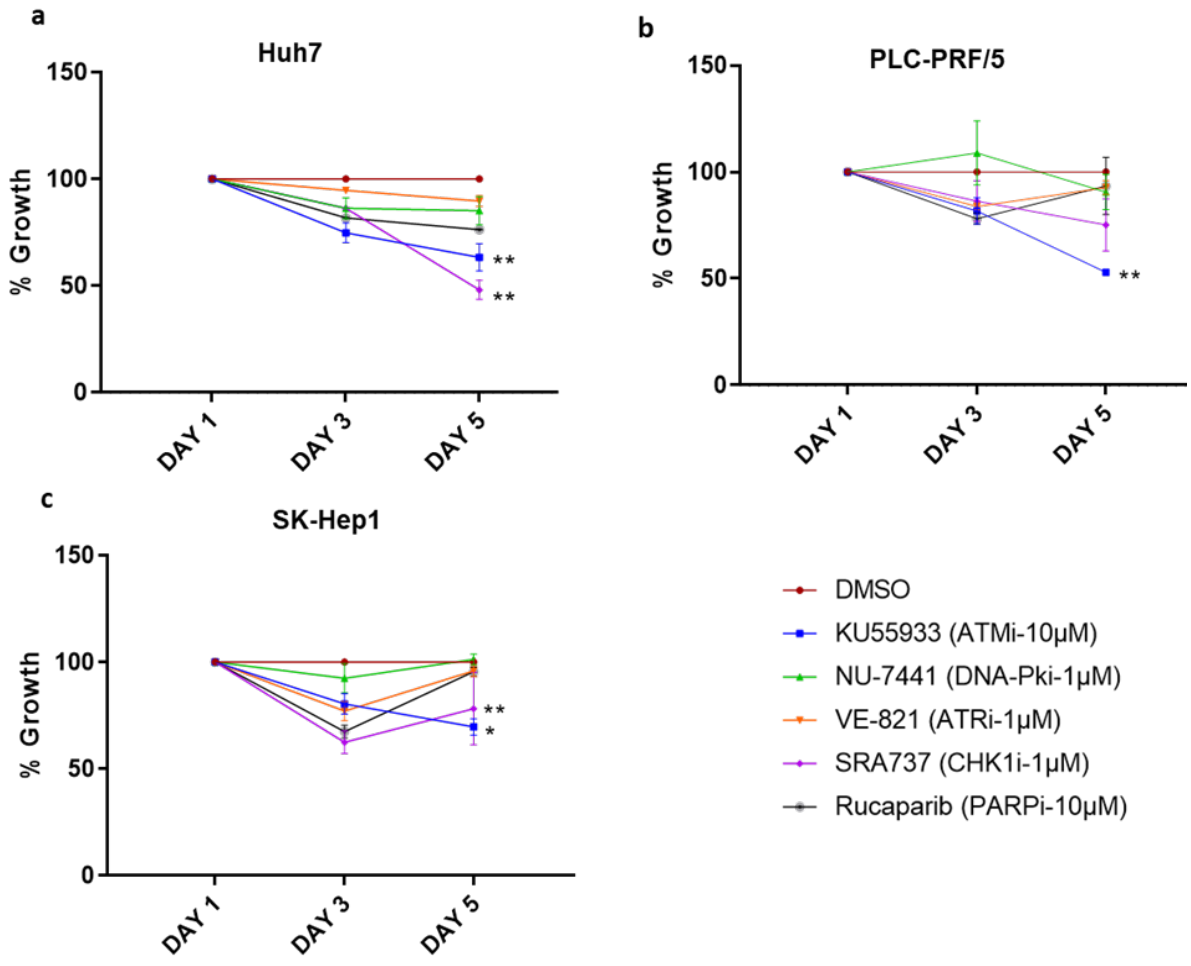

Figure 3.10 *Effect of DNA damage response inhibitors on cell proliferation*

Proliferation of Huh7 (a), PLC-PRF/5 (b) and SK-Hep1 (c) cells treated with DNA damage response inhibitors at indicated doses were measured by SRB staining. Growth percentage calculated by taking DMSO treated cells as 100% (baseline). Error bars indicate mean±SD of two independent experiments.

3.4 Discussion

Our data provide preliminary evidence that L1's are activated before oncogenic transformation in CHC patients, with HCV activated retrotransposition potentially mediating mutagenic consequences leading to HCC development. In addition, we proved that the rate of retrotransposition remains enhanced even after the viral clearance compared to the parental cells with no viral infection. Upregulation of active retrotransposition in cells with HCV infection, could possibly be due to increased DNA damage exerted by HCV infection, with impaired DDR pathways as potential mechanisms.

Upon DNA damage response suppression with small molecule inhibitors, we observed an increase in the retrotransposition efficiency of engineered human L1 in p53 mutant liver cancer cell lines (Huh7 and PLC-PRF/5) particularly upon inhibition of Ataxia Telangiectasia And Rad3-Related Protein (ATR), a serine/threonine kinase involved in DNA replication stress and DNA damage signalling. Together, these data suggest that cellular proteins involved in the DNA damage response, especially the ATR pathway, modulate L1 retrotransposition in liver cancer.

A recent study by Schobel et al observed a restriction of L1 retrotransposition in Huh7 cells in presence of HCV infection and also in Huh7.5 cells carrying a mutation (Thr-55-Iso) in the RIG-1 (DDX58) gene, that was attributed to sequestration of L1ORF1p in HCV-induced stress granules (Schobel *et al.*, 2021). This is in contrast to our observation in this study. But more in line with our own study, they observed an increase in L1 retrotransposition on overexpression of the HCV core protein, with an increase of L1orf1p levels in the cells in the presence of HCV. With both our studies indicating L1s as a source of genomic instability and that cells have evolved several different mechanisms to keep these elements in check, the discrepancies in the rate of L1 retrotransposition observed could be due to the model systems utilized (Goodier, 2016). HCV differentially modulates several L1 inhibitory pathways including activation of autophagy and interferon response pathways, while it impairs DNA damage repair pathways, especially HRR. Hence, there is a delicate balance between these various pathways which determines the success or failure of active retrotransposition in a cell. The viral load can be a major determinant of the extent of dysregulation of these processes and final outcome of the altered rate of active retrotransposition. Besides HCV,

other human oncogenic viruses have been shown to influence L1 retrotransposition- like Kaposi's sarcoma (KS)-associated herpesvirus (KSHV) infection in human patients has been shown to activate L1 retrotransposition (Nakayama *et al.*, 2019), and with contradictory reports about influence of Human immunodeficiency virus (HIV) (Iijima *et al.*, 2013; Jones *et al.*, 2013; Kawano *et al.*, 2018).

Our current study extends the previous work by Shukla et al (Shukla *et al.*, 2013) where the authors have demonstrated active L1 retrotransposition in HCC cases with HCV infection by indicating the possibility of HCV mediated active L1 retrotransposition in preneoplastic liver. Hence suggesting another underlying cause of hepatocarcinogenesis in CHC patients. HCC is a stepwise process and involves a combination of alterations. Since L1 mediated potential driver mutations have been implicated in various cancer types (Rodriguez-Martin *et al.*, 2020), it is worth considering a combination therapy of DAAs with retrotransposition inhibitory drugs such as antiretroviral drugs (reverse transcriptase inhibitors). CHC patients were treated with IFN therapy and type 1 interferons before the DAA therapy was introduced. Type 1 interferons are known to play a role in restricting L1 retrotransposition (Yu *et al.*, 2015). However, influence of DAAs on L1 activity is not known and worth investigating further. In addition, comparison of frequency of L1 mediated genomic rearrangements in HCV-HCC developed in IFN therapy patients versus DAAs therapy patients can shed light further on the role of L1 in HCC development in CHC patients. L1 retrotransposition can be inhibited by anti-retroviral drugs (Jones *et al.*, 2008; Dai *et al.*, 2011). A big advantage in developing this approach is that the drugs are already clinically available such as Nucleoside reverse transcriptase inhibitors (NRTIs) eg. lamivudine for HBV and HIV therapy and Non-nucleoside reverse transcriptase inhibitors (NNRTIs) eg. efavirenz for HIV treatment (Maeda *et al.*, 2019). Hence, it will be worth interrogating the effect of these drugs on the rate of HCC development in patients co-infected with HBV-HCV (1-15% of world population (Mavilia and Wu, 2018)) and HIV-HCV (~6.2% HIV infected individuals (Platt *et al.*, 2016)) and evaluate if the HCC developed in these patients is any different from only HCV-associated HCC, especially in terms of L1 mediated genomic rearrangements.

3.5 Limitations

Immunohistochemical staining for L1-ORF1p on CHC patients liver biopsies were carried out in only 11 cases and further paired samples of pre-HCC liver- and subsequent HCC biopsy were available for only 2 patients, hence the sample size is very low. This is partly because diagnostic biopsy carries significant risks (haemorrhage, tumour seeding) and is avoided. Despite this potential limitation, liver tissue analysis strongly supported our hypothesis and proved activation of L1 elements in the liver of CHC patients before HCC development.

The Huh7-J17 in-vitro model system used in this study is an artificial system where the HCV replication is mimicked using a plasmid based system. Thus, these findings need to be validated in other in vitro systems where cells are directly infected with HCV infection particles such as by using primary hepatocytes or stem cell derived hepatocytes. In addition, studies involving HCV mouse models should be carried out to validate the findings and interrogate the influence of active L1s on HCC progression

Concluding remarks- Our data clearly indicates that L1 activation is common in HCV related HCC. The tissue study clearly helps in understanding the relationship between L1 activation and cancer development and ultimately our hope is that this work can be combined with other molecular markers to identify a cancer predictive signature. Identification of specific active L1 loci will provide further power instead of looking at global L1 activation. Hence we tried to address these points in the next 2 chapters.

Chapter 4: Oncogenic potential of active retrotransposition

4.1 Introduction

Although a majority of the L1 mobilisation occurs in the germline and are presumed to be silenced in adult somatic cells by epigenetic mechanisms, there are several studies which provide evidence of somatic de-novo L1 insertions in mature adult cells. Muotri et al in 2005 were first to show somatic L1 retrotransposition of an engineered human L1 element in mouse brain although they could not demonstrate naturally occuring endogenous L1 retroransposition events in neuronal cells at that time (Muotri A.R. *et al.*, 2005). It was later in 2011 that Baillie et al identified several somatic endogenous L1 insertions by high throughput genomic DNA sequencing in post-mortem human brain cells (Baillie J.K. *et al.*, 2011). Furthermore, Upton et al in 2015 provided preliminary evidence that the L1 insertions in normal neurons were found in protein coding genes showing L1 driven mosaicism (Upton K.R. *et al.*, 2015). L1 insertions were also reported in epithelial somatic tumours and it was Miki et al in 1992 (Miki Y. *et al.*, 1992) that first identified a somatic L1 insertion that disrupted the APC gene in a colorectal cancer (CRC). There were several studies after that which demonstrated potential L1 driver candidates in human epithelial tumours although it was unclear whether these somatic L1 insertions had a causal role in initiation and progression of tumorigenesis (Lee E. *et al.*, 2012; Shukla R. *et al.*, 2013b; Helman E. *et al.*, 2014; Tubio J.M.C. *et al.*, 2014). Ewing et al in 2015 identified somatic L1 insertions in precancerous colonic adenomas using L1-Seq indicating that somatic retrotransposition occurs in very early stages in development of GI tumours (Ewing A.D. *et al.*, 2015). Furthermore, Scott et al in 2016 demostrated by whole genome sequencing that a hot L1 element escaped somatic silencing and mutated an APC allele which along with a existing mutation in second APC allele acted as a driver in initiation of colorectal cancer through classical route of colorectal cancer progression in a human CRC patient (Scott E.C. *et al.*, 2016). Shukla et al have demonstrated active retrotransposition in the non-tumour liver tissue of one HBV-HCC case (Shukla *et al.*, 2013) showing that although not as frequent as in tumour tissues, active retrotransposition can occur in non-tumour tissues (at least in presence of HBV infection). In chapter 3 of this thesis, we have shown that both L1 expression and retrotransposition can be activated in a chronically diseased HCV infected pre-neoplastic liver; however, direct evidence of L1

mediated driver mutations in HCC is still lacking. Thus in this chapter, we **aimed** to develop a cell line model of active L1 retrotransposition in order to check its potential for and frequency of cellular transformation using immortalised human hepatocytes.

4.2 Materials and Methods

4.2.1 Generation of HHL-5-RTN cell lines

A retrotransposition assay was set up in HHL-5 cells using an L1 GFP reporter construct (GFPLRE3), as described in section 2.3.1. Five days after transfection the cells were FACS sorted to select a GFP positive population (cells that have undergone a cycle of retrotransposition) and seeded onto 96 well plates at a density of 1, 50 or 200 cells/well. The plates were then incubated until wells became confluent with cells and then the expanded cell populations were further propagated individually as individual HHL-5-RTN cell lines.

4.2.2 Generation of HHL-5-RTNBlast^{res} cell lines

HHL-5 cells were subjected to the blasticidin resistance-based retrotransposition assay as described in section 2.3.2. Clones that emerged out of the pJJ101/L1.3 plasmid transfected and blasticidin resistant population were harvested individually and propagated as HHL-5-RTNBlast^{res} clones.

4.2.3 Evaluation of cellular transformation

Indicated cell lines were tested for transformed properties *in vitro* by the following assays: cell proliferation in normal and low serum conditions (sec 2.6), colony formation (sec 2.8) and loss of contact inhibition (sec 4.2.4) using non-transfected cell lines as controls.

4.2.4 Contact inhibition assay

Cells were plated on 6 well plates as a 2D monolayer to achieve 70-80% confluence the next day and examined for increased monolayer density over a period of 3 weeks. Phase contrast images were captured on days 2, 5 and 18 to look for transformed foci showing evidence of loss of contact inhibition.

4.3 Results

4.3.1 Generation of HHL-5-RTN cell lines to test for cellular transformation potential of de novo L1 insertion events

A retrotransposition assay was set up with the HHL-5 cell line using GFP as a reporter system (as stated in Materials and Methods section 2.4.1). The live cells that expressed GFP signal were FACS sorted into a 96 well plate (see Materials and Methods section 4.2.1). No colonies developed from wells seeded with single cells. Three lines emerged from wells seeded with 200 cells and other 7 lines emerged from wells seeded with 50 cells. 10 RTN lines that emerged were propagated further and named as HHL-5-RTN1-10. Wild type HHL-5 cells were also subjected to FACS sorting in a similar manner and the lines which emerged out of wells seeded with 50 cells were taken as negative controls (HHL-5-WT1-3). Out of the 10 RTN lines only 4 were positive for GFP insertion (as shown in the agarose gel image fig 4.1). The RTN lines were further tested for potential tumorigenic properties using various *in vitro* assays as described below.

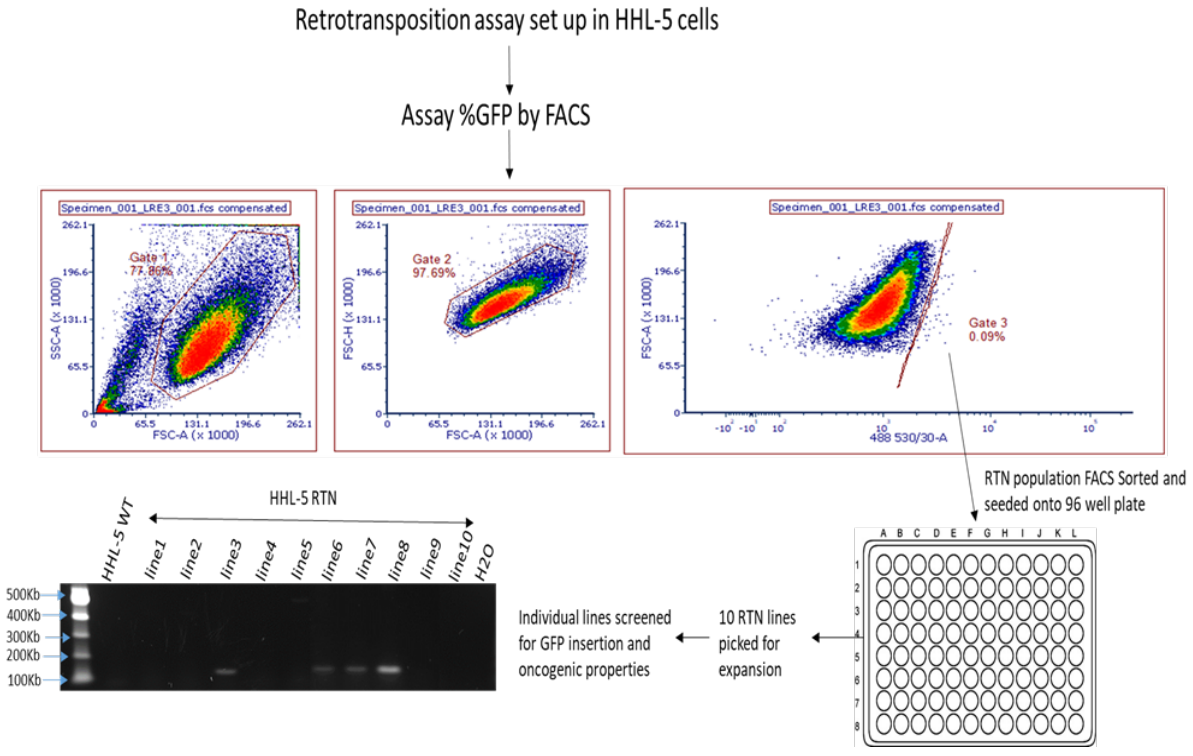

Figure 4.1 *Schematic representation of GFP RTN assay*

Geeneration of HHL-5 RTN cell lines, basis of gating, GFP primer and the expected PCR product size are discussed on Chapter 3- Fig 3.2.

4.3.2 HHL5-RTN-line6 shows cellular transformation properties

Cellular transformation may lead to increased growth rate and increased colony forming ability, hence the HHL-5-RTN lines were assessed for their growth rate compared to the parental HHL-5 cell line using Incucyte imaging in normal and low serum culture conditions. HHL-5-RTN-line6 showed an increased proliferation rate (~1.7-2 fold, Fig 4.2a) compared to all other RTN lines and WT cells in normal culture conditions whereas in low serum (2%) conditions HHL-5-RTN-line6 showed a decrease in proliferation rate (~1.8-2.4 fold, Fig 4.2b) compared to WT cells and other RTN line indicating that the growth factors present in serum are essential for HHL-5-RTN-line6 proliferation. However, this is contrary to what is expected of transformed cells (Rubin H., 2017). The growth rate of a subset of HHL-5-RTN lines and HHL-5-WT lines were also assessed using an SRB assay in normal culture conditions And the result corroborated with incucyte imaging (Fig 4.2c).

HHL-5 cells are immortalised human hepatocytes and exhibit very poor colony forming ability (approximately 2% plating efficiency (PE)). A marked increase in colony forming ability (26% PE, Fig 4.3) was observed for the HHL-5-RTN-line6.

Another important characteristic of cellular transformation is uncontrolled proliferation and loss of contact inhibition. Non-cancerous cells cease to proliferate when they contact each other, whereas cancer cells exhibit uncontrolled proliferation (Pavel M. *et al.*, 2018). The HHL-5-RTN-line6 showed marked loss of contact inhibition (Fig 4.4) compared to the WT and other RTN lines indicating a transformed phenotype.

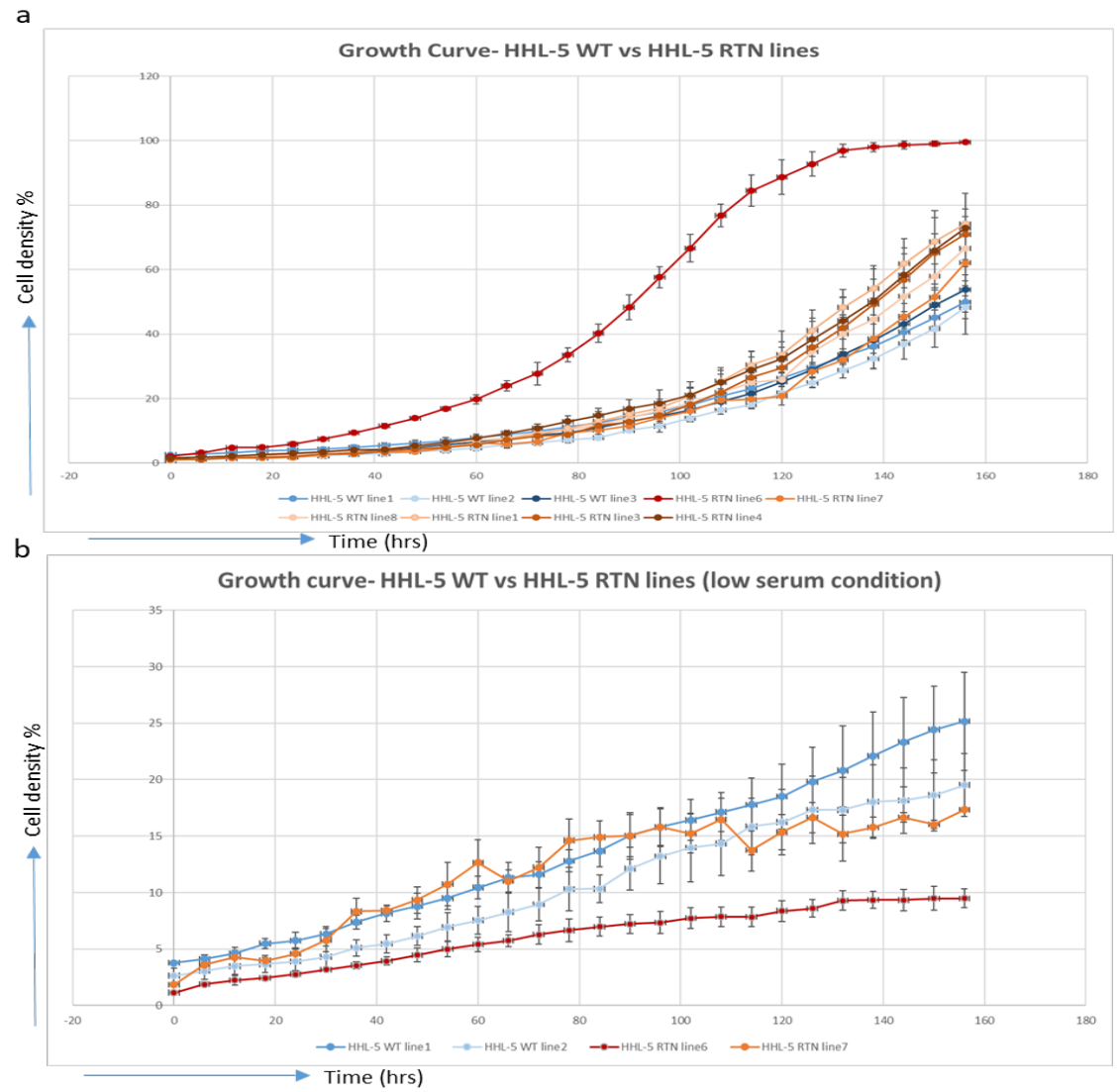

Figure 4.2 *Cell proliferation of HHL5 WT vs RTN lines*

Cell proliferation was measured using the Incucyte live cell analysis system with time-lapse phase contrast images captured every 4 hours for 6 days in complete growth media (a) and in low serum condition (b). Cell proliferation measured by SRB staining (c). Error bars indicate mean±SD of three technical repeats.

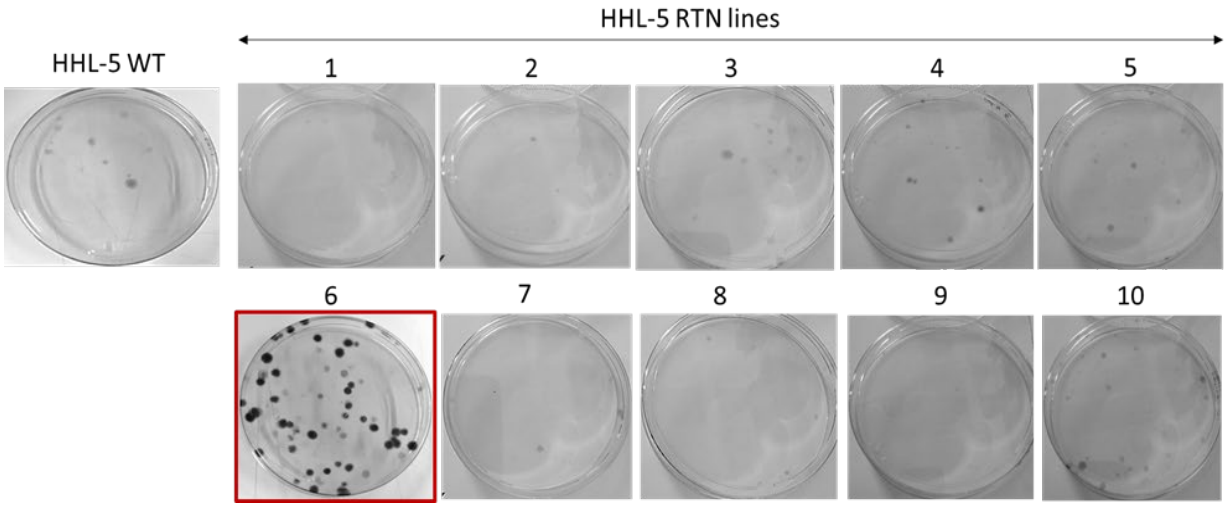

Figure 4.3 *Colony formation- HHL-5 WT vs RTN lines*

Digital images showing colonies produced by HHL-5 WT and the RTN lines following plating of 200 cells and 14 days of incubation.

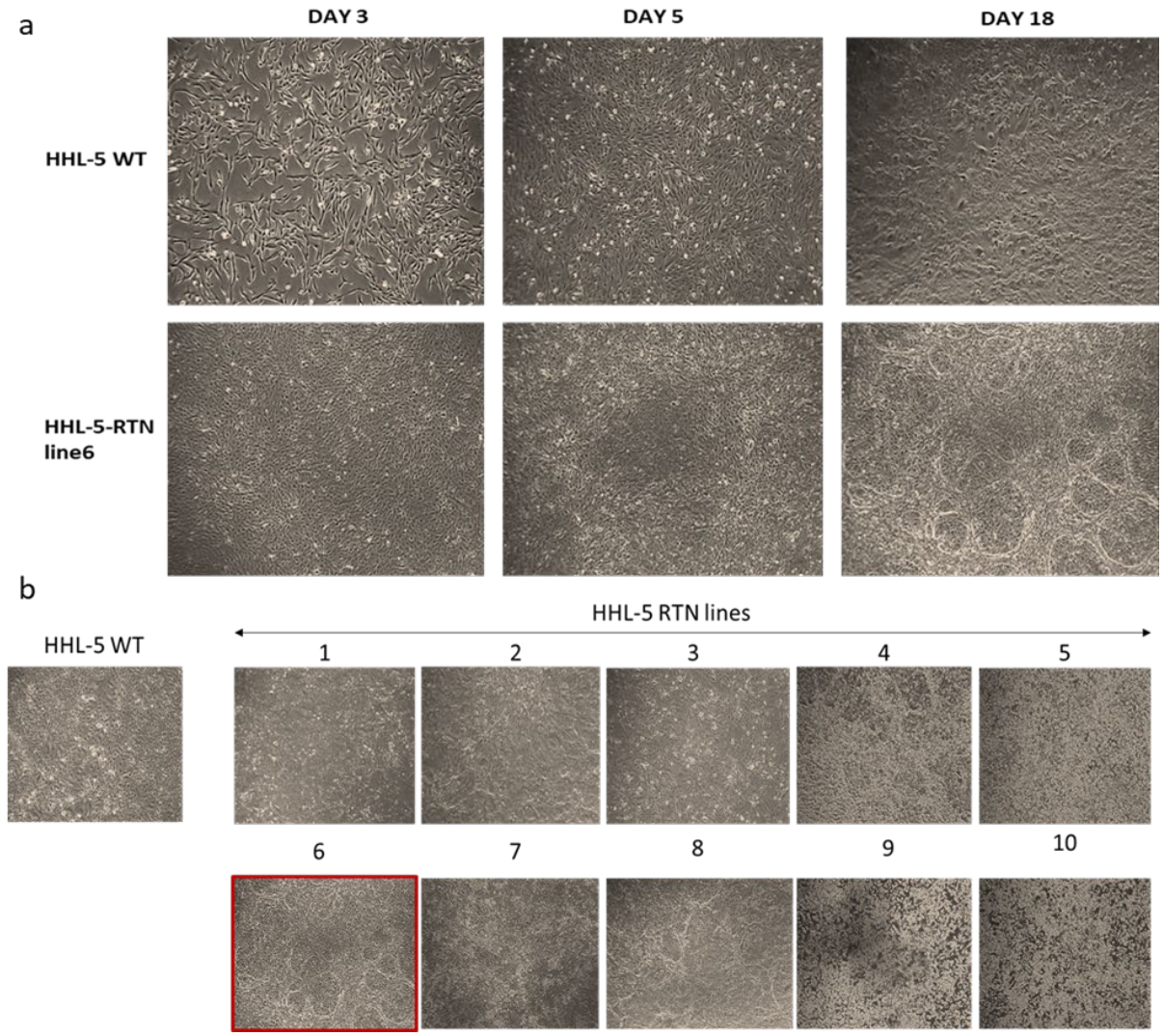

Figure 4.4 *Loss of contact inhibition- HHL-5 WT vs RTN lines*

Phase contrast images of HHL-5 WT and RTN-line6 taken at days 3, 5 and 18 after seeding on a 6 well plate to observe loss of contact inhibition (a) Day 18 phase contrast images of HHL-5 WT and RTN lines (b).

4.3.3 Generation of HHL-5-RTNBlast^{res} clones to test for cellular transformation potential of de novo L1 insertion events

In a second method for generating RTN clones we performed a retrotransposition assay with HHL-5 cells using Blasticidin as selectable reporter construct (as stated in Materials and Methods section 2.4.2). Three clones which emerged Blasticidin resistant were harvested individually using O rings from the 10cm dishes and the population of cells expanded further. The HHL-5-RTNBlast^{res} clones were then tested for transformation properties as described

previously for HHL-5-RTN lines. However, none of the 3 RTNBlast^{res} clones showed any phenotyic changes *in vitro* (Fig 4.5).

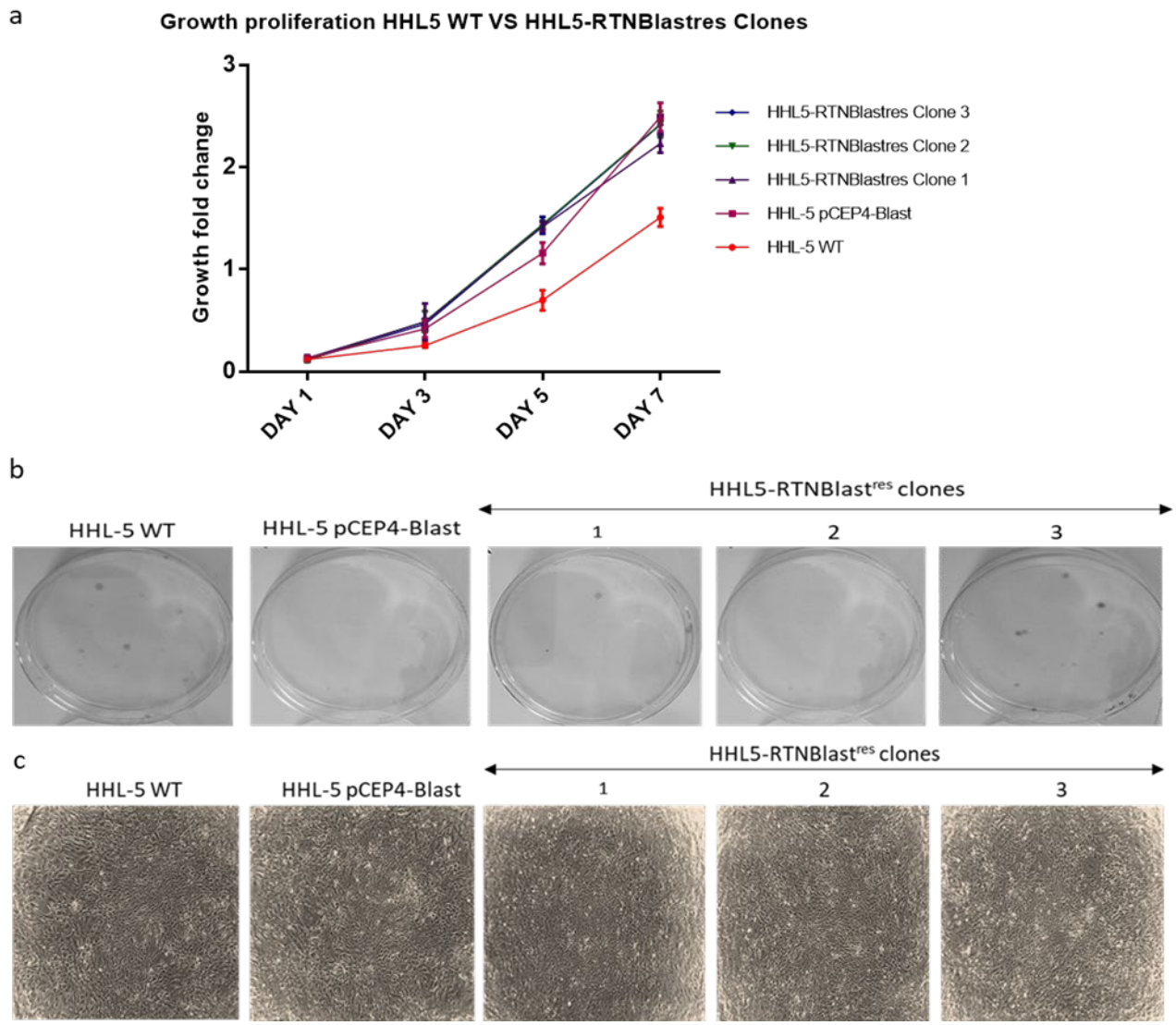

Figure 4.5 *Observation of cellular transformation-HHL-5 WT vs HHL-5-RTNBlast^{res} clones*

Proliferation of HHL-5 WT and HHL-5-RTNBlast^{res} clones measured by SRB staining. Error bars indicate mean±SD of five technical repeats (a). Digital images showing colonies produced by HHL-5 WT and HHL-5-RTNBlast^{res} clones following plating of 200 cells and 14 days of incubation (b). Phase contrast images of HHL-5 WT and HHL-5-RTNBlast^{res} clones taken 18 days after seeding on a 6 well plate to observe loss of contact inhibition (c).

4.4 Discussion

L1 insertion in the *ST18* gene reported by Shukla et al in 2013 was the first report of somatic L1 insertion leading to activation of the gene expression in hepatocellular carcinoma (Shukla R. *et al.*, 2013b). Other HCC specific L1 insertions are being reported in the same study and in a follow up study by the group. However, there is no evidence if any of these L1 insertions acted as a driver mutation and directly contributed in initiating the tumour or whether these events happened in later stages of tumour progression. We designed this study using the HHL-5 cell line to evaluate the potential of L1 retrotransposition to intiate cellular transformation and generate *in vitro* transformed cells, which might support a causal role of L1 retrotransposition in oncogenesis and potentially lead to the identification of novel oncogenes or tumour suppressor pathways. One RTN line out of 13 (GFP and Blasticidin based retrotransposition assays) showed evidence of cellular transformation as observed by *in vitro* assays (Fig 4.2, 4.3 and 4.4). However, to confirm the tumorogenic potential of these transformed cells they should be evaluated *in vivo* by injecting into an immunodeficient mouse model such as NOD-SCID or NSG using an orthotopic approach (Pez F. *et al.*, 2019). The orthotopic approach would allow the assessment of the environment surrounding the tumour which would give more informative results.

Insertion of L1 in a potential oncogene or tumour suppressor gene, where a clear role in tumorigenesis is established could lead to cellular transformation. However, a specific retrotransposition event and the exact point of L1 insertion in the HHL-5-RTN-line6 genome is not known; hence, the mechanism underlying the observed changes in cellular properties can not be predicted. Moreover, a possibility that the cells are spontaneously transformed can not be ruled out. To identify any de novo L1 inserton sites in the cells, an inverse PCR strategy using GFP specific primers, as described by Coufal et al (Coufal N.G. *et al.*, 2011) could be employed on the HHL-5-RTN-line6 genomic DNA. Overall this encouraging pilot study warrants a larger scale screening of RTN lines or clones to identify transformed cells with de novo L1 insertions to identify potential novel mechanisms of L1 retrotransposition driven hepatocytic transformation.

Chapter 5: Development of a CRISPR-Cas9-based system to selectively isolate L1 promoter bound proteins

5.1 Introduction

Nuclear DNA is packed as chromatin in eukaryotic cells with nucleosomes being the fundamental repeating unit consisting of about 150bp of DNA tightly wrapped around a protein octamer (Cheng Y. *et al.*, 2019). The octamer comprises of two copies each of four core histones H2A, H2B, H3 and H4 (Cheng Y. *et al.*, 2019). Eukaryotic gene expression is controlled by regulatory mechanisms that involve changes in chromatin structures that are mediated by interplay between DNA methylation and histone modification (Van Driel R. *et al.*, 2003). There are several methods developed to identify the biochemical changes in chromatin structure. Zhang et al, 1982 were the first to develop a purification method for satellite chromatin from mouse liver nuclei by nuclease digestion (Zhang X.Y. and Hörz W., 1982). An independent study by Workman et al, 1985 developed a nucleoprotein hybridisation technique to selectively isolate specific eukaryotic genes for structural and biochemical study (Workman J.L. and Langmore J.P., 1985). Boffa et al, 1995 developed a biotin labelled complementary peptide nucleic acid (PNA) to bind the CAG repeats in chromatin and employed a method to selectively isolate the chromatin fragments containing the PNA hybrids (Boffa L.C. *et al.*, 1995). Jasinskas et al, 1999 employed the nucleoprotein hybridisation technique to isolate primate satellite chromatin for the analysis of centromere enriched proteins (Jasinskas A. and Hamkalo B.A., 1999). Another study by Griesenbeck et al demonstrated a technique involving a recombinant adapter molecule containing LexA protein, Nuclear localisation signal and tandem affinity purification tag to specifically purify the *PHO5* promoter region from chromatin segments of the yeast *Saccharomyces cerevisiae* (Griesenbeck J. *et al.*, 2003). While all these studies achieved selective isolation of the targeted regions, none of them gave sufficient amount of chromatin and purity to identify the bound factors until a study by Déjardin et al, where they have used a DNA hybridisation probe to isolate the genomic DNA and proteins associated in sufficient quantity to retrieve the protein information. The protocol was described as proteomics of isolated chromatin segments (PICh) (Déjardin J. and Kingston R.E., 2009). They were able to specifically isolate the cross linked chromatin regions and identify the proteins bound to them using Mass Spectrometry (MS) (Déjardin J. and Kingston R.E., 2009). Hoshino et al in 2009 developed

Insertional Chromatin Immunoprecipitation (iChIP) technology to purify genomic regions of interest by immunoprecipitation with antibody against a tag which is fused to the DNA binding domain of the target DNA binding protein (Hoshino A. and Fujii H., 2009). The chromatin complexes isolated by this technique retained the factors interacting with genomic region of interest. Reverse cross linking of the chromatin regions allowed them to purify the DNA, RNA and protein for further characterisation (Hoshino A. and Fujii H., 2009). Although iChIP is a powerful tool, it requires insertion of the recognition sequences of the exogenous DNA binding molecule, which is considered to be a limitation of this technique. Fujita et al, 2013 developed a novel enChIP technique using "engineered DNA binding molecules like zinc-finger proteins, transcription activator-like (TAL) proteins, and the CRISPR (clustered regularly interspersed short palindromic repeats) system consisting of a catalytically inactive form of Cas9 endonuclease (dCas9) and small guide RNA (gRNA)" to tag a specific genomic locus (Fujita T. and Fujii H., 2013). The enChIP using CRISPR had a major advantage over the iChIP technique as it does not need insertion of exogenous DNA sequences (Fujita T. and Fujii H., 2013). Another recent advancement in this field is the dCas9-APEX2 biotinylation at genomic elements by restricted spatial tagging (C-BERST) which enables high throughput identification of proteins localised to the individual loci (Gao *et al.*, 2018).

HCC tumorigenesis is characterised by global hypomethylation, as a consequence of which the L1 elements get activated, since the L1 promoter is predominantly regulated epigenetically via DNA methylation. Studies in the past have identified several transcription factors interacting with the L1 promoter by employing different IP techniques (discussed in detail in the introduction chapter section 1.6). In this chapter we aimed to activate the endogenous L1 elements in immortalised hepatocyte cell line (HHL-5) by manipulating them epigenetically (global demethylation using 5-Aza-2'-Deoxycytidine treatment). Then we aimed to develop an enChIP model to selectively isolate the endogenous L1 loci and proteins bound to them in order to understand the regulation of L1 elements at the promoter level specifically in HCC related cell lines.

5.2 Materials and methods

5.2.1 Single guide RNA (sgRNA) design and cloning

The 3xFlag-Cas9 expression plasmid (pX330-U6-Chimeric BB-CBh-hSpCas9) was a gift from Feng Zhang (Addgene plasmid # 42230 ; http://n2t.net/addgene:42230; RRID:Addgene_42230) (Cong L. *et al.*, 2013). It was subjected to D10A/H840A substitution to create dCas9 plasmid and a 2xTy1 tag was added to the N-terminal of 3xFlag-dCas9 to produce the Ty1-Flag-dCas9 expression plasmid (*generated by Dr Ruchi Shukla*). Three Single Guide RNA's (sgRNA's) were designed to target specific regions of L1 promoter using the tool- http://crispr.mit.edu/CRISPR Design (L1gRNA1, L1gRNA2 and L1gRNA3) (see appendix table 1). L1gRNA1 and L1gRNA2 were cloned into the BbsI site of the Ty1-Flag-dCas9 plasmid to produce the dCas9-L1gRNA1 and dCas9-L1gRNA2 plasmids (*generated by Dr Ruchi Shukla*). L1gRNA3 was ordered as a complete gRNA expression cassette (gBlock) from Integrated DNA Technologies (IDT) (Fig 5.1) and was cloned into a pSup-puro vector using ECOR-I and Bgl-II restriction sites, thus replacing the H1 promoter region with the L1gRNA3 gBlock (Fig 5.2).

TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGGTACCAAGGTCGGGCAGGAAGAG

GGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTT

GACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTT

TTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATAT

ATCTTGTGGAAAGGACGAAACACCGTGCATTTCCATCTGAGGTACGTTTTAGAGCTAGAAATAGCAAGTTA

AAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTT

GTACAAAGTTGGCATTA

Figure 5.1 *L1gRNA3 (gBlock expression system)*

This 455bp fragment bears the components necessary for L1gRNA3 expression namely- U6 promoter + L1gRNA3 sequence + guide RNA scaffold + termination signal. L1gRNA3 expression system was synthesized as a complete expression cassette (gBlock) from IDT.

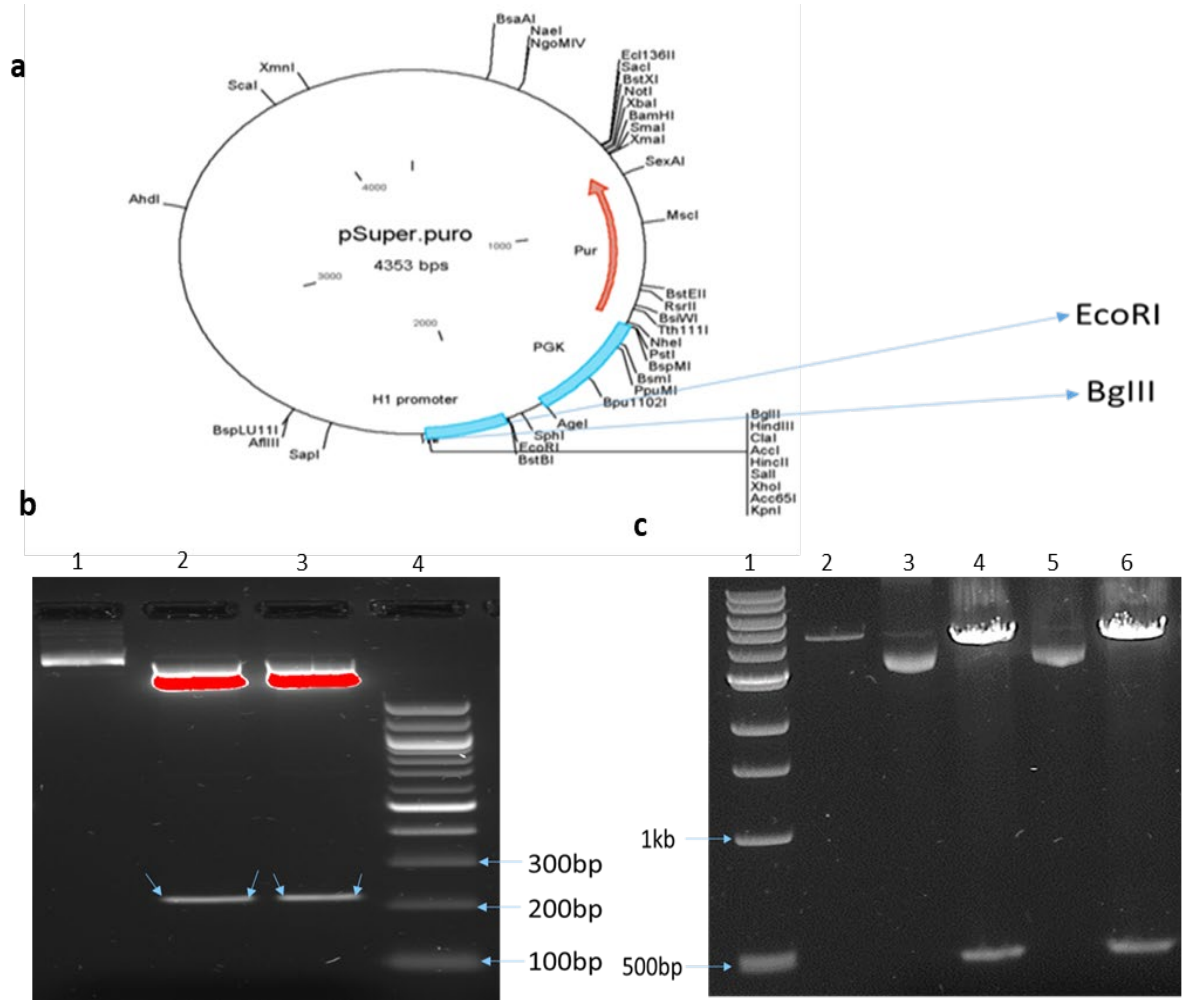

Figure 5.2 *Cloning L1gRNA3 into pSuper-puro vector*

Plasmid map of pSuper-puro vector (OligoEngine™). EcoRI and BglII sites highlighted (a). pSuper-puro vector was subjected to double enzymatic digestion and the digested products were run on agarose gel. 0.8% agarose gel image showing Lane 1- Undigested pSup-puro, Lane 2 and 3- pSuper-puro digested with EcoRI and BglII, Lane 4- 100bp DNA ladder (b). After the successful cloning of pSuper-puro vector with L1gRNA3 insert the digest of the clones were run on 0.8% agarose gel to verify the cloning of L1gRNA3 (500bp insert) into pSuper-puro vector. Lane 1-1kb DNA ladder, Lane 2- pSuper-puro vector positive control, Lane 3 and 5- Undigested DNA, Lane 4 and 6- Midiprep DNA subjected to EcoRI and BglII double enzymatic digestion (c).

5.2.1.1 Cloning L1gRNA3 into pSuper-puro vector

Firstly, the pSuper-puro vector was digested using EcoRI and BglII restriction enzymes (Fig 5.2a). L1gRNA3 was amplified on a thermocycler (L1gRNA3_EcoRI and L1gRNA3_BglII primers) using Platinum SuperFi green PCR master mix. The PCR conditions for amplification of L1gRNA3 were as follows- Initial denaturation for 30sec at 90°C, followed by 30 PCR cycles of 98°C for 10sec, 60°C for 20sec and 72°C for 30sec with a final extension step of 72°C for 5

minutes. 20µl of the PCR product was resolved on a 0.8% agarose gel and the ˜200bp products were excised and purified by Qiagen Minelute PCR purification kit (Cat#28004) (Fig 5.2b). DNA concentrations were measured by Nanodrop. The ligation reaction was setup for pSup-puro digest and L1gRNA3 PCR product using T4 DNA ligase (Promega M180A) followed by transformation into library efficient DH5α competent cells to get recombinant colonies. Individual colonies were picked, Miniprep DNA extracted and subjected to double enzymatic digestion with EcoRI and BglII enzymes. The clone which was positive for the ˜500bp insert was selected and purified using Qiagen Midiprep kit. Midiprep DNA was run on 0.8% agarose gel to verify the cloning of L1gRNA3 into pSuper-puro vector (Fig 5.2c).

All the plasmid vectors were amplified by transformation of Ecoli DH5α competent cells (Invitrogen) and then plasmids were purified using GenElute™ HP Plasmid Midiprep Kit (SIGMA-NA0200-1KT).

5.2.2 Anti-sense based identification of active L1-loci in Huh7 cells

To identify expressed L1 RNA transcripts, Rapid Amplification of cDNA ends (3' RACE) was carried out on 1µg of RNA using RACE primer (5'-GCGAGCACAGAATTAATACGACTCACTATAGGTTTTTTTTTTTT-3'). Then PCR was carried out on RACE primed cDNA input using primers Outer and L1_library (5'-GTGAGATGAACCCGGTACCTCAG-3') using High Fidelity Expand *Taq* (Roche) (Macia A. *et al.*, 2011). The PCR conditions included an initial cycle of 95°C for 2 min, followed by 30 cycles of 30 s at 94°C, 30 s at 54°C, and 90 s at 72°C, with a final step of 72°C for 10 min. Thirty microliters of the PCR product was resolved on 2% agarose gels, and products were excised and purified. DNA was extracted using the QIAquick extraction kit (Qiagen), and products were cloned in pGEMT-Easy (Promega). Approximately 50 clones were randomly sequenced using an SP6 primer and the genes associated with full length L1 transcripts and their chromosomal locations identified by BLAT alignment (http:www.genome.ucsc.edu) to the human reference sequence (Dec. 2013 GRch38/hg38 assembly) to identify a source L1 locus. These different active L1 loci were aligned and compared with the reference L1 sequence and a consensus sequence was identified using CLUSTAL-OMEGA. This was used to design sgRNA (L1gRNA3) sequences targeted at the maximum complementary sequence to different L1 transcripts.

5.2.3 Piggy-bac transposon system

A Piggy-bac transposon system was chosen for plasmid cloning of dCas9-L1gRNA, as it can be used to integrate a large amount of DNA into the host cell genome. This system consists of a Piggy-bac vector and a plasmid encoding a transposase enzyme that recognises the transposon specific inverted terminal repeats and efficiently integrates the DNA sequence in between the terminal repeats into the genome at TTAA chromosomal sites.

5.2.3.1 Cloning strategy of Piggy-bac vector with dCas9 insert

Firstly, the TY1-Flag-dCas9 expression plasmid was subjected to enzymatic digestion at 3 sites- AgeI, SacII and EcoRI (Fig 5.3a) and the digested products were resolved on an 0.8% agarose gel (Fig 5.3c). The ˜4.5kb product (containing Ty1-Flag-dCas9 coding region) was excised out of the gel and purified by QIAquick Gel extraction kit (Cat no.28704). Similarly, the Piggy-bac vector containing a doxycycline-inducible promoter and linked puromycin-resistance selection cassette was subjected to enzymatic digestion at the HpaI site to get a linear product (Piggy-bac plasmid map Fig 5.3b). The purified insert was modified to generate blunt DNA ends (Quick blunting kit) then a ligation reaction was set up with the HpaI-digested Piggy-bac vector using T4 DNA ligase (Promega M180A) followed by transformation into DH5α competent cells and streaked onto ampicillin containing LB agar plates to obtain recombinant colonies. Individual colonies were picked, and DNA purified using Qiagen Miniprep kit. The purified DNA products were subjected to diagnostic double enzymatic digestion at MluI and PstI sites and the products analysed on a 1% agarose gel for verification. Out of 11 colonies picked for DNA purification 2 colonies were positive for the dCas9 insert (Fig 5.3d). The piggy-bac colonies which were positive for the dCas9 insert were also sequence verified and finally a midiprep was carried out to amplify the plasmid (PB-TET-dCas9) for further experimentation.

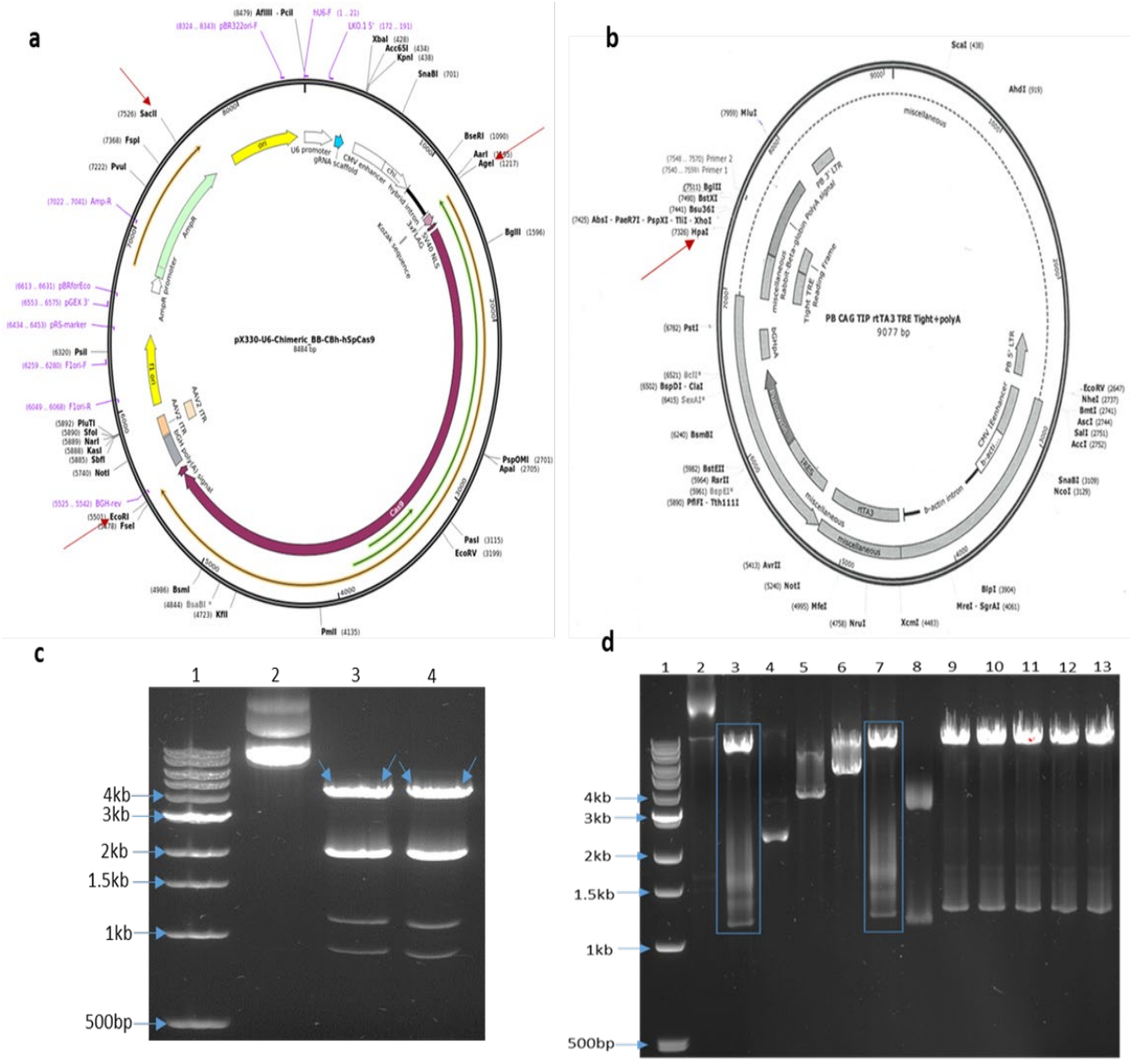

Figure 5.3 *Cloning strategy of Piggy-bac vector with dCas9 insert*

Plasmid map of TY1-Flag-dCas9 vector (Addgene). EcoRI, AgeI and SacII cut sites highlighted (a). Plasmid map of Piggy-bac vector (Addgene). HpaI cut site highlighted (b). 0.8% agarose gel image. Lane 1-10kb DNA ladder, Lane 2- Undigested TY1-Flag-dCas9, Lane 3 and 4- Midiprep TY1-Flag-dCas9 DNA subjected to EcoRI, AgeI and SacII enzymatic digestion (c). 1% agarose gel image. Lane 1- 10kb DNA ladder, Lane 2- Undigested Piggybac-TY1-Flag-dCas9. Lane 3-13-Miniprep Piggybac-TY1-Flag-dCas9 DNA subjected to MluI and PstI double enzymatic digestion. Lane 3 and 7- positive for the dCas9 insert (d).

5.2.3.2 Cloning strategy of PB-TET-dCas9 with L1gRNA insert

The developed PB-TET-dCas9 plasmid DNA was subjected to MluI enzymatic digestion, purified by Qiagen Minelute PCR purification kit (Cat no.28004) and dephosphorylated using an alkaline phosphatase.

L1gRNA was amplified by PCR using Platinum green PCR master mix with Hu6_Forward and Reverse gRNA scaffold primers and dCas9-L1gRNA2 plasmid as a template. The PCR conditions were as follows- Initial denaturation for 2 minutes at 94°C, followed by 30 PCR cycles of 94°C for 30sec, 60°C for 30sec and 72°C for 30sec. 20µl of the PCR product was resolved on a 1.5% agarose gel and the 450bp product was excised and purified by QIAquick Gel extraction kit (Cat no.28704). Purified product was subjected to MluI enzymatic digestion and phosphorylation (Fig 5.3b). The ligation reaction was setup for MluI digested PB-TET-dCas9 and Mlu-1 digested TY1-Flag-dCas9 and L1gRNA PCR product using T4 DNA ligase (Promega M180A) overnight at 4°C followed by transformation into DH5α competent cells and streaked on to ampicillin LB agar plates to obtain recombinant colonies. The individual colonies were picked and screened by colony PCR and positive colonies were purified using a Qiagen Miniprep kit and sequence verified. Finally, a midiprep was carried out to amplify the plasmid (PB-TET-dCas9-L1gRNA) for further experimentation.

5.2.3.3 Development of Huh7 PB-TET-dCas9 and PB-TET-dCas9-L1gRNA stable cell lines

Huh7 wild type cells were seeded on 6 well plates to attain approximately 80% confluency the next day. The cells were transfected with 1.5µg of PB-TET-dCas9 or PB-TET-dCas9-L1gRNA and 0.5µg of transposase plasmids using TRANS-LT1 transfection reagent. Selection with 2µg/ml puromycin was initiated 2 days after transfection and continued for a further 3-5 days until the non-transfected control cells were all dead. The cells were harvested and propagated further in puromycin containing media. To validate the cell lines, cells were induced with 500ng/ml Doxycycline (SIGMA D9891) for 24 hrs and harvested as 2 fractions for total RNA and protein extractions. The protein samples were analysed by western blot to look for Cas9 (Anti-Flag M2 monoclonal antibody, Sigma F3165, 1 in 1000 dilution in 5% BSA dissolved in 1x TBS tween) expression as stated in section 2.9 and the RNA samples were analysed for Cas9 and L1gRNA transcript expression by RT-qPCR as stated in section 2.8, non-induced cells were used as control.

5.2.4 Chromatin immunoprecipitation (ChIP)

ChIP assays were carried out by using 50-100µg cross-linked chromatin, prepared by fixing cells in 1% formaldehyde for 7 minutes followed by quenching with 1.25M glycine for 5 minutes. Cells were washed with ice cold PBS containing a protease inhibitor cocktail (Roche

4693159001) then pelleted and lysed with the following detergent based lysis solutions to liberate the cellular components and to remove the cytosolic proteins (as the protein-DNA interactions primarily occur in the nuclear compartment): Lysis buffer 1 (Hepes-KOH pH 7.5 50 mM, NaCl 140 mM, EDTA 1 mM, glycerol 10%, NP-40 0.5%, Triton X-100 0.25% and the protease inhibitors) incubated at 4°C for 10 minutes, then Lysis buffer 2 (NaCl 200 mM, EDTA 1 mM, EGTA 0.5 mM and 10 mM Tris-HCl pH 8 and protease inhibitors) incubated at 4°C for 5 minutes then the pellets were resuspended in Lysis buffer 3 (EDTA 1 mM, EGTA 0.5 mM, Tris-HCl pH8 100 mM, N-lauroyl-sarcosine 0.5%, Na-deoxycholate 0.1% and protease inhibitors). Finally, the chromatin was sheared or digested by sonicating for 30 minutes using a Diagenode Bioruptor to reach a fragment size of approximately 200-500bp. Magnetic beads (Invitrogen Dyna beads- Protein G, Cat-10003D) were coated with 5µg of anti-flag antibody by incubation at 4°C for 6 to 8 hours. Chromatin (50-100µg) samples were mixed with the antibody-coated beads in ChIP dilution buffer (Triton X-100 1%, EDTA 1Mm, Tris-HCl pH 8 20mM and NaCl 150mM) and incubated overnight at 4°C on a rotating wheel mixer. A fraction of chromatin extract (10%) was taken as input before adding the antibody-coated beads. The chromatin-antibody-bead complexes were extensively washed with Low salt buffer (SDS 0.1%, Triton X-100, EDTA 2mM, Tris-HCl pH 8.1 10mM and NaCl 150mM), High salt buffer (SDS 0.1%, Triton X-100, EDTA 2mM, Tris-HCl pH 8.1 10mM and Nacl 500mM), LiCl buffer (LiCl 0.25M, NP-40 1%, Na-Deoxycholate 1%, EDTA 1mM and Tris-Hcl pH 8.1 10mM) followed by 2 washes with TE buffer (Tris-HCl pH 8.1 10mM, EDTA 1mM). The chromatin samples were then separated from the antibody-coated beads with Elution buffer (Tris-HCl pH 8.0 50mM, EDTA 10mM, SDS 1%). Crosslinks were reversed by extensive heat incubation (65°c for 8 hrs) followed by digestion with RNAase A and Proteinase K to eliminate nucleases from DNA (this step was carried out for the input DNA as well as ChIP eluted DNA). The input and ChIPed DNA samples were purified using Qiagen Minelute PCR purification kit (Cat no.28004).

The purified ChIPed DNA was analysed by qPCR using input DNA as control. Each qPCR reaction (see section 2.8) was performed in quadruples and the analysis was repeated at least two times from independent ChIP experiments. PCR C_t values of the IP samples were normalized to that of the input samples using the formula: $\Delta C_t = C_t - (C_t\,[\text{input}] - DF)$ (Philippe C. *et al.*). DF-Dilution Factor corresponding to the number of cycles based on 10% ChIP input is 3.32. Percent input is calculated as $(\text{primer efficiency})^{-\Delta C_t} \times 100\%$. Final quantifications were

done by ΔΔCt method where fold changes were normalised to dCas9-nogRNA control and the fold enrichment above background (L1-ORF2) was calculated as (primer efficiency)$^{-\Delta\Delta C_t}$.

5.2.5 Silver staining

Silver Stain Plus kit (BIORAD-1610449) was used to detect proteins on a SDS-PAGE gel. Briefly, after electrophoresis the gel was fixed with gentle agitation for 20 minutes with fixative enhancer solution. Then the gel was rinsed with deionised water twice for 10 minutes with gentle agitation followed by staining with Development Accelerator Solution for approximately 20 minutes until protein bands become visible. The staining reaction was stopped using 5% acetic acid solution for 15 minutes. Finally, the gel was rinsed with high purity water and photographed using ChemiDoc MP Imaging System (BIORAD- 17001402).

5.2.6 5-Aza-2'-Deoxycytidine Treatment

To explore the role of L1 in HCC development and identify L1 promoter regulators, a panel of liver cancer cell lines were tested for their basal L1 expression. Very little to no L1 ORF1p protein was detected in HHL-5 cells. Since DNA methylation is known to regulate transposable elements, the response of these cells to 5-Aza-2'deoxycytidine (DAC, global hypo-methylating agent) was investigated and HHL5 cells responded to the stimulus by upregulation of LINE1 expression. Indicated cells were treated with 2.5μM 5-Aza-2'-deoxycytidine (SIGMA A-3656) and 1:1000 dilution of 50% acetic acid as a vehicle control. Starting cell densities were approximately $3x10^5$ cells/ml in a t-25cm^2 flask. Culture medium was changed every day for 3 days with fresh 5-Aza-2'-deoxycytidine medium. Cells were harvested at day 4 post-treatment and pellets were collected for total RNA and protein extraction.

5.3 Results

5.3.1 Active L1 loci in Huh7 cells

In order to identify active L1 loci in Huh7 cells, so as to use their sequence information to design specific gRNAs we used a method employing the anti-sense promoter located within the L1 5'UTR region (Macia A. *et al.*, 2011). Transcripts originating from the L1-antisense promoter were cloned and were identified using flanking genomic sequence information, as described in materials and methods section 5.2.2. 11 loci of active L1s were identified (Table 5.1).

Chromosomal location	L1 Subfamily	L1 Size	Genes upstream of the insertion
chr11:27219704-27225730	L1PA2	Full length Begin in repeat: 127 End in repeat: 6153	BBOX1-AS1
chr3:41279783-41283806	L1PA6	4024bp Begin in repeat: 2126 End in repeat: 6154	ULK-4
chr14:30474045-30480062	L1PA3	Full length Begin in repeat: 129 End in repeat: 6154	RP11-1103G16.1 (G2E3-AS1)
chr21:33925607-33931606	L1PA2	Full length Begin in repeat: 129 End in repeat: 6155	AP000304.12, LINC00649
chr13:74235945-74241970	L1PA3	Full length Begin in repeat: 124 End in repeat: 6155	LINC00402
chr9:20655633-20658802	L1HS	3170bp Begin in repeat: 126 End in repeat: 3295	FOCAD
Chr22:28663284-28669315	L1HS	Full length Begin in repeat: 124 End in repeat: 6155	TTC28
Chr15:55958210-55959465	L1MEf	1267bp Begin in repeat: 2855 End in repeat: 4122	NEDD4
Chr7:111243516-111249546	L1HS	Full length Begin in repeat: 124 End in repeat: 6155	IMMP2L
Chr9:112798108-112804159	L1HS	Full length Begin in repeat: 127 End in repeat: 6155	SNX30
Chr8:134069537-134070669	L1MEd	1133bp Begin in repeat: 870 End in repeat: 2113	Intergenic

Table 5.1 *Summary of active L1 elements expressed in Huh7 cells*

Column 2 indicates the subfamily of L1 according to Repeatmasker (http://www.repeatmasker.org/). Column 3 indicates if a full-length L1 is annotated in the Human Genome Reference Sequence (Dec. 2013 GRch38/hg38 assembly http://genome.ucsc.edu/). Column 4 indicates name of the gene that contains the L1 element.

5.3.2 Development of dCas9+L1gRNA system for L1 promoter enChIP in Huh7 cells

Chromatin Immuno-precipitation (ChIP) was performed using Monoclonal ANTI-FLAG M2 antibody on sheared chromatin from Huh7 cells transfected with dCas9 expressing plasmids along with no gRNA or gRNAs (L1gRNA1, L1gRNA2, L1gRNA3-gBlock and L1gRNA3-vector) targeting the L1 promoter (see methods section 5.2.1). Normal mouse IgG was used as a negative IP control. The purified DNA was analyzed on the QuantStudio ™ 7 flex Real Time PCR system with optimized primers for the region of L1 5'UTR, and primers targeting L1-ORF2 region and H19 single gene locus were used as negative controls. ChIP qPCR analysis showed about 6, 19, 1.75 and 8.36 fold enrichment of L1-5'UTR in the presence of L1gRNA1, L1gRNA2, L1gRNA3-gBlock and L1gRNA3-vector compared to nogRNA control respectively (Fig 5.4a). In terms of enrichment over the background (L1-ORF2) L1-5'UTR was pulled out more specifically in presence of L1gRNA2 (Fig 5.4b). Hence, L1gRNA2 was chosen for further ChIP experiments and is named as dCas9-L1gRNA hereafter.

In order to enrich for cells with positive transfection, puromycin selection was carried out. As expected, the selection improved enrichment and minimised background signal (Fig 5.5). As shown in Fig 5.5e there was a 15-fold enrichment of dCas9-L1gRNA with respect to the nogRNA control on L1 promoter with puromycin selection compared to 5-fold enrichment without selection (Fig 5.5b). Further on, exploring the results in terms of enrichment over the background (L1-ORF2) it was evident that the L1 promoter was pulled out more specifically and the chromatin shearing separated the L1 promoter from the L1-ORF2 region (fig 5.5c and 5.5f).

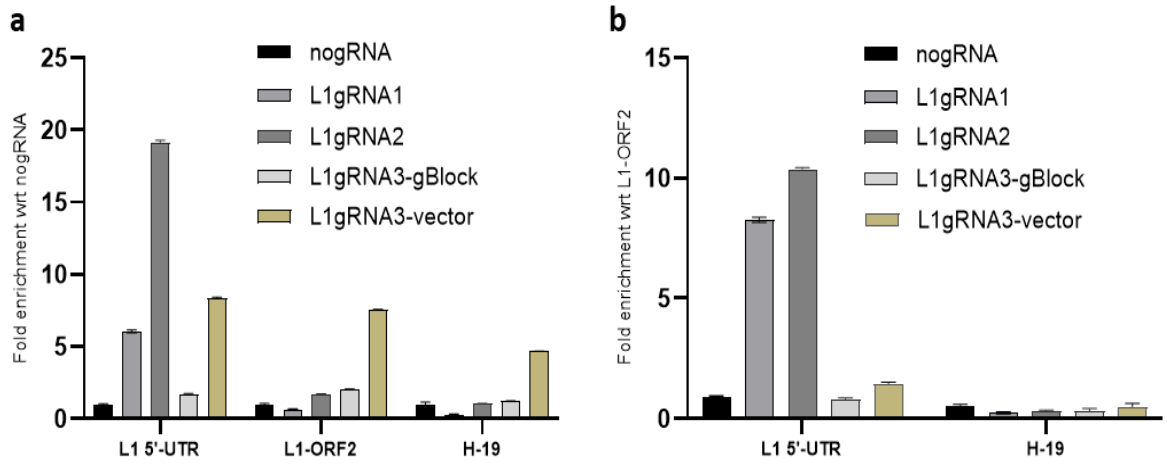

Figure 5.4 *L1 promoter enChIP in Huh7 cells*

Anti-Flag ChIP to immunoprecipitate dcas9 in presence and absence of sgRNA (L1gRNA1, L1gRNA2 L1gRNA3-gBlock and L1gRNA3-vector) against L1-5'UTR in Huh7 cells. Data are presented as fold enrichment of the locus in the presence of L1gRNA, with respect to nogRNA control (a) and with respect to L1-ORF2 (b). Error bars represent mean±SEM of 4 technical repeats.

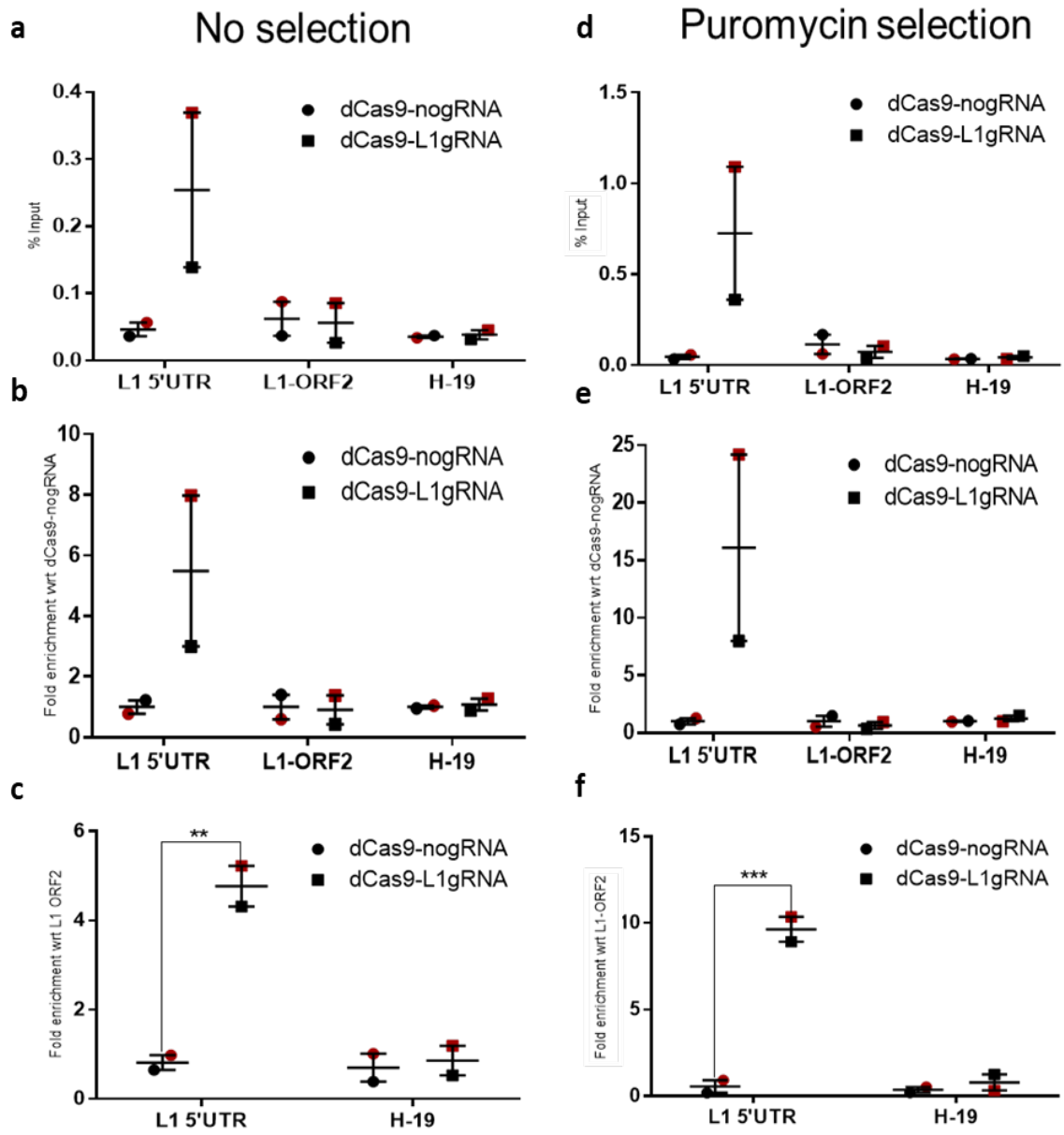

Figure 5.5 *L1 promoter enChIP with and without puromycin selection*

Anti-Flag ChIP to immunoprecipitate dcas9 in presence and absence of sgRNA against L1-5'UTR in Huh7 cells without (a-c) or with puromycin selection (d-f) was carried out followed by real time qPCR analysis of pulled out DNA to quantify isolated L1 5'UTR sequences compared to control regions (L1-ORF2 and H19). Data are presented as percent input (a and d), fold enrichment of the locus in the presence of L1gRNA with respect to dCas9-nogRNA (b and e) and as fold enrichment of the locus with respect to the negative control (L1-ORF2) or

*background signal (c and f). Data presented as mean±SEM of 2 independent experiments indicated in red and black colours. **p value 0.01, ***p value 0.001, Two-way Anova with Tukey's multiple comparisons test.*

5.3.3 Influence of dCas9 binding on L1 expression

Since dCas9 is a large protein (170KDa), binding of it to the L1 promoter may block access of the transcription factors or other chromatin remodellers to the L1 promoter, thereby potentially blocking L1 expression itself. Hence, a diagnostic western blot was performed by extracting proteins from Huh7 cells transfected with dCas9-nogRNA and dCas9-L1gRNA plasmids. Fig 5.6 shows that expression of dCas9 along with L1-promoter targeting gRNA did not hamper L1 expression, hence the system was used further to study promoter bound proteins.

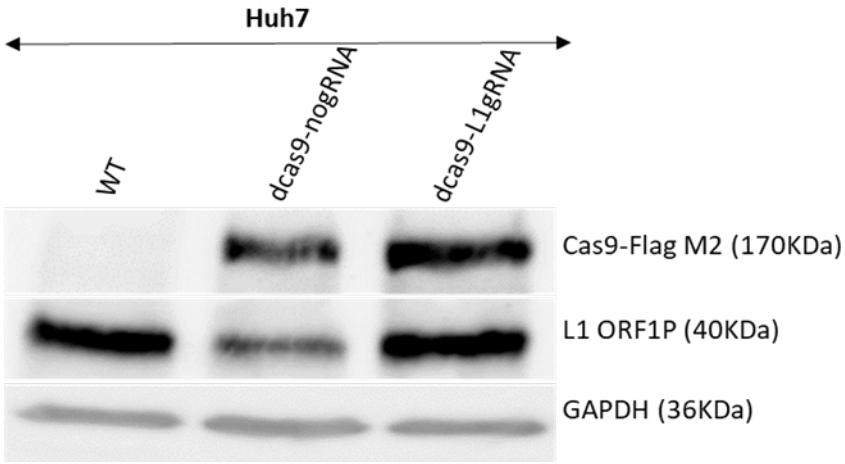

Figure 5.6 *Western Blot image showing dCas9 and L1-ORF1p expression*

Huh7 cells were transiently transfected with dCas9-nogRNA and dCas9-L1gRNA plasmids. The protein lysates were evaluated for dCas9 and L1-ORF1p expression by western blotting. GAPDH was used as a loading control. A representative image of 2 independent experiments with similar results is shown here.

5.3.4 Optimisation of chromatin concentration for proteomic analysis

The goal of this study was to identify proteins bound to an active L1 locus to carry out mass spectrometry analysis of the ChIP selected complex. To estimate whether sufficient chromatin could be selected to carry out the required analysis, a preliminary silver stain was used to visualize the dCas9 protein after separating the proteins associated with the isolated chromatin by SDS-PAGE. For this, 250µg of chromatin was used as starting material and ChIP was carried out using 25µg Anti-flag antibody. Finally, 98% of ChIP selected material was

subjected for in-gel silver staining and the remaining 2% was used on a separate gel for western blotting. The membrane was probed with Anti-flag antibody to detect dCas9 by a western blot. Although the western blot clearly showed the enrichment of dCas9 protein (Fig 5.7a) there was insufficient dCas9 to detect by a silver stain (Fig 5.7b). Hence, it was concluded that much more chromatin as starting material was needed to obtain enough protein for mass spectrometry analysis. Hence, stable cell lines by inserting the dCas9-gRNA construct using a PiggyBac vector system where constructed to overcome the limitation of starting material (described in the next section).

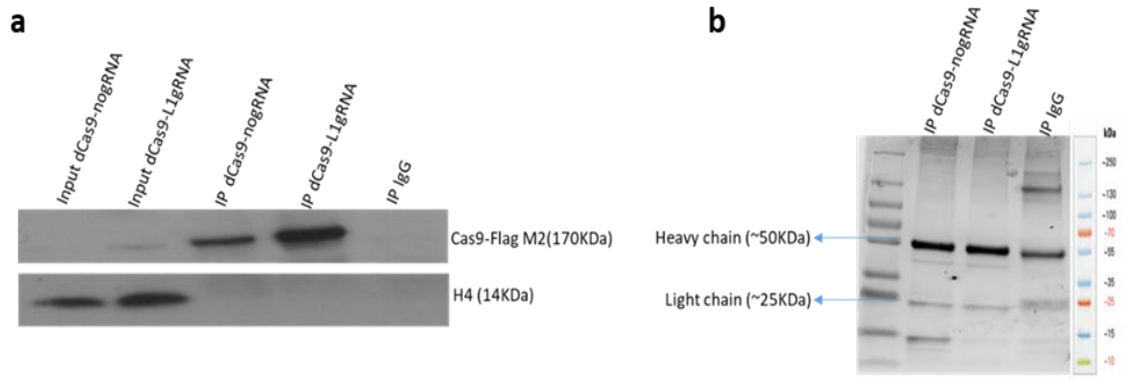

Figure 5.7 *Optimisation of chromatin concentration for proteomic analysis*

Western blot image of ChIP- Input and IP samples (Huh7 cells) showing dCas9 expression. Histone H4 was used as a loading control for input protein (a). Image of the gel with ChIP- IP samples transfected with dCas9-nogRNA, dCas9-L1gRNA plasmids stained by silver stain and PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa (b). A representative image of two independent experiments with similar results are shown here.

5.3.5 Establishment of stable Huh7-PB-TET-dCas9 and Huh7-PB-TET-dCas9-L1gRNA cell lines

The dCas9-L1gRNA construct and dCas9 alone were cloned into the PiggyBac (PB) transposon vector with a Tetracycline/Doxycycline inducible promoter. Huh7 cells were transfected with the PB-TET-dCas9 and PB-TET-dCas9-L1gRNA constructs along with plasmid encoding transposase to integrate the CRISPR contents into the TTAA chromosomal site on the cell line genome. The cells were subjected to puromycin selection for 4 days and a diagnostic western blot was carried out to check the insertion of the constructs in the cells.

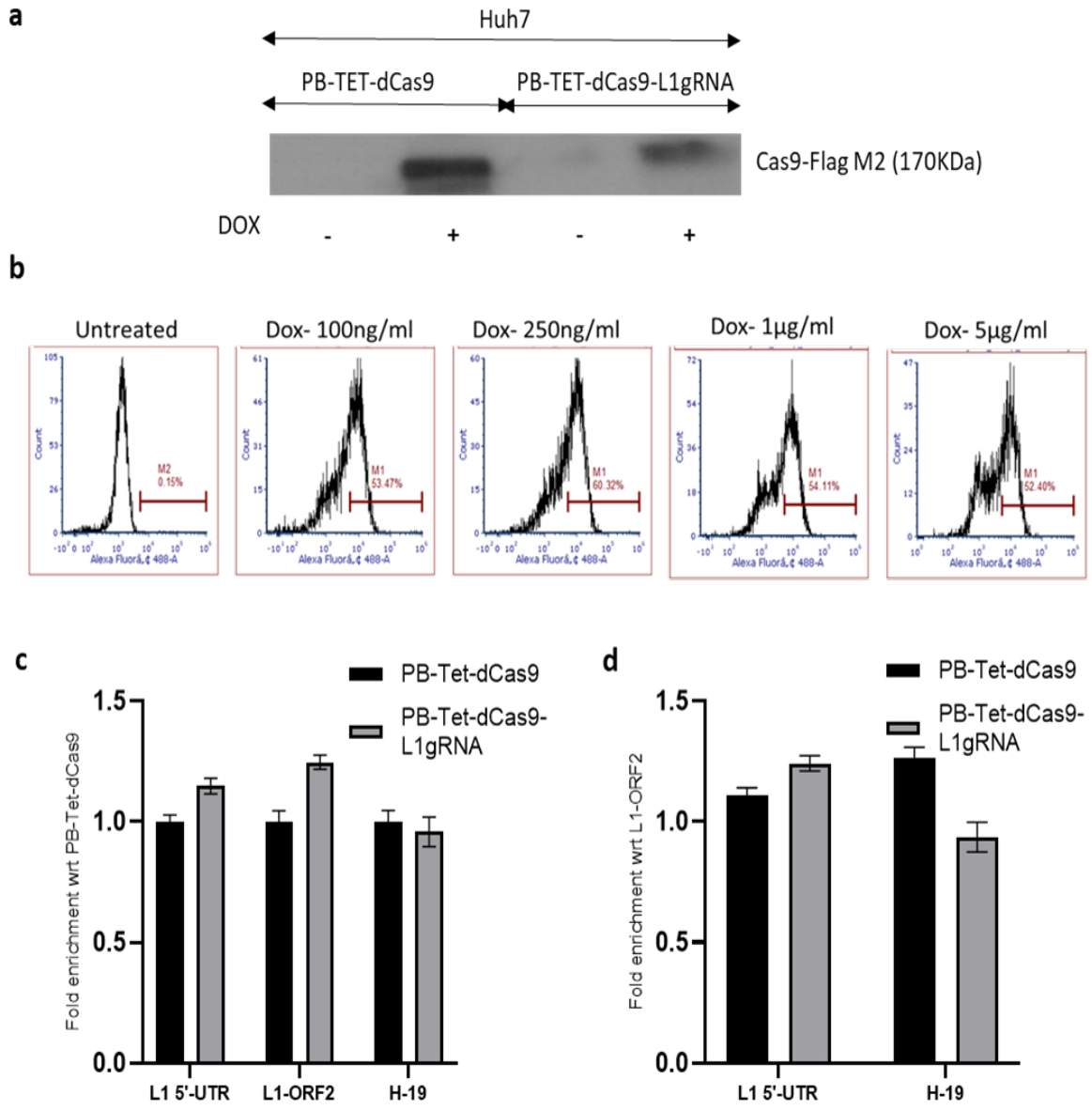

Figure 5.8 *Establishment of stable Huh7-PB-TET-dCas9 and Huh7-PB-TET-dCas9-L1gRNA cell lines*

Western blot image of Huh7 cells stably expressing PB-TET-dCas9 and PB-TET-dCas9-L1gRNA constructs showing Flag-dCas9 expression after induction with Doxycycline (500ng/ml for 24 hours) (a). FACS histograms showing Doxycycline dose-dependent increase in Flag-dCas9 positive cells (indicated doses for 24hrs) in Huh7 PB-TET-dCas9 cells where the Flag-dCas9 expression was assessed by setting a gate for untreated Huh7 PB-TET-dCas9 cells. The shift in the number of fluorescent cells was then assessed for cells treated with different doses of Doxycycline (b). Anti-Flag ChIP to Immunoprecipitate dcas9 in presence and absence of sgRNA against L1-5'UTR in Huh7 PB-TET-dCas9 and PB-TET-dCas9-L1gRNA cell lines. Data are presented as fold enrichment of the antibody signal with respect to (c) no-gRNA control and (d) fold enrichment of the antibody signal with respect to the negative control or background signal (L1-ORF2).

As expected, dCas9 expression was detected in Dox-induced cell lysates (Fig 5.8a) indicating successful integration of the expression-cassette in the cells. To determine optimum doxycycline dose to be used for performing the ChIP assay, doxycycline dose response was evaluated in Huh7 PB-TET-dCas9 cells by carrying out FACS analysis to detect dCas9 expression at single cell level in the whole population. As shown in Fig 5.8b only 50-60% of cells are induced even with highest concentration of Doxycycline revealing a drawback of the cell line that might be overcome by isolating and expanding populations of positive single cell clones. However, before going on to derivation of clones, ChIP was carried out in the cells after induction with Doxycycline (500ng/ml for 24 hours), in the same way as done previously in Huh7 cells after transient plasmid transfection.

However, ChIP with an anti-Flag antibody showed no specific enrichment for the L1 promoter in the Huh7 PB-TET-dCas9-L1gRNA cell line compared to the negative control Huh7 PB-TET-dCas9 cells (Fig 5.8c and 5.8d). This indicated some intrinsic defect in the newly generated stable cell line.

Upon further investigation, it was observed that although the integration of the L1gRNA has happened in the cell line (Fig 5.9a) there is no expression of the gRNA (Fig 5.9b). This explains why there was no specific enrichment of L1-5'UTR in Huh7-dCas9-L1gRNA cells by ChIP.

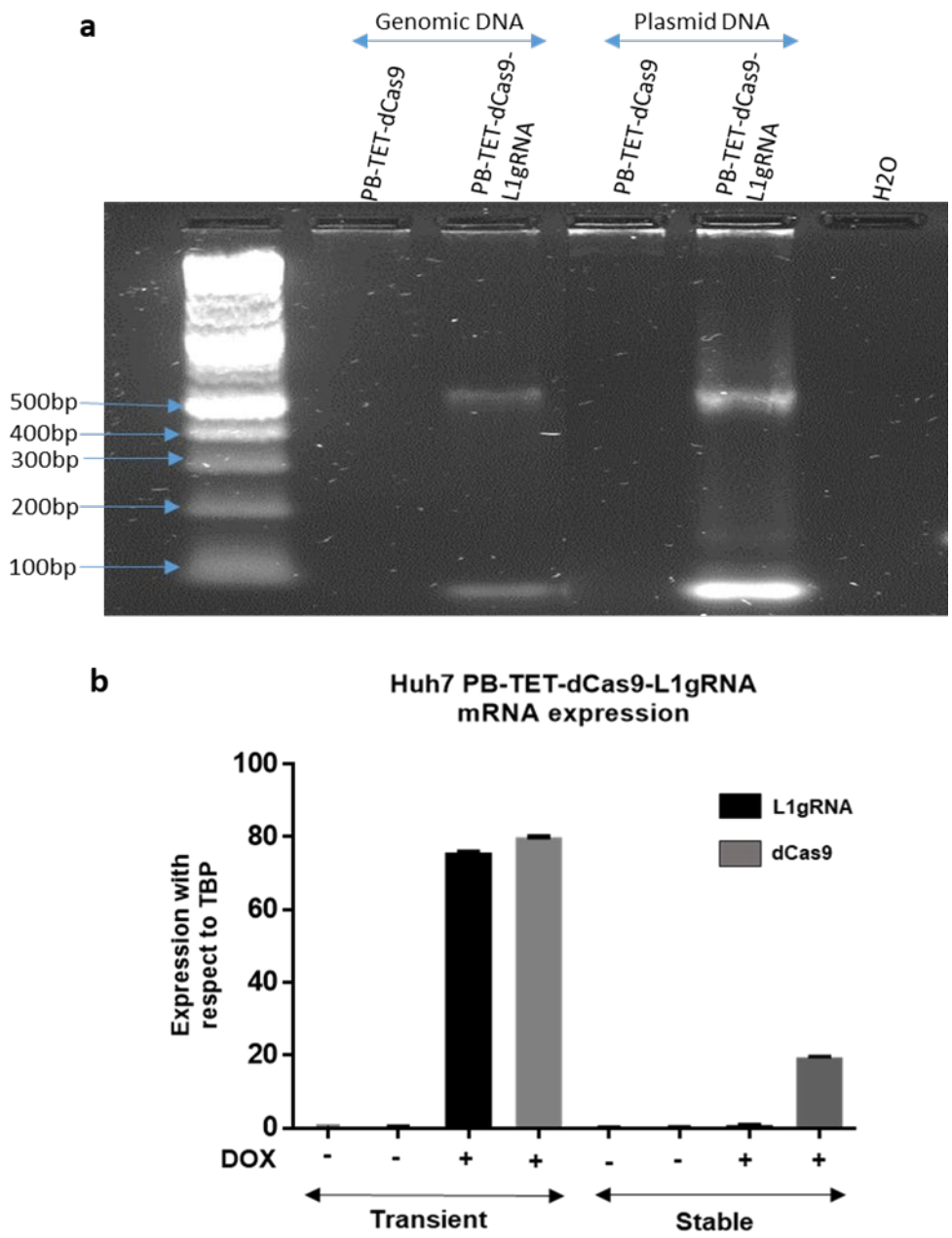

Figure 5.9 *Diagnosis of L1gRNA integration and expression in Huh7 PB-TET-dCas9-L1gRNA cell line*

PCR of genomic DNA isolated from the Huh7 cells stably expressing PB-TET-dCas9-L1gRNA construct revealed the integration of L1gRNA when analysed on a 2% agarose gel. The 90bp product (lane 3) corresponds to the L1gRNA sequence compared to plasmid DNA positive control (lane 5) (a). Results from RT-qPCR analysis showing the transcript expression of L1gRNA and dCas9 with and without Doxycycline (DOX) induction (500ng/ml for 24 hours) in Huh7 cells after transient and stable transfection. To normalise the results amplification was expressed relative to the Ct for TBP (tata binding protein housekeeping control) (b).

To overcome these defects, we devised a strategy to transiently transfect the L1gRNA as RNA oligo into the PB-TET-dCas9 cells. First, L1gRNA transcript expression conditions were optimised in a dose and time dependent manner. As evident from Fig 5.10a and 5.10b, 50nM concentration of L1gRNA gave maximum expression after 24 hrs of transfection. These conditions were hence chosen for further ChIP studies.

L1 promoter ChIP was then performed in Huh7 PB-TET-dCas9 cells with and without L1-gRNA transfection. RT-qPCR of ChIPed DNA confirmed successful pulldown of L1 promoter (Figure 5.11a and 5.11b) however, the efficiency was less than what was observed by transient transfection. Hence, even though the initial idea was to develop stable Huh7 PB-Tet-dCas9-L1gRNA cell line in order to increase efficiency and easily scale up the process for mass spectrometric analysis of L1 promoter bound proteins it was shown not to provide sufficient yields to pursue this strategy.

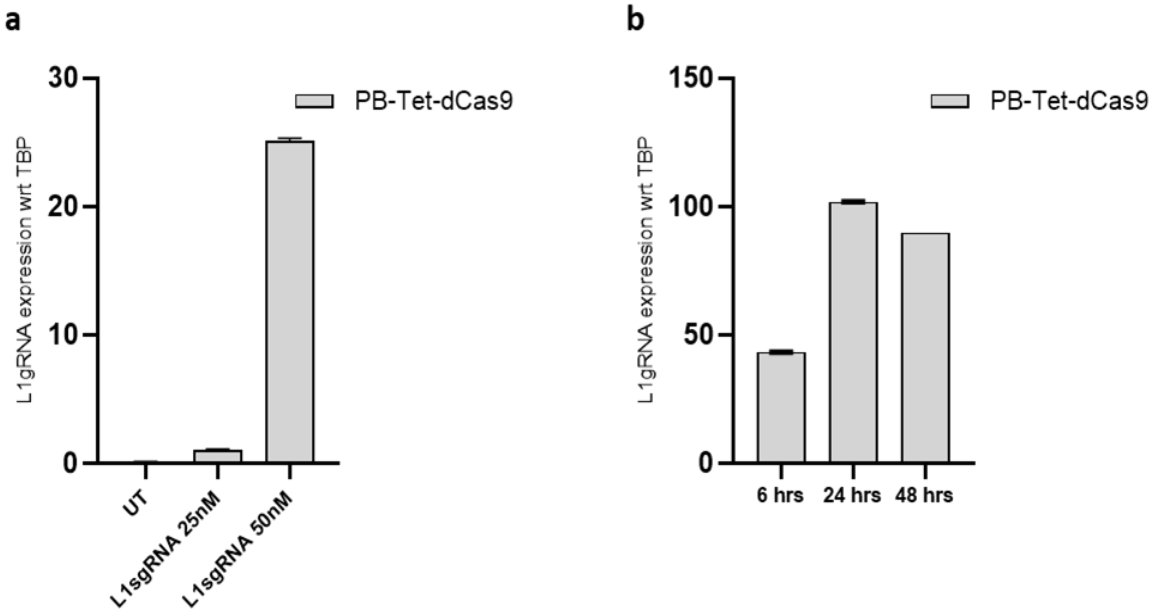

Figure 5.10 *Optimisation of L1gRNA transcript expression in Huh7 PB-TET-dCas9 cells*

Results from qPCR analysis showing the transcript expression of L1gRNA in Huh7 PB-TET-dCas9 cells after transient transfection of L1gRNA as RNA oligo (a). L1gRNA transfection optimisation (50nM concentration) at different time points was carried out and the figure shows L1gRNA transcript expression at 3 different time point. To normalise the results L1gRNA expression relative to the Ct for TBP is used (housekeeping control).

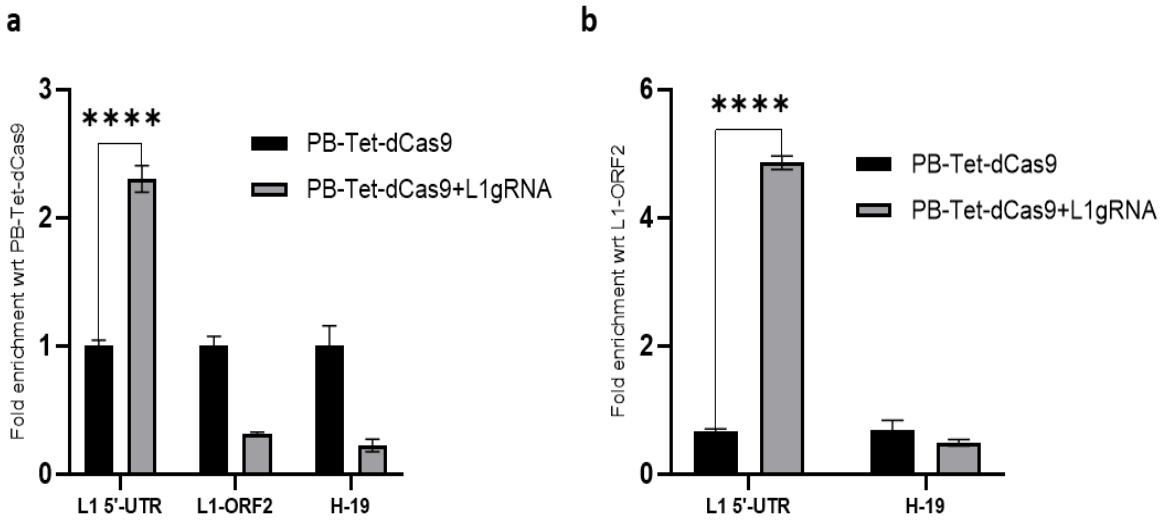

Figure 5.11 *Anti-Flag ChIP to Immuno-precipitate dcas9 in presence and absence of L1gRNA against L1-5'UTR in Huh7 PB-TET-dCas9 cells*

*Data are presented as fold enrichment of the antibody signal with respect to nogRNA control (a) and fold enrichment of the antibody signal with respect to the negative control or background signal (b). Error bars represents mean±SD of 4 technical repeats. ****p value <0.0001, Two-way Anova with Tukey's multiple comparisons test.*

5.3.6 Stimulating endogenous L1 expression by 5-Aza-2'deoxycytidine treatment

To explore the role of L1 in HCC development and identify L1 promoter regulators, a panel of liver cancer cell lines were tested for their basal L1 expression (see chapter 3). Very little to no L1 ORF1p protein was detected in HHL-5 cells. Since DNA methylation is known to regulate transposable elements, the response of these cells to 5-Aza-2'deoxycytidine (DAC, global hypomethylating agent) was investigated and HHL-5 cells was found to respond to the stimulus by upregulation of L1 expression (Fig5.12a and b).

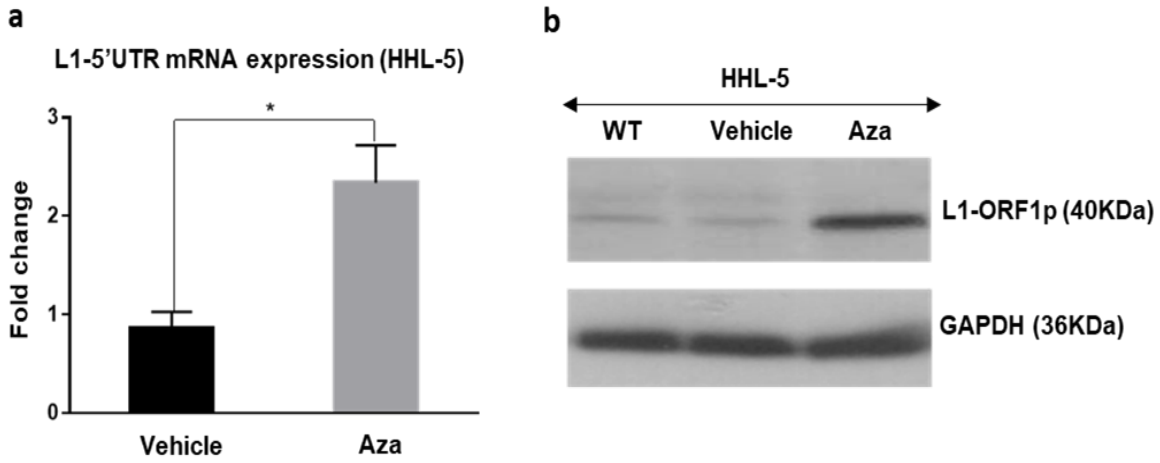

Figure 5.12 *Stimulating endogenous L1 expression by 5-Aza-2'deoxycytidine treatment*

*L1 5'UTR relative m-RNA expression detected by Real time PCR. 5-Aza-2'deoxycytidine treated (2.5µM for 72 hours) cells exhibited increased L1 mRNA expression with respect to TBP taken as the loading control. (*p<0.05. Data presented as mean±SD of 3 independent experiments, n=3.) (a). Western blot image shows L1 ORF1P protein expression of HHL-5 wild type cells treated with 5-Aza-2'deoxycytidine (2.5µM for 72 hours), 50% acetic acid as vehicle control. GAPDH was used as loading control (b). Representative images of three independent experiments with similar results are shown.*

5.3.7 PB-TET-dCas9-L1gRNA system in HHL-5 cells

In parallel to Huh7 cells, stable cell lines expressing Dox-inducible Flag-dCas9 and Dox-inducible FlagdCas9-L1gRNA were generated for HHL-5 cells as well. HHL-5 cells were chosen because they are immortalised normal human hepatocytes and L1 promoter gets activated in them upon treatment with a global hypomethylating agent (Fig 5.12). A diagnostic western blot was carried out to look for dCas9 protein expression and RT-qPCR to look for L1gRNA transcript expression. As shown in Fig 5.13a, there was a significant induction of dCas9 in the cells upon Doxycycline treatment however, the cells exhibited expression of dCas9 even at basal level. This leaky expression was also evident at transcript level (Fig 5.13b). Importantly, unlike Huh7 cells significant expression of L1gRNA was observed (Fig 5.12b). However, enChIP could not be optimised in HHL-5 cell line due to time constraint.

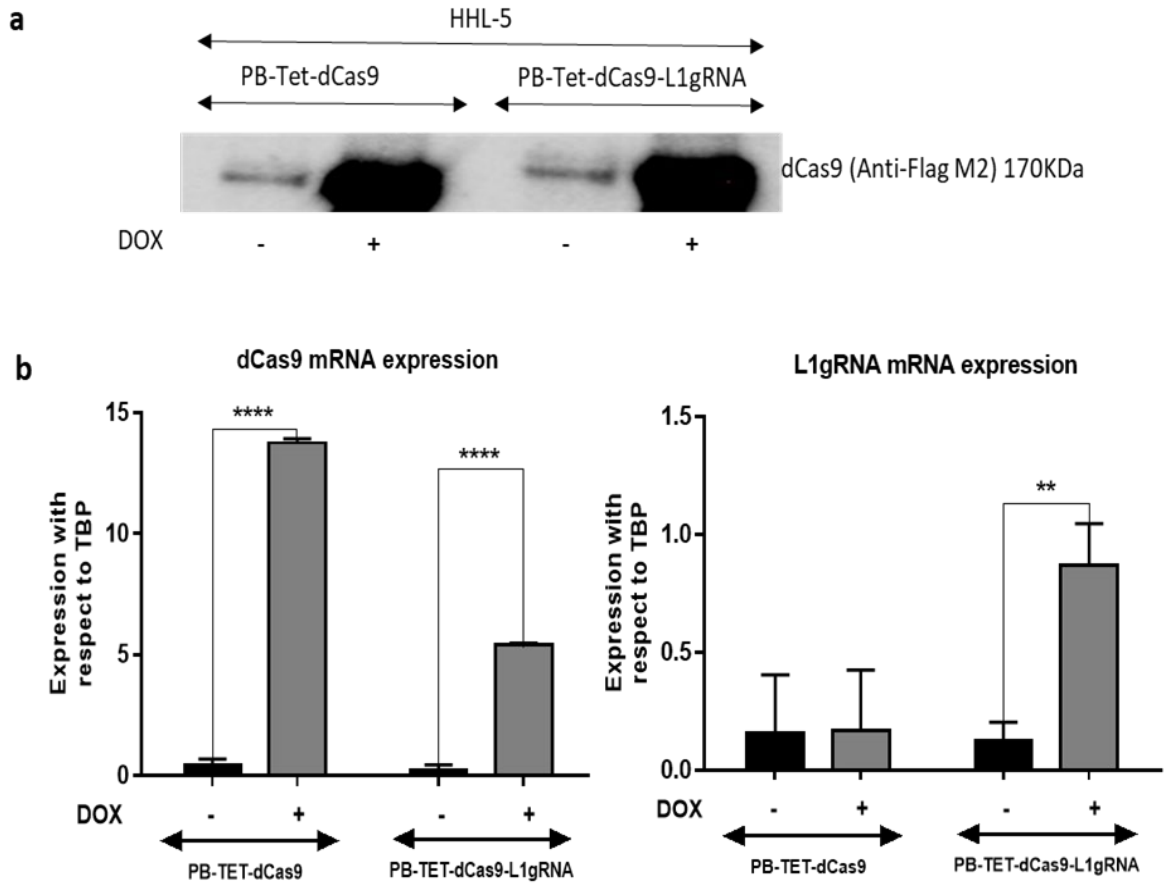

Figure 5.13 *PB-TET-dCas9-L1gRNA system in HHL-5 cells*

*Western blot image showing dCas9 protein expression in HHL-5 cells stably expressing PB-TET-dCas9 and PB-TET-dCas9-L1gRNA constructs after Doxycycline (DOX) induction (500ng/ml for 24 hrs) (a). RT-qPCR results showing dCas9 and L1gRNA transcript expression with respect to TBP (used as the house keeping gene) with and without Doxycycline induction (500ng/ml for 24 hours) in HHL-5 PB-TET-dCas9 and PB-TET-dCas9-L1gRNA cell lines (b). Error bars represents mean±SD of 4 technical repeats. **p value 0.0076 ****p value <0.0001, Two-way Anova with Tukey's multiple comparisons test.*

5.4 Discussion

Engineered DNA-binding molecule-mediated Chromatin Immunoprecipitation (enChIP) is a technique used to isolate specific genomic regions of interest using a CRISPR system containing deactivated Cas9 (dCas9) and specific guide RNAs. This system was first established by Fujita et al to isolate the *IRF-1* gene locus (Fujita T. and Fujii H., 2013). The enChIP system consists of 3xFlag tagged dCas9, locus-specific gRNA and a nuclear localisation signal derived from the SV40 T-antigen. This system is expressed in appropriate cell lines to isolate specific genomic regions of interest and the associated interacting proteins can be characterised by mass spectrometry (Fujita T. and Fujii H., 2013; Fujita T. and Fujii H., 2016). In this study, we developed a similar enChIP system using L1 promoter specific gRNAs in the Huh7 cell line.

L1gRNA1 and L1gRNA2 were designed based on the consensus L1 sequence (accession number L19088.1) in order to target as many L1 promoters as possible in the genome. However, it is worth keeping in mind that most of the L1 copies are truncated and/or defective due to mutations and rearrangements and thus won't be targeted by these gRNAs. BLAT searches for L1gRNA1 and L1gRNA2 returned hundreds of potential targets within the genome. However, not all of these will have the same status i.e. ON or OFF in a given cell. In a cell with high L1 expression status only a handful of L1 sequences at specific loci are transcriptionally active (Philippe C. *et al.*, 2016). Since we aimed to develop a locus-specific proteomics approach using CRISPR-Cas9 system to identify factors bound to 'active' versus 'inactive' L1 promoter, we first identified specific L1 loci that are active in Huh7 cells. Using an anti-sense based identification method we identified 11 active L1 loci in Huh7 cells (Table 5.1). Guide RNAs were then designed to target a majority of these active L1 loci and ordered from IDT as a gblock (named as L1gRNA3). However, no successful pulldown of L1 promoters was observed using this gRNA (Fig 5.4). Since transfecting a linear DNA construct to generate L1gRNA3 transcripts could lead to degradation of the DNA, we cloned the L1gRNA3 construct into a pSuper-puro vector and used it in co-transfection into the Huh7 cells with vector encoding dCas9 and L1gRNA3 vector. Still no significant enrichment of L1-5'UTR was observed by ChIP (Fig 5.4) indicating that more gRNAs needed to be tested as not all the gRNAs work at equal efficiency. L1gRNA1 and L1gRNA2 both worked positively in the enChIP system, as real time PCR analysis showed that the L1-5'UTR was isolated at a significantly higher amount than non-specific genomic regions such as L1-ORF2 and the H-19 locus in the presence of the

L1gRNAs, thus confirming the specificity of the developed system (Fig 5.5). Again, the efficiency of L1gRNA2 was greater than that of L1gRNA1, hence L1gRNA2 was chosen for further optimisations and downstream experiments. However, the developed enChIP system could not be taken further for proteomics mass spectrometry analysis because the yields were found to be too low when investigated by silver staining of SDS-PAGE gels of the immunoprecipitated complexes.

Overall, we identified HCC cells lines with different L1 status (see chapter 3). Then, conditions were optimised to activate L1 promoters in HHL-5 cells (immortalized hepatocytes) using global hypo-methylating agent 5-Aza-2'deoxycytidine (Fig 5.12). However, the enChIP approach using L1gRNA2 could not be optimised in HHL-5 cells because of time constraints. This is something which can be carried out in future to compare, proteins bound to the L1 promoters in untreated HHL-5 cells (L1 promoters OFF) versus 5-Aza-2'deoxycytidine-treated HHL-5 cells (L1 promoters ON). Likewise, conditions to switch-OFF the L1 promoter in Huh7 cells can be identified and then Huh7 cells with different states of L1 promoters can be compared. An alternate approach can be to design gRNA in the genomic region upstream of L1 5'end so as to target a specific L1 locus at a time and then can target an 'active' and 'inactive' L1 locus within one cell. Overall, the study will help to identify key factors involved in reactivation of retrotransposons in somatic cells and might direct towards developing strategies to 'repress' these transposable elements in oncogenically transformed cells.

Chapter 6: Conclusions and Future Prospective

Hepatocellular carcinoma (HCC) is the sixth most common type of cancer and third most frequent cause of cancer-related deaths worldwide (Ferlay J. et al., 2019). The key drivers of HCC in chronic liver disease remain elusive. The purpose of this thesis was to investigate the regulators of active L1 retrotransposition and to understand the factors leading to aberrant activation of L1 elements in the context of HCC.

Chronic infection with hepatitis C virus is the most common cause for HCC worldwide. An estimated 130-170 million people are living with HCV infection globally and more than 350,000 die of HCV related conditions per year ('Global burden of disease (GBD) for hepatitis C,' 2004). With the current advancements in DAA therapy, most patients with HCV infection get a sustained viral response (SVR), however the risk of cancer still persists in patients with advanced fibrosis and cirrhosis, hence they are kept in continued surveillance (Roche et al., 2018; Dash et al., 2020; Ioannou, 2021). HCV infected patients with cirrhosis are currently monitored for HCC by liver imaging and checking for blood alpha-fetoprotein (AFP) levels two times a year indefinitely post-SVR (Jacobson et al., 2017). Hence understanding the risk of progression in patients would have a huge impact to help develop stratified approaches to surveillance, better screening methods and prevention. However progress in this field has been relatively slow. Although the recent genetic studies show promise in polymorphism of some genes such as patatin-like phospholipase domain-containing 3 (PNPLA3) (Liu et al., 2014), transmembrane-6 superfamily member-2 (TM6SF2) (Tang et al., 2019) and programmed death receptor 1 (PDCD1) (Eldafashi et al., 2021). These genetic factors along with other risks such as the patient's age, sex, BMI, alcohol intake and type 2 diabetes, may ultimately play a role in stratified surveillance for chronic liver disease (Bianco et al., 2021). We evaluated the L1 transcript expression in RNAseq dataset (GSE84346) of Chronic HCV Hepatitis (CHC) patients and controls and observed significantly higher L1 transcripts expression in CHC liver compared to controls ($p=0.001$), confirming an association of HCV infection with retrotransposon activation in HCV patients (Chapter 3). Moreover, L1 transcripts were found to be upregulated in the non-tumour liver of HCC patients with viral hepatitis aetiology compared with patients who had HCC with no known risk of HCC (TCGA LIHC RNAseq data). Hence, data in Chapter 3 demonstrates that for CHC patients, L1 activation

can be explored further as a HCC risk factor. L1 activation can be monitored by measuring the methylation status of the active L1 promoter in the circulating cell free DNA isolated from peripheral blood, without the need for analysis of tissue biopsies (Lee K.H. et al., 2019; Ponomaryova A.A. et al., 2020).

In order to evaluate the influence of HCV on the process of active retrotransposition, we employed an in-vitro engineered retrotransposition assay in Huh7 cells in presence and absence of an HCV replicon (Chapter 3). We observed an increase in active retrotransposition in presence of HCV (Huh7-J17 cells) compared to the Huh7 parental control cells. The upregulation of active retrotransposition could be due to the impaired DDR pathways in presence of HCV (validated by HR and NHEJ plasmid rejoining assays) and increased DNA damage exerted by HCV infection (validated by increased γH2AX levels in Huh7-J17 cells compared to control cells). We also observed increase in endogenous L1 protein level in presence of HCV compared to the control (validated by L1-ORF1 protein expression by western blotting). We also examined L1 retrotransposition in three liver cancer related cell lines (Huh7, PLC-PRF/5 and SK-Hep1) in the presence and absence of DNA damage response suppression using small molecule inhibitors towards ATM (KU-55933), DNA-PK (NU-7441), ATR (VE-821), CHK1 (SRA737) and PARP (Rucaparib). Overall, we observed an increase in the retrotransposition efficiency of engineered human L1 in the p53 mutant cell lines particularly upon inhibition of Ataxia Telangiectasia And Rad3-Related Protein (ATR), and Chk1 pathway involved in DNA replication stress and DNA damage signalling.

In Chapter 4, we aimed to check the oncogenic potential of active retrotransposition using cell line model. Retrotransposition assay using EGFP or blasticidin as an indicator cassette in an immortalised human hepatocyte cell line (HHL-5) was set up to evaluate the oncogenic potential of active retrotransposition. Cells that have undergone active retrotransposition were FACS sorted and several lines were developed (HHL-5-RTN) in case of EGFP while blasticidin selection was carried out to obtain resistant clones (HHL-5-RTNBlastres). The RTN lines were characterised against the wild type lines by various in-vitro assays such as colony formation and contact inhibition to assess if the lines are transformed. One of the HHL-5-RTN lines exhibited higher growth rate, increased colony forming ability and loss of contact inhibition compared to other lines. These preliminary results from cell lines support active

role of L1 retrotransposition in cellular transformation. However, the data needs to be confirmed by in vivo tumorigenic assay such as ability to generate xenograft tumours upon subcutaneously injected in a NOD/SCID mouse. This initial pilot study warrants large scale investigation to decipher the role of L1s in cellular transformation and calculate its frequency.

Reactivation of L1s is reported in several diseases and are especially associated with inflammation and cancer. We have also demonstrated activation of L1s in the non-tumour liver of CHC patients (Chapter 3). However, exact molecular factors leading to aberrant activation of retrotransposons are still largely unknown. Thus in chapter 5 we aimed to develop a locus-specific proteomics approach using CRISPR-Cas9 system to identify factors bound to 'active' versus 'inactive' L1 promoter. gRNAs specific to L1 promoter has been optimised in order to target inactive-Cas9 to L1 promoter in Huh7 cells (cell line positive for endogenous L1 expression). However, in a given cell with L1 expression, different L1 promoters are in 'ON' or 'OFF' state depending upon the chromatin context. Since the gRNA is not specific for a particular locus specific L1, the pulled chromatin will be of mixed status. Hence to identify factors/proteins bound specific to an active L1 promoter, identifying conditions to 'switch OFF' L1 promoters in Huh7 cells is essential. Then L1 enChIP should be combined with quantitative proteomics to assess the pulled out proteins from the two cell types. Alternatively, L1 gRNAs specific to 'active L1 promoters' can be designed or a locus specific approach can be applied. However, we have identified a cell line with undetectable level of L1 expression (HHL-5 cells, immortalized human hepatocytes) and optimised conditions to activate L1 promoters in the cells using global hypo-methylating agent (5-Aza-2deoxycytidine). However, enChIP promoter hasn't been optimised on these cells. Hence, L1 promoter regulators have not been identified. In future, these cells can also be potentially used to study regulation of L1 by HCV by generating HHL5-HCV line similar to Huh7-J17 cells. Overall, the developed method can be further optimised and developed to identify key factors involved in reactivation of retrotransposons in cancer, aiming to discover novel therapeutic targets.

Chapter 7: Appendix

Gene target	Primer sequence
L1-5UTR-gRNA2_F	5’-CACCGAAAAGCGCAATATTCGGGT-3’
L1-5UTR-gRNA2_R	5’-AAACACCCGAATATTGCGCTTTTC-3’
L1gRNA3_F	5’-TGTACAAAAAAGCAGGCTTTAAAG-3’
L1gRNA3_R	5’-TAATGCCAACTTTGTACAAGAAAG-3’
L1gRNA3_ECORI	5’-GACGAATTCTGTACAAAAAAGCAGGCTTTAAAG-3’
L1gRNA3_BglII	5’-CGAGATCTTAATGCCAACTTTGTACAAGAAAG-3’
RACE Primer Outer L1_Library	5’-GCGAGCACAGAATTAATACGACTCACTATAGGTTTTTTTTTTT-3’ 5’-GCGAGCACAGAATTAATACGACT-3’ 5’-GTGAGATGAACCCGGTACCTCAG-3’
L1gRNA1_RTPCR_F	5’-CTTTCCACAAGATAAACACCCG-3’
L1gRNA1_RTPCR_R	5’-CGAAACACCGAAAAGCGCAA-3’
L1gRNA2_RTPCR_F	5’-GCAATATTCGGGTGTTTTAGAGC-3’
L1gRNA2_RTPCR_R	5’-CGGTGCCACTTTTTCAAGTTGA-3’
Cas9_F	5’-CTCTGGCCAGGGGAAACAG-3’
Cas9_R	5’-GGCAGGTTCTTATCGAAGTTGG-3’
GAPDH_F	5’-CACTAGGCGCTCACTGTTCT-3’
GAPDH_R	5’-GACCAAATCCGTTGACTCCG-3’
TBP_F	5’-GCAAGGGTTTCTGGTTTGCC-3’
TBP_R	5’-GGGTCAGTCCAGTGCCATAA-3’
HPRT_F	5’-GCTATAAATTCTTTGCTGACCTGCTG-3’
HPRT_R	5’-AATTACTTTTATGTCCCCTGTTGACTGG-3’
18s_F	5’-GTAACCCGTTGAACCCCATT-3’
18s_R	5’-CCATCCAATCGGTAGTAGCG-3’

Table 7.1 *List of Primers used for RT-qPCR in this project*

Genomic DNA target	Primer sequence
L15'UTR primer 1_F	5'-TCCATCTGAGGTACCGGGTT-3'
L15'UTR primer 1_R	5'-GGTGGGAGTGACCCGATTTT-3'
L15'UTR primer 2_F	5'-AAAATCGGGTCACTCCCACC-3'
L15'UTR primer 2_R	5'-AGCAATCAGCGAGATTCCGT-3'
L1 ORF2_F	5'-TGCGGAGAAATAGGAACACTTTT-3'
L1 ORF2_R	5'-TGAGGAATCGCCACATCGACT-3'
H19_F	5'-GGAGTCAAGGGCACAGGA-3'
H19_R	5'-GCTCTTCGAGACACCGATC-3'
GFP_R	5'-GGTGCTCAGGTAGTGGTTGTC-3'
GFP_F	5'-GAAGAACGGCATCAAGGTGAAC-3'
GFP Probe	5'-[6FAM]AGCGTGCAGCTCGCCGACCA[BHQ1]-3'
RNase P	Taqman Copy Number Reference Assay RNaseP VIC-MGB Applied Biosystems, catalog no. 4401631
5S_F	5'-CTCGTCTGATCTCGGAAGCTAAG-3'
5S_R	5'-GCGGTCTCCCATCCAAGTAC-3'

Table 7.2 *shows the list of genomic DNA primers and probes used in the project*

Protein target	Company (Catalogue number)	Dilution used	Milk/BSA
L1 ORF1p	MERCK (MABC1152) Monoclonal anti-mouse	1 in 1000	BSA
p53	DAKO (M7001) Monoclonal anti-mouse	1 in 1000	BSA
p-p53 (Ser 15)	Santa Cruz (SC-101762) Polyclonal anti-rabbit	1 in 1000	BSA
SQSTM1/P62	Santa Cruz (SC-28359) Monoclonal anti-mouse	1 in 1000	BSA
MOV10	Santa Cruz (SC-515722) Monoclonal anti-mouse	1 in 1000	BSA
NS5A	A kind gift of Charles M Rice, Rockefeller University, USA. Monoclonal anti-mouse	1 in 5000	Milk
pCHK1 ser345	Abcam (ab58567) Polyclonal anti-rabbit	1 in 1000	BSA
pCHK1 ser296	Abcam (ab79758) Monoclonal anti-rabbit	1 in 1000	BSA
YH2Ax	Millipore (JBW301) Monoclonal anti-mouse	1 in 500	BSA
Flag M2	Sigma (F3165) Monoclonal anti-mouse	1 in 1000	BSA
GAPDH	Sigma Monoclonal anti-rabbit (SAB2108266)	1 in 4000	BSA

Table 7.3 *shows the list of primary antibodies used to detect the respective protein targets by western blotting*

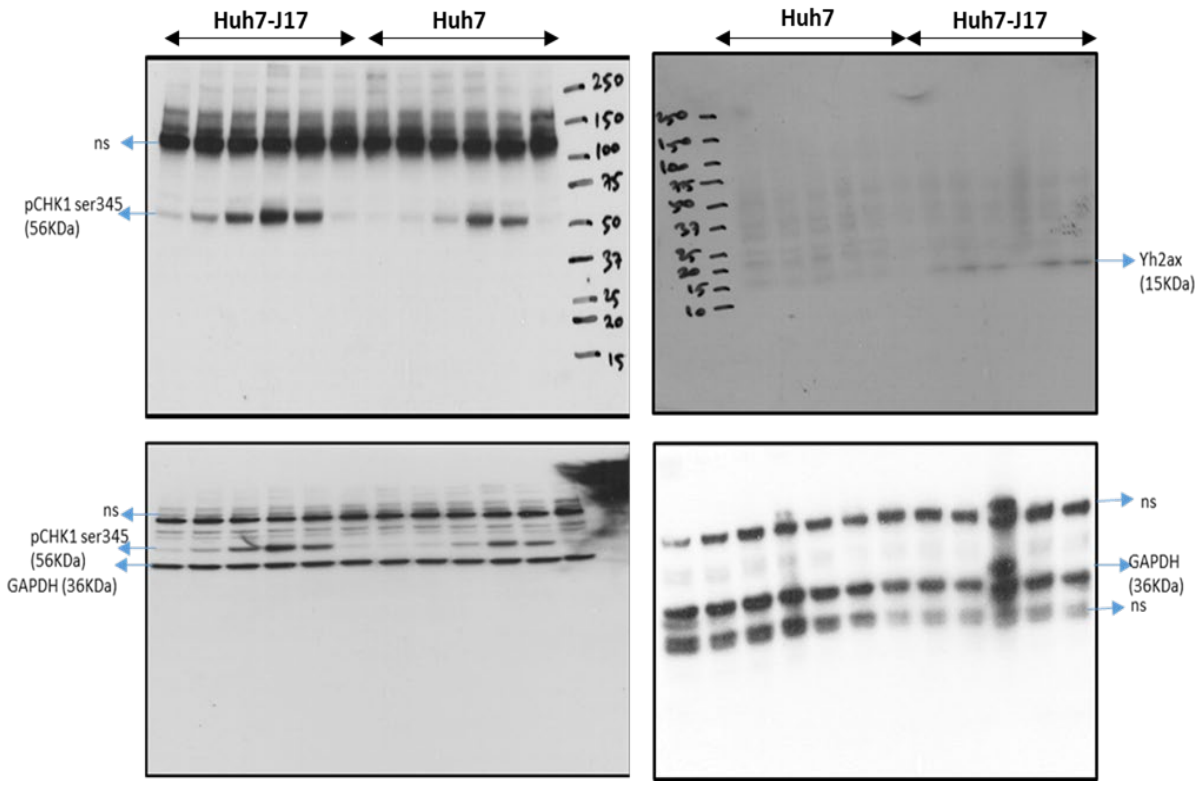

Fig S1. *Whole X-ray film scans of Western blot analysis of cell lysates of Huh7 (J7) and Huh7-J17 cells with indicated antibodies. Nonspecific bands are marked as ns. The loading sequence is same as in Fig 3.5. Expected molecular weights are indicated below each protein.*

Chapter 8: Annexure

List of conferences attended

1. NICR Postgraduate Conference
   Place- Great North Museum
   Date- 16-03-18
   Title- Locus-specific proteomics approach to identify regulators of L1 retrotransposons (Oral presentation)

2. Northeast Postgraduate Conference
   Place- Newcastle Civic Centre
   Date- 09-11-18
   Title- Locus-specific proteomics approach to identify regulators of L1 retrotransposons (Poster presentation)

3. BASL Basic Science Meeting
   Place- The Hayes Conference Centre at Alfreton, Derbyshire
   Date- 05-06-19
   Title- ATR modulates long interspersed element-1 (L1) retrotransposition in human liver cancer cells (Oral presentation)

4. HCC UK Annual Meeting 2020
   Place- Double Tree Hilton, Tower of London
   Date- 12-03-20 and 13-03-20
   Title- L1 retrotransposable elements accumulates in Hepatitis C virus (HCV) infected liver cancer cells (Poster presentation)

5. Chapter 3 of this thesis has been published in Cancers MDPI- (Sudhindar *et al.*, 2021)

Chapter 9: References

Ahmad A.S., Ormiston-Smith N. and Sasieni P.D. (2015) 'Trends in the lifetime risk of developing cancer in Great Britain: comparison of risk for those born from 1930 to 1960', *British Journal of Cancer*, 112(5), pp. 943-947.

Alter H.J. and Seeff L.B. (2000) 'Recovery, persistence, and sequelae in hepatitis C virus infection: a perspective on long-term outcome', *Semin Liver Dis*, 20, pp. 17-35.

Amicone L., Citarella F. and Cicchini C. (2015) 'Epigenetic regulation in hepatocellular carcinoma requires long noncoding RNAs', *BioMed Research International*, 2015(473942).

Anestopoulos I., Voulgaridou G.P., Georgakilas A.G., Franco R., Pappa A. and Panayiotidis M.I. (2015) 'Epigenetic therapy as a novel approach in hepatocellular carcinoma', *Pharmacology and Therapeutics*, 145, pp. 103-119.

Angus, A.G., Loquet, A., Stack, S.J., Dalrymple, D., Gatherer, D., Penin, F. and Patel, A.H. (2012) 'Conserved glycine 33 residue in flexible domain I of hepatitis C virus core protein is critical for virus infectivity', *J Virol*, 86(2), pp. 679-90.

Anstee Q.M., McPherson S. and Day C.P. (2011) 'How big a problem is non-alcoholic fatty liver disease?', *Br. Med. J*, 343, p. d3897.

Anwar S.L., Wulaningsih W. and Lehmann U. (2017) 'Transposable Elements in Human Cancer: Causes and Consequences of Deregulation', *Int J Mol Sci*, 18(5).

Ariumi Y. (2016) 'Guardian of the Human Genome: Host Defense Mechanisms against LINE-1 Retrotransposition', *Front Chem*, 4, p. 28.

Ateeq B., Unterberger A., Szyf M. and Rabbani S.A. (2008) 'Pharmacological Inhibition of DNA Methylation Induces Proinvasive and Prometastatic Genes In Vitro and In Vivo', *Neoplasia*, 10(3), pp. 266-278.

Athanikar J.N., Badge R.M. and Moran J.V. (2004) 'A YY1-binding site is required for accurate human LINE-1 transcription initiation.', *Nucleic Acid Research*, 32(13), pp. 3846–3855.

Baillie J.K., Barnett M.W., Upton K.R., Gerhardt D.J., Richmond T.A., De Sapio F., Brennan P.M., Rizzu P., Smith S., Fell M., Talbot R.T., Gustincich S., Freeman T.C., Mattick J.S., Hume D.A., Heutink P., Carninci P., Jeddeloh J.A. and Faulkner G.J. (2011) 'Somatic retrotransposition alters the genetic landscape of the human brain', *Nature*, 479(7374), pp. 534-7.

Baldick C.J., Wichroski M.J., Pendri A. and et al. (2010) 'A novel small molecule inhibitor of hepatitis C virus entry', *PLos Pathog*, 6(9), p. e1001086.

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, pp. 99-105.

Baumeister S.E., Volzke H., Marschall P., John U., Schmidt C.O. and et al. (2008) 'Impact of fatty liver disease on health care utilization and costs in a general population: a 5-year observation.', *Gastoenterology*, 134, pp. 85-94.

Baumert, T.F., Juhling, F., Ono, A. and Hoshida, Y. (2017) 'Hepatitis C-related hepatocellular carcinoma in the era of new generation antivirals', *BMC Med*, 15(1), p. 52.

Bestor T.H. (2000) 'The DNA methyltransferases of mammals', *Human Molecular Genetics*, 9(16), pp. 2395-2402.

Blank A., Markert C., Hohmann N., Carls A., Mikus G., Lehr T., Alexandrov A., Haag M., Schwab M., Urban S. and Haefeli W.E. (2016) 'First-in-human application of the novel hepatitis B and hepatitis D virus entry inhibitor myrcludex B', *Journal of Hepatology*, 65(3), pp. 483-489.

Boffa L.C., Carpaneto E.M. and Allfrey V.G. (1995) 'Isolation of active genes containing CAG repeats by DNA strand invasion by a peptide nucleic acid', *Proceedings of the National Academy of Sciences*, 92(6), pp. 1901-1905.

Bouchard M.J., Wang L. and Schneider R.J. (2006) 'Activation of Focal Adhesion Kinase by Hepatitis B Virus HBx Protein: Multiple Functions in Viral Replication', *Journal of Virology*, 80(9), pp. 4406-4414.

Bradbury, A., O'Donnell, R., Drew, Y., Curtin, N.J. and Sharma Saha, S. (2020) 'Characterisation of Ovarian Cancer Cell Line NIH-OVCAR3 and Implications of Genomic, Transcriptomic, Proteomic and Functional DNA Damage Response Biomarkers for Therapeutic Targeting', *Cancers (Basel)*, 12(7).

Bray F., Ferlay J. and et al. (2018) 'Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries', *CA: A Cancer Journal for Clinicians*, 68(6), pp. 394-424.

Buendia M. (2015) 'Hepatocellular Carcinoma', *Cold Spring Harb Perspect Med* 5(a021444).

Calland N., Albecka A., Belouzard S. and et al. (2012) 'Epigallocatechin-3-gallate is a new inhibitor of hepatitis C virus entry', *Hepatology*, 55, pp. 720-729.

Calvisi D.F., Ladu S., Gorden A., Farina M. and et al. (2007) 'Mechanistic and prognostic significance of aberrant methylation in the molecular pathogenesis of human hepatocellular carcinoma', *The Journal of clinical investigation*, 117(9), pp. 2713-2722.

Cancer definition. Available at: https://www.cancer.gov (Accessed: 22/08/2020).

Cerami E., Gao J., Dogrusoz U., Gross B.E., Sumer S.O., Aksoy B.A., Jacobsen A., Byrne C.J., Heuer M.L., Larsson E., Antipin Y., Reva B., Goldberg A.P., Sander C. and Schultz N. (2012) 'The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data', *Cancer Discov*, 2(5), pp. 401-4.

Chan S.T. and Ou J.J. (2017) 'Hepatitis C Virus-induced Autophagy and Host Innate Immune Response', *Viruses*, 9(8), p. 224.

Chan, S.T. and Ou, J.J. (2017) 'Hepatitis C Virus-Induced Autophagy and Host Innate Immune Response', *Viruses*, 9(8).

Chen J., Zaidi S., Rao S., Chen J.S., Phan L. and et al. (2018) 'Analysis of Genomes and Transcriptomes of Hepatocellular Carcinomas Identifies Mutations and Gene Expression Changes in the Transforming Growth Factor-β Pathway', *Gastroenterology*, 154(1), pp. 195-210.

Chen Y., Lin M.C., Yao H., Wang H., Zhang A.Q. and et al. (2007) 'Lentivirus-mediated RNA interference targeting enhancer of zeste homolog 2 inhibits hepatocellular carcinoma growth through down-regulation of stathmin', *Hepatology*, 46(1), pp. 200-208.

Cheng P.L., Chang M.H., Chao C.H. and Lee Y.H. (2004) 'Hepatitis C viral proteins interact with Smad3 and differentially regulate TGF-beta/Smad3-mediated transcriptional activation', *Oncogene*, 23(47), pp. 7821-38.

Cheng Y., He C., Wang M., Ma X., Mo F., Yang S., Han J. and Wei X. (2019) 'Targeting epigenetic regulators for cancer therapy: mechanisms and advances in clinical trials', *Signal Transduction and Targeted Therapy*, 4(1), p. 62.

Chevaliez S. and Pawlotsky J.M. (2006) 'HCV Genome and Life Cylce', in SL, I.T. (ed.) *Hepatitis C Viruses: Genomes and Molecular Biology*. Norfolk (UK): Horizon Bioscience.

Choo Q.L., Kuo G., Weiner A.J., Overby L.R., Bradley D.W. and Houghton M. (1989) 'Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome', *Science*, 244, pp. 359-362.

Chu, J.Y.K. and Ou, J.J. (2021) 'Autophagy in HCV Replication and Protein Trafficking', *Int J Mol Sci*, 22(3).

Chung R.T. and Baumert T.F. (2014) 'Curing chronic hepatitis C--the arc of a medical triumph', *N Engl J Med*, 370(17), pp. 1576-8.

Clayton R.F., Rinaldi A., Kandyba E.E., Edward M., Willberg C., Klenerman P. and Patel A.H. (2005) 'Liver cell lines for the study of hepatocyte functions and immunological response', *Liver Int*, 25(2), pp. 389-402.

Cong L., Ran F.A., Cox D., Lin S., Barretto R., Habib N., Hsu P.D., Wu X., Jiang W., Marraffini L.A. and Zhang F. (2013) 'Multiplex genome engineering using CRISPR/Cas systems', *Science*, 339(6121), pp. 819-23.

Cordaux R. and Batzer M.A. (2009) 'The impact of retrotransposons on human genome evolution', *Nat. Rev. Genet*, 10, pp. 691-703.

Coufal N.G., Garcia-Perez J.L., Peng G.E., Marchetto M.C., Muotri A.R., Mu Y., Carson C.T., Macia A., Moran J.V. and Gage F.H. (2011) 'Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells', *Proc Natl Acad Sci U S A*, 108(51), pp. 20382-7.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Marchetto, M.C., Muotri, A.R., Mu, Y., Carson, C.T., Macia, A., Moran, J.V. and Gage, F.H. (2011) 'Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells', *Proc Natl Acad Sci U S A*, 108(51), pp. 20382-7.

Cruickshanks H.A. and Tufarelli C. (2009) 'Isolation of cancer-specific chimeric transcripts induced by hypomethylation of the LINE-1 antisense promoter', *Genomics*, 94(6), pp. 397-406.

Cui, D. and Xu, X. (2018) 'DNA Methyltransferases, DNA Methylation, and Age-Associated Cognitive Function', *Int J Mol Sci*, 19(5).

Dai, L., Huang, Q. and Boeke, J.D. (2011) 'Effect of reverse transcriptase inhibitors on LINE-1 and Ty1 reverse transcriptase activities and on LINE-1 retrotransposition', *BMC Biochem*, 12, p. 18.

Dasarathy S., Dasarathy J., Khiyami A., Joseph R., Lopez R. and McCullough A.J. (2009) 'Validity of real time ultrasound in the diagnosis of hepatic steatosis: a prospective study.', *J. Hepatol.*, 51, pp. 1061-1067.

Dash S., Chava S., Aydin Y., Chandra P.K., Ferraris P., Chen W., Balart L.A., Wu T. and Garry R.F. (2016) 'Hepatitis C Virus Infection Induces Autophagy as a Prosurvival Mechanism to Alleviate Hepatic ER-Stress Response', *Viruses*, 8(5).

Daugherity E.K. and et al. (2012) 'The DNA damage checkpoint protein ATM promotes hepatocellular apoptosis and fibrosis in a mouse model of non-alcoholic fatty liver disease', *Cell cycle*, 11, pp. 1918-1928.

Day C.P. (2000) 'Who gets alcoholic liver disease: nature or nurture?', *J R Coll Physicians Lond.*, 34, pp. 557-562.

De Francesco R. and Migliaccio G. (2005) 'Challenges and successes in developing new therapies for hepatitis C', *Nature*, 436, pp. 953-960.

de Martel C., Maucort-Boulch D., Plummer M. and Franceschi S. (2015) 'World-wide relative contribution of hepatitis B and C viruses in hepatocellular carcinoma', *Hepatology*, 62(4), pp. 1190-200.

Deaton A.M. and Bird A. (2011) 'CpG islands and the regulation of transcription', *Genes Dev*, 25(10), pp. 1010-22.

Déjardin J. and Kingston R.E. (2009) 'Purification of Proteins Associated with Specific Genomic Loci', *Cell*, 136(1), pp. 175-186.

Denli, A.M., Narvaiza, I., Kerman, B.E., Pena, M., Benner, C., Marchetto, M.C., Diedrich, J.K., Aslanian, A., Ma, J., Moresco, J.J., Moore, L., Hunter, T., Saghatelian, A. and Gage, F.H. (2015) 'Primate-specific ORF0 contributes to retrotransposon-mediated diversity', *Cell*, 163(3), pp. 583-93.

Dhondurao-sudhindar.P. (2017) *Evaluating the role of L1 expression on Hepatocellular Carcinoma tumour biology and Doxorubicin resistance.* MRes thesis thesis. Newcastle University.

Diao J., Pantua H., Ngu H., Komuves L., Diehl L., Schaefer G. and Kapadia S.B. (2012) 'Hepatitis C virus induces epidermal growth factor receptor activation via CD81 binding for viral internalization and entry', *J Virol*, 86(20), pp. 10935-49.

Ding W., Tan H., Li X1., Zhang Y., Fang F., Tian Y., Li J. and Pan X. (2018) 'MicroRNA-493 suppresses cell proliferation and invasion by targeting ZFX in human hepatocellular carcinoma', *Cancer Biomarkers*.

Dreux M., Gastaminza P., Wieland S.F. and Chisari F.V. (2009) 'The autophagy machinery is required to initiate hepatitis C virus replication', *Proc Natl Acad Sci U S A*, 106(33), pp. 14046-51.

Duan Y., Tian L., Gao Q., Liang L., Zhang W., Yang Y., Zheng Y., Pan E., Li S. and Tang N. (2016) 'Chromatin remodeling gene ARID2 targets cyclin D1 and cyclin E1 to suppress hepatoma cell progression', *Oncotarget*, 7(29), pp. 45863-45875.

Dustin, L.B., Bartolini, B., Capobianchi, M.R. and Pistello, M. (2016) 'Hepatitis C virus: life cycle in cells, infection and host response, and analysis of molecular markers influencing the outcome of infection and response to therapy', *Clin Microbiol Infect*, 22(10), pp. 826-832.

Ehrlich M. (2002) 'DNA methylation in cancer: too much, but also too little', *Oncogene*, 21(35), pp. 5400-5413.

Einav S., Gerber D., Bryson P.D. and et al. (2008) 'Discovery of a hepatitis C target and its pharmacological inhibitors by microfluidic affinity analysis', *Nat Biotechnol*, 26, pp. 1019-1027.

El-Serag H.B. (2012) 'Epidemiology of viral hepatitis and hepatocellular carcinoma', *J. Gastroenterology*, 142(6), pp. 1264-1273.

Erwin J., Marchetto M.C.N. and Gage F.H. (2014) 'Mobile DNA Elements in the Generation of Diversity and Complexity in the Brain', *Nature Reviews Neuroscience*, 15(8), pp. 497-506.

Esposti D.D., Hernandez-Vargas H., Voegele C., Fernandez-Jimenez N., Forey N., Bancel B., Le Calvez-Kelm F., McKay J., Merle P. and Z., H. (2016) 'Identification of novel long non-coding RNAs deregulated in hepatocellular carcinoma using RNA-sequencing', *Oncotarget*, 7(22), pp. 31862–31877.

Ewing A.D., Gacita A., Wood L.D., Ma F., Xing D., Kim M.S., Manda S.S., Abril G., Pereira G., Makohon-Moore A., Looijenga L.H., Gillis A.J., Hruban R.H., Anders R.A., Romans K.E., Pandey A., Iacobuzio-Donahue C.A., Vogelstein B., Kinzler K.W., Kazazian H.H., J. and Solyom S. (2015) 'Widespread somatic L1 retrotransposition occurs early during gastrointestinal cancer evolution', *Genome Res*, 25(10), pp. 1536-45.

Falkenberg K.J. and Johnstone R.W. (2014) 'Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders', *Nature Reviews Drug Discovery*, 13, pp. 673-691.

Fattovich G., Bortolotti F. and Donato F. (2008) 'Natural history of chronic hepatitis B: special emphasis on disease progression and prognostic factors', *J Hepatol*, 48(2), pp. 335-52.

Finn R.S., Aleshin A., Dering J., Yang P., Ginther C., Desai A. and et al. (2013) 'Molecular subtype and response to dasatinib, an Src/Abl small molecule kinase inhibitor, in hepatocellular carcinoma cell lines in vitro', *Hepatology*, 57, pp. 1838-1846.

Fujimoto A., Furuta M., Shiraishi Y. and et al. (2015) 'Whole-genome mutational landscape of liver cancers displaying biliary phenotype reveals hepatitis impact and molecular diversity', *Nat Commun*, 6, pp. 1-8.

Fujita T. and Fujii H. (2013) 'Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR', *Biochemical and Biophysical Research Communications*, 439(1), pp. 132-136.

Fujita T. and Fujii H. (2016) 'Isolation of Specific Genomic Regions and Identification of Associated Molecules by enChIP', *Journal of Visualized Experiments*, 107, p. p. e53478.

Fukutomi T., Zhou Y., Kawai S., Eguchi H., Wands J.R. and Li J. (2005) 'Hepatitis C virus core protein stimulates hepatocyte growth: correlation with upregulation of wnt-1 expression', *Hepatology*, 41(5), pp. 1096-105.

Gao M., Nettles R.E., Belema M. and et al. (2010) 'Chemical genetics strategy identifies an HCV NS5A inhibitor with a potent clinical effect', *Nature*, 465, pp. 96-100.

Gao S.B., Xu B., Ding L.H., Zheng Q.L., Zhang L. and et al. (2014) 'The functional and mechanistic relatedness of EZH2 and menin in hepatocellular carcinoma', *Journal of Hepatology*, 61(4), pp. 832-839.

Gao X., Qu J., Chang X., Lu Y., Bai W., Wang H. and et al. (2014) 'Hypomethylation of long interspersed nuclear element-1 promoter is associated with poor outcomes for curative resected hepatocellular carcinoma', *Liver International*, 34(1), pp. 136-146.

Gao, X.D., Tu, L.C., Mir, A., Rodriguez, T., Ding, Y., Leszyk, J., Dekker, J., Shaffer, S.A., Zhu, L.J., Wolfe, S.A. and Sontheimer, E.J. (2018) 'C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9-APEX2', *Nat Methods*, 15(6), pp. 433-436.

Gasior, S.L., Wakeman, T.P., Xu, B. and Deininger, P.L. (2006) 'The human LINE-1 retrotransposon creates DNA double-strand breaks', *J Mol Biol*, 357(5), pp. 1383-93.

Giannelli G., Mikulits W., Dooley S., Fabregat I., Moustakas A. and et al. (2016) 'The rationale for targeting TGF-beta in chronic liver diseases', *Eur J Clin Invest*, 46, pp. 349-361.

Giannelli G., Villa E. and Lahn M. (2014) 'Transforming growth factor-β as a therapeutic target in hepatocellular carcinoma', *Cancer Research*, 74(7), pp. 1890-1894.

'Global burden of disease (GBD) for hepatitis C', (2004) *J Clin Pharmacol*, 44(1), pp. 20-9.

Goodier J.L., Ostertag E.M. and Kazazian Jr H.H. (2000) 'Transduction of 3'-flanking sequences is common in L1 retrotransposition', *Human Molecular Genetics*, 9(4), pp. 653-657.

Goodier J.L., Zhang L., Vetter M.R. and Kazazian H.H. (2007) 'LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex', *Molecular and Cellular Biology*, 27(18), pp. 6469-6483.

Goodier, J.L. (2016) 'Restricting retrotransposons: a review', *Mob DNA*, 7, p. 16.

Goossens N., Sun X. and Hoshida Y. (2015) 'Molecular classification of hepatocellular carcinoma: potential therapeutic implications', *Hepat Oncol*, 2, pp. 371-379.

Grant E. (2018) 'Pharmacodynamic response to oral administration of the selective toll-like receptor 8 agonist GS-9688 in healthy volunteers', *Hepatology : Official Journal of the American Association for the Study of Liver Diseases*, 68, p. 270A.

Griesenbeck J., Boeger H., Strattan J.S. and Kornberg R.D. (2003) 'Affinity Purification of Specific Chromatin Segments from Chromosomal Loci in Yeast', *Molecular and Cellular Biology*, 23(24), pp. 9275-9282.

Hamdane, N., Juhling, F., Crouchet, E., El Saghire, H., Thumann, C., Oudot, M.A., Bandiera, S., Saviano, A., Ponsolles, C., Roca Suarez, A.A., Li, S., Fujiwara, N., Ono, A., Davidson, I., Bardeesy, N., Schmidl, C., Bock, C., Schuster, C., Lupberger, J., Habersetzer, F., Doffoel, M., Piardi, T., Sommacale, D., Imamura, M., Uchida, T., Ohdan, H., Aikata, H., Chayama, K., Boldanova, T., Pessaux, P., Fuchs, B.C., Hoshida, Y., Zeisel, M.B., Duong, F.H.T. and Baumert, T.F. (2019) 'HCV-Induced Epigenetic

Changes Associated With Liver Cancer Risk Persist After Sustained Virologic Response', *Gastroenterology*, 156(8), pp. 2313-2329 e7.

Hamdane N., Juhling F., Crouchet E., El Saghire H., Thumann C. and et al. (2019) 'HCV-Induced Epigenetic Changes Associated With Liver Cancer Risk Persist After Sustained Virologic Response', *Gastroenterology*, 156(8), pp. 2313-2329 e7.

Han G., Zhang L., Ni X., Chen Z., Pan X., Zhu Q., Li S., Wu J., Huang X. and Wang X. (2018) 'MicroRNA-873 Promotes Cell Proliferation, Migration, and Invasion by Directly Targeting TSLC1 in Hepatocellular Carcinoma', *Cellular Physiology and Biochemistry*, 46(6), pp. 2261–2270

Hardy T., Oakley F., Anstee Q.M. and Day C.P. (2016) 'Nonalcohplic Fatty Liver Disease: Pathogenesis and Disease Spectrum', *Annual Review of Pathology: Mechanisms of Disease* 11, pp. 451-496.

Harris C.R., Dewan A., Zupnick A., Normart R., Gabriel A., Prives C., Levine A.J. and Hoh J. (2009) 'p53 responsive elements in human retrotransposons', *Oncogene*, 28(44), pp. 3857-65.

Hassan M., Selimovic D., Ghozlan H. and Abdel-kader O. (2009) 'Hepatitis C virus core protein triggers hepatic angiogenesis by a mechanism including multiple pathways', *Hepatology*, 49(5), pp. 1469-82.

Helman E., Lawrence M.S., Stewart C. and et al. (2014) 'Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing', *Genome Research*, 24(7), pp. 1053-1063.

Hernandez-Meza G., von Felden J., Gonzalez-Kozlova E.E., Garcia-Lezana T., Peix J., Portela A., Craig A.J., Sayols S., Schwartz M., Losic B., Mazzaferro V., Esteller M., Llovet J.M. and Villanueva A. (2020) 'DNA methylation profiling of human hepatocarcinogenesis', *Hepatology*.

Hirschfield H., Bian C.B., Higashi T., Nakagawa S., Zeleke T.Z., Nair V.D., Fuchs B.C. and Hoshida Y. (2018) 'In vitro modeling of hepatocellular carcinoma molecular subtypes for anti-cancer drug assessment', *Exp Mol Med*, 50(1), p. e419.

Hodgson A.J., Hyser J.M., Keasler V.V., Cang Y. and Slagle B.L. (2012) 'Hepatitis B virus regulatory HBx protein binding to DDB1 is required but is not sufficient for maximal HBV replication', *Virology*, 426(1), pp. 73-82.

Holmberg S.D., Spradling P.R., Moorman A.C. and Denniston M.M. (2013) 'Hepatitis C in the United States.', *N. Engl. J. Med*, 368, pp. 1859-1861.

Hoshida Y., Nijiman S.M.B., Kobayashi M., Chan J.A. and et al. (2009) 'Integrative Transcriptome Analysis Reveals Common Molecular Subclasses of Human Hepatocellular Carcinoma', *Cancer Research*, 69(15), pp. 7385-7392.

Hoshino A. and Fujii H. (2009) 'Insertional chromatin immunoprecipitation: A method for isolating specific genomic regions', *Journal of Bioscience and Bioengineering*, 108(5), pp. 446-449.

Hsu I.C., Tokiwa T., Bennett W., Metcalf R.A., Welsh J.A., Sun T. and Harris C.C. (1993) 'p53 gene mutation and integrated hepatitis B viral DNA sequences in human liver cancer cell lines', *Carcinogenesis*, 14(5), pp. 987-92.

Igloi Z., Kazlauskas A., Saksela K., Macdonald A., Mankouri J. and Harris M. (2015) 'Hepatitis C virus NS5A protein blocks epidermal growth factor receptor degradation via a proline motif- dependent interaction', *J Gen Virol*, 96(8), pp. 2133-2144.

Iijima, K., Okudaira, N., Tamura, M., Doi, A., Saito, Y., Shimura, M., Goto, M., Matsunaga, A., Kawamura, Y.I., Otsubo, T., Dohi, T., Hoshino, S., Kano, S., Hagiwara, S., Tanuma, J., Gatanaga, H., Baba, M., Iguchi, T., Yanagita, M., Oka, S., Okamura, T. and Ishizaka, Y. (2013) 'Viral protein R of human immunodeficiency virus type-1 induces retrotransposition of long interspersed element-1', *Retrovirology*, 10, p. 83.

Inamura K., Yamauchi M., Nishihara R., Lochhead P., Qian Z.R., Kuchiba A. and et al. (2014) 'Tumor LINE-1 methylation level and microsatellite instability in relation to colorectal cancer prognosis', *Journal of the National Cancer Institute*, 106(9).

Ioannou, G.N. (2021) 'HCC surveillance after SVR in patients with F3/F4 fibrosis', *J Hepatol*, 74(2), pp. 458-465.

Ito S., D'Alessio A.C., Taranova O.V., Hong K., Sowers L.C. and Zhang Y. (2010) 'Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification', *Nature*, 466(7310), pp. 1129-1133.

Iwai A., Takegami T., Shiozaki T. and Miyazaki T. (2011) 'Hepatitis C virus NS3 protein can activate the Notch-signaling pathway through binding to a transcription factor, SRCAP', *PLoS One*, 6(6), p. e20718.

Jacobson I.M., McHutchison J.G., Dusheiko G. and et al. (2011) 'ADVANCE Study Team. Telaprevir for previously untreated chronic hepatitis C virus infection.', *N Engl J Med* 364(2405-2416).

Jasinskas A. and Hamkalo B.A. (1999) 'Purification and Initial Characterization of Primate Satellite Chromatin', *Chromosome Research*, 7(5), pp. 341-354.

Jenkins N.C. and et al. (2011) 'The p16(INK4A) tumor suppressor regulates cellular oxidative stress', *Oncogene*, 30, pp. 265-274.

Jia L., Gao Y., He Y., Hooper J.D. and Yang P. (2020) 'HBV induced hepatocellular carcinoma and related potential immunotherapy', *Pharmacological Research*, 159, p. 104992.

Jiang H., Cao H., Ma N. and et al. (2020) 'Chromatin remodeling factor ARID2 suppresses hepatocellular carcinoma metastasis via DNMT1-Snail axis', *Proceedings of the National Academy of Sciences*, 117(9), pp. 4770-4780.

Jiang Y.Z., Manduchi E., Jiménez J.M. and Davies P.F. (2015) 'Endothelial epigenetics in biomechanical stress: disturbed flow-mediated epigenomic plasticity in vivo and in vitro', *Arterioscler Thromb Vasc Biol*, 35(6), pp. 1317-26.

Jones, R.B., Garrison, K.E., Wong, J.C., Duan, E.H., Nixon, D.F. and Ostrowski, M.A. (2008) 'Nucleoside analogue reverse transcriptase inhibitors differentially inhibit human LINE-1 retrotransposition', *PLoS One*, 3(2), p. e1547.

Jones, R.B., Song, H., Xu, Y., Garrison, K.E., Buzdin, A.A., Anwar, N., Hunter, D.V., Mujib, S., Mihajlovic, V., Martin, E., Lee, E., Kuciak, M., Raposo, R.A., Bozorgzad, A., Meiklejohn, D.A., Ndhlovu,

L.C., Nixon, D.F. and Ostrowski, M.A. (2013) 'LINE-1 retrotransposable element DNA accumulates in HIV-1-infected cells', *J Virol*, 87(24), pp. 13307-20.

Jueliger S., Lyons J., Cannito S., Pata I., Pata P., Shkolnaya M., Lo Re O., Peyrou M., Villarroya F., Pazienza V., Rappa F., Cappello F., Azab M., Taverna P. and Vinciguerra M. (2016) 'Efficacy and epigenetic interactions of novel DNA hypomethylating agent guadecitabine (SGI-110) in preclinical models of hepatocellular carcinoma', *Epigenetics*, 11(10), pp. 709-720.

Kanda M., Sugimoto H. and Kodera Y. (2015) 'Genetic and epigenetic aspects of initiation and progression of hepatocellular carcinoma', *World Journal of Gastroenterology*, 21(37), pp. 10584–10597.

Kannan M., Li J., Fritz S.E., Husarek K.E., Sanford J.C., Sullivan T.L., Tiwary P.K., An W., Boeke J.D. and Symer D.E. (2017) 'Dynamic silencing of somatic L1 retrotransposon insertions reflects the developmental and cellular contexts of their genomic integration', *Mob DNA*, 8, p. 8.

Kawano, K., Doucet, A.J., Ueno, M., Kariya, R., An, W., Marzetta, F., Kuroki, M., Turelli, P., Sukegawa, S., Okada, S., Strebel, K., Trono, D. and Ariumi, Y. (2018) 'HIV-1 Vpr and p21 restrict LINE-1 mobility', *Nucleic Acids Res*, 46(16), pp. 8454-8470.

Kelso T.W.R., Porter D.K., Amaral M.L., Shokhirev M.N., Benner C. and Hargreaves D.C. (2017) 'Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in ARID1A-mutant cancers', *Elife*, 6.

Khan H., Smit A. and Boissinot S. (2006) 'Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates.', *Genome Research*, 16(1), pp. 78-87.

Kim C.W. and Chang K. (2013) 'Hepatitis C virus: virology and life cycle', *Clinical and Molecular Hepatology*, 19, pp. 17-25.

Kitrinos K.M., Corsa A., Liu Y., Flaherty J., Snow-Lampart A., Marcellin P., Borroto-Esoda K. and Miller M.D. (2014) 'No detectable resistance to tenofovir disoproxil fumarate after 6 years of therapy in patients with chronic hepatitis B', *Hepatology*, 59(2), pp. 434-442.

Korkmaz A., Manchester L., Topal T., Ma S., Tan D. and Reiter R. (2011) 'Epigenetic mechanisms in human physiology and diseases', *Journal of Experimental and Integrative Medicine*, 1.

La Coste de A., Romagnolo B., Billuart P. and et al. (1998) 'Somatic mutations of the beta-catenin gene are frequent in mouse and human hepatocellular carcinomas', *Proc Natl Acad Sci U S A*, 95, pp. 8847-8851.

Lachenmayer A., Toffanin S., Cabellos L., Alsinet C., Hoshida Y., Villanueva A., Minguez B., Tsai H.W., Ward S.C., Thung S., Friedman S.L. and Llovet J.M. (2012) 'Combination therapy for hepatocellular carcinoma: additive preclinical efficacy of the HDAC inhibitor panobinostat with sorafenib.', *Journal of Hepatology*, 56(6), pp. 1343-1350.

Lai C.L., Dienstag J., Schiff E., Leung N.W., Atkins M., Hunt C., Brown N., Woessner M., Boehme R. and Condreay L. (2003) 'Prevalence and clinical correlates of YMDD variants during lamivudine therapy for patients with chronic hepatitis B', *Clin Infect Dis*, 36(6), pp. 687-96.

Lam A.M., Espiritu C., Vogel R., Ren S., Lau V., Kelly M., Kuduk S.D., Hartman G.D., Flores O.A. and Klumpp K. (2019) 'Preclinical Characterization of NVR 3-778, a First-in-Class Capsid Assembly Modulator against Hepatitis B Virus', *Antimicrobial Agents and Chemotherapy*, 63(1), pp. e01734-18.

Lan K.H., Sheu M.L., Hwang S.J., Yen S.H., Chen S.Y., Wu J.C., Wang Y.J., Kato N., Omata M., Chang F.Y. and Lee S.D. (2002) 'HCV NS5A interacts with p53 and inhibits p53-mediated apoptosis', *Oncogene*, 21(31), pp. 4801-11.

Lan Y. and Evans T. (2019) 'Epigenetic Regulation of Cardiac Development and Disease through DNA Methylation', *Journal of life sciences (Westlake Village, Calif.)*, 1(2), pp. 1-10.

Lander E.S. and et al. (2001) 'Initial sequencing and analysis of the human genome', *Nature*, 409, pp. 860-921.

Lanford R.E., Guerra B., Chavez D., Giavedoni L., Hodara V.L., Brasky K.M., Fosdick A., Frey C.R., Zheng J., Wolfgang G., Halcomb R.L. and Tumas D.B. (2013) 'GS-9620, an oral agonist of Toll-like receptor-7, induces prolonged suppression of hepatitis B virus in chronically infected chimpanzees', *Gastroenterology*, 144(7), pp. 1508-17, 1517.e1-10.

Lavanchy, D. (2009) 'The global burden of hepatitis C', *Liver Int*, 29 Suppl 1, pp. 74-81.

Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L.J., 3rd, Lohr, J.G., Harris, C.C., Ding, L., Wilson, R.K., Wheeler, D.A., Gibbs, R.A., Kucherlapati, R., Lee, C., Kharchenko, P.V., Park, P.J. and Cancer Genome Atlas Research, N. (2012) 'Landscape of somatic retrotransposition in human cancers', *Science*, 337(6097), pp. 967-71.

Lee E., Iskow R., Yang L., Gokcumen O., Haseley P., Luquette L.J., r., Lohr J.G., Harris C.C., Ding L., Wilson R.K., Wheeler D.A., Gibbs R.A., Kucherlapati R., Lee C., Kharchenko P.V. and Park P.J. (2012) 'Landscape of somatic retrotransposition in human cancers', *Science*, 337(6097), pp. 967-71.

Li H., Yan L., Shi Y., Lv D., Shang J., Bai L. and Tang H. (2020) 'Hepatitis B Virus Infection: Overview', in Tang, H. (ed.) *Hepatitis B Virus Infection: Molecular Virology to Antiviral Drugs*. Singapore: Springer Singapore, pp. 1-16.

Li M., Zhao H., Zhang X. and et al. (2011) 'Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma', *Nat Genet*, 43, pp. 828-829.

Liang H.W., Wang N., Wang Y., Wang F., Fu Z., Yan X., Zhu H., Diao W., Ding Y., Chen X., Zhang C.Y. and Zen K. (2016) 'Hepatitis B virus-human chimeric transcript HBx-LINE1 promotes hepatic injury via sequestering cellular microRNA-122', *Journal of Hepatology*, 64(2), pp. 278-291.

Lim D.H.K. and Maher E.R. (2010) 'DNA methylation: a form of epigenetic control of gene expression', *The Obstetrician and Gynaecologist*, 12, pp. 37-42.

Llovet J.M., Burroughs A. and Bruix J. (2003) 'Hepatocellular Carcinoma', *Lancet*, 362, pp. 1907-1917.

Lovsin N. and Peterlin B.M. (2009) 'APOBEC3 proteins inhibit LINE-1 retrotransposition in the absence of ORF1p binding', *Annals of the Newyork Academy of Sciences*, 1178, pp. 268-275.

Lozada M.E., Chaiteerakij R. and Roberts L.R. (2015) 'Screening for Hepatocellular Carcinoma and Cholangiocarcinoma: Can Biomarkers Replace Imaging? ', *Current Hepatology Reports*, 14(2), pp. 128-138.

Lü J., Zhao Q., Ding X., Guo Y., Li Y., Xu Z., Li S., Wang Z., Shen L., Chen H.W., Yu Z. and Pestell R.G. (2020) 'Cyclin D1 promotes secretion of pro-oncogenic immuno-miRNAs and piRNAs', *Clin Sci (Lond)*, 134(7), pp. 791-805.

Luan D.D., Korman M.H., Jakubczak J.L. and Eickbush T.H. (1993) 'Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition', *Cell*, 72(4), pp. 595-605.

Luger K., Dechassa M.L. and Tremethick D.J. (2012) 'New insights into nucleosome and chromatin structure: an ordered state or a disordered affair?', *Nature Review Molecular Cell Biology*, 13, pp. 436-447.

Ma X.L., Lalezari J. and et al. (2019) 'LBO-06-Interim safety and efficacy results of the ABI-H0731 phase 2a program exploring the combination of ABI-H0731 with Nuc therapy in treatment-naive and treatment-suppressed chronic hepatitis B patients.', *Journal of Hepatology : the Journal of the European Association for the Study of the Liver*, 70(1), p. e130.

Machida, K., McNamara, G., Cheng, K.T., Huang, J., Wang, C.H., Comai, L., Ou, J.H. and Lai, M.M. (2010) 'Hepatitis C virus inhibits DNA damage repair through reactive oxygen and nitrogen species and by interfering with the ATM-NBS1/Mre11/Rad50 DNA repair pathway in monocytes and hepatocytes', *J Immunol*, 185(11), pp. 6985-98.

Macia A., Munoz-Lopez M., Cortes J.L., Hastings R.K., Morell S. and et al. (2011) 'Epigenetic Control of Retrotransposon Expression in Human Embryonic Stem Cells', *MOLECULAR AND CELLULAR BIOLOGY*, 31(2), pp. 300-316.

Maeda, K., Das, D., Kobayakawa, T., Tamamura, H. and Takeuchi, H. (2019) 'Discovery and Development of Anti-HIV Therapeutic Agents: Progress Towards Improved HIV Medication', *Curr Top Med Chem*, 19(18), pp. 1621-1649.

Magri, A., Ozerov, A.A., Tunitskaya, V.L., Valuev-Elliston, V.T., Wahid, A., Pirisi, M., Simmonds, P., Ivanov, A.V., Novikov, M.S. and Patel, A.H. (2016) 'Exploration of acetanilide derivatives of 1-(omega-phenoxyalkyl)uracils as novel inhibitors of Hepatitis C Virus replication', *Sci Rep*, 6, p. 29487.

Maki, A., Kono, H., Gupta, M., Asakawa, M., Suzuki, T., Matsuda, M., Fujii, H. and Rusyn, I. (2007) 'Predictive power of biomarkers of oxidative stress and inflammation in patients with hepatitis C virus-associated hepatocellular carcinoma. ', *Annals of surgical oncology*, 14, pp. 1182-90.

Malone C.D., Brennecke J., Dus M., Stark A., McCombie W.R., Sachidanandam R. and et al. (2009) 'Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. ', *Cell*, 137(3), pp. 522-535.

Mandal P.K., Ewing A.D., Hancks D.C. and Kazazian H.H., J. (2013) 'Enrichment of processed pseudogene transcripts in L1-ribonucleoprotein particles', *Hum Mol Genet*, 22(18), pp. 3730-48.

Mandayam S., Jamal M.M. and Morgan T.R. (2004) 'Epidemiology of alcoholic liver disease', *Semin Liver dIS*, 24, pp. 217-232.

Mann J., Chu D.C., Maxwell A., Oakley F., Zhu N.L., Tsukamoto H. and Mann D.A. (2010) 'MeCP2 controls an epigenetic pathway that promotes myofibroblast transdifferentiation and fibrosis', *Gastroenterology*, 138(2), pp. 705-14, 714.e1-4.

Mao B. and Wang G. (2015) 'MicroRNAs involved with hepatocellular carcinoma ', *Oncology Reports*, 34(6), pp. 2811-2820

Marcellin P., Chang T.T., Lim S.G., Sievert W., Tong M., Arterburn S., Borroto-Esoda K., Frederick D. and Rousseau F. (2008) 'Long-term efficacy and safety of adefovir dipivoxil for the treatment of hepatitis B e antigen–positive chronic hepatitis B', *Hepatology*, 48(3), pp. 750-758.

Matsumura T., Hu Z., Kato T. and et al. (2009) 'Amphipathic DNA polymers inhibit hepatitis C virus infection by blocking viral entry', *Gastroenterology*, 137, pp. 673-681.

Maucort-Boulch D., de Martel C., Franceschi S. and Plummer M. (2018) 'Fraction and incidence of liver cancer attributable to hepatitis B and C viruses worldwide', *Int J Cancer*, 142(12), pp. 2471-2477.

Mavilia, M.G. and Wu, G.Y. (2018) 'HBV-HCV Coinfection: Viral Interactions, Management, and Viral Reactivation', *J Clin Transl Hepatol*, 6(3), pp. 296-305.

McAleer W.J., Buynak E.B., Maigetter R.Z., Wampler D.E., Miller W.J. and Hilleman M.R. (1984) 'Human hepatitis B vaccine from recombinant yeast', *Nature*, 307(5947), pp. 178-80.

McClain C.J., Song Z., Barve S.S., Hill D.B. and Deaciuc I. (2004) 'Recent advances in alcoholic liver disease. IV. Dysregulated cytokine metabolism in alcoholic liver disease', *American journal of physiology Gastrointestinal and liver physiology*, 287, pp. G497-502.

McGivern D.R., Villanueva R.A., Chinnaswamy S., Kao C.C. and Lemon S.M. (2009) 'Impaired replication of hepatitis C virus containing mutations in a conserved NS5B retinoblastoma protein-binding motif', *J Virol*, 83(15), pp. 7422-33.

McHutchison J.G., Everson G.T., Gordon S.C. and et al. (2009) 'Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection', *N Engl J Med*, 360, pp. 1827-1838.

Medvedev, R., Ploen, D. and Hildt, E. (2016) 'HCV and Oxidative Stress: Implications for HCV Life Cycle and HCV-Associated Pathogenesis', *Oxid Med Cell Longev*, 2016, p. 9012580.

Meuleman P., Bukh J., Verhoye L. and et al. (2011) 'In vivo evaluation of the cross-genotype neutralising activity of polyclonal antibodies against hepatitis C virus', *Hepatology*, 53, pp. 755-762.

Miki Y., Nishisho I., Horii A., Miyoshi Y., Utsunomiya J., Kinzler K.W. and et al. (1992) 'Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer', *Cancer Research*, 52, pp. 643-645.

Mita, P., Sun, X., Fenyo, D., Kahler, D.J., Li, D., Agmon, N., Wudzinska, A., Keegan, S., Bader, J.S., Yun, C. and Boeke, J.D. (2020) 'BRCA1 and S phase DNA repair pathways restrict LINE-1 retrotransposition in human cells', *Nat Struct Mol Biol*, 27(2), pp. 179-191.

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D. and Kazazian, H.H. (1996) 'High frequency retrotransposition in cultured mammalian cells', *Cell*, 87(5), pp. 917-27.

Morera L., Lubbert M. and Jung M. (2016) 'Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy', *Clinical Epigenetics*, 8(57).

Morikawa T., Kuchiba A., Liao X., Imamura Y., Yamauchi M., Qian Z.R. and et al. (2012) 'Tumor TP53 expression status, body mass index, and prognosis in colorectal cancer', *International Journal of Cancer*, 131(5), pp. 1169-1178.

Morrish T.A., Garcia-Perez J.L., Stamato T.D., Taccioli G.E., Sekiguchi J. and Moran J.V. (2007) 'Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres', *Nature*, 446(7132), pp. 208-212.

Morrish T.A., Gilbert N., Myers J.S., Vincent B.J., Stamato T.D., Taccioli G.E., Batzer M.A. and Moran J.V. (2002) 'DNA repair mediated by endonuclease-independent LINE-1 retrotransposition', *Nature genetics*, 31(2), pp. 159-165.

Morse B., Rotherg P.G., South V.J., Spandorfer J.M. and Astrin S.M. (1988) 'Insertional mutagenesis of the myc locus by a LINE-1 sequence in a human breast carcinoma', *Nature* 333, pp. 87-90.

Munakata T., Nakamura M., Liang Y., Li K. and Lemon S.M. (2005) 'Down-regulation of the retinoblastoma tumor suppressor by the hepatitis C virus NS5B RNA-dependent RNA polymerase', *Proc Natl Acad Sci U S A*, 102(50), pp. 18159-64.

Muotri A.R., Chu V.T., Marchetto M.C., Deng W., Moran J.V. and Gage F.H. (2005) 'Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition', *Nature*, 435(7044), pp. 903-910.

Nakayama, R., Ueno, Y., Ueda, K. and Honda, T. (2019) 'Latent infection with Kaposi's sarcoma-associated herpesvirus enhances retrotransposition of long interspersed element-1', *Oncogene*, 38(22), pp. 4340-4351.

Nguyen, T.T.T., Park, E.M., Lim, Y.S. and Hwang, S.B. (2018) 'Nonstructural Protein 5A Impairs DNA Damage Repair: Implications for Hepatitis C Virus-Mediated Hepatocarcinogenesis', *J Virol*, 92(11).

Okamoto, Y., Shinjo, K., Shimizu, Y., Sano, T., Yamao, K., Gao, W., Fujii, M., Osada, H., Sekido, Y., Murakami, S., Tanaka, Y., Joh, T., Sato, S., Takahashi, S., Wakita, T., Zhu, J., Issa, J.P. and Kondo, Y. (2014) 'Hepatitis virus infection affects DNA methylation in mice with humanized livers', *Gastroenterology*, 146(2), pp. 562-72.

Orecchini E., Doria M., Antonioni A., Galardi S., Ciafrè S.A., Frassinelli L., Mancone C., Montaldo C., Tripodi M. and Michienzi A. (2017a) 'ADAR1 restricts LINE-1 retrotransposition', *Nucliec Acid Research*, 45(1), pp. 155-168.

Orecchini E., Frassinelli L., Galardi S., Ciafrè S.A. and Michienzi A. (2018) 'Post-transcriptional regulation of LINE-1 retrotransposition by AID/APOBEC and ADAR deaminases', *Chromosome Research*, 26(1-2), pp. 45-59.

Orecchini E., Frassinelli L. and Michienzi A. (2017b) 'Restricting retrotransposons: ADAR1 is another guardian of the human genome.', *RNA Biology*, 22, pp. 1-7.

Pace J. K. and Feschotte C. (2007) 'The evolutionary history of human DNA transposons: Evidence for intense activity in the primate lineage', *Genome Research*, 17, pp. 422-432.

Page A., Paoli P.P., Hill S.J. and et al. (2015) 'Alcohol directly stimulates epigenetic modifications in hepatic stellate cells', *J Hepatol*, 62(2), pp. 388-397.

Patel V. (2017) *Deaths registered in England and Wales: 2016*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsregistrationsummarytables/2016.

Pavel M., Renna M., Park S.J., Menzies F.M., Ricketts T., Füllgrabe J., Ashkenazi A., Frake R.A., Lombarte A.C., Bento C.F., Franze K. and Rubinsztein D.C. (2018) 'Contact inhibition controls cell survival and proliferation via YAP/TAZ-autophagy axis', *Nature Communications*, 9(1), p. 2961.

Pereira Tde A., Witek R.P., Syn W.K. and et al. (2010) 'Viral factors induce Hedgehog pathway activation in humans with viral hepatitis, cirrhosis, and hepatocellular carcinoma', *Lab Invest*, 90(12), pp. 1690-703.

Perez, S., Kaspi, A., Domovitz, T., Davidovich, A., Lavi-Itzkovitz, A., Meirson, T., Alison Holmes, J., Dai, C.Y., Huang, C.F., Chung, R.T., Nimer, A., El-Osta, A., Yaari, G., Stemmer, S.M., Yu, M.L., Haviv, I. and Gal-Tanamy, M. (2019) 'Hepatitis C virus leaves an epigenetic signature post cure of infection by direct-acting antivirals', *PLoS Genet*, 15(6), p. e1008181.

Perz J.F., Armstrong G.L., Farrington L.A., Hutin Y.J. and Bell B.P. (2006) 'The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. ', *Journal of Hepatology*, 45(4), pp. 529-538.

Pez F., Gifu P., Degli-Esposti D., Fares N., Lopez A., Lefrançois L., Michelet M., Rivoire M., Bancel B., Sylla B.S., Herceg Z., Merle P. and Caron de Fromentel C. (2019) 'In vitro transformation of primary human hepatocytes: Epigenetic changes and stemness properties', *Exp Cell Res*, 384(2), p. 111643.

Pezic D., Manakov S.A., Sachidanandam R. and A.A., A. (2014) 'piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells', *Genes and Development*, 28, pp. 1410-28.

Philippe C., Vargas-Landin D.B., Doucet A.J., van Essen D., Vera-Otarola J., Kuciak M., Corbin A., Nigumann P. and Cristofari G. (2016) 'Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci', *Elife*, 5.

Piratvisuth T., Lau G., Chao Y.C., Jin R., Chutaputti A., Zhang Q.B., Tanwandee T., Button P. and Popescu M. (2008) 'Sustained response to peginterferon alfa-2a (40 kD) with or without lamivudine in Asian patients with HBeAg-positive and HBeAg-negative chronic hepatitis B', *Hepatol Int*, 2(1), pp. 102-10.

Platt, L., Easterbrook, P., Gower, E., McDonald, B., Sabin, K., McGowan, C., Yanny, I., Razavi, H. and Vickerman, P. (2016) 'Prevalence and burden of HCV co-infection in people living with HIV: a global systematic review and meta-analysis', *Lancet Infect Dis*, 16(7), pp. 797-808.

Ploss A. and Dubuisson J. (2012) 'New advances in the molecular biology of hepatitis C virus infection: towards the identification of new treatment targets', *Gut*, 61, pp. i25-i35.

Poordad F., McCone J Jr., Bacon B.R. and et al. (2011) 'SPRINT-2 Investigators. Boceprevir for untreated chronic HCV genotype 1 infection', *N Engl J Med* 364(1195-1206).

Prensner J.R. and Chinnaiyan A.M. (2011) 'The emergence of lncRNAs in cancer biology', *Cancer Discovery*, 1(5), pp. 391-407.

Rahbari R., Habibi L., Garcia-Puche J., Badge R. and J., G.-P. (2015) *LINE-1 Retrotransposons and Their Role in Cancer*. Springer, Dordrecht.

Rao C.V., Asch A.S. and Yamada H.Y. (2017) 'Frequently mutated genes/pathways and genomic instability as prevention targets in liver cancer', *Carcinogenesis*, 38(1), pp. 2-11.

Reig, M., Boix, L., Marino, Z., Torres, F., Forns, X. and Bruix, J. (2017) 'Liver Cancer Emergence Associated with Antiviral Treatment: An Immune Surveillance Failure?', *Semin Liver Dis*, 37(2), pp. 109-118.

Reig, M., Marino, Z., Perello, C., Inarrairaegui, M., Ribeiro, A., Lens, S., Diaz, A., Vilana, R., Darnell, A., Varela, M., Sangro, B., Calleja, J.L., Forns, X. and Bruix, J. (2016) 'Unexpected high rate of early tumor recurrence in patients with HCV-related HCC undergoing interferon-free therapy', *J Hepatol*, 65(4), pp. 719-26.

Revill K., Wang T., Lachenmayer A., Kojima K., Harrington A., Li J., Hoshida Y., Llovet J.M. and Powers S. (2013) 'Genome-wide methylation analysis and epigenetic unmasking identify tumor suppressor genes in hepatocellular carcinoma', *Gastroenterology*, 145(6), pp. 1424-35.e1-25.

Richardson S.R., Narvaiza I., Planegger R.A., Weitzman M.D. and Moran J.V. (2014) 'APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition', *eLife*, 3, pp. e02008-e02008.

Rinaldi L., Nevola R., Franci G., Perrella A., Corvino G., Marrone A., Berretta M., Morone M.V., Galdiero M., Giordano M., Adinolfi L.E. and Sasso F.C. (2020) 'Risk of Hepatocellular Carcinoma after HCV Clearance by Direct-Acting Antivirals Treatment Predictive Factors and Role of Epigenetics', *Cancers (Basel)*, 12(6).

Robertson K.D. (2005) 'DNA methylation and human disease', *Nature Review Genetics*, 6, pp. 597-610.

Rocco A., Compare D., Angrisani D., Sanduzzi Zamparelli M. and Nardone G. (2014) 'Alcoholic disease: Liver and beyond', *World J Gastroenterol*, 20(40), pp. 14652-14659.

Rodriguez-Martin, B., Alvarez, E.G., Baez-Ortega, A., Zamora, J., Supek, F., Demeulemeester, J., Santamarina, M., Ju, Y.S., Temes, J., Garcia-Souto, D., Detering, H., Li, Y., Rodriguez-Castro, J., Dueso-Barroso, A., Bruzos, A.L., Dentro, S.C., Blanco, M.G., Contino, G., Ardeljan, D., Tojo, M., Roberts, N.D., Zumalave, S., Edwards, P.A.W., Weischenfeldt, J., Puiggros, M., Chong, Z., Chen, K., Lee, E.A., Wala, J.A., Raine, K., Butler, A., Waszak, S.M., Navarro, F.C.P., Schumacher, S.E., Monlong, J., Maura, F., Bolli, N., Bourque, G., Gerstein, M., Park, P.J., Wedge, D.C., Beroukhim, R., Torrents, D., Korbel, J.O., Martincorena, I., Fitzgerald, R.C., Van Loo, P., Kazazian, H.H., Burns, K.H., Group, P.S.V.W., Campbell, P.J., Tubio, J.M.C. and Consortium, P. (2020) 'Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition', *Nat Genet*, 52(3), pp. 306-319.

Rodriguez-Martin B., Alvarez E.G., Baez-Ortega A., Zamora J., Supek F. and et al. (2020) 'Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition', *Nature Genetics*, 52(3), pp. 306-319.

Roman-Gomez J., Jimenez-Velasco A., Agirre X., Cervantes F., Sanchez J., Garate L., Barrios M. and et al. (2005) 'Promoter hypomethylation of the LINE-1 retrotransposable elementsactivates sense/antisense transcription and marks the progression of chronic myeloid leukemia', *Oncogene*, 24(48), pp. 7213-7223.

Rongrui, L., Na, H., Zongfang, L., Fanpu, J. and Shiwen, J. (2014) 'Epigenetic mechanism involved in the HBV/HCV-related hepatocellular carcinoma tumorigenesis', *Curr Pharm Des*, 20(11), pp. 1715-25.

Ronson G.E., Piberger A.L., Higgs M.R., Olsen A.L., Stewart G.S., McHugh P.J., Petermann E. and Lakin N.D. (2018) 'PARP1 and PARP2 stabilise replication forks at base excision repair intermediates through Fbh1-dependent Rad51 regulation', *Nature Communications*, 9(1), p. 746.

Rubin H. (2017) 'Dynamics of cell transformation in culture and its significance for tumor development in animals', *Proceedings of the National Academy of Sciences*, 114(46), pp. 12237-12242.

Rundle, S., Bradbury, A., Drew, Y. and Curtin, N.J. (2017) 'Targeting the ATR-CHK1 Axis in Cancer Therapy', *Cancers (Basel)*, 9(5).

Ryan, E.L., Hollingworth, R. and Grand, R.J. (2016) 'Activation of the DNA Damage Response by RNA Viruses', *Biomolecules*, 6(1), p. 2.

Sanchez-Luque F.J., Kempen M.H.C., Gerdes P., Vargas-Landin D.B., Richardson S.R., Troskie R.L., Jesuadian J.S., Cheetham S.W., Carreira P.E., Salvador-Palomeque C., García-Cañadas M., Muñoz-Lopez M., Sanchez L., Lundberg M., Macia A., Heras S.R., Brennan P.M., Lister R., Garcia-Perez J.L., Ewing A.D. and Faulkner G.J. (2019) 'LINE-1 Evasion of Epigenetic Repression in Humans', *Mol Cell*, 75(3), pp. 590-604.e12.

Santoro A., Rimassa L., Borbath I., Daniele B., Salvagni S., Van Laethem J.L. and et al. (2013) 'Tivantinib for second-line treatment of advanced hepatocellular carcinoma: a randomised, placebo-controlled phase 2 study', *Lancet Oncology*, 14, pp. 55-63.

Sasaki M., Ikeda H., Itatsu K., Yamaguchi J., Sawada S., Minato H., Ohta T. and Nakanuma Y. (2008) 'The overexpression of polycomb group proteins Bmi1 and EZH2 is associated with the progression and aggressive biological behavior of hepatocellular carcinoma', *Laboratory Investigation*, 88(8), pp. 873-882.

Sato T., Issa J.J. and Kropf P. (2017) 'DNA Hypomethylating Drugs in Cancer Therapy', *Cold Spring Harb Perspect Med.*, 7(5), p. a026948.

Sato Y. and Tsurumi T. (2013) 'Genome guardian p53 and viral infections', *Rev Med Virol*, 23(4), pp. 213-20.

Satoh S., Daigo Y., Furukawa Y. and et al. (200) 'AXIN1 mutations in hepatocellular carcinomas, and growth suppression in cancer cells by virus-mediated transfer of AXIN1', *Nat Genet*, 24, pp. 245-250.

Sawey E.T., Chanrion M., Cai C. and et al. (2011) 'Identification of a therapeutic strategy targeting amplified FGF19 in liver cancer by oncogenomic screening', *Cancer Cell*, 19, pp. 347-358.

Schauer S.N., Carreira P.E., Shukla R., Gerhardt D.J., Gerdes P., Sanchez-Luque F.J., Nicoli P., Kindlova M., Ghisletti S., Santos A.D., Rapoud D., Samuel D., Faivre J., Ewing A.D., Richardson S.R. and Faulkner G.J. (2018) 'L1 retrotransposition is a common feature of mammalian hepatocarcinogenesis', *Genome Res*, 28(5), pp. 639-653.

Schauer, S.N., Carreira, P.E., Shukla, R., Gerhardt, D.J., Gerdes, P., Sanchez-Luque, F.J., Nicoli, P., Kindlova, M., Ghisletti, S., Santos, A.D., Rapoud, D., Samuel, D., Faivre, J., Ewing, A.D., Richardson,

S.R. and Faulkner, G.J. (2018) 'L1 retrotransposition is a common feature of mammalian hepatocarcinogenesis', *Genome Res*, 28(5), pp. 639-653.

Schiano T.D., CHarlton M., Younossi Z. and et al. (2006) 'Monoclonal antibody HCV-AbXTL68 in patients undergoing liver transplantation for HCV: results of a phase 2 randomised study', *Liver Transpl*, 12, pp. 1381-1389.

Schobel, A., Nguyen-Dinh, V., Schumann, G.G. and Herker, E. (2021) 'Hepatitis C virus infection restricts human LINE-1 retrotransposition in hepatoma cells', *PLoS Pathog*, 17(4), p. e1009496.

Schulze K., Imbeaud S., Letouze E. and et al. (2015) 'Exome sequencing of hepatocellular carcinoma identifies new mutations signatures and potential therapeutic targets', *Nat Genet*, 47, pp. 505-511.

Schumann G.G. (2007) 'APOBEC3 proteins: major players in intracellular defence against LINE-1-mediated retrotransposition', *Biochemical Society*, 35(3), pp. 637-642.

Schwenzer N.F., Springer F., Schraml C., Stefan N., Machann J. and Schick F. (2009) 'Non-invasive assessment and quantification of liver steatosis by ultrasound, computed tomography and magnetic resonance', *J. Hepatol.*, 51, pp. 433-445.

Scott E.C., Gardner E.J., Masood A., Chuang N.T., Vertino P.M. and Devine S.E. (2016) 'A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer', *Genome Res*, 26(6), pp. 745-55.

Scourzic L., Mouly E. and Bernard O.A. (2015) 'TET proteins and the control of cytosine demethylation in cancer', *Genome Med*, 7, p. 9.

Shahnazari P., Sayehmiri K., Minuchehr Z. and et al. (2014) 'The Increased Level of Serum p53 in Hepatitis B-Associated Liver Cirrhosis', *Iranian Journal of Medical Sciences*, 39(5), pp. 446-451.

Sharma S., Kelly T.K. and Jones P.A. (2010) 'Epigenetics in cancer', *Carcinogenesis*, 31(1), pp. 27-36.

Shigaki H., Baba Y., Watanabe M., Murata A., Iwagami S. and et al. (2013) 'LINE-1 hypomethylation in gastric cancer, detected by bisulfite pyrosequencing, is associated with poor prognosis', *Gastric Cancer*, 16(4), pp. 480-487.

Shukla, R., Upton, K.R., Muñoz-Lopez, M., Gerhardt, D.J., Fisher, M.E., Nguyen, T., Brennan, P.M., Baillie, J.K., Collino, A., Ghisletti, S., Sinha, S., Iannelli, F., Radaelli, E., Dos Santos, A., Rapoud, D., Guettier, C., Samuel, D., Natoli, G., Carninci, P., Ciccarelli, F.D., Garcia-Perez, J.L., Faivre, J. and Faulkner, G.J. (2013) 'Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma', *Cell*, 153(1), pp. 101-11.

Shukla R., Upton K.R., Munoz-Lopez M., Gerhardt D.J., Fisher M.E., Nguyen T., Brennan P.M., Baillie J.K., Collino A., Ghisletti S. and et al. (2013a) 'Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma', *Cell*, 153(1), pp. 101-111.

Shukla R., Upton K.R., Munoz-Lopez M., Gerhardt D.J., Fisher M.E., Nguyen T. and et al. (2013b) 'Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma', *Cell*, 153(1), pp. 101-111.

Sigurdsson M.I., Smith A.V., Bjornsson H.T. and Jonsson J.J. (2012) 'The distribution of a germline methylation marker suggests a regional mechanism of LINE-1 silencing by the piRNA-PIWI system', *BMC Genetics*, 13(31).

Soifer H.S., Zaragoza A. and et al. (2005) 'A potential role for RNA interference in controlling the activity of the human LINE-1 retrotransposon', *Nucleic Acid Research*, 33(3), pp. 846-856.

Song M.A., Tiirikainen M., Kwee S., Okimoto G., Yu H. and Wong L.L. (2013) 'Elucidating the landscape of aberrant DNA methylation in hepatocellular carcinoma', *PLoS One*, 8(2), p. e55761.

Speek M. (2001) 'Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes', *Mol Cell Biol*, 21(6), pp. 1973-85.

Sporn M.B. and Liby K.T. (2012) 'NRF2 and cancer: the good, the bad and the importance of context', *Nature*, 12, pp. 564-571.

Sudhindar, P.D., Wainwright, D., Saha, S., Howarth, R., McCain, M., Bury, Y., Saha, S.S., McPherson, S., Reeves, H., Patel, A.H., Faulkner, G.J., Lunec, J. and Shukla, R. (2021) 'HCV Activates Somatic L1 Retrotransposition—A Potential Hepatocarcinogenesis Pathway', *Cancers*, 13(20), p. 5079.

Sun X., Wang X., Tang Z., Grivainis M., Kahler D., Yun C., Mita P., Fenyö D. and Boeke J. D. (2018) 'Transcription factor profiling reveals molecular choreography and key regulators of human retrotransposon expression', *Proc Natl Acad Sci U S A*, 115(24), pp. E5526-e5535.

Tardif K., Waris G. and Siddiqui A. (2005) 'Hepatitis C virus, ER stress, and oxidative stress', *Trends in microbiology*, 13, pp. 159-63.

Tchénio T., Casella J.F. and Heidmann T. (2000) 'Members of the SRY family regulate the human LINE retrotransposons.', *Nucleic Acid Research*, 28(2), pp. 411-415.

Teneng I., Montoya-Durango D.E., Quertermous J.L., Lacy M.E. and Ramos K.S. (2011) 'Reactivation of L1 retrotransposon by benzo(a)pyrene involves complex genetic and epigenetic regulation', *Epigenetics*, 6(3), pp. 355-367.

Thiery J.P. and Sleeman J.P. (2006) 'Complex networks orchestrate epithelial-mesenchymal transitions', *Nat Rev Mol Cell Biol*, 7(2), pp. 131-42.

Tian Y., Wong V.W., Wong G.L., Yang W., Sun H., Shen J., Tong J.H. and et al. (2015) 'Histone deacetylase HDAC8 promotes insulin resistance and beta-catenin activation in NAFLD-associated hepatocellular carcinoma', *Cancer Research*, 75(22), pp. 4803–4816.

Totoki Y., Tatsuno K., Covington K. and et al. (2014) 'Trans-ancestry mutational landscape of hepatocellular carcinoma genomes', *Nat Genet*, 46, pp. 1267-1273.

Tourancheau A., Mead E.A., Zhang X. and Fang G. (2020) 'Discovering and exploiting multiple types of DNA methylation from individual bacteria and microbiome using nanopore sequencing', *bioRxiv*, p. 2020.02.18.954636.

Trepo C. (2014) 'A brief history of hepatitis milestones', *Liver Int*, 34 Suppl 1, pp. 29-37.

Tsukuda S. and Watashi Koichi. (2020) 'Hepatitis B virus biology and life cycle', *Antiviral Research*, 182, p. 104925.

Tubio J.M.C., Li Y., Ju Y.S. and et al. (2014) 'Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes', *Science* 345(6196).

Tufarelli C. and Badge R.M. (2017) 'Retrotransposon-Driven Transcription and Cancer', in Cristopari G. (ed.) *Retrotransposon-Driven Transcription and Cancer*. Springer, Cham, pp. 259-273.

Upton K.R., Gerhardt D.J., Jesuadian J.S., Richardson S.R., Sánchez-Luque F.J., Bodea G.O., Ewing A.D., Salvador-Palomeque C., van der Knaap M.S., Brennan P.M., Vanderver A. and Faulkner G.J. (2015) 'Ubiquitous L1 mosaicism in hippocampal neurons', *Cell*, 161(2), pp. 228-39.

Vafadar-Isfahani N., Parr C., McMillan L.E., Sanner J., Yeo Z., Saddington S., Peacock O., Cruickshanks H.A., Meehan R.R., Lund J.N. and Tufarelli C. (2017) 'Decoupling of DNA methylation and activity of intergenic LINE-1 promoters in colorectal cancer', *Epigenetics*, 12(6), pp. 465-475.

Van Driel R., Fransz P.F. and Verschure P.J. (2003) 'The eukaryotic genome: a system regulated at different hierarchical levels', *Journal of Cell Science*, 116(20), pp. 4067-4075.

Vandenbossche J., Jessner W., van den Boer M., Biewenga J., Berke J.M., Talloen W., De Zwart L., Snoeys J. and Yogaratnam J. (2019) 'Pharmacokinetics, Safety and Tolerability of JNJ-56136379, a Novel Hepatitis B Virus Capsid Assembly Modulator, in Healthy Subjects', *Advances in Therapy*, 36(9), pp. 2450-2462.

Vescovo, T., Refolo, G., Vitagliano, G., Fimia, G.M. and Piacentini, M. (2016) 'Molecular mechanisms of hepatitis C virus–induced hepatocellular carcinoma', *Clinical Microbiology and Infection*, 22(10), pp. 853-861.

Villanueva A., Portela A., Sayols S., Battiston C., Hoshida Y., Méndez-González J., Imbeaud S., Letouzé E., Hernandez-Gea V. and et al. (2015) 'DNA methylation-based prognosis and epidrivers in hepatocellular carcinoma', *Hepatology*, 61(6), pp. 1945-1956.

Viré E., Brenner C., Deplus R., Blanchon L., Fraga M., Didelot C., Morey L., Van Eynde A., Bernard D., Vanderwinden J.M., Bollen M., Esteller M., Di Croce L., de Launoit Y. and Fuks F. (2006) 'The Polycomb group protein EZH2 directly controls DNA methylation', *Nature*, 439(7078), pp. 871-4.

Virzì A., Roca Suarez A.A., Baumert T.F. and Lupberger J. (2018) 'Oncogenic Signaling Induced by HCV Infection', *Viruses*, 10(10).

Walter M., Teissandier A., Perez-Palacios R. and Bourc'his D. (2016) 'An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells', *eLife*, 5, p. e11418.

Wang M. and Kaufman R.J. (2014) 'The impact of the endoplasmic reticulum protein-folding environment on cancer development', *Nat Rev Cancer*, 14(9), pp. 581-97.

Wei W., Gilbert N., Ooi S.L., Lawler J.F., Ostertag E.M., Kazazian H.H., Boeke J.D. and Moran J.V. (2001) 'Human L1 retrotransposition: cis preference versus trans complementation', *Mol Cell Biol*, 21(4), pp. 1429-39.

Wenjin Z., Chuanhui P., Yunle W., Lateef S.A. and Shusen Z. (2012) 'Longitudinal fluctuations in PD1 and PD-L1 expression in association with changes in anti-viral immune response in chronic hepatitis B', *BMC Gastroenterol*, 12, p. 109.

'WHO (2017) Global Hepatitis Report', (2017).

Wijetunga, N.A., Pascual, M., Tozour, J., Delahaye, F., Alani, M., Adeyeye, M., Wolkoff, A.W., Verma, A. and Greally, J.M. (2017) 'A pre-neoplastic epigenetic field defect in HCV-infected liver at transcription factor binding sites and polycomb targets', *Oncogene*, 36(14), pp. 2030-2044.

Wilson C.L., Mann D.A. and Borthwick L.A. (2017) 'Epigenetic reprogramming in liver fibrosis and cancer', *Advanced Drug Delivery Reviews*, 121, pp. 124-132.

Wissing S., Montano M., Garcia-Perez J.L., Moran J.V. and Greene W.C. (2011) 'Endogenous APOBEC3B restricts LINE-1 retrotransposition in transformed cells and human embryonic stem cells', *The Journal of Biological Chemistry*, 286(42), pp. 36427-37.

Wolff E.M., Byun H.M., Han H.F., Sharma S., Nichols P.W., Siegmund K.D., Yang A.S., Jones P.A. and Liang G. (2010) 'Hypomethylation of a LINE-1 promoter activates an alternate transcript of the MET oncogene in bladders with cancer', *PLoS Genetics* 6(4).

Wong D.K., Cheung A.M., O'Rourke K., Naylor C.D., Detsky A.S. and Heathcote J. (1993) 'Effect of alpha-interferon treatment in patients with hepatitis B e antigen-positive chronic hepatitis B. A meta-analysis', *Ann Intern Med*, 119(4), pp. 312-23.

Wooddell C.I., Yuen M.F., Chan H.L., Gish R.G., Locarnini S.A., Chavez D., Ferrari C., Given B.D., Hamilton J., Kanner S.B., Lai C.L., Lau J.Y.N., Schluep T., Xu Z., Lanford R.E. and Lewis D.L. (2017) 'RNAi-based treatment of chronically infected patients and chimpanzees reveals that integrated hepatitis B virus DNA is a source of HBsAg', *Sci Transl Med*, 9(409).

Workman J.L. and Langmore J.P. (1985) 'Nucleoprotein hybridization: a method for isolating specific genes as high molecular weight chromatin', *Biochemistry*, 24(25), pp. 7486-7497.

World Health Organisation. Global status report on alcohol and health 2018. (2018). Available at: https://apps.who.int/iris/handle/10665/274603.

Wu X., Li J., Gassa A., Buchner D., Alakus H., Dong Q., Ren N., Liu M., Odenthal M., Stippel D., Bruns C., Zhao Y. and Wahba R. (2020) 'Circulating tumor DNA as an emerging liquid biopsy biomarker for early diagnosis and therapeutic monitoring in hepatocellular carcinoma', *Int J Biol Sci*, 16(9), pp. 1551-1562.

Xiao-Jie L., Hui-Ying X., Qi X., Jiang X. and Shi-Jie M. (2016) 'LINE-1 in cancer: multifaceted functions and potential clinical implications.', *Genetics in medicine : official journal of the American College of Medical Genetics.*, 18(5), pp. 431-439.

Xu R.H., Wei W., Krawczyk M., Wang W., Luo H., Flagg K., Yi S., Shi W., Quan Q., Li K., Zheng L., Zhang H., Caughey B.A., Zhao Q. and et al. (2017) 'Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma', *Nat Mater*, 16(11), pp. 1155-1161.

Yang N. and Kazazian H.H., J. (2006) 'L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells.', *Nature Structural and Molecular Biology*, 13(9), pp. 763-771.

Yang N., Zhang L., Zhang Y. and Jr, K.H.H. (2003) 'An important role for RUNX3 in human L1 transcription and retrotransposition', *Nucleic Acid Research*, 31(16), pp. 4929-40.

Yokosuka O., Takaguchi K., Fujioka S., Shindo M., Chayama K., Kobashi H., Hayashi N., Sato C., Kiyosawa K., Tanikawa K., Ishikawa H., Masaki N., Seriu T. and Omata M. (2010) 'Long-term use of entecavir in nucleoside-naïve Japanese patients with chronic hepatitis B infection', *J Hepatol*, 52(6), pp. 791-9.

Yu, Q., Carbone, C.J., Katlinskaya, Y.V., Zheng, H., Zheng, K., Luo, M., Wang, P.J., Greenberg, R.A. and Fuchs, S.Y. (2015) 'Type I interferon controls propagation of long interspersed element-1', *J Biol Chem*, 290(16), pp. 10191-9.

Zhang H. (2018) 'Safety, pharmacokinetics and anti-viral efficacy of novel core protein allosteric modifier GLS4 in patients with chronic hepatitis B: interim results from a 48 weeks phase 2a study', *Hepatology : Official Journal of the American Association for the Study of Liver Diseases*, 68(6), p. 1454A.

Zhang P., Wen X., Gu F., Deng X., Li J., Dong J., Jiao J. and Tian Y. (2013) 'Methylation profiling of serum DNA from hepatocellular carcinoma patients using an Infinium Human Methylation 450 BeadChip', *Hepatol Int*, 7(3), pp. 893-900.

Zhang X.F. and et al. (2015) 'Overexpression of WWP1 promotes tumorigenesis and predicts unfavorable prognosis in patients with hepatocellular carcinoma', *Oncotarget*, 6, pp. 40920-40933.

Zhang X.Y. and Hörz W. (1982) 'Analysis of highly purified satellite DNA containing chromatin from the mouse', *Nucleic Acids Research*, 10(5), pp. 1481-1494.

Zhang, Z., Rong, L. and Li, Y.P. (2019) 'Flaviviridae Viruses and Oxidative Stress: Implications for Viral Pathogenesis', *Oxid Med Cell Longev*, 2019, p. 1409582.

Zhao L.H., Liu X., Yan H.X. and et al. (2016) 'Genomic and oncogenic preference of HBV integration in hepatocellular carcinoma', *Nature Communications*, 7(1), p. 12992.

Zheng D.L., Zhang L., Cheng N., Xu X., Deng Q., Teng X.M., Wang K.S., Zhang X., Huang J. and Han Z.G. (2009) 'Epigenetic modification induced by hepatitis B virus X protein via interaction with de novo DNA methyltransferase DNMT3A', *Journal of Hepatology*, 50(2), pp. 377-387.

Zheng Y., Huang Q., Ding Z., Liu T., Xue C., Sang X. and J., G. (2016) 'Genome-wide DNA methylation analysis identifies candidate epigenetic markers and drivers of hepatocellular carcinoma.', *Breifings in Bioinformatics*, 19(1), pp. 101-108.

Zhu A.X., Gold P.J., El-Khoueiry A.B., Abrams T.A., Morikawa H., Ohishi N., Ohtomo T. and Philip P.A. (2013) 'First-in-man phase I study of GC33, a novel recombinant humanized antibody against glypican-3, in patients with advanced hepatocellular carcinoma', *Clinical Cancer Research*, 19(4), pp. 920-928.

Zhu A.X., Park J.O., Ryoo B.Y., Yen C.J., Poon R., Pastorelli D. and et al. (2015) 'Ramucirumab versus placebo as second-line treatment in patients with advanced hepatocellular carcinoma following first-line therapy with sorafenib (REACH): a randomised, double-blind, multicentre, phase 3 trial', *Lancet Oncology*, 16, pp. 859-870.

Zhu C., Utsunomiya T., Ikemoto T., Yamada S., Morine Y., Imura S. and et al. (2014) 'Hypomethylation of Long Interspersed Nuclear Element-1 (LINE-1) is Associated with Poor Prognosis via Activation of c-MET in Hepatocellular Carcinoma', *Annals of Surgical Oncology*, 21(4), pp. 729-735.

Zoulim, F. and Locarnini, S. (2009) 'Hepatitis B virus resistance to nucleos(t)ide analogues', *Gastroenterology*, 137(5), pp. 1593-608.e1-2.

Zucman-Rossi J. and et al. (2015) 'Genetic landscape and biomarkers of hepatocellular carcinoma', *Gastoenterology*, 149, pp. 1226-1239.