Understanding the viral diversity of Hepatitis B virus in Saudi Arabia using Next Generation Sequencing (NGS)

Ahmed Abdulkarem Alessa

BSc, MSc

Submitted for the fulfilment of the requirements of degree of Doctor of Philosophy

College of Medical, Veterinary and Life Sciences

MRC-University of Glasgow Centre for Virus Research

Institute of Infection, Immunity, and Inflammation

University of Glasgow

September 2021

ABSTRACT

Next-generation sequencing (NGS) is a powerful method for detecting the viral mutations. The aim of this study was divided into two main sections; initially developed a protocol for Hepatitis B virus (HBV) full genome sequencing using NGS and contributed to the development of in-house bioinformatics tools to analyze drug resistance and vaccine escape mutations in a cohort of HBV-positive samples from Saudi Arabia. Then, performed detailed functional analysis using either *in vitro* infection or replicon system, of selected mutation to gain insights into their role conferring resistance to currently available drugs.

To examine circulating HBV genotypes in Saudi Arabia, 64 patients with chronic hepatitis B infection were enrolled in this study. Plasma samples with known viral load were collected retrospectively from two major hospitals in Saudi Arabia. We used two sequencing approaches: i) Metagenomic approach and ii) Target enrichment. The designed probes were validated using MiSeq® platform from Illumina®. We validated a whole-genome sequencing protocol for HBV using deep sequencing, which enabled us to characterize the prevalence of HBV genotypes. Our results suggest that HBV genotype D is predominant in Saudi Arabia, as observed throughout the Middle East.

A transfection-based *in vitro* system was developed in second part of this study to investigate the effect of changes in the genomes on HBV replication. A variety of cell lines were tested, and protein expression and viral DNA were characterised using several techniques, including ELISA and western blotting. The Huh7 cell line was validated for expressing HBV antigens, and the assessment of several monoclonal antibodies was conducted using transfected cells. This system was used to explore the mutations associated with antiviral treatment resistance observed in our sequence. Two mutations from NGS outcome (rtD134E and rtD134N) utilized for *in-vitro* drug assays to evaluate the efficacy of antiviral drug.

In this study, none of the above mutant strains conferred resistance to ADV and TDF, suggesting the good, sustained antiviral efficacy of these two drugs with regard to inhibiting viral replication. This would provide strong support for the two mutations having a clear role in resistance to these drugs.

Lastly, expression of HBcAg was conducted to confirm if the inhibition of the drugs was linked to cell toxicity. The results support the initial finding with lack of core expression in ADV and TDF treated samples whereas LdT, LAM, and ETV did not demonstrate any effect on HBcAg expression.

Table of contents

ABSTRACT ..... **II**
TABLE OF CONTENTS ..... **III**
LIST OF TABLES ..... **VIII**
LIST OF FIGURES ..... **X**
ACKNOWLEDGMENTS ..... **XIII**
AUTHOR'S DECLARATION ..... **XIV**
DEFINITIONS/ ABBREVIATIONS ..... **XV**
SCIENTIFIC COMMUNICATIONS ..... **XIX**
CHAPTER 1. INTRODUCTION **20**
1.1 The Hepatitis B Virus (HBV) **20**
1.1.1 The discovery of the HBV 20
1.1.2 The classification of the HBV 21
1.1.3 HBV structure 22
1.1.4 Epidemiology 22
1.1.5 HBV transmission 23
1.1.6 The natural course of HBV infection and diagnosis 26
1.2 The HBV genome **26**
1.3 Viral proteins **28**
1.3.1 The hepatitis B surface antigen (HBsAg) 28
1.3.2 The large hepatitis B surface protein / Pre-S1 protein 29
1.3.3 X protein 29
1.3.4 Polymerase 30
1.3.5 Precore/core (C) ORF 32
1.4 The HBV replication cycle **33**
1.4.1 HBV entry into hepatocytes 33
1.4.2 Nucleocapsid trafficking & DNA entry into the nucleus 34
1.4.3 Covalently closed circular DNA formation 35
1.4.4 HBV transcription and translation 35
1.4.5 Capsid formation, reverse transcription, and DNA synthesis 36
1.4.6 Assembly and secretion 36
1.5 HBV serotypes **39**
1.6 Viral mutations **39**
1.6.1 Mutations affecting HBsAg 39

1.6.2 Mutations affecting the HBV polymerase 40
1.6.3 Mutations affecting HBeAg and HBx 41
1.7 HBV genotype 42
1.8 Laboratory investigation of HBV infection 46
1.8.1 Serum HBsAg and anti-HBs 46
1.8.2 Serum HBcAg and anti-HBc 46
1.8.3 Serum HBeAg and anti-HBe 47
1.8.4 HBV viral load assays 47
1.8.5 Quantitative serum HBsAg and HBeAg assays 47
1.8.6 HBV Genotyping 50
1.8.7 Viral mutation analysis 50
1.9 HBV prevention 50
1.9.1 Passive immunisation 50
1.9.2 Active immunisation 51
1.9.3 The post-vaccine Era: a reduction in HBV 52
1.9.4 Vaccine escape 52
1.10 Management of HBV Infection 53
1.10.1 Acute HBV infection 53
1.10.2 Chronic HBV infection 53
1.11 Antiviral therapy 53
1.11.1 Pegylated interferon alpha (PegIFNa) 54
1.11.2 Nucleos(t)ide analogues (NAs) 54
1.12 Guidelines for Currently Approved Medications 58
1.12.1 End Point of Antiviral Drugs 59
1.12.2 The Antiviral Resistance Mechanism 59
1.12.3 Comparative Worldwide Reports on Antiviral Resistance 62
1.12.4 Management of Antiviral Resistance 64
1.12.5 Prospects in the Drugs 66
1.12.6 Future Directions 68
1.13 DNA Sequencing 68
1.13.1 Sanger Sequencing 68
1.13.2 Next Generation Sequencing (NGS) 69
1.14 HBV Sequencing 74
1.14.1 Sequencing of HBV genome using Next-generation Sequencing (NGS) 74
1.14.2 Bioinformatic analysis of NGS data 76

1.15 Current Status in Saudi Arabia and Problem Statements 76
1.16 Aim of the Thesis 81
CHAPTER 2. MATERIALS AND METHODS 82
2.1 Materials 82
2.2 Methods 98
2.2.1 Sample cohort 98
2.2.2 DNA extraction 98
2.2.3 DNA Second Strand synthesis 98
2.2.4 Library Preparation 99
2.2.5 Enrichment Process 102
2.2.6 DNA purification 109
2.2.7 Measuring DNA concentration 111
2.2.8 Bacterial Techniques 112
2.2.9 Sequencing using the MiSeq® Platform 115
2.2.10 Data Analysis and Bioinformatics 118
2.2.11 Genotyping 118
2.2.12 Construction of the phylogenetic trees 120
2.2.13 Determination of Antiviral HBV Resistance and HBV Vaccine Escape Mutations 120
2.2.14 Molecular Cloning 121
2.2.15 Cell Culture 123
2.2.16 Transfection 124
2.2.17 Drug Treatment 125
2.2.18 Immunostaining 126
2.2.19 Western Blot 126
2.2.20 Measurement of HBeAg 129
2.2.21 Statistical Analysis 130
CHAPTER 3. DEVELOPMENT OF AN HBV WHOLE-GENOME SEQUENCING PROTOCOL USING NEXT-GENERATION SEQUENCING (NGS) 131
3.1 Introduction 131
3.2 Results 131
3.2.1 Sample Preparation for Next Generation Sequencing 131
3.2.2 DNA library quality 131
3.2.3 Sequencing quality scores 132
3.2.4 Sequence alignments 132
3.2.5 Metagenomic sequencing 136

3.2.6 Nimbelgen target enrichment 138

3.2.7 DNA Enrichment using host depletion utilising the NEBNext Microbiome..... 142

3.2.8 Statistical analysis of the four sequencing runs: metagenomics, enrichment with NimbleGen, enrichment with NimbleGen/ NEBNext Microbiome (high viral loads), and enrichment with NimbleGen/ NEBNext Microbiome (low viral loads) 148

3.2.9 HBV Phylogenetic analysis 153

3.3 DISCUSSION 159

CHAPTER 4. ESTABLISHMENT OF INFECTIOUS/NON-INFECTIOUS HBV REPLICATION SYSTEMS 163

4.1 Introduction 163

4.1.1 The history of infection/replicon systems for studying HBV 163

4.2 Results 175

4.2.1 Subcloning of the HBV sequences from pTHBV-L$^-$ into the retrovirus transfer vector pQCXIP 175

4.2.2 Subcloning of HBV 1.3mer WT (wild type) sequences in pQCXIP..... 178

4.2.3 Generation and Subcloning of the retrovirus transfer vector pQCXIP pQCXIP carrying 2.7 of length HBV genome 181

4.2.4 Generation of plasmid pTHBV-L$^-$ /Puro 183

4.2.5 Analysis of intracellular HBV core, DNA synthesis and HBc protein expression in Hep38.7-Tet cells 186

4.2.6 Optimisation of transfection efficiency of HBV plasmids 188

4.2.7 Expression of HBcAg in HepG2-hNTCP-C4 cells by Western blot analysis 190

4.2.8 Assessment of various primary antibodies for detection of HBcAg expression by Western blot analysis..... 191

4.2.9 Expression of HBcAg and HBsAg in Huh7 cells transfected with different plasmids. 193

4.2.10 Analysis of HBcAg and HBsAg in transfected Huh7 cells by immunofluorescence microscopy 197

4.2.11 Introduction to HDV as a model for HBV infection 198

4.2.12 Introduction to ELISA assay for HBeAg quantitation 203

4.3 DISCUSSION 212

CHAPTER 5. FUNCTIONAL EVALUATION OF HBV POLYMERASE MUTATIONS ASSOCIATED WITH RESISTANCE TO CLINCIALLY APPROVED ANTIVIRAL DRUGS 215

5.1 Introduction 215

5.2 Results 219

5.2.1 Sub-cloning of custom-synthesized fragments carrying the 3-null or 3-null plus individual pol mutation of interest into pTHBV-L$^-$ /Puro 219

5.2.2 Time course of HBeAg expression 226

5.2.3 Screening pTHBV-L$^-$ null-Puro No (11) /RT mutant construct for HBeAg expression using ELISA 227
5.2.4 Functional evaluation of drug resistance-associated mutations..... 229
5.2.5 Testing of the activity of each drug against pTHBV- L$^-$ 3-null-Puro No 11 and the related RT mutant constructs by Western blot 239
5.3 DISCUSSION 243
CHAPTER 6. CONCLUSIONS AND FUTURE WORK 245
6.1 Background 245
6.2 Summary of the thesis 245
6.3 Research impact 251
6.4 Concluding remarks 251
6.4.1 Future Work 251
6.4.2 Prospective of HBV in Saudi Arabia 252
6.4.3 Prospective of HBV in Worldwide..... 252
APPENDICES 253
BIBLIOGRAPHY 288

LIST OF TABLES

Table 1-1: Hepatitis B virus genotypes. 45
Table 1-2: Hepatitis B virus infection: Clinical course and serological profile. 49
Table 1-3: Summary of therapeutic target of the anti-HBV drugs. 57
Table 1-4: International treatment guidelines for Hepatitis B virus. 59
Table 1-5: Summary of the Recommendations as Practice Guidelines for Patients with Drug-resistant Virus. 65
Table 1-6: Potential HBV screening points in Saudi Arabia. 78
Table 2-1: Kits 82
Table 2-2: Equipment 83
Table 2-3: Cell Culture Reagents Utilised 84
Table 2-4: Multipurpose Chemicals/Reagents used in the Study 85
Table 2-5: Illumina Sequencing Reagents 86
Table 2-6: Buffers made in-house 86
Table 2-7: Transfection Reagents 87
Table 2-8: SDS-PAGE and Western Blot Reagent 87
Table 2-9: Bacterial Propagation 88
Table 2-10: DNA Manipulation 88
Table 2-11: DNA Analysis 89
Table 2-12: Cell Lysis 89
Table 2-13: Cell lines 90
Table 2-14: Plasmids 91
Table 2-15: Primer Sequences 93
Table 2-16: Enzymes 95
Table 2-17: List of Primary Antibodies 96
Table 2-18: List of Secondary Antibodies 97
Table 2-19: Hepatitis B Virus Drugs 97
Table 2-20: The Parameters used for the Physical Fragmentation of DNA 100
Table 2-21: The Parameter of Polymerase chain reaction (PCR) cycles 121
Table 2-22: DNA ligation reaction recipe 122
Table 3-1: Overview of all results generated using the metagenomics approach. 137
Table 3-2: Overview of the results generated using the target enrichment (enriched with NimbleGen) approach. 139
Table 3-3: Target enrichment versus metagenomic (unenriched) sequencing. 140

Table 3-4: Metagenomic (KAPA Biosystems) versus metagenomic (Nextera XT®) versus metagenomic (Nextera XT®) enriched with NEBNext Microbiome DNA. 143

Table 3-5: Overview of results generated using enrichment methods with NEBNext Microbiome and NimbleGen. * Control sample was sequenced and earlier enriched with NimbleGen indicating 0% mapped read and approximately 6% genome coverage. 144

Table 3-6: Overview of results generated using enrichment methods with NEBNext Microbiome and NimbleGen for samples with low viral loads. 145

Table 3-7: Descriptive statistics for viral load (IU/mL) in all NGS runs 149

Table 3-8: Descriptive statistics for total reads in all NGS runs 149

Table 3-9: Descriptive statistics for mapped read in all NGS runs 150

Table 3-10: Descriptive statistics for genotype in all NGS runs 150

Table 3-11: Coverage group crosstabulation 151

Table 3-12: Coverage group for viral load (IU/mL) and total reads 152

Table 3-13: Descriptive statistics for coverage group, coverage nt and coverage percentage 152

Table 3-14: Summary of HBV drug resistant mutations (RAMs) identified from HBV genome sequences 155

Table 5-1: Expected mutation changes in the preS1, preS2 and the overlapping Pol-encoding sequences. 218

LIST OF FIGURES

Figure 1-1: The Hepatitis B Virus Genome. 27
Figure 1-2: HBV polymerase protein structure. 31
Figure 1-3: Hepatitis B lifecycle (Thomas and Liang, 2016). 38
Figure 1-4: Global distribution of hepatitis B virus genotypes. 44
Figure 1-5: Antiviral Resistance Mechanism in HBV. 61
Figure 1-6: Illustration of main approaches of NGS (Houldcroft et al., 2017). 73
Figure 2-1: NGS Library Preparation using KAPA LTP Kit. 106
Figure 2-2: Nextera XT® workflow. 107
Figure 2-3: NEBNext Microbiome DNA Enrichment Kit. 108
Figure 2-4: Workflow for PCR Purification using Agencourt AMPure XP® beads. 110
Figure 2-5: Illumina Sequencing Workflow. 117
Figure 2-6: Bioinformatics pipeline for HBV data analysis. 119
Figure 3-1: DNA fragment size evaluated using a 4200 Tapestation® (Agilent). 133
Figure 3-2: Example Phred score data obtained from a Fastq file. 134
Figure 3-3: Read coverage across the whole HBV genome using Tanoti. 135
Figure 3-4: An illustrative example (HBV-DMM-14) to show the compare Genome coverage in target enrichment versus metagenomic sequencing. 141
Figure 3-5: An illustrative example (HBV-DMM-29) Genome coverage in metagenomic sequencing (KAPA Biosystems and Nextera XT®) versus NEBNext Microbiome DNA metagenomic (enriched). 146
Figure 3-6: Genome coverage and mapped read percentages in metagenomics, target enrichment with NimbleGen, and enrichment with NimbleGen/NEBNext Microbiome. 147
Figure 3-7: Molecular phylogenetic analysis of HBV/D sub genotype using the Maximum Likelihood method. 154
Figure 3-8: Analysis of antiviral therapy resistance mutations. 157
Figure 3-9: Analysis of vaccine escape mutations. 158
Figure 4-1: The history of an *in vitro* systems for studying HBV. 164
Figure 4-2: Structure diagram ptetHBV (Ladner et al., 1997). 169
Figure 4-3: Diagram of pTHBV1.3-L- construction. 174
Figure 4-4: Diagram of Ligation 1 construct. 176
Figure 4-5: Screening of Ligation 1 (pQ-HBV-L$^-$) by Agarose gel electrophoresis. 177
Figure 4-6: Diagram of Ligation 2 construct. 179
Figure 4-7: Screening of Ligation 2 (pQ-HBV1.3mer WT) by Agarose gel electrophoresis. 180
Figure 4-8: Diagram of Ligation 3 construct. 182

Figure 4-9: Screening of Ligation 3 (pQ-HBV2.7) by Agarose gel electrophoresis. 183

Figure 4-10: Diagram of Ligation 4 construct. 184

Figure4-11: Screening of Ligation 4 (pTHBV-L$^-$/ Puro, No 11) by Agarose gel electrophoresis. 185

Figure 4-12: Expression of HBV core in HepAD38-Tet cell line. 187

Figure 4-13: Optimisation of DNA transfection efficiency. 189

Figure 4-14: Expression of HBcAg in HepG2-hNTCP-C4 cells. 190

Figure 4-15: Assessment of different anti-HBV core antibody reactivities in a Western blot assay. 192

Figure 4-16: Expression of HBcAg in transfected cells. 194

Figure 4-17: Expression of HBcAg and HBsAg in transfected cells. 196

Figure 4-18: Expression of HBV proteins as detected by immunofluorescence. 197

Figure 4-19: Infection of Huh7 cell line with HDV particles. 199

Figure 4-20: Infection of NTCP-expressing HepG2 cell line G2N7 with HDV particles. 200

Figure 4-21: Immunofluorescence derived from potential HDV infection rescued by pSVL (D3) + pQ-HBV2.7 (3B4) co-transfection. 201

Figure 4-22: Design of ELISA for the detection of HBeAg. 204

Figure 4-23: ELISA plate design to determine optimal antibody concentration 205

Figure 4-24: Optimization of the Antibodies concentration used in ELISA to screen secretion of HBeAg. 206

Figure 4-25: ELISA validation and plasmid screening. 208

Figure 4-26: Secretion of HBeAg in transfected cells as detected using a commercial ELISA kit. 209

Figure 4-27: Comparison between commercial and in house ELISA assay. 211

Figure 5-1: Agarose gel electrophoresis of pTHBV-L-/Puro (11) and pUC-GW cleaved with MfeI and BstZ17I. 220

Figure 5-2: Agarose gel electrophoresis for 5 ligations (RT D134N, RT D134E, RT R153W, RT M204I, and pTHBV-L$^-$ /Puro (11)). 221

Figure 5-3: Agarose gel electrophoresis for 4 ligations (RT L80I, RT L91I, RT RH126Y, and pTHBV-L$^-$ / Puro (11)). 222

Figure 5-4: Screening of Lig 1 construct by Agarose gel electrophoresis. 223

Figure 5-5: Screening of Ligs 2, 3, 4, 5, 6, 7, and 8 constructs by Agarose gel electrophoresis. 224

Figure 5-6: Screening of Ligs 2, 7, and 8 constructs by Agarose gel electrophoresis. 225

Figure 5-7: Time course of HBeAg secretion. 226

Figure 5-8: Assessment of HBeAg secretion in a panel of plasmid constructs using an in-house ELISA. 228

Figure 5-9: Schematic representation of experimental design for Antiviral drug assay. 229

Figure 5-10: Effect of ADV on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E. 231

Figure 5-11: Effect of TDF on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E. 231

Figure 5-12: Effect of LdT on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E. 232

Figure 5-13: Effect of LAM on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E. 232

Figure 5-14: Effect of ETV on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E. 233

Figure 5-15: Effect of ADV on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8). 233

Figure 5-16: Effect of TDF on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8). 234

Figure 5-17: Effect of LAM on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8). 234

Figure 5-18: Effect of LdT on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8). 235

Figure 5-19: Effect of ETV on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8). 235

Figure 5-20: Normalized effect of ADV on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8). 236

Figure 5-21: Normalized effect of TDF on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8). 236

Figure 5-22: Effect of ADV and TDF on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8). 238

Figure 5-23: Expression of the intracellular Gamma tubulin in presence/ absence of HBV drugs by Western blot analysis. 240

Figure 5-24: Expression of HBcAg in Huh7 cells in presence of HBV drugs by Western blot analysis. 241

Figure 5-25: Expression of HBcAg in Huh7 cells in absence of HBV drugs by Western blot analysis. 242

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor Prof. Arvind Patel for his enormous support and continuous guidance during my Ph.D. Arvind, you have been an excellent mentor for me, and I could not have achieved this success without your guidance and patience. Thanks also to Dr Tamir Abdelrahman, my second supervisor, for his support, patience, comments, and advice. Both of you were incredibly busy but always made yourselves available to pass on your wisdom, for troubleshooting and to proofread abstracts, presentations and, most importantly, this thesis.

I am also extremely grateful to past and present members of Patel's lab. It has been an enjoyable four years in a very productive and friendly working environment. Especially, Ricardo Sanchez, it was a hard time, days and late nights but we did it. Thanks to Meredith, Ruth, Mat, Elena, Elihu, Dany, Chris, and Rajiv. Thanks to everyone at the MRC Centre for Virus Research for their hospitality and willingness to help.

Thanks to my colleagues for their encouragement and contact at all times. Thanks, Dr. Yasser Ad-Dab'bagh, Dr. Mahmood Akhtar, Ibrahim Baydoun, and Ibrahim Abdulla EL-Hassan

Special thanks to my sponsor King Fahad Specialist Hospital - Dammam (KFSH-D) and Saudi government for financial support throughout my PhD.

Many heartfelt thanks to my wife Fatimah who was with me during the pain, thanks for making my journey enjoyable, for your support during a hard time. Thanks for the two beautiful kids in my life Ali inspiration of love and Mohammed, sorry to be away from you in the last few months.

Lastly, a huge THANK YOU to my family: Mom, siblings (Mustafa, Eman, Asmahan, Ali, Hussain and Akram), my parents-in-law for their continuous support. It has already been 4 years since I moved from Saudi Arabia. I really missed you all, and I was in tears many times because of that. Soon, I'll be back.

AUTHOR'S DECLARATION

This work was completed at the University of Glasgow between October 2017- 2021 and has not been submitted for another degree. All work presented in this thesis was obtained by the author's own efforts, unless otherwise stated.

Next-generation sequencing using target enrichment technology was carried out in collaboration with Dr Ana Fillipe and Daniel Mair, NGS facility, MRC- University of Glasgow Centre for Virus Research.

The bioinformatics pipeline used in data analysis was based on scripts developed by Dr Sreenu Vattipally, research bioinformatician in the Bioinformatics unit, MRC- University of Glasgow Centre for Virus Research.

The cloning strategy was designed by my supervisor Prof Arvind Patel, MRC- University of Glasgow Centre for Virus Research.

DEFINITIONS/ ABBREVIATIONS

µg	Microgram
µl	Microliter
µM	Micromolar
AASLD	The American Association for the Study of Liver Disease
ALT	Alanine aminotransferase
Anti-HBc	Antibodies to the HBV core antigen
Anti-HBe	Antibodies to the HBV e antigen
Anti-HBs	Antibodies to the HBV surface antigen
ATP	Adenosine triphosphate
AuAg	Australia antigen
BCP	Basal core promoter
bp	Base pair
C°	Degrees celsius
cccDNA	HBV covalently closed circular DNA
cDNA	Complimentary DNA
CHB	Chronic hepatitis B
cm	Centimeter
DHBV	Duck hepatitis B virus
DMEM	Dulbecco's Modified Eagles Medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DR1	Direct repeat 1
DR2	Direct repeat
dsDNA	Double-stranded linear DNA
E. coli	*Escherichia coli*
EASL	European Association for the Study of the Liver
EDTA	Ethylene diamine tetra-acetic acid
ELISA	Enzyme-linked immunosorbent assay
Enh I	Enhancer 1
Enh II	Enhancer 2
EPI	Expanded Program for Immunization
ER	Endoplasmic reticulum
EtBr	Ethidium Bromide
EtOH	Ethanol
ETV	Entecavir
FBS	Fetal Bovine Serum
FCS	Fetal calf serum
GAPDH	Glyceraldehyde – 3 – phosphate dehydrogenase
HAV	Hepatitis A virus
HBcAg	Hepatitis B core antigen
HBeAg	Hepatitis B e antigen
HBIG	Hepatitis B immunoglobulin
HBsAb	HBsAg antibodies
HBsAg	Hepatitis B virus surface antigen
HBSP	Hepatitis B spliced protein
HBV	Hepatitis B virus

HBx	Hepatitis B X protein
HCC	Hepatocellular carcinoma
HCV	Hepatitis C Virus
HDV	Hepatitis delta virus
HEV	Hepatitis E virus
HIV	Human Immunodeficiency Virus
HT1	Hybridization buffer
IC50	The half maximal inhibitory concentration
Ig	Immunoglobulin
IPA	Isopropanol
IVDU	Intravenous drug use
kb	Kilobase
LAM	Lamivudine
LB	Luria-Bertani
LB2	Cell Lysis Buffer
LdT	Telbivudine
LHBs	Large surface protein
MCS	Multi-cloning site
mg	Milligram
MHBs	Middle surface protein
MHR	Major hydrophilic region
min	Minute
ml	Milliliter
mRNA	Messenger RNA
MTCT	mother-to-child transmission
NA	nucleos(t)ide analogues
NaCl	Sodium Chloride
NaOH	Sodium hydroxide
NF	Nuclear factor
NFH2O	Nuclease free water
ng	Nanogram
NGS	Next-generation sequencing
NPM	Nextera PCR Master Mix
NT	Neutralize Tagment Buffer
nt	Nucleotides
NTCP	Sodium taurocholate co-transporting polypeptide
OD	Optical density
Opti-MEM	Opti-Minimum Essential Medium
ORF	Open reading frame
PBS	Phosphate buffer solution
PC	HBV precore
RT-PCR	Real-time polymerase chain reaction
Peg-IFN	Pegylated interferon
pgRNA	Pregenomic RNA
rcDNA	Relaxed circular DNA
RES	Internal ribosome entry site
RNA	Ribonucleic acid
RNase H	Ribonuclease H
rpm	Revolutions per minute

rRNA	Ribosomal RNA
RT	reverse transcriptase
s	Second
SASLT	Saudi Association for the Study of Liver Disease and Transplantation
SDS	Sodium dodecyl sulphate
SHBs	Small hepatitis B surface protein
SNP	Single nucleotide polymorphisms
SOLiD	Sequencing by Oligonucleotide Ligation and Detection
SP	Spacer region
SVP	Subviral particles
TAE	Tris-acetate-EDTA
TD	Tagment DNA Buffer
TDF	Tenofovir disoproxil fumarate
TMB	3,3’5,5’-tetramethylbenzidine substrate
TP	Terminal protein
U	Unit
v/v	Volume per volume
WHO	World Health Organisation
WT	Wild type
ε	Encapsidation signa

Amino Acid Names Abbreviations

Amino acid	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Scientific communications

This work led to various scientific communications

PUBLICATION

First author: Alessa et al 2021, Development of a new deep sequencing protocol for whole HBV genome using Next-Generation Sequencing (NGS). *Manuscript in preparation*.

PRESENTATIONS

- Poster presentation: Microbiology Annual Conference Online 2021 (26 – 30 April 2021).
- Poster presentation: The International Liver Congress™ 2019, Annual Meeting of the European Association for the Study of the Liver (EASL) .10 to 14 April 2019 in Vienna, Austria.
- Poster presentation: 4th Annual Viral Genomics & Bioinformatics Training Course, 20-24 August 2019, Glasgow, United Kingdom.
- Poster presentation: Third Annual Meeting of the European Virus Bioinformatics Centre, 28-29 March 2019, Glasgow, United Kingdom.
- Poster presentation: 23rd Glasgow Virology Workshop, Saturday 10th of February 2018 in the Sir Charles Wilson Building, University of Glasgow, United Kingdom.

Chapter 1. Introduction

1.1 The Hepatitis B Virus (HBV)

1.1.1 The discovery of the HBV

The history of research on viral hepatitis began in 1963, when Nobel Prize winner Baruch S. Blumberg and colleagues reported for the first time publicly on the discovery of a new antigen named the Australia antigen (AuAg) (Blumberg, 1977; Blumberg, 2002; Gerlich, 2013). In the ensuing years, AuAg become the first specific marker of viral hepatitis, identified as the hepatitis B virus surface antigen (HBsAg). After that, viral hepatitis type B became a driving force for developing modern virus diagnostics and vaccines. Blumberg's discovery was significant for two reasons: firstly, it gave a conclusive answer to the decade-old quest of countless researchers to identify a viral cause of hepatitis. More importantly, it demonstrated that millions who were asymptomatically but chronically infected across the globe are potential candidates for developing liver cirrhosis and even liver cancer (Al-Mahtab et al., 2020).

A century earlier, interest in the basis of this disease had increased when, in 1865, a German physician named Rudolf Virchow described a patient with symptoms of epidemic jaundice, in whom the lower end of the common bile duct was clogged with a plug of mucus. This led to the term "catarrhal jaundice," because the disease was believed to be caused by catarrh due to mucus obstructing the bile duct. This confused the understanding of the aetiology of epidemic jaundice (Gruber and Virchow, 1865). Initially, the diagnosis of hepatitis required the existence of the clinical features of hepatic injury and a corresponding gross diagnosis if a liver biopsy was performed. In addition, evidence consistent with what was then known about the natural history of hepatitis was needed through the patient's history.

In the 19th century, a genuine insight into the transmissibility of viral hepatitis emerged, with epidemics of jaundice among military and civilian populations (Burns and Thompson, 2014).

In 1885, A. Lurman documented an epidemic of "serum hepatitis", showing that more than a thousand German ship workers, initially vaccinated with human lymph to prevent smallpox, subsequently became the victims of jaundice (Lurman, 1885).

During the early 20th century, the research on viral hepatitis was largely descriptive in nature, with only the causes of epidemics of jaundice being identified.

The first documented cases of hepatitis in patients were in 1930 with several serum transmitted outbreaks observed. The next reported case was associated with a massive epidemic of jaundice in American soldiers in 1942. This was one of the largest outbreaks reported among the U.S. Army, and had followed the administration of a yellow fever vaccine containing human serum (Seeff et al., 1987**;** Martin, 2003). As they suspected it was associated with the vaccine, it inspired several studies using human volunteers, and by the end of World War II, it was well recognised that infections of the liver were believed to be caused by several distinct viruses, which led to use of the term "viral hepatitis" (Block, 2016).

In 1947, the terms "hepatitis A" and "hepatitis B" were classified by a British hepatologist named F.O. MacCallum. Hepatitis A was designated as "infectious or epidemic" hepatitis, and hepatitis B was designated as "serum" hepatitis to correspond with their observed transmission routes within this nomenclature.

The development of several liver function tests during the 1950s led to the recognition of anicteric infections and the existence of chronic carriers. However, the causative agents still could not be isolated (Block, 2016). Once Blumberg and colleagues had found a specific viral marker, the vast amount of accumulated epidemiologic and clinical data, together with huge numbers of stored serum samples, enabled rapid progress in understanding hepatitis B, and revealed the existence of a vast population of chronically infected people in Asia and Africa (Block, 2016).

In 1971, using transmission electron microscopy, David Dane and colleagues discovered HBV virions composed of the core antigen, which was surrounded by a surface antigen envelope. At this stage, it was clear that AuAg was the hepatitis B surface antigen (HBsAg) (Reviewed by Gerlich, 2013).

In 1974, William Robinson discovered HBV Deoxyribonucleic acid (DNA), having detected endogenous DNA polymerase activity within the virus (Reviewed by Gerlich, 2013).

1.1.2 The classification of the HBV

The Hepatitis B virus (HBV) is a member of the hepadnavirus family. Hepadnaviruses can be found in both mammals (orthohepadnaviruses) and birds (avihepadnaviruses) (Howard, 2002; Locarnini and Roggendorf, 2013; Karayiannis, 2017). Although these two genera differ in terms of their genetic organisation, structure, and biological properties, all of the

hepadnaviruses can induce persistent infections in their natural hosts (Howard, 1994; Scaglioni et al., 1996). All hepadnaviruses share several common properties, which are a partially double-stranded genome, having a viral polymerase able to repair the gap in the DNA template, producing an excess of subviral lipoprotein particles composed of envelope proteins, and having a narrow host range, infecting only species closely related to their natural host (Howard**,** 1986).

HBV is a species of the genus Orthohepadnavirus (infecting mammals), including non-primate species such as the ground squirrel hepatitis virus, woodchuck hepatitis virus, and arctic squirrel hepatitis virus (Locarnini and Roggendorf, 2013; Karayiannis, 2017). HBV has been identified in non-human primates such as apes (chimpanzees, gorillas, and gibbons) and 'Old World monkeys'. such as macaques (Grethe et al., 2000). The 'New World primate' woolly monkey HBV is the most genetically divergent among the primate hepadnaviruses (Lanford et al., 1998). The transmission of human HBV has been found among chimpanzees who had previously been used in experimental models. More recently, baboons were under consideration as a xenograft source for human liver transplantation (Kedda et al., 2000). Recently, the discovery of hepadnavirus in bats, with a reported zoonotic potential, has sparked public health concerns (Drexler et al., 2013).

1.1.3 HBV structure

HBV is an enveloped virus with a complete virion (also known as the Dane particle) diameter of 42-45nm (Dane et al., 1970). There are two types of HBV viral particles that are usually identified in the serum of an infected individual: sub-viral particles and infectious virions (Liang, 2009; Venkatakrishnan and Zlotnick, 2016; Karayiannis, 2017). The sub-viral particles consist of an outer glycosylated envelope containing surface antigen and are usually non-infectious as they lack a viral genome (Karayiannis, 2017). A mature infectious virion is composed of an outer glycosylated envelope containing a surface antigen surrounding a nucleocapsid enclosing the viral genome and a copy of the viral polymerase (Liang, 2009; Venkatakrishnan and Zlotnick, 2016).

1.1.4 Epidemiology

It is estimated that approximately two billion people worldwide have confirmation of a past or present infection with HBV, and that 248 million individuals are chronic carriers (positive for hepatitis B surface antigen (HBsAg)) (Ott et al., 2012). The overall prevalence of HBsAg is reported to be 3.6%; however, this varies depending upon the geographic area. The

frequency of chronic HBV ranges from <2% in low-prevalence regions (the United States, Canada, and Western Europe) to 2-7% in intermediate-prevalence regions (Mediterranean countries, Japan, Central Asia, the Middle East, and parts of South America) to ≥8% in high-prevalence regions (West Africa) (Ott et al., 2012). In their most recent report, the World Health Organization estimates the global prevalence of chronic hepatitis B infection to be 3.5%, or 257 million people globally, with the majority of the burden (68%) falling on the African and Western Pacific regions (WHO, 2017). In July 2018, the World Health Organisation (WHO) reported that HBV resulted in 887,000 deaths in 2015, mostly from complications including cirrhosis and hepatocellular carcinoma (HCC) (WHO, 2018).

The majority of people with chronic hepatitis B (CHB) are infected in early childhood, with a 25% risk of developing life-threatening complications of cirrhosis or hepatocellular carcinoma as adults (Edmunds et al., 1993; Lavanchy, 2004). There has been a significant reduction in the proportion of children under the age of five years who have become chronically infected with HBV, from 4.7% during the pre-vaccination era to current estimates of 1.3% (WHO, 2017). Despite this, the estimates for this age group in the African region are the highest in the world, at around 3%. Although there has been an overall reduction in HBsAg prevalence, particularly in more developed regions, where universal HBV vaccination has a high coverage, global migration from highly endemic regions, such as sub-Saharan Africa, to low-endemicity countries is changing the local HBV burden, requiring vigilant surveillance by the local public health systems and the implementation of practical measures regarding prevention, screening, monitoring, and treatment availability.

1.1.5 HBV transmission

HBV is a blood-borne virus, present in the blood and bodily fluids (including the saliva, semen, menstrual blood and vaginal fluids) of infected individuals (Croagh and Lubel, 2014). Smaller amounts of the virus have also been demonstrated in perspiration, tears, breast milk, and urine (Lavanchy, 2004). HBV can also survive outside the body for at least seven days on environmental surfaces at room temperature and can transmit the virus with appropriate percutaneous or mucosal exposure even in the absence of visible blood (WHO, 2017). The incubation period of the hepatitis B virus is around 75 days but can vary from 30 to 180 days. It may be identified within 30 to 60 days post-infection and can persist and develop into CHB (Nelson et al., 2016; WHO, 2017).

HBV is transmitted through contact with infected body fluids, and the only natural host is humans. Blood is the primary transmission mode, but other bodily fluids have also been involved, including saliva and semen.

To date, three modes of HBV transmission have been recognized: perinatal (vertical), horizontal, and sexual transmission, with the two former modes predominating. In the areas with the highest endemicity, perinatal and early childhood transmission are the most common, with the mode of transmission varying geographically (Nelson et al., 2016). There is no reliable indication that airborne infections arise, and faeces are not a source of infection (Hou et al., 2005).

In developed countries (and also, countries with a low endemicity of HBV infection), sexual transmission and intravenous drug abuse in adolescence and adult life account for the majority of HBV transmission cases. In countries with intermediate and high levels of endemicity of HBV infection, mother-to-child and child-to-child transmission during the early years of life are the major modes of transmission of HBV infection.

Perinatal transmission involves mother-to-child transmission (MTCT) of the virus. Placental breakdown and the leakage of maternal blood during delivery, in utero infection, post-natal infection through breastmilk, a baby's ingestion of blood and tiny scratches to the baby during birth are the postulated mechanisms of perinatal transmission (Umar et al., 2013). The main factors determining perinatal transmission are the mother's hepatitis B surface antigen (HBsAg) and hepatitis B e-antigen (HBeAg) status (Hwang et al., 1985). HBeAg in the mother's serum correlates with an HBV DNA load and is considered a marker of infectivity. The highest risk factor for MTCT is a chronically infected mothers who is HBeAg positive with a high viral load (>200,000 IU/mL), with the highest risk of the child developing chronic HBV infection being transmission during the new-born period (Ott et al., 2012).). Perinatal transmission is effectively linked with a positive HBeAg and/or a high HBV DNA viral load of childbearing women. A recent report indicated that approximately 90% of babies born to HBeAg positive mothers become chronic HBV carriers compared to 5%–31% of babies born to HBeAg negative mothers (Johannessen et al., 2021).

In the Middle East, the majority of HBV infections occur through childhood and perinatal transmission. Sexual and parenteral transmission are not common modes of infection in the Middle East. Sexual transmission is uncommon on a large scale, probably due to the conservative culture of Middle Eastern societies. Nevertheless, wife-to-husband

transmission was suggested to be important in one study in Tunisia (Coursaget et al., 1994). HBV chronic infection was found in 18% (6/33) of the husbands of HBsAg carrier mothers compared to only 4-6% in the general population, while 75% (3/4) of the husbands of HBeAg positive mothers were also HBsAg positive (Coursaget et al., 1994).

Horizontal transmission, which represents the second most significant route for HBV transmission, that occurs via close contact with infected blood or body fluids, is an important cause of CHB, especially if it occurs in children (Franco et al., 2012; Lemoine and Thursz, 2017). A study conducted in the United States demonstrated, during the pre-vaccination era, the age-related risk of HBV acquisition and established chronicity with approximately 30% of those below five years-old, compared to 16% of five- to ten-year-old children, and 8% of those greater than 30 years (McMahon et al., 1985). Also, another study in Africa showed an even higher rate of 50% chronic infection in children infected before the age of two years (Coursaget et al., 1987). Early childhood horizontal transmission in sub-Saharan Africa is the predominant route, even though vertical transmission still occurs, particularly in the HIV-HBV co-infected population (Spearman et al., 2017). Moreover, HBV can survive on contaminated surfaces and household items, including toothbrushes and razors, and horizontal transmission can occur through exposure to open cuts and scratches (Nelson et al., 2016).

Unsterile injecting medical, surgical and dental equipment, intravenous drug use (IVDU), and tattooing needles are also important sources of horizontal transmission, along with certain cultural practices, such as circumcision (Nelson et al., 2016). Blood transfusions and donated organs are now an uncommon source of HBV infection due to the sensitivity of the screening practices; however, it is still a concern in high-risk healthcare environments, including emergency departments and dialysis units (Lavanchy, 2004; Nelson et al., 2016). Also, the occult hepatitis B infection (OBI) that is described by the lack of HBsAg regardless of the presence of HBV DNA blood serum is a challenging clinical entity (Said, 2011). Recently, OBI presented a significant risk to those receiving organ transplants or blood transfusions because standard donor screening with HBsAg and anti-HBc may return serologically negative results despite the presence of HBV DNA (Said, 2011; Martinez et al., 2015). There are challenges in detecting the OBI in infected individuals without detectable HBsAg due to the very low levels of viral DNA (<200 IU/mL). it has been reported that around 20% of OBI infections for all HBV serological markers even though HBV DNA is present (Said, 2011; Martinez et al., 2015).

1.1.6 The natural course of HBV infection and diagnosis

There is a wide range of clinical presentations of hepatitis B infection, from asymptomatic infection through to the rare occurrence of fulminant hepatitis, and the long-term consequences of chronic infection, such as cirrhosis, decompensated liver disease and HCC (Shepard et al., 2006). The natural course of acute or CHB is determined by a complex interplay of viral factors (genotype, mutations, and viral load), host factors (gender, age, immune response, and family history) and environmental factors (liver toxins, alcohol, and coinfections such as Hepatitis C Virus (HCV), Hepatitis D Virus (HDV), and Human Iimmunodeficiency Virus (HIV) (Croagh, 2014). For example, following acute infection, 90% of neonates develop chronic infection, whereas 20-60% of children aged under five years and 5% of adults do so (WHO, 2017).

1.2 The HBV genome

The HBV genome is a circular, partially double-stranded DNA genome of approximately 3,200 nucleotides in length (Di Bisceglie, 2009). The HBV genome consists of a minus-strand, which spans the entire genome, and a plus-strand of DNA, spanning roughly two-thirds of the genome.

HBV encodes four overlapping open-reading frames (ORFs) that are translated into a surface (HBsAg) protein, core proteins, polymerase/reverse transcriptase (RT), and HBx (Figure 1-1).

There are three forms of HBsAg (large, medium, and small or L-HBsAg, M-HBsAg and S-HBsAg, respectively), which all share an identical 226 amino acid C-terminal region, called the surface domain (Barrera et al., 2005). The M protein contains an additional 55 amino acid N- terminal domain, called the pre-S2 domain. The L protein is further extended in the N-terminal direction with the addition of the 108 or 119 amino acid pre-S1 domain. Posttranslational modifications of all three proteins into glycoproteins contribute to their biological activities (Chouteau et al., 2001; Wu et al., 2018). The differences in posttranslational modification of the different SVP and Dane particles is described later in the text (See section 1.4.6).

Meanwhile, the basal core promoter (BCP) and its adjacent pre-core region are crucial for replicating HBV (Wu et al., 2018). The polymerase gene features four domains as follows: a terminal protein (TP) region involved in priming the viral template, a spacer (SP) region, a catalytic domain with reverse transcriptase (RT) activity, and a C-terminus that has

ribonuclease H (RNase H) activity. Further, the HBx, a protein encoded by the HBV Open Reading Frame (ORF) X gene, is implicated in many intracellular signal pathways that are closely associated with cell proliferation and cell apoptosis (Wu et al., 2018).

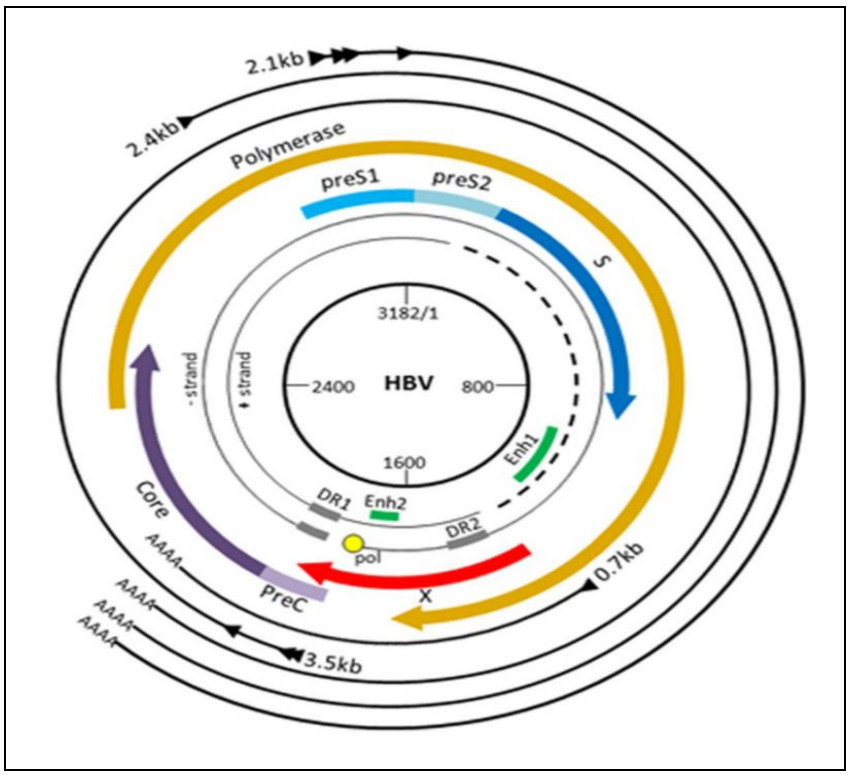


Figure 1-1: The Hepatitis B Virus Genome.

Schematic representation of the HBV genome. The genome is composed of a partially double-stranded DNA; the incomplete positive inner strand is of variable length, whereas the negative outer strand has a definite 5' and 3'-end (He et al., 2002). The genome is organized into four overlapping reading frames: the polymerase protein (coloured yellow), (Pre S1/ ORF Pre S2/ ORF S) the HBsAg protein (coloured blue), X protein (coloured red), and (C/ pre C) the core/pre-core protein (coloured pink) (Jalali and Alavian, 2006). The outer, black arrows represent the viral mRNA transcripts

1.3 Viral proteins

1.3.1 The hepatitis B surface antigen (HBsAg)

HBsAg consists of three, structurally related envelope proteins called the small (S or SHBs), middle (M, MHBs), and large (L, LHBs) surface envelope proteins, all of which exist in two forms that differ regarding the extent of their glycosylation. These mRNAs are transcribed from the Pre-S/S ORF, comprised of 389-400 codons, entirely overlapped by the Pol ORF. The Pre-S/S ORF is divided into three domains by the in-frame start codons: Pre-S1, Pre-S2, and S. They are encoded by two HBV mRNAs, 2.4kb for the L and 2.1kb for the M and S. The translation of these proteins occurs in the endoplasmic reticulum (ER) and exists in glycosylated and un-glycosylated forms (Schadler and Hildt, 2009). The nucleocapsid is surrounded by a lipid bilayer, in which the S, M and L surface proteins are anchored by the S domain to the membrane to form the viral envelope. The HBV surface proteins are also secreted in excess to virions and can bud from the ER, forming empty filamentous and spherical 22nm subviral particles (SVP) (Huovila et al., 1992; Seeger and Mason, 2000).

1.3.1.1 The small hepatitis B surface protein

The small hepatitis B surface protein (SHBs) is 226 amino acids in length and is the most abundant of the three HBV envelope proteins, comprising almost 90% of the spherical particles and 70% of the filamentous forms (Seeger and Mason, 2000). It is encoded by an in-phase start codon within the 2.1kb mRNA and is highly N-glycosylated.

The 'a' determinant within the HBsAg is a loop structure created by cross-linked cysteine residues, and is the major antigenic determinant of the HBsAg, located in the SHBs between amino acids 124 to 147. It is the major neutralisation domain of the antibodies to HBsAg, with amino acid substitutions in this region resulting in conformational changes that affect the binding of the anti-HBs antibodies that permit immune escape (Lada et al., 2006). The minor sub-determinants "d", "y", "w" or "r" vary between HBV strains based on the amino acid profile in this region, designated as the HBV serotype. Determinant d has a lysine at position 122, and y is represented by arginine; thus, the determinants are mutually exclusive. Similarly, determinant w has a lysine at position 160, and r has an arginine (Okamoto et al., 1994).

1.3.1.2 The middle hepatitis B surface (MHBs) / Pre-S2 protein

The middle hepatitis B surface protein (MHBs) is encoded by the 2.1kb RNA transcript and represents about 10% of the HBsAg on subviral particles and virions. It contains the extra pre-S2 domain, which comprises the SHBs plus an N-terminal extension of 55 amino acids (Seeger and Mason, 2000). The membrane topology of the MHBs is the same as that of the SHBs (Schadler and Hildt, 2009). The function of the MHBs is unclear, as it does not appear to be integral to viral infection (Fernholz et al., 1993; Schadler and Hildt, 2009). The Pre S2 region appears to stimulate a more remarkable B cell immune response compared to the SHBs and has been proposed to be an alternative antigen for the HBV vaccine (Milich et al., 1985).

1.3.2 The large hepatitis B surface protein / Pre-S1 protein

The large hepatitis B surface protein (LHBs) encompasses the Pre S1, Pre S2 and S domains, comprising 389-400 amino acids (depending on HBV genotype) (Churin et al., 2015; Cornberg et al., 2017). The LHBs displays a dual membrane topology on the ER, with the different orientation of the Pre S1- Pre S2 domains having important properties, including viral attachment and entry, interaction with the nucleocapsid, and the regulation of HBV replication (Schadler and Hildt, 2009). The Pre-S1 domain of LHBs is crucial for viral entry by binding to the recently discovered cellular receptor, sodium taurocholate co-transporting polypeptide (NTCP) (Seeger and Mason, 2000; Yan et al., 2012).

The N-terminal myristylation of the Pre S1 of the L protein is an essential step for viral infectivity. The N-terminal also has B and T cell epitopes that are important stimulating the cellular and humoral response against HBV infection (Milich et al., 1986; Kuroki et al., 1990).

1.3.3 X protein

The X gene codes HBV X protein (HBx), and its name were obtained due to the lack of a similar protein in the host and other viral proteins. However, similar proteins were later identified in ground squirrel and woodchuck hepatitis virus (Murakami, 1999). HBx, produced from a 0.7kb mRNA, is a polypeptide of 154 amino acids with a molecular weight of 17.5 kDa. It is predominantly localised in the cytosol, with a small amount in the nucleus (Hoare et al., 2001; Kew, 2011; Murakami, 1999). It is a highly conserved protein and its coding sequence overlaps with those of the polymerase and pre-core proteins as well as

enhancer 1 (Enh I) (X promoter) and enhancer 2 (Enh II) (core promoter) regulatory elements (Moolla et al., 2002; Kew, 2011; Geng et al., 2015). The two enhancers (Enh I and Enh II) interact with host transcription factors that bind HBV promoters and positively regulate their transcription. The Direct Repeats DR1 and DR2 present in the viral genome play an essential role in directing the initiation of viral DNA synthesis (Geng et al., 2015). These two repeats bind to form a double-stranded intermediate that allow the circularisation of the genome to enable replication, without the direct repeats replication is aborted.

The solved crystal structure of HBx demonstrated that mutations, including deletions in HBx, interfere with viral replication and transcriptional regulation (Geng et al., 2015).

HBx does not bind directly to DNA; however, it has been proposed that it is involved in transcriptional activation, which is essential for viral replication, cell cycle progression, the modulation of growth regulators, protein degradation, apoptosis, and genetic stability through interaction with the host factors, and strongly associated with hepatocellular carcinogenesis (Murakami, 2001; Block et al., 2007; Kew, 2011; Geng et al., 2015).

1.3.4 Polymerase

The polymerase ORF is the longest reading frame of the HBV genome and it is translated from the viral pregenomic RNA. The encoded polymerase (Pol) is enzymatically active and plays an essential role in viral replication. It consists of four domains (terminal, spacer, reverse transcriptase (RT) and RNase H) (Figure 1-2) (Kann, 2002; Yokosuka and Arai, 2006; Beck and Nassal, 2007)

The terminal region (181 aa long) commences with the terminal protein or primase, and it has three roles: it initiates DNA synthesis by facilitating attachment of the first nucleotide of the viral minus-strand DNA to polymerase enzyme, it facilitates binding and packaging of pgRNA into nucleocapsid, and is also essential for protein priming (Clark and Hu, 2015).

The spacer domain (165 aa long) has been previously considered non-essential for viral replication competency (Jones et al., 2014). However, it has been proven to be a highly variable site serving as a flexible stretch that allows the virus to adapt to conformational changes in response to immune selective pressure given its overlap with pre-S1 domain (Chen et al., 2013). It is located between the primase and RT domains. The envelope ORF is located within the Pol-ORF but in a frame-shifted manner, overlapped by the other ORFs to varying degrees. This is important, as viral mutations within the polymerase gene, that

lead to antiviral drug resistance, may also lead to mutations within the overlapping reading frames (Locarnini and Zoulim, 2010).

The Pol/RT domain (344 aa long) forms the catalytic core of the enzyme and is targeted by nucleos(t)ide analogues to inhibit viral replication (Clark and Hu, 2015). RT is classified into subdomains (termed 'finger', 'palm', and 'thumb') (Figure 1-2) which are further divided into conserved regions annotated A–G (Clark and Hu, 2015). Functions of the conserved regions A–G are as follows; A: contains a phenylalanine residue which discriminates against deoxyribonucleotide triphosphate (dNTP) and nucleoside triphosphate (NTP), and then interacts with regions B and G to form dNTP binding pocket, B: pairs the primer and template strand, C: contains the tyrosine-methionine-aspartate-aspartate ('YMDD') active site, D: has no role in dNTP binding, E: aids in polymerase - RNA epsilon binding and DNA.

The RNase H domain (152 aa long) has three functions as follow: it contains the aspartic acid-glutamic acid-aspartic acid-aspartic acid ('DEDD') motif which coordinates metal ion binding, which is essential for viral replication, it is responsible for the RNA cleavage of the RNA/DNA hybrids generated during reverse transcription.; and it is essential for pgRNA packaging (Kawamoto et al., 1994; Patel et al., 2017).

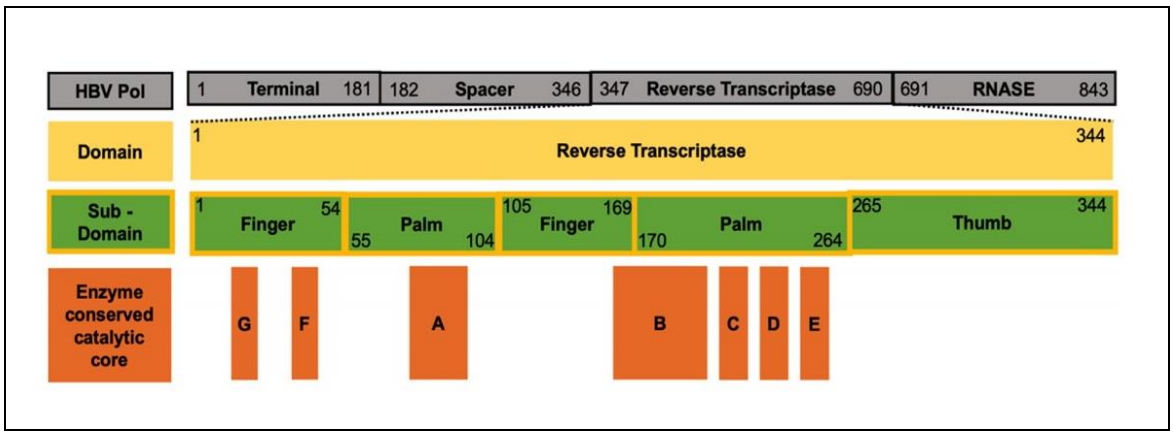


Figure 1-2: HBV polymerase protein structure.

HBV Polymerase gene encodes for the polymerase protein which is divided into four regions: terminal which has 181 amino acids; spacer which has 165 amino acids; reverse transcriptase (RT) (Yellow bar) which has 344 amino acids and RNase H which has 152 amino acids. RT protein is further classified into sub-domains (green bars): Finger, Palm, and Thumb which are further classified into seven active sites of the enzyme regions (A–G) (Wang et al., 2012; Clark and Hu, 2015).

1.3.5 Precore/core (C) ORF

The C ORF has between 212 or 214 codons, depending on the HBV genotype. It encodes two structurally and functionally distinct proteins: the nucleocapsid or core antigen (HBcAg) and the secreted hepatitis B e antigen (HBeAg) (Schödel et al., 1993; Datta et al., 2012).

1.3.5.1 HBcAg/ Core protein

HBcAg (21.5 kDa) is a structural protein and essential for the viral replication cycle. It is produced in the hepatocyte cytosol and consists of 183-185 amino acids (depending on the genotype). The assembly of the core particles requires the dimerisation of the HBcAg subunits via the first 144 amino acids, stabilised by disulfide bonds. HBcAg is divided into three regions: an N-terminal domain (amino acid residues 1–140), a linker peptide (amino acid residue 141–149) and a C-terminal domain (amino acid residue 150–183/185) (Schödel et al., 1993; Zlotnick et al., 2015).

The N-terminal domain and linker peptide are required for capsid assembly, whereas the C-terminal domain is required for pre-genomic RNA (pgRNA) packaging and reverse transcription (Liu et al., 2018). Following dimerisation and capsid formation, the pg-RNA polymerase complex is encapsidated and reverse transcribed within these core particles (replication complexes); the core particles are then enveloped with HBsAg proteins and the newly-synthesised genomic DNA complexes associated with the enveloped core are transported through the endoplasmic reticulum (ER) and Golgi complex to bud as mature HBV virions. The C-terminal region of the HBcAg contains nuclear localisation signals that facilitate the transfer of mature viral particles into the infected cell nucleus as part of the intranuclear HBV DNA amplification pathway. In addition, the nuclear localisation signalling also targets the unassembled protein at the nucleus to form empty core particles (Belloni et al., 2009). The function of these empty core particles remains unclear; however, it is hypothesised that they regulate the transcription by interacting with HBV covalently closed circular DNA (cccDNA) (Belloni et al., 2009).

1.3.5.2 Hepatitis B e antigen (HBeAg)/ Pre core protein

The HBeAg is a product of the Pre-C/C ORF and is transcribed from a distinct 3.5 kb mRNA. The HBV precore (PC) protein is non-structural, overlaps with the core protein by around 90%, sharing the central core domain and C-terminal protamine-rich domain, and

undergoes processing in the ER/Golgi complex, producing a 22 kDa protein, which can either undergo further processing through the cleavage of 10 amino acids from the N terminus to form the secreted 17 kDa HBeAg, or traffic to the cytosol, where it remains localised (Garcia et al., 1988).

HBeAg is highly conserved among hepadnaviruses (Revill et al., 2010) and plays an immunoregulatory function in the HBV life cycle. HBeAg is essential for establishing chronic infection (Milich and Liang, 2003); however, it is not required for viral replication (Beck and Nassal, 2007). It plays a role in reducing the effect of host immune response, resulting in viral persistence and the development of CHB (Walsh and Locarnini, 2012).

Approximately 15-30% of synthesised HBeAg is not secreted; this cytoplasmic HBeAg appears to play an important regulatory role of inhibiting innate immunity by suppressing Toll-like receptor signalling pathways and contribute to host immune exhaustion by forming empty capsids that are enveloped and secreted (Milich and Liang, 2003; Yang et al., 2006; Datta et al., 2012; Kramvis, 2016).

Also, the HBeAg is a crucial protein in paediatric HBV infection. It represents the only HBV protein known to cross the placenta and is proposed to act as an immune system tolerogen that is important for establishing chronic HBV infection in infants born to HBsAg-positive mothers, especially in patient mothers with a high HBV viral load (Ni, 2011).

1.4 The HBV replication cycle

The steps of the HBV replication cycle are illustrated in Figure 1-3.

1.4.1 HBV entry into hepatocytes

There are multiple steps involved in the HBV entry into hepatocyte during infection. Initially, the virus binds to factors that include heparan sulphate proteoglycans (HSPGs); for example, glypican 5 in a low-affinity and non-specific way, thereby attaching itself to the host cell's surface. Subsequently, it interacts with Na^+-taurocholate co-transporting polypeptide (NTCP) receptor in a process that is known as the first step in the replication cycle (Ni et al., 2014; Tsukuda and Watashi, 2020).

NTCP functions as a bile acid transpoter in the liver and as a host receptor for HBV. It has been shown to bind to the amino acids (aa) 2–48 of the preS1 region of the viral L envelope protein (Yan et al., 2012; Ni et al., 2014; Tsukuda and Watashi, 2020). It is belived that virus-receptor interactions initiate virus internalisation into cells receptor-dependent

endocytosis (Hao et al., 2011; Yamada et al., 2012). It was revealed in a recent report that the internalisation of HBV is activated by the receptor tyrosine kinase, epidermal growth factor receptor (EGFR), via its direct interaction with NTCP (Iwamoto et al., 2019, 2020). It is believed that internalisation of the virus in vesicles causes fusion between the cell-derived vesicular membrane and the viral envelope, although the mechanisms remain unknown. Following entry and release, the incoming nucleocapsid in the cytoplasm is directed to the nucleus along the microtubules and then imported into it through the nuclear pore complex in an importin-dependent process (Kann et al., 1999; Rabe et al., 2003; Rabe et al., 2006).

1.4.2 Nucleocapsid trafficking & DNA entry into the nucleus

The nucleocapsid is degraded at the pore and is part of a procedure known as 'uncoating' which means that only the relaxed circular DNA (rcDNA) enters the nucleus. The association between the nuclear import receptors (importin-α and β) and the nuclear localisation signalling in the C-terminal of core protein triggers (or allows for the) transportation (Morikawa et al., 2016; Diogo et al., 2021). The interaction between the adaptor proteins and the nuclear pore complex are increased by the accumulation of nucleocapsid transported along microtubules on the nuclear membrane (Schmitz et al., 2010).

Furthermore, it is apparent that nuclear localisation is dependent on the cell cycle (Ko et al., 2019; Luo et al., 2020). If cells are not in the correct cycle, the capsid is not imported into the nucleus and this prevent the packaging and assembly of the newly forming virions ((Ko et al., 2019; Luo et al., 2020). Exposure of the nuclear localisation signal initiates a structural change in the capsid, and it is generally assumed that it is controlled by and dependent on genome maturation that induces structural changes of capsid. There is evidence suggesting that the complete disassembly of capsids and release of viral genomic DNA into nucleus most possibly occurs in the nuclear basket of nuclear pore complex (Block et al., 2007; Morikawa et al., 2016; Diogo et al., 2021).

Following this, in order to expose the nucleocapsid, the virus is uncoated. The partial double-stranded DNA (rcDNA) within the nucleocapsid which is released into the cytoplasm is covalently bound to polymerase and it occurs prior to it reaching the hepatocyte nucleus through the nuclear pore complex.

At this point, the rcDNA is repaired and creates a complete double-stranded DNA either by repair activity of host proteins or through the HBV polymerase. The nucleocapsid comprises

rcDNA that is covalently cross-linked to the viral polymerase terminal protein domain. It is recognized that the C-terminal arginine-rich of core protein domain accelerates the transportation of the nucleocapsid to a nuclear pore from the cytoplasm (Morikawa et al., 2016).

1.4.3 Covalently closed circular DNA formation

In the HBV life cycle, cccDNA is a crucial stage as a replication template. The rcDNA is converted into complete circle, in a process that creates the cccDNA 'reservoir'. This conversion requires various steps: (i) removal of the of a viral polymerase linked to the 5' end of the minus strand of DNA as well as the short terminal redundant sequence in the 5' end of the negative strand; (ii) removal of the capped oligoribonucleotide RNA primer linked to the 5' end of the positive strand; (iii) after step 1 and 2, the positive strand is completed and ligation of the two viral DNA strands occurs (Yang and Kao, 2014; Morikawa et al., 2016).

With regard to the cccDNA process formation, it has been revealed that several factors have been involved. It has been demonstrated in vitro that a DNA repair enzyme, tyrosyl-DNA phosphodiesterase 2 (TDP2), cleaves the tyrosyl-DNA phosphodiester bond between the rcDNA and pol (Cui et al., 2015; Koniger et al., 2014). Another factor stated which cleaves the flap structure at the 5′ end of the minus strand by a flap structure-specific endonuclease 1 (FEN1), has also been implicated in cellular DNA replication and repair. Furthermore, numerous epigenetic elements are recruited onto the cccDNA; for instance, histones H3 and H4, transcription factors that include histone acetyltransferases and deacetylases as well as chromatin-modifying enzyme; moreover, it is apparent that the HBc and HBx proteins regulate cccDNA transcriptional activity (Morikawa et al., 2016; Karayiannis, 2017; Kitamura et al., 2018).

1.4.4 HBV transcription and translation

cccDNA acts as a template for host RNA polymerase II (recognizes cis acting elements on the viral DNA) for the transcription of the five key HBV mRNAs in the nucleus which are then exported to the cytoplasm. Translation of the viral transcripts take place within the cytoplasm subsequent to nuclear export (Lamontagne et al., 2016).

The following are the five mRNAs:

(i) 3.5 kb pgRNA encodes the viral proteins necessary for nucleocapsid assembly (HBcAg, HBV Pol/RT) which is also the template for the reverse transcription of the HBV DNA genome.

(ii) 3.4 kb precore mRNA that directs the translation of the precore gene product. Precore mRNA is translated and proteolytically processed to yield HBV e antigen (HBeAg).

(iii) 2.4 kb preS1 mRNA encode the envelope small, medium, and large surface proteins.

(iv) 2.1 kb preS2/S mRNA.

(v) The transactivating factor HBx protein is encoded by the smallest 0.7 kb mRNA.

During replication, these transcripts are translated into seven viral proteins which include (polymerase (Pol), capsid protein (Core), HBeAg, X protein, and three envelope proteins: LHBs (L), MHBs (M) and SHBs (S)) using cellular translation mechanism. There is some new evidence to suggest that there is a specific 3.9kb promoter in the genome to allow the transcription of the x gene (Doitsh & Shaul, 2003).

1.4.5 Capsid formation, reverse transcription, and DNA synthesis

Since the viral pgRNA that has an ε-stem-loop (encapsidation signal epsilon structurally contains lower stem, bulge region, upper stem and a tri-loop) is close to the 5′-end, the polymerase and the pgRNA are selectively encapsidated in core particles within the cytoplasm (Flodell, et al., 2002). The reverse transcription-polymerase enables viral DNA synthesis to be activated in each core particle. This is followed by negative-strand synthesis and concomitant degradation of the RNA template by RNase H.

Moreover, the generation and/or synthesis of positive-strand DNA synthesis is initiated by utilising short RNA from non-degraded pgRNA as a primer (Lamontagne et al., 2016; Morikawa et al., 2016). Nucleocapsids containing rcDNA either progress to the subsequent step (envelopment) or are recycled by uncoating and importation into the new rcDNA nucleus as described above.

1.4.6 Assembly and secretion

HBV produces non-infectious particles termed subviral particles (SVPs) which are produced in much more quantity than infectious virion (at least 1,000-fold), this is commonly referred to as a Dane particle. Filamentous and spherical particles have been

observed as the two kinds of SVPs. Where nucleocapsid is present, L- HBsAg interact with the coated genome in order to create a total HBV virion (Dane particle) (Morikawa et al., 2016). Filaments would have fewer L- HBsAg and M-HBsAg than virion, and spheres would be composed by S- HBsAg and M-HBsAg without and L-HBsAg. Filaments and spheres are varied in size and shape from 22 to 28 nm spheres or as variable length tubular filaments of approximately 20 nm wide, both are made of protein dimers organized in a helical symmetry (Jiang and Hildt, 2020).

For a considerable period of time, it has been assumed those viral particles and SVPs (spheres and filaments) self-assembled in the endoplasmic reticulum lumen, were transported into the ER-Golgi intermediate compartment (ERGIC) and subsequently released by the general secretory pathway. Recently, it was shown that the release pathway of HBV viral particles differs from that of spheres (Watanabe et al., 2007).

The egress and budding of HBV rely on intraluminal vesicles of maturing endosomes; multivesicular bodies (MVBs), whereas the endosomal sorting complex is necessary for the transport, the virus requires the endosomal sorting complex required for transport (ESCRT) system (Jiang and Hildt, 2020).

The ESCRT pathway is specified as being the cargoes necessary for transporting membrane proteins into vesicles that bud inwards into the maturing endosomes' lumen (MVBs). The MVBs then either fuse with the plasma membrane to release extracellular exosomes or with lysosomes to release vesicle contents (Jiang and Hildt, 2020).

It is suggested that the HBV virion packaging mechanism comprises these stages: (i) nucleocapsid is transported to the MVB surface by the cellular proteins Nedd4 (neuronal precursor that is cell-expressed developmentally down-regulated) and γ2-adaptin; and (ii) nucleocapsid buds into MVB via the endosomal sorting complex when it is in contact with HBV envelope proteins (Morikawa et al., 2016).

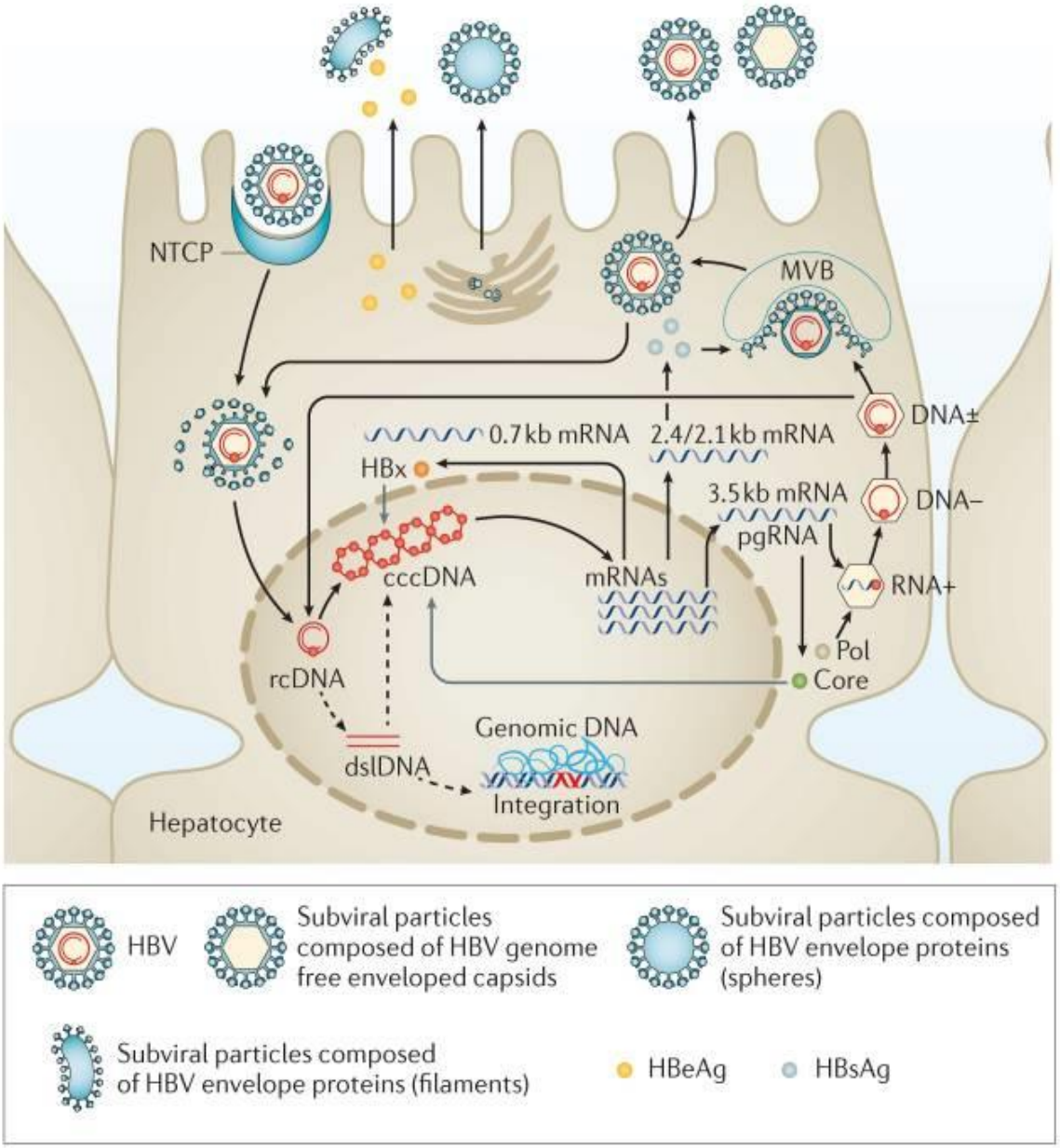


Figure 1-3: Hepatitis B lifecycle (Thomas and Liang, 2016).

HBV life cycle: HBV attaches onto the hepatocyte via sodium taurocholate co-transporting polypeptide (NTCP), and viral capsids are released and consequently directed to the nucleus. Relaxed circular DNA (rcDNA) is converted into covalently closed circular DNA (cccDNA) in the nucleus and serves as a template for viral RNA transcription. Also, double-stranded linear DNA (dsDNA) is produced that can be incorporated into the cellular genome or converted into cccDNA. Viral mRNAs are translated into viral proteins after being transported into the cytoplasm. The pregenomic RNA (pgRNA) is encapsidated together with viral polymerase and is reverse transcribed within the nucleocapsid into progeny rcDNA. Then, mature nucleocapsids are either redirected to the nucleus to establish a cccDNA pool or directed to a multivesicular body (MVB) pathway for envelopment with HBV envelope proteins (Thomas and Liang, 2016).

1.5 HBV serotypes

HBV was initially categorised into nine different serotypes (ayw1, ayw2, ayw3m, ayw4, ayr, adw2, adwq, adr, adrq) , which were based on the immune reaction according to the amino acid pattern at a specific location of "a" determinant present on the surface protein, HBsAg (Norder et al., 1992; Magnius and Norder, 1995). The classification system reflected serological reactivity rather than the phylogenetic relationships between the strains (Ohba et al., 1995; Okamoto et al., 1988) and was eventually replaced by genotyping, as introduced by Okamoto and co-workers (Okamoto et al., 1988).

1.6 Viral mutations

Heterogeneity exists in the HBV genome, primarily due to the lack of a proofreading function in the viral RT, which ensures that HBV exists in the host as a quasispecies. Mutations generated due to the virus and host factors can be selected by the immune response or by the exogenous pressure exerted by vaccination or the use of hepatitis B immunoglobulin (HBIG) and antiviral therapy. These can lead to clinically important virus variants.

1.6.1 Mutations affecting HBsAg

HBsAg is cleared during recovery from acute hepatitis B infection, and an anti-HBs immune response (>10 IU/mL) offers protective against subsequent infection. The diagnosis of HBV infection is based primarily on the detection of HBsAg. However, the presence of a mutation in the S gene had affected the diagnostic sensitivity of the current diagnostic serological test mainly in the "a" determinant, which might lead to false-negative results. The antigenic alterations may also help the virus to escape from the host immune system.

Despite the reported diversity of the "a" determinant, conserved 17 amino acid residues were identified by a Spanish study, in which the fragment encoding the region between amino acids 112–212 of HBsAg was sequenced from sera collected from chronic HBV carriers (Al-Qudari et al., 2016). The implication of the genetic variability in the S gene is well described at both the diagnostic and clinical levels. The mutations lead to the emergence of immune escape variants that are missed by the current diagnostic tests based on the detection of HBsAg, which raises the number of false negatives in our laboratories.

The mutations in the S gene will cause changes in the conformational structure of HBsAg, leading to diagnostic failure, as the antibodies used in the HBsAg detection assays fail to

bind to the distorted antigen. This diagnostic failure could result in the introduction of a mutant HBV variant in the supplies in the blood bank, starting a new global pandemic of HBV due to the risk of post-transfusion infection (Allain et al., 2012).

It was reported that substitution mutations in the S region in HBV variants cause diagnostic failure in 12.5% of the cases where HBV was misdiagnosed, and the mutant variant is also responsible for 6.6% of the cases of failed vaccination, while 9.2% escaping from immunoglobulin therapy (Avellon and Echevarrie, 2006).

Vaccine escape isolates often carry a substitution of glycine to arginine at residue 145 of HBsAg (sG145R) or aspartate to alanine/glutamine at residue 144 (sD144A/E). These variants can be viable and infectious, and the mutation position can also affect the conformation of the L and M envelope proteins as well as that of the S region.

1.6.1.1 Pre-S mutations

Naturally occurring or spontaneous point mutations can occur throughout the HBV genome, including the Pre-S region (Gerken et al., 1991). Several studies have shown that Pre-S deletion variants may be associated with more advanced liver disease in CHB patients, and recent reports suggest there may be an association with the development of HCC (Zhang et al., 2016; Fan et al., 2001; Li et al., 2016). A study reported that 38% of CHB patients with HCC had pre-S2 deletion variants, compared with only 7.14% of those without HCC (Gao et al., 2007). Similarly, another study found that the prevalence of HBV Pre-S deletions in HBV with HCC and non-HCC patients was 29.2% and 14.6%, respectively (Yeung et al., 2011).

It is important to note that the clinical reports vary regarding the significance and prevalence of Pre-S mutations. These differences may be due to variations between the study populations, including their age, HBV genotype, stage of disease, viral load and exposure to antiviral therapy (Choi et al., 2007). It was reported that mutations in the Pre-S are more prevalent in a patient infected with HBV genotype C (Sugauchi et al., 2003).

1.6.2 Mutations affecting the HBV polymerase

Mutations affecting the polymerase protein with reverse transcriptase (RT) often occur in the setting of antiviral therapy with nucleos(t)ide analogues (NA). These mutations will be discussed in further detail in the Antiviral Resistance (Sections 1.12.2 and 1.12.3).

1.6.3 Mutations affecting HBeAg and HBx

The basal core promoter (BCP) region is an element in the HBV genome that initiates the transcription of pgRNA and precore RNA (Okamoto et al., 1994). Precore (PC) and basal core promoter (BCP) mutations can abrogate or reduce HBeAg production, respectively. These mutations can occur either individually or in combination, and this may be influenced by virus genotype.

The most frequent BCP mutations emerge in two adjacent nucleotide positions, A1762T and G1764A, interfering with a transcription factor binding site in this region (Okamoto et al., 1994). Several studies have shown that this double mutation reduces the transcription of mRNA and secretion of HBeAg and enhances viral replication (Buckwold et al., 1996; Laras et al., 2002). At the same time, the most common PC mutation is a single nucleotide substitution of G to A at nucleotide (nt) 1896 (G1896A) and was first described in 1989 (Brunetto et al., 1989; Carman et al., 1989). G1896A induces a TAG stop codon in the PC region, which eliminates the production of HBeAg. Later, it was found that this mutation was also frequent in HBeAg negative patients with low HBV DNA levels. Other mutations in the precore region, e.g., in the start codon or codon 2, may also abolish HBeAg production (Lindh et al., 1996).

The occurrence of substitutions in the BCP region has also been shown to increase viral replication, even in the presence of lamivudine resistance (Tacke et al., 2004). In addition to the mutations at nt 1762 and nt 1764, others appear at nt 1753 and nt 1766.

The double mutation in the BCP region is more frequent in genotype C than genotype B and has been associated with more severe liver inflammation (Lindh et al., 1999) as well as cirrhosis and HCC (Fang et al., 2002; Kao et al., 2003; Takahashi et al., 1998). Despite the high frequency of BCP mutations, particularly of mutations at positions 1762 and 1764, and their strong association with more severe liver damage, the mechanism for their selection remains uncertain. The observation that these mutations evolve in parallel with immunologic activity indicates that they represent some form of escape. Of note, they usually lead to PC mutations and often appear during the HBeAg positive phase, indicating that their selection is driven by other mechanisms rather than PC mutations, even if they have been proposed to represent an alternative way of down-regulating the synthesis of HBeAg. A further mutation at G1899A is frequently found in association with the G1896A mutation. Other changes in the PC region which may prevent the translation of HBeAg include nonsense mutations at nt1814-1816, 1817, 1874 and 1897, and frameshift mutations

at nt1838-988. Mutations at nt1809-1812 may also result in reduced HBeAg expression (Ahn et al., 2003). Many of these changes are genotype dependent. BCP variants predominantly have a double mutation of A1762T and G1764A, resulting in the transcriptional reduction of Pre-C/C mRNA, with up to a 70% decrease in HBeAg production compared to wild type HBV (Hunt et al., 2000).

BCP mutations are more common in genotype C than genotype B (Lindh et al., 1999). They have been associated with an increased risk of developing severe inflammation (Lindh et al., 1999), cirrhosis and HCC (Lindh et al., 1999; Fang et al., 2002; Kao et al., 2003). BCP mutations have also been shown to predict HBeAg seroconversion (Chan et al., 1999; Yamaura et al., 2003). The importance of BCP mutations has been less well studied in European patients. However, one study of patients carrying genotype A and D suggest that they are associated with a worse clinical course (Jardi et al., 2004). In contrast, a report from India failed to find any clinical impact of mutations in this region in the subtypes of genotype D (Chandra et al., 2009). As the PC protein has an inhibitory effect on HBV replication, the emergence of BCP variants has been associated with high viral loads (Tacke et al., 2004).

Significantly, changes in the BCP usually directly affect the overlapping X ORF. The BCP double mutations A1762T & G1764A result in changes to X at codon 130 and 130, which are associated with increase in tumorigenicity and oncogenicity of HBV (Baptista et al., 1999).

1.7 HBV genotype

There are ten different genotypes of HBV (A – J), classified based on the genetic distance and differing by >7.5% in genomic sequence diversity. HBV is further classified into 40 different subtypes (separated by >4% genomic sequence diversity) (Pourkarim et al., 2014). Genotypes are distributed across different geographical regions and can be associated with differences in disease progression, severity and prognosis (Sunbul et al., 2014) (Figure 1-4) (Table 1-1). Genotype A is the most distributed globally and the leading genotype in Europe, North America, Africa and India, whereas B and C are the dominant genotypes in East and Southeast Asia (Norder et al., 1993; Kao and Chen, 2006). Genotype D prevails in the Middle East and Mediterranean region and India, while genotype E is commonly found in sub-Saharan Africa (Mulders et al., 2004; Kramvis et al., 2005), along with certain other continents. Genotypes F and H are closely related and almost exclusively found in the Americas. Thus, genotype F can be found in South and Central America, whereas genotype H is present in Central America and Mexico (Alvarado Mora et al., 2011;

Devesa et al., 2008), while Genotype G has been found primarily in the USA, France, and Germany (Tanwar and Dusheiko, 2012). Genotype I has been detected in Laos, Vietnam and China (Phung et al., 2010; Olinger et al., 2008), while the newest genotype, J, was identified and recovered from an 88-year-old patient in Japan, who had been in Borneo during World War II (Tatematsu et al., 2009).

The criteria for assigning a new subgenotype are; i) a comparison of whole-genome sequences, ii) a genetic distance with inter-genotypic pairwise distance >7.5% and intra-genotypic pairwise distance >4.5%, iii) a well-supported phylogenetic tree with bootstrap values of more than 75%, iv) the identification of any recombination events, v) the identification of fingerprints in the genome sequence, including amino acid motifs, vi) the presence of three clinical isolates representing the new strain, and vii) the validation of all new subgenotype strains by evolutionary and phylogeny analysis (Pourkarim et al., 2014).

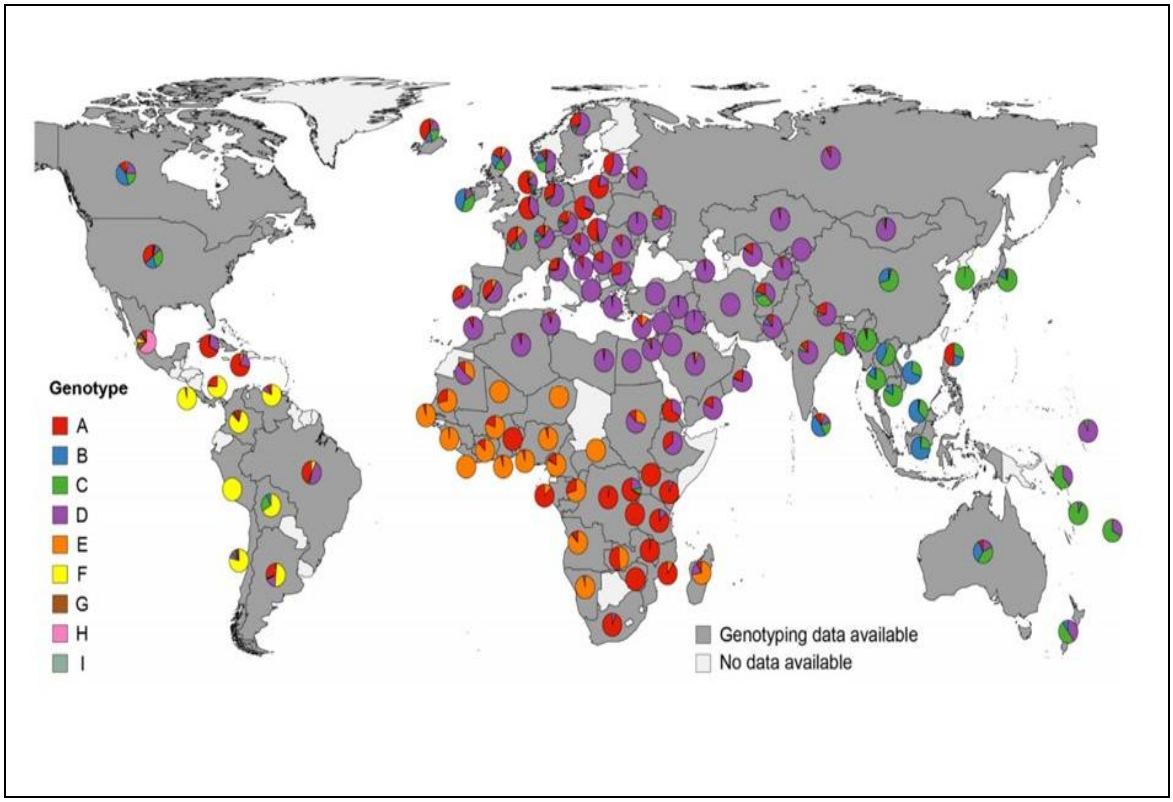


Figure 1-4: Global distribution of hepatitis B virus genotypes.

Global distribution of different HBV genotypes. Pie charts indicate proportional HBV genotype distributions in the respective countries (Velkov et al., 2018).

Table 1-1: Hepatitis B virus genotypes.

Genotype and sub genotypes	Clinical implication	Geographical location
A (1-4)	➢ Higher rates of spontaneous HBsAg seroconversion compared to the other genotypes (Ganem and Prince, 2004; Flink et al., 2006). ➢ Higher HBeAg clearance compared to genotypes B, C and D when on pegylated interferon (Janssen et al., 2005).	Africa, Europe, North America, and Asia (India)
B (1-6)	➢ Lower prevalence of HBeAg positivity and higher rates of HBeAg seroconversion (Hadziyannis and Papatheodoridis, 2006). ➢ Earlier HCC development among younger patients and those with non-cirrhotic HCC (Kao et al. 2000; Yin et al. 2008).	Asia, North America and Greenland
C (1-16)	➢ Associated with chronicity in Chinese individuals (Zhang et al., 2008). ➢ Higher HBeAg seropositivity among children infected with genotype C compared to those infected with genotype B (Ni et al., 2004). ➢ Higher risk of HCC development and severe form of liver disease compared to genotype B (Watanabe et al., 2005; McMahon., 2009).	Asia and Australia
D (1-10)	➢ High risk of developing HCC and liver cirrhosis (Sunbul, 2014). ➢ Associated with acute liver failure (Wai et al., 2005).	Africa, Europe and Asia
E	➢ High rates of HBeAg positivity (Sonderup, 2020)	Africa
F (1-4)	➢ High risk of developing HCC and liver cirrhosis (Sunbul, 2014).	Central and South America
G	➢ Serological diagnosis may prove challenging due to the impaired production of HBsAg (Zaaijer et al., 2011).	Europe and North America
H	➢ None described	Central America
I	➢ None described	Asia
J	➢ None described	Asia

Distribution of hepatitis B virus genotypes across geographical locations and their clinical associations. HBsAg: Hepatitis B virus surface antigen; HBeAg: Hepatitis E antigen; HCC: Hepatocellular carcinoma.

1.8 Laboratory investigation of HBV infection

The serological and virological diagnosis of HBV acute and chronic infection can be diagnosed by a characteristic serological and virological marker profile.

1.8.1 Serum HBsAg and anti-HBs

Serum HBsAg is the serological indication of HBV infection, and this protein is usually detected around 6-12 weeks post-infection, with HBeAg appearing after that (Bowden, 2006).

Clinical symptoms, if they occur at all in the early course of acute infection, become apparent as the serum alanine aminotransferase (ALT) becomes abnormal after the viral antigen titers peak, usually between 2-4 months following exposure to HBV (Table 1-2) (Hui et al., 2008). Recovery from acute HBV infection is indicated by the loss of serum HBsAg, which usually occurs 4-6 months after virus exposure. Anti-HBs can be detected after the loss of serum HBsAg from the systemic circulation. Furthermore, Anti-HBs is a serological marker of successful hepatitis B vaccination.

1.8.2 Serum HBcAg and anti-HBc

Serum HBcAg is not easily detected in the systemic circulation and thus plays no role in hepatitis B diagnosis. However, anti-HBc is often the first marker detected during acute hepatitis B infection (Table 1-2). Anti-HBc IgM may be the only marker of HBV infection during the window period (gap between the disappearance of HBsAg and the appearance of anti-HBs), but may also be detected during flares of CHB, making it serologically indistinguishable as a one-off measurement from acute HBV infection; this represents a challenge in differentiating between newly-acquired HBV and chronic infection (Tan et al., 2015).

Rapidly rising or falling quantitative HBsAg titers suggest acute HBV compared to stable quantitative HBsAg levels in chronic infection. Newly acquired infection is assumed if the patient has an obvious recent risk factor for acquisition, previously negative HBsAg, and high levels of specific antibodies to hepatitis B core protein (anti-HBc IgM) in the absence of prior evidence of HBV infection. Anti-HBc IgG is the hallmark of previous exposure to HBV, persisting in both resolved acute and chronic HBV (Table 1-2) (Tan et al., 2015).

In individuals with positive anti-HBs and negative HBsAg, anti-HBc IgG can be used to differentiate between individuals with vaccination (no anti-HBc) or past infection (Tan et al., 2015).

1.8.3 Serum HBeAg and anti-HBe

Serum HBeAg is an indirect marker of ongoing HBV replication and infectivity. Usually, HBeAg seroconversion occurs before HBsAg seroconversion in patients who recover from an acute infection (Niederau et al., 1996; Lin et al., 2007). HBeAg seroconversion may be delayed for several years in patients with CHB and is considered an essential factor in the natural history of HBeAg-positive patients with CHB, as it is associated with low viral replication, the remission of liver disease, and an increased prospect of HBsAg seroconversion (Niederau et al., 1996; Lin et al., 2007).

1.8.4 HBV viral load assays

The HBV viral load is the key virological marker of the clinical management of CHB, which indicates viral burden and viral replication, with landmark studies showing the important relationship between higher viral loads and the risk of more severe liver disease and HCC development (Chen et al., 2006). Several commercial HBV viral load assays are currently available, each with its fundamental advantages and disadvantages. The World Health Organization (WHO) defined the calibrated measurement of HBV viral load in IU/mL to ensure standardisation across the different assays for patient care, and the results from different laboratories using different assays can now be correlated (Saldanha et al., 2001).

The HBV viral load is an essential factor in determining patient eligibility for antiviral drug and treatment response after that. Additionally, it is essential to monitor the resistance in patients with antiviral drug resistances (Locarnini et al., 2004). Real-time polymerase chain reaction (RT-PCR) tests are now the more sensitive viral load assay and the most common diagnostic assay used in routine clinical practice, exhibiting a broad dynamic range (from 10 up to 10^9 IU/mL) (Bowden, 2006; Song and Kim, 2016).

1.8.5 Quantitative serum HBsAg and HBeAg assays

Most commercial serological assays for both serum HBsAg and HBeAg are qualitative; however, sensitive, reliable assays have been developed to quantify serum HBsAg. These immune enzyme assays are fully automated, with a high throughput capacity, and calibrated to the WHO HBsAg standard, with the results reported in IU/mL (Deguchi et al., 2004). The

role of quantitative HBsAg has garnered considerable attention due to its potential to monitor natural CHB infection as well as the treatment response (Martinot et al., 2013; Honer and Cornberg, 2014). Most commercial HBeAg assay kits are not marketed as quantitative but produce a signal which is linear, within a restricted range. This assay can be further modified and optimized for use as a quantitative assay by reference to an external standard.

Table 1-2: Hepatitis B virus infection: Clinical course and serological profile.

Weeks after exposure	Phase	Signs and symptoms	Serological profile
Initial exposure	Incubation phase Duration: 2 -3 months	No Signs and symptoms	High viral replication
12 weeks	Predicters phase Duration: 1 -7 days	Fever Anorexia Nausea Fatigue Body aches Onset of jaundice	HBsAg present Anti-HBc (IgM) present High level of HBV DNA High level of HBeAg High level of ALT
13 weeks	Icetric phase Duration: 1 -2 weeks	Fever Anorexia Nausea Fatigue Body aches Jaundice	HBsAg clearance/present Anti-HBc(IgM) pesent low level of HBV DNA low level of HBeAg low level of ALT
15 weeks	Convalescent phase Duration: weeks to months	Jaundice resolves	HBsAg clearance and HBV DNA (recovery) or HBsAg Present (Persistent carriage) Anti-HBc(IgM) present Anti-HBc(IgG) present Anti-HBe present Normal level of ALT

Clinical course and serological profile of acute hepatitis B virus infection (Liang, 2009; Liu and Zhang, 2015). HBsAg: HBV surface antigen; Anti-HBc: antibody to HBV core antigen; HBeAg: HBV e antigen; ALT: Alanine Aminotransferase; Anti-HBe: antibody to HBV e antigen.

1.8.6 HBV Genotyping

The HBV genotype is not routinely used as a standard of care. However, it provides clinical support given the known differences in disease progression and response to antiviral therapy, as described later in Sections 1.12.2 and 1.12.3.

The prevalence of antiviral drug resistance and vaccine mutation across different regions is influenced by the predominant genotype and may also relate to different drug or vaccine exposure patterns within the population.

There are various methods of HBV genotyping, including restriction fragment-length polymorphism, reverse hybridisation, genotype-specific PCR assays, and sequencing analysis, that will be described later in this thesis. Genotype-specific PCR is suitable for detecting mixed genotype infections. While they all have costs and benefits, sequence analysis is the recognised gold standard for genotyping and can identify viral recombinants (Downs et al., 2019).

1.8.7 Viral mutation analysis

It is possible through viral sequencing techniques to identify viral mutations that are associated with antiviral drug resistance and vaccine mutation. This is particularly important in the setting where mutation pathways may confer resistance to other nucleos(t)ide analogues (NAs), and mutation analysis can facilitate appropriate salvage therapy decisions. Common mutations associated with polymerase are described in (Sections 1.12.2 and 1.12.3)

1.9 HBV prevention

HBV prevention can be achieved through the widespread implementation of the HBV vaccine (passive or active) and/or ensuring safe practices, such as screening blood and organ donors, applying infection control measures, sterilising medical equipment, introducing needle exchange programs, and enforcing preventative measures regarding households' contact to minimise exposure to the bloodborne pathogen (Liaw, 2009; Shepard, 2006; Chang and Chen, 2015).

1.9.1 Passive immunisation

Hepatitis B immunoglobulin (HBIg) is a preparation of purified IgGs obtained from plasma of individuals with immunity to HBV and they contain high titres of the antibody to HBsAg.

It is administered to confer temporary passive immunity on non-immune individuals with exposure to HBsAg, through either blood or bodily secretions infected with HBV, sexual contact with an HBsAg-positive person or to infants with perinatal exposure born to HBsAg-positive mothers (Habib and Shaikh, 2007; Zuckerman, 2007). HBIg provides immediate protection, but its effects last for only three to six months (Zuckerman, 2007).

In adults, HBIg is provided with the first dose of HBV vaccine within 12 hours of exposure, followed by a second and third doses at 1 and 6 months of age. The HBIG can be repeated at one month if there is a history of previous vaccine non-response. HBIG in exposed new-borns should also be administered with the birth dose of the HBV vaccine, and then the subsequent doses completed as part of the Expanded Program for Immunization (EPI) schedule (Habib and Shaikh, 2007; Zuckerman, 2007). HBIg is not widely available in many settings due to its high cost, dependency on a robust cold chain, and relatively short-term effectiveness (Zuckerman, 2007; Chang and Chen, 2015).

1.9.2 Active immunisation

The HBV vaccine offers active immunity against HBV infection and its associated complications, including liver cirrhosis and hepatocellular carcinoma (Chang and Chen, 2015).

In the late 1970s, the first generation HBV vaccine, containing purified HBsAg obtained from the serum of CHB patients, was developed in the USA (containing 22 nm of SHBs of HBsAg) and France (containing SHBs of HBsAg and small amounts of pre-S) (Shouval, 2003; Shouval et al., 2015).

In the 1980s, a second-generation HBV vaccine, produced in yeast, was developed using advances in recombinant DNA technology due to concerns about the safety of using antigens obtained from blood products (Shouval, 2003; Shouval et al., 2015). This second-generation vaccine composed of SHBs and is widely used as a component of the WHO Expanded Programme on the Immunization (EPI) of infants (Shouval, 2003; Shouval et al., 2015). A third-generation HBV vaccine constructed in transfected mammalian cells has been developed in many countries (Shouval, 2003; Shouval et al., 2015).

The HBV second-generation vaccine is available in both monovalent and combined formulations (Lavanchy, 2012). The combined formation (pentavalent or hexavalent) contains immunogens against HBV, diphtheria, tetanus, pertussis and Haemophilus influenzae type b (Hib), with or without a polio vaccine, and is administered to infants at

the age of six, ten and 14 weeks as part of the WHO EPI (Lavanchy, 2012; Nayagam et al., 2020), aiming to interrupt horizontal transmission (Nayagam et al., 2020). An HBV monovalent vaccine should be administered to all infants within 24 hours of birth, especially those born in high endemic areas (WHO, 2015). Three doses of this monovalent formulation should also be administered to adults in the high-risk groups, including health care professionals, people requiring frequent blood or blood products, people who inject drugs, prisoners, the household and sexual contacts of people with CHB and travellers to endemic areas (Lavanchy, 2012; Chang and Chen, 2015; WHO, 2018).

1.9.3 The post-vaccine Era: a reduction in HBV

Since the widespread introduction of the HBV vaccine, the global rates of HBV infection and chronic carriage have fallen. Furthermore, the important long-term complications of chronic HBV infection, including the rates of HCC, have also significantly decreased. The use of combined active and passive vaccination, given to infants with HBsAg-positive mothers at birth, has been shown to prevent at least 94% of perinatal HBV transmission (Beasley, 1983). In Taiwan, where there was an early implementation of universal infant HBV vaccination in a highly endemic area with high rates of perinatal transmission, a staggering impact was demonstrated at the 20-year follow-up, showing: that the HBsAg prevalence rates in children had decreased from 11% to 1%, an 80-90% protection rate for infants born to HBsAg-positive mothers, a 68% reduction in mortality from fulminant hepatitis in infants, and a 75% decrease in the incidence of HCC in children (Chien, 2006). Other regions of the world have shown similarly impressive trends. However, the burden of CHB in some low prevalence countries has continued to increase due to the migration of people from highly endemic areas, such as Asia and sub-Saharan Africa (Liaw, 2009).

1.9.4 Vaccine escape

Variants in the small surface protein gene have been implicated in cases of vaccine escape. The most widely reported of these is glycine to arginine at amino acid position 145 of the antigenic ‘a’ determinant of the surface antigen, which is the major neutralising domain of the HBsAg (Zanetti, 2008). Changes in the S protein, specifically within the major antigenic determinant, can result in conformational alterations, leading to the evasion of vaccine-induced immunity and HBIg therapy (Cooreman et al., 2001; Caligiuri et al., 2016; Lazarevic et al., 2019). Furthermore, diagnostic escape mutations can return false-negative HBsAg results (Caligiuri et al., 2016; Lazarevic et al., 2019).

Mutations in the HBV polymerase can also lead to amino acid changes in the S protein due to their ORF overlaps in the genome (Caligiuri et al., 2016). As mentioned above, the surface protein mutation G145R is a well-known VEM which, alone, can be responsible for vaccine escape. Experimentally, it fails to bind to the HBsAg antibodies (HBsAb) and has also been isolated from vaccinated children who become infected (Steward et al., 1993; Lazarevic et al., 2019).

In addition, the mutation K141E has been reported to have reduced binding to HBsAb and has also been isolated from vaccinated children who became HBV core antibody positive (a marker of infection) (Steward et al., 1993).

Several other mutations in the S protein have recently been associated with vaccine/HBIg resistance (Lazarevic, 2014; Lazarevic et al., 2019). There exist very limited data for VEMs in the Middle East. However, in other areas with high endemicity, VEMs can be common, as evidenced by a reported prevalence of 28% among vaccinated HBV-infected children in Taiwan (Hsu et al., 1999).

1.10 Management of HBV Infection

1.10.1 Acute HBV infection

The clinical spectrum of acute HBV infection ranges from asymptomatic to fulminant hepatitis, occurring in less than 1% of cases (Liang, 2009). In symptomatic infection, a non-specific prodromal stage is followed by symptoms which may include anorexia, nausea and vomiting, fatigue, malaise, arthralgia, myalgia and headache, that precede the onset of clinical jaundice by one to two weeks. The rate of acute HBV infection resulting in spontaneous resolution is greater than 95% in immunocompetent adults and therapy is not required. Antiviral treatment is commonly given to patients with fulminant hepatitis, but whether it improves the outcome remains unclear.

1.10.2 Chronic HBV infection

The aim of antiviral treatment in CHB management is clearly described in Section 1.11.

1.11 Antiviral therapy

Antiviral therapy aims to improve the quality of life and survival of chronically HBV-infected patients by preventing or delaying disease progression (Grimm et al., 2011). The aim is to suppress the viral replication, which usually results in the reduced histological

activity of chronic hepatitis and biochemical remission. Therefore, the risk of progression to cirrhosis decreases along with the incidence of HCC in non-cirrhotic patients, and also to a lower level in cirrhotic patients (EASL, 2017).

Recently, many clinical studies have demonstrated that the risk of HCC developing significantly decreased when patients underwent successful antiviral therapy compared to patient groups that remained historically untreated (Seeger & Locarnini, 2015). HBV replicates through an RNA intermediate step, despite having a DNA genome. Lack of proofreading capacity of the virus-encoded, RNA-dependent DNA polymerase, when coupled with an extremely high rate of HBV replication, increases the potential for genome-wide mutations. Viruses with mutations that confer a replicative advantage are selected when subjected to selective pressure in the presence of an antiviral agent. They ultimately become the predominant viral species (Richman, 2000). The current therapies for CHB include either interferon – α (IFN-α) and its pegylated form (Peg-IFN-α) or oral NA treatment (Ward et al. 2016; EASL, 2017).

1.11.1 Pegylated interferon alpha (PegIFNa)

PegIFNa was developed in 2005 for CHB treatment (Woo et al., 2017). PegIFNa bind to the IFNAR receptors and triggers the JAK/STAT pathway resulting the upregulation of other IFN, antiviral cytokines and interferon regulated genes (Tan et al., 2018). It has also been shown to augment the immune response against HBV (Tan et al., 2018). PegIFNa is administered through subcutaneous injection once a week for a period of 24 weeks; it is thought to increase host cell-mediated immunity against HBV and also target various parts of the HBV life cycle including decreasing HBV RNA transcription thus inhibiting viral replication (EASL, 2017; Woo et al., 2017). Despite the finite duration of treatment and its potential ability to induce long-term immunological control, only about 30% of individuals achieve a sustained off treatment response (Woo et al., 2017).

Other disadvantages of PegIFNa include its adverse side effect profile, such that it cannot be used in infants below one year of age, pregnant women and individuals with decompensated cirrhosis, hypersplenism, thyroid disease, autoimmune diseases, heart diseases, kidney diseases, psychiatric illness, retinopathy, and blood disorders, and it is also limited by cost (WHO, 2015; EASL, 2017; Woo et al., 2017).

1.11.2 Nucleos(t)ide analogues (NAs)

NAs act directly by inhibiting the reverse transcriptase, all of them target one enzymatic activity and the DNA strand elongation of the HBV polymerase (Ismail et al., 2012).

The first group of NAs contain lamivudine (LAM) and telbivudine (LdT), that are the only L-nucleosides that are approved for use in the United States. In addition, clevudine is only approved for use in Korea (Ghany & Doo, 2009). All of the L-nucleosides have a similar molecular structure and target site, which contributes to their possession of a similar form of antiviral resistance. L-nucleosides act as chain terminators that inhibit the synthesis of both the negative and positive HBV DNA strands (Ghany & Doo, 2009).

LAM was initially developed and widely used for HIV infection, and in 1998 the FDA approved it for treatment of CHB (De Clercq et al., 2010; Lingala et al., 2016; Luo et al., 2018). As mentioned, it is administered orally in its inactive form and then transported into the cell where it is converted to its active metabolite, lamivudine triphosphate by intracellular kinases (Else et al., 2012). In practice, this is able to reduce the level of HBV DNA serum in patients who have been treated. However, in order to sustain inhibition, LAM requires a lengthy course (Kang et al., 2015). It has been demonstrated that this well-accepted drug has a good impact even with patients who suffer from hepatic failure (Koumbi, 2015). Where a long-term LAM therapy course is applied in as many as 50 percent of HBeAg seroconversions, in both HBeAg-negative and HBeAg-positive patient, it sustains low levels of ALT and HBV DNA. Nevertheless, after a period of 12 months of treatment, resistant mutations develop in 20 per cent of cases, whereas after a five-year period of treatment, the figure is 70 percent (Koumbi, 2015). Among the most frequently encountered drug resistant mutations are tyrosine-methionine-aspartate-aspartate mutations at position rt204 (otherwise referred to as the YMDD mutant) (Kang et al., 2015). It is considered safe for use among both children and adults, including during pregnancy (De Clercq et al., 2010).

LdT was approved for the treatment of CHB in 2006 (Long et al. 2017). It is a nucleoside analogue of thymidine which is administered orally and then gets converted to its active form, telbivudine 5'-triphosphate intracellularly by cellular kinases (Amarapurkar, 2007). Its active metabolite, 5-triphosphate of β-L-2-deoxynucleosides, is incorporated into the second strand of the growing DNA (DNA-to-DNA transcription) thus inhibiting replication (Amarapurkar, 2007). The antiviral efficacy of LdT is more potent than that of LAM - the number of patients having undetectable HBV DNA is considerably greater than in those who are given LAM treatment (Kang et al., 2015).

The second group of NAs include adefovir dipivoxil (ADV), tenofovir disoproxil fumarate (TDF) and tenofovir alafenamide (TAF), which comprise of acrylic phosphonates that are approved for use in CHB therapy (EASL, 2017). Structurally, instead of a simple phosphate group, these agents possess a chain terminating phosphonate group that cannot be cleaved by the host esterases. Such compounds have improved access to the HBV polymerase active site because of their structural similarity to the flexible phosphonate linker and the natural substrate dATP (Ghany & Doo, 2009).

ADV was approved for treatment of CHB in 2002 (Lingala et al., 2016). It is a nucleotide analogue administered orally as adefovir which is then converted to ADV in plasma and then converted again to its active metabolite adefovir diphosphate (ADV-DP) intracellularly (Marcellin et al., 2003). ADV-DP inhibits viral replication by competing with the natural substrate deoxyadenosine 5' triphosphate for incorporation into viral DNA thus acting as chain terminator (Marcellin et al., 2003). It is a potent inhibitor of viral replication of both the wild type and LAM resistance HBV. It displays renal toxicity when administered at a daily dose of 30 mg or more, however, no serious effects have been observed at a standard dose of 10 mg/day although close monitoring of renal function is recommended (Kang et al., 2015). Emergence of resistance mutations have been observed in about 30 percent of patients upon long-term treatment (Koumbi, 2015).

TDF was approved for treatment of CHB in 2008 (Lingala et al., 2016). It was initially developed and widely used for HIV treatment and inhibits both HBV and HIV replication at the RT step (Lovett et al., 2017; Kang et al., 2015). It is administered orally as TDF which is then converted to tenofovir (TFV) in plasma and finally to its active ingredient tenofovir diphosphate (TFV-DP) intracellularly (Kearney et al., 2004; Abdul Basit et al., 2017; Byrne et al., 2018). Since TFV-DP competes with the natural substrate nucleotide deoxyadenosine 5´triphosphate for merging with the viral DNA, it inhibits viral replication, thereby terminating DNA chain elongation. TDF can be safely used in pregnancy and is included in some guidelines for HBV PMTCT (EASL, 2017; Terrault et al., 2018). Approved in 2015, TAF a phosphonate prodrug of TFV, was developed to mitigate the risk of renal and bone toxicities associated with TDF (De Clercq, 2016; Byrne et al., 2018) since it can be at administered at a lower dose (25 mgs) and has high plasma stability compared to TDF (Abdul Basit et al., 2017; Byrne et al., 2018).

The third group of NAs comprise the cyclopentane group, that include entecavir (ETV), which is a carbocyclic analogue of 2’deoxyguanosine and demonstrates potent activity against the HBV (EASL, 2017). ETV acts at three stages in the viral replication cycle by

inhibiting the priming as well as the synthesis of both the negative and positive strands of HBV DNA (Ghany & Doo, 2009). ETV was approved for treatment of CHB in 2005 (Lingala et al., 2016). ETV is administered orally then converted to its active metabolite, entecavir triphosphate (ETV-TP) intracellularly (Yan et al., 2006). ETV is carbocyclic analogue of 2' deoxyguanosine which inserts into the first and second strands of the viral DNA inhibiting both RNA-to-DNA and DNA-to-DNA transcription, and also prevents replication by inhibiting the unique protein-linked priming activity (Dimou et al., 2007; Langley et al., 2007). ETV is considered to have potent antiviral activity with a high genetic barrier to resistance (Villamil and Cairo, 2013). ETV cannot be used during pregnancy due to its carcinogenic potential shown in animal studies (Lin et al., 2008; Aslam et al., 2018). A summary of NAs and therapeutic target are presented in (Table 1-3). Details of each drug was cited as presented by provider instructions.

Table 1-3: Summary of therapeutic target of the anti-HBV drugs.

Drug	Therapeutic Target
LAM	Inhibition of the DNA and RNA-dependent polymerase HBV reverse transcriptase activities.
LdT	Inhibition reverse transcriptase by competing with the thymidine 51-triphosphate natural substrate.
ADV	Inhibition of reverse transcriptase by competing with the deoxyadenosine triphosphate natural substrate, thereby resulting in the termination of the DNA chain subsequent to its merging into viral DNA.
TDF	Inhibition of reverse transcriptase activity by competing with the deoxyadenosine 51-triphosphate natural substrate subsequent to incorporation into DNA, through DNA chain termination.
ETV	Inhibition of reverse transcription of the negative DNA strand from the pregenomic messenger RNA and Inhibition of the synthesis of the HBV DNA positive strand.

Lamivudine: LAM; Telbivudine: LdT; Adefovir dipivoxil: ADV; Tenofovir: TDF; Entecavir: ETV

1.12 Guidelines for Currently Approved Medications

There are two international treatments standards published for HBV management at this time, representing the optimal management of HBV infection. These include the European Association for the Study of the Liver (EASL, 2017) and the WHO guidelines for the prevention, care and treatment of persons with CHB (WHO, 2015).

The following two strategies for treatment are given in the 2017 EASL guidelines: long-term treatment with NAs and finite-duration treatment with pegIFN. However, if pegIFN treatment is given, it ought to be applied very carefully because its combination with NA and its contraindications are not recommended.

The strongest agent which is the greatest barrier to drug resistance ought to be the NAs that are applied for finite-duration treatment, whereas the first-line monotherapies are ETV and TDF. It is evident that most patients with monotherapy for a period of greater than three years keep a full virological response (Chang et al., 2010; Heathcote et al., 2011). Furthermore, the WHO guidelines recommend solutions to antiviral treatment failure. A drug switch to entecavir or tenofovir from the initial drug is advised when there is partial response or a primary non-response (Table 1-4) (WHO, 2015). Following the identification of virological breakthrough and the exclusion of patient compliance as a potential cause, a new therapeutic regimen ought to be employed as soon as possible after the HBV DNA loads have been monitored, and the pattern of resistance mutations has been detected.

It is recommended by the WHO that the entecavir and tenofovir, with their high genetic drug resistance barrier, ought to be used in the initial treatment for patients of 12 years of age and over, but that entecavir ought to be used for children between the ages of two and eleven**.** For the second stage of treatment in the same paediatric group, it is recommended that TDF is used. However, it is strongly advised that other NAs, for instance LdT and ADV, are not used. Moreover, since pegIFN involves contraindications and resource-limited settings, it is not regarded as being a suitable treatment option (WHO, 2015).

Table 1-4: International treatment guidelines for Hepatitis B virus.

	EASL 2017	WHO 2015
Recommended antiviral drug	TDF: 245 mg/day; TAF: 25 mg/day; ETV: 0.5 mg/day	TDF:245 mg/day ETV: 0.5 mg/day
Duration of treatment	Until loss of HBsAg (long-term).	Indefinite for all patients with cirrhosis.

TDF: Tenofovir disoproxil fumarate; TAF: Tenofovir alafenamide; ETV: Entecavir; HBsAg: HBV surface antigen.

1.12.1 End Point of Antiviral Drugs

Based on the EASL guidelines, the end point of HBV therapy is found to differ between patient groups and is recognized based on the following three serological sets:

1. Patients with positive HBeAg and negative HBeAg for whom the ideal end point is sustained HBsAg loss, with or without seroconversion to anti-HBs (Grimm et al., 2011).

2. Patients with positive HBeAg and durable seroconversion to anti-HBe, which is a satisfactory end point (Grimm et al., 2011).

3. Patients with positive HBeAg who do not achieve an anti-HBe seroconversion, a maintained undetectable HBV DNA level of treatment with NAs or a sustained undetectable HBV DNA level after interferon-α (IFN-α) therapy, which is the second most desirable end point (Grimm et al., 2011).

1.12.2 The Antiviral Resistance Mechanism

During the disease course, viral mutants are generated which subsequently result in the development of viral quasispecies. Antiviral drug resistance is caused by adaptive mutations within the viral genome, whereas HBV infection has particularly high levels of turnover and virus production, producing over 10^{11} virions each day (Figure 1-5) (Locarnini, 2008; Zoulim & Locarnini, 2009). The high HBV replication rate, together with the high mutation rate (1 in every 10^5 nucleotide substitutions in each replication cycle, due to the error-prone property of reverse transcription) leads to CHB patients having various types of viral quasispecies, each differing in at least one mutation. Furthermore, the likelihood of a mutation being related to the drug resistance that is chosen during therapy is dependent on the potency of that drug (Locarnini, 2008; Zoulim & Locarnini, 2009).

Replication fitness, which is described as the capability of replicating when under pressure, and the replication capacity of resistant isolates are able to shape the mutation pattern that appears in primary mutations. This refers to the amino acid change(s) which lead to a lower susceptibility to an antiviral agent as against secondary and compensatory mutations. This restores replication defects related to primary drug resistance and could be associated with low-level reduced susceptibility. Many evolutionary drug resistant HBV pathways have been identified in patients being treated with NAs (Locarnini, 2008; Zoulim & Locarnini, 2009).

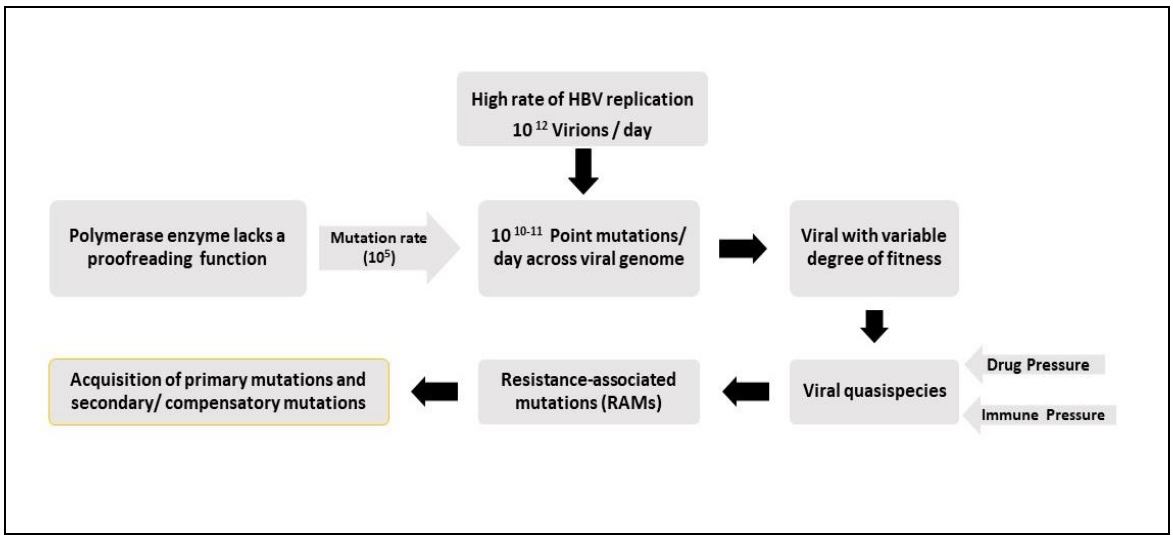
Furthermore, resistance is recognized by the availability of the HBV replication space. However, the liver is able to accommodate new cccDNA transcriptional templates only if uninfected cells are created by hepatocyte proliferation, normal liver growth or the direct loss of cccDNA. When the genetic barrier to the treatment regimen resistance increases, the number of particular mutations needed for drug resistance also increases. Moreover, the host characteristics of the virus-infected hepatocytes, the genetic background and the immune response have an impact on antiviral drug resistance.

Additionally, the effectiveness of the treatment could be influenced by pre-existing antiviral-resistant mutants and mutation patterns which accumulate in the course of time (Ismail et al., 2012). When drug resistance commences with pre-existing mutations, this will lead to an increase in the viral load, and the ALT levels will also increase many weeks and months later which will result in the progression of liver disease (Zoulim & Locarnini, 2009). The potential of increased serum ALT is generally associated with identifiability of the resistant strain in the case of patients having LAM resistance (Lok et al., 2012). Likewise, such patients are particularly vulnerable to ALT flare which could be linked with hepatic decompensation (Lok et al., 2012).

Resistance development is dependent on numerous elements which are associated with the progress of antiviral resistance. The principal factors are the potency and genetic barrier to resistance of the antiviral agent as well viral fitness, which indicates the capability of a virus to replicate within a defined environment (Richman, 2000). It is possible to detect the strength of antiviral NAs by their ability to act as a competitive inhibitor of the HBV polymerase with regard to that of the natural substrate (Richman, 2000). From a clinical perspective, a drug's strength is reflected by how quickly it is able to restrain viral replication, and the more rapidly it does this, then the possibility of it developing antiviral resistance is lower (Richman, 2000). The genetic barrier to resistance indicates the number of mutations that need to be accumulated by the virus to enable effective replication where

the antiviral agent is present. The genetic barrier to resistance partly depends on the antiviral compound's structure, and the constraints imposed by the viral polymerase's ability to tolerate compensatory mutations without impeding its enzymatic activity to any great extent. Consequently, if an agent has a high genetic barrier against the accumulation of mutations, it will clearly have less possibility of developing resistance (Richman, 2000).

Figure 1-5: Antiviral Resistance Mechanism in HBV.



Virus such as HBV and more famously HIV-1, generate quasi-species during replication, this is a population of viruses that have accumulated mutations which may or may not be associated with a fitness cost. As these viruses are able to tolerate high mutations, it can lead to mutations that allow the virus to replicate in the presence of a drug (Rodriguez-Frias et al., 2013). Mutations associated with drug resistance occur because of high viral replication. Some mutations can lead to severe loss of function thus impairing viral fitness and replication capacity, whereas other mutations can lead to the development of variants, which can become dominant with exposure to drug. These variants evolve to acquire compensatory mutations, which restore their replication capacity.

The mutations in the RT domain of the polymerase gene have been given specific names based on the nomenclature introduced by Stuyver et al. (2001). Currently, the resistance mutations are divided into two types; namely, primary resistance mutation and secondary or compensatory mutation (Lok et al., 2003).

The primary mutation, namely LAM resistance mutation, causes an amino acid substitution that results in resistance, such as rtM204V/I, which occurs within the YMDD motif in the C domain of the viral polymerase (Fischer et al., 2001). The rtL180M is the main compensatory change caused by LAM resistance and other compensatory mutations, including rtV173L and rtL80I (Lok et al., 2003; Ismail et al., 2012). The primary LAM resistance mutation also confers cross-resistance to other L-nucleosides, such as LdT and CLV (Lok et al., 2003). The primary mutations are rtN236T in addition to rtA181T/V for

ADV resistance. In addition, ADV-associated mutations display partial cross-resistance to TDF (Lok et al., 2003).

The mutations (rtL180M + rtA194T + rtM204V) were reported to be associate with TDF resistance in HBV-HIV co-infections after dual therapy with TDF and LAM, but not in HBV-infected patients undergoing TDF.

1.12.3 Comparative Worldwide Reports on Antiviral Resistance

An Indian study conducted in 2011 described the amino acids substitutions in 97 patients who had not been prescribed antiviral drugs before. Genotype D was the main genotype amongst those patients, followed by Genotype C and Genotype A (Ismail et al., 2012). One patient carried a sub-genotype D4 virus with a unique substitution rtE/D263S together with rtA329V. A mutation rtI233V, which is associated with ADV resistance, was identified in four of the patients who participated in the study.

Similarly, rtT128N, rtV214A, rtQ215S, rtS219T and rtN238S compensatory mutations were identified separately. Atypical mutations with new amino acid substitutions at positions rtV84, rtT128, rtS213, rtV214, rtQ215, rtS219, rtP237, rtN238 and rtY245 were identified in 19 patients (Ismail et al., 2012). Furthermore, the study showed some amino acid substitutions that are common and specific to certain genotypes; for example, in the Genotype D patients, rtA/P/S/T54H/Y, rtS135Y/N, rtK149Q, rtW257Y, rtT259S, rtI/L/V266R/K, rtI/N/S/T53D and rtH126R were found to be common. Likewise, rtF151Y and rtR153W/Q were only found in the patients with Genotype A. In patients with Genotype C, only the rtH9Y substitution was identified (Ismail et al., 2012).

More recently, Hermans et al. (2016) disclosed data on 1317 patients covering 18 European countries along with their treatment history. Genotype D was found to be the most common genotype, followed by Genotype A and Genotype C. Antiviral resistance was frequently encountered in patients who were exposed to monotherapy with LAM and ADV. Based on this, 1,102 patients were treated with NAs monotherapy – mainly with LAM (972/1317), ADV (59/1317), ETV (50/1317), TDF (18/1317) and LdT (3/1317). A total of 197 patients were exposed to two NAs, either simultaneously or consecutively and mainly received LAM + ADV, LAM + ETV, LAM + TDF, and LAM + LdT. Triple antiviral exposure was present in 17 patients, of whom several received LAM + ADV + ETV, while the rest received LAM + ADV + TDF. Lastly, one patient was exposed to four different NAs (LAM + ADV + TDF + LdT) (Hermans et al., 2016).

The common primary mutation rtM204V/I associated with LAM and LdT found was detected in the results, while the rtA181T/V and rtN236T mutations were identified associated with ADV resistance (Hermans et al., 2016). The mutation of rtM204V/I was accompanied by additional mutations in many cases.

Resistance against ADV was present in 43/193 patients, and against ETV in 36/102 patients. In patients who used ETV as monotherapy without previous exposure to other NAs, intermediate resistance to ETV (rtM204V + rtL180M) was identified in some cases (Hermans et al., 2016).

There are also reports of four different studies based in Middle Eastern countries (Oman, Jordan, Iran and Turkey). The first report, from Oman, reported that of 146 HBsAg-positive patients, a majority of which infected with the Genotype D virus, eight had developed resistance to NAs (Al Baqlani et al., 2014). Amongst these, seven patients treated with LAM, had the ADV resistance mutation rtS85A. In particular, one patient showed four mutations (rtL80V, rtL180M, rtM204V and rtQ215S), and another three (rtL80V, rtL180M and rtM204V) (Al Baqlani et al., 2014). The second study was conducted in Jordan in 2008. It targeted 20 genotype D-infected patients, 4 of which were LAM naïve while the rest had a history of LAM treatment (Masaadeh et al., 2008). Six mutations in the rt domain were detected in five different patients with a history of LAM treatment. Two mutations (rtM204I and rtL180M) were detected in one patient. On the other hand, only one mutation (rtM204V) was detected in a LAM naive patient (Masaadeh et al., 2008).

An Iranian study reported 45 CHB genotype D-infected patients who had not been treated with any antiviral therapy for at least a year. Viral mutations were identified in 23 patients. Specifically, primary resistance mutations rtA181T and rtA181S were identified in two patients, whereas one patient had rtM204I (Hamidi-Fard et al., 2013). Secondary mutations were found in 20 patients, including an rtQ215S mutation in four patients, an rtQ215P mutation in three, an rtQ215H mutation in two and an rtQ215M mutation in one. The rtN238H mutation was detected in two patients (Hamidi-Fard et al., 2013). Lastly, each mutation of rtV214I, rtL80I, rtV214E, rtV214T, rtS219A, rtV207I, rtF221Y and rtF221V was detected in one patient (Hamidi-Fard et al., 2013). To summarize, most of these mutations were associated with ADV resistance and only one case demonstrated a LAM resistance. The last study, that was conducted in 2016 and covered the Middle East was from Turkey. It involved 53 CHB patients who had not received NAs before. The study showed that Genotype D was predominant amongst these patients. There was no primary drug resistance in any of the patients; however, there were compensatory mutations in 19

cases. Mutations associated with LdT were identified in four patients (rtL91I), 13 patients showed mutations with LAM and ADV (rtQ149K, rtQ215S, rtQ215H, rtQ249K, rtV214A), and two patients indicated mutations associated with ADV (rtN238D) (Altındis et al., 2016).

An additional report from China in 2016 studied 139 patients experiencing NAs treatment. The classic primary mutations (rtM204V/I, rtA181V and rtN236T) were identified and other secondary mutations (rtL80I/V, rtV173L, rtL180M and rt250L) were detected (Qian et al., 2016). The mutation of the rtM204V/I was considered the highest amongst the patients who had received the treatment associated with resistance to LAM, LdT and ETV. Mutation in rtM204V/I was the major type in the patients who were found to harbour drug-resistant mutations. Changes in the amino acids at rtM250 were found in some ETV-naïve patients. Amongst these patients, five had received LAM monotherapy, and rtM204I + rtM250L mutations were observed in three cases. The rtL180M + rtM204V + rtM250L mutations were found in one case, and the rtM250L mutation was found in another (Qian et al., 2016). The mutation pattern of rtL180M + rtM204V + rtM250V was detected in one patient receiving LAM who had switched to ADV sequential treatment. Combination mutations in rtA194G + rtS202N were identified in this report, even though this mutation has not been reported before (Qian et al., 2016). However, a previous report showed that the rtA194T mutation was associated with ADV and TDF resistance and that rtS202C/G/I was associated with ETV resistance (Tenney et al., 2007).

1.12.4 Management of Antiviral Resistance

The basic principles of management of antiviral resistance are to change therapy early once virological breakthrough is observed, select an add-on approach over switching, and choose rescue therapy based on the cross-resistance profile and potency of the rescue drug and presence of comorbid conditions. The recommendations guidelines for patients with Drug-resistant Virus are shown in (Table 1-5). The treatment approach for LAM resistance can also be relevant to LdT-resistance because of the shared drug-resistance profile, as recommended by EASL, TDF is primary antiviral drug against LAM and LdT resistant as well as against wild type HBV. For ETV resistance, TDF is proposed to be effective in suppressing the replication of ETV-resistance and combination of TDF plus ETV would be a more appropriate treatment for reducing ETV resistance and improving antiviral efficacy. ETV and TDF have been recommended for HBV patients with ADV-resistance while TDF can also be a treatment option for multidrug resistant HBV (EASL, 2017). Furthermore, there are no recommendations have been suggested in the management of TDF resistance.

A combination of ADV and LAM indicated no better antiviral efficacy than ADV monotherapy for LAM-resistant patients as recommended by EASL (EASL, 2017). However, ADV and LAM combination therapy had been recommended for these patients to prevent the additional development of ADV-resistant HBV mutants (Lim**,** 2017). The potency of TDF is around 30 times higher than that of ADV, which can be used as a substitute for ADV. The responses to TDF and LAM combination therapy are likely to be mediated by TDF alone, as LAM might have minimal or no antiviral efficacy in the presence of LAM-resistant HBV mutants (Lim, 2017). With this combination, continuous LAM treatment is likely only to help to prevent the development of TDF-resistant mutations rather than increase the antiviral potency.

Consequently, the combination of LAM with TDF would be beneficial in terms of providing a minimal risk of TDF resistance (Lim, 2017). In addition, TDF monotherapy delivers similar antiviral efficacy compared to the combination of TDF and ETV. Moreover, no other HBV-resistant mutations emerged during TDF monotherapy for up to 96 weeks (Lim**,** 2017). TDF monotherapy would be a practical choice for treating antiviral resistant patients with CHB, considering the comparable antiviral efficacy, low risks of resistance, lower costs and higher safety potential (Lim, 2017).

Table 1-5: Summary of the Recommendations as Practice Guidelines for Patients with Drug-resistant Virus.

Resistance to:	**EASL 2017**	**WHO 2015**
Primary recommendations		
LAM	TDF	TDF
ETV	TDF or TDF + ETV	TDF
ADV (no LAM exposure)	ETV or TDF	TDF
Multidrug	TDF + nucleoside analogue	TDF
Secondary recommendations		
LAM	ADV + LAM	
ETV	ADV + ETV	
ADV (no LAM exposure)	TDF	

EASL: European Association for the Study of the Liver; WHO: World Health Organization; LAM: Lamivudine; TDF: Tenofovir; ETV: Entecavir; ADV: Adefovir.

1.12.5 Prospects in the Drugs

In the past 20 years, there has been an increase in knowledge and research in the field of hepatitis B. Curative antiviral therapy has the potential to prevent morbidity and mortality associated with CHB infection. While treatments for hepatitis B have improved, CHB patients require lifelong management and prolonged therapies, and these rarely result in virus eradication. In order to end the HBV endemic, there is a need for effective therapies for CHB given that the approved NAs are not curative, they do not commonly result in the clearance of HBsAg, and their indefinite use can result in the emergence of RAMs which are a potentially significant concern for their effectiveness. As described, inhibition of RT by NAs is of limited efficacy owing to their primary mode of action, which has little impact on either cccDNA establishment or viral gene expression. Additional aspects of the HBV replication cycle have been explored over the past 5 years, to identify agents that could address these limitations and achieve antiviral effects that persist after treatment.

The crucial goal for future therapy would be to target viral proteins, and host-directed antivirals, such as entry inhibitors and immunomodulators as well as approaches, such as direct cccDNA targeting, and RNAse H inhibition (Tavis and Lomonosova, 2015).

Similarly, it could be possible to succeed in finding a functional cure soon. This is because, at present, many new agents in pre-clinical or trials of phases 1 and 2 exist, the objective of which is to lead to the loss of HBsAg, thereby enabling the discontinuation of treatment (Dawood et al., 2017).

There are two ways in which such agents can repress viral replication. The first of these is by impeding a particular stage this process, which is known as direct-acting antivirals (DDAs), while the second is by adjusting the host cell function, a procedure referred to as host-targeting agents (Dawood et al., 2017). In their development, the DDAs include: core/capsid inhibitors, cccDNA formation and transcription inhibitors, HBsAg release inhibitors, RT inhibitors, helioxanthin analogues, and antisense oligonucleotides.

As the host-targeting agents develop, they include: cyclophilin inhibitors, entry inhibitors (Myrcludex-B), and multiple immunomodulatory agents (Lin and Kao, 2016; Dawood et al., 2017). MYR GmbH (Gilead Sciences, Inc) is developing bulevirtide (myrcludex B) as a specific inhibitor for NTCP. Its objective is to obstruct the viral pre-S1 attachment by means of high-affinity binding to NTCP (Wedemeyer et al., 2018).

The infectivity of both HDV and HBV is dependent on HBsAg being present in the viral envelope. Therefore, the existing clinical studies of bulevirtide concentrate on patients who have HDV coinfection, and evaluate the virologic responses of HBV and HDV.

As Phase 2 data were performed, they indicated a significant decrease in HDV RNA following 24 weeks of bulevirtide merged with TDF, as well as after 48 weeks of bulevirtide either with or without PegIFNa. In the latter study, the HBsAg decline was greater in the case of bulevirtide being combined with PegIFNa (Wedemeyer et al., 2018). Phase 3 of the studies investigates the extended therapy either with bulevirtide as monotherapy or in conjunction with PegIFNa in patients who have HDV/HBV coinfection, and concentrating on HDV clearance as the principal aim. It would not be anticipated that the entry blockers would impede the HBV cccDNA formation. However, it would be expected that entry inhibition with bulevirtide would protect cells that remained unaffected, thereby contributing to HBV therapy either by being merged with other modalities, or alone (Wedemeyer et al., 2018; Lopatin, 2019).

However, the development of CRISPR-Cas9 and other technologies that is able to edit DNA directly can target cccDNA directly (Kennedy et al., 2015; Lopatin, 2019). Numerous firms have appeared around this interesting platform. Moreover, many academic groups have demonstrated, in non-clinical studies, that this technique has the potential to reduce functional cccDNA successfully. Numerous significant warnings remain, despite the fact that the data are striking. These include: (1) removal of off-target impacts that should be addressed; (2) It would be necessary for the vector for delivery to access all infected hepatocytes in order to eliminate the possibility of reactivation; and (3) Cleavage of such integrants carries a theoretical risk of inducing genomic instability, as well as the concomitant risk of carcinogenesis. This is despite the fact that studies have indicated that integrated genomes may be targeted. It is not yet known if such questions can be addressed prior to entering the clinic (Lopatin, 2019).

Thus, the challenge now is to understand these new insights in the context of developing the next generation treatments for CHB. With advances in the understanding of the HBV life cycle, viral and host immune systems, the goal of a functional cure should be achievable in the not-too-distant future.

1.12.6 Future Directions

Resistance to the drug will remain a critical issue when managing patients with CHB, as long-term therapy with NAs appears to be mandatory in most cases. It is unknown at present whether combination therapy with the currently approved first-line therapies (ETV or TDF) is necessary. This is because minimal or no resistance has been reported after approximately seven to eight years of monotherapy using these agents (Warner & Locarnini, 2014). However, as these patients are neither 'cured' nor 'controlled', the emergence of resistance is always a hidden threat in CHB. Therefore, new therapies against new targets are still necessary in order truly to 'cure' CHB (Warner & Locarnini, 2014).

The field of HBV antiviral therapy is entering a new era, and the research, scientific, medical and industrial communities have demonstrated an enhanced interest in developing new drug concepts for curing HBV infections (Zoulim & Durantel, 2015). A better knowledge of the viral life cycle and its interaction with the liver microenvironment and host immune responses, along with the development of new study models, will provide the right drive for pioneering research in this area. A better understanding and measurement of the major clinical endpoints will also provide better guidance for the preclinical and early clinical evaluation of the treatment concepts, which, in turn, should translate into improved treatment outcomes in the future (Zoulim & Durantel**,** 2015).

1.13 DNA Sequencing

DNA sequencing is not generally employed as a standard diagnostic test for HBV; instead, it is implemented to ascertain patterns of resistance exhibited by the virus and provide further information on its characteristics. Traditionally, HBV antiviral resistance testing is performed using Sanger sequencing. Recently various laboratories introduced deep sequencing, also called next generation sequencing for detection of resistance-associated mutations.

1.13.1 Sanger Sequencing

Sanger sequencing denotes a group of technologies that employ dideoxynucleotides (dNTPs) (chain terminator nucleotides) in a polymerase chain reaction (PCR) to acquire different length DNA fragments. Electrophoresis is used to divide these fragments based on size (Sanger et al., 1977a).

Following a series of improvements, such as transferring from radioactive to dye labelling of nucleotides and utilising capillary electrophoresis rather than slab gels, the Sanger method was made considerably more efficient. Consequently, over the next twenty years, it became the foremost DNA sequencing method and enabled the first complete human genome sequence to be generated in the year 2000 (Lander et al., 2001; Venter et al., 2001). Sanger sequencing has a number of important limitations when applied to viral populations, such as an absence of information on linkage and insufficient sensitivity to detect lesser-known variants.

1.13.2 Next Generation Sequencing (NGS)

Next-generation sequencing (NGS), also known as deep sequencing, massively parallel, or high throughput, has substantially elevated the output of genetic sequencing. However, the power of NGS means it can produce millions of readings of a sequence in a single run. It is a technique that can be employed to identify complete viral genomes. Moreover, it is unselected (and does not require specific PCR primers), which means it can be applied to detect novel viruses. Furthermore, second-generation sequencing technology amplifies and then reads separate strands of DNA before providing information on linkage across the full length of each strand. Its general availability and low cost mean Sanger sequencing continues to be employed in clinical settings to treat hepatitis B. In terms of recent applications, NGS has been applied to identify a new bunyavirus in a patient with "severe fever with thrombocytopenia syndrome" as well as the aetiology of Merkel cell cancer (Feng et al., 2008; Xu et al., 2011). Its metagenomic characteristics means it has also been useful in helping characterise the virome in children presenting with acute diarrhoea and fever (Wylie et al., 2012). However, because the origin of each sequence read is not viral, it remains an inefficient albeit powerful technique. This thesis later describes how target enrichment (magnetic beads adhering to virus-specific oligonucleotides) can be utilised to improve the technique and purify the virus of concern. Also, Nanopore technology can reveal mutations contained within a single viral particle, its directly sequences single molecules of amplified PCR products, producing genome-length reads that can reveal mutations contained within a single viral particle (McNaughton et al. 2019). Though, its use to date has been limited due to high error rate in raw reads of ~5–10% (McNaughton et al. 2019). Generally, Illumina is considered the most robust sequencing platform with the ability to capture within sample diversity with high accuracy (McNaughton et al. 2019). The first second-generation sequencing technology to become commercially available was

Pyrosequencing (Margulies et al., 2005). Instead of sequencing by chain termination, pyrosequencing applies the concept of sequencing-by-synthesis (SBS) (Ronaghi et al., 1998). In this technique, single-stranded DNA fragments are used as a template to manufacture complementary strands and sequence the library. Nucleotides are added sequentially to the plate (Voelkerding et al., 2009). The intensity of detected signal correlates with the number of bases incorporated. However, it is not possible to determine the number of bases accurately when multiple similar nucleotides are incorporated into just one cycle. Consequently, an extremely large number of insertion and deletion errors may arise in regions that contain long homopolymers. For instance, approximately 65%- 75% and 20%-30% of all sequencing errors are thought to be attributable to insertion and deletion errors, respectively (Astrovskaya et al., 2011). Several other technologies emerged (e.g., SOLiD Sequencing and Ion Torrent ™ Technology). However, the Illumina technology had been the main technology implemented in viral population studies.

1.13.2.1 Illumina

Illumina became the standard NGS methods in microbial genomics due to its low cost and high efficiency. It works using sequencing-by-synthesis (SBS), where the four nucleotides along with DNA polymerase are added to a flow cell channel containing primers. PCR is then performed to generate multiple clustered copies of the target.

To detect these, each nucleotide is linked to a distinct fluorescent label (Zhang et al., 2011). The NGS process comprises the following sequence of reactions: i) nucleotides added to elongating strands of DNA; ii) identification of the nucleotides incorporated into each sequenced fragment; and iii) washing by removing the fluorescent labels that block the groups to facilitate detection of the subsequent reaction (Mardis, 2011). Further details on this process are provided in chapter two. Illumina NGS technology had been used in different approaches depending on the objective of the sequencing strategy ranging from metagenomic approached in virus discovery studies, amplicon sequencing strategy where a targeted region is amplified to understand the diversity and recently the target enrichment approach where certain pathogens are targeted in a clinical sample while excluding all other DNA present in the sample. The three main approaches are described in (Figure 1-6).

1.13.2.1.1 Metagenomic Sequencing

Metagenomics has come to the fore as an area of scientific research that facilitates analysis of microbial communities within an environmental sample that is independent of culture

(Culligan et al., 2014). The field has grown exponentially since its emergence in 1998, allowing researchers to develop an unprecedented understanding of the non-culturable microbes both in our environment and within our body (Culligan et al., 2014; Hayes et al., 2017). The field has been strengthened by the introduction of NGS methods that generate a greater volume at a lower cost and in less time than earlier sequencing methods. It is now widely applied in the field of deep sequencing (Thomas et al., 2012).

1.13.2.1.2 Amplicon sequencing approach

The principle of this approach is to enrich the specific viral genome before sequencing. Specific primers utilize for PCR amplification of viral genetic material that are complementary to a recognized nucleotide sequence. This approach has been knowing the most common for enriching small viral genomes, such as HBV and HIV virus (Houldcroft et al., 2017). Amplicon sequencing approach is more successful than metagenomic approach for whole-genome sequencing from samples that have low virus concentrations (Thomson et al., 2016).

1.13.2.1.3 Target enrichment approach

By targeting specific regions only, target enrichment is employed in NGS workflows to eradicate genomic DNA regions that are of no experimental interest (Kozarewa et al., 2015). The substantially increased number of reads specific to the viral genome constitutes the primary advantage of a target enrichment approach over metagenomics. This facilitates the assessment of genetic diversity within hosts. Such data are vital for analysing minority variants that are potential precursors to any changes in pathogenicity (Tsetsarkin et al., 2007).

1.13.2.1.4 Errors and Limitations of NGS using Illumina technology

To perform NGS, a small number of non-specific PCR steps is required, leading to a PCR-based error (Poh et al., 2013). In contrast to traditional PCR-based methods, this error can be reduced by restricting the number of PCR cycles and utilising high-fidelity polymerase enzymes. However, the technique is extremely sensitive and thus susceptible to cross-contamination errors. Furthermore, compared to methods such as Sanger sequencing, the sequence reads generated using the Illumina platform are generally shorter. This is because the NGS read length is restricted by the signal-to-noise ratio (Mardis, 2013). To assess error rates, control samples can be incorporated into each run (Hillier et al., 2008b). Multiple instruments have been created to align substantial amounts of data from short fragment reads

generated by NGS instruments with a reference genome. Various techniques can be employed to identify an over- or under-enriched area, or to define the variants or single nucleotide polymorphisms (SNP).

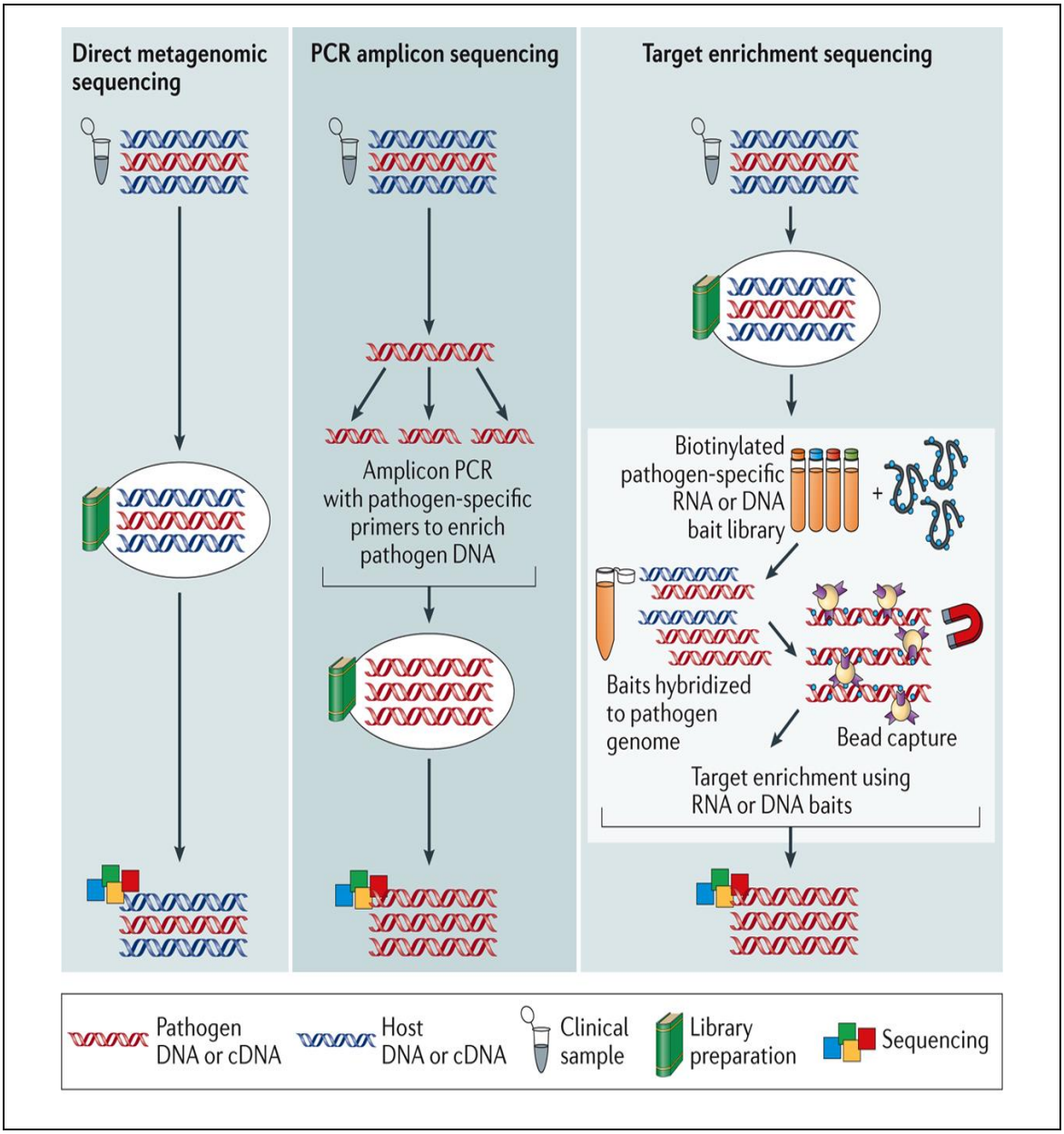


Figure 1-6: Illustration of main approaches of NGS (Houldcroft et al., 2017).

Metagenomic approach offers a simple and accurate description of the sequences in the sample, even though at high sequencing and data analysis. It presents lower required number of PCR cycles and no primer or probe design required. PCR amplicon approach uses many PCR reactions to enrich the viral genome. Most sequencing reads in this approach will be pathogen-specific, which decreases sequencing costs (highly specific) and it offer decent coverage even at low viral load sample (highly sensitive). Target enrichment approach uses virus-specific nucleotide probes that are bound to beads in order to enrich the viral genome in a single reaction (highly specific) and introduce less PCR cycles than amplicon approach.

1.14 HBV Sequencing

HBV whole-genome sequencing offers the opportunity to explore mutations associated with drug and vaccine resistance, diagnostic failure, and different clinical outcomes, as well as to study viral transmission and evolution (Gunther et al., 1995; Chook et al., 2015; Wang et al., 2017). Sanger sequencing has also provided insights into HBV genetic sequencing, HBV diversity and its role in pathogenesis (Gunther et al., 1995). However, this method generates only a consensus-level sequence and cannot capture information on HBV quasispecies (McNaughton et al., 2019). HBV evade the host immune system is through the generation of individual sequences (haplotypes) that change their antigenic profile. Viral haplotypes were explored historically by Sanger sequencing and conventional cloning, which limited the analysis to a small number of samples due to the laborious nature of this approach (Wagner et al., 2021). Viral haplotype reconstruction programs and NGS are now facilitated the deeper characterization of the viral haplotype range. Recently, a haplotype study has identified HBV cell-specific infectivity, mutations that lead to immune escape, and transmission of haplotypes between individuals (Wagner et al., 2021).

1.14.1 Sequencing of HBV genome using Next-generation Sequencing (NGS)

The NGS platforms, such as Illumina, can reveal the full landscape of HBV quasispecies in individual patients, allowing a more detailed analysis of the within- and between-host distribution of resistant mutations (McNaughton et al., 2019). NGS is superior to the routine diagnostic methods for understanding the viral population due to its enhanced ability to detect minority variants with a frequency as low as 1% (Abdelrahman et al., 2015), there have been reports of lower frequencies (McNaughton et al., 2019). NGS assays have been introduced in clinical laboratories over the last decade, based on their ability to provide comprehensive, detailed information. Illumina is considered the most robust sequencing platform due to its ability to capture the level of within-sample diversity to a high degree of accuracy (McNaughton et al., 2019). The most common type of error identified in Illumina is a substitution error. This occurs when poor performance of the de-blocking step results in a cluster falling out of phase or incomplete cleavage of the fluorescence label prior to DNA cycles, which generates interface noise that results in identification of the wrong nucleotide (Mardis, 2013). Nevertheless, Illumina is the preferred choice of platform for most genome-sequencing projects because it generates superior quality data with low error rates (Zhang

et al., 2011). A library comprising millions of DNA fragments is produced by NGS. These fragments can be read from both sides, a technique known as paired-end sequencing that facilitates the identification of unusual fragment lengths. It also provides information on duplicate overlap areas, which leads to an enhanced analysis of NGS data. Utilised to align sequences to the reference genomes, alignment scripts are extremely accurate because they consider the lengths of the synthesised DNA fragments held in the sequence library (Korbel et al., 2007).

The introduction of NGS to the infectious disease field has elucidated the intra-host viral population and enabled scientists to address key questions pertaining to viral genomics and the natural history of diseases (Berry et al., 2020). Detailed analyses of the HBV genome can enhance understanding of disease pathogenesis, help in the development of new therapeutics, enhance treatment, and make accurate predictions of patient outcomes.

Third-generation sequencing techniques, as compared with second-generation methods, are employed to sequence long DNA and RNA molecules. In this domain, the Oxford Nanopore Technologies (ONT) MinION and Pacific Biosciences (PacBio) are now the leaders in commercialised technology (Slatko et al., 2013). The establishment of the ONT MinION and other third-generation platforms allow sequencing where the read length has no theoretical upper limit, thereby permitting sequencing of viral genomes or complete genes in one read (McNaughton et al., 2019; Astbury et al,. 2020). Although Nanopore-based DNA sequencing was proposed during the late 1990s, it is only in recent times that the ONT has succeeded in commercialising it. This was achieved by using a portable MinION (512 nanopore flowcell channels), a high throughput PromethION (in development, 48 flow cells of 3000 nanopores each), and benchtop GridION (5 mimIONs in a single module) (Slatko et al., 2013; McNaughton et al., 2019; Astbury et al., 2020). These sequencers apply protein nanopores in a polymer membrane which is electrically resistant, and causes typical current changes as each nucleotide moves through the detector (Slatko et al., 2013; McNaughton et al., 2019). However, a methodology for the MinION platform has been developed lately. This has been applied to haplotyping and HBV sequencing in order to use the genome's circular properties for the purpose of creating single genome replicates (Slatko et al., 2013; McNaughton et al., 2019; Astbury et al., 2020).

In the PacBio detection procedure and real-time sequencing imaging, it is possible to measure the rate of each nucleotide addition in the synthesis process, otherwise known as the inter-pulse duration (IPD) (Slatko et al., 2013). This is a significant advantage. It is

possible to recognise most nucleotides with base modifications because they change the IPD. These include certain cytosine and adenine methylations (Slatko et al., 2013).

It is possible to identify and record several different modifications for the purpose of epigenetic studies. However, not every modification may be recognised at this point, because of a small adjustment of the IPD (Slatko et al., 2013). Nevertheless, there is a potential for chemical modification of such nucleotides to be identified. Exceptionally lengthy fragments (up to 30–50 kb or longer) may be enabled by PacBio sequencing, otherwise known as Singe Molecule Real Time (SMRT) (Slatko et al., 2013). PacBio SMRT sequencing allows long reads for genome assemblies, and rapid detection of methylation sites for epigenetic studies; therefore, it has several advantages over other techniques. Recently, two workflows were developed in order to create sequencing data from dried serum spot (DSS) sampling by utilising MinION and Sanger platforms, thus acquiring totally analysed sequences from the receipt of samples within less than 48 hours for both techniques (Astbury et al., 2020). For the six samples that were tested, deep sequencing data were created successfully. Additionally, the silico processing produced consensus sequences identical to those generated by Sanger sequencing (Astbury et al., 2020). Following its development, the assay enabled the determination of HBV genotype, as well as the elucidation of HBV resistance associated substitutions to 600 IU/mL by applying a 550bp amplicon. A MinION sequencer that was sequenced a 1.2 kb amplicon produced results that were consistent with Sanger sequencing, permitting the recognition of small populations of variants (Astbury et al., 2020).

1.14.2 Bioinformatic analysis of NGS data

Assembling and then mapping sequence reads to a reference sequence is the most vital stage in NGS data analysis. Given the enormous volume of data produced, two important areas need to be addressed: the error profile and how the data will be used. Using traditional methods, a number of days are needed to map the data to the original reference genomes or use computationally intensive software such as BLAT, BLAST, or Smith-Waterman dynamic programming.

1.15 Current Status in Saudi Arabia and Problem Statements

Chronic HBV infection in Saudi Arabia is a significant problem that requires a concerted, multi-disciplinary research effort to address critical questions relating to the epidemiology, pathogenesis, treatment, and natural history of the disease.

In an epidemiological study published in 1988, the seroprevalence of HBsAg among Saudi children was reportedly 7%, and at least one HBV marker was reported as being positive in 70% of the screened children (Abdo & Sanai, 2015). This study activated a nationwide response, resulting in the introduction of a national universal HBV immunisation programme in 1989. A catch-up vaccination programme followed for children of school age, healthcare workers, and other at-risk groups. These national efforts resulted in a high vaccination uptake rate, with virtually all Saudis aged 27 years or younger being vaccinated either at birth or on starting school.

HBV continues to be a major health problem in Saudi Arabia, even though a vaccination programme has been implemented. Some recent studies in the country suggest that the prevalence rate for HBV is 3.6% (Schweitzer et al., 2015) while others report a lower rate of 1.3%, indicating a gradual decline in HBV infection rates (Aljumah et al., 2019; Alswaidi et al., 2013). However, numerous studies have reported a much greater rate of HBV infection in high-risk populations, such as HIV positive patients and intravenous drug users (Alhuraiji et al., 2014; Alzahrani et al., 2009; Al-Qahtani, 2020).

During the last two decades, a substantial decline has been observed in the frequency of HBV infection in Saudi Arabia; this was evidenced in the decline of HBV prevalence in a study that compared the HBV prevalence between two periods (2002-2005) and (2012-2015) in Saudi Arabia. In this study, the incidence of HBV was assessed amongst couples attending an in vitro fertilisation clinic in a tertiary care hospital. The researchers found that the prevalence of HBV significantly decreased from 4.7 % in 2002-2005 to 1.67% in 2012-2015 (Alzahrani et al., 2009; Aljumah et al., 2019; Al-Qahtani, 2020).

As mentioned above, Saudi Arabia has witnessed a steady decline in HBV infection over the past 30 years (Aljumah et al., 2019). This could be attributable to childhood immunisation against HBV, improved living conditions, increased awareness of safe clinical practices, and the mandated screening programme implemented in the country to identify patients through blood tests (Table 1-6). Regardless of this decline in prevalence, challenges remain concerning CHB management, and chronic liver disease will continue to place enormous pressure on the Saudi healthcare system for many years, with additional challenges arising when a substantial number of undiagnosed cases starts to emerge at advanced stages of the disease.

Table 1-6: Potential HBV screening points in Saudi Arabia.

Category	Recommendation
Pre-marital screening	Mandatory
Pre- employment (Expatriate)	Mandatory
Health care workers and medical students	Mandatory
Pre-employment (Saudi) in military and police	Mandatory
Blood donation units	Mandatory
Dialysis patients	Mandatory
Pregnancy	Recommended
Any contact with infected patients	Recommended
Drug addicts	Recommended
Prison inmates	Recommended
Before a patient's procedure	Done as a part of the hospital regulations
Patients about to start immunotherapy and pre-transplant patients	Done as a part of the hospital regulations

Several mandated and recommended screening forms present in the country aim to identify patients through blood tests (Aljumah et al., 2019).

Over the next two to four decades, an increase is expected in the number of patients with liver cirrhosis who experience complications and will need a more intensive level of care (Alghamdi et al., 2012). To reduce rates of morbidity and mortality within Saudi Arabia, substantive work is needed by all primary stakeholders, including policy makers, in every phase of the HBV care pathway. A crucial goal would be an initiative that encompasses early detection, proper and timely assessment, and successful treatment. However, achieving this will need a fully coordinated effort that involves ensuring liver transplantation is readily available to patients with end-stage liver disease. Moreover, with respect to the epidemiology of viral hepatitis within the Saudi Arabia, a great deal depends on the successful in-depth evaluation of risk factors associated with acquiring both HBV and HCV,

and large epidemiological, community-based studies of HBV prevalence in older, non-vaccinated populations (Abdo et al., 2012).

Regarding the treatment choice for chronic HBV patients, the Saudi Association for the Study of Liver Disease and Transplantation (SASLT) developed a series of HBV practice guidelines that were published in 2014 (Abaalkhail et al., 2014). However, specialists also use other guidelines from the World Health Organization (WHO), the European Association for the Study of the Liver (EASL), and the American Association for the Study of Liver Disease (AASLD) (Aljumah et al., 2019; WHO, 2017; EASL, 2017; Terrault et al., 2016).

The antiviral therapies for treatment recommended in Saudi Arabia for CHB are TDF and ETV. However, others – pegylated interferon (Peg-IFN-2a), ADV, LAM, and LdT - have also been approved. Although all these drugs have received market authorisation, only TDF, ETV, and LAM are widely available and free of charge to Saudi patients when prescribed by a doctor (Aljumah et al., 2019).

The molecular characterisation of HBV is understudied in the Middle East. There is a dearth of studies characterising the circulating variants in Saudi Arabia and antiviral resistance-associated mutations. The available data on the dominant genotype in the Kingdom are based on the sequencing of the S gene rather than whole-genome sequences. However, there is limited data covering the prevalence of resistance-associated mutations in treatment-naïve patients.

In Saudi Arabia, the predominant HBV genotypes are genotype D and E. Knowledge of the HBV serologic and genotypic patterns would inform disease management guidelines and strategies for eradicating virus infection (Asaad et al., 2015).

Phyloepidemiology of HBV is an essential step towards eradicating HBV in Saudi Arabia as the genotype and subgenotype of HBV are related to the clinical outcome, transmission network and treatment outcome (Bui et al., 2017). HBV genotyping is not a routine diagnostic test in microbiology laboratories, due to the limited sequencing facilities and inaccuracy of the available technologies. There has been a recent drive globally to identify circulating variants in the population to improve the clinical decision-making, including creating treatment guidelines, identifying prevention strategies via disease modelling, and ensuring adequate health resource allocation for managing chronic HBV patients (Bui et al., 2017).

The diagnosis of HBV infection is based mainly on the detection of HBsAg. However, the presence of a mutation in the S gene had a profound effect on the diagnostic sensitivity of

the current serological tests, mainly in the “a” determination; basically, this might lead to false-negative results. The antigenic alterations may also help the virus to escape from the host immune system (Bernard, 2005). This poses an excellent aspect for studies to unveil the circulating variants in Saudi Arabia, as there is a reported genetic diversity in the “a” determinant (a short segment of 24 amino acids), even in different geographical regions within the same country (Al Qudari, 2016).

The implication of the genetic variability in the S gene is well-described at both the diagnostic and clinical levels, as the mutations lead to the emergence of immune escape variants that are missed by the current diagnostic tests based on the detection of HBsAg, which increases the number of false negatives in the laboratories.

Avellon et al. (2006) reported that the substitution mutations in the S region in the HBV variants caused diagnostic failure in 12.5% of the cases where HBV was misdiagnosed, and the mutant variant is also responsible for 6.6% of cases of invalid vaccination, while 9.2% escape immunoglobulin therapy.

Next-generation sequencing (NGS) is superior to routine diagnostic methods for understanding the viral population with an enhanced ability to detect minority variants with a frequency as low as 1% (Abdelrahman et al., 2015).

1.16 Aim of the Thesis

The aim of this project is to expand our understanding of the viral diversity of HBV in Saudi Arabia using NGS and to develop an *in vitro* model for investigating HBV replication, which includes setting up the infectious/non-infectious virus replication systems.

Hence, the main objectives of this project were as follows:

1. Design and validation of target enrichment probe set to perform HBV full genome sequencing using NGS technology.
2. Utilize the new protocol to identify the circulating variants in Saudi Arabia with its antiviral resistance profile.
3. Establishing different assays to determine HBV replication levels.
4. Characterizing viral replication in HBV infectious and non-infectious replicons.
5. Generating a panel of mutations potentially associated with drug resistance in the replicon.
6. Functionally analysing these mutants in the presence and absence of the drugs of interest in the replicon system to ascertain their role in drug-resistance, if any.

Chapter 2. Materials and Methods

2.1 Materials

The kits and reagents (Table 2.1) and equipment (Table 2.2) used in this thesis are outlined below with the manufacturer's information supplied. This allows for reproducibility between studies.

Table 2-1: Kits

Kit	Manufacture
NEBNext® mRNA Second Strand Synthesis Module.	New England BioLabs Inc.
KAPA LTP Library Preparation Kit.	KAPA Biosystems Inc.
Nextera XT® DNA Sample Preparation Kit.	Illumina.
NEBNext® Microbiome DNA Enrichment Kit.	New England BioLabs Inc.
Agencourt® Ampure® XP beads.	Beckman Coulter.
Tapestation D1000 ScreenTape System.	Agilent.
Bioanalyzer DNA 7500 Kit.	Agilent.
MagNa Pure® Total Nucleic Acid Extraction kit.	Roche Diagnostics.
Qiagen Plasmid Maxi Kit.	Qiagen.
QIAprep Spin Miniprep Kit.	Qiagen.
QiaQuick Gel Extraction Kit.	Qiagen.
Monarch® DNA Gel Extraction Kit.	New England BioLabs Inc.
MINI-PROTEAN TGX 4-20% 1mm 10 well.	BioRad 4561093
QuickTiter™ Hepatitis B "e" Antigen (HBeAg) ELISA Kit.	Cell Biolabs, Inc.
Human hepatitis B virus e antigen (HBeAg) ELISA Kit.	CUSABIO Technology LLC.

Human HBeAg ELISA Kit (Colorimetric).	Bio-Techne Ltd.
Chemiluminescence Assay Kit.	BioRad.

Table 2-2: Equipment

Equipment	Manufacture
M220 Focused-Ultrasonicator	Covaris Limited.
2200 TapeStation	Agilent Technologies.
Qubit® 2.0 Fluorometer	Thermo Fisher Scientific.
MiSeq System Desktop Sequencer	Illumina.
Nanodrop 1000 spectrophotometer	Thermo Fisher Scientific.
Varioskan microplate reader	Thermo Fisher Scientific.
Odyssey® CLx. Imaging System	LI-COR Biosciences.
EVOS M5000 Imaging System	EVOS.
Applied Biosystems™ 7500 Real-Time PCR Systems	Life Technologies.

The reagents (Table 2.3), chemicals used (Table 2.4) mammalian cells used and something for the in vitro studies undertaken in this thesis are outlined below. The manufacturer or supplier along with the catalogue number is supplied,

Table 2-3: Cell Culture Reagents Utilised

Component	**Manufacture**	**Catalogue number**
Dulbecco's Modified Eagles Medium (DMEM)	Gibco (Invitrogen Life Technologies)	41966-029
DMEM/F-12+GlutaMAX (Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12)	Gibco (Invitrogen Life Technologies)	10565-018
Opti-Minimum Essential Medium (Opti-MEM)	Thermo Fisher Scientific	31985062
10% Fetal calf serum (FCS)	Gibco (Invitrogen Life Technologies)	A4766801
Fetal Bovine Serum (FBS)	Gibco (Invitrogen Life Technologies)	10270-106
100 units/ml Penicillin/streptomycin	Gibco (Invitrogen Life Technologies)	15140-122
1x Trypsin (10x Stock)	Sigma	C100ML-5927C
Versene in PBS	E&O laboratories	BM0400
Phosphate buffer solution (PBS)	Gibco (Invitrogen Life Technologies)	10010015
Non-essential amino acids: MEM NEAA (100X)	Gibco (Invitrogen Life Technologies)	11140-035
L-glutamine	Gibco (Invitrogen Life Technologies)	25030081
1000x Ampicillin (100 mg/ml)	Melford Laboratories	A40040-1.0
1000x G418 (100 mg/ml) – Gentamicin	Melford Laboratories	G64000-1.0
Puromycin	Sigma-Aldrich	P8833
Dimethyl sulphoxide (DMSO)	Sigma	D2660

Table 2-4: Multipurpose Chemicals/Reagents used in the Study

Chemical	Abbreviation	Manufacture
Ethanol	EtOH	Sigma-Aldrich
Nuclease free water	NFH2O	Thermo Fisher Scientific
Sodium hydroxide	NaOH	Thermo Fisher Scientific
Sodium Chloride	NaCl	Thermo Fisher Scientific
Ethidium Bromide	EtBr	Thermo Fisher Scientific
Isopropanol	IPA	Thermo Fisher Scientific
3,3'5,5'-tetramethylbenzidine substrate	TMB	Sigma-Aldrich
Tris-acetate-EDTA	TAE	Thermo Fisher Scientific
Tris-HCl		Sigma-Aldrich
Methanol		Thermo Fisher Scientific
Phosphate Buffered Saline	PBS	Thermo Fisher Scientific
HEPES: HEPES, free acid		Melford
Skimmed Milk Powder		Marvel
S.O.C. Media		Invitrogen
Sodium dodecyl sulphate solution	SDS	Sigma-Aldrich
Tris		Roche
Triton-X-100		Promega
Tween 20		Millipore
Formaldehyde		Thermo Fisher Scientific
Glycerol		BDH
Sucrose		Sigma-Aldrich
PageRulerTM Prestain NIR Protein Ladder		Thermo Scientific

The illumine reagents (Table 2.5), buffer used (Table 2.6), and material for transfection (Tables 2-7 and 2-8) are outlined below. The manufacturer and composition are supplied,

Table 2-5: Illumina Sequencing Reagents

Reagent	Abbreviation
Tagment DNA Buffer	TD
Neutralize Tagment Buffer NT	NT
Nextera PCR Master Mix NPM	NPM
Hybridization buffer	HT1

Table 2-6: Buffers made in-house

Name	Component
Blocking buffer (FBS)	5% filtered FBS in PBS.
Blocking buffer (milk)	5% Skimmed Milk Powder in PBS.
DNA loading buffer	65% Sucrose (w/v), 10 mM Tris, 10 mM EDTA.
PBS-T	0.05% Tween-20 in PBS.
Coomassie Brilliant Blue Stain	0.12% Coomassie Brilliant Blue powder, 50% Methanol, 10% Acetic acid in water.
Coomassie destain Solution	5% Methanol, 7% Acetic acid in water.
Coomassie fix solution	50% Methanol, 10% Acetic acid in water.
TAE Buffer (50x)	40 mM Tris, 1 mM EDTA, 5 mM sodium acetate (pH7.6).
Trypsin-versene	2.5% trypsin (v/v) in versene.

Table 2-7: Transfection Reagents

Reagent	Source
Lipofectamine 3000	Life Technologies
Opti-MEM I	Gibco
Calcium Phosphate Transfection Kit	Sigma-Aldrich

Table 2-8: SDS-PAGE and Western Blot Reagent

Solution	Composition
Sample loading buffer	31% stacking gel buffer, 31% glycerol, 21% SDS (25%), 9% 2-mercaptoethanol, 1μg/ml BPB, 8% H2O.
Gel running buffer	5mM Tris-HCl, 200mM glycine, 0.1% (w/v) SDS.
Resolving buffer	1.5M Tris-HCl (pH 8.9), 0.4% SDS.
Stacking buffer	0.5M Tris-HCl (pH 6.8), 0.4% SDS.
Towbin buffer	25mM Tris-HCl (pH 8.0), 200mM glycine, 20% (v/v) Methanol.
PBS-T	PBS plus 0,05% (v/v) Tween 20.
Blocking buffer	5% (w/v) milk in PBS-T.

Table 2-9: Bacterial Propagation

Solution	Component	Source
Luria-Bertani (LB) - broth	170mM NaCl, 10g/l Bactopeptone, 5g/l yeast extract.	E&O Laboratories
LB-agar	LB-broth plus 1.5% (w/v) agar.	E&O Laboratories

Table 2-10: DNA Manipulation

Solution	Component
DNA loading dye	30 % glycerol, 0.25 % bromophenol blue, 0.25 % xylene blue.
TBE (10x)	8.9 M Tris-borate, 8.9 M boric acid, 0.02 M EDTA (pH 8.0).
TE buffer (Qiagen)	0.01M Tris-HCl (pH 8.0), 0.001M EDTA, ddH2O water, pH adjusted to 8.0.

Table 2-11: DNA Analysis

Solution	Component
TAE Buffer	Tris 245.375g, Acetic acid (17.5%), EDTA dehydrate (18.5g).
Agarose	0.8-1.5% (w/v) diluted in 1x TAE (A9539 Sigma-Aldrich).
DNA loading buffer	0.4% orange G, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 15% Ficoll® 400, 10mM Tris-HCl (pH 7.5) and 50mM EDTA (pH 8.0); Promega.
P1 buffer (Qiagen)	Tris 3.03g, EDTA dihydrate 1.86mg, pH 8.0. RNase was added at 100μg/ml.
P2 buffer (Qiagen)	SDS (20%), NaOH 4g.
P3 buffer (Qiagen)	Potassium acetate 147.25g and Acetic acid until reach pH 5.5.

Table 2-12: Cell Lysis

Solution	Component
Cell Lysis Buffer (LB2)	20 mM Tris-HCl pH 7.4, 20 mM iodoacetamide, 150 mM NaCl, 1 mM EDTA, 0.5 % Triton X-100

Table 2-13: Cell lines

Name	Type	Growth and maintenance media	Growth and maintenance conditions
Huh7	Human Hepatoma	DMEM supplemented with 10% FBS, penicillin, 100 µg/ml streptomycin, 0.1 mM non-essential amino acids and 2 mM glutamine.	37˚C, 5% CO2
HEK-293T	Human Embryonic Kidney	DMEM supplemented with 10% (v/v) FCS, supplemented with 0.1 mM non-essential amino acid.	37˚C, 5% CO2
HepG2-hNTCP-C4	Human liver cancer	DMEM/F-12+GlutaMAX (Invitrogen,10565-018), 10% FBS, 100 U/ml penicillin, 100 µg/ml streptomycin, 5 µg/ml insulin, 10 mM HEPES, 500 ng/ml G418 (actually 500 µg/ml).	37˚C, 5% CO2
HepG2	Human liver cancer	DMEM/F-12+GlutaMAX (Invitrogen,10565-018), 10% FBS, 100 U/ml penicillin, 100 µg/ml streptomycin, 5 µg/ml insulin, 10 mM HEPES.	37˚C, 5% CO2
Hep38.7-Tet	Human liver cancer	Same as that for HepG2-hNTCP-C4 cells above, except also add 400 ng/ml tetracycline (sigma, T8032).	37˚C, 5% CO2
Huh7-hNTCP (Huh7N)	Human Hepatoma	Same as that for HepG2-hNTCP-C4 cells above.	37˚C, 5% CO2
Huh7 express 3B4	Human Hepatoma	Same as that for Huh7 cells above.	37˚C, 5% CO2

Table 2-14: Plasmids

Plasmid	Description	Source
pTHBV- L^-	Plasmid carries 1.3x HBV genotype D genome. It lacks the ATG start codon of L-HBsAg and therefore is L-null (assembly defective) but replication-competent. It can be used at CL2. It expresses the HBV genes efficiently (and also replicates efficiently) in Huh7 cells.	Provided by Professor Arvind Patel
pMD2-G	For expression of VSV envelope protein for pseudotyped lentivirus production.	Provided by Professor Arvind Patel
MLV-Gag/Pol	Encoding MLV gag and pol genes cloned downstream of a human CMV promoter.	Provided by Professor Arvind Patel
L-HBsAg	Expresses HBV Large envelope glycoprotein in mammalian cells.	Addgene (103011)
S-HBsAg	Expresses HBV Small envelope glycoprotein in mammalian cells.	Addgene (103013)
M-HBsAg	Expresses HBV Medium envelope glycoprotein in mammalian cells.	Addgene (103012)
Pol	Expresses HBV polymerase in mammalian cells.	Addgene (65520)
HBV 1.3-mer WT replicon	Plasmid carries the 1.3x HBV genotype D sequence in its WT form. It is replication- and assembly competent, so upon transfection into Huh7 or HepG2 cells tis plasmid initiates HBV DNA replication, and subsequent assembly and release of the virus that could be used in infection studies. Work involving this plasmid should be done in CL3.	Addgene (65459)
pSVL(D3)	Hepatitis delta virus genome from cloned DNA.	Addgene (29335)
pSVL (D2m)	Hepatitis delta virus genome from cloned DNA.	Addgene (29336)
pQ-HBV 1.3 mer wt (2B3)	The entire HBV genomic sequences from pHBV1.3mer WT (from Addgene) subcloned into pQCXIP plasmid which has the puromycin-resistance gene downstream. In theory, this plasmid would allow generation of cell lines stably expressing WT HBV. Work involving this plasmid (or the stable cell line carrying it) should be done in CL3.	Made during PhD
pQ-HBV 2.7 (3B4)	Plasmid carries a shorter form of HBV sequence. It doesn't express HBV core or polymerase and therefore it is replication-defective and non-infectious. It expresses all 3	Made during PhD

	HBsAgs (i.e. L, M, and S) and HBx. It has puromycin-resistance marker so can be used to select a stable cell line.	
pT-HBV- L^- Puro (1 – 6)	The entire HBV genomic sequences from pTHBV- L^- subcloned into pQCXIP plasmid which has the puromycin-resistance gene downstream. In theory, this plasmid would allow generation of cell lines stably expressing HBV1.3- L^-	Made during PhD
pTHBV- L^- Puro (11)	The entire HBV genomic sequences from pTHBV- L^- subcloned into pQCXIP plasmid which has the puromycin-resistance gene downstream. In theory, this plasmid would allow generation of cell lines stably expressing HBV1.3- L^-	Made during PhD
pQ-HBV1.3mer WT (2-8)	The entire HBV genomic sequences from pHBV1.3mer WT (from Addgene) subcloned into pQCXIP plasmid which has the puromycin-resistance gene downstream. In theory, this plasmid would allow generation of cell lines stably expressing WT HBV. Work involving this plasmid (or the stable cell line carrying it) should be done in CL3.	Made during PhD
pTHBV- L^- Puro (2)	The entire HBV genomic sequences from pTHBV- L^- subcloned into pQCXIP plasmid which has the puromycin-resistance gene downstream. In theory, this plasmid would allow generation of cell lines stably expressing HBV1.3- L^-	Made during PhD
pTHBV- L^- Puro (4)	The entire HBV genomic sequences from pTHBV- L^- subcloned into pQCXIP plasmid which has the puromycin-resistance gene downstream. In theory, this plasmid would allow generation of cell lines stably expressing HBV1.3- L^-	Made during PhD

Table 2-15: Primer Sequences

Primer ID	Sequence	
HB1	5’ – CCTTTTACACAATGT - 3’ (Forward)	nt 1026-1040
HB2	5’ – TCACATCACCATAC – 3’ (Forward)	nt 2047-2060
HB3	5’ – CACGAGTCTAGACTCT – 3’ (Reverse)	nt 260-145
HB4	5’ – TTTGCTCTGAAGGCTG – 3’ (Reverse)	nt 2926-2911
HB5	5’ – GTATGGTGATGTGA – 3’ (Reverse)	nt 2061-2047
HBV-6	5’ – GGTATATTATATAAGAGAGA – 3’ (Forward)	nt2771-2790
HBV-7	5’ - TATTGTGAGGATTCTTGT – 3’ (Reverse)	nt240-223
HBV-8	5’ – TTGGATCATCCAAATTACTA – 3’ (Reverse)	nt2140-2121
HBV-9	5’ – CTCAGGAGACTCTAAGGCTT – 3’ (Reverse)	nt 2040-2021
HBV-10	5’ – GGTTATCCTGCTTTAATGCC – 3’ (Forward)	nt 1041-1060
HBV-11	5’ – TAAAACTACCTCTCTTCCAA – 3’ (Reverse)	nt 2240 – 2231
HBV-12	5’ – AATGTGGTTATCCTGCTTTAAT – 3' (Reverse)	nt 1036-1057
HBV-13	5’ – CATGGCCTCTCTGGCCACTGAAGGAAAGAAGTCAGAAG – 3’ (Reverse)	nt 1206-1226
HBV-14	5’ – CATGCTAGCTGTATTCAATCTAAGCA – 5’ (Forward)	nt 2100-2116
HBV-15	5’ – CATGAATTCCACTGCATGGCCTGAGGATGA – 3’ (Reverse)	nt 6-26
HBV-16	5’ – CATGAATTCCACAACCTTCCACCAAACTCT – 3’ (Forward)	nt 3162-3181
HBV-17	5’ – CATCTCGAGAAGATTGACGATAAGGGAGAG – 3’ (Reverse)	nt 3061-3079
HBV-18	5’ – CATCTCGAGGATTGGGGACCCTGCGCTGAACAT – 3’ (Forward)	nt 3030-3052
HB19	5’ – TATAGTGAGTCGTATTAATAGCT – 3’ (Forward)	nt 28 – 50 of pTHBV-L$^-$
HB20	5’ – TACAATTAATACATAACCTTAT – 3’ (Reverse)	nt 135 – 114 of pTHBV-L$^-$
HB21	5’ – ATACACTATTCTCAGAATGACTT – 3’ (Forward)	nt 617 - 143 of pTHBV-L$^-$
HB22	5’ – TTAACGTGAGTTTTCGTTCCAC – 3’ (Forward)	nt 1312 - 1333 of pTHBV-L$^-$

HB23	5’ – GCTTTACACTTTATGCTTCC – 3’ (Forward)	nt 2283 - 2301 of pTHBV-L$^-$
HB24	5’ – GGAAGCATAAAGTGTAAAGC – 3’ (Reverse)	nt 2301 – 2283 of pTHBV-L$^-$)
HB25	5’ – GTGGAACGAAAACTCACGTTAA – 3’ (Reverse)	nt 1333 - 1312 of pTHBV-L$^-$
HB26	5’ – AACCAAGTCATTCTGAGAATAGTGTAT – 3’ (Reverse)	nt 643 - 617 of pTHBV-L$^-$
HB27	5’ – TGAACCTCCTCGTTCGACC – 3’ (Forward)	nt 1501 – 1519 of pQCXIP
HB28	5’ – GGAACATACGTCATTATTGA – 3’ (Reverse)	nt 1687 – 1665 of pQCXIP
HB29	5’ – CAT TGATCACTGCAGGGCCCGTCGACAAGCTT – 3’	nt 2378-2400 of pTHB-L$^-$
HB30	5’– CAT TGATCATAGAATACGAATTCGAGCTCGTATT – 3’	nt 5542-5566 of pHBV1.3-mer-WT
HB31	5’ – CAT TGATCACAATCTCGGGAATCTCAATGTTAGTATT – 3’	nt 101-125 of pHBV1.3-mer WT
HB32	5’ – CATAAGCTTCAATCTCGGGAATCTCAATGTTAGTATT – 3’	nt 101-125 of pHBV1.3-mer WT
HBV2268	5′ – GAGTGTGGATTCGCACTCC-3′ (Forward)	nt 3602-3620 of pTHBV-L$^-$
HBV2372	5′ – GAGGCGAGGGAGTTCTTCT-3′ (Reverse)	nt 37706-3724 of pTHBV-L$^-$
HBV1803	5′ – TCACCAGCACCATGCAAC-3′ (Forward)	nt 6320-6336 of pTHBV-L$^-$
HBV1872	5′ – AAGCCACCCAAGGCACAG-3′ (Reverse)	nt 6393-6410 of pTHBV-L$^-$
ccc-1582	5′ – TGCACTTCGCTTCACCT-3′ (Forward)	nt 6096-6112 of pTHBV-L$^-$
ccc-2316	5′ – GACCACCAAATGCCCCT-3′ (Reverse)	nt 3632-3638 of pTHBV-L$^-$

Table 2-16: Enzymes

Enzyme	Manufacture
RNase H	Sigma-Aldrich Co Ltd
DNase I	Sigma-Aldrich Co Ltd
Proteinase K	Sigma-Aldrich Co Ltd
T5 exonuclease	New England Biolabs (NEB)
End repair enzyme	KAPA Biosystems
A-tailing enzyme	KAPA Biosystems
Mung bean nuclease	New England Biolabs (NEB)
T4 Ligase	New England Biolabs (NEB)

All restriction enzymes were supplied by New England Biolabs (NEB) and used throughout the project to confirm successful cloning by digestion, prior to the sequencing of the cloned plasmids.

Table 2-17: List of Primary Antibodies

Antibody Name	Description	Type	Source / Catalogue number
10-H10M	Monoclonal HBeAg antibody, Anti-HBeAg antibody	Mouse	Fitzgerald Industries Int'l Rdi Div
10-H10N	Monoclonal HBeAg antibody, Anti-HBeAg antibody	Mouse	Fitzgerald Industries Int'l Rdi Div
NTCP Antibody	Polyclonal antibody to NTCP	Rabbit	Biorbyt Ltd (orb13624)
NTCP Antibody	Polyclonal antibody to NTCP	Rabbit	Thermo Fisher Diagnostics Ltd (PA5-80001)
Gamma-Tubulin	Monoclonal Anti-Gamma-Tubulin	Mouse	Sigma-Aldrich Co Ltd
HBeAg	SKU: REC31677-100 Native Antigen Company: Hepatitis B Virus e Antigen (HBeAg)	Mouse	2B SCIENTIFIC
HBV X	Hepatitis B Virus X Monoclonal Antibody (X36C)	Mouse	Life Technologies Ltd
HBV core	Hepatitis B Virus Core Antigen Monoclonal Antibody (1-5)	Mouse	Life Technologies Ltd
HBV S	Hepatitis B Virus Surface Monoclonal Antibody (S 26)	Mouse	Life Technologies Ltd
Hep Pol	SC-81590 Hep Pol (2C8)	Mouse	Insight Biotechnology Ltd
NTCP	SLC10A1 / NTCP Antibody (aa141-154)	Mouse	SOURCE BIOSCIENCE (C312823-100)
Gamma-Tubulin	Gamma tubulin antibody	Mouse	Provided by Prof. Arvind Patel
RC28	Anti-HBV preS1 MAb	Mouse	Provided by Prof. Arvind Patel
R193	Anti-HBV core polyclonal serum	Rabbit	Provided by Prof. Arvind Patel
R143	Anti-HBV preS1 polyclonal serum	Rabbit	Provided by Prof. Arvind Patel
HBV cAg AB-1	Anti-HBV core polyclonal antibodies	Rabbit	Neomarkers Inc (RB-1413-A)
HBsAg [1834]	Hepatitis B Virus Surface Antigen 1834 Monoclonal Antibody 100ug	Mouse	INSIGHT BIOTECHNOLOGY LTD (GTX40707)

Table 2-18: List of Secondary Antibodies

Antibody Name	Species	Source
IRDye® 680RD anti-Rabbit IgG (H + L), 0.5 mg (P/N 926-68073)	Donkey	LICOR UK
IRDye® 800CW anti-Mouse IgG (H + L), 0.5 mg (P/N 926-32212)	Donkey	LICOR UK
IRDye 800CW anti-Mouse IgG2b Specific, 0.5 mg lot# 926-32352	Goat	LICOR UK
Anti-Rabbit IgG H&L (Alexa Fluor® 488) (ab150077)	Goat	Abcam
Anti-Mouse IgG (H+L) (Alexa Fluor® 488) (ab150133)	Goat	Abcam
Anti-Rabbit IgG (H+L) (Alexa Fluor® 488) (A21206)	Donkey	Life Technologies
Anti-Rabbit IgG (H+L) (Alexa Fluor® 647) (A31573)	Donkey	Life Technologies
61-H10K, HBEAG ANTIBODY (HRP), Monoclonal HBeAg antibody.	Mouse	Fitzgerald Industries Int'l Rdi Div
Hepatitis B Virus Surface Antigen HRP Polyclonal Antibody (GTX19990)	Goat	Insight Biotechnology Ltd

Table 2-19: Hepatitis B Virus Drugs

Drug	Source	Catalogue number
Adefovir dipivoxil	MedChem Express	HY-B0255-50 mg
Tenofovir (Disoproxil)	MedChem Express	HY-13782A-50 mg
Telbivudine	MedChem Express	HY-B0017-10 mg
Lamivudine	Cayman Chemical	CAY18514-50 mg
Entecavir (hydrate)	Cayman Chemical	CAY13831-10 mg

2.2 Methods

2.2.1 Sample cohort

Sixty-four samples with known viral loads ranging from 5.59 x 10^3 to 1.17 x10^7 IU/ml were collected from the Microbiology Diagnostic Laboratories in tertiary care centres: King Fahad Specialist Hospital (KFSH) in Dammam, Saudi Arabia, and King Fahd Medical City (KFMC), Riyadh, Saudi Arabia. The Medical Ethics Committee approved the protocol for this study (Institutional Review Board log number 17- 404 at KFMC). The specimens were sent to the MRC-University of Glasgow Centre for Virus Research, Glasgow (CVR), UK, for deep sequencing and analyses. All the samples were anonymised and were supplied without associated clinical data.

2.2.2 DNA extraction

DNA samples were extracted from the plasma samples using MagNA Pure LC® (Roche Diagnostics) using the MagNa Pure® Total NA Extraction kit, following the manufacturer's recommendations, at the Microbiology Diagnostic Laboratories at KFMC. The DNA extracts were submitted to MRC-University of Glasgow, Centre for Virus Research, UK. The samples were stored at -80 degrees until processed for sequencing.

2.2.3 DNA Second Strand synthesis

Double-stranded DNA was synthesised and then purified using AMPure XP beads (Agencourt, Beckman Coulter) (Section 2.2.6.2).

To create double-stranded cDNA, the NEBNext® mRNA Second-Strand Synthesis Module (New England BioLabs) was used. The cDNA products were mixed with 48 µL of nuclease-free water, 8 µL of 10X second-strand synthesis reaction buffer, and 4 µL of second-strand synthesis enzyme. Incubation then took place for 2.5 h at 16°C.

2.2.4 Library Preparation

Two methods were tested to fragment the DNA in order to prepare the libraries to be used for deep sequencing on Illumina-sequencing platforms.

2.2.4.1 Physical or Mechanical Shearing

This method was conducted using the AFA Ultrasonicator® (Covaris M220) with Covaris Adaptive Focused Acoustics technology. The parameters employed for the shearing were taken from the Covaris M220 protocol, with a target fragment size of 400 bp (Table 2.1).

Ultrasonication produces DNA fragments containing staggered overhangs. Library preparation was performed via the KAPA LTP Library Preparation Kit® for Illumina Platforms (KAPA Biosystems), according to the following steps:

2.2.4.1.1 End Repair:

Production of blunt-ended, 5'-phosphorylated fragments. End Repair Enzyme master mix was used containing T4 DNA Polymerase, T4 Polynucleotide Kinase, and 10X End Repair Buffer with dNTPs.

2.2.4.1.2 A-tailing:

dAMP was added to the 3'-ends of the dsDNA library fragments. In brief, was added the volume of each sample was increased to 50 µL by adding 20 µL of 10 mM Tris-HCl pH 8.0. The samples were then washed as described previously. The beads were dried for five minutes and then eluted in 10 mM Tris-HCl pH 8.0. Next, 4 µL of a master mix containing 10x A-tail buffer and A-tail enzyme was added. The solution was thereafter incubated for 60 minutes at 30˚C.

2.2.4.1.3 Adapter Ligation:

Ligating dsDNA adapters containing 3'-dTMP overhangs were fused to the A-tailed library fragments. We diluted the adapter in water to achieve a solution of 0.15 µM. Then, a master mix was prepared for each sample, containing a buffer, T4 ligase, adapter, and water. An aliquot of 12.5 µL of the master mix was added to each 12.5 µL sample, following which the solution was incubated at 20°C for 15 minutes. We next eluted each sample in 14 µL 10 mM Tris-Cl pH 8.0. Finally, a 10 µL sample was transferred to a fresh tube to perform PCR (Figure 2- 1).

2.2.4.1.4 Library amplification:

PCR was carried out to amplify the library fragments that carried suitable appropriate adapter sequences on both ends. To each sample, 45 µL of Solid Phase Reversible Immobilization (SPRI) beads were added at a ratio of 0.9. The beads were then washed, as described in Section 2.2.6.2). The samples were prepared as follows:

a) 12.5 µL 2X KAPA HF Real-time Hot Start Mix.

b) 1.25 µL 10 µM Universal Primer.

c) Aliquot 13.75 µL of the previous master mix to each 10 µL sample and mixing by pipette.

Add 1.25 µL 10 µM Index Primer individually to each sample. Then, PCR was performed in a thermocycler in accordance with the following protocol: 98°C for 45 sec and 98°C for 15 sec, after which the amplified DNA was purified using AMPure XP magnetic beads, as described in Section (2.2.6.2).

Table 2-20: The Parameters used for the Physical Fragmentation of DNA

Parameter	Value
Target BP (Peak)	500
Duty Factor	10%
Peak Incident Power	450
Cycles per Burst	1000
Treatment Time: 8 microTUBE-50 Strip V2	45 Sec & 80 Sec
Temperature (˚C)	7
Water Level	- 2
Sample Volume (µL)	55

2.2.4.2 Enzymatic shearing

The Nextera XT® DNA Sample Preparation kit containing a synthetically manufactured transposome was employed to perform shearing. The DNA strands were cut into ~300 bp

fragments while the DNA was tagged with adapter sequences. In addition to two flow-cell attachment sites, the adapter sequences contained binding sites for dual index sequences that were unique for each sample library (Figure 2-2). To help optimise the protocol, five samples were tested with Nextera XT® in accordance with the manufacturer's instructions.

2.2.4.2.1 Input DNA Quantification and Tagging

The Nextera XT® DNA Sample Preparation library preparation procedure makes use of a biological DNA fragmentation step. Compared to mechanical fragmentation techniques, this is a more sensitive method for DNA fragmentation. Furthermore, it requires only 1 ng input DNA depending on the undetectable starting input in our samples, the tagging time was halved to avoid over-fragmentation. The protocol was optimized by taking the following steps:

Firstly, 10 µL of Tagment DNA (TD) Buffer was added to a 0.2 ml tube, followed by 5 µL of input DNA at 0.2 ng/µL (1 ng total) and then 5 µL of Amplicon Tagment Mix (ATM). The reaction was then mixed and centrifuged at 280 g for one minute at 20°C. The reaction was then placed in a thermocycler at 55°C for two and a half minutes, then maintained at 10°C. At this temperature, 5 µL of the Neutralisation Tagment (NT) Buffer was added to each reaction to initiate neutralization. Each sample was then mixed lightly and centrifuged at 280 g for one minute. The samples were then left to settle at room temperature for five minutes before proceeding to PCR amplification.

2.2.4.2.2 PCR Amplification

For this step, a limited-cycle PCR program was used to amplify the tagmented DNA. Index 1 (i7) and index 2 (i5), along with the sequences required for cluster formation and the PCR step, were then added. The Nextera PCR Master Mix (NPM) and the index primers were then thawed at room temperature for 30 minutes, following which 15 µL of NPM was added to each reaction tube containing 5 µL of each index and lightly mixed. The reaction was then centrifuged at 280 g at 20°C for one minute, following which PCR was carried out in a thermocycler in accordance with the following conditions: 72°C for three minutes, 95°C for 30 sec, 12 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 30 sec, extension at 72°C for 30 sec and 72°C for five minutes, and final stabilisation at 10°C.

2.2.5 Enrichment Process

2.2.5.1 Host depletion using the NEBNext Microbiome DNA Enrichment Kit

This section describes a novel method for enriching microbial DNA in which methylated host DNA, such as human genomic DNA, is selectively bound and separated from microbial DNA prior to the construction of the NGS library. The principle of the kit is to allow microbial DNA from samples containing methylated host DNA to be enriched through the selective binding and removal of the CpG-methylated host DNA. It is striking that the microbial diversity remains intact after enrichment. To selectively bind and remove the CpG-methylated host DNA, the kit uses a rapid, straightforward magnetic bead-based method. Further, the technique utilises the MBD2-Fc protein, comprising the methylated CpG-specific binding protein, MBD2, fused to the Fc fragment of human IgG. The Fc fragment binds easily to Protein A, facilitating an attachment to the Protein A-bound magnetic beads. The MBD2 domain of this protein binds tightly and specifically to CpG-methylated DNA. Applying a magnetic field then extracts the CpG-methylated (eukaryotic) DNA, which leaves non-CpG-methylated (microbial) DNA in the supernatant (Figure 2-3). The following steps were used in the protocol.

2.2.5.1.1 Prebinding MBD2-Fc protein to magnetic beads

First, 1 µL of MBD2-Fc-bound magnetic beads without vortexing was added to a sample containing 6.25 ng of input DNA into a 1.5 ml low-binding tube. Next, 0.1 volume was added MBD2-Fc protein magnetic bead solution. The tube was placed on a rotating mixer for ten minutes. Following incubation, the tube was spun briefly and placed it onto the magnetic rack for two to five minutes until the beads had been collected. At that point, the supernatant was removed without disturbing the beads. Subsequently, 1 ml of 1X of ice-cold Bind/wash buffer was added to the tube to wash the beads. The samples was pipetted up and down a few times to mix. Afterwards, the tubes was placed on a rotating mixer for three minutes at room temperature. Next, the sample was centrifuged, and placed it on the magnetic rack for two to five minutes and removed the supernatant. Lastly, a volume equal to 1X Bind/wash Buffer was added that calculated during the first step to resuspend the beads.

2.2.5.1.2 Capture-methylated host DNA

At the start, 1/4 sample volume of 5X Bind/wash buffer was added. then added the volume of the sample to yield 1 µg of DNA. This was followed by mixing and incubating within a

rotating mixer at room temperature for 3-4 h. Following separation with a magnet (for at least five minutes), the supernatant was transferred to fresh tubes (this is the host-depleted DNA).

Later, 1.8X volume of ampure beads was added to the depleted samples and incubated the mixture for at least 15 minutes with separation via a magnet (at least five minutes) thereafter and removed the supernatant. The remaining solution was washed twice with 800 µL 80% ethanol and air dried for three to seven minutes, taking care not to over-dry the beads. Finally, the mixture was eluted in 9.5 µL of 10 mM pH 8. Tris, at 9 µL, and was used for the dsDNA synthesis.

2.2.5.2 Target-enrichment Method

Using the proprietary tool of NimbleGen, a DNA probe was designed that was placed complementary to the target regions. Enrichment of the in-solution target was conducted following the NimbleGen standard protocol for preparing an Illumina library. The samples for enrichment initially contained relatively high amounts of DNA (based on the pre-enrichment PCR) and, after the first wash, the non-specific DNA was removed, resulting in far lower quantities of DNA. Therefore, to prevent this low amount of DNA from becoming contaminated, these stages were separated specifically in the enrichment room. For this, a reagent was initially prepared in a clean hood. Second, hybridization and pre-capture were performed at the pre-enrichment station. Third, post-capture was conducted at the post-enrichment station. Later, library pools was prepared according to the viral load for each sample and calculated the total volumes so that each pool contained a total of 11 µg DNA (minimum 0.5 µg). If it were less than 0.5 µg, a double capture was considered to be performed. The following optimised capture protocol was adhered to:

2.2.5.2.1 Hybridisation (pre-enrichment station)

Library multiplex was prepared as follows: each component with an equal viral load and a total mass of 1 µg or 0.5 µg. Then, 1 µg of multiplexed library was mixed and prepared in 1.1 with 5 µL of COT DNA.

Thereafter, 2000 pmol (2 µL) of mixed oligos was mixed with the previous DNA mix. Then, added 2X volumes of AMPureXP/ KAPA pure, mixed incubated at room temperature for ten minutes. Next, the tube was placed on a magnetic rack until the beads and solution had fully separated before carefully removing the supernatant. After that, left the solution to air dry for five minutes. During this time, a master mix was prepared for each sample under the

following conditions: 7.5 µL of hybridisation buffer pH 7.5 with 3 µL hybridisation component A 3. Later, 10.5 µL of the master mix was added to bead-bound DNA samples, removed them from the magnet and mixed them thoroughly, and then followed this with incubation at room temperature for 2 minutes. Each tube was placed on a magnetic rack, and eluted 10.5 µL (entire volume) to a new tube containing a SeqCap EZ probe pool, mixed thoroughly, and incubated in a PCR machine at 95°C for five minutes, then cooled to 47°C. Finally, the tubes quickly was transferred to a second PCR machine at 47°C for 72 h.

2.2.5.2.2 Washing and recovering captured multiplex DNA

In this step, reagents was prepared in the clean hood for each capture as follows: two tubes of 20 µL 10x stringent wash buffer with 180 µL of water in each, 10 µL in one tube of 10x wash buffer 1 mixed with 90 µL of water plus 20 µL in the other tube mixed with 180 µL of water using the same buffer. In addition, there were two tubes of 20 µL of 10x wash buffer 2 and buffer 3 mixed with 180 µL of water in each, and, lastly, 200 µL of 2.5x bead wash buffer was mixed with 300 µL of water.

200 µL of the 1x Stringent Wash Buffer and 100 µL of the 1x Wash Buffer 1 was transferred to PCR machine, which was set to 47°C. At that time, the Capture Beads and SeqCap EZ Pure Capture Bead Kit were equilibrated to room temperature for 30 minutes. Then, the beads was vortexed for 15 sec before placing the tubes on a magnet. Later, the liquid was removed and resuspended the beads in 1x original volume of 1x Wash Buffer, followed by transferring 100 µL resuspended beads per capture to a 0.2 ml PCR tube.

Next, the capture beads was added to the 15 µL samples. Then mixed the tube by pipetting 10 times and vortexing 10 times. This step was repeated using an additional 200 µL 1x Stringent Wash Buffer, 200 µL 1x Wash Buffer 1, 200 µL 1x Wash Buffer 2, and 200 µL 1x Wash Buffer 3. Finally, eluted in 20 µL water.

- Amplifying captured multiplex DNA using LM-PCR.

  First, a master mix was and PCR reactions of 25 µL KAPA HiFi HotStart ReadyMix with 5 µL Post-LM-PCR Oligos 1+2 (5 µM) for each sample. Then, aliquoted 15 µL of the prepared reaction into two reaction tubes per capture and briefly vortexed the bead-bound captured DNA. Later, added 10 µL of bead-bound captured DNA to each reaction tube, followed by a PCR run. After that, combined 2 x 25 µL reactions for each pool to make a total of 50 µL. Next, 45 µL ampure XP (ratio 0.9) was added, mixed each sample by pipetting up and down, and then incubated it at room temperature for five minutes. The samples were placed on a magnetic rack until the beads and solution had fully

separated; this was followed by the removal of the supernatant, washing 2x with 200 µL 80% EtOH, eliminating all traces of ethanol, air drying for five minutes, and elution in 17.5 µL of Tris. Finally, the sample was quantified with Qubit® as described in (Section 2.2.7.1) and verified the size with TapeStation, as described in (Section 2.2.7.2).

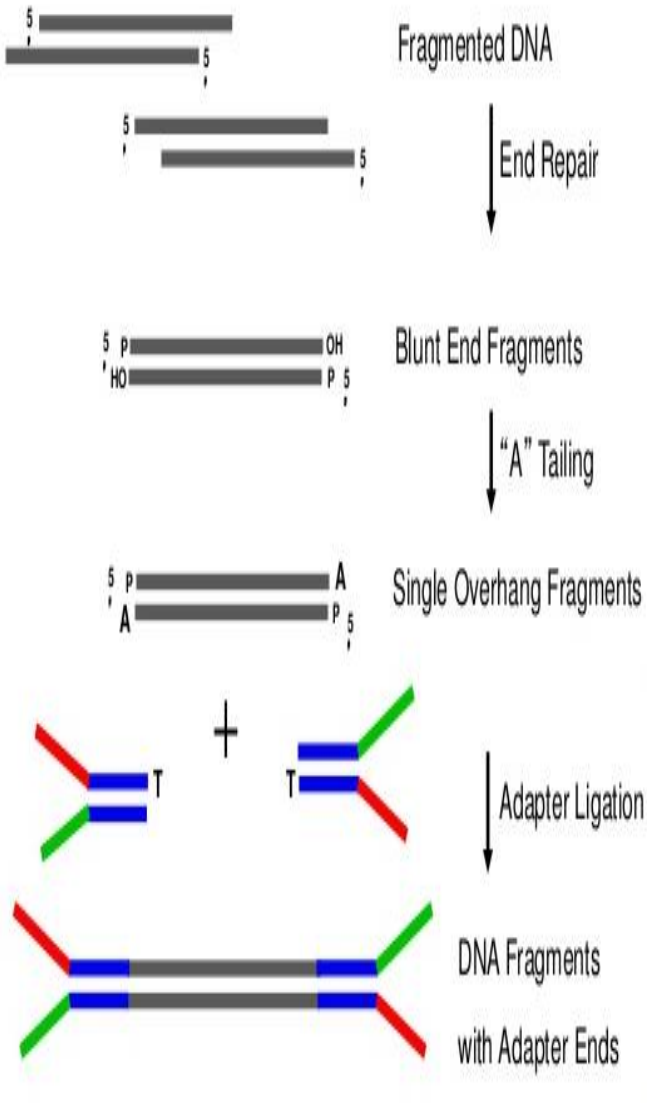


Figure 2-1: NGS Library Preparation using KAPA LTP Kit.

Schematic depiction of the method for DNA library construction using KAPA LTP Library Preparation Kit Illumina® Platforms (KAPA Biosystems).

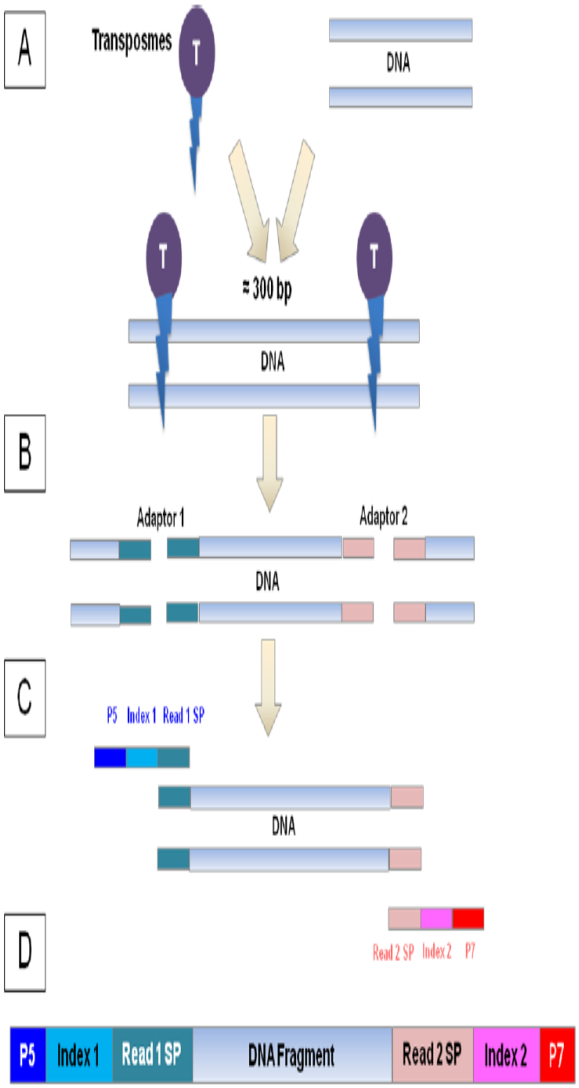


Figure 2-2: Nextera XT® workflow.

A) Nextera XT transposome with adapters combined with template DNA. B) Tagmentation to fragment the DNA and the addition of the adapters. C) Limited cycle PCR to add sequencing primer sequences and indices (SP: sequencing primer). D) Sequence-ready fragment.

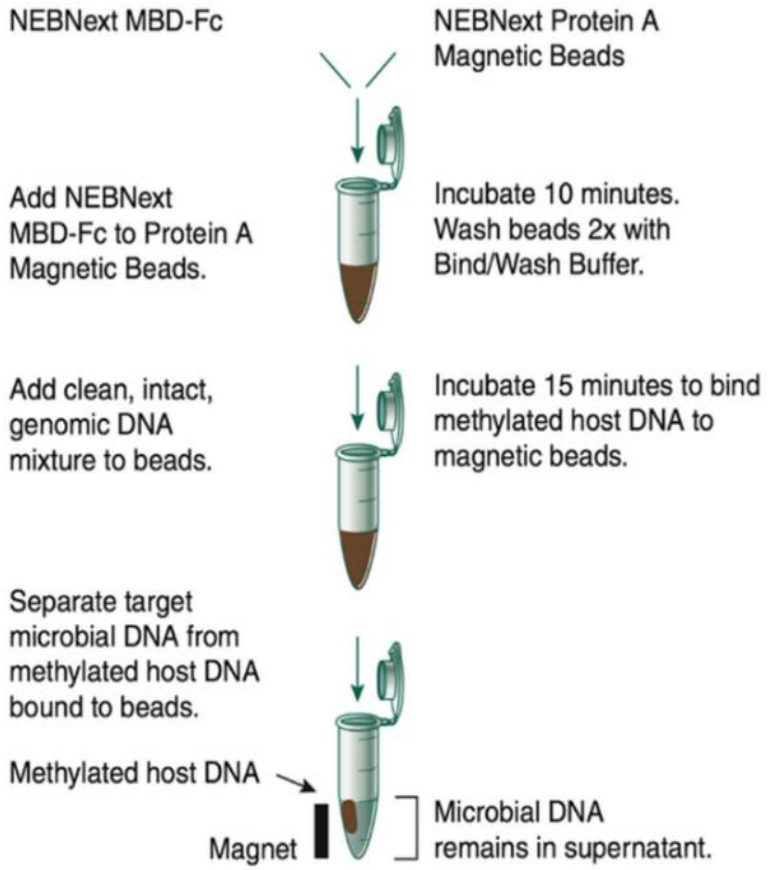


Figure 2-3: NEBNext Microbiome DNA Enrichment Kit.

Schematic illustration of microbiome DNA enrichment using MBD-fc (NEBNext Microbiome protocol, 2018).

2.2.6 DNA purification

2.2.6.1 Isolation and Purification of DNA from Agarose Gels

First, the DNA fragment was excised from an agarose gel, placed in a microcentrifuge tube, and weighed. Binding buffer was then added at a weight/volume ratio of 1:1. The chaotropic agent in the binding buffer denatures proteins, dissolves agarose, and promotes DNA binding to the silica membrane in the column. The solution was incubated at 50°C for 10 minute, or until the gel slice had entirely dissolved and the mixture had turned yellow (which denotes an optimal pH for DNA binding). Next, 800 µL of the solubilized gel solution were placed in a purification column and centrifuged for one minute at 13,000 revolutions per minute (rpm). The flow-through was then discarded, the column returned to the collection tube, and 100 µL of binding buffer was added to the purification column. The column was then centrifuged for one minute, the flow-through discarded, and the column placed back in the collection tube. Next, 700 µL of wash buffer was added to the purification column and centrifuged for one minute, following which the flow-through was disposed of and the column returned to the collection tube. The vacant purification column was then centrifuged for a further one minute to ensure all the residual wash buffer was removed. Finally, 50 µL of elution buffer was added to the center of the purification column membrane, following which the purification column was transferred to a clean 1.5 ml microcentrifuge tube and centrifuged for one minute. The purified DNA was then stored at –20°C.

2.2.6.2 Agencourt AMPure XP® beads

Before use, magnetic beads were warmed to room temperature. DNA was added and mixed thoroughly, and the mix was incubated for three to five minutes at room temperature. To separate the beads from the solution, the reaction was placed on a magnetic plate for five to ten minutes. Once fully separated, the liquid was aspirated from the reaction tube and disposed of. The beads were then washed twice by adding 200 µL of newly prepared 80% ethanol. The DNA was then eluted in 40 µL of nuclease-free water (Figure 2-4).

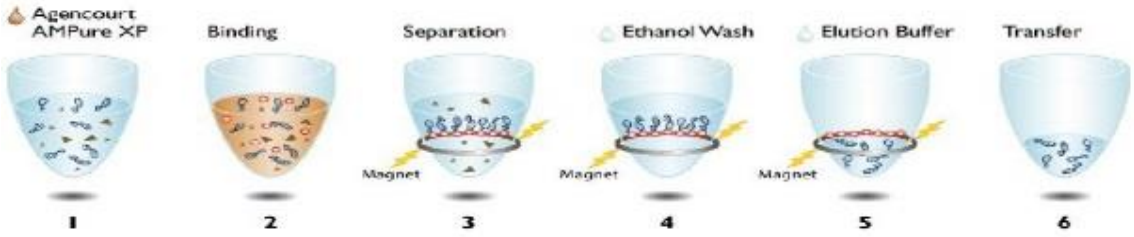


Figure 2-4: Workflow for PCR Purification using Agencourt AMPure XP® beads.

1. AMPure XP. 2. Binding of the DNA fragments to paramagnetic beads. 3. Separation of the beads + DNA fragments from contaminants. 4. Beads + DNA fragments were washed twice using 70% ethanol to remove contaminants. 5. Purified DNA fragments from the beads were eluted. 6. The product was transferred to a new tube (Beckman Coulter User Guide, 2013).

2.2.7 Measuring DNA concentration

2.2.7.1 Measurement of nucleic acid concentration using Qubit®

Using a combination of 200 µL of buffer and 1 µL of dye for every sample, a fluorescent dye working solution was prepared in a plastic tube. Following vortexing, an aliquot of 190 µL of working solution was added to two 0.5 ml Eppendorf tubes. These were labelled standard 1 and standard 2. Ten µL of standard 1 (0 ng/ µL) and standard 2 (10 ng/µL) were then added to the corresponding tube and mixed by vortexing. At the same time, a 198 µL aliquot of working solution in 0.5 ml Eppendorf tubes were mixed with 2 µL of each sample, also by vortexing.

During Qubit analysis, all samples and standards were placed in tubes supplied by the manufacturer. A Qubit® 2.0 Fluorometer (Invitrogen) was used to measure the concentration of DNA while a Qubit® ds DNA High Sensitivity Assay kit was utilized to measure the quantity of DNA. The fluorometer calculates the concentration of nucleic acid in the original suspension and presents the result in ng/µL.

2.2.7.2 Nucleic acid QC using the 2200 TapeStation

An automated electrophoresis platform with flexible throughput capabilities, the Agilent 2200 TapeStation® system can be used for just one sample up to an entire 96-well plate. D1K ScreenTape analysis of the 2200 TapeStation facilitates an analysis of DNA fragments ranging in size from 35 bp to 1000 bp. It can therefore be employed to analyse library preparations. The analysis was performed following the manufacturer's guidelines. To prepare libraries for TapeStation analysis, 1 µL of the sample was mixed with 3 µL of D1K Sample Buffer Mixtures and loaded onto the TapeStation. The protocols for the R6K were similar, although in this case 1 µL High Sensitivity R6K Sample Buffer was mixed with 2 µL of the RNA sample. The samples were then denatured by incubating them at 72°C for three minutes and placing them on ice for two minutes. We then briefly centrifuged the samples in order to retrieve the contents at the base of each tube.

2.2.8 Bacterial Techniques

2.2.8.1 Culturing

To grow and amplify the DNA plasmids, we used three laboratory strains of *Escherichia coli* (*E. coli*) as described in Section 2.2.8.2.1. All bacterial strains were grown in a sterile Luria-Bertani (LB) broth (E&O Laboratories Ltd) that contained an appropriate plasmid selection antibiotic; either 50 µg/ml kanamycin (Kana) (Sigma Aldrich) or 100 µg/ml ampicillin (Amp) (Melford Laboratories Ltd.). Agar plates were prepared by pouring melted sterile LB agar (E&O Laboratories Ltd), along with the appropriate antibiotic, into 10 cm2 Petri dishes (Greiner Bio-One) and allowed to set. Using a plastic spreader (VWR), the bacteria transformed with appropriate plasmid were streaked onto the surface of the cooled agar. The plate was then inverted and incubated overnight (approximately 18 h – 20 h) at 37°C. Single colonies were then selected and inoculated in an LB broth that contained the appropriate antibiotic. Once again, the cultures were incubated overnight at 37°C and subjected to orbital shaking at 225 rpm for approximately 18 h to 20 h. Amplified DNA plasmids were then purified from the bacteria using Miniprep (Section 2.2.8.3.1) or Maxiprep (2.2.8.3.2).

2.2.8.2 High efficiency Transformation Protocol

2.2.8.2.1 NEB 5-alpha competent E. coli®

First, a tube of NEB 5-alpha competent *E. coli* cells was thawed on ice for ten min. Next, 1-5 µl of H_2O containing an average 50 ng of plasmid DNA was added to the cell mixture. The contents of the tube were then mixed by flicking 4-5 times. Next, the mixture was placed on ice for 30 minutes before being subjected to heat-shock at 42°C for 30 sec. The mixture was then placed back ice for five minutes, after which 950 µL was pipetted into the SOC medium and positioned in a shaker (250 rpm) at 37°C for 60 minutes. Upon completion of incubation, 100 ul of the mixture was spread on an agarose plate which contained antibiotics for colony selection and incubated at 30°C for 24-36 h.

2.2.8.2.2 NEB 10-beta Competent *E. coli*®

First, one tube of NEB 10-beta Competent E. coli cells were thawed on ice, following which 50 ng of the plasmid DNA were added in a volume of 1-5 µL. The contents of the tube were then mixed by flicking 4-5 times. Next, the mixture was placed on ice for 30 minutes before being subjected to heat-shock at 42°C for 30 sec. The mixture was then returned to ice for

five min, after which 950 µL was pipetted into the SOC medium and placed on a shaker (250 rpm) at 37°C for 60 minutes. Upon completion of incubation, 200 ul of the mixture was spread on an agarose plate which contained antibiotics for colony selection and incubated at 30°C for 24-36 h.

2.2.8.2.3 One shot® Top 10 cells

For this phase, one shot® Top 10 cells from Invitrogen® were used in accordance with the manufacturer's instructions. First, one vial of One Shot® TOP10 cells chemically competent for each transformation was thawed on ice. Next, approximately 5 µL (depending on the concentration) of DNA (10 pg to 100 ng) was added to a vial of One Shot® cells and lightly mixed. Pipetting up and down was then performed so that the vial containing the DNA and the cells were not mixed. The vial(s) were then incubated on ice for 30 minutes. In the next step, they were subjected to heat-shock for exactly 30 sec at 42°C without shaking, removed from the bath, and placed on ice for two minutes. 250 µL of pre-warmed SOC Medium was then added to each vial aseptically. The vial(s) were then capped tightly and shaken horizontally at 37°C for 1 h at 225 rpm in a shaking incubator. Following this procedure, 150 µL from each transformation was spread on a pre-warmed selective plate and incubated for 24-36 h at 30°C. The remaining transformation mix was then stored at 4°C before proceeding to the downstream step.

2.2.8.3 Bacterial Cloning

2.2.8.3.1 Small scale plasmid preparation from transformed bacteria

To inoculate a 2 ml culture of LB with selective antibiotics, a single colony from a newly streaked agar plate was used. Following culture for 24 h at 37 ºC with vigorous shaking (180 rpm), the bacteria were centrifuged in an Eppendorf microcentrifuge at 13,000 rpm. In accordance with the manufacturer's instructions, the QIAprep®Spin Miniprep Kit (QIAGEN), was then used to remove the DNA from the bacterial pellet. All purification steps were performed at room temperature. Centrifugation was conducted in a microcentrifuge at 13,300 rpm, following which the pelleted cells were resuspended in 250 µL of a resuspension solution. To make sure all the bacteria were completely suspended, up and down vertexing and pipetting were performed until cell clumps were no longer evident. Next, 250 µL of lysis solution was added and mixed thoroughly by inverting the tube 4-6 times until it became transparent and viscous. Next, 350 µL of neutralisation solution were added and mixed thoroughly in the same way. To pellet the cell debris and DNA, the mixture was centrifuged for 10 minutes. Next, 800 µL of the supernatant obtained was applied to a QIAprep 2.0 spin column by decanting or pipetting and then centrifuged for one min. Washing was then performed by adding 700 µL of buffer PE to the QIAprep®Spin column, which was centrifuged for 30-60 sec and the flow-through disposed. To remove all residual wash solution, this process was repeated, and the column centrifuged for an additional one minute. After transferring the QIAprep®Spin column into a fresh 1.5 ml microcentrifuge tube, elution was performed by adding 50 µL of buffer EB (10 mM Tris-Cl, pH 8.5) to the centre of the QIAprep®Spin column membrane. The tube was then incubated for two min at room temperature and centrifuged for two minutes. Finally, the purified plasmid DNA was stored at -20°C.

2.2.8.3.2 Large scale plasmid preparation from transformed bacteria

To inoculate a 500 ml starter culture of LB-selective antibiotics, one colony from a freshly streaked selective agar plate was selected and used. Using a plasmid Maxi kit (Qiagen) in accordance with the manufacturer's instructions, albeit with a small modification, the plasmid DNA from bacteria cultures was isolated on a large scale. First, one colony of transformed bacteria was applied to inoculate 2ml of the starter culture of the LB-selective antibiotic. Following incubation for 2 h, the starter culture was transferred to 200 ml of LB with antibiotics and cultured overnight. Having been incubated for approximately 18 h at 37°C with 250 rpm shaking, the bacteria were harvested by centrifuging at 5,000 x g for 20

minutes at 4°C. The pellet was then fully re-suspended in 10 ml chilled buffer P1 and transferred to a 50 ml universal tube (Corning). The bacteria were then lysed by adding 10 ml of buffer P2 and inverting the tube vigorously six times. The reaction was mixed as before and then incubated for five minutes at room temperature. Next, 10 ml of chilled buffer P3 was added and mixed by inverting the tube another six times. To purify the lysate, a coffee filter was then placed in a QIAGEN-tip 500 through which an 11 ml buffer QBT was passed for equilibration. The lysate was then transferred from the coffee filter to the universal filter. While the cell debris remained, the cleared supernatant holding the DNA passed through the coffee filter into the QIAGEN-tip. Once all supernatants had passed through the resin, the QIAGEN-tip was washed using 60 ml buffer QC. The bound DNA was eluted with 15 ml buffer QF and collected in a clean 50 ml universal tube. To precipitate the DNA, 10.5 ml room temperature isopropanol was added, and the sample was centrifuged at 5,000 x g for 60 minutes at 4°C. To remove all traces of the isopropanol, the pellet was then washed with 5 ml 70 % EtOH and centrifuged. The pellet was then air-dried for approximately ten minutes and the supernatant decanted. The pellet was then re-suspended in 300-500 µL TE buffer (Qiagen), measured as before, and stored at -20°C prior to use.

2.2.8.4 Quantification of Nucleic Acids

A NanoDrop ND-1,000 spectrophotometer (Thermal Fisher Scientific) was utilized to quantify all the nucleic acid (dsRNA, ssRNA and dsDNA) yields. A 2 µL aliquot from each sample was placed on the measurement pedestal and a reading taken of concentration and purity. Samples were deemed ‘pure’ at a 260/280 ratio of ~2.0 for RNA and ~1.8 for DNA. The sequencing of DNA (PCR products or/and plasmid constructs) was outsourced to Eurofins along with appropriate primer. The product requirements recommended by the companies were followed.

2.2.9 Sequencing using the MiSeq® Platform

In this process, the flow cell provided within the MiSeq reagent kit was removed from its buffer and washed using distilled H_2O. To ensure the outlet and inlet ports were clear, and that the surface was dust-free and clean, it was dried with a lint-free wipe, The flow cell was then loaded onto the MiSeq instrument. Next, the chiller compartment was opened and the wash bottle removed and replaced with an incorporation buffer. The reagent cartridge was then loaded into the machine and the run was initiated. A Phi X 174 (phiX) bacteriophage genome library is mixed with the sample pool to offer a calibration control for cross-talk

matrix generation, phasing and prephasing, and also a quality control for cluster generation, sequencing, and alignment. This can be aligned at speed to estimate relevant sequencing using synthesis (SBS) metrics, such as the error rate and phasing. By adding 2 μL of a 10nM stock solution to 3 μL H2Om Phi X control DNA (Illumina®, UK), dilution to 4 nM was achieved. The Phi X library was then denatured by adding 5 μL of 0.2 N NaOH to 5 μL Phi X, which was then vortexed and incubated at room temperature for five minutes. This was subsequently diluted to a concentration of 20 pM by adding 980 μL HT1 buffer, following which it was diluted to the same loading concentration as the DNA library. 100 μL of Phi X was then added to the 90 μL DNA library to create a final library with a 10% concentration of Phi X (Figure 2-5).

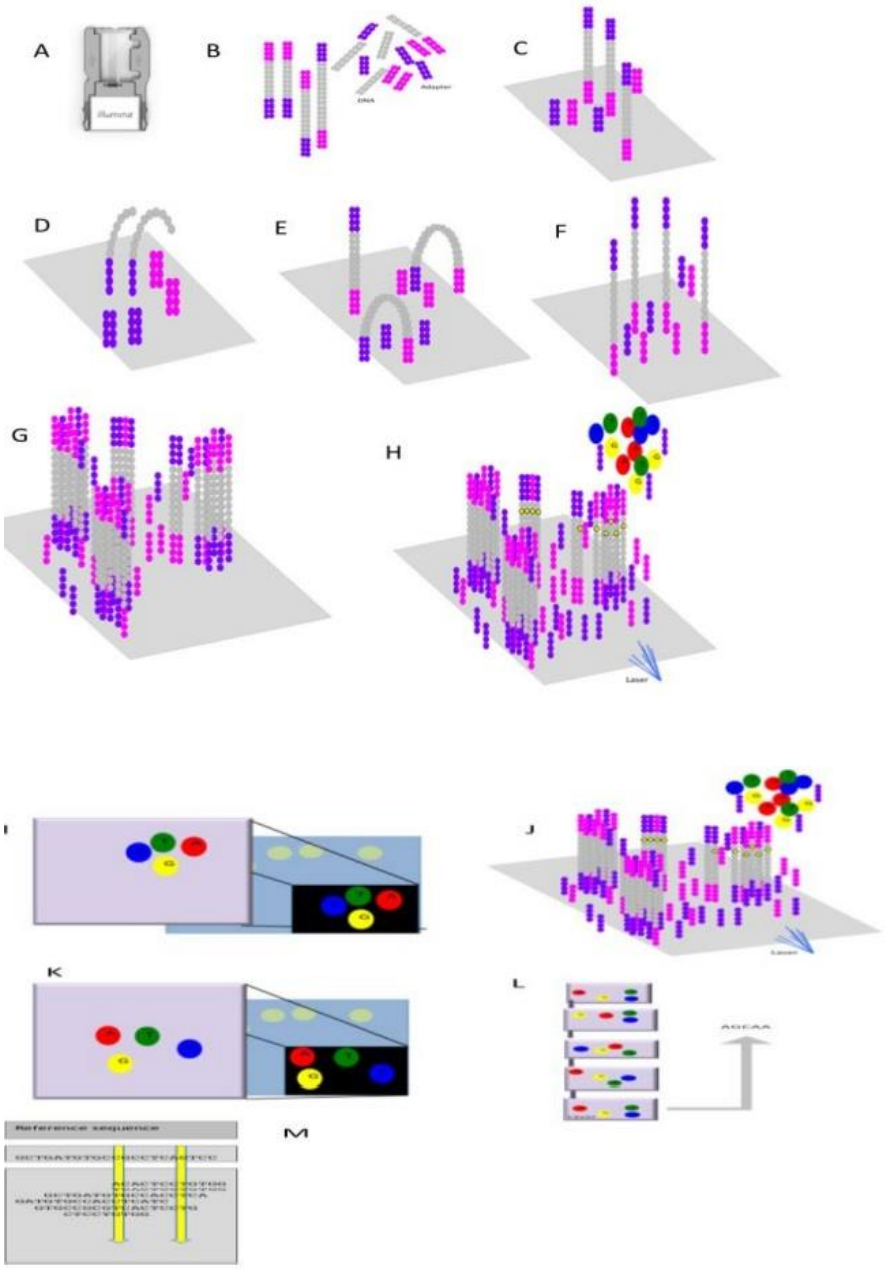


Figure 2-5: Illumina Sequencing Workflow.

A) Flow cell. B) Genomic DNA is fragmented and adaptors ligated to both ends of each fragment. C) DNA adheres to the surface. D) Bridge amplification. E) Fragments become double stranded, bridges of which form on the flow cell. F) Denaturation of the double-stranded molecules. G) Amplification: several million dense clusters of double-stranded DNA are generated. H) To determine the first base, the first sequencing cycle adds four labelled reversible terminators. I) Image of the first base following laser excitation. J) The next cycle repeats the incorporation of four labelled reversible terminators. K) Image of the second cycle following laser excitation; this image is taken in the same way, and the second base is recorded. L) Sequencing over multiple chemistry cycles, one base at a time. M) Alignment of the data. The data are then compared to a reference and differences in sequencing identified.

2.2.10 Data Analysis and Bioinformatics

The samples were multiplexed with 20 samples per run. Using the 500 cycle v2 Reagent Kit, paired-end sequencing was performed on an Illumina MiSeq sequencing platform. The NGS data were analysed using the in-house bioinformatics pipeline (Figure 2-6).

As we carried out paired-end sequencing, each sample generated two files containing both forward and reverse reads. The per-base quality scores were tested using FastQC v. 0.118 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The majority of reads had base quality scores well above the minimum quality cut-off score of Phred 30. Prior to our analysis, the low-quality (<30 Phred score) and short (<50bp) reads were removed from the dataset using the Trim galore! (v. 0.6.3) Program (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). The Fastq files then underwent reference assembly using the in-house assembler Tanoti, a blast-guided mapping program (http://bioinformatics.cvr.ac.uk/tanoti.php) based on a reference HBV genome (GenBank ID: KX357637.1). The mapped files were saved in SAM (sequence alignment/mapping) format for downstream analysis. Using samtools (v 1.10), we calculated the mapping statistics, coverage, and depth of the mapped reads for each sample. We utilized an in-house program to detect the low-frequency mutations and point mutations in the SAM files.

2.2.11 Genotyping

We generated consensus sequences from the SAM files with the SAM2CONSENSUS (https://github.com/vbsreenu/Sam2Consensus) program and submitted them as query sequences using BLASTN (a program available online at http://www.ncbi.nlm.nih.gov/BLAST/). The HBV genotypes were then identified for the samples. The cut-off used to define the success of whole-genome sequencing was >90% coverage of the full HBV genome. To ensure the consensus sequence was robust, a >100-fold mean read depth was highly recommended and samples that did not accomplish >90% in genome were classified as “fail”.

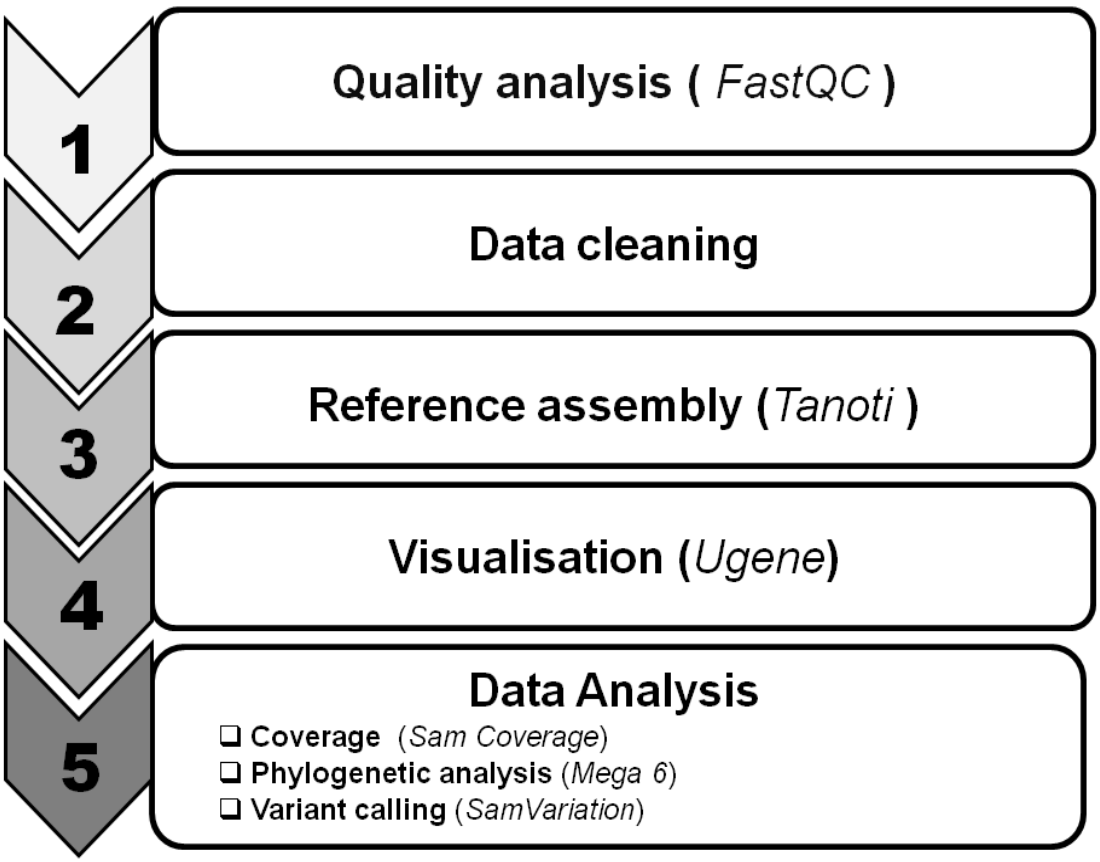


Figure 2-6: Bioinformatics pipeline for HBV data analysis.

Different programmes and scripts were employed in the pipeline; examples are given in brackets.

2.2.12 Construction of the phylogenetic trees

Phylogenetic trees were generated using exhaustive search method, maximum likelihood (ML) trees, ML trees were employed to construct the phylogenetic tree. MUSCLE was used to align the sequences while MEGA 7.0 was used to construct the phylogenetic trees The Tamura-Nei mode parameter was used l with entire nucleotide sequences. Each data set was then tested for the best-fit substitution model. Across our data sets, the best model was the Kimura two-parameter distance for all substitutions. The inferred phylogenies were then tested with 1,000 bootstrap replications.

2.2.13 Determination of Antiviral HBV Resistance and HBV Vaccine Escape Mutations

An extensive literature review of the reported HBV antiviral resistant associated mutations (RAM) using several databases. The main databases were PubMed and Embase using keywords: Lamivudine, Telbivudine, Adefovir dipivoxil, Tenofovir disoproxil, Entecavir and 'resistance'. The abbreviated names of these drugs were also used as keywords: LAM, LdT, ADV, TDF, and ETV. Resistance-associated mutations (RAMs) identified in the literature were listed in the mutation tables (Appendix 1- 5). These refer to the appropriate direct-acting agents (DAAs) and alterations in fold-resistance in comparison to wild-type variants and the reported genotype.

Similarly, a literature review was performed on HBsAg that lead to vaccine escape VEMs (Appendix 6). All of the collected data, along with the literature review on HBV antiviral drugs resistance vaccine escape mutations, were integrated into the database supporting HBV-GLUE software (http://hbv-glue.cvr.gla.ac.uk/#/hbvFastaAnalysisAhmedAlessa), a web-based resource specific to HBV developed by the MRC-University of Glasgow Centre for Virus Research in collaboration with the University of Oxford Nuffield Department of Medicine. GLUE is a data-centric bioinformatics environment for viral sequence data that focuses on evolution, variation, and the interpretation of sequences. This provides basic genotyping, drug resistance and vaccine escape analysis of the submitted FASTA sequences. To detect antiviral therapy resistance mutations, the 846 bp fragment of the reverse transcriptase (rt) domain situated between nt130 to nt1161 coding for amino acid rtM1 to rtQ344 (numbered according to GenBank reference sequence X02763) of the HBV polymerase was utilised to analyse the drug resistant mutations. The sequences were in alignment with the corresponding regions of the reference sequences and translated into amino acids (aa) sequences.

2.2.14 Molecular Cloning

2.2.14.1 Polymerase Chain Reaction (PCR)

A polymerase chain reaction (PCR) was used to amplify specific regions of DNA. As indicated in the table 2-21, different sequences needed specific primers (Table 2-15: Primer Sequences) and reaction cycles. All primers were manufactured by Sigma Aldrich. A 50 µL reaction volume was prepared on ice in 0.2 ml thin-walled PCR tubes (Axygen). PCR was used to amplify the nucleic acids. Briefly, the reaction mix contained 5 µl 10X Advantage cDNA Polymerase buffer, 1 µl Advantage cDNA Polymerase enzyme (1 U), 1 µl dNTPs, 1 µl each of 10 µM forward primer, 10 µM reverse primer, and approximately 50-100 ng template DNA or cDNA. The remaining volume was composed of nuclease-free water. The reactions were prepared on ice and incubated in a PCR thermal cycler (Veriti®, Applied Biosystems). The subsequent standard reactions were cycled as listed in (Table 2-21). Steps 2- 4 were repeated over 35 cycles. To test for contamination, a non-template (H2O) negative control reaction was run in conjunction with the samples for each PCR. Agarose gel electrophoresis (Section 2.2.14.4) was applied to analyze all the PCR products, which were temporarily stored at 4°C before the downstream application was performed.

Table 2-21: The Parameter of Polymerase chain reaction (PCR) cycles

Step	Stage	Temperature (˚C)	Duration
1	Initial denaturation	94	1 minute
2	Denaturation	94	30 seconds
3	Oligo annealing	5 °C below the lowest primer melting temperature	45 seconds
4	Extension	68	23 seconds
5	Final extension	68	3 minutes

2.2.14.2 DNA digestion

Unless specified otherwise, DNA digestion was performed using NEB restriction enzymes and buffers at 37°C for 2 h. The total reaction volume in each case was 50 µL. This included 10 units of restriction enzyme for 10 µg DNA. Unless stated otherwise, the reactions used the supplier cut smart buffer and BSA.

2.2.14.3 DNA ligation

To confirm the anticipated size and determine the concentration, the restriction enzyme-digested DNA fragments of the vector insert, and backbone were examined on 0.8 - 2 % agarose gel and quantified using a Nanodrop spectrophotometer (2.2.8.4). The products were ligated together at an insert fragment to vector backbone ratio of 7:1.

For DNA ligation reactions, a DNA Ligation kit from Thermo Scientific was employed to introduce the intermediate insert into the plasmid vector (Table 2-22). Inserts containing specific mutations were then introduced and analyzed in vitro. The plasmid pQHBV- L^-/ Puro (No 11) (Chapter 5) replicon was utilized as a vector in each case. The samples were then incubated overnight at room temperature.

Table 2-22: DNA ligation reaction recipe

Material	Volume (µL)
Linearised vector DNA	100 ng
Insert DNA	Variable
5 X Rapid Ligation Buffer	4 µL
T4 DNA Ligase, 5 U/µL	1 µL
Water, nuclease-free	Up to 20 µL
Total reaction volume	20 µL

2.2.14.4 Agarose gel electrophoresis

Gel electrophoresis on an agarose gel was performed to fractionate DNA fragments. Agarose gels were prepared in 1 X TAE buffer (40 mM Tris containing 20 mM acetic acid and 1 mM EDTA). Depending on the sizes of the fragments needing separation, the concentration consisted of 0.8-2 % agarose.

To enable the agarose to dissolve, the preparation was heated and cooled until it was warm to the touch. Gel Red® nucleic acid gel stain (Biotium) was then added to a final concentration of approximately 0.5 µg/ml. Gel Red binds to DNA to render the nucleic acids visible under UV light.

To make sure there was an equal distribution of Gel Red, the preparation was swirled and using a comb, was poured into a gel tray and allowed to solidify. The gel tray was then placed into a horizontal gel electrophoresis tank (Bio-Rad), submerged in a 1 x TAE buffer, and the comb taken out.

DNA samples were then mixed with 6x purple loading dye (New England Biolabs) at a 1:5 ratio and placed into the wells next to a 100 bp or 1 kb DNA ladder marker (New England Biolabs, Promega) to provide a reference regarding concentration and fragment size.

To separate the nucleic acid fragments, an electric current was applied to the gel. Depending on the distance required between the bands and the size of the product, gels were run at 100 V from 30 minutes to 2 h. The nucleic acid products were then visualised using an ultraviolet (UV) transilluminator (Bio-Rad). Long wave UV was utilized for gel extraction and shortwave UV for the imaging. Using a scalpel, DNA fragments were cut from the gel and purified using a Qiagen gel extraction kit (Section 2.2.6.1) in accordance with the manufacturer's instructions.

2.2.15 Cell Culture

2.2.15.1 Maintenance, Growth, and Passaging of Cells

The cells were carefully thawed from frozen cryovials stock in a 37°C water bath and added to 9 mL pre-warmed cell culture medium as described in (Table 2-13: Cell lines). They were then pelleted, resuspended in a fresh medium and added to a suitable cell culture flask.

Unless otherwise stated, different cell cultures were maintained in T75 tissue culture flasks (Nunc Fisher Scientific UK Ltd) contained in suitable media (Table 2-13: Cell lines) at 37 ˚C with 5% CO2. Passage was carried out when the cells were approximately 80-90%

confluent, following which they were washed 2-3 times with 3-4 ml versene. This removed dead cells and any residual foetal bovine serum, Incubation then took place with 2 ml trypsin-versene until the cells were separated from the flasks. An aliquot of the cells was then either counted prior to seeding for maintenance or plating for experimentation or re-suspended in 5 ml of a suitable media prior to being seeded in another flask.

2.2.15.2 Seeding of Cells

Using a Neubauer haemocytometer counting chamber, the number of cells was determined under a light microscope. All cells were seeded in 6 -, 12-, or 24-well plates. Seeding density was dependent on cell type, with typical seeding density ranging from 1.5 x 10^5 cells/well in a 24-well plate to 3 x 10^5 cells/well in 12-well plates and 5 x 10^5 cells/well in six-well plates. Cells in the 24-well plates were covered with 1 ml of appropriate media, cells in the 12-well plates with 1.5 ml media, and cells in the six-well plates with 3 ml media. Cells were incubated under the necessary conditions overnight before undergoing further manipulation.

2.2.16 Transfection

2.2.16.1 Transfection of cells with Lipofectamine 3000

Transfection of DNA into cells was performed using the Lipofectamine®3000 transfection kit (Invitrogen, ref no L3000-008) in accordance with the manufacturer's guidelines. Unless otherwise stated, Huh7 cells at a confluence of 80% were utilized for transfection. These cells were seeded on a variety of well plates. Lipofectamine 3000 reagent was then diluted in an Opti-MEM medium (Gibco, ref no 31985-047) and thoroughly mixed. By diluting the amount of plasmid into an Opti-MEM medium, a master mix of DNA was prepared, to which Lipofectamine 3000 reagent was added and mixed by inversion. The diluted lipofectamine 3000 reagent was then combined with the DNA master mix and incubated at RT for 10 to 15 minutes. Upon completion of this process, the requisite number of DNA-lipid complexes was added to each well (10 µL/well in a 96-well plate, 50 µL/well in a 24-well plate, 125 µL/well in a 12-well plate, and 250 µL/well in a 6-well plate). These were then incubated overnight. at 37°C. On the following day, the medium was changed for fresh medium. Co-transfection of a GFP expression plasmid was utilised to determine the transfection efficiency by viewing the cell population under a microscope and counting the percentage of GFP-positive cells. The transfection efficiency was generally found to be approximately 90%.

2.2.16.2 Calcium Phosphate transfection of HEK cells (Generation of Lentivirus)

To introduce plasmid DNA into HEK-293T cells, the calcium phosphate transfection method was applied. This technique forms a precipitate containing DNA and calcium phosphate. Approximately 24 h before transfection, the cells were seeded into 90 mm tissue culture dishes in 15 ml complete DMEM. Using the Sigma-Aldrich Calcium Phosphate Transfection Kit, the cells were co-transfected on the following day with three different plasmids: (i) pHIV-Luc (8 µg), (ii) pVSV-G (3 µg), and (iii) pCMV- S (8 µg) expressing lentivirus gag-pol. The plasmid DNA was mixed with dH2O and 100 µL 2.5 M CaCl2 in a sterile 1.5 ml Eppendorf at a volume of 500 µL. Next, 500 µL of 2x HEPES-Buffered Saline (HEPES) pH 7.05 was added to a second sterile 1.5 ml tube. To prepare the precipitate, the HEPES solution containing sodium phosphate was gradually mixed with the Calcium chloride (CaCl2) solution containing the DNA. This involved slowly bubbling the HEPES using an automatic pipette pump attached to a 1 ml sterile serological pipette while adding the CaCl2/DNA solution in drops using a sterile pipette tip. Prior to being distributed in drops over the cells in the culture dish using a sterile pipette tip, the precipitate was incubated for 20 min at room temperature. Gentle agitation was then applied to mix the solution. The DNA-calcium phosphate co-precipitate attaches to the cell surface and is presumably taken up by the cell through endocytosis. At 24 h post-transfection, the medium was replaced with a suitable 8 ml fresh medium. At 72 h post-transfection, the culture media containing lentivirus particles was filtered through a 0.45 µm pore-sized membrane for harvesting. It was then used to infect the target cells or stored at 4° C or cells were then lysed with LB2 on a shaker for 10 min. Using SDS-PAGE and western blotting, the target protein was continually monitored at the transduction endpoint (Section 2.2.19).

2.2.17 Drug Treatment

Huh7 cells were seeded in various sizes of the well tissue culture dishes and transfected using lipofectamine 3000 protocol. The following day, the medium was replaced with a fresh medium containing various concentrations of drugs. Drugs were purchased from MedChem Express (Adefovir dipivoxil, Tenofovir disoproxil, and Telbivudine) and (Lamivudine, and Entecavir) from Cayman Chemical.

All five drugs were provided as pure substances (purity > 98%). These chemicals were dissolved in dimethyl sulfoxide (DMSO) as manufacture instructions. Stock solutions them

were prepared as shown below and stored at -20°C in aliquots. After drug treatment for 3 days, the supernatants were collected, and cells were lysed for analysis.

Drug	Molecular weight	Stock concentration (mM)	Solvent
Adefovir dipivoxil	229.3	10	DMSO
Tenofovir	635.51	10	DMSO
Telbivudine	242.23	10	DMSO
Lamivudine	229.3	20	DMSO
Entecavir	295.3	10	DMSO

2.2.18 Immunostaining

To examine the expression of HBV proteins, transfected/infected cells on coverslips were fixed in methanol, washed with PBS, blocked for 10 min with PBS containing 2 % FCS and incubated at room temperature for 60-minute with primary antibody at the desired dilution. were washed three times with PBS before a 60-minute incubation with the secondary antibody.

The secondary antibodies used were Alexa 488 goat anti-rabbit IgG, Alexa 488 donkey anti-rabbit IgG, Alexa 647, and Alexa 488 goat anti- mouse IgG. Coverslips were washed three times with PBS followed by three washes with distilled water. Plates were allowed to dry before being examined under EVOS M5000 Imaging System.

2.2.19 Western Blot

2.2.19.1 Sample Preparation

The appropriate cell lines were seeded in well plates and infected or transfected as required. They were then incubated as necessary, a process followed by the removal of the growth media, then we directly washed the cell twice with 1X PBS. Cells were lysed in varying volumes of lysis buffer 2 (LB2) depending on the well size i.e., 100 µL per well for 24-well dish, 150 µL per well of a 12-well dish, and 300 µL per well of a six-well dish. Once lysis was complete, the lysate was subjected to centrifugation at 13,000 rpm for 10 min. The supernatant with clarified lysate was then placed in a clean 1.5 ml tube and either used directly or stored at -20°C for future use.

2.2.19.2 Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Unless stated, otherwise most of the SDS-PAGE gels were prepared in-house, and a 10% or 15% gel was created using Bio-Rad's Mini-Protean II apparatus. These comprised two elements: the lower resolving gel and the upper stacking gel. The former was prepared by combining distilled H_2O, acrylamide/bis-acrylamide (40%), 1.5 M Tris-HCl (pH 8.8), and 10% SDS per gel. Before use, polymerisation was activated by adding 10% APS and TEMED. The entirety of the previous solution was then added in various volumes, depending on the percentage required. The solution was then placed between two glass plates situated approximately 0.75 mm apart. Between the top of the plates and the top of the gel, there was a gap of approximately 2 cm. To ensure the gel was set at the required level, this was filled with H2O. Approximately 20 min later, once the gel had set, the H_2O layer was disposed of and replaced with the stacking gel. This composed differing volumes of distilled H_2O, acrylamide/bis-acrylamide (40 %), 0.5 M Tris-HCl (pH 6.8), and 10% SDS per gel. Like the resolving gel, polymerisation was activated by adding TEMED and 10% APS immediately before the transfer to the two gel plates on top of the resolving gel. A 10- or 15-well comb was then inserted to form the lanes, following which the gel polymerized for approximately 30 min. We also used commercial Mini-PROTEAN precast gels, including Mini-PROTEAN TGX™ and TGX Stain-Free™, Tris/tricine, TBE, and TBE/urea polyacrylamide gels.

The gel was then placed into an electrophoresis tank and covered with a suitable running buffer (1 x SDS-PAGE running buffer (25 mM Tris, 0.25 M Glycine (Calbiochem), 0.1% SDS). Next, 20 or 25 µL of protein lysate were mixed with 5 µL SDS-PAGE reducing buffer (composed of 200 mM Tris-HCl, pH 6.7; 0.5 % SDS; 5 % β-mercaptoethanol; 10 % glycerol, 1 µg/ml bromophenol blue). The protein samples were then boiled at 100°C for 5 min and, loaded on the gel along with a 5 µL PageRule Plus pre-stained protein ladder protein molecular weight marker (Thermo Scientific), and the electrophoresis performed at 120 V till the bromophenol blue dye front migrated off the resolving gel.

2.2.19.3 Protein Transfer

The gel was then extracted from between the two glass plates and the stacking gel removed. Two protein transfer techniques were then applied: semi-dry transfer for smaller proteins and wet transfer for large proteins > 100 kDa. Unless stated otherwise, wet transfer was implemented. This involved assembling a transfer cassette submerged in a tray flooded with

1 x wet transfer buffer (50 mM Tris, 375 mM Glycine, 20% Methanol (VWR)). From the cathode to the anode, the transfer cassette contained the gel, a sponge, an extra thick Whatman paper (Bio-Rad, a Hybond ECL nitrocellulose membrane (GE Healthcare) (previously activated in distilled H2O) or a polyvinylidene difluoride (PVDF) membrane that needs to be activated in methanol (Immobilon®-FL, Merck Millipore), a second sponge, and a second Whatman paper. Having been assembled, the transfer cassette was positioned within the electrophoresis tank (Bio-Rad) in such a way as to transfer the proteins from the gel to the membrane. An ice pack was then added, and a 1 x wet transfer buffer used to fill the tank. A current of 250 mA was then applied for 3 h. Semi-dry transfer was performed by soaking two extra-thick Whatman papers, cut to the size of the gel, in a semi dry transfer buffer (48 mM Tris, 39 mM Glycine, 0.0375 % SDS, 20 % Methanol). The first paper was placed on the blotting plate of the semi-dry blotting machine (Bio- Rad), followed by the Hybond ECL (GE Healthcare) nitrocellulose membrane, cut to the appropriate size and soaked in distilled water, the gel, and the second Whatman paper. A constant voltage of 25 V was then applied for 30 min.

2.2.19.4 Immuno-Detection

Following transfer, the blotted membranes were blocked post-transfer with ODYSSEY® blocking buffer and placed at 4°C overnight or at room temperature for 1 h on a rotating shaker.

The membranes were washed three times with PBS-T prior to being covered with a suitable volume of primary antibody, which was diluted in an appropriate buffer. The antibodies used are described in (Table 2-17: List of Primary Antibodies). All of the antibody incubations of the membranes were performed at room temperature for at least 1 h under constant agitation. All antibody dilutions were performed in western blot washing buffer and incubated in a sealed box.

Once incubation was complete, the membranes were washed and covered with fluorescent IRdye® secondary antibodies (

Table 2-18) diluted in PBS-T solution. They were then incubated at room temperature for approximately 1 h, following which the membrane was washed one more time.

The membranes were washed three times with PBS-tween for ten min followed by one final wash for ten min in either 1X PBS or distilled water. Because of the light sensitivity of IRdye® LI-COR secondary antibodies, the membranes were kept in darkness during the

washing. The membranes were then imaged and analysed using two-colour fluorescent western blotting on the Odyssey infrared imaging system (Oddysey® CLx scanner, LI-COR® Biosciences, Cambridge, UK) and image studio® software (LI-COR® Biosciences, Cambridge, UK).

As an alternative method, after washing three times with PBS-tween for 5 min, the membrane was incubated on a rotating shaker at room temperature for 1 h with an appropriate horseradish peroxidase (HRP)-conjugated secondary antibody (1:1000) diluted in a blocking solution. The membranes were then washed three times for five min with PBS-Tween, following which a signal was developed with enhanced chemiluminescence (ThermoFisher, Renfrew, UK, ref no 32106) for 1 min or with ECL plus (ThermoFisher, Renfrew, UK, ref no 32132) for 5 min and then exposed to high-performance chemiluminescence film (GE Health, ref no 28906835).

2.2.19.5 Coomassie staining of SDS-PAGE gels

The gels required for Coomassie staining analysis were fixed with Coomassie fixing solution for five min (Table 2-6). The Coomassie fixing solution was washed off with distilled water, and Coomassie Brilliant Blue Stain was added to the gel (Table 2-6). This solution was incubated for three min before being washed off once more with distilled water. The gel was then rinsed in 100% methanol and left in a Coomassie destain solution at room temperature, where it was subjected to continual shaking until the bands could be seen.

2.2.20 Measurement of HBeAg

2.2.20.1 Enzyme-linked immunosorbent assay (ELISA)

An in-house ELISA assay to detect and quantify HBeAg secreted from transfected or HBV-infected cells was developed during this project. HBeAg is an effective marker of certain events in the life cycle of a virus as described earlier. Immulon 96 well flat-bottom ELISA plates (Dynatech Laboratories) were coated with 95 µL/well Anti-HBeAg Mouse Monoclonal antibody (10-H10M and 10-H10N; Fitzgerald Industries International, USA) in PBS and incubated at 4°C overnight. The plates were washed three times with PBS containing 0.1% (v/v) Tween (Sigma) (PBS-T) and then wells blocked with 100 µL/well of PBS containing 5% milk and 1% (v/v) Tween for 1 h.

The plate was washed three times with PBS-T, following which anti-HBeAg mouse monoclonal conjugated to horseradish peroxidase (HRP) (61-H10K; Fitzgerald Industries

Fitzgerald Industries International, USA) in PBS-T was added and incubated for 2 h (dilution 1: 8000 in PBS-T).

The plate was washed three times with PBS-T and 100 µl were added to each well of the substrate 3,3′,5,5′-tetramethylbenzidine (TMB; Sigma-Aldrich) which develops a soluble blue reaction product (Sigma-Aldrich). The reaction was terminated by adding 50 µl of 0.5 M H_2SO_4 after 10 to 15 min and optical density was assessed at 650 nm with a reference wavelength of 490 nm.

2.2.21 Statistical Analysis

All Statistical analyses were performed using GraphPad Prism Version 7.02 (GraphPad Software, Inc., California, USA) and Statistical Package for the Social Sciences (SPSS v. 23, IBM, New York) for the most of this project. Where other programs were used, this has been stated in the individual chapters.

For SPSS, the statistical analysis included quantitative descriptive analysis and summary statistics (means, median, percentages, standard deviations, etc.); $p < 0.05$ was considered statistically significant.

Chapter 3. Development of an HBV whole-genome sequencing protocol using Next-Generation Sequencing (NGS)

3.1 Introduction

Next-generation sequencing (NGS) is a robust tool for clinical virology with different protocols available. The objective of this research was to develop a new deep sequencing protocol for whole hepatitis B virus genomes (HBV) that can be employed to detect mutations in drug resistance and vaccine escape in HBV (McNaughton et al., 2019). The NGS protocols require minimal viral DNA concentrations for optimal results and sufficient coverage to detect and analyze virus variants.

DNA enrichment method is necessary for any NGS sample (Flavia et al., 2020); target enrichment improves the capture of specific reads and reduces non-specific genome data, mainly from host DNA (Flavia et al., 2020). We, therefore, describe a sequencing strategy using NimbleGen® target enrichment probes (Roche, Basel, Switzerland) for whole-genome sequencing of HBV. This strategy has been successfully applied to sequence whole genomes of HCV and HEV (Bowyer and Sim, 2000; Davis et al., 2021). In this study, HBV genomes were enriched directly from DNA extracts (human plasma) using a panel of custom-designed 120-mer RNA baits (also known as probes) complementing all Hepatitis B full genome sequences retrieved from Genebank®. Genomic data were then used to describe HBV diversity in Saudi Arabia.

3.2 Results

3.2.1 Sample Preparation for Next Generation Sequencing

The methods employed are described in detail in Chapter 2. In brief, the frozen extract DNA was thawed in a water bath at 37°C and 10 µL was used as the template in a reverse transcription reaction. Second strand synthesis was then conducted, and the resulting DNA was clarified using AmpureXP®.

3.2.2 DNA library quality

Following AmpureXP® magnetic bead purification, DNA fragment quality was examined and visualized using a 2200 Tapestation (Agilent Technologies), as shown in Figure 3-1.

The quality and fragment size of each library was measured using a 2200 TapeStation platform (Agilent Technologies), as described in Section 2.2.7.2, following AmpureXP® bead purification. Fragment lengths were dependent on the concentration of input DNA used in the KAPA LTP reaction. We initially used 1 ng DNA, as recommended by the manufacturer, but noted a shift towards longer fragment length as DNA concentration increased, reaching saturation at 10 ng (Figure 3-1).

3.2.3 Sequencing quality scores

NGS using the Illumina MiSeq® platform is an extremely sensitive method but one that is subject to sequencing error. To assess the quality of NGS, a metric method called Phred score was utilized to report base-calling accuracy. The quality of Fastq sequences was examined using FastQC® (Babrahman Bioinformatics), which measured the Phred score for each raw Fastq file (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Before our analysis, low-quality reads were trimmed and removed to increase the average quality of the Phred score to a minimum of Q30 (1 base call error in 1000 bases or 99.9 accuracy) (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).

The large amount of data produced by Illumina sequencing can make downstream analyses cumbersome; therefore, the quality score was used to remove all low-quality reads and limit the downstream sequence analysis to high-quality score reads. High-quality sequence data are essential for accurate alignments and other downstream analyses. Because we performed paired-end sequencing, each sample generated two files containing forward and reverse reads. The per-base quality scores were tested using FastQC. The majority of reads had base quality scores well above the minimum quality cut-off score of Phred 30 (Figure 3-2).

3.2.4 Sequence alignments

Fastq sequences were assembled and aligned to several HBV reference sequences using an in-house mapping program (Tanoti) developed in-house by Dr. Sreenu Vattipally, as described in Section 2.2.10. The trimmed Fastq reads were aligned to several HBV reference sequences using the mapping program Tanoti (https://github.com/vbsreenu/Tanoti) with default parameters. In our comparison, we found that Tanoti is more sensitive than other alignment programs such as BWA and Bowtie.

An example of the coverage obtained following mapping with Tanoti is presented in Figure 3-3. An HBV consensus sequence was calculated (using SAM2CONSENSUS) from SAM files along with its coverage.

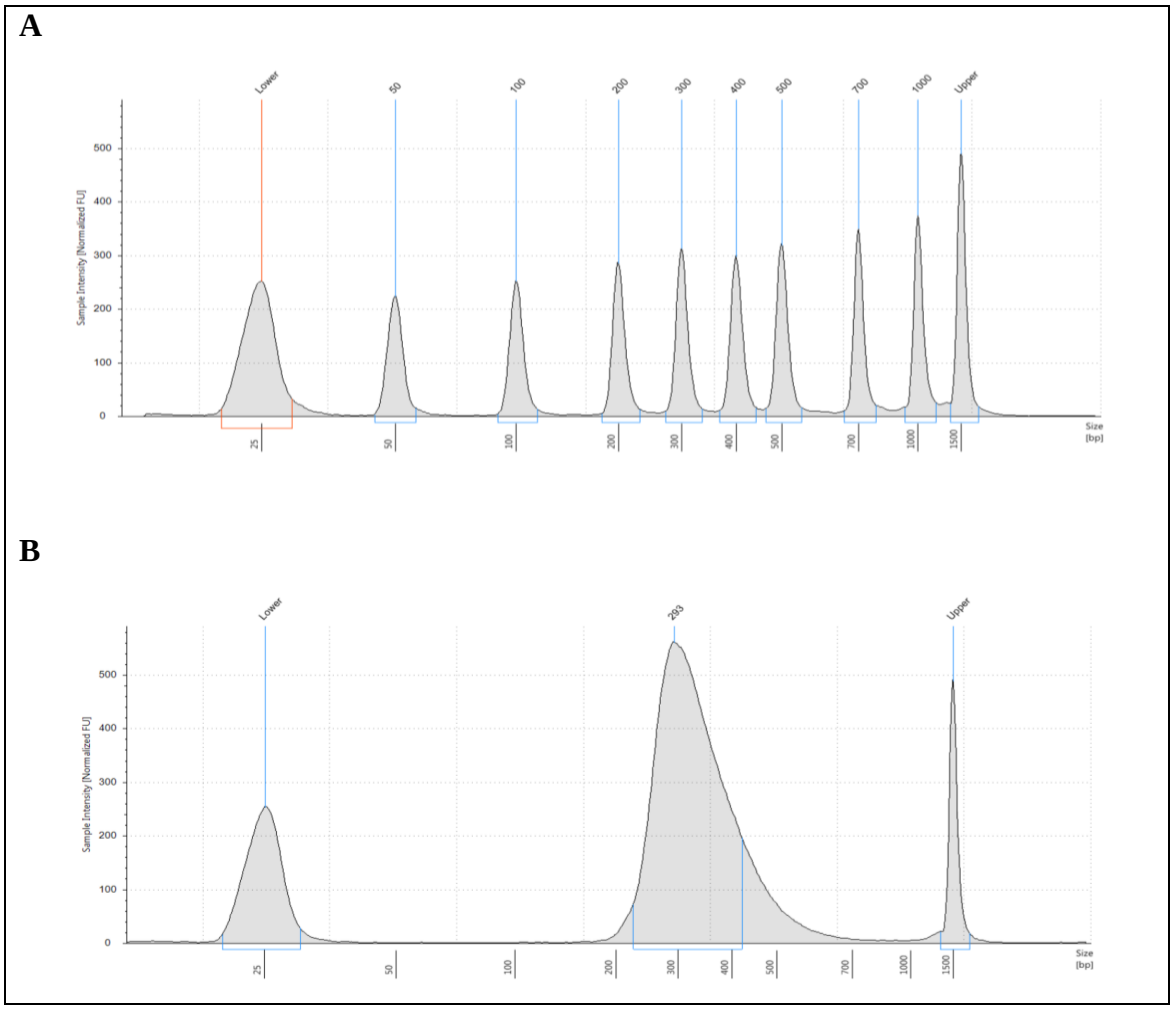


Figure 3-1: DNA fragment size evaluated using a 4200 Tapestation® (Agilent).

Library size was measured using the D1000 DNA screen tape (TapeStation platform). A) DNA ladder. B) DNA input.

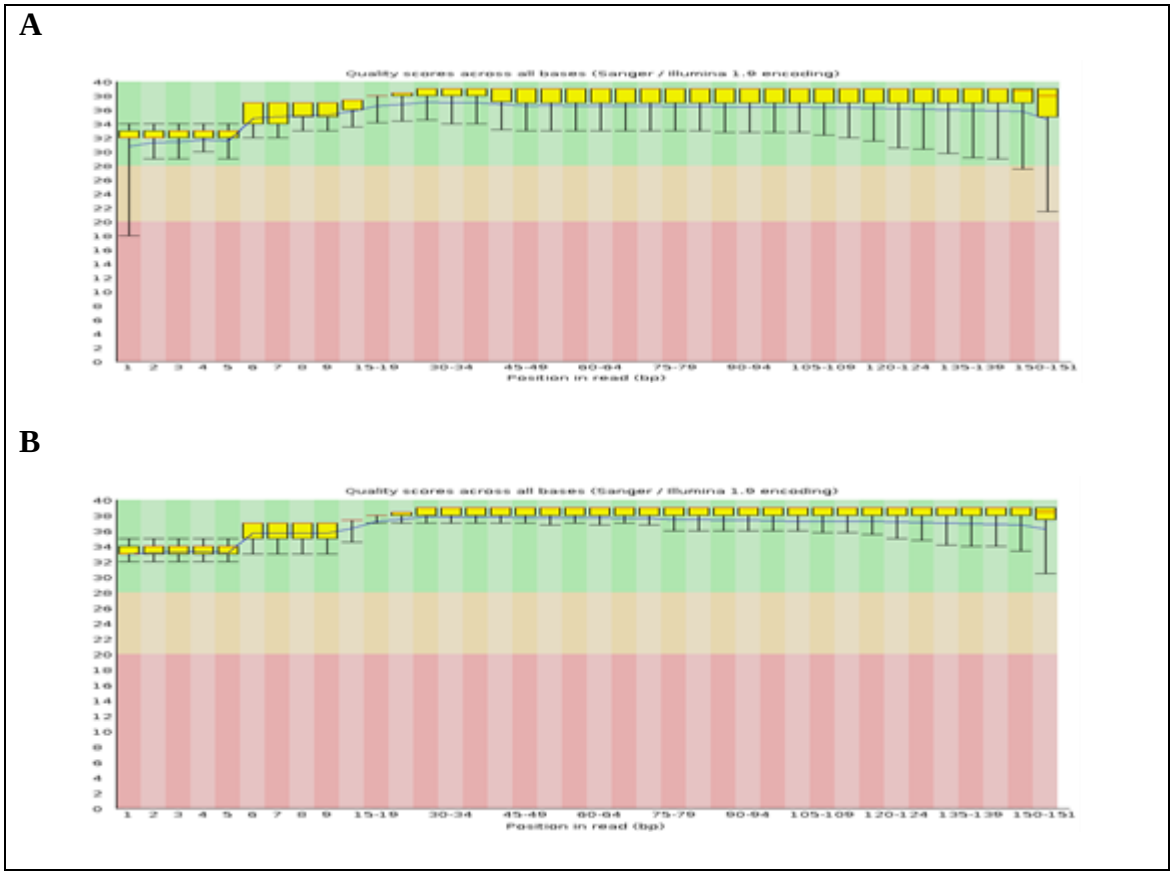


Figure 3-2: Example Phred score data obtained from a Fastq file.

The quality scores of sequenced reads visualised with FastQC before and after quality trimming A) Raw reads (first read) generated by MiSeq run. B) Cleaned reads (first read) cleaned using the in-house script weeCleaner. The yellow boxes plot the interquartile range around the median with black whiskers extending to the outer range limits.

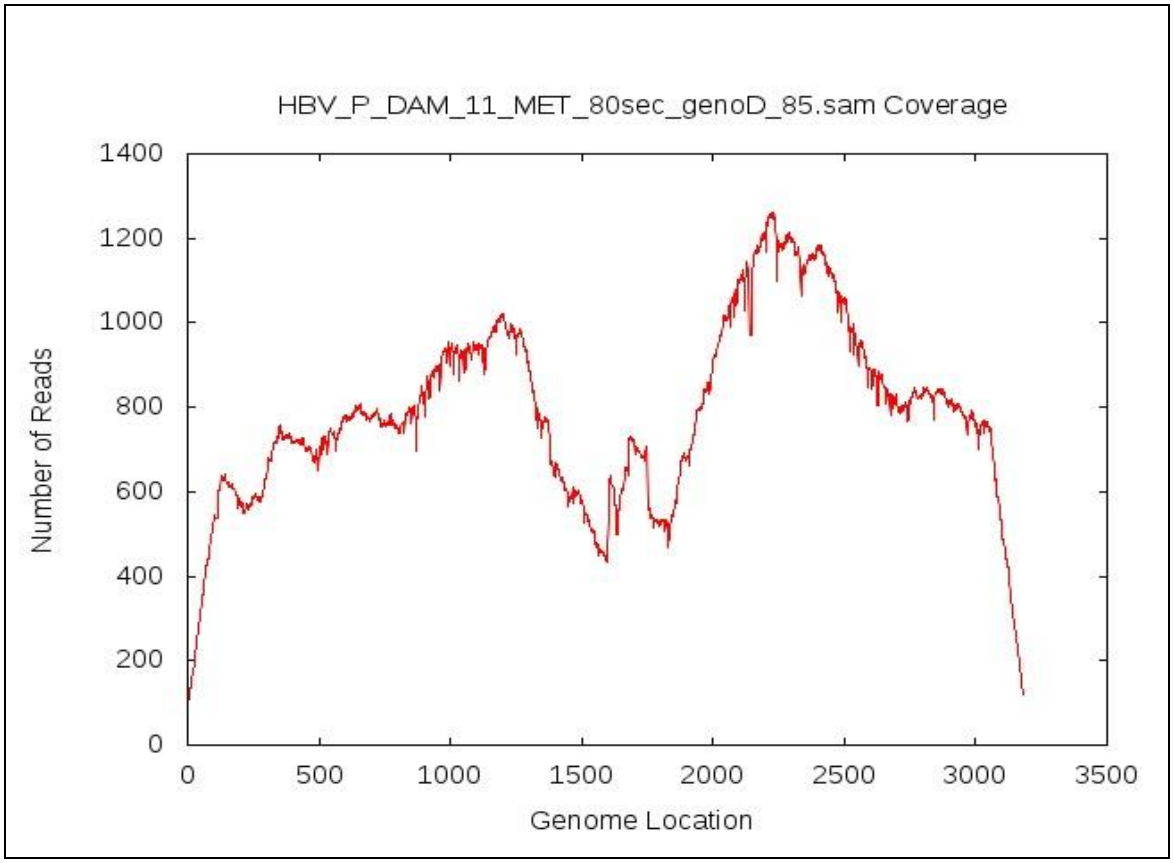


Figure 3-3: Read coverage across the whole HBV genome using Tanoti.

Sam coverage carried to show the coverage across the whole HBV genome from a single sample analysed using Tanoti.

3.2.5 Metagenomic sequencing

We began sequencing whole HBV genomes using a metagenomic sequencing approach applied on Illumina® MiSeq platform. The central aim was to provide a detailed method for the construction of unbiased metagenomic libraries to design a specific probe for HBV, as described in Section 2.2.5.2.

A total of seven samples with high viral loads (>5000000 IU/mL) were selected for sequencing. We generated 14 SAM files from these samples as part of protocol optimisation (Table 3.1). Library preparation was performed using the KAPA LTP Library Preparation Kit for Illumina Platforms (KAPA Biosystems®). For optimisation purposes, several parameters were tested when using physical fragmentation. Samples were processed in several runs of the Illumina MiSeq®. The number of total reads generated from each sample after quality trimming ranged from 206,898 - 2,235,798 (and after quality trimming, 1365100- 4673166, mean = 2877682). Genotype D was identified in all samples using an in-house tool, as described in Section 2.2.11.

	Sample	Fragmentation Condition (Metagenomic)	Viral load (IU/mL)	Total Read	Genotype	Mapped Read	Convergent	Coverage %	Average depth reads/site	GC%
1	HBV-DMM-10	Treatment time (45 sec) Target BP (500)	>110,000,000	1365100	D	51623 (3.78%)	3182	100	2303	48.59
2	HBV-DMM-10	Treatment time (80 sec) Target BP (500)	>110,000,000	1477588	D	62039 (4.20%)	3182	100	2793	48.59
3	HBV-DMM-10	Treatment time (160 sec) Target BP (200)	>110,000,000	3755300	D	66683 (1.78%)	3182	100	3329	47.32
4	HBV-DMM-11	Treatment time (45 sec) Target BP (500)	>110,000,000	2208900	D	14525 (0.66%)	3182	100	660	48.46
5	HBV-DMM-11	Treatment time (80 sec) Target BP (500)	>110,000,000	2143152	D	16958 (0.79%)	3182	100	773	48.46
6	HBV-DMM-11	Treatment time (160 sec) Target BP (200)	>110,000,000	3318658	D	111466 (3.36%)	3182	100	5215	47.7
7	HBV-DMM-14	Treatment time (80 sec) Target BP (500)	5,783,259	2999700	D	1512 (0.05%)	3182	100	70	48.01
8	HBV-DMM-14	Treatment time (160 sec) Target BP (200)	5,783,259	4178098	D	437 (0.01%)	3182	100	20	48.21
9	HBV-DMM-20	Treatment time (80 sec) Target BP (500)	57,267,116	3027902	D	284 (0.01%)	3182	100	13	47.1
10	HBV-DMM-20	Treatment time (160 sec) Target BP (200)	57,267,116	3142972	D	2003 (0.06%)	3182	100	93	48.63
11	HBV-DMM-26	Treatment time (80 sec) Target BP (500)	32,670,435	4673166	D	97321 (2.08%)	3182	100	4472	46.8
12	HBV-DMM-26	Treatment time (160 sec) Target BP (200)	32,670,435	3978126	D	105620 (2.66%)	3182	100	4948	47.28
13	HBV-DMM-29	Treatment time (45 sec) Target BP (500)	>110,000,000	2005170	D	23574 (1.18%)	3182	100	1061	48.93
14	HBV-DMM-41	Treatment time (45 sec) Target BP (500)	>110,000,000	2013714	D	108977 (5.41%)	3182	100	4911	48.71

Table 3-1: Overview of all results generated using the metagenomics approach. Summary of the number of reads in the two files generated for each sample, along with the calculated per-base sequence coverage using the metagenomics approach. Nt represents the number of nucleotides in each sequence. GC% is the GC content in each sequence. All samples had high viral loads. Sample HBV-DMM-10 and sample HBV-DMM-11 were sequenced three times using various fragmentation conditions (KAPA LTP Library Preparation Kit). Three samples (HBV-DMM-14, 20, and 26) were sequenced using two different treatment times with two target base pairs. Samples HBV-DMM-29 and HBV-DMM-41 were sequenced with one treatment time (45 sec) and one target base pair (500). All samples indicated 100% genome coverage with high viral read**.**

3.2.6 Nimbelgen target enrichment

Following the same steps employed in the metagenomic sequencing approach, 22 samples were sequenced using designed probes as described in Section 2.2.5.2.

The probes exhibited a high percentage of mapped reads and average depth reads per site only in samples with high viral loads (> 100,000) (Table 3-2). These samples demonstrated approximately 100% genome coverage with genotype D observed in each sample. Only one sample with viral load (< 100,000) exhibited 96.35% genome coverage with average depth reads/site (212). This was considered a successful sample based on the defined cut-off value for whole-genome sequencing described in 2.2.11.

In addition, two samples were sequenced using the metagenomic approach and NimbleGen target enrichment to observe the effect of enrichment in the preparation stage, both samples had a high viral load and exhibited a substantial effect on mapped reads and average depth reads/site when enriched with NimbleGen, compared with non-enriched library. Total reads indicated a reliable effect on one sample only (Table 3-3). Full genome coverage was observed in these two samples (Figure 3-4) and genotype D was most dominant.

Regrettably, the remaining samples either did not meet the criteria for a successful run or did not provide any valuable data in terms of genome coverage. They did provide sufficient nucleotide coverage but the average depth for most was extremely low. Furthermore, there was an extremely low percentage of mapped reads, mainly for low-viral load samples.

	Sample	Viral load (IU/mL)	Total Read	Genotype	Mapped Read	Converge nt	Coverage %	Average depth reads/site	GC%
1	HBV-DMM-20*	57,267,116	502832	D	339551(67.53%)	3182	100	15934	49.14
2	HBV-DMM- 39	26,011,855	12464510	D	672692(5.40%)	3182	100	31405	49.34
3	HBV-DMM-14*	5,783,259	3394436	D	116200(3.42%)	3182	100	5451	48.35
4	HBV-DMM- 18	2,178,551	36147438	D	232336(0.64%)	3182	100	10860	49.35
5	HBV-DMM- 30	649,607	7864084	D	341084(4.34%)	3182	100	15980	49.57
6	HBV-DMM- 21	303,959	9299014	D	158011(1.70%)	3182	100	7388	51.25
7	HBV-DMM- 47	114,378	25821508	D	63326(0.25%)	3124	98.18	3031	48.56
8	HBV-DMM- 27	96,442	16988168	-	2(0.00%)	196	6.16	1	51
9	HBV-DMM- 09	56,651	12156632	D	4378(0.04%)	3066	96.35	212	50.23
10	HBV-DMM- 06	41,641	1739228	-	6138(0.35%)	196	6.16	4696	50.97
11	HBV-DMM- 16	37,785	14637474	-	325(0.00%)	150	4.71	324	47.05
12	HBV-DMM- 02	30,040	5522534	-	4(0.00%)	301	9.46	2	50.5
13	HBV-DMM- 01	29,037	4485828	-	7(0.00%)	800	25.14	1	44.82
14	HBV-DMM- 07	28,305	141382858	-	580(0.00%)	1833	57.61	47	47.48
15	HBV-DMM- 37	27,266	-	-	-	-	-	-	-
16	HBV-DMM- 28	22,662	2219350	D	254(0.01%)	1951	61.31	19	48.74
17	HBV-DMM- 40	18,552	-	-	-	-	-	-	-
18	HBV-DMM- 12	17,489	-	-	-	-	-	-	-
19	HBV-DMM- 19	15,295	4485828	-	12(0.00%)	1038	32.62	1	46.77
20	HBV-DMM- 36	7,501	-	-	-	-	-	-	-
21	HBV-DMM-44	6,417	7362718	-	67(0.00%)	513	16.12	19	48.49
22	HBV-DMM-34	6,386	3183136	-	297(0.01%)	1892	59.46	23	49.99

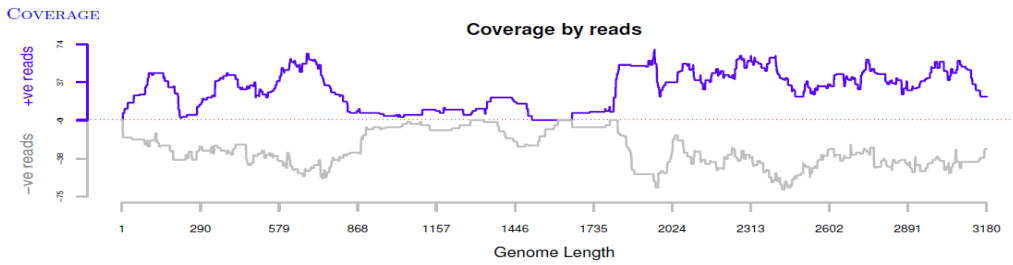
Table 3-2: Overview of the results generated using the target enrichment (enriched with NimbleGen) approach. Summary of the number of reads in the two files generated for each sample along with the calculated per-base sequence coverage using the target enrichment approach (NimbleGen) along with a KAPA LTP Library Preparation Kit. Nt represents the number of nucleotides in each sequence. GC% is the GC content in each sequence. * Repeated sample from the metagenomics approach. Samples (1-6) with high viral loads (> 100,000) exhibited nearly 100% genome coverage with good average depth reads/site and total viral reads. Only one sample (HBV-DMM-09) with a viral load < 100,000 exhibited 96.35% genome coverage with average depth reads/site (212). This was considered a successful sample based on our cut-off value for whole genome sequencing. The remaining samples either did not meet the criteria for a successful run or did not provide any valuable data.

	Sample	Method	Viral load (IU/mL)	Total Read	Genotype	Mapped Read	Convergent	Coverage %	Average depth reads/site	GC%
1	HBV-DMM- 14	Metagenomic	5,783,259	2,999,700	D	1512 (0.05%)	3182	100	70	48.01
		Enriched with NimbleGen		3,394,436	D	672692 (5.40%)	3182	100	31405	49.34
2	HBV-DMM- 20	Metagenomic	57,267,116	3,142,972	D	2003 (0.06%)	3182	100	93	48.63
		Enriched with NimbleGen		502,832	D	339551 (67.53%)	3182	100	15934	49.14

Table 3-3: Target enrichment versus metagenomic (unenriched) sequencing.

Two samples with high viral loads were tested with a metagenomics approach using a KAPA LTP Library Preparation Kit and enriched with NimbleGen to assess the effect of probes. Nt represents the number of nucleotides in each sequence. Both samples exhibited a massive effect in terms of mapped reads and average depth reads/site when enriched compared with when unenriched. Total reads indicated a reliable effect in one sample that was enriched (HBV-DMM-14) but not in the other. Full genome coverage was observed with genotype D in all runs.

Metagenomic approach



Target Enrichment approach (Nimblegen)

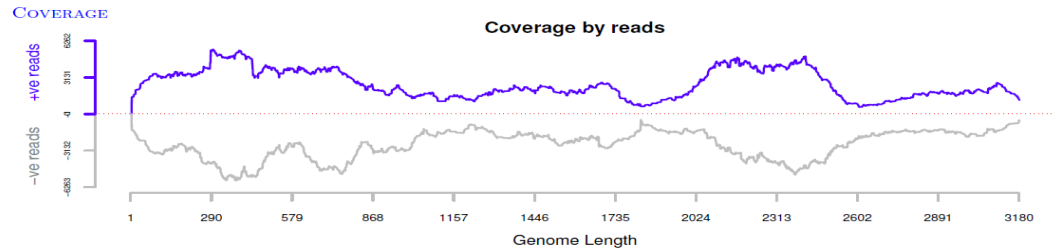


Figure 3-4: An illustrative example (HBV-DMM-14) to show the compare Genome coverage in target enrichment versus metagenomic sequencing.

Genome coverage using metagenomic and target enrichment NimbleGen approach for the same sample. 100% coverage was observed using both methods.

3.2.7 DNA Enrichment using host depletion utilising the NEBNext Microbiome

Because a range of unreliable results were generated using the target enrichment approach, we decided to optimise our protocol and included host depletion using the NEBNext Microbiome kit, as explained earlier in Section 2.2.5.1.

Additionally, we included the Nextera XT® DNA Sample Preparation kit described in Section 2.2.4.2) to evaluate the outcome and compare it with that using the KAPA LTP Library Preparation Kit.

As an initial stage, we tested the NEBNext Microbiome by sequencing two samples that had been tested previously in a metagenomic run using the KAPA Biosystems kit with and without the host depletion kit. The results indicated a remarkable change in mapped reads in both samples using Nextera XT® and enriched with the NEBNext Microbiome DNA kit. This gave us the opportunity to try the aforementioned kit in addition to target enrichment (NimbleGen) (Table 3-4). All samples exhibited 100% genome coverage using the three methods.

Subsequently, we sequenced 30 samples using our probe together with the Microbiome Host Depletion kit from the NGS protocol. Due to time and cost considerations, we used only the KAPA LTP Library Preparation Kit with Covaris for the fragmentation condition. This is because KAPA LTP library preparation is less time-consuming and cheaper than Nextera XT® library preparation. Moreover, we observed higher viral reads, mapped reads, and average depth reads/site using KAPA LTP compared with Nextera XT®, as shown in Table 3-5. In addition, sample HBV-DMM-27 was included as a control that showed 0% mapped reads in the earlier run with less than 6% coverage when enriched only with NimbleGen and 100% genome coverage when enriched with NimbleGen in conjunction with NEBNext Microbiome DNA.

To conclude, we sequenced 12 libraries that included eight samples with low viral loads (mostly below $1x10^4$ IU/mL) and four samples where the full method had not previously been applied (tested with capture but not host depletion) (Table 3-6).

A summary of the substantial enhancement in genome coverage and mapped reads when the NEBNext Microbiome kit was included is presented in Figure 3-6 and Figure 3-7. As stated previously, we were able to include one sample only in both runs

	Sample	Method	Viral load (IU/mL)	Total Read	Genotype	Mapped Read	Convergent	Coverage %	Average depth reads/site	GC%
1	HBV-DMM- 29	Metagenomic (KAPA Biosystems)	>110,000,000	2,005,170	D	23574 (1.18%)	3182	100	1061	48.93
		Metagenomic (Nextera XT®)		377,860	D	2415 (0.64%)	3182	100	102	48.46%
		Metagenomics (Nextera XT®) + NEBNext Microbiome DNA		426,020	D	5664 (1.33%)	3182	100	246	48.93%
2	HBV-DMM- 41	Metagenomic (KAPA Biosystems)	>110,000,000	2,013,714	D	108977 (5.41%)	3182	100	4911	48.71
		Metagenomic (Nextera XT®)		343,324	D	30120 (8.77%)	3182	100	1248	48.33%
		Metagenomic (Nextera XT®) + NEBNext Microbiome DNA		335,080	D	37581 (11.22%)	3182	100	1622	48.71%

Table 3-4: Metagenomic (KAPA Biosystems) versus metagenomic (Nextera XT®) versus metagenomic (Nextera XT®) enriched with NEBNext Microbiome DNA.

Two samples with a viral load >110,000,000 were tested with three different methods to optimise the NGS protocol. High viral reads and average depth reads/site were observed using the KAPA Biosystems kit in both samples (HBV-DMM-29 and HBV-DMM-41), compared with the other two methods. Mapped reads were higher in both samples using Nextera XT® and enriched with the NEBNext Microbiome DNA kit. Samples exhibited 100% genome coverage using the three methods.

	Sample	Viral load (IU/mL)	Total Read	Genotype	Mapped Read	Converge nt	Coverage %	Average depth reads/site	GC%
1	HBV-RYD-14	>110,000,000	514912	D	260(0.05%)	3013	94.69%	11	48.69%
2	HBV-RYD-15	>110,000,000	178678	A	2347 (1.31%)	3212	100.00%	106	48.41%
3	HBV-RYD-05	>110,000,000	2858246	D	2272500 (79.51%)	3182	100.00%	104588	49.25%
4	HBV-RYD-06	>110,000,000	1641752	D	3211 (0.20%)	3001	94.31%	156	48.29%
5	HBV-RYD-48	62,296,440	215022	D	936 (0.44%)	3167	95.47%	43	48.35%
6	HBV-RYD-13	62,034,425	504706	E	244365 (48.42%)	3182	100.00%	146	48.51%
7	HBV-RYD-49	43,213,615	7212738	D	6777302 (93.96%)	3182	100.00%	312469	49.15%
8	HBV-RYD-11	42,226,024	930314	D	1332 (0.14%)	3182	100.00%	60	48.77%
9	HBV-RYD-50	34,654,734	120220	D	916 (0.76%)	3182	100.00%	41	48.69%
10	HBV-RYD-03	29,003,877	7669620	D	7453288 (97.18%)	3182	100.00%	342123	48.25%
11	HBV-RYD-02	24,662,163	1457638	D	1239412 (85.03%)	3182	100.00%	57254	49.00%
12	HBV-DMM-27*	16,988,168	3235628	D	13191 (0.41%)	3182	100.00%	612	48.25%
13	HBV-RYD-08	10,026,852	822378	D	688782 (83.75%)	3212	100.00%	30819	48.55%
14	HBV-RYD-23	5,874,169	237152	D	95 (0.04%)	2641	83.00%	4	48.59%
15	HBV-RYD-19	2,789,978	1708796	E	815774 (47.74%)	3099	97.39%	37753	48.84%
16	HBV-RYD-24	1,883,203	1648852	D	29233 (1.77%)	2957	92.93%	1357	45.19%
17	HBV-RYD-29	1,155,656	1115912	D	278044 (24.92%)	3106	97.61%	146	47.85%
18	HBV-RYD-51	1,043,095	1711294	D	159296 (9.31%)	3182	100.00%	7429	47.88%
19	HBV-RYD-26	581,665	1628312	D	100654 (6.18%)	3129	98.33%	4741	48.54%
20	HBV-RYD-42	562,038	2626016	D	384129 (14.63%)	2802	100.00%	17849	48.43%
21	HBV-RYD-53	182,070	406034	E	1425 (0.35%)	2069	65.02%	101	46.88%
22	HBV-RYD-10	161,674	3374730	D	1792447 (53.11%)	3182	100.00%	82075	48.55%
23	HBV-RYD-34	158,477	306808	D	44347 (14.45%)	3182	100.00%	2021	48.68%
24	HBV-RYD-01	157,946	7402036	E	6252927 (84.48%)	3182	100.00%	282475	48.77%
25	HBV-RYD-41	142,777	789098	D	556493 (70.52%)	3182	100.00%	25718	48.40%
26	HBV-RYD-36	141,022	427004	D	121985 (28.57%)	3173	99.72%	5651	48.51%
27	HBV-RYD-07	137,855	3558086	D	499650 (14.04%)	3182	100.00%	23152	48.74%
28	HBV-RYD-20	111,328	1468512	D	69712 (4.75%)	3070	96.48%	3220	48.90%
29	HBV-RYD-27	109,082	942038	D	3949 (0.42%)	1348	42.36%	184	48.89%
30	HBV-RYD-25	94,536	3456208	D	5389 (0.16%)	3164	98.41%	244	48.65%

Table 3-5: Overview of results generated using enrichment methods with NEBNext Microbiome and NimbleGen. * Control sample was sequenced and earlier enriched with NimbleGen indicating 0% mapped read and approximately 6% genome coverage.

	Sample	Viral load (IU/mL)	Total Read	Genotype	Mapped Read	Converge nt	Coverage %	Average depth reads/site	GC%
1	HBV-DMM- 18*	2,180,000	9758406	D	8237593 (84.42%)	3182	100.00%	367953	49.18%
2	HBV-DMM- 21*	304,000	10717122	D	1616696(15.09%)	3212	100.00%	70434	49.03%
3	HBV-DMM- 47*	114,000	18829392	D	13569877(72.07%)	3191	99.35%	582342	49.14%
4	HBV- RYD- 37	79,800	5209690	D	3383072(64.94%)	3182	99.97%	143420	48.78%
5	HBV-DMM- 09*	56,700	8059626	C	485540(6.02%)	3215	100.00%	21442	48.68%
6	HBV-RYD- 43	56,100	18033444	E	13170256(73.03%)	3182	99.97%	578774	48.90%
7	HBV-RYD- 38	36,700	7942610	D	2648716(33.35%)	3211	99.97%	108844	49.15%
8	HBV-RYD- 16	24,500	14543538	E	9892620(68.02%)	3212	99.97%	432533	48.92%
9	HBV-RYD- 12	11,000	21185952	D	173704(0.82%)	3212	99.97%	7548	49.42%
10	HBV-DMM- 15	6,270	4339900	E	16005(0.37%)	3035	95.38%	704	48.47%
11	HBV-DMM- 48	6,190	5900428	D	151671(2.57%)	3182	95.47%	6703	48.58%
12	HBV-DMM- 46	5,590	3577944	E	6053(0.17%)	3212	95.11%	267	48.22%

Table 3-6: Overview of results generated using enrichment methods with NEBNext Microbiome and NimbleGen for samples with low viral loads.

Summary of the number of reads in the two files generated for each sample along with the calculated per-base sequence coverage using the probe together with the Microbiome Host Depletion kit in the NGS protocol. * Repeated sample was sequenced earlier only enriched with NimbleGen.

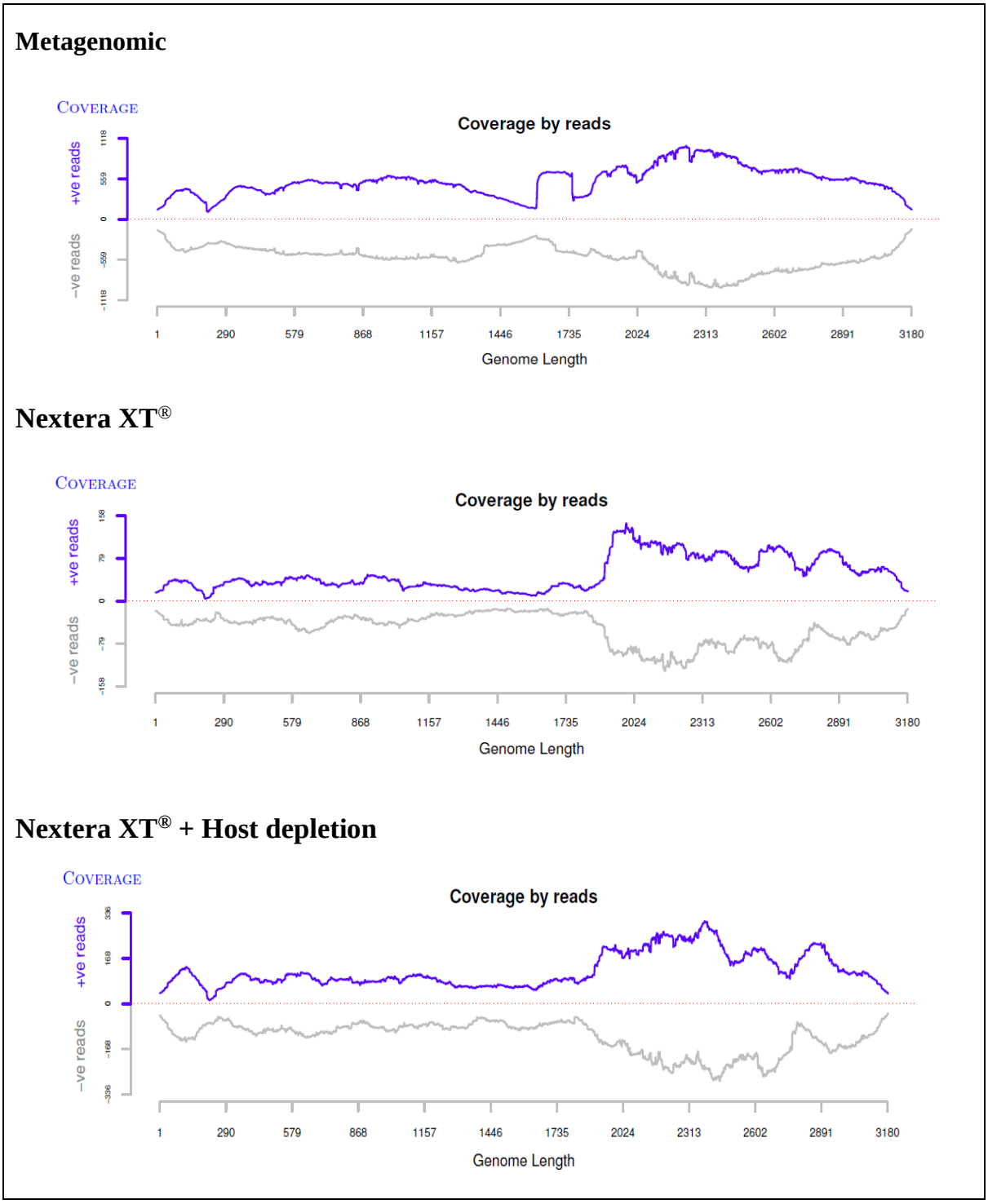


Figure 3-5: An illustrative example (HBV-DMM-29) Genome coverage in metagenomic sequencing (KAPA Biosystems and Nextera XT®) versus NEBNext Microbiome DNA metagenomic (enriched).

Genome coverage using metagenomic and NEBNext Microbiome DNA (enriched) for the same samples. 100 % coverage was observed in the using the three methods.

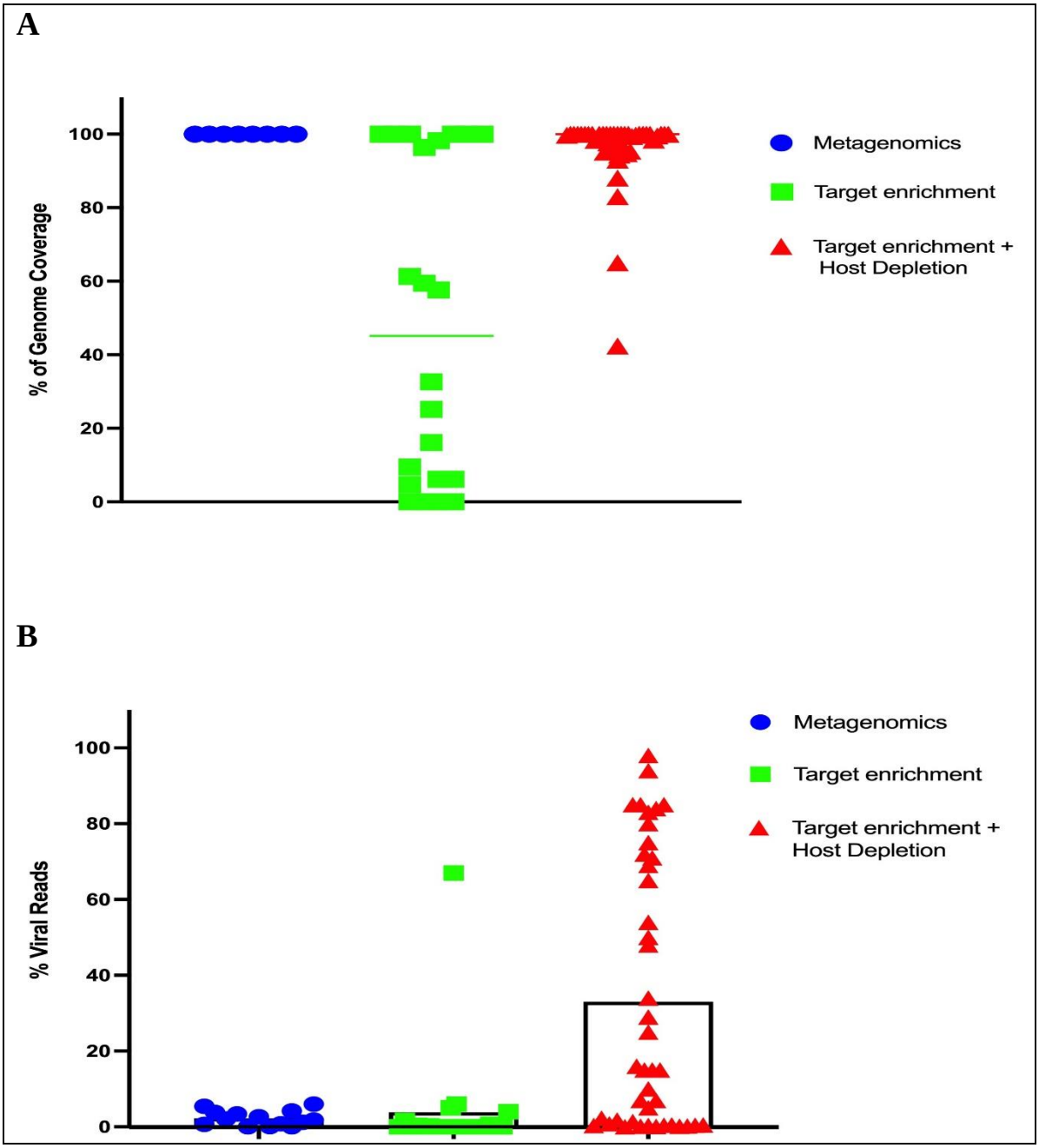


Figure 3-6: Genome coverage and mapped read percentages in metagenomics, target enrichment with NimbleGen, and enrichment with NimbleGen/NEBNext Microbiome.

Improvement when employing the target-enrichment method, compared with the metagenomics approach. A) Percentage of genome coverage for the three sequencing approaches. B) Percentage of mapped reads for the three sequencing approaches. Bars represent the mean.

3.2.8 Statistical analysis of the four sequencing runs: metagenomics, enrichment with NimbleGen, enrichment with NimbleGen/ NEBNext Microbiome (high viral loads), and enrichment with NimbleGen/ NEBNext Microbiome (low viral loads)

In this study, we generated 78 SAM files out of 64 HBV samples separated into four sequencing runs, as described earlier. As stated in Chapter Two, we defined the success of whole genome sequencing as a cut-off point of >90% coverage of the full HBV genome. To ensure the consensus sequence was robust, a >100-fold mean read depth was strongly recommended. Samples or files that did not accomplish >90% genome coverage were classified as "fail".

We conducted analysis of variance (ANOVA) for the 4 sequencing runs versus all variables and cross-tabulation of sequencing runs versus genotypes and coverage (full coverage/failure). We covered the main parameter as follows:

3.2.8.1 Sequencing method versus viral load (IU/mL) and total read

As presented in Table 3-7, great variation was observed in the means. The mean of viral load in the metagenomics approach was the highest, whereas the lowest mean (240070.83) was found in the NimbleGen/NEBNext Microbiome (low V/L) approach. These variations were highly significant ($p = 0.00001$). In addition, great variation was observed in the means for total reads; the mean in the target enrichment approach was the highest, whereas the lowest mean was in the NimbleGen/NEBNext Microbiome (high V/L) approach. These variations were highly significant ($p = 0.013$) (Table 3-8).

Table 3-7: Descriptive statistics for viral load (IU/mL) in all NGS runs

Method	Number of samples	Mean	Std. Deviation	Minimum	Maximum
Metagenomics	14	82245830.00	47441770.164	5783259	120000000
Target enrichment (NimbleGen)	22	4215917.91	13100076.520	6386	57267116
NimbleGen/ NEBNext Microbiome (high V/L)	30	27346428.97	41370239.158	94536	120000000
NimbleGen/ NEBNext Microbiome (low V/L)	12	240070.83	616523.324	5590	2180000
Total	78	26505968.37	43429065.501	5590	120000000

Table 3-8: Descriptive statistics for total reads in all NGS runs

Method	Number of samples	Mean	Std. Deviation	Minimum	Maximum
Metagenomics	14	2877681.85	1035560.68	1365100	4673166
Target enrichment (NimbleGen)	18	17203198.66	32307827.40	502832	141382858
NimbleGen/ NEBNext Microbiome (high V/L)	30	2005624.66	2118003.88	120220	7669620
NimbleGen/ NEBNext Microbiome (low V/L)	12	10674837.66	6063320.47	3577944	21185952
Total	74	7273133.97	17085447.14	120220	141382858

3.2.8.2 Sequencing method versus mapped read and genotype

As shown in Table 3-9, great variation was observed in means with the highest in the NimbleGen/NEBNext Microbiome (low V/L) approach and the lowest in the metagenomics approach. These variations were highly significant (p =0.000).

Genotype D was dominant 54 (69%) as shown in Table 3-10. The target enrichment approach method also exhibited the highest failure (59.1%). These variations were significant (Chi-square = 55.141, p =0.000).

Table 3-9: Descriptive statistics for mapped read in all NGS runs

Method	Number of samples	Mean	Std. Deviation	Minimum	Maximum
Metagenomics	14	47358.71	44331.786	284	111466
Target enrichment (NimbleGen)	18	107514.67	183526.885	2	672692
NimbleGen/ NEBNext Microbiome (high V/L)	30	993779.7	2058706.242	95	7453288
NimbleGen/ NEBNext Microbiome (low V/L)	12	4445983.5	5287047.906	6053	13569877
Total	74	1158965.9	2863710.941	2	13569877

Table 3-10: Descriptive statistics for genotype in all NGS runs

Method	A	C	D	E	Failed	Total
Metagenomics	0	0	14	0	0	14
	0.0%	0.0%	100.0%	0.0%	0.0%	100.0%
Target enrichment (NimbleGen)	0	0	9	0	13	22
	0.0%	0.0%	40.9%	0.0%	59.1%	100.0%
NimbleGen/ NEBNext Microbiome (high V/L)	1	0	25	4	0	30
	3.3%	0.0%	83.3%	13.3%	0.0%	100.0%
NimbleGen/ NEBNext Microbiome (low V/L)	0	1	7	4	0	12
	0.0%	8.3%	58.3%	33.3%	0.0%	100.0%
Total	1	1	54	9	13	78
	1.3%	1.3%	69.2%	11.5%	16.7%	100.0%

In summary, Chi-square analysis indicated that target enrichment (NimbleGen) exhibited the highest failure (63.6%), compared with other approaches, as shown in Table 3-11. Both runs of NimbleGen/NEBNext Microbiome revealed full and significant coverage (100%, p =0.000).

Table 3-11: Coverage group crosstabulation

		Coverage group		Total
		failed	Full coverage	
Groups	Target enrichment (NimbleGen)	14	8	22
		63.6%	36.4%	100.0%
	NimbleGen/ NEBNext Microbiome (high V/L)	3	27	30
		10.0%	90.0%	100.0%
	NimbleGen/ NEBNext Microbiome (low V/L)	0	12	12
		0.0%	100.0%	100.0%
Total		17	61	78
		21.8%	78.2%	100.0%

Additionally, we conducted analysis of variance (ANOVA) for coverage parameter (full or failure) versus all variables.

3.2.8.3 Coverage group versus viral load (IU/mL), total reads coverage nt, and coverage percentage

As presented in Table 3-12 for viral load, the mean for those failures (385302.29) was lower than for full coverage (33785498.26, p = 0.004). The mean total read for failures was greater but not significant (p = 0.051).

The mean coverage nt for failures was lower than for full coverage (p = 0.000). Similarly, the coverage % mean for failures was lower than for full coverage (p = 0.000) (Table 3-13).

Table 3-12: Coverage group for viral load (IU/mL) and total reads

		N	Mean	Std. Deviation	Minimum	Maximum
Viral load (IU/mL)	failed	17	385302.29	1415194.66	6386	5874169
	Full coverage	61	33785498.26	46612766.77	5590	120000000
	Total	78	26505968.37	43429065.501	5590	120000000
Total reads	failed	13	15660949.69	38130981.68	237152	141382858
	Full coverage	61	5485566.68	6772999.74	120220	36147438
	Total	74	7273133.97	17085447.14	120220	141382858

Table 3-13: Descriptive statistics for coverage group, coverage nt and coverage percentage

		N	Mean	Std. Deviation	Minimum	Maximum
Coverage nt	failed	13	1148.3077	858.64762	150	2641
	Full coverage	61	3159.1967	70.60638	2802	3215
	Total	74	2805.9324	847.87661	150	3215
Coverage %	failed	13	36.0869	26.98508	4.71	83
	Full coverage	61	99.0985	1.80167	92.93	100
	Total	74	88.0289	26.55607	4.71	100

3.2.9 HBV Phylogenetic analysis

As indicated previously, we obtained 47 consensus sequences from 47 HBV-positive samples from Saudi Arabia that exhibited full genome coverage (>90). These samples were eligible for inclusion in the phylogenetic analysis. The results indicated that HBV genotype D is predominant in Saudi Arabia, followed by genotype E, genotype A, and genotype C. Although, there was only one sample that was genotype C specific it represents ~2% of our samples and thus should not be discounted. Also, Phylogenetic analysis confirmed that sub genotype D1 is the predominant compared to another sub genotype, (Figure 3-7).

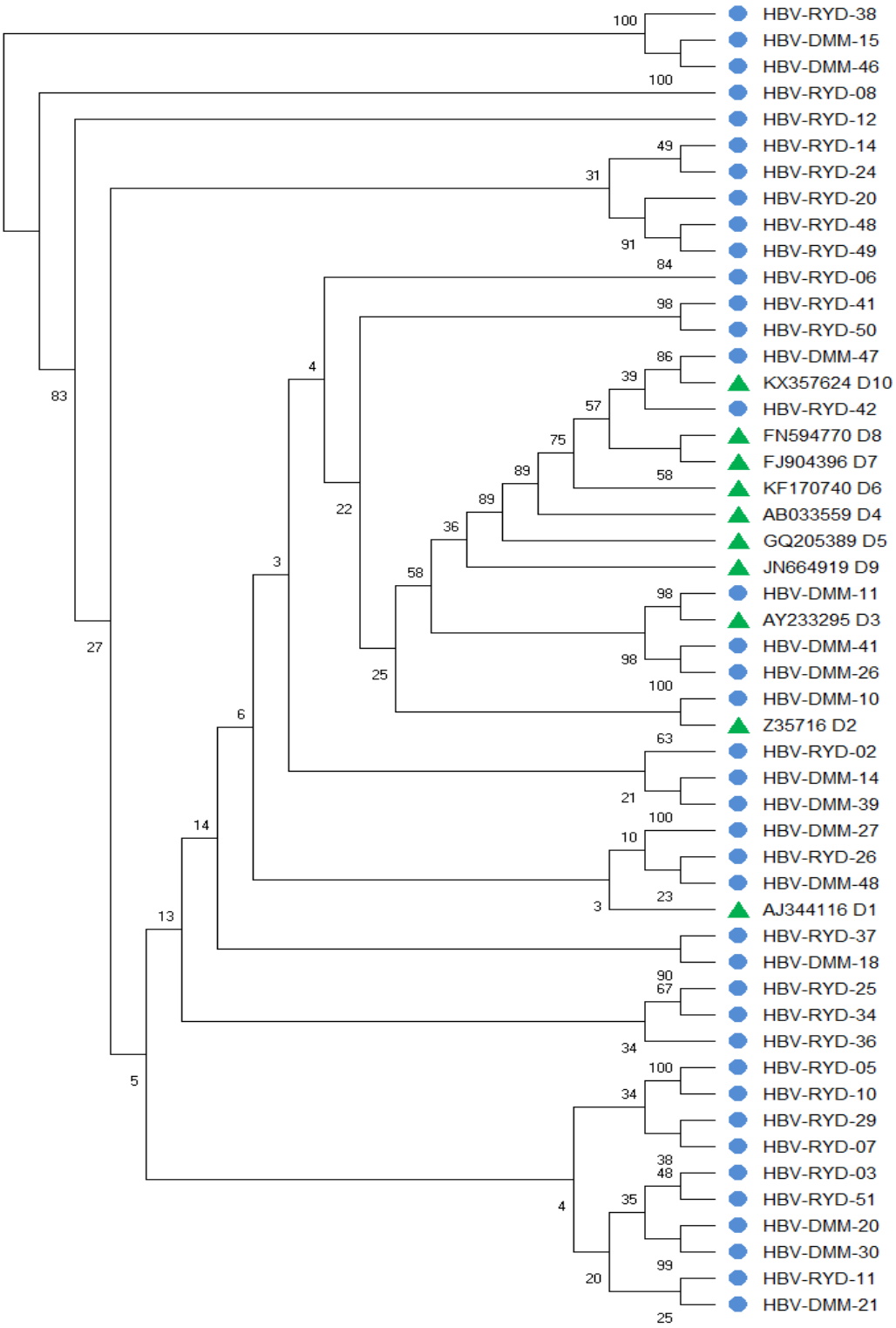


Figure 3-7: Molecular phylogenetic analysis of HBV/D sub genotype using the Maximum Likelihood method. The Maximum Likelihood method based on the Tamura-Nei model and used the entire nucleotide sequences of the 10 sub genotypes D1-D10 (green dots). Saudi HBV/D sequences were determined in the present study and are indicated with a blue circle. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 51 nucleotide sequences. Evolutionary analyses were conducted using MEGA7.

Determination of Antiviral HBV Resistance (NAs) and HBV Vaccine Escape Mutations (VEMs)

3.2.9.1 Antiviral HBV Resistance (NAs)

The mutational analysis indicated that 6 HBV samples exhibited NA resistance. Drug resistance mutations were detected at positions rt80, rt91, rt134, rt153, rt204, rt215, and rt221 (Table 3-14). Notably, one sample with genotype D (HBV-DMM-27) exhibited three mutations (L80I, L91I and M204I) (Figure 3-8). The N236T adefovir dipivoxil resistance mutation and resistance mutations against other antiviral drugs (e.g., entecavir and telbivudine) could not be detected in any HBV sample. In the patient's samples, in order to determine if they had resistance to a specific antiviral it had to be at a 50% frequency in the genome sequence.

Table 3-14: Summary of HBV drug resistant mutations (RAMs) identified from HBV genome sequences

Sample	Mutation	NA resistance	Nature of the resistance	Genotype
HBV-RYD- 38	Q215S	Lamivudine resistance	putative drug resistance.	D
HBV-RYD- 29	D134N	Lamivudine resistance	related to HCC outcomes.	D
	F221Y	Adefovir dipivoxil	putative drug resistance, associated with the progression of severe liver diseases.	D
HBV-DMM- 48	M204V	Lamivudine resistance	primary drug resistance.	D
HBV-DMM-47	R153W	tenofovir resistance	secondary resistance mutation.	D
HBV-DMM-27	L80I	Lamivudine resistance	related to HCC outcomes.	D
	I91L	Lamivudine resistance	related to HCC outcomes.	D
	M204I	Lamivudine resistance	primary drug resistance.	D
HBV-DMM-21	D134E	Tenofovir resistance	associated with the progression of severe liver diseases.	D

Primary drug resistance: mutation refer to amino acid change that result in reduced susceptibility to an antiviral agent.; Secondary resistance mutation: restore replication defects associated with primary drug resistance and may be associated with low level reduced susceptibility; Putative resistance: might be selected under drug pressure, however, it reported without sufficient evidence in vitro to classify the functional relevancy.

3.2.9.2 HBV Vaccine Escape Mutations (VEMs)

HBsAg is a surface antigen targeted by antibodies present in vaccinated people and by antibodies binding to HBsAg in serological immunoassays (Caligiuri et al., 2015). It is the major envelope protein in HBV, which is composed of 226 amino acids. The amino acid positions between 99 and 169 are termed the major hydrophilic region (MHR) (Caligiuri et al., 2015). This contains the 'a' determinant (residues 124–147), which was originally defined as the antigenic region shared by all serological variants of HBV. It is the region primarily associated with the induction of a protective humoral immune response and the immunodominant region of HBsAg (Guangxi et al., 2015). Amino acid substitutions and multiple changes in the 'a' determinant can modify the antigenicity and immunogenicity of the hepatitis B virus (HBV), resulting in a failure to react in immunoassays and antibody escape (Guangxi et al., 2015). In the current study, we identified different amino acid substitutions involving 12 mutations detected in the major hydrophilic loop (I110L, P120S, P120T, T126S, T131N, M133I, M133T, Y134F, Y134K, S143T, F161Y and A168V). Six mutations had amino acid substitutions (T126S, T131N, M133I/T, Y134F and S143T) within the 'a' determinant region (a.a 124-147) of the MHR (Figure 3-9). The classical vaccine escape mutations D144E and G145R, which are reported to be closely associated with genotype D, could not be detected. However, the well-described vaccine escape mutation P120T was found in one sequence of an HBV sample.

Additionally, substitution I110L was detected in two samples, whereas F161Y was found only in one sample; both mutations lead to vaccine escape (Coppola et al., 2015). Similarly, substitutions P120S, P120T, T126S and S143T were detected and may lead to problems in both diagnostic assays and vaccine escape (Guangxi et al., 2015). Substitution T131N, which is responsible for the rescue of virion secretion, was also found. In addition, M133T was detected and may lead to problems in diagnostic assays; however, it has yet to be identified as VEMs (Yan et al., 2017). Furthermore, the Y134K mutation was found, which had not been detected in the usual escape setting (Ziaee et al., 2016).

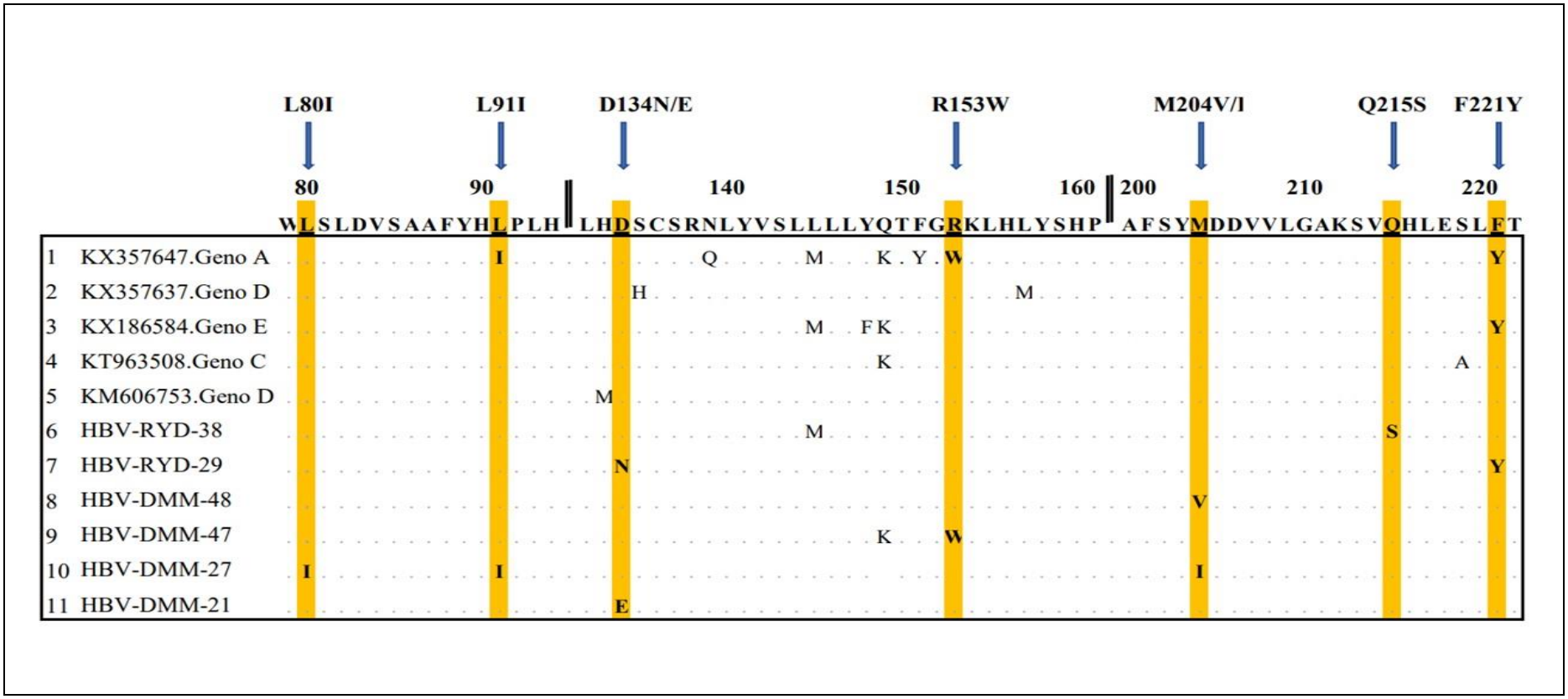


Figure 3-8: Analysis of antiviral therapy resistance mutations.

The amino acid sequences of the HBV polymerase (reverse transcriptase region; aa rtM1 to rtQ344) were aligned with the corresponding region of the reference sequences.

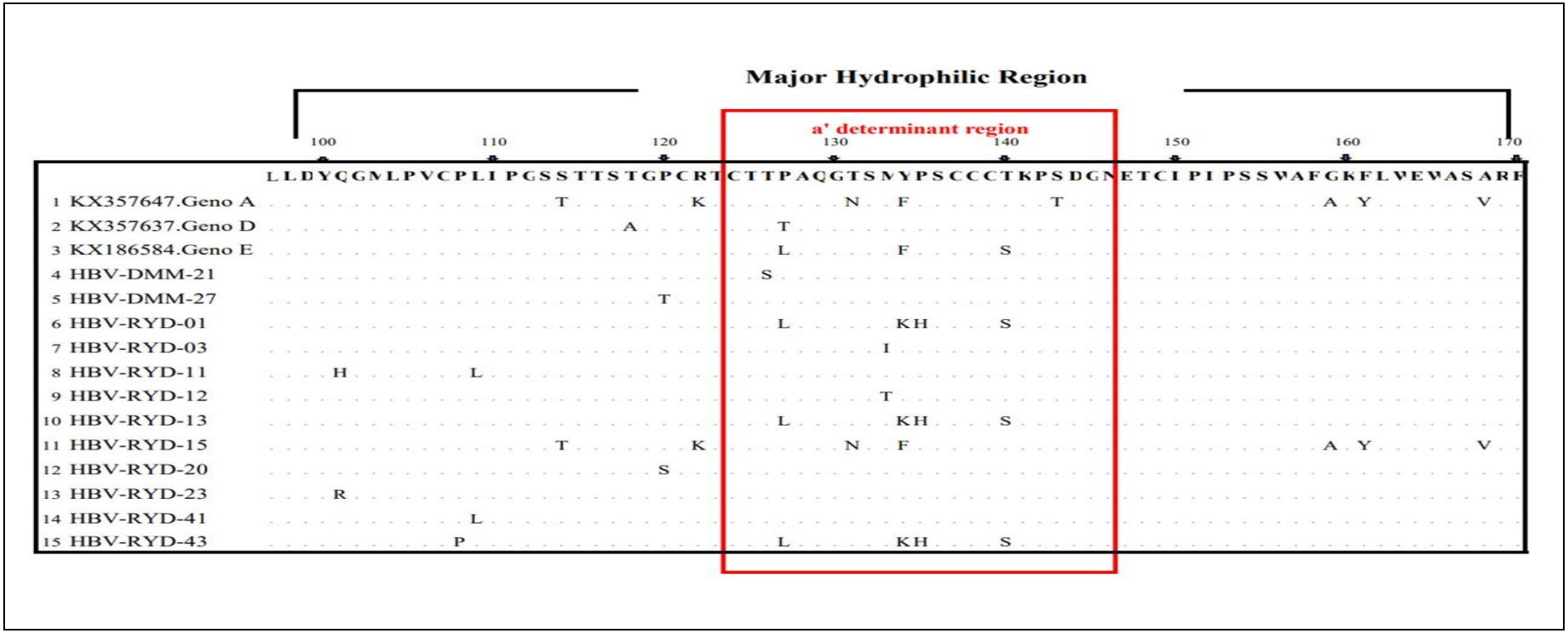


Figure 3-9: Analysis of vaccine escape mutations.

Distribution of amino acid substitutions (mutations) detected within the major hydrophobic region (MHR) of the HBV S-gene. The amino acid sequences in the ‘‘a’’ determinant region were aligned with the corresponding region of the reference sequences (HBV-genotypes A, D, and E).

3.3 DISCUSSION

Chronic hepatitis B infection (CHB) is a prominent public health issue for the 21st century and a major health burden in Saudi Arabia and other parts of the world (Lau et al., 2007). HBV is the most common cause of chronic viral hepatitis in Saudi Arabia (Tamura et al., 2011). However, little is known about the prevalence or distribution of HBV genotypes in this region, and few studies have been conducted to determine the genotype variation of HBV in the Kingdom with the limitation of the sequencing method leading to partial sequencing rather than whole-genome sequencing. In 2006, one investigation (Al Ashgar et al., 2006) assessed the distribution of HBV genotypes in Saudi Arabia using the INNO-LiPA methodology (LiPA, INNO-LiPA HBV genotyping assay, Innogenetics NV, Ghent, Belgium) and identified HBV genotyping among patients at a tertiary referral centre, which serves population groups resident in different regions of Saudi Arabia. Serum samples of 54 CHB patients and reported the prevalence of 85% had D genotype; while 5.7% has A genotype and E has 1.4% respectively. Recent report using what method conducted by Alshabi et al. (2021) reported the prevalence of HBV genotype out of 50 patients in the southwestern region in the Saudi Arabia as 90% of D genotype, 4% of each A and H genotype and 2% of E genotype (Alshabi et al., 2021). Another previous study using this method conducted by Abdo et al. (2006) included 70 patients showed that genotype D was in 57 patients (81.4%), 4 patients had genotype E (5.7%), 1 patient had genotype A (1.4%), 1 patient had genotype C (1.4%), and 7 patients had mixed genotype (10%) (Abdo et al., 2006).

At the beginning of this thesis, the study of HBV strains and drug resistance was via Sanger sequencing and the use of enrichment protocols and NGS was limited. This maybe in part due to cost. The analysis of multiple areas of the genome in a single reaction is enabled by NGS, while Sanger sequencing only permits the analysis of one region at a time (Slatko et al., 2013). Furthermore, in order to identify low-frequency variants, it gives greater sensitivity by presenting more discovery power to detect novel or rare variants. It also allows a more rapid turnaround time for high-sample volumes which have full genomic coverage (Slatko et al., 2013; McNaughton et al., 2019).

The beauty of NGS and target enrichment allow us to identify potentially new genotypes, subgenotype and mixed infections with greater confidence and reliability. This study validated a protocol for deep whole-genome sequencing for HBV using the target enrichment method. To the best of our knowledge, this is the first description of whole genome HBV sequencing using this method. The data indicated that NGS is a reliable tool for genotype prediction, which makes it useful for clinical evaluation; however, a larger number of samples is needed to evaluate the method for detection of other genotypes. Our findings using NGS were in in agreement with previously published studies implemented in the Saudi Arabia to identify the circulating genotypes. Our results confirmed that genotype D is predominant strain in Saudi Arabia, followed by genotype E, genotype A, and genotype C. Recently, many publications and strategies are in development to bring HBV genome sequence analysis the clinical virology laboratory.

Teng et al. (2018) have developed NGS platform for quantitative detection of pre-S mutants in plasma samples from patients with HCC. They have compared the new NGS platform to the TA cloning-based approach that is well established for isolation of specific pre-S gene DNA from the PCR products. Their data indicated that the NGS approach sensitivity higher in detecting pre-S deletion quantitatively which was indicating that this approach could enhance the accuracy of detecting pre-S deletion that led to improving the outcome prediction of patients with HCC (Teng et al., 2018).

Another approach that has developed by Barbosa et al. (2020) validated the use of NGS as an approach to monitor viral polymorphism and HBV genotypes that are circulating within the populations and within an individual. They defined an assay based on amplification in only three overlapping regions of the full-length HBV genome. This may be critical in chronically infected patients where treatments are failing (Barbosa et al., 2020).

A Recent project conducted by Ishii et al. (2020) developed an approach for NGS using Target-capture sequencing method (SeqCap EZ Pure Capture (Roche)). They performed analysis of the HBV genome integration in Huh7 cells. , this is relevant in the context of this thesis as we use Huh7 in subsequent chapters for the replicon system. They have accomplished functional analysis and demonstrated the expression of some HBV proteins of HBV integrants in transfected Huh7 cells with DNA sequences. They did not used a clinical sample which was the major limitation on their study (Ishii et al., 2020).

Colson et al. (2020) have also applied NGS approach in clinical microbiology laboratory for the purpose of diagnosing of HBV. DNA was extracted from plasma sample utilizing Qiagen

kit and then directly sequenced on Illumina instrument using Nextera XT® protocol without prior HBV DNA amplification with PCR. They have used their approach only in two patients and this was one limitation on their study (Colson et al., 2020). For first patient, they observed two mutation (sD144A and sG145R) that were associated with vaccine escape while for second patient two mutations (rtL180M/rtM204V) that conferring lamivudine and entecavir resistance were detected. Genotype A was observed for both patients (Colson et al., 2020).

The NGS approach also has advantage over traditional methods as there are issues with detecting mixed infection as confirmed by a study mixing different ratios of the different genotype (Mercier et al., 2011), NGS results enables a comparison result and allows sensitivity of the two methods; hybridization and sanger, to be determined from the same sample to be confirmed (Lowe et al., 2016). Some of the challenges encountered during the study included the need to deplete host reads, which could involve enrichment and amplification steps. The cost of the system and reagents as well as the interpretation of NGS data require substantial bioinformatics support and adaptation for different genomic configurations.

HBV requires methods refined for a circular and partial dsDNA genome. In addition, it presents several bioinformatics challenges that include the circular genome, overlapping open reading frames, and the different genome lengths of the genotypes. Moreover, large gaps in our understanding remain regarding the relationship between HBV genome structure, the replication cycle, diversity, transmission, and clinical outcomes. Recent sequencing advances offer an enormous opportunity to generate datasets that will help to address some of these questions. A repository (or database) of standardised reference genomes of all HBV genotypes subtypes has been defined and it has facilitated the consistent assembly and analysis required to develop insights into current and future epidemiology, inform better clinical assessment, improve deployment of current antiviral drugs and vaccines, and drive the discovery of new antiviral agents (McNaughton et al., 2019).

The method validated in this study paves the way for a better understanding of HBV epidemiology by offering a new tool for large scale whole genome sequencing in reference laboratories or research settings. Further developments to shorten the sequencing time and automate the bioinformatics pipeline would facilitate the introduction of this method into diagnostic laboratories and public health settings.

We contributed to the development of the in-house bioinformatics tools in the centre to analyse drug resistance and vaccine escape mutations amongst the Saudi patient samples involved in my work. Data collected along with the literature review on HBV antiviral drugs and resistance

were shared via GLUE software, which is specifically developed for HBV drugs currently in the development phase at the MRC-University of Glasgow Centre for Virus Research.

GLUE is a data-centric bioinformatics environment for viral sequence data with a focus on variation, evolution, and sequence interpretation. Additionally, this is alpha software, which is still undergoing testing prior to its official release. Creating an HBV sequence database to identify resistance-associated variants and inform any future treatment plans would be of tremendous value, this allows a genotype to phenotype study of arising mutations to be undertaken.

Several limitations towards our study are the following: the need to deplete host reads, which could involve enrichment and amplification steps, the cost of the system and reagents, as well as the interpretation of NGS data, require substantial bioinformatics support and adaptation for different genomic configurations.

In conclusion, the results from this new NGS protocol provide valuable information for the identification of the mutant surface antigens and polymerase genes of HBV-infected patients. It is hoped that the project will contribute to national guidelines for the eradication of HBV in Saudi Arabia. Additionally, a selected set of drug-resistant mutations were used to generate a cell culture-based HBV system in which we measured the replication rate by HBeAg secretion in the presence and absence of appropriate drugs, as described later in Chapter 5.

Chapter 4. Establishment of infectious/non-infectious HBV replication systems

4.1 Introduction

The *in vitro* HBV cell culture system is an essential tool for studying the biological properties of HBV and for screening anti-HBV drugs (Iwamoto et al., 2017; Lin et al., 2016; Xu et al., 2021). The host and tissue specificity of HBV means that the availability of a stable and reliable *in vitro* cell culture system is a key factor affecting research on the mechanism of HBV action (Iwamoto et al., 2017; Lin et al., 2016; Xu et al., 2021). Therefore, creating an active cell culture system supporting HBV infection has become vital in studying HBV and the development of effective therapeutic drugs.

4.1.1 The history of infection/replicon systems for studying HBV

Several HBV *in vitro* infection systems have been established recently. Although these systems have weaknesses, they are useful in the study of aspects of the virus life cycle and they represent an important tool in the development and evaluation of anti-HBV drugs (Xu et al., 2021). The following history reflects the major milestones in the development of cell culture systems for HBV, focusing on the cells lines that support HBV replication as well as the cell lines that can be infected with HBV (Figure 4-1).

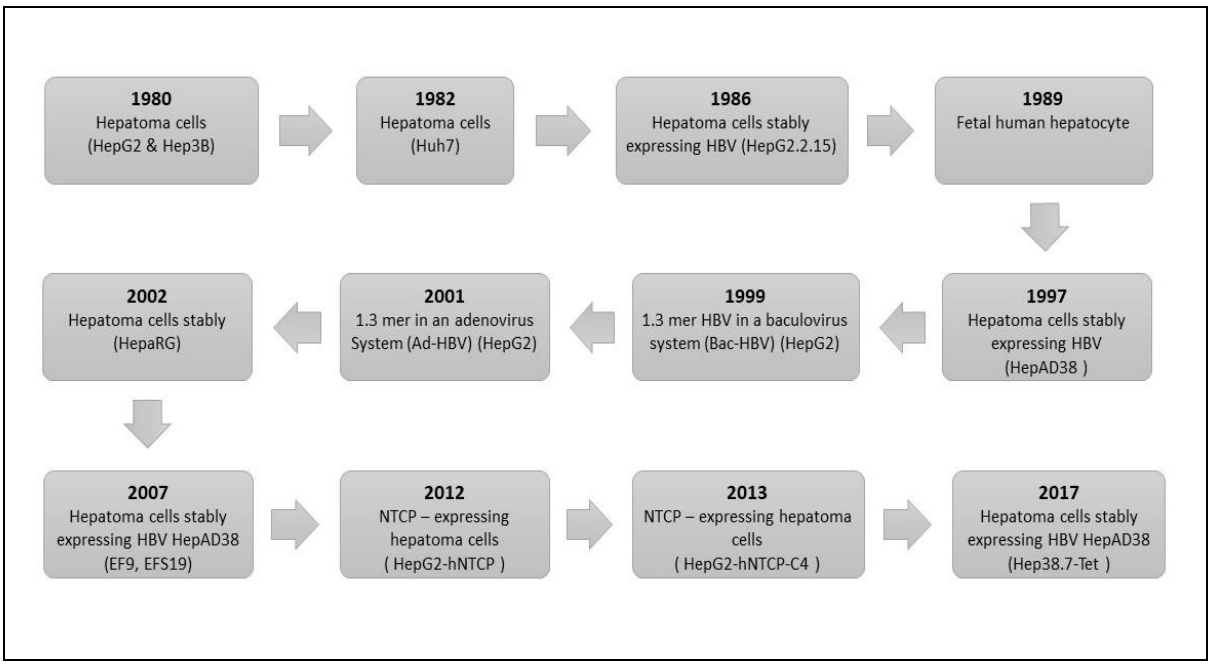


Figure 4-1: The history of an *in vitro* systems for studying HBV.

Timeline the substantial research achievements in HBV cell culture systems *in vitro*. The discovery of the receptor of HBV (NTCP) in 2012 has advanced the understanding of HBV biology. Then, several liver cancer cells overexpressing NTCP that support HBV infection have been established, opening a new door for studying HBV infection.

HepG2 and Hep3B were originally established by Aden et al. (1979). They were isolated from liver biopsy specimens of a 15-year-old Caucasian male from Argentina with primary primary hepatoblastoma, or an 8-year-old black male from the US with primary HCC, respectively (Aden et al., 1979; Knowles et al., 1980). Both cell lines contained distinctive rearrangements of chromosome 1, and other abnormal chromosomes. But they differ in the number of chromosomes per cell as HepG2 cells contain an average of 55 (50–56) chromosomes per cell whereas Hep3B cells, 60 (Knowles et al., 1980). In addition, HepG2 was HBV negative and non-tumorigenic while Hep3B was HBV positive and tumorigenic (Knowles et al., 1980). Analysis of the cell culture fluid from two new human hepatoma-derived cell lines showed that 17 of the major human plasma proteins were synthesized and secreted by these cells (Knowles et al., 1980).

Hep 3B was also produced the two major polypeptides of the HBV surface antigen. Metastatic hepatocellular carcinomas were appeared when Hep 3B was injected into athymic mice. Experimental models were provided from these two-cell line for investigation of plasma protein biosynthesis and the relation of the HBV genome to tumorigenicity (Knowles et al., 1980).

In 1982, the human hepatoma cell line Huh7 was established from male hepatoma tissue obtained from a 57-year-old Japanese male patient with hepatocellular carcinoma. Huh7 was found to replicate continuously in a chemically defined medium when the medium was supplemented with Sodium Selenite (Na_2SeO_3). In this medium, the cells were grown better than in serum-containing medium without any adaptation period (Nakabayashi et al., 1982). Several human plasma proteins were observed to be produced in this cell line (albumin, alpha 1-antitrypsin, hemopexin, ceruloplasmin, fibrinogen, beta-lipoprotein, haptoglobin, and alpha 1-fetoprotein). Moreover, these proteins were detected during periods of serial cultivation over 9 months under the above culture conditions. The cell line was developed in the fully defined synthetic medium which offered a new approach for investigating the growth and metabolism of human hepatoma cells in vitro (Nakabayashi et al., 1982).

Sells et al. (1986) developed a recombinant vector pDoLT-HBV-1 comprising of two head-to-tail dimers of HBV genome in a tail-to-tail orientation. This plasmid was together with a plasmid encoding the neomycin resistance gene to co-transfect into HepG2 cells. Following selection and cloning in the presence of G418, individual clones were tested for their ability to synthesize and secrete HBsAg and HBeAg. One such clone of the cells, designated HepG2.2.15, produced hight levels of both antigens (Sells et al., 1986). As reported, the cell line carries a complete and an incomplete HBV DNA along with extra-chromosomal and chromosomally integrated DNA and cccDNA (Sells et al., 1986). This line can also generate a variety of HBV-specific mRNAs (3.5 kb, 2.5 kb, 2.1 kb). It can also produce all viral markers, and stably secrete Dane particles, HBsAg and HBeAg for a considerable period of time (Sells et al., 1988). HepG2.2.15 is thus a widely used cell line as it produces infectious virions and supports continuous virus replication. The HepG2.2.15 cell line is an efficient tool for the study of the function, structure, the first screening of anti-HBV drugs *in vitro*, as well as the gene expression and regulation of HBV DNA. However, this cell line has limitations as it lacks the viral receptor NTCP and thus is insensitive to direct infection with HBV. Therefore, it cannot be utilised for studying cellular entry, HBV adsorption or virus uncoating (reviewed by Xu et al., 2021).

A technique centred on human embryonic hepatocytes with similar attributes to those of adult hepatocytes was developed by Ochiya (1989). A considerable number of mononuclear polyhedral hepatocytes – which were in trabeculae from fetal liver tissue at a gestation period of between 20 and 24 weeks – were obtained by Ochiya et al. (1989). Glycogen, glucose-6-phosphatase and transpeptidase were identified in the medium after these were cultured for approximately a week. The albumin that was identified on the second day after the cells had

been plated continued being secreted until the sixteenth day. The biochemical attributes and the regular morphologies were retained for a minimum period of two weeks of subsequent culture (Ochiya et al., 1989). Following infection with HBV of primary human fetal hepatocytes, viral replication indexes were identified both intracellularly and within the culture medium. Fetal human hepatocyte culture *in vitro* has the ability to imitate the biological role of hepatocytes within the human body. Furthermore, in comparison with human primary adult hepatocytes, the cells have an improved survivability, differentiation and proliferation (Ochiya et al., 1989). It has several benefits as a tool for studying HBV infection: (i) the cells have the potential to be infected by sera containing HBV particles; (ii) all known viral protein, DNAs and RNAs that are detected in HBV infected livers *in vivo* can also be produced in infected fetal human hepatocytes *in vitro*; (iii) infectious viral particles are released by the cells; (iv) the cells are capable of producing cccDNA. Nevertheless, the fact that productive infection of these cells is sustained only for a period of between 16 and 18 days remains a major limitation (Reviewed by Lin et al., 2007).

Ladner et al. (1997) transfected HepG2 cells with the plasmid pTet-HBV. This plasmid carries the tetracycline-responsive CMV-IE promoter fused with a 1.1 copy of ayw subtype of the HBV genome. It was used to generate the HepAD38 cell line capable under inducible condition of high level of HBV production and viral replication (Ladner et al., 1997). HepAD38 cell line produces about 11 times more HBV DNA than HepG2.2.15 cells (Ladner et al., 1997).

Moreover, tetracycline can be applied in order to regulate HBV replication in the cell line. HBV is not synthesised in the presence of medium containing tetracycline because pgRNA synthesis is inhibited (Ladner et al., 1997). The cells immediately express pgRNAs, cccDNA and HBV subsequent to the removal of tetracycline. Due to the low sensitivity of direct cccDNA recognition and the fact that its detection is susceptible to interference by rcDNA signals, the HBeAg secreted by HepAD38 cells can be utilized as the major surrogate marker of cccDNA (Ladner et al., 1997).

The integration of the HBV genome enables the HepAD38 and HepG2.2.15 cell lines to secrete HBV particles continuously. In fact, these cells, are frequently used as the virus source for HBV infection in cell culture systems and are broadly utilised in associated studies. Furthermore, transient transfection with plasmids carrying more than genome-length HBV sequences have also been a useful strategy to study aspects of virus life cycle.

Compared to HepG2.2.15 cells, in HepAD38 cells the initiation of HBV replication can be controlled accurately thus enabling greater levels of virus production. However, like

HepG2.2.15, the HepAD38 cell line is limited in that it is inappropriate for the study of the interaction between host and virus cells during the early stage of HBV infection as they also lack NTCP. Nonetheless, this cell line has been used for the study of anti-HBV drug screening and HBV replication process (reviewed by Xu et al., 2021).

The biochemical attributes and the regular morphologies were retained for a minimum period of two weeks of subsequent culture (Ochiya et al., 1989). Following the application of primary human fetal hepatocytes with HBV infection, viral replication indexes were identified both intracellularly and within the culture medium. Fetal human hepatocyte culture *in vitro* has the ability to imitate the biological role of hepatocytes within the human body. Furthermore, in comparison with human primary adult hepatocytes, the cells have an improved survivability, differentiation, and proliferation (Ochiya et al., 1989).

It has several benefits as a tool for studying HBV infection: (i) the cells have the potential to be infected by sera containing HBV particles; (ii) all known viral protein, DNAs and RNAs that are detected in HBV infected livers *in vivo* can also be produced in infected fetal human hepatocytes *in vitro*; (iii) infectious viral particles are released by the cells; (iv) the cells are capable of producing cccDNA. Nevertheless, the fact that productive infection of these cells is sustained only for a period of between 16 and 18 days remains a major limitation (Reviewed by Lin et al., 2007).

Ladner et al. (1997) transfected HepG2 cells with the plasmid pTet-HBV. A mammalian promoter derived (pCMV) from a high-copy-number pUC-based plasmid was used to permit protein expression in mammalian systems in this vector, the expression is driven by the human cytomegalovirus (CMV) immediate early promoter to promote constitutive expression of cloned inserts in a wide variety of cell lines. A translation initiation sequence must be incorporated if the DNA fragment to be cloned does not have an initiating ATG codon or an optimal sequence for initiating translation since the pCMV- vector does not contain an ATG initiation codon. It has fifteen unique restriction enzyme recognition sites in the multiple cloning site (MCS) that arrange with alternating 5´ and 3´ overhangs to allow serial exonuclease III/mung bean nuclease deletions. In order to construct plasmid pTet-HBV, the cytomegalovirus immediate-early (CMV-IE) promoter from plasmid pCMV-HBV was removed and subsequently fused with the ayw subtype of the HBV genome (Figure 4-2). It was replaced by the tetracycline-responsive CMV-IE promoter for the purpose of acquiring the HepAD38 cell line that offers a high level of HBV production along with the precisely regulation in the initiation of viral replication (Ladner et al., 1997). HepAD38 cells were selected due to the low level of virus replication in the previous cell line such as HepG2.2.15.

The expression of the 1.1 copies of the HBV genome contained within the HepAD38 cell genome is regulated by the inducible CMV-IE promoter. The pre-core gene disruption causes the HepAD38 cell line to produce HBV DNA about 11 times more than HepG2.2.15 cells (Ladner et al., 1997).

Moreover, tetracycline can be applied in order to regulate HBV replication in the cell line. HBV is not synthesised in the presence of medium containing tetracycline because pgRNA synthesis is inhibited (Ladner et al., 1997). The cells immediately express pgRNAs, cccDNA and HBV subsequent to the removal of tetracycline. Due to the low sensitivity of direct cccDNA recognition and the fact that the detection results are susceptible to interference by rcDNA signals, the HBeAg secreted by HepAD38 cells can be utilized as the major replacement marker of cccDNA; thus, the estimation level of HBeAg can be used to measure the replication level of cccDNA (Ladner et al., 1997).

The integration of the HBV genome enables the HepAD38 and HepG2.2.15 cell lines to secrete HBV particles continuously. In fact, HepAD38 cells and HepG2.2.15 cells, are frequently used as the virus source for HBV infection in cell culture systems and are broadly utilised in associated studies. Furthermore, transient transfection with plasmids carrying more than genome-length HBV sequences have also been a useful strategy to study aspects of virus life cycle.

Compared to HepG2.2.15 cells, HepAD38 cells produce greater levels of HBV than HepG2.2.15 cells do and are also able to control the initiation of viral replication accurately. However, like HepG2.2.15, the HepAD38 cell line is limited in that it is inappropriate for the study of the interaction between host and virus cells during the early stage of HBV infection as they also lack NTCP. Nonetheless, this cell line has been used for the study of anti-HBV drug screening and HBV replication process (reviewed by Xu et al., 2021).

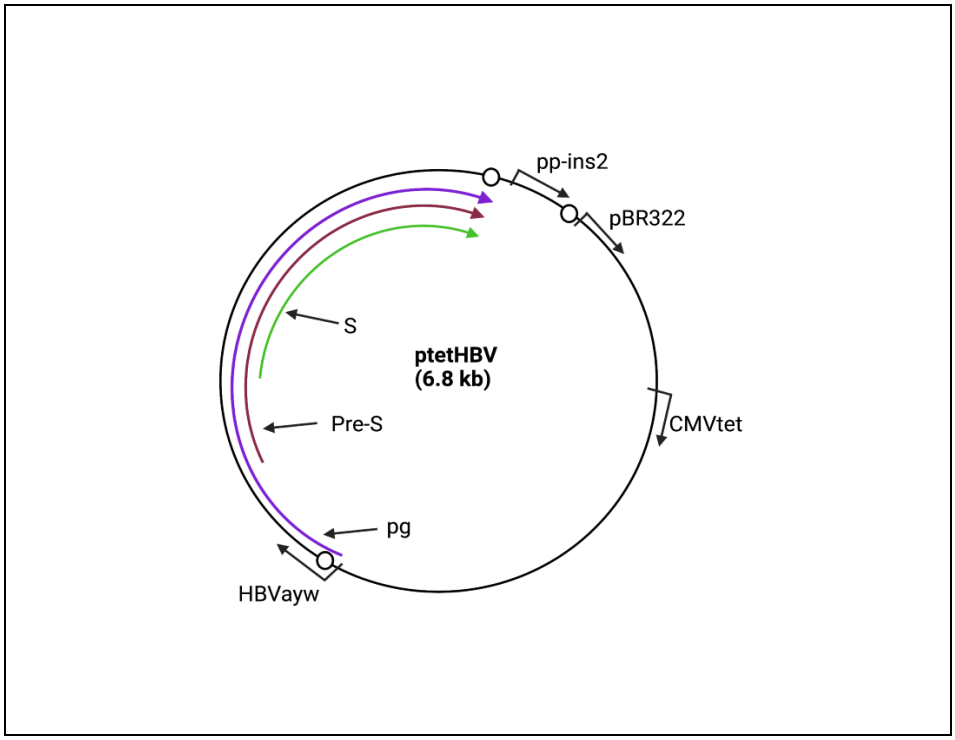


Figure 4-2: Structure diagram ptetHBV (Ladner et al., 1997).

A map of the plasmid pTet-HBV map; this design was adapted from Ladner et al. (1997). Position of three major viral transcripts, pregenomic RNA, Pre-S, and surface are indicated as coloured arrows. The position of the CMVtet promoter is also indicted in addition to HBV sequences. Plasmid pBR322 was used as vector and polyadenylation signal was introduced by the pp-ins-2 segment which is derived from pre-proinsulin gene 2 (Ladner et al., 1997).

Guo et al. (2007) generated HepDE19 and HepDES19 cells that were similar to HepAD38. with some modifications. Briefly, HepG2 cells were transfected with plasmid pTet-off that expresses the Tet-responsive transcriptional activator and plasmid pTREHBVDE in which HBV pgRNA expression is controlled by a CMV early promoter with a tetracycline-responsive element. Transfected HepG2 cells were selected with 500 μg/ml G418 in the presence of 1 μg/ml tetracycline. G418-resistant colonies were picked and expanded into cell lines. HBV replication was induced by culturing cells in tetracycline-free medium, and the levels of viral DNA replicative intermediates were determined by Southern blot hybridization. The cell lines with high levels of HBV replication were chosen and designated as HepDE19 and HepDES19 (Guo et al., 2007).

HepDE19 contains a 1.1 mer HBV transgene mutated in its 5' pre-core ATG, while the 3' pre-core ATG remains unaltred. The expression of HBeAg in this cell line emanated from the episomal DNA rather than integrated DNA, thereby supplying a platform for a large-scale screening of cccDNA-targeting drugs (Guo et al., 2007; Cai et al., 2012). Although HepDE19 cells accomplish the same function as do HepAD38 cells, but the dependency connection between secreted cccDNA and HBeAg is closer than the dependence in the HepAD38 cell line (Guo et al., 2007). Subsequently, a cccDNA reporter cell line, known as HepBHAe82 which is 'second generation' was developed. In this case, an in-frame haemagglutinin (HA) epitope tag was implemented into the pre-core HBeAg domain's open reading frame in the transgene of HepBHAe82 cells with no disruption to any cis-element, crucial for HBeAg secretion and HBV replication (Cai et al., 2016). Although such developments have enabled HBV production in vitro, their application regarding studying the regulation of HBV replication is limited because some carry a neomycin resistance gene and a greater-than genome length HBV. Therefore, the search to establish stable HBV-expressing cell lines continues, but it will encourage the study of the connection between host genes and HBV (Wose et al., 2020).

Iwamoto et al. (2017) isolated Hep38.7-Tet cell line, a clonal derivative of HepAD38 cells, exhibiting greater cccDNA levels and HBV replication. They also produced approximately 3 times higher HBs and 3–5 times higher levels of cccDNA. These findings indicated that HBV replicates more efficiently in Hep38.7-Tet cells than in its parental cells. Consequentially, the Hep38.7-Tet cells were screened with Nocodazole which is known to disrupt cellular microtubules and to arrest cell cycle. Mainly, once the cells reached confluency level tetracycline was removed to induce the replication of HBV and the cells were then treated with Nocodazole up to six days. Culture supernatant was collected for HBV quantification and cells were harvested for cell viability. The results indicated that HBV DNA decreased to less than

20% in the supernatant without reducing cell viability. Likewise, the HBV capsid assembly in microtubule-disrupted cells was remarkably reduced without substantial change in pre-assembly process (Iwamoto et al., 2017). The data thus suggested a significant role of microtubules in capsid formation during HBV replication. (Iwamoto et al., 2017).

In order to introduce a replication competent HBV genome into HepG2 cells, Dalaney et al. (1999) utilised baculovirus-base system, the purpose of which was to establish the HBV recombinant baculovirus/HepG2. The recombinant baculovirus contained 1.3-genome length HBV construct originating from pTHBV1.3 which had previously been shown to drive high levels of HBV replication in the livers of transgenic mice. Following infection of HepG2 cells with this recombinant baculovirus, the presence of high levels of HBsAg and HBeAg was demonstrated and shown to be sustained for a minimum period of 35 days (Delaney et al., 1998). The recombinant baculovirus/HepG2 technique is unique in that expression of cccDNA and rcDNA can be detected relatively easily; consequently, it may be applied to quantify the impact of antiviral agents on nuclear HBV DNA (Abdelhamed et al., 2002). Furthermore, it is amenable to studying viral resistance to nucleoside analogues (Delaney et al., 1999).

Sprinzl et al. (2001) generated an adenovirus (Ad) vector carrying a 1.3-fold overlength HBV or duck HBV genome following transfection into HEK-293 cells, The recombinant Ad-HBV1.3 release from HEK-293 cells was subsequently used to infect the HepG2 cells (Sprinzl et al., 2001). The E1 region was substituted from adenovirus by a reporter gene and replication-competent HBV genomes. Following transduction into hepatoma cells and primary hepatocytes viral proteins and nucleic acids were detected and infectious HBV and duck HBV virions were shown to be released.

These adenovirus vectors were also shown to be suitable for *in vivo* experiments (Sprinzl et al., 2001). Among all known delivery vectors, adenovirus vectors are specifically the most efficient in the transfer of exogenous DNA into the livers of experimental animals of choice as.no species barrier exists for the Ad HBV to transduce target tissue cells, thus enabling it to attain replication in hepatocytes of non-specific host (Sprinzl et al., 2001).

Gripon et al. (2002), isolated hepatocytes from liver tumor tissue of HCV-infected patients. After many passages, the cells first acquired an undifferentiated morphology. Subsequently, the authors added hydrocortisone and DMSO to the medium for the purpose of differentiating into cells that had the functional attributes of biliary cells and mature hepatocytes. Eventually, this process led to the generation of HepaRG cells that following differentiation supported the

infection and relocation cycle of HBV (Gripon et al., 2002; 2005). Furthermore, HepaRG cells have been broadly used to study toxicity and drug metabolism (Marion et al., 2010).

Yan et al. (2012) utilised the primary Tupaia hepatocytes from tree shrews to conduct photo-cross-linking experiments with a synthetic pre-S1 peptide. These experiments enabled identification of NTCP as a receptor for HDV and HBV (Yan et al., 2012). Tree shrews, except for chimpanzees, are the only known animals that can be infected by HBV. Primary hepatocytes derived from tree shrews are susceptible to infection by HBV (Walter et al., 1996). Yan et al. (2012) showed that expression of NTCP is essential for infection of HepG2 or Huh7 cells. They established stable cell lines, Huh7-hNTCP and HepG2-hNTCP, which were susceptible to infection with both HBV. Such cell lines have the potential for continuous propagation and spread of HBV, but they are more useful in the study of early events of virus infection.

Iwamoto et al. (2013) established HepG2-hNTCP-C4 cell line, a clonal derivative of HepG2-NTCP, which were more efficient at virus infection compared to the polyclonal pool of parent cells. HepG2-hNTCP-C4 cells were shown to be susceptible to infection by blood–borne and cell culture-derived HBV. HBV infection in this cell line was facilitated by pre-treating cells with 3% dimethyl sulfoxide (DMSO) enabling nearly 50% of the cells to be infected with HBV. In comparison, only 7% to 20% of HepaRG cells were infected (Iwamoto et al., 2013). The infection of HBV was blocked by an anti-HBV surface protein neutralizing antibody, by compounds known to inhibit NTCP transporter activity, and by cyclosporin A and its derivatives. Also, further chemical screening identified oxysterols, oxidized derivatives of cholesterol, as inhibitors of HBV infection (Iwamoto et al., 2013).

In our study, we exploited a previously developed plasmid carrying assembly-defective over-length HBV genome for assessing the effectiveness of antiviral drugs. Specifically, the plasmid construct, pTHBV1.3-L$^-$, carrying 1.3 genome length of HBV (*ayw* subtype) DNA with a point mutation in the initiation codon (ATG to ACG) of preS1 was used. This plasmid was kindly provided to my supervisor back in 2001 by Professor Heinz Schaller, University of Heidelberg. My supervisor had more recently had it sequenced with the help of the CVR NGS team. It confirmed that this plasmid carries nt 1068-3182/1-1990 of HBV 1.3 genome with the point mutation in preS1 initiation codon as described above. The complete sequence of pTHBV1.3-L$^-$ is shown in Appendix 7. It includes a 59 nt terminal redundancy encompassing enhancers I and II, the origin of replication (DR1 and DR2), pregenomic/core promoter regions,

transcription initiation site of the pgRNA, the unique polyadenylation site, and the entire X ORF (Figure 4-3). This plasmid has been shown to be viral DNA replication-competent but assembly-deficient and as such can be used at BSL2 level. Following transfection into hepatoma cells, the plasmid is capable of producing replication-competent pre-genomic RNA and cccDNA (Ko et al., 2018) and as is amenable to studying viral replication processes. As described below, we made further modifications to pTHBV-L$^-$ and WT HBV 1.3 length genome for use in our studies to evaluate the efficacy of different NA drugs on genomes carrying drug-resistance-associated mutations of interest. These molecular tools will also be of value in future virus-related research in our laboratory. Specifically, the aim was to construct plasmids carrying HBV sequences of interest and a drug resistance gene to enable generation of stable cell lines.

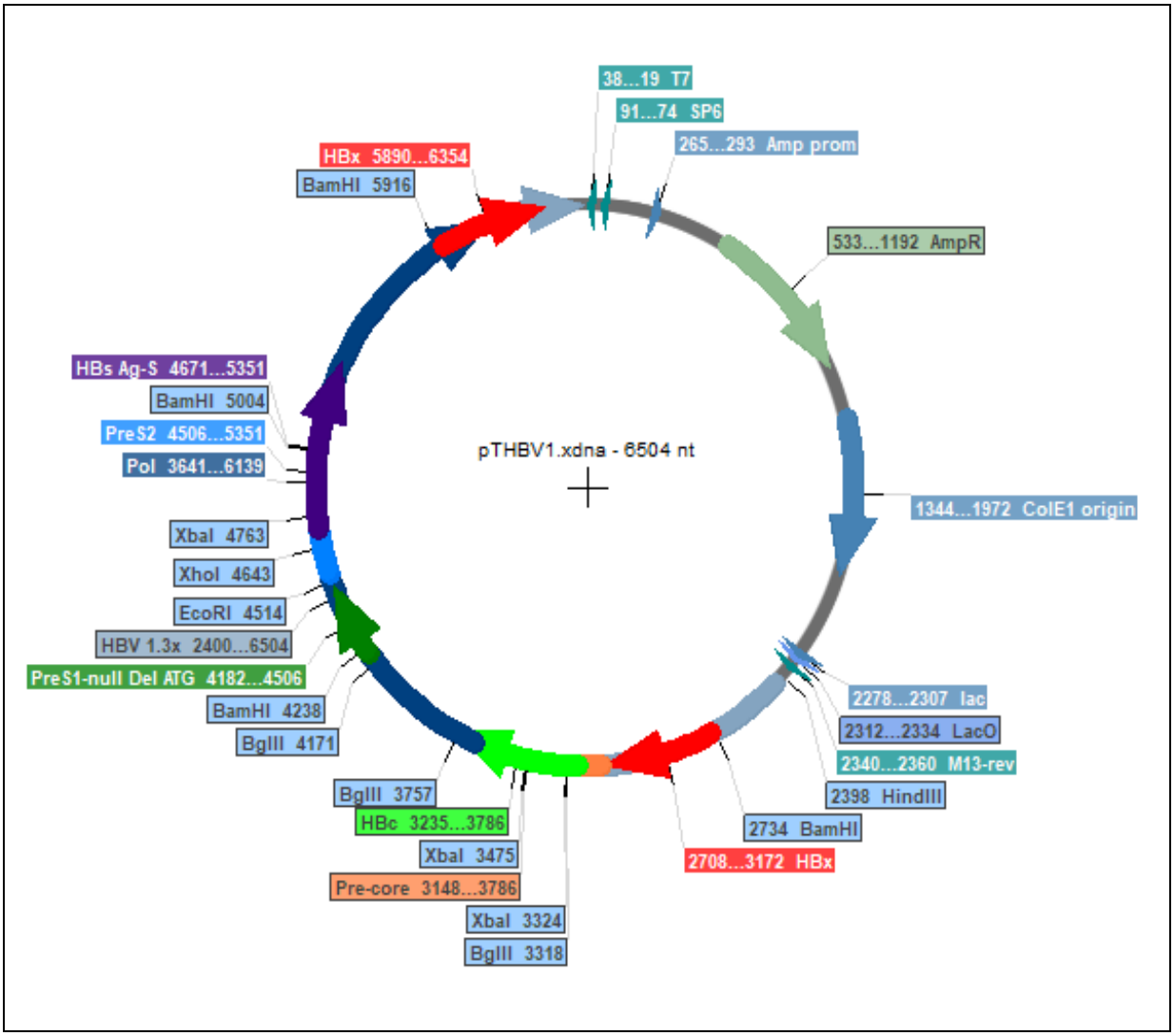


Figure 4-3: Diagram of pTHBV1.3-L- construction.

Diagram of pTHBV1.3-L$^-$ carrying a 1.3-fold-overlength genome of HBV (derived from pHBV1.3) genotype D and subtype *ayw*. The HBV core (HBc), Pre-Core, X, S Pre-S, and Pol sequences are indicated.

4.2 Results

4.2.1 Subcloning of the HBV sequences from pTHBV-L$^-$ into the retrovirus transfer vector pQCXIP

The entire HBV genomic sequences from pT-HBV- L$^-$ subcloned into the retrovirus transfer vector pQCXIP (Clonetech, see Appendix 8 for the sequence) which has the puromycin-resistance gene downstream. In theory, this plasmid would allow generation of VSV-G enveloped retroviral particles carrying the HBV genome following co-transfection into HEK-293T cells of plasmids expressing VSV-G, MLV-Gag-Pol, and pQCXIP carrying the HBV genome described below. The VSV-G-enveloped particles generated thus could then be used produce cell lines of interest stably expressing HBV1.3-L-. First, nt 2381-4519 of pT-HBV- L$^-$ carrying nt 1086 to 3182 of the viral genome were PCR-amplified using the forward primer HB29 (CAT**TGATCA** CTGCAGGGCCCGTCGACAAGCTT with Bcl I site) and the reverse primer HB15 (CAT**GAATTC**CACTGCATGGCCTGAGGATGA with EcoRI site). Separately, the 2 kbp EcoRI to NheI fragment of pT-HBV-L$^-$ (nt 4514 to 14 carrying nt 1 to 1990 of HBV genome) was gel-extracted (as per the protocol described in Section 2.2.6.1). The above PCR product was cleaved with BclI and EcoRI and ligated together with the EcoRI-NheI fragment into pQCXIP cut with BglII to XbaI just upstream of the CMV promoter which drives the IRES-puromycin expression (Figure 4-4). This ligation reaction was called Lig 1.

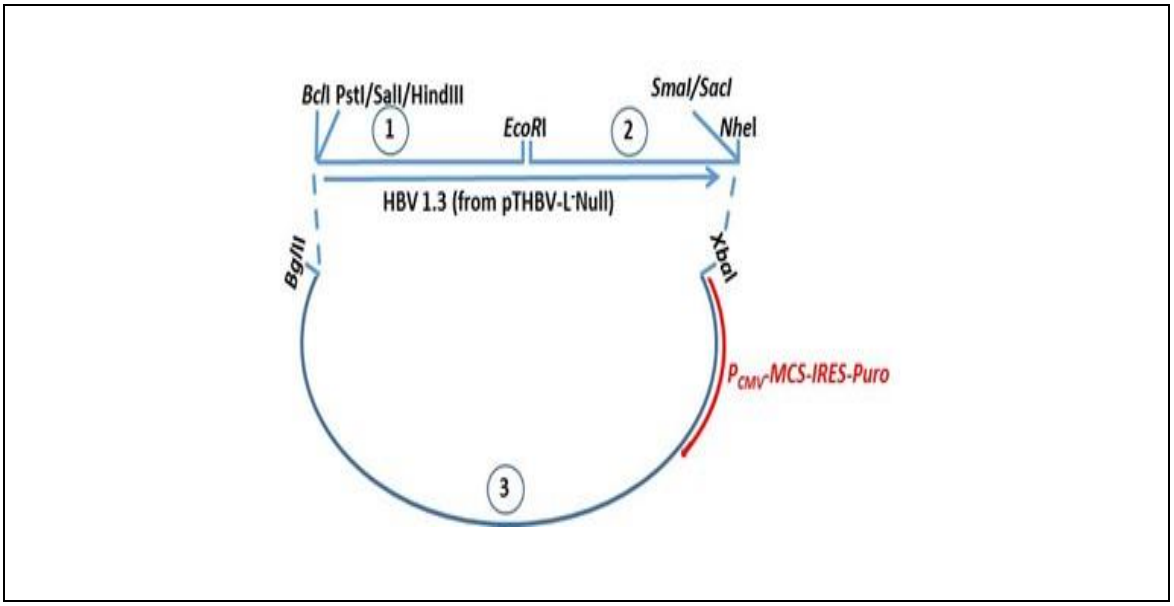


Figure 4-4: Diagram of Ligation 1 construct.

1) PCR up sequences upstream from the HBV genome in pTHBV-L$^-$ to the EcoRI site at nt 3182. Primers HB29 + HBV-15 were used and pTHBV-L$^-$ as template to PCR up a fragment of ~2kbp and cut it with *Bcl*I + *Eco*RI. 2) Cut pTHBV-L$^-$ with EcoRI and NheI and gel was extracted the ~2kp 3' fragment. 3) pQCXIP: cut with *Bgl*II (NE3.1) and *Xba*I (add CiP also) and then gel was extracted linear plasmid fragment. Ligated the three fragments to generate a construct (carrying HBV nt 1068-3182/2-1990) that carries the HBV L-minus 1.3 genome in a retrovirus transfer construct expressing replication-competent/assembly-deficient HBV genome following VSVpp transduction, and a stable cell line can be made following selection on puromycin.

An aliquot of Lig 1 was transformed into E. coli Stbl3 (NEB) cells. Following selection on L-agar ampicillin, 12 colonies (named 1-1 to 1-12) were screened to identify clones carrying the desired sequences. These colonies were grown up in small cultures and plasmids prepare using the Qiagen Miniprep kit. An aliquot of these plasmids were treated with restriction enzymes HindIII and EcoRI. The sizes of the fragments resulting from this digestion were expected to be 3872, 2694, 2116, 1169, 859, 348, and 247 bp. As shown in Figure 4-5, plasmids 1-1, 1-4, 1-6, 1-9 and 1-10 gave the expected profile (the 348 and 247 bp fragments were undetectable due to their small size). The plasmid 1-6, named pQ-HBV-L$^-$ (1-6), was selected for further analysis and was sent off for nucleotide sequencing which confirmed that it carried the expected HBV sequence (See Appendix 9).

Figure 4-5: Screening of Ligation 1 (pQ-HBV-L$^-$) by Agarose gel electrophoresis.

As described in the text, the cleaved plasmids were electrophoresed through 0.8% agarose gel, and fragments indicated in the gel were extracted. Several samples indicated the fragment profile from ligation 1. Only sample (1-6) was selected for further investigation with fragment sizes (3872, 2694, 2116, 1169, 859, 348, and 247 bp). Also, plasmids 1-1, 1-4, 1-6, 1-9 and 1-10 gave the expected profile (the 348 and 247 bp fragments were undetectable due to their small size). All products were run alongside with a 1kb DNA ladder.

4.2.2 Subcloning of HBV 1.3mer WT (wild type) sequences in pQCXIP

The plasmid pHBV1.3mer overlength WT HBV genotype D construct (obtained from Addgene) contains 1.3 units of HBV genome, spanning nt 1072-3182/1-1990 of HBV ayw subtype (Wang et al., 2009). It is replication- and assembly competent, so upon transfection into Huh7 or HepG2 cells, this plasmid initiates HBV DNA replication and the subsequent assembly and release of the virus that could be used in infection studies. Here, pQCXIP carrying WT 1.3 length HBV genome was constructed by combining relevant fragments from pHBV1.3mer WT and pT-HBV-L$^-$ effectively restored the ATG start codon in preS1. Work involving this plasmid were conducted in BSL3 level.

Briefly, nt 1072 to 3182 of the WT viral genome of pHBV1.3mer WT were PCR-amplified using the forward primer HB30 (CAT**TGATCA**CTGCAGGGCCCGTCGACAAGCTT with Bcl I site) and the reverse primer HB15 (CAT**GAATTC**CACTGCATGGCCTGAGGATGA with EcoRI site). Separately, the 2 kbp EcoRI to NheI fragment of pT-HBV-L$^-$ (nt 4514 to 14 carrying nt 1 to 1990 of HBV genome) was gel-extracted (as per the protocol described in Section 2.2.6.1.The above PCR product was cleaved with BclI and EcoRI and ligated together with the EcoRI-NheI fragment of pT-HBV-L$^-$ (described above in 4.1.2) into pQCXIP cut with BglII to XbaI just upstream of the CMV promoter which drives the IRES-puromycin expression (Figure 4-6). This ligation reaction was called Lig 2. The resulting construct is expected to contain chimeric sequences derived from pHBV1.3mer WT (nt 1072-3182) and pT-HBV-L$^-$ (nt 1-1990) reconstituting 1.3 length WT.

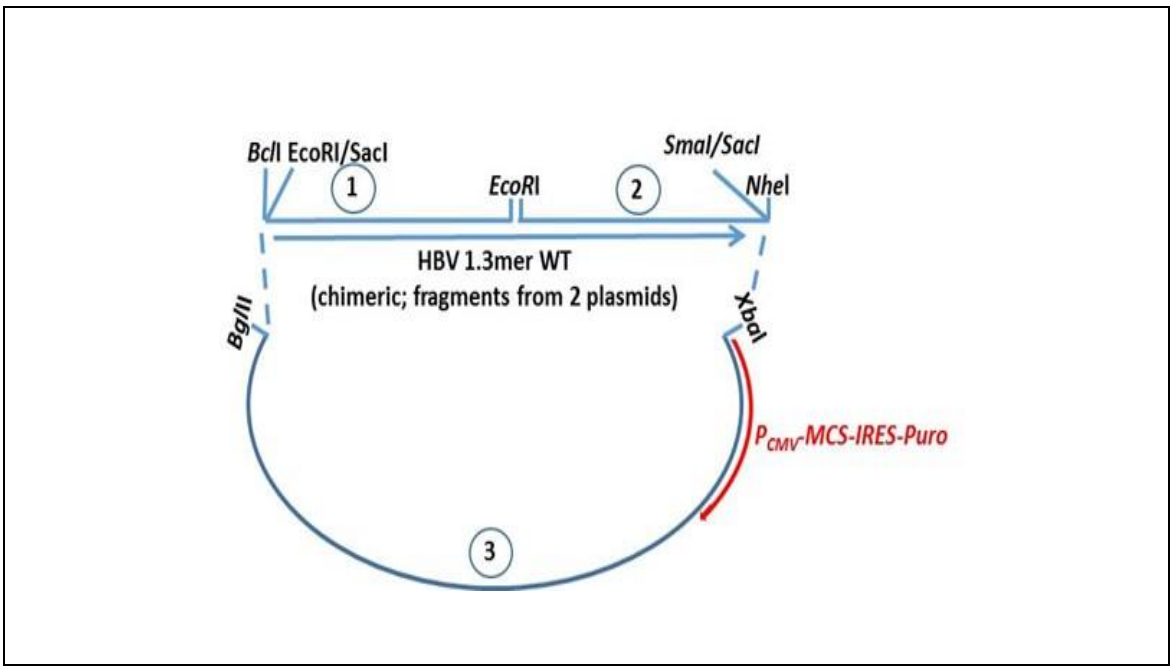


Figure 4-6: Diagram of Ligation 2 construct.

1) PCR up sequences upstream from the HBV genome in pHBV1.3mer WT to the EcoRI site at nt 3182. Use primers HB30 + HBV-15 and pHBV1.3mer WT as template to PCR up a fragment of ~2kbp and cut it with *Bcl*I + *Eco*RI. 2) Cut pTHBV-L$^-$ with EcoRI and NheI and gel extract the ~2kp 3' fragment. 3) pQCXIP: cut with *Bgl*II (NE3.1) and *Xba*I (add CiP also) and then gel extract linear plasmid fragment. Ligate the three fragments to generate a construct (carrying HBV nt 1072-3182/2-1990) that carries the 'chimeric' HBV WT 1.3 genome in a retrovirus transfer construct expressing replication-/assembly-competent HBV genome following VSVpp transduction, and a stable cell line can be made following selection on puromycin.

An aliquot of Lig 2 was transformed into E. coli Stbl3 (NEB) cells. Following selection on L-agar ampicillin, 8 colonies (named 2A1 to 2A4, 2B1 to 2B4) were screened to identify clones carrying the desired sequences. These colonies were grown up in small cultures and plasmids prepare using the Qiagen Miniprep kit. An aliquot of these plasmids were treated with restriction enzyme EcoRI. The sizes of the fragments resulting from this digestion were expected to be 6486, 2694, and 2122 bp. As shown in Figure 4-7, plasmid 2B3 had the expected profile. This plasmid, named pQ-HBV1.3mer WT, was selected for further analysis and was sent off for nucleotide sequencing which confirmed that it carried the expected HBV sequence (See Appendix 10).

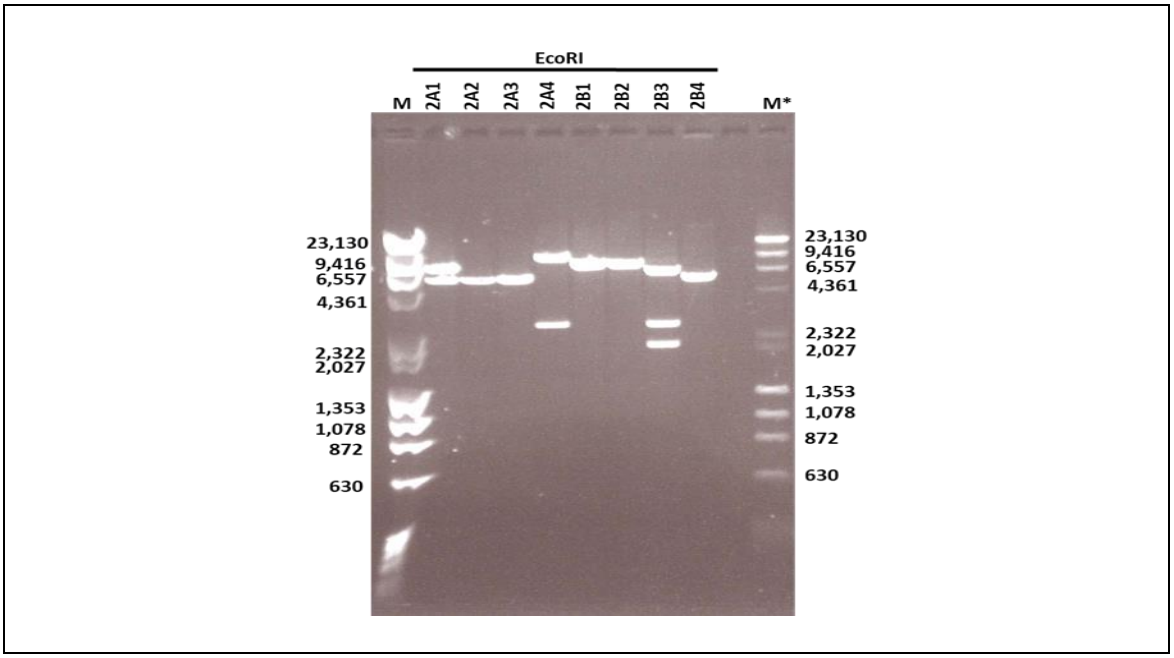


Figure 4-7: Screening of Ligation 2 (pQ-HBV1.3mer WT) by Agarose gel electrophoresis.

As described in the text, the cleaved plasmids were electrophoresed through 0.8% agarose gel, and fragments indicated in the gel were extracted. Several samples indicated the fragment profile from ligation 2. Only sample (2B3) was selected for further investigation with fragment sizes (6486, 2694, and 2122 bp) and was sent off for nucleotide sequencing which confirmed that it carried the expected HBV sequence. All products were run alongside with a 1kb DNA ladder.

4.2.3 Generation and Subcloning of the retrovirus transfer vector pQCXIP pQCXIP carrying 2.7 of length HBV genome

To generate a cell line expessing the L-, M- and S-HBsAg, pQCXIP carrying a shorter 2.7 kb HBV sequence (compared to the above plasmids) was generated. Similar constructs have been described previously (Ni et al., 2019; Lemp et al., 2019). This construct is designed to express all 3 HBsAg (i.e., L, M, and S) and HBx from their native promoter/enhancer. It does not, however, express the HBV core or polymerase and, therefore, is replication-defective and non-infectious. Furthermore, pQCXIP has a puromycin-resistance marker, so it can be used to select a stable cell line. Indeed, my supervisor Prof. Arvind Patel has used this plasmid to generate a cell line, Huh7-3B4, which stably expresses the L-HBsAg. This plasmid, and the stable cell line, can be used to generate HDV particles following transfection or co-transfection with HDV genome-expressing pSVL (D3) or pSVL(D2m).

Briefly, nt 2425 to 3182 + 1 to 1990 of the WT viral genome of pHBV1.3mer WT were PCR-amplified using the forward primer HB31 (CAT**TGATCA** CAATCTCGGGAATCTCAATGTTAGTATT with Bcl I site) and the reverse primer HB15 (CAT**GAATTC**CACTGCATGGCCTGAGGATGA with EcoRI site). Separately, the 2 kbp EcoRI to NheI fragment of pT-HBV- L^- (nt 4514 to 14 carrying nt 1 to 1990 of HBV genome) was gel-extracted (as per the protocol described in Section 2.2.6.1). The above PCR product was cleaved with BclI and EcoRI and ligated together with the EcoRI-NheI fragment of pT-HBV- L^- (described above in 4.2.1) into pQCXIP cut with BglII to XbaI just upstream of the CMV promoter which drives the IRES-puromycin expression (Figure 4-8). This ligation reaction was called Lig 3. The resulting construct is expected to contain chimeric sequences derived from pHBV1.3mer WT (nt 2425-3182) and pT-HBV-L- (nt 1-1990) capable of expressing the viral surface antigens under their native promoter/enhancer.

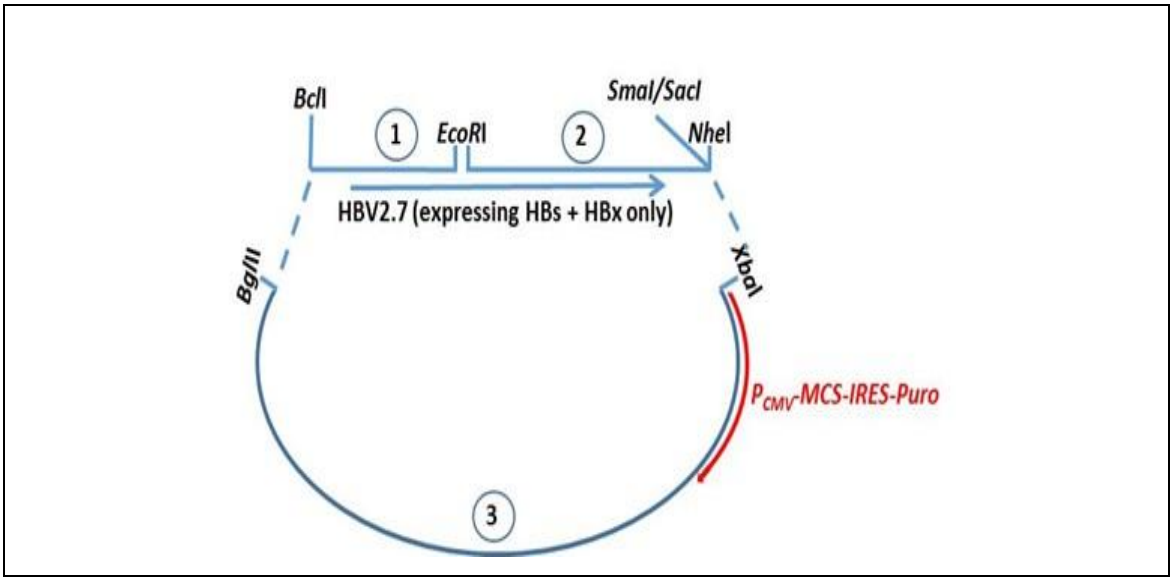


Figure 4-8: Diagram of Ligation 3 construct.

1) PCR up nt 1366 to 2123 (to the EcoRI site) in the HBV genome of pHBV1.3mer WT (representing nt 2425-3182 of HBV genome). Use primers HB31 + HBV-15 and pHBV1.3mer WT as template to PCR up a fragment of ~750bp and cut it with *Bcl*I + *Eco*RI. 2) Cut pTHBV-L^{-} with EcoRI and NheI and gel extract the ~2kp 3' fragment. 3) pQCXIP: cut with *Bgl*II (NE3.1) and *Xba*I (add CiP also) and then gel extract linear plasmid fragment (this was done previously and not shown below in the gel). Ligate the three fragments to generate a construct (carrying HBV nt 2425-3182/2-1990) that carries the 'chimeric' replication-defective HBV WT 1.3 genome sequences in a retrovirus transfer construct expressing L-, M-, S- HBsAg & HBx from own promoters following VSVpp transduction, and a stable cell line can be made following selection on puromycin.

An aliquot of Lig 3 was transformed into E. coli Stbl3 (NEB) cells. Following selection on L-agar ampicillin, 4 colonies (named 3A1, 3A2, 3A4, and 3B4) were screened to identify clones carrying the desired sequences. These colonies were grown up in small cultures and plasmids prepare using the Qiagen Miniprep kit. An aliquot of these plasmids was treated with restriction enzyme EcoRI. The sizes of the fragments resulting from this digestion were expected to be 7279 and 2694 bp. As shown in Figure 4-9, plasmids 3A2, 3A4 and 3B4 had the expected profile. The plasmid 3B4, named pQ-HBV2.7, was selected for further analysis and was sent off for nucleotide sequencing which confirmed that it carried the expected HBV sequence (See Appendix 11).

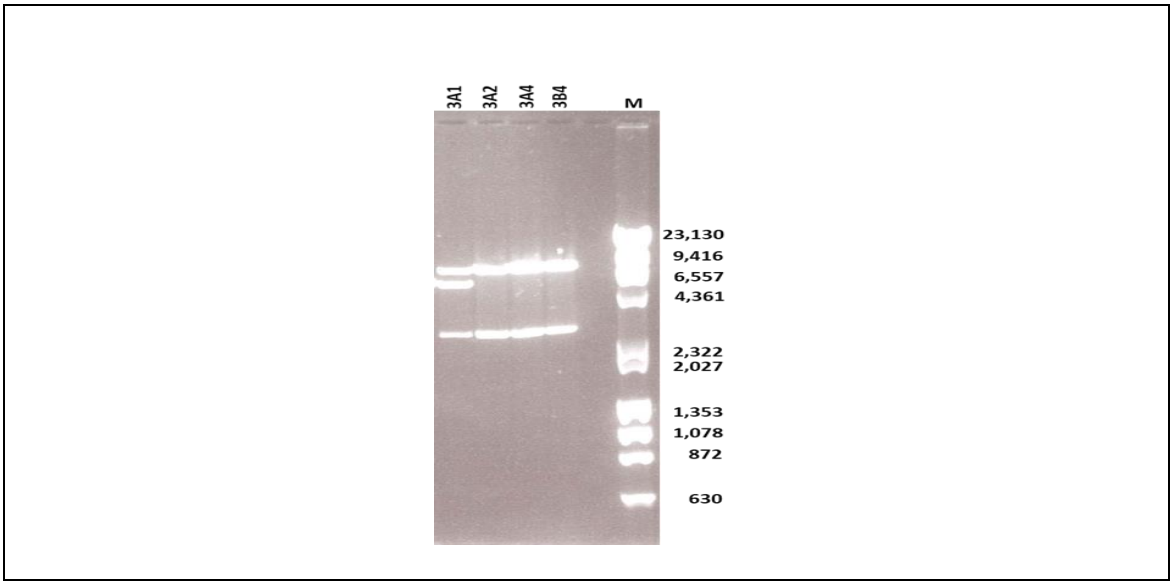


Figure 4-9: Screening of Ligation 3 (pQ-HBV2.7) by Agarose gel electrophoresis.

Sample (3B4) was selected for further investigation. All products were run on a 0.8% agarose gel alongside a 1kb DNA ladder.

4.2.4 Generation of plasmid pTHBV-L$^-$ /Puro

Lastly, pT-HBV- L$^-$ carrying the puromycin-resistance gene was constructed. Specifically, the XhoI to SmaI fragment carrying nt 129 to 1990 of pT-HBV- L$^-$ with a XhoI to EcoRV fragment of pQ-HBV-L- (plasmid 1-6, see 4.2.1) carrying nt 129 to 1990 + CMV promoter-EMCV IRES-puromycin resistance gene sequences (Lig 4). The construct was named L$^-$ /Puro (Figure 4-10).

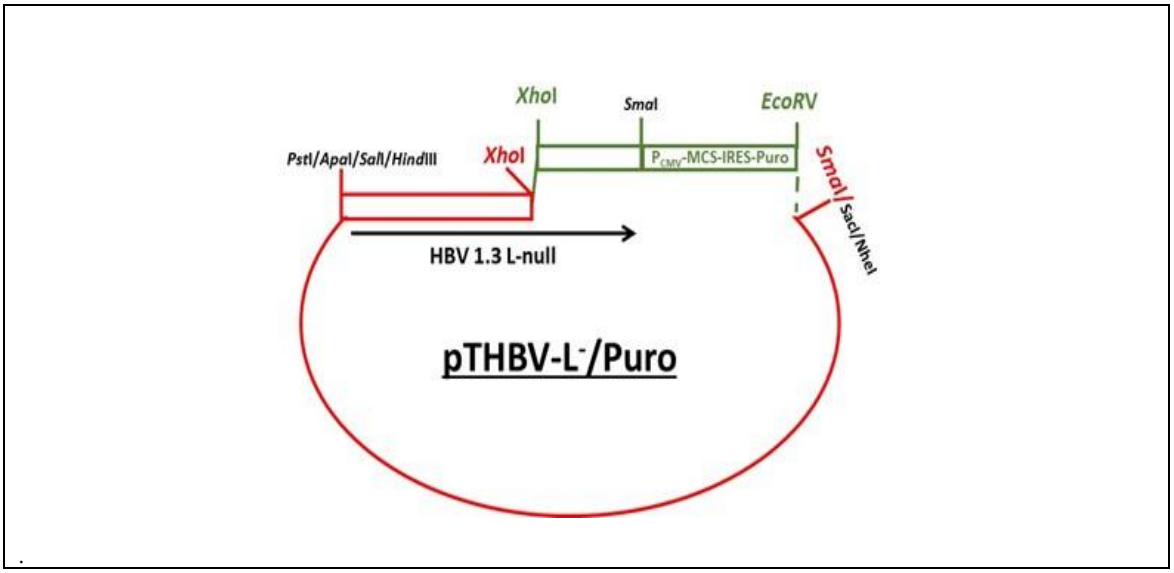


Figure 4-10: Diagram of Ligation 4 construct.

1) Cut pTHBV-L$^-$ with XhoI and SmaI. The digest should give 2 fragments (4640 and 1864bp) and gel extract the large 4640 bp fragment. 2) Cut pQ-HBV-L$^-$ (No 1-6) with XhoI and EcoRV. This should give 2 fragments (7252 and 4053bp) and gel extract the small 4053bp fragment. Ligate the two extracted fragments to generate a construct expressing the L-null HBV nt 1068-3182/1-1990 – P_{CMV}-IRES-Puro.

An aliquot of Lig 4 was transformed into *E. coli* Stbl3 (NEB) cells. Following selection on L-agar ampicillin, 24 colonies were initially screened to identify clones carrying the desired sequences. These colonies were grown up in small cultures and plasmids prepare using the Qiagen Miniprep kit. An aliquot of these plasmids were treated with restriction enzyme EcoRI + NheI, and only one plasmid, No 11, had the expected profile, the initial screen was done by my supervisor which found no 11 to be of potential interest. This plasmid was transformed into *E. coli* and plasmid DNA from 6 colonies was analysed by digestion with EcoRI + NheI, and separately with NheI + XhoI. The sizes of the fragments resulting from this digestion were expected to be 4500, 2694, and 1500 bp (EcoRI/NheI) or 4640 and 4055 bp (NheI/XhoI). As shown in Figure 4-11, all 6 plasmids had the expected respective profiles. The plasmid, named pTHBV-L-/Puro, No 11, was selected for further analysis and was sent off for nucleotide sequencing which confirmed that it carried the expected HBV sequence (See Appendix 12).

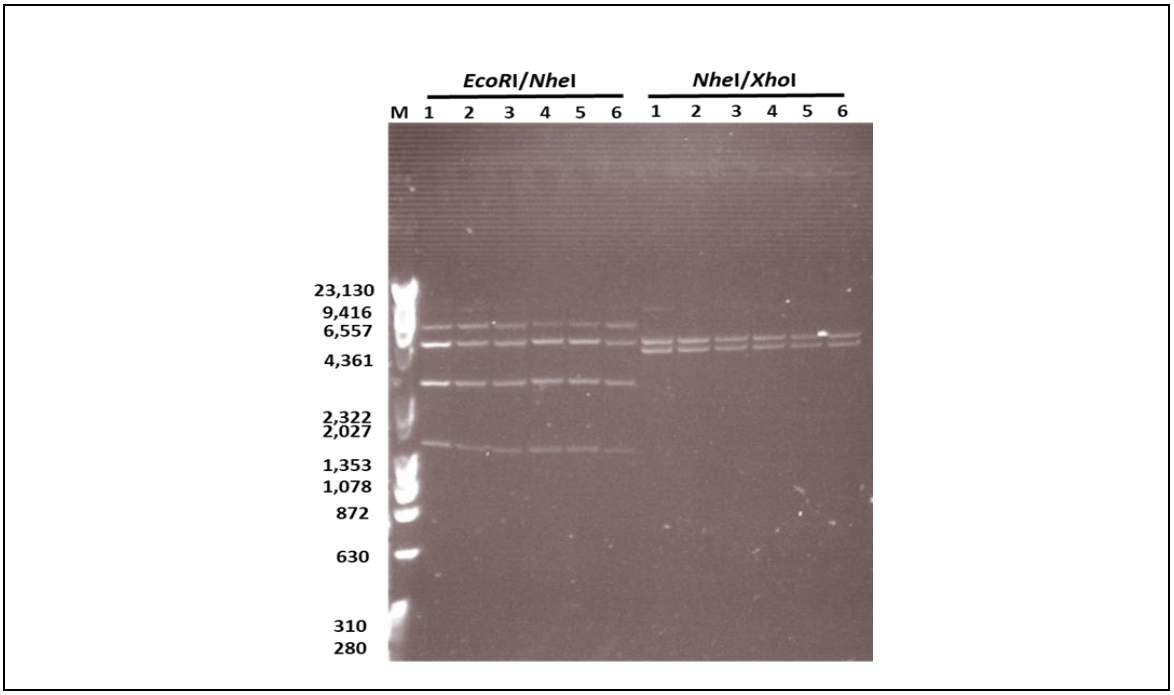


Figure4-11: Screening of Ligation 4 (pTHBV-L^{-}/ Puro, No 11) by Agarose gel electrophoresis.

All products were run on a 0.8% agarose gel alongside a 1kb DNA ladder.

4.2.5 Analysis of intracellular HBV core, DNA synthesis and HBc protein expression in Hep38.7-Tet cells

The establishment of infection and viral persistence are both dependent on the formation of cccDNA during the HBV replication cycle (Naoki et al., 2014). The half-life of cccDNA is longer than other viral nucleic acids, ranging from days to months in animal and tissue culture models. As described above, Iwamoto et al. (2017) developed Hep38.7-Tet, an optimized derivative of the HepAD38 cell line, that supports HBV replication under tetracycline inducible conditions. In Hep38.7-Tet cells, the production of secreted HBeAg is predominantly cccDNA-dependent and thus acts as a surrogate marker of cccDNA formation. Here, we aimed to set up this inducible HBV expression system with a view to studying aspects of virus life cycle. Hep38.7-Tet cells, obtained from Drs Masashi Iwamoto and Koichi Watashi, were maintained and seeded in a 6-well dish with medium containing 400 ng/mL tetracycline as described by Iwamoto et al. (2017).

Following 36 h incubation at 37°C, tetracycline was removed from the medium of one well of cells to induce HBV replication whereas the control cells were incubated in the presence of the antibiotic. Following further incubation at 37°C for 7 days, the cells were lysed with LB2 lysis buffer and cytoplasmic and nuclear fraction prepared (see Methods, Section 2.2.19.1) and analysed by Western blot for the presence of HBV core using an in-house generated rabbit polyclonal serum R193 (Clayton, R., 2000, PhD thesis). In parallel, hepG2-hNTCP-C4 cells (Iwamoto et al., 2013) were also transfected with pTHBV-L$^-$ using the lipofectamine 3000 protocol, as described in (Section 2.2.16.1) and incubated at 37°C for 4 days. The transfected cells were then lysed with LB2 lysis buffer and cytoplasmic/nuclear fractions prepared and analysed by Western blot. As shown in Figure 4-12, HBV core was detected predominantly in the cytoplasmic extracts and to a lesser extent in the nuclear extracts of induced cells (i.e. minus tetracycline; lanes 2 and 3, and 5 an 6, respectively) but not in uninduced culture (with tetracycline; lanes 1 and 4, respectively). The small amount of core seen in the nuclear extracts is likely due to contamination of cytoplasmic components. As expected, the core protein was also seen in the cytoplasmic extract of pTHBV-L$^-$ -transfected HepG2-NTCP-C4 cells (lane 7).

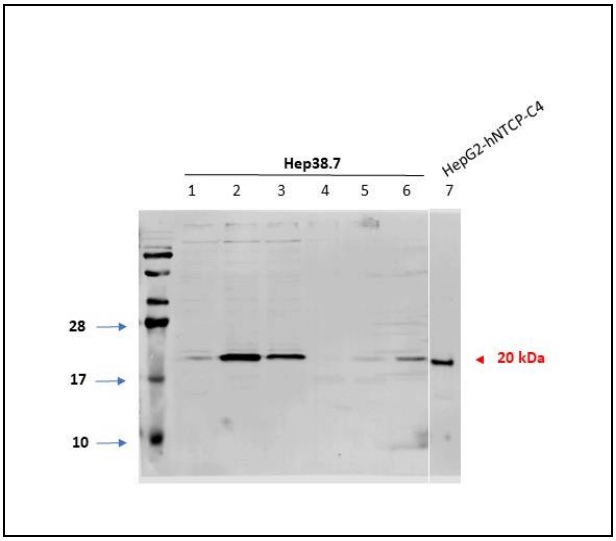


Figure 4-12: Expression of HBV core in HepAD38-Tet cell line.

As described in the text, HepAD38-Tet cells were incubated for 7 days in the presence or absence of tetracycline. Cells were then lysed and their cytoplasmic and nuclear extracts analysed by Western immunoblotting using 1:500 dilution of the polyclonal serum R193. Samples were loaded in the gel as follows; Lane 1 – 25 ul cytoplasmic extract (tet+), Lanes 2 and 3- 25 µl and 25 µl, respectively, of cytoplasmic extract (tet-), Lane 4 – 25 µl nuclear extract (tet+), Lanes 5 and 6 -25 µl and 25 µl, respectively, of nuclear extract (tet-), Lane 7 -25 µl of cytoplasmic extract of pT-HBV-L$^-$ -transfected HepG2-NTCP-C4 cells.

4.2.6 Optimisation of transfection efficiency of HBV plasmids

To obtain optimal conditions for determining viral protein expression by transient transfection assay, Huh7 and HEK293T cells seeded into a 6-well plate and a 24-well plate were tested using lipofection or calcium phosphate transfection protocols. Huh7 cells were transfected with plasmids 1 μg each of pTHBV-L- , pQ-HBV- L- (1-6), and pQ-HBV2.7 (3B4) in a 24-well plate and two plasmids (pQ-HBV-L- (1-6) and pQ-HBV2.7 (3B4)) using the lipofectamine 3000 protocol (see Methods, section 2.2.16.1). HEK-293T cells were transfected with only two plasmids (pQ-HBV-L- (1-6) and pQ-HBV2.7 (3B4)) in a 6-well plate using the Calcium Phosphate transfection protocol, (see Methods, section 2.2.16.2). At 4 d post-transfection, cells were lysed with LB2 lysis buffer, and their cytoplasm extract (CE) analysed on a 15% polyacrylamide gel for SDS-PAGE followed by western blotting to verify the expression of HBcAg using the in-house developed rabbit polyclonal anti-core antiserum R193 (Clayton, R., 2000, PhD thesis). As shown in Figure 4-13, HBV core expression was observed only in Huh7 cells transfected in 6-well dish with pT-HBV-L- and pQ-HBV-L-. There was no core expression observed in HEK-293T cells. Thus, this experiment establishes that pQ-HBV-L- as well as pTHBV-L- are competent for viral protein expression and likely by inference replication with the detection of the core protein. The failure of HBV core expression in HEK-293T cells may be due to lack of viral DNA replication (as would be expected of this cell line).

Another reason for this failure may be due to lack of transfection, although this is unlikely as the Calcium Phosphate method of DNA transfection is highly effective in HEK-293T cells (see figure 4-13). As we had shown that HEK-293T did not support HBV replication with wildtype virus, this result was expected.

.

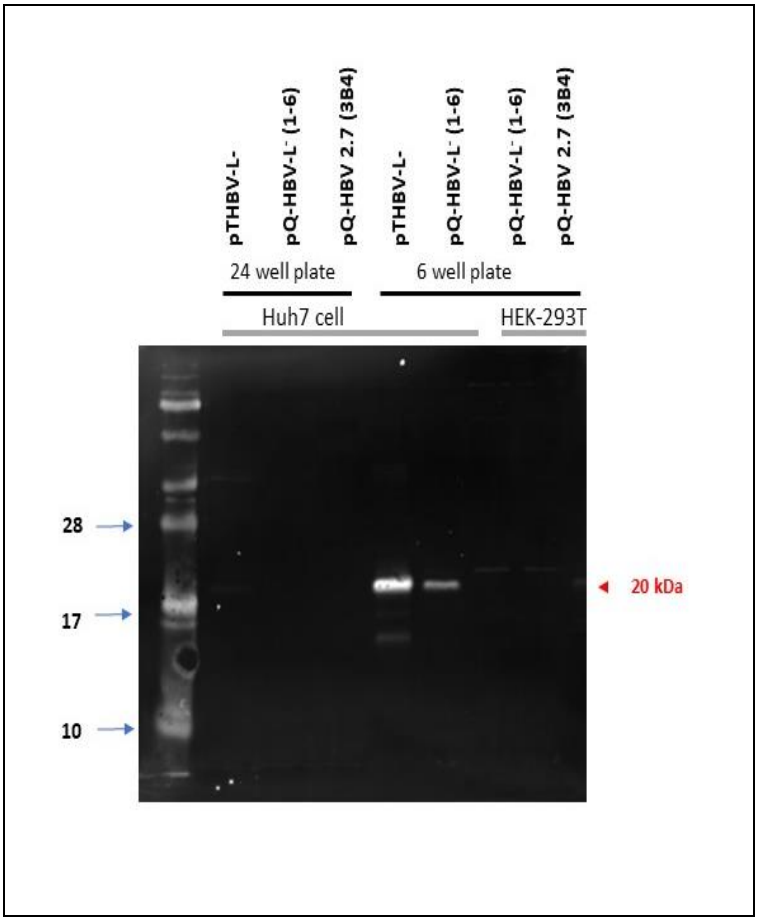


Figure 4-13: Optimisation of DNA transfection efficiency.

Huh7 and HEK-293T cells seeded into a 24-well or 6-well dish were transfected with the indicated plasmids using Lipofectamine 3000 or Calcium Phosphate method, respectively. At 4 d post-transfection, the cytoplasmic extracts of cells were subjected to 15% SDS-PAGE followed by immunoblotting to detect HBV core protein using the anti-core polyclonal antiserum R193 at the dilution of 1:500. The bound primary antibody was recognised by a secondary IRDye® 680RD anti-Rabbit IgG (H + L) Licor (1:2000).

4.2.7 Expression of HBcAg in HepG2-hNTCP-C4 cells by Western blot analysis

We analysed another human hepatoma cell line, HepG2-hNTCP-C4, for expression of HBcAg in a transient transfection assay. These cells were seeded in a 6-well plate and transfected with 5 µg of pTHBV-L$^-$ using the Lipofectamine 3000 protocol. Then, the cells were lysed with LB2 and analysed on a 15% SDS-PAGE to verify the expression of HBcAg in the cytoplasm and the nucleus. HBcAg of approximately 20 kDa was detected strongly in the cytoplasmic extract and weakly in the nuclei extract (Figure 4-14). The latter is likely due to cross-contamination of the cytoplasmic content in the cellular nuclear extract.

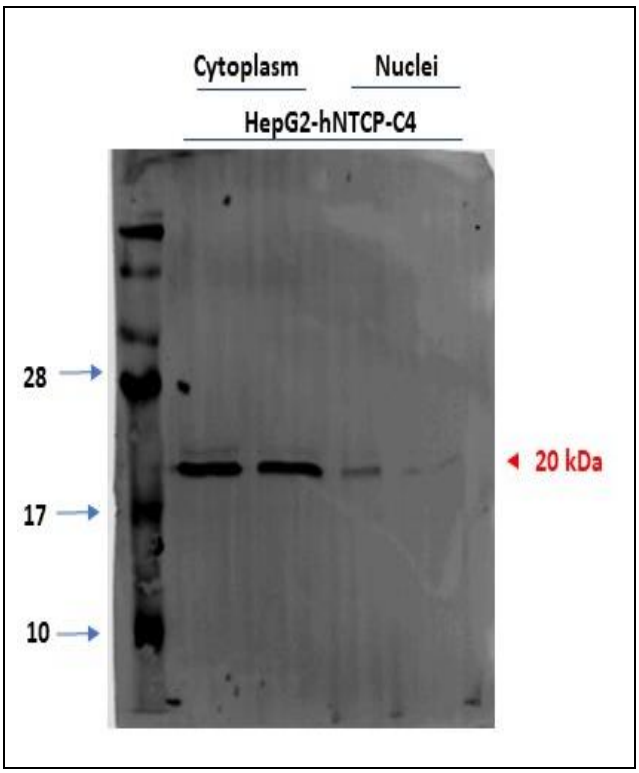


Figure 4-14: Expression of HBcAg in HepG2-hNTCP-C4 cells.

HepG2-NTCP-C4 cells in a 6-well plate were transfected with pTHBV-L$^-$. At 4 days post-transfection, cytoplasmic and nuclear fraction of the cells were prepared and analysed on a 15% Polyacrylamide gel for SDS-PAGE followed by Western blotting to verify the expression of HBcAg using 1:1000 diluted rabbit polyclonal serum R193 adding the R193. The bound primary antibody was detected using a secondary IRDye® 680RD anti-Rabbit IgG (H + L) Licor (1:2000).

4.2.8 Assessment of various primary antibodies for detection of HBcAg expression by Western blot analysis

To compare the relative reactivity of our anti-core polyclonal serum R193 with that of commercial anti-core antibodies, Huh7 cells in a 6-well plate were transfected with 5 ug each of pTHBV-L$^-$ and pQ-HBV2.7 (3B4) using the Lipofectamine 3000 protocol. The cells were lysed with LB2, and the cytoplasmic cell extracts were analysed on a 15% SDS-PAGE followed by Western blotting to verify the expression of HBcAg using R193 (1:1000), another polyclonal antibody PAb cAg Ab1 (Neomarkers Inc (RB-1413-A)), HepB cAg 1-5 (Life Technologies Ltd (MA1-7607); 1:1000), and mouse monoclonal Anti-HepB cAg Antibody (C1-5) (Santa Cruz (sc-23945); 1:1000).

As shown in Figure 4-15, the HBV cAg Ab-1 polyclonal antibody picked up the viral core protein of the same molecular weight (albeit very weakly) as that recognised by R193. As expected, no protein was recognised in cells transfected with the negative control plasmid pQ-HBV-2.7 (3B4) as this plasmid is not expected to express HBV core. The other two antibodies failed to recognise the core protein in transfected cells. Together, these data indicate that R193 is appropriate for use in our studies.

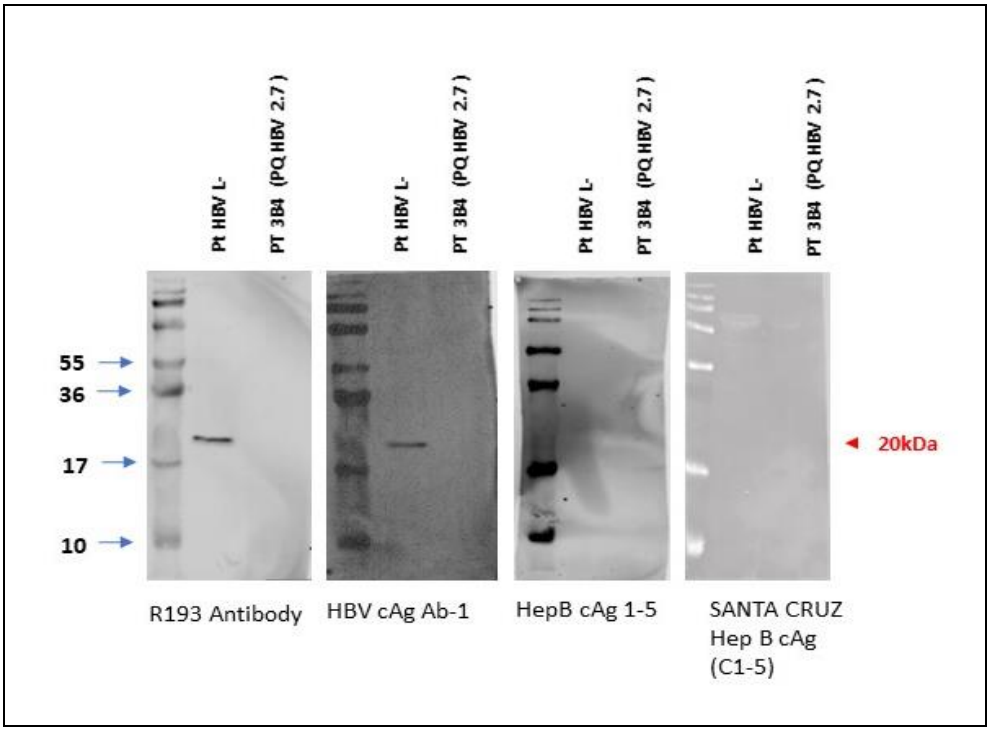


Figure 4-15: Assessment of different anti-HBV core antibody reactivities in a Western blot assay.

Huh7 cells in a 6-well dish were transfected with pTHBV-L$^-$ or pQ-HBV2.7 (3B4). Following incubation at 37°C for 4 days, cells were lysed, and cytoplasmic extracts subjected to 15% SDS-PAGE followed by Western blotting using 1:1000 diluted polyclonal antibodies R193 or HBV cAg Ab-1, or monoclonal antibodies HepB cAg 1-5 or HepB cAg Antibody (C1-5). The bound primary antibodies were detected using secondary IRDye® 680RD anti-Rabbit IgG (H + L) Licor (1:2000) or secondary IRDye® 800CW anti-Mouse IgG (H + L) Licor (1:2000). The 20 kDa band representing HBV core is indicated by an arrow.

4.2.9 Expression of HBcAg and HBsAg in Huh7 cells transfected with different plasmids.

Huh7 cells in a 6-well plate were first transfected with 5 μg each of pTHBV1.3mer WT, pTHBV-L$^-$, pQ-HBV2.7 (3B4), or pQ-HBV1.3mer WT (2B3). Following incubation at 37 ºC for 4 days, cells were lysed and their cytoplasmic fraction subjected to Western blotting using the anti-core antibody R193 and the in-house-developed mouse monoclonal antibody MAb RC28 to the preS1 domain of L-HBsAg (Clayton, R., 2000, PhD thesis). HBV core protein was seen in cells transfected with pTHBV1.3mer WT or pTHBV- L$^-$, but not with pQ-HBV1.3mer WT (2B3) nor with the negative control pQ-HBV2.7 (3B4) (Figure 4-16 A). Probing the Western blot with MAb RC28 revealed the presence of L-HBsAg of the expected molecular weight of ~40 kDa in cells transfected with pTHBV1.3mer WT, pQ-HBV2.7 (3B4), or pQ-HBV1.3mer WT (2B3), but not as expected with the negative control pTHBV-L$^-$ (Figure 4-16 B). It is not clear why there was no core protein seen in cells transfected with pQ-HBV1.3mer WT (2B3) but given that it expressed L-HBsAg it would be reasonable to assume that the plasmid is functional.

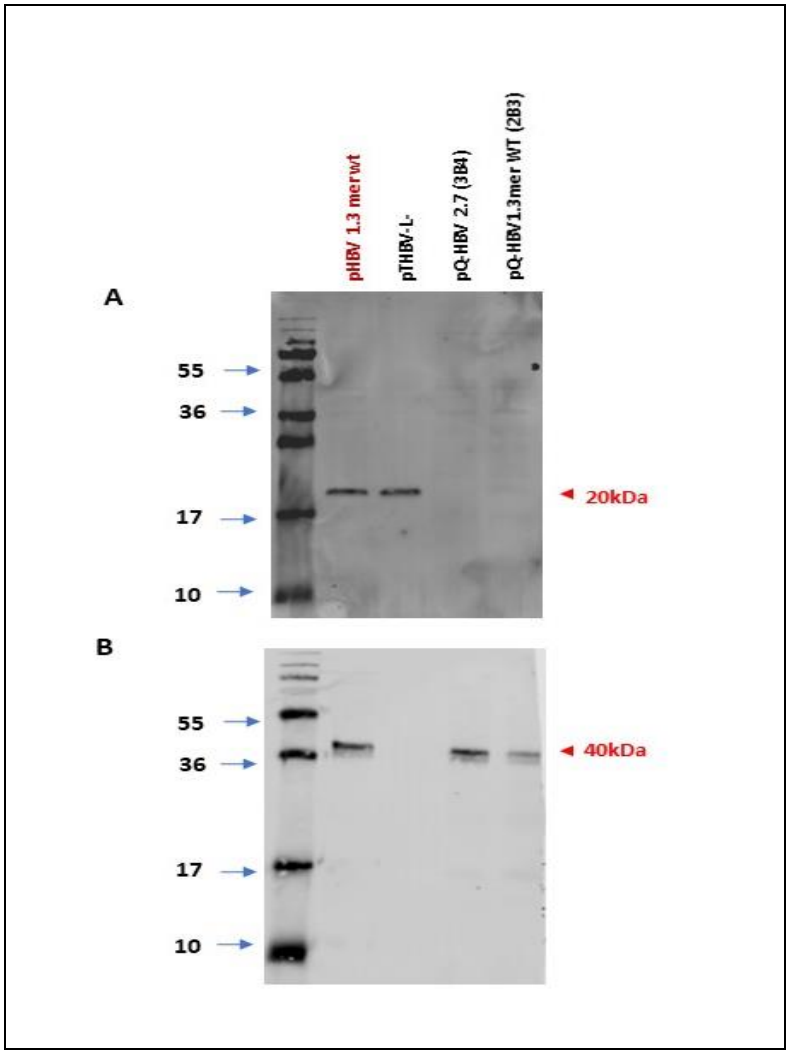


Figure 4-16: Expression of HBcAg in transfected cells.

Huh7 cells transfected with pTHBV1.3mer WT, pTHBV-L$^-$, pQ-HBV2.7 (3B4), or pQ-HBV1.3mer WT (2B3) were incubated at 37°C for 4 days after which they were lysed and the cytoplasmic fraction anlasyed by Western immunoblotting using the (A) 1:1000 diluted anti-core antibody R193 or (B) neat medium of the mouse hybridoma expressing the anti-L-HBsAg MAb RC28. The respective bound antibody was detected using 1:2000 dilution of the secondary IRDye® 680RD anti-Rabbit IgG (H + L) Licor (1:2000) or secondary IRDye® 800CW anti-Mouse IgG (H + L) Licor. The presence of the HbcAg and HBsAg are indicated by an arrow.

Nevertheless, this experiment was repeated, this time also including the plasmid pQ-HBV- L$^-$ Puro (No 11). Furthermore, the plasmid pQ-HBV1.3mer WT (2B3) was linearization with the restriction enzyme NheI to assess if that would improve the levels of HBV proteins. Following transfection of Huh7 cells with 5 µg each of pTHBV1.3mer WT, pQ-HBV1.3mer WT (2B3) (both circular and linearised form), pTHBV- L$^-$, pQ-HBV- L$^-$ /Puro (No 11), or pQ-HBV2.7 (3B4), and incubation at 37°C for 4 days, the cells were lysed and their cytoplasmic fraction subjected to Western blotting using the anti-core antibody R193 and the anti-L-HBsAg MAb RC28. Separately, as loading control, the same amounts of cell lysates were analysed for the

cellular protein γ-tubulin using the monoclonal anti -γ- tubulin antibody produced in mouse (Sigma-Aldrich Co Ltd, T6557). As shown in Figure 4-17 A, the anti-core antibody R193 identified core protein the cells transfected with pTHBV1.3mer WT, pTHBV- L^-, or pQ-HBV-L^-/Puro (No 11), but not with the negative control pQ-HBV2.7 (3B4). Again, no core expression was seen in cells transfected with both the circular and linear forms of pQ-HBV1.3mer WT (2B3). As expected, L-HBsAg expression was seen in cells transfected with pTHBV1.3mer WT and pQ-HBV2.7, but not with pTHBV- L^-, or pQ-HBV- L^-/Puro (No 11) (Figure 4-17 B). Furthermore, the L-HBsAg was also present in cells transfected with circular pQHBV1.3mer WT (2B3) but not with the linearised form of the plasmid (Figure 4-17 B). Western blot analysis using the anti-γ-tubulin MAb showed that this host protein was present in all extracts in comparable amounts (Figure 4-17 C).

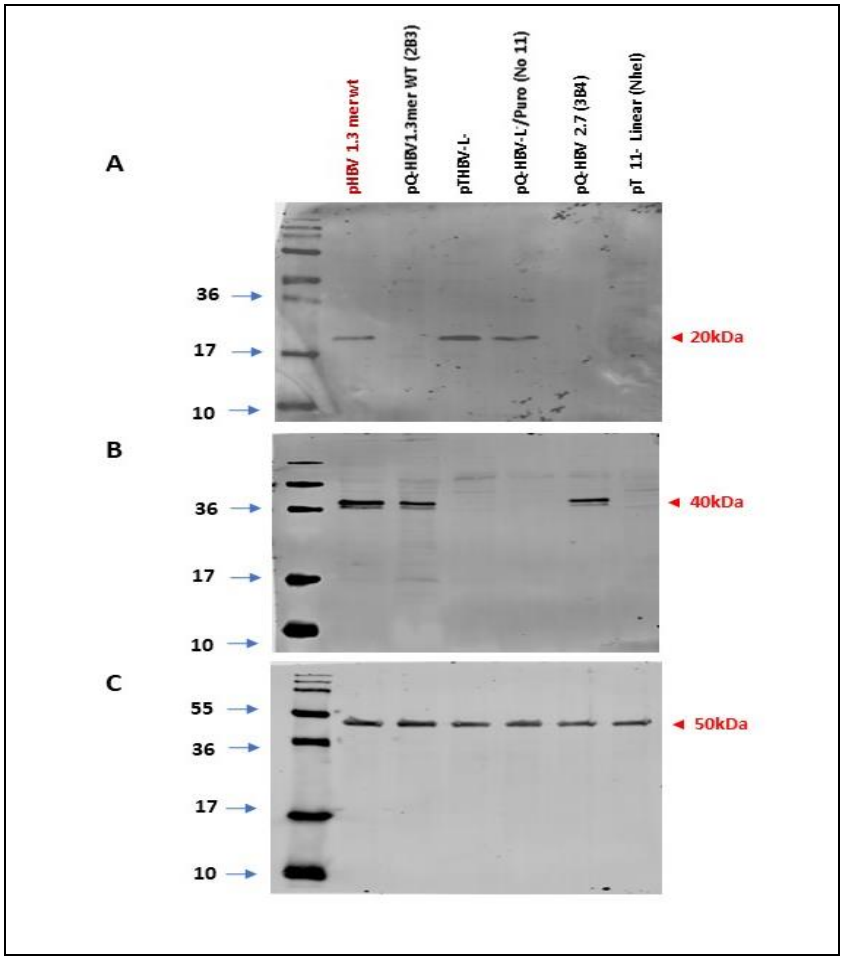


Figure 4-17: Expression of HBcAg and HBsAg in transfected cells.

Huh7 cells transfected with pTHBV1.3mer WT, pQHBV1.3mer WT (circular), pTHBV-L$^-$, pQ-HBV-L$^-$/Puro (No 11), pQ-HBV2.7 (3B4), or linearised pQ-HBV1.3mer WT (2B3) were incubated at 37°C for 4 days after which they were lysed and the cytoplasmic fraction analysed by Western immunoblotting using the (A) 1:1000 diluted anti-core antibody R193, (B) neat medium of the mouse hybridoma expressing the anti-L-HBsAg MAb RC28, (C) 1:1000 diluted anti-*γ-tubulin MAb*. The respective bound antibody was detected using 1:2000 dilution of the secondary IRDye® 680RD anti-Rabbit IgG (H + L) Licor (1:2000) or secondary IRDye® 800CW anti-Mouse IgG (H + L) Licor. The presence of the HBcAg, HBsAg or *γ-tubulin* is indicated by an arrow.

4.2.10 Analysis of HBcAg and HBsAg in transfected Huh7 cells by immunofluorescence microscopy

Huh7 cells in a 24-well dish were transfected with 1 µg each of pTHBV L$^-$, pQ-HBV L$^-$ /Puro (1-6), or pQ-HBV L$^-$ /Puro (1-6) Linear (*EcoRV*-HF®), using the Lipofectamine 3000 protocol. The cells were probed with the rabbit anti-core R193 antibody (dilution 1:1000), followed by a goat polyclonal Alexa Fluor® 488-labelled secondary anti-Rabbit IgG - H&L (Abcam*)*. Separately, *c*ells were transfected with 1 ug each of pQ-HBV 2.7(3B4), pQ-HBV2.7 (3B4) Linear (*EcoRV*-HF®), or phCMV-L (a plasmid expressing L-HBsAg previously generated in our lab) using the Lipofectamine 3000 protocol. They were probed with the anti-L-HBsAg MAb RC28 (dilution 1:1000), followed by goat polyclonal Alexa Fluor® 488-conjugated secondary anti-mouse IgG- H&L *(*Abcam*)*. As shown in Figure 4-18 A and B, all constructs expressed the relevant protein as expected.

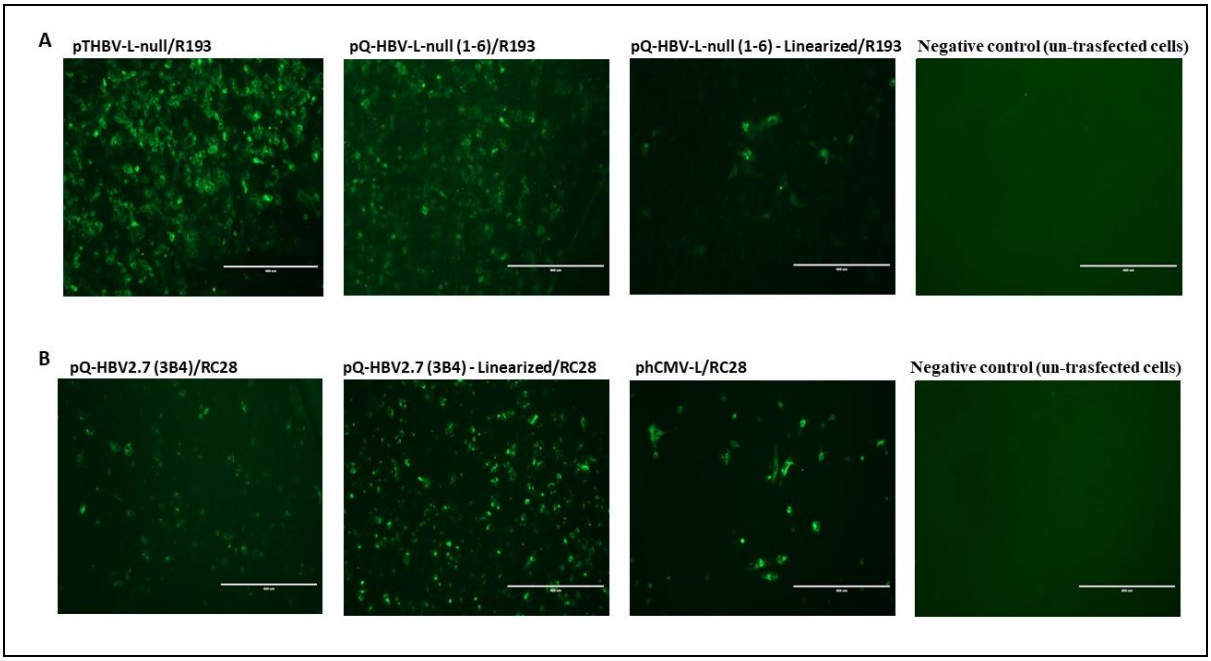


Figure 4-18: Expression of HBV proteins as detected by immunofluorescence.

Huh7 cells on coverslips in 24-well dish were transfected with various plasmids as shown. Following incubation at 37°C for 3 days, cells were fixed with ice-cold methanol, washed and the probed with antibodies as follows: (A) Cells transfected with pTHBV-L$^-$, pQ-HBV L$^-$ /Puro (1-6), and pQ-HBV L$^-$ /Puro (1-6) Linear (*EcoRV*-HF®) were probed with 1:100 dilution of the anti-core antibody R193 antibody followed by Alexa Fluor® 488-conjugated goat polyclonal secondary antibody to *r*abbit IgG - *H&L*. (B) Huh7 cells transfected with pQ-HBV 2.7(3B4), pQ-HBV2.7 (3B4) Linear (*EcoRV*-HF®), phCMV-L as above and then probed with 1:1000 diluted anti-L-HBsAg MAb RC28 followed by Alexa Fluor® 488-conjugated goat polyclonal secondary antibody to mouse IgG- H&L. The probed cells were analysed and image at 10x magnification on EVOS M5000 Imaging System.

4.2.11 Introduction to HDV as a model for HBV infection

Hepatitis delta virus (HDV) is an excellent model for studying the early events leading to HBV infection. HDV is a satellite virus that uses HBsAg envelope proteins produced by co-infecting HBV to spread and cause disease in patients. As mentioned earlier, previous work with HBV and HDV (Yan et al., 2012) identified NTCP as the viral receptor for both viruses. Using this knowledge, we tried to generate the surrogate HDV infection model with a view to test a small panel of HBV preS1-specific MAbs that had previously been developed in our laboratory for their potential to neutralise HBV infection of cultured cells.

4.2.11.1 Reverse genetics and rescue of HDV

To produce the HDV particles, two commercial plasmids were acquired from Addgene: pSVL (D3) (Addgene plasmid # 29335) and pSVL (D2m) (Addgene plasmid # 29336). pSVL (D3) contains a head-to-tail trimer of full-length HDV cDNA, whereas pSVL (D2m) contains a dimer of mutated L-HDAg sequence encoding only S- HDAg, and thus is not able to form HDV particles as the L-HDAg is essential for assembly. These plasmids were directly transfected into Huh7-3B4 stable cell line expressing the HBV L-HBsAg that developed previously by my supervisor. The plasmid used to develop this cell line, pQHBV2.7 (3B4), the generation and characterisation of which was performed as described in this Chapter. In addition to using this stable cell line, attempts were also made to generate HDV particles by co-transfection of human hepatoma cells with the pSVL plasmids and pQHBV2.7. In both cases, we expected to see HDV particle generation. Transfection was performed using Lipofectamine 3000 protocol and cells were maintained at 37 ˚C with 5% CO2 for 12 days. Every three-days, medium was collected and replenished with the fresh medium. The collected medium was used to test the presence of infections surrogate HDV particles as described below.

It should be noted that normal hepatoma cell lines such as Huh7 or HepG2 are not susceptible for HBV entry on account of the lack of NTCP receptor in them. Therefore, to establish the HDV entry model we had obtained the HepG2-NTCP-C4 cells from Masashi Iwamoto and Koichi Watashi (Iwamoto et al., 2017). Furthermore, my supervisor had also generated a polyclonal Huh7, and a panel of clonal HepG2 cell lines stably expressing the human NTCP. Some of these cell lines as described were used in this study to establish a working HDV infection model.

4.2.11.2 Infection of cell lines by rescued surrogate HDV virus

To test the generation of infectious HDV, the supernatant collected at day 12 of transfected cells was concentrated in presence of 4% PEG8000 as described by Iwamoto et al. (2013). The PEG-precipitated particles were resuspended in 3 ml of medium, and 500 ul of this was used to infect four HepG2 cell lines expressing the NTCP receptor (G2N1, G2N7, G2N15, and G2N21) and the Huh7-NTCP (7N) cell line. Huh7 was also infected as a negative control as it lacks NTCP. Medium from infected cells were removed after 16 hr, and cells were washed and incubated further in for 7 days at 37˚C. The cells were then methanol-fixed, washed with PBS and probed with 1:5000 diluted rabbit polyclonal antiserum to HDV delta antigen (kindly provided by Masashi Iwamoto, 2019). The bound antibody was detected using a secondary anti-rabbit (Alexa Fluor 488 Goat anti Rabbit IgG (H+L)) at a dilution of 1/2000 and the cells analysed by immunofluorescence on EVOS M5000 Imaging System. As expected, no HDV delta antigen was detected in Huh7 cells (Figure 4-19). This will require further confirmation with an appropriate positive control for HDV, we only had a negative control. In contrast, some fluorescent cells were observed in cell lines bearing the NTCP receptor such as the representative cell line G2N7 shown in Figure 4-20. It should be noted that the stable Huh7-3B4 transfected with pSVL (D3) or pSVL (D2m) failed to produce HDV particles as no infection was observed from the inoculum derived from it.

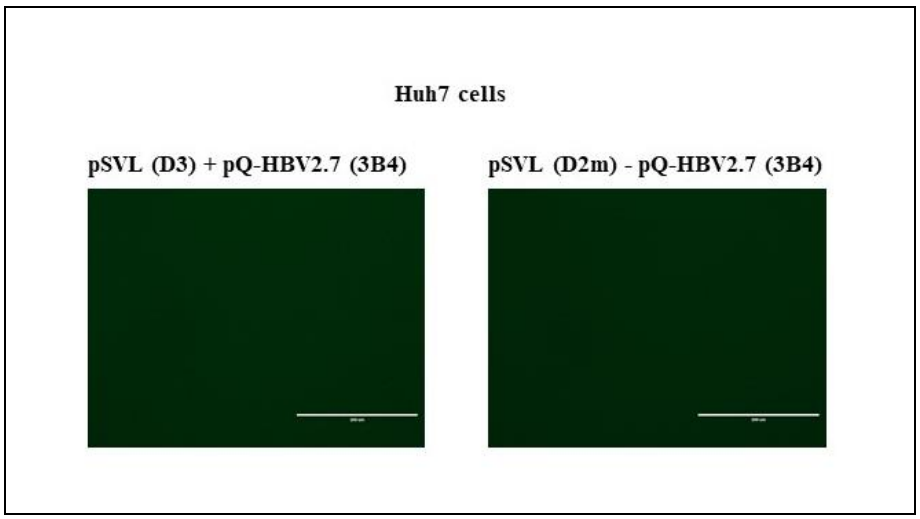


Figure 4-19: Infection of Huh7 cell line with HDV particles.

The Huh7 cells were infected with HDV inoculum derived from Huh7 cells co-trasnfected with pQHBV2.7 (3B4) and pSVL (D3) or pSVL (D2m). At 7 days post-infection cells were fixed and probed with the anti-HDV delta antigen antibody (1:5000) followed by Anti-Rabbit IgG (H+L) (Alexa Fluor® 488) (Life Technologies), no HDV delta antigen was detected in Huh7 cells.

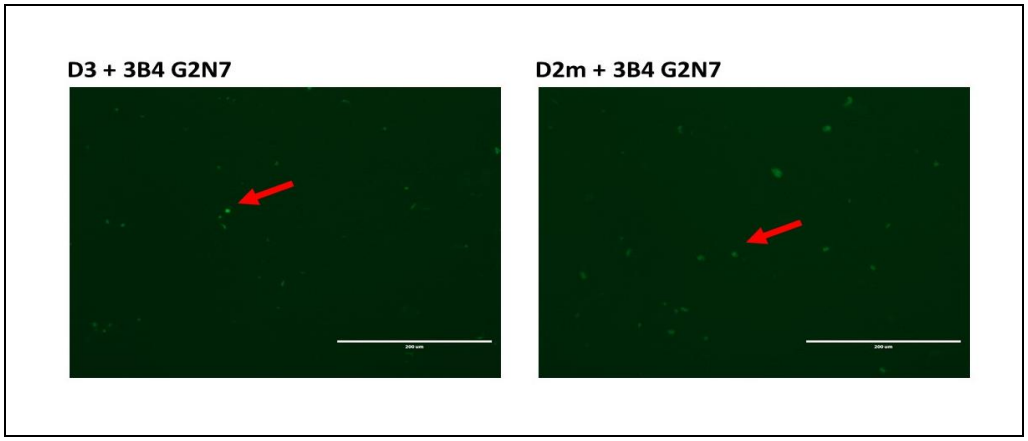


Figure 4-20: Infection of NTCP-expressing HepG2 cell line G2N7 with HDV particles.

The G2N7 cells were infected with HDV inoculum derived from Huh7 cells co-trasnfected with pQHBV2.7 (3B4) and pSVL (D3) or pSVL (D2m). At 7 days post-infection cells were fixed and probed with the angti-HDV delta antigen antibody (1:5000) followed by Anti-Rabbit IgG (H+L) (Alexa Fluor® 488) (Life Technologies). The above Immunofluorescence are of G2N7 cells infected with rescued virus derived from pSVL (D3) + pQ-HBV2.7 (3B4) or pSVL (D2m) + pQ-HBV2.7 (3B4) co-transfection, or transfection of pSVL (D2m) and pSVL (D3) in Huh7-3B4 stable cell line.

After analysis of the pictures, it was decided that pSVL (D3) + pQ-HBV2.7 (3B4) co-transfection produced the highest HDV titre. This combination was then used to infect an additional group of 8 NTCP-expressing cell lines along with HepG2 and Huh7.

The immunofluorescence shown by the cell lines is shown in figure 4-21. The panel cell lines included a negative control, Huh7 cells. As expected, this cell line did not show any fluorescence as it lacks an adequate receptor. In comparison, some immunofluorescence was detected in HepG2-hNTCP-C4 and HepG2, suggesting these cell lines support the HDV viral replication (Figure 4-21).

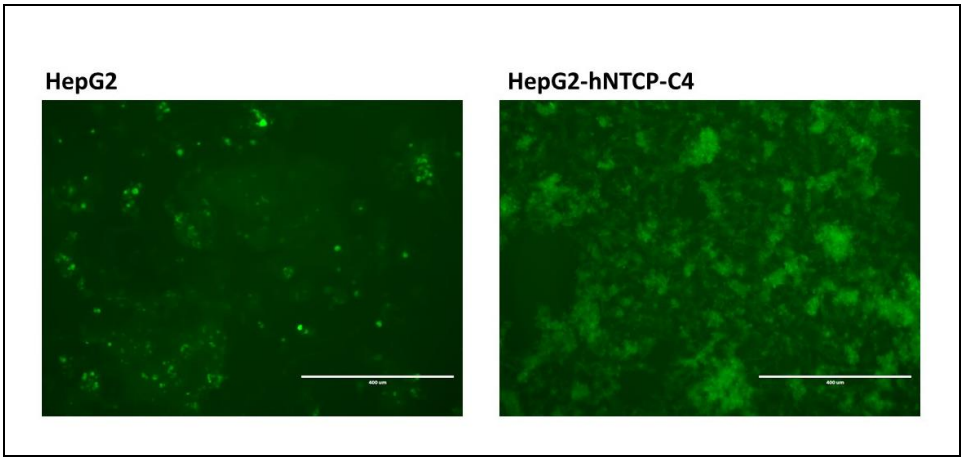


Figure 4-21: Immunofluorescence derived from potential HDV infection rescued by pSVL (D3) + pQ-HBV2.7 (3B4) co-transfection.

The panel of cells were infected with HDV inoculum derived from Huh7 cells co-trasnfected with pQHBV2.7 (3B4) and pSVL (D3). At 7 days post-infection cells were fixed and probed with the angti-HDV delta antigen antibody (1:5000) followed by Anti-Rabbit IgG (H+L) (Alexa Fluor® 488) (Life Technologies). Scale bars represent 400 µm.

The results shown in these experiments suggest that the generation of HDV is possible by co-transfection of pSVL (D3) and pQ-HBV2.7 (3B4) plasmids in the Huh7 cell line. The rescued virus particles were concentrated using PEG8000 and then used to infect a series of cell lines containing the NTCP receptor. The presence of infectious HDV was tested in immunofluorescent studies in various cell lines. The generation of HDV can be useful for the indirect investigation of HBV entry process. In the future, this HDV system could be used to develop a drug screening assay that could introduce new drugs against HBV. Furthermore, it is also amenable for testing anti-HBV neutralising antibodies. Indeed, our intention was to test our in-house-developed anti-HBV preS1 MAbs for their virus neutralisation potential as their epitopes have previously been mapped to the region in preS1 that has been proposed to be involved in virus entry (Clayton, R., 2000, PhD thesis). However, due to COVID-19-related lock-down this work could not be performed.

Taken together, these data confirm that all the plasmids constructed in this study are functional in that they produce viral proteins as expected upon transfection into human hepatoma cells and as such they are amenable to future molecular studies on aspects of the HBV life cycles. Indeed, my intention was to perform some of these studies during the remainder of my PhD project. Specifically, the intention was to further develop and explore the HDV surrogate entry system and the HepaAD38-Tet- and HepG2-NTCP-C4-based models to investigate the possible potential of RC28 and related anti-preS1 MAbs that had previously been developed in our laboratory (Clayton, R., 2000, PhD thesis) to neutralise HBV. Furthermore, the plasmids

developed in this project would have been excellent tools to study aspects of HBV replication and cccDNA formation. However, sadly this was not possible due to the imposition of COVID-19 pandemic-related country-wide lock-down. Given that the relaxation of the restrictions came towards the tail-end of my PhD project, it was decided to use the short time available to focus only on establishing, as a proof-of-concept, a cell culture model based on the plasmid pQ-HBV- L^-/Puro (No 11) to evaluate the drug resistance mutations in the RT domain of the viral polymerase that were identified in our cohort of Saudi Arabian patients chronically infected with HBV (Chapter 3). Towards this, we first developed an in-house HBeAg ELISA to quantitate HBeAg as a surrogate marker of HBV replication in cells transfected with pQ-HBV- L^-/Puro (No 11). This is described in the following section.

4.2.12 Introduction to ELISA assay for HBeAg quantitation

HBeAg is a useful marker for the study of specific viral stages, such as viral DNA replication and cccDNA formation in *in vitro* infection systems. Therefore, an assay allowing rapid quantitation of HBeAg secreted into the medium of cells replicating HBV DNA would be ideal to evaluate drug resistance conferred by mutant viral genomes.

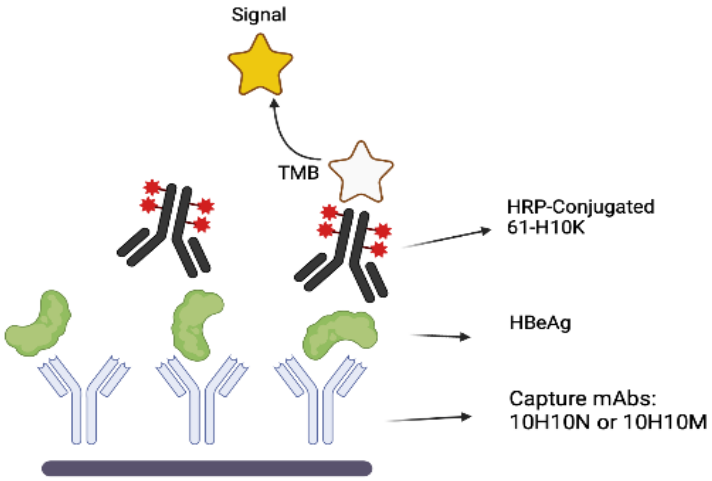
Commercial ELISA kits for the recognition of HBeAg are available, however they are expensive and not easily amenable to large-scale screens. Instead, to test for the expression and secretion of HBeAg, a cost-effective in-house ELISA was developed. The ELISA was validated using commercially available HBeAg for the recognition of the viral antigen. Moreover, a commercially available ELISA kit was acquired and directly compared with our in-house developed ELISA, to provide further validation of our newly made assay.

4.2.12.1 Design and development of in-house ELISA for HBeAg

The in-house HBeAg ELISA format was essentially based on protocols described by Morey J. D. (2004); Goddard et al. (2017); and Guo, H. and Cuconati, A. (eds.). These protocols were customised and optimised as described below.

Two anti-HBeAg mouse monoclonals antibodies were commercially acquired to compare their binding affinity to capture HBeAg. As a detection antibody, an anti-HBeAg mouse monoclonal, conjugated with horseradish peroxidase (HRP) (Fitzgerald industries), was employed. The addition of an HRP-conjugated anti-HBeAg antibody alleviated the use of a secondary anti-mouse antibody that would serve to detect the HBeAg-bound primary antibody, which is also of mouse origin and therefore cause a high background signal. The assay design is presented in (Figure 4-22).

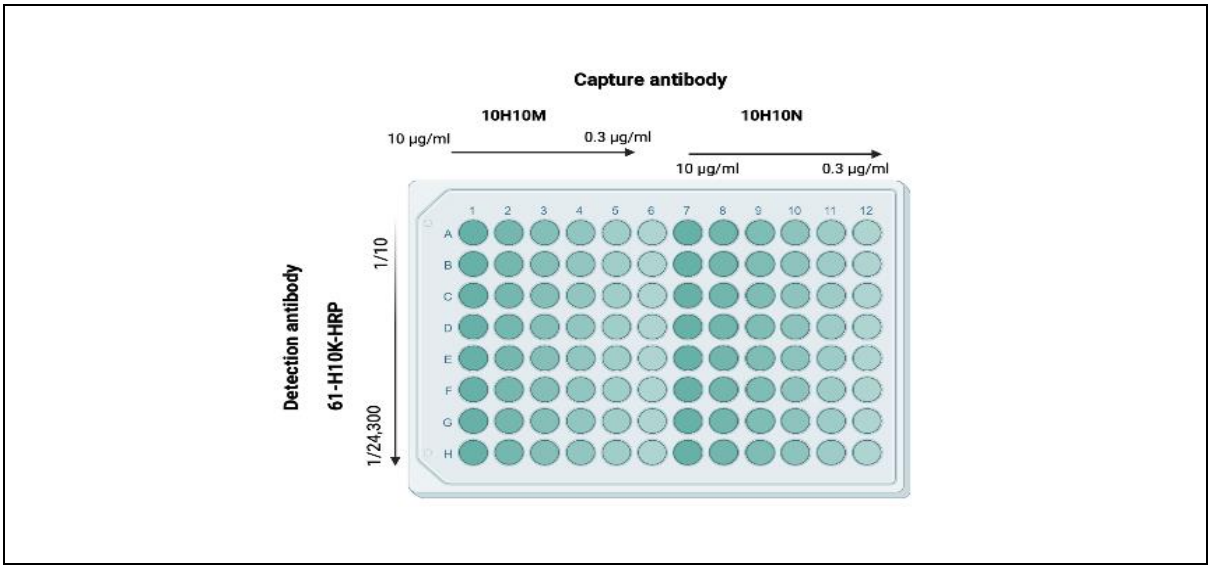
Figure 4-22: Design of ELISA for the detection of HBeAg



To test the designed ELISA, capture antibodies (mouse monoclonal HBeAg antibody, anti-HBeAg (10H10N) Fitzgerald Industries or mouse monoclonal HBeAg antibody, anti-HBeAg (10H10M) Fitzgerald Industries) were used to coat a plate, then recombinant HBeAg was added followed by an HRP-conjugated 61-H10K detection antibody. A signal was generated upon the addition of substrate 3,3′,5,5′-tetramethylbenzidine (TMB; Sigma-Aldrich).

To test and validate the ELISA, a recombinant HBeAg was acquired from Cell Biolabs, Inc. The antigen was used to test the binding affinity of the two commercially acquired detection antibodies. To identify the best detection antibody for the development of the ELISA, the capture antibodies, mouse monoclonal HBeAg antibody, anti-HBeAg (10H10N) Fitzgerald Industries and mouse monoclonal HBeAg antibody, anti-HBeAg (10H10M) Fitzgerald Industries, were incubated in a 96-well plate in six point 2-fold serial dilutions ranging from 10.0 to 0.125 µg/ml in PBS, overnight at room temperature. The following day, the plate was blocked with 5% non-fat skimmed milk in PBST for 2 hours at room temperature. After plate washing with PBST, 100 µl of recombinant HBeAg was added at a concentration of 5 ng/ml for 2 hours at room temperature. To test the optimal concentration of the detection antibody, mAb 61-H10K-HRP, three-fold serial dilutions of the antibody (ranging from 1/100 to 1/24,300) were added to the plate. The plate design to test the optimal concentration of both the capture and detection antibodies is shown in (Figure 4-23). For assay development, the plate was washed with PBST prior to the addition of TMB as a substrate, followed by sulfuric acid (H2SO4) to stop the reaction. The optical density was measured at 450nm, using a Varioskan plate reader.

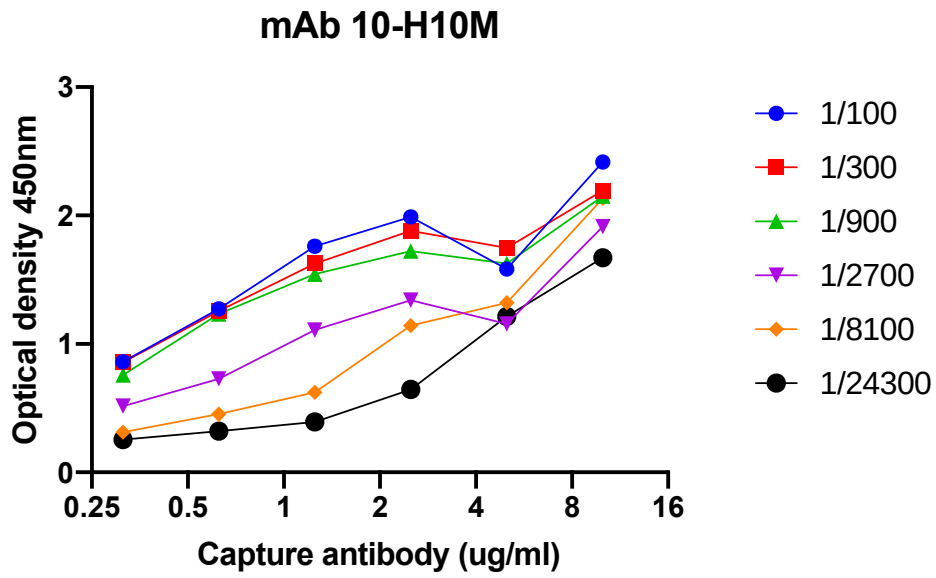
Figure 4-23: ELISA plate design to determine optimal antibody concentration



To find the optimal concentration of the capture and detection antibody, a variety of combinations as described in the text were tested using different concentrations of antibodies to detect a fixed concentration (5 ng/ml) of recombinant HBeAg.

To determine the optimal concentration of primary and secondary antibody to use, the optical density values shown in (Figure 4-24) were analysed. The ideal concentrations should produce a positive signal of close to 2 and a background that is ten times lower (0.2). As shown in (Figure 4-24A), using a dilution of 1/8100 of the HRP-conjugated detection antibody together with 10 µg/ml of the capture antibody mAb 10-H10M produced a strong signal of 2. As the capture antibody concentration is decreased, which represents the interpretation of a sample expressing less antigen, the optical density is decreased proportionally. It was found that the capture mAb 10-H10M paired with mAb 61-H10K-HRP produced a stronger signal than the other capture mAb 10-H10N (Figure 4-24B). Therefore, it was concluded that the optimal concentration of antibody to use was 10 ug/ml for the capture antibody mAb 10-H10M and 1/8100 for the HRP-conjugated antibody.

A



B

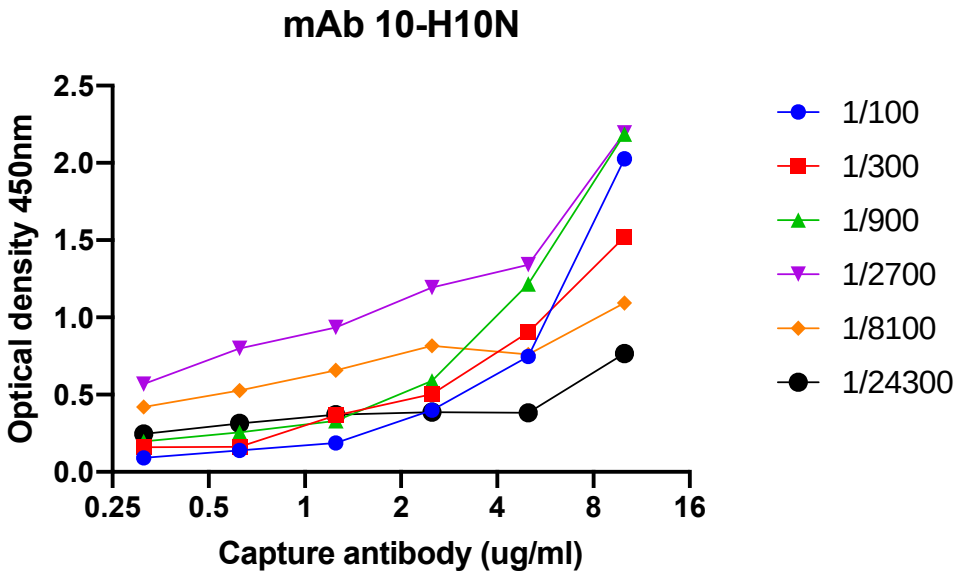


Figure 4-24: Optimization of the Antibodies concentration used in ELISA to screen secretion of HBeAg.

4.2.12.2 Assay validation and plasmid screening

The in-house ELISA here developed was used to test the expression of HBeAg by the HBV genome-carrying plasmids described above. Huh7 cells were transfected with 5 ug each of pHBV1.3mer WT and pQ-HBV2.7 (3B4), using Lipofectamine 3000, and the cells incubated at 37°C for four days. Then, the medium was collected and adjusted with the buffer LB2 before being serially diluted and tested for the presence of HBeAg using the ELISA described above. Medium of untransfected cells was used as a negative control, whereas HBeAg antigen (Cell Biolabs) was use in dilution range of 2.5 μg/ml to 20 ng/ml in a 2-fold dilutions as a positive control to generate a standard curve to quantify the antigen in the sample. As shown in (Figure 4-25 A and B), the HBeAg antigen formed a concentration-dependent standard curve, which also validates the assay here developed. Transfection with pHBV1.3mer WT also showed a strong secretion of HBeAg, reaching a similar pattern to the standard curve. In contrast, there was no evidence of HBeAg in the medium of cells transfected with plasmid pQ-HBV2.7 (3B4) which was expected as this plasmid does not express the protein. In summary, the ELISA here developed was validated using a standard antigen curve and used to screen the secretion of HBeAg in two plasmids in Huh7 cells. By comparing the absorbance value of the plasmid 1.3mer with the standard curve, it is possible to calculate the amount of HBeAg expressed after cell transfection. The neat supernatant of the transfected cells produced an absorbance unit of approximately 3. This roughly corresponds to the signal produced by the standard curve antigen at a concentration of 2.5 ug per ml.

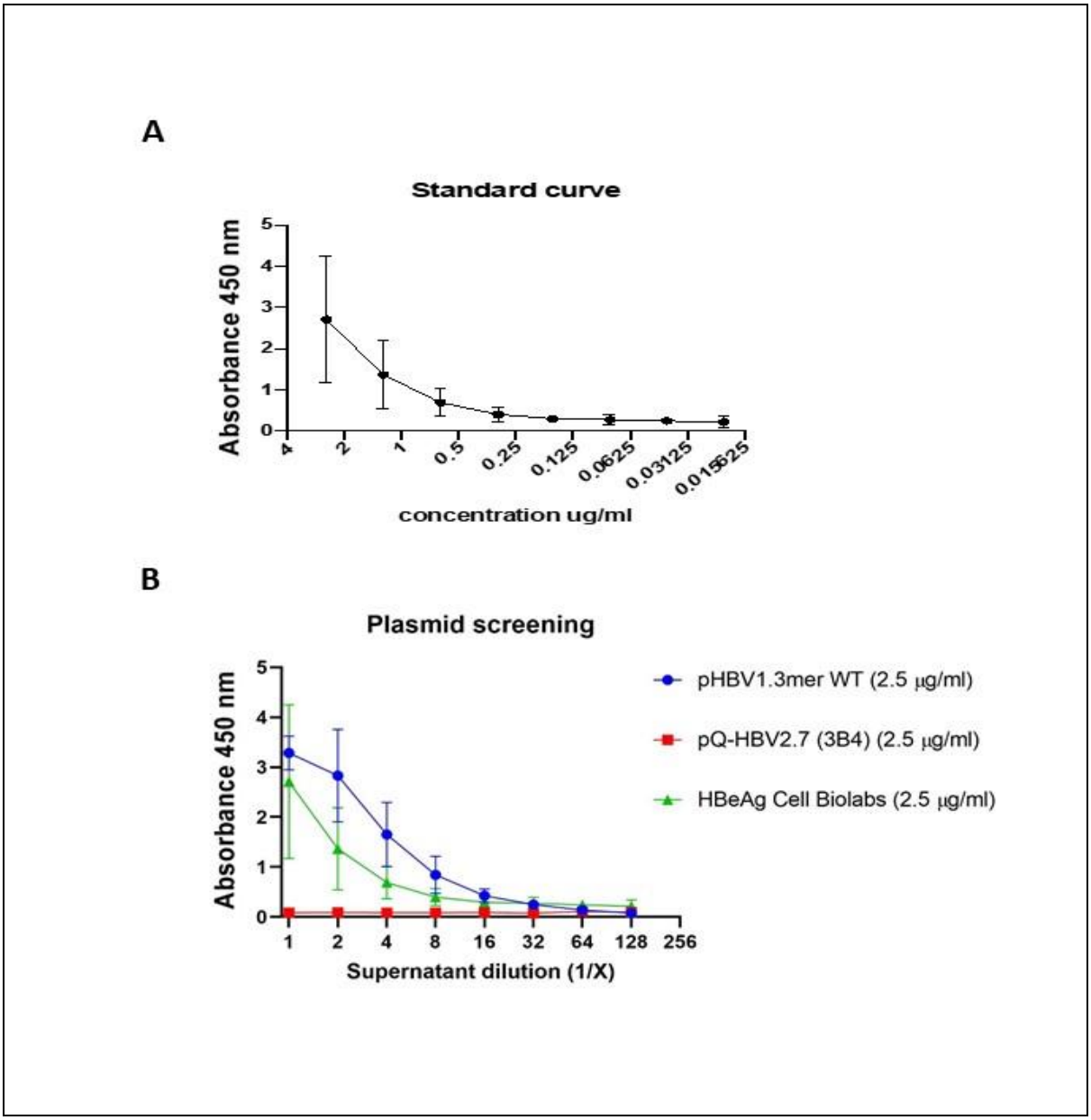


Figure 4-25: ELISA validation and plasmid screening.

A) positive control standard curve of HBeAg with known concentrations on the x-axis. B) The ELISA was used to screen the expression of HBeAg by plasmids pHBV1.3mer WT and pQ-HBV2.7 (3B4). A standard antigen curve, obtained from Cell Biolabs, was used as a control of HBeAg shown in the top figure. Error bars represent standard deviation from two independent experiments performed in duplicate.

4.2.12.3 Commercial kit for the analysis screening of additional plasmids

A commercially available ELISA kit was used to screen the HBeAg expression of another set of plasmids for the expression and secretion of HBeAg, including pHBV1.3mer WT, pTHBV-L$^-$, pQ-HBV L$^-$ /Puro (11), and pQ-HBV2.7 (3B4). Similar to the results obtained from the in-house assay (Figure 4-19), pHBV1.3mer WT showed a positive signal that was inversely proportional to the dilution of the supernatant, as shown in (Figure 4-26). For plasmid pQ-HBV2.7 (3B4), there was no evidence of a positive signal at any supernatant dilution. In addition, the pTHBV- L$^-$ and pQ-HBV L$^-$/Puro (11) both expressed HBeAg in the medium in a dose-dependent fashion. Together, these data confirm that the commercial kit, although expensive, was useful for the screening of HBeAg expression upon transfection.

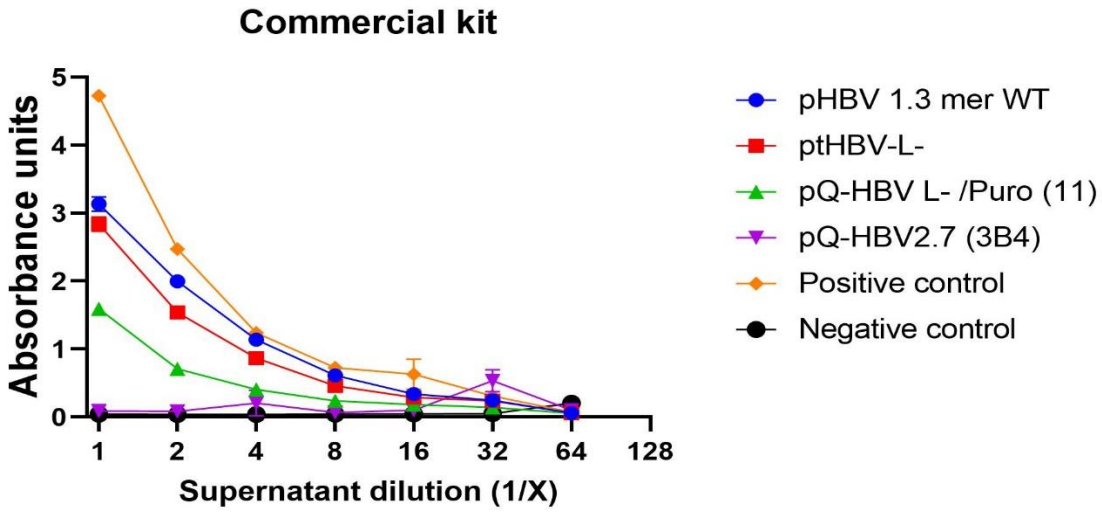


Figure 4-26: Secretion of HBeAg in transfected cells as detected using a commercial ELISA kit.

The secretion of the plasmids was tested using a commercially available ELISA kit from Cusabio. A positive and a negative control were included as instructed and provided by the manufacturer. Error bars represent the standard deviation of one experiment done in duplicate. Absorbance at 450nm was measured in an ELISA plate reader.

4.2.12.4 Direct comparison between commercial kit and in-house developed ELISA

To understand if the data obtained from the recently developed in-house ELISA provides reliable and accurate data, it was compared with the commercial kit. Commercially available ELISA kits are not cost effective as it is only possible to test a limited number of samples and the reagents are not usually replaceable.

To perform this comparison, medium Huh7 cells transfected with pHBV1.3mer WT and pQ-HBV2.7 (3B4) from the experiment described in Figure 4-19 was used in a 3-fold dilution series. Hundred ul of the diluted samples were applied to wells of the commercial ELISA plate or to those coated with our capture antibody mouse monoclonal HBeAg antibody, anti-HBeAg (10H10M) Fitzgerald Industries. The assay was then performed as per the manufacturer's instruction (commercial kit (Cusabio)) or as described above for our in-house assay. As shown in Figure 4-27, dose-dependent antigen expression was seen with both formats in the medium of pHBV1.3mer WT-transfected cells, but not in in those transfected with pQ-HBV2.7 (3B4). Interestingly, a higher signal was detected in our assay as compared to the commercial one.

In summary, the in-house ELISA developed in this project is equally effective as commercially available ELISAs for the screening of plasmids encoding the expression of HBeAg. It is possible that this assay may also be adapted for other uses, such as diagnostics and neutralisation assays. The development of the in-house ELISA is an invaluable tool for the study of HBV, particularly to test the expression of HBeAg in transfected cells. As HBeAg is a marker of virus replication, the use of it as the basis of our ELISA allows it be used as a proxy for the replicon. This assay will be later used for the virus-specific inhibition of antigen expression upon the addition of drugs. Importantly, mutations in HBV may confer antiviral resistance. However, the characterisation of such mutations in the context of viral inhibition can be assessed using transfection coupled with this ELISA. Together, this removes the need to work with wild type viruses in category 3 laboratories.

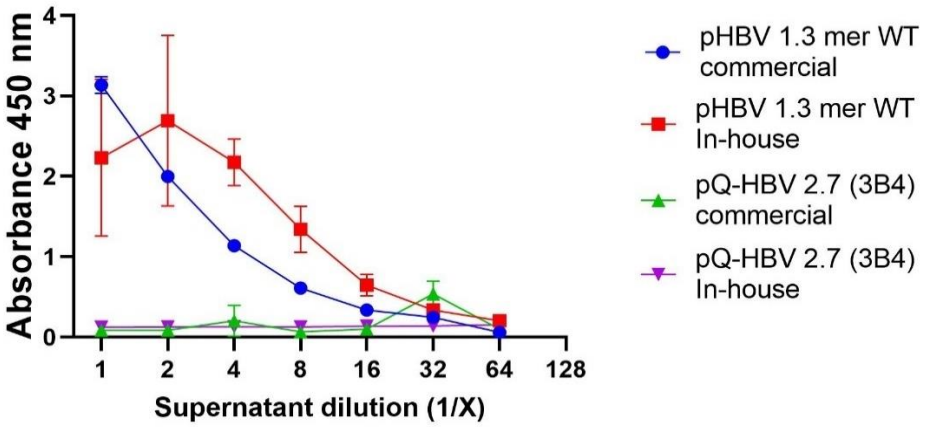


Figure 4-27: Comparison between commercial and in house ELISA assay.

The commercial kit was performed as instructed by the manufacturer. The in-house ELISA was developed using a concentration of 10 ug/ml capture antibody H-10M and 1/8000 concentration of HRP-conjugated detection antibody. Absorbance at 450nm was measured in an ELISA plate reader.

4.3 DISCUSSION

In this chapter, we aimed to develop and validate HBV replicon system to enable the investigation of drug resistance at BSL2 rather than BSL3. Currently, the study of HBV requires BSL3 facilities, which can be time consuming, require specialised training and facilities and does not necessarily allow for high throughput of drug screens. The approach used in this thesis enable the successful generation of a replicon system with the output measured using two different viral antigens (E and core), providing a robust measure of replicon. Indeed, this system may also enable the different stages of the HBV genome replication to be studied. We also attempted the development of a HDV replicon system, while the initial development and data appeared promising, the SARS-CoV-2 (COVID-19) outbreak and subsequent lockdown stopped/ hampered further validation of this new approach.

The replication cycle of HBV remains poorly understood, partially due to the limitations of the existing in vivo and in vitro models. In recent years, the development of new cells and combination of new technologies have made it possible to characterise more effectively the replication cycle of HBV and its interaction with the host's antiviral immune system (Chen et al., 2015). Advances in understanding the details of virus replication and the mechanisms involved in the insufficient induction of anti-HBV immune responses have guided the development of strategies aimed at achieving a functional cure for HBV infection (Chen et al., 2015).

The existing HBV cell culture systems have performed a crucial role in reviewing the pathogenesis of HBV infection, immune mechanisms, and screening of HBV antiviral drugs. The different cell culture systems have made a considerable impact on the research of the HBV infection process and the development of HBV vaccine (Xu et al., 2021). A caveat of these systems is that the studies are reliant on different strains which may influence or unduly bias the results particularly when they contain mutation that may provide a replicate advantage or contain uncharacterised drug resistant mutant.

This Chapter describes the successful development of an *in vitro* replicon methods for investigating HBV replication as wells as having critical use in drug screens. We generated various HBV genomic DNA-carrying constructs with the aim to examine virus replication in the context of expression of the viral HBcAg, HBeAg and HBsAg under the control of their native promoters/enhancers. As described earlier, HBs proteins and nucleocapsid play a central role in the viral life cycle, and the primary component of the HBV envelope is the preS1 domain of L-HBsAg protein which is the primary entry point of infectious viral particles while the core

protein is the major component of the viral nucleocapsids that protect the viral genome, in addition to its role in capsid assembly (Chen et al., 2015).

After confirming the nucleotide sequences of the different constructs, we performed their functional validation by their transfection in several different cell lines (data unshown) and found that human hepatoma cell line (Huh7) was able to support the replicon system. Initially we validated expression of the proteins from the plasmid alone in Huh7 cell line by Western blotting and ELISA, using several monoclonal antibodies.

We were able to validate expression of HBcAg and HBsAg in Huh7 cells while expression of the proteins in HEK-293T cells when transfected with the same plasmid (pTHBV-L^-, pQ-HBV-L^- (1-6), pHBV1.3mer WT, pQ-HBV2.7 (3B4), and pQ-HBV- L^- /Puro (No 11) for core and L-HBsAg) showed no expression.

The ability to measure the release of the viral antigens into the cell culture supernatant is a critical measure of virus replication for the replicon system. To date ELISA have been previously developed for the detection HBeAg by others (Morey J. D, 2004; Goddard et al., 2017). The current commercially available kits are expensive and thus adapting the replicon to a high-throughput screen method would be non-viable. Therefore, we were able to design, develop, and validate an in-house ELISA assay that we then used to test the expression of HBeAg from media of the Huh7 cell that had been transfected with a variety of plasmids commercially and obtained from our internal plasmid constructions. Comparison of our inhouse ELISA to the commercially available kit, showed no difference in sensitivity or specificity of the assay. Thus, we can be confident in the limit of detection when comparing different drug regimens or examine the different stage of replication. Also, the results obtained from our ELISA suggested that the Huh7 was the ultimate cell line for observing HBeAg secretion, which will be crucial in the investigation of HBV in the presence and absence of appropriate drugs designed to combat RT mutant constructs.

To further confirm our ELISA data, we examined the production of two HBV antigens expressed in vitro. The use of ELISA and western blot allowed for two different outputs as a measure of the replicon system output. A further validation would have been the identification of the different genomic species generated by HBV during its life cycle. In order to detect the different species, we had purchased a commercial kit for detecting the presence of cccDNA intermediate of replicon. At the time that these experiments were planned and were being validated, the research was interrupted by the COVID-19 pandemic. While promising, the

preliminary data from this experiment was not sufficient to be included in this thesis. It will be critical for this to be followed up.

As an alternative replicon system that only contained the S antigen of HBV, we had developed a pseudo virus particle using the HDV backbone as this model would enable entry and neutralisation of HBV. We had generated the appropriate clones, and assay but as mentioned earlier this work was unable to continue due to COVID-19 restriction.

Chapter 5. Functional evaluation of HBV polymerase mutations associated with resistance to clincially approved antiviral drugs

5.1 Introduction

As described in chapter 1, the viral polymerase protein is key protein involved in HBV genome replication. It consists of terminal protein (following a spacer), reverse transcriptase (RT), and RNAse H domains. The RT domain is responsible for the minus strand DNA synthesis occurring from the pregenomic RNA template, which is the first step in HBV replication. Hence, RT is a main target for anti-HBV drugs.

The antiviral drugs act to improve the quality of life and survival of chronically HBV-infected patients by preventing or delaying disease progression (Grimm et al., 2011). The goal is to suppress viral replication, which usually results in the reduced histological activity of chronic hepatitis and biochemical remission. Despite the development of nucleos(t)ide analogue (NA) drugs, the emergence of resistance remains a major clinical concern. The prolonged use of these antiviral compounds will contribute to the constant elimination of HBV, along with the increasing emergence of resistance mutants that may markedly reduce the initial effects of these treatments (Fu et al., 2020). All the clinically approved HBV drugs target the activity of viral RT and all are reported to have viral resistance (as explained in Chapter 1) due to specific mutations in the RT domain, which encourages the development of novel anti-HBV agents that target non-polymerase viral or host proteins. Also, the emergence of drug-resistant mutants is inevitable due to heterogeneity of the HBV genome and the development of drug resistance is associated with a poor prognosis. Problems arising from drug resistance include hepatitis flares, the reversion of histologic improvement, and sometimes the severe exacerbation of illness, hepatic decompensation, or death (Park et al., 2019).

As defined in Chapter 3, a total of 64 samples were sequenced in this project and nine known resistance mutations were detected, at positions (rt80, rt91, rt134, rt153, rt204, rt215 and rt221). Among them, six mutations were associated with resistance to LAM resistance, two with TDF resistance (D134E and R153W), and one with ADV (F221Y). Remarkably, one sample presented three mutations (L80I, L91I and M204I) associated with LAM resistance. All these mutations were found in the viral genotype D isolates.

In this chapter, we aim to test the drug-resistant mutations identified in this study and measure their effect on virus replication in the presence and absence of selected drugs using the cell-

based assay described in Chapter 4. To do this, new constructs in pT-HBV-L$^-$/Puro background were generated. The rationale for their generation and the strategy used is described below.

The plasmids pTHBV-L$^-$ and pTHBV-L$^-$/Puro (11) described above contain one codon change in the HBV genome – the ATG start codon of preS1 domain of L-HBsAg changed to ACG resulting in no initiation of the L-HBsAg. This change doesn't affect the amino acid sequence encoded by the HBV Pol ORF which overlaps that of L-HBsAg although the codon is changed from CAT to CAC (both coding for histidine). Sprinzl et al. (2001) have described another L-HBsAg-null HBV construct in which stop codons at the 5' ends of the L and M protein ORFs were introduced at HBV nt 1003 and 1279 (numbering starting from the core initiation codon with the A residues being nt 1). As an added safety measure, we decided to introduce two stop codons at the same L and M amino acid positions in pTHBV-L$^-$/Puro. This would amount to changing nt 4234 and 4510 to 'A', respectively, of pTHBV-L$^-$/Puro (See Appendix 11). This means changing 4233-TTG-4235 and 4509-TGG-4511 each to TAG. This will result in the preceding overlapping Pol codons 4232-GTT-4234 and 4508-GTG-4510 changed to GTA in both cases but the coding amino acid 'V' of Pol will remain the same. Thus, effectively this construct will have 3 null mutations, two in the L ORF and one in the M ORF of HBV. The construct carrying these mutations was named pTHBV-L$^-$ 3-null/Puro. The desired individual Pol-RT mutation(s) were generated in this vector background. The strategy to design these constructs were developed by Prof Arvind Patel and their cloning, isolation and characterisation was performed by me.

Briefly, nt 3959 to 5342 (covering preS1/S2/Pol gene sequences of interest) with the native MfeI and BstZ17I at the 5' and 3' end, respectively, of pTHBV-L$^-$/Puro (11) carrying the above-described STOP codons, and also separately additional RT mutations of interest as shown below were custom synthesized by Genewiz.

3-Null

3 Null + RT L80I

3 Null + RT L91I

3 Null + RT H126Y

3 Null + RT D134N

3 Null + RT D134E,

3 Null + RT R153W

3 Null + RT M204I

3 Null + RT M204V

3 Null + RT Q215S

3 Null + RT F221Y

The strategy was to cleave the commercially synthesized fragments with MfeI and BstZ17I, gel purify, and individually sub-clone to replace the native MfeI/BstZ17I of pTHBV-L^- /Puro to generate the plasmids carrying the above mutations. Table 5-1 below provides the relevant information of the expected mutation changes in the preS1-, preS2- and the overlapping Pol-encoding sequences.

Table 5-1: Expected mutation changes in the preS1, preS2 and the overlapping Pol-encoding sequences.

	Codon change			Amino acid change				Conferring	
Plasmid name	preS1	preS2	Pol/RT	preS1	preS2	RT (Pol aa)	Overlapping ORF	Resistance to	Sensitivity to
pTHBV- L$^-$ 3Null-Puro-11	TTG - TAG	TGG - TAG	N/A	L19 - STOP	W3 - STOP	N/A	No change	N/A	
pTHBV- L$^-$ 3Null-Puro-11-L80I	"	"	TTA - ATA	"	"	L80I (425)	"	LAM, LdT, TDF	
pTHBV- L$^-$ 3Null-Puro-11-L91I	"	"	CTT - ATT	"	"	L91I (436)	"	LAM	
pTHBV- L$^-$ 3Null-Puro-11-H126Y	"	"	CAC - TAC	"	"	H126Y (471)	"	TDF	ADV
pTHBV- L$^-$ 3Null-Puro-11-D134N	"	"	GAC - AAC	"	"	D134N (479)	HBsAg M125I		
pTHBV- L$^-$ 3Null-Puro-11-D134E	"	"	GAC - GAG	"	"	D134E (479)	HBsAg T126S	TDF	
pTHBV- L$^-$ 3Null-Puro-11-R153W	"	"	CGG - TGG	"	"	R153W (498)	No change	LAM	LAM
pTHBV- L$^-$ 3Null-Puro-11-M204I	"	"	ATG - ATT	"	"	M204I (549)	HBsAg W196L	LAM, LdT, ADV, ETV	
pTHBV- L$^-$ 3Null-Puro-11-M204V	"	"	ATG - GTG	"	"	M204V (549)	HBsAg I195M	LAM, LdT, ADV, ETV	
pTHBV- L$^-$ 3Null-Puro-11-Q215S	"	"	CAG - TCG	"	"	Q215S (560)	HBsAg S207R	LAM, ADV, LAM	
pTHBV- L$^-$ 3Null-Puro-11-F221Y	"	"	TTT - TAT	"	"	F221Y (566)	HBsAg L213Y	ADV	

The aim was then to use these constructs in Huh7-based transfection assays to evaluate the effect of the mutations on virus replication (i.e. HBeAg secretion and expression of HBcAg) in the presence or absence of selected drugs.

5.2 Results

5.2.1 Sub-cloning of custom-synthesized fragments carrying the 3-null or 3-null plus individual pol mutation of interest into pTHBV-L$^-$ /Puro

As described in the Introduction section above, sequences representing nt 3959 to 5342 on an MfeI to BstZ17I were custom synthesized via Genewiz into their vector pUC57-GW. The synthetic fragments from pUC57-GW were cleaved with MfeI + BstZ17I, gel-extracted as described in Section 2.2.6.1. Separately, pTHBV-L$^-$ /Puro was cleaved with MfeI + BstZ17I in the presence of Calf-Intestine Phosphatase (CiP) and the large fragment carrying the plasmid backbone plus the HBV sequences flanking the MfeI and BstZ17I was gel-extracted. The latter was then ligated individually with the gel-extracted fragment carrying the desired mutation as shown below:

Lig 1. pTHBV-L$^-$ /Puro large fragment 3 Null fragment

Lig 2. pTHBV-L$^-$ /Puro large fragment 3 Null + RT L80I fragment

Lig 3. pTHBV-L$^-$ /Puro large fragment 3 Null + RT L91I fragment

Lig 4. pTHBV-L$^-$ /Puro large fragment 3 Null + RT H126Y fragment

Lig 5. pTHBV-L$^-$ /Puro large fragment 3 Null + RT D134N fragment

Lig 6. pTHBV-L$^-$ /Puro large fragment 3 Null + RT D134E fragment

Lig 7. pTHBV-L$^-$ /Puro large fragment 3 Null + RT R153W fragment

Lig 8. pTHBV-L$^-$ /Puro large fragment 3 Null + RT M204I fragment

Lig 9. pTHBV-L$^-$ /Puro large fragment 3 Null + RT M204V fragment

Lig 10. pTHBV-L$^-$ /Puro large fragment 3 Null + RT Q215S fragment

Lig 11. pTHBV-L$^-$ /Puro large fragment 3 Null + RT F221Y fragment

Representative images of the gel-extracted fragments of MfeI-BstZ1iI- cleaved pTHBV-L$^-$ /Puro and pUC57GW-3-Null or pUC57GW-3-Null carrying each of the RT mutation above are shown in Figures 5-1, 5-2 and 5-3.

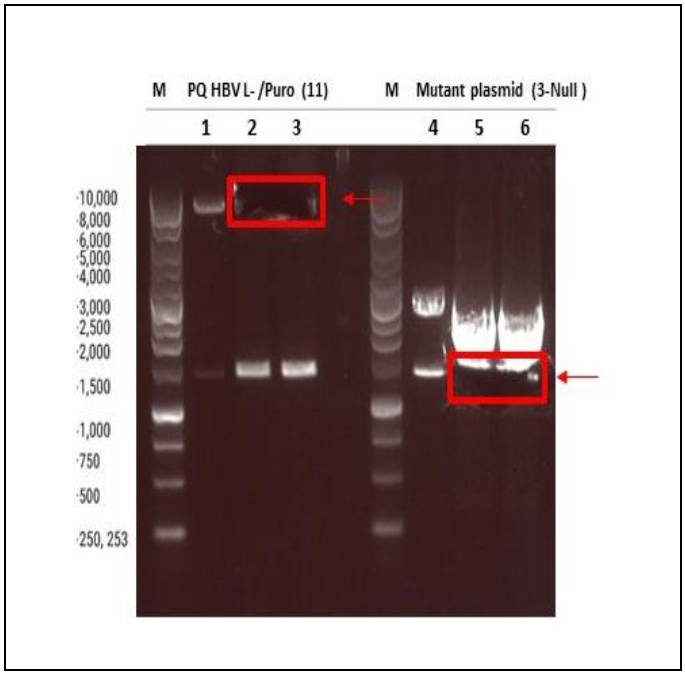


Figure 5-1: Agarose gel electrophoresis of pTHBV-L-/Puro (11) and pUC-GW cleaved with MfeI and BstZ17I.

As described in the text, the cleaved plasmids were electrophoresed through 0.8% agarose gel, and fragments indicated in the gel were extracted. An aliquot each of the extracted fragment was used in a ligation reaction Lig 1 to generate pTHBV-L-3-Null/Puro. A 1kb DNA ladder marker lane is shown. Samples were loaded in the gel as follow; Lane 1: 2 µl of pTHBV-L$^-$/Puro (11) + 10 µl water + 2 µl loading dye, Lanes 2 and 3: 25 µl of pTHBV-L$^-$/Puro (11) + 5 µl loading dye, Lane 4: 2 µl of pUC-GW cleaved with MfeI and BstZ17I + 10 µl water + 2 µl loading dye, Lanes 5 and 6: 25 µl of pUC-GW cleaved with MfeI and BstZ17I + 5 µl loading dye.

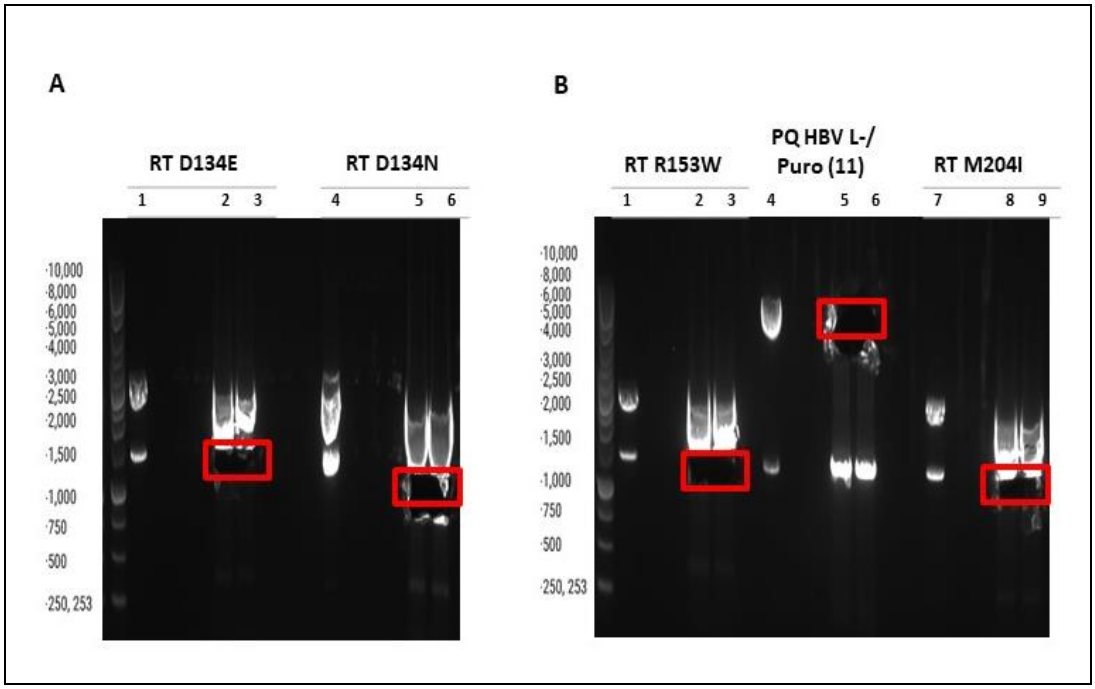


Figure 5-2: Agarose gel electrophoresis for 5 ligations (RT D134N, RT D134E, RT R153W, RT M204I, and pTHBV-L$^-$ /Puro (11)).

Five constructs were presented A) The desired bands for RT D134E and RT D134N were excised, recovered from the gel, and DNA was purified from gel using the purification kit. Samples were loaded in the gel as follows; Lanes 1 and 4: 2 µl of pUC-GW cleaved with MfeI and BstZ17I + 10 µl water + 2 µl loading dye, Lanes 2, 3, 5 and 6: 25 µl of pUC-GW cleaved with MfeI and BstZ17I + 5 µl loading dye. B) The desired bands for RT R153W, RT M204I and control sample were excised, recovered from the gel, and DNA was purified from gel using the purification kit. Samples were loaded in the gel as follows, Lanes 1 and 7: 2 µl of pUC-GW cleaved with MfeI and BstZ17I + 10 µl water + 2 µl loading dye, Lanes 2, 3, and 9: 25 µl of pUC-GW cleaved with MfeI and BstZ17I + 5 µl loading dye, Lane 4: 2 µl of pTHBV-L$^-$ / Puro (11) + 10 µl water + 2 µl loading dye, Lanes 5 and 6: 25 µl of pTHBV-L$^-$ /Puro (11) + 5 µl loading dye, All products were run on a 0.8% agarose gel alongside a 1kb DNA ladder marker.

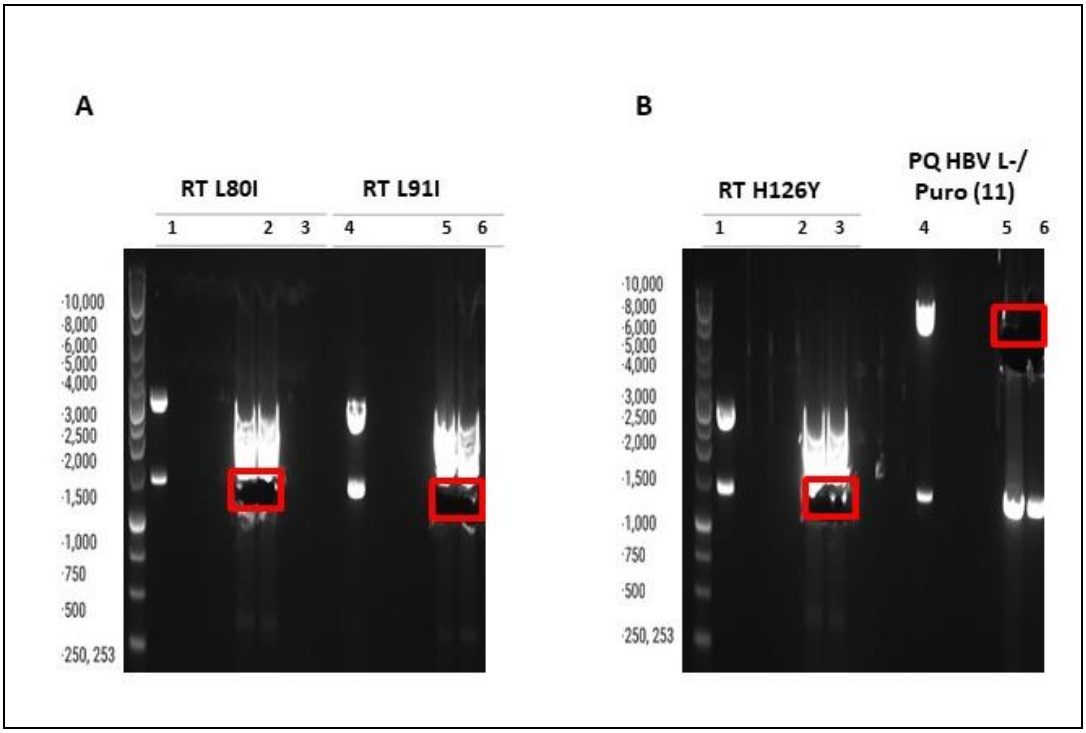


Figure 5-3: Agarose gel electrophoresis for 4 ligations (RT L80I, RT L91I, RT RH126Y, and pTHBV-L$^-$ / Puro (11)).

Four constructs were presented A) The desired bands for RT L80I and RT L91I were excised, recovered from the gel, and DNA was purified from gel using the purification kit. Samples were loaded in the gel as follows; Lanes 1 and 4: 2 µl of pUC-GW cleaved with MfeI and BstZ17I + 10 µl water + 2 µl loading dye, Lanes 2, 3, 5 and 6: 25 µl of pUC-GW cleaved with MfeI and BstZ17I + 5 µl loading dye. B) The desired bands for RT H126Y and control sample were excised, recovered from the gel, and DNA was purified from gel using the purification kit. Samples were loaded in the gel as follows, Lanes 1 and 7: 2 µl of pUC-GW cleaved with MfeI and BstZ17I + 10 µl water + 2 µl loading dye, Lanes 2, 3, and 9: 25 µl of pUC-GW cleaved with MfeI and BstZ17I + 5 µl loading dye, Lane 4: 2 µl of pTHBV-L$^-$ / Puro (11) + 10 µl water + 2 µl loading dye, Lanes 5 and 6: 25 µl of pTHBV-L$^-$ / Puro (11) + 5 µl loading dye. All products were run on a 0.8% agarose gel alongside a 1kb DNA ladder marker.

An aliquot each of the above ligation reactions was transformed into *E. coli* Stbl3 cells and the transformations selected on ampicillin L-agar plates. Plasmids from several colonies arising from each transformation were extracted and analyzed by agarose gel electrophoresis following cleavage with EcorI/NheI. As a positive control, pTHBV-L$^-$/Puro (11) was digested with the same enzymes was electrophoresed side-by-side.

As shown in (Figure 5-4), six plasmids (no 1, 2, 4, 6, 7, and 8) from Lig 1 displayed digestion profile similar to that of the control plasmid pTHBV-L$^-$/Puro (11). Likewise, plasmid number 5 arising from Lig 4 (RT R126Y), plasmids 1 and 2 of Lig 5 (RT D134N), and plasmid 1 of Lig 6 (RT D134E) had restriction digestion profile similar to that of pTHBV-L$^-$/Puro (11) (Figure 5-5 B and C, respectively). Unfortunately, no plasmids of interest were detected from other ligation reactions (Figures 5-5A and D, Figure 5-6). Due to limited time available, it was decided not to perform further screens to obtain the whole panel of mutant plasmids. Instead, we focused on working with a limited panel of constructs as described below.

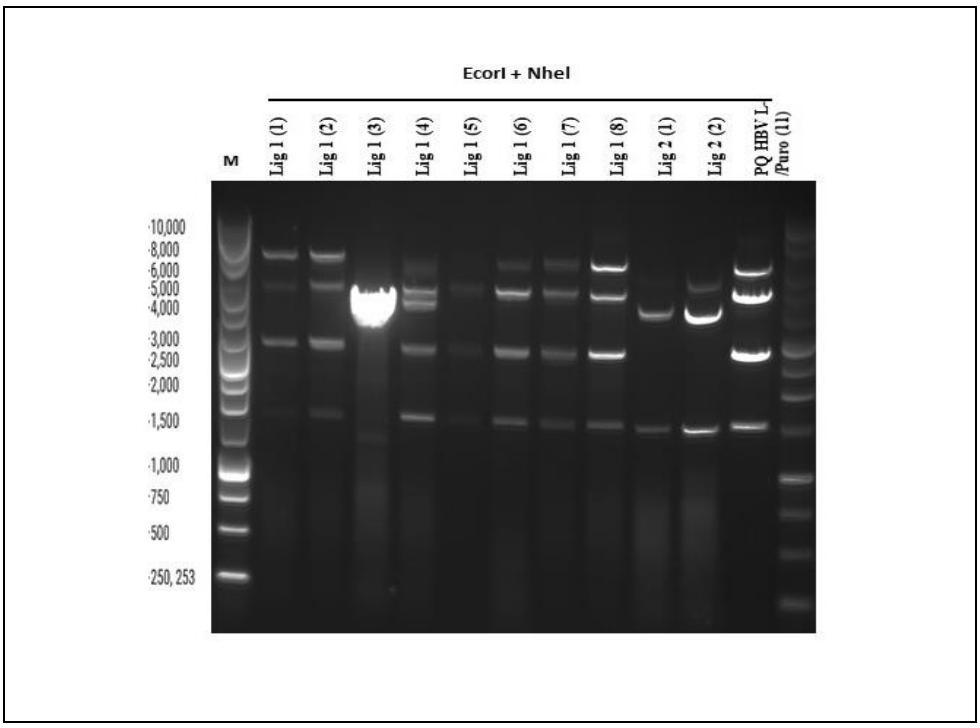


Figure 5-4: Screening of Lig 1 construct by Agarose gel electrophoresis.

Eights and two plasmids (no 1 to 8, and 1 and 2, respectively) arising from Ligs 1 and 2 were digested with NheI-EcoRI along with the positive control pTHBV-L$^-$/Puro (11). The digests were analysed on a 0.8% agarose gel alongside a 1kb DNA ladder marker.

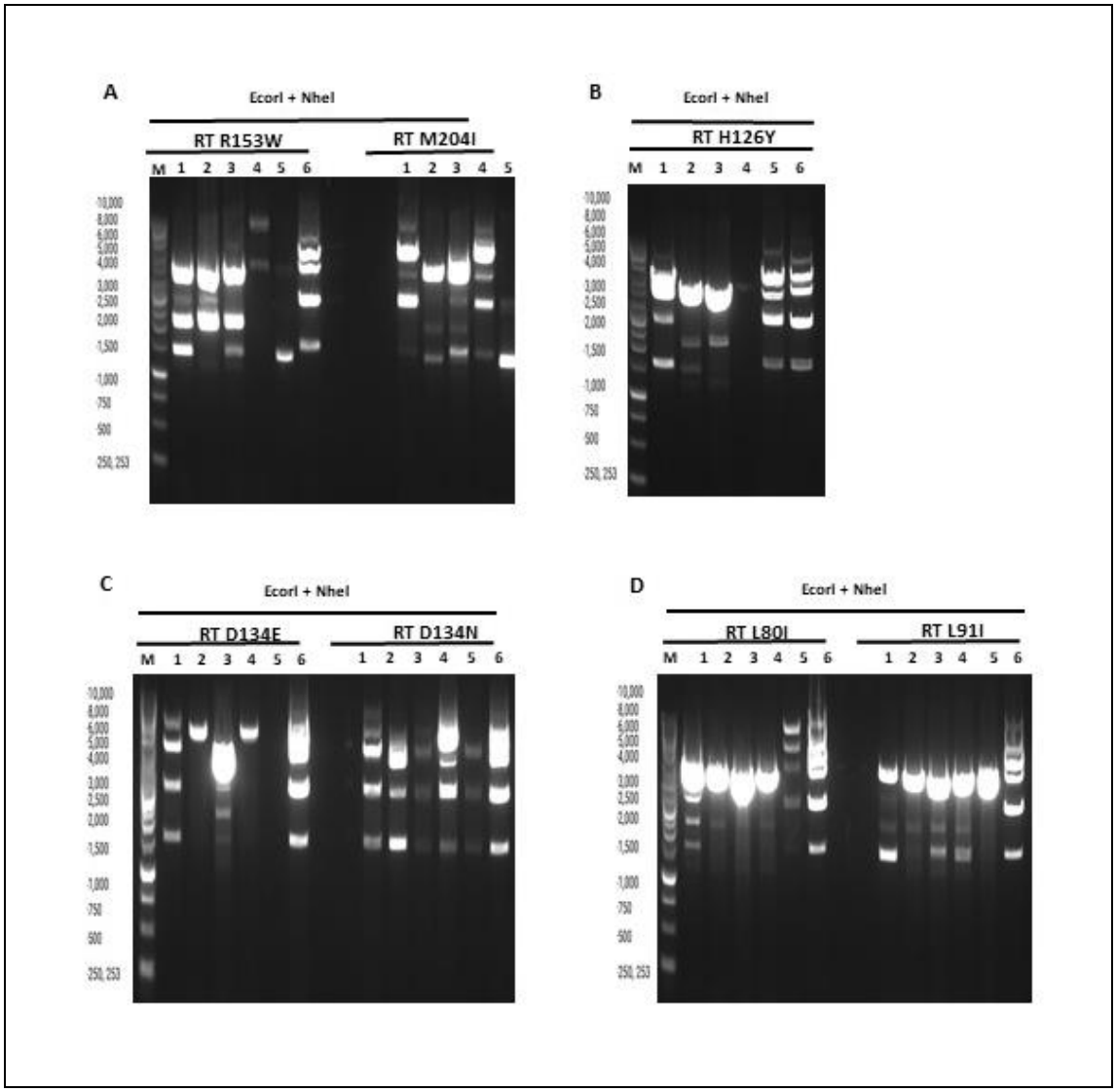


Figure 5-5: Screening of Ligs 2, 3, 4, 5, 6, 7, and 8 constructs by Agarose gel electrophoresis.

A) No plasmids of interest were detected from (RT R153W and RT M204I). B) Plasmid no 5 arising from Lig 4 (RT R126Y) had restriction digestion profile similar to that of pTHBV-L$^-$ /Puro (11). C) plasmids 1 and 2 of Lig 5 (RT D134N), and plasmid 1 of Lig 6 (RT D134E) had restriction digestion profile similar to that of pTHBV-L$^-$/Puro (11). D) No plasmids of interest were detected from (RT L80I and RT L91I).

All Ligs were digested with NheI-EcoRI along with the positive control pTHBV-L$^-$/Puro (11). The digests were analysed on a 0.8% agarose gel alongside a 1kb DNA ladder marker.

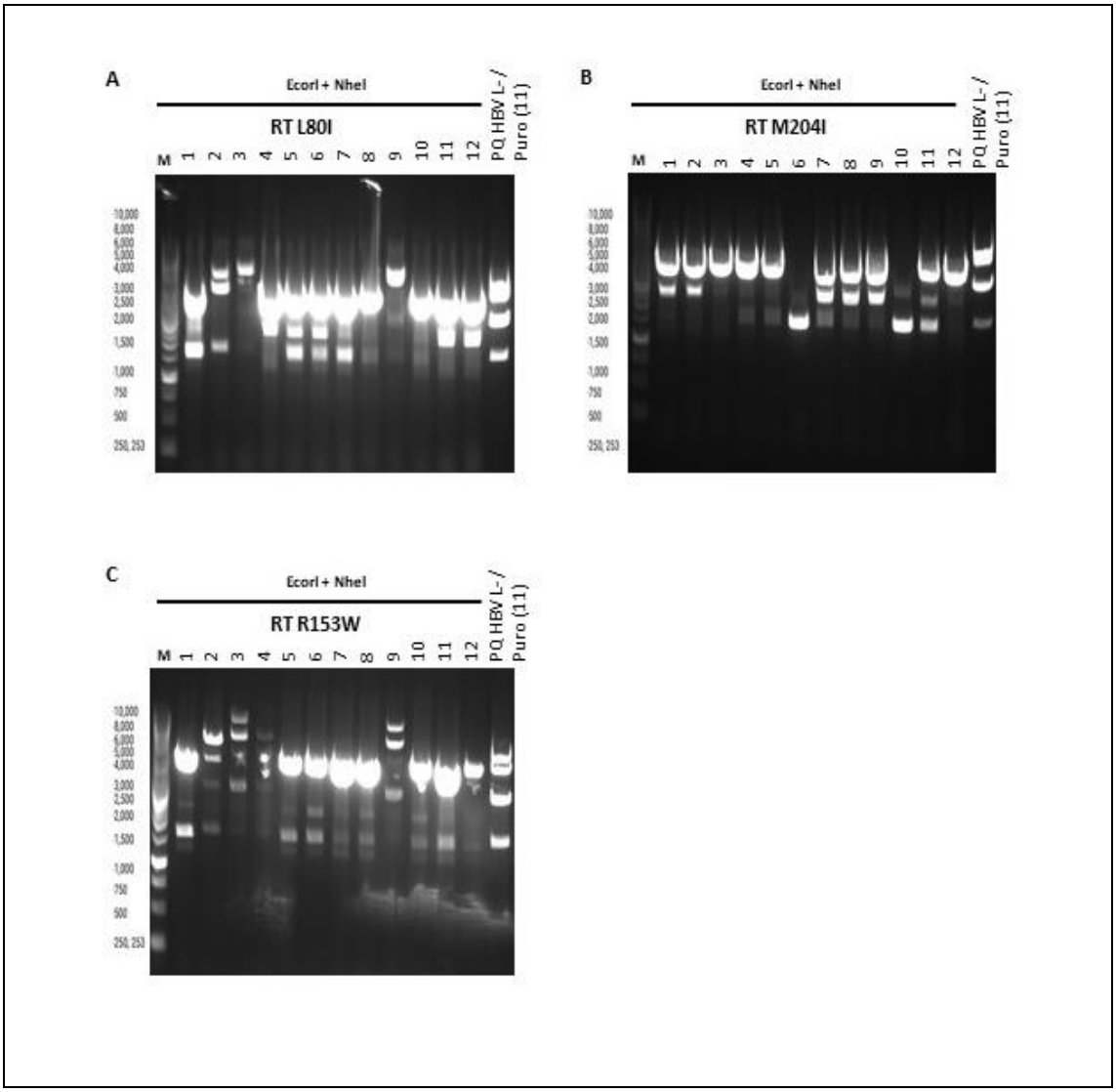


Figure 5-6: Screening of Ligs 2, 7, and 8 constructs by Agarose gel electrophoresis.

No plasmids of interest were detected from ligation reactions similar to that of pTHBV-L$^-$/Puro (11). All Ligs were digested with NheI-EcoRI along with the positive control pTHBV-L$^-$/Puro (11). The digests were analysed on a 0.8% agarose gel alongside a 1kb DNA ladder marker.

5.2.2 Time course of HBeAg expression

We first performed a time course of HBeAg expression in order to determine optimal condition for sample collection for use in our drug efficacy evaluation assay. Huh7 cell in 6-well tissue culture dish were transfected with 5 ug of pTHBV-L$^-$ / Puro (11) or pQ-HBV2.7 (3B4) using Lipofectamine 3000 and incubated at 37°C for five days. At every 24 h, an aliquot of the medium was collected (and replaced with equal amount of fresh medium) for analysis of HBeAg by the in-house ELISA, described above.

As shown in Figure 5-7, there was a time-dependent increase in the expression of HBeAg in the medium of pTHBV-L$^-$ / Puro (11)- transfected cells over the 5-day period. As expected, no HBeAg was detected in the medium of pQ-HBV2.7 (3B4)- transfected cells. Based on these results, it was decided to assess HBeAg levels at day 4 or day 5 post-transfection for our drug evaluation assay described below.

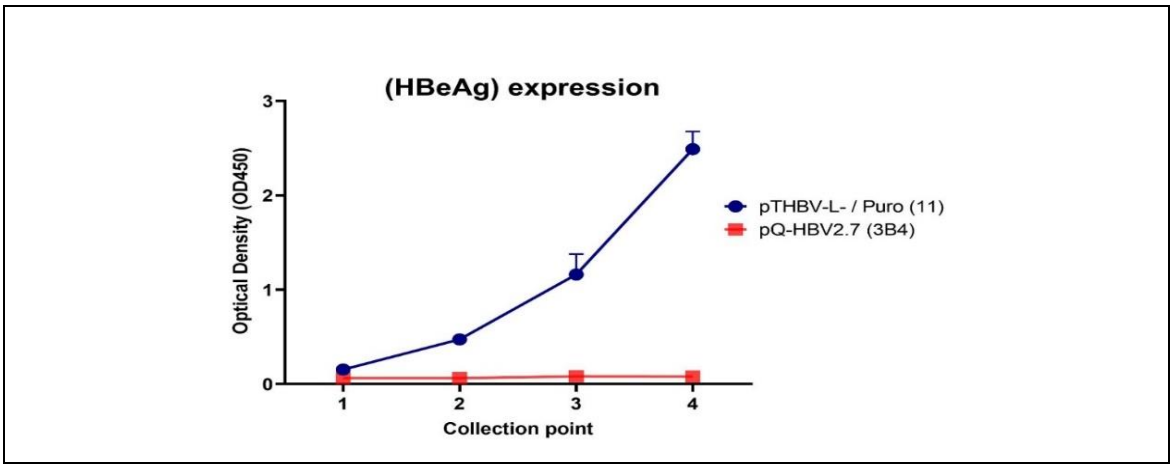


Figure 5-7: Time course of HBeAg secretion.

As described in the text, Huh7 cells were transfected with the indicated plasmids and incubated at 37°C for 5 days. Samples collected at each day post-transfection were analysed by our HBeAg ELISA. Briefly, HBeAg from cell medium was captured on coated ELISA plate and the captured protein detected using the primary antibody at concentration of 10 µg/ml anti-HBeAg MAb (10H10M) (Fitzgerald Industries). The bound antibody was then detected using the secondary HRP-conjugated MAb (61-H10K) (Fitzgerald Industries) at 1: 8000 dilutions. Subsequently, TMB was added for colour development and the readings taken at OD450 in the Varioscan ELISA reader. Error bars represent standard deviation from three independent experiments each performed in triplicate.

5.2.3 Screening pTHBV-L$^-$ null-Puro No (11) /RT mutant construct for HBeAg expression using ELISA

We next tested all the '3-null' plasmids selected from the restriction enzyme digestion screen above (Figures 5-5 and 5-6) for HBeAg expression using ELISA. The plasmid pHBV1.3mer WT was used as a positive control and pQ-HBV2.7 (3B4) as a negative control. Both plasmids were verified earlier (Section 4.2.12.2, Figure 4-19). Huh7 cells in 6-well dish were transfected with 5 ug of pHBV1.3 mer WT, 5 ug of pQ-HBV2.7 (3B4), 5 ug of pTHBV-L$^-$ / Puro (11), or the '3-null' plasmids and cells incubated for 5 days. At, 4-day post-transfection the cell medium was collected and tested for the presence of HBeAg. As shown in (Figure 5-8), the plasmids 3, 4, and 5 from the 3-null-D134N (Lig 5) ligation produced a modest signal; however, it was not twice as high as the negative control and, therefore, considered negative.

In contrast, plasmids 1 and 2 from the 3-null-D134N (Lig 5) produced a considerably higher OD450nm value, suggesting robust expression and secretion of HBeAg. Of the '3-null' constructs without RT mutations (Lig 1), only plasmids 6 and 8 were positive for HBeAg expression, whereas plasmids 1, 2, 4, and 7 produced no HBeAg. Only one plasmid (No 1) from 3-null-D134E ligation (Lig 6) was screened, and it produced a higher OD450nm value, suggesting a strong expression and secretion of HBeAg.

Furthermore, three plasmids were screened from the 3-null-L91I ligation and one from 3-null-H126Y ligation. None of them expressed HBeAg. However, they were screened one more time to confirm their negativity.

Thus, based on the data from this ELISA experiment it was decided to use for further experiments. The plasmid pTHBV-L$^-$ 3null /Puro/D134E (no 1) was sent off for nucleotide sequencing using primers HBV07 and AP314. This confirmed to have the expected 3-null and the D134E mutation and no other spurious mutations were observed in the sequence (nt 3690 to 5847) obtained using these primers. The mutations in the other plasmids remain to be confirmed by nucleotide sequencing. Given the time constraints it was nevertheless decided to proceed with the drug evaluation assays using these plasmids as the aim was to establish a proof-of-concept platform.

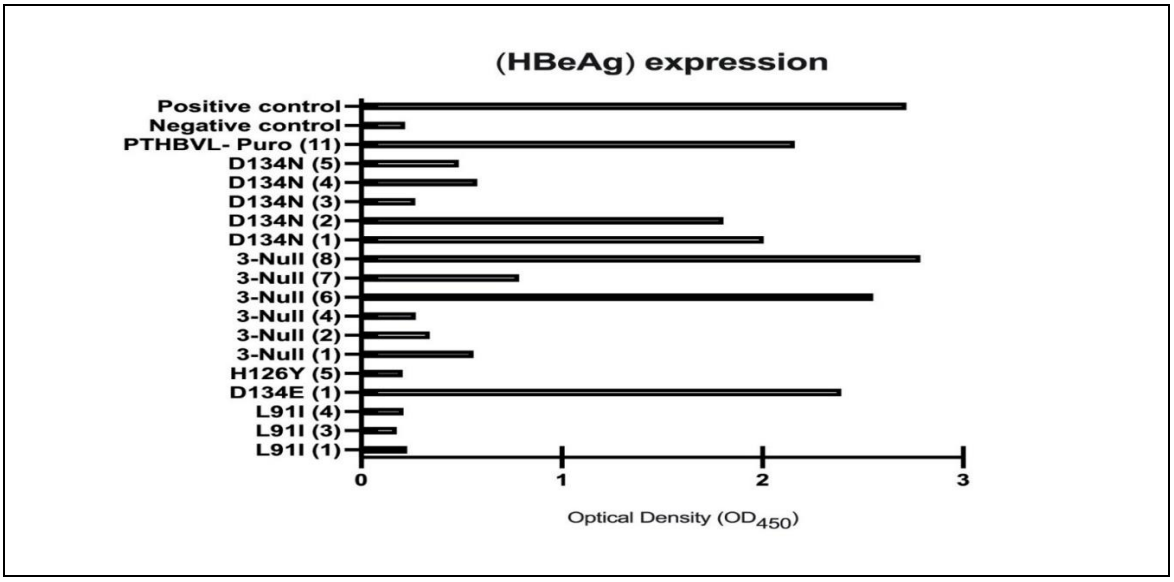


Figure 5-8: Assessment of HBeAg secretion in a panel of plasmid constructs using an in-house ELISA.

As described in the text, Huh7 cells were transfected with the indicated plasmids and incubated at 37°C for 5 days. The positive control was ptHBV1.3mer WT and the negative control was pQ-HBV2.7 (3B4). Briefly, HBeAg from cell medium was captured on coated ELISA plate and the captured protein detected using the primary antibody at a concentration of 2.5 µg/ml anti-HBeAg MAb (10H10M) (Fitzgerald Industries). The bound antibody was then detected using the secondary HRP-conjugated MAb (61-H10K) (Fitzgerald Industries) at 1: 8000 dilutions. Subsequently, TMB was added for colour development and the readings taken at OD450 in the Varioscan ELISA reader.

5.2.4 Functional evaluation of drug resistance-associated mutations

As described in the above section, pTHBV-L$^-$ 3null /Puro (no 8), pTHBV-L$^-$ 3null /Puro/D134E (no 1), and pTHBV-L$^-$ 3null /Puro/D134N (no 1) were selected for drug efficacy evaluation. A panel of drugs were tested to determine if they interfered with the expression of HBeAg bearing the two mutations in the transfected cells. The rationale behind this experiment is to understand if the mutations identified can confer drug resistance. Five drugs were commercially acquired (ADV, TDF, LdT, LAM, and ETV) and tested using the in-house ELISA as an indicator of HBeAg expression.

Initially, only plasmids pTHBV-L$^-$ 3null /Puro/D134E (no 1), and pTHBV-L$^-$ 3null /Puro/D134N (no 1) were chosen to test the effect of the above drugs on HBeAg expression. Huh7 cells seeded in a 6-well plate format were transfected as above with 5 ug of these plasmids and cells incubate for 1 day. Then, the medium was replaced with fresh medium containing serial diluted drugs with their concentration ranging from 0 to 50 uM or 0 to 200 uM (see Figure 5-9 for plate layouts). The cells were incubated for further 4 days. The medium was collected and tested for HBeAg by the capture ELISA. In short, cell supernatant was added to a previously coated plate with 10-H10M at 10 ug/ml. Then a secondary-HRP conjugated antibody was added at a dilution of 1/8000. Next, TMB was added for ten minutes, and the reaction was stopped with 0.1M sulphuric acid. The absorbance values were measured using a plate reader. The reason for testing the drugs at two different concentration ranges was to determine the optimal range to use in subsequent tests.

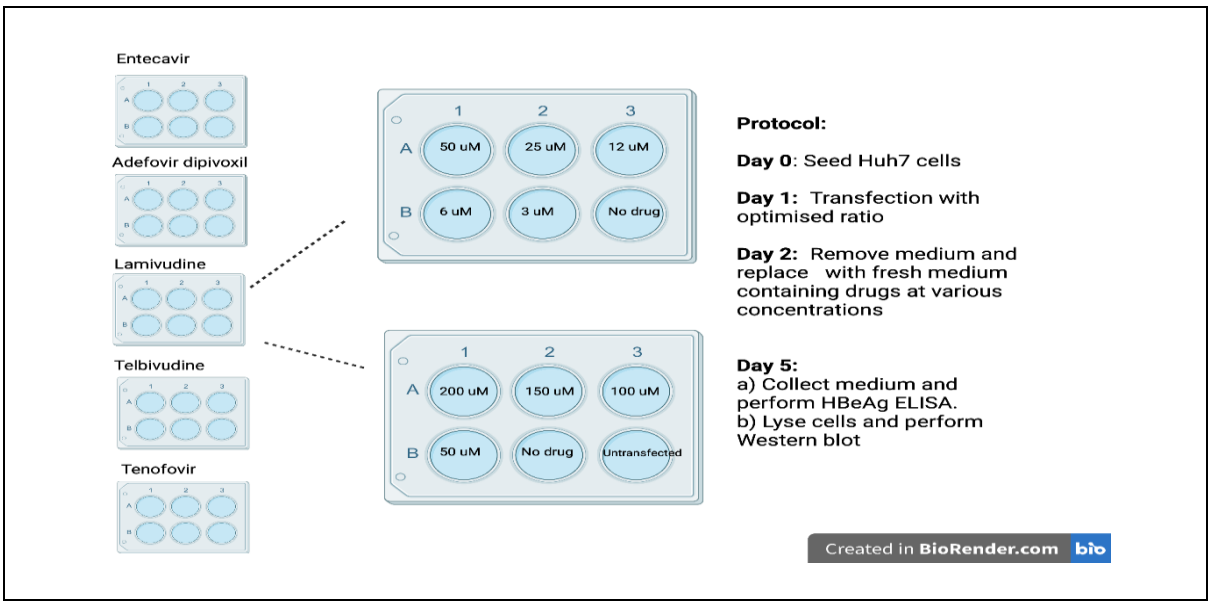


Figure 5-9: Schematic representation of experimental design for Antiviral drug assay.

As shown in (Figure 5-10), a dose-dependent inhibition of pTHBV-L$^-$ 3null /Puro/D134E (no 1), and pTHBV-L$^-$ 3null /Puro/D134N (no 1)- expressed HBeAg was observed with ADV, with an IC50 of around 30 μM in both cases. This result suggests that HBeAg bearing these mutations is sensitive to ADV. In contrast, the other four drugs failed to inhibit the expression of HBeAg by these plasmids which encode the variant sequences Figure 5-11, 5-12, 5-13 and 5-14). Although we had a negative control, due the COVID-19 pandemic, we were unable to confirm the sequence of the wild type plasmid for these initial drug screen assays. However, these outcomes were expected based on the previous finding as reported earlier (Yamani et al., 2017; Cho et al., 2018; Park et al., 2019).

To investigate if these drugs had an effect at a higher concentration, the experiment was repeated using a concentration range of 50-200 μM. This time, Huh7 cells were transfected with the same 2 mutant plasmids but also pTHBV-L$^-$ 3null /Puro (no 8) and pTHBV-L$^-$ (positive control) were used as controls to monitor the expression of HBeAg. At one day post-transfection, the medium of cells was replaced with fresh medium containing the appropriate drug at a concentration ranging from 0 to 200 uM. After incubation for a further 4 days, the medium of cells was assessed for HBeAg. This showed that not only ADV, but also TDF displayed a dose-dependent inhibition, as shown in (Figures 5-15 and 5-16) which is in concordance with Park et al. (Park et al., 2019). The other three drugs (LdT, LAM, and ETV) did not indicate any effect in terms of HBe expression (Figures 5-18 to 5-20). ADV showed an IC50 value between 50-75 μM, while TDF showed an IC50 of 118-128 μM. Little difference was observed between the IC50 of ADV for pTHBV-L$^-$ 3null /Puro/D134E (no 1) and pTHBV-L$^-$ 3null /Puro/D134N (no 1)- expressed HBeAg. Similarly, the IC50 of TDF against HBeAg expressed by the pTHBV-L$^-$ 3null /Puro/D134E (no 1) and pTHBV-L$^-$ 3null /Puro/D134N were comparable.

The IC50 for each of the drugs tested was statistically calculated using inhibitor vs. normalized response using the OD values from the ELISA. This allowed for use to determine if there was a shift in the IC50 due to the specific mutations. The IC50 of ADV calculated for each plasmid was (pTHBV-L$^-$ 3null /Puro/D134E (no 1): 74.35 μM, pTHBV-L$^-$ 3null /Puro/D134N (no 1): 51.45 μM, and pTHBV-L$^-$ 3null /Puro (no 8): 69.27 μM) (Figure 5-21). Similar results were observed for TDF with IC50 values of (pTHBV-L$^-$ 3null /Puro/D134E (no 1): 128.0 μM, pTHBV-L$^-$ 3null /Puro/D134N (no 1): 123.0 μM, and pTHBV-L$^-$ 3null /Puro (no 8): 118.2 μM) (Figure 5-21).

Importantly, all the IC50 calculated have an appropriate goodness of fit with an R square of between 0.8-0.9 (13 degrees of freedom) in ADV (Figure 5-20) and between 0.7-0.8 (8 degrees

of freedom) in TDF (Figure 5-21). We would have expected that the wild-type was more sensitive to the antivirals test than what is depicted, the assays require further optimisation and repeats to be undertaken as we only had significate experiments.

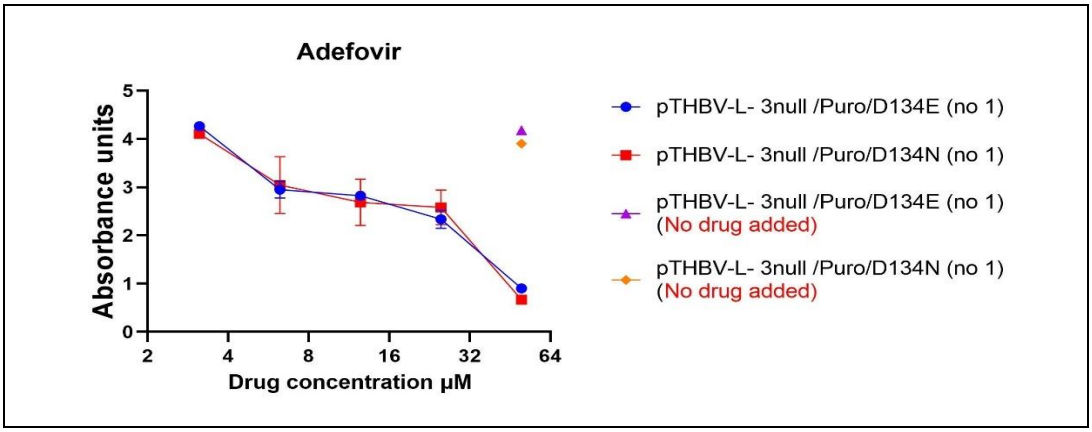


Figure 5-10: Effect of ADV on HBeAg expressed by pTHBV-L$^-$ 3-null/Puro carrying the RT mutation D134N or D134E.

Huh7 cells were transfected with 5 ug each of the two plasmid and cells incubated in the presence of absence of ADV added at the concentration range of 0 to 50 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

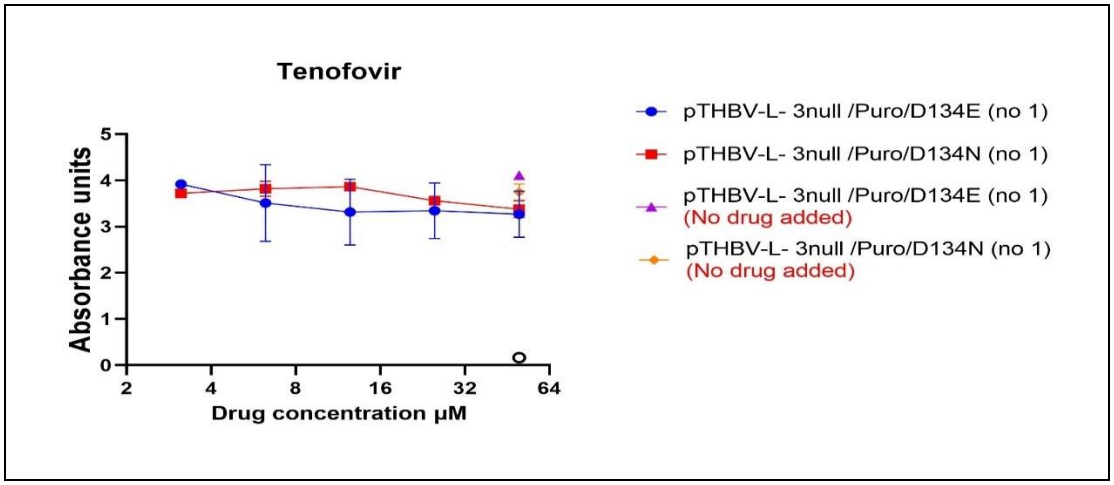


Figure 5-11: Effect of TDF on HBeAg expressed by pTHBV-L$^-$ 3-null/Puro carrying the RT mutation D134N or D134E.

Huh7 cells were transfected with 5 ug each of the two plasmid and cells incubated in the presence of absence of TDF added at the concentration range of 0 to 50 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

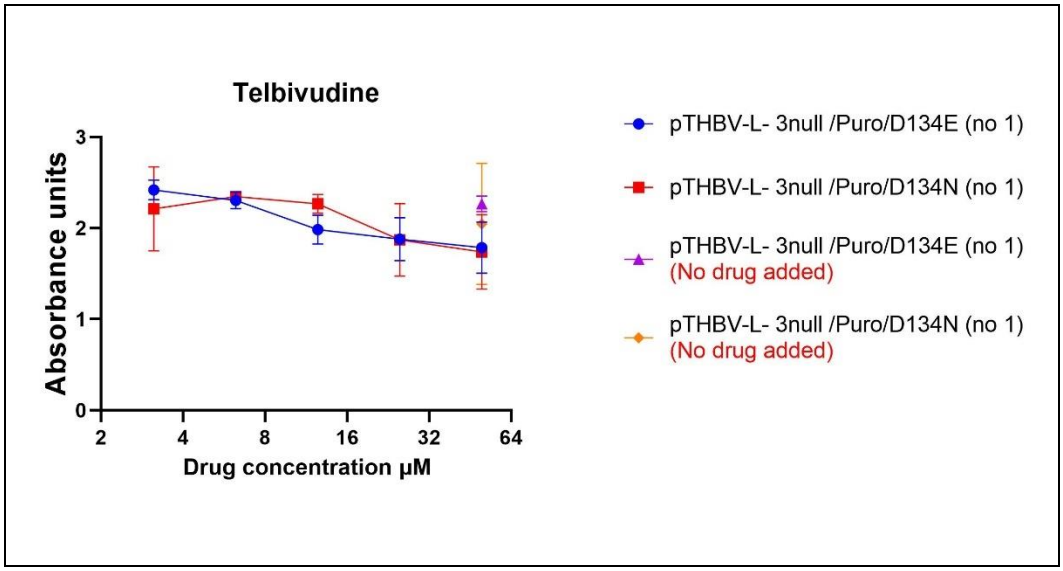


Figure 5-12: Effect of LdT on HBeAg expressed by pTHBV-L$^-$ 3-null/Puro carrying the RT mutation D134N or D134E.

Huh7 cells were transfected with 5 ug each of the two plasmid and cells incubated in the presence of absence of LdT added at the concentration range of 0 to 50 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

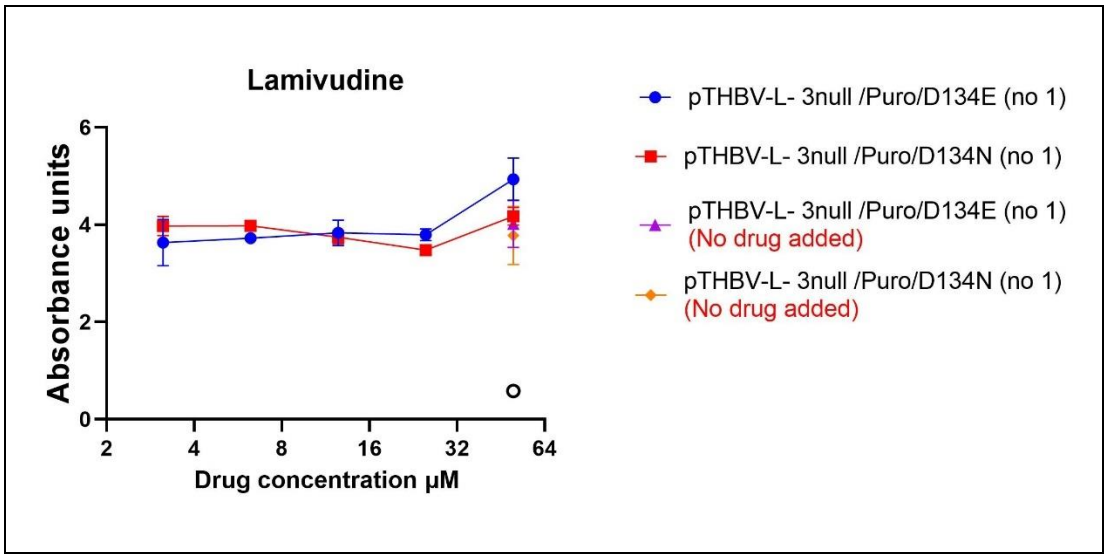


Figure 5-13: Effect of LAM on HBeAg expressed by pTHBV-L$^-$ 3-null/Puro carrying the RT mutation D134N or D134E.

Huh7 cells were transfected with 5 ug each of the two plasmid and cells incubated in the presence of absence of LAM added at the concentration range of 0 to 50 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

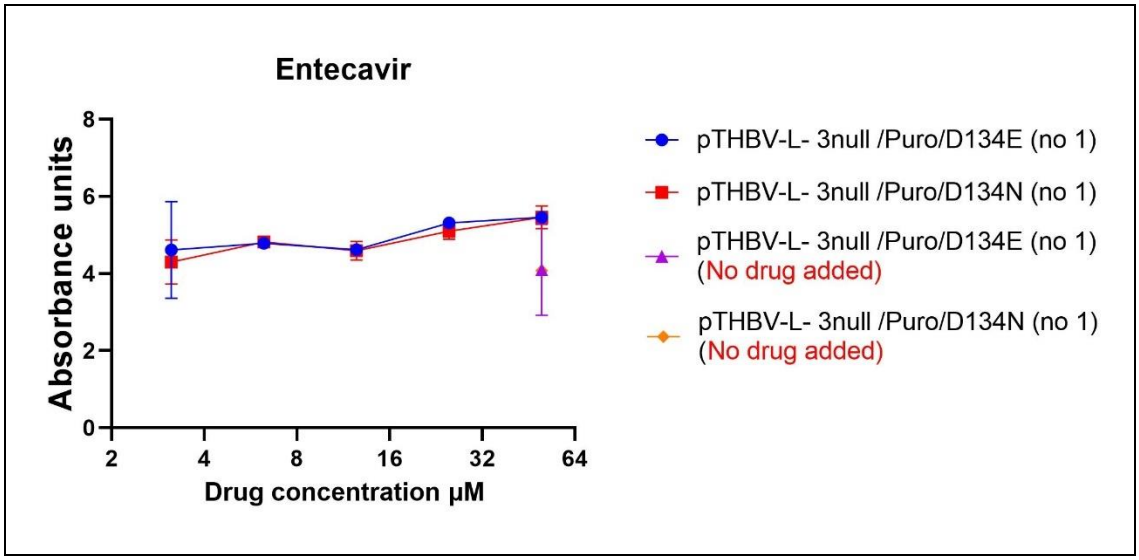


Figure 5-14: Effect of ETV on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E.

Huh7 cells were transfected with 5 ug each of the two plasmid and cells incubated in the presence of absence of ETV added at the concentration range of 0 to 50 µM . At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

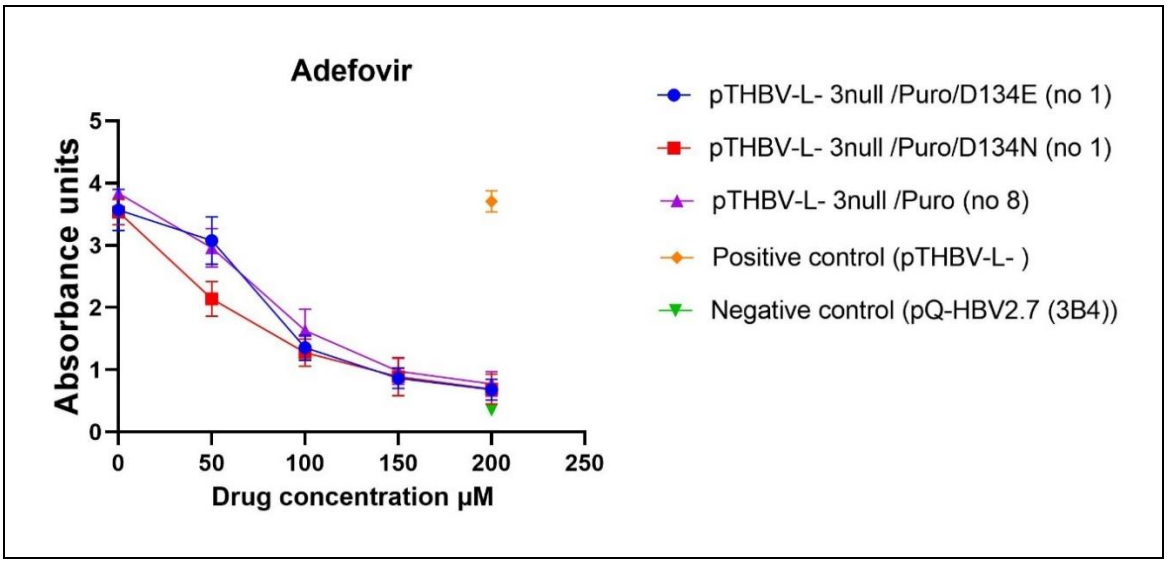


Figure 5-15: Effect of ADV on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8).

Huh7 cells were transfected with 5 ug each plasmid and cells incubated in the presence of absence of ADV added at the concentration range of 0 to 200 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

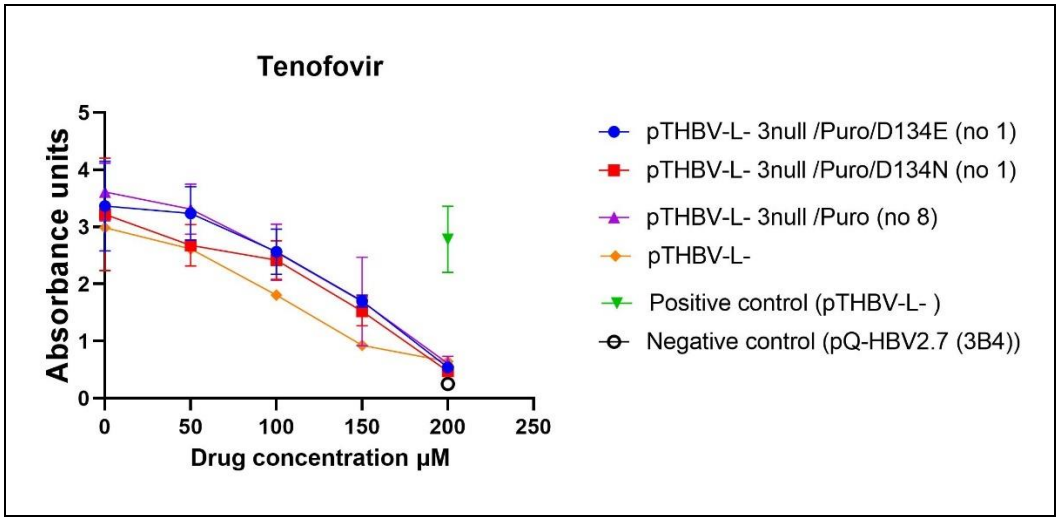


Figure 5-16: Effect of TDF on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8).

Huh7 cells were transfected with 5 ug each plasmid and cells incubated in the presence of absence of TDF added at the concentration range of 0 to 200 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

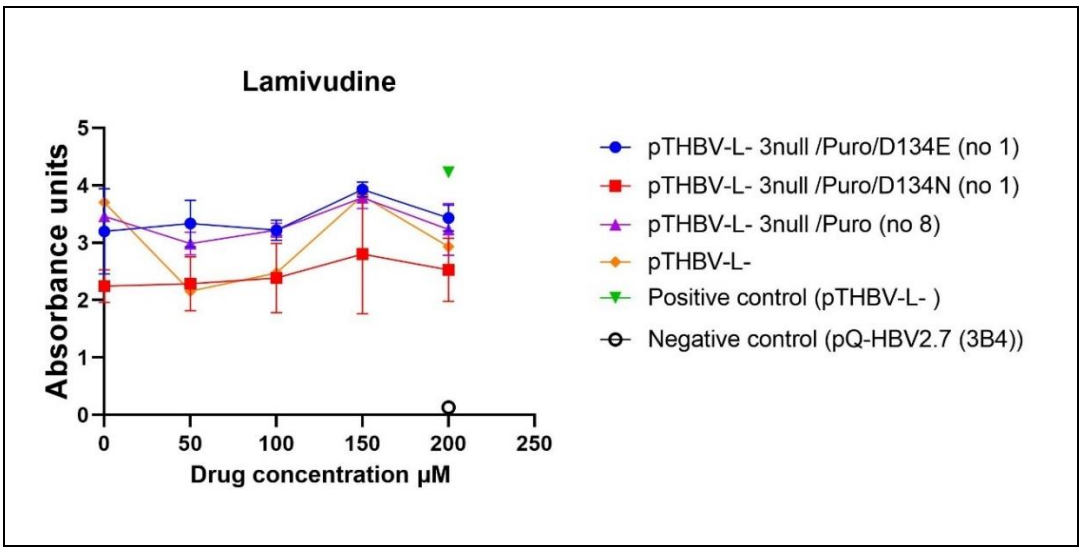


Figure 5-17: Effect of LAM on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8).

Huh7 cells were transfected with 5 ug each plasmid and cells incubated in the presence of absence of LAM added at the concentration range of 0 to 200 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

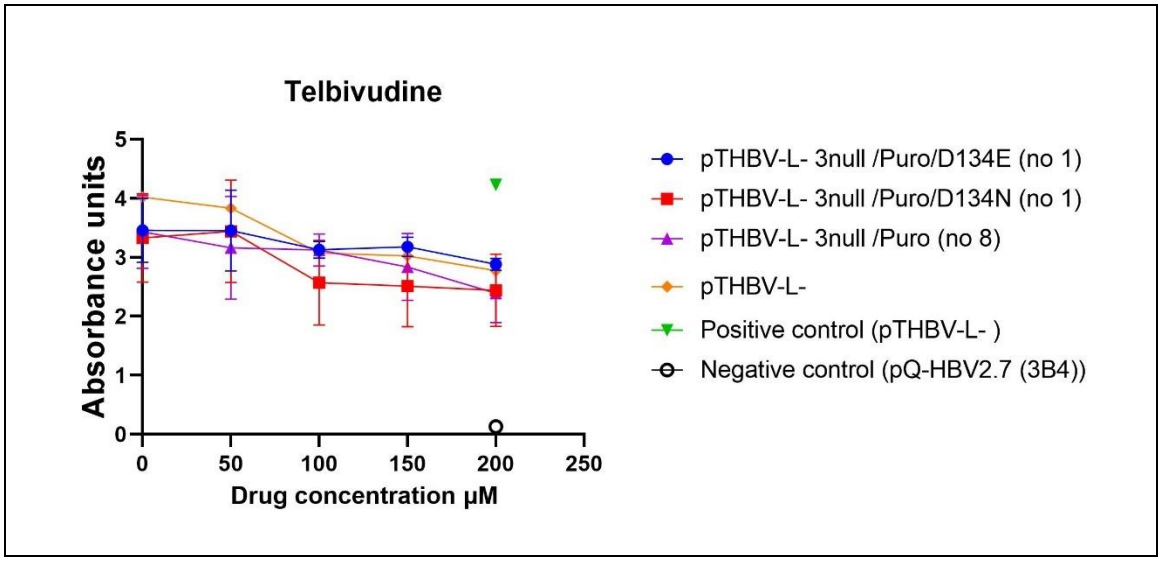


Figure 5-18: Effect of LdT on HBeAg expressed by pTHBV-L$^-$ 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L$^-$ 3null /Puro (no 8).

Huh7 cells were transfected with 5 ug each plasmid and cells incubated in the presence of absence of LdT added at the concentration range of 0 to 200 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

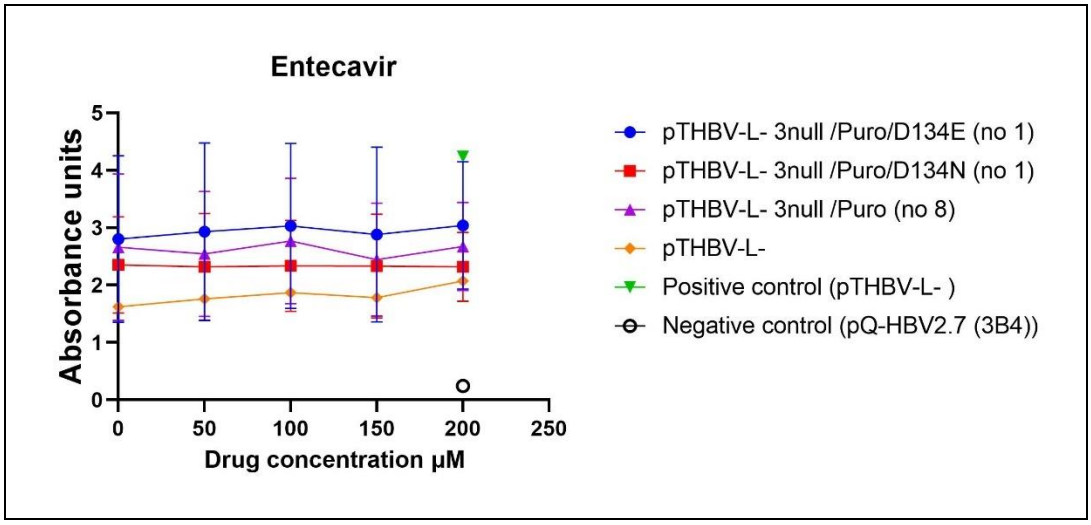


Figure 5-19: Effect of ETV on HBeAg expressed by pTHBV-L$^-$ 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L$^-$ 3null /Puro (no 8).

Huh7 cells were transfected with 5 ug each plasmid and cells incubated in the presence of absence of ETV added at the concentration range of 0 to 200 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

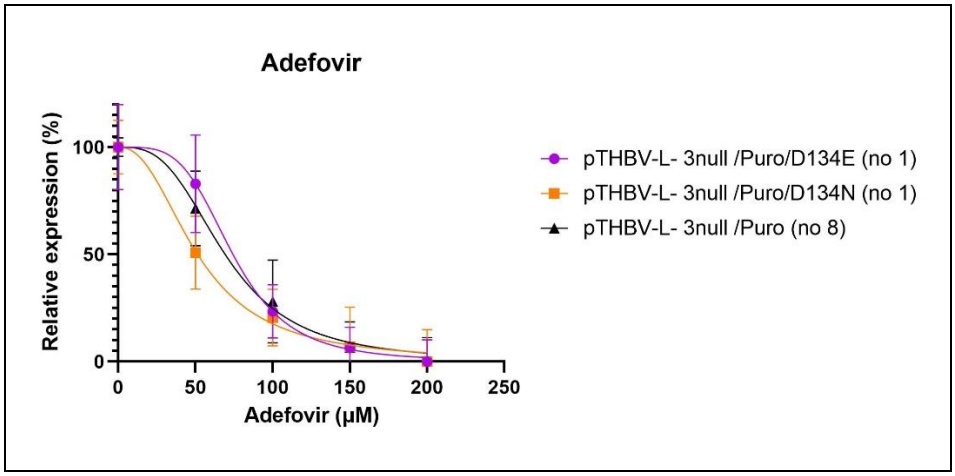


Figure 5-20: Normalized effect of ADV on HBeAg expressed by pTHBV-L- 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L- 3null /Puro (no 8).

Huh7 cells were transfected with 5 ug each plasmid and cells incubated in the presence of absence of ADV added at the concentration range of 0 to 200 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

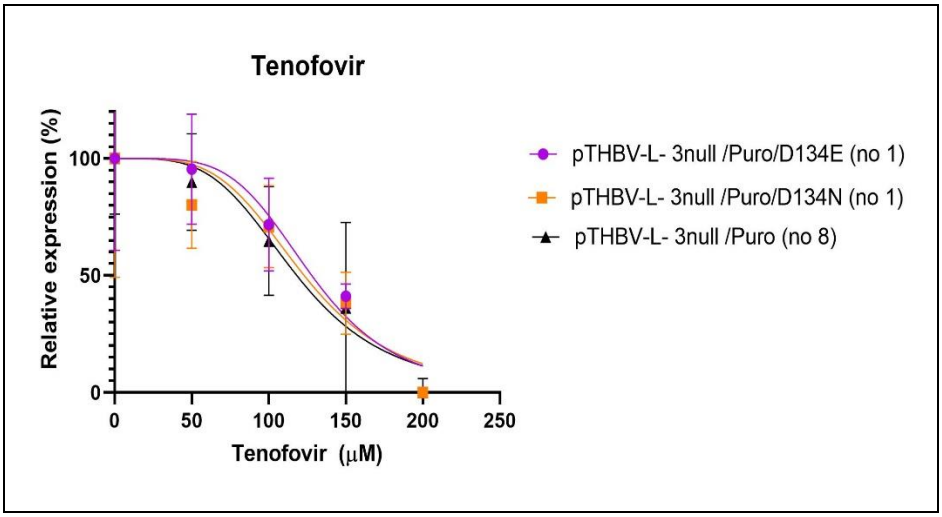


Figure 5-21: Normalized effect of TDF on HBeAg expressed by pTHBV-L$^-$ 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L$^-$ 3null /Puro (no 8).

Huh7 cells were transfected with 5 ug each plasmid and cells incubated in the presence of absence of TDF added at the concentration range of 0 to 200 µM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD450 measured as described in the text. Error bars represent standard deviation from three independent experiments each performed in singlicate.

As shown above, the secretion of HBeAg expressed by mutants (pTHBV-L$^-$ 3null /Puro/D134E (no 1), pTHBV-L$^-$ 3null /Puro/D134N (no 1), and pTHBV-L$^-$ 3null /Puro (no 8)) were inhibited by high concentrations of ADV and TDF (>50 µM), as shown by the reduction of the OD450nm value in (Figure 5-15 and Figure 5-16). The inhibitory activity of these drugs in the different constructs (bearing specific mutations (pTHBV-L$^-$ 3null /Puro/D134E (no 1), pTHBV-L$^-$ 3null /Puro/D134N (no 1), and pTHBV-L$^-$ 3null /Puro (no 8)) were normalised the signal derived from a standard curve from known concentrations of HBeAg, which were assayed alongside the samples. In few words, HBeAg ranging from 10 µg/ml to 0.12 ng/ml in 5-fold dilution intervals were assayed in the ELISA. Then, a line was fitted to these values and were extrapolated for larger antigen concentrations. The signal derived from the expression of the plasmids were normalised relative to this regression line for comparison and quantification of HBeAg expression.

As shown in (Figure 5-22), ADV inhibits the expression of transiently transfected HBeAg from all constructs in Huh7 cells. It is, however, important to consider that the construct pTHBV-L$^-$ 3null /Puro/D134N (no 1) is more susceptible than pTHBV-L$^-$ 3null /Puro/D134E (no 1) and pTHBV-L$^-$ 3null /Puro (no 8) to ADV. On the contrary, no evidence was found regarding drug resistance to TDF, associated with mutations in the pTHBV-L$^-$ 3null /Puro/D134N (no 1) or pTHBV-L$^-$ 3null /Puro/D134E (no 1) or pTHBV-L$^-$ 3null /Puro (no 8). Combined, these results suggest that ADV and TDF can inhibit the expression of HBeAg antigen, although some drug resistance against ADV is shown in construct bearing pTHBV-L$^-$ 3null /Puro/D134E (no 1) and pTHBV-L$^-$ 3null /Puro (no 8).

The normalisation of the data permits the quantification of the inhibition in terms of the relative expression of the HBeAg antigen, which may translate into how the drug would inhibit the replication of the WT virus. Additional experimentation to measure the drug's efficacy in terms of inhibiting replication in susceptible cell lines is required to confirm these results. However, the results here presented provide essential information on the potential drug resistance effect that certain mutations such as pTHBV-L$^-$ 3null /Puro/D134N (no 1), pTHBV-L$^-$ 3null /Puro/D134E (no 1) and pTHBV-L$^-$ 3null /Puro (no 8) could have upon treatment with ADV. The study of the mutations and their respective susceptibility to different drugs are very relevant to providing the best personalised antiviral treatment depending on the mutations present in the patient's quasispecies.

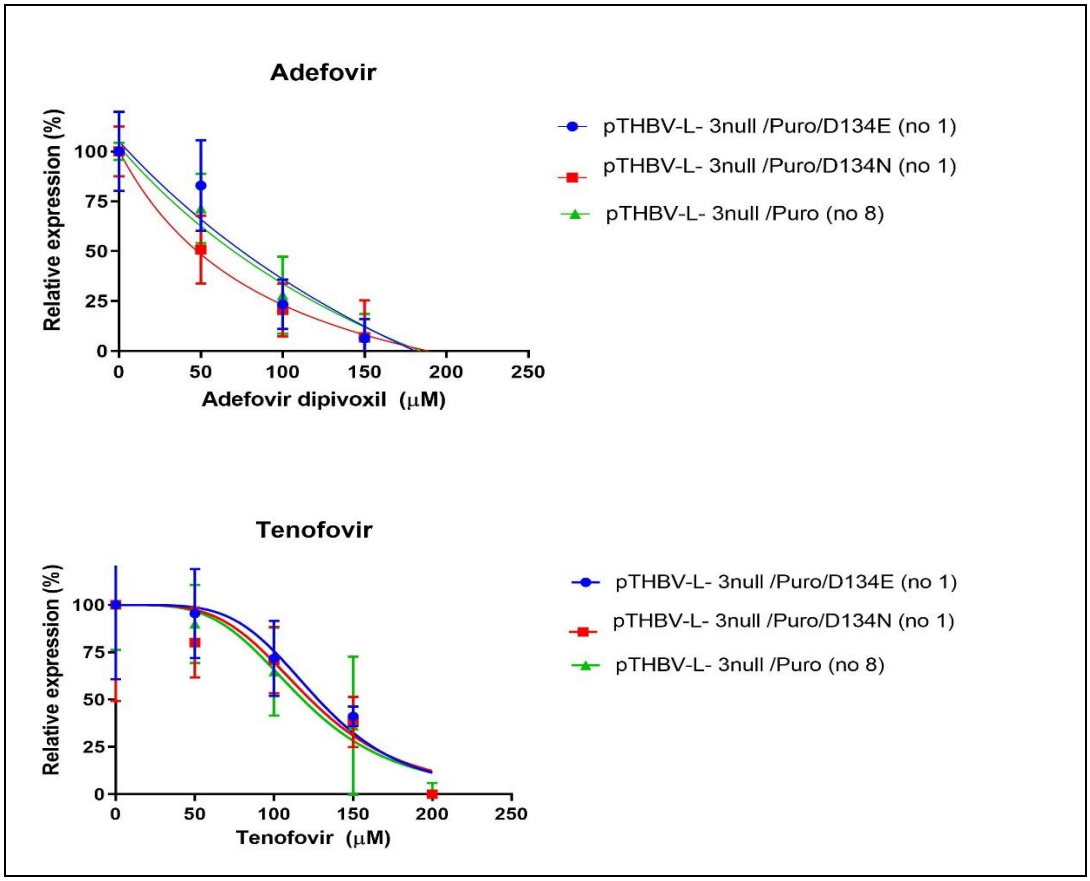


Figure 5-22: Effect of ADV and TDF on HBeAg expressed by pTHBV-L$^-$ 3-null/Puro carrying the RT mutation D134N or D134E and pTHBV-L$^-$3null /Puro (no 8).

Huh7 cells were transfected with 5 ug each plasmid and cells incubated in the presence of absence of drugs added at the concentration range of 0 to 200 μM. At 4-day post-transfection, the medium of cells was tested for HBeAg expression by our capture ELISA and OD 450 measured as described in the text. The OD450nm values were normalised to relative expression % by considering the OD450nm of mock treated as 100% expression. Error bars represent standard deviation from three independent experiments each performed in singlicate.

5.2.5 Testing of the activity of each drug against pTHBV- L$^-$ 3-null-Puro No 11 and the related RT mutant constructs by Western blot

To test if the inhibition of the drugs was linked to cell toxicity, Huh7 cells in 6-well-plates were transfected with pTHBV-L$^-$ 3null /Puro/D134E (no 1), pTHBV-L$^-$ 3null /Puro/D134N (no 1), pTHBV-L$^-$ 3null /Puro (no 8), pQ-HBV-L$^-$/Puro (No 11), and pQ-HBV2.7 (3B4), incubated with a single drug concentration (200 μM) and incubated using similar parameters as previously used for the ELISA.

Then, the cells were lysed with the LB2 buffer, and the cell extracts were subjected to 15% SDS-PAGE followed by Western blotting for HBcAg using our polyclonal serum R193 (Dilution 1:1000). In parallel, and as loading control, the extracts were also probed for tubulin which also served as a marker for cell viability. The cellular tubulin was detected using monoclonal *anti-gamma-tubulin* (Sigma-Aldrich) at 1:1000 dilution.

As shown in the Figure 5-23, none of the drugs prevented the expression of tubulin, although its levels varied in some cases (e.g. LdT, LAM). In keeping with the ELISA data above (Figure 5-22), there was a reduced or complete lack of HBcAg expression in the presence of ADV or TDF in cells transfected with all the plasmids (Figure 5-24A). In contrast, the other drugs failed to inhibit HBcAg expression in cells transfected with the mutant or the parent 3-null plasmid (Figure 5-24 B and C, Figure 5-25). While the HBcAg levels appear variable in some cases, overall, these results are in keeping with our ELISA results described above.

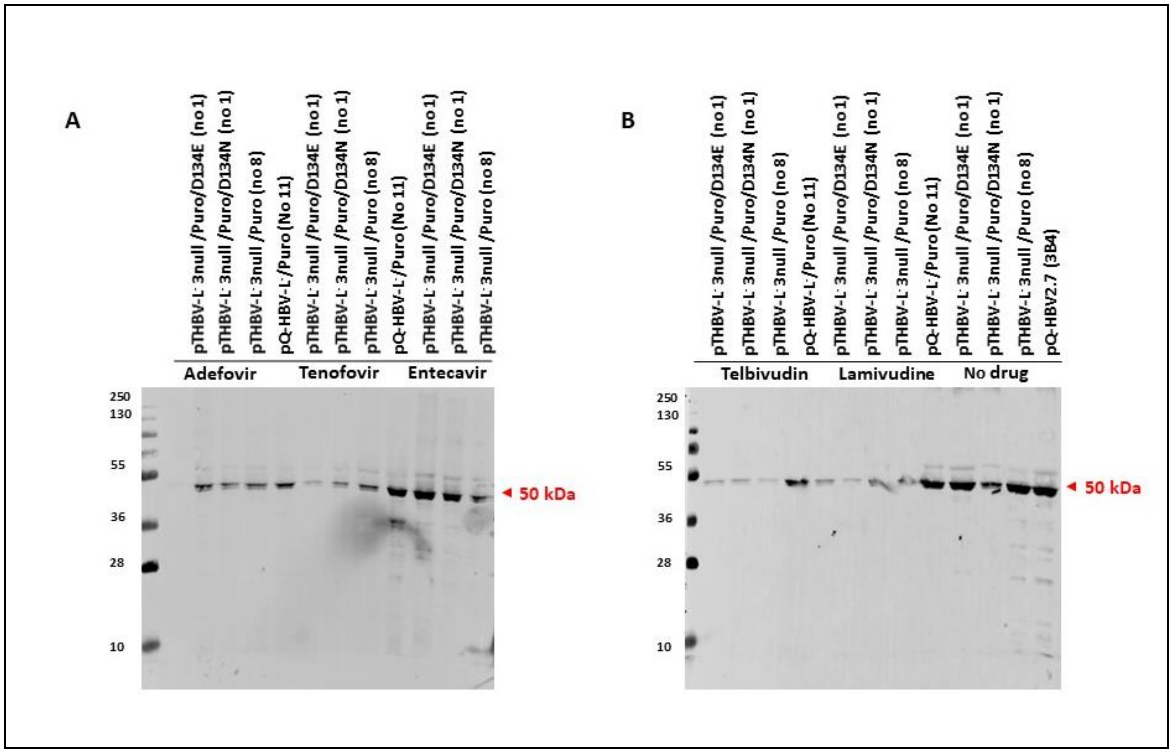


Figure 5-23: Expression of the intracellular Gamma tubulin in presence/ absence of HBV drugs by Western blot analysis.

Intracellular expression HBV proteins were detected in the immunoblot by expression of control protein gamma tubulin (*γ-tubulin)* in (Dilution 1:1000). Huh7 cells in 6 – wells plate was transfected with various plasmids presence/ absence of drugs and analysed on a 15% Polyacrylamide gel for SDS-PAGE to verify the expression of HBV Intracellular proteins. A) Potential targets recognition was achieved in all samples in presence of 200 μM of ADV, TDF, and ETV. B) Potential targets recognition was achieved in all samples in presence of 200 μM of LdT, LAM, and in absence of the drugs. The size expected for Intracellular protein was approximately 50 kDa. The primary antibody was recognised by secondary IRDye® 800CW anti-Mouse IgG (H + L) Licor. (Dilution 1:2000). An arrow indicates the potential protein band weight.

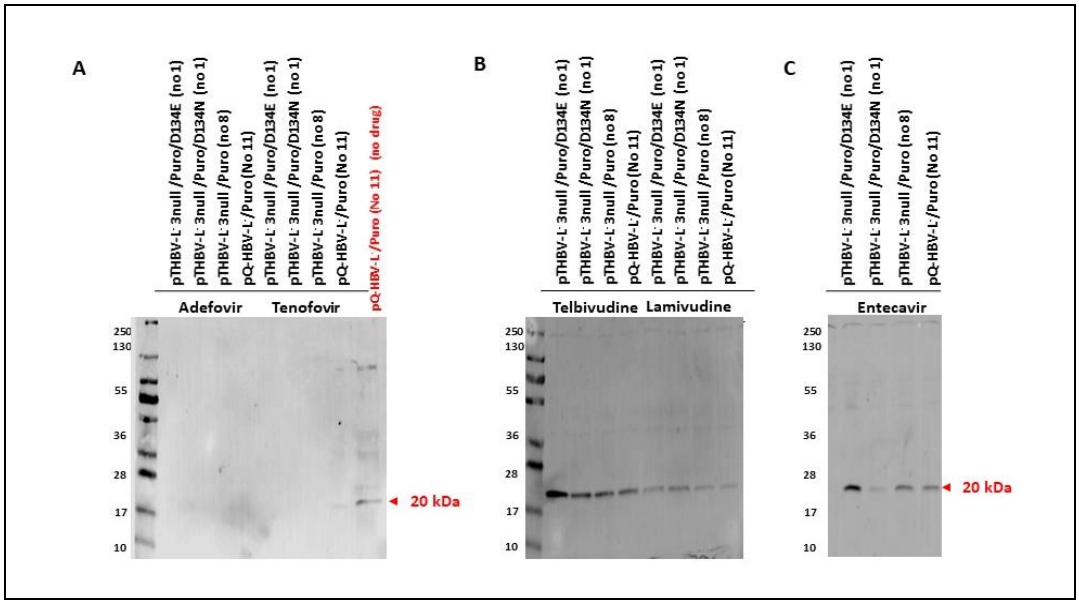


Figure 5-24: Expression of HBcAg in Huh7 cells in presence of HBV drugs by Western blot analysis.

HBcAg was detected by the binding of R193 primary rabbit monoclonal against the HBcAg regions. Huh7 cells in 6 – wells plate were transfected with various plasmids in presence of 200 μM of each drug and analysed on a 15% Polyacrylamide gel for SDS-PAGE to verify the expression of HBcAg. A) ADV and TDF showed a reduced or complete lack of core expression, shown by the 20 kDa protein indicated in the control sample without drug added (pTHBV-L$^-$ Puro No (11)). B) Potential targets recognition were achieved in experiment by adding R193 antibody against HBcAg (Dilution 1:1000) in all samples, LdT and LAM showed no effect in HBcAg expression. C) Potential targets recognition was achieved in experiment by adding R193 antibody against HBcAg (Dilution 1:1000) in all samples, ETV showed no effect in HBcAg expression. The primary antibody was recognised by secondary anti-rabbit IRDye® 680RD anti-Rabbit IgG (H + L) Licor (Dilution 1:2000). An arrow indicates the potential protein band weight.

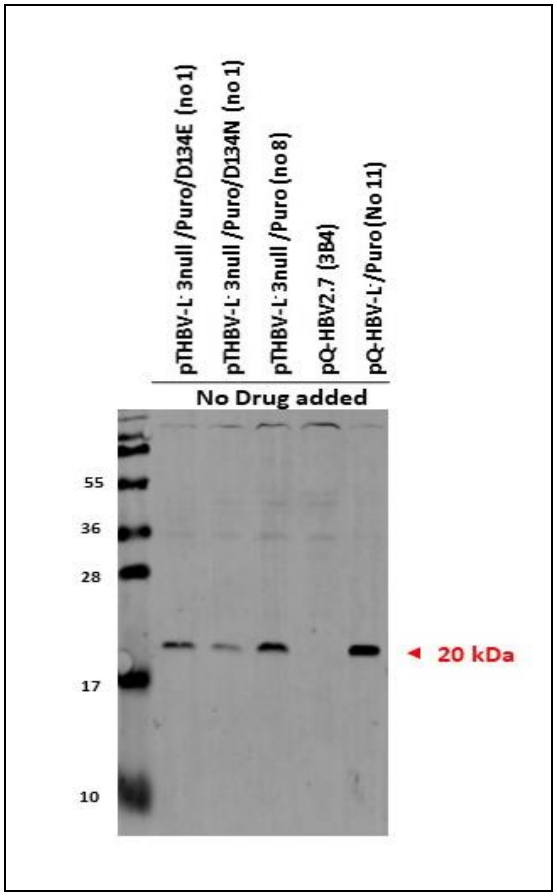


Figure 5-25: Expression of HBcAg in Huh7 cells in absence of HBV drugs by Western blot analysis.

HBcAg was detected by the binding of primary rabbit monoclonal against the HBcAg regions. Huh7 cells in 6 – wells plate were transfected with various plasmids in absence drug and analysed on a 15% Polyacrylamide gel for SDS-PAGE to verify the expression of HBcAg. Potential target recognition was achieved in all samples by adding R193 antibody (Dilution 1:1000) primary rabbit monoclonal antibody cross-reacting against HBcAg. pQ-HBV2.7 (3B4) was used as a negative control as this plasmid should not express HBcAg. The size expected for core antigen was approximately 20 kDa. The primary antibody was recognised by an anti-rabbit secondary antibody (Dilution 1:2000). An arrow indicates the potential protein band weight.

5.3 DISCUSSION

The aim of this chapter was to characterise mutations in the RT domain of the polymerase that we had detected in our Saudi Arabian samples by the NGS as described in Chapter 3 in respect to their susceptibility to different nucleoside analogues. In this chapter, we were able to generate three constructs which contained the specific mutation and had good protein expression, these mutations included rtD134E and rtD134N, both in the 3 Null background. Unfortunately, the HBeAg expression by constructs with rtH126Y and rtL91I mutations was very low. The low protein expression could be due to a number of factors including transfection efficiency, the mutations causing misfolding of the protein and the cell turning it over, and the protein maybe toxic to the cells with these mutations. There was no time to isolate these and the other RT mutants due to the restriction imposed by COVID-19 pandemic.

All 3 constructs expressed high levels of HBeAg in Huh7 cell. As the ELISA developed in Chapter 4 was used as an output of our replicon system, the expression of HBeAg is an essential for the investigation of antiviral drugs effect. Besides, it can be used as a marker of infection to determine the efficiency of drugs or other antiviral therapies in terms of combatting HBV. It is useful marker for the study of specific viral stages of in vitro infection systems as explained previously. It is critical to have an ELISA assay that detects the protein within a linear range of the standard curve, since if expression is too low or too high, the effect of the drug may not be properly quantitated. Indeed, the ELISA used enabled antigen detection in the linear range using a commercial antigen as well as the cell-expressed antigen. The results with the cell-expressed antigen and the commercially sourced HBeAg were comparable, both in detection limits and range. This provides greater confidence in the results of our drug screen assays.

The two RT mutations D134E and D134N analysed in this study were reported previously as pre-existing antiviral drug-resistant mutations (LAM) in patients with CHB, as were found prior to exposure to therapy (Zhang et al., 2019; Choi et al., 2018). Also, they were reported in patients with genotype D associated with the progression of severe liver disease, such as HCC and cirrhosis (Choi et al., 2018; Panigrahi et al., 2013). Recent report conducted by Mokaya et al. (2020) described D134E, as a primary drug resistance mutation, associated with causing resistance to LAM and ETV (Mokaya et al., 2020), whereas D134N was reported to be associated with LAM resistance (Choi et al., 2018).

To validate the robustness of the in-vitro drug assays developed in this chapter, we examined whether we observed the same decrease in susceptibility of the; pTHBV-L- 3null /Puro/D134E (no 1) and pTHBV-L- 3null /Puro/D134N (no 1), in comparison to wild-type pTHBV-L- 3null

/Puro (no 8) using the nucleoside analogues LdT, ETV and LAM. Using both the ELISA measuring HBeAg and a Western blot detecting HBcAg, both pTHBV-L- 3null /Puro/D134E (no 1) and pTHBV-L- 3null /Puro/D134N (no 1) were confirmed to have antiviral resistance using our assays. These three drugs (LdT, LAM, and ETV) did not appear to exert any effect on HBeAg expression, clearly indicating resistance by these mutations to these drugs. This is in keeping with previous studies (Choi et al., 2018; Mokaya et al., 2020). Although requiring further optimisation particularly with the lack of sensitivity of the wildtype to know inhibitors of HBV, our date validates the replicon system developed here as an alternate to using live virus to study potential new small molecule inhibitors and resistance markers. Our replicon system enables single points to be studying in a homogenous background without any other confiding mutations within the genome. This is critical in monitoring changes in the genome to different drugs.

We have validated an in vitro model for screening antiviral drugs under BSL2 conditions. Importantly we were able to show in our study that none of the mutant variants carrying (D134E and D134N) conferred resistance to alternative inhibitors including ADV and TDF. This is in concordance with the previously published studies.

Due to COVID-19-related time restriction and unavailability of kits, it was not possible to perform cell toxicity assays for each drug. As an alternative method, expression of control protein *gamma tubulin* was examined to check if the inhibition of the drugs was linked to cell toxicity. This was performed by Western blot. This would support our hypothesis that expression was affected by the inhibitor and not due to the 'death' of the cell due to the addition of the drug. Also, as shown in results (ADV and TDF) indicated a complete lack of core expression while (LdT and LAM, and ETV) showed no effect in HBcAg expression. Importantly, the expression of core matched the output from the ELISA for each mutation in the presence or absence of the different drugs. Furthermore, none of these drugs prevented the expression of tubulin that served as a marker for cell viability.

In summary, the development of the tools in this chapter provides alternate, safer means to undertake high throughput drug screens with drug resistant strains of HBV. Any hits identified in this preliminary screen could then be tested in a live virus assay. Furthermore, the findings described in this chapter provide novel insights into the evolution of antiviral resistance. The plasmids and cell systems described in this chapter also represent invaluable tools for future studies on HBV replicative processes.

Chapter 6. Conclusions and Future work

6.1 Background

The World Health Organization (WHO) estimates that two billion people have been exposed to the HBV and despite the availability of efficacious vaccines and treatments, there are 240 million people today suffering from CHB. Approximately 25% of these chronically infected individuals will go on to develop cirrhosis and/or liver cancer. With the advent of molecular testing for HBV DNA, it has been shown that there is a positive correlation between HBV viral load and the incidence of cirrhosis and liver cancer. NA therapy including LAM was initially approved in 1998 and can lower the HBV viral load reducing the incidence of liver injury and frequency of liver cancer.

WHO sustainable development goals have set a target aiming to combat viral hepatitis as a public health threat by the year 2030 by upscaling prevention and treatment strategies which include ensuring: (i) 90% coverage of universal hepatitis B immunisation of infants; (ii) 90% coverage of hepatitis B birth dose vaccine; (iii) 90% coverage in the diagnosis of HBV and HCV; (iv) 80% coverage in treatment of HBV and HCV; (v) 100% coverage in the screening of blood donations in a quality-assured manner; (vi) 90% coverage of injections given with safety-engineered devices; and (vi) distribution of at least 300 sterile needles and syringes to each person who injects drugs every year (WHO, 2016a). However, HBV drug resistance-associated mutations (RAMs) and/or vaccine escape mutations (VEMs) are potential barriers to the achievement of this target.

Though NA therapy alone cannot achieve a complete or virological cure, defined as loss of all HBV markers, including the key replicative intermediate, cccDNA in the hepatocyte nucleus. Long term NA therapy can be associated with the emergence of resistant variants which results in treatment failure. Development of mutations selected via resistance pathways has been implicated in hepatocarcinogenesis, with studies showing the incidence of HCC significantly higher in patients with LAM resistance compared to treatment naïve CHB patients.

6.2 Summary of the thesis

This project aimed to establish a new sequencing protocol using NGS technology and then apply this protocol to detect the resistance associated viral mutations. Subsequently, *in vitro* assay was developed, and antiviral resistance –associated mutations were tested against NAs.

Currently, NGS and bioinformatics have been utilized magnificently in a huge selection of analyses for infectious disease research of public health relevance. For example, these two approaches have been used to discover outbreak origins, track transmissions, determine etiological agents of disease, investigate epidemic dynamics and discover novel human pathogens (Maljkovic et al., 2020). However, many challenges have grown in establishing of high-quality genomics surveillance in public health laboratories and research setting including the limitation in setting up sequencing facilities and its downstream supportive services like Bioinformatics analysis. These challenges mostly include the choice of the sequencing platform and the sequencing approach, the choice of bioinformatics methodologies, availabilities of suitable computational power and information technology (IT) infrastructure. Another major challenge is recruiting and retaining personnel with the specialized skills and experience in this field.

Among the various approaches of deep sequencing, the target enrichment approach remains the most complicated to implement in diagnostic laboratories. There are various attempts to develop target enrichment methods to improve the probability of capturing pathogen-derived genomes because metagenomic approaches produce high level of background noise (Maljkovic et al., 2020). In this project, we were able to design specific probes for HBV based on the initial sequence data generated to develop a target enrichment protocol that allowed large-scale whole HBV genome sequencing. A protocol that has been successfully used for HCV and HEV sequencing (Thomson et al., 2016; Davis et al., 2021).

A major improvement in the yield of generated HBV-specific reads was achieved by an additional approach to decrease the quantity of human genetic material within a sample which can be achieved by exploiting the genetic differences between the host and the pathogen DNA. We utilized the method that allow microbial DNA to be enriched through the selective binding and removal of the CpG-methylated host DNA from HBV sample containing methylated host DNA. Using the target enrichment approach, we produced full genome sequences from most of the studied samples using the target enrichment approach. Also, we were able to optimise our protocol and produce full genome coverage from most of our samples. The method validated in this study paves the way for a better understanding of HBV epidemiology by offering a new tool for large scale whole-genome sequencing to achieve cost-effective testing in reference laboratories or research settings. Further developments to shorten the sequencing time and automate the bioinformatics pipeline would facilitate the introduction of this method to diagnostic laboratories and public health settings.

The prevalence of HBV infection varies to a great extent in different areas and countries, depending on several considerations that include implementation of the HBV vaccination program, screening of blood samples from blood donors, or education of the population about the risk factors for HBV infection. Saudi Arabia, a country with an intermediate HBV carrier prevalence of 3.6%, has implemented HBV vaccination of all new-borns in the country for the population free of charge in 1989.

As explained earlier, the global distribution of HBV genotypes seems to have an obvious pattern. Genotype A is the most prevailing in Western Europe, genotypes B, C, I, and J in East Asia, genotype E in Africa, genotype F in America, and genotypes G and H in Europe and Japan, whereas genotype D is widespread globally. Genotype D is distributed worldwide and is the most dominant HBV genotype in the Middle East and neighbouring countries or those in close contact to Saudi Arabia (Alfaresi et al., 2010; Gasim, 2013; Sallam and William Tong, 2004). It has been isolated from almost all HBV carriers in Iran, Afghanistan, Pakistan, and Turkey and it is predominant in neighbouring countries as Republic of Yemen (90%), but also prevalent in Egypt, Turkey, Iran, and Tunisia (Bahri et al., 2006; Al-Qahtani et al., 2020). We have shown in our study the distribution of HBV genotypes, resistance associated mutations, and vaccine escape mutations in HBV patients in Saudi Arabia. Our findings will have a major impact on therapy management and diagnostics of chronic HBV infections in Saudi Arabia to control HBV infection in this intermediate HBV-endemic country. Phylogenetic analysis demonstrates genotype D was predominant of our study in Saudi Arabia, which is in line with previous molecular epidemiology studies in Saudi Arabia (Al-Qahtani et al., 2020). Various pattern is reported in geographical distribution for HBV subgenotypes. As reported; the dominant subgenotype in the Mediterranean region is D1, whereas D2, D3, D4, D5, D6, D7, and D8 subgenotypes are mainly isolated in Eastern Europe, India, Australia, Indonesia, and Africa (Pourkarim et al., 2014; Al-Qahtani et al., 2020). Our analysis showed that the majority of subtype were categorized as D1 and a few cases belonged to the subgenotypes D2, D3, and D10 which is in line with molecular epidemiology studies conducted previously in Saudi Arabia. The data indicated that NGS is a reliable tool for genotype prediction, especially with respect to genotype D, which makes it useful for clinical evaluation.

The current antiviral five NAs have been introduced worldwide for treatment of targeting the viral polymerase (RT domain) in addition to interferon. Even though, the introduction of NAs for therapy of CHB significantly decreases the progression of the disease and efficiently controls HBV replication, unfortunately, NAs can select resistance mutations in the RT-domain which is a major limiting factor for long-term viral suppression (Tujios and Lee, 2013).

As indicated, the prevalence of resistance mutations rapidly rises to 80% after four years of treatment especially in LAM. While ADV, ETV, LdT, and TDF, demonstrate higher genetic barriers with a resistance development. For instance, resistance to TDF arise in 0% to 5.9% after three years of therapy (Locarnini and Yuen, 2010)

Little is known about genetic diversity in Saudi Arabia, the prevalence of antiviral resistance, and surface antigen vaccine escape mutations. Therefore, we determined the prevalence of HBV genotypes among different sequences, we explored the prevalence of antiviral treatment resistance mutations and the prevalence of ‘‘a’’ determinant vaccine escape mutant. It is hoped that the project will contribute to national guidelines for the eradication of HBV in Saudi Arabia. Though, a surveillance program to observe the situation of circulating HBV escape mutations in Saudi Arabia is needed to determine the prevalence of HBV escape mutations more comprehensively.

In Saudi Arabia, LAM is widely used and ADV is an alternative drug of choice. Therefore, the possibility of detecting pre-existing HBV variants resistant to these drugs is more when compared to other drugs introduced rather recently. The mutational analysis showed drug resistance identified in ten different mutations that are associated mainly with LAM resistance. All mutations were identified in samples with genotype D. From this analysis, drug resistance-associated mutations (RAMs) were more likely to be present in genotype D sequences; this finding is in concordance with a previous report (Al-Qahtani et al., 2020). The mutational analysis indicated that 6 samples exhibited NA resistance, detected in nine different mutations that linked mainly with LAM resistance (rtQ215S, rtD134N/E, rtF221Y, rtR153W, rtM204V/I, rtL80I, rtI91L). These RAMs were selected because they are primary RAMs or have strong evidence in causing resistance to LAM, ADV and or TDF. Remarkably, one sample presented three mutations (rtL80I, rtL91I and rtM204I). The N236T ADV resistance mutation and resistance mutations against other antiviral drugs (e.g., ETV and LdT) could not be detected in any sample.

RTM204I/V mutations were detected and are one of the best-recognised drug resistance motifs in HBV within our whole sequence database. The prevalence of rtM204I/V as 4.9% among >12,000 treatment-naïve individuals were estimated in a previous meta-analysis (Zhang et al., 2015) and another review reported a prevalence of rtM204I/V of 5.9% among 8,435 treatment-naïve individuals (Choi et al., 2018). RTQ215S mutation was detected, and it frequently occurs during LAM therapy to restore the replication capability. It is also called secondary or compensatory mutations (Ismail et al., 2012). Also, rtR153W mutation was identified and reported previously in association with TDF (Mokaya et al., 2020). The Seven mutations in rt

region (rtL80I, rtI91L, rtD134N/E, rtM204I/V, and rtF221Y) were reported substantially associated with the progression of HCC (Choi et al., 2018).

In addition, we have described all various vaccine/immune and diagnostic escape mutations which are detectable mainly in the most antigenic MHR located in the region of aa124– aa147. As reported; mutations G145R, D144E/A, and P120T mutations are the best-described and stable vaccine escape mutations (Locarnini et al., 2003; Bock et al., 2002) while G145R substitution is the most stable and frequent amino acid change in HBsAg worldwide. This is known to cause an immune escape during vaccination and is recorded to be closely associated with genotype D viruses (Al-Qudari et al., 2016). Although our phylogenetic analysis demonstrated that most of the samples in Saudi Arabia were of genotype D, none of these identified HBV genomes had G145R mutation.

The mutations in HBsAg are mainly caused by the following aspects: vaccine and/or hepatitis B immunoglobulin (HBIG) administration, natural selection by the host immune response, and antiviral therapy pressure inducing antiviral resistance mutants which alter simultaneously the amino acid composition of the surface antigen (Ghany et al., 10998; Khedive et al., 2013). Furthermore, the mutations in the 'a' determinant in HBsAg can lead to several dilemmas include;(i) Missed diagnosis and the current serological screening assays miss the target antigen; (ii) conformational changes with altered antigenicity and failure to neutralize the virus; (iii) associated with liver disease progression like fulminant hepatitis cirrhosis and HCC (Al Baqlani et al., 2014). In our study, we detected 12 mutations in the MHR (I110L, P120S, P120T, T126S, T131N, M133I, M133T, Y134F, Y134K, S143T, F161Y and A168V). Six mutations had amino acid substitutions (T126S, T131N, M133I/T, Y134F and S143T) within the 'a' determinant region (aa 124-147) of the MHR. Our result is in good agreement with the latest reports showing a prevalence of HBsAg vaccine escape mutations (Al-Qudari et al., 2016; El-Kafrawy et al., 2018). Well-described vaccine escape mutation P120T was found only in one sample while the classical vaccine escape mutations D144E and G145R, which were reported to be closely associated with genotype D, could not be detected in any sample. The mutations (I110L and F161Y) were identified and both mutations lead to vaccine escape. Likewise, substitutions (P120S, P120T, T126S and S143T) were also detected and may lead to problems in both diagnostic assays /vaccine escape (Coppola et al., 2015; Guangxi et al., 2015). M133T mutation was found and may lead to problems in diagnostic assays; however, it has yet to be categorized as VEMs (Yan et al., 2017). Y134K mutation was also found and had not been identified in the usual escape setting (Ziaee et al., 2016).

However, the study had few limitations. the lack of sample metadata including patient identifiers (e.g. age, gender, etc) and the drug history of the respective patients, the lack of these data led to difficulty in linking the genomic data to clinical outcomes. Another limitation was difficulty in obtaining geographically representative samples from the western region of Saudia Arabia to ensure a representative sample for the prevalence of RAMs and S gene mutations on clinical outcomes of the disease.

In order to maintain the genomic surveillance of HBV, a repository (or database) of standardised reference genomes of all HBV genotypes and subtypes is needed to facilitate the consistent assembly and analysis required to develop insights into current and future epidemiology, inform better clinical assessment, improve deployment of current antiviral drugs and vaccines, and drive the discovery of new antiviral agents.

In the second part of the thesis, we established a successful an *in vitro* HBV replicon-based drug testing platform. This tool has been demonstrated to be useful for testing of future antiviral drugs as the *in vitro* HBV cell culture scheme is a crucial tool for anti-HBV drugs screening and for exploring the natural properties of HBV (Xu et al., 2021). Several constructs were established using pTHBV1.3-L^- that is carrying 1.3 genome length of HBV (*ayw* subtype). As this plasmid viral replication competent. Various modifications were introduced into this plasmid with the purpose to study the effectiveness of NA on genomes carrying drug-resistance-associated mutations of interest.

Replication of HBV in Huh7 cell line using integrated or transfected HBV genomes as templates has been achieved also well as shown occur with various constructs developed during this project, evidence of virus DNA replication was confirmed by expression of HBcAg and HBsAg in cell lysates using Western blot analysis to HBeAg in cell medium using ELISA method.

In order to increase the throughput of the replication assay, in this study, an ELISA assay was developed and verified for the expression of HBeAg using monoclonal antibodies and antibody conjugates. When the in house HBeAg ELISA system was compared with the other commercial kit, it was shown that our system corresponded with the results of negative and positive samples at similar ratio using pQ-HBV2.7 (3B4) and pHBV1.3mer WT. Suggesting that the in-house ELISA developed in this project is equally effective as commercially available ELISAs for the screening of plasmids encoding the expression of HBeAg. Thus, this system may be a useful tool for screening purposes and outbreak investigations.

Then we investigated the mutations were identified previously in chapter 3 for antiviral drugs screening using the developed protocol. We explored only two mutations in reverse transcriptase domain due to time constraint (rtD134E and rtD134N) which were reported previously to be antiviral resistance-associated mutations. Three constructs were developed in chapter 5 in order to screen the effectiveness of the antiviral using the expression HBeAg as a marker of replication. The result indicated that both mutations (rtD134E and rtD134N) did not confer resistance to ADV and TDF while it confer resistance to the other agents (LdT, LAM, and ETV). These results were comparable to the previous studies. Additionally, the expression of HBcAg was validated as a marker to confirm the antiviral drug in the transfection cells. These outcomes provide unique insights into the progress of antiviral resistance.

6.3 Research impact

We validated a protocol for deep whole-genome sequencing for HBV using the target enrichment method. To the best of our knowledge, this is the first description of whole-genome HBV sequencing using this method. Our study will serve as one of the initial efforts to garner a better comprehension of the dynamics of HBV circulating variants and link it to the epidemiology of the infection in Saudi Arabia.

Likewise, the *in vitro* system that has been established in this thesis could support the antiviral screening of the future drugs while we were able to target and measure the viral proteins such as HBeAg and HBcAg.

6.4 Concluding remarks

6.4.1 Future Work

All PhD, open up further research and optimisation. The onset of the COVID-19 pandemic affected the finalisation of a number of sections of this thesis. The following should be considered;

- Further optimisation and validation of the drug screening capacity of the replicon-based system. The assays were only undertaken as significates and the wildtype constructs were not as sensitive as we expected.
- Once optimised, test other known drugs/antiviral that inhibit HBV replication and test new potential antiviral for their activity using a safe method at CL2.

- Generate new constructs to validate antiviral resistant marker such as we were unable to confirm the HBeAg expression in remaining constructs.
- Comparison of the NGS with MinION to generate a cheaper alternate approach to screen for potential genotype, and drug resistant markers

6.4.2 Prospective of HBV in Saudi Arabia

Future studies should be performed with large sample size covering wider geographic distribution with a linkage to medical meta data to provide a correlation beween genomic data and the disease outcome. Epidemiological and genotyping mapping for HBV should be implemented and updated. Studies on genotype D and improvement in liver disease should be carried out and the association between clinical presentation and antiviral therapy should be shown in future studies.

6.4.3 Prospective of HBV in Worldwide

In order to drive progress towards 2030 elimination targets, sustainable long-term investment is required to drive improvements in clinical services that include a screening of infection, appropriate targeted clinical monitoring for treated patients with monitoring of drug resistance, provide universal access to antenatal screening, expand consistent drug and vaccine supply, improve the education of the public and health-care workers on HBV prevention, diagnosis and treatment.

Finally, if drug resistance emerges as a substantial clinical problem, there will be a need for consideration of synergistic drug regimens, new agents that inhibit a target other than viral rt, and for the development of new therapeutic strategies that can bring about the cure.

APPENDICES

APPENDIX 1: Reported resistance mutations to Lamivudine (LAM)

Mutation	Subdomain	Genotype	References
A181T	B	A, B, C, D	Strasfeld & Chou, 2010 ; Liu et al., 2010 ; Lim, 2017; Choi et al., 2018
A181V	B	A, B, C, D	Strasfeld & Chou, 2010 ; Liu et al., 2010 ; Salpini et al., 2011 ; Choi et al., 2018, Lim, 2017 ; Qian et al., 2016
A200V	C	B, C	Choi et al., 2018
C256G	E	B, C	Choi et al., 2018
D134E		B, C, D	Cho et al., 2018
F166L	B		Choi et al., 2018
I91L	A	B, C, D	Choi et al., 2018
L180M	C	A, B, C, D	Salpini et al., 2011 ; Liu et al., 2010 ; He et al., 2015; Lok et al., 2017; Choi et al., 2018; Hermans et al., 2016
L229F	B	B, C, D	Choi et al., 2018
L229G	B	B, C	Choi et al., 2018 ; Liu et al., 2010
L229V	B	B, C	Choi et al., 2018; Liu et al., 2010
L229W	B	B, C	Choi et al., 2018; Liu et al., 2010
L80I	A	A, B, C, D	Lok et al., 2017; Strasfeld & Chou, 2010; Choi et al., 2018
L80V	A	A, B, C, D	Strasfeld & Chou, 2010; Al Baqlani et al., 2014; Choi et al., 2018, Yamani et al., 2017
L82M	A	C	Liu et al., 2010; Choi et al., 2018
L91I	A	C	Wasityastuti et al., 2016
M204I	C	A, B, C, D	Zoulim & Locarnini, 2009; Hermans et al., 2016; Lim, 2017; Ciftci et al., 2014; He et al., 2015; Al Baqlani et al., 2014; Choi et al., 2018; Hermans et al., 2016
M204S	C	A, B, C, D	Zoulim & Locarnini, 2009; Hermans et al., 2016; Lim, 2017; Ciftci et al., 2014; He et al., 2015; Al Baqlani et al., 2014; Choi et al., 2018
M204V	C	A, B, C, D	Zoulim & Locarnini, 2009; Hermans et al., 2016; Lim, 2017; Ciftci et al., 2014; He et al., 2015; Al Baqlani et al., 2014; Choi et al., 2018
N236A	D	B	He et al., 2015
N236T	B		Lim, 2017
Q215E	C	A, B, C, D	Choi et al., 2018
Q215H	C	A, B, C, D	Choi et al., 2018

Mutation	Subdomain	Genotype	References
Q215P	C	A, B, C, D	Choi et al., 2018; Liu et al., 2010
Q215S	C	D	Liu et al., 2010; Salpini et al., 2011; Al Baqlani et al., 2014
R153Q		C	Wasityastuti et al., 2016; Yamani et al., 2017
S202C	C	A, B, C, D	Lok et al., 2017; Choi et al., 2018
S202G	C	A, B, C, D	Choi et al., 2018
S202I	C	A, B, C, D	Choi et al., 2018
S213T		A, B, C, D	He et al., 2015
S246A			Wasityastuti et al., 2016
S256G	E	B, C	Choi et al., 2018; Wasityastuti et al., 2016; Yamani et al., 2017
S53N		B, C	Ciftci et al., 2014; Liu et al., 2010; Choi et al., 2018
T184A	B	A, B, C, D	Choi et al., 2018; He et al., 2015; Lok et al., 2017; Yamani et al., 2017
T184C	B	A, B, C, D	Choi et al., 2018; He et al., 2015; Lok et al., 2017; Yamani et al., 2017
T184F	B	A, B, C, D	Choi et al., 2018; He et al., 2015; Lok et al., 2017; Yamani et al., 2017
T184G	B	A, B, C, D	Choi et al., 2018; He et al., 2015; Lok et al., 2017; Yamani et al., 2017
T184I	B	A, B, C, D	Choi et al., 2018; He et al., 2015; Lok et al., 2017; Yamani et al., 2017
T184L	B	A, B, C, D	Choi et al., 2018; He et al., 2015; Lok et al., 2017; Yamani et al., 2017
T184M	B	A, B, C, D	Choi et al., 2018; He et al., 2015; Lok et al., 2017; Yamani et al., 2017
V173L	C	A, B, C, D	Salpini et al., 2011; He et al., 2015; Qian et al., 2016; Choi et al., 2018
V191I		B, C, D	Liu et al., 2010; Salpini et al., 2011; Choi et all
V207G	C	C	Wang et al., 2017
V207I	C	A, D, B, C	Wang et al., 2017; Choi et al., 2018
V207L	C	C	Wang et al., 2017
V207M	C		He et al., 2015
V214A		D, C	He et al., 2015; Taramasso et al., 2011
W153E		B, C, D	Choi et al., 2018
W153Q		B, C, D	Choi et al., 2018
W153R		B, C, D	Choi et al., 2018

APPENDIX 2: Reported resistance mutations to Telbivudine (LdT)

Mutation	Subdomain	Genotype	References
A181T	B	A, B, C, D	Strasfeld & Chou, 2010; Liu et al., 2010; Lim, 2017; Choi et al., 2018
A181V	B	A, B, C, D	Strasfeld & Chou, 2010; Liu et al., 2010; Salpini et al., 2011; Choi et al., 2018
H238T	D		He et al., 2015
I169T	B		Zoulim & Locarnini, 2009
L180I	B		He et al., 2015
L180M	B	A, C, D	Salpini et al., 2011; Liu et al., 2010; He et al., 2015; Wang et al., 2017; Zoulim & Locarnini, 2009; Hermans et al., 2016; Lim, 2017
L80I	A	B	You & Jia, 2013
L80V	A	B	You & Jia, 2013
M204I	C	A, B, C, D	You & Jia, 2013; Zoulim & Locarnini, 2009; Hermans et al., 2016; Ciftci et al., 2014; Wang et al., 2017; Qian et al., 2016
M204S	C	A, B, C, D	Ciftci et al., 2014; Wang et al., 2017; Zoulim & Locarnini, 2009
M204V	C	A, B, C, D	Zoulim & Locarnini, 2009; Choi et al., 2018
M204V	C	A, B, C, D	Zoulim & Locarnini, 2009; He et al., 2015; Lim, 2017; Ciftci et al., 2014; Wang et al., 2017; Choi et al., 2018
N236A	D	B	He et al., 2015
S202G	C		Zoulim & Locarnini, 2009
S202I	C		Zoulim & Locarnini, 2009
S213T	C		He et al., 2015
T184A	B		He et al., 2015
V173L	B		Zoulim & Locarnini, 2009
V207M	C		He et al., 2015
V214A	C	D, C	He et al., 2015

APPENDIX 3: Reported resistance mutations to Adefovir (ADV)

Mutation	Subdomain	Genotype	References
A181T	B	A, B, C, D	He et al., 2015; Lee et al., 2014; Wang et al., 2017; Choi et al., 2018
A181V	B	A, B, C, D	Liu et al., 2010; Lee et al., 2014; Wang et al., 2017; Qian et al., 2016; Choi et al., 2018
A194T	B	B, C, D	Choi et all
B236T	D		Lee et al., 2014
B238D	D		Lee et al., 2014
B238T	D		Lee et al., 2014
E218D	C	A, B, C, D	Choi et al., 2018; Jiang et al., 2014
E218G	C		Liu et al., 2010
F221Y		A, B, C, D	Choi et al., 2018; Jiang et al., 2014
H238T	D	C	He et al., 2015; Wang et al., 2017
I233V	D	A, C, D	Choi et al., 2018, Yamani et al., 2017
L180I	B		He et al., 2015
L180M	B		He et al., 2015
L217R		A, B, C, D	Choi et al., 2018; Jiang et al., 2014
M204I		A, B, C, D	Choi et al., 2018
M204S		A, B, C, D	Choi et al., 2018
M204V		A, B, C, D	Choi et al., 2018
N236A	D	B, C	He et al., 2015
N236T	D	B, C	He et al., 2015; Zoulim & Locarnini, 2009; Qian et al., 2016
N238D	D	B, C	Choi et al., 2018
N238S	D	B, C	Choi et al., 2018
N238T	D	B, C	Choi et al., 2018
N248H		B, C	Wasityastuti et al., 2016; Yamani et al., 2017
P237H	D	B, C	Lee et al.,2014; Choi et al., 2018
Q215E		A, B, C, D	Choi et al., 2018

Mutation	Subdomain	Genotype	References
Q215H		A, B, C, D	Choi et al., 2018
Q215P		A, B, C, D	Choi et al., 2018
Q215S		C	Wang et al., 2017; Micco et all
S213T			He et al., 2015
S256G	E	C	Wasityastuti et al., 2016
S85A	A		Jiang et al.,2014; Al Baqlani et al., 2014
S85A		B, C	Choi et al., 2018
T184A	B	A	He et al., 2015; Lok et al., 2017
T54N		A, B, C, D	Choi et al., 2018
V191I		B, C, D	Choi et al., 2018
V207M	C		He et al., 2015
V214A		B, C, D	He et al., 2015; Wang et al., 2017; Choi et al., 2018
V84M		B, C	Choi et al., 2018
Y126C		A, B, C, D	Choi et al., 2018
Y126H		A, B, C, D	Choi et al., 2018
Y126R		A, B, C, D	Choi et al., 2018
Y221F		C	Wasityastuti et al., 2016; Yamani et al., 2017
Y245H		B, C	Choi et al., 2018

APPENDIX 4: Reported resistance mutations to Entecavir (ETV)

Mutation	Subdomain	Genotype	References
A181V	B	A	Lok et al., 2017
A194G	B		Qian et al., 2016
C256G	E	B, C	Choi et al., 2018
H238T	D		He et al., 2015
I169T	A	B, C	Ciftci et al., 2014; Locarnin & Devi, 2013; Zoulim & Locarnini, 2009; Tenney et al., 2007; Choi et al., 2018
L180I	B	A, B, C, D	He et al., 2015
L180M	B	A, B, C, D	Ghany & Doo, 2009; Locarnin & Devi, 2013; He et al., 2015; Zoulim & Locarnini, 2009; Tenney et al., 2007; Choi et al., 2018
M204I	C	A, B, C, D	He et al., 2015; Zoulim & Locarnini, 2009; Lim, 2017
M204S	C	A, B, C, D	Choi et al., 2018; He et al., 2015
M204V	C	A, B, C, D	He et al., 2015; Zoulim & Locarnini, 2009
M250I	E	A, B, C, D	Choi et al., 2018
M250L	E	B	Ciftci et al., 2014; Locarnin & Devi, 2013
M250V	E	B	Ghany & Doo, 2009; Baldick et al., 2008; Ciftci et al., 2014; Locarnin & Devi, 2013
N236A	D	B	He et al., 2015; Baldick et al., 2008
S202C	C	C	Locarnin & Devi, 2013; Zoulim & Locarnini, 2009; Lok et al., 2017
S202G	C	A, B, C, D	Locarnin & Devi, 2013; Zoulim & Locarnini, 2009; Baldick et al., 2008; Choi et al., 2018
S202I	C	A, B, C, D	Ghany & Doo, 2009; Zoulim & Locarnini, 2009; Zoulim and Locarnini. Baldick et al., 2008; Choi et al., 2018
S202N	C		Qian et al., 2016
S213T	C	A, B, C, D	He et al., 2015; Baldick et al., 2008; Choi et al., 2018
S256G	E	B, C	Yamani et al., 2017; Choi et al., 2018
T184A	B	A, B, C, D	Ciftci et al., 2014; Locarnin & Devi, 2013; Z., 2008ulim & Locarnini, 2009; Tenney et al.,2007; Baldick et al., 2008
T184C	B	A, B, C, D	Ciftci et al., 2014; Locarnin & Devi, 2013; Zoulim & Locarnini, 2009; Tenney et al., 2007
T184F	B	A, B, C, D	Ciftci et al., 2014; Locarnin & Devi, 2013; Zoulim & Locarnini, 2009; S.Baldick et al., 2008
T184G	B	A, B, C, D	Ciftci et al., 2014; Baldick et al., 2008; Ghany & Doo, 2009; Locarnin & Devi, 2013; Zoulim & Locarnini, 2009

Mutation	Subdomain	Genotype	References
T184I	B	A, B, C, D	Ciftci et al., 2014; Locarnin & Devi, 2013; He et al., 2015; Baldick et al., 2008; Lok et al., 2017; Zoulim & Locarnini, 2009
T184L	B	A, B, C, D	Ciftci et al., 2014; Locarnin & Devi, 2013; Zoulim & Locarnini, 2009; Tenney et al., 2007; Baldick et al., 2008
T184S	B	A, B, C, D	Ciftci et al., 2014; Locarnin & Devi, 2013; Zoulim & Locarnini, 2009; Tenney et al., 2007; Baldick et al., 2008
V173L	C	D, B, C	He et al., 2015
V207M	C		He et al., 2015
V214A	C	D, C	He et al., 2015

APPENDIX 5: Reported resistance mutations to Tenofovir (TDF)

Mutation	Subdomain	Genotype	References
A181V	B	D, C	Strasfeld & Chou, 2010; Liu et al., 2010; Salpini et al., 2011; Choi et al., 2018
A194T	B	B, C, D	Choi et al., 2018; Yamani et al., 2017
D134E		B, C, D	Cho et al., 2018
H126Y			Cho et al., 2018
H238T	D		He et al., 2015
L180I	C	D	He et al., 2015
L180M	C	A, D	Salpini et al., 2011; Liu et al., 2010; He et al., 2015; Zoulim & Locarnini, 2009; Hermans et al., 2016; Lim, 2017
L80I	A	B	You & Jia, 2013
L80V	A		You & Jia, 2013
M204I	C	A, B, C, D	Choi et al., 2018; Zoulim & Locarnini, 2009; Hermans et al., 2016; Lim, 2017
M204S	C	A, B, C, D	Ciftci et al., 2014; Zoulim & Locarnini, 2009; Hermans et al., 2016; Lim, 2017
M204V	C	A, B, C, D	Zoulim & Locarnini, 2009; Hermans et al., 2016; Lim, 2017; Ciftci et al., 2014
N236A	D	B	He et al., 2015
N236T	D	B, C	Choi et al., 2018
S106C	A		Choi et al., 2018
S213T	C		He et al., 2015
T184A	B		He et al., 2015
V207M	C		He et al., 2015
V214A	C	D, C	He et al., 2015; Taramasso et al., 2011

APPENDIX 6: Reported resistance mutations affecting HBsAg

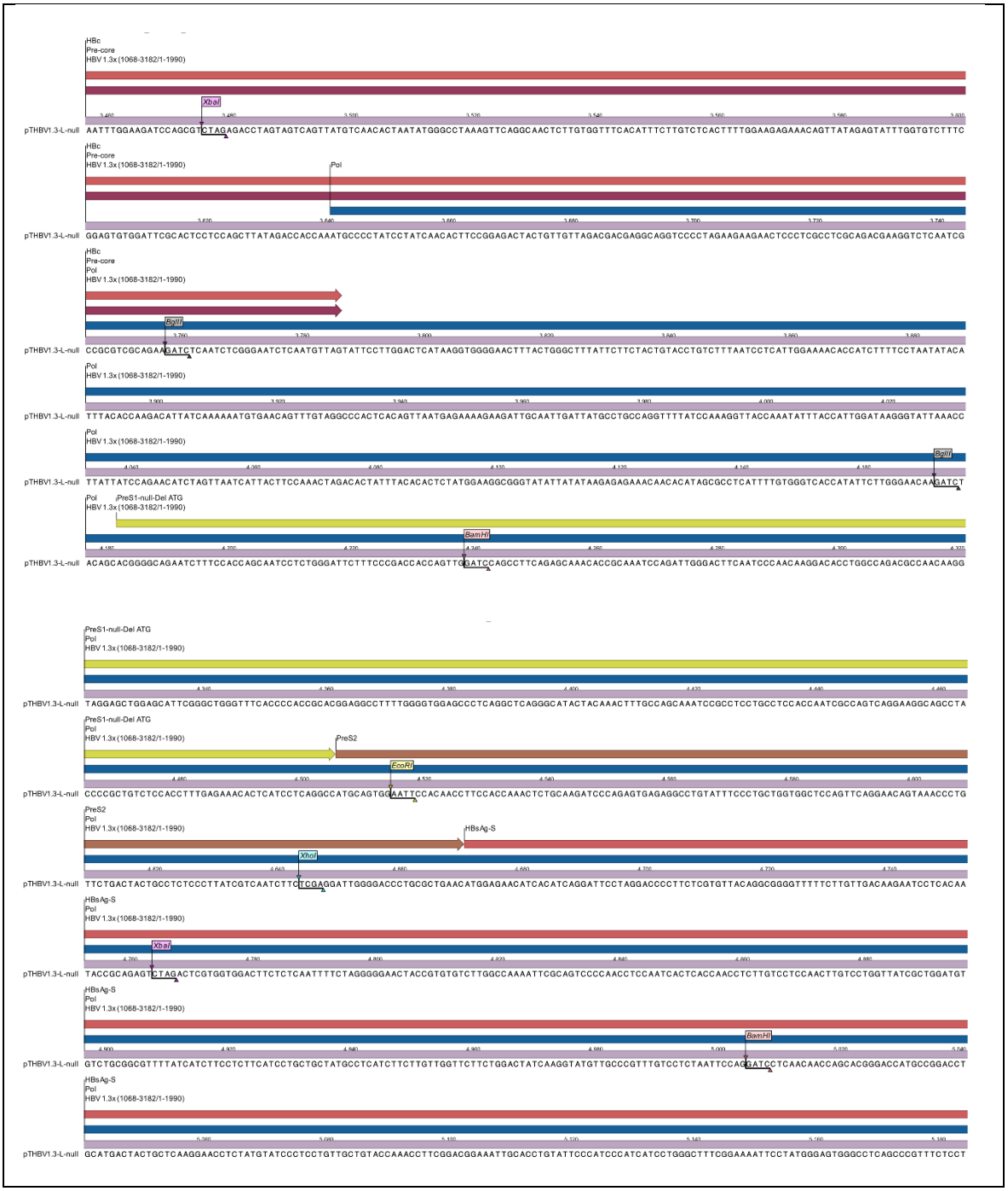
Amino Acid Position	Native Amino Acid	Escape Mutation	References	Note
3	N	S	Wang et al., 2017	
4	W	P/R	Caligiuri et al., 2016	
4	T	K/A/V	Wang et al., 2017	
5	A	T/S	Wang et al., 2018	
5	T	A	Wang et al., 2019	
10	G	Q/R/E	Wang et al., 2020	
21	L	S	Wang et al., 2021	
46	P	H	Perazzo et al., 2015	
47	T	K	Wang et al., 2017	
48	C	G	Caligiuri et al., 2016	major hydrophilic loop
49	L	H	Perazzo et al., 2015	
69	C	Y	Perazzo et al., 2016	
90	C	Y	Perazzo et al., 2017	
95	L	W	Perazzo et al., 2018	
98	L	V/P	Wang et al., 2017	
98	L	V	Wang et al., 2018	
100	Y	C	Sayan et al., 2009	
103	M	I	Caligiuri et al., 2016; Ambachew et al., 2017	outside a determinant region, major hydrophilic loop
107	C	R	Perazzo et al., 2015	
109	L	I	Caligiuri et al., 2016; Sayan et al., 2009	major hydrophilic loop
109	L	R	Jaramillo et al., 2017	
110	I	V	Sayan et al., 2009	
110	I	V/B	Sayan et al., 2010	
110	L	L	Perazzo et al., 2015	
110	I	L	Ambachew et al., 2017; Mokaya et al., 2018	HBIg, outside a determinant region
114	T	T	Perazzo et al., 2015	
115	T	I	Ambachew et al., 2017	outside a determinant region
116	T	N	Caligiuri et al., 2016; Mokaya et al., 2018	VEM, a determinant region
117	S	I/N/S/T	Sayan et al., 2009	
118	T	A/P	Ambachew et al., 2017	outside a determinant region
118	T	K	Caligiuri et al., 2016	major hydrophilic loop
119	G	R	Mokaya et al., 2018	HBIg
119	G	E/R	Ambachew et al., 2017	outside a determinant region

Amino Acid Position	Native Amino Acid	Escape Mutation	References	Note
120	P	S/E/T/Q/A	Mokaya et al., 2018; Caligiuri et al., 2016; Mokaya et al., 2018; Jaramillo et al., 2017; Perazzo et al., 2015; Sayan et al., 2009	VEM, a determinant region, major hydrophilic loop
121	I	L	Perazzo et al., 2015	
124	C	R/Y	Perazzo et al., 2015	
125	T	A	Perazzo et al., 2015	
126	T	S/I/N	Perazzo et al., 2015; Mokaya et al., 2018	
126	I	S/T	Perazzo et al., 2015	
126	I/T	A/F/N/S	Mokaya et al., 2018	VEM + HBIg
126	I	T	Mokaya et al., 2018	HBIg
126	T	I/N	Ambachew et al., 2017	a determinant region
126	I/T	A/N/I/S	Caligiuri et al., 2016	a determinant region
126	T/A	S	Wang et al., 2017	
126	I/S	N/V/T	Wang et al., 2018	
126	I/S	T	Wang et al., 2019	
126	T	A/S	Jaramillo et al., 2017	
127	P	Tb	Sayan et al., 2009	
129	N	R	Perazzo et al., 2015	
129	Q	H	Perazzo et al., 2015	
129	Q	R	Mokaya et al., 2018	VEM
129	Q	H/R	Caligiuri et al., 2016	a determinant region
129	Q	H/R/L	Wang et al., 2017	
130	G	E/R/N	Jaramillo et al., 2017; Sayan et al., 2009; Perazzo et al., 2015	genotype F
131	N	T	Ambachew et al., 2017	a determinant region
131	T/N	I/A	Wang et al., 2017	a determinant region
131	T	P/I/N/A	Wang et al., 2017; Mokaya et al., 2018; Perazzo et al., 2015	HBIg, a determinant region
132	S	A	Sayan et al., 2009	
133	M	I/K/L/T/V	Sayan et al., 2009	genotype D, E
133	M	K	Perazzo et al., 2015	
133	M	L	Perazzo et al., 2015	
133	M	T	Mokaya et al., 2018	genotype D
133	M	L	Mokaya et al., 2018	VEM
133	M	T	Mokaya et al., 2018	HBIg, genotype E
133	M	V	Ambachew et al., 2017	a determinant region
133	M	L	Caligiuri et al., 2016	a determinant region
134	Y	N	Sayan et al., 2009	
134	F/P	I/V	Mokaya et al., 2018	VEM
134	F	Y	Ambachew et al., 2017	a determinant region
134	Y	H	Caligiuri et al., 2016	major hydrophilic loop

Amino Acid Position	Native Amino Acid	Escape Mutation	References	Note
137	C	G/L	Sayan et al., 2009	
138	C	R	Perazzo et al., 2015	
139	N	R	Perazzo et al., 2015	
139	C	Y	Perazzo et al., 2015	
140	S	T	Perazzo et al., 2015	
141	K	E	Caligiuri et al., 2016; Mokaya et al., 2018	VEM, a determinant region
142	P	S	Caligiuri et al., 2016; Mokaya et al., 2018	VEM, a determinant region
143	S	L	Perazzo et al., 2015	
143	T	N	Mokaya et al., 2018	VEM
143	S	T	Mokaya et al., 2018	HBIg
143	T	L	Ambachew et al., 2017	a determinant region
143	S	L	Caligiuri et al., 2016; Sayan et al., 2009	major hydrophilic loop
144	D	A/E/H/G/V	Mokaya et al., 2018; Perazzo et al., 2015; Caligiuri et al., 2016; Jaramillo et al., 2017; Sayan et al., 2009	VEM + HBIg, a determinant region , major hydrophilic loop, genotype D, E
145	S	A	Perazzo et al., 2015	
145	G	A/E/K/R/X	Mokaya et al., 2018; Wang et al., 2017; Jaramillo et al., 2017; Perazzo et al., 2015, Sayan et al., 2009	VEM + HBIg, genotype D, E
149	C	R	Mokaya et al., 2018	HBIg
152	I	F	Perazzo et al., 2015	
153	P	T	Perazzo et al., 2016	
158	L	F	Perazzo et al., 2017	
161	F	H/L	Romanò et al., 2014	
161	Y	F/S	Ambachew et al., 2017; Romanò et al., 2013	outside a determinant region
164	E	D/G	Ambachew et al., 2017; Romanò et al., 2014	outside a determinant region
168	A	T	Perazzo et al., 2015	
168	V	A	Ambachew et al., 2017	outside a determinant region
171	S	F	Caligiuri et al., 2016	major hydrophilic loop
172	W	L	Romanò et al., 2014	
173	V	L	Perazzo et al., 2015	
173	L	F	Romanò et al., 2013	
175	L	F	Romanò et al., 2014	
175	L	S/F	Wang et al, 2017; Romanò et al., 2014; Caligiuri et al., 2016	major hydrophilic loop
176	C	R	Perazzo et al., 2015	

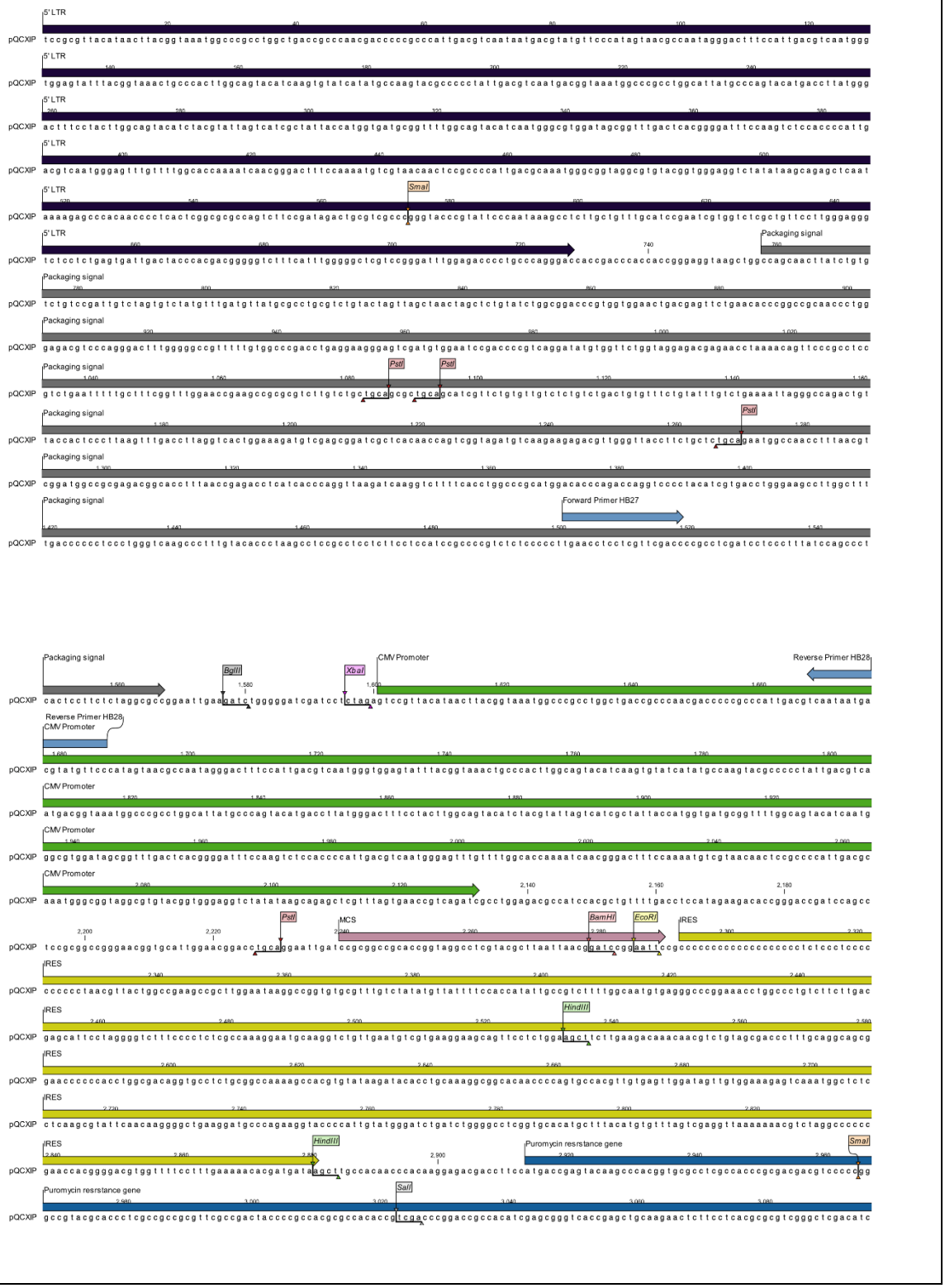
Amino Acid Position	Native Amino Acid	Escape Mutation	References	Note
176	L	V	Romanò et al., 2014	
177	V	A	Ambachew et al., 2017	
178	N	S	Perazzo et al., 2015	
184	V	A	Ambachew et al., 2017	
185	G	E	Caligiuri et al., 201; Perazzo et al., 2015	major hydrophilic loop
189	T	I	Ambachew et al., 2017	
190	V	A	Caligiuri et al., 2016	major hydrophilic loop
190	V	A	Caligiuri et al., 2017	major hydrophilic loop
193	S	F/L	Romanò et al., 2014	
194	A	T	Romanò et al., 2015	
194	V	S	Romanò et al., 2016	
195	I	M	Romanò et al., 2017	
196	W	S/L	Romanò et al., 2018	
202	G	R	Perazzo et al., 2015	
204	N	S	Mokaya et al., 2018	HBIg
209	L	V	Perazzo et al., 2015	
213	L	I/S	Wang et al., 2017	
224	V	A/G	Wang et al., 2017	

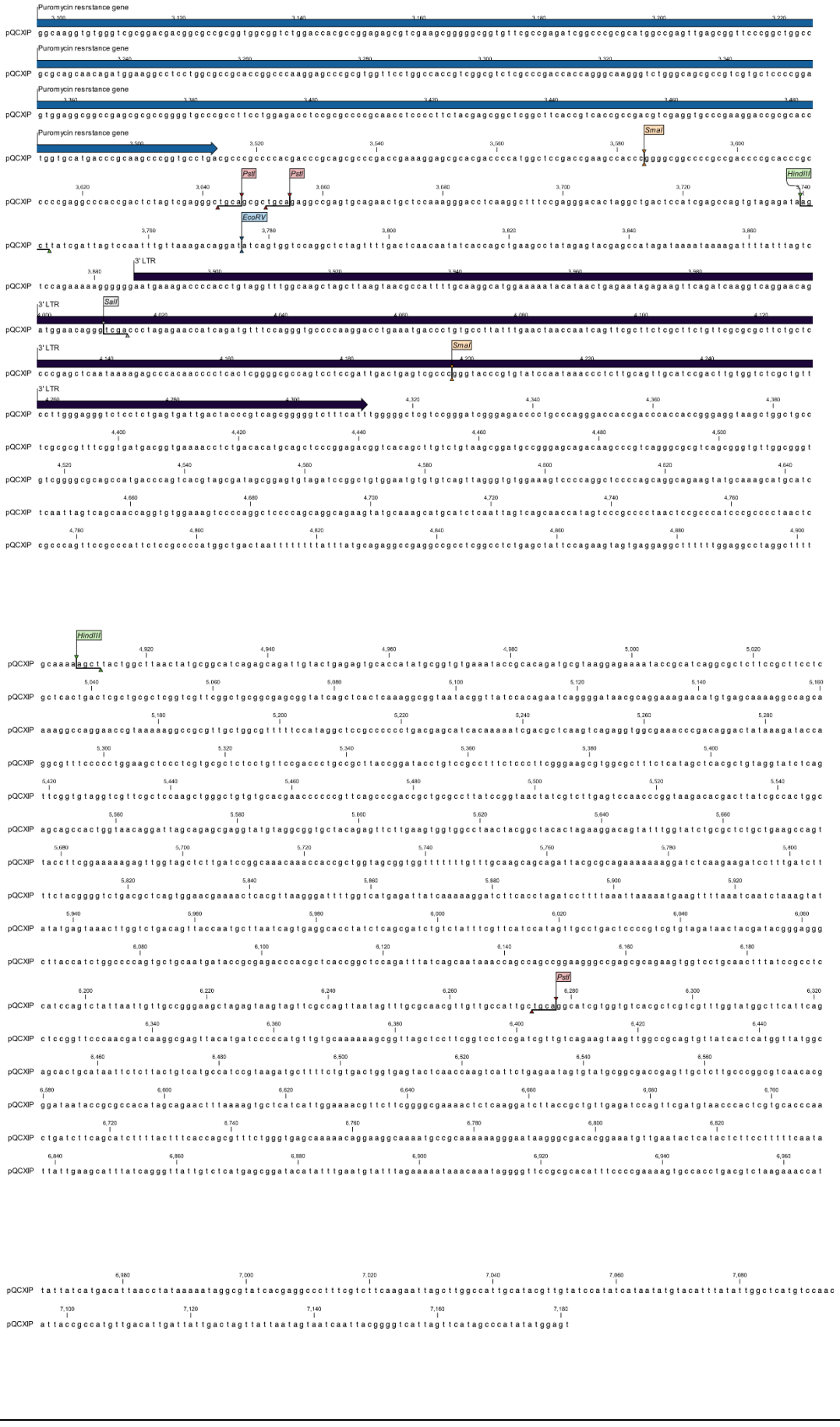
APPENDIX 7: pTHBV1.3-L^{-} sequences


HBc
Pre-core
HBV 1.3x (1068-3182/1-1990)
XbaI
pTHBV1.3-L-null AATTTGGAAGATCCAGCGTCTAGAGACCTAGTAGTCAGTTATGTCAACACTAATATGGGCCTAAAGTTCAGGCAACTCTTGTGGTTTCACATTTCTTGTCTCACTTTTGGAAGAGAAACAGTTATAGAGTATTTGGTGTCTTTC
HBc
Pre-core
HBV 1.3x (1068-3182/1-1990)
Pol
pTHBV1.3-L-null GGAGTGTGGATTCGCACTCCTCCAGCTTATAGACCACCAAATGCCCCTATCCTATCAACACTTCCGGAGACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAGACGAAGGTCTCAATCG
HBc
Pre-core
Pol
HBV 1.3x (1068-3182/1-1990)
BglII
pTHBV1.3-L-null CCGCGTCGCAGAAGATCTCAATCTCGGGAATCTCAATGTTAGTATTCCTTGGACTCATAAGGTGGGGAACTTTACTGGGCTTTATTCTTCTACTGTACCTGTCTTTAATCCTCATTGGAAAACACCATCTTTTCCTAATATACA
Pol
HBV 1.3x (1068-3182/1-1990)
pTHBV1.3-L-null TTTACACCAAGACATTATCAAAAAATGTGAACAGTTTGTAGGCCCACTCACAGTTAATGAGAAAAGAAGATTGCAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACC
Pol
HBV 1.3x (1068-3182/1-1990)
BglII
pTHBV1.3-L-null TTATTATCCAGAACATCTAGTTAATCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGGTCACCATATTCTTGGGAACAAGATCT
Pol PreS1-null-Del ATG
HBV 1.3x (1068-3182/1-1990)
BamHI
pTHBV1.3-L-null ACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGATCCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGG
PreS1-null-Del ATG
Pol
HBV 1.3x (1068-3182/1-1990)
pTHBV1.3-L-null TAGGAGCTGGAGCATTCGGGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGCCTCCACCAATCGCCAGTCAGGAAGGCAGCCTA
PreS1-null-Del ATG
Pol
HBV 1.3x (1068-3182/1-1990)
PreS2
EcoRI
pTHBV1.3-L-null CCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTGGAATTCCACAACCTTCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTG
PreS2
Pol
HBV 1.3x (1068-3182/1-1990)
HBsAg-S
XhoI
pTHBV1.3-L-null TTCTGACTACTGCCTCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAA
HBsAg-S
Pol
HBV 1.3x (1068-3182/1-1990)
XbaI
pTHBV1.3-L-null TACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGT
HBsAg-S
Pol
HBV 1.3x (1068-3182/1-1990)
BamHI
pTHBV1.3-L-null GTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCT
HBsAg-S
Pol
HBV 1.3x (1068-3182/1-1990)
pTHBV1.3-L-null GCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCT

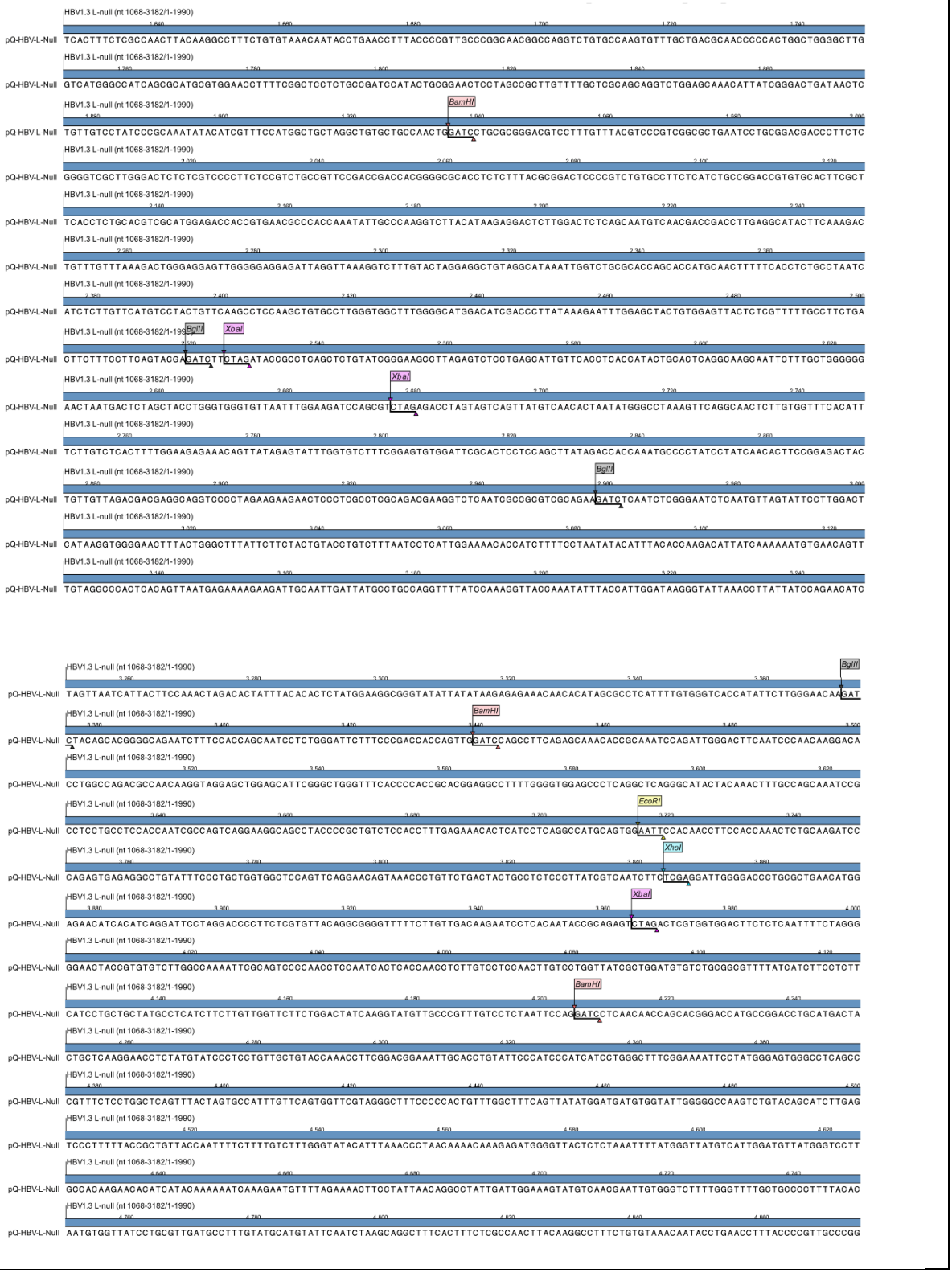

HBsAg-S
Pol
HBV 1.3x (1068-3182/1-1990)
pTHBV1.3-L-null
GGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCT
TTTGTCTTTGGGTATACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTATGTCATTGGATGTTATGGGTCCTTGCCACAAGAACACATCATACAAAAAATCAAAGAATGTTTTAGAAAACT
TCCTATTAACAGGCCTATTGATTGGAAAGTATGTCAACGAATTGTGGGTCTTTTGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCTC
GCCAACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTCATGGGCCATCAGCGCATGCGTGGAAC
CTTTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGTTTCCATGGCTGCTAGGCTG
TGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTC
TCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCA
GCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAAC
TTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTG
ACTTCTTTCCTTCAGTACGAGATC

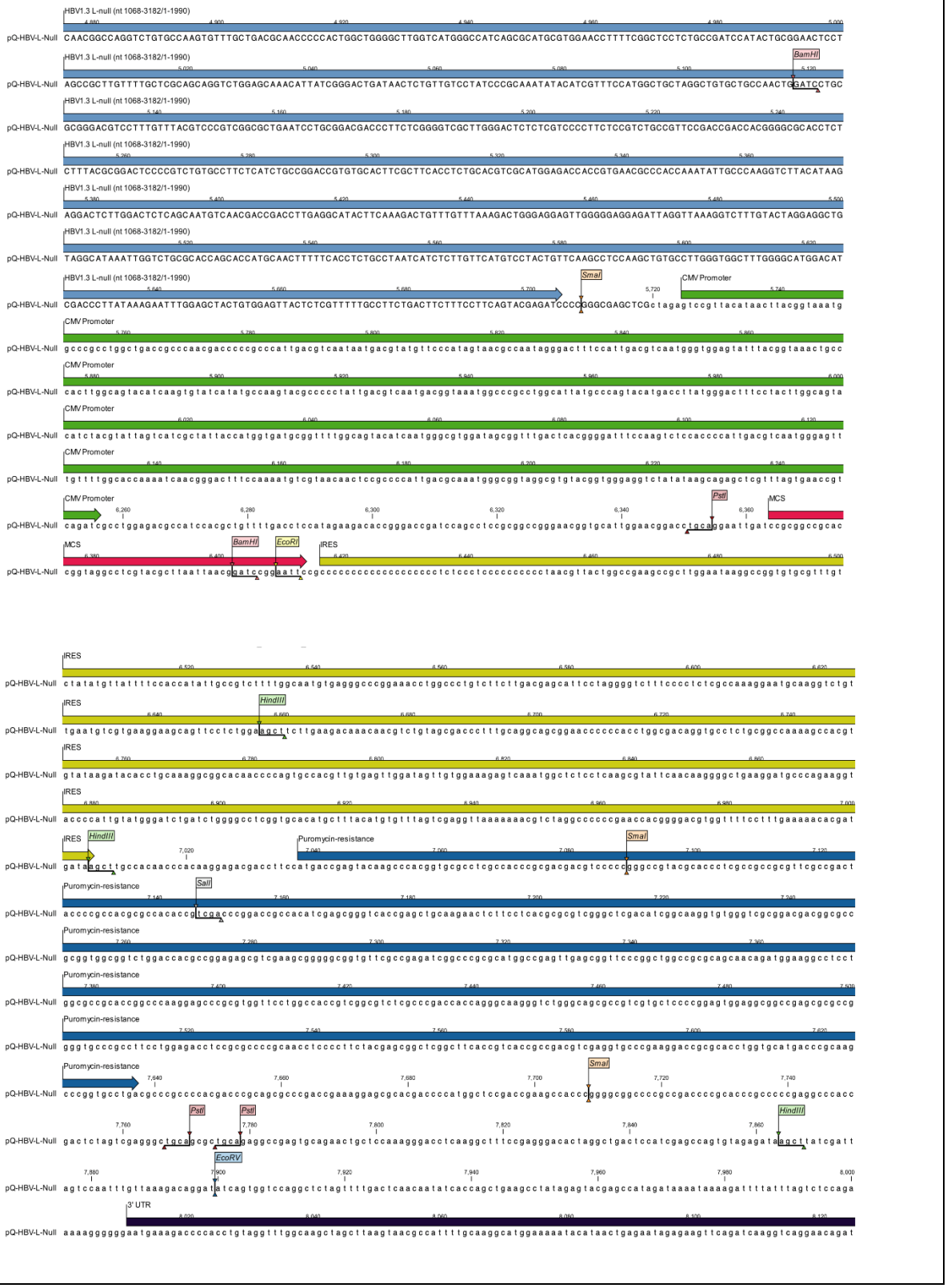
APPENDIX 8: pQCXIP sequences

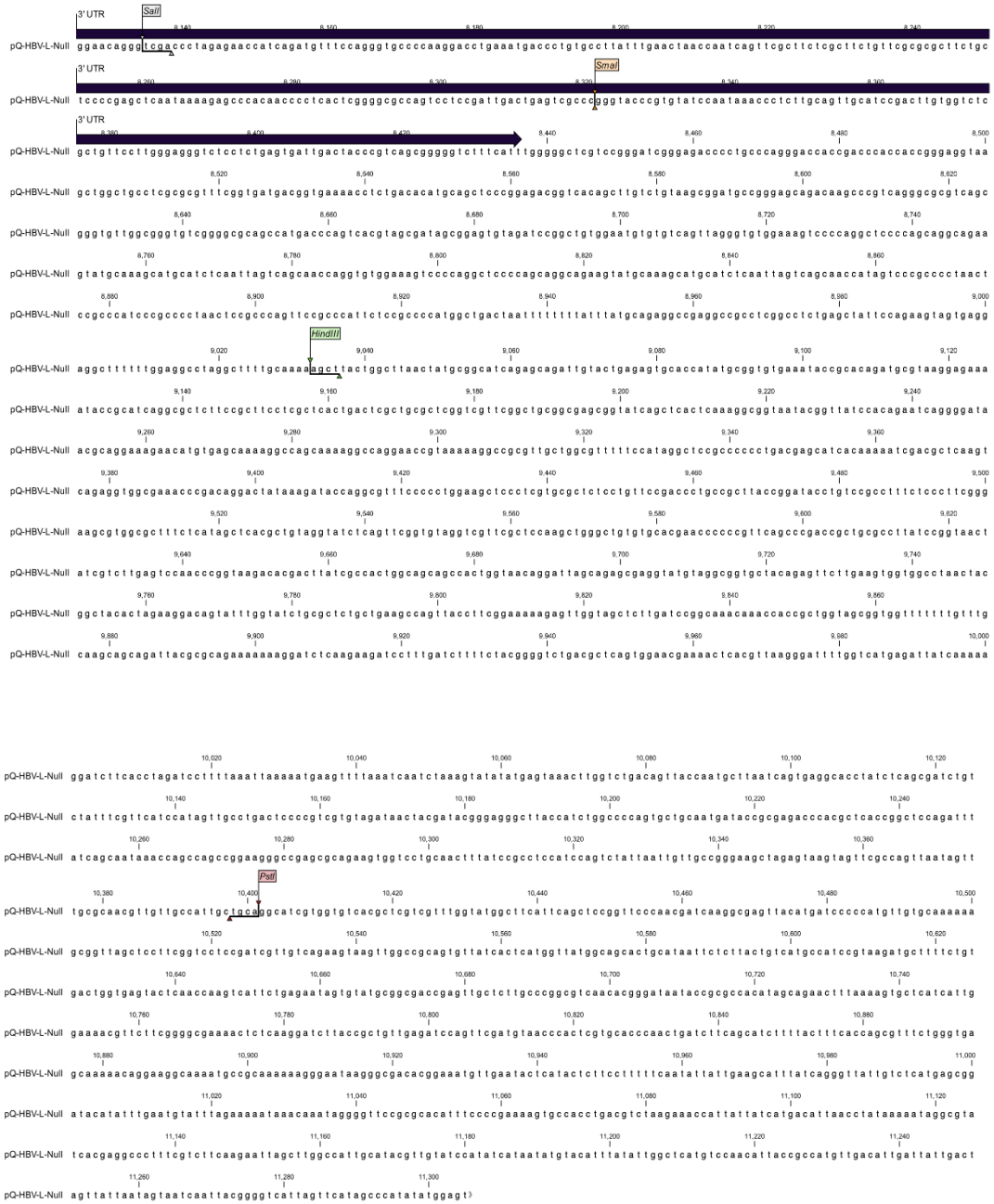



Puromycin resrstance gene
pQCXIP
SmaI
PstI
PstI
HindIII
EcoRV
3' LTR
SalI
SmaI
HindIII
PstI

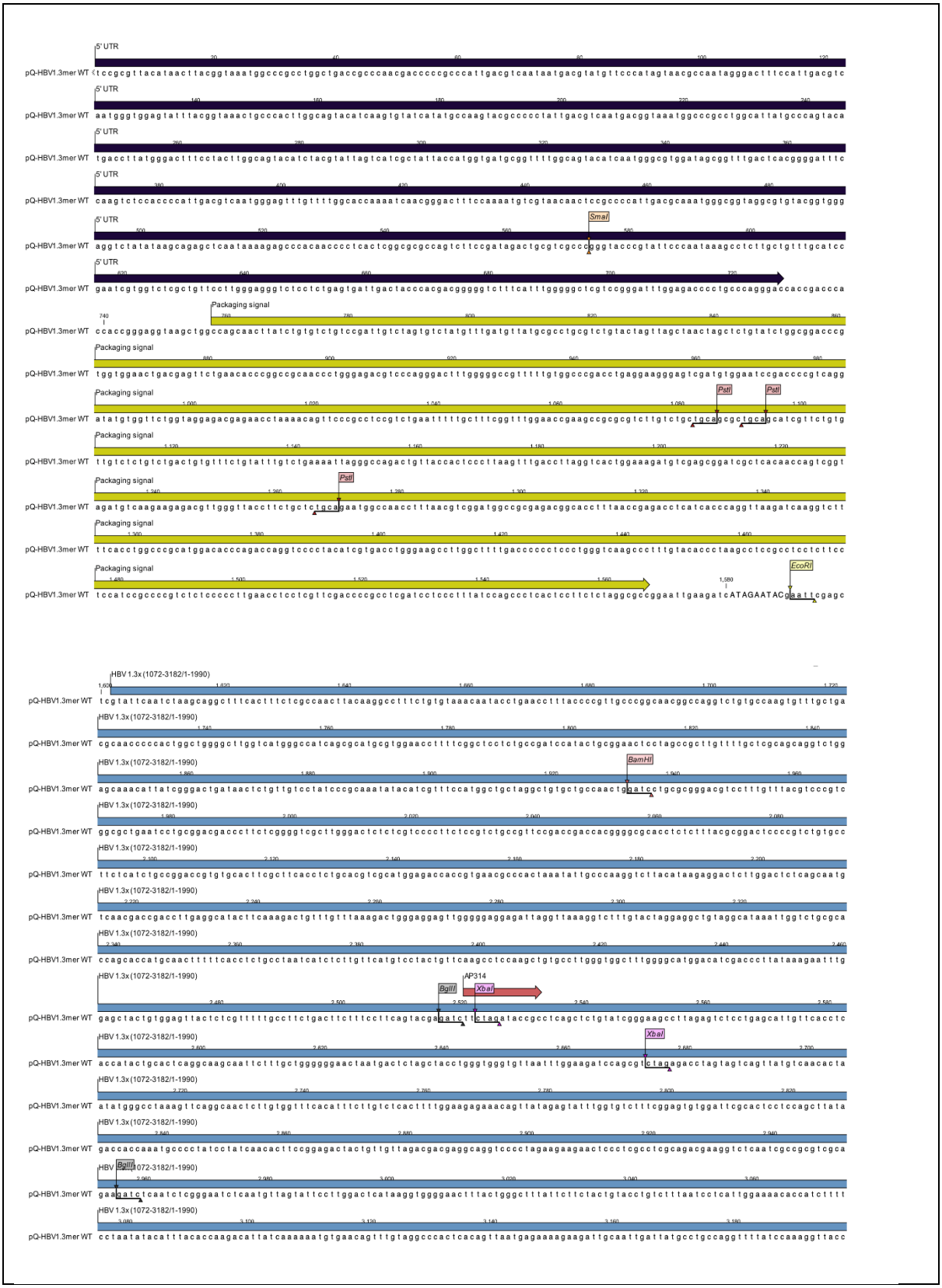
APPENDIX 9: pQ-HBV-L-Null sequences




HBV1.3 L-null (nt 1068-3182/1-1990)
pQ-HBV-L-Null
BamHI
SmaI
CMV Promoter
PstI
MCS
BamHI
EcoRI
IRES
HindIII
Puromycin-resistance
SmaI
SalI
PstI
PstI
HindIII
EcoRV
3' UTR

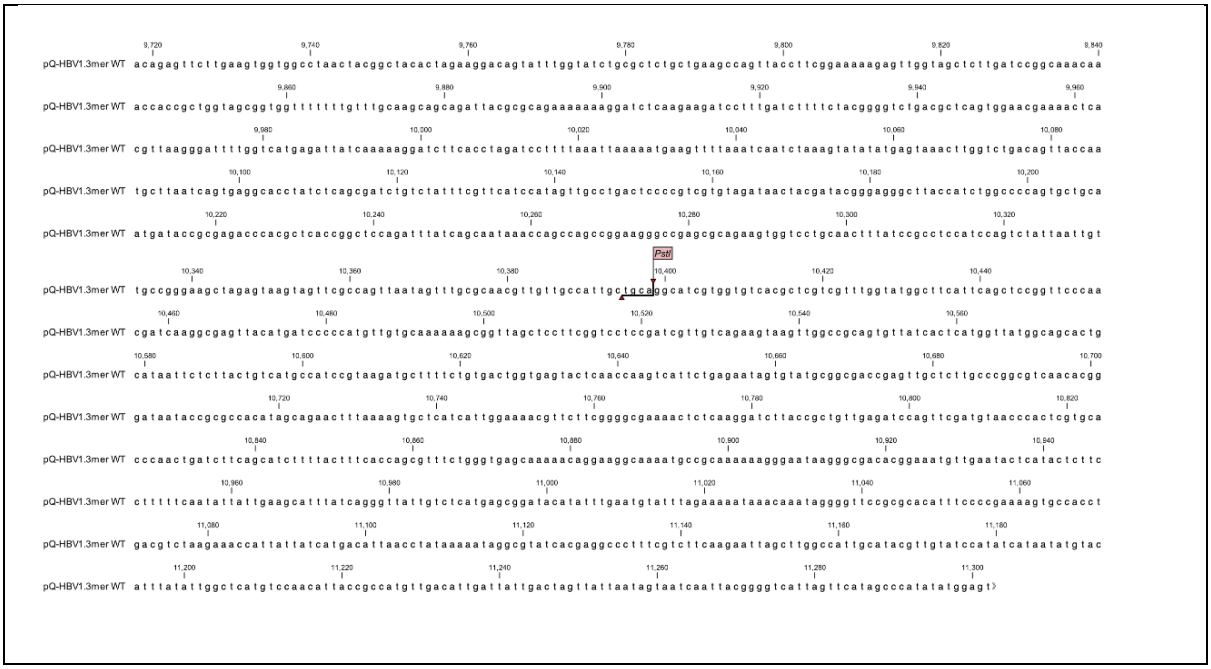

3' UTR
SalI
pQ-HBV-L-Null ggaacagggtcgaccctagagaaccatcagatgtttccagggtgccccaaggacctgaaatgaccctgtgccttatttgaactaaccaatcagttcgcttctcgcttctgttcgcgcgcttctgc
3' UTR
SmaI
pQ-HBV-L-Null tccccgagctcaataaaagagcccacaacccctcactcggggcgccagtcctccgattgactgagtcgcccgggtacccgtgtatccaataaaccctcttgcagttgcatccgacttgtggtctc
3' UTR
pQ-HBV-L-Null gctgttccttgggagggtctcctctgagtgattgactacccgtcagcgggggtctttcatttgggggctcgtccgggatcgggagacccctgcccagggaccaccgacccaccaccgggaggtaa
pQ-HBV-L-Null gctggctgcctcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcagggcgcgtcagc
pQ-HBV-L-Null gggtgttggcgggtgtcggggcgcagccatgacccagtcacgtagcgatagcggagtgtagatccggctgtggaatgtgtgtcagttagggtgtggaaagtccccaggctccccagcaggcagaa
pQ-HBV-L-Null gtatgcaaagcatgcatctcaattagtcagcaaccaggtgtggaaagtccccaggctccccagcaggcagaagtatgcaaagcatgcatctcaattagtcagcaaccatagtcccgcccctaact
pQ-HBV-L-Null ccgcccatcccgcccctaactccgcccagttccgcccattctccgccccatggctgactaatttttttatttatgcagaggccgaggccgcctcggcctctgagctattccagaagtagtgagg
HindIII
pQ-HBV-L-Null aggctttttggaggcctaggcttttgcaaaaagctactggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaaataccgcacagatgcgtaaggagaaa
pQ-HBV-L-Null ataccgcatcaggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggata
pQ-HBV-L-Null acgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagt
pQ-HBV-L-Null cagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcggg
pQ-HBV-L-Null aagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaact
pQ-HBV-L-Null atcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactac
pQ-HBV-L-Null ggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttg
pQ-HBV-L-Null caagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaa
pQ-HBV-L-Null ggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgt
pQ-HBV-L-Null ctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagattt
pQ-HBV-L-Null atcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtt
PstI
pQ-HBV-L-Null tgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaa
pQ-HBV-L-Null gcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgt
pQ-HBV-L-Null gactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaacacgggataataccgcgccacatagcagaactttaaaagtgctcatcattg
pQ-HBV-L-Null gaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtga
pQ-HBV-L-Null gcaaaaacaggaaggcaaaatgccgcaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcgg
pQ-HBV-L-Null atacatatttgaatgtatttagaaaaataaacaaatagggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaatagggcgta
pQ-HBV-L-Null tcacgaggcccttcgtcttcaagaattagcttggccattgcatacgttgtatccatatcataatatgtacatttatattggctcatgtccaacattaccgccatgttgacattgattattgact
pQ-HBV-L-Null agttattaatagtaatcaattacggggtcattagttcatagcccatatatggagtt

APPENDIX 10: pQ-HBV1.3mer WT sequences

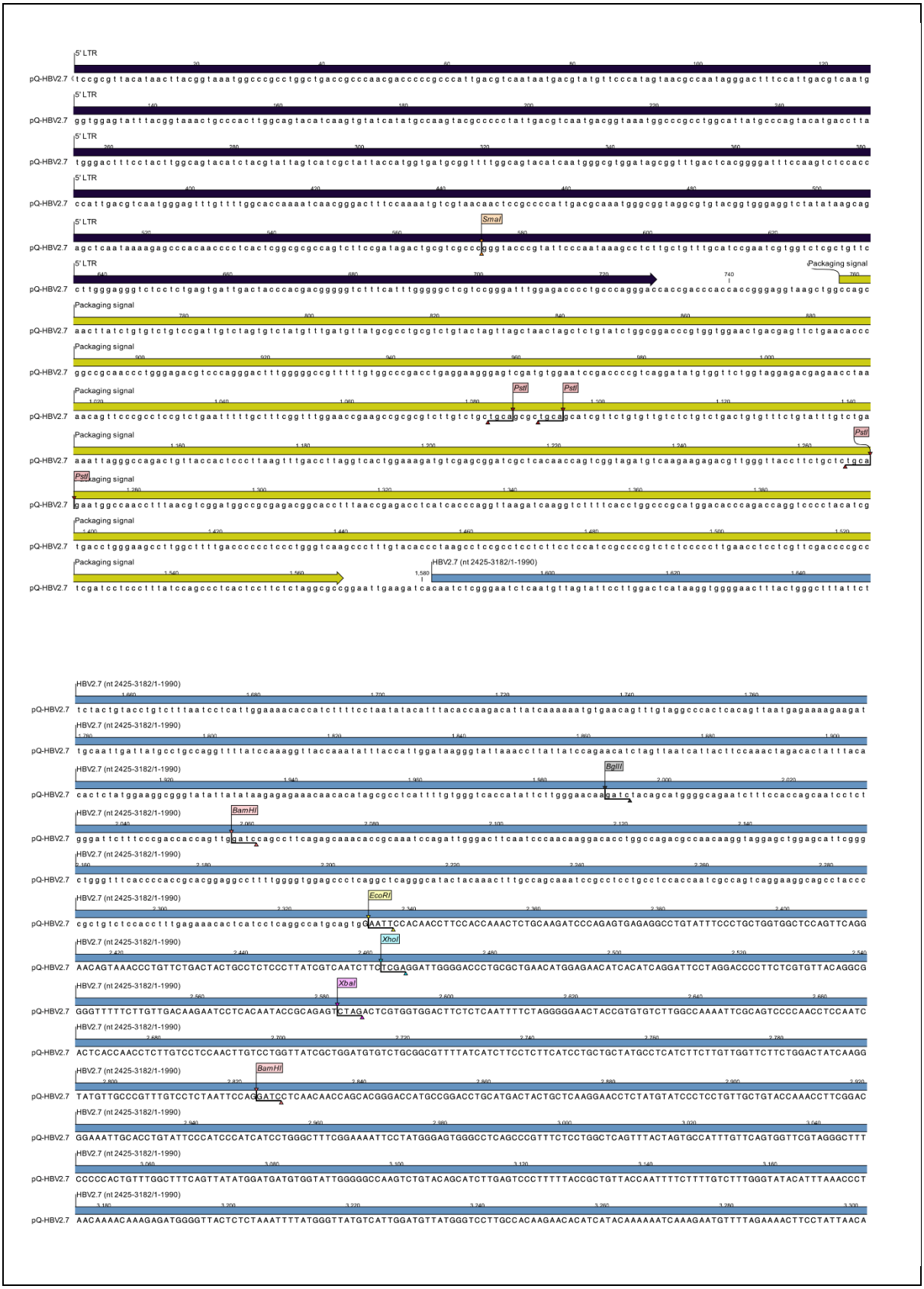


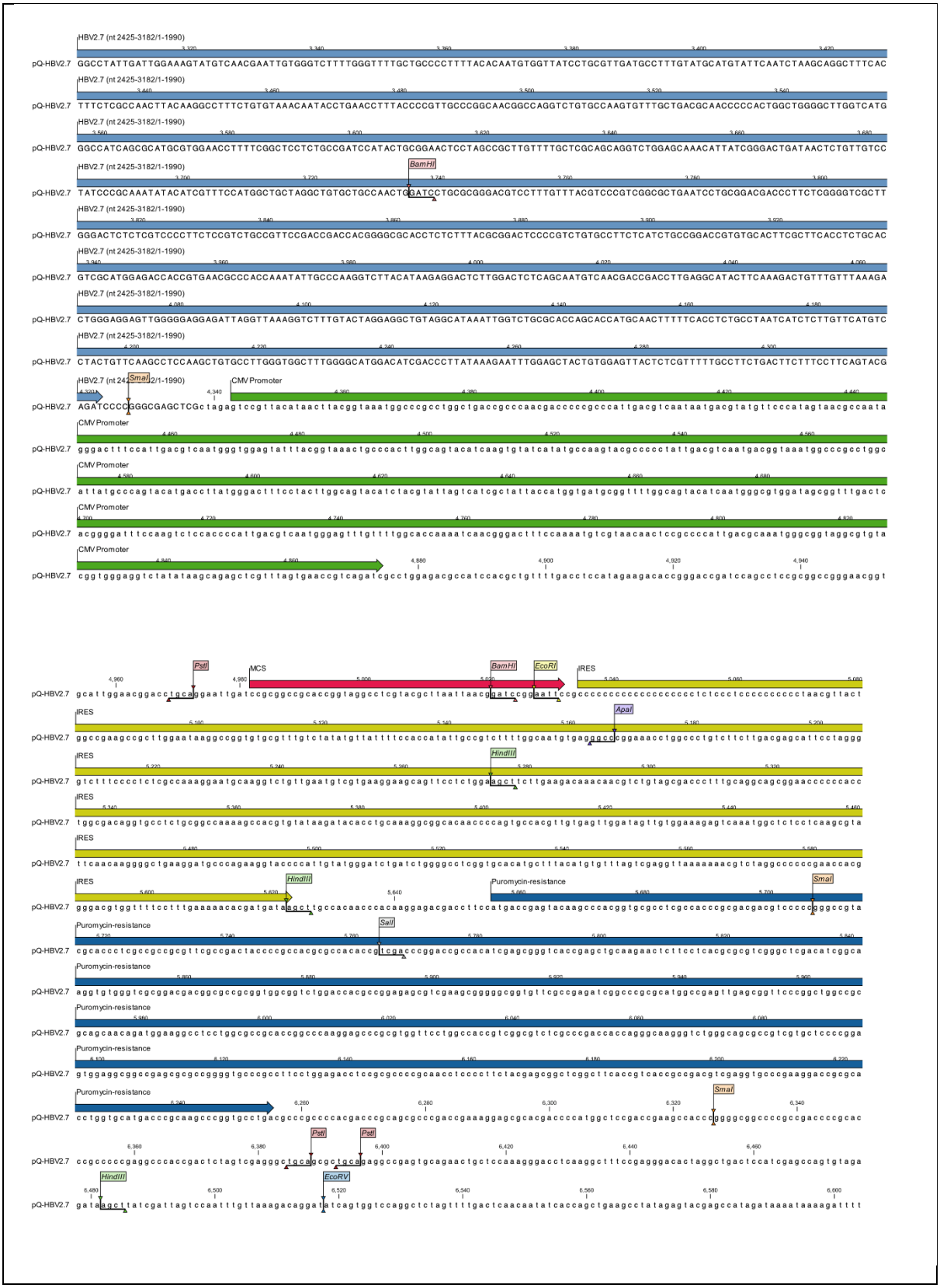

HBV 1.3x (1072-3182/1-1990)
pQ-HBV1.3mer WT
BglII
BamHI
EcoRI
XhoI
XbaI
BamHI
BamHI
AP314
SmaI
CMV Promoter

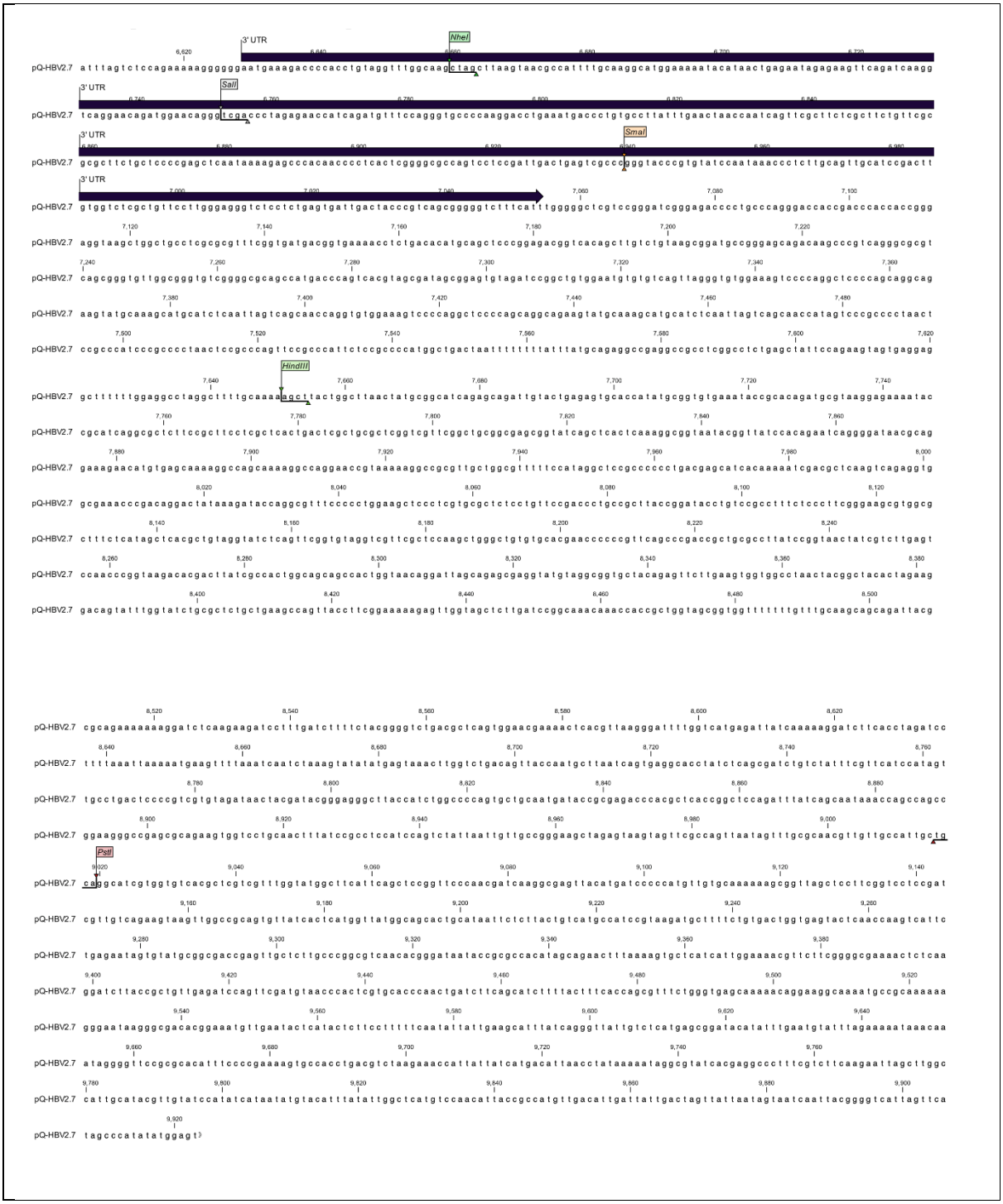

PstI
MCS
BamHI
EcoRI
IRES
HindIII
Puromycin-resistance
SmaI
SalI
EcoRV
3' UTR
pQ-HBV1.3mer WT


PstI
pQ-HBV1.3mer WT
9,720 9,740 9,760 9,780 9,800 9,820 9,840
acagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaa
9,860 9,880 9,900 9,920 9,940 9,960
accaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactca
9,980 10,000 10,020 10,040 10,060 10,080
cgttaagggatttttggtcatgagattatcaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaa
10,100 10,120 10,140 10,160 10,180 10,200
tgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgca
10,220 10,240 10,260 10,280 10,300 10,320
atgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgt
10,340 10,360 10,380 10,400 10,420 10,440
tgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctgcaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaa
10,460 10,480 10,500 10,520 10,540 10,560
cgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactg
10,580 10,600 10,620 10,640 10,660 10,680 10,700
cataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaacacgg
10,720 10,740 10,760 10,780 10,800 10,820
gataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgca
10,840 10,860 10,880 10,900 10,920 10,940
cccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttc
10,960 10,980 11,000 11,020 11,040 11,060
cttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggtttccgcgcacatttccccgaaaagtgccacct
11,080 11,100 11,120 11,140 11,160 11,180
gacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggccctttcgtcttcaagaattagcttggccattgcatacgttgtatccatatcataatatgtac
11,200 11,220 11,240 11,260 11,280 11,300
atttatattggctcatgtccaacattaccgccatgttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatatggagtt

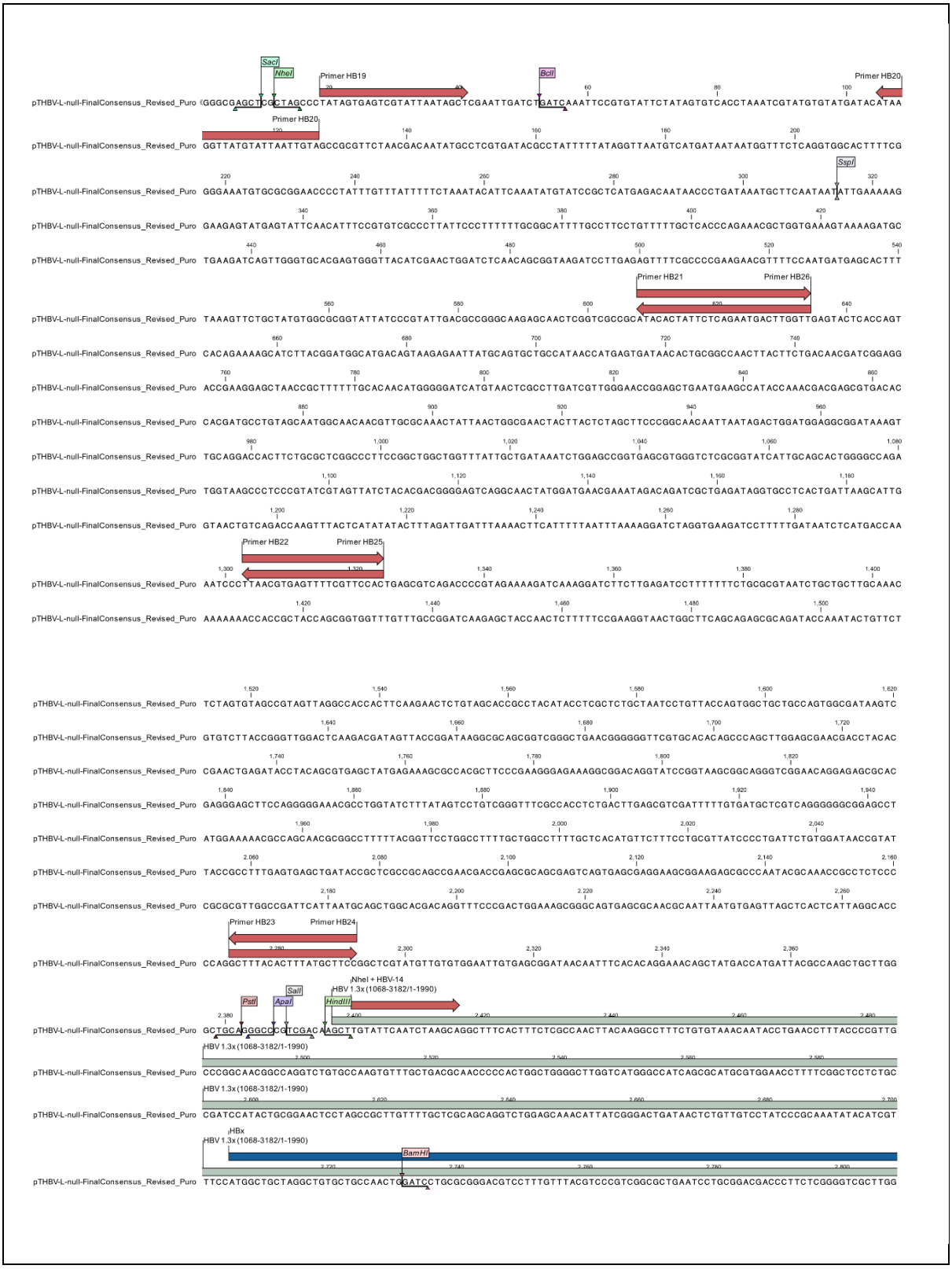
APPENDIX 11: pQ-HBV2.7 sequences

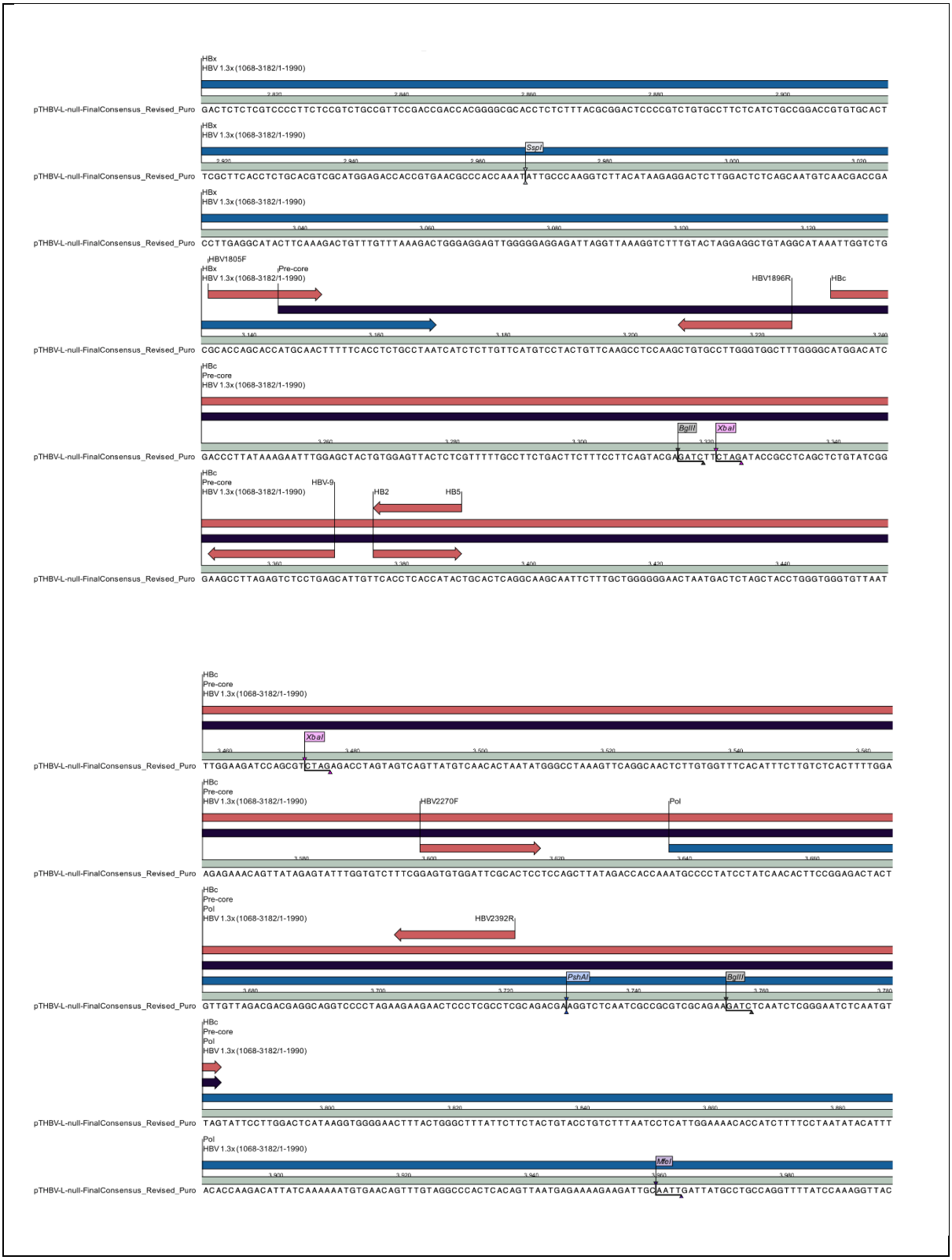


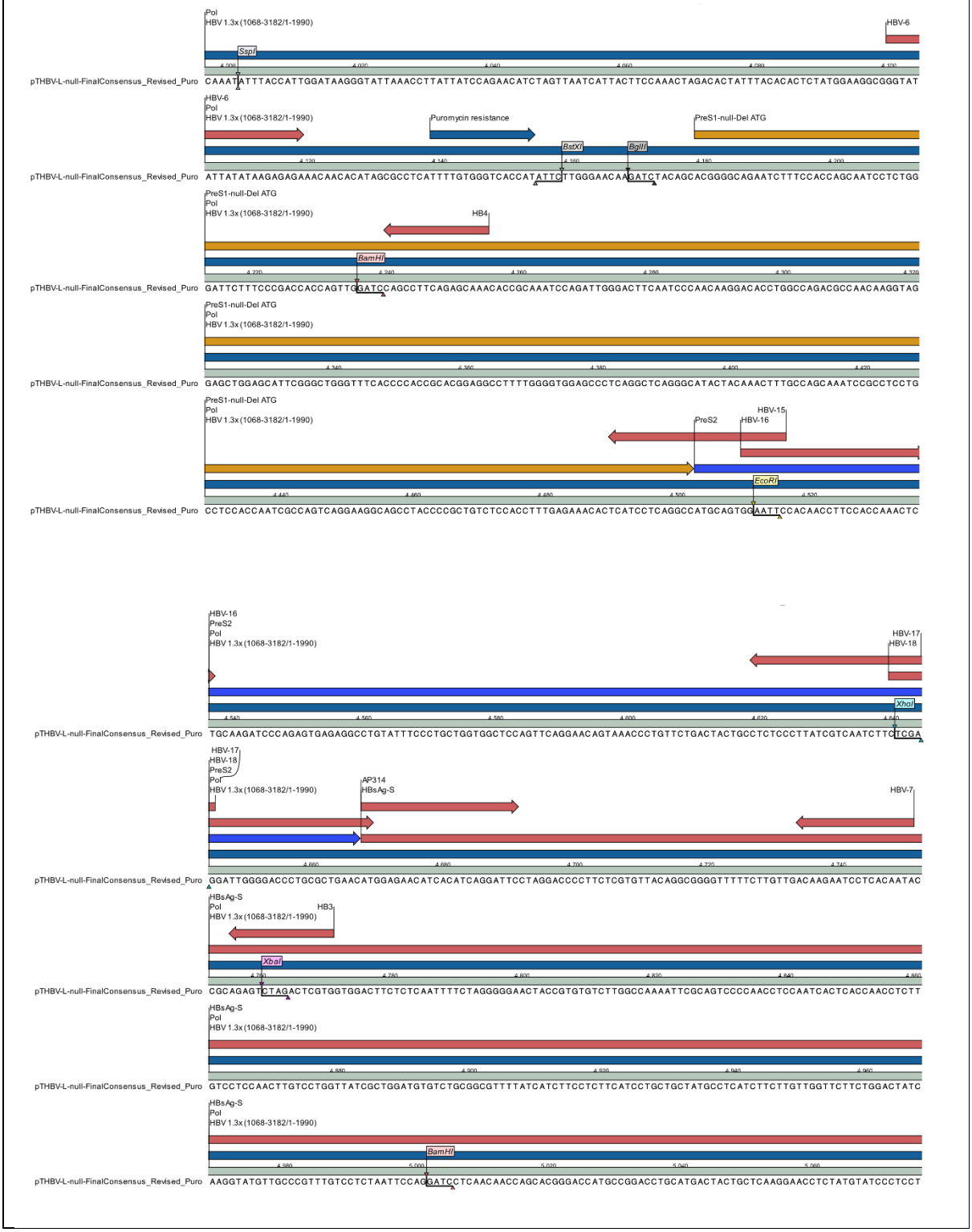

pQ-HBV2.7
BamHI
SmaI
CMV Promoter
PstI
MCS
EcoRI
IRES
ApaI
HindIII
Puromycin-resistance
SalI
EcoRV

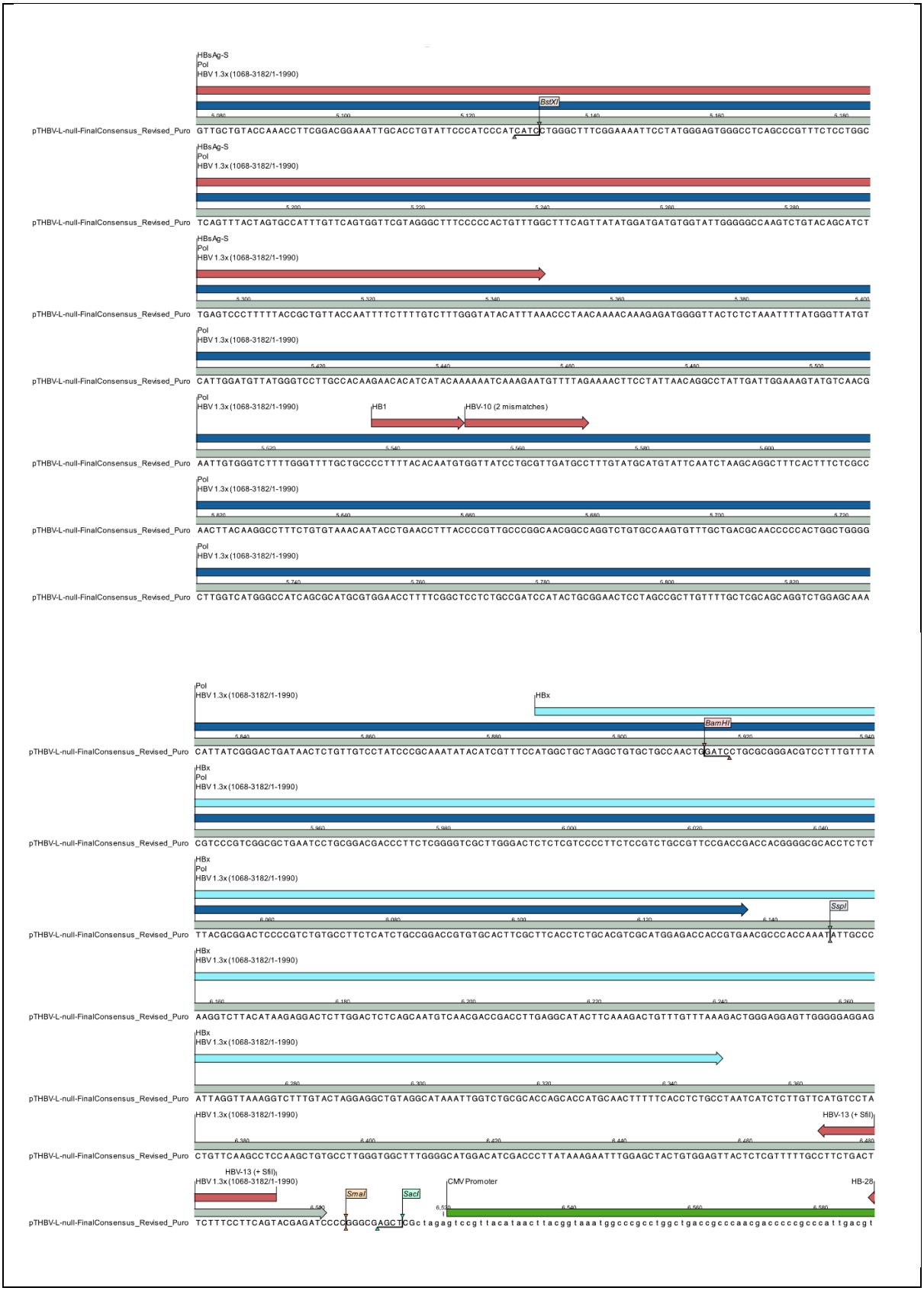

3' UTR
NheI
SalI
SmaI
HindIII
PstI
pQ-HBV2.7

APPENDIX 12: pTHBV-L^{-} /Puro, No 11 sequences




HBx
HBV 1.3x (1068-3182/1-1990)
pTHBV-L-null-FinalConsensus_Revised_Puro
GACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACT
SspI
TCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGA
CCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTG
HBV1805F
Pre-core
HBV1896R
HBc
CGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATC
BglII
XbaI
GACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATCTTCTAGATACCGCCTCAGCTCTGTATCGG
HBV-9
HB2
HB5
GAAGCCTTAGAGTCTCCTGAGCATTGTTCACCTCACCATACTGCACTCAGGCAAGCAATTCTTTGCTGGGGGGAACTAATGACTCTAGCTACCTGGGTGGGTGTTAAT
TTGGAAGATCCAGCGTCTAGAGACCTAGTAGTCAGTTATGTCAACACTAATATGGGCCTAAAGTTCAGGCAACTCTTGTGGTTTCACATTTCTTGTCTCACTTTTGGA
AGAGAAACAGTTATAGAGTATTTGGTGTCTTTCGGAGTGTGGATTCGCACTCCTCCAGCTTATAGACCACCAAATGCCCCTATCCTATCAACACTTCCGGAGACTACT
HBV2270F
Pol
GTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAGACGAAAGGTCTCAATCGCCGCGTCGCAGAAGATCTCAATCTCGGGAATCTCAATGT
HBV2392R
PshAI
BglII
TAGTATTCCTTGGACTCATAAGGTGGGGAACTTTACTGGGCTTTATTCTTCTACTGTACCTGTCTTTAATCCTCATTGGAAAACACCATCTTTTCCTAATATACATTT
MfeI
ACACCAAGACATTATCAAAAAATGTGAACAGTTTGTAGGCCCACTCACAGTTAATGAGAAAAGAAGATTGCAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTAC


Pol
HBV 1.3x (1068-3182/1-1990)
HBV-6
SspI
pTHBV-L-null-FinalConsensus_Revised_Puro
CAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAATCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTAT
HBV-6
Pol
HBV 1.3x (1068-3182/1-1990)
Puromycin resistance
PreS1-null-Del ATG
BstXI
BglII
pTHBV-L-null-FinalConsensus_Revised_Puro
ATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGGTCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGG
PreS1-null-Del ATG
Pol
HBV 1.3x (1068-3182/1-1990)
HB4
BamHI
pTHBV-L-null-FinalConsensus_Revised_Puro
GATTCTTTCCCGACCACCAGTTGGATCCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAG
PreS1-null-Del ATG
Pol
HBV 1.3x (1068-3182/1-1990)
pTHBV-L-null-FinalConsensus_Revised_Puro
GAGCTGGAGCATTCGGGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTG
PreS1-null-Del ATG
Pol
HBV 1.3x (1068-3182/1-1990)
HBV-15
PreS2
HBV-16
EcoRI
pTHBV-L-null-FinalConsensus_Revised_Puro
CCTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTGGAATTCCACAACCTTCCACCAAACTC
HBV-16
PreS2
Pol
HBV 1.3x (1068-3182/1-1990)
HBV-17
HBV-18
XhoI
pTHBV-L-null-FinalConsensus_Revised_Puro
TGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCCTCTCCCTTATCGTCAATCTTCTCGA
HBV-17
HBV-18
PreS2
Pol
HBV 1.3x (1068-3182/1-1990)
AP314
HBsAg-S
HBV-7
pTHBV-L-null-FinalConsensus_Revised_Puro
GGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATAC
HBsAg-S
Pol
HBV 1.3x (1068-3182/1-1990)
HB3
XbaI
pTHBV-L-null-FinalConsensus_Revised_Puro
CGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTT
HBsAg-S
Pol
HBV 1.3x (1068-3182/1-1990)
pTHBV-L-null-FinalConsensus_Revised_Puro
GTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATC
HBsAg-S
Pol
HBV 1.3x (1068-3182/1-1990)
BamHI
pTHBV-L-null-FinalConsensus_Revised_Puro
AAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCT


HBsAg-S
Pol
HBV 1.3x (1068-3182/1-1990)
BstXI
pTHBV-L-null-FinalConsensus_Revised_Puro
GTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGC
TCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCT
TGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTATGT
CATTGGATGTTATGGGTCCTTGCCACAAGAACACATCATACAAAAAATCAAAGAATGTTTTAGAAAACTTCCTATTAACAGGCCTATTGATTGGAAAGTATGTCAACG
HB1
HBV-10 (2 mismatches)
AATTGTGGGTCTTTTGGGTTTTGCTGCCCCTTTTACACAATGTGGTTATCCTGCGTTGATGCCTTTGTATGCATGTATTCAATCTAAGCAGGCTTTCACTTTCTCGCC
AACTTACAAGGCCTTTCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGG
CTTGGTCATGGGCCATCAGCGCATGCGTGGAACCTTTTCGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAA
HBx
BamHI
CATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATCGTTTCCATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTA
CGTCCCGTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCT
SspI
TTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATGGAGACCACCGTGAACGCCCACCAAATATTGCCC
AAGGTCTTACATAAGAGGACTCTTGGACTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAG
ATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTA
HBV-13 (+ SfiI)
CTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACT
SmaI
SacI
CMV Promoter
HB-28
TCTTTCCTTCAGTACGAGATCCCCGGGCGAGCTCGctagagtccgttacataacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccattgacgt


CMV Promoter
HB-28
pTHBV-L-null-FinalConsensus_Revised_Puro
SacI
PstI
MCS
BamHI
EcoRI
EMCV IRES
ApaI
HindIII
Puromycin resistance
SmaI
SalI
PshAI

APPENDIX 11: 3-Null - pTHBV-L^{-} 3Null-Puro-11

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT
TTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
ACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTA
CTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAG
CATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

APPENDIX 12: RT L80I - pTHBV-L^{-} 3Null-Puro-11-L80I

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGaTATCGCTGGATGTGTCTGCGGCGTT
TTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
ACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTA
CTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAG
CATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

APPENDIX 13: RT L91I - pTHBV-L^{-} 3Null-Puro-11-L91I

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT
TTATCATaTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
ACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTA
CTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAG
CATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

APPENDIX 14: RT H126Y - pTHBV-L^{-} 3Null-Puro-11-H126Y

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT
TTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGtACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
ACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTA
CTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAG
CATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

APPENDIX 15: RT D134N - pTHBV-L- 3Null-Puro-11-D134N

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT
TTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATaACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
ACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTA
CTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAG
CATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

APPENDIX 16: RT D134E - pTHBV-L- 3Null-Puro-11-D134E

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT
TTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGAgTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
ACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTA
CTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAG
CATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

APPENDIX 17: RT R153W - pTHBV-L- 3Null-Puro-11-R153W

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT
TTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
AtGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTAC
TAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGC
ATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

APPENDIX 18: RT M204I - pTHBV-L- 3Null-Puro-11-M204I

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT
TTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
ACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTA
CTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATtGATGATGTGGTATTGGGGGCCAAGTCTGTACAGC
ATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

APPENDIX 19: RT M204V - pTHBV-L^{-} 3Null-Puro-11-M204V

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT
TTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
ACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTA
CTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATgTGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGC
ATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

APPENDIX 20: RT Q215S - pTHBV-L^{-} 3Null-Puro-11-Q215S

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT
TTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
ACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTA
CTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTAtcGC
ATCTTGAGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

APPENDIX 21: RT F221Y - pTHBV-L^{-} 3Null-Puro-11-F221Y

CAATTGATTATGCCTGCCAGGTTTTATCCAAAGGTTACCAAATATTTACCATTGGATAAGGGTATTAAACCTTATTATCCAGAACATCTAGTTAA
TCATTACTTCCAAACTAGACACTATTTACACACTCTATGGAAGGCGGGTATATTATATAAGAGAGAAACAACACATAGCGCCTCATTTTGTGGG
TCACCATATTCTTGGGAACAAGATCTACAGCACGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTaGGATCCAG
CCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACCTGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCG
GGCTGGGTTTCACCCCACCGCACGGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAATCCGCCTCCTGC
CTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCTTTGAGAAACACTCATCCTCAGGCCATGCAGTaGAATTCCACAACCT
TCCACCAAACTCTGCAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTGCC
TCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACA
GGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAACTACCGTGT
GTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTT
TTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCAGG
ATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGG
ACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTA
CTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAG
CATCTTGAGTCCCTTTaTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATAC

BIBLIOGRAPHY

ABAALKHAIL, F., ELSIESY, H., ALOMAIR, A., ALGHAMDI, M.Y., ALALWAN, A., ALMASRI, N. & AL-HAMOUDI, W. 2014. SASLT practice guidelines for the management of hepatitis B virus. *Saudi J Gastroenterol*, 20, 5-25.

ABDELHAMED, M., KELLEY, M., MILLER, G., FURMAN, A., & ISOM, C. 2002. Rebound of hepatitis B virus replication in HepG2 cells after cessation of antiviral treatment. *Journal of virology*, *76*, 8148-8160.

ABDELRAHMAN, T., HUGHES, J., MAIN, J., MCLAUCHLAN, J., THURSZ, M. & THOMSON, E. 2015. Next-generation sequencing sheds light on the natural history of hepatitis C infection in patients who fail treatment. *Hepatology*, 61, 88-97.

ABDO, A., & SANAI, F. 2015. Viral hepatitis in Saudi Arabia. An unfinished story. Saudi medical journal, *36*, 785-786.

ABDO, A., SANAI, M. & AL-FALEH, Z. 2012. Epidemiology of viral hepatitis in Saudi Arabia: Are we off the hook? *Saudi J Gastroenterol*, 18, 349-57.

ABDUL BASIT, S., DAWOOD, A., RYAN, J., & GISH, R. 2017. Tenofovir alafenamide for the treatment of chronic hepatitis B virus infection. *Expert review of clinical pharmacology*, *10*, 707-716.

ACHARYA, K., MADAN, K., DATTAGUPTA, S. & PANDA, K. 2006. Viral hepatitis in India. *Natl Med J* India, 19, 203-217.

ADEN, D., FOGEL, A., PLOTKIN, S., DAMJANOV, I., & KNOWLES, B. 1979. Controlled synthesis of HBsAg in a differentiated human liver carcinoma-derived cell line. *Nature*, *282*, 615-616.

AHN, S. H., KRAMVIS, A., KAWAI, S., SPANGENBERG, H. C., LI, J., KIMBI, G., KEW, M., WANDS, J., & TONG, S. 2003. Sequence variation upstream of precore translation initiation codon reduces hepatitis B virus e antigen production. *Gastroenterology*, 125, 1370-1378.

AL ASHGAR, H. I., IMAMBACCUS, H., PEEDIKAYIL, M. C., AL THAWADI, S., AL QUAIZ, M., AL FADDA, M., AL KAHTANI, K., KAGEVI, I., & KHAN, M. Q. 2008. Prevalence of hepatitis B virus genotype in Saudi Arabia: a preliminary report. *Indian journal of gastroenterology: official journal of the Indian Society of Gastroenterology*, *27*, 81-82.

AL BAQLANI, S. A., SY, B. T., RATSCH, B. A., AL NAAMANI, K., AL AWAIDY, S., BUSAIDY, S. A., PAULI, G., & BOCK, C. T. 2014. Molecular epidemiology and genotyping of hepatitis B virus of HBsAg-positive patients in Oman. *PloS one*, *9*, e97759.

ALFARESI, M., ELKOUSH, A., ALSHEHHI, H., ALZAABI, A., & ISLAM, A. 2010. Hepatitis B virus genotypes and precore and core mutants in UAE patients. *Virology journal*, *7*, 160.

ALGHAMDI, A. S., SANAI, F. M., ISMAIL, M., ALGHAMDI, H., ALSWAT, K., ALQUTUB, A., ALTRAIF, I., SHAH, H., ALFALEH, F. Z., & SAUDI ASSOCIATION FOR THE STUDY OF LIVER DISEASES AND TRANSPLANTATION. 2012. SASLT practice guidelines: Management of hepatitis C virus infection. *Saudi J Gastroenterol*, 18 (suppl 1), 1-32.

ALHURAIJI, A., ALARAJ, A., ALGHAMDI, S., ALRBIAAN, A. & ALRAJHI, A.A. 2014. Viral hepatitis B and C in HIV-infected patients in Saudi Arabia. *Ann Saudi Med*, 34, 207.

ALJARALLAH B. M. 2006. Hepatitis B genotyping and its clinical implications. *Saudi journal of gastroenterology: official journal of the Saudi Gastroenterology Association*, *12*, 146-148.

ALJUMAH, A. A., BABATIN, M., HASHIM, A., ABAALKHAIL, F., BASSIL, N., SAFWAT, M., & SANAI, F. M. 2019. Hepatitis B care pathway in Saudi Arabia: Current situation, gaps and actions. *Saudi J Gastroenterol*, 25, 73-80.

ALLAIN, J., ABOL, H., ARSLAN, O., AYOB, Y., BELKHIRI, D., BIANCO, L., BIRD, A., BON, E., BOUKEF, K., BROJER, E., CROOKES, R., DODD, R., ECHEVARRIA, J., CHAAR, M., EL-EKIABY, M., SIDDIG, A., ERIKSTRUP., & FOGLIENI, B., &LAPERCHE, S. 2012. Hepatitis B virus in transfusion medicine: still a problem?. Biologicals, 40, 180-186.

AL-MAHTAB, M., ROY, P. P., KHAN, M., & AKBAR, S. M. 2020. Nobel Prize for the Discovery of Hepatitis B and C: A Brief History in Time. *Euroasian journal of hepato-gastroenterology*, *10*, 98-100.

AL-QAHTANI, A. A., POURKARIM, M. R., TROVÃO, N. S., VERGOTE, V., LI, G., THIJSSEN, M., ABDO, A. A., SANAI, F. M., DELA CRUZ, D., BOHOL, M., AL-ANAZI, M. R., & AL-AHDAL, M. N. 2020. Molecular epidemiology, phylogenetic analysis and genotype distribution of hepatitis B virus in Saudi Arabia: Predominance of genotype D1. *Infection, Genetics and Evolution*, 77, 1567-1348.

AL-QAHTANI, A., POURKARIM, M., TROVÃO, N., VERGOTE, V., LI, G., THIJSSEN, M., ABDO, A., SANAI, F., DELA CRUZ, D., BOHOL, M., AL-ANAZI, M., & AL-AHDAL, M. 2020. Molecular epidemiology, phylogenetic analysis and genotype distribution of hepatitis B virus in Saudi Arabia: Predominance of genotype D1. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases*, *77*, 104051.

AL-QUDARI, A. Y., AMER, H. M., ABDO, A. A., HUSSAIN, Z., AL-HAMOUDI, W., ALSWAT, K., & ALMAJHDI, F. N. 2016. Surface gene variants of hepatitis B Virus in Saudi Patients. *Saudi journal of gastroenterology: official journal of the Saudi Gastroenterology Association*, *22*(2), 133-138.

ALSHABI, A., FATIMA, N., MARWAN, A., ORAIBI, K. G., QUBAISI, E. A., ARIF, H. O., DAGHRIRI, E. M., ZELAI, N. A., & ALI KHAN, I. 2021. Epidemiology screening and genotyping analysis for Hepatitis B virus in Southwestern region of Saudi Arabia. *Journal of infection and public health*, *14*, 187-192.

ALSWAIDI, F.M. & O'BRIEN, S. 2010. Is there a need to include HIV, HBV and HCV viruses in the Saudi premarital screening program on the basis of their prevalence and transmission risk factors? *J Epidemiol Community Health*, 64, 989-997.

ALTINDIS, M., ASLAN, F.G., KOROĞLU, M., EREN, A., DEMIR, L., USLAN, M.İ., ASLAN, S., OZDEMI, R M., & BAYKAN, M. 2016. Hepatitis B virus carrying drug-resistance compensatory mutations in chronically infected treatment-naive patients. *Viral Hepat J*, 22, 103-107.

ALVARADO MORA, M. V., ROMANO, C. M., GOMES-GOUVÊA, M. S., GUTIERREZ, M. F., BOTELHO, L., CARRILHO, F. J., & PINHO, J. R. 2011. Molecular characterization of the Hepatitis B virus genotypes in Colombia: a Bayesian inference on the genotype F. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases*, *11*, 103-108.

ALZAHRANI, A.J., CRUZ, D.M.D., OBEID, O.E., BUKHARI, H.A., AL-QAHTANI, A.A. & AL-AHDAL, M.N. 2009. Molecular detection of hepatitis B, hepatitis C, and torque teno viruses in drug users in Saudi Arabia. *J. Med. Virol*, 81, 1343-1347.

AMARAPURKAR D. N. 2007. Telbivudine: a new treatment for chronic hepatitis B. *World journal of gastroenterology*, *13*, 6150-6155.

ASAAD, A.M., AL-AYED, M.S., ALERAKY, M., & QURESHI, M.A. 2015. Hepatitis B virus genotyping in chronic hepatitis B patients in southwestern Saudi Arabia. Brazilian J. Infect. Dis, 19, 525-528.

ASLAM, A., CAMPOVERDE REYES, K. J., MALLADI, V. R., ISHTIAQ, R., & LAU, D. 2018. Management of chronic hepatitis B during pregnancy. *Gastroenterology report*, *6*, 257-262.

ASTBURY, S., COSTA NUNES SOARES, M. M., PEPRAH, E., KING, B., JARDIM, A., SHIMIZU, J. F., JALAL, P., SAEED, C. H., SABEER, F. T., IRVING, W. L., TARR, A. W., & MCCLURE, C. P. 2020. Nanopore sequencing from extraction-free direct PCR of dried serum spots for portable hepatitis B virus drug-resistance typing. *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology*, *129*, 104483.

ASTROVSKAYA, I., TORK, B., MANGUL, S., WESTBROOKS, K., MANDOIU, I., BALFE, P. & ZELIKOVSKY, A. 2011. Inferring viral quasispecies spectra from 454 pyrosequencing reads. *BMC Bioinformatics*, 12 Suppl 6, p.S1.

AVELLON, A. & ECHEVARRIA, J.M. 2006. Frequency of hepatitis B virus 'a' determinant variant in unselected Spanish chronic carriers. *J Med Virol*, 78, 24-36.

BAHRI, O., CHEIKH, I., HAJJI, N., DJEBBI, A., MAAMOURI, N., SADRAOUI, A., MAMI, N. B., & TRIKI, H. 2006. Hepatitis B genotypes, precore and core promoter mutants circulating in Tunisia. *Journal of medical virology*, *78*, 353-357.

BAPTISTA, M., KRAMVIS, A. & KEW, M.C. 1999. High prevalence of 1762(T) 1764(A) mutations in the basic core promoter of hepatitis B virus isolated from black Africans with hepatocellular carcinoma compared with asymptomatic carriers. *Hepatology,* 29, 946-953.

BARRERA, A., GUERRA, B., NOTVALL, L., & LANFORD, R. E. 2005. Mapping of the hepatitis B virus pre-S1 domain involved in receptor recognition. *Journal of virology*, *79,* 9786-9798.

BEASLEY, R. P., HWANG, L. Y., LEE, G. C., LAN, C. C., ROAN, C. H., HUANG, F. Y., & CHEN, C. L. 1983. Prevention of perinatally transmitted hepatitis B virus infections with hepatitis B immune globulin and hepatitis B vaccine. *Lancet (London, England)*, 2, 1099-1102.

BECK, J. & NASSAL, M. 2007. Hepatitis B virus replication. *World J Gastroenterol*, 13, 48-64.

BELLONI, L., POLLICINO, T., DE NICOLA, F., GUERRIERI, F., RAFFA, G., FANCIULLI, M., RAIMONDO, G., & LEVRERO, M. 2009. Nuclear HBx binds the HBV minichromosome and modifies the epigenetic regulation of cccDNA function. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 19975-19979.

BENTLEY, D. R., BALASUBRAMANIAN, S., SWERDLOW, H. P., SMITH, G. P., MILTON, J., BROWN, C. G., HALL, K. P., EVERS, D. J., BARNES, C. L., BIGNELL, H. R., BOUTELL, J. M., BRYANT, J., CARTER, R. J., CHEETHAM, R. K., COX, A. J., ELLIS, D. J., FLATBUSH, M. R., GORMLEY, N. A., HUMPHRAY, S. J., IRVING, L. J., KARBELASHVILI, M. S., KIRK, S. M., LI, H., LIU, X. H., MAISINGER, K. S., MURRAY, L. J., OBRADOVIC, B., OST, T., PARKINSON, M. L., PRATT, M. R., RASOLONJATOVO, I. M. J., REED, M. T., RIGATTI, R., RODIGHIERO, C., ROSS, M. T., SABOT, A., SANKAR, S. V., SCALLY, A., SCHROTH, G. P., SMITH, M. E., SMITH, V. P., SPIRIDOU, A., TORRANCE, P. E., TZONEV, S. S., VERMAAS, E. H., WALTER, K., WU, X. L., ZHANG, L., ALAM, M. D., ANASTASI, C., ANIEBO, I. C., BAILEY, D. M. D., BANCARZ, I. R., BANERJEE, S., BARBOUR, S. G., BAYBAYAN, P. A., BENOIT, V. A., BENSON, K. F., BEVIS, C., BLACK, P. J., BOODHUN, A., BRENNAN, J. S., BRIDGHAM, J. A., BROWN, R. C., BROWN, A. A., BUERMANN, D. H., BUNDU, A. A., BURROWS, J. C., CARTER, N. P., CASTILLO, N., CATENAZZI, M. C. E., CHANG, S., COOLEY, R. N., CRAKE, N. R., DADA, O. O., DIAKOUMAKOS, K. D., DOMINGUEZ-FERNANDEZ, B., EARNSHAW, D. J., EGBUJOR, U. C., ELMORE, D. W., ETCHIN, S. S., EWAN, M. R., FEDURCO, M., FRASER, L. J., FAJARDO, K. V. F., FUREY, W. S., GEORGE, D., GIETZEN, K. J., GODDARD, C. P., GOLDA, G. S., GRANIERI, P. A., GREEN, D. E., GUSTAFSON, D. L., HANSEN, N. F., HARNISH, K., HAUDENSCHILD, C. D., HEYER, N. I., HIMS, M. M., HO, J. T., HORGAN, A. M., et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature, 456, 53-59.

BERRY, I., MELENDREZ, M. C., BISHOP-LILLY, K. A., RUTVISUTTINUNT, W., POLLETT, S., TALUNDZIC, E., MORTON, L., & JARMAN, R. G. 2020. Next Generation Sequencing and Bioinformatics Methodologies for Infectious Disease Research and Public Health: Approaches, Applications, and Considerations for Development of Laboratory Capacity. *The Journal of infectious diseases*, *221*(Suppl 3), S292-S307.

BLOCK, T. M., GUO, H., & GUO, J. T. 2007. Molecular virology of hepatitis B virus for clinicians. *Clinics in liver disease*, *11*, 685-706.

BLOCK, T.M., BLOCK, H.J., ALTER, LONDON, W.T, & BRAY, M. 2016. A historical perspective on the discovery and elucidation of the hepatitis B virus. *Antiviral Research*, 131, 109-123.

BLOCK, T.M., GUO, H. & GUO, J.T. 2007. Molecular Virology of hepatitis B virus for clinicians. *Clin Liver Dis*, 11, 685-706.

BLUMBERG B. S. 2002. The discovery of the hepatitis B virus and the invention of the vaccine: a scientific memoir. *Journal of gastroenterology and hepatology*, *17 Suppl*, S502-S503.

BLUMBERG, B.S. 1977. Australia antigen and the biology of hepatitis B. *Science,* 197, 17-25.

BOCK, C. T., TILLMANN, H. L., TORRESI, J., KLEMPNAUER, J., LOCARNINI, S., MANNS, M. P., & TRAUTWEIN, C. 2002. Selection of hepatitis B virus polymerase mutants with enhanced replication by lamivudine treatment after liver transplantation. *Gastroenterology*, *122*, 264-273.

BOETTLER, T., MARJOT, T., NEWSOME, P. N., MONDELLI, M. U., MATICIC, M., CORDERO, E., JALAN, R., MOREAU, R., CORNBERG, M., & BERG, T. 2020. Impact of COVID-19 on the care of patients with liver disease: EASL-ESCMID position paper after 6 months of the pandemic. *JHEP reports: innovation in hepatology*, *2*, 100169.

BOWDEN, D.S. & THOMPSON, A.J. 2008. New developments in HBV molecular diagnostics and quantitative serology. *Hepatol Int*, 2 Supplement 1, 3-11.

BOWDEN, S. 2006. Serological and molecular diagnosis. *Semin Liver Dis*, 26, 97-103.

BOWYER, S. & SIM, J. 2000. Relationships within and between genotypes of hepatitis B virus at points across the genome: Footprints of recombination in certain isolates. *Journal of General Virology*, 81, 379-392.

BOYD, A., LACOMBE, K., LAVOCAT, F., MAYLIN, S., MIAILHES, P., LASCOUX-COMBE, C., DELAUGERRE, C., GIRARD, P.M., & ZOULIM, F., 2016. Decay of ccc-DNA marks persistence of intrahepatic viral DNA synthesis under tenofovir in HIV-HBV co-infected patients. J. Hepatol. 65, 683-691.

BRAGG, LM., STONE, G., BUTLER, MK., HUGENHOLTZ, P., & TYSON, GW. 2013. Shining a Light on Dark Sequencing: Characterising Errors in Ion Torrent PGM Data. PLOS Computational Biology, 9, e1003031.

BRUNETTO, M.R., STEMLER, M., SCHODEL, F., WILL, H., OTTOBRELLIA., RIZZETTO, M., VERME, G & BONINO, F. 1989. Identification of HBV variants which cannot produce precore derived HBeAg and may be responsible for severe hepatitis. Italian Journal of *Gastroenterol Hepatol*, 21, 151-154.

BRUSS, V. & GANEM, D. 1991. The role of envelope proteins in hepatitis B virus assembly. *Proc Natl Acad Sci USA*, 88, 1059-1063.

BRUSS, V., GERHARDT, E., VIELUF, K. AND WUNDERLICH, G. 1996. Functions of the large hepatitis B virus surface protein in viral particle morphogenesis. *Intervirology*, 39, 23-31.

BUCKWOLD, V.E., XU, Z., CHEN, M., YEN, T.S. & OU, J.H. 1996. Effects of a naturally occurring mutation in the hepatitis B virus basal core promoter on precore gene expression and viral replication. *J Virol*, 70, 5845-5851.

BUI, T., TRAN, T. T., NGHIEM, M. N., RAHMAN, P., TRAN, T., DINH, M., LE, M. H., NGUYEN, V., THWAITES, G., & RAHMAN, M. 2017. Molecular characterization of hepatitis B virus in Vietnam. *BMC infectious diseases*, *17*, 601.

BURNS, G. S., & THOMPSON, A. J. 2014. Viral hepatitis B: clinical and epidemiological characteristics. *Cold Spring Harbor perspectives in medicine*, *4*(12), a024935.

BYRNE, R., CAREY, I., & AGARWAL, K. 2018. Tenofovir alafenamide in the treatment of chronic hepatitis B virus infection: rationale and clinical trial evidence. *Therapeutic advances in gastroenterology*, *11*, 1756284818786108.

CAI, D., MILLS, C., YU, W., YAN, R., ALDRICH, C. E., SAPUTELLI, J. R., MASON, W. S., XU, X., GUO, J. T., BLOCK, T. M., CUCONATI, A., & GUO, H. 2012. Identification of disubstituted sulfonamide compounds as specific inhibitors of hepatitis B virus covalently closed circular DNA formation. *Antimicrobial agents and chemotherapy*, *56*, 4277-4288.

CAI, D., WANG, X., YAN, R., MAO, R., LIU, Y., JI, C., CUCONATI, A., & GUO, H. 2016. Establishment of an inducible HBV stable cell line that expresses cccDNA-dependent epitope-tagged HBeAg for screening of cccDNA modulators. *Antiviral research*, *132*, 26-37.

CALIGIURI, P., CERRUTI, R., ICARDI, G., & BRUZZONE, B. 2016.Overview of hepatitis B virus mutations and their implications in the management of infection. *World J Gastroenterol*, 22, 145-154.

CARMAN, W. F., JACYNA, M. R., HADZIYANNIS, S., KARAYIANNIS, P., MCGARVEY, M. J., MAKRIS, A., & THOMAS, H. C. 1989. Mutation preventing formation of hepatitis B e antigen in patients with chronic hepatitis B infection. *Lancet (London, England)*, *2*, 588-591.

CHAN, H.L., HUSSAIN, M., & LOK, A.S. 1999. Different hepatitis B virus genotypes are associated with different mutations in the core promoter and precore regions during hepatitis B e antigen seroconversion. *Hepatology*, 29, 976-984.

CHANDRA, P.K., BISWAS, A., DATTA, S., BANERJEE, A., PANIGRAHI, R., CHAKRABARTI, S., DE, B.K. AND CHAKRAVARTY, R. .2009. Subgenotypes of hepatitis B virus genotype D (D1, D2, D3 and D5) in India: differential pattern of mutations, liver injury and occult HBV infection. *J Viral Hepatm,* 16, 749-756.

CHANG, M. H., & CHEN, D. S. 2015. Prevention of hepatitis B. *Cold Spring Harbor perspectives in medicine*, 5, p.a021493.

CHANG, T. T., LAI, C. L., KEW YOON, S., LEE, S. S., COELHO, H. S., CARRILHO, F. J., POORDAD, F., HALOTA, W., HORSMANS, Y., TSAI, N., ZHANG, H., TENNEY, D. J., TAMEZ, R., & ILOEJE, U. 2010. Entecavir treatment for up to 5 years in patients with hepatitis B e antigen-positive chronic hepatitis B. *Hepatology (Baltimore, Md.)*, *51*, 422-430.

CHEN, C. J., YANG, H. I., SU, J., JEN, C. L., YOU, S. L., LU, S. N., HUANG, G. T., ILOEJE, U. H., & REVEAL-HBV STUDY GROUP. 2006. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. *JAMA*, *295*, 65-73.

CHEN, G., LIN, W., SHEN, F., ILOEJE, U.H., LONDON, W.T. & EVANS, A.A. 2006. Past HBV viral load as predictor of mortality and morbidity from HCC and chronic liver disease in a prospective study. *Am J Gastroenterol*, 101(8), 1797-1803.

CHEN, P., GAN, Y., HAN, N., FANG, W., LI, J., ZHAO, F., HU, K., & RAYNER, S. 2013. Computational evolutionary analysis of the overlapped surface (S) and polymerase (P) region in hepatitis B virus indicates the spacer domain in P is crucial for survival. *PloS one*, *8*, e60098.

CHENG, H., ZHANG, H. Z., SHEN, W. A., LIU, Y. F., & MA, F. C. 2003. Expression of RNase H of human hepatitis B virus polymerase in Escherichia coli. *World journal of gastroenterology*, *9*, 513-515.

CHIEN, Y. C., JAN, C. F., KUO, H. S., & CHEN, C. J. 2006. Nationwide hepatitis B vaccination program in Taiwan: effectiveness in the 20 years after it was launched. *Epidemiologic reviews*, 28, 126-135.

CHO, W.H., LEE, H.J., BANG, K.B., KIM, S.B. & SONG, I.H. 2018. Development of tenofovir disoproxil fumarate resistance after complete viral suppression in a patient with treatment-naïve chronic hepatitis B: A case report and review of the literature. *World Journal of Gastroenterology*, 24, 1919-1924.

CHOI, M. S., KIM, D. Y., LEE, D. H., LEE, J. H., KOH, K. C., PAIK, S. W., RHEE, J. C., & YOO, B. C. 2007. Clinical significance of pre-S mutations in patients with genotype C hepatitis B virus infection. *J Viral Hepat*, 14, 161-168.

CHOI, Y. M., LEE, S. Y., & KIM, B. J. 2018. Naturally occurring hepatitis B virus reverse transcriptase mutations related to potential antiviral drug resistance and liver disease progression. *World journal of gastroenterology*, *24*, 1708-1724.

CHOOK, J. B., TEO, W. L., NGEOW, Y. F., TEE, K. K., NG, K. P., & MOHAMED, R. 2015. Universal Primers for Detection and Sequencing of Hepatitis B Virus Genomes across Genotypes A to G. *J Clin Microbiol*, 53, 1831-1835.

CHOU, S.F., TSAI, M.L., HUANG, J.Y., CHANG, Y.S., & SHIH, C. 2015. The dual role of an ESCRT-0 component HGS in HBV transcription and naked capsid secretion. PLoS Pathog, 11, e1005123.

CHURIN, Y., RODERFELD, M. & ROEB, E. 2015. Hepatitis B virus large surface protein: function and fame. *Hepatobiliary Surg Nutr,* 4, 1-10.

CIFTCI, S., KESKIN, F., CAKIRIS, A., AKYUZ, F., PINARBASI, B., ABACI, N., DINCER, E., BADUR, S., KAYMAKOGLU, S., & USTEK, D. 2014. Analysis of potential antiviral resistance mutation profiles within the HBV reverse transcriptase in untreated chronic hepatitis B patients using an ultra-deep pyrosequencing method. *Diagnostic microbiology and infectious disease*, *79*, 25-30.

CLARK, D.N. & HU, J. 2015. Unveiling the roles of HBV polymerase for new antiviral strategies. *Future Virology*, 10, 283-295.

COLSON, P., BORENTAIN, P., RAVAUX, I., AHERFI, S. 2020. Hepatitis B Virus Genomics Knocking at the Door of Routine Diagnostic Laboratories, *The Journal of Infectious Diseases*, 221, 1026-1029.

COOREMAN, M.P., LEROUX-ROELS, G., & PAULIJ, W.P. 2001Vaccine- and hepatitis B immune globulin-induced escape mutations of hepatitis B virus surface antigen. *Journal of biomedical science*, *8*, 237-247.

CORNBERG, M., WONG, V. W., LOCARNINI, S., BRUNETTO, M., JANSSEN, H., & CHAN, H. L. 2017. The role of quantitative hepatitis B surface antigen revisited. *Journal of hepatology*, *66*, 398-411.

COURSAGET, P., GHARBI, Y., KHROUF, N., DEPRIL, N., BOUKHRIS, N., FRITZELL, B., & KASTALLY, R. 1994. Familial clustering of hepatitis B virus infections and prevention of perinatal transmission by immunization with a reduced number of doses in an area of intermediate endemicity (Tunisia). *Vaccine*, *12*, 275-278.

COURSAGET, P., YVONNET, B., CHOTARD, J., VINCELOT, P., SARR, M., DIOUF, C., CHIRON, J. P., & DIOP-MAR, I. 1987. Age- and sex-related study of hepatitis B virus chronic carrier state in infants from an endemic area (Senegal). *Journal of medical virology*, *22*, 1-5.

CROAGH, C. M., & LUBEL, J. S. 2014. Natural history of chronic hepatitis B: phases in a complex relationship. *World journal of gastroenterology*, *20*, 10395-10404.

CULLIGAN, E. P., SLEATOR, R. D., MARCHESI, J. R., & HILL, C. 2014. Metagenomics and novel gene discovery: promise and potential for novel therapeutics. *Virulence*, *5*, 399-412.

DANE, D.S., CAMERON, C.H. & BRIGGS, M. 1970. Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet*, 1, 695-8.

DATTA, S., CHATTERJEE, S., VEER, V., & CHAKRAVARTY, R. 2012. Molecular biology of the hepatitis B virus for clinicians. *Journal of clinical and experimental hepatology*, *2*, 353-365.

DAVIS, C., HAYWOOD, B., VATTIPALLY, S., FILIPE, A., ALSAEED, MARIAM., SMOLLET, K., BAYLIS, S., IJAZ, S., & TEDDER, R., THOMSON, E., & ABDELRAHMAN, T. 2021. Hepatitis E virus: Whole genome sequencing as a new tool for understanding HEV epidemiology and phenotypes. *Journal of Clinical Virology*. 139. 104738.

DAWOOD, A., ABDUL BASIT, S., JAYARAJ, M., & GISH, R. G. 2017. Drugs in Development for Hepatitis B. *Drugs*, *77*, 1263-1280.

DE CLERCQ E. 2016. Tenofovir alafenamide (TAF) as the successor of tenofovir disoproxil fumarate (TDF). *Biochemical pharmacology*, *119*, 1-7.

DE CLERCQ, E., FÉRIR, G., KAPTEIN, S., & NEYTS, J. 2010. Antiviral treatment of chronic hepatitis B virus (HBV) infections. *Viruses*, *2*, 1279-1305.

DEGUCHI, M., YAMASHITA, N., KAGITA, M., ASARI, S., IWATANI, Y., TSUCHIDA, T., IINUMA, K., & MUSHAHWAR, I. K. 2004. Quantitation of hepatitis B surface antigen by an

automated chemiluminescent microparticle immunoassay. *Journal of virological methods*, *115*, 217-222.

DELANEY, W. E., 4TH, & ISOM, H. C. 1998. Hepatitis B virus replication in human HepG2 cells mediated by hepatitis B virus recombinant baculovirus. *Hepatology (Baltimore, Md.)*, *28*, 1134-1146.

DELANEY, W. E., 4TH, MILLER, T. G., & ISOM, H. C. 1999. Use of the hepatitis B virus recombinant baculovirus-HepG2 system to study the effects of (-)-beta-2',3'-dideoxy-3'-thiacytidine on replication of hepatitis B virus and accumulation of covalently closed circular DNA. *Antimicrobial agents and chemotherapy*, *43*, 2017-2026.

DEVESA, M., LOUREIRO, C. L., RIVAS, Y., MONSALVE, F., CARDONA, N., DUARTE, M. C., POBLETE, F., GUTIERREZ, M. F., BOTTO, C., & PUJOL, F. H. 2008. Subgenotype diversity of hepatitis B virus American genotype F in Amerindians from Venezuela and the general population of Colombia. *Journal of medical virology*, *80*, 20-26.

DEVI, U., & LOCARNINI, S. 2013. Hepatitis B antivirals and resistance. *Current opinion in virology*, *3*, 495-500.

DI BISCEGLIE, M. 2009. Hepatitis B and hepatocellular carcinoma. *Hepatology*, 49, S56-S60.

DIMOU, E., PAPADIMITROPOULOS, V., & HADZIYANNIS, S. J. 2007. The role of entecavir in the treatment of chronic hepatitis B. *Therapeutics and clinical risk management*, *3*, 1077-1086.

DIOGO DIAS, JOÃO, NAZIM SARICA, & CHRISTINE NEUVEUT. 2021. Early Steps of Hepatitis B Life Cycle: From Capsid Nuclear Import to cccDNA Formation. *Viruses* 13, no. 5: 757.

DOITSH, G., & SHAUL, Y. 2003. A long HBV transcript encoding pX is inefficiently exported from the nucleus. *Virology*, *309*, 339-349.

DOWNS, L. O., SMITH, D. A., LUMLEY, S. F., PATEL, M., MCNAUGHTON, A. L., MOKAYA, J., ANSARI, M. A., SALIH, H., VÁRNAI, K. A., FREEMAN, O., CRIPPS, S., PHILLIPS, J., COLLIER, J., WOODS, K., CHANNON, K., DAVIES, J., BARNES, E., JEFFERY, K., & MATTHEWS, P. C. 2019. Electronic Health Informatics Data To Describe Clearance Dynamics of Hepatitis B Surface Antigen (HBsAg) and e Antigen (HBeAg) in Chronic Hepatitis B Virus Infection. *mBio*, *10*, e00699-19.

DRESSMAN, D., YAN, H., TRAVERSO, G., KINZLER, K.W., & VOGELSTEIN, B. 2003. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 8817-8822.

DREXLER, J. F., GEIPEL, A., KÖNIG, A., CORMAN, V. M., VAN RIEL, D., LEIJTEN, L. M., BREMER, C. M., RASCHE, A., COTTONTAIL, V. M., MAGANGA, G. D., SCHLEGEL, M., MÜLLER, M. A., ADAM, A., KLOSE, S. M., CARNEIRO, A. J., STÖCKER, A., FRANKE, C. R., GLOZA-RAUSCH, F., GEYER, J., ANNAN, A., & DROSTEN, C. 2013. Bats carry pathogenic hepadnaviruses antigenically related to hepatitis B virus and capable of infecting human hepatocytes. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 16151-16156.

EDMUNDS, W. J., MEDLEY, G. F., NOKES, D. J., HALL, A. J., & WHITTLE, H. C. 1993. The influence of age on the development of the hepatitis B carrier state. *Proceedings. Biological sciences*, *253*, 197-201.

EL-KAFRAWY, S., JAMJOOM, G., AKBAR, H., FALLATAH, H., EL-DALY, M., QARI, Y., ALGHAMDI, A., BABATIN, M., ALSAEDI, M., OTHMAN, N., AL-SUBHI, T., ABDEL-HAMID, M., & AZHAR, E. 2018. Analysis of Hepatitis B virus (HBV) mutations in patients from Western Saudi Arabia with chronic disease. *Journal of infection in developing countries*, *12*, 557-567.

ELSE, L. J., JACKSON, A., PULS, R., HILL, A., FAHEY, P., LIN, E., AMARA, A., SICCARDI, M., WATSON, V., TJIA, J., EMERY, S., KHOO, S., BACK, D. J., & BOFFITO, M. 2012. Pharmacokinetics of lamivudine and lamivudine-triphosphate after administration of 300 milligrams and 150 milligrams once daily to healthy volunteers: results of the ENCORE 2 study. *Antimicrobial agents and chemotherapy*, *56*, 1427-1433.

EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER. 2017. EASL clinical practice guidelines: Management of chronic hepatitis B virus infection. *J Hepatol*, 67, 370-398.

FAN, Y. F., LU, C. C., CHEN, W. C., YAO, W. J., WANG, H. C., CHANG, T. T., LEI, H. Y., SHIAU, A. L., & SU, I. J. 2001. Prevalence and significance of hepatitis B virus (HBV) pre-S mutants in serum and liver at different replicative stages of chronic HBV infection. *Hepatology (Baltimore, Md.)*, *33*, 277-286.

FANG, Z.L., YANG, J., GE, X., ZHUANG, H., GONG, J., LI, R., LING, R. & HARRISON, T.J. 2002. Core promoter mutations (A(1762)T and G(1764)A) and viral genotype in chronic hepatitis B and hepatocellular carcinoma in Guangxi, China. *J Med Virol*, 68, 33-40.

FENG, H., SHUDA, M., CHANG, Y. & MOORE, P.S. 2008. Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science*, 319, 1096-1100.

FERNHOLZ, D., GALLE, P. R., STEMLER, M., BRUNETTO, M., BONINO, F., & WILL, H. 1993. Infectious hepatitis B virus variant defective in pre-S2 protein expression in a chronic carrier. *Virology*, *194*, 137-148.

FISCHER, K.P., GUTFREUND, K.S. & TYRRELL, D.L. 2001. Lamivudine resistance in hepatitis B: mechanisms and clinical implications. *Drug Resist Updat*, 4, 118-127.

FLINK, H.J., VAN ZONNEVELD, M., HANSEN, B. E., DE MAN, R. A., SCHALM, S. W., JANSSEN, H. L., & HBV 99-01 STUDY GROUP. 2006. Treatment with Peg-interferon alpha-2b for HBeAg-positive chronic hepatitis B: HBsAg loss is associated with HBV genotype. *The American journal of gastroenterology*, *101*, 297-303.

FLODELL, S., SCHLEUCHER, J., CROMSIGT, J., IPPEL, H., KIDD-LJUNGGREN, K., & WIJMENGA, S. 2002. The apical stem-loop of the hepatitis B virus encapsidation signal folds into a stable tri-loop with two underlying pyrimidine bulges. *Nucleic acids research*, *30*, 4803-4811.

FRANCO, E., BAGNATO, B., MARINO, M. G., MELELEO, C., SERINO, L., & ZARATTI, L. 2012. Hepatitis B: Epidemiology and prevention in developing countries. *World journal of hepatology*, *4*, 74-80.

FUKANO, K., TSUKUDA, S., OSHIMA, M., SUZUKI, R., AIZAKI, H., OHKI, M., PARK, S.Y., MURAMATSU, M., WAKITA, T., SUREAU, C., OGASAWARA, Y., WATASHI, K., 2018. Troglitazone impedes the oligomerization of sodium taurocholate cotransporting polypeptide and entry of hepatitis B virus into hepatocytes. Front. Microbiol. 9, 3257.

GANEM, D. & PRINCE, A.M. 2004. Hepatitis B Virus Infection - Natural History and Clinical Consequences. *N Engl J Med*, 350, 1118-1129.

GAO, Z.Y., LI, T., WANG, J., MEI DU, J., JUAN, L., LI, J., MINLU, F., & ZHUANG, H. 2007. Mutations in preS genes of genotype C hepatitis B virus in patients with chronic hepatitis B and hepatocellular carcinoma. *J Gastroenterol*, 42, 761-768.

GARCIA, P. D., OU, J. H., RUTTER, W. J., & WALTER, P. 1988. Targeting of the hepatitis B virus precore protein to the endoplasmic reticulum membrane: after signal peptide cleavage translocation can be aborted and the product released into the cytoplasm. *The Journal of cell biology*, *106*, 1093-1104.

GASIM G. I. 2013. Hepatitis B virus in the Arab world: where do we stand?. *Arab journal of gastroenterology: the official publication of the Pan-Arab Association of Gastroenterology*, *14*, 35-43.

GENG, M., XIN, X., BI, L.-Q., ZHOU, L.-T. & LIU, X.-H. 2015. Molecular mechanism of hepatitis B virus X protein function in hepatocarcinogenesis. *World J Gastroenterol*, 21, 10732-10738.

GERKEN, G., KREMSDORF, D., CAPEL, F., PETIT, M. A., DAUGUET, C., MANNS, M. P., MEYER ZUM BÜSCHENFELDE, K. H., & BRECHOT, C. 1991. Hepatitis B defective virus with rearrangements in the preS gene during chronic HBV infection. *Virology*, *183*, 555-565.

GERLICH, W.H. 2013. Medical Virology of hepatitis B: how it began and where we are now. *Virol J*., 10, 239.

GHANY, M. G., AYOLA, B., VILLAMIL, F. G., GISH, R. G., ROJTER, S., VIERLING, J. M., & LOK, A. S. 1998. Hepatitis B virus S mutants in liver transplant recipients who were reinfected despite hepatitis B immune globulin prophylaxis. *Hepatology (Baltimore, Md.)*, *27*1, 213-222.

GHANY, M.G. & DOO, E.C. 2009. Antiviral resistance and hepatitis B therapy. *Hepatology (Baltimore, Md.)*, 49, S174-S184.

GLEBE, D., ALIAKBARI, M., KRASS, P., KNOOP, E. V., VALERIUS, K. P., & GERLICH, W. H. 2003. Pre-s1 antigen-dependent infection of Tupaia hepatocyte cultures with human hepatitis B virus. *Journal of virology*, *77*, 9511-9521.

GODDARD, C. D., BILDRICI-ERTEKIN, L., WANG, X., & CUCONATI, A. 2017. Microtiter-Format Assays for HBV Antigen Quantitation in Nonclinical Applications. *Methods in molecular biology (Clifton, N.J.)*, *1540*, 203-209.

GRETHE, S., HECKEL, J.O., RIETSCHEL, W. &HUFERT, F.T. 2000. Molecular epidemiology of hepatitis B virus variants in nonhuman primates. *J Virol*., 74, 5377-5381.

GRIMM, D., THIMME, R., & BLUM, H.E. 2011. HBV life cycle and novel drug targets. *Hepatology International*, 5, 644-653.

GRIPON, P., CANNIE, I., & URBAN, S. 2005. Efficient inhibition of hepatitis B virus infection by acylated peptides derived from the large viral surface protein. *Journal of virology*, *79*, 1613-1622.

GRIPON, P., DIOT, C., & GUGUEN-GUILLOUZO, C. 1993. Reproducible high-level infection of cultured adult human hepatocytes by hepatitis B virus: effect of polyethylene glycol on adsorption and penetration. *Virology*, *192*, 534-540.

GRUBER, W., *&* VIRCHOW, R. *18*65. Occurrence and diagnosis of the hepatic, and in particular, the catarrhal icterus. Virchows Arch, 32, 117-125.

GÜNTHER, S., LI, B. C., MISKA, S., KRÜGER, D. H., MEISEL, H., & WILL, H. 1995. A Novel Method for Efficient Amplification of Whole Hepatitis B Virus Genomes Permits Rapid Functional Analysis and Reveals Deletion Mutants in Immunosuppressed Patients. *J Virol*, 69, 5437-5444.

GUO, H., & CUCONATI, A. 2017. *Hepatitis B Virus: Methods and Protocols*. New York, NY: Humana Press, vol. 1540.

GUO, H., JIANG, D., ZHOU, T., CUCONATI, A., BLOCK, T. M., & GUO, J. T. 2007. Characterization of the intracellular deproteinized relaxed circular DNA of hepatitis B virus: an intermediate of covalently closed circular DNA formation. *Journal of virology*, *81*, 12472-12484.

HABIB, S. & SHAIKH, O.S. 2007. Hepatitis B immune globulin. *Drugs Today (Barc)*, 43, 379-394.

HADZIYANNIS, S. J., & PAPATHEODORIDIS, G. V. 2006. Hepatitis B e antigen-negative chronic hepatitis B: natural history and treatment. *Seminars in liver disease*, *26*, 130–141.

HADZIYANNIS, S.J. 2011. Natural history of chronic hepatitis B in Euro-Mediterranean and African countries. *J Hepatol*, 55, 183-91.

HAMIDI-FARD, M., MAKVADI, M., SAMARBAF-ZADEH, A., HAJIANI, E., SHAYESTEH, A. &, A. 2013. Mutation analysis of hepatitis B virus reverse transcriptase region among untreated chronically infected patients in Ahvaz city (South-West of Iran). *Indian J Med Microbiol, 31,* 360-365.

HAO, X., SHANG, X., WU, J., SHAN, Y., CAI, M., JIANG, J., HUANG, Z., TANG, Z. & WANG, H. 2011, Single-Particle Tracking of Hepatitis B Virus-like Vesicle Entry into Cells. *Small (Weinheim an der Bergstrasse, Germany)*, 7, 1212-1218.

HAYES, S., MAHONY, J., NAUTA, A., & VAN SINDEREN, D. 2017. Metagenomic Approaches to Assess Bacteriophages in Various Environmental Niches. *Viruses*, *9*, 127.

HE, X., WANG, F., HUANG, B., CHEN, P., & ZHONG, L. 2015. Detection and analysis of resistance mutations of hepatitis B virus. International Journal of Clinical and Experimental Medicine, 8, 9630-9639.

HEATHCOTE, E. J., MARCELLIN, P., BUTI, M., GANE, E., DE MAN, R. A., KRASTEV, Z., GERMANIDIS, G., LEE, S. S., FLISIAK, R., KAITA, K., MANNS, M., KOTZEV, I., TCHERNEV, K., BUGGISCH, P., WEILERT, F., KURDAS, O. O., SHIFFMAN, M. L., TRINH, H., GUREL, S., SNOW-LAMPART, A., BORROTO-ESODA, K., MONDOU, E., ANDERSON, J., SORBEL, J, & ROUSSEAU, F. 2011. Three-year efficacy and safety of tenofovir disoproxil fumarate treatment for chronic hepatitis B. *Gastroenterology*, *140*, 132-143.

HEBELER-BARBOSA, F., WOLF, I. R., VALENTE, G. T., MELLO, F., LAMPE, E., PARDINI, M., & GROTTO, R. 2020. A New Method for Next-Generation Sequencing of the Full Hepatitis B Virus Genome from A Clinical Specimen: Impact for Virus Genotyping. *Microorganisms*, *8*, 1391.

HERMANS, L. E., SVICHER, V., PAS, S. D., SALPINI, R., ALVAREZ, M., BEN ARI, Z., BOLAND, G., BRUZZONE, B., COPPOLA, N., SEGUIN-DEVAUX, C., DYDA, T., GARCIA, F., KAISER, R., KÖSE, S., KRARUP, H., LAZAREVIC, I., LUNAR, M. M., MAYLIN, S., MICHELI, V., MOR, O., … HEPVIR WORKING GROUP OF THE EUROPEAN SOCIETY FOR TRANSLATIONAL ANTIVIRAL RESEARCH 2016. Combined Analysis of the Prevalence of Drug-Resistant Hepatitis B Virus in Antiviral Therapy-Experienced Patients in Europe (CAPRE). *The Journal of infectious diseases*, *213*, 39-48.

HILLIER, L. W., MARTH, G. T., QUINLAN, A. R., DOOLING, D., FEWELL, G., BARNETT, D., FOX, P., GLASSCOCK, J. I., HICKENBOTHAM, M., HUANG, W., MAGRINI, V. J., RICHT, R. J., SANDER, S. N., STEWART, D. A., STROMBERG, M., TSUNG, E. F., WYLIE, T., SCHEDL, T., WILSON, R. K. & MARDIS, E. R. 2008. Whole-genome sequencing and variant discovery in C. elegans. Nat Methods, 5, 183-8.

HOARE, J., HENKLER, F., DOWLING, J. J., ERRINGTON, W., GOLDIN, R. D., FISH, D., & MCGARVEY, M. J. 2001. Subcellular localisation of the X protein in HBV infected hepatocytes. *Journal of medical virology*, *64*, 419-426.

HONER ZU SIEDERDISSEN, C. & CORNBERG, M. 2014. The role of HBsAg levels in the current management of chronic HBV infection. *Ann Gastroenterol*, 27, 105-112.

HOU, J., LIU, Z., & GU, F. 2005. Epidemiology and Prevention of Hepatitis B Virus Infection. *International journal of medical sciences*, *2*, 50-57.

HOULDCROFT, C. J., BEALE, M. A., & BREUER, J. 2017. Clinical and biological insights from viral genome sequencing. *Nature reviews. Microbiology*, *15*, 183-192.

HOWARD, C. R. 1986. The Biology of Hepadnaviruses. *Journal of General Virology*, 67, 1215-1235.

HOWARD, C.R. 1994. Classification and taxonomy of hepadnaviruses - current status. In: Viral Hepatitis and Liver Disease. Edited by: Nishioka, K., Suzuki, H., Mishiro, S. and Oda, T. Tokyo, Japan., Springer-Verlag, 54-56.

HOWARD, C.R. 2002. Hepatitis viruses: a pandora's box? *J Gastroenterol Hepatol*, 17 Suppl., S464-7.

HSU, H. Y., CHANG, M. H., LIAW, S. H., NI, Y. H., & CHEN, H. L. 1999. Changes of hepatitis B surface antigen variants in carrier children before and after universal vaccination in Taiwan. *Hepatology*, 30, 1312-7.

HUI, C. K., ZHANG, H. Y., BOWDEN, S., LOCARNINI, S., LUK, J. M., LEUNG, K. W., YUENG, Y. H., WONG, A., ROUSSEAU, F., YUEN, K. Y., NAOUMOV, N. N., & LAU, G. K. 2008. 96 weeks combination of adefovir dipivoxil plus emtricitabine vs. adefovir dipivoxil monotherapy in the treatment of chronic hepatitis B. *Journal of hepatology*, *48*, 714-720.

HUNT, C.M., MCGILL, J.M., ALLEN, M.I. & CONDREAY, L.D. 2000. Clinical relevance of hepatitis B viral mutations. *Hepatology*, 31, 1037-1044.

HUOVILA, A.P., EDER, A.M. & FULLER, S.D. 1992. Hepatitis B surface antigen assembles in a post-ER, pre-Golgi compartment. *J Cell Biol*, 118, 1305-1320.

HWANG, L.Y., ROGGENDORF, M., BEASLEY, P. & DEINHARDT, F. 1985. Perinatal transmission of hepatitis B virus: Role of maternal HBeAg and anti-HBc IgM. *Journal of Medical Virology*, 15, 265-269.

ISHII, T., TAMURA, A., SHIBATA, T., KURODA, K., KANDA, T., SUGIYAMA, M., MIZOKAMI, M., & MORIYAMA, M. 2020. Analysis of HBV Genomes Integrated into the Genomes of Human Hepatoma PLC/PRF/5 Cells by HBV Sequence Capture-Based Next-Generation Sequencing. *Genes*, *11*, 661.

ISMAIL, A. M., SAMUEL, P., EAPEN, C. E., KANNANGAI, R., & ABRAHAM, P. 2012. Antiviral resistance mutations and genotype-associated amino acid substitutions in treatment-naive hepatitis B virus-infected individuals from the Indian subcontinent. *Intervirology*, *55*, 36-44.

IWAMOTO, M., CAI, D., SUGIYAMA, M., SUZUKI, R., AIZAKI, H., RYO, A., OHTANI, N., TANAKA, Y., MIZOKAMI, M., WAKITA, T., GUO, H., & WATASHI, K. 2017. Functional association of cellular microtubules with viral capsid assembly supports efficient hepatitis B virus replication. *Scientific reports*, *7*, 10620.

IWAMOTO, M., SASO, W., NISHIOKA, K., OHASHI, H., SUGIYAMA, R., RYO, A., OHKI, M., YUN, J. H., PARK, S.Y., OHSHIMA, T., SUZUKI, R., AIZAKI, H., MURAMATSU, M., MATANO, T., IWAMI, S., SUREAU, C., WAKITA, T. & WATASHI, K. 2020. The machinery for endocytosis of epidermal growth factor receptor coordinates the transport of incoming hepatitis B virus to the endosomal network. J. Biol. Chem. 295, 800–807.

IWAMOTO, M., SASO, W., SUGIYAMA, R., ISHII, K., OHKI, M., NAGAMORI, S., SUZUKI, R., AIZAKI, H., RYO, A., YUN, J.H., PARK, S.Y., OHTANI, N., MURAMATSU, M., IWAMI, S., TANAKA, Y., SUREAU, C., WAKITA, T., & WATASHI, K. 2019. Epidermal growth factor receptor is a host-entry cofactor triggering hepatitis B virus internalization. Proc. Natl. Acad. Sci. U. S. A, 116, 8487-8492.

IWAMOTO, M., WATASHI, K., TSUKUDA, S., ALY, H. H., FUKASAWA, M., FUJIMOTO, A., SUZUKI, R., AIZAKI, H., ITO, T., KOIWAI, O., KUSUHARA, H., & WAKITA, T. 2014. Evaluation and identification of hepatitis B virus entry inhibitors using HepG2 cells overexpressing a

membrane transporter NTCP. *Biochemical and biophysical research communications*, *443*, 808-813.

JANSSEN, H. L., VAN ZONNEVELD, M., SENTURK, H., ZEUZEM, S., AKARCA, U. S., CAKALOGLU, Y., SIMON, C., SO, T. M., GERKEN, G., DE MAN, R. A., NIESTERS, H. G., ZONDERVAN, P., HANSEN, B., SCHALM, S. W., HBV 99-01 STUDY GROUP, & ROTTERDAM FOUNDATION FOR LIVER RESEARCH 2005. Pegylated interferon alfa-2b alone or in combination with lamivudine for HBeAg-positive chronic hepatitis B: a randomised trial. *Lancet (London, England)*, *365*, 123-129.

JARDI, R., RODRIGUEZ, F., BUTI, M., COSTA, X., VALDES, A., ALLENDE, H., SCHAPER, M., GALIMANY, R., ESTEBAN, R. & GUARDIA, J. 2004. Mutations in the basic core promoter region of hepatitis B virus. Relationship with precure variants and HBV genotypes in a Spanish population of HBV carriers. *J Hepatol,* 40, 507-514.

JIANG, B., & HILDT, E. 2020. Intracellular Trafficking of HBV Particles. *Cells*, *9*, 2023.

JIANG, B., HIMMELSBACH, K., REN, H., BOLLER, K., & HILDT, E., 2015. Subviral hepatitis B virus filaments, like infectious viral particles, are released via multivesicular bodies. J. Virol, 90, 3330-3341.

JIANG, S. W., YAO, L. P., HU, A. R., HU, Y. R., CHEN, S. X., XIONG, T., GAO, G. S., LIANG, X. Y., DING, S. X., & WENG, P. J. 2014. Resistant mutants induced by adefovir dipivoxil in hepatitis B virus isolates. *World journal of gastroenterology*, *20*, 17100-17106.

JOHANNESSEN, A., MEKASHA, B., DESALEGN, H., ABERRA, H., STENE-JOHANSEN, K., & BERHE, N. 2021. Mother-to-Child Transmission of Hepatitis B Virus in Ethiopia. *Vaccines*, *9*, 430.

JONES, S.A. AND HU, J. 2013. Hepatitis B virus reverse transcriptase: diverse functions as classical and emerging targets for antiviral intervention. *Emerging Microbes & Infections*, 2, 1-11.

KANG, L., PAN, J., WU, J., HU, J., SUN, Q., & TANG, J. 2015. Anti-HBV Drugs: Progress, Unmet Needs, and New Hope. *Viruses*, *7*, 4960-4977.

KANN, M. 2002. Structural and molecular *Virology*. In: Lai, C.L. & Locarnini, S., eds., *Hepatitis B virus*. London: International Medical Press, 9-21.

KANN, M., SODEIK, B., VLACHOU, A., GERLICH, W.H., & HELENIUS, A., 1999. Phosphorylation-dependent binding of hepatitis B virus core particles to the nuclear pore complex. J. Cell Biol, 145, 45-55.

KAO, J & CHEN, D. 2006. HBV genotypes: Epidemiology and implications regarding natural history. Current Hepatitis Reports, 5, 5-13.

KAO, J.H., CHEN, P.J., LAI, M.Y. & CHEN D.S. 2003. Basal core promoter mutations of hepatitis B virus increase the risk of hepatocellular carcinoma in hepatitis B carriers. *Gastroenterology,* 124, 327-334.

KARAYIANNIS, P. 2017. Hepatitis B virus: Virology, molecular biology, life cycle and intrahepatic spread. *Hepatol Int*., 11, 500-508.

KAWAMOTO, S., UEDA, K., MITA, E., & MATSUBARA, K. 1994. The packaging signal in hepatitis B virus pregenome functions only at the 5' end. *Journal of virological methods*, *49*, 113-127.

KEARNEY, B. P., FLAHERTY, J. F., & SHAH, J. 2004. Tenofovir disoproxil fumarate: clinical pharmacology and pharmacokinetics. *Clinical pharmacokinetics*, *43*, 595-612.

KEDDA, M. A., KRAMVIS, A., KEW, M. C., LECATSAS, G., PATERSON, A. C., ASPINALL, S., STARK, J. H., DE KLERK, W. A., & GRIDELLI, B. 2000. Susceptibility of chacma baboons (Papio ursinus orientalis) to infection by hepatitis B virus. *Transplantation*, *69*, 1429-1434.

KENNEDY, E. M., BASSIT, L. C., MUELLER, H., KORNEPATI, A., BOGERD, H. P., NIE, T., CHATTERJEE, P., JAVANBAKHT, H., SCHINAZI, R. F., & CULLEN, B. R. 2015. Suppression of hepatitis B virus DNA accumulation in chronically infected cells using a bacterial CRISPR/Cas RNA-guided DNA endonuclease. *Virology*, *476*, 196-205.

KEW, M.C. 2011. Hepatitis B virus x protein in the pathogenesis of hepatitis B virus induced hepatocellular carcinoma. *J Gastroenterol Hepatol*, 26, 144-152.

KHEDIVE, A., NOROUZI, M., RAMEZANI, F., KARIMZADEH, H., ALAVIAN, S. M., MALEKZADEH, R., MONTAZERI, G., NEJATIZADEH, A., ZIAEE, M., ABEDI, F., ATAEI, B., YARAN, M., SAYAD, B., SOMI, M. H., SARIZADEH, G., SANEI-MOGHADDAM, I., MANSOUR-GHANAEI, F., RAFATPANAH, H., POURHOSSEINGHOLI, M. A., KEYVANI, H., … JAZAYERI, S. M. 2013. Hepatitis B virus surface protein mutations clustered mainly in CTL immune epitopes in chronic carriers: results of an Iranian nationwide study. *Journal of viral hepatitis*, *20*, 494-501.

KITAMURA, K., QUE, L., SHIMADU, M., KOURA, M., ISHIHARA, Y., WAKAE, K., NAKAMURA, T., WATASHI, K., WAKITA, T., & MURAMATSU, M. 2018. Flap endonuclease 1 is involved in cccDNA formation in the hepatitis B virus. PLoS Pathog. 14, e1007124.

KNOWLES, B. B., HOWE, C. C., & ADEN, D. P. 1980. Human hepatocellular carcinoma cell lines secrete the major plasma proteins and hepatitis B surface antigen. *Science (New York, N.Y.)*, *209*, 497-499.

KO, C., BESTER, R., ZHOU, X., XU, Z., BLOSSEY, C., SACHERL, J., VONDRAN, F., GAO, L., & PROTZER, U. 2019. A New Role for Capsid Assembly Modulators To Target Mature Hepatitis B Virus Capsids and Prevent Virus Infection. *Antimicrobial agents and chemotherapy*, *64*, e01440-19.

KO, C., CHAKRABORTY, A., CHOU, W. M., HASREITER, J., WETTENGEL, J. M., STADLER, D., BESTER, R., ASEN, T., ZHANG, K., WISSKIRCHEN, K., MCKEATING, J. A., RYU, W. S., & PROTZER, U. 2018. Hepatitis B virus genome recycling and de novo secondary infection events maintain stable cccDNA levels. *Journal of hepatology*, *69*, 1231-1241.

KONIGER, C., WINGERT, I., MARSMANN, M., ROSLER, C., BECK, J., & NASSAL, M., 2014. Involvement of the host DNA-repair enzyme TDP2 in formation of the covalently

closed circular DNA persistence reservoir of hepatitis B viruses. Proc. Natl. Acad. Sci. U. S. A. 111, E4244-E4253.

KORBEL, J.O., URBAN, A.E., AFFOURTIT, J.P., GODWIN, B., GRUBERT, F., SIMONS, J.F., KIM, P.M., PALEJEV, D., CARRIERO, N.J., DU, L., TAILLON, B.E., CHEN, Z., TANZER, A., SAUNDERS, A.C., CHI, J., YANG, F., CARTER, N.P., HURLES, M.E., WEISSMAN, S.M., HARKINS, T.T., GERSTEIN, M.B., EGHOLM, M. & SNYDER, M. 2007. Paired-end mapping reveals extensive structural variation in the human genome. *Science*, 318, 420-426.

KOUMBI, L. 2015. Current and future antiviral drug therapies of hepatitis B chronic infection. *World journal of hepatology*, *7*, 1030-1040.

KOZAREWA, I., ARMISEN, J., GARDNER, A. F., SLATKO, B. E., & HENDRICKSON, C. L. 2015. Overview of Target Enrichment Strategies. *Current protocols in molecular biology*, *112*, 7.21.1-7.21.23.

KRAMVIS, A. 2016. The clinical implications of hepatitis B virus genotypes and HBeAg in pediatrics. *Rev Med Virol*, 26, 285-303.

KRAMVIS, A., RESTORP, K., NORDER, H., BOTHA, J. F., MAGNIUS, L. O., & KEW, M. C. 2005. Full genome analysis of hepatitis B virus genotype E strains from South-Western Africa and Madagascar reveals low genetic variability. *Journal of medical virology*, *77*, 47-52.

KUO, M. Y., CHAO, M., & TAYLOR, J. 1989. Initiation of replication of the human hepatitis delta virus genome from cloned DNA: role of delta antigen. *Journal of virology*, *63*, 1945-1950.

KUROKI, K., FLOREANI, M., MIMMS, L. T., & GANEM, D. 1990. Epitope mapping of the PreS1 domain of the hepatitis B virus large surface protein. *Virology*, *176*, 620-624.

LADA, O., BENHAMOU, Y., POYNARD, T., & THIBAULT, V. 2006. Coexistence of hepatitis B surface antigen (HBs Ag) and anti-HBs antibodies in chronic hepatitis B virus carriers: influence of "a" determinant variants. *Journal of virology*, *80*, 29682975.

LADNER, S. K., OTTO, M. J., BARKER, C. S., ZAIFERT, K., WANG, G. H., GUO, J. T., SEEGER, C., & KING, R. W. 1997. Inducible expression of human hepatitis B virus (HBV) in stably transfected hepatoblastoma cells: a novel system for screening potential inhibitors of HBV replication. *Antimicrobial agents and chemotherapy*, *41*, 1715-1720.

LAMONTAGNE, R. J., BAGGA, S., & BOUCHARD, M. J. 2016. Hepatitis B virus molecular biology and pathogenesis. *Hepatoma research*, *2*, 163-186.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., … INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM. 2001. Initial sequencing and analysis of the human genome. *Nature*, *409*, 860-921.

LANFORD, R. E., CHAVEZ, D., BRASKY, K. M., BURNS, R. B., 3RD, & RICO-HESSE, R. 1998. Isolation of a hepadnavirus from the woolly monkey, a New World

primate. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 5757-5761.

LANGLEY, D. R., WALSH, A. W., BALDICK, C. J., EGGERS, B. J., ROSE, R. E., LEVINE, S. M., KAPUR, A. J., COLONNO, R. J., & TENNEY, D. J. 2007. Inhibition of hepatitis B virus polymerase by entecavir. *Journal of virology*, *81*, 3992-4001.

LARAS, A., KOSKINAS, J. & HADZIYANNIS, S.J. 2002. In vivo suppression of precore mRNA synthesis is associated with mutations in the hepatitis B virus core promoter. *Virology,* 295, 86-96.

LAU, G., MARCELLIN, P., & PETERS, M. 2007. Chronic hepatitis B: a global health problem requiring coherent worldwide treatment strategies. *Hepatology international*, *1*, 316-325.

LAVANCHY, D. 2004. Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. *J Viral Hepat*, 11, 97-107.

LAVANCHY, D. 2012. Viral hepatitis: Global goals for vaccination. *J ClinVirol*, 55, 296-302.

LAZAREVIC, I. 2014. Clinical implications of hepatitis B virus mutations: recent advances. *World journal of gastroenterology*, *20*, 7653-7664.

LAZAREVIC, I., BANKO, A., MILJANOVIC, D., & CUPIC, M. 2019. Immune-escape hepatitis B virus mutations associated with viral reactivation upon immunosuppression. *Viruses*, 11, 778.

LEE, A. J., LEE, C. H., & JEON, C. H. 2014. Analysis of reverse transcriptase gene mutations in the hepatitis B virus at a university hospital in Korea. *Annals of laboratory medicine*, *34*3, 230-234.

LEMOINE, M., & THURSZ, M. R. 2017. Battlefield against hepatitis B infection and HCC in Africa. *Journal of hepatology*, *66*, 645–654.

LEMPP, F., SCHLUND, F., RIEBLE, L., NUSSBAUM, L., LINK, C., ZHANG, Z., NI, Y., & URBAN, S. 2019. Recapitulation of HDV infection in a fully permissive hepatoma cell line allows efficient drug evaluation. *Nature Communications,* 10, 2265.

LI, Y.W., YANG, F.C., LU, H.Q. & ZHANG, J.S. 2016. Hepatocellular carcinoma and hepatitis B surface protein. *World Journal of Gastroenterology*, 22, 1943-1952.

LIANG, T.J. 2009. Hepatitis B: the virus and disease. *Hepatology*, 49, 13-21.

LIAW, Y.F. & CHU, C.M. 2009. Hepatitis B virus infection. *Lancet*, 373, 582-592.

LIM, Y. S. 2017. Management of Antiviral Resistance in Chronic Hepatitis B. *Gut and liver*, *11*, 189-195.

LIN, C. L., & KAO, J. H. 2016. Review article: novel therapies for hepatitis B virus cure - advances and perspectives. *Alimentary pharmacology & therapeutics*, *44*, 213-222.

LIN, L., YANG, C., & KAO. 2016, Hepatitis B virus: new therapeutic perspectives. Liver Int, 36, 85-92.

LIN, M., CHEN, Q., YANG, L. Y., LI, W. Y., CAO, X. B., WU, J. R., PENG, Y. P., & CHEN, M. R. 2007. Hepatitis B virus infection and replication in primarily cultured human fetal hepatocytes. *World journal of gastroenterology*, *13*, 1027-1031.

LIN, P. F., NOWICKA-SANS, B., TERRY, B., ZHANG, S., WANG, C., FAN, L., DICKER, I., GALI, V., HIGLEY, H., PARKIN, N., TENNEY, D., KRYSTAL, M., & COLONNO, R. 2008. Entecavir exhibits inhibitory activity against human immunodeficiency virus under conditions of reduced viral challenge. *Antimicrobial agents and chemotherapy*, *52*, 1759-1767.

LIN, S. M., YU, M. L., LEE, C. M., CHIEN, R. N., SHEEN, I. S., CHU, C. M., & LIAW, Y. F. 2007. Interferon therapy in HBeAg positive chronic hepatitis reduces progression to cirrhosis and hepatocellular carcinoma. *Journal of hepatology*, *46*, 45-52.

LINDH, M., HANNOUN, C., DHILLON, A.P., NORKRANS, G. & HORAL, P. 1999. Core promoter mutations and genotypes in relation to viral replication and liver damage in East Asian hepatitis B virus carriers. *J Infect Dis*, 179, 775-782.

LINDH, M., HORAL, P., DHILLON, A.P., FURUTA, Y. & NORKRANS, G. 1996. Hepatitis B virus carriers without precore mutations in hepatitis B e antigen negative stage show more severe liver damage. *Hepatology*, 24, 494-501.

LINGALA, S., LAU, D. T., KOH, C., AUH, S., GHANY, M. G., & HOOFNAGLE, J. H. 2016. Long-term lamivudine therapy in chronic hepatitis B. *Alimentary pharmacology & therapeutics*, *44*, 380-389.

LIU, B. M., LI, T., XU, J., LI, X. G., DONG, J. P., YAN, P., YANG, J. X., YAN, L., GAO, Z. Y., LI, W. P., SUN, X. W., WANG, Y. H., JIAO, X. J., HOU, C. S., & ZHUANG, H. 2010. Characterization of potential antiviral resistance mutations in hepatitis B virus reverse transcriptase sequences in treatment-naïve Chinese patients. *Antiviral research*, *85*3, 512-519.

LIU, H.-Y. & ZHANG, X.-Y. 2015. Innate immune recognition of hepatitis B virus. *World J* Hepatol, 7, 2319-2322.

LIU, K., LUCKENBAUGH, L., NING, X., XI, J., & HU, J. 2018. Multiple roles of core protein linker in hepatitis B virus replication. *PLoS pathogens*, *14*, e1007085.

LIU, L. J., WANG, J. H., DU, S. C., TIAN, J. H., YANG, R. F., & WEI, L. 2010. rtE218G, a novel hepatitis B virus mutation with resistance to adefovir dipivoxil in patients with chronic hepatitis B. *Journal of viral hepatitis*, *17 Suppl 1*, 66-72.

LOCARNINI S. 2008. Primary resistance, multidrug resistance, and cross-resistance pathways in HBV as a consequence of treatment failure. *Hepatology international*, *2*, 147-151.

LOCARNINI, S. & ZOULIM, F. 2010. Molecular genetics of HBV infection. *Antivir Ther*, 15 Suppl 3, 3-14.

LOCARNINI, S. 2004. Molecular Virology of hepatitis B virus. *Semin Liver Dis*, 24 Suppl 1, 3-10.

LOCARNINI, S. A., & YUEN, L. 2010. Molecular genesis of drug-resistant and vaccine-escape HBV mutants. *Antiviral therapy*, *15*, 451-461.

LOCARNINI, S., HATZAKIS, A., HEATHCOTE, J., KEEFFE, E. B., LIANG, T. J., MUTIMER, D., PAWLOTSKY, J. M., & ZOULIM, F. 2004. Management of antiviral resistance in patients with chronic hepatitis B. *Antiviral therapy*, *9*, 679-693.

LOCARNINI, S., MCMILLAN, J., & BARTHOLOMEUSZ, A. 2003.The hepatitis B virus and common mutants. Semin Liver Dis, 23, 5-20.

LOCARNINI, S.A. & ROGGENDORF, M. 2013. Other Hepadnaviridae (Avihepadnaviridae (DHBV) and Orthohepadnaviridae (WHV)). In Viral Hepatitis (eds H.C. Thomas, A.S. Lok, S.A. Locarnini and A.J. Zuckerman).

LOK, A. S., GANOVA-RAEVA, L., CLOONAN, Y., PUNKOVA, L., LIN, H. S., LEE, W. M., GHANY, M. G., & HEPATITIS B RESEARCH NETWORK (HBRN). 2017. Prevalence of hepatitis B antiviral drug resistance variants in North American patients with chronic hepatitis B not receiving antiviral treatment. *Journal of viral hepatitis*, *24*, 1032-1042.

LOK, A. S., LAI, C. L., LEUNG, N., YAO, G. B., CUI, Z. Y., SCHIFF, E. R., DIENSTAG, J. L., HEATHCOTE, E. J., LITTLE, N. R., GRIFFITHS, D. A., GARDNER, S. D., & CASTIGLIA, M. 2003. Long-term safety of lamivudine treatment in patients with chronic hepatitis B. *Gastroenterology*, *125*, 1714-1722.

LOK, A. S., TRINH, H., CAROSI, G., AKARCA, U. S., GADANO, A., HABERSETZER, F., SIEVERT, W., WONG, D., LOVEGREN, M., COHEN, D., & LLAMOSO, C. 2012. Efficacy of entecavir with or without tenofovir disoproxil fumarate for nucleos(t)ide-naïve patients with chronic hepatitis B. *Gastroenterology*, *143*, 619-628.

LONG, J., WANG, M., CHEN, B., ZHANG, J., & WANG, B. 2017. Chronic hepatitis B: a critical discussion. *Future Virology,* 12, 3.

LOPATIN U. 2019. Drugs in the Pipeline for HBV. *Clinics in liver disease*, *23*, 535-555.

LOVETT, G. C., NGUYEN, T., ISER, D. M., HOLMES, J. A., CHEN, R., DEMEDIUK, B., SHAW, G., BELL, S. J., DESMOND, P. V., & THOMPSON, A. J. 2017. Efficacy and safety of tenofovir in chronic hepatitis B: Australian real world experience. *World journal of hepatology*, *9*, 48-56.

LOWE, C. F., MERRICK, L., HARRIGAN, P. R., MAZZULLI, T., SHERLOCK, C. H., & RITCHIE, G. 2016. Implementation of Next-Generation Sequencing for Hepatitis B Virus Resistance Testing and Genotyping in a Clinical Microbiology Laboratory. *Journal of clinical microbiology*, *54*, 127-133.

LUO, A., JIANG, X., & REN, H. 2018. Lamivudine plus tenofovir combination therapy versus lamivudine monotherapy for HBV/HIV coinfection: a meta-analysis. *Virology journal*, *15*, 139.

LUO, J., XI, J., GAO, L., & HU, J. 2020. Role of Hepatitis B virus capsid phosphorylation in nucleocapsid disassembly and covalently closed circular DNA formation. *PLoS pathogens*, *16*, e1008459.

LURMAN, A. 1885. Eine icterus epidemic. *Berl Klin Woschenschr*, 22, 20-23.

MAGNIUS, L.O. & NORDER, H. 1995. Subtypes, genotypes and molecular epidemiology of the hepatitis B virus as reflected by sequence variability of the S-gene. *Intervirology*, 38, 24-34.

MAHMOOD, M. & ANWAR, M.A. 2017. Analysis of resistant mutations in reverse transcriptase domain of hepatitis B virus from patients from Islamabad, Pakistan. *J Unexplored Med Data*, 2, 60-64.

MARCELLIN, P., CHANG, T. T., LIM, S. G., TONG, M. J., SIEVERT, W., SHIFFMAN, M. L., JEFFERS, L., GOODMAN, Z., WULFSOHN, M. S., XIONG, S., FRY, J., BROSGART, C. L., & ADEFOVIR DIPIVOXIL 437 STUDY GROUP. 2003. Adefovir dipivoxil for the treatment of hepatitis B e antigen-positive chronic hepatitis B. *The New England journal of medicine*, *348*, 808-816.

MARDIS, E.R. 2008. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet*, 9, 387-402.

MARDIS, E.R. 2011. A decade's perspective on DNA sequencing technology. *Nature*, 470, 198-203.

MARDIS, E.R. 2013. Next-generation sequencing platforms. *Annu Rev Anal Chem (Palo Alto Calif),* 6, 287-303.

MARGULIES, M., EGHOLM, M., ALTMAN, W. E., ATTIYA, S., BADER, J. S., BEMBEN, L. A., BERKA, J., BRAVERMAN, M. S., CHEN, Y. J., CHEN, Z., DEWELL, S. B., DU, L., FIERRO, J. M., GOMES, X. V., GODWIN, B. C., HE, W., HELGESEN, S., HO, C. H., IRZYK, G. P., JANDO, S. C., ALENQUER, M. L., JARVIE, T. P., JIRAGE, K. B., KIM, J. B., KNIGHT, J. R., LANZA, J. R., LEAMON, J. H., LEFKOWITZ, S. M., LEI, M., LI, J., LOHMAN, K. L., LU, H., MAKHIJANI, V. B., MCDADE, K. E., MCKENNA, M. P., MYERS, E. W., NICKERSON, E., NOBILE, J. R., PLANT, R., PUC, B. P., RONAN, M. T., ROTH, G. T., SARKIS, G. J., SIMONS, J. F., SIMPSON, J. W., SRINIVASAN, M., TARTARO, K. R., TOMASZ, A., VOGT, K. A., VOLKMER, G. A., WANG, S. H., WANG, Y., WEINER, M. P., YU, P., BEGLEY, R. F., & ROTHBERG, J. M. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature, 437(7057):376-80

MARION, M. J., HANTZ, O., & DURANTEL, D. 2010. The HepaRG cell line: biological properties and relevance as a tool for cell biology, drug metabolism, and virology studies. *Methods in molecular biology (Clifton, N.J.)*, *640*, 261-272.

MARTIN, N.A. 2003. The discovery of viral hepatitis: A military perspective. *J R Army Med Corps*, 149, 121-124.

MARTINEZ, M. C., KOK, C. C., BALERIOLA, C., ROBERTSON, P., & RAWLINSON, W. D. 2015. Investigation of occult hepatitis B virus infection in anti-hbc positive patients from a liver clinic. *PloS one*, *10*, e0117275.

MARTINOT-PEIGNOUX, M., LAPALUS, M., ASSELAH, T. & MARCELLIN, P. 2013, The role of HBsAg quantification for monitoring natural history and treatment outcome. Liver Int, 33, 125-132.

MASAADEH, H.A., HAYAJNEH, W.A & ALQUDAH, E.A. 2008. Hepatitis B virus genotypes and lamivudine resistance mutations in Jordan. *World J Gastroenterol,* 14, 7231-7234.

MAXAM, A.M. & GILBERT, W. 1977. A new method for sequencing DNA. *Proc Natl Acad Sci USA*, 74, 560-564.

MCMAHON B. J. 2009. The influence of hepatitis B virus genotype and subgenotype on the natural history of chronic hepatitis B. *Hepatology international*, *3*, 334-342.

MCMAHON, B. J., ALWARD, W. L., HALL, D. B., HEYWARD, W. L., BENDER, T. R., FRANCIS, D. P., & MAYNARD, J. E. 1985. Acute hepatitis B virus infection: relation of age to the clinical expression of disease and subsequent development of the carrier state. *The Journal of infectious diseases*, *151*, 599-603.

MCNAUGHTON, A. L., ROBERTS, H. E., BONSALL, D., DE CESARE, M., MOKAYA, J., LUMLEY, S. F., GOLUBCHIK, T., PIAZZA, P., MARTIN, J. B., DE LARA, C., BROWN, A., ANSARI, M. A., BOWDEN, R., BARNES, E., & MATTHEWS, P. C. 2019. Illumina and Nanopore methods for whole genome sequencing of hepatitis B virus (HBV). *Scientific reports*, *9*, 7081.

MERCIER, M., LAPERCHE, S., GIRAULT, A., SUREAU, C., & SERVANT-DELMAS, A. 2011. Overestimation of incidence of hepatitis B virus mixed-genotype infections by use of the new line probe INNO-LiPA genotyping assay. *Journal of clinical microbiology*, *49*, 1154-1156.

METZKER, M.L. 2010. Sequencing technologies - the next generation. *Nat Rev Genet*, 11, 31-46.

MICCO, L., FIORINO, S., LOGGI, E., LORENZINI, S., VITALE, G., CURSARO, C., RIILI, A., BERNARDI, M., & ANDREONE, P. 2009. Polymorphism rtQ215H in primary resistance to adefovir dipivoxil in hepatitis B virus infection: a case report. *BMJ case reports*, *2009*, bcr06, 2008.0287.

MILICH, D. & LIANG, T.J. 2003. Exploring the biological basis of hepatitis B e antigen in hepatitis B virus infection. *Hepatology*, 38, 1075-1086.

MILICH, D. R., MCLACHLAN, A., CHISARI, F. V., KENT, S. B., & THORTON, G. B. 1986. Immune response to the pre-S (1) region of the hepatitis B surface antigen (HBsAg): a pre-S (1)-specific T cell response can bypass nonresponsiveness to the pre-S (2) and S regions of HBsAg. *Journal of immunology (Baltimore, Md.: 1950)*, *137*, 315-322.

MILICH, D. R., THORNTON, G. B., NEURATH, A. R., KENT, S. B., MICHEL, M. L., TIOLLAIS, P., & CHISARI, F. V. 1985. Enhanced immunogenicity of the pre-S region of hepatitis B surface antigen. *Science (New York, N.Y.)*, *228*, 1195-1199.

MOHAMMADNEJAD, L., FARAJNIA, S., PARIVAR, K., NAGHILI, B., & YOUSEFZADEH KHEIRNAGSH, R. 2012. Hepatitis B virus genotypes in eastern Azerbaijan, Northwest Iran. *Archives of Iranian medicine*, *15*, 446-448.

MOKAYA, J., MCNAUGHTON, A. L., BESTER, P. A., GOEDHALS, D., BARNES, E., MARSDEN, B. D., & MATTHEWS, P. C. 2020. Hepatitis B virus resistance to tenofovir: fact or fiction? A

systematic literature review and structural analysis of drug resistance mechanisms. *Wellcome open research*, *5*, 151.

MOOLLA, N., KEW, M. & ARBUTHNOT, P. 2002. Regulatory elements of hepatitis B virus transcription. *J Viral Hepat*, 9, 323-331.

MORIKAWA, K., SUDA, G., & SAKAMOTO, N. 2016. Viral life cycle of hepatitis B virus: Host factors and druggable targets. *Hepatol Res*, 46, 871- 877.

MORREY J. D. 2004. Transgenic hepatitis B virus mouse model in the study of chemotherapy. *Methods in molecular medicine*, *96*, 239–252.

MULDERS, M. N., VENARD, V., NJAYOU, M., EDORH, A. P., BOLA OYEFOLU, A. O., KEHINDE, M. O., MUYEMBE TAMFUM, J. J., NEBIE, Y. K., MAIGA, I., AMMERLAAN, W., FACK, F., OMILABU, S. A., LE FAOU, A., & MULLER, C. P. 2004. Low genetic diversity despite hyperendemicity of hepatitis B virus genotype E throughout West Africa. *The Journal of infectious diseases*, *190*, 400-408.

MURAKAMI, S. 1999. Hepatitis B virus X protein: structure, function and biology. *Intervirology*, 42, 81-99.

MURAKAMI, S., 2001. Hepatitis B virus X protein: a multifunctional viral regulator. *J Gastroenterol*, 36, 651-660.

NAKABAYASHI, H., TAKETA, K., MIYANO, K., YAMANE, T., & SATO, J. 1982. Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. *Cancer research*, *42*, 3858-3863.

NAYAGAM, S., SHIMAKAWA, Y. & MAUD, L. 2020. Mother-to-child transmission of hepatitis B: What more needs to be done to eliminate it around the world? J *Vir Hepatol,* 27, 342-349.

NELSON, N. P., EASTERBROOK, P. J., & MCMAHON, B. J. 2016. Epidemiology of Hepatitis B Virus Infection and Impact of Vaccination on Disease. *Clinics in liver disease*, *20*, 607-628.

NEWBOLD, J. E., XIN, H., TENCZA, M., SHERMAN, G., DEAN, J., BOWDEN, S., & LOCARNINI, S. 1995. The covalently closed duplex form of the hepadnavirus genome exists in situ as a heterogeneous population of viral minichromosomes. *Journal of virology*, *69*, 3350-3357.

NI, Y., LEMPP, F.A., MEHRLE, S., NKONGOLO, S., KAUFMAN, C., FALTH, M., STINDT, J., KONIGER, C., NASSAL, M., KUBITZ, R., SULTMANN, H. & URBAN, S. 2014. Hepatitis B and D viruses exploit sodium taurocholate co-transporting polypeptide for species-specific entry into hepatocytes. *Gastroenterology*, 146, 1070-1083

NI, Y., ZHANG, Z., ENGELSKIRCHER, L., VERCH, G., TU, T., LEMPP, F., & URBAN, S. 2019. Generation and characterization of a stable cell line persistently replicating and secreting the human hepatitis delta virus. *Scientific Reports*; 9, 10021.

NI, Y.H. 2011. Natural history of hepatitis B virus infection: pediatric perspective. *J Gastroenterol*, 46, 1-8.

NIEDERAU, C., HEINTGES, T., LANGE, S., GOLDMANN, G., NIEDERAU, C. M., MOHR, L., & HÄUSSINGER, D. 1996. Long-term follow-up of HBeAg-positive patients treated with interferon alfa for chronic hepatitis B. *The New England journal of medicine*, *334*, 1422-1427.

NORDER, H., EBERT, J.W., FIELDS, H.A., MUSHAHWAR, I.K. & MAGNIUS, L.O. 1995. Complete sequencing of a gibbon hepatitis B virus genome reveals a unique genotype distantly related to the chimpanzee hepatitis B virus. *Virology*, 218, 214-23.

NORDER, H., HAMMAS, B., LEE, S. D., BILE, K., COUROUCÉ, A. M., MUSHAHWAR, I. K., & MAGNIUS, L. O. 1993. Genetic relatedness of hepatitis B viral strains of diverse geographical origin and natural variations in the primary structure of the surface antigen. *The Journal of general virology*, *74,* 1341-1348.

NORDER, H., HAMMAS, B., LÖFDAHL, S., COUROUCÉ, A. M., & MAGNIUS, L. O. 1992. Comparison of the amino acid sequences of nine different serotypes of hepatitis B surface antigen and genomic classification of the corresponding hepatitis B virus strains. *The Journal of general virology*, *73,* 1201-1208.

OCHIYA, T., TSURIMOTO, T., UEDA, K., OKUBO, K., SHIOZAWA, M., & MATSUBARA, K. 1989. An in vitro system for infection with hepatitis B virus that uses primary human fetal hepatocytes. *Proceedings of the National Academy of Sciences of the United States of America*. 86,1875-1879.

OHBA, K., MIZOKAMI, M., OHNO, T., SUZUKI, K., ORITO, E., LAU, J.Y., INA, Y., IKEO, K. & GOJOBORI, T. 1995. Relationships between serotypes and genotypes of hepatitis B virus: genetic classification of HBV by use of surface genes. *Virus Research*, 39, 25-34.

OKAMOTO, H., TSUDA, F., AKAHANE, Y., SUGAI, Y., YOSHIBA, M., MORIYAMA, K., TANAKA, T., MIYAKAWA, Y., & MAYUMI, M. 1994. Hepatitis B virus with mutations in the core promoter for an e antigen-negative phenotype in carriers with antibody to e antigen. *Journal of virology*, *68*, 8102-8110.

OKAMOTO, H., TSUDA, F., SAKUGAWA, H., SASTROSOEWIGNJO, R.I., IMAI, M., MIYAKAWA, Y. & MAYUMI, M. 1988. Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. *J Gen Virol*, 69, 2575-83.

OLINGER, C. M., JUTAVIJITTUM, P., HÜBSCHEN, J. M., YOUSUKH, A., SAMOUNTRY, B., THAMMAVONG, T., TORIYAMA, K., & MULLER, C. P. 2008. Possible new hepatitis B virus genotype, southeast Asia. *Emerging infectious diseases*, *14*, 1777-1780.

OON, C. J., LIM, G. K., YE, Z., GOH, K. T., TAN, K. L., YO, S. L., HOPES, E., HARRISON, T. J., & ZUCKERMAN, A. J. 1995. Molecular epidemiology of hepatitis B virus vaccine variants in Singapore. *Vaccine*, *13*, 699-702.

OTT, J.J., STEVENS, G.A. & WIERSMA, S.T. 2012. The risk of perinatal hepatitis B virus transmission: hepatitis B e antigen (HBeAg) prevalence estimates for all world regions. *BMC Infect Dis*, 12, 131.

OU, J. H., LAUB, O., & RUTTER, W. J. 1986. Hepatitis B virus gene function: the precore region targets the core antigen to cellular membranes and causes the secretion of the e antigen. *Proceedings of the National Academy of Sciences of the United States of America*, *83*, 1578-1582.

PANIGRAHI, R., BISWAS, A., DE, B., CHAKRABATI, S., & CHALKRAVARTY, R. 2013. Characterization of antiviral resistance mutations among the Eastern Indian Hepatitis B virus infected population. *Virol J,* 10, 56.

PATEL, N., WHITE, S. J., THOMPSON, R. F., BINGHAM, R., WEIß, E. U., MASKELL, D. P., ZLOTNICK, A., DYKEMAN, E., TUMA, R., TWAROCK, R., RANSON, N. A., & STOCKLEY, P. G. 2017. HBV RNA pre-genome encodes specific motifs that mediate interactions with the viral core protein that promote nucleocapsid assembly. *Nature microbiology*, *2*, 17098.

PHUNG, T. B., ALESTIG, E., NGUYEN, T. L., HANNOUN, C., & LINDH, M. 2010. Genotype X/C recombinant (putative genotype I) of hepatitis B virus is rare in Hanoi, Vietnam--genotypes B4 and C1 predominate. *Journal of medical virology*, *82*, 1327-1333.

POH, W.T., XIA, E., CHIN-INMANU, K., WONG, L.P., CHENG, A.Y., MALASIT, P., SURIYAPHOL, P., TEO, Y.Y. & ONG, R.T. 2013. Viral quasispecies inference from 454 pyrosequencing. *BMC Bioinformatics*, 14, 355.

POURKARIM, M.R., AMINI-BAVIL-OLYAEE, S., KURBANOV, F., VAN RANST, M. & TACKE, F. 2014. Molecular identification of hepatitis B virus genotypes/subgenotypes: revised classification hurdles and updated resolutions. *World J Gastroenterol*, 20, 7152-7168.

QIAN, F., QIN, J., LI, D., MA, Z., ZHANG, H., JIN, F. & WANG, W. 2016. Monitoring of genotypic resistance profile in chronic hepatitis B patients receiving nucleos(t)ide analogues in Huzhou, China. *J. of Infect Dev Ctries,* 10, 996-1002.

RABE, B., GLEBE, D., & KANN, M. 2006. Lipid-mediated introduction of hepatitis B virus capsids into nonsusceptible cells allows highly efficient replication and facilitates the study of early infection events. J. Virol, 80, 5465-5473.

RABE, B., VLACHOU, A., PANTE, N., HELENIUS. A., & KANN, M. 2003. Nuclear import of hepatitis B virus capsids and release of the viral genome. Proc. Natl. Acad. Sci. USA, 100, 9849-9854.

REVILL, P., YUEN, L., WALSH, R., PERRAULT, M., LOCARNINI, S. & KRAMVIS, A. 2010. Bioinformatic analysis of the hepadnavirus e-antigen and its precursor identifies remarkable sequence conservation in all orthohepadnaviruses. *J Med Virol*, 82, 104-115.

RICHMAN, D.D. 2000. The impact of drug resistance on the effectiveness of chemotherapy for chronic hepatitis B. *Hepatology*, 32, 866-867.

RODRIGUEZ-FRIAS F, BUTI M, TABERNERO D, HOMS M. 2013. Quasispecies structure, cornerstone of hepatitis B virus infection: mass sequencing approach. *World J Gastroenterol*, 19: 6995-7023.

RONAGHI, M., UHLÉN, M., & NYRÉN, P. 1998. A sequencing method based on real-time pyrophosphate. *Science (New York, N.Y.)*, *281*, 363-365.

ROTHBERG, J.M., HINZ, W., REARICK, T.M., SCHULTZ, J., MILESKI, W., DAVEY, M., LEAMON, J.H., JOHNSON, K., MILGREW, M.J., EDWARDS, M., HOON, J., SIMONS, J.F., MARRAN, D., MYERS, J.W., DAVIDSON, J.F., BRANTING, A., NOBILE, J.R., PUC, B.P., LIGHT, D., CLARK, T.A., HUBER, M., BRANCIFORTE, J.T., STONER, I.B., CAWLEY, S.E., LYONS, M., FU, Y., HOMER, N., SEDOVA, M., MIAO, X., REED, B., SABINA, J., FEIERSTEIN, E., SCHORN, M., ALANJARY, M., DIMALANTA, E., DRESSMAN, D., KASINSKAS, R., SOKOLSKY, T., FIDANZA, J.A., NAMSARAEV, E., MCKERNAN, K.J., WILLIAMS, A., ROTH, G.T. & BUSTILLO, J. 2011. An integrated semiconductor device enabling nonoptical genome sequencing. *Nature*, 475, 348-352.

SAADAH, O. I., SINDI, H. H., BIN-TALIB, Y., AL-HARTHI, S., & AL-MUGHALES, J. 2012. Entecavir treatment of children 2-16 years of age with chronic hepatitis B infection. *Arab journal of gastroenterology: the official publication of the Pan-Arab Association of Gastroenterology*, *13*, 41-44.

SAID Z. N. 2011. An overview of occult hepatitis B virus infection. *World journal of gastroenterology*, *17*, 1927-1938.

SALDANHA, J., GERLICH, W., LELIE, N., DAWSON, P., HEERMANN, K., HEATH, A., & WHO COLLABORATIVE STUDY GROUP. 2001. An international collaborative study to establish a World Health Organization international standard for hepatitis B virus DNA nucleic acid amplification techniques. *Vox sanguinis*, *80*, 63-71.

SALLAM, T. A., & WILLIAM TONG, C. Y. 2004. African links and hepatitis B virus genotypes in the Republic of Yemen. *Journal of medical virology*, *73*, 23-28.

SALPINI, R., SVICHER, V., CENTO, V., GORI, C., BERTOLI, A., SCOPELLITI, F., MICHELI, V., CAPPIELLO, T., SPANÒ, A., RIZZARDINI, G., DE SANCTIS, G. M., SARRECCHIA, C., ANGELICO, M., & PERNO, C. F. 2011. Characterization of drug-resistance mutations in HBV D-genotype chronically infected patients, naïve to antiviral drugs. *Antiviral research*, *92*, 382-385.

SANGER, F., AIR, G.M., BARRELL, B.G., BROWN, N.L., COULSON, A.R., FIDDES, C.A., HUTCHISON, C.A., SLOCOMBE, P.M., & SMITH, M. 1977B. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature*, 265, 687-695.

SANGER, F., NICKLEN, S., & COULSON, A.R. 1977A. DNA sequencing with chainterminating inhibitors. *Proc Natl Acad Sci USA*, 74, 5463-5467.

SARIN, S. K., KUMAR, M., ESLAM, M., GEORGE, J., AL MAHTAB, M., AKBAR, S., JIA, J., TIAN, Q., AGGARWAL, R., MULJONO, D. H., OMATA, M., OOKA, Y., HAN, K. H., LEE, H. W., JAFRI, W., BUTT, A. S., CHONG, C. H., LIM, S. G., PWU, R. F., & CHEN, D. S. 2020. Liver diseases in the Asia-Pacific region: a Lancet Gastroenterology & Hepatology Commission. *The lancet. Gastroenterology & hepatology*, *5*, 167-228.

SCAGLIONI, P. P., MELEGARI, M., & WANDS, J. R. 1996. Recent advances in the molecular biology of hepatitis B virus. *Bailliere's clinical gastroenterology*, *10*, 207-225.

SCHADLER, S. & HILDT, E. 2009. HBV life cycle: entry and morphogenesis. *Viruses*, 1, 185-209.

SCHAEFER S. 2007. Hepatitis B virus taxonomy and hepatitis B virus genotypes. *World J Gastroenterol.* 13, 14-21.

SCHMITZ, A., SCHWARZ, A., FOSS, M., ZHOU, L., RABE, B., HOELLENRIEGEL, J., STOEBER, M., PANTÉ, N., & KANN, M. 2010. Nucleoporin 153 arrests the nuclear import of hepatitis B virus capsids in the nuclear basket. *PLoS pathogens*, *6*, e1000741.

SCHÖDEL, F., PETERSON, D., ZHENG, J., JONES, J. E., HUGHES, J. L., & MILICH, D. R. 1993. Structure of hepatitis B virus core and e-antigen. A single precore amino acid prevents nucleocapsid assembly. *The Journal of biological chemistry*, *268*, 1332-1337.

SCHULZE, A., MILLS, K., WEISS, T. S., & URBAN, S. 2012. Hepatocyte polarization is essential for the productive entry of the hepatitis B virus. *Hepatology (Baltimore, Md.)*, *55*, 373-383.

SCHWEITZER, A., HORN, J., MIKOLAJCZYK, R.T., KRAUSE, G. & OTT, J.J. 2015. Estimations of worldwide prevalence of chronic hepatitis B virus infection: A systematic review of data published between 1965 and 2013. *Lancet,* 386, 1546-1555.

SEEFF, L. B., BEEBE, G. W., HOOFNAGLE, J. H., NORMAN, J. E., BUSKELL-BALES, Z., WAGGONER, J. G., KAPLOWITZ, N., KOFF, R. S., PETRINI, J. L., JR, & SCHIFF, E. R. 1987. A serologic follow-up of the 1942 epidemic of post-vaccination hepatitis in the United States Army. *The New England journal of medicine*, *316*, 965-970.

SEEGER, C. & MASON, W.S. 2000. Hepatitis B virus biology. *Microbiol Mol Biol Rev*, 64, 51-68.

SELLS, MA., CHEN, ML., & ACS, G. 1987. Production of hepatitis B virus particles in Hep G2 cells transfected with cloned hepatitis B virus DNA. Proc Natl Acad Sci U S A, 84, 1005-1009.

SEN, N., CAO, F., & TAVIS, E. 2004. Translation of duck hepatitis B virus reverse transcriptase by ribosomal shunting. J. Virol, 78, 11751-11757.

SHEPARD, C. W., SIMARD, E. P., FINELLI, L., FIORE, A. E., & BELL, B. P. 2006. Hepatitis B virus infection: epidemiology and vaccination. *Epidemiologic reviews*, *28*, 112-125.

SHOUVAL, D. 2003. Hepatitis B vaccines. *J Hepatol*, 39, 70-76.

SHOUVAL, D., ROGGENDORF, H. & ROGGENDORF, M. 2015. Enhanced immune response to hepatitis B vaccination through immunization with a Pre-S1/Pre-S2/S Vaccine. *Med Microbiol Immunol*, 204, 57-68.

SLATKO, B. E., GARDNER, A. F., & AUSUBEL, F. M. 2018. Overview of Next-Generation Sequencing Technologies. *Current protocols in molecular biology*, *122*, e59.

SONDERUP, M. W., DUSHEIKO, G., DESALEGN, H., LEMOINE, M., TZEUTON, C., TAYLOR-ROBINSON, S. D., & SPEARMAN, C. W. 2020. Hepatitis B in sub-Saharan Africa-How many patients need therapy?. *Journal of viral hepatitis*, *27*, 560-567.

SONG, J.E. & KIM, D.Y. 2016. Diagnosis of hepatitis B. *Ann Transl Med*, 4, 338.

SOZZI, V., WALSH, R., LITTLEJOHN, M., COLLEDGE, D., JACKSON, K., WARNER, N., YUEN, L., LOCARNINI, S.A. AND REVILL, P.A. 2016. In Vitro Studies Show that Sequence Variability Contributes to Marked Variation in Hepatitis B Virus Replication, Protein Expression, and Function Observed across Genotypes. *J Virol* 90, 10054-10064.

SPEARMAN, C. W., AFIHENE, M., ALLY, R., APICA, B., AWUKU, Y., CUNHA, L., DUSHEIKO, G., GOGELA, N., KASSIANIDES, C., KEW, M., LAM, P., LESI, O., LOHOUÈS-KOUACOU, M. J., MBAYE, P. S., MUSABEYEZU, E., MUSAU, B., OJO, O., RWEGASHA, J., SCHOLZ, B., & SHEWAYE, A. GASTROENTEROLOGY AND HEPATOLOGY ASSOCIATION OF SUB-SAHARAN AFRICA (GHASSA). 2017. Hepatitis B in sub-Saharan Africa: strategies to achieve the 2030 elimination targets. *The lancet. Gastroenterology & hepatology*, *2*, 900-909.

SPRINZL, M. F., OBERWINKLER, H., SCHALLER, H., & PROTZER, U. 2001. Transfer of hepatitis B virus genome by adenovirus vectors into cultured cells and mice: crossing the species barrier. *Journal of virology*, *75*, 5108-5118.

STEWARD, M. W., PARTIDOS, C. D., D'MELLO, F., & HOWARD, C. R. 1993. Specificity of antibodies reactive with hepatitis B surface antigen following immunization with synthetic peptides. *Vaccine*, 11, 1405-1414.

STRASFELD, L., & CHOU, S. 2010. Antiviral Drug Resistance: Mechanisms and Clinical Implications. Infectious Disease Clinics of North America, 24, 413-437.

STUYVER, L. J., LOCARNINI, S. A., LOK, A., RICHMAN, D. D., CARMAN, W. F., DIENSTAG, J. L., & SCHINAZI, R. F. 2001. Nomenclature for antiviral-resistant human hepatitis B virus mutations in the polymerase region. *Hepatology (Baltimore, Md.)*, *33*, 751-757.

SUGAUCHI, F., OHNO, T., ORITO, E., SAKUGAWA, H., ICHIDA, T., KOMATSU, M., KURAMITSU, T., UEDA, R., MIYAKAWA, Y., & MIZOKAMI, M. 2003. Influence of hepatitis B virus genotypes on the development of preS deletions and advanced liver disease. *J Med* Virol, 70, 537-44.

SUNBUL, M. 2014. Hepatitis B virus genotypes: Global distribution and clinical importance. *World J Gastroenterol,* 20, 5427-5434.

TACKE, F., GEHRKE, C., LUEDDE, T., HEIM, A., MANNS, M.P. & TRAUTWEIN C. 2004. Basal core promoter and precore mutations in the hepatitis B virus genome enhance replication efficacy of Lamivudine-resistant mutants. *J Virol*, 78, 8524-8535.

TAKAHASHI, M., NISHIZAWA, T., GOTANDA, Y., TSUDA, F., KOMATSU, F., KAWABATA, T., HASEGAWA, K., ALTANKHUU, M., CHIMEDREGZEN, U., NARANTUYA, L., HOSHINO, H., HINO, K., KAGAWA, Y. AND OKAMOTO, H. 2004. High prevalence of antibodies to hepatitis A and E viruses and viremia of hepatitis B, C, and D viruses among apparently healthy populations in Mongolia. *Clin Diagn Lab Immunol*, 11, 392-398.

TAN, A., KOH, S., & BERTOLETTI, A. 2015. Immune Response in Hepatitis B Virus Infection. *Cold Spring Harbor perspectives in medicine*, *5*, a021428.

TAN, G., SONG, H., XU, F., & CHENG, G. 2018. When Hepatitis B Virus Meets Interferons. *Frontiers in microbiology*, *9*, 1611.

TANWAR, S., & DUSHEIKO, G. 2012. Is there any value to hepatitis B virus genotype analysis?. *Current gastroenterology reports*, *14*, 37-46.

TAPPER, E. B., & ASRANI, S. K. 2020. The COVID-19 pandemic will have a long-lasting impact on the quality of cirrhosis care. *Journal of hepatology*, *73*(2), 441-445.

TARAMASSO, L., CALIGIURI, P., DI BIAGIO, A., BRUZZONE, B., ROSSO, R., ICARDI, G., & VISCOLI, C. 2011. Lamivudine resistance mutations in European patients with hepatitis B and patients co-infected with HIV and hepatitis B. *Journal of medical virology*, *83*, 1905-1908.

TATEMATSU, K., TANAKA, Y., KURBANOV, F., SUGAUCHI, F., MANO, S., MAESHIRO, T., NAKAYOSHI, T., WAKUTA, M., MIYAKAWA, Y., & MIZOKAMI, M. 2009. A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. *Journal of virology*, *83*, 10538-10547.

TAVIS, J. E., & LOMONOSOVA, E. 2015. The hepatitis B virus ribonuclease H as a drug target. *Antiviral research*, *118*, 132-138.

TENG, C. F., LI, T. C., HUANG, H. Y., LIN, J. H., CHEN, W. S., SHYU, W. C., WU, H. C., PENG, C. Y., SU, I. J., & JENG, L. B. 2020. Next-Generation Sequencing-Based Quantitative Detection of Hepatitis B Virus Pre-S Mutants in Plasma Predicts Hepatocellular Carcinoma Recurrence. *Viruses*, *12*, 796.

TENNEY, D. J., ROSE, R. E., BALDICK, C. J., LEVINE, S. M., POKORNOWSKI, K. A., WALSH, A. W., FANG, J., YU, C. F., ZHANG, S., MAZZUCCO, C. E., EGGERS, B., HSU, M., PLYM, M. J., POUNDSTONE, P., YANG, J., & COLONNO, R. J. 2007. Two-year assessment of entecavir resistance in Lamivudine-refractory hepatitis B virus patients reveals different clinical outcomes depending on the resistance substitutions present. *Antimicrobial agents and chemotherapy*, *51*, 902-911.

TERRAULT, N. A., BZOWEJ, N. H., CHANG, K. M., HWANG, J. P., JONAS, M. M. & MURAD, M. H. 2016. AASLD guidelines for treatment of chronic hepatitis B. *Hepatology (Baltimore, Md.),* 63, 261-283.

TERRAULT, N. A., LOK, A., MCMAHON, B. J., CHANG, K. M., HWANG, J. P., JONAS, M. M., BROWN, R. S., JR, BZOWEJ, N. H., & WONG, J. B. 2018. Update on Prevention, Diagnosis, and Treatment of Chronic Hepatitis B: AASLD 2018 Hepatitis B Guidance. *Clinical liver disease*, *12*, 33-34.

THOMAS, E., & LIANG, T. J. 2016. Experimental models of hepatitis B and C - new insights and progress. *Nature reviews. Gastroenterology & hepatology*, *13*, 362-374.

THOMAS, T., GILBERT, J., & MEYER, F. 2012. Metagenomics - a guide from sampling to data analysis. *Microbial informatics and experimentation*, *2*, 3.

THOMSON, E., IP, C. L., BADHAN, A., CHRISTIANSEN, M. T., ADAMSON, W., ANSARI, M. A., BIBBY, D., BREUER, J., BROWN, A., BOWDEN, R., BRYANT, J., BONSALL, D., DA SILVA FILIPE, A., HINDS, C., HUDSON, E., KLENERMAN, P., LYTHGOW, K., MBISA, J. L., MCLAUCHLAN, J., MYERS, R., PIAZZA, P., ROY, S., TREBES, A., SREENU, V., WITTEVELDT., BARNES, E., & SIMMONDS, P. 2016. Comparison of Next-Generation Sequencing Technologies for Comprehensive Assessment of Full-Length Hepatitis C Viral Genomes. *Journal of clinical microbiology*, *54*, 2470-2484.

TSETSARKIN, K. A., VANLANDINGHAM, D. L., MCGEE, C. E., & HIGGS, S. 2007. A single mutation in chikungunya virus affects vector specificity and epidemic potential. *PLoS pathogens*, *3*, e201.

TSUKUDA, S., & WATASHI, K. 2020. Hepatitis B virus biology and life cycle. *Antiviral research*, *182*, 104925.

TUJIOS, S. R., & LEE, W. M. 2013. Update in the management of chronic hepatitis B. *Current opinion in gastroenterology*, *29*, 250-256.

UMAR, M., HAMAMA-TUL-BUSHRA, UMAR, S., & KHAN, H. A. 2013. HBV perinatal transmission. *International journal of hepatology*, *2013*, 875791.

VELKOV, S., OTT, J. J., PROTZER, U., & MICHLER, T. 2018. The Global Hepatitis B Virus Genotype Distribution Approximated from Available Genotyping Data. *Genes*, *9*, 495.

VENKATAKRISHNAN, B. & ZLOTNICK, A. 2016. The Structural Biology of Hepatitis B Virus: Form and Function. *Annu Rev Virol*., 3, 429-451.

VENTER, J. C., ADAMS, M. D., MYERS, E. W., LI, P. W., MURAL, R. J., SUTTON, G. G., SMITH, H. O., YANDELL, M., EVANS, C. A., HOLT, R. A., GOCAYNE, J. D., AMANATIDES, P., BALLEW, R. M., HUSON, D. H., WORTMAN, J. R., ZHANG, Q., KODIRA, C. D., ZHENG, X. H., CHEN, L., SKUPSKI, M., & ZHU, X. 2001. The sequence of the human genome. *Science*, 291, 1304-1351.

VILLAMIL, F.G. & CAIRO, F.M. 2013, Hepatitis B virus: Prevention of recurrent infection. Clinical Liver Disease, 2: 169-172.

VOELKERDING, K.V., DAMES, S.A. & DURTSCHI, J.D. 2009. Next-generation sequencing: from basic research to diagnostics. *Clin Chem*, 55, 641-658.

WAGNER, J., YUEN, L., LITTLEJOHN, M., SOZZI, V., JACKSON, K., SURI, V., TAN, S., FEIERBACH, B., GAGGAR, A., MARCELLIN, P., BUTI FERRET, M., JANSSEN, H.L., GANE, E., CHAN, H.L., COLLEDGE, D., ROSENBERG, G., BAYLISS, J., HOWDEN, B.P., LOCARNINI, S.A., WONG, D., THOMPSON, A.T. AND REVILL, P.A. 2021. Analysis of Hepatitis B Virus Haplotype Diversity Detects Striking Sequence Conservation Across Genotypes and Chronic Disease Phase. Hepatology, 73: 1652-1670

WAI, C. T., FONTANA, R. J., POLSON, J., HUSSAIN, M., SHAKIL, A. O., HAN, S. H., DAVERN, T. J., LEE, W. M., LOK, A. S., & US ACUTE LIVER FAILURE STUDY GROUP. 2005. Clinical outcome and virological characteristics of hepatitis B-related acute liver failure in the United States. *Journal of viral hepatitis*, *12*, 192-198.

WALSH, R. & LOCARNINI, S. 2012. Hepatitis B precore protein: pathogenic potential and therapeutic promise. *Yonsei Med J*, 53, 875-885.

WALTER, E., KEIST, R., NIEDERÖST, B., PULT, I., & BLUM, H. E. 1996. Hepatitis B virus infection of tupaia hepatocytes in vitro and in vivo. *Hepatology (Baltimore, Md.)*, *24*, 1-5.

WANG, H.C., HUANG, W., LAI, M.D. & SU, I.J. 2006. Hepatitis B virus pre-S mutants, endoplasmic reticulum stress and hepatocarcinogenesis. *Cancer Sci*, 97, 683-688.

WANG, L., HAN, F., DUAN, H., JI, F., YAN, X., FAN, Y., & WANG, K. 2017. Hepatitis B virus pre-existing drug resistant mutation is related to the genotype and disease progression. *Journal of infection in developing countries*, *11*, 727-732. YOU, H & JIA, J. 2013. Telbivudine treatment in chronic hepatitis B: experience from China. J Viral Hepat, 20 suppl 1, 3-8.

WANG, X., XU, L., CHEN, Y., LIU, A., WANG, L., XU, P., LIU, Y., LI, L., & MENG, F. 2017. Integrating nested PCR with high-throughput sequencing to characterize mutations of HBV genome in low viral load samples. *Medicine*, *96*, e7588.

WANG, Y. X., LUO, C., ZHAO, D., BECK, J., & NASSAL, M. 2012. Extensive mutagenesis of the conserved box E motif in duck hepatitis B virus P protein reveals multiple functions in replication and a common structure with the primer grip in HIV-1 reverse transcriptase. *Journal of virology*, *86*, 6394-6407.

WARD, H., TANG, L., POONIA, B., & KOTTILIL, S. 2016. Treatment of hepatitis B virus: an update. *Future microbiology*, *11*, 15811597.

WARNER, N. & LOCARNINI, S. 2014. Mechanisms of hepatitis B virus resistance development. *Intervirology*, 57, 218-224.

WASITYASTUTI, W., YANO, Y., WIDASARI, D. I., YAMANI, L. N., RATNASARI, N., HERIYANTO, D. S., OKADA, R., TANAHASHI, T., MURAKAMI, Y., AZUMA, T., & HAYASHI, Y. 2016. Different Variants in Reverse Transcriptase Domain Determined by Ultra-deep Sequencing in Treatment-naïve and Treated Indonesian Patients Infected with Hepatitis B Virus. *The Kobe journal of medical sciences*, *62*, E1-E8.

WATANABE, K., TAKAHASHI, T., TAKAHASHI, S., OKOSHI, S., ICHIDA, T., & AOYAGI, Y. 2005. Comparative study of genotype B and C hepatitis B virus-induced chronic hepatitis in relation to the basic core promoter and precore mutations. *Journal of gastroenterology and hepatology*, *20*, 441-449.

WATANABE, T., SORENSEN, E. M., NAITO, A., SCHOTT, M., KIM, S., & AHLQUIST, P. 2007. Involvement of host cellular multivesicular body functions in hepatitis B virus budding. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 10205-10210.

WEBER B. 2005. Genetic variability of the S gene of hepatitis B virus: clinical and diagnostic impact. *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology*, *32*, 102-112.

WEDEMEYER. H., SCHONEWEIS, K., BOGOMOLOV,P., VORONKOVA, N., CHULANOV, V., STEPANOVA, T., BREMER, B., LEHMANN, P., RAUPACH., R., ALLWEISS, L., DANDIR, M., CIESEK, S., DIITMER, U., HAEFELI, W., ALEXANDROV, A., & URBAN , S. 2018. Interim results of a multicenter, open-label phase 2 clinical trial (MYR203) to assess safety and efficacy of Myrcludex B in combination with PEG-IFNa in patients with chronic HBV/HDV co-infection. *The Liver Meeting*. San Francisco, CA, November 9-13.

WOO, A., KWOK, R., & AHMED, T. 2017. Alpha-interferon treatment in hepatitis B. *Annals of translational medicine*, *5*, 159.

WOODDELL, C. I., YUEN, M. F., CHAN, H. L., GISH, R. G., LOCARNINI, S. A., CHAVEZ, D., FERRARI, C., GIVEN, B. D., HAMILTON, J., KANNER, S. B., LAI, C. L., LAU, J., SCHLUEP, T., XU, Z., LANFORD, R. E., & LEWIS, D. L. 2017. RNAi-based treatment of chronically infected patients and chimpanzees reveals that integrated hepatitis B virus DNA is a source of HBsAg. *Science translational medicine*, *9*, eaan0241.

WORLD HEALTH ORGANISATION. 2007. Hepatitis B. Factsheet [Online]. Available: http://www.who.int/mediacentre/factsheets/fs204/en/ [Accessed 10/06/2021].

WORLD HEALTH ORGANISATION. 2015. *Guidelines for the Prevention Care and Treatment of Persons with Chronic Hepatitis B Infection*. [online] Available at: https://www.who.int/hiv/pub/hepatitis/hepatitis-b-guidelines/en/[Accessed 10/06/2021].

WORLD HEALTH ORGANISATION. 2017. Guidelines on hepatitis B and C testing. [online]. Available at: https://www.who.int/hepatitis/publications/guidelines-hepatitis-c-b-testing/en/ [Accessed 10/06/2021].

WORLD HEALTH ORGANISATION. 2018.Immunization, Vaccine and Biologicals: Hepatitis B. [online]. Available at: http://www.who.int/immunization/diseases/hepatitisB/en/ [Accessed 10/06/2021].

WOSE, K., CONSTANCE, N., NIMISHA, B., & ANNA, K. 2020. In Vitro Systems for Studying Different Genotypes/Sub-Genotypes of Hepatitis B Virus: Strengths and Limitations. *Viruses*, 12, 353.

WU, I.C., LIU, W.C. & CHANG, T.T. 2018. Applications of next-generation sequencing analysis for the detection of hepatocellular carcinoma-associated hepatitis B virus mutations. *J Biomed Sci*, 25, 51.

WYLIE, K.M., MIHINDUKULSASURIYA, K.A., SODERGREN, E., WEINSTOCK, G.M. & STORCH, G.A. 2012. Sequence analysis of the human virome in febrile and afebrile children. *PLoS One*, 7, e27735.

XU, B., LIU, L., HUANG, X., MA, H., ZHANG, Y., DU, Y., WANG, P., TANG, X., WANG, H., KANG, K., ZHANG, S., ZHAO, G., WU, W., YANG, Y., CHEN, H., MU, F. AND CHEN, W. 2011. Metagenomic analysis of fever, thrombocytopenia and leukopenia syndrome (FTLS) in Henan Province, China: discovery of a new bunyavirus. *PLoS Pathog*, 7, e1002369.

XU, R., HU, P., LI, Y., TIAN, A., LI, J., & ZHU, C. 2021. Advances in HBV infection and replication systems *in vitro*. *Virol J,* 18, 105.

YAMADA, M., OEDA, A., JUNG, J., IIJIMA, M., YOSHIMOTO, N., NIIMI, T., JEONG, S.Y., CHOI, E. K., TANIZAWA, K., & KURODA, S. 2012. Hepatitis B virus envelope L protein-derived bio nanocapsules: mechanisms of cellular attachment and entry into human hepatic cells. J. Control Release, 160, 322-329.

YAMANI, L. N., YANO, Y., UTSUMI, T., WASITYASTUTI, W., RINONCE, H. T., WIDASARI, D. I., JUNIASTUTI, LUSIDA, M. I., SOETJIPTO, & HAYASHI, Y. 2017. Profile of Mutations in the Reverse Transcriptase and Overlapping Surface Genes of Hepatitis B Virus (HBV) in Treatment-Naïve Indonesian HBV Carriers. *Japanese journal of infectious diseases*, *70*, 647-655.

YAMAURA, T., TANAKA, E., MATSUMOTO, A., ROKUHARA, A., ORII, K., YOSHIZAWA, K., MIYAKAWA, Y. & KIYOSAWA, K. 2003. A case-control study for early prediction of hepatitis B e antigen seroconversion by hepatitis B virus DNA levels and mutations in the precore region and core promoter. *J Med Virol*, 70, 545-552.

YAN, H., ZHONG, G., XU, G., HE, W., JING, Z., GAO, Z., HUANG, Y., QI, Y., PENG, B., WANG, H., FU, L., SONG, M., CHEN, P., GAO, W., REN, B., SUN, Y., CAI, T., FENG, X., SUI, J., & LI, W. 2012. Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. *eLife*, *1*, e00049.

YANG, C.Y., KUO, T. H. & TING, L.P. 2006. Human hepatitis B viral e antigen interacts with cellular interleukin-1 receptor accessory protein and triggers interleukin-1 response. *J Biol Chem*, 281, 34525-34536.

YANG, H.C. & KAO, J.H. 2014. Persistence of hepatitis B virus covalently closed circular DNA in hepatocytes: molecular mechanisms and clinical significance. *Emerg Microbes Infect*, 3, e64.

YEUNG, P., WONG, D.K., LAI, C.L., FUNG, J., SETO, W.K. & YUEN, M.F. 2011. Association of hepatitis B virus pre-S deletions with the development of hepatocellular carcinoma in chronic hepatitis B. *J Infect Dis*, 203, 646-654.

YIN, Y., HE, K., WU, B., XU, M., DU, L., LIU, W., LIAO, P., LIU, Y., & HE, M. 2019. A systematic genotype and subgenotype re-ranking of hepatitis *B virus* under a novel classification standard. *Heliyon*, *5*10, e02556.

YOKOSUKA, O. & ARAI, M. 2006. Molecular biology of hepatitis B virus: effect of nucleotide substitutions on the clinical features of chronic hepatitis B. *Med Mol Morphol*, 39, 113-120.

ZAAIJER, H. L., BOOT, H. J., VAN SWIETEN, P., KOPPELMAN, M. H., & CUYPERS, H. T. 2011. HBsAg-negative mono-infection with hepatitis B virus genotype G. *Journal of viral hepatitis*, *18*, 815-819.

ZANETTI, A.R., VAN DAMME, P. & SHOUVAL, D. 2008. The global impact of vaccination against hepatitis B: a historical overview. *Vaccine*, 26, 6266-6273.

ZHANG, H. W., YIN, J. H., LI, Y. T., LI, C. Z., REN, H., GU, C. Y., WU, H. Y., LIANG, X. S., ZHANG, P., ZHAO, J. F., TAN, X. J., LU, W., SCHAEFER, S., & CAO, G. W. 2008. Risk factors for

acute hepatitis B and its progression to chronic hepatitis in Shanghai, China. *Gut*, *57*, 1713-1720.

ZHANG, J., CHIODINI, R., BADR, A. & ZHANG, G. 2011. The impact of next generation sequencing on genomics. J Genet Genomics, 38, 95-109.

ZHANG, Q., LIAO, Y., CAI, B., LI, Y., LI, L., ZHANG, J., AN, Y., & WANG, L. 2015. Incidence of natural resistance mutations in naïve chronic hepatitis B patients: a systematic review and meta-analysis. *Journal of gastroenterology and hepatology*, *30*, 252-261.

ZHANG, Z.H., WU, C.C., CHEN, X.W., LI, X., LI, J., & LU, M.J. 2016. Genetic variation of hepatitis B virus and its significance for pathogenesis. *World Journal of Gastroenterology*, 22, 126-44.

ZIAEE, M., JAVANMARD, D., SHARIFZADEH, G., HASAN NAMAEI, M., & AZARKAR, G. 2016. Genotyping and Mutation Pattern in the Overlapping MHR Region of HBV Isolates in Southern Khorasan, Eastern Iran. *Hepatitis monthly*, *16*, e37806.

ZLOTNICK, A., VENKATAKRISHNAN, B., TAN, Z., LEWELLYN, E., TURNER, W., & FRANCIS, S. 2015. Core protein: A pleiotropic keystone in the HBV lifecycle. *Antiviral research*, *121*, 82-93.

ZOULIM, F. & LOCARNINI, S. 2009. Hepatitis B virus resistance to nucleos(t)ide analogues. *Gastroenterology*, 137, 1593-1608.

ZOULIM, F., & DURANTEL, D. 2015. Antiviral therapies and prospects for a cure of chronic hepatitis B. *Cold Spring Harbor perspectives in medicine*, *5*, a021501.

ZOULIM, F., LEBOSSE, F., & LEVRERO, M. 2016. Current treatments for chronic hepatitis B virus infections. *Current opinion in virology*, *18*, 109-116.

ZUCKERMAN, J.N. 2007. Review: hepatitis B immune globulin for prevention of hepatitis B infection. *J Med Virol*, 79, 919-921.