



Role of retrotransposons in hepatocellular carcinoma origin and progression – potential biomarker and therapeutic target?

Bassier Zadran

PhD Thesis

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

November 2021

Abstract

Introduction

Hepatocellular carcinoma (HCC) has the fourth highest cancer-related mortality worldwide and is associated with a poor 5-year survival. Current therapies are limited and provide short median survival. Thus, there is an urgent need to develop novel therapies. Long interspersed class elements-1 (LINE1 or L1) activation has been demonstrated in several cancers including HCC and can inhibit tumour suppressor genes or activate oncogenes. However, the role of L1 in hepatocarcinogenesis is still unknown.

Methods

L1 expression was evaluated in the RNAseq data of HCC (n=372) from the cancer genome atlas liver hepatocellular carcinoma (TCGA LIHC) study and formalin-fixed paraffin-embedded (FFPE) patient biopsies (n=48) from our own biobank. RNAseq data was analysed for L1 counts and their distribution was assessed in different HCC subclasses based on previously known molecular classifications and associations with clinical parameters were explored. Likewise, FFPE samples were stained for L1orf1p using an automated immunohistochemistry machine and were scored by a pathologist; associations between L1orf1p expression in HCC and clinical parameters like cirrhosis, tumour stage, albumin, bilirubin, alpha fetoprotein (AFP) and survival were explored. The role of L1 was further characterised in different liver cancer cell lines utilising L1 knockdown and overexpression systems. L1 knockdown was achieved using a lentivirus-based shRNA expression vector targeting L1orf1 in Huh-7 cells. The influence of L1-knockdown on functional properties such as proliferation, migration and invasion of the cells were investigated by comparing L1-knockdown cells with wild type and non-targeted controls. RNAseq evaluated the influence of L1-knockdown on whole transcriptome. Transient L1 full-length overexpression and conditional L1orf1p overexpression were used to further validate the influence of L1 on cell signalling pathways.

Results

L1 expression was elevated in HCC both at transcript and protein level compared to adjacent non-tumour tissues. L1 transcripts correlated with high AFP, TP53 mutation, macrovascular invasion and activated TGF- β signalling. Likewise, L1orf1p expression correlated with AFP, activated TGF- β signalling and poorly differentiated tumours. A positive association between L1orf1p and pSMAD3 confirmed the relationship between L1 expression and TGF- β signalling in HCC. L1 knockdown in Huh-7 cells led to decrease in migratory and invasion capacity of the cells compared to control cell lines. Furthermore, gene set enrichment analysis (GSEA) of the RNAseq data demonstrated downregulation of TGF- β pathway in Huh7-L1knockdown cells compared to non-

targeting control cells, which was confirmed by Pai1-luciferase reporter assay. Conversely, L1 overexpression (full-length and L1orf1 alone) increased TGF- β signalling as confirmed by Pai1-luciferase reporter assay, RT-qPCR and FACS analysis in HepG2, PLC/PRF-5 and HHL5 cell lines.

Conclusion

L1 is upregulated in human HCC and associated with high AFP, TP53 mutation and activated TGF- β signalling. Further in vitro studies demonstrated a crosstalk between L1orf1p and TGF β -signalling. Overall, our data demonstrates a causal link between L1orf1p and TGF β signalling, which presents a novel therapeutic avenue and potential treatment stratification biomarker for HCC.

Acknowledgements

I would like to thank my supervisor Ruchi Shukla for her kind support and help. I am very grateful for the time she dedicated supervising my working, meetings, and guidance. In addition I would like to thank my other supervisors John Lunec and Helen Reeves for help and support.

Furthermore, I would like to thank Yvonne Bury for immunohistochemical pathological scoring and Huw Thomas for technical help to implant HCC tumours into mice. I would also like to thank my colleagues for help and support.

Finally I would like to thank my parents for help and support.

Table of Content

| | |
|-------------------------------------------------------------------------------------------------------|-----------|
| Abstract | i |
| Acknowledgements | iii |
| Table of Contents..... | iv |
| List of figures..... | vii |
| List of tables | xi |
| List of Abbreviations..... | xii |
| | |
| Chapter 1. Introduction | 1 |
| 1.1. Hepatocellular Carcinoma | 1 |
| 1.1.1. Epidemiology..... | 1 |
| 1.1.2. Risk factors | 3 |
| 1.1.3. Diagnosis of HCC.. .. | 8 |
| 1.1.4. Staging of HCC | 11 |
| 1.1.5. Treatment approaches..... | 12 |
| 1.2. Molecular mechanisms of HCC..... | 20 |
| 1.3.L1 elements..... | 32 |
| 1.3.1. Structure of L1..... | 33 |
| 1.3.2.Mechanism ofL retrotransposition | 37 |
| 1.3.3. Mechanisms of L1 Repression | 39 |
| 1.3.4 L1 and diseases especially cancer..... | 42 |
| 1.3.5L1 retrotransposition in cancer..... | 44 |
| 1.3.6. L1 influence besides insertional mutagenesis..... | 46 |
| 1.4 Aims of the thesis..... | 49 |
| | |
| Chapter 2. L1 expression and correlation with HCC subclasses..... | 50 |
| 2.1 Introduction..... | 50 |
| 2.1 Aims..... | 53 |
| 2.3 Material and Methods..... | 54 |
| 2.3.1 Ethical Approval..... | 54 |
| 2.3.2 Patient cohort for immunohistochemical analysis..... | 54 |
| 2.3.3 Immunohistochemistry..... | 54 |
| 2.3.4 L1orf1p and pSMAD3 scoring..... | 55 |
| 2.3.5 L1orf1p algorithm scoring..... | 55 |
| 2.3.6 HCC RNAseq analysis..... | 56 |
| 2.3.7 Statistical analysis | 57 |
| 2.4 Results..... | 58 |
| 2.4.1 L1 5'UTR counts are elevated in HCC tumours..... | 59 |
| 2.4.2 HCC L1 5'UTR counts correlated with high pathological stage, recurrence and poor prognosis..... | 59 |
| 2.4.3 L1 5'UTR counts and correlation with HCC-related signalling pathways..... | 62 |
| 2.4.4 Expression of L1orf1 protein in HCC tissues..... | 59 |
| 2.4.5 L1orf1 protein expression is elevated in HCC..... | 68 |
| 2.4.6 L1 protein expression correlated with poorly differentiated tumours and increased AFP | 71 |
| 2.4.7 L1 high correlated with poor survival in TACE treated patients..... | 71 |
| 2.4.8 L1 correlated with TGF- β activation..... | 74 |
| 2.5Discussion..... | 76 |
| 2.6 Summary..... | 78 |

| | |
|-----------------------------------------------------------------------------------------------------------------|------------|
| Chapter 3. Effect of L1 knockdown on Huh7 cells | 79 |
| 3.1 Introduction..... | 79 |
| 3.11 TGF β signalling pathway..... | 79 |
| 3.12 TGF- β and HCC..... | 80 |
| 3.13. L1 and TGF- β | 83 |
| 2.3.2 Aims..... | 84 |
| 3.3 Materials and Methods..... | 85 |
| 3.31 Reagents..... | 85 |
| 3.32 Cell Culture..... | 85 |
| 3.33 Huh7-L1KD lentiviral particles..... | 85 |
| 3.34 Transduction into Huh7 cell line..... | 86 |
| 3.35 Western blot..... | 86 |
| 3.36 FACS | 87 |
| 3.37 Cell cycle by PI staining | 87 |
| 3.38 SMAD3 luciferase assay | 87 |
| 3.39 RNA isolation..... | 88 |
| 3.310 Reverse transcription | 88 |
| 3.311 Real-time PCR..... | 88 |
| 3.312 Cell proliferation..... | 92 |
| 3.313 Clonogenic assay | 92 |
| 3.314 Invasion assay | 93 |
| 3.315 RNA seq..... | 93 |
| 3.316 Mouse xenograft experiment | 94 |
| 3.317 Statistics..... | 94 |
| 3.4 Results | 95 |
| 3.4.1 L1 expression in various liver cancer cell lines..... | 95 |
| 3.4.2 Generation of Huh7-L1 knockdown (Huh7-L1KD) and Huh7-non-target (Huh7-NT) control cell lines..... | 98 |
| 3.4.3 Effect of L1 knockdown on phenotypic and functional properties of cells..... | 99 |
| 3.4.4 Influence of L1 knockdown on cellular transcriptome - RNAseq analysis of Huh7-NT and Huh7-L1KD cells..... | 104 |
| 3.4.5 Effect of L1 knockdown on the response of cells to TGF- β treatment..... | 114 |
| 3.4.6 Effect of L1 knockdown on TGF- β mediated cell growth inhibition..... | 116 |
| 3.4.7 Effect of L1 knockdown on doxorubicin sensitivity of Huh7 cells..... | 117 |
| 3.5 Discussion..... | 119 |
| 3.5.1 L1 expression in HCC cell lines..... | 119 |
| 3.5.2 Influence of L1 knockdown on Huh7 cells..... | 120 |
| 3.5.3 Huh7-L1KD and doxorubicin sensitivity..... | 121 |
| 3.6 Summary..... | 122 |
| | |
| Chapter 4: Effect of L1 overexpression on liver cancer cell lines | 123 |
| 4.1 Introduction..... | 123 |
| 4.12 L1orf1p and its interaction with different cellular pathways..... | 125 |
| 4.2 Aims..... | 126 |
| 4.3 Material and Methods..... | 127 |
| 4.3.1 Cell culture..... | 127 |
| 4.3.2 L1orf1p doxycycline (dox) inducible transfection..... | 127 |
| 4.3.3 SMAD3 luciferase assay..... | 127 |
| 4.3.4 FACS-dual L1orf1p and vimentin staining..... | 128 |
| 4.3.5 Plasmids..... | 128 |
| 4.3.6 Statistics..... | 128 |

| | |
|------------------------------------------------------------------------------------------------------------------------|------------|
| 4.4 Results..... | 129 |
| 4.4.1 Full length L1 overexpression activates TGF- β signalling..... | 129 |
| 4.4.2 L1 increases TGF- β signalling via L1orf1p..... | 131 |
| 4.5 Discussion..... | 141 |
| 4.6 Summary..... | 142 |
| Chapter 5: Discussion..... | 143 |
| 5.1 L1 transcript and protein expression in HCC and correlations with clinical parameters and signalling pathways..... | 143 |
| 5.2 L1orf1p activates TGF- β signalling in liver cancer cell lines..... | 146 |
| 5.3 Further studies | 149 |
| 5.4 Summary..... | 151 |
| References..... | 152 |

Lists of figures and tables

List of Figures

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1.1 Global disease burden of primary liver cancer (numbers are per 100,000 person)..... | 2 |
| Figure 1.2 Hepatitis B viral cycle..... | 4 |
| Figure 1.3 Hepatitis C viral cycle | 5 |
| Figure 1.4 BCLC staging system. The staging system predict prognosis and recommends treatment strategies by assessing patient tumour and liver related factors..... | 12 |
| Figure 1.5 Strategy of HCC treatments in countries with different resources..... | 17 |
| Figure 1.6 Pathogenesis of HCC; HD=high dysplastic, LD=low dysplastic | 21 |
| Figure 1.7 TGF- β pathway during early and late carcinogenesis. The pathway has tumour suppressive effects during early carcinogenesis and pro-tumour effects during later stages..... | 27 |
| Figure 1.8 Chromatin structure regulation and effect on gene expression | 30 |
| Figure 1.9 DNA methylation changes during carcinogenesis..... | 32 |
| Figure 1.10 structure of an active L1 element..... | 33 |
| Figure 1.11 Structure of L1orf1p consisting of a N-terminal domain, coiled coil domain, RNA recognition motif and a C-terminal domain. | 35 |
| Figure 1.12 Process of retrotransposition..... | 39 |
| Figure 1.13 L1 and different cellular mechanisms inhibiting L1 retrotransposition..... | 41 |
| Figure 1.14 L1 and its role during carcinogenesis | 46 |
| Figure 2.1 Molecular classification of HCC using S clusters | 50 |
| Figure 2.2 L1Ta transcript correlates with L1 Ta 5'UTR promoter counts..... | 58 |
| Figure 2.3 L1 5'UTR counts of tumour and corresponding non-tumour patient samples...59 | |
| Figure 2.4 L1 5'UTR counts were elevated in high pathological tumours and associated with vascular invasion | 61 |
| Figure 2.5 L1 5'UTR counts correlated with patient recurrence and poor survival | 62 |
| Figure 2.6 L1 5'UTR counts correlated negatively with TP53 expression..... | 63 |

| | |
|---------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 2.7 mRNA expression, miRNA expression, DNA copy number, DNA methylation, and RPPA were previously clustered into three main iClusters..... | 64 |
| Figure 2.8 L1 5'UTR counts correlated with active TGF- β status..... | 65 |
| Figure 2.9 L1HS correlates with L1PA2 (r=0.866, p<0.001) L1PA3 (r=0.684, p<0.001) and L1Pba1 (r=0.605, p<0.001) spearman correlation..... | 65 |
| Figure 2.10 Negative and positive control for L1orf1p antibody using colorectal carcinoma tissue..... | 68 |
| Figure 2.11 L1orf1p levels in tumour and corresponding non-tumour samples..... | 69 |
| Figure 2.12 FFPE tumour and non-tumour biopsies were stained for L1orf1p using immunohistochemistry..... | 70 |
| Figure 2.13 Median survival between L1_low and L1_high in tumour | 73 |
| Figure 2.14 FFPE HCC tumour and non-tumour biopsies were stained for pSMAD3 by immunohistochemistry..... | 75 |
| Figure 3.1 TGF- β pathway..... | 80 |
| Figure 3.2 TGF- β pathway in early and late carcinogenesis..... | 82 |
| Figure 3.3 Melting temperature of different qPCR primers..... | 90 |
| Figure 3.4 L1 transcript expression in different HCC related cell lines..... | 96 |
| Figure 3.5 L1orf1p expression in different HCC related cell lines..... | 97 |
| Figure 3.6 L1orf1p FACS staining in different HCC related cell lines..... | 97 |
| Figure 3.7 Huh7-WT, Huh7-NT and Huh7-L1KD transcript and protein..... | 99 |
| Figure 3.8 Huh7-WT, Huh7-NT and Huh7-L1KD cell morphology..... | 100 |
| Figure 3.9 Growth proliferation curve between Huh7-WT WT, Huh7-NT and Huh7-L1KD..... | 100 |
| Figure 3.10 Huh7-WT, Huh7-NT and Huh7-L1KD clonogenic assay..... | 101 |
| Figure 3.11 Mean colony number and size in Huh7-WT , Huh7-NT and Huh7-L1KD..... | 101 |
| Figure 3.12 Migration assay of Huh-7, NT-shRNA and L1-shRNA..... | 102 |
| Figure 3.13 Huh7-WT, Huh7-NT and Huh7-L1KD cell invasion..... | 103 |
| Figure 3.14 Huh7-WT, Huh7-NT and Huh7-L1KD cells were subcutaneously injected into nude mice, 5×10^6 cells..... | 104 |

| | |
|-----------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 3.15 Volcano plot of differentially expressed genes between Huh7-NT and Huh7-L1KD..... | 105 |
| Figure 3.16 TGF- β enrichment plot using Hallmark analysis..... | 110 |
| Figure 3.17 SMAD3, TGF β 1, SMAD2, SMAD4 and SMAD7 transcript expression in Huh7-WT, Huh7-NT and Huh7-L1KD..... | 112 |
| Figure 3.18 SMAD3 and SMAD 4 protein expression in Huh7-WT, Huh7-NT and Huh7-L1KD..... | 113 |
| Figure 3.19 SMAD 3 luciferase assay in Huh7-WT, Huh7-NT and Huh7-L1KD..... | 113 |
| Figure 3.20 RT-qPCR analysis of WT, Huh7-NT and Huh7-L1KD..... | 115 |
| Figure 3.21 SMAD3 and L1orf1p expression in Huh7-WT, Huh7-NT and Huh7-L1KD in absence and present of TGF- β induction for 72 hours..... | 116 |
| Figure 3.22 Growth proliferation curves of Huh7-WT, Huh7-NT and Huh7-L1KD with TGF- β induction..... | 117 |
| Figure 3.23 Cell cycle analysis of WT, Huh7-NT and Huh7-L1KD after 48 hour doxorubicin induction..... | 118 |
| Figure 4.1 Crystal structure of L1orf1p (blue spheres= positive N-terminal) (Khazina et al., 2018)..... | 124 |
| Figure 4.2 Effect of L1 overexpression on TGF- β signalling in Huh-7 (n=4), HepG2 (n=4) and PLC/PRF5 (n=1)..... | 129 |
| Figure 4.3 L1orf1p and vimentin cell distribution in PLC/PRF5 pcDNA and L1coneo..... | 130 |
| Figure 4.4 L1 overexpression and TGF- β signalling..... | 131 |
| Figure 4.5 L1orf1p codon optimised transcript expression in HepG2 dox-L1orf1 with dox induction..... | 132 |
| Figure 4.6 L1orf1p expression in HepG2 dox-control and HepG2 dox-ORF1p with dox induction for 48 hours..... | 133 |
| Figure 4.7 L1orf1p FACS staining in HepG2 dox-control and HepG2 dox-ORF1p with the following dox inductions..... | 134 |
| Figure 4.8 L1orf1p and vimentin cell distribution in HepG2 dox-control and L1orf1p overexpression..... | 135 |
| Figure 4.9 SMAD3 and L1orf1p protein expression in PLC/PRF5 dox-control and L1orf1p with dox induction for 48 hours..... | 136 |
| Figure 4.10 L1orf1p and vimentin cell distribution in PLC/PRF5 dox-control and L1orf1p overexpression..... | 137 |

Figure 4.11 L1orf1p and vimentin cell distribution in HHL5 dox-control and L1orf1p overexpression.....138

Figure 4.12 HepG2 dox-control and HepG2 dox-ORF1p with dox induction: untreated, 250ng/ml and 500ng/ml.....140

List of Tables

| | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 1.1 Different HCC biomarker and diagnostic values..... | 9 |
| Table 1.2 Prevalence of HCC mutations in patients. Prevalence statistics were obtained from cBioportal..... | 22 |
| Table 2.1 Molecular classification of iClusters in HCC..... | 51 |
| Table 2.2 Aperio algorithm parameters used for H-score quantification in HCC FFPE samples..... | 56 |
| Table 2.3 L1 5'UTR counts and correlation with different aetiological factors in 196 HCC patients..... | 60 |
| Table 2.4 pathological stage group..... | 60 |
| Table 3.1 qPCR conditions..... | 89 |
| Table 3.2 qRT-PCR primer sequences..... | 89 |
| Table 3.3 Primer efficacy of L1 5'UTR, L1orf1p, L1orf2p, TBP and HPRT..... | 91 |
| Table 3.4 L1orf1p mean fluorescence intensity (MFI) and expression (%) in different HCC related cell lines..... | 98 |
| Table 3.5 Number of reads in control (Huh7-NT) and L1 knockdown (Huh7-L1KD)..... | 104 |
| Table 3.6 Top 50 most significant differentially expressed genes in Huh7-L1KD compared to Huh7-NT..... | 105 |
| Table 3.7 TCGA analysis for Hallmark pathways on top 50 significant differentially expressed genes between Huh7-L1KD versus Huh7-NT cells..... | 108 |
| Table 3.8 TCGA analysis for Hallmark pathways on all differentially expressed gene at logFC 0.5, p<0.05 between Huh7-L1KD versus Huh7-NT cells (Top 10 pathways are shown only)..... | 109 |
| Table 3.9 Keggs pathway analysis on all differentially expressed gene at logFC 0.5, p<0.05..... | 110 |

List of Abbreviations

| | |
|--------------------|----------------------------------------------|
| 4-HNE | hydroxynonenol |
| 5mC | 5'-methylcytosine |
| AFP | Alpha Fetoprotein |
| ASGR1 | asialoglycoprotein receptor 1 |
| BCLC | Barcelona Clinic Liver Cancer |
| CAR T cell therapy | Chimeric antigen receptor |
| cccDNA | covalently closed circular DNA |
| CC domain | conserved coiled coil |
| CDKN1A | p14ARF1 |
| CDKN2A | p16INK4A |
| CDKs | cyclin-dependent kinases |
| CIK cells | Cytokine-induced killer |
| CTD | C-terminal domain |
| DMEM | Dulbecco's Modified Eagle Medium |
| EGF | Epidermal growth factor |
| EN | endonuclease |
| ENi | endonuclease independent |
| FBP1 | fructose-1,6-bisphosphatase |
| FFPE | Formalin fixed and paraffin embedded patient |
| FBS | fetal bovine serum |

| | |
|-------------------|-------------------------------------|
| GP73 | Golgi protein-73 |
| GPC3 | Glypican-3 |
| GSEA | Gene Set Enrichment Analysis |
| HAT | histone acetyltransferases |
| HBeAg | pre-core antigen |
| HD | high dysplastic |
| HDAC | histone deacetylases |
| HepG2 dox-empty | HepG2 control |
| HepG2 dox-L1orf1p | HepG2 L1orf1p overexpression |
| HER | hydroxyethyl radicals |
| HGDP | Human Genome Diversity Panel |
| HHL5 dox-empty | HHL5 control |
| HHL5 dox-L1orf1p | HHL5 L1orf1p overexpression |
| HP1 | heterochromatin protein 1 |
| Huh7-L1KD | Huh7-L1 knockdown |
| Huh7-NT | Huh7-non-target |
| JSH | Japan Society of Hepatology |
| L1 5'UTR | L1Ta 5'UTR promoter |
| LINE-1 or L1 | Long interspersed nuclear element-1 |
| L1 ASP | antisense promoter |
| LD | low dysplastic |
| LSD1 | lysine-specific demethylase 1 |
| MBD | methyl binding domain |
| MCC | Mutated in Colorectal Cancers |

| | |
|----------------------|-----------------------------------------|
| MDR1 | ATP-dependent efflux pump |
| MeCP | methyl-CpG binding proteins |
| MFI | mean fluorescence intensity |
| miRs | MicroRNAs |
| MMLV | moloney's mouse leukaemia virus |
| NAD | nicotinamide adenine dinucleotide |
| NAFLD | non-alcoholic fatty liver disease |
| NAHR | non-allelic homologous recombination |
| NASH | non-alcoholic steatohepatitis |
| NRES | National Research Ethics Service |
| PDGF | Platelet-derived growth factor |
| PDPKs | proline-directed protein kinases |
| PLC/PRF5 dox-empty | PLC/PRF5 control |
| PLC/PRF5 dox-L1orf1p | PLC/PRF5 L1orf1p overexpression |
| PRC | Polycomb repressive complex |
| Rb | retinoblastoma protein |
| RC-seq | retrotransposon capture sequencing |
| RNP | L1 ribonucleoprotein |
| ROS | reactive oxygen species |
| RPPA | reverse phase protein lysate microarray |
| RRM | RNA recognition motif |
| RT | reverse transcription activity |
| SIRT | Selective internal radiation therapy |
| SNU recurrence | Seoul National University |

| | |
|----------------------|--------------------------------------------------------|
| SWI/SNF | switch/sucrose non-fermenting |
| TACE | Transarterial chemoembolization |
| TCGA-LIHC | The Cancer Genome Atlas Hepatocellular carcinoma study |
| TERT | telomerase reverse transcriptase |
| TFIIH | the transcriptional factor II H |
| TGF- β pathway | transforming growth factor β |
| TIL | tumour-infiltrating lymphocytes |
| TNF- α | tumour necrosis factor- α |
| TPRT | Target primed reverse transcription |
| VEGF | vascular endothelial growth factor |
| XPB | xeroderma pigmentosum B |
| XPB | xeroderma pigmentosum D |

Chapter 1: Introduction

1.1 Hepatocellular carcinoma

1.1.1 Epidemiology

Liver cancer is the fourth leading cause of cancer related deaths worldwide. 85 to 90% of primary malignant liver cancers are hepatocellular carcinoma (HCC) (Diaz-Gonzalez et al., 2016, Ferlay et al., 2013). HCC is associated with a poor 5 year survival, causing about 781,600 deaths worldwide annually (Ferlay et al., 2019). The poor prognosis is associated with late patient diagnosis as the disease usually develops on a background of chronic liver disease and has no distinct specific symptoms in early stages. Thus more than 50% of patients are diagnosed in advanced stage HCC with metastasis (Altekruse et al., 2009).

HCC incidence rates are varied in different parts of the world (**Figure 1.1**). 80% cases have been observed in sub-Saharan Africa and East Asia due to the prevalence of Hepatitis B and C viruses (Zhu et al., 2016, Yang et al., 2019). In recent years, these rates are decreasing due to effective vaccination for Hepatitis B (Goldstein et al., 2005). However, in North America and Europe, the incidence rate are increasing due to higher prevalence of Hepatitis C infection, obesity and alcohol consumption (Hajarizadeh et al., 2013). In the USA alone, incidence levels have more than trebled in the last two decades. (El-Serag, 2011). Similarly in the UK, incidence have increased by 2.5 in the last two decades (CancerResearchUK, 2017).

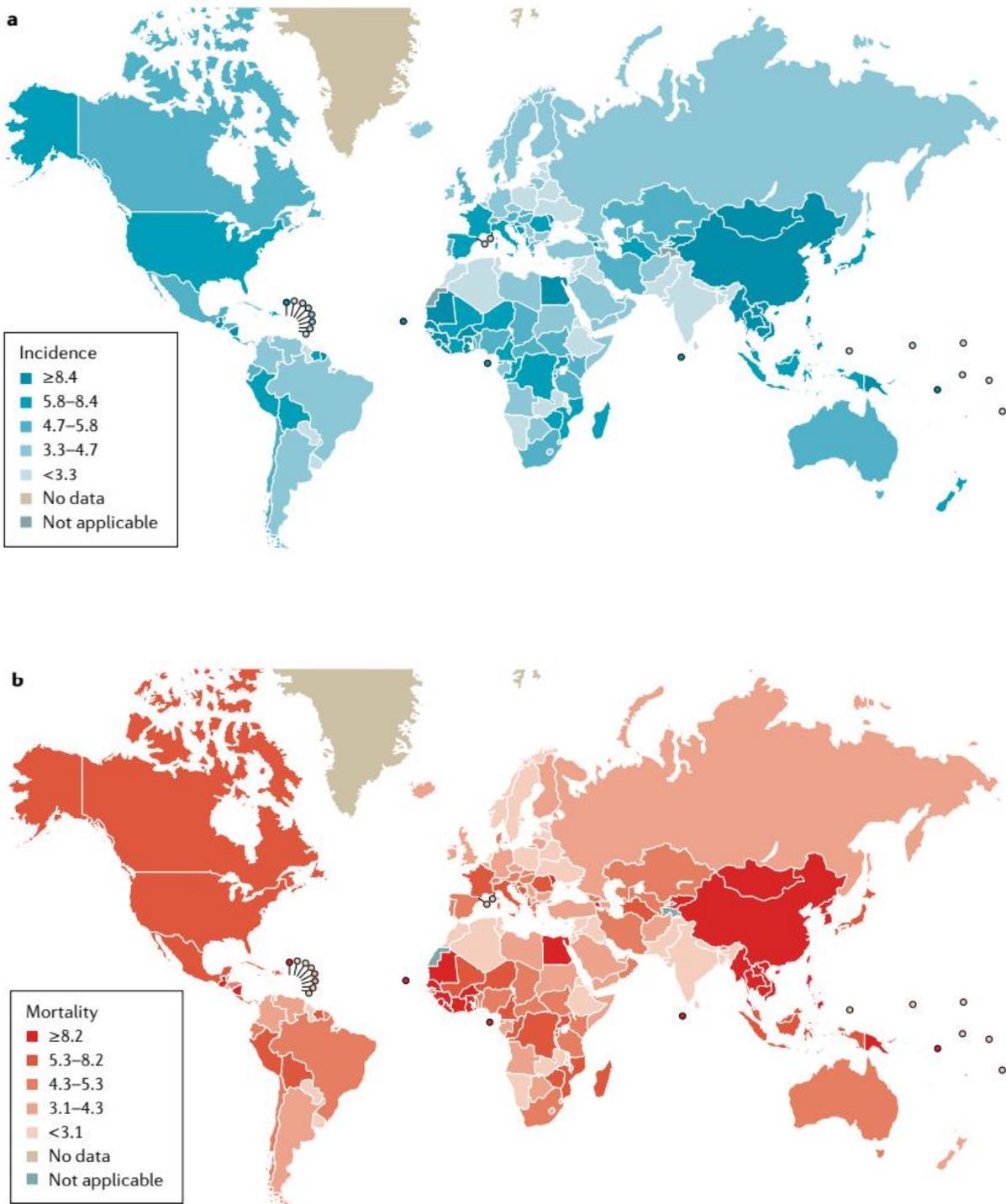


Figure 1.1 Global disease burden of primary liver cancer (numbers are per 100,000 person) (a) incidence (b) mortality during 2018 (Yang et al., 2019)

1.1.2 Risk factors

Several risk factors are associated with HCC. The most prevalent risk factors are chronic infection with Hepatitis B and Hepatitis C viruses. Other major risk factors are alcohol consumption, aflatoxin B1 exposure and obesity.

1.1.2.1 Hepatitis B virus

Hepatitis B virus (HBV) is the most common risk factor for HCC and accounts for more than 55% of all cases (Kew, 2010, Yang et al., 2019). Chronic hepatitis B patients particularly have a 10%–25% risk of developing HCC (McMahon, 2009).

HBV is a partially double-stranded DNA and has four overlapping open reading frames (ORFs): pre-C/C, pre-S/S, P and X (Tsuge et al., 2010, Lucifora et al., 2011, Zoulim et al., 1994). HBV virus enters hepatocytes through the sodium taurocholate cotransporting polypeptide receptor, **Figure 1.2** (Yan et al., 2013). In the cytoplasm, the nucleocapsid releases the semi-circular DNA. The semi-circular DNA is exported into the nucleus and converted into covalently closed circular DNA (cccDNA). cccDNA serves as template for four viral RNAs, which contain polyadenylated tail but do not undergo splicing. The 3.5kB precore RNA is translated into pre-core antigen (HBeAg). The 3.5kB pregenomic RNA is encapsulated with protein kinase and viral polymerase into core particle. The 2.4kB transcript is translated into HBsAg and 0.8kb transcript is translated into HBx. (Tang et al., 2005, Bock et al., 2001). The viral DNA polymerase contains a reverse transcriptase domain but lacks proofreading capacity and has a high frequency of mutation particularly substitution mutations. Subsequently, it causes genetic diversity and several viral genotypes can emerge leading to drug resistance.

HBV can also integrate into the host genome and can affect gene expression including inactivation of tumour suppressor genes or activation of oncogenes (Jiang et al., 2012). High-throughput sequencing-based method has shown that HBV integration tend to occur commonly in chromosome 17. Its integration is particularly reported in two HCC relevant genes: fibronectin 1 and telomerase reverse transcriptase and also in SMAD5, PHACTR4 and RBFOX1 genes (Ding et al., 2012a).

HBV can also promote hepatocarcinogenesis via HBx protein as it can influence various cell signalling pathways. For example, HBx protein can induce promoter methylation of several genes by increasing DNMT1 and DNMT3a expression (Jung et al., 2007).

p16INK4A promoter methylation is particularly associated with HBx protein expression leading to increased cell cycle progression (Zhu et al., 2010). HBx can also indirectly inhibit E-cadherin expression through DNMT1. E-cadherin inhibition is associated with reduced cell-cell contacts and increased cancer invasion. Its inhibition can also activate Wnt signalling pathway, as it cannot sequester β -catenin and inhibit translocation into the nucleus. Thus the HBx protein induces EMT and increases cell proliferation (Lee et al., 2005a). HBx can also reduce p53 nuclear expression by sequestering it in the cytoplasm (Wang et al., 1995). Furthermore, it can also bind to the C-terminal of p53 and inhibits binding of xeroderma pigmentosum B (XPB) and xeroderma pigmentosum B (XPD). Both XPB and XPD are components of the transcriptional factor II H (TFIIH) and have an important role in inducing apoptosis. Thus, XPB and XPD inhibition leads to reduce apoptosis. HBx protein can also upregulate survivin expression, which belongs to the apoptosis-inhibitor protein family (Zhang et al., 2005). Also, a crosstalk between HBx and NF-kappaB pathway has been reported (Shukla et al., 2011).

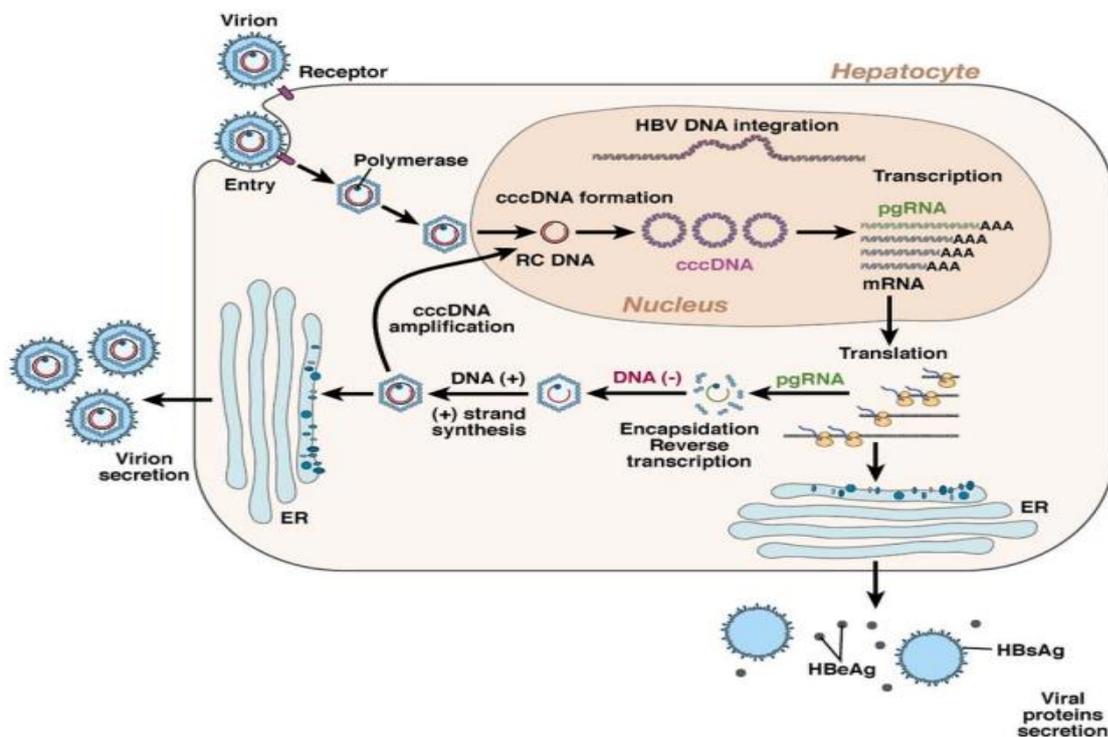


Figure 1.2 Hepatitis B viral cycle (Zoulim and Locarnini, 2009)

1.1.2.2 Hepatitis C virus

Hepatitis C virus (HCV) infection is the second most common risk factor for HCC and accounts for 10-25% of all HCC cases (Huang et al., 2011). It is estimated that 57 million people have chronic hepatitis C worldwide and 10-20% have liver complications including cirrhosis and HCC (Heffernan et al., 2019, Hajarizadeh et al., 2013).

HCV is a positive-sense single-stranded RNA and cannot integrate into the host genome as it does not transcribe via reverse transcription enzyme. The virus tends to harbour in the endoplasmic reticulum of hepatocytes (**Figure 1.3**). In the endoplasmic reticulum, it replicates its RNA and translates key structural and non-structural proteins such as HCV Core, p7 E1 and E2; and (NS2, NS3, NS4A, NS4B, NS5A and NS5B) (Kim et al., 2007). HCV core, E2, NS5A, and NS5B proteins can also activate proliferating pathways such as the E2F1 pathway and RAF/MAPK/ERK kinase pathways. These pathways are associated with poorly differentiated HCC tumour phenotype.

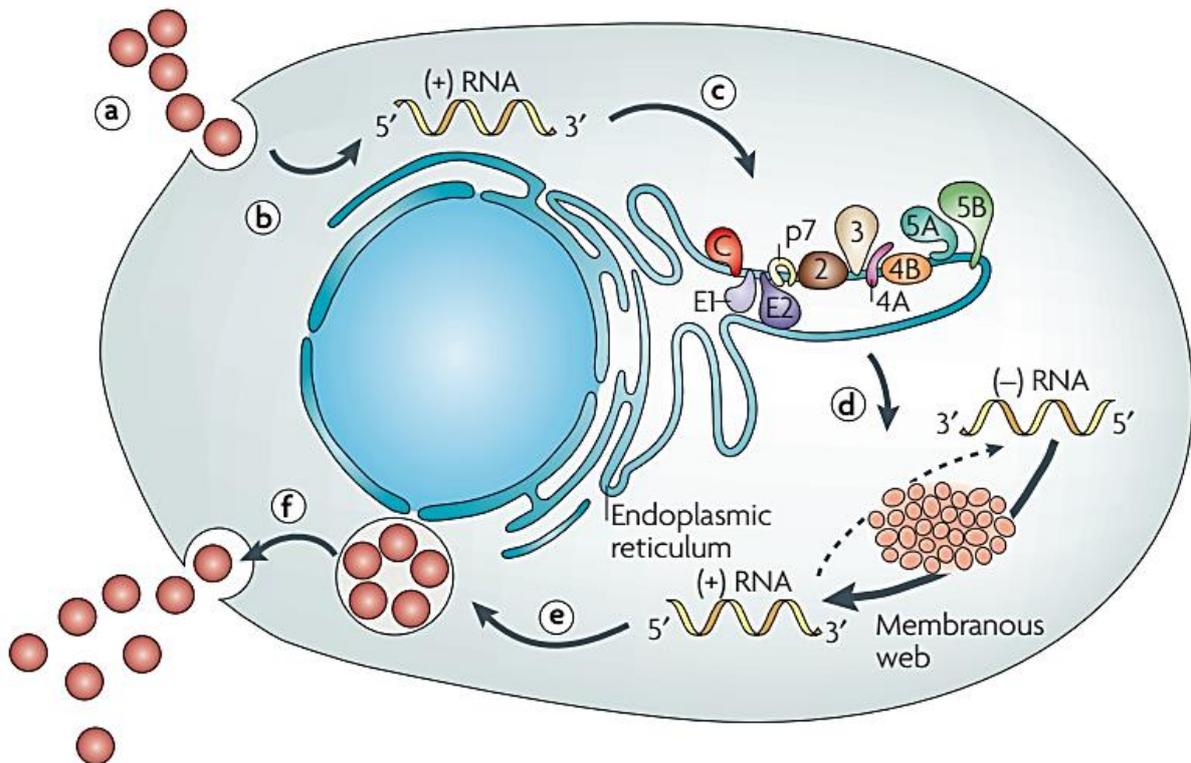


Figure 1.3 Hepatitis C viral cycle (Moradpour et al., 2007)

In particular, HCV core protein influences several cellular pathways. It can suppress tumour suppressor genes such as RB1, TP53 and TP73 (Machida et al., 2009, Kao et al., 2004, Alisi et al., 2003). Furthermore, HCV core protein is a negative regulator of

CDKN1A and CDKN2A. CDKN1A (p21) activates TP53 by blocking its inhibitor MDM2. Thus its downregulation increases MDM2 expression leading to increased TP53 degradation. CDKN2A (p16^{INK4A}) is a cell cycle inhibitor and blocks cyclin-dependent kinases, and inhibits S-phase cell cycle progression (Hayashi et al., 2000, Lim et al., 2012). HCV core protein also induces the production of ROS, which may induce an inflammatory microenvironment in the liver and recruit profibrogenic mediators such as TGF- β (Bataller et al., 2004). The HCV core protein can also inhibit NF-kappa B mediated pathways leading to a reduced innate immune and anti-tumour response (Joo et al., 2005). *In vivo* studies have shown that HCV core protein can induce insulin resistance, steatosis and HCC in transgenic mice. Similarly, HCV core protein can co-localise with apolipoprotein A2 on the surface of triglyceride and hence regulates lipid metabolism (Moriya et al., 1998, Barba et al., 1997).

Besides HCV core protein, NS5A and NS5B can also contribute to hepatocarcinogenesis. NS5A can inhibit TP53 signalling pathway progression (Majumder et al., 2001). It can also activate β -catenin/WNT and PI3K/AKT pathways and evade cell apoptosis by inhibiting caspase-3 (Street et al., 2005). In addition, NS5A can induce the production of ROS and causes an inflammatory liver microenvironment (Baek et al., 2006). NS5B protein can also inhibit the tumour suppressor gene RB1 (Munakata et al., 2005). These findings demonstrate that HCV proteins have diverse roles in cell proliferation, lipid metabolism, cell survival that support HCC carcinogenesis.

1.1.2.3 Alcohol

Alcohol is third most common risk factor in HCC worldwide however, accounts for 40 to 50% of all HCC in Europe (Jewell and Sheron, 2010). A meta analysis has shown that consumption of alcohol three or more drinks per day increases the risk of HCC by 16% and consumption of more than 6 drinks increases the risk by 22% (Turati et al., 2014). A positive association has been observed between alcohol consumption, obesity, Hepatitis B, Hepatitis C and risk of HCC (Donato et al., 2002). Particularly, heavy alcohol consumption (210–560 g/week) and Hepatitis B infection increases the risk of HCC by three fold when compared to Hepatitis B infection only (Farazi and DePinho, 2006).

Alcohol is metabolised in the liver by the enzyme alcohol dehydrogenase. Alcohol dehydrogenase metabolises ethanol and synthesis nicotinamide adenine dinucleotide (NAD). Subsequently, NAD produces free radical species which induces oxidative stress

in the liver and induces the formation of protein and DNA adducts leading to DNA damage (McKillop and Schrum, 2005). Free radical species can also activate the innate immune system particularly Kupffer cells. Kupffer cells in turn release several inflammatory cytokines such as tumour necrosis factor- α (TNF- α) (Mandrekar and Szabo, 2009) (Farazi and DePinho, 2006). Chronic alcohol consumption can also increase cytochrome CYP2E1 levels by 10 to 20 fold. CYP2E1 induction by ethanol is associated with formation of hydroxyethyl radicals (HER) leading to lipid peroxidation (Dupont et al., 1998, Albano et al., 1993). Furthermore, CYP2E1 induction increases ethanol reactive species and 4-hydroxynonenol (4-HNE). 4-HNE is associated with mutation at codon 243 of TP53 gene. CYP2E1 can also activate procarcinogen such as aflatoxin and polycyclic hydrocarbons (Hu et al., 2002, Aleynik et al., 1998), thus promoting hepatocarcinogenesis.

1.1.2.4 Obesity

Obesity is another major risk factor and a meta analysis has shown that overweight and obese individual have a 17% and 89% increased risk of HCC (Larsson and Wolk, 2007). Furthermore, a separate meta-analysis study showed an increased risk of 25% for each 5 kg/m² increase of BMI (Renehan et al., 2008). Obesity is caused by excess body fat leading to the accumulation of triglycerides in the liver. As the condition progresses, it can develop into non-alcoholic fatty liver disease (NAFLD). NAFLD is characterised by ballooning of hepatocytes and cell death. A strong association have been observed between NAFLD, obesity, hypertension and insulin resistance. The disease may progress into non-alcoholic steatohepatitis (NASH) which includes inflammation. NASH patients have a 5 year risk of 11.7% for developing HCC (Bhala et al., 2011).

1.1.2.5 Aflatoxin B1

Aflatoxin B1 (AFLB1) is synthesised by the fungi *Aspergillus flavis* and *Aspergillus Parasitans*. AFLB1 is metabolised in the liver and the metabolite can cause alkylation of nucleic acids in the genome. Several studies have observed somatic mutations associated with aflatoxin especially in codon 249 of TP53, which affects its tumour suppressor function (Gouas et al., 2009). Thus, areas with high prevalence of AFLB1 such as sub-Saharan Africa and eastern Asia are associated with high prevalence of HCC. Exposure of AFLB1 to Hepatitis B positive patients further increases the risk of HCC development (Kew, 2003).

1.1.2.6 Hereditary Haemochromatosis

Hereditary haemochromatosis is an autosomal recessive disorder caused by mutation in the HFE gene. The mutation causes excess iron absorption and storage in several organs including liver. In the liver, it causes liver damage and fibrosis (Powell et al., 2000). Several studies have shown an association between hereditary haemochromatosis and HCC (Fracanzani et al., 2001, Elmberg et al., 2003).

1.1.3 Diagnosis of HCC

Several techniques have been developed to detect HCC. These include liver biopsy, ultrasonography and certain biomarkers. Liver biopsy is the optimum diagnostic test and has a sensitivity of about 70% but is lower in tumours smaller than 2cm due to difficulties in distinguishing well differentiated HCC from dysplastic lesions. Several guidelines have been suggested to identify major histological features of HCC such as increased cell intensity, diffuse fatty changes, pseudo-glandular pattern and stromal invasion (Jain, 2014).

Ultrasonography is a non-invasive alternative, and an effective tool in monitoring HCC in cirrhotic patients. Lesions larger than 1cm should be analysed by dynamic contrast enhanced MRI or quadruple-phase CT. HCC lesions will be brighter than the surrounding liver in the arterial phase and are less bright compared to the surrounding parenchyma in the venous phase. The phenomenon is known as delayed washout and has a sensitivity of 89% and a specificity of 96% for HCC (Shinmura et al., 2005, van der Pol et al., 2019). The use of ultrasonography can be subjective as it depends upon the operator (Atiq et al., 2017). In cirrhotic patients, lesions larger than 2cm correlate with a 95% risk of HCC (Frazer, 1999).

Smaller lesion less than 1 cm are less likely to be malignant particularly in patients with cirrhosis (Iwasaki et al., 1998). Nonetheless, routine surveillance is required 3-6 months to monitor any malignant growth (Bruix and Sherman, 2005). Several studies have observed better survival during routine surveillance in patients with underlying liver conditions. For example, in a randomised controlled trial, Hepatitis B patients were categorised into two groups, screening (9373 patients) and control group (9443 patients). The screening group were monitored both by α -fetoprotein test and an ultrasonography examination every 6 months. The mortality group in the screening group was significantly

lower (83.2 per 100000) compared to (131.5 per 100000) in the control group (Duffy, 2010).

Several biomarkers have been developed to detect HCC (**Table 1.1**) but they all have some limitations.

Table 1.1 Different HCC biomarker and diagnostic values

| Biomarker | Sensitivity, % | Specificity, % |
|------------------------------------|-----------------------|-----------------------|
| AFP (Marrero et al., 2009) | 53.0 | 90.0 |
| AFP-L3 (Marrero et al., 2009) | 28.0 | 97.0 |
| Golgi protein-73 (Xu et al., 2014) | 74.6 | 97.4 |
| Glypican-3 (Sawada et al., 2012) | 55.1 | 97.0 |
| miR-15b (Liu et al., 2012) | 98.3 | 15.3 |
| miR-130b (Liu et al., 2012) | 87.7 | 81.4 |
| miR-21 (Tomimaru et al., 2012) | 87.3 | 92.0 |

- Alpha Fetoprotein (AFP) is a 70kDa glycoprotein and is usually produced during the first 12 to 16 weeks of neonatal development at a level of about 3g/L in serum. Afterwards, AFP level rapidly declines and reach trace levels. In response to injury, hepatocyte proliferate and release AFP into the circulation (Mizejewski, 2001). Elevated AFP levels have been associated with HCC particularly serum levels higher than 200 ng/L (Bruix and Sherman, 2005). However, AFP has a low sensitivity of 53% and a specificity of 90% for detecting HCC (Marrero et al., 2009). False positive may occur as AFP levels can increase in other disorders such as cirrhosis and Hepatitis. A more specific biomarker is AFP-L3. AFP-L3 is one of the three glycoforms of AFP and is specifically elevated in HCC (Spangenberg et al., 2006). Similarly, elevated AFP-L3 levels are associated with poorly differentiated tumours. Its sensitivity is 28% but has a specificity of 97% (Marrero et al., 2009). As the sensitivity is low, it has significant limitation as a biomarker for HCC.

- Golgi protein-73 (GP73) is a type II Golgi-specific membrane protein and expressed in normal epithelial cells but not in normal hepatocytes. It is detected in the serum during liver disease particularly in HCC (Kladney et al., 2002). A previous study has shown that GP73 levels were significantly higher in patients with Hepatitis B or HCC compared to healthy adults (Mao et al., 2010). The sensitivity and specificity is 74.6% and 97.4% respectively (Xu et al., 2014).
- Glypican-3 (GPC3) belongs to the heparan sulfate proteoglycan family and is involved in cell proliferation and survival. In normal hepatocytes, it is usually repressed (Filmus and Capurro, 2008). In HCC, GPC3 is elevated and associated with increased Wnt signalling pathway leading to increased cell proliferation. As it is expressed in tumour tissues and not in benign tissue, it provides the potential as a tissue biomarker (Sung et al., 2003). Furthermore, GPC3 can be released in the peripheral blood vessels as sGPC3 and be detected in the serum. The sensitivity and specificity is 55.1% and 97.0% respectively for sGPC3 (Sawada et al., 2012).
- MicroRNAs (miRs) are small non-coding RNAs and can bind to the 3' untranslated region of transcripts and induce transcript degradation. miRs are able to affect several genes and thus affect cellular processes such as proliferation and differentiation, and can also affect tumour suppressor or oncogene functions (Ferracin et al., 2010). Several studies have explored these as potential biomarkers for cancers including HCC (Bracken et al., 2016, Lin and Gregory, 2015). For example, patient serum samples from a cohort of 30 healthy controls, 57 patients with HCC and 29 hepatitis B individual were analysed for miR-15b and miR-130b. miR-15b and miR-130b were both elevated in the HCC cohort. miR-15b had a high sensitivity rate of 98.3% but a poor specificity of 15.3% for detecting HCC. miR-130b had a sensitivity rate of 87.7% and specificity of 81.4% (Liu et al., 2012). miR-21 is also an important potential biomarker as it can distinguish chronic hepatitis and HCC. It had a sensitivity and specificity of 61.1% and 83.3% in detecting HCC compared to chronic hepatitis patients in a cohort of 30 patients (4 with Hepatitis B and 26 with Hepatitis C). The sensitivity and specificity increased to 87.3% and 92.0% in detecting HCC (n=126) compared to healthy control, n=50 (Tomimaru et al., 2012). The data is encouraging but more studies with larger cohorts are needed to validate the findings and establish clinical usage of these as potential HCC detection biomarkers.

Current data suggest that a single biomarker is not sufficient for detecting HCC particularly in early stages of the disease. Several studies have shown that combination of different diagnostic markers may improve the diagnostic rate. Hence, early detection biomarkers are an unmet need of HCC patients.

1.1.4 Staging of HCC

Staging is an important tool to assess the extent of the cancer and identify the best treatment strategies. It generally involves evaluation of liver function and patient performance. The Barcelona Clinic Liver Cancer (BCLC) staging is a common staging system in the world (Burak and Kneteman, 2010). It was introduced in 1999 and further adapted in 2008 to include Sorafenib treatment (Llovet et al., 2008a, Llovet et al., 1999a). The staging system predicts patient prognosis and suggest treatment strategies (**Figure 1.4**). Patients are classified into four main groups: early, intermediate, advance and terminal stage. Patients with child score of A or B and 3 lesions less than 3cm are classified as early stage. The 5-year survival is usually between 50-75%. Intermediate stage patients have a child score of A and B and multifocal HCC with no macrovascular invasion. The 3-year survival is usually less than 50%. Advance stage patients have a child score of A and B, macrovascular invasion and extrahepatic spread. Terminal stage patients have a child score of C and extensive tumours. The survival is usually less than 3 months.

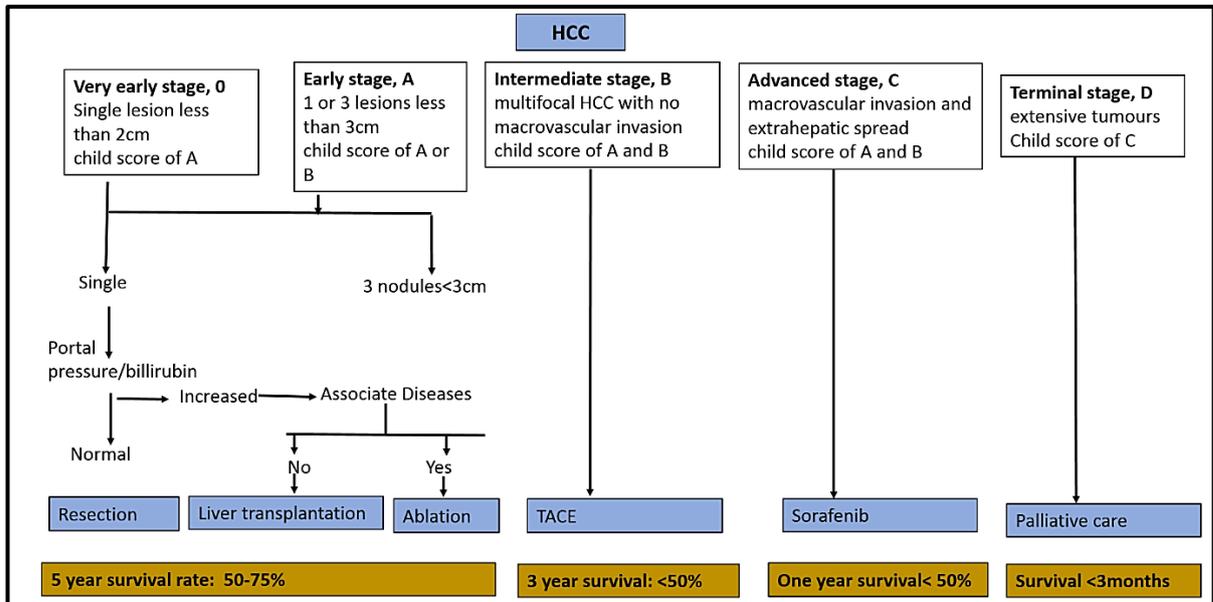


Figure 1.4 BCLC staging system. The staging system predict prognosis and recommends treatment strategies by assessing patient tumour and liver related factors.

The BCLC has been used to assess HCC staging but prescribed treatment options are limited. Some centres may use their own guidelines or use other staging systems such as the Alberta HCC and the Japan Society of Hepatology (JSH) algorithms. Alberta HCC and JSH both extends the BCLC system by providing more treatment options and re-classifying some treatment options to more advance HCC patients. For example in these staging system, early stage HCC can be treated with radiofrequency ablation, and Child-Pugh class C patients can receive liver transplantation, which is not present in BCLC staging (Burak and Kneteman, 2010). Nonetheless as the disease progresses, treatment strategies become limited and are associated with a poor patient prognosis.

1.1.5 Treatment approaches

Several treatment options exist to treat HCC. Though as the disease progresses, treatment options become limited (**Figure 1.4**), the options also depend upon the available resources (**Figure 1.5**).

1.1.5.1 Liver resection

Liver resection is the primary treatment for HCC patients. It is offered to patients with normal liver function, normal bilirubin levels, no portal hypertensions and no liver cirrhosis. In addition, tumour lesions must have the following criteria: single tumour lesion less than 5 cm or up to 3 tumour lesions less than 3 cm in size, Child Pugh Score A or B or performance score of 0 (Llovet et al., 1999b). Liver resection also provides the capability of measuring the likelihood of disease recurrence using markers such as microvascular invasion. Nonetheless, liver resection has a high patient recurrence. 70% of patients will usually get tumour recurrence within 5 years (Forner et al., 2012) and 10 year survival is 7-15%. Furthermore, patients with non-cirrhotic NAFLD HCC have a high 20% mortality rate with liver resection, which is also observed in patients with liver cirrhosis. Thus liver resection can only be performed in certain patients (Piscaglia et al., 2016).

1.1.5.2 Liver transplant

In many cases, resection may not be possible in patients with chronic liver disease. These patients may receive a liver transplant. Currently, there are several patient criteria before a transplant can be performed. Milan criteria is the common liver transplant criteria used in patients. In Milan criteria, HCC tumour nodule must be less than 5 cm or 3 nodules each less than 3 cm. The 5-year survival usually exceed 70% and tumour recurrence is about 20%. Furthermore, liver transplantation removes the risk of any unrecognised intrahepatic metastasis. However, there is a short supply of donor livers available for transplants and patients require life-long immunosuppressive drugs (Zimmerman et al., 2008). Loco-regional treatment may be performed such as liver resection or ablation until a suitable donor is found.

1.1.5.3 Ablation

Tumour ablation is an indirect therapy to remove HCC tumours and has a median overall survival of 60 months and a 5 year recurrence of 50-70%. Several ablation techniques exist including chemical, electrical or thermal methods. Chemical ablation uses ethanol to induce tissue ischaemia in tumour tissues. Radiofrequency ablation is another type of ablation

which uses heat induced by radiofrequency (Yang and Roberts, 2010). Recently it has been suggested as a primary treatment for HCC tumours less than 2cm. Several studies have demonstrated improved overall survival rate with radiofrequency ablation compared with untreated. In a separate study, a specific cohort of patients demonstrated similar overall survival between radiofrequency ablation and liver resection (Diaz-Gonzalez et al., 2016).

Radiation

During radiation treatment, HCC tumours were treated with external radiation beams to reduce or eliminate small tumours. A pooled analysis was performed on 102 patients with unresectable HCC tumours and Child Pugh score A. Half of patients were treated with photon stereotactic body radiation therapy from 24 to 54G. Patients with 54G demonstrated an improved overall response rate of 54% and an overall survival of 17 months (Bujold et al., 2013).

1.1.5.4 Transarterial chemoembolization (TACE)

TACE is a therapy commonly used in patients with unresectable tumours. Gelatine microspheres embolised with chemotherapeutic reagent are occluded onto the hepatic artery to induce tumour necrosis. The hepatic artery is the primary artery of HCC and nearly all its vascularisation occurs through this artery. In contrast, the hepatic artery supplies 30% of blood supply to the normal hepatic parenchyma. Thus, it is a selective target particularly for TACE.

As tumour is reduced, more curative therapy options such as liver transplant may become available (Belghiti et al., 2008, Heckman et al., 2008). Several studies have shown significant survival benefit with TACE therapy compared to best supportive care or Tamoxifen treatment. The overall survival with TACE treatment was 37 months in random controlled studies (Liu et al., 2018). In a separate study, TACE reported a treatment related mortality of just 0.6% (Lencioni et al., 2016). However, 5-year survival do not usually exceed 50%, and is not suitable in advance HCC patients as it may induce liver failure(Diaz-Gonzalez et al., 2016).

A previous study investigated the effect of treatment with TACE before and after liver resection. The study concluded that prior HCC treatment may lead to a poorer prognosis.

Postoperative TACE treatment did not affect survival (Sasaki et al., 2006, Schwartz et al., 2002). In a separate study, TACE treatment survival was measured in patients with different stages of HCC. Early stage (BCLC 0-A) demonstrated a survival of 16-45 months, intermediate stage (BCLC B) 15.6-26.3 months and late stage (BCLC C) 6.8-13.6 months {Llovet, 2016 #331}.

1.1.5.5 Selective internal radiation therapy (SIRT)

Selective internal radiation therapy (SIRT) is used in patients with unresectable tumours. Implantable radioactive microspheres are typically 35 µm in diameter compared with 100-500 µm in TACE (Volk et al., 2008). The most common radiation agent used is Y90 b-emitting isotope and directly delivered into the artery of the tumour. SIRT radiation induces localised radiation without causing ischaemia to the liver or tumour due to the small size of microspheres (Lau et al., 2013). In a study consisting of 463 patients (122 of patients received TACE and 123 receive SIRT), SIRT had an improved response rate and longer time for tumour progression compared to TACE therapy (13.3 months vs 8.4 months). Though the median survival was not significant between these two groups (20.5 months vs 17.5 months) (Salem et al., 2011). SIRT is commonly applied to patients which are not good candidate for TACE therapy due to the larger tumour burden and vascular invasion (Lau et al., 2013).

1.1.5.6 Medical Therapies

Several chemotherapeutic agents have been tested for their efficacy against HCC. Systematic doxorubicin, PIAF (interferon, platinum, doxorubicin and 5-FU) regime have not reported any survival advantage, and in some cases it was associated with additional toxicity. Furthermore, no survival benefits were observed with the following drugs in phase 3 trials: sunitinib, erlotinib, epidermal growth factor receptor, brivanib, linifanib, ramucirumab, everolimus. These phase 3 drugs acts on various molecular targets including VEGFR, VEGFR2 fibroblast growth factor receptor, c-Kit, epidermal growth factor

receptor. Several reasons were associated with poor efficacy including liver toxicity, low potency or flaws in the design of the study (Sun and Cabrera, 2018).

However, sorafenib has been shown to increase median survival across all different HCC aetiologies. Sorafenib is a multikinase inhibitor and is used to treat advanced progressed HCC. It inhibits the (Raf/MEK/ERK) signalling pathway, specifically VEGF-2, VEGF-3 and PDGF, leading to reduced proliferation and increased cell apoptosis. In a phase 3 clinical trial involving 602 advanced HCC patients who had no previous systematic treatments, Sorafenib has improved survival by 7.9 to 10.7 months. (Llovet et al., 2008b). Side effects were manageable and included skin reactions and diarrhoea. Further studies demonstrated that combination therapy with TACE and sorafenib have not resulted in any clinical benefits.

Another tyrosine kinase inhibitor is Lenvatinib. Lenvatinib is an oral tyrosine kinase inhibitor and blocks fibroblast growth factor receptors 1–4, platelet-derived growth factor receptor, VEGFR1–3, KIT and RET. In a phase 3 trial study of 954 patients with unresectable HCC, Lenvatinib improved survival by 13.6 months compared with Sorafenib 12.3 months. However the complete response is poor as <1% had a complete response, 18% had a partial response for Lenvatinib. The complete response rate for Sorafenib was <1% and 6.5% had a partial response (Kudo et al., 2018). Both Sorafenib and Lenvatinib are first-line treatments and are FDA approved for HCC. However, the response rate is poor and further secondary line treatments are needed.

Regorafenib is an oral multikinase inhibitor targeting VEGF1-3 and other kinases leading to inhibition of cell proliferation, metastasis, angiogenesis and tumour immunity (Wilhelm et al., 2011). Its efficacy was tested in 573 patients who had been treated previously with Sorafenib but showed tumour progression after treatment. Regorafenib improved survival by 10.6 months compared to 7.8 months with placebo (Bruix et al., 2017). Cabozantinib is another second line treatment and inhibits VEGFR2, MET and RET. In a phase 3 study consisting of 707 patients assigned in a 2:1 ratio of Cabozantinib and placebo, Cabozantinib improved survival by 10.2 months compared to 8 months with placebo (Abou-Alfa et al., 2018).

5-FU is a drug which inhibits cell cycle progression at S-phase and upregulates p53 expression. 5-FU and cisplatin combination therapy was particularly shown to extend

survival to 14 months compared to HCC patients who did not receive the combination therapy (5.2 months) (Nouso et al., 2013).

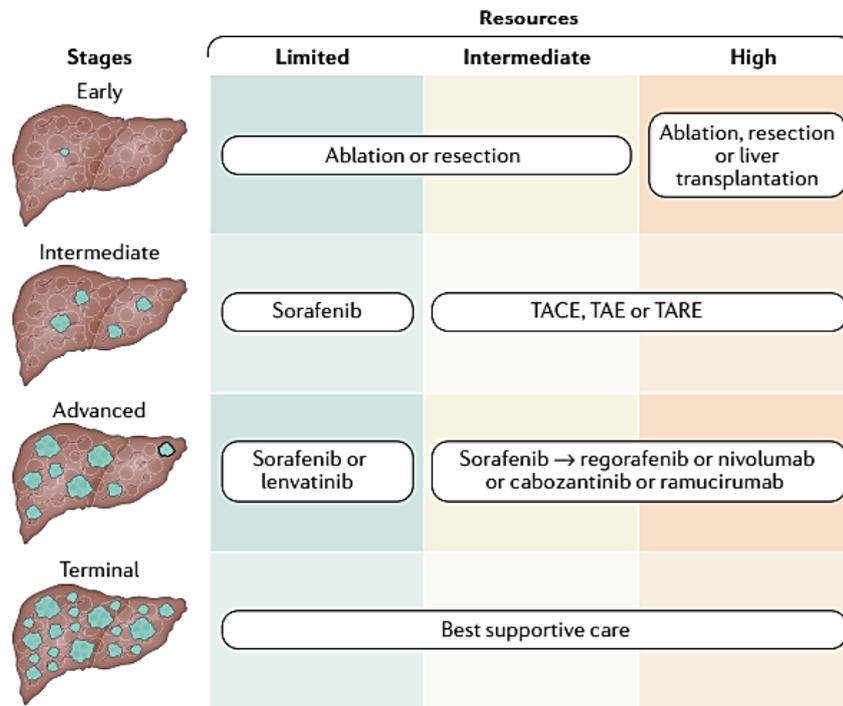


Figure 1.5 Strategy of HCC treatments in countries with different resources (Yang et al., 2019)

1.1.5.7 Novel/Experimental therapies

Novel treatment approaches are required to target tumours whilst not affecting normal tissues. As the immune system has the capability of targeting tumour cells, several different strategies have been developed to increase the anti-tumour immune response.

Adoptive Cell Therapy

Adoptive cell therapy is an immunotherapeutic approach, which stimulates patient derived lymphocytes with cytokines *ex vivo*. Cells are re-infused into the patient to increase the patient anti-tumour immune response. Several cell types have been used such as T cells, natural killer cells, and tumour-infiltrating lymphocytes (Yeku et al., 2017, Baruch et al., 2017).

Chimeric antigen receptor (CAR) T cell therapy has been used *in vitro* and *in vivo* studies. In a previous study, genetically modified T lymphocytes were used which were customised to the tumour antigen GPC3. Survival was significantly higher with CAR T cell therapy (n=6) compared to untreated in HCC mice xenografts (n=6) (Gao et al., 2014a). Similarly, CAR T cells customised to GPC3 and asialoglycoprotein receptor 1 (ASGR1) had an elevated anti-tumour immune response both *in vitro* and *in vivo* (n=5 for each group) (Chen et al., 2017). Nonetheless, CAR T cells can have severe complications such as cytokine release syndrome and tumour lysis syndrome. Further clinical trials are required to test the efficacy and safety of CAR T cells therapy (Teachey et al., 2016, Xu and Tang, 2014).

Cytokine-induced killer (CIK) cells are peripheral mononuclear cells consisting of CD3⁻CD56⁺ cells, CD3⁺CD56⁺ and CD3⁺CD56⁻. Cells are stimulated by cytokines such as IL-12, IL-1, IL-15 and interferon γ *ex vivo*, and then re-infused into the patient (Mata-Molanes et al., 2017, Gao et al., 2017). In a meta analysis consisting of 693 HCC patients, TACE and dendritic cells CIK cells combination therapy increased patient survival by 1 to 2 years (Su et al., 2016). Similarly in a mice model, CIK cells and valproate were infused. Valproate is a drug used to treat epilepsy and headaches. Recent studies has shown that Valproate can inhibit histone deacetylase (HDAC). HDAC inhibition can increase the transcription of several genes and enhance innate immune response by specifically increasing the NKG2D ligand. The NKG2D ligand is expressed in tumour cells and has an important role in the induction of cytotoxic lymphocytes. Valproate has shown to increased MIC-A and MIC-B mRNA expression and making HCC more susceptible to innate immune system particularly through cytotoxic NK cells. CIK cells and Valproate had a synergistic effect and reduced tumour growth (Lee et al., 2017).

Lymphokine-activated killer cells may provide the potential for novel therapy. However, most lymphokine-activated killer cells lack tumour antigen specificity. In contrast, tumour-infiltrating lymphocytes (TIL) contain tumour antigen-specific T cells as they are derived from tumour tissues (Toh et al., 2006). In a previous study, TIL cells were infused in the hepatic artery and were able to migrate to HCC tumour sites. TIL cells were also associated with lower recurrence rate in HCC patients (Xie et al., 2012).

Immune Checkpoint Inhibitors

Immune checkpoint regulators are overexpressed in several cancers and inhibit the anti-tumour immune response. Different immune checkpoint inhibitors have been developed to increase the anti-tumour immune response.

CTLA-4 inhibitor prevents the binding of CTLA-4 to B7-1 and B7-2 leading to T cell activation (Vesely et al., 2011). In a phase II study involving patients with advanced HCC and hepatitis C, anti-CTLA-4 monoclonal antibody (Tremelimumab) decreased the Hepatitis C viral load and no significant toxicity was observed. Moreover, the survival was improved by 6.48 months (Sangro et al., 2013). Similarly in a separate study, combination therapy of radiofrequency ablation and Tremelimumab decreased Hepatitis C viral load but also increased CD8+T cells in HCC tumours. Thus CTLA-4 inhibitor may be a potential successful therapy particularly in Hepatitis related HCC patients (Duffy et al., 2017).

PD-1 immune checkpoint inhibitors can inhibit the binding of PD-L1 and PD-L2 to PD-1. PD-L1 are expressed in cancer cells and PD-1 in immune cells. If the binding is inhibited, it leads to T-cells, B-cells, natural killer cells, dendritic cells activation (Dai et al., 2014, Shi et al., 2013). In a phase I/II study, 41 advanced HCC patients were treated with the PD-1 inhibitor Nivolumab. In 72% of patients, the survival significantly increased by at least 6 months (El-Khoueiry et al., 2015).

Furthermore in a randomised phase 3 trial, the combination therapy of atezolizumab and bevacizumab was tested (558 patients). Both atezolizumab and bevacizumab are monoclonal antibodies. Atezolizumab targets PDL1 and bevacizumab targets VEGF. The combination therapy improved survival compared to sorafenib 19.2 months vs. 13.4 months. Thus, immune therapies may provide a better treatment option for HCC (Finn et al., 2020).

HCC Vaccines

Cancer vaccination can be used to reduce tumour load or prevent tumour relapse. A range of HCC vaccines exists. Antigen Peptide Vaccines uses tumour specific peptides such as AFP and GPC3. As AFP is overexpressed in many HCC tumours, it can be a potential vaccine target. However, the anti-tumour immune response is reduced due to acquired immune tolerance in early development (Pardee and Butterfield, 2012). GPC3 provides a

potential target as it is highly expressed in HCC cells. In a phase II trial, 25 patients received 10 GPC3 peptide derived vaccination over the course of 12 months. Patients' recurrence was significantly lower in patients with surgery and GPC3 peptide derived vaccination than surgery alone (at year 1: 24% versus 48%; at year 2: 52.4% versus 61.9%) (Sawada et al., 2016).

Vaccination can also be induced using dendrite cells. Dendritic cells vaccination is developed by inducing dendrite cells with cytokine such as rhGM-CSF and then sensitising to autologous tumour antigens (Sun et al., 2015). Several studies have used dendrite vaccination to treat HCC. In a previous study, dendrite cells were pulsed with autologous tumour lysates and then infused into 31 patients with advanced HCC. Four patients had partial response, seventeen patients had stable disease and ten had progressive disease, but the overall 1-year survival was improved in all patients (Lee et al., 2005b). Similarly in a separate study, dendritic cells were sensitised by tumour cell lysate. Both dendrite cells and nifuroxazide (STAT3 inhibitor, increases dendritic cells maturation) were then infused into mice with HCC. The combination therapy increased the anti-tumour immune response leading to decreased tumour growth and improved survival (Zhao et al., 2017).

1.2 Molecular mechanisms of HCC

HCC is a multistep process, involving multiple changes such as chronic inflammation, cirrhosis leading to genomic changes and HCC development.

Chronic inflammation is usually the first step in the development of HCC (**Figure 1.6**). Previous studies have shown that more than 90% HCC patients had a background of inflammation, irrespective of the cause of liver disease. Chronic inflammation is usually caused by liver injury induced by toxins, autoimmune related causes or infection. The inflammatory response consists of proinflammatory cytokines such as interleukin 1B and Tumour necrosis factor α , which further induce inflammation and hepatocellular damage. This may induce liver regeneration in order to repair the damage. Usually the inflammation subsides if the initial trigger is removed (Ramadori et al., 2008, Rebouissou and Nault, 2020). However, chronic inflammation may trigger recruitment of myofibroblasts. Myofibroblasts are derived from bone marrow precursor hepatic stellate cells and induce wound healing including the deposition of extracellular matrix leading to hepatic fibrosis (Brenner, 2009).

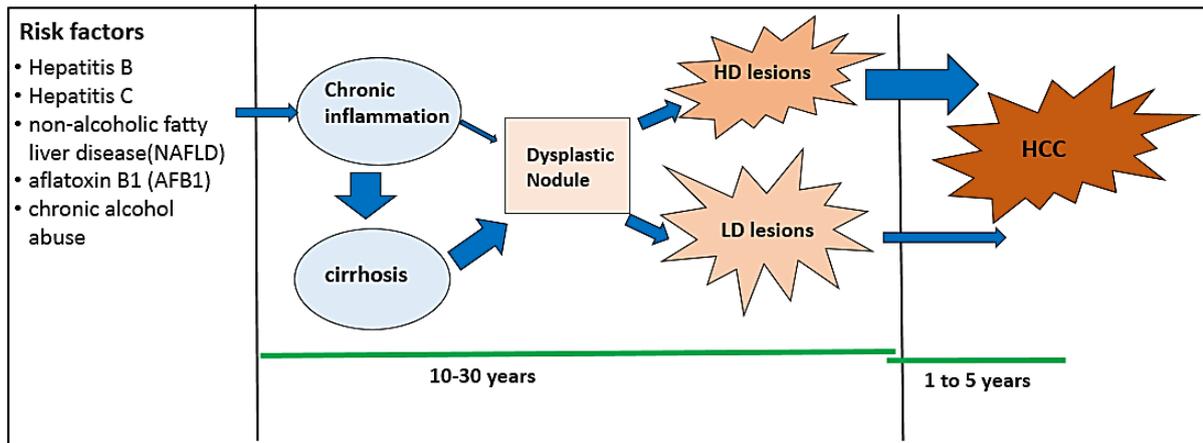


Figure 1.6 Pathogenesis of HCC; HD=high dysplastic, LD=low dysplastic

Chronic liver injury and regeneration may lead to liver cirrhosis. Cirrhosis is characterised by irreversible abnormal liver nodules and are surrounded by collagen deposits. Liver function may also be impaired as reduced blood flow may lead to tissue hypoxia or portal hypertension. Persistent chronic inflammation induces the formation of dysplastic hepatocytic lesions. Dysplastic hepatocyte lesions are abnormal immature hepatocytes and have irreversible structural alterations in gene expression, chromosome structure and are considered to be precancerous lesions. These are categorised as low or high depending upon the morphological structure and presence of atypia. Although both low and high dysplastic lesions can develop into HCC, high dysplastic lesions have a greater risk. High dysplastic lesion may develop into malignant tumour in 30% of patients between 1 to 5 years (Kojiro, 2009).

The presence of stromal invasion usually differentiates between dysplastic lesions and early HCC. HCC is a vascularised and heterogeneous tumour. It may constitute of a single nodule or multiple nodules. Hepatocytes in HCC usually have a prominent nucleolus, high nuclear-cytoplasmic ratio and a granular eosinophilic cytoplasm. However, the cell morphology can be diverse. It can have a pseudo glandular, trabecular, scirrhous or compact tumour pattern. In its early stages, HCC is well differentiated but it can evolve to poorly differentiated tumours with time. Several different driver mutations have been identified in the development of HCC. The alteration of key pathways and epigenetic changes are similar in most HCC but individual pathways may differ depending upon aetiology and tumour stage.

1.2.1 Common mutations and of signalling pathways in HCC

Below are some of the common mutations seen in HCC (**Table 1.2**):

Table 1.2 Prevalence of HCC mutations in patients. Prevalence statistics were obtained from cBioportal (Soumerai et al., 2018).

| Mutation | Pathway | Prevalence, % |
|------------------------|-------------------------------|---------------|
| TERT promoter mutation | Telomerase maintenance | 53.5 |
| WNT/ β -Catenin | Wnt/ β -catenin pathway | 34.6 |
| TP53 | Cell cycle regulation | 30.7 |
| ARID1A and ARID2 | Chromatin remodelling | 13.4 and 4.7 |
| RB | Cell cycle regulation | 3.9 |

Telomerase maintenance

Telomeres are located at the tip of the chromosome and protect against end-to-end fusion and degradation by nucleases and ligases. Telomerase is an enzymatic complex consisting of telomerase reverse transcriptase and the telomerase RNA component. The complex maintains telomeres length by synthesising specific telomeres DNA sequences and attaching them towards the end of the chromosome. In normal cells, the process is primarily suppressed. During each cell division, the telomeres become shorter ultimately leading to cell apoptosis. However in cancer cells, telomerase stabilisation pathways are elevated allowing cells to proliferate indefinitely. In HCC, TERT promoter mutation is the most common mutation observed and occur in 30-60% of HCC. Mutation usually result in the formation of novel ETS transcription factor binding site upstream of the TERT start site. This results in increased TERT transcript expression. TERT promoter mutation can also occur in about 25% of cirrhotic preneoplastic lesions (Nault et al., 2013). Thus, TERT promoter mutation may have a key role or act as a driver mutation in HCC.

Wnt/ β -catenin pathway

The Wnt/ β -catenin pathway has an important role in liver development, metabolism and growth. In HCC, the pathway is usually dysregulated through several mutations, predominantly affecting the CTNNB1 gene. CTNNB1 encodes for β -catenin and most

HCC patients have a point mutation in exon 3. The point mutation inhibits the binding of GSK3 β , which inhibits β -catenin phosphorylation and degradation. β -catenin stabilisation activates the transcription factor LEF-TCF which in turn increases the expression of cell cycle progressive genes such as cyclin D and myc (Lu et al., 2014). The presence and absence of CTNNB1 mutation with Wnt signalling activation demonstrates distinct HCC phenotypes. CTNNB1 mutations with Wnt signalling activation are usually present in low grade tumours and are associated with a better prognosis (Pez et al., 2013, Lachenmayer et al., 2012). In contrast, HCC without CTNNB1 mutations and Wnt signalling activation are usually seen in HBV patients. For this subset, cancers have an aggressive phenotype and high level of genetic instability (Hoshida et al., 2009a, Pez et al., 2013).

Other somatic mutations have been observed specifically targeting negative regulators of the Wnt/ β -catenin pathway. These include missense mutations in Axin 1 and Axin 2. The inactivation of these genes inhibited β -catenin degradation (Guichard et al., 2012). Furthermore a functional study has shown that Wnt signalling pathway may interact with TGF- β pathway in an HCC subset, which had no CTNNB1 mutation and Wnt signalling activation (Hoshida et al., 2009a).

TP53 pathway

TP53 is an important tumour suppressor protein and found to be mutated in about 30% of HCC. It plays a central role in several cellular responses such as cell cycle arrest, DNA repair and apoptosis. TP53 can also activate other tumour suppressor and apoptosis related proteins such as p21 and BAX. Elevated p21 expression inhibits cyclin-dependent kinases and thus hinders cell cycle progression. BAX is an intrinsic apoptosis protein and activates the intrinsic apoptosis pathway leading to cell apoptosis. Low level or mutated TP53 have been observed in many cancers and its mutations are linked to large tumour size and poorly differentiated tumours. In HCC, TP53 mutation correlated with patient aetiology. Several studies have shown a strong association between TP53 mutation, aflatoxin exposure and chronic Hepatitis B infection in development of HCC (Madden et al., 2002). Furthermore in aflatoxin induced HCC, specific TP53 mutation have been observed at codon 249 (Ozturk, 1991).

Retinoblastoma protein and dysregulation of cell cycle regulators

Cell cycle is regulated by cyclin-dependent kinases (CDKs) and activators such as cyclin D and cyclin-dependent kinase inhibitors including p16 and p21. During cell cycle

progression, cyclin D becomes active and binds to CDK4/6. The resulting cyclin/CDK complex phosphorylates the retinoblastoma protein (Rb) resulting in its inactivation. Rb protein is a tumour suppressor and has a growth repressive role. Its gene is deleted in several cancers. Specifically, its deletions lead to increased DNA synthesis and cell division which may support or lead to tumorigenesis (Williams and Stoeber, 2012). Cyclin-dependent kinase hypermethylation have also been observed. For example in HCC, CDKNA2 promoter hypermethylation reduces p16 expression and leads to increased cell cycle progression and proliferation (Wong et al., 1999).

Oxidative stress

Oxidative stress or elevation of reactive oxygen species (ROS) can cause damage to DNA and proteins. In HCC, oxidative stress pathways were elevated in 12% of HCC patients (Schulze et al., 2015). Recent studies have shown that continued oxidative stress is a key characteristic during carcinogenesis. Specifically, the NRF2-KEAP1 pathway is affected and several somatic mutations have been observed. The NRF2-KEAP1 pathway induces a cytoprotective response to ROS. NRF2 acts as a transcription factor in this pathway and mediates the stress oxidative response. In contrast, KEAP1 negatively regulate NRF2. In HCC, somatic mutation of NRF2 and inactivating mutation of KEAP1 have been observed (Menegon et al., 2016).

Akt/mTor and map kinase pathway

The Akt/mTor and map kinase pathway aberration has shown to occur in about 50% of HCC cases. The pathway is induced by ligand binding and phosphorylation of EGF (Epidermal growth factor) and VEGF (Vascular endothelial growth factor) receptors. Subsequently, the pathway activates proto-oncogene cFos and transcription factor AP-1/c-Jun leading to increased cell proliferation (Niu et al., 2016).

TGF- β pathway

The TGF- β pathway plays an important role in the liver by controlling liver architecture and regeneration, but can also contribute to pathological condition such as fibrosis, cirrhosis and HCC (Karkampouna et al., 2012).

A recent transcriptome and genomic analysis were performed on 488 HCC samples from the cancer genome atlas. 38% of HCC samples contained at least one mutation in gene related to the TGF- β pathway. Overall, HCC patients were classed into three major subgroups based on TGF- β signalling pathway status in the tumour tissue compared to normal surrounding liver tissue: activated, inactivated and normal. TGF- β dysregulation was particularly associated with somatic mutation in DNA repair proteins such as RAD51, TP53BP, ATR, FANCD2, FANCM, FAN1 and RAD51. Patients with active TGF- β pathway had increased DNA damage response and activation of sirtuin signalling pathways. Although overall survival was significantly better in patients with TGF- β activated group compared to patients with TGF- β inactivated group ($p=0.013$), overall disruption of TGF- β pathway was associated with poorer patient outcome as patients without TGF- β disruption (normal group) exhibited significantly better survival than in patients with activated or inactivated TGF- β status (Chen et al., 2018). The study highlights the importance of the balance of TGF- β signalling in HCC.

Several studies have shown that TGF- β has a dual role and can act both as a tumour suppressor and as a tumour promoter (**Figure 1.7**). In early stages of HCC, TGF- β acts as a tumour suppressor and inhibits c-myc and certain cyclin dependent kinase inhibitors. As HCC progresses, TGF- β can act as an autocrine and paracrine molecule and can activate stromal fibroblasts. Stromal fibroblasts can activate regulatory T cells and tumour initiating cells. Regulatory T cells are a subtype of T-cells which have immunosuppressive effects on the immune system and decrease the anti-tumour immune response (Giannelli et al., 2014). Tumour initiating cells, also known as cancer stem cells, are a subpopulation of highly tumorigenic cells. These cells can self-renew and are resistant to many conventional therapies (Wu et al., 2012). Elevated tumour initiating and regulatory T cells are both associated with a poor prognosis.

In late stages of HCC, tumour cells can inhibit TGF- β tumour suppressive functions by inhibiting genes such as T β RII. T β RII receptor expression is reduced particularly in malignant hepatocytes compared to surrounding non-malignant tissue (Yamazaki et al., 2011). Similar observations were seen in cell lines with late TGF- β response. These cell lines had no T β R1 receptor expression and low level of T β RII receptor expression (Yamazaki et al., 2011, Matsuzaki, 2013, Nagata et al., 2009). Similarly patients with T β RII receptor mutations are associated with a poor prognosis (Yamazaki et al., 2011). TGF- β dysregulation can also affect several pathways such as epidermal growth factor (EGF) and

Platelet-derived growth factor (PDGF) mediated pathways. The two pathways are mitogenic signalling pathways and have an important role in cell survival and proliferation. EGF has an important role in cell survival and stimulates proliferation via the MAPK/ERK and PI3K-Akt Pathway. PDGF leads to the accumulation of nuclear β -catenin and increases cell proliferation. Both growth factors are elevated in late stages of carcinogenesis and can inhibit TGF- β suppressive functions and increase tumour promoting functions.

Several *in vivo* assays were performed to investigate the role of TGF- β signalling in HCC. TGF- β signalling was investigated *in vivo* by hepatocyte-specific deletion of TAK1 in Albumin-Cre recombinase transgenic mice. TAK1 deficient mice developed spontaneous inflammation, fibrosis and HCC. TAK1 can interfere with R-SMAD activation by binding to the SMAD protein MH2 domain (Hoffmann et al., 2005) or increase the expression of TGF- β inhibitor SMAD7 (Dowdy et al., 2003). Interestingly, TAK-1 null mice developed fibrosis and HCC. Further studies were performed to develop double knockout mice model for TAK1 and SMAD4; TAK1 and TGFBR2. Both double knockout mouse models had a lesser incidence of fibrosis and HCC, particularly in the TAK1 and TGFBR2 knockout mouse model (Yang et al., 2013a). Here TAK1 acts as a possible tumour suppressor gene in HCC and disruption of the TGF- β pathway reduced tumour incidence. In another study, a TP53 knockout mouse model was developed. The mouse model had increased expression of TGF- β 1, Pai1 and AFP. Furthermore, mice with TP53 and T β RII knockout had reduced HCC and cholangiocarcinoma (Morris et al., 2012). Thus, TGF- β signalling may interact with TP53 signalling during HCC formation.

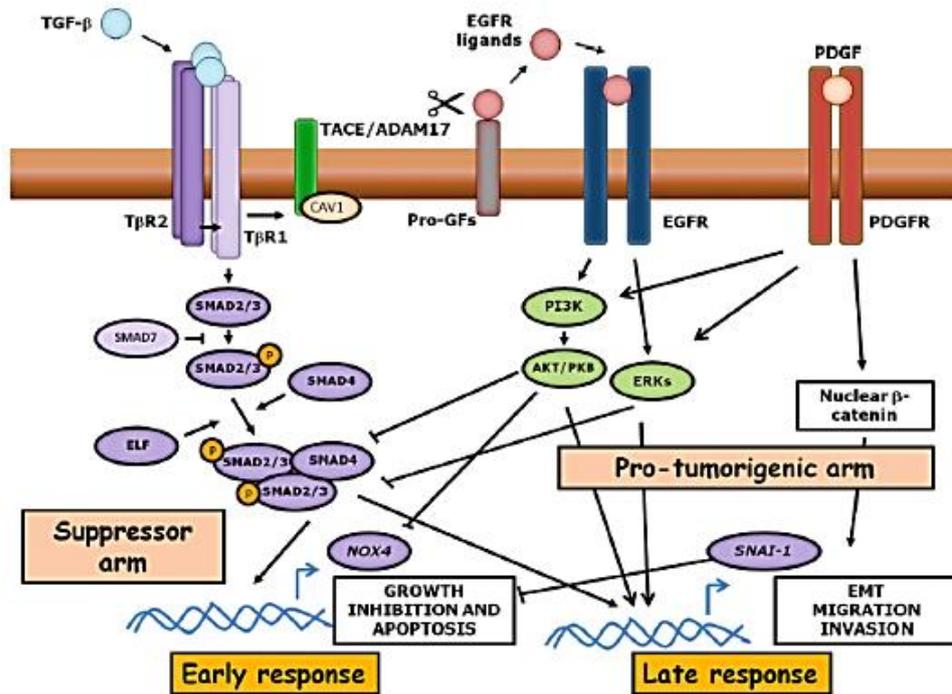


Figure 1.7 TGF- β pathway during early and late carcinogenesis. The pathway has tumour suppressive effects during early carcinogenesis and pro-tumour effects during later stages (Fabregat et al., 2016).

1.2.2 Epigenetic modifications in HCC

Epigenetic modifications are changes in DNA that affect gene expression without affecting the DNA sequence. Several biological processes are controlled by epigenetic modifications including cell signalling pathways, cell division and apoptosis. Thus, epigenetic modifications have an important role in carcinogenesis, and consist of several mechanisms including histone modification, chromatin remodelling and DNA methylation.

Histone modifications

Histone modification is a key mechanism that affects the interaction between histone and DNA strand and its accessibility to transcription factors and regulators.

DNA units are coiled with histone proteins to form a nucleosome. A nucleosome has two copies each of H2A, H2B, H3, and H4, and is wrapped by 147 base pair of DNA. The tightness of the configuration can be affected specifically by the histone N-terminal. The

N-terminal can be methylated, acetylated, ubiquitinated or phosphorylated. The changes can alter the binding between histone proteins and DNA (Cutter and Hayes, 2015, Jun et al., 2012).

Histone acetylation affects the positive charged residues on histone proteins. The process is regulated by two key enzymes: histone acetyltransferases (HAT) and histone deacetylases (HDAC). HAT add acetyl moieties to lysine residues and increases the histone positive charge inducing stronger histone and DNA interaction leading to transcriptional inactivation. HDAC removes acetyl groups and reduces the histone positive charge and thus reduces histone and DNA interaction leading to transcriptional activation. In several cancers, HDAC overexpression has been observed and can affect variety of pathways. HDAC1/2 affect glucose metabolism by inhibiting fructose-1,6-bisphosphatase (FBP1) leading to increased lactate production in HCC (Yang et al., 2017).

Histone methylation can also alter DNA transcription factor binding sites. It can occur in three different levels: mono, di or trimethylation and predominantly occur on arginine and lysine residues. Histone methylation can either be gene activating or gene inhibitory depending upon the histone site. H3K4, H3K79 and H3K36 methylation lead to transcriptional activation. In contrast H4K20, H3K27 and H3K9 methylation lead to transcriptional silencing. Several transcription factors have shown to influence histone methylation. SETDB1 is a methyltransferase and can influence histone methylation. It induces H3K9 methylation leading to gene silencing. SETDB1 knockdown cells had reduced cell migration and invasion capacity in several HCC cell lines such as MHCC97L and Bel-7402 (Zhang et al., 2018). Tiam 1 expression was also downregulated indicating a correlation between SETDB1 and Tiam 1 expression (Zhang et al., 2018, Karanth et al., 2017). Tiam1 is associated with metastasis in several cancers including colorectal cancer, breast cancer and lung cancer, and linked with poor prognosis in HCC (Ding et al., 2009, Izumi et al., 2019, Li et al., 2016, Zhu et al., 2019). Furthermore the transcription factor SNAIL can recruit lysine-specific demethylase 1 (LSD1). LSD1 demethylate histones on H3K4me2 and induces transcriptional repression of E-cadherin (Lin et al., 2010). E-cadherin gene silencing is associated with a poor prognosis in HCC (Kwon et al., 2005).

Chromatin remodelling

Chromatin modellers are important class of proteins that regulate chromatin accessibility to transcription factors and complexes thus can regulate gene expression, DNA repair and apoptosis (**Figure 1.8**). Many of these play key roles in carcinogenesis for example, switch/sucrose non-fermenting (SWI/SNF) complex consists of several key proteins including ARID1A and ARID2A. ARID1A encodes for a BAF250a subunit and has both tumour suppressor and oncogene functions. ARID1A mutations are present in 13.4% of HCC patients. In early stages of hepatocarcinogenesis, it has tumour promoting functions, but its functions are repressed during HCC. Reduced ARID1A expression are associated with tumour progression and metastasis (He et al., 2015). ARID2A is part of the polybromo associated BRG1 associated factor complex and has an important role in nucleotide excision repair of DNA that can be caused by carcinogenic reagents or UV light. It is mutated in 4.7% of HCC patients. ARID1A inhibition increases expression of cell cycle proteins such as cyclin D1 and cyclin E1 leading to increased cell growth. ARID2A restoration decreases cell proliferation in hepatoma cells and decrease tumour growth. Thus, both ARID1A and ARID2A may have tumour suppressive effects (Fujimoto et al., 2012, Li et al., 2011).

The Polycomb repressive complex (PRC) has an important role in chromatin structuring and consist of PRC1 and PRC2. PRC1 interacts with ubiquitin ligases and modifies histone N-terminals. PRC2 induces histone H3K27 methylation (Simon and Kingston, 2009). PRC2 also contains the EZH2 subunit which is elevated in HCC and associated with poorly differentiated tumours, metastasis and poor survival (Sudo et al., 2005). Particularly, EZH2 can induce gene silencing of several Wnt antagonists leading to Wnt signalling activation. Furthermore EZH2 inhibition increases re-expression of natural killer cell ligands in HCC cells leading to increased natural killer cell anti-tumour immune response. Similarly, *in vivo* studies have shown EZH2 inhibition reduced tumour growth in diffuse large B-cell lymphoma, lung cancer and HCC (Cheng et al., 2011, Sudo et al., 2005, Vaswani et al., 2016, Serresi et al., 2018, Zhao et al., 2013).

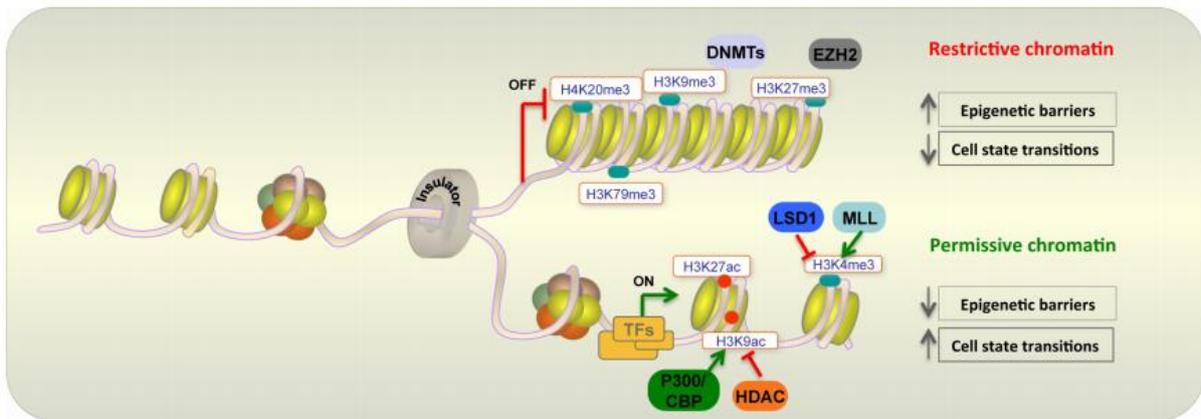


Figure 1.8 Chromatin structure regulation and effect on gene expression (Nebbio et al., 2018)

DNA Methylation

DNA methylation is a chemical modification of DNA wherein the transfer of S-adenosyl methionine to cytosine carbon 5 of CpG dinucleotides takes place. The majority of CpG dinucleotides are methylated and located within the heterochromatin. Unmethylated CpG dinucleotides tend to cluster in CpG islands. CpG islands are at least 200 base pairs DNA sequences and tend to cluster in the promoter region. Promoter regions have at least 50% CpG islands. DNA methylation can affect transcription by two ways: (1) by directly inhibiting the binding of transcription factors and regulators to DNA, or (2) interact with the methyl binding domain (MBD) protein family and induce the recruitment of chromatin modifier such as histone deacetylases leading to chromatin condensation and gene inactivation (Klose and Bird, 2006).

Several molecules are involved in regulating DNA methylation particularly DNMT1, DNMT3A and DNMT3B. DNMT1 has a high affinity for hemimethylated DNA and maintains DNA methylation pattern particularly during DNA replication. DNMT3A and DNMT3B can induce de novo DNA methylation and have no preference to hemimethylated DNA (Klose and Bird, 2006). In contrast, ten-eleven translocation enzymes (TET1, TET2 and TET3) removes 5'-methylcytosine (5mC) via methylcytosine dioxygenase activity.

Methylation patterns are essential in regulating cell signalling, growth and development. In cancer cells, methylation patterns are dysregulated leading to regional hypermethylation and

global hypomethylation (**Figure 1.9**). Regional hypermethylation occur in CpG islands and are associated with increased DNMTs expression. DNMTs are elevated and correlated with poor survival in HCC (Saito et al., 2001, Dong and Wang, 2014). Several types of DNMTs can promote carcinogenesis. For example, DNMT1 can induce hypermethylation of tumour suppressor genes such as MYOCD, LHX9 and PANX2 (Xie et al., 2015). Other key tumour suppressor genes are also hypermethylated in early stages of HCC such as CDKN2A, SOCS1, APC, HIC1, GSTP1, RUNX3, PRDM2 and RASSF1. Therefore, regional methylation could be an early step or driver of hepatocarcinogenesis (Nishida et al., 2012). Furthermore, DNMT3a and DNMT3b have shown to affect several oncogenic pathways. For example, DNMTs can regulate MTA1 gene in patients with HBV HCC. Particularly, HBV X protein can interact with DNMT3a and DNMT3b to induce TP53 promoter methylation leading to increased MTA1 transcription. MTA1 can support carcinogenesis by inducing angiogenesis, invasion, metastasis, and survival (Sen et al., 2014). In a separate study, DNMT3a depletion caused demethylation of tumour suppressor promoter PTEN leading to reduced cell proliferation in HCC cell lines (Zhao et al., 2010). The transcription factor SNAIL can induce histone modification to inhibit E-cadherin as described earlier but it can also interact with DNMT3a and DNMT3b to induce E-cadherin promoter methylation (Lin et al., 2010). E-cadherin promoter methylation has been associated with invasion and metastasis. DNMT inhibitors have been demonstrated to reduce metastasis in HCC (Ding et al., 2012b). Besides affecting tumour suppressors and oncogenes, global hypomethylation can promote chromosomal and genetic instability. A key group of active retrotransposable elements are long interspersed class elements-1 (LINE-1 or L1). L1 hypomethylation have been demonstrated in several cancers including HCC (Zhu *et al.*, 2014).

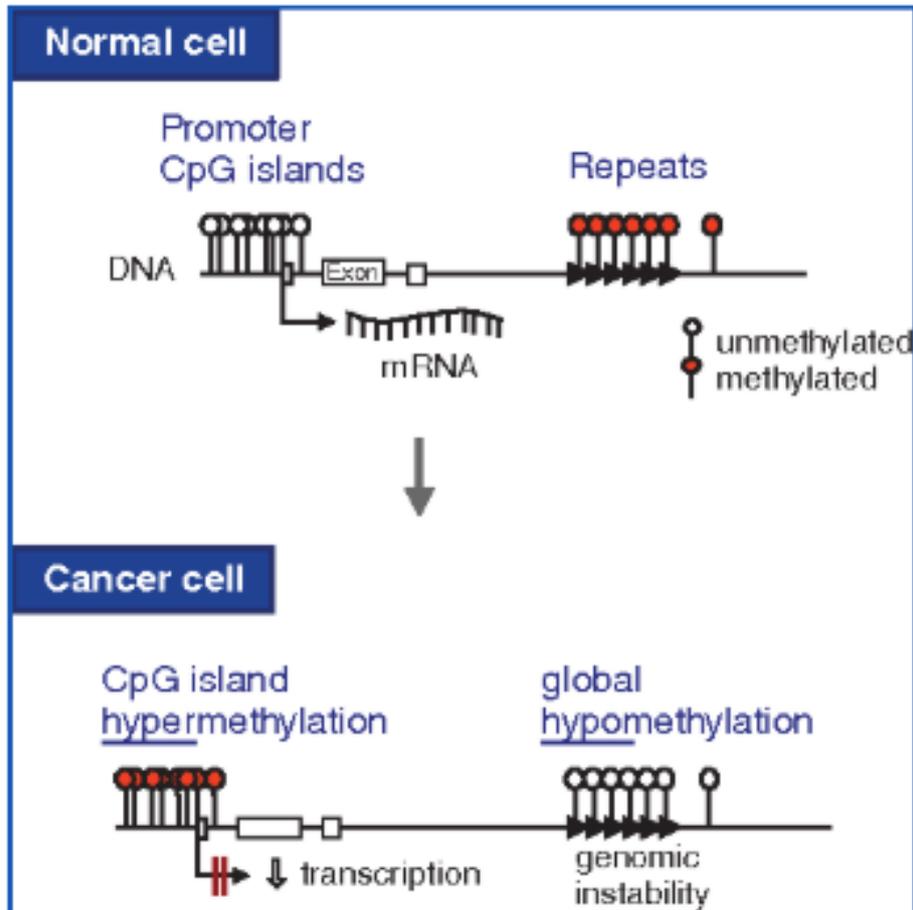


Figure 1.9 DNA methylation changes during carcinogenesis (Sharma et al., 2010, Gerhauser et al., 2015).

1.3 L1 elements

L1 retrotransposons are autonomous non-long terminal repeat elements and constitute 17% of the human genome. Collectively, there are about 500,000 L1 elements in the genome and they consist of 50 different families and subfamilies. However, in humans the only active L1 family is L1Hs (L1 human specific). The L1Hs elements are clustered into pre-Ta and Ta subfamilies. The Ta subfamily is subclustered further as Ta-0 and Ta-I. The vast majority of L1 are inactive and about 80-100 L1 elements remain active in the genome. Most of these active L1 elements belong to the Ta subfamily. Several studies have demonstrated that retrotransposons can lead to human diseases. (Kazazian et al., 1988) have observed the first L1 insertion in X-linked gene, specifically L1 insertion in the F8

exon. The exon is coding for coagulation factor VIII and was observed in a haemophiliac A patient. L1 insertions have also been demonstrated to affect developmental processes and influence behaviour by influencing multiple gene products. Thus these studies demonstrate that L1 are still active in the human genome.

Furthermore, active L1s may vary in frequency in individuals and populations. Fosmid end resequencing and mapping identified 68 novel L1 retrotransposons insertion polymorphisms which were differentially present in individuals but not present in the human genome reference sequence. The majority of the 68 novel L1 retrotransposition insertions were active in a L1 retrotransposition cultured cell based assay. 37 were considered as active L1 elements. Furthermore, L1 genotyping revealed that a subset of L1 were specific to Africans and some were absent from the Human Genome Diversity Panel (HGDP)(Beck et al., 2010).

1.3.1 Structure of L1

Active L1 elements are 6kb in length and consist of a 5'UTR, three open reading frames L1orf0, L1orf1 and L1orf2 and a 3'UTR poly(A) signal, and are flanked by two target site duplications (TSDs) (Furano, 2000), **Figure 1.10**



Figure 1.10 structure of an active L1 element. On the sense strand, it has two key open reading frames, L1orf1p and L1orf2p. L1orf1p is a RNA binding protein, L1orf2p has endonuclease (EN) and reverse transcription activity (RT). ORF0 is located within the antisense promoter.

L1 5'UTR

L1 5'UTR untranslated region is about 900 nucleotide long and contains CpG island which is heavily methylated in normal cells (Furano, 2000, Woodcock et al., 1997). It contains both sense and an antisense promoter (L1 ASP). The sense promoter is an internal promoter and transcribes the full L1 RNA. Antisense promoter transcribes chimeric transcripts. These chimeric transcripts have a portion of the L1 5'UTR and genomic sequences which

flanks the end of L1 5'UTR. L1 chimeric transcripts have been previously observed in cancer specimens (Cruickshanks and Tufarelli, 2009). Transcriptional profiling revealed that L1 5'UTR ASP are involved in diverse cellular processes including mitosis, intracellular protein transport, protein modification and vesicle mediated transport. A recent study has also demonstrated that L1 ASP activation induces cell cycle progression and cell growth (Criscione et al., 2016). Furthermore, chimeric transcripts may promote the formation of double stranded L1 RNA and regulate L1 retrotransposition through RNA interference mechanisms (Yang and Kazazian, 2006).

L1orf0p

ORF0 is located on the antisense strand and codes for a 71 amino acid peptide. The protein product is located in promyelocytic leukaemia-adjacent nuclear bodies. Currently its functions is unknown but some studies have suggested that it may enhance L1 mobility (Denli et al., 2015).

L1orf1p

L1orf1p is a 40kDa protein and has 338 amino acids. It has both RNA binding and nucleic acid chaperone activity. L1orf1p structure was analysed using atomic force microscopy and shown to have a dumbbell structure (**Figure 1.11**). L1orf1p crystal structure analysis demonstrated it can trimerize via its N-terminal with other L1orf1p molecules. The structure is semi stable, which can partially open indicating L1orf1p can form larger assemblies on L1RNA transcript (Khazina et al., 2011).

The N-terminal consist of a highly variable N-terminal domain and a conserved coiled coil (CC) domain. The CC domain is predicted to contain a leucine zipper, which is important in protein-to-protein interaction. The C-terminal consist of an RNA recognition motif (RRM) and a C-terminal domain (CTD). The CTD also contains residues that have high affinity for RNA binding and nucleic acid chaperone activity (Martin, 2010). For example, mutations on the CTD residues Tyr282–Ser287 disrupted L1orf1p binding to L1RNA and inhibited L1 ribonucleoprotein (RNP) formation. Similarly, mutations in Arg206, Arg210, and Arg211 of the RRM, and Lys133, Lys137, Lys140, and Arg141 in the CC domain reduced RNA binding affinity. These studies indicate that both RRM and CC domain are important in RNA binding function of L1orf1p (Khazina et al., 2011, Martin et al., 2005). Both the RRM and CTD form an intervening cleft that is highly positively charged and is likely to interact with the backbone of single-stranded RNA.



Figure 1.11 Structure of L1orf1p consisting of a N-terminal domain, coiled coil domain, RNA recognition motif and a C-terminal domain.

L1orf1p has several phosphorylation sites that regulate its function. These are predominantly serine and threonine phosphorylation sites. Some of these sites are docking motifs for proline-directed protein kinases (PDPKs) such as cyclin-dependent kinases, mitogen-activated protein kinases and glycogen synthase 3. These kinases phosphorylate serine or threonine residues with a proline residue at position +1 (S/T-P motifs) (Khazina and Weichenrieder, 2009, Khazina et al., 2011). L1orf1p has four S/T-P motifs: S18P19, S27P28, T203P204, and T213P214. T203 and T213 motifs have been predicted to be key PDPK sites for L1orf1p function. These two sites are flanked by three conserved arginines (R206, 210 and 211) and mediate RNA binding (Khazina and Weichenrieder, 2009, Khazina et al., 2011). T203 or T213 phosphorylation can also lead to hydrogen bond formation within the guanidino group of the three arginines. Subsequently, it may affect L1orf1p RNA binding (Mandell et al., 2007). Moreover, L1orf1p has several PDPK docking motifs such as T241 and T250 are protein kinase A binding sites. These are highly conserved and substitution mutation of a non-phosphorylated residue reduced L1 retrotransposition by 0-40% (Cook et al., 2015).

PDPK motifs are also regulated by other proteins such as Pin1. Pin1 is a prolyl isomerase and induces proline isomerization in phosphorylated S/T-P motifs (Lu et al., 2002, Liou et al., 2011). Proline isomerization induces a conformational change and enhances phosphorylation (Litchfield et al., 2015, Nishi et al., 2014). Pin1 has shown to interact with various phosphorylated S/T-P motifs in L1orf1p such as S18 and S27 sites. Pin1 protects

S18 and S27 phosphorylation state by inhibiting cis trans prolyl-sensitive phosphatases. S18 or S27 mutation have shown to reduce retrotransposition between 60-80%. Further analysis also revealed that PDPK phosphorylation is not essential for L1orf1p annealing RNA but are important for retrotransposition. Interestingly, mutation in the PDPK phospho-acceptor lead to L1 retrotransposition inhibition (Cook et al., 2015). Further research is required to investigate how the different phosphorylation sites may interfere with L1orf1p function and its interaction with different pathways.

L1orf2p

L1orf2p is a 150kDa protein and has 1278 amino acids. It has both endonuclease and reverse transcription activities. It consists of four key domains: N-terminal domain, seven subdomains of reverse transcriptase, Z-domain, C-terminal domain and a cysteine rich domain. The N-terminal domain consist of an Apurinic/Apyrimidinic endonuclease domain (Feng et al., 1996, Saxton and Martin, 1998.). The C-terminal domain has an important role in nucleic acid binding during retrotransposition (Moran et al., 1996). The reverse transcription domain and Z region are important for reverse transcriptase function. The 3' cysteine rich motif has been shown to have an important role in retrotransposition by supporting nucleic acid binding (Feng et al., 1996, Saxton and Martin, 1998.).

Normally L1 proteins are cis-acting i.e. the proteins favourably binds to L1 transcripts (Wei et al., 2001) however, they can bind to any transcript and induce cDNA conversion. Non-autonomous elements such as Short interspersed nuclear elements (SINE) use the L1 reverse transcriptase and endonuclease domain for its integration into the genome (Schmid and Maraia, 1992). SINEs are nonautonomous retrotransposons which do not encode any proteins but rather hijack L1 proteins. SINE consist between 85 to 500 bps and are located in gene rich regions, whereas L1 are located in intergenic regions (Dewannieux et al., 2003). Furthermore, SINEs are transcribed by RNA polymerase III (pol III) which is initiated at its internal promoter (Kramerov and Vassetzky, 2011).

L1 3'UTR

3'UTR region contains about 200 nucleotides and is attached to a conserved G-rich polypurine tract. The G-rich polypurine tract forms intrastrand tetraplexes (Howell and Usdin, 1997). The G-rich polypurine tract function is currently unknown. Some studies suggested that it may interact with nuclear export factor 1 (Lindtner et al., 2002). L1 also contains a poly A tail. The poly A tail has a weak termination codon and read-through can

occur. In silico studies have estimated that 15% of L1 have read-through past the termination codon, leading to 3' flanking DNA and creation of about 19-30.5 Mb of new DNA to novel L1 DNA sites (Goodier et al., 2000).

1.3.2 Mechanism of L1 retrotransposition

L1 retrotransposition has several steps. These include transcription, translation, formation of L1 ribonucleoprotein and reverse transcription for genomic integration (**Figure 1.12**).

- **Transcription**

As the TATA box is absent in the internal L1 promoter, several cis-acting sequences have been identified which can regulate L1 transcription. For example, RUNX3 and YY1 (Ying-Yang 1) transcription factor binding sites have been identified as positive regulators of L1. These binding sites are located at nucleotide positions +13 to +21 of L1 5'UTR (Nigumann et al., 2002, Yang et al., 2003). Particularly, the YY1 binding site has an important role in transcript initiation and mediates accurate initiation at the +1 site within 5'UTR (Seto et al., 1991). YY1 may interact with other trans-acting factors located downstream of the promoter to induce RNA polymerase recruitment. L1 transcription is mediated by RNA polymerase II (Ostertag and Kazazian, 2001, Hirose and Manley, 1998). L1 transcripts are then exported into the cytoplasm by an unknown mechanism. Some studies suggested that the nuclear export factor 1 (NXF1) pathway might be involved (Lindtner et al., 2002).

- **L1 proteins translation**

L1orf1p and L1orf2p are translated in the cytoplasm. Several studies have demonstrated that L1orf1p and L1orf2p initiation and translation occurs by scanning ribosomes (McMillan and Singer, 1993, Dmitriev et al., 2007). Ribosomes bind to the L1 5'UTR end and transcribe until they reach a AUG sequence and stop translation (Kozak, 1989). As most eukaryotic translation occurs as monocistronic, the translation of L1orf2p may occur in a several ways. For example, L1 transcript might have an internal ribosome entry site or a translation re-initiation site post L1orf1p stop codon (Dmitriev et al., 2007, Kozak, 1989). Nonetheless, L1orf1p translation is more efficient than L1orf2p translation (McMillan and Singer, 1993). Once L1orf1p are translated, they form a coiled trimer complex with other L1orf1p along the N-terminal domain (Martin, 2006, Hohjoh and Singer, 1996). L1orf1p

trimer complex has a cis preference and colocalises with its own transcript and L1orf2p to form a L1 RNA particle. The cis preference reduces the risk of incompetent L1 transcripts forming ribonucleoprotein (Wei et al., 2001).

L1 RNP translocate back into the nucleus for L1 retrotransposition to occur. As the L1orf1p trimer is 120Kb in size and L1orf2p is 150kb, passage must be either through nuclear pores or by cell division. An early study demonstrated that L1 retrotransposition can occur without nuclear membrane breakdown. Furthermore, Mita et al., (2018) investigated the mechanism of L1 RNP translocation into the nucleus. Immunofluorescence staining for L1orf1p, geminin (marker of S/G2/M phase) and Cdt (marker of G1 phase) revealed nuclear localisation of L1orf1p in Cdt positive cells while geminin positive cells had cytoplasmic L1orf1p localisation. L1orf2p nuclear expression remained unaltered. Similarly, L1 retrotransposition was inhibited in cells which were arrested at G1, S, G2, or M phase (Shi et al., 2007). L1 retrotransposition was also inhibited in primary human fibroblasts during cellular senescence (Kuilman et al., 2010). It suggests that L1RNP can translocate into the nucleus when the nuclear membrane is degraded during cell division. Advanced imaging techniques showed that most retrotransposition occur during the S phase. Hence, L1-RNP enters the nucleus during cell division (Mita et al., 2018).

- **L1 integration**

L1 integration can occur by two mechanisms: Target primed reverse transcription (TPRT) or endonuclease independent (ENi) L1 retrotransposition (**Figure 1.12**). However, L1 integration predominantly occurs through TPRT (Morrish et al., 2002). During TPRT, L1orf2p cleaves the DNA at one strand, which is often at a 5'-TTTTAA-3' consensus sequence. The 3' hydroxyl group is used as a primer for reverse transcription. The L1RNA poly A tail then binds to the 3' hydroxyl group and reverse transcription is initiated to generate the complementary strand. Previous *in vitro* studies were performed on R2 elements to identify TPRT. R2 elements are non-long terminal repeat retrotransposons and can insert themselves into the 28S rRNA genes of several animal species. *In vitro* experiments on the R2 TPRT demonstrated that cleavage of the second DNA strand occurs after reverse transcription. Second strand synthesis follows by an unknown mechanism. Consequently, it leads to L1 cDNA integration. TPRT often results in 5' truncations or inversion (Ostertag and Kazazian, 2001).

The second mechanism of L1 integration is endonuclease independent (ENi). ENi rather than TPRT utilised pre-existing nicks in the DNA and does not require endonuclease activity. It commonly occurs in cells with dysfunctional telomeres, which may be caused either by loss of DNA-PKc or expression of TRF2{Sen, 2007 #333}. Thus this may cause the L1 retrotransposition machinery to use the 3'OH group in the dysfunctional telomeres as primers to initiate endonuclease independent (ENi) L1 retrotransposition.

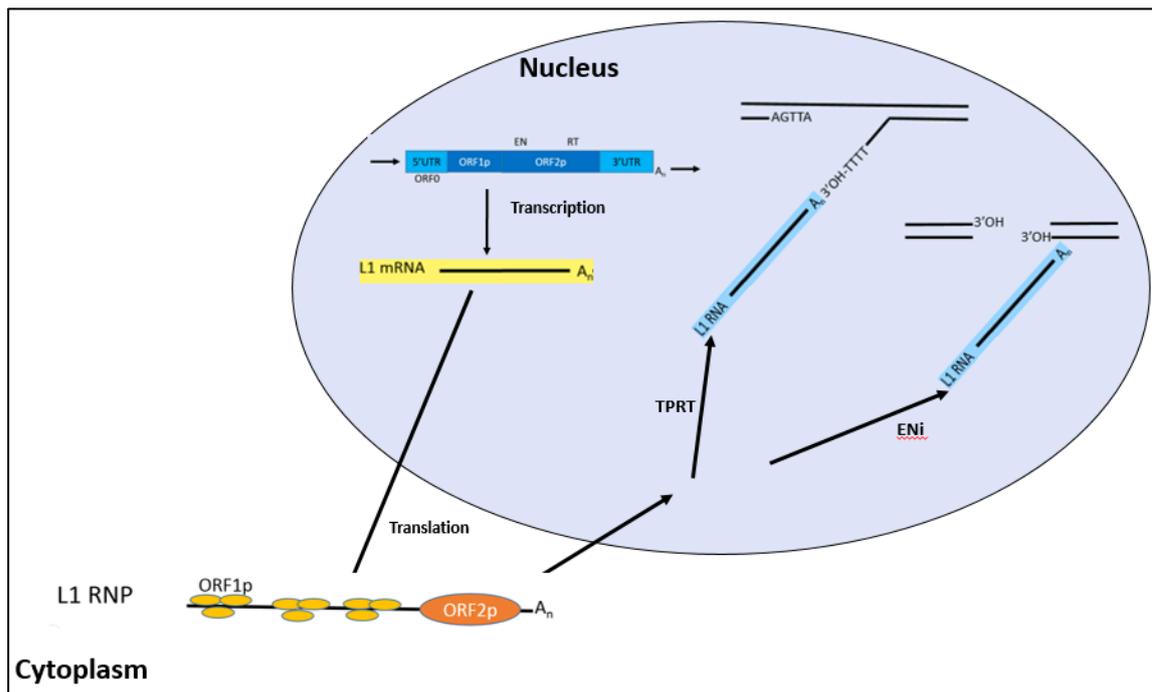


Figure 1.12 Process of retrotransposition. L1 integration can occur by two mechanisms - TPRT and ENi. TPRT requires endonuclease activity for L1 RNA to integrate. In contrast, ENi does not require endonuclease activity. Instead, it can integrate into pre-existing DNA nicks.

1.3.3 Mechanisms of L1 Repression

Since L1 retrotransposition can lead to chromosomal instability and genetic aberrations, cells have evolved several mechanisms to repress the elements and keep retrotransposition in check. Several different mechanisms exist to prevent L1 retrotransposition in cells at promoter, transcript as well as protein level (**Figure 1.13**).

- **DNA methylation**

L1 promoter contains 29 CpG sites in the first 460 base pairs. DNA methylation of the CpG island in L1 promoter can inhibit the transcription of L1 elements. L1 expression is usually elevated in embryogenesis when there is a globally hypomethylated state of the genome but then L1 promoters acquire methylation during early development and become silent. L1 5'UTR methylation has been negatively correlated with L1 expression and thus it is a key L1 inhibitory process. L1 5'UTR methylation is regulated by several epigenetic factors such as DNMT3L. DNMT3L regulates methylation by recruiting DNMT3A and DNMT3B (Liang et al., 2002). DNMT3A and DNMT3B can associate with heterochromatin protein 1 (HP1), methyl-CpG binding proteins (MeCP) leading to L1 promoter methylation (Yu et al., 2001). Further studies revealed that Sirtuin 6 can repress L1 motility through its interaction with the heterochromatin. Sirtuin 6 is a mono-ADP ribosyltransferase and protein deacetylase and binds to L1 5'UTR causing ribosylation of KAP1. KAP1 can then interact with the heterochromatin factor HP1 α and represses L1 transcription (Van Meter et al., 2014). Similarly, promyelocytic leukaemia zinc finger can mediate DNA methylation and induce L1 5'UTR methylation both in germ and progenitor

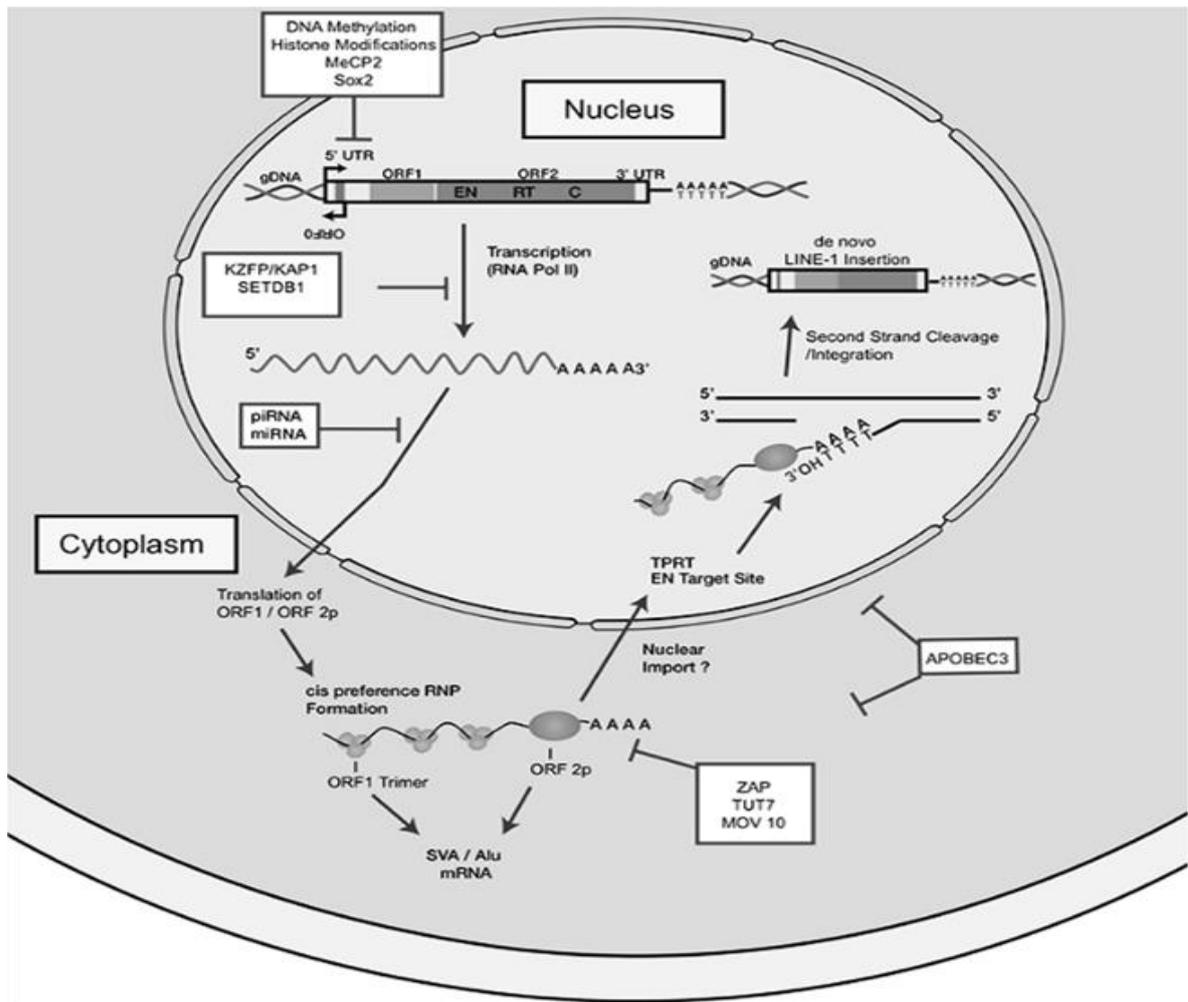


Figure 1.13 L1 and different cellular mechanisms inhibiting L1 retrotransposition (Saleh et al., 2019)

cells (Puszyk et al., 2013). Recently the conserved YY1 transcription factor binding site has also been proved to mediate L1 promoter DNA methylation leading to silencing of these elements both in pluripotent and differentiated cells (Sanchez-Luque et al., 2019).

- **Small RNA molecules**

L1 retrotransposition can also be inhibited by RNA interference molecules. The main RNA interference molecules are siRNA, microRNA, and piRNAs: *in vitro* studies have shown that cells treated with diced L1 siRNAs and L1-related endo-siRNAs had increased promoter hypermethylation and reduced L1 retrotransposition (Chen et al., 2012). Similarly, mir-128 inhibited the integration of L1 RNA in HeLa cells by inducing RNA degradation (Hamdorf et al., 2015). The Piwi-interacting RNA silencing pathway can also

inhibit L1 retrotransposition. Particularly, the pathway induces PIWI protein endonuclease-slicer activity and de novo methylation of transposons (Hamdorf et al., 2015).

- **RNA editing proteins**

APOBEC proteins are part of the cytidine deaminases. In humans, there are seven types of APOBEC3 proteins (A3A, A3B, A3C, A3D, A3F, A3G and A3H) that can inhibit and degrade viruses. Recently, it has been shown that these proteins can inhibit L1 retrotransposition. A3A and A3B are particularly effective in inhibiting retrotransposition (Niewiadomska et al., 2007). Similarly, MOV10 is a potential RNA helicase and inhibits L1 retrotransposon by restricting AGO2. AGO2 is part of a complex consisting of an RNA-induced silencing complex and L1 ribonucleoprotein particle. Its interaction leads to L1 transcript inhibition and degradation (Arjan-Odedra et al., 2012). SAMHD1 can also sequester L1 RNP within the stress granules and thus inhibit L1 retrotransposition (Hu et al., 2015).

- **TP53 pathway**

In normal human cells, TP53 has shown to suppress L1 retrotransposition. Moreover, in cancer cells with TP53 mutation, L1 retrotransposons are found to be significantly elevated. Thus TP53 might be a negative regulatory of L1 (Wylie et al., 2016).

1.3.4 L1 and diseases especially cancer

Despite tight regulation of L1 elements by repressive epigenetic mechanisms, several environmental factors can induce genetic alterations and are capable of inducing L1 retrotransposition. Carcinogens such as benzopyrene are risk factors in several cancers such as: colon cancer, lung cancer and breast cancer, and are also shown to induce L1 retrotransposition in HeLa cells (Stribinskis and Ramos, 2006). Likewise, exposure to specific metals such as nickel can also increase L1 retrotransposition (Kale et al., 2005). Increased L1 retrotransposition is observed in presence of oxidative stress causing chromosomal instability (Giorgi et al., 2011).

L1 insertion can influence expression of genes through different mechanisms depending upon its insertion context with respect to the gene. For example, L1 insertion into an exon can disrupt the coding sequence leading to missense or nonsense mutation. It can also alter the regulatory sequence or induce exon skipping leading to new splicing sites. Gene breaks

can occur, if L1 integrates into an intron in an antisense orientation. In this process, the antisense promoter and polyadenylation tail can split the transcript into two parts and create new transcription start sites at both directions. L1 insertion can also induce cryptic splice sites or induce ectopic recombination such as in non-allelic homologous recombination (NAHR) (Scott et al., 2016) (Burns, 2017).

124 disease causing retrotransposon insertions have been discovered (Carreira et al., 2014). L1 3' end truncation has been observed in the dystrophin gene at exon 67 causing Duchenne muscular dystrophy although the L1 5'UTR was truncated and there was no recognisable sequence (Awano et al., 2010). In factor IX gene causing Haemophilia A, NAHR occurred between two L1 tandem repeats resulting in factor IX gene deletion (Wu et al., 2014). Furthermore, several cancers have increased L1 expression which may lead to genetic alterations. L1 retrotransposition is a hallmark in many cancers especially cancers of epithelial origin (Carreira et al., 2014, Barchitta et al., 2014, Rodriguez-Martin et al., 2020).

Since L1 CpG promoter hypomethylation is a key mechanism of L1 activation, several studies have investigated the association between L1 hypomethylation in normal and cancer tissues and observed higher L1 hypomethylation in malignancies. For example, a meta-analysis consisting of 19 articles was performed on 6107 samples of both tumour and non-tumour samples. L1 hypomethylation was elevated in tumour samples particularly in patients with colorectal and gastric cancer compared to normal tissues (Barchitta et al., 2014). In a separate study, colorectal liver metastases had elevated L1 hypomethylation compared to primary colorectal samples. However, primary and liver metastasis biopsies had similar level of L1 hypomethylation, indicating level of L1 hypomethylation remain similar during tumour progression (Murata et al., 2013). L1 hypomethylation also correlated with poor survival in several cancers. For example, in lung and colorectal cancer, it was associated with a poor prognosis. Similarly in HCC, L1 hypomethylation was associated with a poor prognosis and disease recurrence after resection (Gao et al., 2014b). Interestingly, in melanoma L1 hypomethylation was associated with a favourable prognosis (Sigalotti et al., 2011). These differences in survival may be caused by the different histological tissue types.

1.3.5 L1 promoter activation and retrotransposition in cancer

L1 somatic insertions were observed in several cancers. A previous study has observed elevated L1 somatic insertions in several epithelial cancers by whole genome analysis. The analysis was performed on whole genome paired end sequencing data both from tumour tissues and corresponding blood samples in 45 patients with five different cancer types: colorectal, ovarian, prostate, blood, and brain cancer using a single nucleotide resolution analysis. From the analysis, 183 L1 insertion were identified and each tumour type had an average between 0 to 29 insertions. L1 insertions were present in epithelial cancers only but not in multiple myeloma or Glioblastoma (Lee et al., 2012). Similarly in another study, 290 cancers were analysed for somatic L1 retrotransposition. De novo L1 insertions were predominantly found in heterochromatin regions or in intergenic regions of the genome and showed diverse level of expression in different individuals and tissue types. 53% patients had at least one somatic L1 insertion. L1 retrotransposition events were particularly elevated in colorectal cancer (93%) and lung cancer (75%) patients (Tubio et al., 2014). In a separate study, 1389 distinct L1 insertions were observed in 60 samples (20 primary non-small cell lung cancers and 20 corresponding normal adjacent tissues, 10 human brain tumours and 10 corresponding blood leukocyte controls). In total, 650 were putative novel L1 insertions. Of those, 9 were L1 insertion in lung cancer and 6 were somatic L1 insertions (Iskow et al., 2010).

In HCC, several somatic mutations have been reported. For example in one study, whole-genome sequencing and retrotransposon capture sequencing (RC-seq) was performed on mice and human HCC samples (Schauer et al., 2018). The analysis identified four somatic L1 insertions in 12 mice. One of the somatic insertions belonged to the Tf subfamily and also had a 3'UTR transduction. This L1 insertion had full retrotransposition capacity in cancer cell line. Furthermore, 25 HCC patients with alcohol abuse had 8 tumour specific L1 insertions. Similarly, in 10 intrahepatic cholangiocarcinoma patients, 3 tumour specific L1 insertions were observed. One of the L1 insertion was traced to chromosome 22, which is highly active in cancers. Previously the same group had demonstrated active L1 retrotransposition in HBV and HCV related HCC (Shukla et al., 2013).

These findings indicate elevated L1 insertions in different cancers particularly in epithelial cancers. In fact, one study investigated L1 insertions in pre-cancerous and malignant tissues. Here 4 normal colon, corresponding colonic polyps and colorectal cancer tissues

were analysed by L1-seq and then validated by PCR and Sanger sequencing. In this study, L1 insertion were observed both in colonic polyp and colorectal cancer tissues indicating that L1 may be an early event during colon carcinogenesis and could be a potential driver in carcinogenesis (Ewing et al., 2015).

L1 can support carcinogenesis by several mechanisms including inhibiting tumour suppressor functions. The first bona-fide tumour-related retrotransposon insertion was observed in the APC gene. The somatic insertion was further characterised and had 5' truncation and a 8base pair duplication at the site of insertion, which are characteristics of TPRT (Miki et al., 1992). Scott et al (2016) also identified novel somatic L1 insertion in APC in colorectal cancer using whole genome sequencing. The insertion was complementary to another point mutation in the second allele (Scott et al., 2016). These findings suggest L1 insertion may drive carcinogenesis through a classic two-hit colorectal cancer. Similarly, L1 de novo insertion has been observed in the telomerase reverse transcriptase (TERT) gene. The insertion inhibited the progressive shortening of telomeres and cell apoptosis. Other examples of somatic L1 insertions which potentially drive carcinogenesis are in an exon of the PTEN gene (Helman et al., 2014), in the intron 14 of the RB gene, causing early mRNA splicing (Rodriguez-Martin et al., 2016) and in the intron of ST18 gene, inhibited a negative feedback loop that inhibited its binding to its enhancer thus insertion inhibited the function of ST18 inhibitor leading to the activation of the gene (Shukla et al., 2013). These findings suggest that L1 insertion can affect several tumour suppressor genes and oncogenes thus encourage carcinogenesis.

Besides somatic retrotransposition, germline L1 insertion in Mutated in Colorectal Cancers (MCC) was observed in 4/19 patients. MCC is expressed in the liver and inhibits the β -catenin/Wnt signalling pathway. Its inhibition increased β -catenin/Wnt signalling pathway (Shukla et al., 2013).

L1s can also create chimeric transcripts (**Figure 1.14**). A chimeric L1-MET transcript was observed between an intronic L1 sequence and c-met in HCC. L1-MET correlated with c-MET transcript expression and poor survival (Zhu et al., 2014). Similarly, Hepatitis B virus can integrate into L1 on chromosome 8p11 producing an oncogenic HBx-LINE1 chimeric transcript, and was observed in 23.3% HCC patients (Lau et al., 2014). HBx-LINE1 chimeric transcript expression was associated with increased WNT signalling pathway. Further studies revealed that knockdown of HBx-LINE1 reduced migration and invasion

whereas HBx-LINE1 expression increased migration and invasion in HCC cells (Whittaker et al., 2010, Lau et al., 2014).

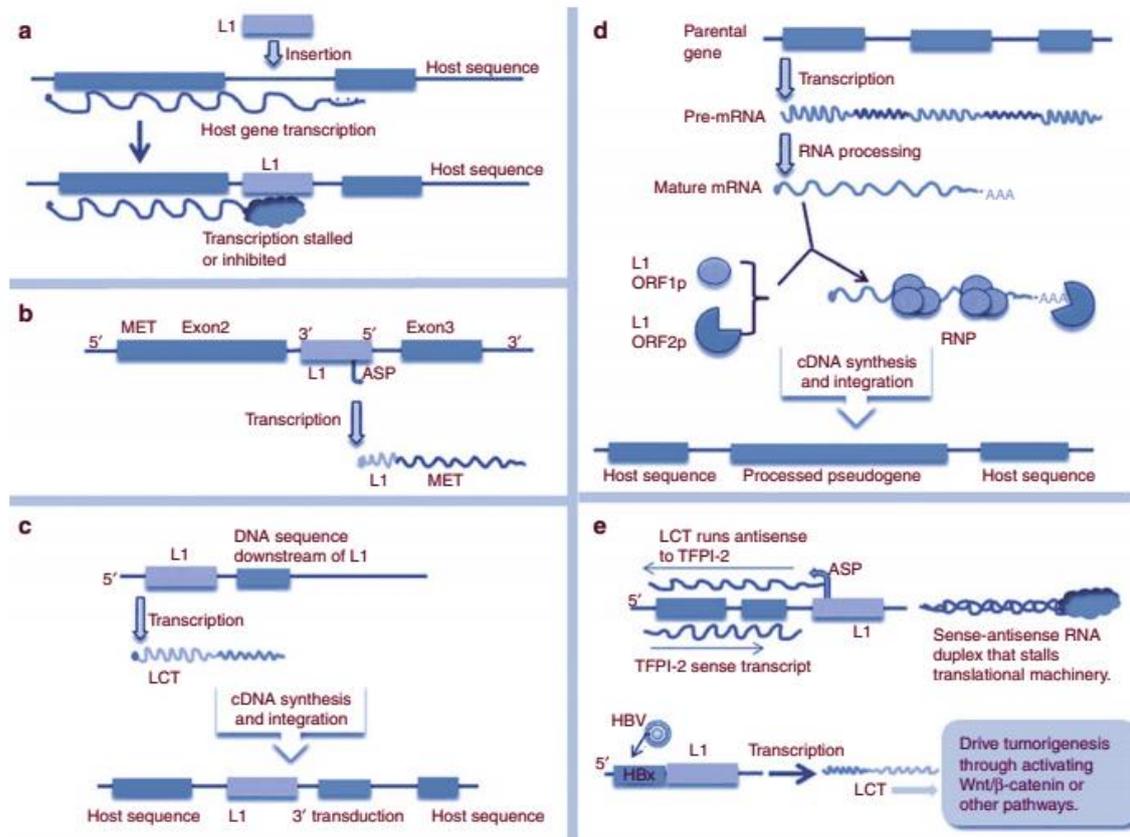


Figure 1.14 L1 and its role during carcinogenesis (Xiao-Jie et al., 2016)

1.3.6 L1 encoded proteins influence on cellular pathways

Besides active retrotransposition, L1 can influence host cell via its increased proteins - L1orf1p and L1orf2p expression. L1orf1p expression were measured in several cancers (1027 patients) by tissue microarray staining. The staining was positive in 47% of the tumours (482 cases) and particularly expressed in the cytoplasm. It was highly expressed in epithelial cancers such as breast carcinoma (97%, 66 of 68 were positive), high-grade ovarian carcinomas (93.5%, 29 of 31 were positive) and pancreatic ductal adenocarcinomas (89%, 56 of 63 were positive). High grade tumours such as sarcomas, pancreatic carcinomas, lymphoma and secondary glioblastomas were also highly positive. Other cancers such as oesophagus, bladder, head and neck, colon, lung, endometrium and biliary tract were 22.6–76.7% positive for L1orf1p. In HCC, L1orf1p expression was positive in 24% of cases. In contrast, corresponding preneoplastic lesion had none or low L1

expression (Rodic et al., 2014). Similarly, high L1orf1p expression were observed in different epithelial cancers such as liver, renal, ovarian, lung and prostate carcinoma (Barchitta et al., 2014).

Further studies demonstrated L1orf1p staining intensity correlated with tumour malignancy and invasion. Thus, malignant tumour cells may produce higher level of L1orf1p and can potentially be a malignancy marker (Asch et al., 1996). However in some cancers, nuclear localisation of L1orf1p has been observed. For example in invasive breast carcinoma, L1orf1p and L1orf2p nuclear localisation correlated with metastasis and poor survival. L1orf1p and L1orf2p nuclear localisation might be more prevalent in certain cancers and may correlate with a worse prognosis (Harris et al., 2010).

Increased L1orf1p expression can affect several oncogenic pathways. For example in prostate cancer, L1orf1p coiled-coil domain can interact with the androgen receptor leading to its translocation to the nucleus. Androgen receptor translocation increases the gene expression of proliferating and anti-apoptotic genes such as VEGF (Lu et al., 2013). L1orf1p can also interact with the ETS-1 transcription factor. ETS-1 is an important transcription factor during development and carcinogenesis and acts downstream of the c-Met signalling pathway. It activates several proliferating and invasion genes such as MMP-1, MMP-9, c-Met, Cyclin D1 and u-PA. In colon and breast cancer, L1orf1p increased ETS-1 transcriptional activity leading to increased cell proliferation and invasion (Bu et al., 2006) (Yang et al., 2013). In a separate study, L1 knockdown cells reduced telomeres length and increased cell apoptosis. Further analysis revealed that telomerase related transcription factors cMyc and KLF-4 transcript and protein expression were reduced (Aschacher et al., 2016). L1orf1p can also affect the TGF- β pathway. An immunoprecipitation analysis has shown that L1orf1p forms a complex with Smad4. The formation of the complex inhibited Smad4 nuclear translocation (Zhu et al., 2013). Role of L1orf1p in drug resistance has also been demonstrated. For example in oesophageal squamous cell carcinoma, L1orf1p can increase gene expression of ATP-dependent efflux pump (MDR1)(Zhu et al., 2015). MDR1 encodes the membrane bound drug transporter P-glycoprotein. The increased expression of p-glycoprotein increases level of drugs being pumped from the cells. Increased levels have been shown to induce doxorubicin and paclitaxel drug resistance. In HepG2 (HCC cell line), L1orf1p overexpression led to epirubicin and cisplatin drug resistance (Feng et al., 2013). Further study revealed that L1orf1p overexpression in HepG2 reduced apoptosis when treated with epirubicin,

cisplatin and paclitaxel due to increased level of BCL-2 (an anti-apoptotic protein). Co-immunoprecipitation also revealed an interaction between L1orf1p and cisplatin-resistance associated proteins CROP or LUC7L3 (Slotkin and Martienssen, 2007). In contrast, HepG2 L1orf1p siRNA cells had significantly reduced IC50 of epirubicin, cisplatin and paclitaxel by up to 9-fold. These findings suggest L1orf1p increases drug resistance and might be a potential therapeutic target or treatment stratification biomarker (Feng et al., 2013).

As L1 expression is elevated in epithelial cancers, it is important to investigate the role of L1 in carcinogenesis. HCC has particularly poor prognosis and several studies have indicated active L1 elements. Though, the role of L1 in HCC carcinogenesis is unknown and requires further studies.

1.4 Aims of the thesis

Purpose of this thesis is to investigate the role of L1 retrotransposons in HCC, specifically the relationship between L1 activation and HCC progression. Overall, the goal is to explore L1 as a therapeutic target or a molecular classifier for HCC..

Specific Aims are as follows

- 1) Explore L1 expression correlations with clinical and pathological features of HCC
- 2) Characterisation of L1 elements in different HCC cell lines and develop an in vitro model with a stable knockdown of L1 encoded protein expression
- 3) Investigate the effect of L1 expression knockdown on cell proliferation, invasion, migration and doxorubicin sensitivity
- 4) Investigate the molecular role of L1 expression in HCC, in relation to the TGF- β pathway.

Chapter 2: L1 expression and correlation with HCC subclasses

2.1 Introduction

HCC is a heterogeneous disease and several studies have focused on genetic and molecular expression patterns to categorise patients into molecular clusters. In a transcriptomes meta-analysis study, 603 HCC patient tumours were categorised into three main clusters: S1, S2 and S3 according to gene signature, somatic DNA alteration and subclasses (**Figure 2.1**). S1 consisted of 28–31% HCC patients and correlated with a cholangioma-like gene signature. A strong association was also observed with steatohepatic HCC, TGF- β activation and TP53 mutation. S2 consisted of 23–24% HCC patients and was prevalent in patients with Hepatitis B positive but no cirrhosis. Patients were positive for several stemness markers such as AFP and Epcam, GPC3, c-myc and had elevated TP53 mutations. S2 cluster was also associated with activated WNT signalling, poorly differentiated tumour and high recurrence. In contrast, S3 tumours were associated with well differentiated tumours and elevated CTNNB1 mutations. Furthermore, S3 cluster was associated with lower tumour recurrence and better survival compared to S1 and S2 clusters (Schmidt et al., 2016, Hoshida et al., 2009b, Hoshida et al., 2010, Goossens et al., 2015).

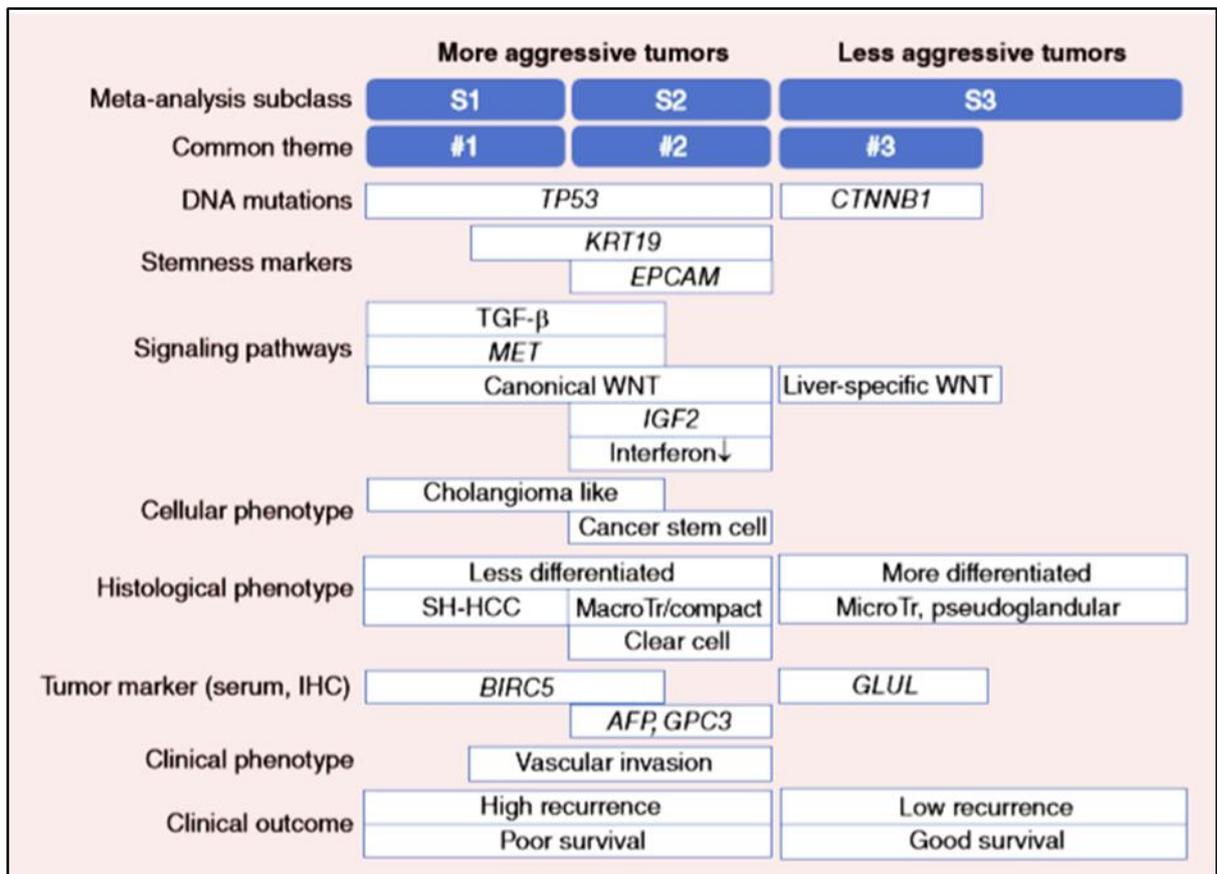


Figure 2.1 Molecular classification of HCC using S clusters (Goossens et al., 2015)

A recent comprehensive analysis integrated transcriptional profiles (RNAseq), microRNA expression, DNA methylation, copy number variations, somatic mutations and reverse phase protein lysate microarray (RPPA) of HCC samples by the TCGA consortium revealed 3 clusters - iCluster1-3 (**Table 2.1**)(Wheeler, 2017). iCluster 1 was associated with young age, female gender, normal body weight and Asian ethnicity. In addition, iCluster 1 was associated with high grade tumour and macrovascular invasion. iCluster 2 was associated with low-grade tumours and less microvascular invasion. iCluster 3 was associated with higher degree of chromosomal instability particularly chromosome 17p loss, hypomethylation of several CpG sites and higher frequency of TP53 mutations. iClusters were also compared with other pre-existing molecular clusters such as Hoshida clustering. A prediction signature was used and demonstrated a correlation between iCluster 1 and Hoshida C2 patients; and iCluster 3 and Hoshida C3 patients.

Table 2.1 Molecular classification of iClusters in HCC

| | iCluster 1 | iCluster 2 | iCluster 3 |
|--------------------------------|-----------------------------------------------------------|-------------------|--------------------------------------|
| Macrovascular invasion | High | Low | - |
| Tumour grade | High | Low | - |
| Chromosomal instability | Low | Low | High |
| CDKN2A silencing | Low | High | High |
| CTNNB1 | Low | High | High |
| TERT promoter mutation | Low | High | High |
| TP53 mutation | Low | Low | High |
| Hoshida clustering | Hoshida S2 | - | Hoshida S3 |
| Key association | young age, female, normal body weight and Asian ethnicity | - | Hypomethylation of several CpG sites |

Besides identifying these molecular subclasses, it is also essential to develop robust biomarkers to identify these subclasses in a clinical setting for prognostic or treatment stratification purposes. Majority of studies have focused on individual genes or gene signatures, their expression, mutation and DNA methylation status however, systematic analysis of repeat elements in this context is still lacking. A recent study investigated L1orf1 protein (L1orf1p) expression in 1027 patient biopsies of different cancer types. Formalin fixed and paraffin embedded patient (FFPE) biopsies were stained for L1orf1p by immunohistochemistry. Staining was positive in 47% of tumours (482 cases). It was particularly positive in epithelial cancers such as breast carcinoma (97%, 66 of 68 were positive), high-grade ovarian carcinomas (93.5%, 29 of 31 were positive) and pancreatic ductal adenocarcinomas (89%, 56 of 63 were positive). However in HCC, L1orf1p expression was positive in about 24% of cases only. Other cancers such as oesophagus, bladder, head and neck, colon, lung, endometrium and biliary tract were 22.6–76.7% positive for L1orf1p. Corresponding preneoplastic lesions had none or low L1orf1p positive cells. Furthermore, high L1orf1p expression also correlated with TP53 mutation. (Rodic et al., 2014). Similar findings were also observed in a meta-analysis wherein high L1orf1p expression was observed in epithelial cancers such as liver, renal, ovarian, lung and prostate carcinoma (Barchitta et al., 2014). In summary, about 50% of tumours of epithelial origin have L1 activation and L1 activation correlates with TP53 mutation.

One study demonstrated an interplay between L1 transcription and P53. L1 contains several P53 DNA binding sites. Furthermore Abrams et al demonstrated that P53 can bind to certain sites in the L1 5'UTR promoter and cause the deposition of restrictive histone marks, thus inhibiting L1 transcription. Interestingly P53 removal or P53 binding site removal in L1 5'UTR promoter caused increased L1 activity and *de novo* transcription. However, Harris et al demonstrated a positive correlation between P53 and L1 transcription leading to cell apoptosis (Harris et al., 2009). These differences may be due to the experimental procedures used. Harris used transient cotransfections to overexpress P53 whereas Abrams used integrated reporters. In addition, Harris's reporter lacked ~100 bp of the L1 promoter. Thus P53 has an important role in interacting with L1 through its promoter, however further studies are required to identify specific regulators in this process.

2.2 Aims

In this chapter, the role of L1 as a potential biomarker and prognostic marker will be investigated in HCC samples.

Specific aims are as follow:

- Patient HCC RNAseq dataset (The Cancer Genome Atlas Hepatocellular carcinoma study (TCGA-LIHC)) will be analysed for key transcriptional changes between L1 5'UTR counts and clinical features and pathway changes.
- Patient formalin fixed and paraffin embedded (FFPE) HCC samples will be stained for L1orf1p to characterise level of staining between tumour and corresponding non-tumour tissues.
- Several HCC aetiological features will be compared with L1 staining. Further clinical features such as tumour grade, liver function and survival will also be compared.
- Patient FFPE HCC samples will be stained for pSMAD3 to measure TGF- β signalling and correlate with L1orf1p staining.

2.3 Material and Methods

2.3.1 Ethical Approval

Ethical approval was obtained for the use of FFPE HCC patient biopsies by the National Research Ethics Service (NRES) Committee North East (12/NE/0395).

2.3.2 Patient cohort for immunohistochemical analysis

FFPE HCC diagnostic biopsies were obtained from our own biobank for patients visited between 2002 to 2018 and who consented to the use of their tissues surplus to diagnostic requirements for research purposes. Clinical features were obtained from patient medical records. A total of 48 patients were obtained, 37 males and 11 females ranging from 49 to 85 years of age (median 71). Most patients were overweight or obese (median BMI: 30) and were diabetic (62.5% patients). Cirrhosis was present in 39.6% patients and NAFLD was the most common aetiology, which was present in 45.8% patients. In terms of tumour histology, tumour size had a median of 4cm and most patients had a BCLC stage of C (45.8%). Furthermore, most patients had been treated with TACE (50%).

2.3.3 Immunohistochemistry

4µm FFPE sections were stained using an automatic immunohistochemistry machine (Ventana Discovery XT). In short, slides were de-paraffinized in EZ prep solution and antigen retrieval was performed by heat retrieval at 95°C for 44 minutes, pH of 7.8 in Tris-EDTA buffer. Primary L1orf1p antibody was added (Gift from Prof K Burns Rodric 2014, 1:3000 dilution in TBST) and incubated for 1 hour. Primary pSMAD3 (Santa Cruzsc-517575) was added with TBST at dilution of 1:50. Secondary antibody was added (OMAP anti- Ms HRP secondary antibody) and incubated for 20 minutes. Slides were counterstained with haematoxylin, dehydrated and then mounted with DPX. Resection tissues of colorectal cancer metastasised to liver were used as a positive control. Colorectal cancer tissues without primary antibody, and corresponding secondary antibody were used as a negative control.

2.3.4 L1orf1p and pSMAD3 scoring

L1orf1p slides were scored by a pathologist and categorised based on staining intensity (1-5) and percentage of positive cells. Scores were combined to finally categorise the samples into two groups: L1_low and L1_high.

pSMAD3 protein expression were examined by immunohistochemistry in a cohort of 17 HCC FFPE patient samples. Slides were scored and classified into the following groups: None (50% positive), low (25% positive), and high (> 25% positive).

2.3.5 L1orf1p algorithm scoring

FFPE slides were digitally scanned onto the Aperio software and analysed using an in-house optimised algorithm to measure cytoplasmic staining using Aperio® Image analysis (**Table 2.2**). The algorithm produced cytoplasmic H-score based on percentage of positive cells and staining intensity. The cytoplasmic H-score was calculated on the software using the following formula: 3x percentage of strongly stained cytoplasm+2x percentage of moderately stained cytoplasm+1x percentage of low stained cytoplasm. Cytoplasmic H-score ranged between 0 to 300, with 0 being totally negative and 300 being all hepatocytes highly positive for L1orf1p. The selection of normal versus tumour areas was supervised by a liver pathologist. Researchers were blinded until the study endpoint.

Table 2.2 Aperio algorithm parameters used for H-score quantification in HCC FFPE samples

| Algorithm parameters | Values |
|--------------------------------------------------------------------------------|----------------|
| Version | 2.0 |
| Algorithm type | Cytoplasmic v2 |
| Height | 1250 |
| Width | 1250 |
| Overlap size | 100 |
| Averaging Radius (μm) | 1 |
| Cytoplasmic Distance (μm) | 6 |
| Max Cell Dimension(μm) | 25 |
| Pixel Area (millimeter-squared) | 2.502e-007 |
| Classifier Neighbourhood | 25 |
| Min Nuclear Area (μm^2) | 20 |
| Maximum allowable distance for cytoplasm surrounding nucleus (μm) | 6 |
| Minimum threshold | 0 |
| Maximum threshold | 240 |
| (1+) Threshold | 210 |
| (2+) Threshold | 180 |
| (3+) Threshold | 150 |
| OD for Red | 0.696858 |
| OD for Green | 0.643073 |
| OD for Blue | 0.317563 |
| OD for positive Red | 0.244583 |
| OD for positive Green | 0.509334 |
| OD for positive Blue | 0.825081 |

2.3.6 HCC RNAseq analysis

RNAseq dataset was downloaded from the TCGA-LIHC study and reanalysed to assess L1 promoter counts (L1-5'UTR, L1Ta subfamily). The reads were mapped to the human active L1-Ta sequence (Genbank: L19092) by BLAT alignment. Counts corresponding full length L1-Ta and just the 5'UTR (promoter region) were obtained and were normalised by the total number of reads in each library.

2.3.7 Statistical analysis

Statistical analysis was carried out on IBM SPSS Statistics Version 24. Wilcoxon signed-rank test was performed for Tumour and Non-tumour samples. Mann-Whitney test was performed for gender, hepatitis B, hepatitis C, NAFLD, alcohol liver disease, cirrhosis, vascular invasion, SNUR, HB16 and TP53 mutation (TCGA samples), and H-score and pathologist scoring (FFPE samples). Spearman's rank correlation was performed for Age, TP53 gene expression targets (TCGA) and pSMAD3 and L1orf1p staining (FFPE). Kruskal-Wallis Test was performed for pathological stage, iCluster and TGF- β status. Survival analysis was performed using Kaplan Meier analysis.

2.4 Results

L1Ta full-length and L1Ta 5'UTR (L1 5' UTR) counts were analysed in HCC patients RNAseq dataset (**Figure 2.2**) and demonstrate a strong correlation ($r=0.597$, $p<0.001$). L1 5'UTR counts is a more reliable measure of legitimate L1 transcripts due presence of L1 promoter in the region. As both demonstrate a strong correlation, L1 5'UTR counts were taken further to evaluate correlations between L1 expression and different clinical factors.

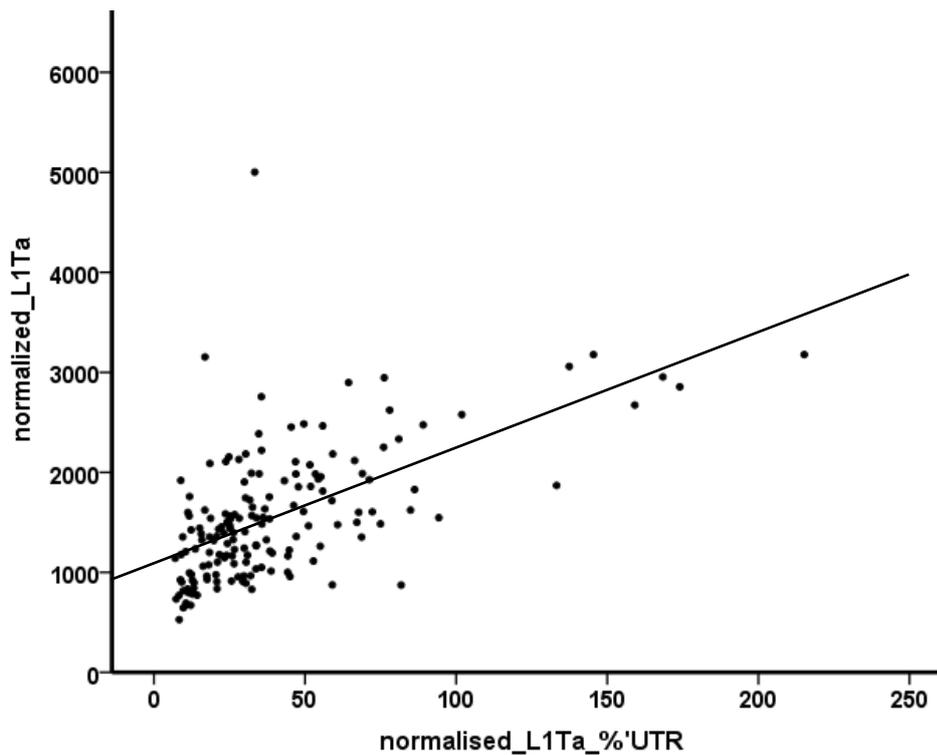


Figure 2.2 L1Ta transcript correlates with L1 Ta 5'UTR promoter counts ($r=0.597$, $p<0.001$ Spearman correlation).

2.4.1 L1 5'UTR counts are elevated in HCC tumours

L1 5'UTR counts were found to be significantly elevated in tumour samples (median: 26.73) compared to non-tumour samples (median:16.25), $p < 0.0001$ Wilcoxon rank test (Figure 2.3).

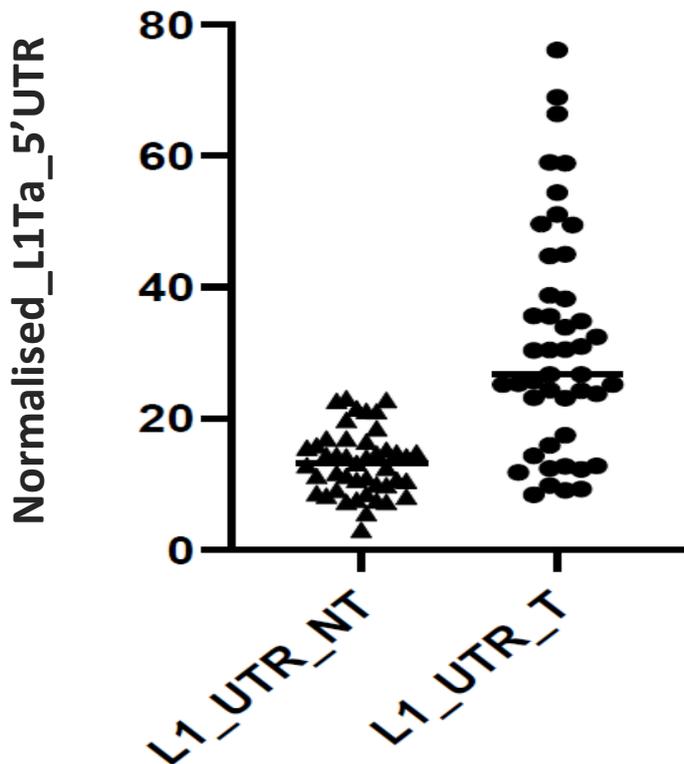


Figure 2.3 L1 5'UTR counts of tumour and corresponding non-tumour patient samples, $p < 0.0001$ Wilcoxon rank test (n=45).

2.4.2 HCC L1 5'UTR counts correlated with high pathological stage, recurrence and poor prognosis

L1 Ta 5'UTR counts were compared with different aetiological factors (Table 2.3). No correlations were observed with gender ($p = 0.381$, Mann-Whitney test), age (spearman $r = 0.12$, $p = 0.093$), Hepatitis B ($p = 0.990$), Hepatitis C ($p = 0.162$, Mann-Whitney test), NAFLD ($p = 0.408$, Mann-Whitney test), alcohol liver disease ($p = 0.652$, Mann-Whitney test) and cirrhosis ($p = 0.677$, Mann-Whitney test).

Table 2.3 L1 5'UTR counts and correlation with different aetiological factors in 196 HCC patients. Mean L1 5'UTR counts and standard error was calculated for each aetiology. Number in brackets represents patient count. Mann-Whitney test was performed for gender, Hepatitis B, Hepatitis C, NAFLD, alcohol liver disease and cirrhosis. Spearman regression analysis was performed for Age.

| Normalised_L1Ta_5'UTR | | | |
|------------------------------|--------------------------|-------------------------|--------------------|
| Clinical features | | Mean ± SE | p value |
| Gender | Male (129) / Female (67) | 36.4 ± 2.9 / 38.3 ± 3.8 | 0.38 |
| Age | | 60.2 ± 1.0 | r=0.12 p= 0.093 |
| Risk factor | Hepatitis B (44) | 40.1 ± 9.0 | 0.99 |
| | Hepatitis C (35) | 31.3 ± 4.5 | 0.16 |
| | NAFLD (11) | 39.2 ± 7.1 | 0.41 |
| | ALD (64) | 37.1 ± 3.6 | 0.65 |
| Cirrhosis | No (142) / Yes (48) | 36.7 ± 2.7 / 39.6 ± 5.2 | 0.68 |

Table 2.4 pathological stage group

| Stage group | Description |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | A single tumour larger than 2cm but has not invaded a blood vessel. The cancer has not spread to any lymph node or sites |
| 2 | Single tumour larger than 2cm which has invaded a blood vessel or a single tumour larger than 5cm. The cancer has not spread to any lymph node or sites |
| 3 | More than 1 tumour larger than 5cm or any tumour size that has grown into the portal or hepatic vein. The cancer has not spread to any lymph node or sites |

Next, we investigated the correlations between HCC L1 expression and tumour features like pathological staging based on tumour number, size and invasion (see **Table 2.4**). L1 5'UTR counts were significantly elevated in Stage 2 (mean+SE 44.8±5.9) and Stage 3 (43.3± 4.9) tumours compared to Stage 1 tumours (29.2±2.5), $p<0.05$, Kruskal-Wallis Test (**Figure 2.4A**). Likewise, L1 5'UTR counts were significantly higher in patients with vascular invasion than without (with= mean+SE 45.8± 5.6 and without 33.7±, 2.7 $p=0.015$ Mann-Whitney test), **Figure 2.4B**.

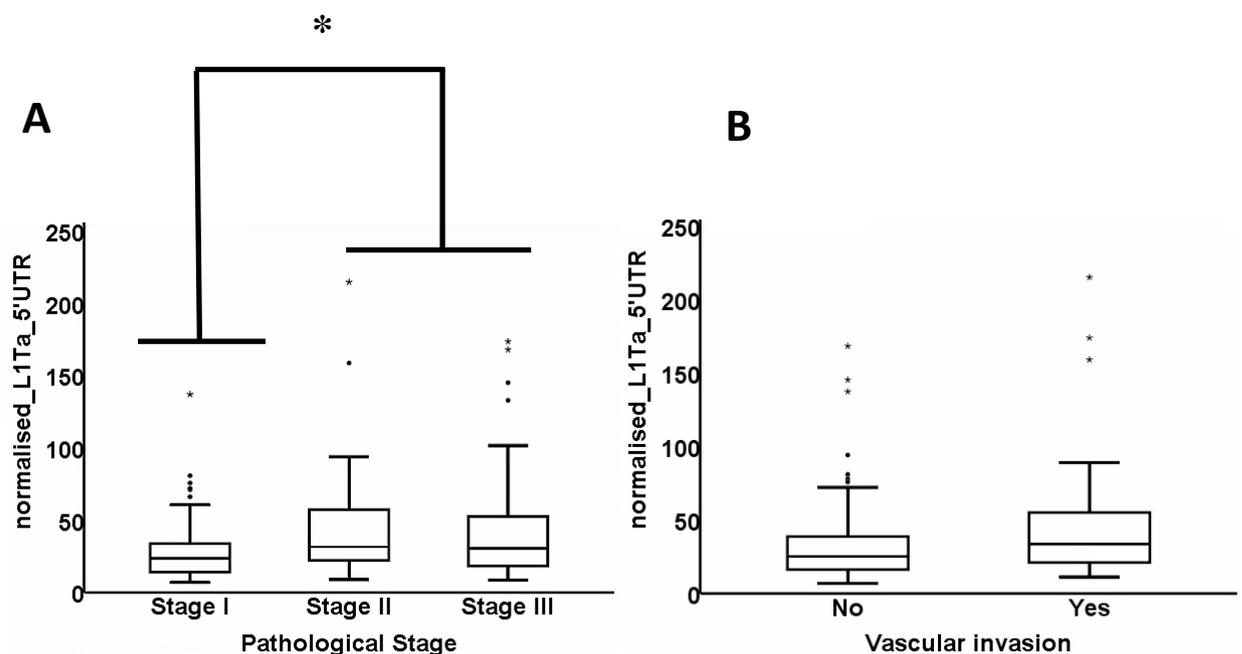


Figure 2.4. L1 5'UTR counts were elevated in high pathological tumours and associated with vascular invasion (box plot). (A) L1 5'UTR were significantly elevated in high pathological stage tumours $p=0.025$, Kruskal-Wallis Test. (B) L1 5'UTR counts were associated with vascular invasion, $p=0.015$ Mann-Whitney test. Values represent mean±SE.

Cancer recurrence can be assessed by Seoul National University (SNU) recurrence signature gene cluster. SNU classifies tumours either as low (low recurrence) or high (high recurrence). L1 5'UTR counts were significantly elevated in SNU high compared to SNU low (mean+SE Low=33.7±2.8, High=42.8±3.9 $p=0.007$ Mann-Whitney test, **Figure 2.5A**). Patient prognosis was assessed by the HB16 gene cluster and were classified either as C1 (well differentiated tumours and good prognosis) or C2 (poorly differentiated tumours and poor prognosis). Again, L1 5'UTR counts were significantly elevated in C2 group implying

advanced tumour stage, metastasis and poor prognosis (mean+SE C1=29.7±2.0, C2=50.3±5.0, p<0.001 Mann-Whitney test), **Figure 2.5B**.

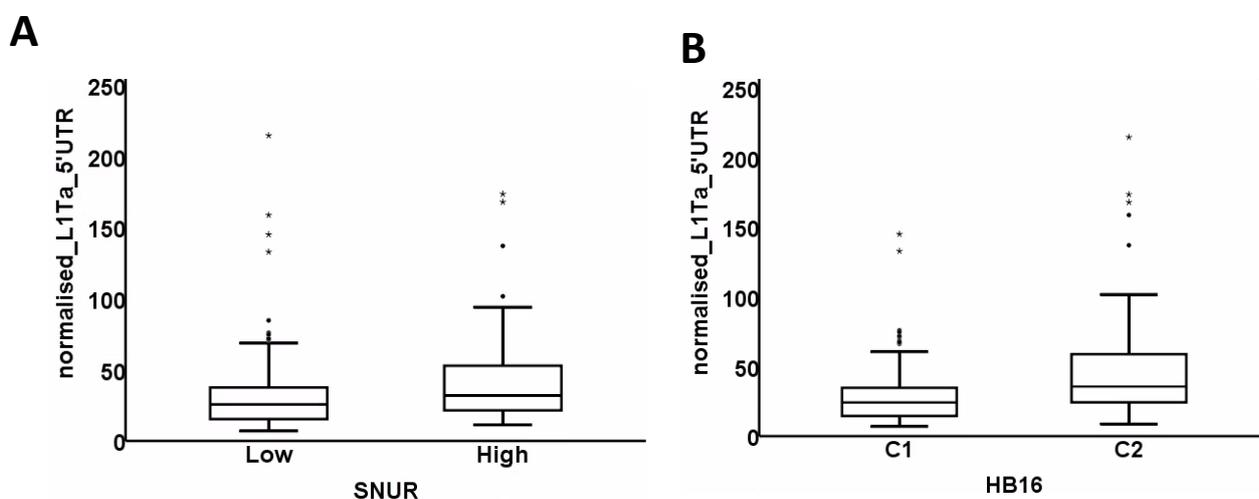


Figure 2.5. L1 5'UTR counts correlated with patient recurrence and poor survival (box plot). (A) L1 5'UTR counts were elevated in SNU high gene cluster p=0.007, Mann-Whitney test (B) L1 5'UTR counts were elevated in HB16 C2 gene cluster p<0.001, Mann-Whitney test. Values represent mean±SE.

2.4.3 L1 5'UTR counts and correlation with HCC-related signalling pathways

Several genes and pathways were altered between tumour and non-tumour samples such as TP53, CTNNB, TGFB, etc. HCC samples were classified into two groups based on their TP53 status - functional TP53 (WT) and mutated. L1 5'UTR counts were elevated significantly in TP53 mutated subgroup compared to TP53 WT (mean+SE TP53 WT:30.6±2.0 and TP53 mutated: 49.1±5.9 p<0.001 Mann-Whitney test), **Figure 2.6A**. Similarly, TP53 gene expression targets correlated negatively with L1 5'UTR counts (Spearman correlation r=-0.166, p=0.022) **Figure 2.6B**. CTNNB1, which has an important role in Wnt signalling, demonstrated no significant correlation with L1 5'UTR counts, **Figure 2.6C**.

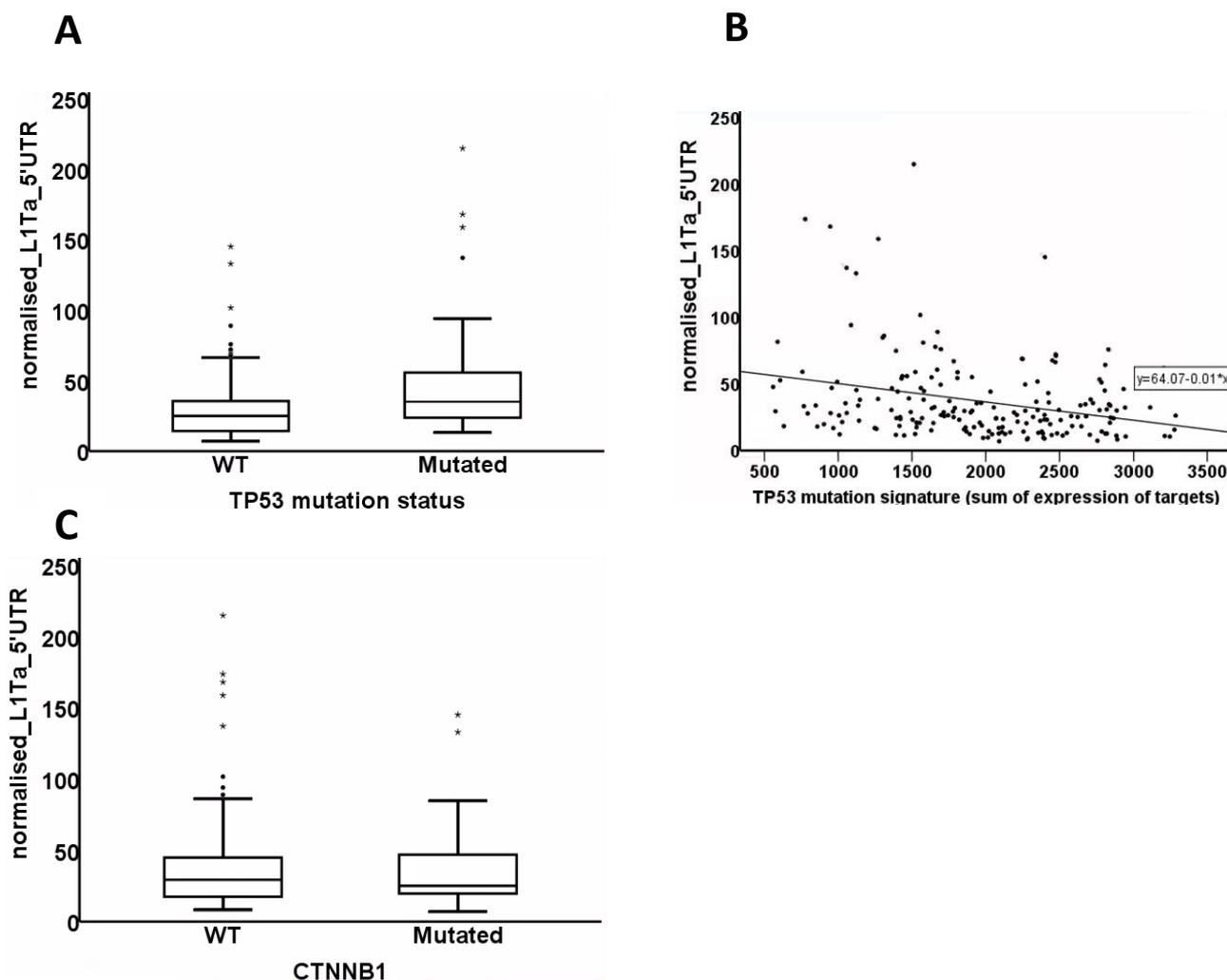


Figure 2.6 L1 5'UTR counts correlated negatively with TP53 expression. (A) L1 5'UTR counts were elevated in TP53 mutated patients TP53 WT, $p < 0.001$, Mann-Whitney test (box plot) (B) L1 5'UTR counts correlated negatively with TP53 gene expression targets $p = 0.022$, Spearman correlation (scatter plot), (C) L1 5'UTR demonstrate no significant difference between normal and mutated CTNNB1 $p = 0.689$, Mann-Whitney U test (box plot).

Furthermore, L1 5'UTR counts were significantly elevated in iCluster 3 (mean \pm SE 53.5 \pm 4.9) compared to iCluster 1 (33.3 \pm 3.4) and iCluster 2 (22.7 \pm 1.9), implying higher chromosomal instability and TP53 mutation (**Figure 2.7**). These iClusters were derived previously by integration of mRNA expression, miRNA expression, DNA copy number, DNA methylation, and RPPA (Wheeler et al, 2017).

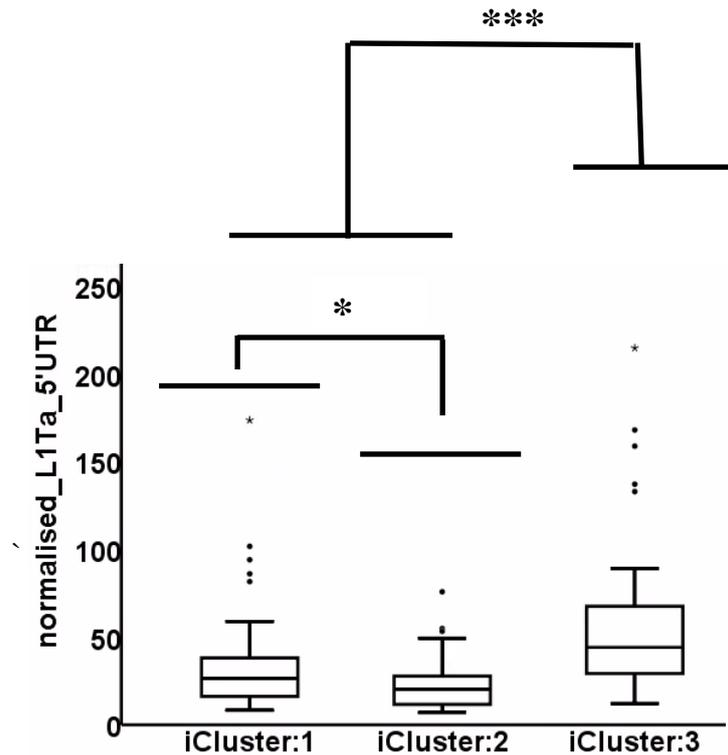


Figure 2.7 mRNA expression, miRNA expression, DNA copy number, DNA methylation, and RPPA were previously clustered into three main iClusters. L1 5'UTR counts were significantly elevated in iCluster 3 compared to iCluster 1 $p < 0.001$ and iCluster 2 $p < 0.001$, iCluster 1 and 2 $p = 0.035$, Kruskal-Wallis Test.

As TGF- β has an important role in HCC and can influence several biological pathways, the pathway was investigated in these samples. TGF- β pathway status were previously assessed and classified into the following groups: activated, normal and inactivated (Chen et al., 2018). L1 5'UTR counts positively correlated with activated TGF- β status (42.1 ± 3.1) compared to Normal (21.0 ± 2.1), and inactivated (mean \pm SE 26.1 ± 5.1), $p < 0.01$, Kruskal-Wallis Test) **Figure 2.8**.

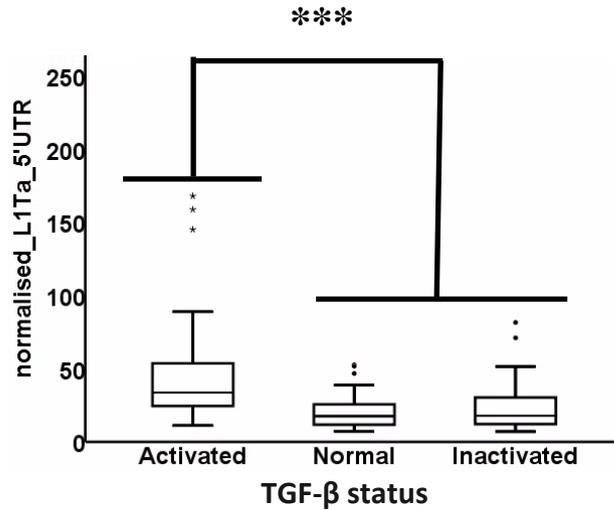


Figure 2.8 L1 5'UTR counts correlated with active TGF-β status. L1 5'UTR counts correlated significantly with activated TGF-β status compared to Normal and inactivated $p < 0.01$, Kruskal-Wallis Test.

There are different algorithms available to evaluate L1 transcripts from RNAseq data. In order to validate our method with another reported pipeline {Chen, 2018 #117} we compared the two datasets. In this study, the authors have analysed 372 HCC cases from the TCGA-LIHC dataset. Our L1Ta UTR counts exhibited a strong positive correlation with the L1HS counts ($r=0.597$, $p < 0.001$ Spearman correlation) reported in the other study for the same samples (**Figure 2.2**). Furthermore, the study also reported counts for other L1 subfamilies such as L1PA2, L1PA3 and L1PBa1. Hence, we evaluated the correlation between different L1 subfamilies. L1HS, L1PA2, L1PA3 and L1PBa1 demonstrated a strong correlation ($p < 0.001$) (**Figure 2.9**).

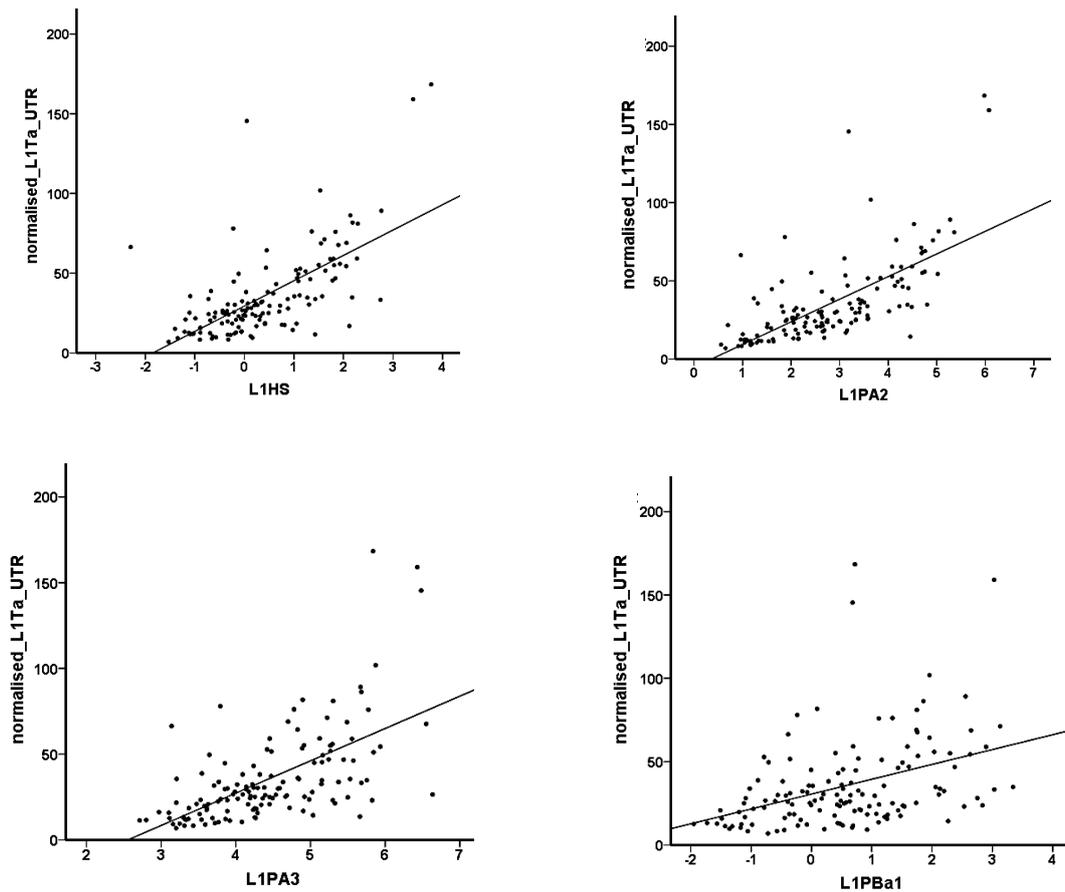


Figure 2.9 L1Ta_5'UTR correlates with L1HS ($r=0.597$, $p<0.001$), L1PA2 ($r=0.866$, $p<0.001$) L1PA3 ($r=0.684$, $p<0.001$) and L1PBa1 ($r=0.605$, $p<0.001$) Spearman correlation.

Several clinical factors were compared to L1 subgroups, and significance was observed with HB16, macrovascular invasion, TP53 mutation signature, iCluster and TGF- β cluster. These findings are consistent with findings observed earlier in which L1 5'UTR (L1HS) is associated with poorly differentiated tumours and poor prognosis (HB16 and presence of macrovascular invasion) and TGF- β activation.

Table 2.5. Clinical parameters were compared in L1HS, L1PA2, L1PA3 and L1PBA1 by measuring L1 full length transcripts. Mann-Whitney U test was performed for continuous data. iCluster and TGF- β were analysed using Kruskal-Wallis test. Error bars represent standard error.

| | L1HS | L1PA2 | L1PA3 | L1PBA1 |
|-------------------------|------------------------------------------------------------------|------------------------------------------------------------------|------------------------------------------------------------------|-----------------------------------------------------------------|
| HB16 | C1 58.1 \pm 0.2 C2 85.3 \pm 0.6 p=0.000 | C1 59.0 \pm 0.2 C2 83.8 \pm 3.4 p=0.000 | C1 60.4 \pm 0.1 C2 81.4 \pm 4.9 P=0.003 | C1 62.7 \pm 0.2 C2 77.4 \pm 0.7 P=0.035 |
| Macrovascular invasion | No 54.0 \pm 0.2 Yes 60.9 \pm 0.1 p=0.310 | No 54.1 \pm 0.2 Yes 60.6 \pm 2.6 p=0.339 | No 52.0 \pm 0.2 Yes 65.9 \pm 4.3 p=0.038 | No 54.3 \pm 0.2 Yes 60.1 \pm 0.5 p=0.394 |
| TP53 mutation signature | r= -0.296 p=0.000 | r= -0.237 p=0.006 | r= -0.241 p=0.005 | r= -0.251 p=0.003 |
| iCluster | iCluster 1 60.4 iCluster 2 56.0 iCluster 3 77.5 p=0.022 | iCluster 1 58.2 iCluster 2 54.8 iCluster 3 81.4 p=0.002 | iCluster 1 60.2 iCluster 2 55.8 iCluster 3 77.9 p=0.018 | iCluster 1 59.1 iCluster 2 56.5 iCluster 3 78.4 p=0.15 |
| TGF- β cluster | Activated 73.2 Normal 46.8 Inactivated 66.0 p=0.004 | Activated 75.2 Normal 47.7 Inactivated 54.6 p=0.001 | Activated 73.6 Normal 50.8 Inactivated 56.8 p=0.010 | Activated 69.7 Normal 52.1 Inactivated 73.3 p=0.062 |

2.4.4 Expression of L1orf1 protein in HCC tissues

L1orf1 protein (L1orf1p) expression was assessed by immunohistochemistry in 48 HCC patient samples. Tissue sample of colorectal carcinoma metastasised to liver was used for assay optimisation and as a positive control (**Figure 2.10**).

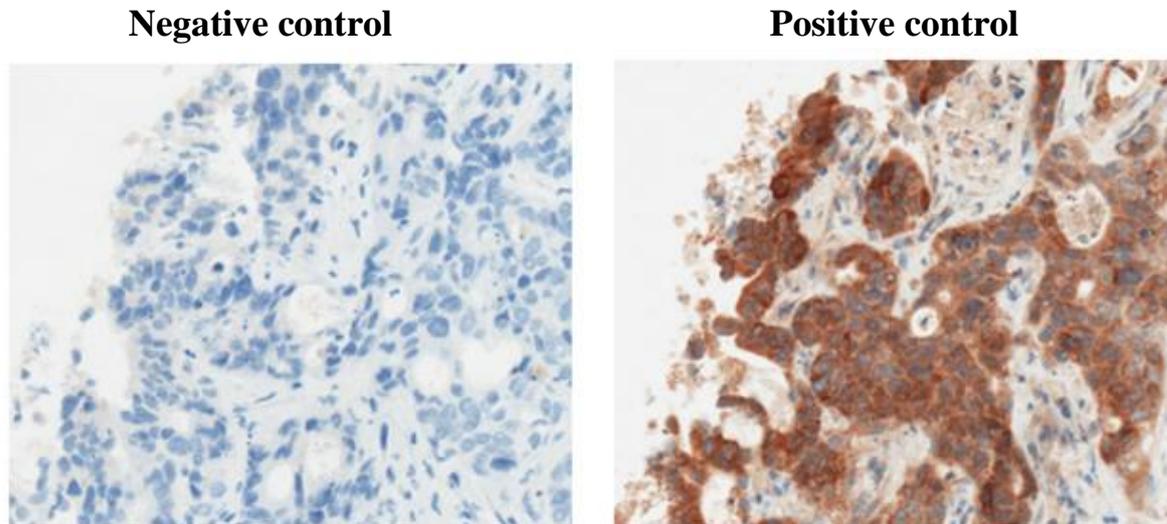


Figure 2.10 Negative and positive control for L1orf1p antibody using colorectal carcinoma tissue. Negative control was stained in the absence of primary antibody, and positive control was stained with primary and secondary antibody (200x magnification).

2.4.5 L1orf1 protein expression is elevated in HCC

L1orf1p level was assessed and compared between tumour and corresponding non-tumour liver samples. L1orf1p stain was predominantly cytoplasmic. Stained slides were scored based on staining intensity and percentage positivity, as negative or grade 1 to 5 (**Figure 2.12**). Tumour biopsies had significantly elevated L1orf1p staining intensity compared to non-tumour biopsies (**Figure 2.11**) ($p=0.001$, Wilcoxon signed ranks test). Based on this scoring, samples were classified into the following groups: L1_low (score 0-2) and L1_high (score 3-5) for tumour; L1_absent (0), L1_low (1), L1_high (2-4) for non-tumour for exploring the correlations between L1 expression and clinical features.

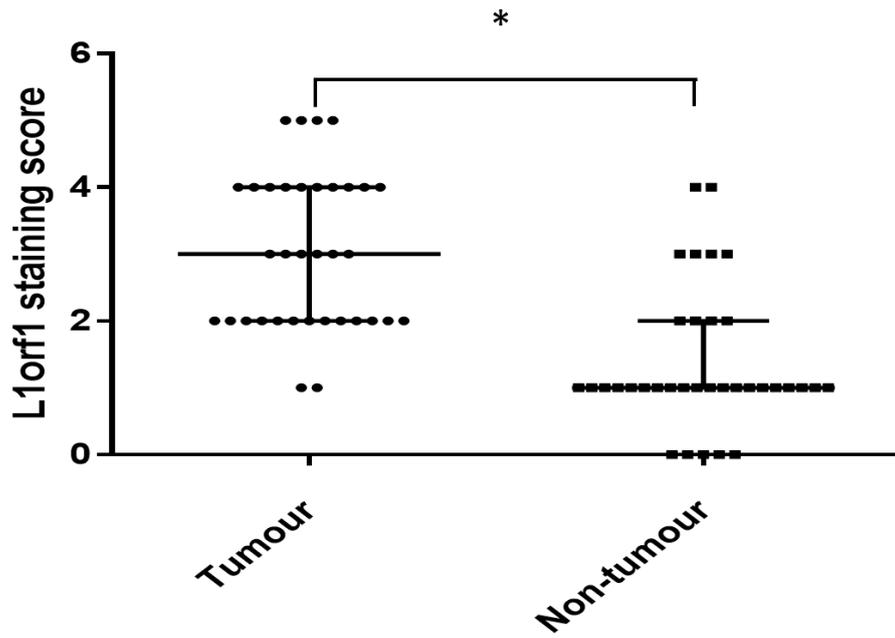


Figure 2.11 L1orf1p levels in tumour and corresponding non-tumour samples (box plot). Tumour tissues had significantly elevated L1orf1p staining intensity compared to non-tumour tissues (p=0.001, Wilcoxon signed ranks test).

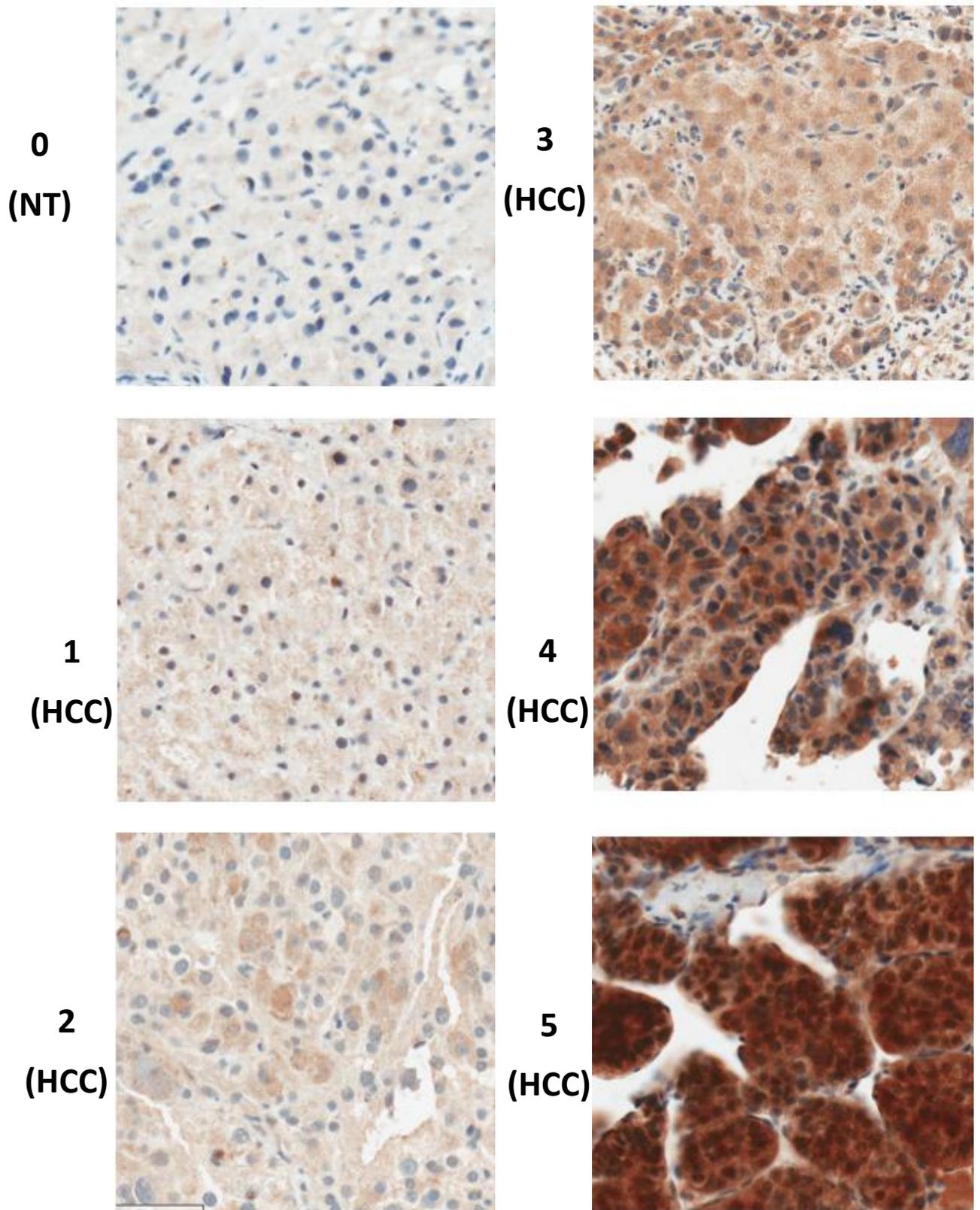


Figure 2.12 FFPE tumour and non-tumour biopsies were stained for L1orf1p using immunohistochemistry. Stained slides were assessed and scored based on staining intensity and percentage positivity: 0 (NT tumour case) or grade 1 to 5 (HCC tumour cases), x200 magnification.

2.4.6 L1 protein expression correlated with poorly differentiated tumours and increased AFP

HCC L1orf1p expression status (L1_low versus L1_high) were compared with different clinical factors (**Table 2.5**). No correlations were observed between L1orf1p and patient age ($p=0.838$), gender ($p=0.52$), type two diabetes ($p=0.34$), cirrhosis ($p=0.643$) and BCLC stage ($p=0.72$). However, high L1orf1p expression correlated with poorly differentiated tumours ($p=0.036$, Chi-square test). Patient liver function indicators such as albumin and bilirubin levels were not significantly different between the 2 groups however, AFP level was significantly elevated in L1_high compared to L1_low (L1_low 6ng/ml vs L1_high 39ng/ml, $p=0.039$, Mann–Whitney U test).

2.4.7 L1 high correlated with poor survival in TACE treated patients

The median survival was measured using Kaplan Maier analysis between L1_low and L1_high patient groups. No significant difference was observed between L1_low (19.3mo \pm 2.9, n= 24) and L1_high (13.2 mo \pm 5.0, n=26) $p=0.756$. Similar findings were observed in patients excluding resection patients L1_low (19.3mo \pm 2.9, n= 24) and L1_high (13.2 mo \pm 5.0, n=26) $p=0.756$. However, a significant difference in survival was observed in patients with TACE treatment (L1_low 20.89 months vs L1_high 10.8 months, $p=0.025$, Log-rank Mantel-Cox test) **Table 2.5** and **Figure 2.13**. However, the number of patients are very low and the finding needs to be validated in a larger cohort.

Table 2.5. Clinical parameters were compared between L1_low and L1_high patients in tumour samples. Chi-square test was performed for categorical data and Mann-Whitney U test for continuous data. Survival analysis was performed using Kaplan Meier analysis.

| Clinical Features | | N (%) | | p-value |
|--------------------------------------------|--------------------------------------|------------------|-----------------|--------------|
| | | L1-low 22 (46) | L1-high 26 (56) | |
| Age | Median [range] | 69 [55-85] | 71 [49-80] | 0.838 |
| Gender | Male/Female (F%) | 17/05 (23) | 20/06 (23) | 0.977 |
| Risk Factor | NAFLD/ALD/Viral/Others/No known risk | 12/1/0/3/6 | 10/4/0/5/6 | 0.273 |
| BMI | Median [range] | 33.9 [25.4-42.0] | 25 19-45.6.0] | 0.021 |
| Diabetes | No/Yes (Y%) | 7/15 (68) | 11/15 (58) | 0.454 |
| Cirrhosis | No/Yes (Y%) | 13/9 (41) | 16/10 (39) | 0.863 |
| Ascites | No/Yes (Y%) | 12/1(8) | 14/2(13) | 0.672 |
| Albumin (g/l) | Median [range] | 40[31-67] | 39[28-48] | 0.431 |
| Bilirubin (µmol/l) | Median [range] | 10[5-28] | 14[4-39] | 0.186 |
| AFP (ng/mL) | Median [range] | 6[1-632] | 29.5[1-50000] | 0.032 |
| Tumour and Histopathologic Features | | | | |
| Tumour number | Median [range] | 1 [1-5] | 1.55 [1-20] | 0.45 |
| Tumour size | Median [range] | 3.6 [1.3-18] | 4[1-18] | 0.329 |
| PVT | No/Yes (Y%) | 16/6 (27) | 19/8 (30) | 0.978 |
| EHD | No/Yes (Y%) | 18/4(18) | 22/3(12) | 0.553 |
| TNM Stage | I/II/III/IV (IV%) | 6/2/7/5 (25) | 10/4/8/3 (12) | 0.637 |
| Grade Well Differentiated | No/Yes (Y%) | 7/15 (68) | 18/7 (28) | 0.046 |
| BCLC Stage | A/B/C/D | 4/5/12/1 | 4/5/17/1 | 0.709 |
| Therapy | OLTx/Res/Ablation/TACE/Med/ BSC | 0/3/0/12/2/5 | 0/2/7/12/0/4 | 0.057 |
| Survival post biopsy Months median | | | | |
| All patients | N = 48 | 18.2± 3.0 | 17.5± 4.4 | 0.701 |
| Treated patients | N = 36 | 16.0± 4.0 | 9.8± 3.1 | 0.281 |
| TACE treated | N = 24 | 19.3± 3.0 | 9.8 ± 3.1 | 0.036 |

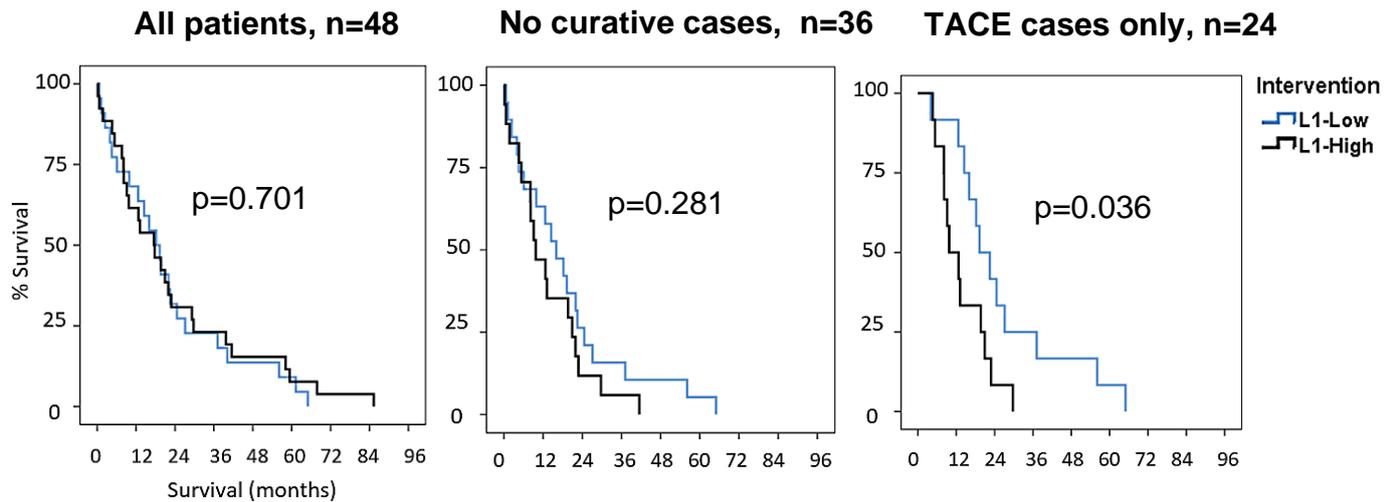


Figure 2.13 Median survival between L1_low and L1_high in tumour (Kaplan Meier curve) **(A)** All patients, L1_low (18.2mo \pm 3.0, n= 22) and L1_high (17.5mo \pm 4.4, n=26), p=0.756 **(B)** all patients except ablation or resection, L1_low (16.0mo \pm 4.0, n= 19) and L1_high (9.8mo \pm 3.1, n=17), p=0.249 **(C)** TACE patients only, L1_low (19.3mo \pm 3.0, n= 12) and L1_high (9.8 mo \pm 3.1, n=12), p=0.036 Log-rank Mantel-Cox test.

Likewise, to explore significance of L1 expression in non-tumour tissues, non-tumour cases were categorised into the following groups: L1_absent, L1_low and L1_high; however, no significant correlation were observed with any of the parameters except cirrhosis. A positive association between liver cirrhosis and L1orf1p expression in non-tumour tissues was observed (**Table 2.6**).

Table 2.6. Clinical parameters were compared between L1_low and L1_high patients in non-tumour samples. Chi-square test was performed for categorical data and Mann-Whitney U test for continuous data. Survival analysis was performed using Kaplan Meier analysis.

| Clinical Features | | N (%) | | | p-value |
|-------------------------------------|--------------------------------------|--------------------|----------------|----------------|--------------|
| | | L1-absent 5(14) | L1-low 23 (62) | L1-high 9 (24) | |
| Age | Median [range] | 61 [52-70] | 73 [56-85] | 71 [41-76] | 0.263 |
| Gender | Male/Female (F%) | 3/2(40) | 19/04 (17) | 8/1 (11) | 0.398 |
| Risk Factor | NAFLD/ALD/Viral/Others/No known risk | 4/0/0/0/1 | 13/2/0/0/3 | 2/4/0/0/1 | 0.116 |
| BMI | Median [range] | 37.8[30-45.58] | 28[23.07-37] | 32[25-40] | 0.438 |
| Diabetes | No/Yes (Y%) | 2/3 (60) | 10/13 (57) | 3/6 (67) | 0.871 |
| Cirrhosis | No/Yes (Y%) | 4/1 (20) | 15/8 (35) | 2/7 (78) | 0.046 |
| Ascites | No/Yes (Y%) | 2/0(0) | 12/1(8) | 6/1(14) | 0.724 |
| Albumin (g/l) | Median [range] | 42.5[39-45] | 41.5[28-67] | 39[34-47] | 0.469 |
| Bilirubin (µmol/l) | Median [range] | 7[5-41] | 11.5[4-39] | 13[7-26] | 0.601 |
| AFP (ng/mL) | Median [range] | 3 | 16[1-18000] | 4[2-2733] | 0.198 |
| Tumour and Histopathologic Features | | | | | |
| Tumour number | Median [range] | 1[1-2] | 1[1-4] | 1[1-20] | 0.848 |
| Tumour size | Median [range] | 3.7±[2-19] | 4[1.3-18] | 3.7[2-11.5] | 0.509 |
| PVT | No/Yes (Y%) | 4/1 (20) | 18/5 (22) | 8/1 (11) | 0.880 |
| EHD | No/Yes (Y%) | 4/1(20) | 19/4(17) | 9/0(0) | 0.390 |
| TNM Stage | I/II/III/IV (IV%) | 2/0/1/1 (25) | 8/3/5/3 (16) | 3/3/2/0 (0) | 0.680 |
| Grade-Well Differentiated | No/Yes (Y%) | 27/19 (41) | 9/12 (57) | 18/7 (28) | 0.046 |
| BCLC Stage | A/B/C/D | 8/9/29/2 | 4/4/12/1 | 4/5/17/1 | 0.945 |
| Therapy | OLTx/Res/Ablation/TACE/Med/ BSC | 0/3/0/0/1/1 | 0/4/3/11/3/2 | 0/2/3/4/0/0 | 0.196 |
| Survival post biopsy Months median | | | | | |
| All patients | N = 36 | 40.1± 20.4 | 19.3± 5.3 | 17.5± 6.4 | 0.607 |
| Treated patients | N = 21 | No patients | 12.7 ± 2.9 | 9.1 ± 2.3 | 0.535 |
| TACE treated | N = 15 | No patients | 14.5 ± 4.1 | 9.1± 2.3 | 0.123 |

2.4.8 L1 correlated with TGF-β activation

As TCGA HCC transcript analysis demonstrated a significant correlation between L1 5'UTR counts and active TGF-β status (**Figure 2.8**), we decided to validate the finding in our FFPE HCC samples at protein level. TGF-β activation was measured by measuring its downstream target pSMAD3 specifically the Ser 425 phosphorylated SMAD3. pSMAD3 protein expression were examined by immunohistochemistry in a cohort of 17 HCC FFPE patient samples for which L1orf1p status has already been determined (L1_low n=8, L1_high n=9).

Slides were scored and classified into the following groups based on staining intensity by a pathologist: None, low, and high. pSMAD3 was expressed in the nucleus only and highly expressed in HCC tumour cells compared to non-tumour cells (**Figure 2.14A**). pSMAD3 expression demonstrated a positive association with L1orf1p expression in HCC ($r=0.502$, $p=0.04$, $n=17$, Spearman correlation). (**Figure 2.14B**)

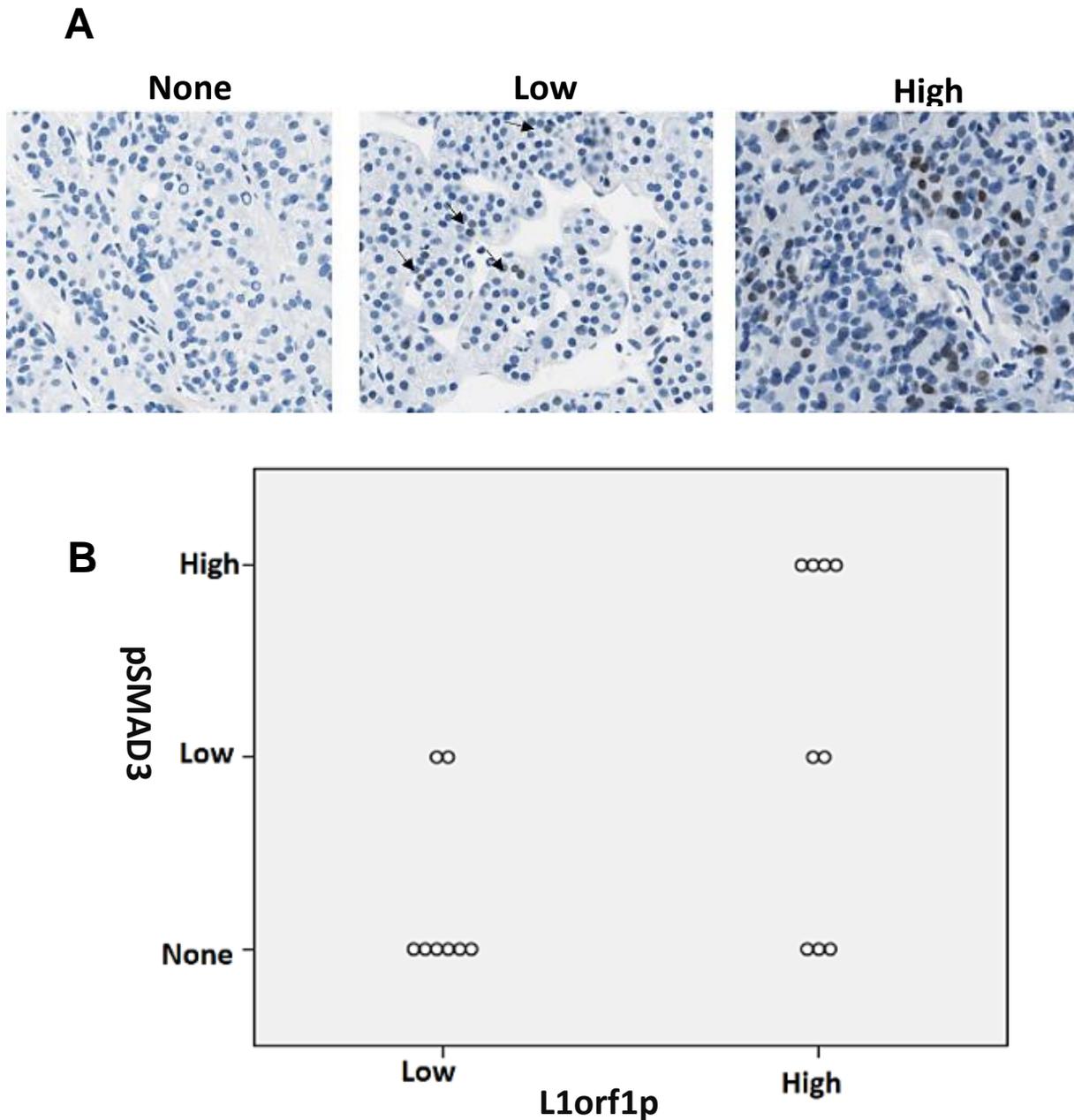


Figure 2.14 FFPE HCC tumour and non-tumour biopsies were stained for pSMAD3 by immunohistochemistry. (A) Stained slides were assessed and scored based on staining intensity and percentage positivity as None, Low or High. (B) pSMAD3 correlated significantly with L1orf1p FFPE staining ($r=0.502$, $p=0.04$, $n=17$, Spearman correlation).

2.5 Discussion

In this chapter, TCGA HCC patient samples were re-analysed to include L1Ta 5'UTR promoter (L1 5'UTR) counts. In addition, L1 subfamilies L1HS, L1PA2, L1PA3 and L1PA4 exhibited strong correlation with each other. L1 counts were elevated in high pathological stage tumours and correlated with vascular invasion. Vascular invasion has been linked with satellite nodule, intrahepatic metastases and portal vein obstruction leading to liver damage and failure (Shi et al., 2010). Furthermore, L1 counts were associated with high tumour recurrence (SNU cluster) and poor prognosis (HB16 cluster). These findings suggest that elevated L1 may contribute to tumour progression and poor prognosis in HCC patients.

Several signalling pathways were also elevated and associated with high L1 counts. Particularly, L1 counts were elevated in iCluster 3 implying higher chromosomal instability and were positively correlated with TP53 mutation and activated TGF- β signalling. TP53 is an important cell cycle regulator and its reduced expression leads to increased cell cycle progression. Similarly, TGF- β signalling has an important role in HCC and has both tumour suppressive and promoting functions. During late stage carcinogenesis, tumour cells can inhibit TGF- β mediated tumour-suppressive functions and increase tumour promoting factors including the induction of EMT. Therefore, L1 may support tumorigenesis by increasing the TGF- β pathway in a background of mutated TP53.

Besides L1 transcripts, we confirmed L1orf1 protein expression in HCC biopsies and overall L1orf1p was elevated in HCC tumours compared to non-tumour tissues. L1orf1p was present in the cytoplasm, but low nuclear positivity was also observed. Similar findings were reported in *in vitro* studies using retrotransposition assay and patient FFPE biopsies (Belgnaoui et al., 2006, Rodic et al., 2014). However, some studies have shown a correlation between L1orf1p nuclear staining and key clinical parameters. For example, in invasive breast carcinoma, L1orf1p and L1orf2p nuclear localisation correlated with metastasis and poor survival (Chen et al., 2012). L1orf1p and L1orf2p nuclear localisation might be more prevalent in certain cancers and correlate with a worse prognosis.

L1 expression (transcripts as well as L1orf1p) was also significantly associated with poorly differentiated tumours, and poor survival in patients receiving TACE therapy. These data

suggest that L1 is significantly elevated in tumour tissues and might support resistance to TACE therapy. As well as being associated with early mortality, L1 expression correlated with other indicators of adverse outcome including elevated AFP levels. During liver injury or hepatocarcinogenesis, hepatocyte release AFP into the circulation (Steiner et al., 1966). Elevated AFP levels have been observed in 60% of HCC patients (Taketa, 1990). Cirrhosis is also another important risk factor for HCC and present in 80% of HCC patients. Here, L1 expression positively correlated with cirrhosis in non-tumour tissues. Previous studies have shown that several HCC related aetiologies have late diagnosis of HCC. For example, in NAFLD, 22.8% related HCC were detected by surveillance. In alcohol related HCC, it was 32% and in Hepatitis C it was 46.2% (Mittal et al., 2015). These data suggest that L1 might be a potential biomarker particularly in early stage of cancer such as in cirrhotic patients.

In terms of cell signalling pathways, L1 transcript expression positively correlated with TP53 mutation and TGF- β signalling in HCC. The association between TGF- β signalling and L1 expression was further verified at the protein level as well using pSMAD3 as a surrogate for TGF- β signalling activation. This is in line with previous reports where L1 expression is found to be positively correlated with TP53 mutation in other epithelial cancers such as lung carcinoma, pancreatic carcinoma, ovarian carcinoma and secondary glioblastoma (Rodic et al., 2014). There are reports showing an association between L1 expression and TGF- β signalling and indicates either regulation of L1 by TGF- β (Reyes-Reyes et al., 2017) or regulation of TGF- β by L1 (Zhu et al., 2013). Hence, the relationship between L1 and TGF- β signalling can be context dependent. There can be a number of potential explanations that could explain the association between L1 and TGF- β signalling activation in HCC: (i) TGF- β signalling activates L1 expression, (ii) L1 drives activation of TGF- β signalling, (iii) there is a feedback loop between TGF- β signalling and L1, or (iv) these are independent events and are only non-causally linked in cancer cells through chance. To answer this systematically, we utilised HCC-related cell line models described in Chapter 2 and 3.

Summary

L1 was elevated in tumours both at transcript and protein level in HCC patients. High L1 correlated with poorly differentiated advanced stage tumours with vascular invasion and poorer patient outcome. Both transcript and protein analysis demonstrated a significant

association between L1 and TGF- β signalling. Further studies are required to investigate the mechanism of how L1 influences TGF- β signalling and explore its potential as a treatment stratification marker and/or therapeutic target.

Chapter 3: Effect of L1 knockdown on Huh7 cells

3.1 Introduction

The transforming growth factor β (TGF- β) pathway has an important role in HCC and can affect several cellular processes such as proliferation, migration and invasion. Previously in Chapter 1, analysis of RNAseq data and FFPE biopsies demonstrated a positive relationship between L1 expression and activated TGF- β signalling in HCC patients. However, it remains unclear if there is a causal relationship between L1 and TGF- β signalling in HCC.

3.11 TGF- β signalling pathway

The TGF- β superfamily consist of 30 different structural regulatory proteins, which includes TGF- β , activins, BMP, inhibins, growth and differentiation factors (Padua and Massague, 2009). The TGF- β ligand is an important molecule in this pathway. It is a polypeptide cytokine and has three main isoforms: TGF β 1, TGF β 2 and TGF β 3. TGF β 1 is the most abundant isoform and regulates several cellular functions such as cell proliferation and apoptosis (Massague, 1998, Heldin et al., 1997).

Active TGF- β ligand has a dimer configuration and binds to both Type 1 and Type 2 transmembrane serine/threonine kinase receptors (**Figure 3.1**). Upon binding, these two receptors form a heteromeric complex. The Type 2 receptor has constitutively active kinases, which phosphorylate the Type 1 cytoplasmic juxtamembrane glycine-serine domain. Subsequently Type 1 receptor phosphorylates two serine residues located at the end of its carboxyl-terminal in the cytoplasmic domain. Phosphorylation of these sites creates specific binding sites for SMAD proteins (R-SMAD) particularly SMAD2 and SMAD3. The adaptor protein SARA facilitates the binding. Both SMAD2 and SMAD3 are then phosphorylated by the two receptors and dissociate from it. These two proteins form a heterotrimeric complex with SMAD4 and translocate to the nucleus. SMAD4 accumulates in the nucleus and interacts with several transcription factors and target genes (Tang et al., 2018).

The TGF- β ligand can also activate non-canonical pathways, which do not involve SMAD proteins. For example, it can activate the Erk/MAPK (PI3K)/Akt and MAPK cell signalling pathways (Moustakas and Heldin, 2005).

The TGF- β pathway is usually inhibited by SMAD6 and SMAD7. Both SMAD6 and SMAD7 inhibit TGF- β signalling but affect different points of the pathway. For example, SMAD6 inhibits BMP signalling and SMAD1–Co-SMAD interaction. SMAD7 interferes with Type 1 receptor and blocks ligand binding and thus the receptor cannot get activated. It can also recruit the E3 ubiquitin ligase Smurf to the receptor and target it for degradation (Ebisawa et al., 2001).

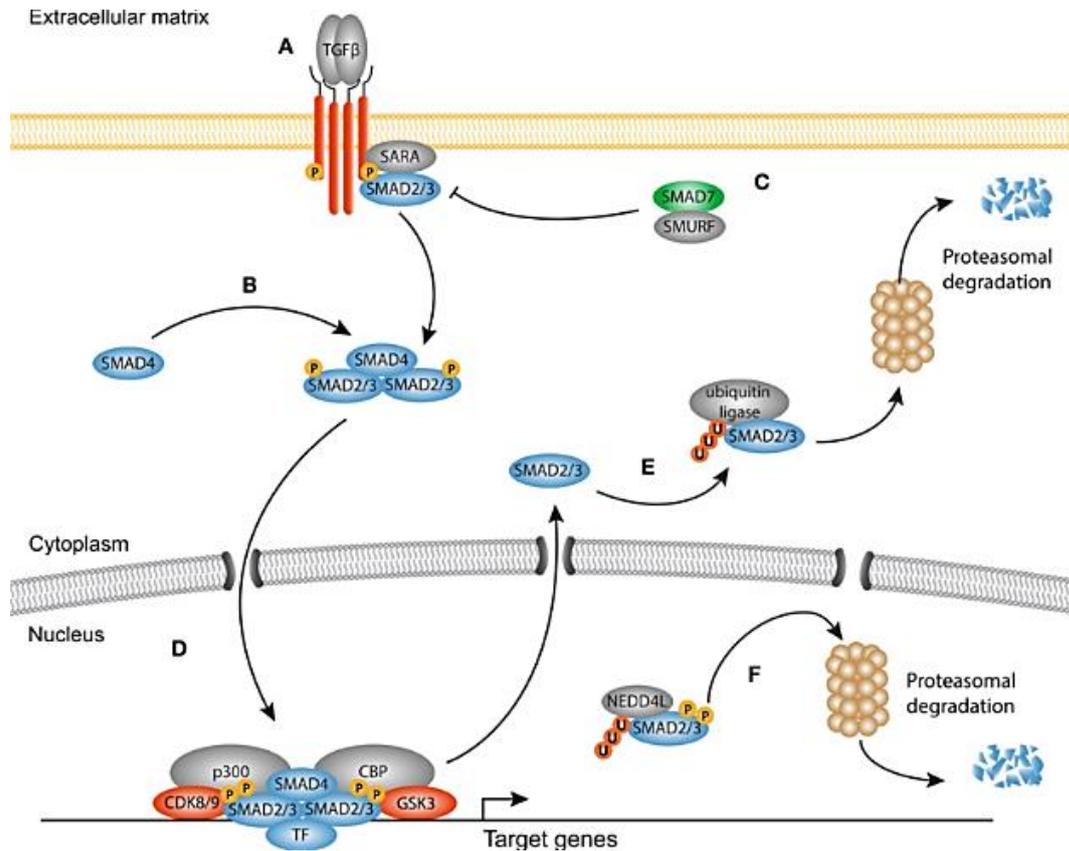


Figure 3.1 TGF- β pathway (Piersma et al., 2015).

3.12 TGF- β and HCC

The TGF- β pathway has an important role in the liver by controlling liver architecture and regeneration, but can also contribute to pathological condition such as fibrosis, cirrhosis and HCC (Karkampouna et al., 2012).

TGF- β can have different functions and can act either as a tumour suppressor or tumour promoter (**Figure 3.2**). In early stages of HCC, TGF- β acts as a tumour suppressor and inhibits c-myc and certain cyclin dependent kinase inhibitors. In turn, it reduces cell

proliferation and induces cell apoptosis. As HCC progresses, TGF- β can act as an autocrine and paracrine molecule, and also activate stromal fibroblasts. Stromal fibroblasts can activate regulatory T and cancer stem cells. Regulatory T cells are a subtype of T-cells, and have an immunosuppressive effect on the immune system (Giannelli et al., 2014). Cancer stem cells are a subpopulation of tumour cells which can self-renew and are resistant to many conventional therapies. Both regulatory T and cancer stem cells contribute to tumorigenesis (Wu et al., 2012, Gallimore and Simon, 2008).

In late stages of HCC, tumour cells become less responsive to the TGF- β suppressive effects by inhibiting key genes in the pathway such as T β RII. T β RII expression has shown to decrease particularly in malignant cancers. In malignant hepatocytes, T β RII expression was reduced by more than 25% compared to surrounding non-malignant tissues (Yamazaki et al., 2011). Similar observations were seen in cell lines with late TGF- β response. These cell lines lacked the T β R1 receptor and had low level of T β RII receptors (Yamazaki et al., 2011, Matsuzaki, 2013, Nagata et al., 2009). T β RII receptor mutations are particularly associated with a poor patient prognosis (Yamazaki et al., 2011).

TGF- β dysregulation can affect several pathways such as epidermal growth factor (EGF) and Platelet-derived growth factor (PDGF) mediated pathways. Both pathways are mitogenic signalling pathways and affect cell survival and proliferation. EGF has an important role in cell survival and stimulates proliferation via the MAPK/ERK and PI3K-Akt pathways. PDGF leads to the accumulation of nuclear β -catenin leading to increased cell proliferation. Both EGF and PDGF are elevated in late stages of HCC and can inhibit TGF- β mediated tumour suppressive functions and increase tumour promoting functions.

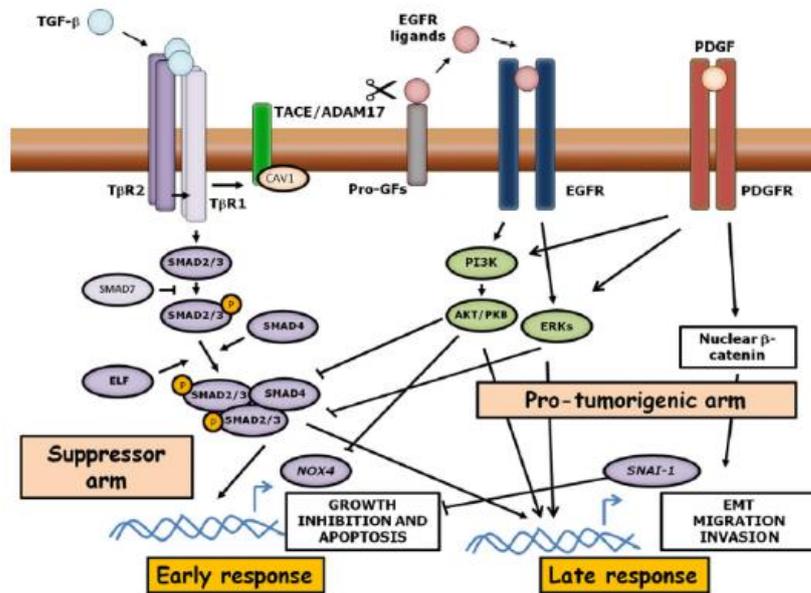


Figure 3.2 TGF- β pathway in early and late carcinogenesis (Fabregat et al., 2016). In early carcinogenesis the pathway has tumour suppressive effects but has tumour supporting effects in late carcinogenesis.

EMT can also increase during late stages of HCC. A meta-analysis has shown that the TGF- β pathway converges on several pro-EMT inducers such as transcription factors (TWIST1, TWIST2), transcriptional repressors SNAIL (SNAI1), SLUG (SNAI2), ZEB1 and ZEB2. Collectively these were overexpressed in 49.6% of 1334 HCC patients (Wan et al., 2017). EMT markers were also increased at protein level such as Twist1 (60.3% of patients), Snail (51.9% of patients), Zeb1 (43.6% of patients) and Zeb2 (50.3% of patients).

TGF- β pathway can also increase the expression of Axl through its ligand Gas6 in malignant hepatoma cells. Axl is a tyrosine kinase and high level of Axl has been associated with increased invasion, tumour recurrence and poor survival in HCC patients. Axl can also interact with 14-3-3 zeta. 14-3-3 zeta can increase the expression of several TGF- β mediated EMT genes such as PAi1, SNAIL and MMP-9 in HCC, which are associated tumour progression (Reichl et al., 2015). Similarly, TGF- β downregulation decreased migration and invasion in HCC cell lines (Fabregat and Giannelli, 2017).

Several mouse model studies have also investigated a correlation between TGF- β dysregulation and HCC. For example, one study investigated the role of TAK1 and HCC using TAK-1 null mice (Yang et al., 2013a). TAK1 is a negative regulator of the TGF- β pathway and has two mechanisms of inhibition. It can interfere with R-SMAD activation

by binding to the MH2 domain of SMAD proteins (Hoffmann et al., 2005). TAK1 can also increase SMAD7 expression and thus decrease TGF- β signalling (Dowdy et al., 2003). Interestingly, TAK-1 null mice developed fibrosis and HCC, indicating increased TGF- β signalling led to increased HCC. Furthermore, two double knockout mice models - TAK-1 with SMAD4 and TAK-1 with T β RII receptor - had lesser incidence of fibrosis and HCC, particularly in TAK1 and T β RII receptor knockout mouse model. Furthermore, SMAD4 and T β RII receptor null mutation had also reduced tumour incidence possibly due to decreased TGF- β signalling (Yang et al., 2013a). In another study, TP53 knockout mice had increased expression of TGF- β 1, Pail and Afp. Furthermore, mice with TP53 and T β RII knockout had reduced incidence of HCC and cholangiocarcinoma (Morris et al., 2012).

3.13 L1 and TGF- β

Relationship between L1 and TGF- β signalling has been reported in literature. Reyes-Reyes et al. have demonstrated activation of L1 in human bronchial epithelial cells upon TGF- β treatment while knockdown of L1 using L1orf1 targeting siRNA did not alter TGF- β pathway. Thus, L1 may act downstream of the TGF- β pathway (Reyes-Reyes et al., 2017). On the other hand, Zhu et al., 2013, have shown that L1orf1p can also affect the TGF- β pathway in HepG2. An immunoprecipitation analysis has shown that ORF1p forms a complex with Smad4. The formation of the complex inhibited Smad4 nuclear translocation. The observations indicate that the relationship between L1 and TGF- β pathway might be different in cancerous versus non-cancerous cells. Hence, in this chapter we focused on the primary question – if L1 has potential to activate TGF- β pathway in HCC using stable knockdown of L1 in an HCC cell line.

3.2 Specific aims

- Characterisation of L1 expression in different HCC cell lines and identify relevant cell line for L1 knockdown
- Stable knockdown of L1 using lentivirus-based shRNA constructs in selected cells and evaluate the role of L1 in cell proliferation, migration and invasion
- RNAseq analysis of Huh7-L1KD and Huh7-NT (control) cell lines to evaluate influence of L1 knockdown on the cellular transcriptome
- Validation of RNAseq findings especially modulated signalling pathways by functional assays like luciferase reporter assay and at protein level by western blotting.

3.3 Materials and Methods

3.31 Reagents

SDS lysis buffer: Tris HCL (pH 6.8) 12.5ml, 2g SDS, 10ml glycerol, 67.5ml distilled water

Loading dye (western blot): 2.5ml 0.5M Tris/HCL pH 6.8, 0.4g SDS, 1ml β -mercaptoethanol, 2ml glycerol, 1ml 0.1% bromophenol blue, 13.5ml distilled water

Running buffer (western blot) 1000ml distilled water, 14.4g glycine, 3g Tris base, 1g SDS

Transfer buffer (western blot): 800ml distilled water, 200ml methanol, 0.3g Tris base and 1.41g glycine

TBST: 1000ml distilled water, 2.42g Tris base, 8g NaCl, 6, 1ml Tween20 and pH=7.

Carnoy fixative: 3 parts acetic acid, 1 part methanol

3.32 Cell Culture

HepG2, Hep3B, SK-Hep1, Huh1, Huh7, SNU182, SNU475, and PLC/PRF5 were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS). HHL5 was grown in DMEM-high glucose media supplemented with 10% FBS, 1% L-Glutamine and 1% non-essential Amino Acid Solution. Cells were incubated at 37°C and 5% CO₂, and sub-cultured every 6 days.

3.33 Huh7-L1KD lentiviral particles

pLKO.1 lentival-based plasmids containing either non-target-shRNA (CAACAAGATGAAGAGCACCAA) or L1orf1-shRNA (AAATGAAGCGAGAAGGGAAGT) were obtained from Prof Gael Critofari, Institute for Cancer Research and Aging, Nice, France (Pizarro and Cristofari et al., 2016). Lentiviral particles were prepared as follows: 500 μ l serum free DMEM/F-12 media were mixed with the following plasmids: pLKO.1 shRNA plasmid (6 μ g), envelope plasmid MD2.G (1.5 μ g), packaging plasmid PAX2 (4.5 μ g) and TransIT-LT1 transfection reagent (36 μ l, Geneflow) and incubated for 30 minutes at room temperature. HEK293T were then transfected with these plasmids (cells were previously seeded at 4×10^6 in a 10cm² petri dish and left for 24 hours to attach). Culture supernatant was collected at 48 and 96 hours post-transfection. Condition media was then centrifuged at 2000g for 5 minutes and filtered

through a 0.45µM filter to remove cell debris. Cleared supernatant containing the packaged lentiviral particles (non-targeted or Huh7-L1KD) were then stored at -80°C.

3.34 Transduction into Huh7 cell line

Wild-type Huh7 (Huh7-WT) cells were seeded in a 6 well plate at 2×10^5 per well. The next day, cell media was removed from cells and then transduced with 50µl non-target or Huh7-L1KD lentiviral supernatant in presence of 8µg/ml polybrene. 450 µl complete media (DMEM/F-12 media+10% FBS) was also added to each well to make a final volume of 500µl per well. After 24 hours, cell media was replaced. The next day, cell media was replaced with media containing 2µg/ml puromycin. Puromycin and media were replaced every 3 days. Selection was performed for 2 weeks before confirming status of knockdown by RT-qPCR and western blot.

3.35 Western blot

Cell pellets were isolated using trypsin and frozen dry at -20°C. Pellets were thawed and solubilised in 50µl SDS lysis buffer. Lysates were then heated at 100°C for 10 minutes followed by sonication for 3 minutes (30sec ON and 30 sec OFF, 3 rounds). After sonication, samples were centrifuged at 13000g for 10 minutes. The supernatant was then taken further for protein estimation. Protein estimation was performed using the Pierce BCA Protein Assay Kit (ThermoFisher). After estimation, protein samples were adjusted to 30µg and a loading dye was added at a 1 to 1 ratio. Samples were heated for 10 minutes at 100°C and then loaded onto a 4–15% Mini-PROTEAN TGX Precast Protein Gels (Biorad). Gel electrophoresis was performed at 100V for 90 minutes, and wet transfer was performed using a nitrocellulose membrane for 60 minutes at 100V.

After transfer, nitrocellulose membrane was blocked with 5% bovine serum albumin (diluted in TBST) for 1 hour at room temperature. The membrane was incubated overnight with the primary antibody at 4°C and then washed three times with TBST, 5 minutes each. The following primary antibodies were used: L1orf1p clone 4H1, 1:1000 (merckmillipore); GAPDH 1:5000 (Sigma); Vimentin 1:1000 (Abcam); SMAD3 1:1000 (Abcam). The membrane was then stained with the appropriate secondary antibodies (Dako) at room temperature for 1 hour and visualised using the ECL kit (GE Healthcare).

3.36 FACS

Cells were harvested from a 70% confluent T75 flask with trypsin, and then fixed in 1% formalin (diluted in PBS) overnight. Following fixation, cells were washed in wash-perm buffer (BD bioscience) twice at 300g for 7 minutes and stained with L1orf1p clone 4H1 1:4000 (merckmillipore) for 1 hour at room temperature. Cells were then washed twice with wash-perm buffer (BD bioscience) and incubated with anti-mouse secondary antibody Alexfluor 488, 1:8000 (Invitrogen) for 1 hour. After secondary incubation, cells were washed with wash perm buffer and PBS, and analysed on the FACS Calibur.

3.37 Cell cycle by PI staining

Cells were harvested from a 70% confluent flask with trypsin and fixed in cold 70% ethanol (4°C) for at least 24 hours. Samples were then centrifuged at 450g for 7 minutes and the supernatant was discarded. To each sample, the following reagents were added: 234µl PBS, 10µl propodium iodide solution (1mg/ml, Sigma), 6µl RNase A (1mg/ml, Sigma) and then incubated for 20 minutes at room temperature in the dark. After incubation, samples were analysed on the FACS Calibur (BD bioscience).

3.38 SMAD3 luciferase assay

Huh7-WT, Non-target-shRNA (Huh7-NT) and Huh7-L1 knockdown (Huh7-L1KD) were seeded at 0.5×10^5 in a 12 well plates in triplicates. The next day, 200µl serum free media were incubated with the following plasmids (one well): 400ng pai1-luc reporter plasmid, 40ng Renilla plasmid and 1.2µl LT1 transfection reagent and incubated for 20 minutes at room temperature. After incubation, cells were transfected, incubated for 48 hours at 37°C, and then washed with PBS and frozen dry at -20°C. Dual-Luciferase Reporter Assay System (Promega) was used to perform the luciferase assay. Plate was thawed to room temperature and 100µl of 5x lysis buffer was added to each well, and then incubated in a plate shaker for 15 minutes at room temperature. 20µl cell lysate was removed and added onto a well of a 96 well white polystyrene microplate plate. 100µl luciferase reagent was added to each well and analysed on the Omega Plate reader for luminescence. Subsequently, 100µl stop and glow solution was added to measure Renilla luminescence. Samples were then normalised to Renilla.

3.39 RNA isolation

RNA was isolated using the Qiagen RNeasy kit according to the instructions. RNA was eluted in 30µl nuclease-free water and concentrations were quantified using the ND-1000 Spectrophotometer. Purified RNA was further treated with the Turbo DNase-I kit as instructed by the manufacturer (Invitrogen).

3.310 Reverse transcription

Reverse transcription was performed using the Promega reverse transcription system. 1µg of RNA was diluted in 10µl nuclease-free water. RNA was heated at 70°C for 10 minutes and placed on ice. The following mastermixes were added for oligo(dT) 15 Primer and gene specific primer cDNA conversions:

- The following reaction mixture were added for **oligo (dT)15 Primer cDNA conversion**: MgCl₂ 25mM (4µl), Reverse transcription buffer 10X (2µl), dNTP Mixture 10mM (2µl), Recombinant RNasin Ribonuclease Inhibitor (0.5µl), AMV Reverse transcriptase (0.5µl) and oligo(dT)15 Primer (1µl). 10µl reaction mixture was added onto each 10 µl RNA sample. The combined 20µl reaction mixture was incubated at 42°C for 15 minutes, 95°C for 5 minutes and 4°C for 5 minutes. 80µl nuclease-free water was added and cDNA was stored at -20°C.
- The following reaction mixture were added for **gene specific primer cDNA conversion**: MgCl₂ 25mM (4µl), Reverse transcription buffer 10X (2µl), dNTP Mixture 10mM (2µl), Recombinant RNasin Ribonuclease Inhibitor (0.5µl), AMV Reverse transcriptase (0.5µl) and L1-RT (1µl of 100µM), 18S-RT and TBP RT primers (1µl of 100µM). 10µl of the reaction mixture was added onto each 10 µl RNA sample. The combined 20µl reaction mixtures were incubated at 42°C for 30 minutes, 50°C for 30 minutes, 95°C for 5 minutes and 4°C at 5 minutes. 80µl nuclease-free water was added to each sample and stored at -20°C.

3.311 Real-time PCR

qPCR was performed in a 10µl reaction per well using the following reagents: 5µl Platinum™ SYBR™ Green qPCR SuperMix-UDG with ROX (Thermofisher), 2.8µl

Nuclease-Free Water, 0.2µl of primer (forward and reverse, 10µM), 2µl cDNA. Samples were loaded onto a 384 well plate and analysed on the QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems). The qPCR cycle conditions are shown in Table 1. qPCR primer sequences are shown in Table 2.

Table 3.1 qPCR conditions

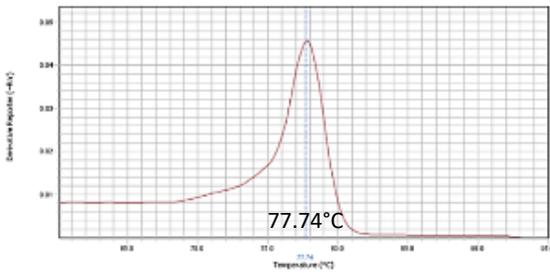
| qPCR condition | | | | | | | | |
|----------------------|-----|------|-----------------------|-----|-------|------------------|-----|--------|
| Hold stage(1 cycle) | | | PCR stage (40 cycles) | | | Melt curve stage | | |
| 50°C | for | 2min | 95°C | for | 10min | 95°C | for | 15 sec |
| | | | | | | 60°C | for | 1min |
| | | | | | | | | 95°C |
| | | | | | | | | 15 sec |

Table 3.2 qRT-PCR primer sequences

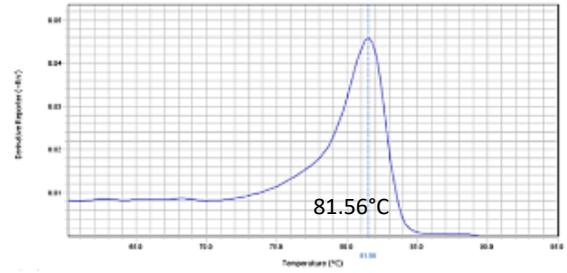
| Gene | Type of Oligo | Sequence of oligo |
|---------|---------------|---------------------------|
| L1'5UTR | Forward | GAATGATTTTGACGAGCTGAGAGAA |
| L1'5UTR | Reverse | GTCCTCCCGTAGCTCAGAGTAATT |
| L1orf1p | Forward | AGTGCTTAAAGGAGCTGATGG |
| L1orf1p | Reverse | GCTGATACCCTTTCTTCCAGTT |
| L1orf2p | Forward | CAAACACCGCATATTCTCACTCA |
| L1orf2p | Reverse | CTTCCTGTGTCCATGTGATTCA |
| TBP | Forward | GCAAGGGTTTCTGGTTTGCC |
| TBP | Reverse | GGGTCAGTCCAGTGCCATAA |
| HPRT | Forward | TTGCTTTCCTTGGTCAGGCA |
| HPRT | Reverse | ATCCAACACTTCGTGGGGTTC |
| SMAD2 | Forward | TGCTCTGAAATTTGGGGACTGA |
| SMAD2 | Reverse | CGACCATCAAGAGACCTGGTT |
| SMAD3 | Forward | GAGGAGAAATGGTGCAGAGAA |
| SMAD3 | Reverse | GCGGCAGTAGATGACATGAG |
| SMAD4 | Forward | TTTGAGGGACAGCCATCGTT |
| SMAD4 | Reverse | ATACTGGCAGGCTGACTTGTG |
| SMAD7 | Forward | TCCTCGGAAGTCAAGAGGCT |
| SMAD7 | Reverse | TGGACAGTCTGCAGTTGGTT |
| Epcam | Forward | GAACACTGCTGGGGTCAGAA |

| | | |
|--------------|---------|-----------------------|
| Epcam | Reverse | CTGAAGTGCAGTCCGCAAAC |
| Vimentin | Forward | TGCAGGAGGCAGAAGAATGG |
| Vimentin | Reverse | AAGGGCATCCACTTCACAGG |
| P15 (CDKN2B) | Forward | GGGACTAGTGGAGAAGGTGC |
| P15 (CDKN2B) | Reverse | CATCATCATGACCTGGATCGC |

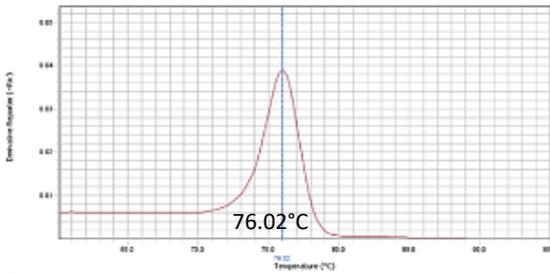
L1 5'UTR



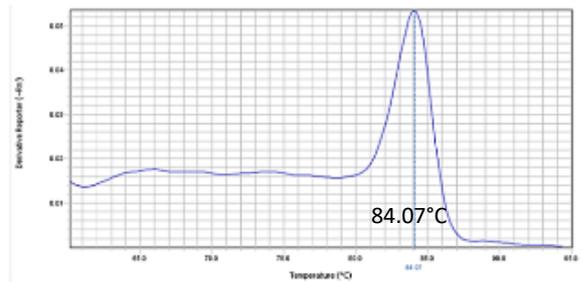
ORF1p



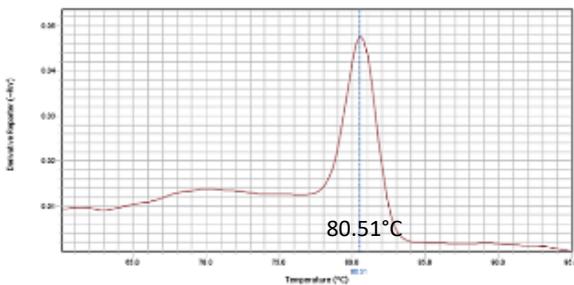
ORF2p



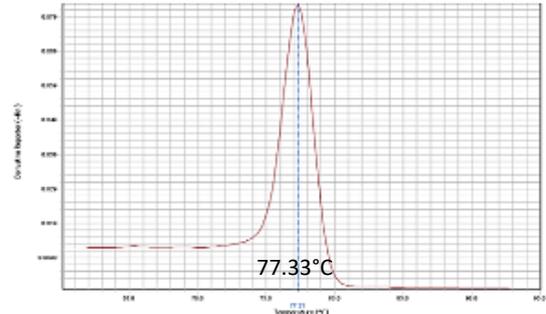
TBP



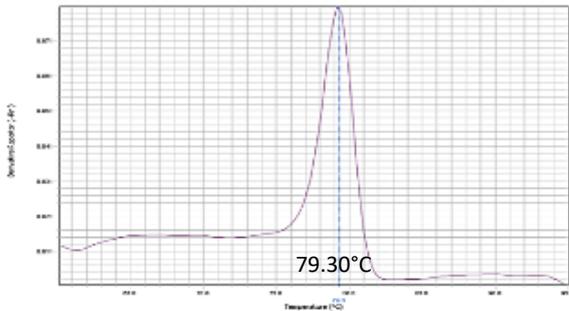
HPRT



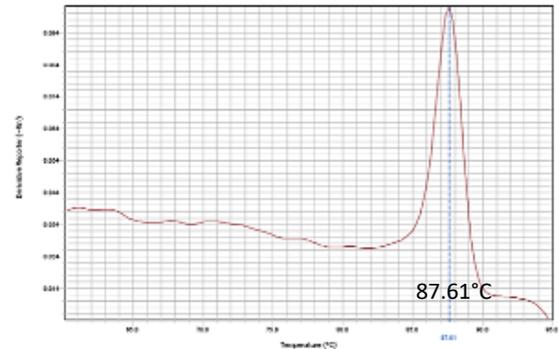
TGFβ



SMAD2



SMAD3



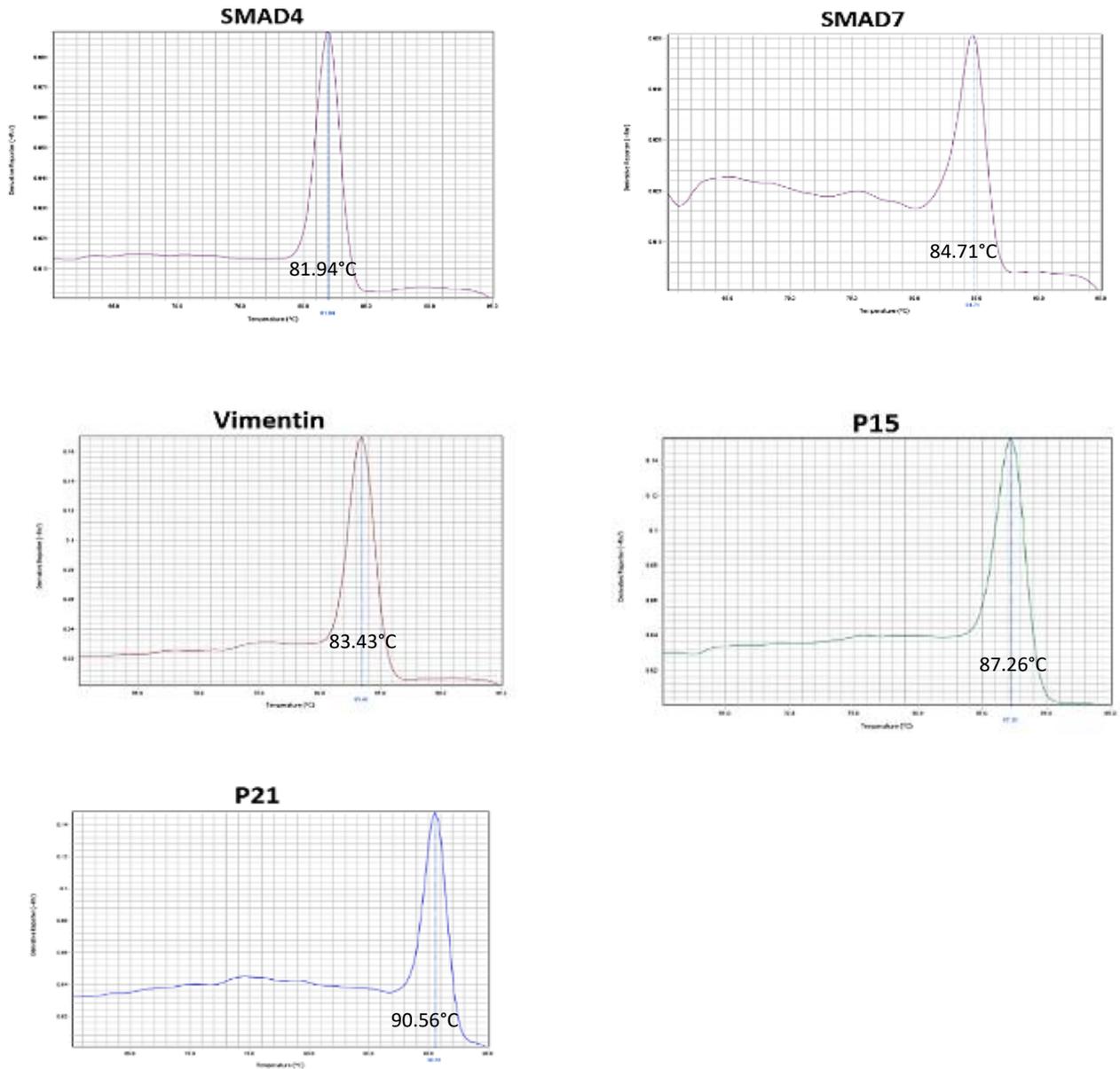


Figure 3.3 Melting temperature of different qPCR primers. Each primer had a single peak indicating no primer dimers.

Table 3.3 Primer efficacy of L1 5'UTR, L1orf1p, L1orf2p, TBP and HPRT

| Gene | Slope | Efficiency | r ² |
|----------|---------------|------------|----------------|
| L1 5'UTR | -3.40 ± 0.043 | 97.0% | 0.99 |
| L1orf1p | -3.24 ± 0.031 | 103.5% | 0.99 |
| L1orf2p | -3.40 ± 0.055 | 96.8% | 0.99 |
| TBP | -3.18 ± 0.033 | 106.1% | 0.97 |
| HPRT | -3.48 ± 0.089 | 93.8% | 0.99 |

3.312 Cell proliferation

Incucyte

Huh7-WT, Huh7-NT and Huh7-L1KD cells were seeded onto a 96 well plate, 500 cells per well. Images were taken every 6 hours using incucyte imaging. Cell proliferation was then measured by a cell specific algorithm.

MTT

Huh7-WT, Huh7-NT and Huh7-L1KD cells were seeded onto a 96 well plate, 500 cells per well. Plates were treated with MTT at the following time points: day 2, 4, 6, 8 and 10. For each well, 20µl MTT reagent was added and then incubated for 5 hours at 37°C in the dark. Subsequently, media was removed and 150µl DMSO was added onto each well. The plate was incubated for 10 minutes at room temperature in a shaker. Readings were then taken at 570nm.

SRB

Huh7-WT, Huh7-NT and Huh7-L1KD cells were seeded onto a 96 well plate, 500 cells per well. Cells were incubated with 200µl complete media and incubated at the following time points: day 1, 2, 3, 4, 5, 6 and 7. Afterwards, cells were fixed with 50µl Carnoy's fixative for 24 hours at 4°C. The plate was then rinsed 5 times with deionised water and air-dried. 200µl sulforhodamine B solution (0.4% diluted in 1% acetic acid) was added to each well and incubated for 40 minutes. Following the incubation, the plate was rinsed 5 times with 1% acetic acid and air-dried. 200 µL of 10 mM Tris base was added to each well and incubated for 30 minutes at room temperature in a shaker. Readings were then taken at 570nm.

3.313 Clonogenic assay

Huh7-WT, Huh7-NT and Huh7-L1KD were seeded at 4000 and 6000 cells in a 10cm² Petri dish in triplicates. Plates were incubated at 37°C and 5% CO₂ for 21 days, and then washed with PBS and fixed with cold Carnoy fixative for 10 minutes. Plates were washed with distilled water and stained with 0.4% crystal violet (diluted in distilled water) for 30 minutes. Plates were then washed with distilled water and colony size and number were

counted using an automatic colony counter. Colonies smaller than 0.5mm were excluded from the analysis.

3.314 Invasion assay

Invasion assay was performed using Boyden chambers (Merckmillipore). Huh7-WT, Huh7-NT and Huh7-L1KD were seeded in 1.0×10^6 cells/ml serum free media. 300µl cell suspension was added at the top of the chamber. On the bottom of the chamber, 500µl complete media was added and incubated for 72 hours. After incubation, non-invasive cells were removed from inside of the chamber using a cotton stick. Invasive cells on the outside of the chamber were stained with 500 µl staining solution for 20 minutes. Residual staining was removed by gently submerging in distilled water. Phase-contrast images were taken and quantified using Image J.

3.315 RNA seq

RNA isolation

Total RNA was isolated from Huh7-NT and Huh7-L1KD from 3 biological repeats. Sample integrity was checked using a bioanalyser and all samples had good quality.

Library preparation and RNA seq

Samples were then sent for sequencing to the Genomic Facility, Newcastle University. Illumina Tru-seq paired end strand-specific sequencing was performed using the NextSeq50.

Bioinformatics

Post trimming quality control was performed with FastQC (version: 1.0.0). The resulting FastQ files were mapped on to the human reference genome using RNAseq alignment tool (V1.1.1) on Illumina BaseSpace software. DESeq2 analysis (version: 1.1.0) was performed to identify differentially expressed genes in Huh7-L1KD compared to Huh7-NT. Differentially expressed genes were defined as those with a p value of less than 0.05. Gene Set Enrichment Analysis (GSEA) was carried out using Broad institute's GSEA software.

3.316 Mouse xenograft experiment

Huh7-WT, Huh7-NT and Huh7-L1KD cells were injected subcutaneously into 8 weeks old nude mice. 5 mice were used for each cell line and injected with 5×10^6 cells.

3.317 Statistics

Statistical analysis was performed using SPSS (Version 24) and Graph Prism (Version 8). Independent t-test was performed for migration assay, One Way ANOVA for invasion assay.

3.4 Results

3.4.1 L1 expression in various liver cancer cell lines

L1 expression was analysed in the following liver cancer cell lines : HepG2, Hep3B, Huh1, Huh7, SNU182, SNU475, PLC/PRF5; SK-Hep1 (hepatic adenocarcinoma cell line, endothelial origin) and HHL5 (human hepatocyte cell line, (Clayton et al., 2005).

Transcript analysis was performed by RT-qPCR using primers targeting different L1 regions: L1-5'UTR, ORF1 and ORF2 and normalised against TBP and HPRT expression. As expected, all three primers had a strong correlation with each other (L1 5'UTR and ORF1($r=0.946$, $p=0.0001$); L1 5'UTR and ORF2 ($r=0.868$, $p=0.002$); ORF1 and ORF2 ($r=0.928$, $p=0.000314$). Overall, L1 expression showed large variation amongst the cell lines, being highest in Huh7-WT and Hep3B, medium expression was observed in Huh-1, HepG2, SK-Hep1, SNU182 and SNU475, and lowest expression was observed in HHL5 and PLC/PRF5 (**Fig 3.4A-4B**).

Furthermore, full length L1 transcripts were measured using cDNA, prepared using L1 3'UTR specific primer and carrying out PCR using primers targeting L1 5'UTR. Relative L1 5'UTR transcript expression were significantly lower compared to when using oligo dt primers in all cell lines (**Figure 3.4C**). However, both oligo dt and gene-specific cDNA had a similar distribution.

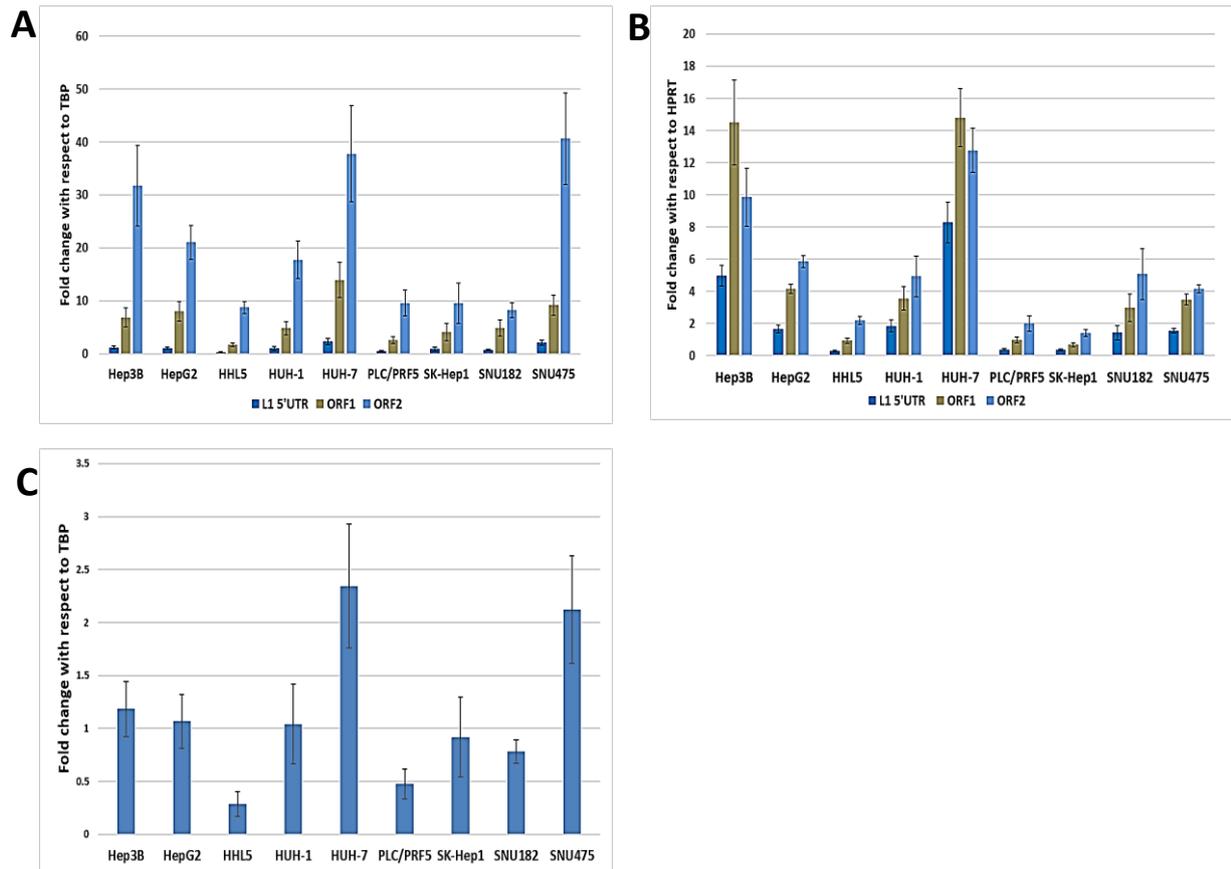


Figure 3.4 L1 transcript expression in different HCC related cell lines: **(A)** L1 5'UTR, ORF1 and ORF2 transcript expression using oligo (dT)15 cDNA conversion, normalised to TBP **(B)** L1 5'UTR, ORF1 and ORF2 transcript expression using oligo (dT)15 cDNA conversion, normalised to HPRT, **(C)** L1 5'UTR transcript expression using gene specific cDNA conversion, normalised to TBP.

Next, L1 protein expression was analysed by measuring L1orf1p levels using western blotting and FACS. L1orf1p was strongly expressed in Huh7, SK-Hep1, SNU182, SNU475, moderately expressed in SNU182 and Huh-1, and low expression was observed in HHL5 and HepG2 (**Figure 3.5**). FACS staining was performed by measuring L1orf1p positive cells (%) and mean fluorescence intensity. Both variables showed a similar trend and results corroborated with western blotting data. Huh7 had the highest L1orf1p expression and lowest expression were observed in HepG2, HHL5 and Huh1 (**Figure 3.6, Table 3.4**). Hence, overall there was good corroboration between transcript and protein expression data in the cell lines with an exception in Hep3B and HepG2 cells, which exhibited high expression of L1 transcript level but very low protein expression.

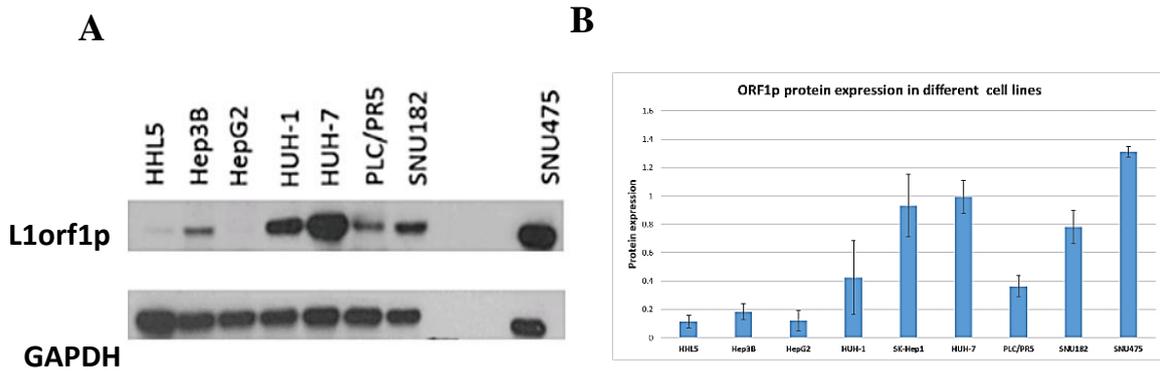


Figure 3.5 L1orf1p expression in different HCC related cell lines. Protein expression was analysed by western blotting. **(A)** Representative image of western blotting **(B)** western blot quantification by image J, normalised to GAPDH (n=3). Error bars represent standard error.

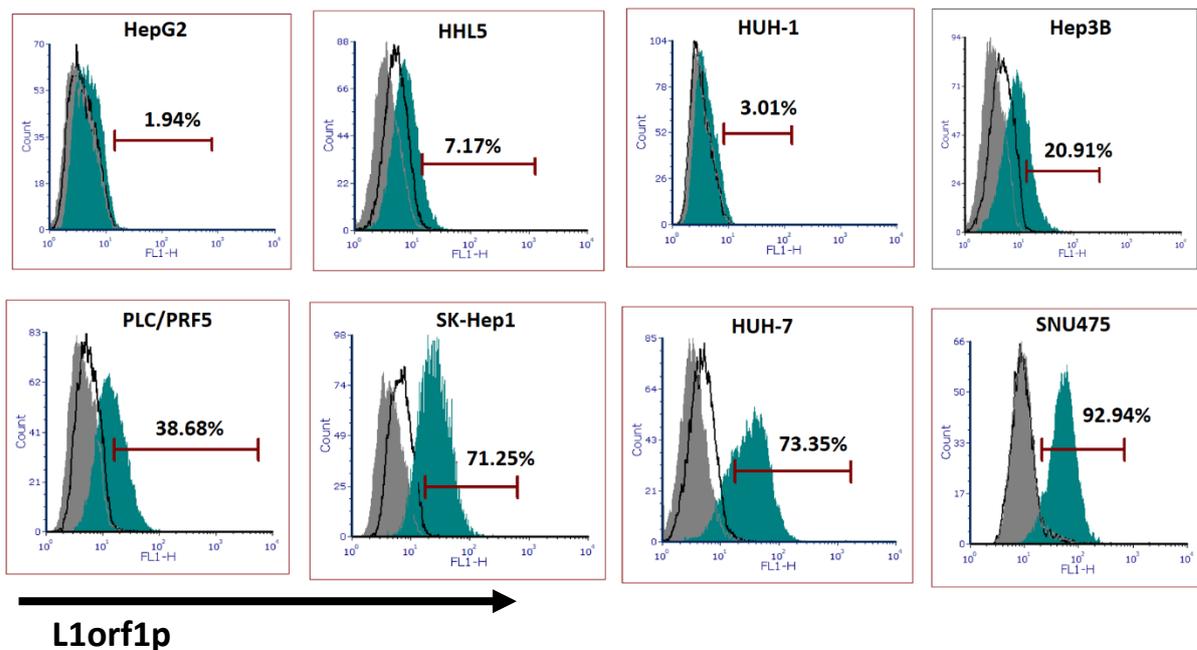


Figure 3.6 L1orf1p FACS staining in different HCC related cell lines. Grey lines represent unstained, black lines represent secondary staining only and green line represent primary and secondary antibody staining. L1orf1p expression was determined by setting the gate at 0.5% of secondary cells only.

Table 3.4 L1orf1p mean fluorescence intensity (MFI) and expression (%) in different HCC related cell lines.

| Cell line | ORF1p MFI | ORF1p expression % |
|-----------|-----------|--------------------|
| HUH-1 | 3.90 | 3.01 |
| HepG2 | 4.70 | 1.94 |
| HHL5 | 7.30 | 7.17 |
| Hep3B | 8.90 | 20.91 |
| PLC/PRF5 | 13.80 | 38.68 |
| SK-Hep1 | 24.60 | 71.25 |
| HUH-7 | 28.60 | 73.35 |
| SNU475 | 49.30 | 92.94 |

3.4.2 Generation of Huh7-L1 knockdown (Huh7-L1KD) and Huh7-non-target (Huh7-NT) control cell lines

Since, Huh7 cells exhibited highest transcript and protein expression, this cell line was selected to generate L1-knockdown transgenic cells to decipher the role of L1 in liver cancer cells. NT and L1KD lentiviral constructs were transduced into Huh7-WT cells and selected with puromycin (2µg/ml) for two weeks before confirming L1 knockdown.

Knockdown of L1 in Huh7-L1KD was confirmed both at transcript and protein levels. Transcript analysis was performed using primers previously optimised for an active L1 region in Huh7 cells (**Figure 3.7A**). As expected, L1 transcripts were downregulated by 7-fold in Huh7-L1KD compared to Huh7-WT and Huh7-NT ($p=0.0096$, One Way ANOVA). Moreover, L1orf1p expression was reduced to undetectable levels in Huh7-L1KD compared to Huh7-WT and Huh7-NT when measured by western blotting (**Figure 3.7B**). FACS analysis also confirmed the downregulation of L1orf1p expression in Huh7-L1KD (37.50%) compared to Huh7-WT (92.18%) and Huh7-NT (92.01%), (**Figure 3.7C-7D**). Hence, L1 knockdown in Huh7-WT was successful, and Huh7-L1KD cells were taken forward for functional analysis.

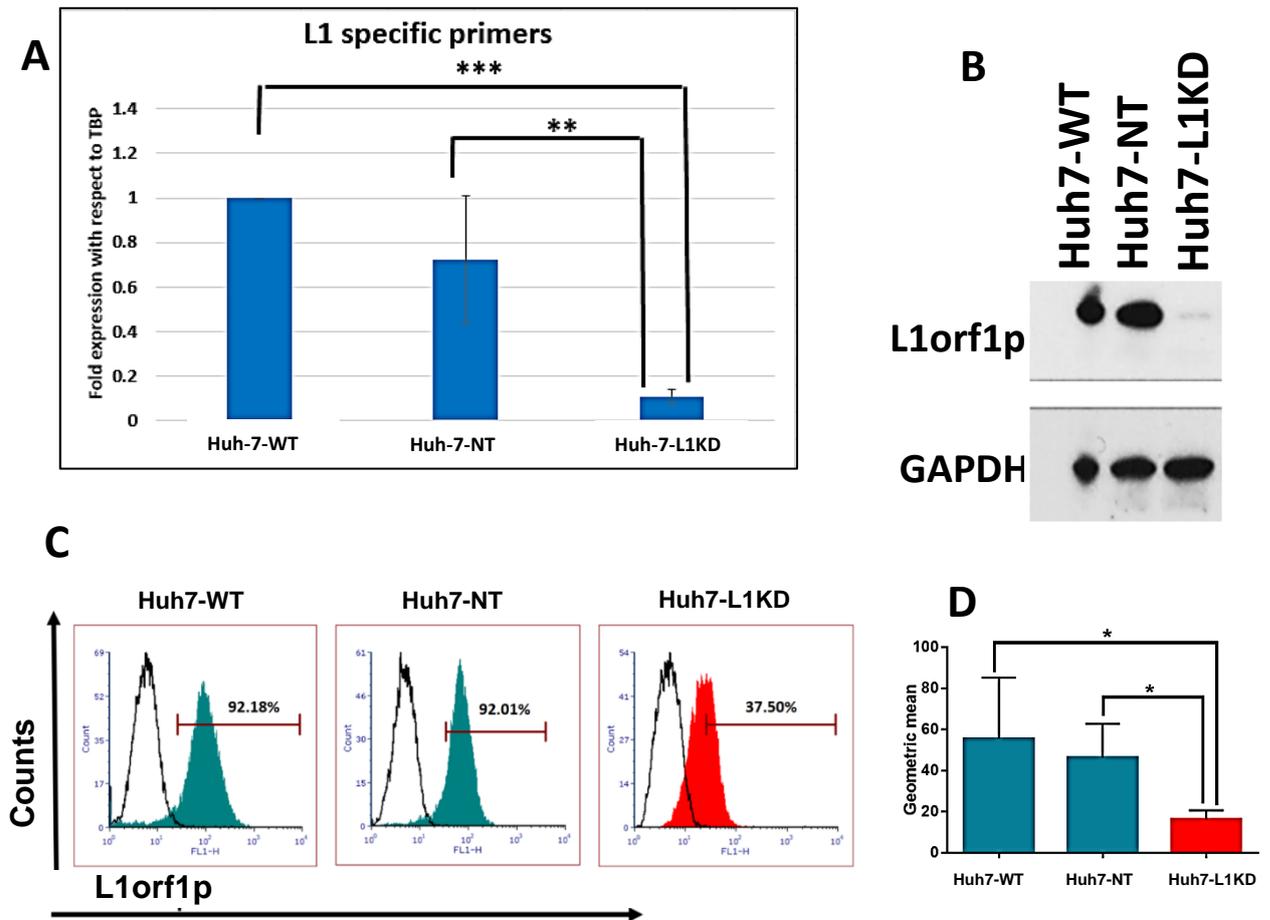


Figure 3.7 Huh7-WT, Huh7-NT and Huh7-L1KD transcript and protein expression (A) L1 transcript expression using active L1 specific primers, Huh7-WT vs Huh7-L1KD $p=0.0001$, Huh7-NT vs Huh7-L1KD $p=0.0096$ (One way ANOVA). Error bars represent standard error ($n=4$) (B) Western blot of L1orf1p and GAPDH, (C) L1orf1p FACS staining. Grey lines represent unstained, black lines represent secondary staining only, and green or red line represent both primary and secondary antibody staining. L1orf1p expression was determined by setting the gate at 0.5% of secondary positive cells only. (D) L1orf1p FACS geometric mean: Huh7-WT ($n=3$) vs Huh7-NT ($n=6$) $p=0.7486$, Huh7-WT vs Huh7-L1KD ($n=6$) $p=0.0255$, Huh7-NT vs Huh7-L1KD $p=0.0463$, One Way ANOVA.

3.4.3 Effect of L1 knockdown on phenotypic and functional properties of cells

No striking difference in cell morphology was observed in Huh7 cells upon L1 knockdown when analysed under phase contrast microscope and by Incucyte imaging (Figure 3.8). Cell proliferation was measured by Incucyte imaging, sulforhodamine B and MTT (Figure 3.9). The doubling time was calculated using Incucyte imaging. Measurements were taken from the exponential phase ($n=3$). Huh7-WT had a doubling time of 60 hours, Huh7-NT

and Huh7-L1KD had a doubling time of 41 hours Hence, Huh7-NT and Huh7-L1KD had a similar proliferation rate.

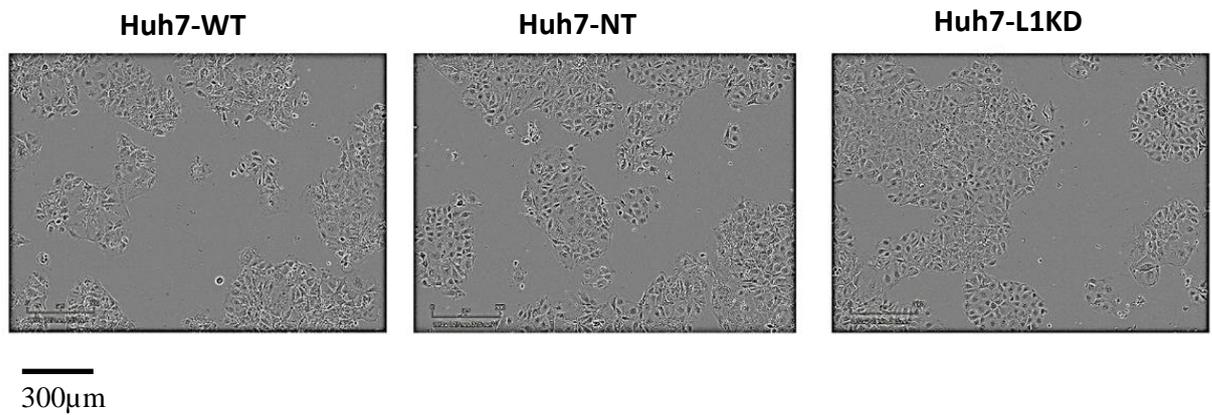


Figure 3.8 Huh7-WT, Huh7-NT and Huh7-L1KD cell morphology. Images were taken by Incucyte imaging from a 24well plate.

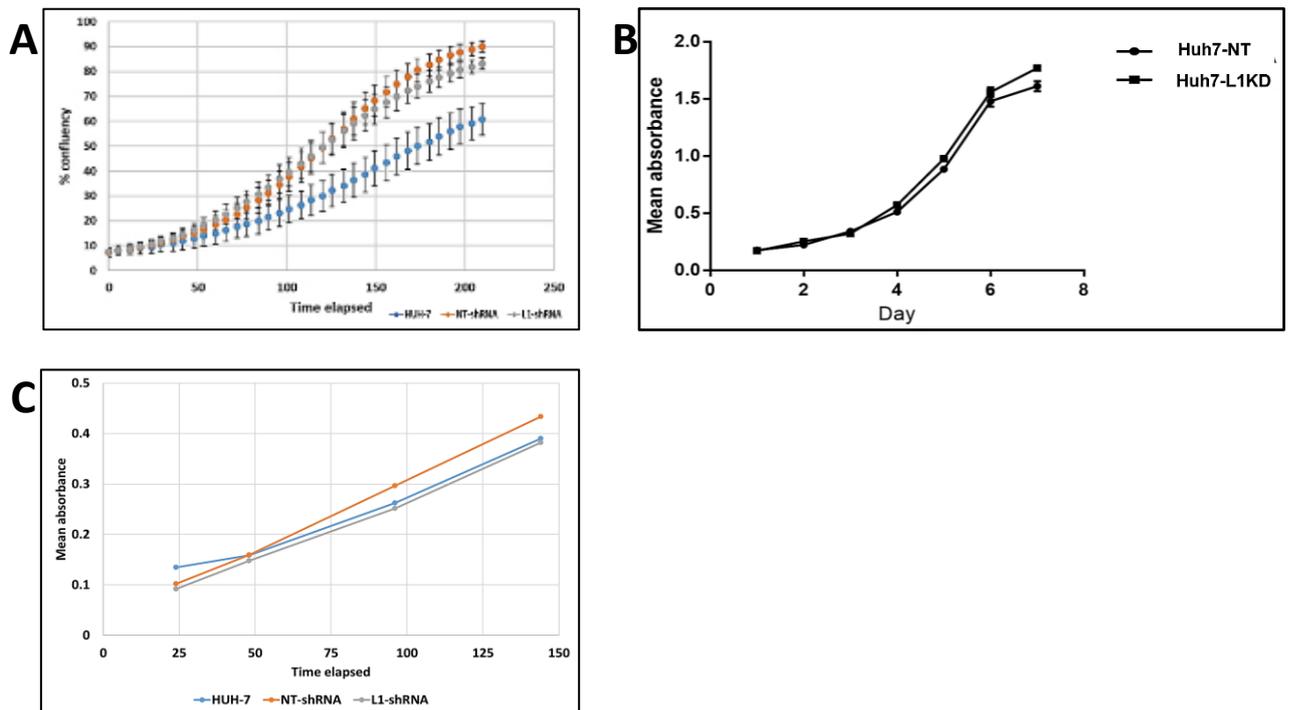


Figure 3.9 Growth proliferation curve between Huh7-WT WT, Huh7-NT and Huh7-L1KD. (A) Incucyte imaging (n=3). Error bars represent standard error (B) SRB assay (n=1) (C) MTT assay (n=1).

Colony formation ability was measured in a 10cm² Petri dish using the following cell densities: 4000 and 6000. After 3 weeks, cells were fixed and stained with crystal violet.

Colony size and number were measured using an automatic cell counter for each cell line. Huh7-L1KD had larger colonies but colony number were similar compared to Huh7-WT and Huh7-NT (**Figure 3.10** and **Figure 3.11**).

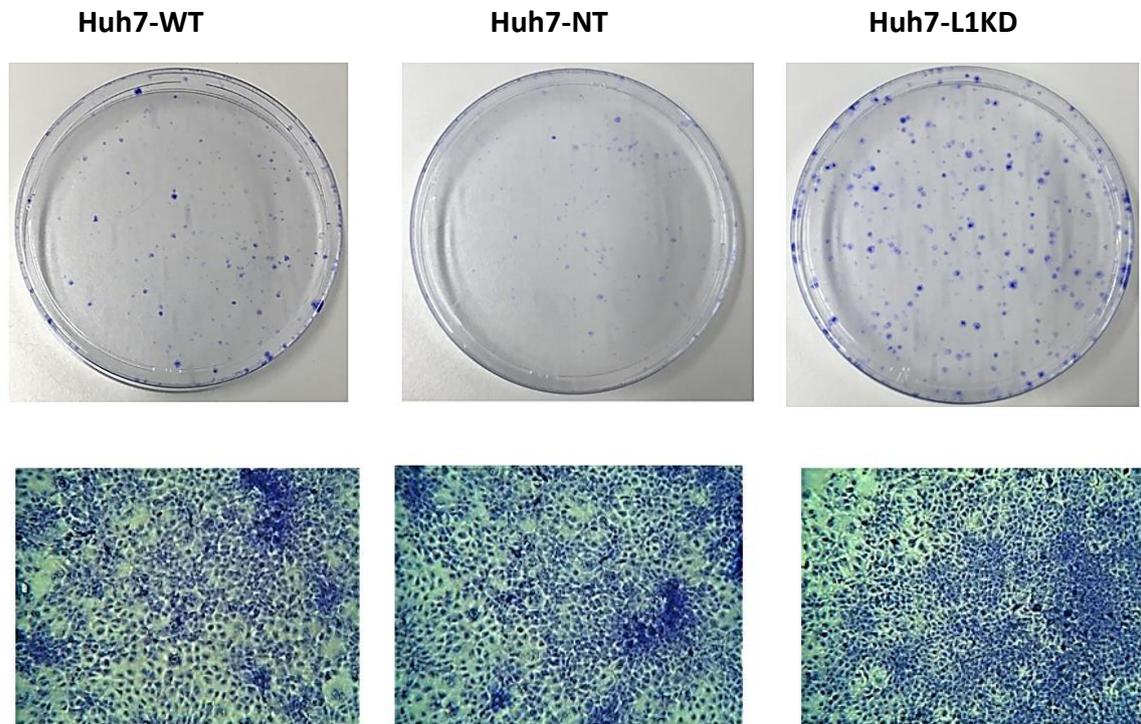


Figure 3.10 Huh7-WT, Huh7-NT and Huh7-L1KD clonogenic assay. Cells were seeded at 4000 cells per petri dish and fixed 21 days after. Top image represents petri dish with colonies and underneath 10x magnification of colonies. Huh7-L1KD colonies were more dense than Huh7-WT and Huh7-NT.

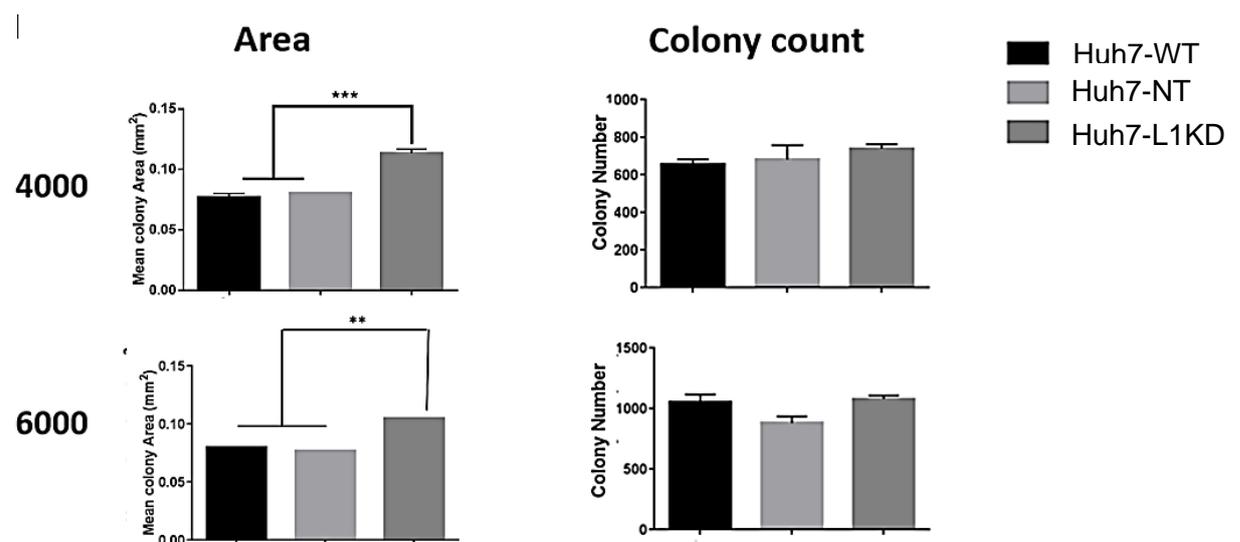


Figure 3.11 Mean colony number and size in Huh7-WT, Huh7-NT and Huh7-L1KD. Clonogenic assay was performed with the following cell densities: 4000 and 6000. Plates were then analysed

using an automatic colony counter. A 0.5mm minimum cut off was applied for each colony. Error bars represent standard error from three technical repeats. **Area (4000)**=Huh7-WT vs Huh7-NT $p=0.68$; Huh7-WT vs Huh7-L1KD $p=0.0002$; Huh7-NT vs Huh7-L1KD $p=0.0003$ **Area (6000)**: Huh7-WT vs Huh7-NT $p=0.77$; Huh7-WT vs Huh7-L1KD $p=0.0031$; Huh7-NT vs Huh7-L1KD $p=0.0022$, One way ANOVA.

Huh7-WT, Huh7-NT and Huh7-L1KD migration capacity were measured using Incucyte (**Figure 3.12**). Cells were seeded onto a 96 well plate and a scratch was applied using the Incucyte Wound Maker 96-pin. Measurements were taken at regular 3-hour intervals. Huh7-L1KD were migrating less than WT and Huh7-NT. Invasion capacity was measured using Boyden chambers. Cells were seeded onto these chambers, and after 72 hours invasive cells were stained and then measured using phase-contrast microscope. Huh7-L1KD cells were significantly less invasive than Huh7-WT and Huh7-NT (**Figure 3.13**). Thus, Huh7-L1KD had reduced cell migration and invasion compared to Huh7-WT and Huh7-NT.

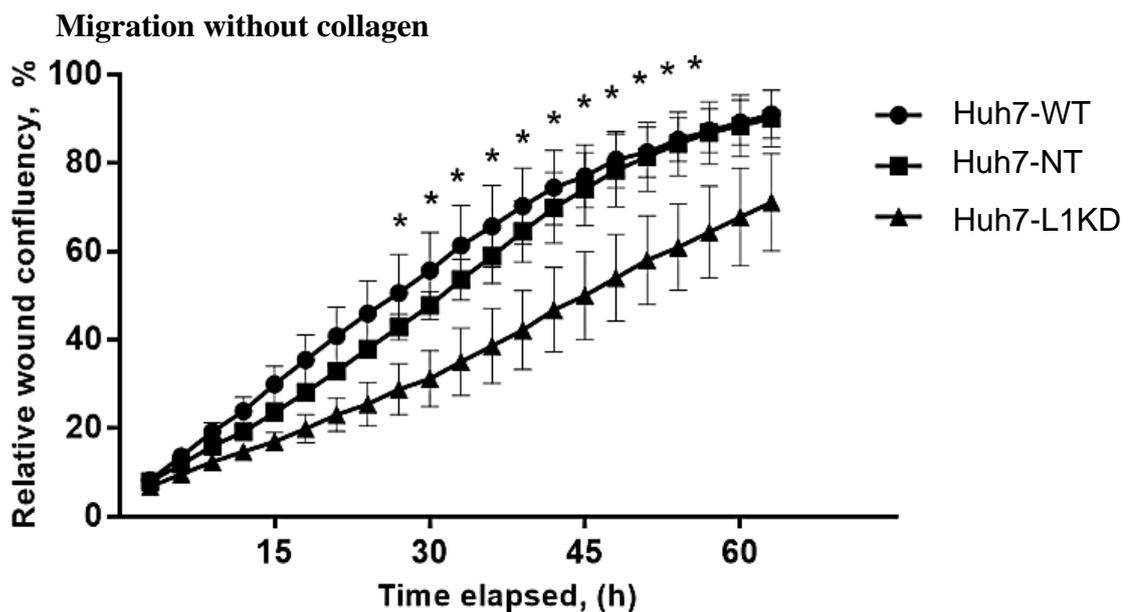


Figure 3.12 Migration assay of Huh7-WT, Huh7-NT and Huh7-L1KD (A) absence of collagen: Huh-7 vs L1-shRNA $p<0.05$ (27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57 hours). Huh-7 vs NT-shRNA $p<0.05$ (39, 42,45,48,51,54,57 hours), 3 independent experiments. Error bars represent standard error from 3 independent experiments

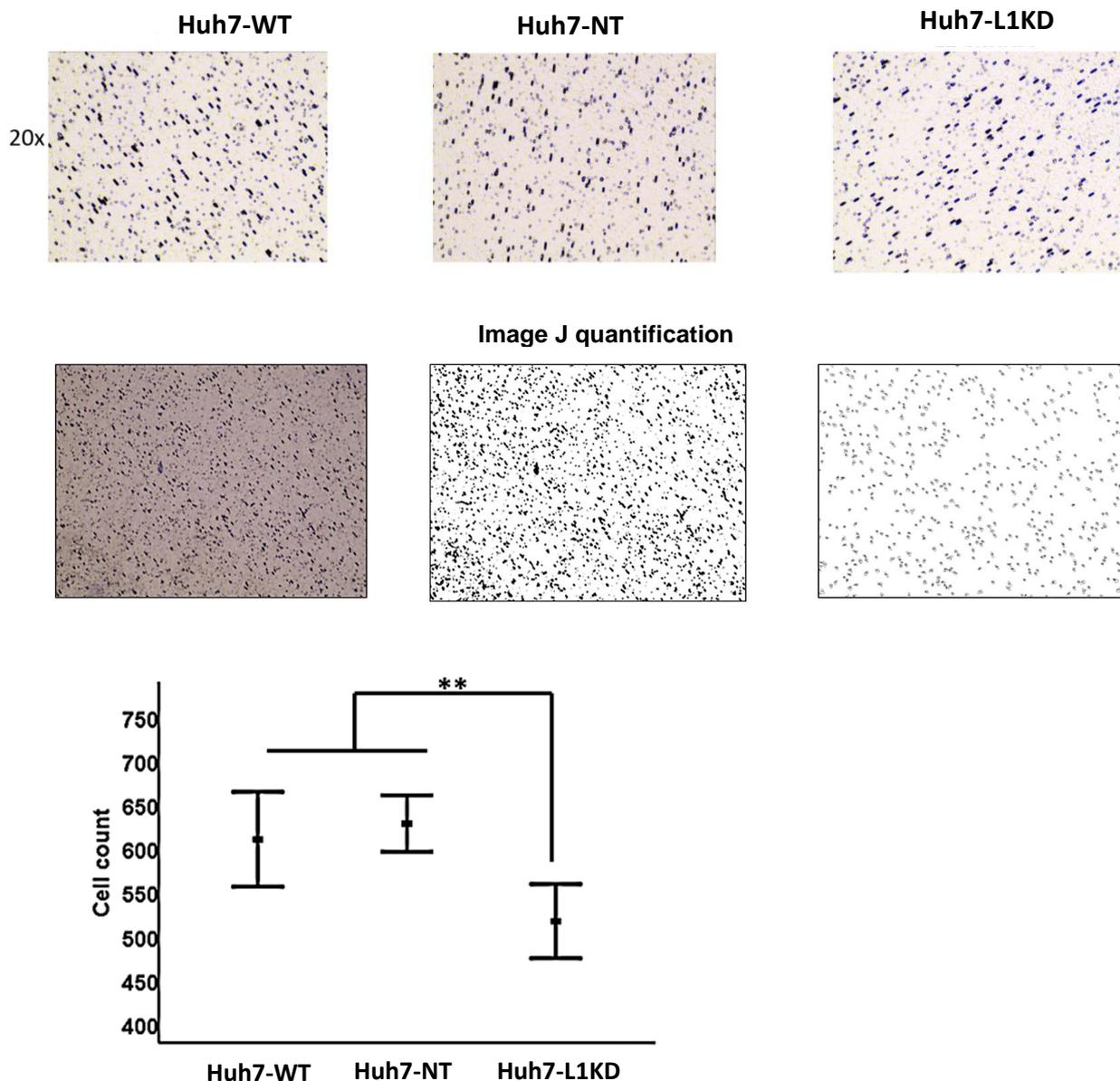


Figure 3.13 Huh7-WT, Huh7-NT and Huh7-L1KD cell invasion. Cells were incubated in a Boyden chamber for 72 hours and stained for invasive cells. Images were taken and quantified using Image J. Statistical analysis was performed using One Way ANOVA **Cell count:** Huh7-WT vs Huh7-NT $p=0.773$, Huh7-WT and Huh7-L1KD $p=0.009$, Huh7-NT and Huh7-L1KD $p=0.001$ ($n=4$, independent experiments).

Further studies were performed to investigate the role of L1 inhibition on tumour growth *in vivo*. 5 nude mice were used for each cell line and injected with 5×10^6 cells onto the dorsal lateral thorax region. From each group, two mice had tumour engraftment and growth was measured until tumour volume reached 500mm^3 . Once tumour engraftment

occurred, rapid tumour growth was observed (**Figure 3.14**). Here, Huh7-L1KD had a delayed tumour response compared to WT and Huh7-NT.

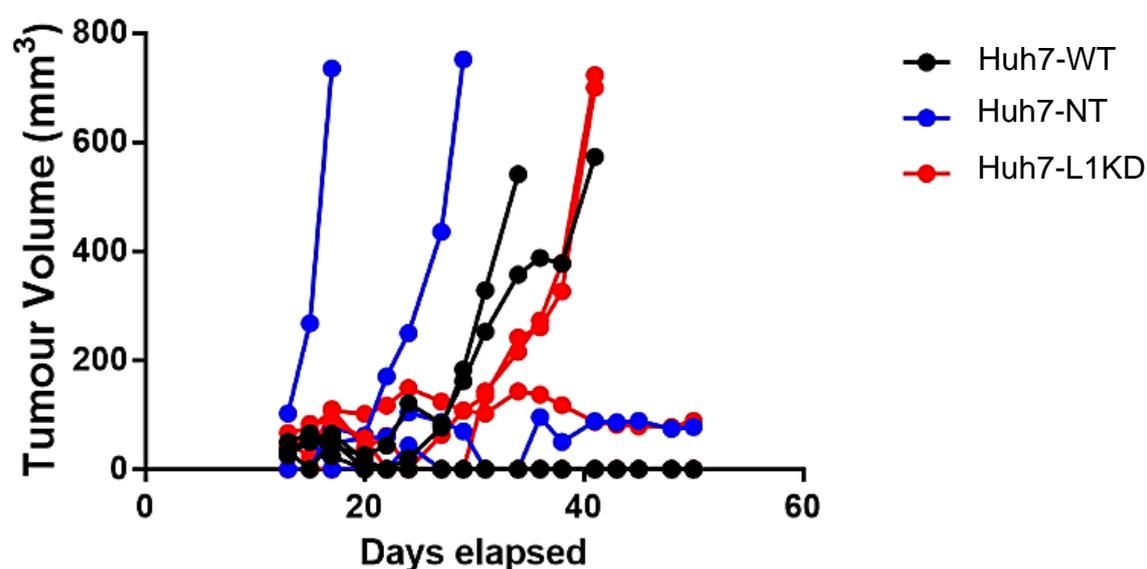


Figure 3.14 Huh7-WT, Huh7-NT and Huh7-L1KD cells were subcutaneously injected into nude mice, 5×10^6 cells. Tumour size were measured until day 50, or terminated early when tumour size was larger than 500mm^3 . Huh7-L1KD (n=2) had a delayed tumour response compared to WT (n=2) and Huh7-NT (n=2).

3.4.4 Influence of L1 knockdown on cellular transcriptome - RNAseq analysis of Huh7-NT and Huh7-L1KD cells

RNAseq was performed between Huh7-NT and Huh7-L1KD to identify differentially expressed genes (**Table 3.5**).

Table 3.5 Number of reads in control (Huh7-NT) and L1 knockdown (Huh7-L1KD)

| Group | Case ID | Number of mapped reads | of mapped reads (%) |
|--------------|-------------|------------------------|---------------------|
| Control | Huh7-NT 1 | 18,104,070 | 90.91 |
| | Huh7-NT 2 | 16,565,680 | 88.21 |
| | Huh7-NT 3 | 14,447,823 | 87.37 |
| L1 knockdown | Huh7-L1KD 1 | 18,766,946 | 91.78 |
| | Huh7-L1KD 2 | 18,578,050 | 92.54 |
| | Huh7-L1KD 3 | 18,586,416 | 91.65 |

DeSeq2 analysis identified 1512 differentially expressed (DE) genes (950 downregulated and 561 upregulated) at $\log_2FC0.5$, $padj < 0.05$. At \log_2FC1 , 334 genes were differentially expressed (242 downregulated and 92 upregulated), **Figure 3.15**. See **Table 3.6** for Top 50 significantly DE genes.

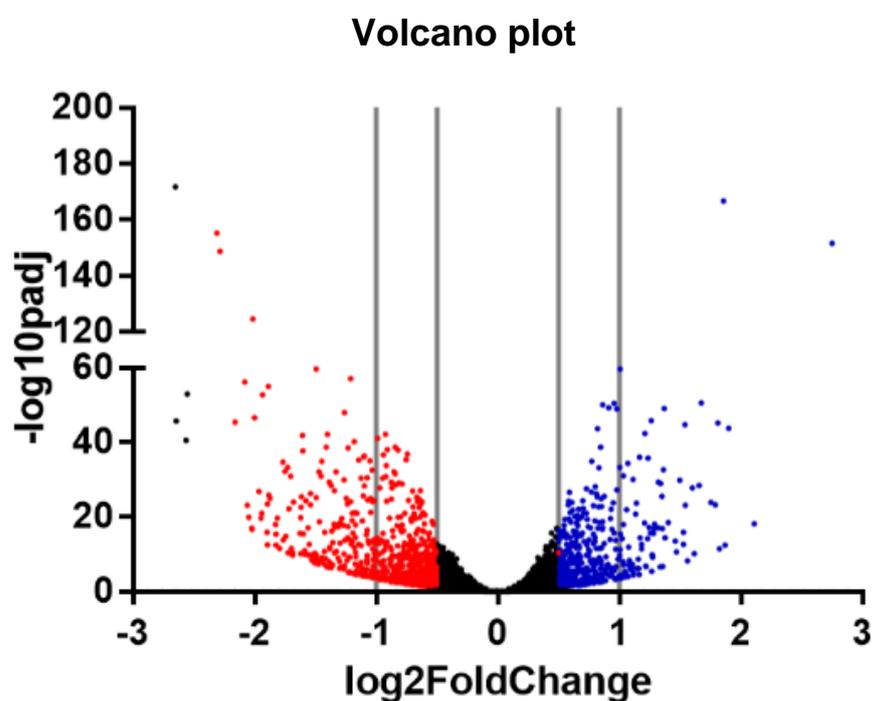


Figure 3.15 Volcano plot of differentially expressed genes between Huh7-NT and Huh7-L1KD. (A) \log_2FC1 , $padj < 0.05$, 334 genes, (B) $\log_2FC0.5$, $padj < 0.05$, 1512 genes.

Table 3.6 Top 50 most significant differentially expressed genes in Huh7-L1KD compared to Huh7-NT

| Gene Name | Base Mean | \log_2 Change | Fold Padj |
|-----------|-----------|--------------------|--------------|
| DPYSL3 | 963.5415 | -2.64835 | 2.47E-172 |
| LYZ | 3045.921 | 1.856762 | 2.14E-167 |
| NPNT | 1197.392 | -2.30917 | 8.09E-156 |
| CPS1 | 792.7656 | 2.750509 | 3.14E-152 |

| | | | |
|-----------------|----------|----------|-----------|
| TMSB4X | 1184.812 | -2.28281 | 2.47E-149 |
| CCDC80 | 1287.429 | -2.01328 | 3.47E-125 |
| PEG10 | 30603.02 | 1.167453 | 2.59E-106 |
| FSTL1 | 1275.984 | -1.65069 | 2.66E-91 |
| PLG | 1154.34 | 1.927129 | 2.11E-85 |
| RGS4 | 200.8243 | -3.2902 | 1.49E-73 |
| NRP1 | 3421.405 | -1.16135 | 8.99E-68 |
| ACBD5 | 1484.936 | -1.31361 | 3.95E-64 |
| RB1CC1 | 321.3372 | -2.33573 | 4.54E-64 |
| FBN2 | 384.2992 | -2.19886 | 5.03E-63 |
| APOB | 64467.16 | 1.00801 | 2.14E-60 |
| MICAL2 | 1667.648 | -1.49296 | 2.42E-60 |
| MYO1C | 1852.899 | -1.20775 | 1.03E-57 |
| CTSE | 405.8982 | -2.0768 | 6.31E-57 |
| SH3BGRL2 | 496.5474 | -1.88495 | 1.12E-55 |
| ALPK2 | 218.4504 | -2.55147 | 1.12E-53 |
| UGT2A3 | 409.9555 | -1.93155 | 1.94E-53 |
| HS3ST3B1 | 764.0174 | 1.673973 | 2.77E-51 |
| ODC1 | 3557.192 | 0.956994 | 3.40E-51 |
| FGB | 16940.52 | 0.864077 | 9.23E-51 |
| AHSG | 15111.9 | 0.91399 | 5.97E-50 |
| FST | 1104.414 | 1.370336 | 1.02E-49 |
| NQO1 | 3128.073 | 0.981947 | 1.15E-49 |
| ITGA5 | 1040.746 | -1.25984 | 1.14E-48 |
| CPT1A | 320.9832 | -1.99917 | 3.24E-47 |
| CDK6 | 993.5035 | 1.263358 | 1.72E-46 |
| GABRA4 | 158.811 | -2.64048 | 2.02E-46 |
| MYO18B | 261.4476 | -2.15778 | 4.45E-46 |
| MEP1A | 483.1994 | 1.811325 | 6.85E-46 |
| PRTG | 566.3422 | 1.54081 | 2.08E-45 |
| DCDC5 | 340.4034 | 1.901649 | 1.82E-44 |
| SCFD1 | 9384.83 | 0.822101 | 2.14E-44 |

| | | | |
|-----------------|----------|----------|----------|
| UBD | 994.9204 | 1.212558 | 4.78E-43 |
| CYP26B1 | 670.0063 | -1.39879 | 7.83E-43 |
| CDH2 | 3089.313 | -0.92206 | 8.20E-43 |
| LOXL2 | 500.9349 | -1.60645 | 1.60E-42 |
| CRAT | 2080.158 | -0.98527 | 9.05E-42 |
| UPK1B | 177.0285 | -2.56235 | 2.95E-41 |
| PTRF | 980.5382 | -1.17841 | 6.69E-41 |
| MYRF | 3885.1 | -0.84294 | 2.06E-39 |
| ST6GAL1 | 8493.857 | 0.846946 | 2.31E-39 |
| MAP1B | 606.4116 | -1.40888 | 2.42E-39 |
| EPS8L2 | 821.6556 | -1.22815 | 3.76E-39 |
| PKM | 12509.91 | -0.90621 | 1.04E-38 |
| SERPINA5 | 3679.159 | -0.8182 | 1.38E-38 |
| COL2A1 | 433.3259 | -1.60156 | 1.94E-38 |

Padj=P.value adjusted

Gene Set Enrichment Analysis (GSEA) was performed to identify pathway differences in Huh7-L1KD compared to Huh7-NT. Initial analysis was performed on the top 50 most significant differentially expressed genes in Huh7-L1KD compared to Huh7-NT. Analysis was performed for Hallmark pathways and identified 8 out of 50 genesets significantly enriched in the genelist (**Table 3.7**). EMT, angiogenesis and Hedgehog signalling were the most significant enriched pathways ($p < 0.001$, $FDR < 0.001$). Both EMT and angiogenesis gene sets were decreased in Huh7-L1KD indicating downregulation of these pathways.

Further analysis was performed by including all the differentially expressed genes with a $\log_2FC > 0.5$, $padj < 0.05$ (1512 genes). Hallmark pathways identified 44 gene sets (**Table 3.8**). Again, EMT geneset was the top most significantly enriched pathway, the list included TGF- β signalling as well. Furthermore, GSEA analysis for Keggs pathways identified 83 out of 186 gene sets significantly enriched in the genelist (**Table 3.9**). The most significant enriched pathways were EMT related pathways such as Focal Adhesion ($k/K = 0.201$), ECM Receptor adhesion ($k/K = 0.2262$) and TGF- β signalling ($k/K = 0.200$), $p < 0.001$, $FDR < 0.001$. Furthermore, hallmark enrichment plot demonstrated TGF- β downregulation in Huh7-L1KD (**Figure 3.16**). Thus, GSEA analysis demonstrate reduced TGF- β signalling and EMT in Huh7-L1KD.

Table 3.7 TCGA analysis for Hallmark pathways on top 50 significant differentially expressed genes between Huh7-L1KD versus Huh7-NT cells.

| Gene Set Name | Description | k/K | p-value | FDR q-value | Differentially expressed Genes |
|------------------------------------------|-------------------------------------------------------------------------------------------------|--------|----------|-------------|------------------------------------------------------|
| EPITHELIAL MESENCHYMAL TRANSITION | Genes defining epithelial-mesenchymal transition, as in wound healing, fibrosis and metastasis. | 0.035 | 8.22E-09 | 4.11E-07 | FSTL1, RGS4, ITGA5, CDH2, DPYSL3, FBN2, LOXL2 |
| ANGIOGENESIS | Genes up-regulated during formation of blood vessels (angiogenesis). | 0.0833 | 1.48E-05 | 2.46E-04 | FSTL1, NRPI, SERPINA5, |
| HEDGEHOG SIGNALING | Genes up-regulated by activation of hedgehog signaling. | 0.0833 | 1.48E-05 | 2.46E-04 | NRPI, PLG, CDK6, |
| IL2 STAT5 SIGNALING | Genes up-regulated by STAT5 in response to IL2 stimulation. | 0.02 | 1.41E-04 | 1.41E-03 | NRPI, MYO1C, ODC1, SH3BGRL2 |
| MYOGENESIS | Genes involved in development of skeletal muscle (myogenesis). | 0.02 | 1.41E-04 | 1.41E-03 | MYO1C, CRAT, NQO1,FST |
| COAGULATION | Genes encoding components of blood coagulation system; also up-regulated in platelets. | 0.0217 | 8.07E-04 | 6.52E-03 | PLG, MEPIA, CTSE |
| UV RESPONSE DN | Genes down-regulated in response to ultraviolet (UV) radiation. | 0.0208 | 9.13E-04 | 6.52E-03 | RGS4, NRPI, MAP1B |
| FATTY ACID METABOLISM | Genes encoding proteins involved in metabolism of fatty acids. | 0.019 | 1.19E-03 | 7.45E-03 | ODC1,CRAT, CPT1A |

Differentially expressed genes: red=downregulated, blue=upregulated

k/K= Genes in Overlap (k)/ Genes in Gene Set (K); FDR= False Discovery Rate

Table 3.8 TCGA analysis for Hallmark pathways on all differentially expressed gene at logFC 0.5, p<0.05 between Huh7-L1KD versus Huh7-NT cells (Top 10 pathways are shown only).

| Gene Set Name | Description | k/K | p-value | FDR q-value |
|------------------------------------------|-------------------------------------------------------------------------------------------------|--------|----------|-------------|
| EPITHELIAL_MESENCHYMAL_TRANSITION | Genes defining epithelial-mesenchymal transition, as in wound healing, fibrosis and metastasis. | 0.315 | 3.39E-39 | 1.70E-37 |
| COAGULATION | Genes encoding components of blood coagulation system; also up-regulated in platelets. | 0.2899 | 7.65E-24 | 1.91E-22 |
| TNFA_SIGNALING_VIA_NFKB | Genes regulated by NF-kB in response to TNF [GeneID=7124]. | 0.215 | 5.61E-20 | 9.35E-19 |
| IL2_STAT5_SIGNALING | Genes up-regulated by STAT5 in response to IL2 stimulation. | 0.21 | 3.84E-19 | 3.84E-18 |
| MYOGENESIS | Genes involved in development of skeletal muscle (myogenesis). | 0.21 | 3.84E-19 | 3.84E-18 |
| UV_RESPONSE_DN | Genes down-regulated in response to ultraviolet (UV) radiation. | 0.2431 | 2.52E-18 | 2.10E-17 |
| ESTROGEN_RESPONSE_EARLY | Genes defining early response to estrogen. | 0.2 | 1.65E-17 | 9.15E-17 |
| HYPOXIA | Genes up-regulated in response to low oxygen levels (hypoxia). | 0.2 | 1.65E-17 | 9.15E-17 |
| XENOBIOTIC_METABOLISM | Genes encoding proteins involved in processing of drugs and other xenobiotics. | 0.2 | 1.65E-17 | 9.15E-17 |
| ESTROGEN_RESPONSE_LATE | Genes defining late response to estrogen. | 0.175 | 1.13E-13 | 5.67E-13 |

Table 3.9 Keggs pathway analysis on all differentially expressed gene at logFC 0.5, p<0.05 (Top 10 pathways are shown only).

| Gene Set Name | Description | k/K | p-value | FDR q-value |
|-----------------------------------------------------|----------------------------------------------|--------|----------|-------------|
| FOCAL_ADHESION | Focal adhesion | 0.201 | 1.37E-17 | 2.55E-15 |
| COMPLEMENT_AND_COAGULATION_CASCADES | Complement and coagulation cascades | 0.3188 | 1.02E-14 | 9.48E-13 |
| AXON_GUIDANCE | Axon guidance | 0.2016 | 5.77E-12 | 3.58E-10 |
| PEROXISOME | Peroxisome | 0.2564 | 1.52E-11 | 7.07E-10 |
| PATHWAYS_IN_CANCER | Pathways in cancer | 0.1231 | 2.08E-10 | 7.72E-09 |
| ECM_RECEPTOR_INTERACTION | ECM-receptor interaction | 0.2262 | 5.06E-10 | 1.57E-08 |
| ABC_TRANSPORTERS | ABC transporters | 0.2955 | 8.39E-09 | 2.23E-07 |
| TGF_BETA_SIGNALING_PATHWAY | TGF-beta signaling pathway | 0.2 | 2.94E-08 | 6.83E-07 |
| METABOLISM_OF_XENOBIOTICS_BY_CYTOCHROME_P450 | Metabolism of xenobiotics by cytochrome P450 | 0.2 | 4.83E-07 | 9.98E-06 |
| SMALL_CELL_LUNG_CANCER | Small cell lung cancer | 0.1786 | 8.84E-07 | 1.64E-05 |

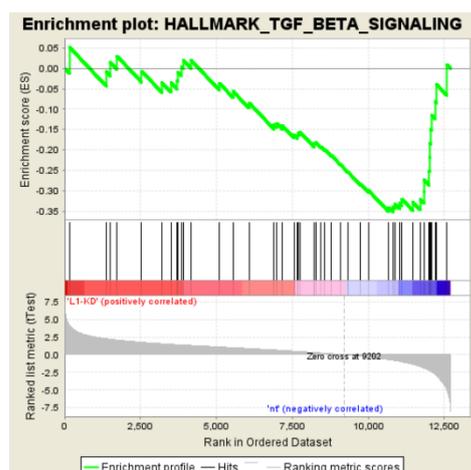


Figure 3.16 TGF- β enrichment plot using Hallmark analysis. Analysis was performed on complete genelist using L1 expression as a continuous variable, demonstrating TGF- β signalling downregulation in Huh7-L1KD.

As RNA-seq analysis demonstrated TGF- β downregulation in Huh7-L1KD, transcript and protein analysis were performed to confirm the finding. Transcript analysis was performed by RT-qPCR for the following TGF- β associated genes: SMAD3, TGF β 1, SMAD2, SMAD4 and SMAD7. SMAD3 and TGF β 1 were both downregulated in Huh7-L1KD compared to WT and Huh7-NT (**Figure 3.17**). SMAD2, SMAD4 and SMAD7 were unaffected.

Protein expression of SMAD3 and SMAD4 was analysed by western blotting (**Figure 3.18**). SMAD3 protein expression was downregulated in Huh7-L1KD, and SMAD4 protein expression was unaffected. In order to check the functionality of SMAD3, a luciferase reporter assay using Pai1 promoter-luc plasmid (containing SMAD3 binding site) was utilised. Again, Huh7-L1KD cells exhibited significantly lower luciferase signal hence, reduced SMAD3 activity and reduced TGF- β activity compared to Huh7-WT and Huh7-NT cells (**Figure 3.19**). Overall, the data confirms reduction in endogenous TGF- β signalling in Huh7 cells upon L1 knockdown.

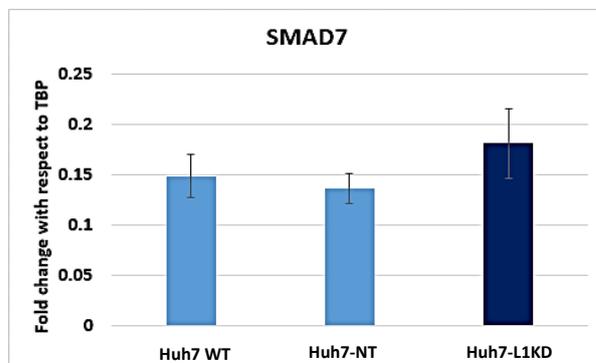
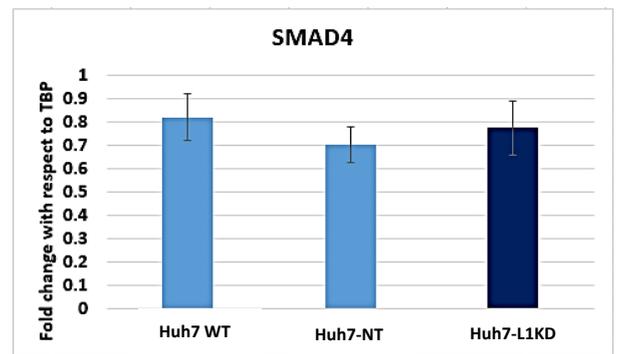
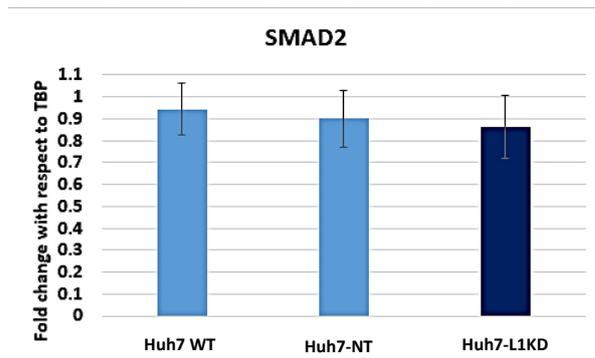
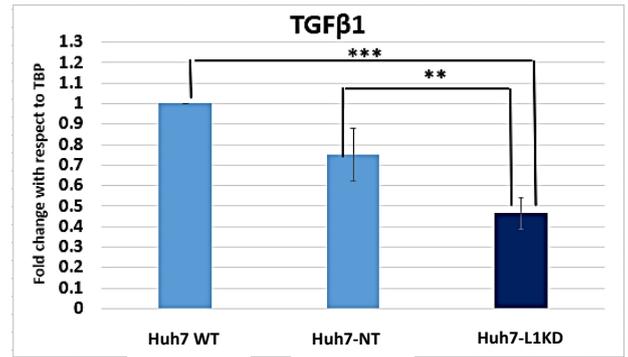
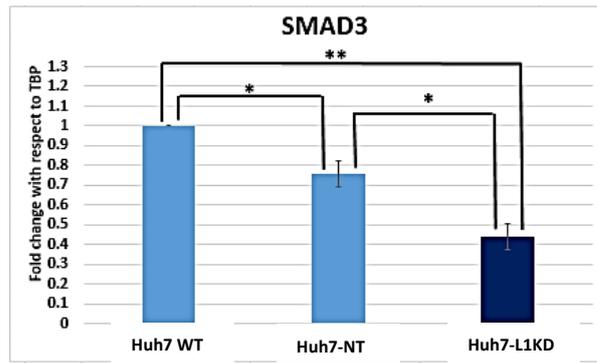


Figure 3.17 SMAD3, TGFβ1, SMAD2, SMAD4 and SMAD7 transcript expression in Huh7-WT, Huh7-NT and Huh7-L1KD. SMAD3 (n=9) and TGFβ1 (n=7) expression were both downregulated in Huh7-L1KD. Statistical analysis was performed using One way ANOVA, **SMAD3**: Huh7-WT vs Huh7-L1KD p= 0.0001; Huh7-NT vs Huh7-L1KD p=0.0096. **TGFβ1**: Huh7-WT vs Huh7-L1KD p=0.0023; Huh7-WT vs Huh7-NT p=0.0435; Huh7-NT vs Huh7-L1KD p=0.0365. SMAD2, SMAD4 and SMAD7 transcript expression were unaltered (n=1).

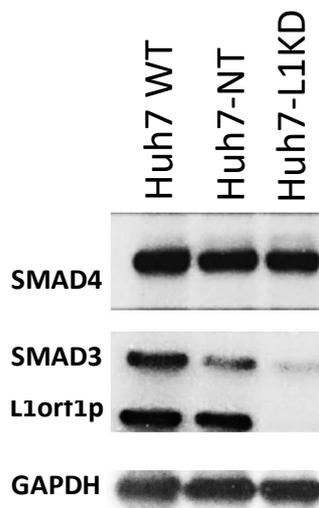


Figure 3.18 SMAD3 and SMAD4 protein expression in Huh7-WT, Huh7-NT and Huh7-L1KD. GAPDH was used as a loading control.

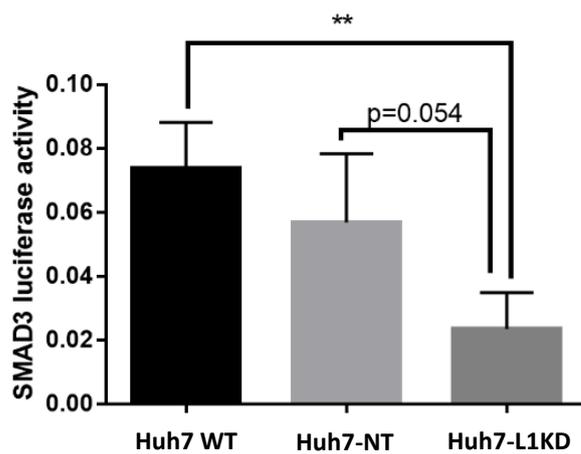


Figure 3.19 SMAD 3 luciferase assay in Huh7-WT, Huh7-NT and Huh7-L1KD: Huh7-WT vs Huh7-L1KD= $p=0.0012$; Huh7-NT vs Huh7-L1KD= $p=0.054$, Independent t-test ($n=3$).

3.4.5 Effect of L1 knockdown on the response of cells to TGF- β treatment

The effect of TGF- β induction was investigated at transcript and protein levels in Huh7-WT, Huh7-NT and Huh7-L1KD (**Figure 3.20 and Figure 3.21**).

Transcript analyses were performed with the following TGF- β inductions: Untreated, 1ng/ml and 10ng/ml, 72 hrs after treatment. SMAD3, TGF β 1 and vimentin transcript expression were analysed to measure TGF- β signalling. All the three cell lines exhibited dose-dependent increase in SMAD3, TGF β 1 and vimentin transcripts. At 10ng/ml TGF- β induction, at least 7-fold increase in TGF β 1 and a 12-fold increase in SMAD3 transcript level expression were observed in all three cell lines. Similarly vimentin expression increased by 4-fold compared to untreated.

There was no significant difference between the cell lines in terms of cell cycle regulators (p15 and p21) at basal level however, an induction in p15 and p21 levels was observed in all the 3 cell lines upon TGF- β stimulation. Hence, overall the cell lines show similar response to activation of TGF- β targets upon treatment with TGF- β treatment implementing that L1 knockdown only reduced basal TGF- β signalling in the cells.

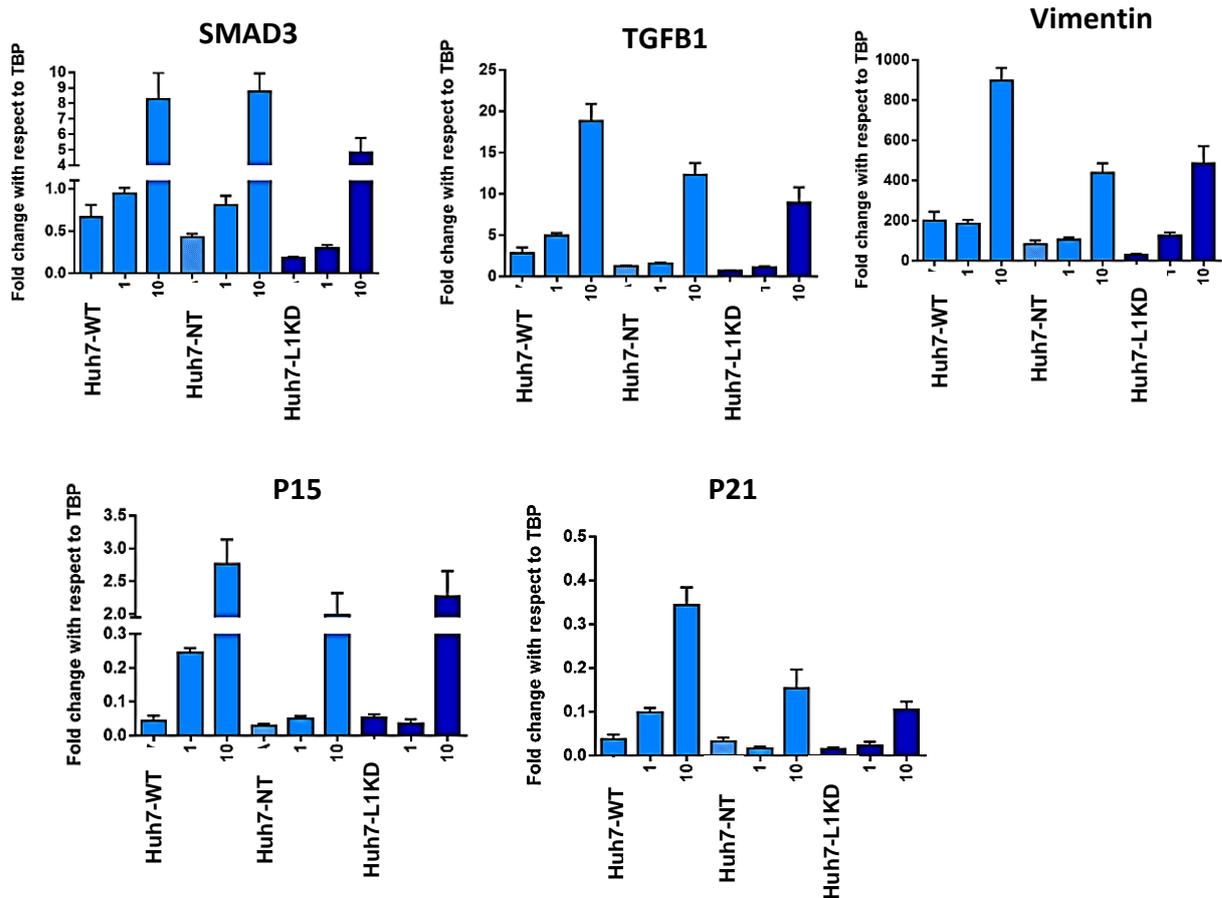


Figure 3.20 RT-qPCR analysis of WT, Huh7-NT and Huh7-L1KD. Cells were induced with TGF- β for 72 hours and analysed for SMAD3, TGF β 1, vimentin, P15 and P21 transcript expression. Error bars represent standard error from three technical repeat. TBP was used as a housekeeping gene.

Furthermore, we measured SMAD3 protein level in Huh7-WT, Huh7-NT and Huh7-L1KD 72 hours after induction with 1ng/ml TGF- β . Similar to transcript expression, SMAD3 protein expression increased in all the three cell lines upon TGF- β induction however, SMAD3 expression remained lower in Huh7-L1KD cells compared to Huh7-WT and Huh7-NT cells as was the case for the cells at basal level (**Figure 3.21**). Also, L1orf1p expression remained unaffected by TGF- β treatment.

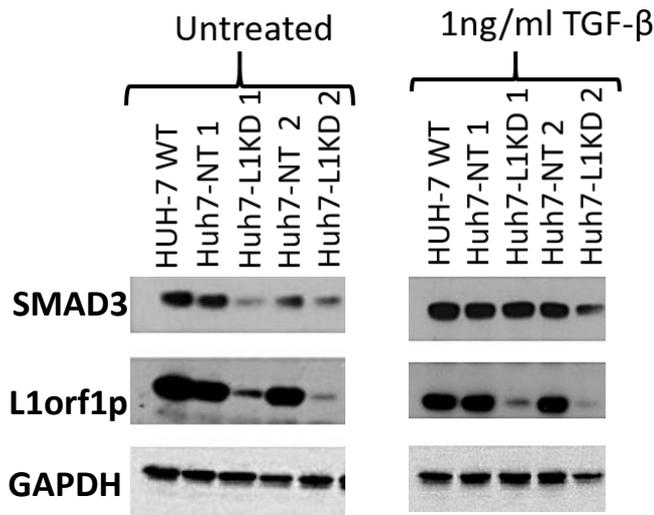


Figure 3.21 SMAD3 and L1orf1p expression in Huh7-WT, Huh7-NT and Huh7-L1KD in absence and present of TGF- β induction for 72 hours. Protein expression was measured by western blotting. (Huh7-NT 1, Huh7-L1KD 1) and (Huh7-NT 2, Huh7-L1KD 2) were from independent lentiviral transduction in Huh7-WT. GAPDH was used as a loading control.

3.4.6 Effect of L1 knockdown on TGF- β mediated cell growth inhibition

Further studies were performed to investigate TGF- β sensitivity in Huh7-WT, Huh7-NT and Huh7-L1KD in terms of cell proliferation. Cells were seeded onto a 96 well plate and treated with TGF- β the next day: Untreated, 0.01ng/ml, 0.05ng/ml and 0.1ng/ml. Cell proliferation was then measured using Incucyte imaging. In absence of TGF- β induction, Huh7-NT and Huh7-L1KD had a similar proliferation rate however, Huh7-WT had a significant lower proliferation rate (**Figure 3.22**). 0.01 and 0.05ng/ml TGF- β induction reduced proliferation in all three cell lines. At 0.01ng/ml, Huh7-WT and Huh7-L1KD had significant lower proliferation than Huh7-NT: WT (47.7% decrease), Huh7-NT (16.4% decrease) and Huh7-L1KD (46.2% decrease), compared to corresponding untreated (at 150 hours). At 0.05ng/ml, Huh7-L1KD had reduced proliferation compared to WT and Huh7-NT: WT (48.7% decrease), Huh7-NT (32.0% decrease) and Huh7-L1KD (59.5% decrease), compared to corresponding untreated (150 hours). At 0.1ng/ml, complete growth inhibition was observed in all three cell lines. Hence, Huh7-L1KD cells show increased susceptibility to TGF- β treatment at lower doses when compared to Huh7-NT cells in terms of percentage of growth inhibition but the response is equivalent to Huh7-WT cells. All the three cell

lines show complete growth inhibition at 0.1ng/ml TGF- β treatment, hence susceptibility of Huh7 cells remained unchanged upon L1 knockdown.

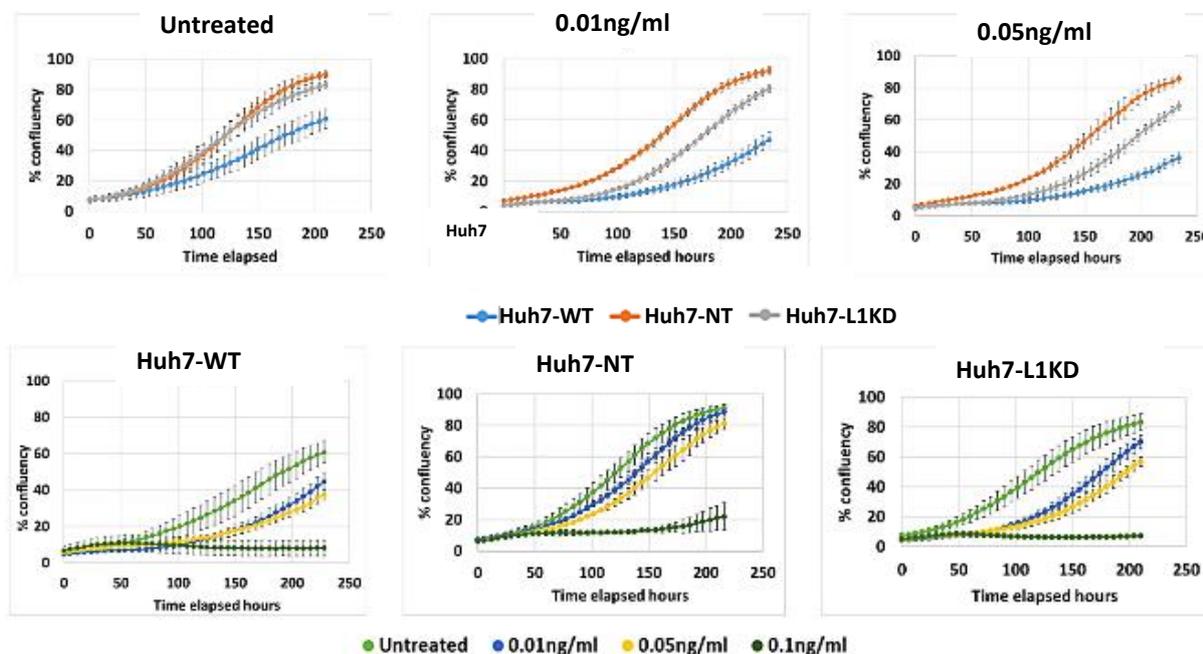


Figure 3.22 Growth proliferation curves of Huh7-WT, Huh7-NT and Huh7-L1KD with TGF- β induction. Cells were seeded onto a 96well plate and treated once with TGF- β : untreated, 0.01ng/ml, 0.05ng/ml and 0.1ng/ml. Cell proliferation was then measured by Incucyte imaging. **Untreated**= Huh7-WT vs Huh7-NT: From 102 to 228 hours. Huh7-WT vs Huh7-L1KD: 150 to 228 hours, $p < 0.05$. **0.01ng/ml**= Huh7-WT vs Huh7-NT: From 66 to 228 hours. Huh7-WT vs Huh7-L1KD: From 126 to 228 hours, Huh7-NT vs Huh7-L1KD: From 78 to 228 hours, $p < 0.05$. **0.05ng/ml**= Huh7-WT vs Huh7-NT: From 84 to 228hours. Huh7-WT vs Huh7-L1KD: From 144 to 228 hours. Huh7-NT vs Huh7-L1KD: From 96 hours to 228 hours, $p < 0.05$. Two Way ANOVA. Error bars represent standard error from three independent replicates.

3.4.7 Effect of L1 knockdown on doxorubicin sensitivity of Huh7 cells

TGF- β signalling can mediate drug resistance particularly doxorubicin resistance (Bhagyaraj et al., 2009, Akhurst 2017). Doxorubicin is a common drug used to treat HCC but its potency is reduced due to emergence of drug resistance. As L1 inhibition reduced TGF- β signalling, the effect of L1 inhibition on doxorubicin sensitivity was investigated. Huh7-WT, Huh7-NT and Huh7-L1KD were treated with 50nM doxorubicin for 48 hours and cells were then fixed and analysed for cell cycle distribution of the population (**Figure**

3.23). G2 cell cycle arrest was observed in all cell lines but was significantly elevated in Huh7-L1KD cells indicating higher doxorubicin sensitivity (WT=45.2%; Huh7-NT=48.1%, and Huh7-L1KD=60.2%).

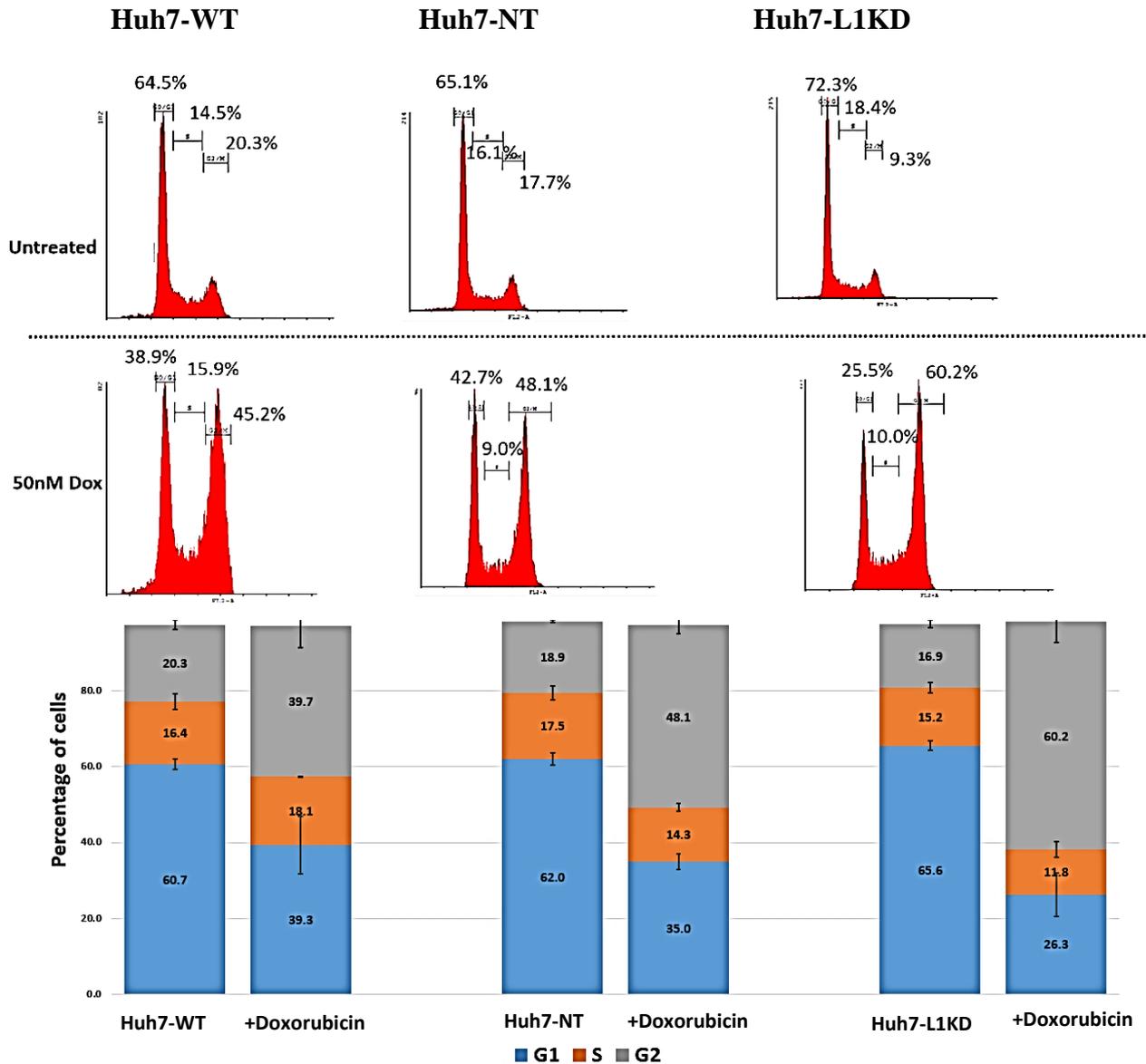


Figure 3.23 Cell cycle analysis of WT, Huh7-NT and Huh7-L1KD after 48 hours doxorubicin induction. G2 arrest is observed in all three cell lines but was significantly higher in Huh7-L1KD compared to WT and Huh7-NT. Values represents mean of three independent replicates. After doxorubicin treatment: **Pre G2:** Huh7-WT vs Huh7-NT $p=0.8844$, Huh7-WT vs Huh7-L1KD $p=0.7668$, Huh7-NT vs Huh7-L1KD $p=0.9648$. **Post G2:** Huh7-WT vs Huh7-NT $p=0.2473$, Huh7-WT vs Huh7-L1KD $p=0.0051$, Huh7-NT vs Huh7-L1KD $p=0.0514$, (n=3) Two Way ANOVA

3.5 Discussion

L1 expression was elevated in HCC patients and highly expressed in poorly differentiated tumours. However, the role of L1 and its interactions with cellular pathways remains largely unexplored. Here, the effect of L1 inhibition in HCC was investigated, specifically its interaction with TGF- β signalling pathway.

3.5.1 L1 expression in HCC cell lines

L1 expression has been measured in different liver cancer cell lines both at transcript and protein level. Transcript analysis was performed for L1 5'UTR, ORF1 and ORF2 using RT-qPCR. All three primers correlated with each other. Highest L1 expression was observed in Huh7 and Hep3B, medium expression was observed in Huh1, HepG2, SK-Hep1, SNU182 and SNU475, and lowest expression was observed in HHL5 (immortalised human hepatocytes) and PLC/PRF5 (**Figure 3.4**).

Protein expression was analysed by measuring L1orf1p, due to its higher expression relative to L1orf2p (Taylor et al., 2013). Similar to transcript, L1orf1p expression also varied in different liver cancer cell lines. Highest expression was observed in Huh7, SK-Hep1, SNU182, SNU475; moderately expressed in SNU182 and Huh-1, and low expression was observed in HHL5 and none in HepG2 (**Figure 3.5**). Previous studies have shown high L1 expression in HepG2 (Reyes-Reyes et al., 2016). However we observed very low expression of L1orf1p in HepG2 cells. Moreover, there were discrepancies between L1 transcript and protein expression in some cell lines for example, Huh-1, PLC/PRF5, SNU182 and SK-Hep1 have similar L1orf1 transcript expression however, L1orf1p expression varied significantly in these cell lines. Huh-1 and PLC/PRF5 have lower L1orf1p expression compared to SNU182 and SK-Hep1. This could be due to L1 posttranscriptional modifications and translational repression differences between the cell lines. Posttranscriptional modification can degrade transcripts either by siRNA, microRNA and piRNA. In particular, microRNA mir128 can inhibit retrotransposition in Hela cells by inducing L1 RNA transcripts degradation (Hamdorf et al., 2015). Furthermore, restriction factors can also affect and inhibit L1 retrotransposons such as APOBEC proteins. APOBEC proteins have several sub proteins but particularly A3A and A3B are effective inhibiting retrotransposition (Niewiadomska et al., 2007). The regulation of these pathways may affect L1 translation and thus its expression level in HCC cell lines.

3.5.2 Influence of L1 knockdown on Huh7 cells

As Huh-7 had high transcript and protein expression, a L1 knockdown transgenic cell line (Huh7-L1KD) was generated from these using shRNA targeting L1-orf1. The knockdown was successful and had reduced L1 transcript and protein expression (**Figure 3.7**). Functional studies were performed to investigate the effect of L1 inhibition on proliferation, colony formation, migration and invasion abilities of the cells. Overall, L1 inhibition had no effect on proliferation, but reduced cell migration and invasion were observed (**Figure 3.12 and 3.13**). *In vivo* experiments demonstrated delayed tumour growth in Huh7-L1KD cells in mice compared to WT and Huh7-NT control cell lines (**Figure 3.14**). Similar findings are reported for other cancer cell lines (Li et al., 2014, Yang et al., 2013b).

RNAseq analysis followed by gene set enrichment analysis revealed downregulation of EMT, angiogenesis, focal adhesion and ECM in Huh7-L1KD cell compared to Huh7-NT cells (**Table 3.8**). Several pathways have been shown to interact with L1. For example in adult brain tissues, L1 were able to increase transiently and mobilise in neuronal progenitor cells. The increase was associated with a switch of the Sox2/HDAC1 repressor complex to the TCF/LEF) activator complex. TCF/LEF activator complex is an important pathway in Wnt signalling and its expression can impact gene expression and function, and potentially increase L1 expression (Muotri et al., 2010).

Overall, all these processes are known to be regulated by TGF- β signalling pathways which was also found to be significantly downregulated in Huh7-L1KD cells. Hence, RNAseq analysis revealed a causal link between L1 expression and TGF- β signalling. The TGF- β pathway has an important role in HCC and can act either as a tumour suppressor or as a tumour promoter. In normal and early stages of HCC, TGF- β acts as a tumour suppressor and inhibits cell proliferation and induces apoptosis. However in late stages, tumour cells can inhibit TGF- β tumour suppressive functions by several mechanisms. For example, several downstream targets such as SMAD4 and TGF- β EMT markers can be affected. Hence, we characterised the TGF- β pathway components in Huh-7, Huh-7-NT control and Huh-7 L1-KD at both transcript and protein levels. In particular, downregulation of SMAD3 was observed in Huh7-L1KD cells compared to control cell lines (**Figure 3.17 and 3.18**). Furthermore, Pai1-luciferase reporter assay revealed reduced SMAD3 activity in Huh7-L1KD compared to WT and Huh7-NT cells (**Figure 3.19**). Hence, L1 knockdown reduced basal TGF- β signalling in Huh7 cells.

To evaluate if L1 knockdown affects response of cells upon stimulation with exogenous TGF- β treatment, Huh-7, Huh-7 NT-control and Huh-7 L1-KD were then induced with TGF- β . TGF- β induction increased SMAD3 and TGF β 1 transcript expression (**Figure 3.20**). Furthermore, increased SMAD3 protein expression was observed in all the 3 cell lines, indicating increased TGF- β signalling (**Figure 3.21**). Moreover, key TGF- β targets such as P15 and P21 were also upregulated in all the 3 cell lines upon TGF- β induction (**Figure 3.20**). P15 and P21 are both cyclin-dependent kinase inhibitors which can interact with TGF- β and affect EMT. P15 inhibits cyclin D-dependent kinase 4/6 mediated phosphorylation of pRB and G1/S progression (Gonzalez et al., 2006). P21 inhibits CDK2 activity, RB phosphorylation and DNA synthesis (Zhu et al., 2005). A recent study has shown that P15 and P21 upregulation can downregulate EMT marker and induce cellular senescence (Senturk et al., 2010b). In line with these results, inhibition of cell proliferation was observed in all the 3 cell lines upon stimulation with TGF- β (**Figure 3.22**). Hence, L1 knockdown reduced basal TGF- β signalling in Huh7 cells but the response to exogenous TGF- β stimulus was maintained in the cells probably due to upregulation of SMAD3 expression once the cells were stimulated with TGF- β .

3.5.3 Huh7-L1KD and doxorubicin sensitivity

TGF- β pathway can induce drug resistance including doxorubicin resistance (Bhagyaraj et al., 2009, Akhurst 2017). Doxorubicin is a chemotherapeutic agent used to treat HCC during TACE therapy. It inhibits DNA repair mechanisms by inhibiting topoisomerase II. Topoisomerase II inhibition blocks DNA and RNA synthesis leading to DNA fragmentation (Tewey et al., 1984). However, it has become increasingly less potent due to the emergence of drug resistance. Several studies have shown that TGF- β signalling can induce doxorubicin resistance. For example, doxorubicin resistant HCT116 cells had higher levels of EMT, plasma membrane glycoprotein and TGF- β signalling. Furthermore, SMAD4 inhibition reduced TGF- β signalling and plasma membrane glycoprotein, reversed EMT leading to increased doxorubicin sensitivity (Li et al., 2015). Other studies have observed similar findings. In an MDA-MB-231 orthotopic xenograft, doxorubicin and TGF- β antagonist treatment reduced EMT and decreased tumour growth. Similarly, in A549 human lung adenocarcinoma cells, reduced TGF- β signalling led to increased doxorubicin cytotoxic effects.

We observed higher G2 arrest in Huh7-L1KD compared to Huh7-WT and Huh7-NT upon treatment with doxorubicin (**Figure 3.23**). Hence, L1 knockdown rendered increased sensitivity to doxorubicin in Huh7 cells probably due to reduced TGF- β signalling. Previously, L1 knockdown in HepG2 cells has been reported to increase sensitivity of the cells to reduced IC₅₀ for epirubicin, cisplatin and paclitaxel and demonstrated a reduced IC₅₀ by up to 9 fold. The decrease was likely due to decreased BCL-2 (anti-apoptotic) expression in L1 knockdown (Feng et al., 2013). Likewise, decrease in BCL-2 expression (FC -0.55, adj p-value 0.00099) was observed in Huh7 cells upon L1 knockdown hence the increased sensitivity could be because of reduced BCL-2 levels.

The observation is in line with our human data, where patients with higher L1 expression have poor prognosis when treated with TACE (Chapter 1, **Figure 3.10C**). Further studies are needed to confirm the findings and understand the mechanism of increase in doxorubicin sensitivity upon L1 knockdown.

3.6 Summary

L1 expression has been characterised in different liver cancer cell lines. Highest expression was observed in Huh-7, and stable knockdown of L1 was performed on this cell line using lentivirus based Huh-7 L1-KD construct. RNAseq analysis demonstrated reduced TGF- β signalling in Huh7-L1KD and was verified at transcript, protein and functional (Pai1 luciferase reporter assay) levels. TGF- β functional properties such as migration, invasion and doxorubicin resistance were also reduced in Huh7-L1KD. However, TGF- β induction increased SMAD3 and TGF β 1 expression in Huh7-L1KD cells leading to similar response in terms of TGF- β targets activation and inhibition in cell proliferation to what was observed in Huh7-WT and Huh7-NT control cells upon TGF- β induction. Hence, TGF- β signalling is reduced in Huh7-L1KD but the pathway is not impaired.

Chapter 4: Effect of L1 overexpression on liver cancer cell lines

4.1 Introduction

L1 promoter is activated in several cancers and can increase the expression of L1 proteins - L1orf1p and L1orf2p. Both proteins are transcribed from an internal sense-promoter within the 5' untranslated region and are essential for retrotransposition. L1orf1p has RNA binding and nucleic acid chaperone activity, while L1orf2p has reverse transcriptase and endonuclease activities. Several studies have demonstrated increased L1orf1p expression in cancers (Barchitta et al., 2014). Furthermore, a correlation between L1 insertion and L1orf1p expression was also observed (Rodic et al., 2015).

L1orf1p is a 40kDA protein and has 338 amino acids (**Figure 4.1**). Atomic force microscopy analysis demonstrated a dumbbell structure (Martin et al. 2003). The N-terminal consist of a highly variable N-terminal domain and a conserved coiled coil (CC) domain. The C-terminal has an RNA recognition motif (RRM) and a C-terminal domain. The C-terminal domain also contains residues which have a high affinity for RNA binding and nucleic acid chaperone activity (Martin, 2010). Upon translation, ORF1 proteins form a coiled trimer complex along the variable coiled coil domain. (Martin, 2006, Hohjoh and Singer, 1996). L1orf1p trimer complex has a cis preference and binds with its own transcript and ORF2p to form a L1 RNA particle (Wei et al., 2001). Both RRM and CC domains are important for the RNA binding function of L1orf1p (Khazina et al., 2011, Martin et al., 2005).

L1orf1p structure

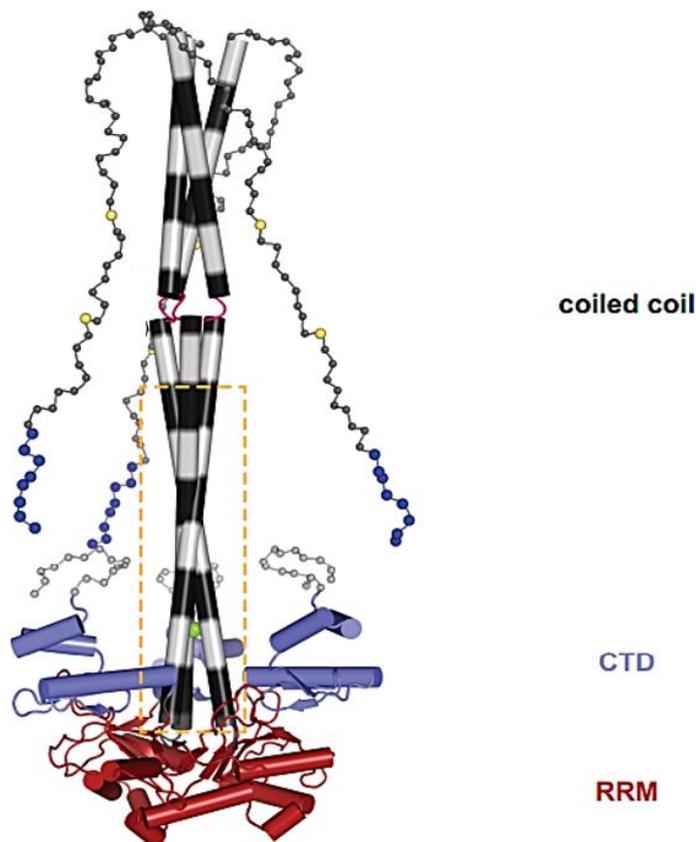


Figure 4.1 Crystal structure of L1orf1p (blue spheres= positive N-terminal) Khazina et al., 2018)

L1orf1p does not contain any enzymatic domain but has several phosphorylation sites that can regulate its function. These sites are predominantly serine and threonine phosphorylation sites. Some of these sites are docking motifs for proline-directed protein kinases (PDPK) such as cyclin-dependent kinases, mitogen-activated protein kinases and glycogen synthase 3. PDPK can interact with several cellular pathways including cell differentiation, cell division, inflammation and cancer, and specifically phosphorylate serine or threonine residues with a proline residue at position +1 (S/T-P motifs). L1orf1p contains four S/T-P motifs: S18P19, S27P28, T203P204, and T213P214. Moreover, it has several PDPK docking motifs as well such as T241 and T250. These two sites are protein kinase A sites. All 6 sites are highly conserved and are essential for retrotransposition, as substitution mutation of a non-phosphorylated residue can reduce L1 retrotransposition to 0-40% (Cook et al., 2015). In addition, T203 and T213 flank three conserved arginines (R206, 210 and 211) that may mediate RNA binding (Khazina and Weichenrieder, 2009, Khazina et al., 2011). T203 or T213 phosphorylation can lead to the formation of hydrogen bonding with the guanidino group of the three arginines, which in turn affect L1orf1p RNA

binding (Mandell et al., 2007). PDPK motifs are also regulated by a number of other factors such as Pin1. Pin1 is a prolyl isomerase and induces proline isomerization in phosphorylated S/T-P motifs (Lu et al., 2002a, Liou et al., 2011). S18 and S27 are key binding sites for Pin1 and protects the phosphorylation state by inhibiting cis trans prolyl-sensitive phosphatases. *In vitro* studies were performed by inducing PDPK mutants and revealed that phosphorylation by PDPKs were not required for RNA binding and did not affect the structural configuration of L1orf1p for its function (Cook et al., 2015).

4.12 L1orf1p and its interaction with different cellular pathways

L1orf1p can interact with several cellular pathways, which may promote carcinogenesis. For example, Lu et al. 2013 et al have demonstrated an interaction between the coiled-coil domain of L1orf1p and cytoplasmic androgen receptor in prostate cancer cells leading to its translocation to the nucleus. Androgen receptor translocation increases the expression of proliferating and anti-apoptotic genes such as vascular endothelial growth factor (VEGF) (Lu et al., 2013). L1orf1p can also increase E26 transformation-specific sequence-1 (ETS-1) transcriptional activity in colon and breast cancer (Bu et al., 2006). ETS-1 is an important transcription factor during development and carcinogenesis. It acts downstream of the c-Met signalling pathway and activates several proliferating and invasive genes such as MMP-1, MMP-9, c-Met, Cyclin D1 and u-PA (Bu et al., 2006) (Yang et al., 2013).

L1orf1p can also induce drug resistance through several pathways. For example, L1orf1p overexpression in HepG2 cells leads to development of resistance to epirubicin and cisplatin compared with empty vector control cells. In the same study, HepG2 L1orf1p siRNA had reduced IC₅₀ for epirubicin, cisplatin and paclitaxel by up to 9 fold (Feng et al., 2013). The increased drug resistance was likely due to increased BCL-2 (anti-apoptotic) expression in L1orf1p overexpression (Feng et al., 2013). Likewise, an immunoprecipitation analysis demonstrated an interaction between L1orf1p and cisplatin-resistance associated proteins CROP or LUC7L3 (Slotkin and Martienssen, 2007). L1orf1p can also increase ATP-dependent efflux pump gene expression in esophageal squamous cell carcinoma, and thus inducing doxorubicin and paclitaxel drug resistance (Zhu et al., 2015). L1 is also shown to affect TGF- β signalling. An immunoprecipitation analysis has shown that L1orf1p forms a complex with Smad4. The complex inhibits Smad4 nuclear translocation in HepG2 cells (Zhu et al., 2013). Furthermore, human bronchial epithelial cells treated with TGF- β had increased ORF1 and ORF2 transcript expression. EMT

markers such as vimentin was increased, and epithelial marker such as E-cadherin was reduced. L1orf1p expression was then inhibited using a siRNA transfection but TGF- β pathway remained unaltered. Thus, L1 may act downstream of the TGF- β pathway in these immortalised cells (Reyes-Reyes et al., 2017). However, we have observed reduction in TGF- β signalling and several associated targets such as reduced migration, invasion and reduced drug sensitivity upon L1 knockdown (**Chapter 3**). The difference could be cell type specific or difference between transformed and non-transformed cells.

4.2 Aims

Here in this chapter, the influence of L1 overexpression on TGF- β signalling pathway was investigated, specifically L1-full length and L1orf1p overexpression and its correlation with the TGF- β signalling pathway.

Specific aims were as follows:

- Full length L1 overexpression in liver cancer cell lines were used to investigate the effect of L1 on TGF- β signalling pathway
- L1orf1p overexpression cell lines were developed to investigate the effect of L1orf1p on the TGF- β signalling pathway

4.3 Methods

4.3.1 Cell culture

All cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS). HHL5 was grown in DMEM-high glucose media supplemented with 10% FBS, 1% L-glutamine and 1% non-essential Amino Acid Solution. Cells were incubated at 37°C and 5% CO₂, and sub-cultured every 6 days.

4.3.2 L1orf1p doxycycline (dox) inducible transfection

HepG2, PLC/PRF5 and HHL5 were seeded in a 6 well plate at 4×10^5 per well. The next day, cell media was removed from cells and replaced with new media. The plasmid transfection mixes were prepared as follows: 200µl serum free media, 1µg PB-dox empty or PB-dox L1orf1p, 0.5µg transposase plasmid and 4.5µl LT1 transfection reagent and incubated for 30 minutes at room temperature. Cells were transfected with these plasmids and incubated for 48 hours at 37°C and 5% CO₂. After 48 hours, new media with 2µg/ml puromycin (HHL5: 0.3 µg/ml puromycin) was added to all cell lines. Puromycin and media were replaced every 3 days.

4.3.3 SMAD3 luciferase assay

Huh-7 and HepG2 were seeded onto a 24well plate at 1×10^5 and 0.5×10^5 per well in triplicates. The next day, cell media was removed from cells and replaced with new media. The plasmid transfection mixes were prepared as follows: 200µl serum free media 500ng pcDNA (control plasmid) or L1 coneo (L1 overexpression plasmid), 400ng Pai1-promoter-luc reporter plasmid (containing SMAD3 binding sites), 40ng Renilla plasmid and 1.5µl LT1 transfection reagent. After 20 minutes incubation at room temperature, plasmid mix was transfected onto cells and incubated for 48 hours at 37°C. After 48 hours, Dual-Luciferase Reporter Assay System (Promega) was performed as per manufacturer's instructions.

4.3.4 FACS-dual L1orf1p and vimentin staining

Cells were harvested from a 70% confluent T75 flask with trypsin, and then fixed in 1% formalin (diluted in PBS) overnight. Following fixation, cells were washed with wash-perm buffer (BD bioscience) twice at 300g for 7 minutes, and stained with L1orf1p clone 4H1 1:4000 (Merckmillipore) for 1 hour at room temperature. Cells were then washed twice with wash-perm buffer and incubated with anti-mouse secondary antibody Alexfluor 488, 1:8000 (Thermofisher) for 1 hour. After primary incubation, cells were washed with wash perm buffer and then incubated with Alexfluor647 conjugated vimentin antibody (sc6260, Santa Cruz Biotechnology) 1 in 50 dilution for 1 hour at room temperature. Cells were then washed with wash-perm buffer and PBS and then analysed on the FACS Calibur (BD Biosciences). Gatings were adjusted relative to controls, cells stained with 5µl of isotype control (Alexfluor 647, Santa Cruz Biotechnology) only.

4.3.5 Plasmids

- Transient L1 overexpression was performed using a codon-optimised full length L1 plasmid (pBSKS-L1-CO, gift from Dr Jose Luis Garcia-Perez, Edinburgh University, UK)
- L1orf1 conditional overexpression plasmid was developed by subcloning L1orf1 from pBSKS-L1-CO plasmid under a dox-on promoter into a piggybac vector containing a puromycin resistance cassette by Dr Ruchi Shukla.
- Reverse transcriptase mutant plasmid L1.3-D702F (mutant for L1 reverse transcription domain) and L1.3 wildtype (control plasmid) also contained blasticidine-based retrotransposition reporter cassette and were kind gifts from Dr Jose Luis Garcia-Perez, Edinburgh University, UK.

4.3.6 Statistics

Statistical analysis was performed using SPSS (Version 24) and GraphPad Prism (Version 8).

4.4 Results

4.4.1 Full length L1 overexpression activates TGF- β signalling

The effect of L1 overexpression was investigated in HepG2, PLC/PRF5 and Huh-7 cell lines. All the cell lines were transfected with a control (pcDNA) and full L1 plasmid (L1coneo).

HepG2 had a 70.5% increase, Huh-7 had a 89.3% increase and PLC/PRF5 had a 56.4% increase in SMAD3 luciferase signal upon transfection with L1coneo relative to pcDNA control transfected cells (**Figure 4.2**). Although, these cell lines have different endogenous L1 expression - HepG2 and PLC/PRF5 have lower L1orf1p expression compared to Huh-7 (**Figure 3.5, Chapter 2**) still L1 overexpression increased SMAD3 luciferase activity in all three cell lines.

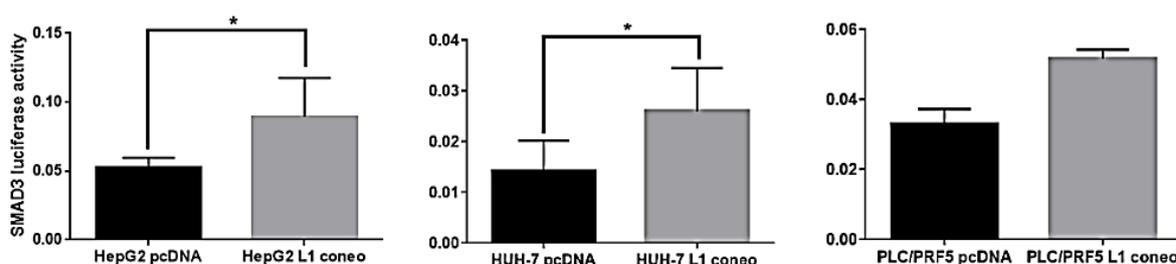


Figure 4.2 Effect of L1 overexpression on TGF- β signalling in Huh-7 (n=4), HepG2 (n=4) and PLC/PRF5 (n=1). Cells transiently co-transfected with pcDNA (control) or L1 coneo (L1 overexpression), with PAi1-promoter-luc and renilla plasmids were used to measure SMAD3 activity. Values represent mean \pm standard error either from 4 individual experiments (Huh-7 and HepG2) or 4 technical repeats (PLC/PRF5). * $p < 0.05$, Unpaired student-test.

Further analysis was performed by staining cells with L1orf1p and TGF- β downstream target vimentin. PLC/PRF5 were transiently transfected with pcDNA and L1 coneo for 48 hours and then stained for L1orf1p and vimentin. The dual staining demonstrated that about 50% of PLC/PRF5 cells transfected with control plasmid were vimentin positive with no L1orf1p positivity while all cells overexpressing L1orf1p (within the population transfected with L1coneo) were also positive for vimentin (**Figure 4.3**). Likewise, full length L1 overexpression increased vimentin expression in HepG2 cells compared to control indicating L1 increases TGF- β signalling (**Figure 4.4A**).

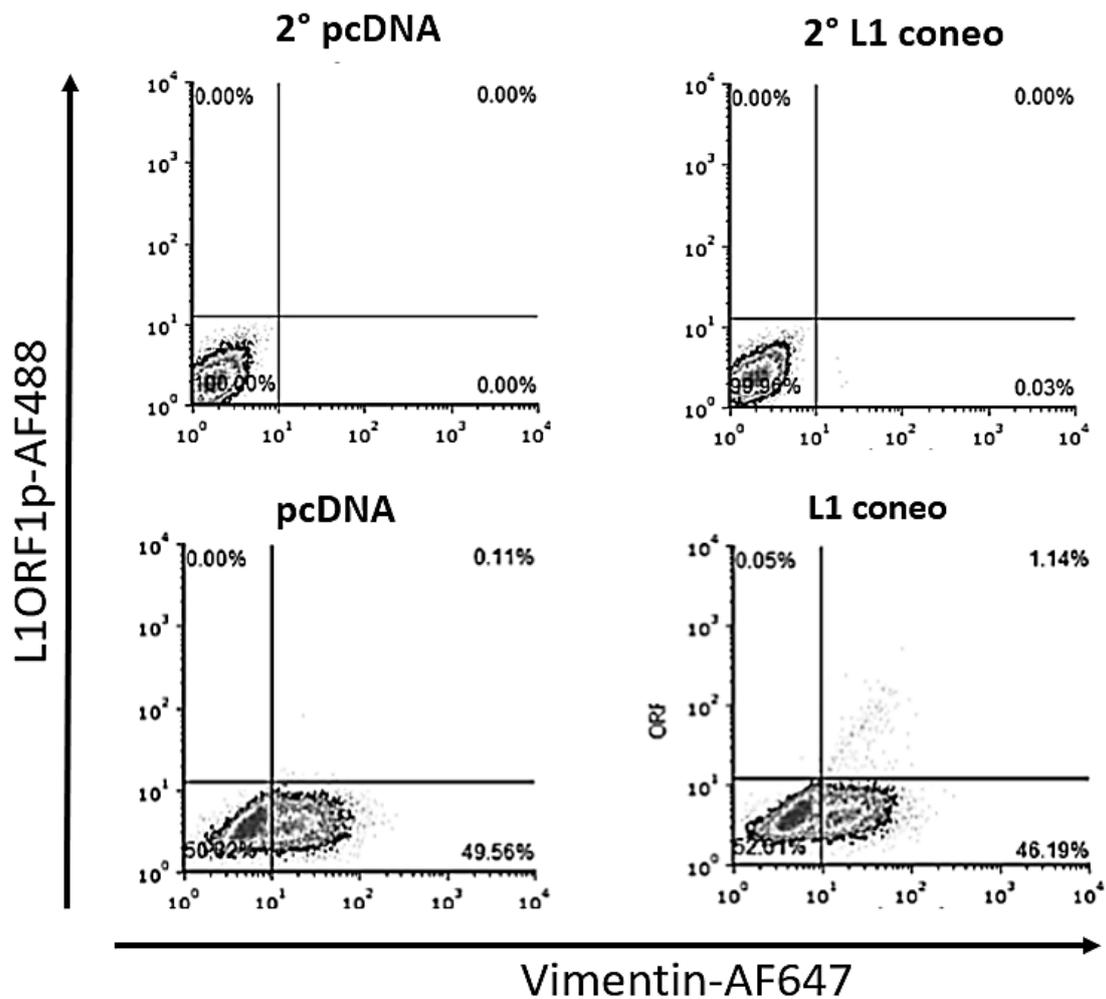


Figure 4.3 L1orf1p and vimentin cell distribution in PLC/PRF5 pcDNA and L1 coneo. Cells were dual stained for L1orf1p and vimentin and analysed by FACS staining. Gatings were adjusted relative to controls, 2° (Cells stained with secondary antibody only). (n=1).

Further studies were performed to investigate the importance of L1orf2p reverse transcriptase activity for the mechanism of L1's influence on TGF- β signalling. HepG2 were transfected with L1.3 wildtype (wt) or D702F mutant (D702F mutation inhibits reverse transcriptase activity of L1orf2p thus hinders L1 retrotransposition) (**Figure 4.4B**). TGF- β signalling was analysed by measuring its downstream target vimentin. Both plasmids demonstrated an equivalent increase in vimentin expression and thus L1 interaction with TGF- β is not dependent upon L1 reverse transcriptase activity and active retrotransposition process.

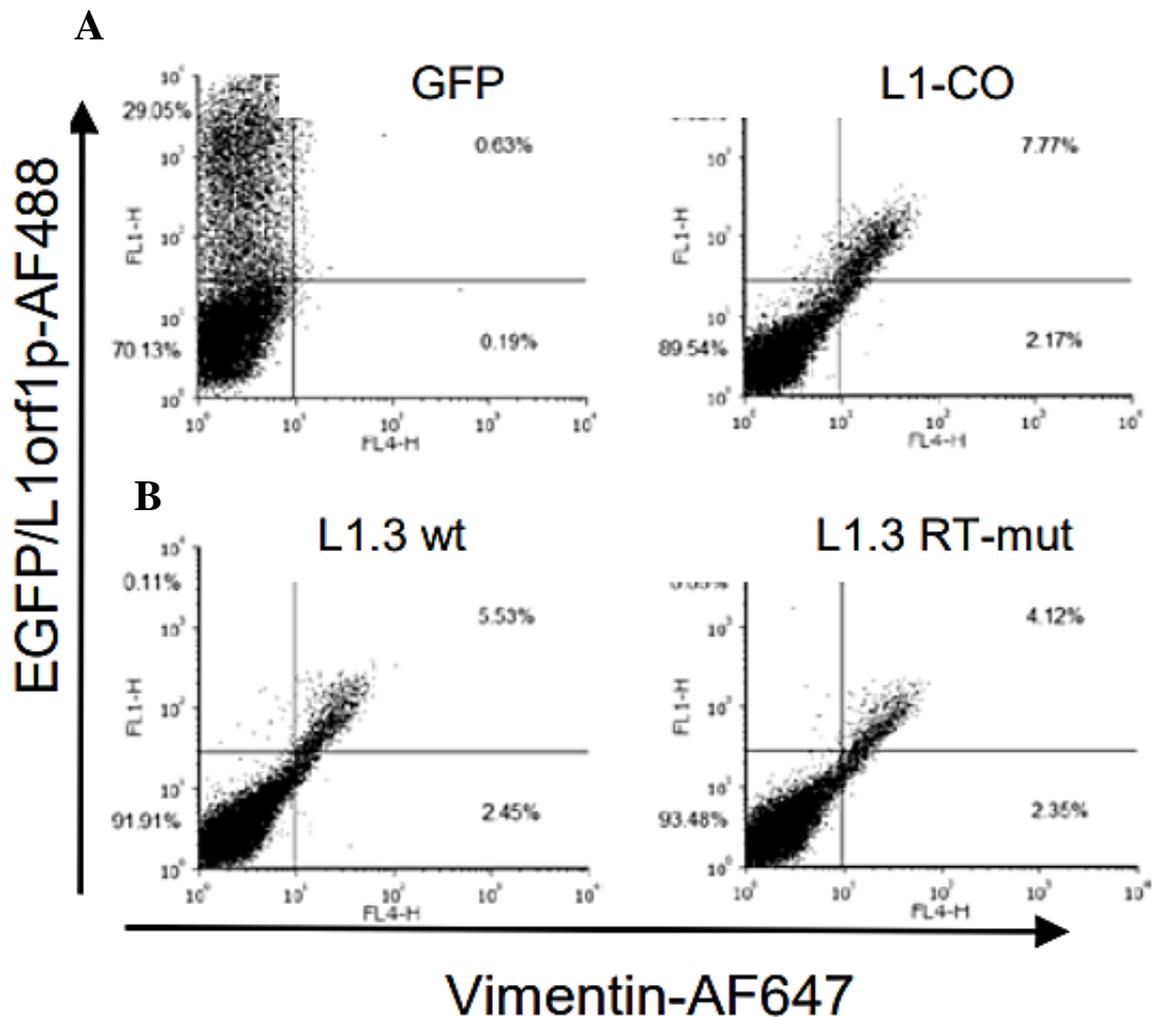


Figure 4.4 L1 overexpression and TGF- β signalling (A) HepG2 were transfected with a GFP (control) or L1coneo (L1 overexpression) plasmid and incubated for 48hours. FACS dual staining was performed for L1orf1p and vimentin. **(B)** HepG2 were transfected with a L1.3 wt (L1 overexpression plasmid) or L1.3 RT-mut (mutant for L1 reverse transcription) for 48 hours. FACS dual staining was performed for L1orf1p and vimentin.

4.4.2 L1 increases TGF- β signalling via L1orf1p

As L1orf2p reverse transcriptase activity is not essential for L1 mediated activation of TGF- β signalling, further studies were performed to investigate the role of L1orf1p in modulating TGF- β signalling. An L1orf1p stable cell line was developed using a piggybac system where L1orf1p was expressed under control of an inducible doxycycline promoter. As HepG2 had the lowest L1orf1p expression amongst the analysed liver cancer cell lines (HepG2, Hep3B, SK-Hep1, HUH-1, Huh-7, SNU182, SNU475, and PLC/PRF5), HepG2 cells were transfected with a control (HepG2 dox-empty) or L1orf1p plasmid (HepG2 dox-

L1orf1p). Puromycin selection was performed for at least 2 weeks before confirming L1orf1p overexpression both at transcript and protein level. Transcript analysis demonstrated a dose-dependent increase of L1-orf1 in HepG2 dox-L1orf1 cells with dox induction for 48 hours (**Figure 4.5**). Likewise, western blot demonstrated similar findings (**Figure 4.6**). No L1orf1 transcript and protein expression was observed in HepG2 dox-control.

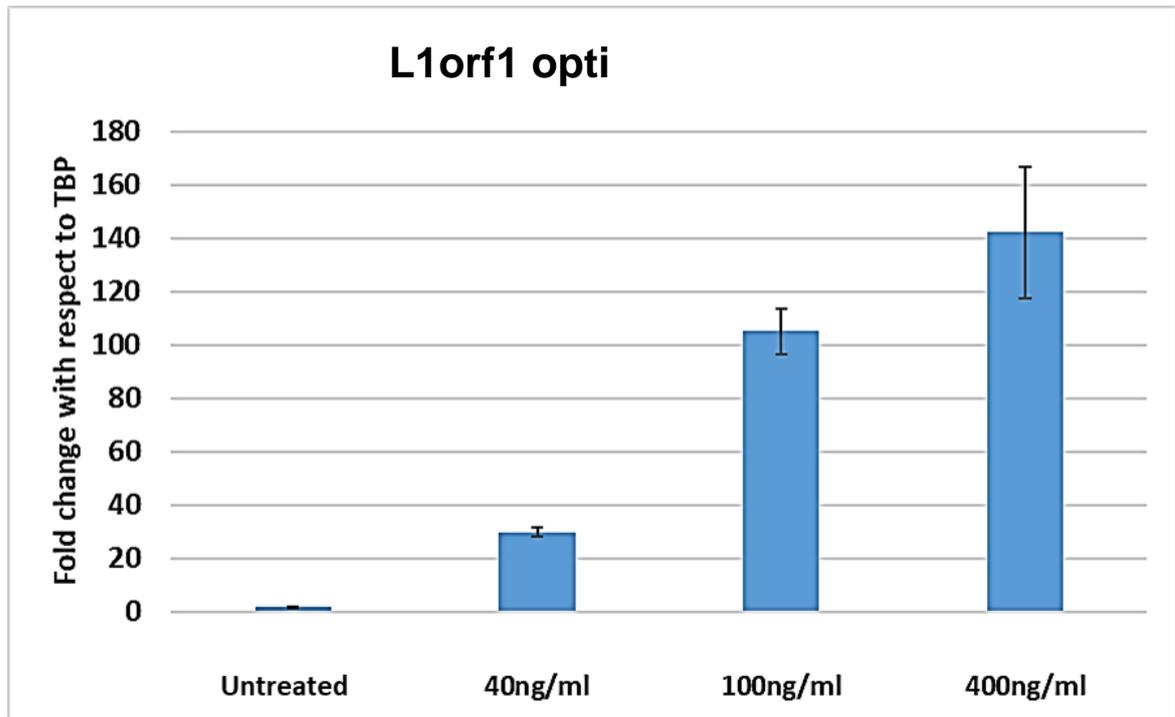


Figure 4.5 L1orf1p codon optimised transcript expression in HepG2 dox-L1orf1 with dox induction, mean: untreated, 40, 100 and 400ng/ml for 48 hours. HepG2 dox-control untreated and 400ng/ml had no L1orf1 transcript expression. Error bars represent standard error from three technical repeats.

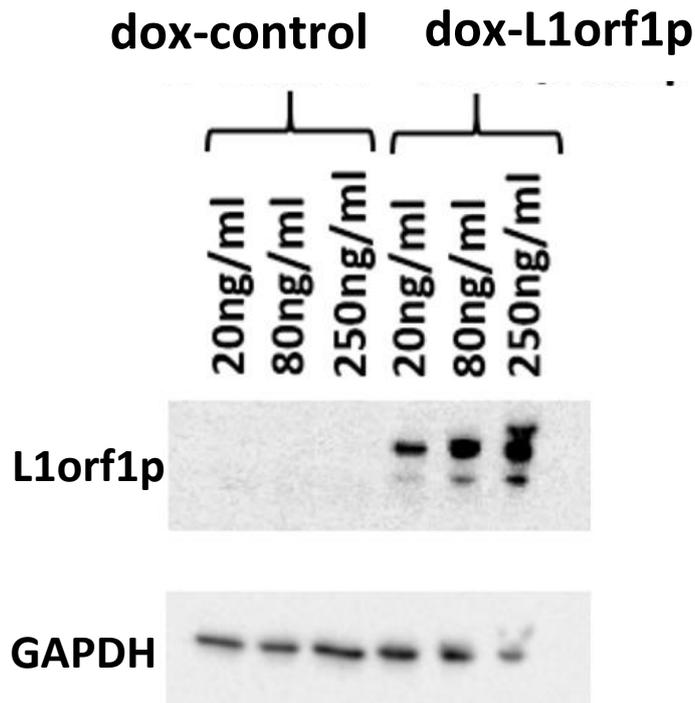


Figure 4.6 L1orf1p expression in HepG2 dox-control and HepG2 dox-ORF1p with dox induction for 48 hours. Protein expression was analysed by western blotting. GAPDH was used as a loading control.

FACS analysis was performed to analyse L1orf1p positive cell distribution in HepG2 dox-control and HepG2 dox-ORF1p. HepG2 dox-control were induced with the following dox doses: untreated and 400ng/ml; and HepG2 dox-L1orf1p: untreated, 40ng/ml, 100ng/ml, 400ng/ml. After 48 hours induction, cells were isolated and stained by FACS (**Figure 4.7**). HepG2 dox-L1orf1p demonstrated a dose dependent increase of L1orf1p positive cells reaching 65% positive at 400ng/ml, while no induction was observed in HepG2 dox-control. Furthermore, dual staining was performed for L1orf1p and TGF- β downstream target vimentin (**Figure 4.8**). The dual staining demonstrated increased vimentin expression in cells with positive L1orf1p expression indicating increased TGF- β signalling

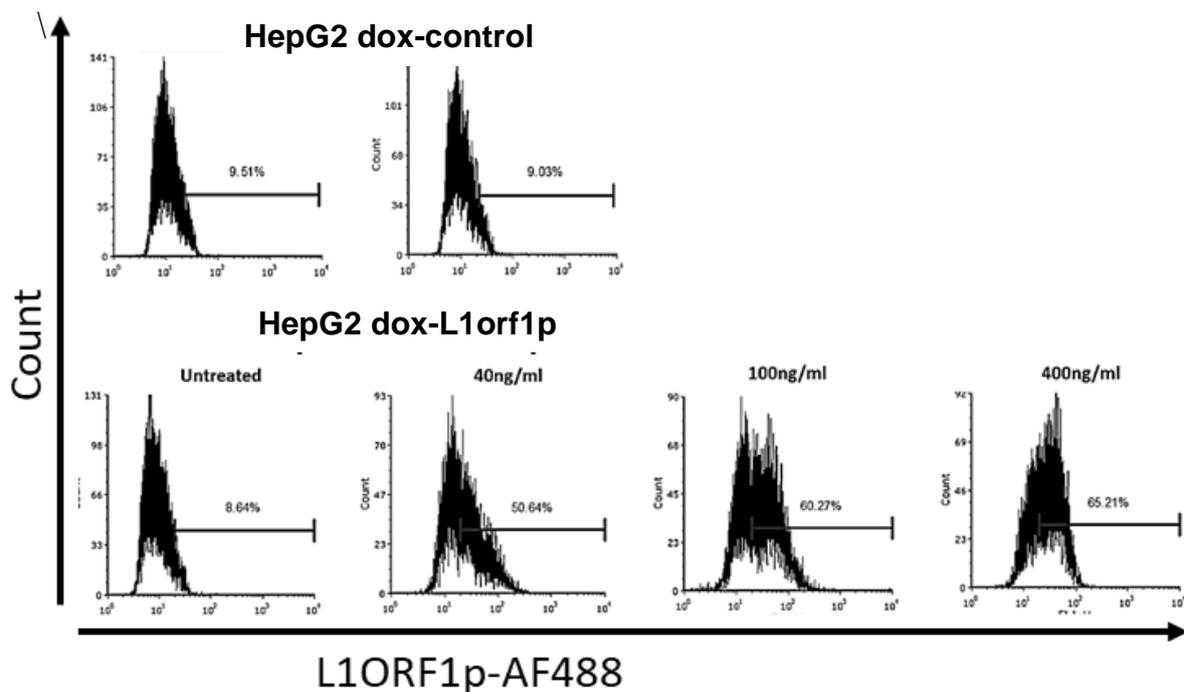


Figure 4.7 L1orf1p FACS staining in HepG2 dox-control and HepG2 dox-L1orf1p with the following dox inductions: untreated, 40ng/ml, 100ng/ml and 400ng/ml for 48 hours. Gatings were adjusted relative to controls (Cells stained with secondary antibody only). Percentage of cells and geometric mean fluorescence intensity were measured (n=1).

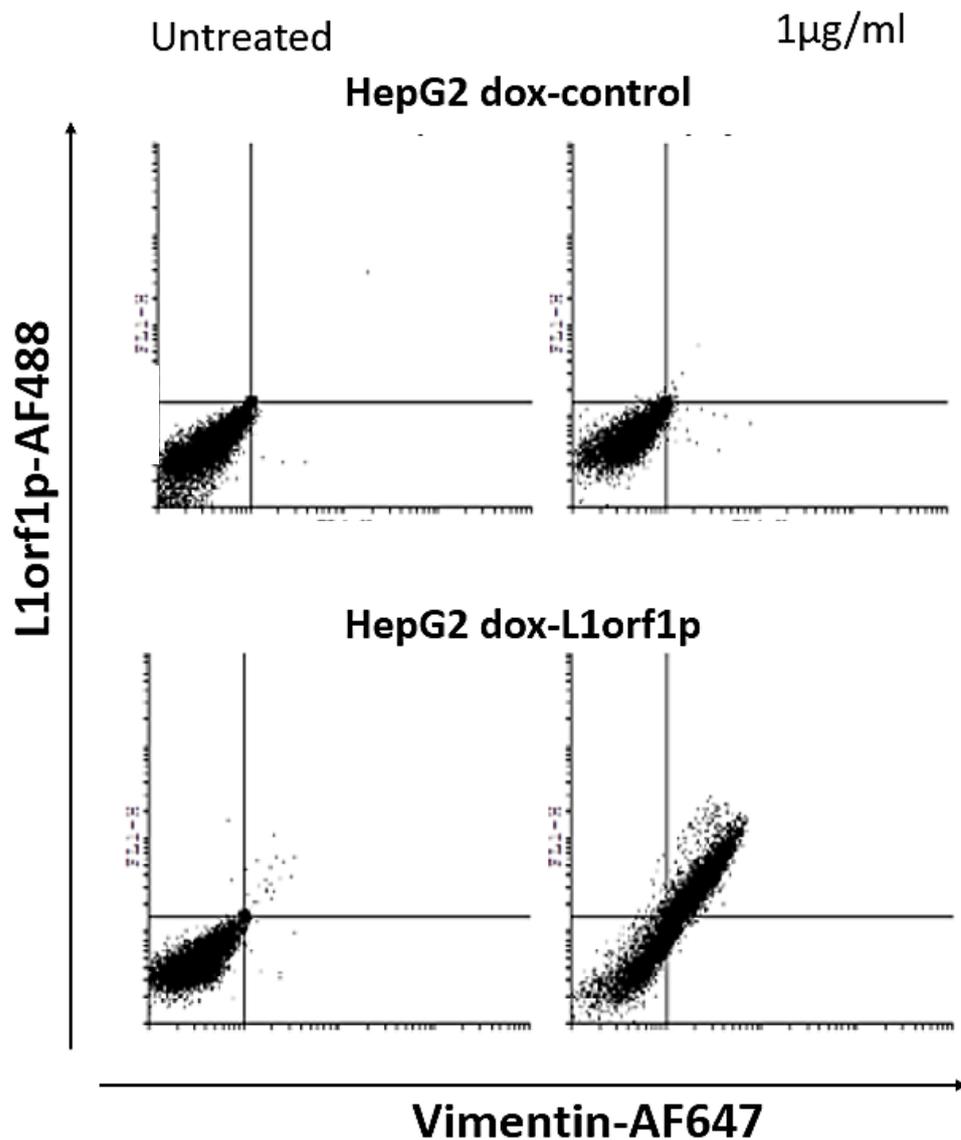


Figure 4.8 L1orf1p and vimentin cell distribution in HepG2 dox-control and L1orf1p overexpression. Cells were dual stained for L1orf1p and vimentin and then analysed by FACS staining. Gatings were adjusted relative to controls (Cells stained with secondary antibody only). Percentage of cells and geometric mean fluorescence intensity were measured.

The finding was further confirmed by developing inducible L1orf1p overexpression cell lines in PLC/PRF5 and HHL5 cells in a similar manner as HepG2 cells. As expected, PLC/PRF5 dox-L1orf1 exhibited dox-inducible L1orf1p expression with no basal L1orf1p expression in control cells however, PLC/PRF5 dox-L1orf1 cells exhibit basal L1orf1p expression (**Figure 4.9A**). This could be due to leaky expression of the promoter in these cells. The leaky expression of L1orf1p was also confirmed by FACS (**Figure 4.9B**). SMAD3 protein expression was also measured in both the cell lines by western blotting. In PLC/PRF5 dox-control untreated and induced, SMAD3 protein expression were similar.

However, SMAD3 protein expression was significantly higher in PLC/PRF5 dox-L1orf1p compared to control cells, which increased further upon dox induction (**Figure 4.9A**).

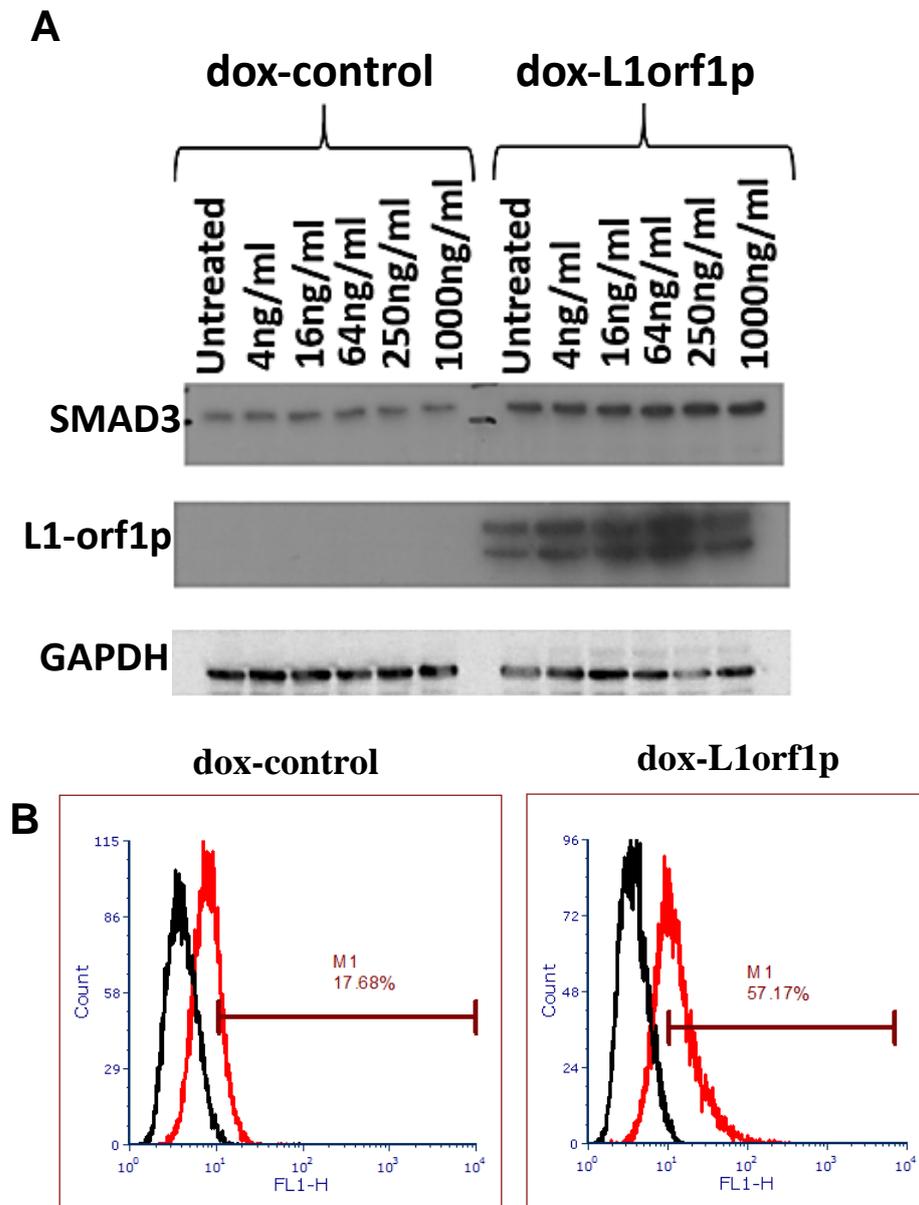


Figure 4.9 SMAD3 and L1orf1p protein expression in PLC/PRF5 dox-control and L1orf1p with dox induction for 48 hours. **(A)** Protein expression was analysed by western blotting. GAPDH was used as a loading control. **(B)** L1orf1p FACS staining in PLC/PRF5 dox-control and L1orf1p without dox induction. Black lines represent secondary staining only and red line represent primary and secondary antibody staining. L1orf1p expression was determined by setting the gate at 0.5% of secondary only cells (n=1).

As PLC/PRF5 dox-L1orf1p had increased SMAD3 protein expression, the TGF- β downstream target vimentin was measured by FACS staining (**Figure 4.10**). However, no significant increase in Vimentin intensity was observed in either PLC/PRF5 dox-control or

PLC/PRF5 dox-L1-ORF1p cells between vimentin only positive versus L1orf1p and vimentin double positive populations (PLC/PRF5 dox-L1orf1p: vimentin only=36.7; L1orf1p and vimentin: 45.0; PLC/PRF5 dox-L1orf1p: vimentin only= 32.9; L1orf1p and vimentin=35.6). This indicates that the leaky levels of L1orf1p is not sufficient to upregulate vimentin expression in the cells.

Even HHL5 dox-L1orf1p demonstrated leaky L1orf1p expression in about 30% of cells (**Figure 4.11**). Interestingly, vimentin geometric mean intensity demonstrated an increase in vimentin expression in L1orf1p and vimentin double positive population compared to vimentin only positive cells. **Control:** vimentin only=76.8, L1orf1p and vimentin: 119.1, **L1orf1p:** vimentin only= 80.4, L1orf1p and vimentin=103.3. Thus, L1orf1p upregulates vimentin expression in HHL5 cell line (derived from normal human hepatocytes and has been immortalised with moloney’s mouse leukaemia virus (MMLV) expressing E6E7 oncoproteins (Clayton et al., 2005)).

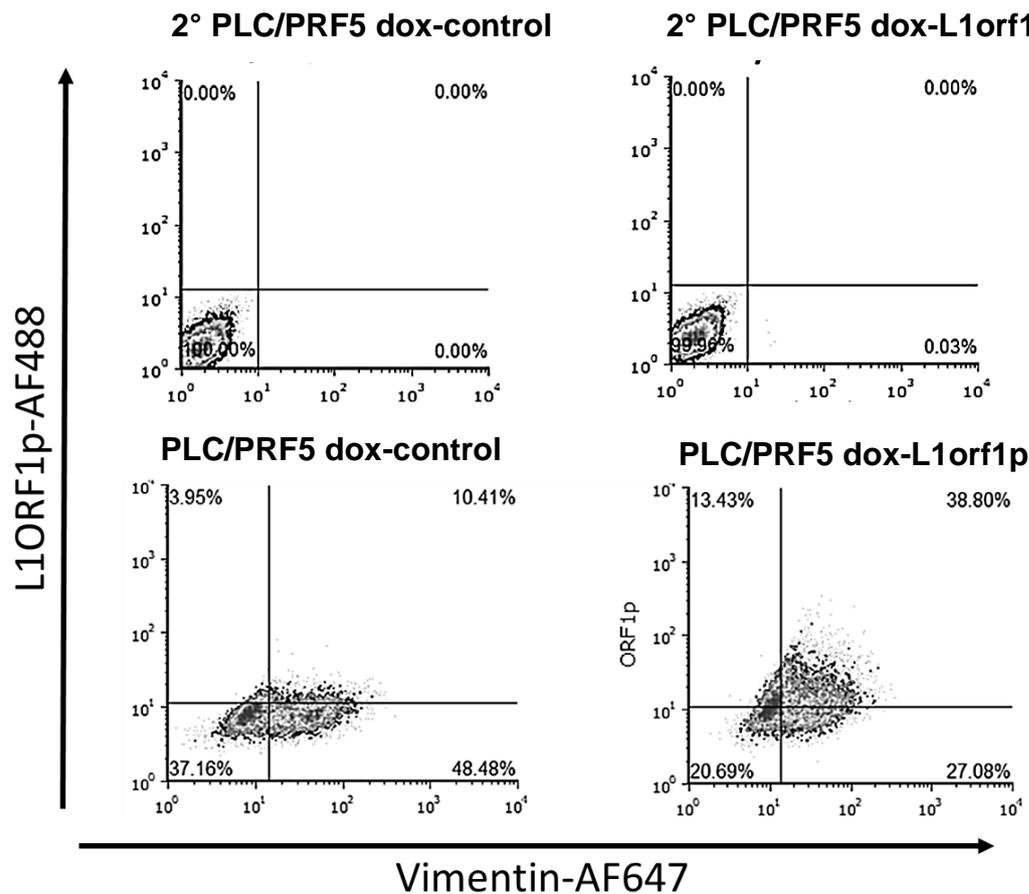


Figure 4.10 L1orf1p and vimentin cell distribution in PLC/PRF5 dox-control and L1orf1p overexpression. Cells were dual stained for L1orf1p and vimentin and then analysed by FACS

staining. Gating were adjusted relative to controls (Cells stained with secondary antibody only). Percentage of cells and geometric mean fluorescence intensity were measured.

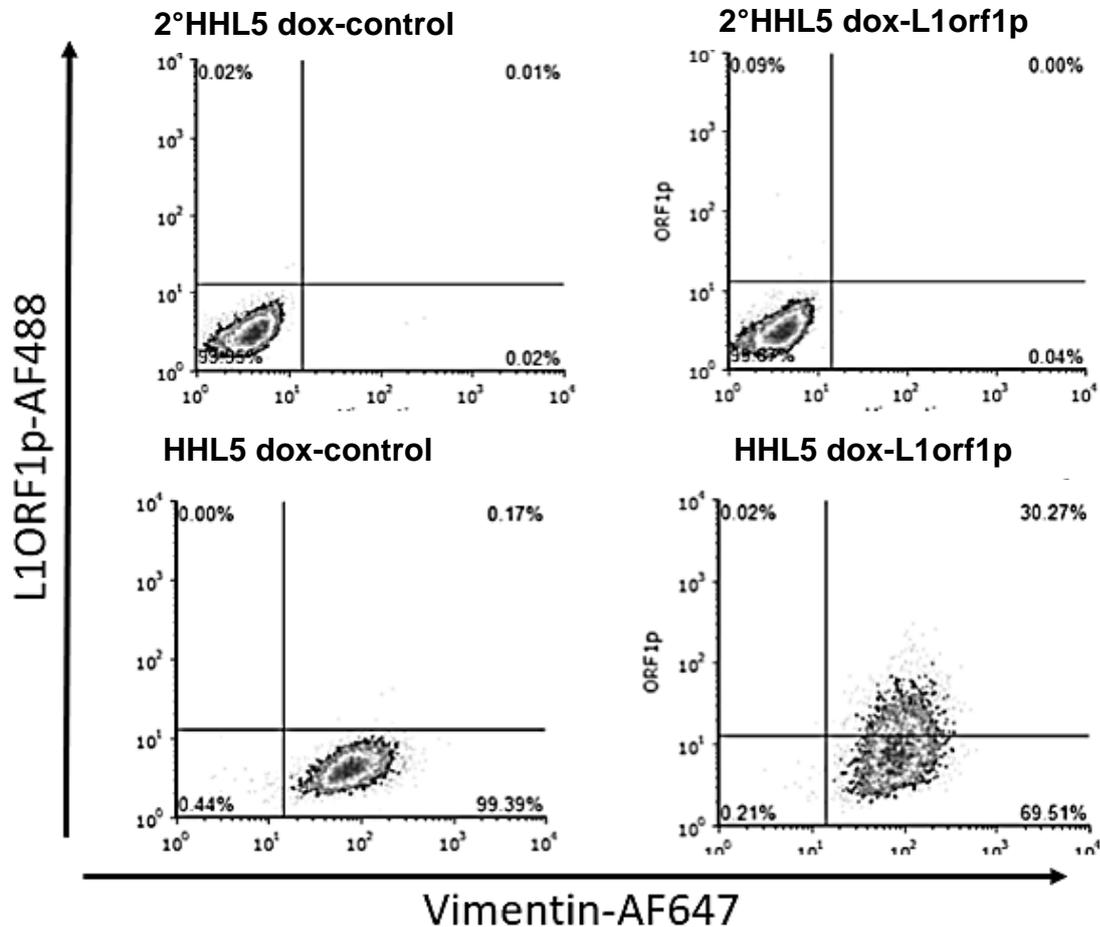
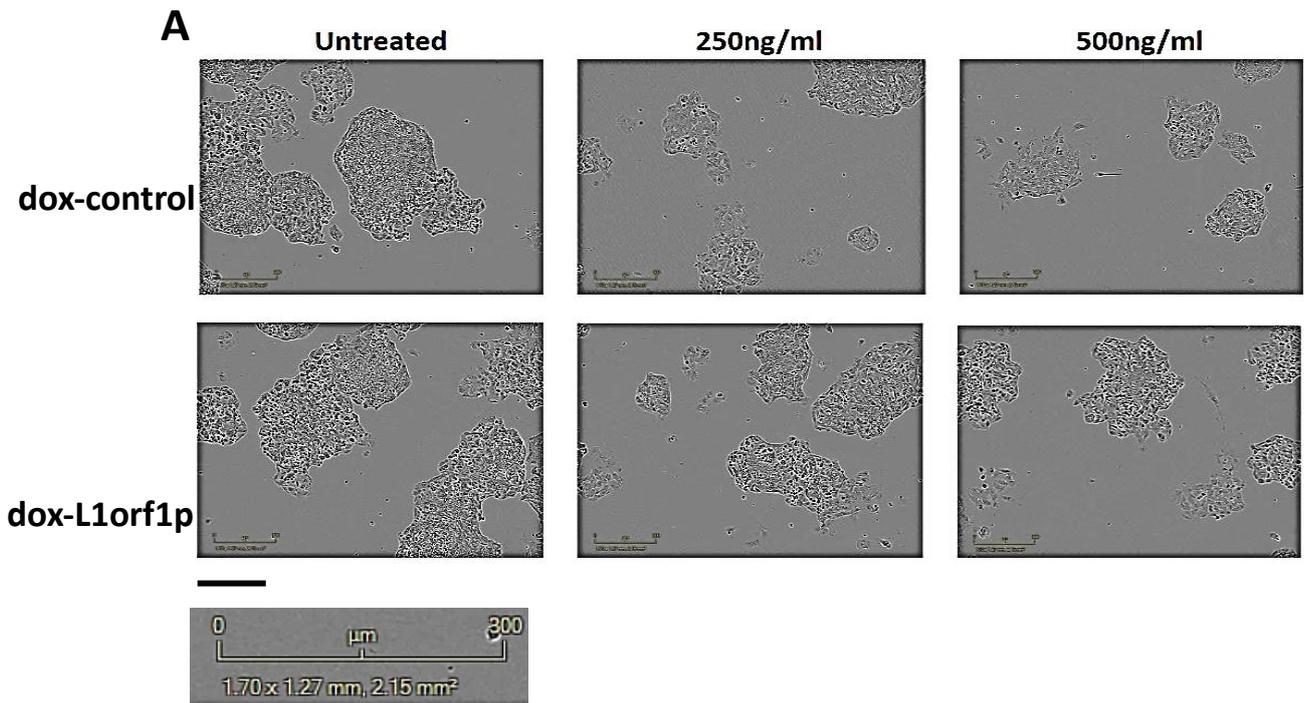


Figure 4.11 L1orf1p and vimentin cell distribution in HHL5 dox-control and L1orf1p overexpression. Cells were dual stained for L1orf1p and vimentin and then analysed by FACS staining. Gatings were adjusted relative to controls, 2⁰ (Cells stained with secondary antibody only). Percentage of cells and geometric mean fluorescence intensity were measured (n=1).

Subsequent studies were performed to investigate the effect of L1orf1p overexpression on cell morphology and proliferation. Morphology was assessed by Incucyte imaging (**Figure 4.12A**). HepG2 dox-control and HepG2 dox-ORF1p have both similar cell morphology in absence and presence of dox.

Cell proliferation was then measured by Incucyte (**Figure 4.12B**) and MTT assay (**Figure 4.12C**). Cells were seeded onto a 96well plate and induced with dox: no dox, 250ng/ml and 500ng/ml dox. Incucyte images were taken every 6 hours and analysed using an existing

cell algorithm. Dox induction reduced proliferation in both cell lines. However, the growth inhibition effect was higher in HepG2 dox-control compared to HepG2 dox-L1orf1p. MTT assay demonstrated similar findings in which HepG2 dox-control had reduced proliferation compared to HepG2 dox-L1orf1p after dox induction. Thus, L1orf1p overexpression overcomes the toxic effect of doxycycline and supports cell proliferation. Since, doxycycline treatment was toxic to cells no further functional assays were performed in HepG2 cells.



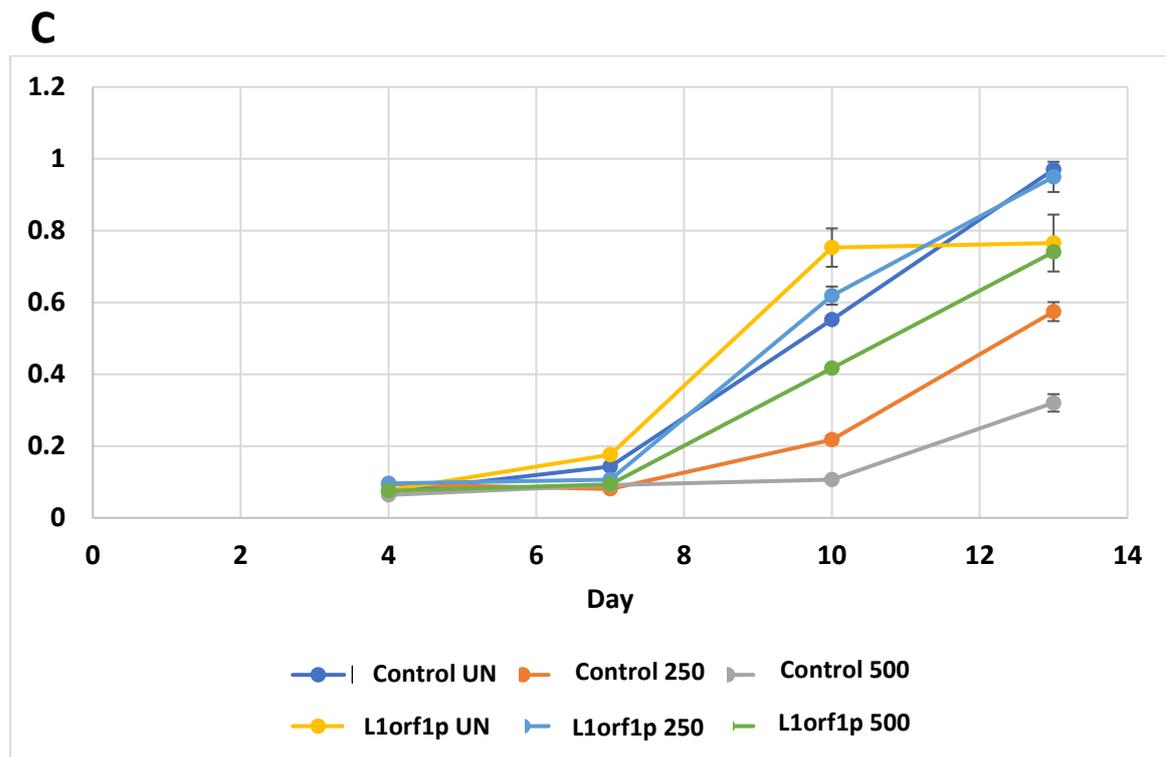
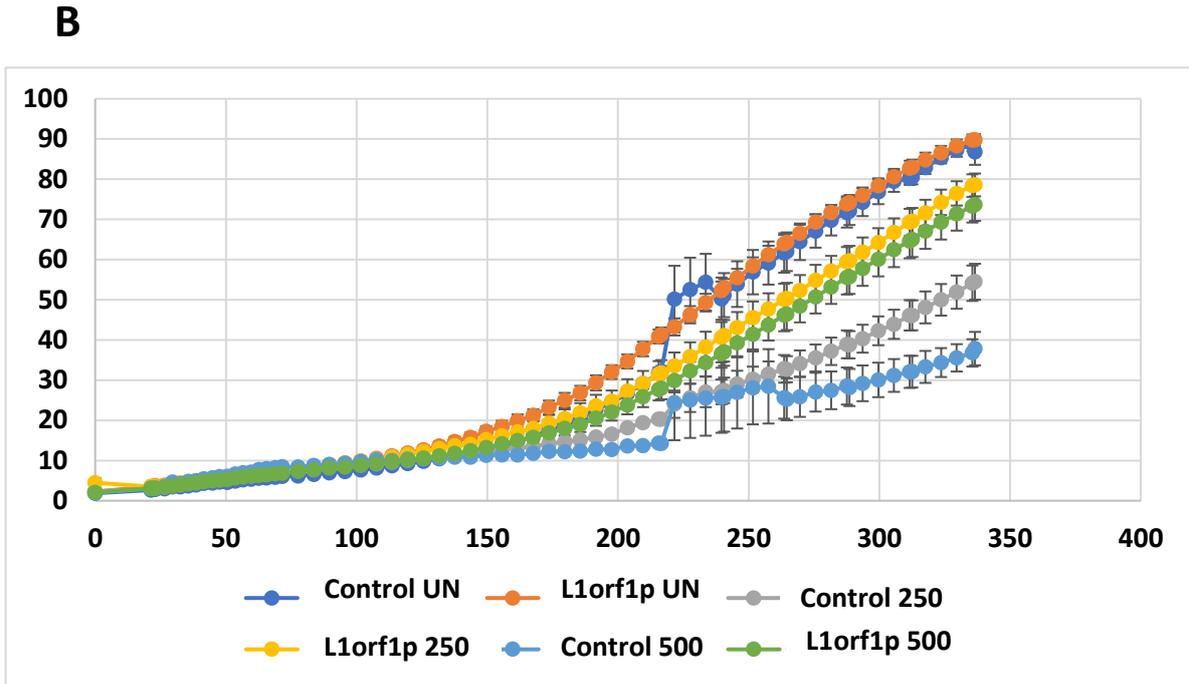


Figure 4.12 HepG2 dox-control and HepG2 dox-ORF1p with dox induction: untreated, 250ng/ml and 500ng/ml. (A) Cell morphology images were taken by Incucyte imaging. (B) Growth proliferation curve between HepG2 dox-control and HepG2 dox-L1orf1p with dox induction (ng/ml) using Incucyte image (average) (C) and MTT assay (average). Error bars represent standard error from five technical repeats (n=1).

4.5 Discussion

Initial studies were performed to overexpress full length L1 in HepG2, Huh-7 and PLC/PRF5 by transient transfection and its influence on the TGF- β signalling pathway was measured using Pai1 promoter (containing SMAD3 binding sites) luciferase reporter assay. All three cell lines exhibited an increase in Pai1-luciferase signal compared to control validating the potential of L1 to upregulate TGF- β signalling.

Furthermore, a meta-analysis has shown that TGF- β pathway converges on several pro EMT inducers such as transcription factors (TWIST1, TWIST2), transcriptional repressors SNAIL (SNAI 1) and SLUG (SNAI2) and ZEB1 and ZEB2. These inducers are associated with increased expression of mesenchymal markers such as vimentin and N-cadherin. Here FACS analysis demonstrated increased L1orf1p and vimentin positive cells indicating L1 increases TGF- β signalling leading to increased vimentin expression.

Further studies were performed to investigate the importance of L1 retrotransposition on TGF- β signalling. Both WT (L1.3 wildtype) and D702F mutant (L1 retrotransposition mutant) demonstrated an increase in vimentin and thus L1 interaction with TGF- β is L1 retrotransposition independent. The potential of L1orf1p for activating TGF- β signalling was then investigated in several cell lines (HepG2, PLC/PRF5 and HHL5) by developing conditional (DOX-ON) L1orf1 overexpressing cells. Functional studies demonstrated PLC/PRF5 and HHL5 both had leaky L1orf1p expression i.e. increased L1orf1p expression was observed in dox-L1orf1 cells without dox induction. PLC/PRF5 dox-L1orf1p demonstrated increased SMAD3 expression indicating increased TGF- β signalling however, no significant shift in vimentin expression was observed in L1orf1p overexpressing cells. However, upregulation of vimentin expression was observed in HHL5 dox-L1orf1p cells. Thus, L1orf1p can interact with the TGF- β pathway leading to increased vimentin expression.

The growth inhibition effect was higher in HepG2 dox-control compared to HepG2 dox-L1orf1p, indicating doxycycline toxicity is significantly higher in the dox-control vector. Several studies have observed similar doxycycline toxicity in HCC cell lines. In one study, HepG2 and PLC/PRF5 were treated with the following doses of doxycycline: untreated, 1000ng/ml, 5000ng/ml and 10000ng/ml. Cell growth was measured by MTT and doxycycline mediated growth suppression was observed as early as 3 days (Meng et al., 2014). Thus, L1overexpression may decrease doxycycline toxicity.

4.6 Summary

L1 overexpression increased TGF- β signalling in several liver cancer cell lines and associated with increased SMAD3 and vimentin expression. Furthermore, L1 retrotransposition inhibition using RT-mutant had no effect on the influence of L1 on TGF- β signalling. Thus, L1orf2p reverse transcriptase activity is not required for the L1-mediated TGF- β signalling activation. L1orf1p overexpression increased SMAD3 and vimentin expression. Hence, L1 overexpression increased TGF- β signalling, via L1orf1p.

Chapter 5: Discussion

L1 expression is elevated in several epithelial cancers and may have an essential role in carcinogenesis. L1 can support carcinogenesis by various means such as inhibition of tumour suppressor genes, and activation of oncogenes. L1 proteins can also interact with cellular proteins involved in various signalling pathways and support carcinogenesis. Currently, the role of L1 in HCC is unknown and requires further research. The purpose of this thesis was to investigate the role of L1 retrotransposons in HCC and to evaluate their potential as a biomarker and potential therapeutic target.

5.1 L1 transcript and protein expression in HCC and correlations with clinical parameters and signalling pathways

Whether L1 activation is a random and bystander event of tumourigenesis or whether it is linked with particular molecular pathways and can be used as a biomarker for tumour classification is still an open question. To address this, HCC patient samples were analysed for L1 expression at both transcript and protein levels. As expected, L1 transcript and protein expression were both upregulated in HCC compared to non-tumour tissues.

Transcript analysis was performed using a publically available RNAseq data of 372 HCC samples. The data is part of the cancer genome atlas liver hepatocellular carcinoma (TCGA-LIHC) study. TCGA consortium has further carried out a comprehensive analysis of a core set of 196 patients combining RNAseq, microRNA expression, DNA methylation, copy number variations, somatic mutations and reverse phase protein lysate microarray datasets and HCC samples were grouped into 3 iClusters (Wheeler et al, 2017). Likewise, the TCGA-LIHC data has been analysed specifically for TGF- β signalling pathway and identified distinct HCC groups either with activated, inactivated or unaltered TGF- β pathway (Chen et al., 2017), and immune signatures to classify the samples into immune subtypes (Sia et al., 2017). However, expression of repeat elements has not been explored. Hence, we analysed the data to obtain normalised L1Ta 5'UTR promoter (L1 5'UTR) counts and explored the correlations between L1 expression in HCC and clinical features of the patients and HCC molecular subclasses. Transcript analysis demonstrated a positive correlation with tumour pathological stage and vascular invasion. Vascular invasion has been linked with satellite nodule, intrahepatic metastases and portal vein obstruction leading to liver damage and failure (Shi et al., 2010). Furthermore, L1 expression correlated with high tumour recurrence (SNU cluster) and poor prognosis (HB16 cluster). These

findings suggest that elevated L1 contributes to tumour progression and poor prognosis in HCC patients. Similar findings were observed in other cancers. For example, in lung and colorectal cancer, L1 hypomethylation is associated with a poor prognosis (Saito et al., 2010, Saito et al., 2008). L1 hypomethylation was also associated with a poor prognosis in HCC and disease recurrence after resection (Gao et al., 2014b).

In terms of molecular subclasses, L1 expression was significantly elevated in iCluster 3 HCC molecular subclass implying higher chromosomal instability and TP53 mutation. L1 5'UTR counts correlated negatively with TP53 gene expression targets indicating reduced TP53 tumour suppressor function. TP53 is an important tumour suppressor protein and has a central role in several cellular responses such as cell cycle arrest, DNA repair and apoptosis. It can also activate other tumour suppressor and apoptosis related genes such as p21 and BAX. Low or mutated TP53 have been observed in many cancers and its mutations are linked to large tumour size and poorly differentiated tumours (Madden et al., 2002). Furthermore, increased L1 expression has been linked with TP53 mutation in other cancer types as well (Rodic et al., 2014, Jung et al., 2018). Recently, TP53 binding sites were observed in the L1 promoter and {Tiwari, 2020 #312} demonstrated that P53 binding to L1 promoter can lead to downregulation of L1 expression. However, (Harris et al., 2009) demonstrated a positive correlation between P53 and L1 transcription leading to cell apoptosis. The differences could be due to difference in experimental approaches or cell types.

L1orf1 protein level was analysed by staining HCC diagnostic formalin-fixed, paraffin-embedded (FFPE) biopsy samples by immunohistochemistry. Staining was mainly cytoplasmic, although some nuclear staining was also observed. These findings are in accordance with previous reports evaluating L1orf1p such as *in vitro* studies using retrotransposition assay and patient FFPE biopsies (Belgnaoui et al., 2006, Rodic et al., 2014). However, in some cancers L1orf1p nucleus observation has been observed. For example in invasive breast carcinoma, L1orf1p and L1orf2p nuclear localisation correlated with metastasis and poor survival (Chen et al., 2012). L1orf1p and L1orf2p nuclear localisation might be more prevalent in certain cancers and correlate with a worse prognosis.

Similar to transcripts data, a significant upregulation of L1orf1p expression was observed in HCC compared to corresponding non-tumour tissues. Previous studies have also observed elevated L1 staining in several cancers. Rodic et al., 2014 measured L1orf1p expression in several cancers (1027 patients) by tissue microarray staining. Staining was positive in 47% of the tumours (482 cases) and highly expressed in epithelial cancers such as breast carcinoma (97%, 66 of 68 were positive), high-grade ovarian carcinomas (93.5%, 29 of 31 were positive) and pancreatic ductal adenocarcinomas (89%, 56 of 63 were positive). High-grade tumours such as sarcomas, pancreatic carcinomas, lymphoma and secondary glioblastomas were also highly positive. Other cancers such as oesophagus, bladder, head and neck, colon, lung, endometrium and biliary tract were 22.6–76.7% positive for L1orf1p. In HCC, L1orf1p expression was positive in 24% of cases. In contrast, corresponding preneoplastic lesion had none or low L1 expression (Rodic et al., 2014). Likewise, Barchitta et al, reported high L1orf1p expression in different epithelial cancers such as liver, renal, ovarian, lung and prostate carcinoma (Barchitta et al., 2014).

Several clinical factors were assessed and compared between L1_low and L1_high HCC categories. Higher L1 expression in HCC was significantly associated with poorly differentiated tumours and poor median survival in patients with TACE therapy. L1orf1p expression positively correlated with differentiated HCC and elevated alpha fetoprotein (AFP) levels. AFP is a 70kDa glycoprotein and present at trace levels in adults. In response to injury, hepatocyte proliferate and release AFP into the circulation (Mizejewski, 2001). In addition, higher L1orf1p was observed in patients with cirrhosis in non-tumour tissues. Previous studies have shown that several HCC related aetiologies have late diagnosis of HCC. For example in NAFLD, 22.8% related HCC were detected by surveillance. In alcohol related HCC, it was 32% and in Hepatitis C it was 46.2% (Mittal et al., 2015). These data suggest that L1 might be a potential tissue biomarker particularly in early stage of cancer, but further studies are required to confirm these findings in a larger patient cohort.

In terms of signalling pathways, L1 transcript expression was significantly higher in HCC patients with activated TGF- β status compared to normal or inactive TGF- β signalling. The relationship was investigated in FFPE tissues at protein level, by carrying out IHC for pSMAD3 (as a surrogate for active TGF- β signalling). L1orf1p staining correlated with pSMAD3 confirming the positive relationship between L1 and TGF- β signalling. TGF- β is an important pathway in the liver both during normal functioning of the organ and

hepatocarcinogenesis. The role of TGF- β in HCC is complex and has both tumour suppressive and promoting functions. However, most carcinomas become refractory to the tumour-suppressive functions such as reduced growth. Its tumour promoting function can induce epithelial-mesenchymal transition (EMT). Our data suggest that L1 supports tumour progression by activating TGF- β signalling in the background of TP53 mutated HCC.

5.2 L1orf1p activates TGF- β signalling in liver cancer cell lines

In vitro studies were performed to investigate the role of L1 in liver cancer cell lines. First, L1 expression (at transcript and protein levels) was measured in 9 cell lines: HepG2, Hep3B, Huh-1, Huh-7, SNU182, SNU475, PLC/PRF5 (liver cancer cell lines); SK-Hep1 (hepatic adenocarcinoma cell line, endothelial origin) and HHL5 (human immortalised hepatocyte cell line). Similar to HCC tissues, the cell lines exhibited a range of L1 expression. Overall, cell lines having ‘late TGF- β signature’ had higher L1 expression than cell lines with ‘early TGF- β signature’. Late TGF- β signature is associated with higher migration and invasion compared to cell lines with an early TGF- β response (Coulouarn et al., 2008).

Amongst the epithelial cell lines, highest L1 expression was observed in Huh-7 both at transcript and protein levels thus this cell line was selected to generate L1 knockdown transgenic line using lentivirus-based shRNA construct targeting L1orf1p. Huh7-WT (wildtype), Huh7-NT (non-target shRNA control) and Huh7-L1KD (L1 knockdown) cells were compared to each other to study the effect of L1 knockdown on the cells. L1 inhibition had no effect on cell proliferation as such, however Huh7-L1KD cells formed larger colonies compared to Huh7-WT and Huh7-NT cells in a clonogenic assay thus indicating increased rate of proliferation during colony formation starting from single cells. Moreover, reduced cell migration and invasion were observed in Huh7-L1KD cells than control cell lines. *In vivo* experiments also demonstrated delayed tumour growth from Huh7-L1KD cells in mice compared to Huh7-WT and Huh7-NT cells. Several studies have observed similar findings in other cancers (Bu et al., 2006, Li et al., 2014, Yang et al., 2013b). For

example, L1 knockdown cells had reduced telomeres length and increased cell apoptosis (Aschacher et al., 2016).

RNAseq was performed between Huh7-NT and Huh7-L1KD to evaluate influence of L1 knockdown on cellular transcriptome. DeSeq2 analysis revealed 1512 differentially expressed (DE) genes (950 downregulated and 561 upregulated) at $\log_2FC0.5$, $padj < 0.05$. At \log_2FC1 , 334 genes were differentially expressed (242 downregulated and 92 upregulated). Further analysis was performed using GSEA to identify pathway differences. The analysis were performed in all differentially expressed genes with a $\log_2FC0.5$, $padj < 0.05$ (1512 genes). GSEA analysis using hallmark pathways and Keggs pathways both demonstrated reduced TGF- β signalling and downstream features such as EMT, focal adhesion and ECM were reduced in Huh7-L1KD cells compared to Huh7-NT cells.

The downregulation of TGF- β pathway upon L1 knockdown was further confirmed by functional assays. Transcript and protein analysis both confirmed TGF- β pathway downregulation in Huh7-L1KD cells especially downregulation of SMAD3. In addition, Pai1 luciferase reporter assay demonstrated downregulation of SMAD3 function in Huh7-L1KD compared to WT and Huh7-NT cells. In addition, EMT marker like vimentin had reduced expression in Huh7-L1KD compared to WT and Huh7-NT cells. Hence, basal TGF- β signalling pathway was reduced in Huh7 cells upon L1 knockdown. TGF- β pathway can also induce drug resistance such as doxorubicin resistance (Li et al., 2015). We observed higher G2 arrest in Huh7-L1KD compared to WT and Huh7-NT upon exposure to doxorubicin however, no significant difference in the IC50 dose of the drug was observed. Hence, the change in drug sensitivity was very subtle in an *in vitro* setting.

Induction with TGF- β increased SMAD3 transcript and protein expression in Huh7-L1KD cells, and thus the cells responded to the exogenous TGF- β stimulus similar to Huh7-WT and Huh7-NT cells in terms of upregulation of its targets like P15 and P21. A recent study has shown that P15 and P21 upregulation can downregulate EMT and induce cellular senescence (Senturk et al., 2010a). Here, TGF- β induction increased P15 and P21 expression in all three cell lines leading to inhibition of cell proliferation. Hence, eventhough TGF- β signalling pathway was downregulated in Huh7-L1KD cells, the response to TGF- β treatment remained intact in the cells.

As L1 inhibition reduced TGF- β signalling in Huh7 cells, the effect of L1 overexpression on the TGF- β pathway was measured in liver cancer cell lines with low endogenous L1

expression. Initial studies were performed to overexpress full length L1 in HepG2, Huh-7 and PLC/PRF5. All three cell lines had increased SMAD3 luciferase activity compared to control. Thus, confirming the causal relationship between L1 overexpression and TGF- β signalling. Since TGF- β can affect several pathways including EMT, which is associated with increased expression of mesenchymal markers such as vimentin and N-cadherin; we carried out FACS analysis to evaluate influence of L1orf1p expression on vimentin. Increased vimentin expression was observed on L1orf1p positive cells in HepG2 transfected with full-length L1 thus indicating L1 increases TGF- β signalling leading to increased vimentin expression. Further studies were performed to investigate the importance of L1 retrotransposition on TGF- β signalling. Since both WT (L1.3 wildtype) and D702F mutant (L1 retrotransposition mutant) demonstrated an increase in vimentin thus L1 interaction with TGF- β pathway is L1 retrotransposition independent.

The role of L1orf1p on TGF- β signalling was then investigated in several cell lines by developing conditional (DOX-ON) L1orf1p overexpressing transgenic lines. The overexpression cell line was successful and demonstrated increased L1orf1p transcript and protein expression with doxycycline (dox) induction. Functional studies demonstrated increased vimentin expression in L1orf1p positive population versus L1orf1p negative population in HepG2 dox-L1orf1p cells upon doxycycline induction thus confirming the positive influence of L1orf1p on TGF- β signalling. However, reduced cell proliferation in both HepG2 dox-control and HepG2 dox-L1orf1p was observed upon doxycycline induction. The growth inhibition effect was higher in HepG2 dox-control compared to HepG2 dox-L1orf1p, indicating doxycycline toxicity is significantly higher in the dox-control vector. Several studies have observed similar doxycycline toxicity in HCC cell lines (Meng et al., 2014). Reduction in doxycycline toxicity upon L1orf1p overexpression could be due to upregulation of MDR1 (ATP-dependent efflux pump) expression by L1orf1p (Zhu et al., 2015) or by its influence on TGF- β signalling pathway.

However, since doxycycline by itself is toxic to HepG2 cells we explored other liver cancer cell lines to identify cells with lower sensitivity to doxycycline-induced toxicity. PLC/PRF5 and HHL5 both demonstrated leaky L1orf1p expression i.e. increased L1orf1p expression in the absence of any induction in dox-L1orf1p cells compared to dox-control cells. PLC/PRF5 L1orf1p untreated had increased SMAD3 expression indicating increased TGF- β signalling compared to PLC/PRF5 dox-control untreated. Though, no increased vimentin expression was observed in dox-L1orf1 cells. However, increased vimentin

expression was observed in HHL5 dox-L1orf1p cells. Thus, L1orf1p can interact with the TGF- β pathway leading to increased vimentin expression.

Future studies

Future studies can be performed to improve the experimental design and expand the research further.

FFPE HCC samples were analysed for L1orf1p expression and higher level of expression was observed in tumour tissues compared to non-tumour tissues. More diverse samples can be analysed including normal liver, liver fibrosis and cirrhotic tissues to evaluate if L1 can get activated in preneoplastic liver as well or not ? In addition, a larger cohort of HCC FFPE samples with different aetiologies can be analysed to investigate whether L1 activation is more prevalent in certain aetiologies. In parallel, the study can be extended by measuring L1 insertions. L1 insertions can be measured by L1-seq/RC-seq and validated by PCR and Sanger sequencing. Thus, both L1 insertion and L1orf1p expression can be measured and compared in both pre-HCC and HCC with different aetiologies. In addition, HCC blood samples can be analysed for L1 promoter methylation in circulating cell free DNA to identify L1 potential as an early biomarker.

Hep3B and HepG2 demonstrated a large variation between L1 transcript and protein expression. Northern blot could be performed to verify transcript findings. Furthermore, L1 expression was predominantly measured using antibody against L1orf1p. L1orf2p is expressed significantly lower level than L1orf1p and antibody detection is usually poor (Smits et al., 2013, Taylor et al., 2013). Instead, L1orf2 expression can be analysed in cell lines using L1 element amplification protocol (LEAP). The LEAP assay measures the ability of L1orf2p to reverse transcribe L1RNA *in vitro* and thus it gives an representation of L1orf2p level and L1 activity in the cells (Kopera et al., 2016).

Several possible mechanisms exist to downregulate L1 in cells. Here, a L1-shRNA transduction was used, as it provides a stable knockdown of L1. However, the experimental method may have some limitations and can be improved. Huh-7 NT and L1-KD were obtained from a heterogeneous cell population and experimental variation were observed between these two cell lines. Further studies can be performed by isolating several single cell colonies from Huh-7, NT-KD and L1-KD. Once sufficiently expanded, these

single colonies can be screened for L1 expression. As these cells would be a homogeneous population, the experimental variation would be less and would provide more reproducible findings. In addition, L1-KD can also be performed in cell lines with low L1 expression such as HepG2 and HUH-1. Due to its low endogenous L1 level, these cell lines may have significantly larger phenotypic changes associated with L1 reduction compared to HUH-7.

TGF- β inhibitors could be an indirect pathway of inhibiting L1 based upon this study. Several TGF- β inhibitors exist such as SB525334. SB525334 is a selective inhibitor of TGF- β Type 1 receptor and has an IC₅₀ of 14.3nM in cell free assay. Previous studies have demonstrated that SB525334 blocks TGF- β -induced Smad activation and decreased migration and invasion was observed in oesophageal cancer (Yue et al., 2015) and ovarian cancer (Wen et al., 2020), and decreased invasion was observed in hepatocellular carcinoma (Zhuang et al., 2017). Further research can be performed by investigating TGF- β inhibition and its effect on L1 expression and functional changes.

L1orf1p doxycycline system should be investigated further to measure level of toxicity of doxycycline. Several techniques exist in measuring cell toxicity including membrane integrity, mitochondrial functionality, oxidative stress and cell death. Annexin V staining is an apoptotic marker and has a high sensitivity. MTT is another assay which measures mitochondrial functionality (Tabernilla et al., 2021). Both assays could be used to measure doxycycline toxicity in HepG2 and determine the maximum doxycycline induction with minimal toxicity.

Current findings have indicated that L1orf1p may have a key role in upregulating TGF- β pathway, particularly through SMAD3. Further research can be performed using an immunoprecipitation assay to confirm these findings. In addition, an L1orf2p inducible cell line can be developed to confirm that L1orf2p has definitely no role in increasing TGF- β signalling.

5.3 Summary

L1 was elevated both at transcript and protein level in HCC patients. Higher L1 expression particularly correlated with poorly differentiated advanced stage tumours with vascular invasion and poorer patient outcome. L1 demonstrated a significant association with TP53 mutation, high AFP and activated TGF- β signalling. *In vitro* studies confirmed a causal link between L1 and TGF- β signalling. As L1 knockdown reduced TGF- β signalling in

Huh7 cells. TGF- β functional properties such as migration and invasion were also reduced in Huh7-L1KD. Furthermore, L1 overexpression increased TGF- β signalling. Further analysis demonstrated that L1orf1p by itself is sufficient to upregulate TGF- β signalling and its downstream EMT marker such as vimentin. Thus, L1orf1p may have a direct role in tumour biology and warrants its possibility as a therapeutic target. Moreover, L1 can serve as a biomarker to guide use of TGF- β -inhibitors.

6.0 References

- ABOU-ALFA, G., MEYER, T., CHENG, A.-L., EL-KHOUEIRY, A., RIMASSA, L., RYOO, B.-Y., CICIN, I., MERLE, P., CHEN, Y., PARK, J., BLANC, J.-F., BOLONDI, L., KLÜMPEN, H.-J., CHAN, S., ZAGONEL, V., PRESSIANI, T., RYU, M.-H., VENOOK, A., HESSEL, C. & KELLEY, R. 2018. Cabozantinib in Patients with Advanced and Progressing Hepatocellular Carcinoma. *The New England journal of medicine*, 379, 54-63.
- ALBANO, E., GORIA-GATTI, L., CLOT, P., JANNONE, A. & TOMASI, A. 1993. Possible role of free radical intermediates in hepatotoxicity of hydrazine derivatives. *Toxicol Ind Health*, 9, 529-38.
- ALEYNIK, S. I., LEO, M. A., ALEYNIK, M. K. & LIEBER, C. S. 1998. Increased circulating products of lipid peroxidation in patients with alcoholic liver disease. *Alcohol Clin Exp Res*, 22, 192-6.
- ALISI, A., GIAMBARTOLOMEI, S., CUPELLI, F., MERLO, P., FONTEMAGGI, G., SPAZIANI, A. & BALSANO, C. 2003. Physical and functional interaction between HCV core protein and the different p73 isoforms. *Oncogene*, 22, 2573-80.
- ALTEKRUSE, S. F., MCGLYNN, K. A. & REICHMAN, M. E. 2009. Hepatocellular carcinoma incidence, mortality, and survival trends in the United States from 1975 to 2005. *J Clin Oncol*, 27, 1485-91.
- ARJAN-ODEDRA, S., SWANSON, C. M., SHERER, N. M., WOLINSKY, S. M. & MALIM, M. H. 2012. Endogenous MOV10 inhibits the retrotransposition of endogenous retroelements but not the replication of exogenous retroviruses. *Retrovirology*, 9, 53.
- ASCH, H. L., ELIACIN, E., FANNING, T. G., CONNOLLY, J. L., BRATTHAUER, G. & ASCH, B. B. 1996. Comparative expression of the LINE-1 p40 protein in human breast carcinomas and normal breast tissues. *Oncol Res*, 8, 239-47.
- ASCHACHER, T., WOLF, B., ENZMANN, F., KIENZL, P., MESSNER, B., SAMPL, S., SVOBODA, M., MECHTCHERIAKOVA, D., HOLZMANN, K. & BERGMANN, M. 2016. LINE-1 induces hTERT and ensures telomere maintenance in tumour cell lines. *Oncogene*, 35, 94-104.
- ATIQ, O., TIRO, J., YOPP, A. C., MUFFLER, A., MARRERO, J. A., PARIKH, N. D., MURPHY, C., MCCALLISTER, K. & SINGAL, A. G. 2017. An assessment of benefits and harms of hepatocellular carcinoma surveillance in patients with cirrhosis. *Hepatology*, 65, 1196-1205.
- AWANO, H., MALUEKA, R. G., YAGI, M., OKIZUKA, Y., TAKESHIMA, Y. & MATSUO, M. 2010. Contemporary retrotransposition of a novel non-coding gene induces exon-skipping in dystrophin mRNA. *J Hum Genet*, 55, 785-90.
- BAEK, K. H., PARK, H. Y., KANG, C. M., KIM, S. J., JEONG, S. J., HONG, E. K., PARK, J. W., SUNG, Y. C., SUZUKI, T., KIM, C. M. & LEE, C. W. 2006. Overexpression of hepatitis C virus NS5A protein induces chromosome instability via mitotic cell cycle dysregulation. *J Mol Biol*, 359, 22-34.
- BARBA, G., HARPER, F., HARADA, T., KOHARA, M., GOULINET, S., MATSUURA, Y., EDER, G., SCHAFF, Z., CHAPMAN, M. J., MIYAMURA, T. & BRECHOT, C. 1997. Hepatitis C virus core protein shows a cytoplasmic localization and associates to cellular lipid storage droplets. *Proc Natl Acad Sci U S A*, 94, 1200-5.
- BARCHITTA, M., QUATTROCCHI, A., MAUGERI, A., VINCIGUERRA, M. & AGODI, A. 2014. LINE-1 hypomethylation in blood and tissue samples as an epigenetic marker for cancer risk: a systematic review and meta-analysis. *PLoS One*, 9, e109478.
- BARUCH, E. N., BERG, A. L., BESSER, M. J., SCHACHTER, J. & MARKEL, G. 2017. Adoptive T cell therapy: An overview of obstacles and opportunities. *Cancer*, 123, 2154-2162.
- BATALLER, R., PAIK, Y. H., LINDQUIST, J. N., LEMASTERS, J. J. & BRENNER, D. A. 2004. Hepatitis C virus core and nonstructural proteins induce fibrogenic effects in hepatic stellate cells. *Gastroenterology*, 126, 529-40.

- BECK, C. R., COLLIER, P., MACFARLANE, C., MALIG, M., KIDD, J. M., EICHLER, E. E., BADGE, R. M. & MORAN, J. V. 2010. LINE-1 retrotransposition activity in human genomes. *Cell*, 141, 1159-70.
- BELGHITI, J., CARR, B. I., GREIG, P. D., LENCIONI, R. & POON, R. T. 2008. Treatment before liver transplantation for HCC. *Ann Surg Oncol*, 15, 993-1000.
- BHALA, N., ANGULO, P., VAN DER POORTEN, D., LEE, E., HUI, J. M., SARACCO, G., ADAMS, L. A., CHARATCHAROENWITTHAYA, P., TOPPING, J. H., BUGIANESI, E., DAY, C. P. & GEORGE, J. 2011. The natural history of nonalcoholic fatty liver disease with advanced fibrosis or cirrhosis: an international collaborative study. *Hepatology*, 54, 1208-16.
- BOCK, C. T., SCHWINN, S., LOCARNINI, S., FYFE, J., MANNIS, M. P., TRAUTWEIN, C. & ZENTGRAF, H. 2001. Structural organization of the hepatitis B virus minichromosome. *J Mol Biol*, 307, 183-96.
- BRACKEN, C. P., SCOTT, H. S. & GOODALL, G. J. 2016. A network-biology perspective of microRNA function and dysfunction in cancer. *Nature Reviews Genetics*, 17, 719-732.
- BRENNER, D. A. 2009. Molecular pathogenesis of liver fibrosis. *Transactions of the American Clinical and Climatological Association*, 120, 361-368.
- BRUIX, J., QIN, S., MERLE, P., GRANITO, A., HUANG, Y. H., BODOKY, G., PRACHT, M., YOKOSUKA, O., ROSMORDUC, O., BREDER, V., GEROLAMI, R., MASI, G., ROSS, P. J., SONG, T., BRONOWICKI, J. P., OLLIVIER-HOURMAND, I., KUDO, M., CHENG, A. L., LLOVET, J. M., FINN, R. S., LEBERRE, M. A., BAUMHAUER, A., MEINHARDT, G. & HAN, G. 2017. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet*, 389, 56-66.
- BRUIX, J. & SHERMAN, M. 2005. Management of hepatocellular carcinoma. *Hepatology*, 42, 1208-36.
- BU, S., YAMANAKA, M., PEI, H., BIELAWSKA, A., BIELAWSKI, J., HANNUN, Y. A., OBEID, L. & TROJANOWSKA, M. 2006. Dihydrospingosine 1-phosphate stimulates MMP1 gene expression via activation of ERK1/2-Ets1 pathway in human fibroblasts. *Faseb j*, 20, 184-6.
- BUJOLD, A., MASSEY, C. A., KIM, J. J., BRIERLEY, J., CHO, C., WONG, R. K., DINNIWELL, R. E., KASSAM, Z., RINGASH, J., CUMMINGS, B., SYKES, J., SHERMAN, M., KNOX, J. J. & DAWSON, L. A. 2013. Sequential phase I and II trials of stereotactic body radiotherapy for locally advanced hepatocellular carcinoma. *J Clin Oncol*, 31, 1631-9.
- BURAK, K. W. & KNETEMAN, N. M. 2010. An evidence-based multidisciplinary approach to the management of hepatocellular carcinoma (HCC): the Alberta HCC algorithm. *Can J Gastroenterol*, 24, 643-50.
- BURNS, K. H. 2017. Transposable elements in cancer. *Nature Reviews Cancer*, 17, 415-424.
- CANCERRESEARCHUK. 2017. *Liver Cancer (ICD-10 C22), European Age-Standardised Incidence Rates, UK, 1993-2017* [Online]. Available: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer/incidence#heading-Two> [Accessed].
- CARREIRA, P. E., RICHARDSON, S. R. & FAULKNER, G. J. 2014. L1 retrotransposons, cancer stem cells and oncogenesis. *Febs j*, 281, 63-73.
- CHEN, C., LI, K., JIANG, H., SONG, F., GAO, H., PAN, X., SHI, B., BI, Y., WANG, H., WANG, H. & LI, Z. 2017. Development of T cells carrying two complementary chimeric antigen receptors against glypican-3 and asialoglycoprotein receptor 1 for the treatment of hepatocellular carcinoma. *Cancer Immunol Immunother*, 66, 475-489.
- CHEN, J., ZAIDI, S., RAO, S., CHEN, J. S., PHAN, L., FARCI, P., SU, X., SHETTY, K., WHITE, J., ZAMBONI, F., WU, X., RASHID, A., PATTABIRAMAN, N., MAZUMDER, R., HORVATH, A., WU, R. C., LI, S., XIAO, C., DENG, C. X., WHEELER, D. A., MISHRA, B., AKBANI, R. & MISHRA, L. 2018. Analysis of Genomes and Transcriptomes of Hepatocellular Carcinomas

- Identifies Mutations and Gene Expression Changes in the Transforming Growth Factor- β Pathway. *Gastroenterology*, 154, 195-210.
- CHEN, L., DAHLSTROM, J. E., LEE, S. H. & RANGASAMY, D. 2012. Naturally occurring endo-siRNA silences LINE-1 retrotransposons in human cells through DNA methylation. *Epigenetics*, 7, 758-71.
- CHENG, A. S., LAU, S. S., CHEN, Y., KONDO, Y., LI, M. S., FENG, H., CHING, A. K., CHEUNG, K. F., WONG, H. K., TONG, J. H., JIN, H., CHOY, K. W., YU, J., TO, K. F., WONG, N., HUANG, T. H. & SUNG, J. J. 2011. EZH2-mediated concordant repression of Wnt antagonists promotes beta-catenin-dependent hepatocarcinogenesis. *Cancer Res*, 71, 4028-39.
- CLAYTON, R. F., RINALDI, A., KANDYBA, E. E., EDWARD, M., WILLBERG, C., KLENERMAN, P. & PATEL, A. H. 2005. Liver cell lines for the study of hepatocyte functions and immunological response. *Liver Int*, 25, 389-402.
- COOK, P. R., JONES, C. E. & FURANO, A. V. 2015. Phosphorylation of ORF1p is required for L1 retrotransposition. *Proc Natl Acad Sci U S A*, 112, 4298-303.
- CRISCIONE, S. W., THEODOSAKIS, N., MICEVIC, G., CORNISH, T. C., BURNS, K. H., NERETTI, N. & RODIĆ, N. 2016. Genome-wide characterization of human L1 antisense promoter-driven transcripts. *BMC Genomics*, 17, 463.
- CRUICKSHANKS, H. A. & TUFARELLI, C. 2009. Isolation of cancer-specific chimeric transcripts induced by hypomethylation of the LINE-1 antisense promoter. *Genomics*, 94, 397-406.
- CUTTER, A. R. & HAYES, J. J. 2015. A brief review of nucleosome structure. *FEBS Letters*, 589, 2914-2922.
- DAI, S., JIA, R., ZHANG, X., FANG, Q. & HUANG, L. 2014. The PD-1/PD-Ls pathway and autoimmune diseases. *Cell Immunol*, 290, 72-9.
- 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. & GAGE, F. H. 2015. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. *Cell*, 163, 583-93.
- DEWANNIEUX, M., ESNAULT, C. & HEIDMANN, T. 2003. LINE-mediated retrotransposition of marked Alu sequences. *Nature Genetics*, 35, 41-48.
- DIAZ-GONZALEZ, A., REIG, M. & BRUIX, J. 2016. Treatment of Hepatocellular Carcinoma. *Dig Dis*, 34, 597-602.
- DING, D., LOU, X., HUA, D., YU, W., LI, L., WANG, J., GAO, F., ZHAO, N., REN, G., LI, L. & LIN, B. 2012a. Recurrent targeted genes of hepatitis B virus in the liver cancer genomes identified by a next-generation sequencing-based approach. *PLoS Genet*, 8, e1003065.
- DING, W., DANG, H., YOU, H., STEINWAY, S., TAKAHASHI, Y., WANG, H. G., LIAO, J., STILES, B., ALBERT, R. & ROUNTREE, C. B. 2012b. miR-200b restoration and DNA methyltransferase inhibitor block lung metastasis of mesenchymal-phenotype hepatocellular carcinoma. *Oncogenesis*, 1, e15-e15.
- DING, Y., CHEN, B., WANG, S., ZHAO, L., CHEN, J., DING, Y., CHEN, L. & LUO, R. 2009. Overexpression of Tiam1 in hepatocellular carcinomas predicts poor prognosis of HCC patients. *International Journal of Cancer*, 124, 653-658.
- DMITRIEV, S. E., ANDREEV, D. E., TERENIN, I. M., OLOVNIKOV, I. A., PRASSOLOV, V. S., MERRICK, W. C. & SHATSKY, I. N. 2007. Efficient translation initiation directed by the 900-nucleotide-long and GC-rich 5' untranslated region of the human retrotransposon LINE-1 mRNA is strictly cap dependent rather than internal ribosome entry site mediated. *Mol Cell Biol*, 27, 4685-97.
- DONATO, F., TAGGER, A., GELATTI, U., PARRINELLO, G., BOFFETTA, P., ALBERTINI, A., DECARLI, A., TREVISI, P., RIBERO, M. L., MARTELLI, C., PORRU, S. & NARDI, G. 2002. Alcohol and hepatocellular carcinoma: the effect of lifetime intake and hepatitis virus infections in men and women. *Am J Epidemiol*, 155, 323-31.

- DONG, Y. & WANG, A. 2014. Aberrant DNA methylation in hepatocellular carcinoma tumor suppression (Review). *Oncology letters*, 8, 963-968.
- DOWDY, S. C., MARIANI, A. & JANKNECHT, R. 2003. HER2/Neu- and TAK1-mediated up-regulation of the transforming growth factor beta inhibitor Smad7 via the ETS protein ER81. *J Biol Chem*, 278, 44377-84.
- DUFFY, A. G., ULAHANNAN, S. V., MAKOROVA-RUSHER, O., RAHMA, O., WEDEMEYER, H., PRATT, D., DAVIS, J. L., HUGHES, M. S., HELLER, T., ELGINDI, M., UPPALA, A., KORANGY, F., KLEINER, D. E., FIGG, W. D., VENZON, D., STEINBERG, S. M., VENKATESAN, A. M., KRISHNASAMY, V., ABI-JAOUDEH, N., LEVY, E., WOOD, B. J. & GRETEN, T. F. 2017. Tremelimumab in combination with ablation in patients with advanced hepatocellular carcinoma. *J Hepatol*, 66, 545-551.
- DUFFY, M. J. 2010. Use of Biomarkers in Screening for Cancer. *EJIFCC*, 21, 1-12.
- DUPONT, I., LUCAS, D., CLOT, P., MÉNEZ, C. & ALBANO, E. 1998. Cytochrome P4502E1 inducibility and hydroxyethyl radical formation among alcoholics. *J Hepatol*, 28, 564-71.
- EBISAWA, T., FUKUCHI, M., MURAKAMI, G., CHIBA, T., TANAKA, K., IMAMURA, T. & MIYAZONO, K. 2001. Smurf1 interacts with transforming growth factor-beta type I receptor through Smad7 and induces receptor degradation. *J Biol Chem*, 276, 12477-80.
- EL-KHOUEIRY, A. B., MELERO, I., CROCENZI, T. S., WELLING, T. H., YAU, T. C., YEO, W., CHOPRA, A., GROSSO, J., LANG, L., ANDERSON, J., CRUZ, C. M. D. & SANGRO, B. 2015. Phase I/II safety and antitumor activity of nivolumab in patients with advanced hepatocellular carcinoma (HCC): CA209-040. *Journal of Clinical Oncology*, 33, LBA101-LBA101.
- EL-SERAG, H. B. 2011. Hepatocellular carcinoma. *N Engl J Med*, 365, 1118-27.
- ELMBERG, M., HULTCRANTZ, R., EKBOM, A., BRANDT, L., OLSSON, S., OLSSON, R., LINDGREN, S., LOOF, L., STAL, P., WALLERSTEDT, S., ALMER, S., SANDBERG-GERTZEN, H. & ASKLING, J. 2003. Cancer risk in patients with hereditary hemochromatosis and in their first-degree relatives. *Gastroenterology*, 125, 1733-41.
- 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., JR. & SOLYOM, S. 2015. Widespread somatic L1 retrotransposition occurs early during gastrointestinal cancer evolution. *Genome Res*, 25, 1536-45.
- FABREGAT, I. & GIANNELLI, G. 2017. The TGF-beta pathway: a pharmacological target in hepatocellular carcinoma? *Hepat Oncol*, 4, 35-38.
- FARAZI, P. A. & DEPINHO, R. A. 2006. Hepatocellular carcinoma pathogenesis: from genes to environment. *Nat Rev Cancer*, 6, 674-87.
- FENG, F., LU, Y.-Y., ZHANG, F., GAO, X.-D., ZHANG, C.-F., MEREDITH, A., XU, Z.-X., YANG, Y.-T., CHANG, X.-J., WANG, H., QU, J.-H., ZENG, Z., YANG, J.-L., WANG, C.-P., ZHU, Y.-F., CUI, J.-J. & YANG, Y.-P. 2013. Long interspersed nuclear element ORF-1 protein promotes proliferation and resistance to chemotherapy in hepatocellular carcinoma. *World journal of gastroenterology*, 19, 1068-1078.
- FENG, Q., MORAN, J. V., KAZAZIAN, H. H., JR. & BOEKE, J. D. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell*, 87, 905-16.
- FERLAY, J., COLOMBET, M., SOERJOMATARAM, I., MATHERS, C., PARKIN, D. M., PIÑEROS, M., ZNAOR, A. & BRAY, F. 2019. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*, 144, 1941-1953.
- FERLAY, J., STELIAROVA-FOUCHER, E., LORTET-TIEULENT, J., ROSSO, S., COEBERGH, J. W., COMBER, H., FORMAN, D. & BRAY, F. 2013. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer*, 49, 1374-403.
- FERRACIN, M., VERONESE, A. & NEGRINI, M. 2010. Micromarkers: miRNAs in cancer diagnosis and prognosis. *Expert Rev Mol Diagn*, 10, 297-308.

- FILMUS, J. & CAPURRO, M. 2008. The role of glypican-3 in the regulation of body size and cancer. *Cell Cycle*, 7, 2787-90.
- FINN, R. S., QIN, S., IKEDA, M., GALLE, P. R., DUCREUX, M., KIM, T. Y., KUDO, M., BREDER, V., MERLE, P., KASEB, A. O., LI, D., VERRET, W., XU, D. Z., HERNANDEZ, S., LIU, J., HUANG, C., MULLA, S., WANG, Y., LIM, H. Y., ZHU, A. X. & CHENG, A. L. 2020. Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. *N Engl J Med*, 382, 1894-1905.
- FORNER, A., LLOVET, J. M. & BRUIX, J. 2012. Hepatocellular carcinoma. *Lancet*, 379, 1245-55.
- FRACANZANI, A. L., CONTE, D., FRAQUELLI, M., TAIOLI, E., MATTIOLI, M., LOSCO, A. & FARGION, S. 2001. Increased cancer risk in a cohort of 230 patients with hereditary hemochromatosis in comparison to matched control patients with non-iron-related chronic liver disease. *Hepatology*, 33, 647-51.
- FRAZER, C. 1999. Imaging of hepatocellular carcinoma. *J Gastroenterol Hepatol*, 14, 750-6.
- FUJIMOTO, A., TOTOKI, Y., ABE, T., BOROEVIK, K. A., HOSODA, F., NGUYEN, H. H., AOKI, M., HOSONO, N., KUBO, M., MIYA, F., ARAI, Y., TAKAHASHI, H., SHIRAKIHARA, T., NAGASAKI, M., SHIBUYA, T., NAKANO, K., WATANABE-MAKINO, K., TANAKA, H., NAKAMURA, H., KUSUDA, J., OJIMA, H., SHIMADA, K., OKUSAKA, T., UENO, M., SHIGEKAWA, Y., KAWAKAMI, Y., ARIHIRO, K., OHDAN, H., GOTOH, K., ISHIKAWA, O., ARIIZUMI, S., YAMAMOTO, M., YAMADA, T., CHAYAMA, K., KOSUGE, T., YAMAUE, H., KAMATANI, N., MIYANO, S., NAKAGAMA, H., NAKAMURA, Y., TSUNODA, T., SHIBATA, T. & NAKAGAWA, H. 2012. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat Genet*, 44, 760-4.
- FURANO, A. V. 2000. The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. *Prog Nucleic Acid Res Mol Biol*, 64, 255-94.
- GALLIMORE, A. M. & SIMON, A. K. 2008. Positive and negative influences of regulatory T cells on tumour immunity. *Oncogene*, 27, 5886-93.
- GAO, H., LI, K., TU, H., PAN, X., JIANG, H., SHI, B., KONG, J., WANG, H., YANG, S., GU, J. & LI, Z. 2014a. Development of T cells redirected to glypican-3 for the treatment of hepatocellular carcinoma. *Clin Cancer Res*, 20, 6418-28.
- GAO, X., MI, Y., GUO, N., XU, H., XU, L., GOU, X. & JIN, W. 2017. Cytokine-Induced Killer Cells As Pharmacological Tools for Cancer Immunotherapy. *Front Immunol*, 8, 774.
- GAO, X. D., QU, J. H., CHANG, X. J., LU, Y. Y., BAI, W. L., WANG, H., XU, Z. X., AN, L. J., WANG, C. P., ZENG, Z. & YANG, Y. P. 2014b. Hypomethylation of long interspersed nuclear element-1 promoter is associated with poor outcomes for curative resected hepatocellular carcinoma. *Liver Int*, 34, 136-46.
- GERHAUSER, C., HEILMANN, K. & PUDENZ, M. 2015. Genome-Wide DNA Methylation Profiling in Dietary Intervention Studies: a User's Perspective. *Current Pharmacology Reports*, 1, 31-45.
- GIANNELLI, G., VILLA, E. & LAHN, M. 2014. Transforming growth factor-beta as a therapeutic target in hepatocellular carcinoma. *Cancer Res*, 74, 1890-4.
- GIORGI, G., MARCANTONIO, P. & DEL RE, B. 2011. LINE-1 retrotransposition in human neuroblastoma cells is affected by oxidative stress. *Cell Tissue Res*, 346, 383-91.
- GOLDSTEIN, S. T., ZHOU, F., HADLER, S. C., BELL, B. P., MAST, E. E. & MARGOLIS, H. S. 2005. A mathematical model to estimate global hepatitis B disease burden and vaccination impact. *Int J Epidemiol*, 34, 1329-39.
- GONZALEZ, S., KLATT, P., DELGADO, S., CONDE, E., LOPEZ-RIOS, F., SANCHEZ-CESPEDES, M., MENDEZ, J., ANTEQUERA, F. & SERRANO, M. 2006. Oncogenic activity of Cdc6 through repression of the INK4/ARF locus. *Nature*, 440, 702-6.
- GOODIER, J. L., OSTERTAG, E. M. & KAZAZIAN, H. H., JR. 2000. Transduction of 3'-flanking sequences is common in L1 retrotransposition. *Hum Mol Genet*, 9, 653-7.

- GOOSSENS, N., SUN, X. & HOSHIDA, Y. 2015. Molecular classification of hepatocellular carcinoma: potential therapeutic implications. *Hepatic oncology*, 2, 371-379.
- GOUAS, D., SHI, H. & HAINAUT, P. 2009. The aflatoxin-induced TP53 mutation at codon 249 (R249S): biomarker of exposure, early detection and target for therapy. *Cancer Lett*, 286, 29-37.
- GUICHARD, C., AMADDEO, G., IMBEAUD, S., LADEIRO, Y., PELLETIER, L., MAAD, I. B., CALDERARO, J., BIOULAC-SAGE, P., LETEXIER, M., DEGOS, F., CLEMENT, B., BALABAUD, C., CHEVET, E., LAURENT, A., COUCHY, G., LETOUZE, E., CALVO, F. & ZUCMAN-ROSSI, J. 2012. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet*, 44, 694-8.
- HAJARIZADEH, B., GREBELY, J. & DORE, G. J. 2013. Epidemiology and natural history of HCV infection. *Nat Rev Gastroenterol Hepatol*, 10, 553-62.
- HAMDORF, M., IDICA, A., ZISOULIS, D. G., GAMELIN, L., MARTIN, C., SANDERS, K. J. & PEDERSEN, I. M. 2015. miR-128 represses L1 retrotransposition by binding directly to L1 RNA. *Nat Struct Mol Biol*, 22, 824-31.
- HARRIS, C. R., DEWAN, A., ZUPNICK, A., NORMART, R., GABRIEL, A., PRIVES, C., LEVINE, A. J. & HOH, J. 2009. p53 responsive elements in human retrotransposons. *Oncogene*, 28, 3857-65.
- HARRIS, C. R., NORMART, R., YANG, Q., STEVENSON, E., HAFFTY, B. G., GANESAN, S., CORDON-CARDO, C., LEVINE, A. J. & TANG, L. H. 2010. Association of nuclear localization of a long interspersed nuclear element-1 protein in breast tumors with poor prognostic outcomes. *Genes Cancer*, 1, 115-24.
- HAYASHI, J., AOKI, H., KAJINO, K., MORIYAMA, M., ARAKAWA, Y. & HINO, O. 2000. Hepatitis C virus core protein activates the MAPK/ERK cascade synergistically with tumor promoter TPA, but not with epidermal growth factor or transforming growth factor alpha. *Hepatology*, 32, 958-61.
- HE, F., LI, J., XU, J., ZHANG, S., XU, Y., ZHAO, W., YIN, Z. & WANG, X. 2015. Decreased expression of ARID1A associates with poor prognosis and promotes metastases of hepatocellular carcinoma. *J Exp Clin Cancer Res*, 34, 47.
- HECKMAN, J. T., DEVERA, M. B., MARSH, J. W., FONTES, P., AMESUR, N. B., HOLLOWAY, S. E., NALESNIK, M., GELLER, D. A., STEEL, J. L. & GAMBLIN, T. C. 2008. Bridging locoregional therapy for hepatocellular carcinoma prior to liver transplantation. *Ann Surg Oncol*, 15, 3169-77.
- HEFFERNAN, A., COOKE, G. S., NAYAGAM, S., THURSZ, M. & HALLETT, T. B. 2019. Scaling up prevention and treatment towards the elimination of hepatitis C: a global mathematical model. *Lancet*, 393, 1319-1329.
- HELDIN, C. H., MIYAZONO, K. & TEN DIJKE, P. 1997. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature*, 390, 465-71.
- HELMAN, E., LAWRENCE, M. S., STEWART, C., SOUGNEZ, C., GETZ, G. & MEYERSON, M. 2014. Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. *Genome Res*, 24, 1053-63.
- HIROSE, Y. & MANLEY, J. L. 1998. RNA polymerase II is an essential mRNA polyadenylation factor. *Nature*, 395, 93-6.
- HOFFMANN, A., PREOBRAZHENSKA, O., WODARCZYK, C., MEDLER, Y., WINKEL, A., SHAHAB, S., HUYLEBROECK, D., GROSS, G. & VERSCHUEREN, K. 2005. Transforming growth factor-beta-activated kinase-1 (TAK1), a MAP3K, interacts with Smad proteins and interferes with osteogenesis in murine mesenchymal progenitors. *J Biol Chem*, 280, 27271-83.
- HOHJOH, H. & SINGER, M. F. 1996. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *Embo j*, 15, 630-9.
- HOSHIDA, Y., NIJMAN, S. M., KOBAYASHI, M., CHAN, J. A., BRUNET, J. P., CHIANG, D. Y., VILLANUEVA, A., NEWELL, P., IKEDA, K., HASHIMOTO, M., WATANABE, G., GABRIEL, S.,

- FRIEDMAN, S. L., KUMADA, H., LLOVET, J. M. & GOLUB, T. R. 2009a. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res*, 69, 7385-92.
- HOSHIDA, Y., NIJMAN, S. M. B., KOBAYASHI, M., CHAN, J. A., BRUNET, J.-P., CHIANG, D. Y., VILLANUEVA, A., NEWELL, P., IKEDA, K., HASHIMOTO, M., WATANABE, G., GABRIEL, S., FRIEDMAN, S. L., KUMADA, H., LLOVET, J. M. & GOLUB, T. R. 2009b. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer research*, 69, 7385-7392.
- HOSHIDA, Y., TOFFANIN, S., LACHENMAYER, A., VILLANUEVA, A., MINGUEZ, B. & LLOVET, J. M. 2010. Molecular classification and novel targets in hepatocellular carcinoma: recent advancements. *Seminars in liver disease*, 30, 35-51.
- HOWELL, R. & USDIN, K. 1997. The ability to form intrastrand tetraplexes is an evolutionarily conserved feature of the 3' end of L1 retrotransposons. *Mol Biol Evol*, 14, 144-55.
- HU, S., LI, J., XU, F., MEI, S., LE DUFF, Y., YIN, L., PANG, X., CEN, S., JIN, Q., LIANG, C. & GUO, F. 2015. SAMHD1 Inhibits LINE-1 Retrotransposition by Promoting Stress Granule Formation. *PLoS Genet*, 11, e1005367.
- HU, W., FENG, Z., EVELEIGH, J., IYER, G., PAN, J., AMIN, S., CHUNG, F. L. & TANG, M. S. 2002. The major lipid peroxidation product, trans-4-hydroxy-2-nonenal, preferentially forms DNA adducts at codon 249 of human p53 gene, a unique mutational hotspot in hepatocellular carcinoma. *Carcinogenesis*, 23, 1781-9.
- HUANG, Y. T., JEN, C. L., YANG, H. I., LEE, M. H., SU, J., LU, S. N., ILOEJE, U. H. & CHEN, C. J. 2011. Lifetime risk and sex difference of hepatocellular carcinoma among patients with chronic hepatitis B and C. *J Clin Oncol*, 29, 3643-50.
- ISKOW, R. C., MCCABE, M. T., MILLS, R. E., TORENE, S., PITTARD, W. S., NEUWALD, A. F., VAN MEIR, E. G., VERTINO, P. M. & DEVINE, S. E. 2010. Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell*, 141, 1253-61.
- IWASAKI, M., FURUSE, J., YOSHINO, M., RYU, M., MORIYAMA, N. & MUKAI, K. 1998. Sonographic appearances of small hepatic nodules without tumor stain on contrast-enhanced computed tomography and angiography. *J Clin Ultrasound*, 26, 303-7.
- IZUMI, D., TODEN, S., URETA, E., ISHIMOTO, T., BABA, H. & GOEL, A. 2019. TIAM1 promotes chemoresistance and tumor invasiveness in colorectal cancer. *Cell Death & Disease*, 10, 267.
- JAIN, D. 2014. Tissue diagnosis of hepatocellular carcinoma. *Journal of clinical and experimental hepatology*, 4, S67-S73.
- JEWELL, J. & SHERON, N. 2010. Trends in European liver death rates: implications for alcohol policy. *Clin Med (Lond)*, 10, 259-63.
- JIANG, Z., JHUNJHUNWALA, S., LIU, J., HAVERTY, P. M., KENNEMER, M. I., GUAN, Y., LEE, W., CARNEVALI, P., STINSON, J., JOHNSON, S., DIAO, J., YEUNG, S., JUBB, A., YE, W., WU, T. D., KAPADIA, S. B., DE SAUVAGE, F. J., GENTLEMAN, R. C., STERN, H. M., SESHAGIRI, S., PANT, K. P., MODRUSAN, Z., BALLINGER, D. G. & ZHANG, Z. 2012. The effects of hepatitis B virus integration into the genomes of hepatocellular carcinoma patients. *Genome Res*, 22, 593-601.
- JOO, M., HAHN, Y. S., KWON, M., SADIKOT, R. T., BLACKWELL, T. S. & CHRISTMAN, J. W. 2005. Hepatitis C virus core protein suppresses NF-kappaB activation and cyclooxygenase-2 expression by direct interaction with IkappaB kinase beta. *J Virol*, 79, 7648-57.
- JUN, H. J., KIM, J., HOANG, M. H. & LEE, S. J. 2012. Hepatic lipid accumulation alters global histone h3 lysine 9 and 4 trimethylation in the peroxisome proliferator-activated receptor alpha network. *PLoS One*, 7, e44345.
- JUNG, H., CHOI, J. K. & LEE, E. A. 2018. Immune signatures correlate with L1 retrotransposition in gastrointestinal cancers. *Genome Res*, 28, 1136-1146.

- JUNG, J. K., ARORA, P., PAGANO, J. S. & JANG, K. L. 2007. Expression of DNA methyltransferase 1 is activated by hepatitis B virus X protein via a regulatory circuit involving the p16INK4a-cyclin D1-CDK 4/6-pRb-E2F1 pathway. *Cancer Res*, 67, 5771-8.
- KALE, S. P., MOORE, L., DEININGER, P. L. & ROY-ENGEL, A. M. 2005. Heavy metals stimulate human LINE-1 retrotransposition. *Int J Environ Res Public Health*, 2, 14-23.
- KAO, C. F., CHEN, S. Y., CHEN, J. Y. & WU LEE, Y. H. 2004. Modulation of p53 transcription regulatory activity and post-translational modification by hepatitis C virus core protein. *Oncogene*, 23, 2472-83.
- KARANTH, A. V., MANISWAMI, R. R., PRASHANTH, S., GOVINDARAJ, H., PADMAVATHY, R., JEGATHEESAN, S. K., MULLANGI, R. & RAJAGOPAL, S. 2017. Emerging role of SETDB1 as a therapeutic target. *Expert Opin Ther Targets*, 21, 319-331.
- KARKAMPOUNA, S., TEN DIJKE, P., DOOLEY, S. & JULIO, M. K. 2012. TGFbeta signaling in liver regeneration. *Curr Pharm Des*, 18, 4103-13.
- KAZAZIAN, H. H., WONG, C., YOUSOUFIAN, H., SCOTT, A. F., PHILLIPS, D. G. & ANTONARAKIS, S. E. 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature*, 332, 164-166.
- KEW, M. C. 2003. Synergistic interaction between aflatoxin B1 and hepatitis B virus in hepatocarcinogenesis. *Liver Int*, 23, 405-9.
- KEW, M. C. 2010. Epidemiology of chronic hepatitis B virus infection, hepatocellular carcinoma, and hepatitis B virus-induced hepatocellular carcinoma. *Pathol Biol (Paris)*, 58, 273-7.
- KHAZINA, E., TRUFFAULT, V., BUTTNER, R., SCHMIDT, S., COLES, M. & WEICHENRIEDER, O. 2011. Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. *Nat Struct Mol Biol*, 18, 1006-14.
- KHAZINA, E. & WEICHENRIEDER, O. 2009. Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc Natl Acad Sci U S A*, 106, 731-6.
- KIM, K. M., KWON, S. N., KANG, J. I., LEE, S. H., JANG, S. K., AHN, B. Y. & KIM, Y. K. 2007. Hepatitis C virus NS2 protein activates cellular cyclic AMP-dependent pathways. *Biochem Biophys Res Commun*, 356, 948-54.
- KLADNEY, R. D., CUI, X., BULLA, G. A., BRUNT, E. M. & FIMMEL, C. J. 2002. Expression of GP73, a resident Golgi membrane protein, in viral and nonviral liver disease. *Hepatology*, 35, 1431-40.
- KLOSE, R. J. & BIRD, A. P. 2006. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci*, 31, 89-97.
- KOJIRO, M. W., IR ; ALVES, V ; BADVE, S ; BALA-BAUD, C ; BEDOSA, P ; BHATHAL, P ; BIOULAC-SAGE, P ; BRUNT, EM ; BURT, AD ; CRAIG, JR ; DHILLON, A ; FERRELL, L ; GELLER, SA ; GOODMAN, ZD ; GOUW, ASH ; GUIDO, M ; GUINDI, M ; HYTIROGLOU, P ; KAGE, M ; KONDO, F ; KUDO, M ; LAUWERS, GY ; NAKANO, M ; PARADIS, V ; PARK, YN ; QUAGLIA, A ; RONCALLI, M ; ROSKAMS, T ; RUEBNER, B ; SAKAMOTO, M ; SAX-ENA, R ; THEISE, ND ; THUNG, S ; TONIAKOS, D 2009. Pathologic diagnosis of early hepatocellular carcinoma: a report of the international consensus group for hepatocellular neoplasia. *Hepatology*, 49, 658-64.
- KOPERA, H. C., FLASCH, D. A., NAKAMURA, M., MIYOSHI, T., DOUCET, A. J. & MORAN, J. V. 2016. LEAP: L1 Element Amplification Protocol. In: GARCIA-PÉREZ, J. L. (ed.) *Transposons and Retrotransposons: Methods and Protocols*. New York, NY: Springer New York.
- KOZAK, M. 1989. The scanning model for translation: an update. *J Cell Biol*, 108, 229-41.
- KRAMEROV, D. A. & VASSETZKY, N. S. 2011. Origin and evolution of SINEs in eukaryotic genomes. *Heredity (Edinb)*, 107, 487-95.
- KUDO, M., FINN, R. S., QIN, S., HAN, K. H., IKEDA, K., PISCAGLIA, F., BARON, A., PARK, J. W., HAN, G., JASSEM, J., BLANC, J. F., VOGEL, A., KOMOV, D., EVANS, T. R. J., LOPEZ, C., DUTCUS, C., GUO, M., SAITO, K., KRALJEVIC, S., TAMAI, T., REN, M. & CHENG, A. L. 2018. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable

- hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet*, 391, 1163-1173.
- KUILMAN, T., MICHALOGLOU, C., MOOI, W. J. & PEEPER, D. S. 2010. The essence of senescence. *Genes Dev*, 24, 2463-79.
- KWON, G. Y., YOO, B. C., KOH, K. C., CHO, J. W., PARK, W. S. & PARK, C. K. 2005. Promoter methylation of E-cadherin in hepatocellular carcinomas and dysplastic nodules. *J Korean Med Sci*, 20, 242-7.
- LACHENMAYER, A., ALSINET, C., SAVIC, R., CABELLOS, L., TOFFANIN, S., HOSHIDA, Y., VILLANUEVA, A., MINGUEZ, B., NEWELL, P., TSAI, H. W., BARRETINA, J., THUNG, S., WARD, S. C., BRUIX, J., MAZZAFERRO, V., SCHWARTZ, M., FRIEDMAN, S. L. & LLOVET, J. M. 2012. Wnt-pathway activation in two molecular classes of hepatocellular carcinoma and experimental modulation by sorafenib. *Clin Cancer Res*, 18, 4997-5007.
- LARSSON, S. C. & WOLK, A. 2007. Overweight, obesity and risk of liver cancer: a meta-analysis of cohort studies. *British journal of cancer*, 97, 1005-1008.
- LAU, C. C., SUN, T., CHING, A. K., HE, M., LI, J. W., WONG, A. M., CO, N. N., CHAN, A. W., LI, P. S., LUNG, R. W., TONG, J. H., LAI, P. B., CHAN, H. L., TO, K. F., CHAN, T. F. & WONG, N. 2014. Viral-human chimeric transcript predisposes risk to liver cancer development and progression. *Cancer Cell*, 25, 335-49.
- LAU, W. Y., SANGRO, B., CHEN, P. J., CHENG, S. Q., CHOW, P., LEE, R. C., LEUNG, T., HAN, K. H. & POON, R. T. 2013. Treatment for hepatocellular carcinoma with portal vein tumor thrombosis: the emerging role for radioembolization using yttrium-90. *Oncology*, 84, 311-8.
- LEE, D. H., NAM, J. Y., CHANG, Y., CHO, H., KANG, S. H., CHO, Y. Y., CHO, E., LEE, J.-H., YU, S. J., KIM, Y. J. & YOON, J.-H. 2017. Synergistic effect of cytokine-induced killer cell with valproate inhibits growth of hepatocellular carcinoma cell in a mouse model. *Cancer biology & therapy*, 18, 67-75.
- 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. 2012. Landscape of somatic retrotransposition in human cancers. *Science*, 337, 967-71.
- LEE, J.-O., KWUN, H. J., JUNG, J. K., CHOI, K. H., MIN, D. S. & JANG, K. L. 2005a. Hepatitis B virus X protein represses E-cadherin expression via activation of DNA methyltransferase 1. *Oncogene*, 24, 6617-6625.
- LEE, W. C., WANG, H. C., HUNG, C. F., HUANG, P. F., LIA, C. R. & CHEN, M. F. 2005b. Vaccination of advanced hepatocellular carcinoma patients with tumor lysate-pulsed dendritic cells: a clinical trial. *J Immunother*, 28, 496-504.
- LENCIONI, R., DE BAERE, T., SOULEN, M. C., RILLING, W. S. & GESCHWIND, J. F. 2016. Lipiodol transarterial chemoembolization for hepatocellular carcinoma: A systematic review of efficacy and safety data. *Hepatology*, 64, 106-16.
- LI, J., LIU, H., YU, J. & YU, H. 2015. Chemoresistance to doxorubicin induces epithelial-mesenchymal transition via upregulation of transforming growth factor beta signaling in HCT116 colon cancer cells. *Mol Med Rep*, 12, 192-8.
- LI, M., ZHAO, H., ZHANG, X., WOOD, L. D., ANDERS, R. A., CHOTI, M. A., PAWLIK, T. M., DANIEL, H. D., KANNANGAI, R., OFFERHAUS, G. J., VELCULESCU, V. E., WANG, L., ZHOU, S., VOGELSTEIN, B., HRUBAN, R. H., PAPADOPOULOS, N., CAI, J., TORBENSON, M. S. & KINZLER, K. W. 2011. Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma. *Nat Genet*, 43, 828-9.
- LI, M. Y., ZHU, M., FENG, F., CAI, F. Y., FAN, K. C., JIANG, H., WANG, Z. Q. & LINGHU, E. Q. 2014. Long interspersed nucleotide acid element-1 ORF-1 protein promotes proliferation and invasion of human colorectal cancer LoVo cells through enhancing ETS-1 activity. *Genet Mol Res*, 13, 6981-94.

- LI, Z., LIU, Q., PIAO, J., HUA, F., WANG, J., JIN, G., LIN, Z. & ZHANG, Y. 2016. Clinicopathological implications of Tiam1 overexpression in invasive ductal carcinoma of the breast. *BMC cancer*, 16, 681-681.
- LIANG, G., CHAN, M. F., TOMIGAHARA, Y., TSAI, Y. C., GONZALES, F. A., LI, E., LAIRD, P. W. & JONES, P. A. 2002. Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol Cell Biol*, 22, 480-91.
- LIM, J. S., PARK, S. H. & JANG, K. L. 2012. Hepatitis C virus Core protein overcomes stress-induced premature senescence by down-regulating p16 expression via DNA methylation. *Cancer Lett*, 321, 154-61.
- LIN, S. & GREGORY, R. I. 2015. MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer*, 15, 321-33.
- LIN, T., PONN, A., HU, X., LAW, B. K. & LU, J. 2010. Requirement of the histone demethylase LSD1 in Snai1-mediated transcriptional repression during epithelial-mesenchymal transition. *Oncogene*, 29, 4896-904.
- LINDTNER, S., FELBER, B. K. & KJEMS, J. 2002. An element in the 3' untranslated region of human LINE-1 retrotransposon mRNA binds NXF1(TAP) and can function as a nuclear export element. *Rna*, 8, 345-56.
- LIOU, Y. C., ZHOU, X. Z. & LU, K. P. 2011. Prolyl isomerase Pin1 as a molecular switch to determine the fate of phosphoproteins. *Trends Biochem Sci*, 36, 501-14.
- LITCHFIELD, D. W., SHILTON, B. H., BRANDL, C. J. & GYENIS, L. 2015. Pin1: Intimate involvement with the regulatory protein kinase networks in the global phosphorylation landscape. *Biochim Biophys Acta*, 1850, 2077-86.
- LIU, A. M., YAO, T. J., WANG, W., WONG, K. F., LEE, N. P., FAN, S. T., POON, R. T., GAO, C. & LUK, J. M. 2012. Circulating miR-15b and miR-130b in serum as potential markers for detecting hepatocellular carcinoma: a retrospective cohort study. *BMJ Open*, 2, e000825.
- LIU, Y.-S., LIN, C.-Y., CHUANG, M.-T., LIN, C.-Y., TSAI, Y.-S., WANG, C.-K. & OU, M.-C. 2018. Five-year outcome of conventional and drug-eluting transcatheter arterial chemoembolization in patients with hepatocellular carcinoma. *BMC Gastroenterology*, 18, 124.
- LLOVET, J. M., BRU, C. & BRUIX, J. 1999a. Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin Liver Dis*, 19, 329-38.
- LLOVET, J. M., DI BISCEGLIE, A. M., BRUIX, J., KRAMER, B. S., LENCIONI, R., ZHU, A. X., SHERMAN, M., SCHWARTZ, M., LOTZE, M., TALWALKAR, J. & GORES, G. J. 2008a. Design and endpoints of clinical trials in hepatocellular carcinoma. *J Natl Cancer Inst*, 100, 698-711.
- LLOVET, J. M., FUSTER, J. & BRUIX, J. 1999b. Intention-to-treat analysis of surgical treatment for early hepatocellular carcinoma: Resection versus transplantation. *Hepatology*, 30, 1434-1440.
- LLOVET, J. M., RICCI, S., MAZZAFERRO, V., HILGARD, P., GANE, E., BLANC, J. F., DE OLIVEIRA, A. C., SANTORO, A., RAOUL, J. L., FORNER, A., SCHWARTZ, M., PORTA, C., ZEUZEM, S., BOLONDI, L., GRETEN, T. F., GALLE, P. R., SEITZ, J. F., BORBATH, I., HAUSSINGER, D., GIANNARIS, T., SHAN, M., MOSCOVICI, M., VOLIOTIS, D. & BRUIX, J. 2008b. Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med*, 359, 378-90.
- LU, K. P., LIOU, Y. C. & ZHOU, X. Z. 2002. Pinning down proline-directed phosphorylation signaling. *Trends Cell Biol*, 12, 164-72.
- LU, L. C., SHAO, Y. Y., LEE, Y. H., HSIEH, M. S., HSIAO, C. H., LIN, H. H., KAO, H. F., MA, Y. Y., YEN, F. C., CHENG, A. L. & HSU, C. H. 2014. beta-catenin (CTNNB1) mutations are not associated with prognosis in advanced hepatocellular carcinoma. *Oncology*, 87, 159-66.
- LU, Y., FENG, F., YANG, Y., GAO, X., CUI, J., ZHANG, C., ZHANG, F., XU, Z., QV, J., WANG, C., ZENG, Z., ZHU, Y. & YANG, Y. 2013. LINE-1 ORF-1p functions as a novel androgen receptor co-

- activator and promotes the growth of human prostatic carcinoma cells. *Cell Signal*, 25, 479-89.
- LUCIFORA, J., ARZBERGER, S., DURANTEL, D., BELLONI, L., STRUBIN, M., LEVRERO, M., ZOULIM, F., HANTZ, O. & PROTZER, U. 2011. Hepatitis B virus X protein is essential to initiate and maintain virus replication after infection. *J Hepatol*, 55, 996-1003.
- MACHIDA, K., LIU, J. C., MCNAMARA, G., LEVINE, A., DUAN, L. & LAI, M. M. 2009. Hepatitis C virus causes uncoupling of mitotic checkpoint and chromosomal polyploidy through the Rb pathway. *J Virol*, 83, 12590-600.
- MADDEN, C. R., FINEGOLD, M. J. & SLAGLE, B. L. 2002. Altered DNA mutation spectrum in aflatoxin b1-treated transgenic mice that express the hepatitis B virus x protein. *J Virol*, 76, 11770-4.
- MAJUMDER, M., GHOSH, A. K., STEELE, R., RAY, R. & RAY, R. B. 2001. Hepatitis C virus NS5A physically associates with p53 and regulates p21/waf1 gene expression in a p53-dependent manner. *J Virol*, 75, 1401-7.
- MANDELL, D. J., CHORNY, I., GROBAN, E. S., WONG, S. E., LEVINE, E., RAPP, C. S. & JACOBSON, M. P. 2007. Strengths of hydrogen bonds involving phosphorylated amino acid side chains. *J Am Chem Soc*, 129, 820-7.
- MANDREKAR, P. & SZABO, G. 2009. Signalling pathways in alcohol-induced liver inflammation. *J Hepatol*, 50, 1258-66.
- MAO, Y., YANG, H., XU, H., LU, X., SANG, X., DU, S., ZHAO, H., CHEN, W., XU, Y., CHI, T., YANG, Z., CAI, J., LI, H., CHEN, J., ZHONG, S., MOHANTI, S. R., LOPEZ-SOLER, R., MILLIS, J. M., HUANG, J. & ZHANG, H. 2010. Golgi protein 73 (GOLPH2) is a valuable serum marker for hepatocellular carcinoma. *Gut*, 59, 1687-93.
- MARRERO, J. A., FENG, Z., WANG, Y., NGUYEN, M. H., BEFELER, A. S., ROBERTS, L. R., REDDY, K. R., HARNOIS, D., LLOVET, J. M., NORMOLLE, D., DALHGREN, J., CHIA, D., LOK, A. S., WAGNER, P. D., SRIVASTAVA, S. & SCHWARTZ, M. 2009. Alpha-fetoprotein, des-gamma carboxyprothrombin, and lectin-bound alpha-fetoprotein in early hepatocellular carcinoma. *Gastroenterology*, 137, 110-8.
- MARTIN, S. L. 2006. The ORF1 protein encoded by LINE-1: structure and function during L1 retrotransposition. *J Biomed Biotechnol*, 2006, 45621.
- MARTIN, S. L. 2010. Nucleic acid chaperone properties of ORF1p from the non-LTR retrotransposon, LINE-1. *RNA Biol*, 7, 706-11.
- MARTIN, S. L., CRUCEANU, M., BRANCIFORTE, D., WAI-LUN LI, P., KWOK, S. C., HODGES, R. S. & WILLIAMS, M. C. 2005. LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. *J Mol Biol*, 348, 549-61.
- MASSAGUE, J. 1998. TGF-beta signal transduction. *Annu Rev Biochem*, 67, 753-91.
- MATA-MOLANES, J. J., SUREDA GONZÁLEZ, M., VALENZUELA JIMÉNEZ, B., MARTÍNEZ NAVARRO, E. M. & BRUGAROLAS MASLLORENS, A. 2017. Cancer Immunotherapy with Cytokine-Induced Killer Cells. *Targeted Oncology*, 12, 289-299.
- MATSUZAKI, K. 2013. Smad phospho-isoforms direct context-dependent TGF-beta signaling. *Cytokine Growth Factor Rev*, 24, 385-99.
- MCKILLOP, I. H. & SCHRUM, L. W. 2005. Alcohol and liver cancer. *Alcohol*, 35, 195-203.
- MCMAHON, B. J. 2009. The natural history of chronic hepatitis B virus infection. *Hepatology*, 49, S45-55.
- MCMILLAN, J. P. & SINGER, M. F. 1993. Translation of the human LINE-1 element, L1Hs. *Proc Natl Acad Sci U S A*, 90, 11533-7.
- MENEGON, S., COLUMBANO, A. & GIORDANO, S. 2016. The Dual Roles of NRF2 in Cancer. *Trends Mol Med*, 22, 578-93.
- MENG, J., SUN, B., ZHAO, X., ZHANG, D., ZHAO, X., GU, Q., DONG, X., ZHAO, N., LIU, P. & LIU, Y. 2014. Doxycycline as an inhibitor of the epithelial-to-mesenchymal transition and vasculogenic mimicry in hepatocellular carcinoma. *Mol Cancer Ther*, 13, 3107-22.

- MIKI, Y., NISHISHO, I., HORII, A., MIYOSHI, Y., UTSUNOMIYA, J., KINZLER, K. W., VOGELSTEIN, B. & NAKAMURA, Y. 1992. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res*, 52, 643-5.
- MIZEJEWSKI, G. J. 2001. Alpha-fetoprotein structure and function: relevance to isoforms, epitopes, and conformational variants. *Exp Biol Med (Maywood)*, 226, 377-408.
- MORADPOUR, D., PENIN, F. & RICE, C. M. 2007. Replication of hepatitis C virus. *Nat Rev Microbiol*, 5, 453-63.
- MORAN, J. V., HOLMES, S. E., NAAS, T. P., DEBERARDINIS, R. J., BOEKE, J. D. & KAZAZIAN, H. H., JR. 1996. High frequency retrotransposition in cultured mammalian cells. *Cell*, 87, 917-27.
- MORIYA, K., FUJIE, H., SHINTANI, Y., YOTSUYANAGI, H., TSUTSUMI, T., ISHIBASHI, K., MATSUURA, Y., KIMURA, S., MIYAMURA, T. & KOIKE, K. 1998. The core protein of hepatitis C virus induces hepatocellular carcinoma in transgenic mice. *Nat Med*, 4, 1065-7.
- MORRIS, S. M., BAEK, J. Y., KOSZAREK, A., KANNGURN, S., KNOBLAUGH, S. E. & GRADY, W. M. 2012. Transforming growth factor-beta signaling promotes hepatocarcinogenesis induced by p53 loss. *Hepatology*, 55, 121-31.
- MORRISH, T. A., GILBERT, N., MYERS, J. S., VINCENT, B. J., STAMATO, T. D., TACCIOLI, G. E., BATZER, M. A. & MORAN, J. V. 2002. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat Genet*, 31, 159-65.
- MOUSTAKAS, A. & HELDIN, C. H. 2005. Non-Smad TGF-beta signals. *J Cell Sci*, 118, 3573-84.
- MUNAKATA, T., NAKAMURA, M., LIANG, Y., LI, K. & 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, 18159-64.
- MUOTRI, A. R., MARCHETTO, M. C., COUFAL, N. G., OEFNER, R., YEO, G., NAKASHIMA, K. & GAGE, F. H. 2010. L1 retrotransposition in neurons is modulated by MeCP2. *Nature*, 468, 443-6.
- MURATA, A., BABA, Y., WATANABE, M., SHIGAKI, H., MIYAKE, K., ISHIMOTO, T., IWATSUKI, M., IWAGAMI, S., SAKAMOTO, Y., MIYAMOTO, Y., YOSHIDA, N., NOSHO, K. & BABA, H. 2013. Methylation levels of LINE-1 in primary lesion and matched metastatic lesions of colorectal cancer. *Br J Cancer*, 109, 408-15.
- NAGATA, H., HATANO, E., TADA, M., MURATA, M., KITAMURA, K., ASECHI, H., NARITA, M., YANAGIDA, A., TAMAKI, N., YAGI, S., IKAI, I., MATSUZAKI, K. & UEMOTO, S. 2009. Inhibition of c-Jun NH2-terminal kinase switches Smad3 signaling from oncogenesis to tumor-suppression in rat hepatocellular carcinoma. *Hepatology*, 49, 1944-53.
- NAULT, J. C., MALLET, M., PILATI, C., CALDERARO, J., BIOULAC-SAGE, P., LAURENT, C., LAURENT, A., CHERQUI, D., BALABAUD, C. & ZUCMAN-ROSSI, J. 2013. High frequency of telomerase reverse-transcriptase promoter somatic mutations in hepatocellular carcinoma and preneoplastic lesions. *Nat Commun*, 4, 2218.
- NEBBIOSO, A., TAMBARO, F. P., DELL'AVERSANA, C. & ALTUCCI, L. 2018. Cancer epigenetics: Moving forward. *PLoS Genet*, 14, e1007362.
- NIEWIADOMSKA, A. M., TIAN, C., TAN, L., WANG, T., SARKIS, P. T. & YU, X. F. 2007. Differential inhibition of long interspersed element 1 by APOBEC3 does not correlate with high-molecular-mass-complex formation or P-body association. *J Virol*, 81, 9577-83.
- NIGUMANN, P., REDIK, K., MATLIK, K. & SPEEK, M. 2002. Many human genes are transcribed from the antisense promoter of L1 retrotransposon. *Genomics*, 79, 628-34.
- NISHI, H., SHAYTAN, A. & PANCHENKO, A. R. 2014. Physicochemical mechanisms of protein regulation by phosphorylation. *Front Genet*, 5, 270.
- NISHIDA, N., KUDO, M., NAGASAKA, T., IKAI, I. & GOEL, A. 2012. Characteristic patterns of altered DNA methylation predict emergence of human hepatocellular carcinoma. *Hepatology*, 56, 994-1003.
- NIU, Z. S., NIU, X. J. & WANG, W. H. 2016. Genetic alterations in hepatocellular carcinoma: An update. *World J Gastroenterol*, 22, 9069-95.

- NOUSO, K., MIYAHARA, K., UCHIDA, D., KUWAKI, K., IZUMI, N., OMATA, M., ICHIDA, T., KUDO, M., KU, Y., KOKUDO, N., SAKAMOTO, M., NAKASHIMA, O., TAKAYAMA, T., MATSUI, O., MATSUYAMA, Y. & YAMAMOTO, K. 2013. Effect of hepatic arterial infusion chemotherapy of 5-fluorouracil and cisplatin for advanced hepatocellular carcinoma in the Nationwide Survey of Primary Liver Cancer in Japan. *Br J Cancer*, 109, 1904-7.
- OSTERTAG, E. M. & KAZAZIAN, H. H., JR. 2001. Biology of mammalian L1 retrotransposons. *Annu Rev Genet*, 35, 501-38.
- OZTURK, M. 1991. p53 mutation in hepatocellular carcinoma after aflatoxin exposure. *Lancet*, 338, 1356-9.
- PADUA, D. & MASSAGUE, J. 2009. Roles of TGFbeta in metastasis. *Cell Res*, 19, 89-102.
- PARDEE, A. D. & BUTTERFIELD, L. H. 2012. Immunotherapy of hepatocellular carcinoma: Unique challenges and clinical opportunities. *Oncoimmunology*, 1, 48-55.
- PEZ, F., LOPEZ, A., KIM, M., WANDS, J. R., CARON DE FROMENTEL, C. & MERLE, P. 2013. Wnt signaling and hepatocarcinogenesis: molecular targets for the development of innovative anticancer drugs. *J Hepatol*, 59, 1107-17.
- PIERSMA, B., BANK, R. A. & BOERSEMA, M. 2015. Signaling in Fibrosis: TGF-beta, WNT, and YAP/TAZ Converge. *Front Med (Lausanne)*, 2, 59.
- PISCAGLIA, F., SVEGLIATI-BARONI, G., BARCHETTI, A., PECORELLI, A., MARINELLI, S., TIRIBELLI, C. & BELLENTANI, S. 2016. Clinical patterns of hepatocellular carcinoma in nonalcoholic fatty liver disease: A multicenter prospective study. *Hepatology*, 63, 827-38.
- POWELL, L. W., SUBRAMANIAM, V. N. & YAPP, T. R. 2000. Haemochromatosis in the new millennium. *J Hepatol*, 32, 48-62.
- PUSZYK, W., DOWN, T., GRIMWADE, D., CHOMIENNE, C., OAKEY, R. J., SOLOMON, E. & GUIDEZ, F. 2013. The epigenetic regulator PLZF represses L1 retrotransposition in germ and progenitor cells. *Embo j*, 32, 1941-52.
- RAMADORI, G., MORICONI, F., MALIK, I. & DUDAS, J. 2008. Physiology and pathophysiology of liver inflammation, damage and repair. *J Physiol Pharmacol*, 59 Suppl 1, 107-17.
- REBOUSSOU, S. & NAULT, J. C. 2020. Advances in molecular classification and precision oncology in hepatocellular carcinoma. *J Hepatol*, 72, 215-229.
- REICHL, P., DENGLER, M., VAN ZIJL, F., HUBER, H., FUHRLINGER, G., REICHEL, C., SIEGHART, W., PECK-RADOSAVLJEVIC, M., GRUBINGER, M. & MIKULITS, W. 2015. Axl activates autocrine transforming growth factor-beta signaling in hepatocellular carcinoma. *Hepatology*, 61, 930-41.
- RENEHAN, A. G., TYSON, M., EGGER, M., HELLER, R. F. & ZWAHLEN, M. 2008. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*, 371, 569-78.
- RODIC, N., SHARMA, R., SHARMA, R., ZAMPELLA, J., DAI, L., TAYLOR, M. S., HRUBAN, R. H., IACOBUZIO-DONAHUE, C. A., MAITRA, A., TORBENSON, M. S., GOGGINS, M., SHIH IE, M., DUFFIELD, A. S., MONTGOMERY, E. A., GABRIELSON, E., NETTO, G. J., LOTAN, T. L., DE MARZO, A. M., WESTRA, W., BINDER, Z. A., ORR, B. A., GALLIA, G. L., EBERHART, C. G., BOEKE, J. D., HARRIS, C. R. & BURNS, K. H. 2014. Long interspersed element-1 protein expression is a hallmark of many human cancers. *Am J Pathol*, 184, 1280-6.
- 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., PUIGGRÒS, 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.,

- AKDEMIR, K. C., ALVAREZ, E. G., BAEZ-ORTEGA, A., BEROUKHIM, R., BOUTROS, P. C., BOWTELL, D. D. L., BRORS, B., BURNS, K. H., CAMPBELL, P. J., CHAN, K., CHEN, K., CORTÉS-CIRIANO, I., DUESO-BARROSO, A., DUNFORD, A. J., EDWARDS, P. A., ESTIVILL, X., ETEMADMOGHADAM, D., FEUERBACH, L., FINK, J. L., FRENKEL-MORGENSTERN, M., GARSEED, D. W., GERSTEIN, M., GORDENIN, D. A., HAAN, D., HABER, J. E., HESS, J. M., HUTTER, B., IMIELINSKI, M., JONES, D. T. W., JU, Y. S., KAZANOV, M. D., KLIMCZAK, L. J., KOH, Y., KORBEL, J. O., KUMAR, K., LEE, E. A., LEE, J. J.-K., LI, Y., LYNCH, A. G., MACINTYRE, G., MARKOWETZ, F., MARTINCORENA, I., MARTINEZ-FUNDICHELY, A., MEYERSON, M., MIYANO, S., NAKAGAWA, H., NAVARRO, F. C. P., OSSOWSKI, S., PARK, P. J., PEARSON, J. V., PUIGGRÒS, M., et al. 2020. Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. *Nature Genetics*, 52, 306-319.
- RODRIGUEZ-MARTIN, C., CIDRE, F., FERNANDEZ-TEIJEIRO, A., GOMEZ-MARIANO, G., DE LA VEGA, L., RAMOS, P., ZABALLOS, A., MONZON, S. & ALONSO, J. 2016. Familial retinoblastoma due to intronic LINE-1 insertion causes aberrant and noncanonical mRNA splicing of the RB1 gene. *J Hum Genet*, 61, 463-6.
- SAITO, Y., KANAI, Y., SAKAMOTO, M., SAITO, H., ISHII, H. & HIROHASHI, S. 2001. Expression of mRNA for DNA methyltransferases and methyl-CpG-binding proteins and DNA methylation status on CpG islands and pericentromeric satellite regions during human hepatocarcinogenesis. *Hepatology*, 33, 561-568.
- SALEH, A., MACIA, A. & MUOTRI, A. R. 2019. Transposable Elements, Inflammation, and Neurological Disease. *Front Neurol*, 10, 894.
- SALEM, R., LEWANDOWSKI, R. J., KULIK, L., WANG, E., RIAZ, A., RYU, R. K., SATO, K. T., GUPTA, R., NIKOLAIDIS, P., MILLER, F. H., YAGHMAI, V., IBRAHIM, S. M., SENTHILNATHAN, S., BAKER, T., GATES, V. L., ATASSI, B., NEWMAN, S., MEMON, K., CHEN, R., VOGELZANG, R. L., NEMCEK, A. A., RESNICK, S. A., CHRISMAN, H. B., CARR, J., OMARY, R. A., ABECASSIS, M., BENSON, A. B., 3RD & MULCAHY, M. F. 2011. Radioembolization results in longer time-to-progression and reduced toxicity compared with chemoembolization in patients with hepatocellular carcinoma. *Gastroenterology*, 140, 497-507.e2.
- 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. & FAULKNER, G. J. 2019. LINE-1 Evasion of Epigenetic Repression in Humans. *Mol Cell*, 75, 590-604.e12.
- SANGRO, B., GOMEZ-MARTIN, C., DE LA MATA, M., INARRAIRAEGUI, M., GARRALDA, E., BARRERA, P., RIEZU-BOJ, J. I., LARREA, E., ALFARO, C., SAROBE, P., LASARTE, J. J., PEREZ-GRACIA, J. L., MELERO, I. & PRIETO, J. 2013. A clinical trial of CTLA-4 blockade with tremelimumab in patients with hepatocellular carcinoma and chronic hepatitis C. *J Hepatol*, 59, 81-8.
- SASAKI, A., IWASHITA, Y., SHIBATA, K., OHTA, M., KITANO, S. & MORI, M. 2006. Preoperative transcatheter arterial chemoembolization reduces long-term survival rate after hepatic resection for resectable hepatocellular carcinoma. *European Journal of Surgical Oncology (EJSO)*, 32, 773-779.
- SAWADA, Y., YOSHIKAWA, T., NOBUOKA, D., SHIRAKAWA, H., KURONUMA, T., MOTOMURA, Y., MIZUNO, S., ISHII, H., NAKACHI, K., KONISHI, M., NAKAGOHRI, T., TAKAHASHI, S., GOTOHDA, N., TAKAYAMA, T., YAMAO, K., UESAKA, K., FURUSE, J., KINOSHITA, T. & NAKATSURA, T. 2012. Phase I trial of a glypican-3-derived peptide vaccine for advanced hepatocellular carcinoma: immunologic evidence and potential for improving overall survival. *Clin Cancer Res*, 18, 3686-96.

- SAWADA, Y., YOSHIKAWA, T., OFUJI, K., YOSHIMURA, M., TSUCHIYA, N., TAKAHASHI, M., NOBUOKA, D., GOTOHDA, N., TAKAHASHI, S., KATO, Y., KONISHI, M., KINOSHITA, T., IKEDA, M., NAKACHI, K., YAMAZAKI, N., MIZUNO, S., TAKAYAMA, T., YAMAO, K., UESAKA, K., FURUSE, J., ENDO, I. & NAKATSURA, T. 2016. Phase II study of the GPC3-derived peptide vaccine as an adjuvant therapy for hepatocellular carcinoma patients. *Oncoimmunology*, 5, e1129483-e1129483.
- SAXTON, J. A. & MARTIN, S. L. 1998. Recombination between subtypes creates a mosaic lineage of LINE-1 that is expressed and actively retrotransposing in the mouse genome. *J Mol Biol*, 280, 611-22.
- 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. & FAULKNER, G. J. 2018. L1 retrotransposition is a common feature of mammalian hepatocarcinogenesis. *Genome Res*, 28, 639-653.
- SCHMID, C. & MARAIA, R. 1992. Transcriptional regulation and transpositional selection of active SINE sequences. *Curr Opin Genet Dev*, 2, 874-82.
- SCHMIDT, B., WEI, L., DEPERALTA, D. K., HOSHIDA, Y., TAN, P. S., SUN, X., SVENTEK, J. P., LANUTI, M., TANABE, K. K. & FUCHS, B. C. 2016. Molecular subclasses of hepatocellular carcinoma predict sensitivity to fibroblast growth factor receptor inhibition. *International journal of cancer*, 138, 1494-1505.
- SCHULZE, K., IMBEAUD, S., LETOUZE, E., ALEXANDROV, L. B., CALDERARO, J., REBOUISSOU, S., COUCHY, G., MEILLER, C., SHINDE, J., SOYSOUVANH, F., CALATAYUD, A. L., PINYOL, R., PELLETIER, L., BALABAUD, C., LAURENT, A., BLANC, J. F., MAZZAFERRO, V., CALVO, F., VILLANUEVA, A., NAULT, J. C., BIOULAC-SAGE, P., STRATTON, M. R., LLOVET, J. M. & ZUCMAN-ROSSI, J. 2015. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genet*, 47, 505-11.
- SCHWARTZ, J. D., SCHWARTZ, M., MANDELI, J. & SUNG, M. 2002. Neoadjuvant and adjuvant therapy for resectable hepatocellular carcinoma: review of the randomised clinical trials. *The Lancet Oncology*, 3, 593-603.
- SCOTT, E. C., GARDNER, E. J., MASOOD, A., CHUANG, N. T., VERTINO, P. M. & DEVINE, S. E. 2016. A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer. *Genome Res*, 26, 745-55.
- SEN, N., GUI, B. & KUMAR, R. 2014. Role of MTA1 in cancer progression and metastasis. *Cancer metastasis reviews*, 33, 879-889.
- SENTURK, S., MUMCUOGLU, M., GURSOY-YUZUGULLU, O., CINGOZ, B., AKCALI, K. C. & OZTURK, M. 2010a. Transforming growth factor-beta induces senescence in hepatocellular carcinoma cells and inhibits tumor growth. *Hepatology*, 52, 966-74.
- SENTURK, S., MUMCUOGLU, M., GURSOY-YUZUGULLU, O., CINGOZ, B., AKCALI, K. C. & OZTURK, M. 2010b. Transforming growth factor-beta induces senescence in hepatocellular carcinoma cells and inhibits tumor growth. *Hepatology*, 52, 966-974.
- SERRESI, M., SITEUR, B., HULSMAN, D., COMPANY, C., SCHMITT, M. J., LIEFTINK, C., MORRIS, B., CESARONI, M., PROOST, N., BEIJERSBERGEN, R. L., VAN LOHUIZEN, M. & GARGIULO, G. 2018. Ezh2 inhibition in Kras-driven lung cancer amplifies inflammation and associated vulnerabilities. *J Exp Med*, 215, 3115-3135.
- SETO, E., SHI, Y. & SHENK, T. 1991. YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. *Nature*, 354, 241-5.
- SHARMA, S., KELLY, T. K. & JONES, P. A. 2010. Epigenetics in cancer. *Carcinogenesis*, 31, 27-36.
- SHI, L., CHEN, S., YANG, L. & LI, Y. 2013. The role of PD-1 and PD-L1 in T-cell immune suppression in patients with hematological malignancies. *J Hematol Oncol*, 6, 74.
- SHI, X., SELUANOV, A. & GORBUNOVA, V. 2007. Cell divisions are required for L1 retrotransposition. *Mol Cell Biol*, 27, 1264-70.

- SHINMURA, R., MATSUI, O., KOBAYASHI, S., TERAYAMA, N., SANADA, J., UEDA, K., GABATA, T., KADOYA, M. & MIYAYAMA, S. 2005. Cirrhotic nodules: association between MR imaging signal intensity and intranodular blood supply. *Radiology*, 237, 512-9.
- 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., 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. & FAULKNER, G. J. 2013. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. *Cell*, 153, 101-11.
- SHUKLA, R., YUE, J., SIOUDA, M., GHEIT, T., HANTZ, O., MERLE, P., ZOULIM, F., KRUTOVSKIKH, V., TOMMASINO, M. & SYLLA, B. S. 2011. Proinflammatory cytokine TNF- α increases the stability of hepatitis B virus X protein through NF- κ B signaling. *Carcinogenesis*, 32, 978-85.
- SIGALOTTI, L., FRATTA, E., BIDOLI, E., COVRE, A., PARISI, G., COLIZZI, F., CORAL, S., MASSARUT, S., KIRKWOOD, J. M. & MAIO, M. 2011. Methylation levels of the "long interspersed nucleotide element-1" repetitive sequences predict survival of melanoma patients. *J Transl Med*, 9, 78.
- SIMON, J. A. & KINGSTON, R. E. 2009. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol*, 10, 697-708.
- SLOTKIN, R. K. & MARTIENSSEN, R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet*, 8, 272-85.
- SMITS, A. H., JANSEN, P. W., POSER, I., HYMAN, A. A. & VERMEULEN, M. 2013. Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. *Nucleic Acids Res*, 41, e28.
- SOMERAI, T. E., DONOGHUE, M. T. A., BANDLAMUDI, C., SRINIVASAN, P., CHANG, M. T., ZAMARIN, D., CADOO, K. A., GRISHAM, R. N., O'CEARBHAILL, R. E., TEW, W. P., KONNER, J. A., HENSLEY, M. L., MAKKER, V., SABBATINI, P., SPRIGGS, D. R., TROSO-SANDOVAL, T. A., CHAREN, A. S., FRIEDMAN, C., GORSKY, M., SCHWEBER, S. J., MIDDHA, S., MURALI, R., CHIANG, S., PARK, K. J., SOSLOW, R. A., LADANYI, M., LI, B. T., MUELLER, J., WEIGELT, B., ZEHIR, A., BERGER, M. F., ABU-RUSTUM, N. R., AGHAJANIAN, C., DELAIR, D. F., SOLIT, D. B., TAYLOR, B. S. & HYMAN, D. M. 2018. Clinical Utility of Prospective Molecular Characterization in Advanced Endometrial Cancer. *Clin Cancer Res*, 24, 5939-5947.
- SPANGENBERG, H. C., THIMME, R. & BLUM, H. E. 2006. Serum markers of hepatocellular carcinoma. *Semin Liver Dis*, 26, 385-90.
- STREET, A., MACDONALD, A., MCCORMICK, C. & HARRIS, M. 2005. Hepatitis C virus NS5A-mediated activation of phosphoinositide 3-kinase results in stabilization of cellular beta-catenin and stimulation of beta-catenin-responsive transcription. *J Virol*, 79, 5006-16.
- STRIBINSKIS, V. & RAMOS, K. S. 2006. Activation of human long interspersed nuclear element 1 retrotransposition by benzo(a)pyrene, an ubiquitous environmental carcinogen. *Cancer Res*, 66, 2616-20.
- SU, Y., YANG, Y., MA, Y., ZHANG, Y., RAO, W., YANG, G. & KOU, C. 2016. The Efficacy and Safety of Dendritic Cells Co-Cultured with Cytokine-Induced Killer Cell Therapy in Combination with TACE-Predominant Minimally-Invasive Treatment for Hepatocellular Carcinoma: a Meta-Analysis. *Clin Lab*, 62, 599-608.
- SUDO, T., UTSUNOMIYA, T., MIMORI, K., NAGAHARA, H., OGAWA, K., INOUE, H., WAKIYAMA, S., FUJITA, H., SHIROUZU, K. & MORI, M. 2005. Clinicopathological significance of EZH2 mRNA expression in patients with hepatocellular carcinoma. *Br J Cancer*, 92, 1754-8.
- SUN, T. Y., YAN, W., YANG, C. M., ZHANG, L. F., TANG, H. L., CHEN, Y., HU, H. X. & WEI, X. 2015. Clinical research on dendritic cell vaccines to prevent postoperative recurrence and metastasis of liver cancer. *Genet Mol Res*, 14, 16222-32.

- SUN, W. & CABRERA, R. 2018. Systemic Treatment of Patients with Advanced, Unresectable Hepatocellular Carcinoma: Emergence of Therapies. *Journal of gastrointestinal cancer*, 49, 107-115.
- SUNG, Y. K., HWANG, S. Y., PARK, M. K., FAROOQ, M., HAN, I. S., BAE, H. I., KIM, J. C. & KIM, M. 2003. Glypican-3 is overexpressed in human hepatocellular carcinoma. *Cancer Sci*, 94, 259-62.
- TABERNILLA, A., DOS SANTOS RODRIGUES, B., PIETERS, A., CAUFRIEZ, A., LEROY, K., VAN CAMPENHOUT, R., COOREMAN, A., GOMES, A. R., ARNESDOTTER, E., GIJBELS, E. & VINKEN, M. 2021. In Vitro Liver Toxicity Testing of Chemicals: A Pragmatic Approach. *International journal of molecular sciences*, 22, 5038.
- TANG, H., DELGERMAA, L., HUANG, F., OISHI, N., LIU, L., HE, F., ZHAO, L. & MURAKAMI, S. 2005. The transcriptional transactivation function of HBx protein is important for its augmentation role in hepatitis B virus replication. *J Virol*, 79, 5548-56.
- TAYLOR, M. S., LACAVALA, J., MITA, P., MOLLOY, K. R., HUANG, C. R., LI, D., ADNEY, E. M., JIANG, H., BURNS, K. H., CHAIT, B. T., ROUTH, M. P., BOEKE, J. D. & DAI, L. 2013. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. *Cell*, 155, 1034-48.
- TEACHEY, D. T., LACEY, S. F., SHAW, P. A., MELENHORST, J. J., MAUDE, S. L., FREY, N., PEQUIGNOT, E., GONZALEZ, V. E., CHEN, F., FINKLESTEIN, J., BARRETT, D. M., WEISS, S. L., FITZGERALD, J. C., BERG, R. A., APLENC, R., CALLAHAN, C., RHEINGOLD, S. R., ZHENG, Z., ROSE-JOHN, S., WHITE, J. C., NAZIMUDDIN, F., WERTHEIM, G., LEVINE, B. L., JUNE, C. H., PORTER, D. L. & GRUPP, S. A. 2016. Identification of Predictive Biomarkers for Cytokine Release Syndrome after Chimeric Antigen Receptor T-cell Therapy for Acute Lymphoblastic Leukemia. *Cancer Discov*, 6, 664-79.
- TEWEY, K. M., ROWE, T. C., YANG, L., HALLIGAN, B. D. & LIU, L. F. 1984. Adriamycin-induced DNA damage mediated by mammalian DNA topoisomerase II. *Science*, 226, 466-8.
- TOH, U., FUJII, T., SEKI, N., NIIYA, F., SHIROUZU, K. & YAMANA, H. 2006. Characterization of IL-2-activated TILs and their use in intrapericardial immunotherapy in malignant pericardial effusion. *Cancer Immunol Immunother*, 55, 1219-27.
- TOMIMARU, Y., EGUCHI, H., NAGANO, H., WADA, H., KOBAYASHI, S., MARUBASHI, S., TANEMURA, M., TOMOKUNI, A., TAKEMASA, I., UMESHITA, K., KANTO, T., DOKI, Y. & MORI, M. 2012. Circulating microRNA-21 as a novel biomarker for hepatocellular carcinoma. *J Hepatol*, 56, 167-75.
- TSUGE, M., HIRAGA, N., AKIYAMA, R., TANAKA, S., MATSUSHITA, M., MITSUI, F., ABE, H., KITAMURA, S., HATAKEYAMA, T., KIMURA, T., MIKI, D., MORI, N., IMAMURA, M., TAKAHASHI, S., HAYES, C. N. & CHAYAMA, K. 2010. HBx protein is indispensable for development of viraemia in human hepatocyte chimeric mice. *J Gen Virol*, 91, 1854-64.
- TUBIO, J. M. C., LI, Y., JU, Y. S., MARTINCORENA, I., COOKE, S. L., TOJO, M., GUNDEM, G., PIPINIKAS, C. P., ZAMORA, J., RAINE, K., MENZIES, A., ROMAN-GARCIA, P., FULLAM, A., GERSTUNG, M., SHLIEN, A., TARPEY, P. S., PAPAEMMANUIL, E., KNAPPSKOG, S., VAN LOO, P., RAMAKRISHNA, M., DAVIES, H. R., MARSHALL, J., WEDGE, D. C., TEAGUE, J. W., BUTLER, A. P., NIK-ZAINAL, S., ALEXANDROV, L., BEHJATI, S., YATES, L. R., BOLLI, N., MUDIE, L., HARDY, C., MARTIN, S., MCLAREN, S., O'MEARA, S., ANDERSON, E., MADDISON, M., GAMBLE, S., FOSTER, C., WARREN, A. Y., WHITAKER, H., BREWER, D., EELES, R., COOPER, C., NEAL, D., LYNCH, A. G., VISAKORPI, T., ISAACS, W. B., VEER, L. V., CALDAS, C., DESMEDT, C., SOTIRIOU, C., APARICIO, S., FOEKENS, J. A., EYFJORD, J. E., LAKHANI, S. R., THOMAS, G., MYKLEBOST, O., SPAN, P. N., BORRESEN-DALE, A. L., RICHARDSON, A. L., VAN DE VIJVER, M., VINCENT-SALOMON, A., VAN DEN EYNDEN, G. G., FLANAGAN, A. M., FUTREAL, P. A., JANES, S. M., BOVA, G. S., STRATTON, M. R., MCDERMOTT, U. & CAMPBELL, P. J. 2014. Mobile DNA in cancer. Extensive transduction

- of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science*, 345, 1251343.
- TURATI, F., GALEONE, C., ROTA, M., PELUCCHI, C., NEGRI, E., BAGNARDI, V., CORRAO, G., BOFFETTA, P. & LA VECCHIA, C. 2014. Alcohol and liver cancer: a systematic review and meta-analysis of prospective studies. *Ann Oncol*, 25, 1526-35.
- VAN DER POL, C. B., LIM, C. S., SIRLIN, C. B., MCGRATH, T. A., SALAMEH, J. P., BASHIR, M. R., TANG, A., SINGAL, A. G., COSTA, A. F., FOWLER, K. & MCINNES, M. D. F. 2019. Accuracy of the Liver Imaging Reporting and Data System in Computed Tomography and Magnetic Resonance Image Analysis of Hepatocellular Carcinoma or Overall Malignancy-A Systematic Review. *Gastroenterology*, 156, 976-986.
- VAN METER, M., KASHYAP, M., REZAZADEH, S., GENEVA, A. J., MORELLO, T. D., SELUANOV, A. & GORBUNOVA, V. 2014. SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nat Commun*, 5, 5011.
- VASWANI, R. G., GEHLING, V. S., DAKIN, L. A., COOK, A. S., NASVESCHUK, C. G., DUPLESSIS, M., IYER, P., BALASUBRAMANIAN, S., ZHAO, F., GOOD, A. C., CAMPBELL, R., LEE, C., CANTONE, N., CUMMINGS, R. T., NORMANT, E., BELLON, S. F., ALBRECHT, B. K., HARMANGE, J. C., TROJER, P., AUDIA, J. E., ZHANG, Y., JUSTIN, N., CHEN, S., WILSON, J. R. & GAMBLIN, S. J. 2016. Identification of (R)-N-((4-Methoxy-6-methyl-2-oxo-1,2-dihydropyridin-3-yl)methyl)-2-methyl-1-(1-(1-(2,2,2-trifluoroethyl)piperidin-4-yl)ethyl)-1H-indole-3-carboxamide (CPI-1205), a Potent and Selective Inhibitor of Histone Methyltransferase EZH2, Suitable for Phase I Clinical Trials for B-Cell Lymphomas. *J Med Chem*, 59, 9928-9941.
- VESELY, M. D., KERSHAW, M. H., SCHREIBER, R. D. & SMYTH, M. J. 2011. Natural innate and adaptive immunity to cancer. *Annu Rev Immunol*, 29, 235-71.
- VOLK, M. L., VIJAN, S. & MARRERO, J. A. 2008. A novel model measuring the harm of transplanting hepatocellular carcinoma exceeding Milan criteria. *Am J Transplant*, 8, 839-46.
- WANG, X. W., GIBSON, M. K., VERMEULEN, W., YEH, H., FORRESTER, K., STÜRZBECHER, H. W., HOEIJMAKERS, J. H. & HARRIS, C. C. 1995. Abrogation of p53-induced apoptosis by the hepatitis B virus X gene. *Cancer Res*, 55, 6012-6.
- WEI, W., GILBERT, N., OOI, S. L., LAWLER, J. F., OSTERTAG, E. M., KAZAZIAN, H. H., BOEKE, J. D. & MORAN, J. V. 2001. Human L1 retrotransposition: cis preference versus trans complementation. *Mol Cell Biol*, 21, 1429-39.
- WEN, H., QIAN, M., HE, J., LI, M., YU, Q. & LENG, Z. 2020. Inhibiting of self-renewal, migration and invasion of ovarian cancer stem cells by blocking TGF- β pathway. *PLoS One*, 15, e0230230.
- WHEELER, D. A. 2017. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell*, 169, 1327-1341.e23.
- WHITTAKER, S., MARAIS, R. & ZHU, A. X. 2010. The role of signaling pathways in the development and treatment of hepatocellular carcinoma. *Oncogene*, 29, 4989-5005.
- WILHELM, S. M., DUMAS, J., ADNANE, L., LYNCH, M., CARTER, C. A., SCHÜTZ, G., THIERAUCH, K. H. & ZOPF, D. 2011. Regorafenib (BAY 73-4506): a new oral multikinase inhibitor of angiogenic, stromal and oncogenic receptor tyrosine kinases with potent preclinical antitumor activity. *Int J Cancer*, 129, 245-55.
- WILLIAMS, G. H. & STOEBER, K. 2012. The cell cycle and cancer. *J Pathol*, 226, 352-64.
- WONG, I. H., LO, Y. M., ZHANG, J., LIEW, C. T., NG, M. H., WONG, N., LAI, P. B., LAU, W. Y., HJELM, N. M. & JOHNSON, P. J. 1999. Detection of aberrant p16 methylation in the plasma and serum of liver cancer patients. *Cancer Res*, 59, 71-3.
- WOODCOCK, D. M., LAWLER, C. B., LINSSENMEYER, M. E., DOHERTY, J. P. & WARREN, W. D. 1997. Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. *J Biol Chem*, 272, 7810-6.

- WU, K., DING, J., CHEN, C., SUN, W., NING, B. F., WEN, W., HUANG, L., HAN, T., YANG, W., WANG, C., LI, Z., WU, M. C., FENG, G. S., XIE, W. F. & WANG, H. Y. 2012. Hepatic transforming growth factor beta gives rise to tumor-initiating cells and promotes liver cancer development. *Hepatology*, 56, 2255-67.
- WU, X., LU, Y., DING, Q., YOU, G., DAI, J., XI, X., WANG, H. & WANG, X. 2014. Characterisation of large F9 deletions in seven unrelated patients with severe haemophilia B. *Thromb Haemost*, 112, 459-65.
- WYLIE, A., JONES, A. E., D'BROT, A., LU, W. J., KURTZ, P., MORAN, J. V., RAKHEJA, D., CHEN, K. S., HAMMER, R. E., COMERFORD, S. A., AMATRUDA, J. F. & ABRAMS, J. M. 2016. p53 genes function to restrain mobile elements. *Genes Dev*, 30, 64-77.
- XIAO-JIE, L., HUI-YING, X., QI, X., JIANG, X. & SHI-JIE, M. 2016. LINE-1 in cancer: multifaceted functions and potential clinical implications. *Genet Med*, 18, 431-9.
- XIE, C.-R., SUN, H., WANG, F.-Q., LI, Z., YIN, Y.-R., FANG, Q.-L., SUN, Y., ZHAO, W.-X., ZHANG, S., ZHAO, W.-X., WANG, X.-M. & YIN, Z.-Y. 2015. Integrated analysis of gene expression and DNA methylation changes induced by hepatocyte growth factor in human hepatocytes. *Molecular medicine reports*, 12, 4250-4258.
- XIE, F., ZHANG, X., LI, H., ZHENG, T., XU, F., SHEN, R., YAN, L., YANG, J. & HE, J. 2012. Adoptive immunotherapy in postoperative hepatocellular carcinoma: a systemic review. *PLoS one*, 7, e42879-e42879.
- XU, W. J., GUO, B. L., HAN, Y. G., SHI, L. & MA, W. S. 2014. Diagnostic value of alpha-fetoprotein-L3 and Golgi protein 73 in hepatocellular carcinomas with low AFP levels. *Tumour Biol*, 35, 12069-74.
- XU, X. J. & TANG, Y. M. 2014. Cytokine release syndrome in cancer immunotherapy with chimeric antigen receptor engineered T cells. *Cancer Lett*, 343, 172-8.
- YAMAZAKI, K., MASUGI, Y. & SAKAMOTO, M. 2011. Molecular pathogenesis of hepatocellular carcinoma: altering transforming growth factor-beta signaling in hepatocarcinogenesis. *Dig Dis*, 29, 284-8.
- YAN, H., PENG, B., HE, W., ZHONG, G., QI, Y., REN, B., GAO, Z., JING, Z., SONG, M., XU, G., SUI, J. & LI, W. 2013. Molecular determinants of hepatitis B and D virus entry restriction in mouse sodium taurocholate cotransporting polypeptide. *J Virol*, 87, 7977-91.
- YANG, J., JIN, X., YAN, Y., SHAO, Y., PAN, Y., ROBERTS, L. R., ZHANG, J., HUANG, H. & JIANG, J. 2017. Inhibiting histone deacetylases suppresses glucose metabolism and hepatocellular carcinoma growth by restoring FBP1 expression. *Sci Rep*, 7, 43864.
- YANG, J. D., HAINAUT, P., GORES, G. J., AMADOU, A., PLYMOTH, A. & ROBERTS, L. R. 2019. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol*, 16, 589-604.
- YANG, J. D. & ROBERTS, L. R. 2010. Hepatocellular carcinoma: A global view. *Nat Rev Gastroenterol Hepatol*, 7, 448-58.
- YANG, L., INOKUCHI, S., ROH, Y. S., SONG, J., LOOMBA, R., PARK, E. J. & SEKI, E. 2013a. Transforming growth factor-beta signaling in hepatocytes promotes hepatic fibrosis and carcinogenesis in mice with hepatocyte-specific deletion of TAK1. *Gastroenterology*, 144, 1042-1054.e4.
- YANG, N. & KAZAZIAN, H. H., JR. 2006. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol*, 13, 763-71.
- YANG, N., ZHANG, L., ZHANG, Y. & KAZAZIAN, H. H., JR. 2003. An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucleic Acids Res*, 31, 4929-40.
- YANG, Q., FENG, F., ZHANG, F., WANG, C., LU, Y., GAO, X., ZHU, Y. & YANG, Y. 2013b. LINE-1 ORF-1p functions as a novel HGF/ETS-1 signaling pathway co-activator and promotes the growth of MDA-MB-231 cell. *Cell Signal*, 25, 2652-60.
- YEKU, O., XINGHUO LI & BRENTJENS, R. J. 2017. Adoptive T-Cell Therapy for Solid Tumors. *American Society of Clinical Oncology Educational Book*, 193-204.

- YU, F., ZINGLER, N., SCHUMANN, G. & STRATLING, W. H. 2001. Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. *Nucleic Acids Res*, 29, 4493-501.
- YUE, D., ZHANG, Z., LI, J., CHEN, X., PING, Y., LIU, S., SHI, X., LI, L., WANG, L., HUANG, L., ZHANG, B., SUN, Y. & ZHANG, Y. 2015. Transforming growth factor-beta1 promotes the migration and invasion of sphere-forming stem-like cell subpopulations in esophageal cancer. *Experimental Cell Research*, 336, 141-149.
- ZHANG, X., DONG, N., YIN, L., CAI, N., MA, H., YOU, J., ZHANG, H., WANG, H., HE, R. & YE, L. 2005. Hepatitis B virus X protein upregulates survivin expression in hepatoma tissues. *Journal of medical virology*, 77 3, 374-81.
- ZHANG, Y., HUANG, J., LI, Q., CHEN, K., LIANG, Y., ZHAN, Z., YE, F., NI, W., CHEN, L. & DING, Y. 2018. Histone methyltransferase SETDB1 promotes cells proliferation and migration by interacting with Tiam1 in hepatocellular carcinoma. *BMC cancer*, 18, 539-539.
- ZHAO, H., XU, Y., MAO, Y. & ZHANG, Y. 2013. Effects of EZH2 gene on the growth and migration of hepatocellular carcinoma HepG2 cells. *Hepatobiliary surgery and nutrition*, 2, 78-83.
- ZHAO, T., JIA, H., CHENG, Q., XIAO, Y., LI, M., REN, W., LI, C., FENG, Y., FENG, Z., WANG, H. & ZHENG, J. 2017. Nifuroxazide prompts antitumor immune response of TCL-loaded DC in mice with orthotopically-implanted hepatocarcinoma. *Oncol Rep*, 37, 3405-3414.
- ZHAO, Z., WU, Q., CHENG, J., QIU, X., ZHANG, J. & FAN, H. 2010. Depletion of DNMT3A suppressed cell proliferation and restored PTEN in hepatocellular carcinoma cell. *Journal of biomedicine & biotechnology*, 2010, 737535-737535.
- ZHU, C., UTSUNOMIYA, T., IKEMOTO, T., YAMADA, S., MORINE, Y., IMURA, S., ARAKAWA, Y., TAKASU, C., ISHIKAWA, D., IMOTO, I. & SHIMADA, M. 2014. Hypomethylation of long interspersed nuclear element-1 (LINE-1) is associated with poor prognosis via activation of c-MET in hepatocellular carcinoma. *Ann Surg Oncol*, 21 Suppl 4, S729-35.
- ZHU, G., ZHANG, Y., WANG, Q., CHE, S., YANG, Y., CHEN, L. & LIN, Z. 2019. The prognostic value of Tiam1 correlates with its roles in epithelial-mesenchymal transition progression and angiogenesis in lung adenocarcinoma. *Cancer management and research*, 11, 1741-1752.
- ZHU, J., LING, Y., XU, Y., LU, M.-Z., LIU, Y.-P. & ZHANG, C.-S. 2015. Elevated expression of MDR1 associated with Line-1 hypomethylation in esophageal squamous cell carcinoma. *International journal of clinical and experimental pathology*, 8, 14392-14400.
- ZHU, R. X., SETO, W. K., LAI, C. L. & YUEN, M. F. 2016. Epidemiology of Hepatocellular Carcinoma in the Asia-Pacific Region. *Gut Liver*, 10, 332-9.
- ZHU, W., ABBAS, T. & DUTTA, A. 2005. DNA replication and genomic instability. *Adv Exp Med Biol*, 570, 249-79.
- ZHU, Y.-Z., ZHU, R., SHI, L.-G., MAO, Y., ZHENG, G.-J., CHEN, Q. & ZHU, H.-G. 2010. Hepatitis B virus X protein promotes hypermethylation of p16(INK4A) promoter through upregulation of DNA methyltransferases in hepatocarcinogenesis. *Experimental and molecular pathology*, 89, 268-275.
- ZHU, Y., FENG, F., YU, J., SONG, B., HU, M., GAO, X., WANG, Y. & ZHANG, Q. 2013. L1-ORF1p, a Smad4 interaction protein, promotes proliferation of HepG2 cells and tumorigenesis in mice. *DNA Cell Biol*, 32, 531-40.
- ZHUANG, R., LU, D., ZHUO, J., ZHANG, X., WANG, K., WEI, X., WEI, Q., WANG, W., XIE, H., ZHOU, L., XU, X. & ZHENG, S. 2017. CR6-interacting factor 1 inhibits invasiveness by suppressing TGF- β -mediated epithelial-mesenchymal transition in hepatocellular carcinoma. *Oncotarget* [Online], 8. Available: <http://europepmc.org/abstract/MED/29212264>

<https://doi.org/10.18632/oncotarget.21925>

<https://europepmc.org/articles/PMC5706910>

<https://europepmc.org/articles/PMC5706910?pdf=render> [Accessed 2017/11//].

- ZIMMERMAN, M. A., GHOBRIAL, R. M., TONG, M. J., HIATT, J. R., CAMERON, A. M., HONG, J. & BUSUTTIL, R. W. 2008. Recurrence of hepatocellular carcinoma following liver transplantation: a review of preoperative and postoperative prognostic indicators. *Arch Surg*, 143, 182-8; discussion 188.
- ZOULIM, F. & LOCARNINI, S. 2009. Hepatitis B virus resistance to nucleos(t)ide analogues. *Gastroenterology*, 137, 1593-608.e1-2.
- ZOULIM, F., SAPUTELLI, J. & SEEGER, C. 1994. Woodchuck hepatitis virus X protein is required for viral infection in vivo. *J Virol*, 68, 2026-30.