The Interaction of APOBEC3 proteins with Hepatitis B virus and their inhibition of HBV genome replication

Thesis Submitted for the Degree of Doctor of Philosophy

by

Ibrahim Mohamed Hasan Elbusifi

MBBS, MSc (Gastroenterology)

Imperial College London

Department of Infectious Diseases

St Mary's Campus

Declaration of Originality

I, Ibrahim Mohamed Hasan Elbusifi, hereby certify that this thesis, which is submitted for assessment as part of a PhD research project, has been written by me and is a record of work carried out by me unless otherwise stated.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution 4.0 International Licence (CC BY). Under this licence, you may copy and redistribute the material in any medium or format for both commercial and non-commercial purposes. You may also create and distribute modified versions of the work. This on the condition that you credit the author. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

Chronic HBV infection leads to an estimated 820,000 deaths yearly. It has no cure and new treatments need to be developed. APOBEC3 proteins can restrict the replication of viruses (e.g., HIV) and retrotransposons. This study aimed to elucidate the role of APOBEC3 proteins in Hepatitis B virus replication. It was observed that APOBEC3G proteins colocalise with HBV polymerase proteins in areas of the cytoplasm using immunofluorescence. It was shown that USF1 could upregulate APOBEC3 gene expression and that USF1 enhanced HBV replication. qPCR used to measure levels of APOBEC3 mutated viruses that do not survive the mutations in HepG2.2.15 cells, these results showed that APOBEC3G and APOBEC3DE could inhibit extracellular virion-associated HBV DNA levels by 48% and 16% respectively and could inhibit cytoplasmic HBV DNA by 65% and 39% respectively. The numbers and percentages of G to A mutations in the HBV X region detected by Phusion U polymerase-based 3D-PCR (modified 3D-PCR) were lower than those detected by Taq polymerase-based 3D-PCR (3D-PCR) in empty vector and APOBEC3G vector transfected HepG2.2.15 cells. However, numbers and percentages of G to A mutations in the HBV X region detected by modified 3D-PCR were higher than those detected by 3D-PCR in APOBEC3DE transfected HepG2.2.15 cells. Primer-ID NGS results demonstrated a higher number of G to A and C to U mutations in the HBV S region compared to the HBV X region. Primer-ID NGS also showed that these mutations already existed in the HBV pre-core, X, and S regions in HepG2.2.15 cells that were not transfected with APOBEC3 plasmids. These novel approaches may help to measure more accurately the effects of APOBEC3 proteins on the HBV virus in vitro and in vivo and could be important to evaluate the possible role of APOBEC3 proteins in treatments for HBV.

Publications and Presentations

- Effects of APOBEC3 family members on Hepatitis B virus replication. **Ibrahim Elbusifi**, Esther Bankole, Jonathan Hoare and Michael McGarvey. Microbiology Society Annual Conference (Virology Workshop: Gene expression and replication), Belfast, UK, April 2019.

- Effects of APOBEC3 proteins on Hepatitis B virus replication. **Ibrahim Elbusifi** and Michael McGarvey. PhD Away Day, Imperial College, London, UK, November 2018.

- Effects of APOBEC3 proteins on Hepatitis B virus replication. **Ibrahim Elbusifi** and Michael McGarvey, Young Scientist Day Symposium, Imperial College, London, UK, February 2016.

Dedication

This thesis is dedicated to my mother, my father, and my brothers in recognition of their continual love, encouragement, and support.

Acknowledgements

My most profound appreciation goes to my supervisor, Dr Michael J McGarvey, for his time, effort, and understanding in helping me succeed in my studies. I am deeply grateful for his support, guidance, and feedback throughout my PhD study. His vast wisdom and wealth of experience have inspired me throughout my studies. He is very kind, caring and helpful, and I will treasure knowing him. I would like to thank Dr Jonathan Hoare, for his help and advice. I want to thank the Libyan Government (Ministry of Higher Education and Scientific Research) for sponsoring me; without their generous financial support, this work couldn't happen. I'm also grateful to Friends, lab mates, colleagues, and staff. I feel deep gratitude for my parents and brothers, who formed part of my vision and taught me the good things that matter in life. I am glad to be one of them.

Table of Contents

Title *1*
Declaration of Originality *2*
Copyright Declaration *2*
Abstract *3*
Publications and Presentations *4*
Dedication *5*
Acknowledgements *6*
Table of Contents *7*
List of Figures *10*
List of Tables *13*
List of Abbreviations *15*
1 Chapter One: Introduction *20*
1.1 Hepatitis B Virus **20**
1.1.1 Overview (History) 20
1.1.2 Classification and Epidemiology 21
1.1.3 Diagnosis and Prevention 24
1.1.4 Treatment of HBV infection 25
1.1.5 Virion Morphology 28
1.1.6 HBV Genome structure 30
1.1.7 HBV life cycle and replication 33
1.2 APOBEC proteins are Host restriction Factors **39**
1.2.1 Interferon (IFN) 39
1.2.2 Apolipoprotein B mRNA editing enzyme catalytic polypeptide-like (APOBEC) 41
1.2.3 Effects of APOBEC3 proteins on HBV replication 44
1.2.4 Sequencing approach for the analysis of APOBEC3 proteins induced mutations in HBV DNA .. 49
1.2.5 Upregulation of APOBEC3 gene expression 49
1.3 Hypothesis & Objectives **50**
2 Chapter Two: Materials and Methods *51*
2.1 Materials **51**
2.1.1 Cell Culture 51
2.1.2 Transformation and Culturing of Competent *E. coli* cells 52
2.1.3 Plasmid Purification and Extraction 53
2.1.4 Plasmid Transfection 53
2.1.5 Immunofluorescence 53
2.1.6 Western Blot Reagents 54
2.1.7 Antibodies 54
2.1.8 Molecular Biology Reagents 55
2.1.9 IFN Antiviral Assay Reagents 56
2.1.10 Restriction Enzymes and Buffers 56

2.1.11 Oligonucleotides 57
2.1.12 Plasmids (Appendix A.1, A.2, A.3, A.4, and A.5): 60
2.1.13 Chemicals and Plasticware 61

2.2 Methods 62
2.2.1 Construction of APOBEC3DE-GFP and APOBEC3G-GFP Mammalian Expression Vector 62
2.2.2 Restriction Digestion of the APOBEC3 Coding ORF 62
2.2.3 Restriction Digestion and Dephosphorylation of the GFP-tagged Destination Vector 63
2.2.4 Agarose Gel Electrophoresis of Digested Plasmids 63
2.2.5 Agarose Gel Extraction 64
2.2.6 DNA Ligations 65
2.2.7 DNA Sequencing of Plasmid Inserts 65
2.2.8 Plasmid Purification and Extraction 66
2.2.9 Mammalian Cell Culture Techniques 70
2.2.10 Fugene HD Mediated plasmid Transfection 72
2.2.11 Detection of plasmid Expression by Immunofluorescence 72
2.2.12 Detection of plasmid Expression by Western Blotting 73
2.2.13 RNA Extraction 74
2.2.14 cDNA Reverse Transcription 75
2.2.15 Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR) 76
2.2.16 Extraction of cytoplasmic Core-associated HBV DNA 77
2.2.17 Extraction of supernatant virion-associated HBV DNA 78
2.2.18 Quantification of HBV DNA by Real-Time PCR 79
2.2.19 Differential DNA Denaturation Polymerase chain reaction (3D-PCR) 80
2.2.20 Next-Generation Sequencing (NGS) 83

3 Chapter Three: Results 87

3.1 Construction of APOBEC3 Mammalian Expression Vectors 88

3.2 The Intracellular localisation of APOBEC3 proteins and the colocalisation of HBV polymerase protein with APOBEC3G 96
3.2.1 The Intracellular localisation of APOBEC3 proteins 96
3.2.2 The colocalisation of HBV polymerase protein with APOBEC3G protein 99
3.2.3 The colocalisation of HBV polymerase protein with DDX3 102

3.3 APOBEC3 proteins inhibit HBV replication 104
3.3.1 APOBEC3G protein expression in HepG2 cells over different time-points 105
3.3.2 APOBEC3 proteins inhibit cytoplasmic HBV DNA 106
3.3.3 APOBEC3 proteins inhibit extracellular capsid-associated HBV DNA 109
3.3.4 APOBEC3 proteins synergistically inhibit cytoplasmic HBV DNA 112

3.4 APOBEC3 proteins edit the HBV genome 115
3.4.1 3D-PCR analysis of the frequency and extent of APOBEC3 proteins mutations in HBV 117
3.4.2 Next-Generation Sequencing approach to the analysis of the frequency and extent of APOBEC3 proteins mutations in HBV DNA 132

3.5 Human Upstream stimulatory factor 1(USF1) effect on APOBEC3 proteins expression levels and the effect on HBV replication. 138
3.5.1 Human Upstream stimulatory factor 1(USF1) upregulates APOBEC3DE and APOBEC3G proteins 138
3.5.2 Human Upstream stimulatory factor 1(USF1) enhance HBV replication 140

4 Chapter Four: Discussion 142

4.1 APOBEC3 proteins inhibit HBV replication 145

4.2 APOBEC3 proteins edit the HBV genome 148

4.3 Human Upstream stimulatory factor 1 (USF1) upregulate APOBEC3 family members and enhances HBV replication 156

4.4 Conclusions 157

4.5 Possible Future Work 158
5 References:..... 159
6 Appendix 174

List of Figures

Figure 1. The global prevalence of Chronic HBV Infection. (Centers for Disease Control and Prevention 2019).....23

Figure 2. The phases of chronic HBV infection and natural history and patient assessment are based on HBV and liver disease markers (European Association For The Study Of The Liver 2017). 25

Figure 3. HBV Virion Morphology. (Luangsay, Lyon University et al. 2013).....29

Figure 4. HBV Genome Structure. 32

Figure 5. HBV life cycle.....37

Figure 6. HBV replication.....38

Figure 7. APOBEC3 proteins' deamination of deoxycytidine to deoxyuridine. 42

Figure 8. Functions of DNA Editing by APOBEC3 proteins Cytidine deamination. 44

Figure 9. Structure of the Precision Shuttle subcloning procedure of A3DE and A3G. 89

Figure 10. APOBEC3-GFP plasmids were confirmed by sequencing. 91

Figure 11. Exogenous expression of APOBEC3DE and APOBEC3G proteins in HEK293T cells.....93

Figure 12. The western blot image shows the protein expression of Empty-GFP, APOBEC3A, APOBEC3DE and APOBEC3G proteins in HEK293T cells.....95

Figure 13. The subcellular localisation of APOBEC3 Proteins in HEK293T cells.....98

Figure 14. Localization and co-localization of HBV polymerase and APOBEC3G.....100

Figure 15. Localization and co-localization of HBV polymerase and APOBEC3G.....101

Figure 16. Localization and co-localization of HBV polymerase and DDX3.....103

Figure 17. The experimental approach to evaluate the effect of APOBEC3 proteins on HBV DNA replication. 104

Figure 18. APOBEC3G overexpression in HepG2 cells at different time points.....105

Figure 19. The effect of individual APOBEC3 proteins on the cytoplasmic HBV DNA levels 107
Figure 20. The effect of individual APOBEC3 proteins on the cytoplasmic HBV DNA levels (percentages of cytoplasmic HBV DNA inhibition). 108
Figure 21. The effect of individual APOBEC3 proteins on the extracellular virion-associated HBV DNA levels (extracellular virion-associated HBV DNA copy number).....110
Figure 22. The effect of individual APOBEC3 proteins on the extracellular virion-associated HBV DNA levels (percentages of extracellular virion-associated HBV DNA inhibition). 111
Figure 23. The synergistic effect of APOBEC3 proteins on the cytoplasmic HBV DNA levels (cytoplasmic HBV DNA copy number). 113
Figure 24. The synergistic effect of APOBEC3 proteins on the cytoplasmic HBV DNA levels (percentages of cytoplasmic HBV DNA inhibition). 114
Figure 25. The diagram illustrates the APOBEC3 proteins edition of GC-rich regions in the HBV genomic DNA. 116
Figure 26. Diagram explains how APOBEC3 proteins can edit single-strand DNA during reverse transcription. 116
Figure 27. The agarose gel image shows the 3D-PCR products of the X region of HBV extracted from the HepG2.2.15 cell line.....118
Figure 28. G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the empty vector transfected HepG2.2.15 cell line. 120
Figure 29. G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the APOBEC3G transfected HepG2.2.15 cell line.....121
Figure 30 G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the APOBEC3DE transfected HepG2.2.15 cell line. 122

Figure 31. G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the empty vector transfected HepG2.2.15 cells. 126
Figure 32. G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the APOBEC3G transfected HepG2.2.15 cells. 127
Figure 33. G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the APOBEC3DE transfected HepG2.2.15 cells. 128
Figure 34. Next-Generation Sequencing approach to analysing the frequency and extent of APOBEC3 protein mutations in HBV DNA in Hep2.2.15 cells. 132
Figure 35. The agarose gel image shows the NGS library products of the HBV PreCore, S, and X regions extracted from APOBEC3 transfected HepG2.2.15 cells. 133
Figure 36. Human Upstream stimulatory factor 1(USF1) overexpression in HepG2.2.15 cells. 139
Figure 37. Human Upstream stimulatory factor 1 (USF1) upregulates APOBEC3DE AND APOBEC3G in HepG2.2.15 cells..... 139
Figure 38. Human Upstream stimulatory factor 1(USF1) enhance HBV replication in HepG2.2.15 cells. 140
Figure 39. Human Upstream stimulatory factor 1(USF1) enhance HBV replication in HepG2.2.15 cells. 141

List of Tables

Table 1. The Global burden of hepatitis B infection. (World Health Organization 2022).....23

Table 2. Advantages and disadvantages of nucleos(t)ides analogues and PEG-IFN-alpha in treating Chronic Hepatitis B (Phillips, Jagatia *et al.* 2022). 27

Table 3. The Goal of HBV therapy (Lok, Zoulim *et al.* 2017, Phillips, Jagatia *et al.* 2022). .27

Table 4. The IFN-stimulated genes (ISGs) inhibitory effects on the HBV life cycle. 40

Table 5. The inhibition of cytoplasmic HBV DNA levels by individual APOBEC3 proteins. 109

Table 6. The inhibition of extracellular virion-associated HBV DNA levels by individual APOBEC3 proteins. 112

Table 7. The synergistic inhibition of cytoplasmic HBV DNA levels by APOBEC3 proteins. 114

Table 8. shows the length of the HBV (ayw) X sequence and the number of Adenine, Guanine, Cytosine and Thymine in the HBV (ayw) X sequence. 118

Table 9. Nucleotide Substitutions of Clones Amplified by Taq polymerase-based 3D-PCR in HepG2.2.15 cells. 124

Table 10. Nucleotide Substitutions of Clones Amplified by Phusion U polymerase-based 3D-PCR in HepG2.2.15 cells. 130

Table 11. a summarised comparison of the number and percentages of G to A mutations in the X region of the extracted cytoplasmic HBV genomic DNA. 131

Table 12. Nucleotide substitutions of the HBV X region sequence amplified by Next Generation Sequencing. 134

Table 13. Nucleotide substitutions of the HBV S region sequence amplified by Next Generation Sequencing. 135

Table 14. Nucleotide substitutions of the HBV PreCore region sequence amplified by Next Generation Sequencing.....136
Table 15. a summarised comparison of the number of G to A and the number of C to T substitutions in the HBV X, S, and PreCore regions sequences amplified by Next Generation Sequencing.....137

List of Abbreviations

A3A	Apolipoprotein B mRNA editing catalytic polypeptide 3A
A3B	Apolipoprotein B mRNA editing catalytic polypeptide 3B
A3DE	Apolipoprotein B mRNA editing catalytic polypeptide 3DE
A3F	Apolipoprotein B mRNA editing catalytic polypeptide 3F
A3G	Apolipoprotein B mRNA editing catalytic polypeptide 3G
AID	Activation-induced deaminase
ALT	Alanine aminotransferase
APOBEC3	Apolipoprotein B mRNA editing catalytic polypeptide 3
AST	Asparatate aminotransferase
ADV	adefovir dipivoxil
A	Adenine
C	Cytosine
cccDNA	Covalently closed circular DNA
CDA	Cytidine deaminase active
cDNA	Complementary DNA
CaMV	Cauliflower mosaic virus
CHB	Chronic Hepatitis B
Da	Daltons
DAPI	4',6 Diamidino -2-phenylindole dihydrochloride
dC	2'-Deoxycytidine 5'-triphosphate
DHBV	Duck hepatitis B virus
DMEM	Dulbecco's modified eagle's medium

DMSO	Dimethyl Sulphoxide
DNA	Deoxynucleic acid
DNase	Deoxyribonuclease
DDX3	The DEAD-box helicase family member 3
3D-PCR	Differential DNA Denaturation PCR
dNTP	Deoxyribonucleotide triphosphate
DR1	Direct repeats 1
DR2	Direct repeats 2
EDTA	Ethylenediaminetetraacetic acid
E. coli	Escherichia coli
EnhI	Enhancer I
ETV	entecavir
FBS	Fetal Bovine Serum
FITC	Fluorescein isothiocyanate
G	Guanine
GFP	Green fluorescent protein
GSHV	Ground squirrel hepatitis virus
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
HA	Influenza hemagglutinin
HBcAg	Hepatitis B core antigen
HBeAb	Hepatitis B e antibody
HBeAg	Hepatitis B e antigen
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus
HBx	Hepatitis B X protein

HCC	Hepatocellular carcinoma
HHBV	Heron hepatitis B virus
HIV-1	Human immunodeficiency virus-1
Hsp40	Heat Shock Protein 40
Hsp70	Heat Shock Protein 70
IDO	indoleamine 2,3-dioxygenase
IFN-α	Interferon α
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IPTG	Isopropyl β-D-thiogalactopyranoside
KDa	Kilo dalton
LB	Luria broth
LHBs	Large hepatitis B surface protein
LTR	Long terminal repeat
LAM	lamivudine
MHBs	Middle hepatitis B surface protein
mRNAs	Messenger RNA
MXA	The human myxovirus resistance protein 1
MyD88	myeloid differentiation primary response protein 88
NA	nucleoside or nucleotide
NC	Nucleocapsid
NGS	Next Generation Sequencing
ORFs	Open reading frames
PBS	Phosphate buffered saline
pCMV	CMV promoter

PCR	Polymerase chain reaction
qPCR	Quantitative Polymerase Chain Reaction
PEG	Polyethylene glycol
pgRNA	Pregenomic RNA
PVDF	Polyvinylidene difluoride
rcDNA	Relaxed circular DNA
RdRp	RNA-dependent RNA polymerase
RH	Ribonuclease H
RNA	Ribonucleic acid
RNase	Ribonuclease
RNP	Ribonucleoprotein
RT-PCR	Reverse transcriptase PCR
SDS	Sodium dodecyl sulphate
STAT	signal transducer and activator of transcription
SDS-PAGE	SDS polyacrylamide gel
SHBs	Small hepatitis B surface protein
shRNA	short hairpin RNA
SOCS	Suppressor of cytokine signalling
T	Thymine
TAE	Tris acetate buffer
TAF	tenofovir alafenamide
TBE	Tris borate buffer
TBV	telbivudine
TDF	tenofovir disoproxil fumarate
TE	Tris EDTA buffer

TNF-α	Tumour necrosis factor-α
TRIM22	the tripartite motif 22
TSHV	Tree squirrel hepatitis virus
UDG	Uracil-DNA glycosylase
UV	Ultra violet
USF1	Human Upstream stimulatory factor 1
Vif	Virion infectivity factor
WHO	World health organisation
WHV	Woodchuck hepatitis virus

1 Chapter One: Introduction

1.1 Hepatitis B Virus

1.1.1 Overview (History)

Hepatitis is a term that describes inflammation of the liver. The causes of the inflammation can be viruses, autoimmune responses, or a consequence of alcohol, toxins, or medications. Viral hepatitis is a common cause of hepatitis. The causative agents of viral hepatitis are Hepatitis A, B, C, D, and E viruses. The Hepatitis B virus (HBV) is a significant cause of acute and chronic hepatitis (Fields, Knipe *et al.* 2013).

The history and evolution of the hepatitis B virus go back thousands of years. Barbara Mühlemann's (2018) study showed that Hepatitis B virus infection was present in the Bronze Age. This study showed that the HBV genomes were approximately 4,500 years old (Muhlemann, Jones *et al.* 2018). Patterson Ross and colleagues (2018) examined the remains of a child from the mid-16th century using shotgun sequencing, and the sequence data revealed that the child had been infected with the hepatitis B virus (Patterson Ross, Klunk *et al.* 2018). Lurman (1885) recorded the earliest known HBV epidemic when people were vaccinated against smallpox with human lymph-containing vaccines contaminated with HBV. People who received this smallpox vaccine developed serum hepatitis (jaundice). In 1947, MacCallum introduced a classification of viral hepatitis into Hepatitis A (previously known as infective hepatitis) and Hepatitis B (formerly known as serum hepatitis) (Purcell 1993).

The Golden Age of hepatitis B research began in 1964 with the discovery of the Australian antigen. Baruch Blumberg and Harvey Alter discovered this when studying genetic polymorphisms of serum proteins in the blood samples from Australian aborigines (Blumberg

1964, Blumberg, Sutnick *et al.* 1968, Purcell 1993). The next milestone in Hepatitis research was when Dane and colleagues (1970) were the first to visualise the Hepatitis B virus virions and sub-viral particles by electron microscopy while examining Australian antigen particles. This milestone in HBV research renamed the spherical virus particles as Dane particles, and the Australia antigen was renamed the hepatitis B virus surface antigen (Dane, Cameron *et al.* 1970). Biochemical studies revealed that the virus had an inner nucleocapsid comprised of HBV core protein and an outer membrane containing HBV surface protein. Soon after, Robinson and colleagues discovered a DNA polymerase activity associated with HBV surface antigen (Kaplan, Greenman *et al.* 1973) and that HBV had a DNA genome (Robinson, Clayton *et al.* 1974). The cloning and sequencing of the HBV genome (subtype ayw) followed in 1979 (Galibert, Mandart *et al.* 1979). A study by Will (1982) found that injecting cloned HBV DNA into a chimpanzee's liver caused hepatitis (Will, Cattaneo *et al.* 1982). Concurrently, Summers and Mason proposed the replication of the DNA genome of the hepatitis B virus by reverse transcription of an RNA intermediate. Developing and testing a vaccine for HBV was shown to be protective against HBV and safe to use (McLean, Hilleman *et al.* 1983). All these advances have led to a new era of Hepatitis B virus research.

1.1.2 Classification and Epidemiology

Hepadnaviridae is a family of viruses that infect hepatocytes and share virion and genomes similarities among their members. The human hepatitis B virus (HBV) is the prototype of Hepadnaviridae. Due to differences in their genomes, the Hepadnaviridae family members are classified into two genera. The Orthohepadnavirus genus includes the Human hepatitis B virus (HBV), the Ground squirrel hepatitis virus (GSHV), the Woodchuck hepatitis virus (WHV), and the Tree squirrel hepatitis virus (TSHV). The Avihepadnavirus genus infects birds (Duck hepatitis B virus and heron hepatitis B virus). There are also other similar viruses in Fish (white sucker hepatitis B virus), Old World monkeys (Hepatitis B likem virus), frogs (Bluegill

hepadnavirus), and bats (Hepatitis B like-virus). The Hepadnaviridae share a similar virion size and genome structure of overlapping reading frames and unique replication through a pregenomic RNA intermediate by reverse transcription. The reverse transcription strategy of the hepatitis B virus means that it is also referred to as a pararetrovirus and is closely related to the replication mechanism of retroviruses (HIV) and other para-retroviruses (e.g. Caulimoviridae; Cauliflower mosaic virus). (Fields, Knipe *et al.* 2013, Hohn 2013, Hu and Seeger 2015).

There are eight genotypes A to H of the Hepatitis B virus, with sequence divergence in their genomes ranging between 8% and 17% over the entire HBV genome. Also, two other genotypes, I and J, do not show the same percentage range of pairwise differences in their genomes. The HBV has also been classified into four main serotypes (adr, adw, ayr, ayw) based on differences in HBV amino acid sequences of the surface proteins. This classification was used to study the evolution and geographical distribution of the Hepatitis B virus and as a determinant of the hepatitis severity, treatment response, and vaccine effectiveness (Magnius and Norder 1995, Kramvis, Kew *et al.* 2005, Fields, Knipe *et al.* 2013).

There are currently 296 million people who are chronically infected with the Hepatitis B virus (HBV) worldwide (Table 1), with 1.5 million new infections each year, according to the WHO report (World Health Organization 2022). HBV complications, including liver cirrhosis and liver cancer, lead to more than 820 thousand deaths yearly. The pattern of the geographical prevalence of chronic HBV varies from one part of the world to another (Figure 1). The South Pacific and Asia regions show the highest sero-prevalence with 8%, and the lowest sero-prevalence is in Europe and Central and South American countries with less than 2% prevalence. However, the Mediterranean and Sub-Saharan Africa regions and Alaska and India

exhibit a sero-prevalence ranging between 2% and 7%(Valsamakis 2007, MacLachlan and Cowie 2015).

Table 1. The Global burden of hepatitis B infection. (World Health Organization 2022)

Western Pacific Region	116 million people
African Region	81 million people
Eastern Mediterranean Region	60 million people
South-East Asia Region	18 million people
European Region	14 million people
Region of the Americas	5 million people

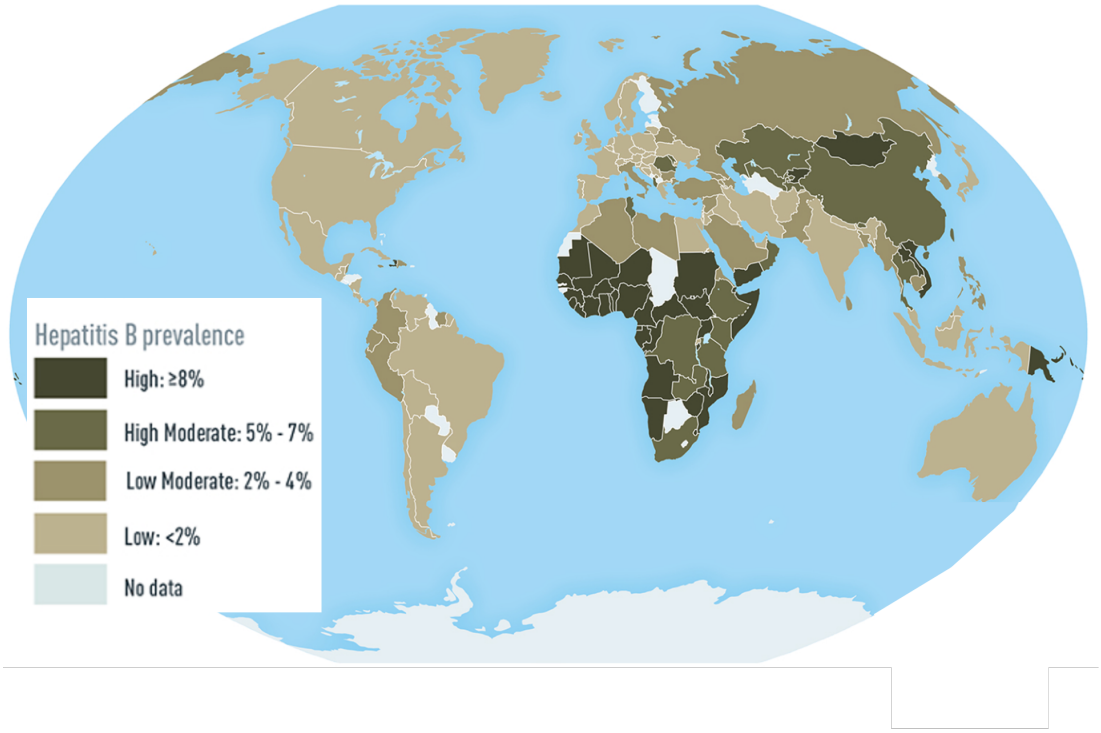


Figure 1. The global prevalence of Chronic HBV Infection. (Centers for Disease Control and Prevention 2019)

1.1.3 Diagnosis and Prevention

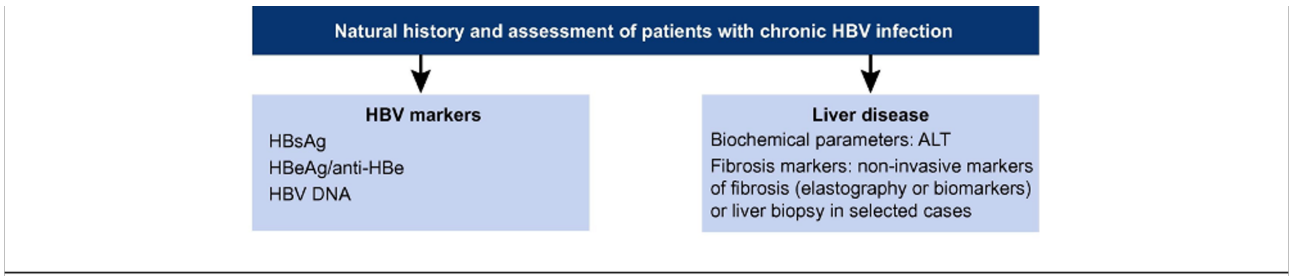
The serological tests include HBsAg, anti-HBs, HBeAg, anti-HBe, anti-HBc, and Serum HBV DNA test. The detection of HBsAg is the key serological marker of Acute and Chronic hepatitis B infection. HBsAg is the standard test to confirm HBV infection. A positive HBsAg test six months after the first test confirms the diagnosis of chronic HBV infection. A negative HBsAg test and a positive anti-HBs indicate an HBV vaccination, recovery from acute HBV infection, or Chronic HBV infection HBsAg seroconversion. HBeAg is used to distinguish between HBeAg-positive and HBeAg-negative chronic HBV infections. The detection of Immunoglobulin M (IgM) anti-HBc in acute HBV infection, or during chronic HBV exacerbations. Immunoglobulin M (IgM) anti-HB is the only serological marker detectable in fulminant acute HBV. Serum HBV DNA test directly assesses viral replication and monitors treatment efficacy. Also, the assessment of the fibrosis and cirrhosis in chronic HBV infection using liver biopsy (invasive test) or include transient elastography (non-invasive test). These tests determine prognosis and management(Fields, Knipe *et al.* 2013, Yuen, Chen *et al.* 2018).

Hepatitis B has no cure but can be prevented by a vaccine. The vaccine for HBV has been developed in the 1980s. HBV vaccine is safe and effective and is usually given in three doses over a period of six months. The two types of vaccines that were developed are plasma-derived vaccines and Recombinant DNA-derived vaccines. The risk of developing chronic hepatitis B virus infection is more than 95% for infants and early childhood. However, the risk for adults to develop chronicity is less than 5%. This is explained by most HBV infections being acquired through mother-to-child transmission. The WHO elimination of HBV infection 2030 included a 90% coverage of HBV vaccination. This is by recommending universal vaccination for HBV by giving the first dose of vaccine to all infants after birth or within 24 hours after delivery. It also suggests that vaccination should also be provided to the high-risk group which includes

healthcare providers, laboratory technicians, and immunocompromised patients. Antiviral prophylaxis using tenofovir disoproxil fumarate (TDF) prophylaxis (safe in pregnancy) in highly viraemic mothers is recommended (Fields, Knipe *et al.* 2013, Yuen, Chen *et al.* 2018).

1.1.4 Treatment of HBV infection

EASL 2017 Clinical Practice Guidelines recommended updated management guidelines for HBV infection. These have shown that chronic HBV infection can be classified into 1) HBeAg-positive chronic infection phase, 2) HBeAg-positive chronic hepatitis phase, 3) HBeAg-negative chronic infection phase, 4) HBeAg-negative chronic hepatitis phase and 5) HBsAg-negative phase (European Association For The Study Of The Liver 2017). These Five phases classification based on both infection and hepatitis. Using these classifications considering the markers of HBV replication (HBeAg, HBV DNA levels, alanine aminotransferase (ALT)) and liver inflammation (Figure 2).



	HBeAg positive		HBeAg negative	
	Chronic infection	Chronic hepatitis	Chronic infection	Chronic hepatitis
HBsAg	High	High/intermediate	Low	Intermediate
HBeAg	Positive	Positive	Negative	Negative
HBV DNA	>10^7 IU/ml	10^4-10^7 IU/ml	<2,000 IU/ml°°	>2,000 IU/ml
ALT	Normal	Elevated	Normal	Elevated*
Liver disease	None/minimal	Moderate/severe	None	Moderate/severe
Old terminology	Immune tolerant	Immune reactive HBeAg positive	Inactive carrier	HBeAg negative chronic hepatitis

Figure 2. The phases of chronic HBV infection and natural history and patient assessment are based on HBV and liver disease markers (European Association For The Study Of The Liver 2017).

*Persistently or intermittently. °°HBV DNA levels can be between 2,000 and 20,000 IU/ml in some patients without signs of chronic hepatitis.

Currently, there are two main treatment options for CHB patients: treatment with a nucleoside or nucleotide (NA) analogue or with IFNa, currently pegylated (PegIFNa) (Table 2). The NAs that have been approved in Europe for HBV treatment include lamivudine (LAM), telbivudine (TBV), entecavir (ETV), adefovir dipivoxil (ADV), tenofovir disoproxil fumarate (TDF) and tenofovir alafenamide (TAF) and can be classified into those associated with a low barrier against HBV resistance (LAM, ADV, TBV) and those with a high barrier to HBV resistance (ETV, TDF, TAF). LAM, TBV, and ETV are nucleoside analogues, while ADV, TDF and TAF are nucleotide analogues.

The main advantage of treatment with a potent NA with a high barrier to resistance (i.e., ETV, TDF, TAF) is its predictable high long-term antiviral efficacy leading to undetectable HBV DNA levels in the vast majority of compliant patients as well as its favourable safety profile. These drugs can be safely used in any HBV-infected patient and represent the only treatment option for several patient subgroups, including those with decompensated liver disease, liver transplants, extrahepatic manifestations, acute hepatitis B or severe chronic HBV exacerbation. The objective of chronic hepatitis B treatment outcomes can be defined as 1) *Partial cure, 2) Functional cure, 3) Complete cure, or Sterilising cure* (Table 3). The goal of treatment for chronic hepatitis B is to suppress HBV replication, prevent the progression of liver disease and, thereby the development of cirrhosis, liver failure and HCC(European Association For The Study Of The Liver 2017, Lok, Zoulim *et al.* 2017).

Table 2. Advantages and disadvantages of nucleos(t)ides analogues and PEG-IFN-alpha in treating Chronic Hepatitis B (Phillips, Jagatia *et al.* 2022).

	Nucleos(t)ides analogues	**PEG-IFN-alpha**
Advantages	Oral Good tolerability No Contraindications	Finite duration Higher rate of HBeAg seroconversions (30%) Higher rate of HBsAg loss (5-8%) Loss of cccDNA may occur
Disadvantages	Indefinite duration HBsAg loss <2% No effect on cccDNA.	Low tolerability (Significant side effects) Subcutaneous injections

Table 3. The Goal of HBV therapy (Lok, Zoulim *et al.* 2017, Phillips, Jagatia *et al.* 2022).

	Partial cure	**Functional cure**	**Complete cure**	**Sterilising cure**
Serum HBV DNA	Undetected	Undetected	Undetected	Undetected
HBsAg	Detected	Undetected	Undetected	Undetected
Anti-HBsAg	Undetected	±	Detected	Detected
cccDNA	Detected	Detected	Undetected	Undetected
Integrated HBV DNA	Present	Present	Present	Absent

1.1.5 Virion Morphology

The human hepatitis B virus appears under electron microscopy as a 42-nanometer infectious virion particle (Dane particle) and non-infectious 22-nanometer filamentous and ~20-nanometer sphere-shaped particles (Figure 3). The infectious HBV Dane double-shelled particle consists of an outer envelope (4-7 nm) made up of endoplasmic derived-lipid bilayer and transmembrane viral surface proteins (HBsAg) (large, middle, and small). This outer envelope surrounds the inner icosahedral (HBcAg) capsid (~22-27 nm). HBV capsid contains a single partially double-stranded relaxed circular (RC) DNA molecule about 3.2 kb in length, covalently linked to the viral polymerase at the 5' end of the negative strand of HBV DNA.

The structure of the Hepatitis B Virus capsid *in vitro* and *in vivo* was studied intensively by Crowther *et al.* 1994, Bottcher *et al.* 1997, Conway *et al.* 1997, Wynne *et al.* 1999 and Dryden 2006 using Electron Cryomicroscopy and X-ray crystallography to determine the three-dimensional structures of the Hepatitis B Virus capsid also known as Core Particle, the capsid was shown to be comprised of dimers of α-helical coiled coils of HBV core protein which form spikes. The HBV surface proteins span the virus lipid envelope surrounding the capsid and project from the virus's surface. The capsid contains 240 protein subunits with a triangulation number T = 4 or 180 with a triangulation number T = 3. Crowther *et al.* 1994 and Bottcher *et al.* 1997 also observed that the Hepatitis B Virus capsid has holes, which may allow access for small molecules needed for the viral replication (Crowther, Kiselev *et al.* 1994, Bottcher, Wynne *et al.* 1997, Dryden, Wieland *et al.* 2006, Prasad and Schmid 2012, Fields, Knipe *et al.* 2013).

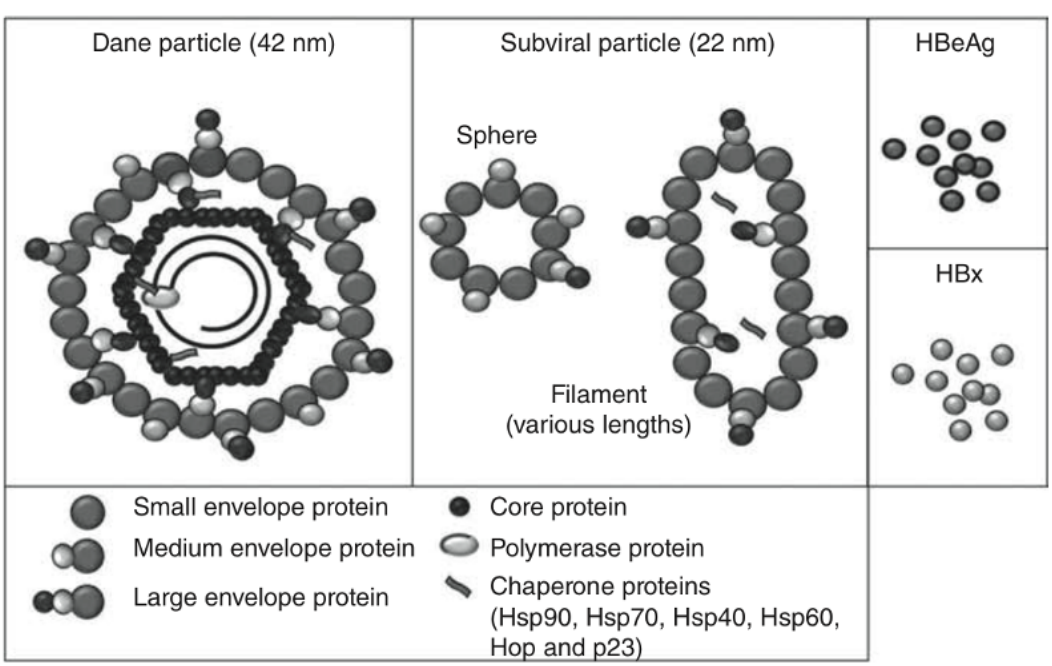


Figure 3. HBV Virion Morphology. (Luangsay, Lyon University et al. 2013)

Schematic representation of the types of HBV Dane particle and subviral particles. The Dane particle with the three envelope proteins (L, M and S) the core protein, the polymerase (RT, RNAseH and TP) and the partially double stranded DNA genome. The empty subviral particles (sphere and filaments) with the three envelope proteins (L, M and S). The HBeAg and HBx proteins.

HBV capsids also contain cellular proteins, including serine kinases mediating HBV core protein phosphorylation during viral infection (Albin and Robinson 1980, Daub, Blencke *et al.* 2002, Dryden, Wieland *et al.* 2006, Bruss 2007, Patient, Hourioux *et al.* 2009, Fields, Knipe *et al.* 2013).

The titer of the infectious virions of hepatitis B virus in the blood of infected patients can vary significantly. The most abundant particles released into the blood from infected liver cells are the non-infectious subviral particles. The number of 22-nanometer non-infectious subviral spheres can reach as high as 10^{12}/ml. These subviral particles comprise 40% lipid and dimers of middle and small HBV surface proteins with traces of large HBV surface proteins. These non-infectious subviral particles vastly outnumber the infectious virions of the hepatitis B virus. A proposed possibility of the role of the non-infectious subviral particles is to bind neutralizing anti-HBs antibodies allowing the infectious virions to evade these virus-neutralizing antibodies. Dane particles have abundant HBV surface proteins needed to attach to HBV receptors on the liver cell membranes. This abundance of large HBV surface proteins in Dane particles gives them an advantage in infecting liver cells compared to the non-infectious subviral particles that only contain middle and small HBV surface proteins (Fields, Knipe *et al.* 2013, Seeger and Mason 2015).

1.1.6 HBV Genome structure

The HBV genome is a small, partially double-stranded relaxed circular (RC) DNA molecule consisting of a complete coding negative DNA strand that is 3,200 base pairs in length with terminal redundancy of 7-9 nucleotides and an incomplete non-coding positive DNA strand that is of variable length. The two strands are held in circular morphology by 11 nucleotide repeats at strands 5' ends called DR1 and DR2; these two repeats play an essential role in the HBV replication (Nassal 1999, Seeger and Mason 2000) (Figure 4).

The HBV genome replicates through an RNA intermediate, the pre-genomic RNA, by reverse transcription (Seeger and Mason 2000, Lamontagne, Bagga *et al.* 2016). The complete coding minus strand of the HBV genome contains four overlapping open reading frames (ORF) that encode four proteins: a DNA polymerase for viral replication (P), a surface protein (S), a core protein (C) and an X protein (Figure 4). ORF P is the longest among the four ORFs, and it consists of four domains that serve multiple functions of DNA polymerase during HBV viral replication. These domains encode an N-terminal protein (TP), spacer protein, polymerase, reverse transcriptase and Ribonuclease H activities. The TP protein can form a covalent bond with the 5' end of the negative DNA strand. The formation of this covalent bond is the priming step for minus-strand DNA synthesis and is vital for pre-genomic RNA encapsidation (Zoulim and Seeger 1994). The spacer domain is after the TP in ORF P, which separates TP from the C-terminal domain. The C-terminal domain contains a reverse transcriptase (RT) synthesising the DNA strand from the 3.5kb pre-genomic RNA and an RNase H degrading pgRNA. ORF pre-S/S encodes the HBsAg (large, medium, and small envelope glycoproteins), and completely overlaps ORF P. ORF C partially overlaps with ORF P and encodes the HBeAg and HBc (core protein). ORF X also partially overlaps with ORF P and directs the synthesis of the small HBx protein. This ORF is only found in mammalian hepadnaviruses. The HBV genome also contains two enhancers (EN1&EN2) and four promoters that control transcription of the HBV genome (Seeger and Mason 2015, Lamontagne, Bagga *et al.* 2016).

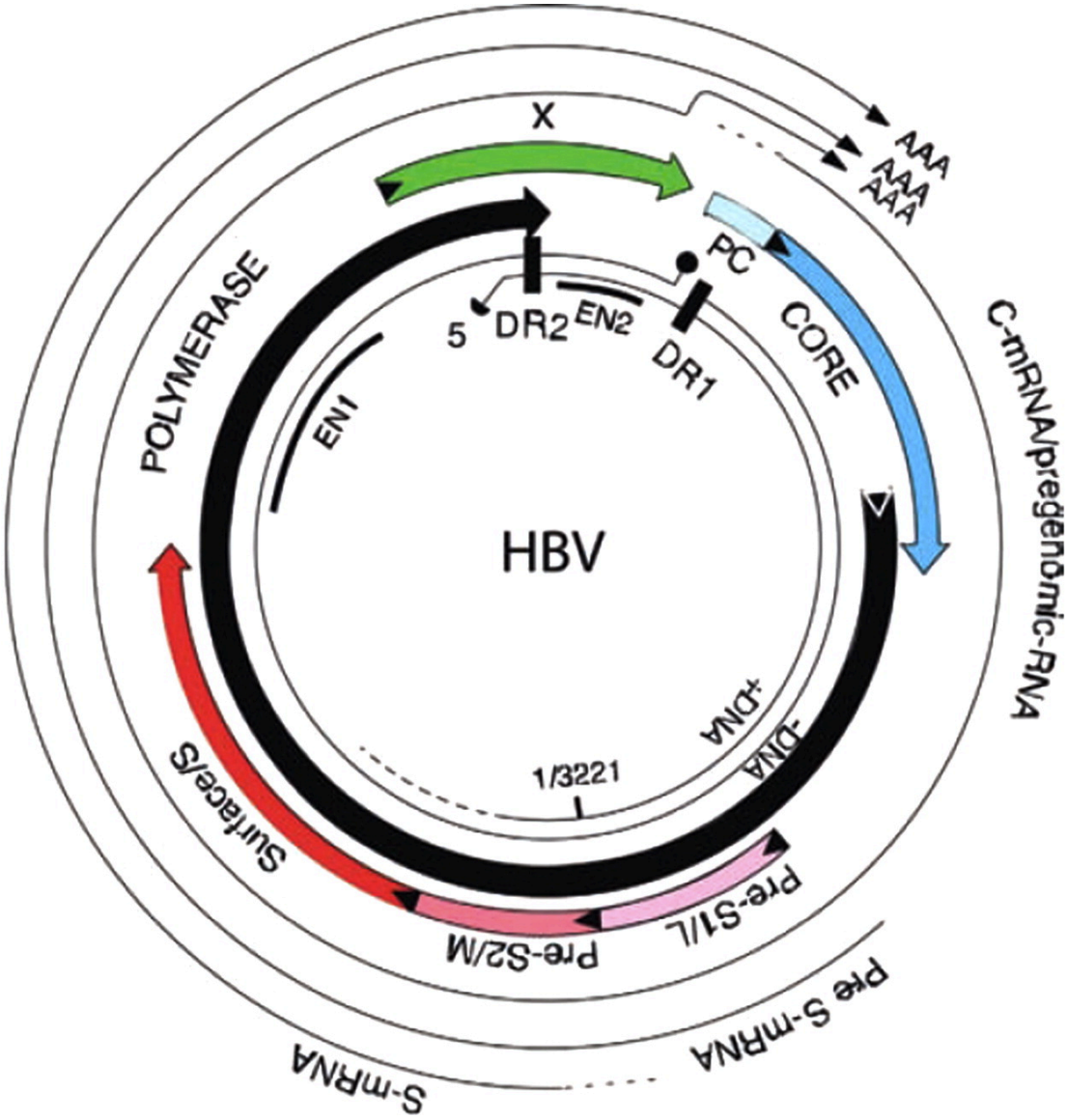


Figure 4. HBV Genome Structure.

Schematic structure of the circular, partially double-stranded HBV DNA genome (3.2 kb). It shows the HBV four overlapping open reading frames (ORF); HBV PreCore and Core (PreC, C) ORF (Blue arrow), HBV polymerase (P) ORF (Black arrow). HBV surface proteins (PreS1, PreS2 & S) ORF (red arrow). HBV X ORF (green arrow). It also shows two enhancer regions, Enh1 and Enh2, and two direct repeats, DR1, and DR2, core mRNA /pregenomic RNA, pre-S mRNA and S mRNA (Seeger and Mason 2015).

1.1.7 HBV life cycle and replication

1.1.7.1 HBV attachment & entry:

HBV infects hepatocytes, and the first components of the virus that interact with these cells are the viral surface proteins (L, M and S). For many years the early phase of HBV infection, including attachment, was poorly understood. This gap in understanding the early stage of HBV infection may be because undifferentiated hepatoma cell lines do not express the receptor for HBV (Glebe and Urban 2007), which is present in intact liver hepatocytes. However, recently, there has been an improvement in understanding how HBV attaches and enters hepatocytes as a result of advances in cell culture and the availability of new HBV research models; primary human hepatocytes (PHH) (Glebe and Urban 2007), transgenic mice with human liver, HepaGR (immortalized and non-transformed cell line) and primary Tupaia hepatocytes (PTH) (Kock, Nassal *et al.* 2001, Gripon, Rumin *et al.* 2002). It is thought that the preS domain of HBV large surface protein plays an essential role in the viral attachment; about half of the preS domain of the large surface protein is translocated to the exterior of the HBV particle during the budding process to act as an attachment site (Kock, Nassal *et al.* 2001). HBV attachment and subsequent entry into the host hepatocyte involve more than one step. The belief is that the virion is captured by heparin sulfate proteoglycan that binds to the HBV small HBsAg protein (Schulze, Gripon *et al.* 2007). Then it attaches to the high-affinity receptors gp180 (carboxypeptidase D) and glycosaminoglycans that facilitate the entry of the virus (Kuroki, Eng *et al.* 1995, Leistner, Gruen-Bernhard *et al.* 2008). Most recently, the sodium-dependent taurocholate co-transporting polypeptide (NTCP) has been identified as a human liver-specific multiple trans-membrane transporter proteins and a receptor that attaches to the preS domain of the HBV surface (Yan, Zhong *et al.* 2012, Seeger and Mason 2013) (Figure 5).

1.1.7.2 HBV nucleocapsid nuclear localization:

Following the HBV-specific attachment to receptors, as mentioned earlier and possibly others, the nucleocapsid release from the endosome into the cytoplasm might be mediated by a translocation motif near the N-terminal preS1 of DHBV. Then the nucleocapsid is transported by microtubules to the hepatocyte nucleus (Stoeckl, Funk *et al.* 2006). The phosphorylation of the C-terminal domain of the core proteins exposes the nuclear localization signal, which interacts with and binds to the Importins Alpha and Beta, facilitating the transport of the nucleocapsid to the nuclear pore complex (Kann, Sodeik *et al.* 1999). It appears that the viral dsDNA induces changes in the nucleocapsid and as a result, leads to the expression of nuclear localization sequences. By contrast, if the nucleocapsids contain immature viral pre-genomic RNA, they will be held in the nuclear basket until maturation. The mature capsid releases the relaxed circular DNA into the karyoplasm, where the cellular machinery repairs to form the cccDNA (Schmitz, Schwarz *et al.* 2010) (Figure 5).

1.1.7.3 Covalently closed circular DNA (cccDNA) formation & transcription

The mature viral nucleocapsid releases the relaxed circular DNA (rcDNA) into the nucleus where it forms an episomal covalently closed circular DNA (cccDNA). There is no clear mechanism for the formation of cccDNA. However, a possible explanation is that the virus depends on cellular repair mechanisms to form its episomal cccDNA. Furthermore, this requires several steps: the removal of the viral DNA Polymerase linked to the 5' of minus-strand DNA, the removal of the capped RNA oligomer region linked to the 5' end of the incomplete positive-strand DNA, and the synthesis of DNA to complete this strand. Following this, the 5' and 3' ends of both the minus and positive DNA strands are covalently ligated to form a plasmid-like DNA called cccDNA (Nassal 2008, Sohn, Litwin *et al.* 2009, Fields, Knipe *et al.* 2013, Guo and Guo 2015, Seeger and Mason 2015). HBV cccDNA binds histones (H3,

H4) and non-histone proteins to form a viral mini-chromosome (chromatin-like structure). The acetylation and methylation of cccDNA-bound H3/H4 regulate and control the transcription of HBV (Pollicino, Belloni *et al.* 2006). The mini-chromosome also contains the HBV core protein that reduces the spacing of HBV nucleoproteins (Bock, Schwinn *et al.* 2001). The number of cccDNA copies in hepatocytes varies, as was shown in patient liver biopsies and animal models. This number ranges from 1 to 50 cccDNA copies per hepatocyte (Allweiss and Dandri 2017) (Figure 5).

Viral cccDNA works as a template for cellular RNA polymerase II to generate viral RNAs (sub-genomic & genomic RNAs), which are capped at 5′end and terminally poly-adenylated at the 3′ end. A 2.4kb and 2.1kb viral sub-genomic RNAs act as mRNAs for S, preS1 and preS2 (small, middle and large viral proteins), and 0.7 kb sub-genomic RNAs act as an mRNA for the X protein, while a 3.5kb pre-genomic RNA (pgRNA) act as mRNAs for core and polymerase proteins. The polymerase protein is then covalently bound to pgRNA, and both are packaged and encapsidated by core proteins, upon which pgRNA is reverse transcribed to produce relaxed circular DNA. The pg RNA contains the full-length sequence of the viral genome bounded by a direct repeat sequence (DR1), a second repeat region (DR2), an encapsidation signal (ε), a 5' cap and a poly-A tail (Beck and Nassal 2007, Fields, Knipe *et al.* 2013, Seeger and Mason 2015). Inactive P protein undergoes folding changes that produce the protein's active form. Using ATP, the active P protein binds to the hairpin encapsidation ε RNA signal and the chaperones Hsp40 & Hsp70. The P protein then transcribes 3-4 oligonucleotides on ε at the 5' end of the minus-strand DNA. The oligonucleotides are covalently linked to a tyrosine residue in the TP domain. The P protein is then translocated to the 3' direct repeat (DR1) to act as a primer to minus DNA strand completion, and the RNA is degraded by Ribonuclease H (RH). Then, an RNA primer is translocated to DR2 and the

positive strand is synthesised towards the 5' prime end. The 3' prime end of the positive strand DNA grows to elongate the R region in the (-) strand. Then circularization followed by elongation of the (-) strand DNA to form a relaxed circular DNA (Nassal 2008, Seeger and Mason 2015) (Figure 5 and 6).

1.1.7.4 HBV assembly and secretion

The mechanism by which the virus is assembled and secreted is still unclear. However, the newly formed nucleocapsid containing the partially relaxed circular DNA is transported to the endoplasmic reticulum or to the nucleus to enrich its viral cccDNA reservoir. In the endoplasmic reticulum, the nucleocapsid proteins interact with the N terminal domain of the abundant surface protein to facilitate the mature virion budding through Multi-vesicular bodies. There is a suggestion that once the envelopment of the virus is complete, the virus interacts with gamma2-adaptin (Clathrin adaptor protein) & ubiquitin ligase Nedd4 and transports to the budding site for release via the Golgi apparatus. The mechanism by which the non-infectious HBV particles are assembled is still unknown (Watanabe, Sorensen *et al.* 2007, Patient, Hourioux *et al.* 2009, Fields, Knipe *et al.* 2013) (Figure 5).

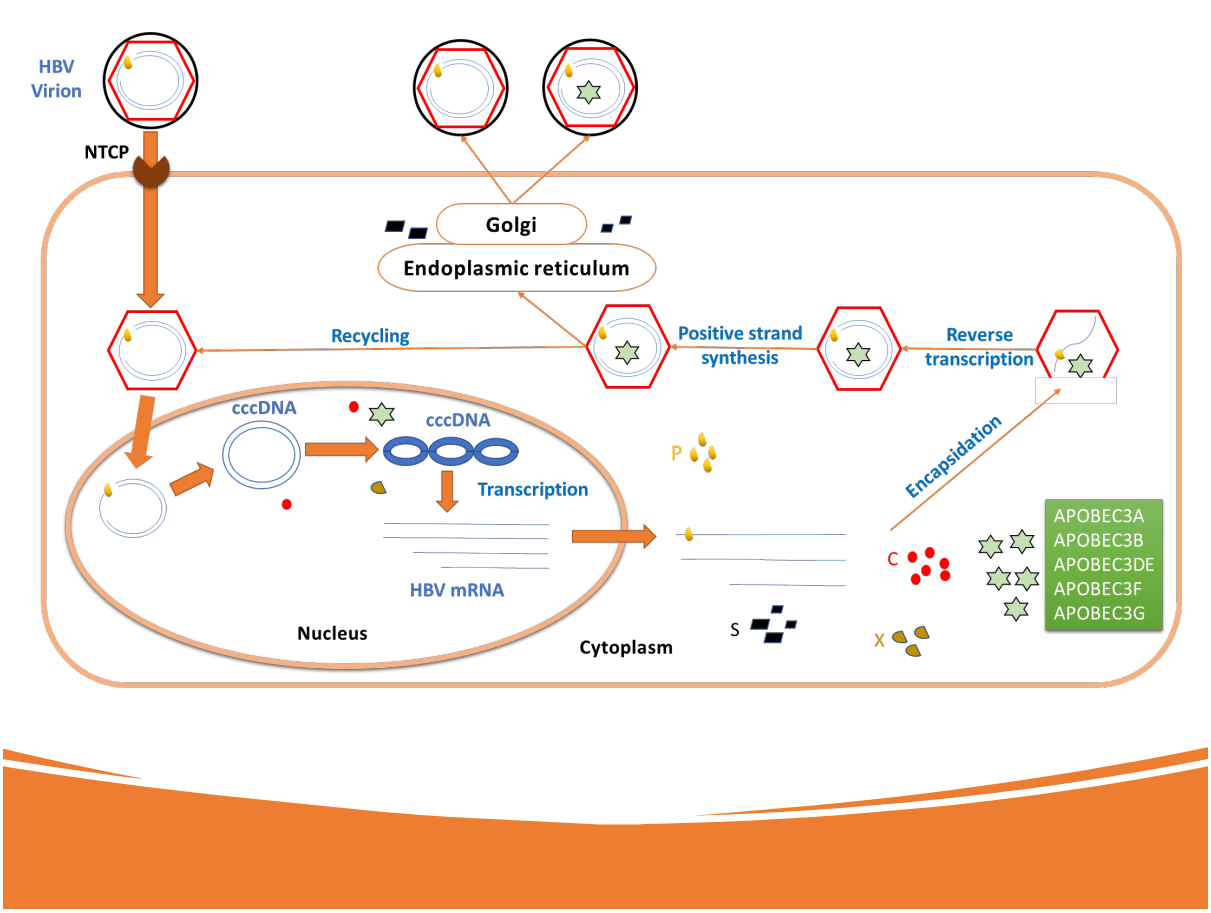


Figure 5. HBV life cycle.

The diagram shows the HBV life cycle in hepatocyte and APOBEC3 proteins' role in HBV replication. 1) HBV attachment & entry. 2) HBV nucleocapsid nuclear localization. 3) Covalently closed circular DNA (cccDNA) formation & transcription. 4) HBV reverse transcription. 5) HBV assembly and secretion.

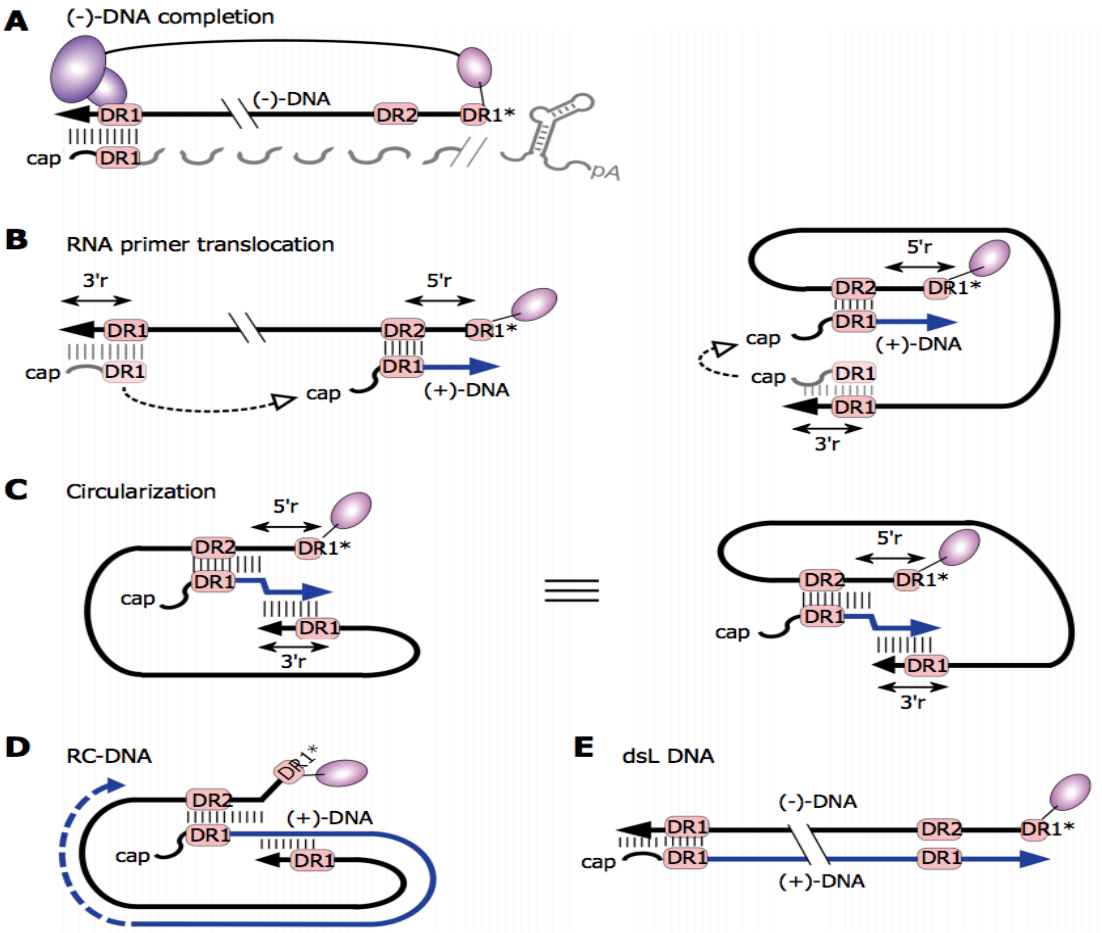


Figure 6. HBV replication.

The diagram shows HBV reverse transcription and elongation of minus strand DNA and then synthesis of the incomplete positive strand DNA. Adapted from (Beck and Nassal 2007)

1.2 APOBEC proteins are Host restriction Factors

1.2.1 Interferon (IFN)

The discovery of the antiviral agent Interferon (IFN) by Isaacs and Lindenmann in 1957 (Isaacs and Lindenmann 1957) during studies on virus interference has led to many studies exploring the members of the interferon family and their effects on viruses.

The interferon family can be divided into subcategories; type I IFN (antiviral interferon), types II IFN and types III IFN. Type I interferon is a group of three members; IFN-α (further subdivided into 13 members), IFN-Beta and IFN-ω, while type II IFN is also called IFN-γ (Samuel 2001). Type I interferon genes are on chromosome 9.

Type I interferon (alpha, beta, and omega) binds to cognate multi-subunit trans-membrane specific cellular receptors, which are IFNAR1 and IFNAR2 (Mogensen, Lewerenz *et al.* 1999), upon binding these receptors to heterodimerize and then activate the JAK-STAT pathways through phosphorylation of the signal transducer and activator of transcription (STAT) family (which includes; Stat-1, Stat-2, Stat-3, Stat-4, Stat-5a, Stat-5b, and Stat-6) by Janus family (which includes; Jak-1, Jak-2, Jak-3, and Tyk-2) of tyrosine kinase (JAK). Upon IFN-alpha and Beta, the phosphorylated STAT1 and STAT2 form a dimer that subsequently binds to interferon regulatory factor 9. This Interferon-stimulated factor complex translocates to the nucleus, stimulating the transcription of interferon-stimulated genes through interferon-stimulated response elements (ISRE) (Samuel 2001). These interferon-stimulated genes (ISGS) produce proteins that exert antiviral activity and play an integral role in the defence against pathogens.

Several studies have reported that IFNs and ISGS (Table 4) inhibit HBV replication through various mechanisms. It was reported that interferon-alpha cause hypo-acetylation of histone

binding to cccDNA minichromosomes, which modifies its epigenetics and suppresses its transcription to pregenomic and subgenomic RNAs (Belloni, Allweiss *et al.* 2012, Liu, Campagna *et al.* 2013). Also, the activity of enhancer or promoter, the RNA stability, and nucleocapsid formation are suppressed by Interferon (Pei, Chen *et al.* 2014).

Table 4. The IFN-stimulated genes (ISGs) inhibitory effects on the HBV life cycle.

APOBEC	The APOBEC3 (apolipoprotein B mRNA editing enzyme catalytic polypeptide-like) proteins are cellular restriction factors. which have been shown to defend cells against viruses' infections through restricting viral replication for several viruses e.g., HIV, and HBV as well as retrotransposons. Apolipoprotein B mRNA editing enzyme family members are cytidine to uridine deaminases that can alter the coding sequence and insert premature stop codons lead to viral dysfunction or even degradation. This family of proteins also interfere with HBV nucleocapsid formation and maturation as well as the regulation of HBV promoter and enhancer activity (Janahi and McGarvey 2013).
MxA	It is an interferon-induced dynamin-like GTP that reduces the level of the encapsidated HBV pregenomic RNA and interfere with nucleocapsid formation (Li, Zhang *et al.* 2012).
MyD88	It is myeloid differentiation primary response protein 88 was shown to reduce the export of the viral RNAs and the stability of pre-genomic RNA (Li, Lin *et al.* 2010).
IDO	It is indoleamine 2,3-dioxygenase, is induced by interferon-gamma (type II IFN) and suppresses HBV replication through tryptophan deprivation (Mao, Zhang *et al.* 2011).

TRIM22	It is a member of the tripartite motif (TRIM) family that is induced upon interferon stimulation and was shown to inhibit HIV, influenza A virus and encephalomyocarditis virus. It is reported that TRIM22 inhibits the replication of HBV by controlling the activity of viral core promoter (Gao, Duan *et al.* 2009).

1.2.2 Apolipoprotein B mRNA editing enzyme catalytic polypeptide-like (APOBEC)

The Apolipoprotein B mRNA editing enzyme catalytic polypeptide-like (APOBEC) family of proteins can edit DNA &/or RNA. These cytidine deaminases are a member of a larger family of deaminases. In mammals, the members of the APOBEC family are activation-induced deaminase (AID), APOBEC1, APOBEC2, the seven members of the APOBEC3 group (A, B, C, D, F, G & H) and APOBEC4 (Janahi and McGarvey 2013, Vieira and Soares 2013). Their structure consists of one or two catalytic domains (CD1& CD2), each of which contains a zinc-dependent cytidine or deoxycytidine deaminase motif that has a conserved sequence of amino acids [H-X-E-X (23−28)-P-C-X (2−4)-C] where X represents a variety of amino acids (Vasudevan, Smits *et al.* 2013). Each of the APOBEC proteins has one of these catalytic zinc-dependent motifs except for APOBEC3A, DE, F and G, which have two catalytic zinc-dependent motifs. APOBEC3 proteins also have two critical aromatic residues that are needed for RNA substrate binding, and they have also a glutamate residue, which is important to shuttle protons during deamination. The cytidine deamination is produced when water donates the hydroxyl group, followed by removing an amino group at position 4 of a cytidine or deoxycytidine to generate uridine or deoxyuridine and ammonia (Wedekind, Dance *et al.* 2003) (Figure 7).

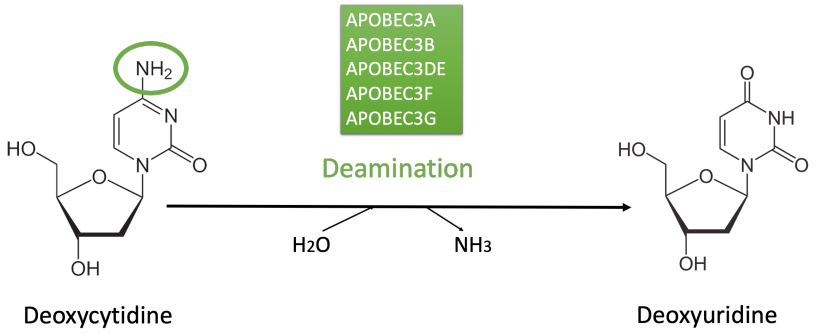


Figure 7. APOBEC3 proteins' deamination of deoxycytidine to deoxyuridine.
Deamination is the hydrolysis reaction that converts cytosine into uridine by the removal of an amine group and the release of ammonia.

The first member of the cytidine deaminases family to be identified was APOBEC1 (A1), after which the family was named. APOBEC1 gene is located in chromosome 12. A1 edits RNA (Teng, Burant *et al.* 1993) as well as DNA (Harris, Petersen-Mahrt *et al.* 2002) in both the cytoplasm (Yang and Smith 1997) and nucleus (Lau, Xiong *et al.* 1991, Sowden, Ballatori *et al.* 2002). In the mammalian gastrointestinal tract, the expression of A1 deaminates cytosine to uridine at position 6666 of the apolipoprotein B 100 mRNA forms a premature stop codon (UAA) (Teng, Burant *et al.* 1993). As a result, a shorter apolipoprotein B (ApoB48) is produced, which exhibits a different function in lipid absorption and transport to the ApoB100 (Anant and Davidson 2001).

Activation-induced deaminase (AID) is another member of the APOBEC family that deaminate single-strand DNA (ssDNA) (Bransteitter, Pham *et al.* 2003, Dickerson, Market *et al.* 2003), and its gene is located in chromosome 12. Although it is mainly in the cytoplasm, it travels back and forth between the nucleus and the cytoplasm (Rada, Jarvis *et al.* 2002, Brar, Watson *et al.* 2004, Ito, Nagaoka *et al.* 2004). The AID expression in germ centre B cells is required for antibody diversification through class switch recombination and somatic

hypermutation (Muramatsu, Sankaranand *et al.* 1999, Muramatsu, Kinoshita *et al.* 2000). It is also reported that AID has a role in the demethylation of DNA (Popp, Dean *et al.* 2010).

APOBEC2 (A2) expression is specific for the heart, and skeletal muscles (Liao, Hong *et al.* 1999) and its gene is in chromosome 6. Although the physiological role of A2 is still unclear, it is suggested that it is essential in developing muscle (Sato, Probst *et al.* 2010). In mammals, the APOBEC4 (A4) gene is located in chromosome 1; its function is still unclear. A4 is mainly expressed in the testis, where it is thought to edit mRNA in the spermatogenesis (Rogozin, Basu *et al.* 2005, Lada, Krick *et al.* 2011).

APOBEC3 subfamily includes seven members, their genes form a tandem cluster located on chromosome 22, and they are expressed in mammals to produce seven proteins. These proteins are APOBEC3A (A3A), APOBEC3B (A3B), APOBEC3C (A3C), APOBEC3DE (A3DE), APOBEC3F (A3F), APOBEC3G (A3G), and APOBEC3H (A3H) (Jarmuz, Chester *et al.* 2002, Belloni, Allweiss *et al.* 2012).

APOBEC3 proteins can restrict retroviruses (e.g., human immunodeficiency virus type 1 (HIV-1)) through viral restriction (Figure 8). G to A mutations induced by APOBEC3 proteins during the reverse transcription step of the viral replication cycle leads to the inhibition of retroviruses HIV-1 lacking its viral infectivity factor (Vif) (Hultquist, Lengyel *et al.* 2011, Malim and Bieniasz 2012, Desimmie, Delviks-Frankenberrry *et al.* 2014). HIV-1 Vif protein counter mechanism to APOBEC3 effects on HIV-1 through polyubiquitination and degradation APOBEC3G (Conticello, Harris *et al.* 2003, Marin, Rose *et al.* 2003). HIV-1 Vif protein also reduces the expression and inhibits the packaging of APOBEC3G (Kao, Khan *et al.* 2003). APOBEC3 proteins editing may cause viral diversification of HIV-1 (Armitage, Deforche *et*

al. 2014). APOBEC3 proteins also can restrict retrotransposons during reverse transcription. There are long terminal repeat (LTR) retrotransposons and non-LTR retrotransposons, LINE-1(Arias, Koyama *et al.* 2012). The data cancer genome shows the APOBEC3 signature. APOBEC3A and APOBEC3B expression in cancer and the link of APOBEC3A and APOBEC3B to Breast Cancer contribute to breast cancer was established (Nik-Zainal, Alexandrov *et al.* 2012, Burns, Lackey *et al.* 2013, Burns, Temiz *et al.* 2013, Nik-Zainal, Wedge *et al.* 2014) (Figure 8).

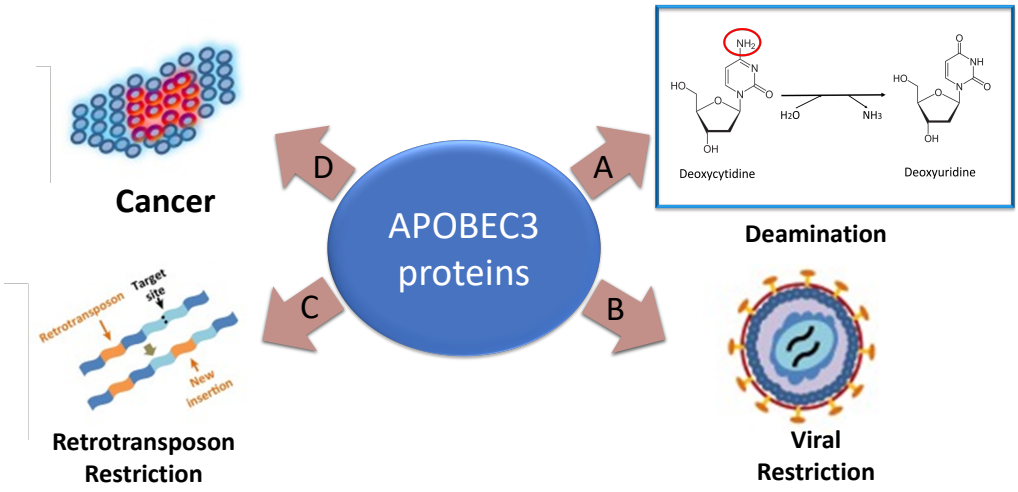


Figure 8. Functions of DNA Editing by APOBEC3 proteins Cytidine deamination.
This illustration shows how APOBEC3 proteins exert their deamination and their role in viral restriction, retrotransposon, and Cancer. Adapted from (Knisbacher, Gerber *et al.* 2016).

1.2.3 Effects of APOBEC3 proteins on HBV replication

APOBEC3 family members showed an ability to inhibit HBV in several studies. This was demonstrated by Turelli *et al.* (Turelli, Mangeat *et al.* 2004) when a fifty-fold reduction in HBV DNA was achieved by co-transfection of Huh7 cells with an HBV-expressing plasmid and an A3G expressing plasmid. It also showed a reduction in the core-associated RNA. However, the A3G inhibition of HBV was eliminated when the HIV-1 VIF protein was co-expressed. In the same study, it was shown that the use of an A3G mutant, which has an inactive cytidine deaminase, produced the same reduction in the HBV DNA as an active one; this

suggested that inhibition of HBV replication by A3G might also be due to mechanisms other than cytidine deamination. These mechanisms could be inhibiting pre-genomic RNA packaging and interfering with the reverse transcription (Rosler, Kock *et al.* 2005, Nguyen, Gummuluru *et al.* 2007).

Another study that investigated the effect of A3G in both Huh7 and HepG2 found inhibition in HBV replication and demonstrated that A3G could mutate, replicating HBV. However, these mutations were significantly higher in the HBV genome produced from HepG2 than that from Huh7 cells (Rösler, Köck *et al.* 2004).

In a study by Suspene *et al.* , they transfected both HBV expressing into Huh7 and HepG2 with A3G, A3C, A3F and A3B plasmids, then amplified the AT-rich DNA by sensitive 3D-PCR. This study found hyper-mutations in viral genomes in the negative DNA strand that cause G to A changes in the positive strand. Moreover, C to T mutations have been observed in the (+) DNA strand. Also, patients' serum samples showed a hyper-mutated viral genome (Suspene, Guetard *et al.* 2005).

In APOBEC3G, both deaminase domains CD1 & CD2 can cause HBV replication inhibition in Huh7 and HepG2 hepatoma cell lines, as well when their constructs were delivered to mice, where the HBV serum level was reduced 50-fold (Lei, Tian *et al.* 2006). Although both A3B domains play a role in HBV DNA reduction, only the C-terminal domain showed deamination activity. A3B lost its hyper-mutation ability when the C-terminal domain of A3B was inactivated by a point mutation, and subsequently its ability to inhibit the viral replication was diminished by 60%. In contrast, A3B did not lose its ability to produce hypermutation in the HBV genome when a point mutation inactivated its N-terminal domain at the H66R position.

However, it caused a 37% reduction in the inhibition of the HBV replication (Bonvin and Greeve 2007).

It was suggested that A3G exhibits its inhibitory effects on the early stages of HBV reverse transcription by showing that the number of longer (-) strand HBV DNA molecules was decreased by the inhibition of A3G in comparison to shorter (-) strand DNAs (Nguyen, Gummuluru *et al.* 2007).

Another mechanism that is independent of the deamination activity of APOBEC3 was shown to inhibit HBV gene expression and the synthesis of HBV core proteins through the interaction between APOBEC3B and the heterogeneous nuclear ribonucleoprotein K (hnRNP-K), a positive regulator of HBV expression, to prevent the latter from binding with HBV enhancer II (Zhang, Zhang *et al.* 2008).

Several studies have shown that APOBEC3G, APOBEC3F and APOBEC3C bind to the viral nucleocapsid. It was demonstrated that A3G was incorporated into the replication-competent HBV nucleocapsids. As A3G was not digested after protease treatment, it was suggested that A3G was inside the viral nucleocapsid (Nguyen and Hu 2008). This A3G incorporation into the nucleocapsid is suggested to be by direct binding to core protein as was shown through an examination using fluorescence resonance energy transfer (FRET) and surface Plasmon resonance and completely independent of both cellular and viral RNAs (Zhao, Wang *et al.* 2010). Co-transfection of HepG2 cells that expressed HBV with a plasmid encoding Core-A3C fusion protein showed a significant decrease in both intracellular & extracellular HBV DNA levels and a large increase in several G-to-A hypermutations (Li, Liu *et al.* 2011).

The expression of APOBEC proteins was observed in the liver of healthy humans at low to moderate levels (Jarmuz, Chester *et al.* 2002, Bonvin, Achermann *et al.* 2006). However, the expression of these proteins can be induced by interferon types II and I. One study showed that when primary human hepatocytes (PHH) and Huh7 and HepG2 cells are treated with interferon-alpha, this increases APOBEC3G expression by an 80-fold but not a large increase in the expression of APOBEC3B, APOBEC3C & APOBEC3F (Tanaka, Marusawa *et al.* 2006). In contrast, another study showed an approximately 14-fold increase in APOBEC3G, APOBEC3B, APOBEC3C and APOBEC3F in primary human hepatocyte (PHH) and a decrease in virus replication levels (Bonvin, Achermann *et al.* 2006). Similarly, the use of interferon-gamma induced the A3B, A3F and A3G (Jost, Turelli *et al.* 2007), and as a consequence, HBV levels were reduced by 50% (Proto, Taylor *et al.* 2008). However, it was found that APOBEC3 enzymes are not essential for the anti-viral effects of both interferon-alpha and gamma on HBV replication in HepG2 cells and HBV mouse model. Using RNA interference and HIV-VIF protein or knocking down only one APOBEC (mA3) in mice proved this and suggested other mechanisms by which Interferon alpha or gamma induce their antiviral effects on HBV replication in these models (Jost, Turelli *et al.* 2007, Proto, Taylor *et al.* 2008, Turelli, Liagre-Quazzola *et al.* 2008).

A study by Lucifora *et al.* in 2014 (Lucifora, Xia *et al.* 2014) showed that using Interferon-α and lymphotoxin-β receptor activation to up-regulate A3A and A3B, respectively, these APOBECs directly interacted with HBV core in the host cell. After that, they translocate to the nucleus, where they are brought into close contact with cccDNA by the HBV core. The APOBECs could then deaminate cccDNA that is transiently rendered single-stranded during transcription. However, Chisari *et al.* raised several technical questions. Firstly, were qPCR assays used cccDNA specific? They suggested using Hirt extraction to isolate cccDNA from

other HBV DNAs and to normalized it to the mitochondrial DNA to show the specificity of the measurements carried out. The second point was whether the non-cytopathic cccDNA loss was due to cell death or regeneration or mediated by IFN-α or LTβR. (Chisari, Mason *et al.* 2014). Previous studies showed that tumour cell apoptosis (Hu, Zimmerman *et al.* 2013) and proliferation (Tumanov, Koroleva *et al.* 2009) were triggered by LTβR-agonism.

The effects of APOBEC3 protein are not only inhibitory. One study showed that the deamination activity of APOBEC3 on HBV could lead to an escape mutation in the virus, consequently leading to an increase in HBV production (Vartanian, Henry *et al.* 2010). This was also shown in another study on uracil DNA glycosylase (UNG), which excises the mis-incorporated uracil residue from DNA. When this base excision/ repair enzyme was inhibited, it led to hypermutation in HBV DNA and decreased cccDNA in the cells that co-express HBV and A3G. This suggested that uracil DNA glycosylase (UNG) removes the C-to-G mutations HBV DNA after A3G deamination, which may lead to an increase in viral replication (Kitamura, Wang *et al.* 2013).

DDX3 is a member of the highly conserved DEAD-box RNA helicases. DDX3 unwinds RNA in an ATPase-dependent manner. DDX3 has a role in RNA transcription regulation, splicing, export, and translation. It was reported that DDX3 inhibits HBV genome replication, and it was also shown that DDX3 interacts with HBV polymerase (Wang, Kim *et al.* 2009). In this study, DDX3 was used as positive control in using immunofluorescence to examine interactions between APOBEC3G protein and HBV polymerase.

1.2.4 Sequencing approach for the analysis of APOBEC3 proteins induced mutations in HBV DNA

Several studies, as described above, have shown that APOBEC3 proteins cause mutations in HBV DNA when the analysis is carried out using 3D-PCR. However, the exact frequency and extent of these APOBEC3-induced mutations in HBV DNA to inhibit HBV is still unclear. Next-generation sequencing (NGS) can give a more accurate measure of the frequency of mutations in HBV DNA caused by APOBEC3 proteins. The greater depth of NGS sequence information over traditional 3D-PCR sequencing methods allowed greater precision in determining the bias of particular mutations caused by APOBEC3 proteins. NGS could also be used as a diagnostic tool to show that APOBEC3 proteins act on the HBV DNA by revealing the mutation bias pattern typical to these restriction factors. This work will include the use of NGS to help demonstrate the effect of APOBEC3 proteins on the HBV virus *in vitro*. This would be important if the effects of APOBEC3 proteins on HBV DNA are tested in a future clinical setting to evaluate the possible role of APOBEC3 proteins in treatments for HBV infection.

1.2.5 Upregulation of APOBEC3 gene expression

APOBEC3 gene expression can be upregulated by type I (IFN-α) (Bonvin, Achermann *et al.* 2006) and type II interferons (IFN-γ) (Proto, Taylor *et al.* 2008), and lymphotoxin-β-receptor agonist (Lucifora, Xia *et al.* 2014). The upregulation of APOBEC3 proteins was examined by Zeng *et al.* , who showed that USF1 binds to the E-box in the promotor of the APOBEC3G gene, which subsequently leads to APOBEC3G upregulation (Zeng, Li *et al.* 2016).

1.3 Hypothesis & Objectives

HBV can be prevented by currently available vaccines that are considered safe and effective. In current treatments against HBV, nucleoside analogs (NA) can reduce virus titers. However, they cannot cure HBV infection (Seeger 2018). Although, as described, members of the APOBEC3 family of mutator proteins (e.g., A3B, A3F and A3G) have been shown to reduce the levels of HBV production in infected HBV culture systems it is not clear to what extent APOBEC3 proteins can target replicating HBV genomes in the cytoplasm. This project, therefore, investigated whether APOBEC3 proteins can directly mutate and lead to the destruction of HBV genomes in the cytoplasm in cell culture systems and whether the frequency and extent of these APOBEC3 mutations in HBV DNA may play a role in preventing the rebound of HBV infection following the cessation of reverse transcriptase inhibitor therapy and the clearance of HBV infection.

Aim: Study the effects of APOBEC3 proteins on HBV replication:

1) Investigate the impact of APOBEC3 family members on cytoplasmic and extracellular virion-associated HBV DNA levels.
2) Analysis of the frequency and extent of APOBEC3 mutations in HBV DNA.
3) Investigate interactions between APOBEC3G protein and HBV polymerase.
4) Investigate the effect of the Human Upstream stimulatory factor 1 (USF1) on APOBEC3 protein expression levels and its impact on HBV replication.

2 Chapter Two: Materials and Methods

2.1 Materials

2.1.1 Cell Culture

Cell culture medium:

500 ml of DMEM (Dulbecco's Modified Eagle's Medium)	Thermo Fisher Scientific	Cat. No: 41966-052
500 ml of RPMI-1640 (Roswell Park Memorial Institute)	Thermo Fisher Scientific	Cat. No: 21875-091
50 ml of Heat-Inactivated Fetal Bovine Serum	Thermo Fisher Scientific	Cat. No: 10082147
5 ml of Penicillin G Sodium (5.0 U/ml) + Streptomycin Sulphate (5.0 µg/ml in 0.85% saline)	Thermo Fisher Scientific	Cat. No: 15070063

Phosphate-Buffered Saline (PBS) 1x, PH 7.4, 500 ml	Thermo Fisher Scientific	Cat. No: 10010023
Trypsin-EDTA (0.05%), phenol red	Thermo Fisher Scientific	Cat. No: 25300054
DMSO: Dimethyl Sulphoxide	Sigma	Cat. No: 276855
Trypan Blue	Thermo Fisher Scientific	Cat. No: 15250061
Collagen, Type I solution from rat tail	Sigma	Cat. No: C3867

Selective antibiotics:

Geneticin disulfate salt solution (G-418 Sulphate)	Sigma	Cat. No: G8168
Doxycycline Hydrochloride, Ready-Made Solution	Sigma	Cat. No: D3072

2.1.2 Transformation and Culturing of Competent *E. coli* cells

$CaCl_2$ solution for the preparation of competent bacteria cells:

60 mM $CaCl_2$, 15% glycerol and 10 mM PIPES [piperazine-N,N′-bis(2-hydroxypropanesulfonic acid)], pH 7. Filter sterilize using a disposable filter unit or autoclave and store at room temperature (stable for years).

Competent *E. coli* Cells: (DH5α, Top10)	Thermo Fisher Scientific	Cat. No: K2040-01
IPTG Solution, ready-to-use	Thermo Fisher Scientific	Cat. No: R1171
X-Gal Solution, ready-to-use	Thermo Fisher Scientific	Cat. No: R0941

Luria Broth (LB): 1% (w/v) Bactotryptone, 0.5% (w/v) yeast extract and 1% (w/v) NaCl in de-ionised H2O (autoclaved at 120°C for 30 minutes).

LB Agar: 1.5% (w/v) Bacto agar in LB broth, and 500-1000 µl of the appropriate antibiotic.

Antibiotic Stock Solutions:

Antibiotics used for bacterial cell culture were purchased from Sigma. Antibiotics were made as stock solutions, filter sterilised (0.2 µm) and stored in aliquots at -20°C and before the addition of antibiotics to the autoclaved media, the media were cooled to 50°C.

Ampicillin	100 mg/ml in distilled water.
Kanamycin	30 mg/ml in distilled water.
Chloramphenicol	25 mg/ml in absolute ethanol.

2.1.3 Plasmid Purification and Extraction

QIAprep Spin Miniprep Kit (Qiagen, Cat No./ID: 27104)

QIAGEN Plasmid Plus Midi Kit (cat. nos. 12943)

2.1.4 Plasmid Transfection

FuGENE® HD Transfection Reagent	Promega	Cat. No: E2311
Lipofectamine 2000 Transfection Reagent	Thermo Fisher Scientific	Cat. No: 11668019
Opti-MEM™ I Reduced Serum Medium	Thermo Fisher Scientific	Cat. No: 31985047

2.1.5 Immunofluorescence

Pierce™ 16% Formaldehyde (w/v), Methanol-free, Thermo Fisher Scientific.

1x PBS (+), Thermo Fisher Scientific.

ProLong™ Gold Antifade Mountant, Thermo Fisher Scientific.

Slides Humidity Chamber (Fisher Scientific).

2.1.6 Western Blot Reagents

Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific).

RIPA lysis buffer (Sigma)

Dichlorodiphenyltrichloroethane (DDT, SIGMA)

Chameleon Duo Pre-Stained Protein Ladder (LI-COR, USA)

10x Running premixed electrophoresis Buffer (BIO-RAD,): Pkg of 1, 5 L cube, contains 25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3 following dilution to 1x with water before use.

4–15% Mini-PROTEAN TGX Precast Protein Gels (BIO-RAD)

Odyssey Blocking Buffer (PBS), 500 mL from (LI-COR, USA)

Odyssey CLx Imaging System (LI-COR).

2.1.7 Antibodies

Donkey anti-mouse secondary antibody conjugated with IRDye680LT (LI-COR)

Goat anti-rabbit secondary antibody conjugated with IRDye800CW (LI-COR).

Mouse monoclonal Monoclonal anti-FLAG® M2, Clone M2, Sigma (#F1804)

Mouse monoclonal anti-DDK (FLAG) monoclonal antibody, OriGene (#TA50011- 100)

Rabbit anti-TurboGFP Polyclonal Antibody, Thermo Fisher Scientific (#PA5-22688)

Mouse monoclonal anti-hepatitis B Virus Antibody, core Antigen (ayw), Sigma (#MAB16989)

Rabbit anti-C-MYC Antibody, Sigma

Rabbit anti-HA Antibody, Sigma (#H6908)

Mouse monoclonal anti-GAPDH antibody [6C5], Abcam (#ab8245)

2.1.8 Molecular Biology Reagents

HBV lysis buffer Composition:	100mM Tris-HCl, pH 8, and 0.2% (v/v) NP-40
5x HBV Buffer:	250mM Tris-HCl, pH 8, and 5% (w/v) SDS.
ice cold 100% ethanol	
ice cold 70% ethanol	

1M Magnesium Acetate	Sigma
1M $MgCl_2$	Sigma
DNase I (20 µg/µl)	Sigma
RNase A, DNase and protease-free (10 mg/mL)	Thermo Fisher Scientific
UltraPure™ 0.5M EDTA, pH 8.0	Thermo Fisher Scientific
20% SDS	Thermo Fisher Scientific
5 M NaCl	Sigma
Proteinase K (20 mg/ml)	Sigma
phenol: chloroform: isoamyl alcohol (125:24:1)	Sigma
3M Sodium Acetate	Sigma
Glycogen (20 µg/µl)	Thermo Fisher Scientific
UltraPure™ TBE Buffer, 10x	Thermo Fisher Scientific

TriDye™ 100 bp DNA Ladder (Size range: 100 bp to 1,517 bp)	New England Biolabs
Quick-Load® Purple 1 kb DNA Ladder (Size range: 500 bp to 10 kb)	New England Biolabs
SYBR Safe DNA Gel Stain	Thermo Fisher Scientific
UltraPure™ Agarose	Thermo Fisher Scientific

2.1.9 IFN Antiviral Assay Reagents

Interferon alpha-2b Solution: 1.8 x 10^7 U/ml solution was prepared from lyophilized IFN-α (Roche), aliquoted and stored at -80ºC.

2.1.10 Restriction Enzymes and Buffers

Restriction Enzymes and Buffers were purchased from New England Biolabs or Thermo Fisher Scientific.

2.1.11 Oligonucleotides

APOBEC3 qPCR primers:

Primer	Sequence (5' - 3')
APOBEC3A Forward	GACAATGGCACCTCGGTCAAGA
APOBEC3A Reverse	CCAACTGCAAAGAAGGAACCAGG
APOBEC3B Forward	CCTCTATGGTCGGAGCTACACT
APOBEC3B Reverse	GAGGAAGCACATTTCTGCGTGG
APOBEC3DE Forward	ACGTCAGTCGAATCACAGGCAG
APOBEC3DE Reverse	CAAACCAGGTGATCTGGAAGCG
APOBEC3F Forward	GGAATACCGTCTGGCTGTGCTA
APOBEC3F Reverse	GAGGAAGCACATTTCTGCGTGG
APOBEC3G Forward	ATGACACCTGGGTCCTGCTGAA
APOBEC3G Reverse	GAATCACGTCCAGGAAGCACAG
GAPDH Forward	CCCAATACGACCAAATCCG
GAPDH Reverse	TCTCTGCTCCTCCTGTTCGAC

HBV qPCR primers and probe:

Primers	Sequence (5' - 3')
HBV Forward	GTGTCTGCGGCGTTTTATCA
HBV Reverse	GACAAACGGGCAACATACCTT
HBV probe	FAM-CCTCT(T/G)CATCCTGCTGCTATGCCTCATC-TAMRA

Differential DNA Denaturation Polymerase chain reaction (3D-PCR) primers:

(Appendix A.8)

First Round Primers	**Sequence (5' - 3')**
HBV Xout Forward	CGCAAATATACATCGTATCCAT
HBV Xout Reverse	AAGAGTCCTCTTATGTAAGACCTT
HBV PreCout Forward	GTACTAGGAGGCTGTAGGCATA
HBV PreC-out Reverse	AGAGCTGAGGCGGTATCTAGAA
HBV Sout Forward	CGGCGTTTTATCATCTTCCTCTTCAT
HBV Sout Reverse	CATCCATATAACTGAAAGCCAAACAGT
Second Round Primers	**Sequence (5' - 3')**
HBV Xin Forward	ATGGCTGCTAGGCTGTGCTGCCAA
HBV Xin Reverse	AAGTGCACACGGTCCGGCAGAT
HBV PreCin Forward	TAAATTGGTCTGCGCACCAGCA
HBV PreCin Reverse	GATCTCGTACTGAAGGAAAGAA
HBV Sin Forward	TCTTCATCCTGCTGCTATGCCTCAT
HBV Sin Reverse	AAAGCCCTACGAACCACTGAACAAAT

Next-Generation Sequencing (NGS) primers ID : (Appendix A.9)

First Round Primers	**Sequence (5' - 3')**
HBV X Forward	TGTCCAGCACGCTTCAGGCTNNNNTNNNNTNNNNATGGCTGCTAGGCTGTGCTGCCAA
HBV X Reverse	AAGTGCACACGGTCCGGCAGAT
HBV PreC Forward	TGTCCAGCACGCTTCAGGCTNNNNTNNNNTNNNNTAAATTGGTCTGCGCACCAGCA
HBV PreC Reverse	GATCTCGTACTGAAGGAAAGAA
HBV S Forward	TGTCCAGCACGCTTCAGGCTNNNNTNNNNTNNNNTCTTCATCCTGCTGCTATGCCTCAT
HBV S Reverse	AAAGCCCTACGAACCACTGAACAAAT
Second Round Primers	**Sequence (5' - 3')**
Primer ID out Forward	TGTCCAGCACGCTTCAGGCT
HBV X Reverse	AAGTGCACACGGTCCGGCAGAT
HBV PreC Reverse	GATCTCGTACTGAAGGAAAGAA
HBV S Reverse	AAAGCCCTACGAACCACTGAACAAAT

2.1.12 Plasmids (Appendix A.1, A.2, A.3, A.4, and A.5):

Plasmids Name	Source
pDONR223-A3B	Dharmacon (#OHS5893-202504326)
pDONR223-A3F	Dharmacon (#OHS5893-202504153)
pTRIP.CMV.IVSB.iTagRFP-APOBEC3G	Was a gift from Prof. Charles M. Rice Lab
pTRIP.CMV.IVSB.iTagRFP-APOBEC3B	*
pTRIP.CMV.IVSB.iTagRFP-APOBEC3F	*
pcDNA3.1-VSVG	Was a gift from Prof. Charles M. Rice Lab
pcDNA3.1-HIVgag/pol	Was a gift from Prof. Charles M. Rice Lab
pCMV6-APOBEC3A-GFP	OriGene (RG220995)
pCMV6-APOBEC3B-GFP	OriGene (RG222712)
pCMV6-APOBEC3DE-GFP	*
pCMV6-APOBEC3F-GFP	OriGene (RC229185)
pCMV6-APOBEC3G-GFP	*
pCMV6-GFP (empty)	OriGene (PS100010)
pLenti-APOBEC3G-myc-DDK	OriGene (RC206821L1)
pCMV6-APOBEC3DE-myc-DDK	OriGene (RC229185)
pCMV6-myc-DDK (empty)	was a gift from Prof. Georgios Giamas, University of Sussex.
pGEM4Z-HBV1.3 mere (WT)	was a gift from Wang-Shick Ryu (Addgene plasmid # 65459)
pcDNA3-3XFlag-HBVpol	was a gift from Wang-Shick Ryu (Addgene plasmid # 65520)
pGEM4Z (empty)	Promega (product Cat No: P216)

pcDNA3-MYC-DDX3-HA	was a gift from Robin Reed (Addgene plasmid #44975)
pCMV3-USF1-myc	Sino Biological (Cat No: HG15301-CM)

* Constructed in this laboratory (Ibrahim Elbusifi, Department of Medicine, Imperial College London).

2.1.13 Chemicals and Plasticware

Plasticware and standard laboratory chemicals were purchased from Greiner Bio-one, VWR or Sigma, except where stated.

2.2 Methods

2.2.1 Construction of APOBEC3DE-GFP and APOBEC3G-GFP Mammalian Expression Vector

The shuttle cloning method was used to create APOBEC3DE, and APOBEC3G GFP tagged destination vectors. In this method, APOBEC3DE and APOBEC3G coding regions were transferred from entry vectors to destination vectors with C-terminal-tagged GFP. Origene company has developed entry (kanamycin selected) and destination vectors (ampicillin selected) containing two rare-cutting restriction endonucleases (*Sgf*I and *Mlu*I). Using *Sgf*I and *Mlu*I sites enables shuttling APOBEC3 protein-coding regions between vectors to create APOBEC3DE and APOBEC3G GFP tagged destination vectors that can be used for localisation studies.

2.2.2 Restriction Digestion of the APOBEC3 Coding ORF

0.5 µg of the shuttle vectors (containing APOBEC3DE or APOBEC3G) or 1 µg of the destination vector, 3 µl of the 10x restriction digest buffer (CutSmart Buffer, New England Biolabs), 0.6 µl *Sgf*I (20 units/ µl) (New England Biolabs), 0.6 µl *Mlu*I (20 units/ µl) (New England Biolabs) and distilled water to bring the final volume to 30 µl was loaded into an Eppendorf tube, mixed thoroughly and kept on ice. The tube's contents were quick-spun for several seconds at table centrifuge and then incubated for 3 hours at 37°C. Then, the tubes were incubated at 80°C for 15 minutes to inactivate the enzymes. Again, the contents of the tubes were quickly spun at the table centrifuge.

2.2.3 Restriction Digestion and Dephosphorylation of the GFP-tagged Destination Vector

2 µl (200 ng/µl) of the GFP tagged destination vector, 2 µl of the 10x restriction digest buffer (CutSmart Buffer, New England Biolabs), 0.6 µl SgfI (New England Biolabs), 0.6 µl *Mlu*I (New England Biolabs) and distilled water to bring the final volume to 20 µl were loaded into an Eppendorf tube, mixed thoroughly and kept on ice. The tube's contents were quick-spun for several seconds in a table centrifuge and then incubated for 3 hours at 37ºC. Then, 0.5 µl Alkaline Phosphatase, Calf Intestinal (CIP) (that non-specifically catalyzes the dephosphorylation of 5′ and 3′ ends of DNA and RNA phosphomonoesters) was added to the digestion reaction, and the tube was incubated at 37ºC for an additional 30 minutes. After that, the tube was incubated at 80°C for 15 minutes to inactivate the enzymes. Again, the contents of the tubes were quickly spun in a table centrifuge. The digestion was purified using a Qiaquick PCR purification Kit (Qiagen, Cat No./ID: 28104); the Qiagen PCR purification protocol was followed except for the elution, which was in 20 µl water.

2.2.4 Agarose Gel Electrophoresis of Digested Plasmids

10 µl of the digested samples was mixed with 2 µl of Gel Loading Dye, Purple (6x) (New England Biolabs). Then the digested samples and 5 µl of Quick-Load Purple 1 kb DNA Ladder (New England Biolabs) were run on an agarose gel. UltraPure™ Agarose gels were formed using a Horizon gel electrophoresis apparatus (Biorad). 0.8 - 2.0% agarose (w/v) in 1x UltraPure™ TBE buffer was melted in a microwave oven and allowed to cool down to 65°C when a 1x of SYBR Safe DNA Gel Stain was added and mixed before pouring it into the assembled apparatus, taking care not to introduce air bubbles. Once the gel had been set, it was covered with the appropriate UltraPure™ TBE buffer, and the combs were removed to reveal the wells. Generally, agarose of 1% (w/v) was used, but for separation of larger fragments (>5 kbp), 0.8 % (w/v) was used, and for smaller fragments (<1 kbp) up to 2% (w/v) was used. The

gel was run at a constant current of 60mA and 150 volts for 70-90 minutes or until the sample (as seen by the appearance of the tracking dye) had migrated the required distance along with the gel. Finally, the gel was visualised on a UV transilluminator, photographed, and the appropriate bands were excised.

2.2.5 Agarose Gel Extraction

The bands corresponding to the APOBEC3DE and APOBEC3G coding regions were excised from agarose gels. The DNA was extracted using a Qiaquick gel extraction kit (Qiagen, Cat No./ID: 28704) using the manufacturer's protocol. UV exposure time was kept to a minimum to prevent DNA damage. The excised gel slices were placed in Eppendorf tubes and weighed. Three volumes of buffer QG were added to 1 volume of the gel slice (100 mg ~ 100 µl). The tubes were incubated at 50ºC for 10 minutes with vortexing every 2-3 minutes during the incubation. After the gel slice was dissolved entirely, one gel volume of isopropanol was added and mixed thoroughly. The mixture was transferred to a QIAquick spin column, assembled in a 2 ml collection tube, and then centrifuged at 17,900 x g (13,000 rpm) for 1 minute at room temperature to let the lysate pass through the column. The flow-through was discarded, and the column was washed with 0.5 ml of buffer QG and centrifuged at 17,900 x g (13,000 rpm) for 1 minute at room temperature. Next, the column was washed with 0.75 ml of buffer PE, incubated for 5 minutes at room temperature, and centrifuged at 17,900 x g (13,000 rpm) for 1 minute at room temperature. The empty column was then centrifuged at 17,900 x g (13,000 rpm) for 1 minute to dry the column matrix, as it is essential to remove ethanol from the column. Finally, the column was placed in a clean 1.5 ml Eppendorf tube, 30 µl of sterile deionised water or EB buffer was directly added to the column matrix, and, after 1-minute incubation at room temperature, the tube was centrifuged at 17,900 x g (13,000 rpm) for 1 minute to elute the DNA.

2.2.6 DNA Ligations

The extracted APOBEC3DE and APOBEC3G coding region fragments were ligated into the pre-digested and dephosphorylated GFP-tagged destination vector. This was by adding 2 µl (33ng) of extracted DNA from the APOBEC3DE ORF clone or 2 µl (33ng) of extracted DNA from the APOBEC3G ORF clone with 1 µl of 10x ligation buffer, 3 µl (20ng) of the pre-digested and dephosphorylated GFP tagged destination vector, 1 µl of ExpressLink™ T4 DNA Ligase, and 3 µl of distilled water in an Eppendorf tube, mixed gently and incubated at 25°C overnight for two hours. The ligation products were used to transform *E. coli* (Top10), and the transformation reaction was used to prepare 1:10 and 1:100 serial dilutions. These dilutions were plated onto ampicillin LB plates (100µg/ml ampicillin). After overnight growth, four colonies resistant to ampicillin were picked from each plated reaction, grown in 5 ml of LB and plasmid DNA was subsequently extracted from the bacteria using a QIAprep Spin Miniprep Kit (Qiagen, Cat No./ID: 27104) and sent for sequencing.

2.2.7 DNA Sequencing of Plasmid Inserts

Sequencing the cloned APOBEC3 expressing plasmids was carried out commercially by sending the plasmids to the GENEWIZ sequencing company in London, UK, using the T7 sequencing primers. In a 0.2 ml tube, 200-500 ng of plasmid DNA and 5 µl of 5 pmol/ µl (µM) of sequencing primer were mixed with distilled water to bring the final volume to 10 µl.

2.2.8 Plasmid Purification and Extraction

2.2.8.1 Bacterial Culture Techniques

Chemically competent cells were either prepared or purchased and used to transform plasmids and produce a high quantity of these plasmids.

2.2.8.2 Preparation of competent bacterial cells

Chemically competent *E. coli* (DH5α) was prepared as follows: first, *E. coli* (DH5α) cells were plated onto a prewarmed 10 cm LB agar plate (without antibiotic) and incubated overnight at 37ºC. The next day, a single colony of *E. coli* cells was inoculated into 50 ml LB medium and was grown overnight at 37°C with moderate shaking (250 rpm). 4 ml of the culture were inoculated into 400 ml LB medium in a sterile 2-litre flask and was grown at 37°C, shaking at 250 rpm, to an OD590 of 0.375. The culture was aliquoted into eight 50-ml prechilled, sterile polypropylene tubes and left on ice for 5 to 10 min. Immediately after, the cells were centrifuged for 7 min at 1600 × g, 4°C. The supernatants were poured off, and each pellet was resuspended in 10 ml ice-cold $CaCl_2$ solution (Resuspension should be performed very gently, and all cells should be kept on ice). The cells were centrifuged for 5 min at 1100 × g, 4°C. The supernatants were discarded, and each pellet was resuspended in a 10 ml ice-cold $CaCl_2$ solution. The resuspended cells were kept on ice for 30 min. Then, the supernatants were discarded, and each pellet was resuspended in a 2 ml ice-cold $CaCl_2$ solution. The cells were dispensed into prechilled, sterile polypropylene tubes (250 µl aliquots) and immediately stored at −80°C. The next day, one of these aliquots was transformed with a known Plasmid and plated in LB agar to test the competent cell.

2.2.8.3 Transformation of Competent E. coli cells

Competent cells were used to transform plasmids and produce a large quantity of these plasmids. In this process, the product of plasmid DNA cloning and existing APOBEC3 plasmids or HBV plasmids were transformed into competent bacterial cells. First, the tube containing 50 µl of competent cells was taken out of -80ºC and thawed on ice for approximately 30 minutes. Then, 2-4 µl of plasmid were added into 50 µl of competent bacterial cells and gently mixed by flicking the bottom of the tube with the tip of the finger a few times. The competent cell/DNA mixture was incubated on ice for 30 minutes. Each transformation tube was heat-shocked by placing the bottom 2/3 of the tube into a 42°C heat block. Immediately after, the transformation tube was put on ice for 3 minutes. Then, 250 µl of pre-heated LB Broth media (without antibiotic) was added to the bacteria and was grown in a 37°C shaking incubator (225rpm) for 1 hour. Two volumes of the transformation were then plated; Two hundred and seventy microliters and thirty microliters (diluted 1:10) were plated onto a prewarmed 10 cm LB agar plate containing the appropriate antibiotic (Ampicillin, Kanamycin, etc.). These LB agar plates were incubated overnight at 37ºC. The next day, a single colony was picked up to amplify these plasmids using Mini-prep and Midi-prep.

2.2.8.4 Plasmid Purification and Extraction (Mini-Prep)

The picked-up colonies from the transformed competent bacterial cells were grown in 5 ml of LB media containing an appropriate amount of selective antibiotic and grown overnight at 37°C in an orbital shaker (225 rpm). The next day, 0.5 ml of the overnight-grown bacteria culture was used for glycerol stock. This was done by mixing 0.5 ml of 50 % glycerol solution with 0.5 ml of the overnight-grown bacteria culture and storing it at -80°C. QIAprep Spin Miniprep Kit (Qiagen, Cat No./ID: 27104) was used for Miniprep extraction, 4 ml of the overnight grown bacteria culture were pelleted by centrifugation at >8000 rpm (6800 x g) for 15 minutes at room temperature. The pelleted bacterial cells were resuspended in 250 µl Buffer

P1 (containing added LyseBlue reagent and mixed with RNase A solution) and transferred to a microcentrifuge tube. 250 µl Buffer P2 was added and mixed thoroughly by inverting the tube 4–6 times until the solution became blue. After less than 5 minutes of incubation with P2 lysis buffer, 350 µl Buffer N3 were added and mixed immediately and thoroughly by inverting the tube 4–6 times until the solution turned colourless. The mixture was centrifuged for 10 min at 13,000 rpm (~17,900 x g) in a table-top microcentrifuge. 800 µl of the supernatant from the previous step was applied to the QIAprep 2.0 spin column by pipetting, centrifuged for 60 seconds, and the flow-through was discarded. The QIAprep 2.0 spin column was washed by adding 0.5 ml Buffer PB, centrifuged for 60 seconds and the flow-through was discarded. The washing step of the QIAprep 2.0 spin column was repeated by adding 0.75 ml Buffer PB, centrifuging for 60 seconds, and the flow-through was discarded. The QIAprep 2.0 spin column was centrifuged for 1 minute to remove residual wash buffer. The QIAprep 2.0 column was placed in a clean 1.5 ml microcentrifuge tube. The plasmid DNA was eluted by adding 50 µl Buffer EB (10 mM TrisCl, pH 8.5) or water to the center of the QIAprep 2.0 spin column; it was let stand for 1 min and was centrifuged for 1 min at 13,000 rpm (~17,900 x g). 2 µl of the purified plasmid DNA were analysed by agarose gel electrophoresis on a 1 % agarose gel. Plasmid DNA concentration was measured using a Nanodrop 1000 spectrophotometer. Then, the purified plasmid DNA was stored at -20°C.

2.2.8.5 Plasmid Purification and Extraction (Midi-prep)

The picked-up colonies from the transformed competent bacterial cells were grown in 2 ml of LB media containing an appropriate amount of selective antibiotic and grown for 8 hours at 37°C in an orbital shaker (225 rpm). 250 µl of the grown bacteria culture were inoculated (at dilution of 1:1000) into 250 ml of LB media containing an appropriate amount of selective antibiotic and grown overnight at 37°C in an orbital shaker (225 rpm). The QIAGEN Plasmid Plus Midi Kit (cat. nos. 12943) was used for Plasmid purification, 25-50 ml of the overnight

grown bacteria culture were pelleted by centrifugation at 6000 x g for 15 min at 4°C. The pelleted bacterial cells were resuspended in 4 ml of Buffer P1 (containing added LyseBlue reagent and mixed with RNase A solution) and transferred to a microcentrifuge tube. 4 ml Buffer P2 were added and mixed thoroughly by inverting the tube 4–6 times until the solution became blue. The QIAfilter Cartridge was placed into a new and suitable tube, allowing space for the addition of Buffer BB. After no more than 3 minutes of incubation with P2 lysis buffer, 4 ml Buffer N3 were added and mixed immediately and thoroughly by inverting the tube 4–6 times until the solution turned colourless. The lysate was transferred to an assembled QIAfilter Cartridge and incubated at room temperature for 10 min. During incubation, QIAGEN Plasmid Plus spin columns were placed into the QIAvac 24 Plus and Tube Extenders were inserted into each column. The plunger was gently inserted into the QIAfilter Cartridge, and the cell lysate was filtered into the tube. 2 ml Buffer BB were added to the cleared lysate and mixed by inverting 4–6 times. The lysate was transferred to a QIAGEN Plasmid Plus spin column on the QIAvac 24 Plus. Approximately –300 mbar vacuum was applied until the liquid had been drawn through all columns. To wash the DNA, 0.7 ml Buffer ETR was added, and the vacuum was applied again until the liquid had been drawn through all columns. To further wash the DNA, 0.7 ml Buffer PE was added, and the vacuum was applied again until the liquid had been drawn through all columns. The columns were centrifuged at 10,000 x g (9,700 rpm) for 1 minute to remove the residual wash buffer completely. The QIAGEN Plasmid Plus spin columns were placed into clean 1.5 ml tubes. The DNA was eluted by adding 200 µl Buffer EB or water to the center of the QIAGEN Plasmid Plus spin column and was let stand for ≥1 min. It was centrifuged for 1 min. 2 µl of the purified plasmid DNA were analysed by agarose gel electrophoresis on a 1 % agarose gel. Plasmid DNA concentration was measured using a Nanodrop 1000 spectrophotometer. Then, the purified plasmid DNA was stored at -20°C.

2.2.9 Mammalian Cell Culture Techniques

HepG2, HepG2.2.15, HEK293T, and HEK29A were used in this project. HepG2 is a hepatoblastoma cell line (ATCC® HB-8065™). HepG2.2.15 cells are HepG2 cells with head-to-tail HBV genome integrated into the cellular genome, constitutively producing HBV virus (Sells, Chen *et al.* 1987). In this project, HepG2 cells were either used as a control for HBV producing cell line (HepG2.2.15) or transiently transfected with HBV-expressing plasmids and APOBEC3-expressing plasmids. HEK293 cell lines are human embryonic kidney cancer cells that are easily transfectable and used for testing the expression and localisation of plasmids (Graham, Smiley *et al.* 1977).

2.2.9.1 Maintenance and Handling

HepG2, HepG2.2.15, HEK293T, and HEK29A cells were propagated using a complete culture medium (RPMI or DMEM) plus 10% Fetal Bovine Serum and 1% of Penicillin G Sodium (5.0 U/ ml) + Streptomycin Sulphate (5.0 µg/ ml in 0.85% saline). Complete RPMI medium was used for HepG2 and HepG2.2.15 cells, while Complete DMEM medium was used to propagate HEK293T and HEK29A cells.

2.2.9.2 Cell thawing and sub-culturing:

Thawing cells were done by removing cryovials from liquid nitrogen storage tanks and placing them in a 37°C water bath for 30 seconds. Immediately after, cells were transferred to 15 ml Falcon tubes containing 10 ml of the pre-warmed medium. Cells were centrifuged at 500 x g (Thermo ALC PK130R Refrigerated Centrifuge). The supernatant was discarded, and cell pellets were resuspended in a 5 ml pre-warmed medium. The 5 ml of cells were then transferred to 25 cm^2 flasks. The flasks were rocked gently and placed in a 37°C incubator overnight. The next day, the cell was checked under an inverted phase microscope. The medium was discarded and replaced with a fresh medium. The flasks were placed in a 37°C incubator for another two days, and after that, cells were sub-cultured routinely every two to three days; HepG2 and

HepG2.2.15 cells were sub-cultured at a dilution of 1:4, while HEK293T and HEK29A cells were sub-cultured at a dilution of 1:10.

2.2.9.3 Cell Freezing and long-term storage

Cell freezing medium was prepared (70% DMEM, 20% FBS and 10% DMSO). Near confluent T175 cm^2 flasks of cells were used for the freezing. The cells were washed with 3 ml of PBS (Gibco, Invitrogen, USA). 3 ml of 0.05% trypsin was added for each T175 cm^2 flask. The cells were incubated at 37°C for 2-3 min. Immediately after the cells became detached, one volume (3 ml) of fresh complete DMEM medium was added. The mixtures were then transferred to a 15 ml tube. The concentration and viability of cells were assayed using the Trypan blue exclusion method, where trypan blue is taken up only by non-viable cells. 10 µl of 0.4% trypan blue was added to 10 µl of cell suspension and incubated for 5 minutes to allow dye uptake. Then the concentration and viability of the cells were analysed using a disposable haemocytometer and an inverted phase microscope. 10 µl of the cells were added to the haemocytometer. Under the 10x objective lens, the blue-stained cells and the total number of cells were counted in the large, central gridded square (1 mm^2). The concentration and viability of cells were calculated as follows: The number of viable cells per ml of culture = number of viable cells in the large, central gridded square x 10^4 x dilution factor. % viable cells = [1.00 - (Number of blue cells ÷ Number of total cells)] x 100. Then the cells were centrifuged at 500 x g (Thermo ALC PK130R Refrigerated Centrifuge) for 5 minutes. The supernatant was discarded, and the cell pellets were re-suspended in 10 ml of freezing medium. 1 ml of cells in the freezing medium was dispensed into each cryovial. Finally, these cryovials were frozen at -80˚C for 24 hours. The next day, these cryovials were transferred to liquid nitrogen for long-term storage.

2.2.10 Fugene HD Mediated plasmid Transfection

HepG2 and HepG2.2.15 cells were seeded into 6-well tissue culture plates at a density of $1x10^5$ cells per well in complete media (2 ml) and incubated at 37°C in a 5% CO_2 for 24-48 hours until they reached 60-70% confluency. Cells were then transfected using Fugene HD: for each well, 2 µg of plasmid was added to 100 µl of OptiMem 1 reduced serum medium in an Eppendorf tube and then vortexed. 6 µl of Fugene HD reagent was added to the mixture at a ratio of (3:1 FuGENE® HD Transfection Reagent: DNA); the mixture was mixed by pipetting (7 to 10 times) and then incubated for 15 minutes at room temperature. Immediately after the incubation, 100 µl of transfection mixture was added dropwise to each well containing cells. The tissue culture plate was then rocked gently for a few seconds and placed at 37°C in a 5% CO_2 incubator. The transfection medium was replaced with a complete fresh medium 6-8 hours post-transfection, and the cells were incubated again at 37°C in 5% CO_2 for 24-72 hours. Once the cells had been transfected in the above manner, several different assays were performed to test the effect of the transfected genes on the cells.

2.2.11 Detection of plasmid Expression by Immunofluorescence

For cell preparation for immunofluorescence experiments, 16 x 16 mm coverslips (sterilized by dipping in 70% ethanol) were placed in each well of a 12-well tissue culture plate. 50-100 thousand HepG2, HepG2.2.15, or HEK293A cells were seeded per coverslip containing a well of a 12-well plate in 1 ml of culture medium and incubated at 37°C in a 5% CO_2. The next day, cells were transfected with 0.5-1 µg of each plasmid. The transfection medium was replaced with a complete fresh medium 6-8 hours post-transfection. 48-72 hours post cell transfection, cells were washed 2 times using 500 µl of 1x PBS (+). To Fix the cells, 500 µl of 4% paraformaldehyde (pH 7.4) was added to each well and incubated at 37°C for 15 min. The fixation solution was removed, and the cells were washed 3 times with 500 µl of 1x PBS (+). Permeabilization of cells was done by adding 500 µl of (0.5% Triton x-100 in 1x PBS) to each

well, and the cells were incubated at room temperature for 10min on a slow speed shaker. After removing the permeabilization solution, cells were washed 3 times with 500 µl of 1x PBS (+). The blocking step was done by adding 500 µl of 10% FBS in 1x PBS and incubating the cells at room temperature for 60 min. Immunostaining of cells was done by adding 50-100 µl of the diluted1:5000 Rabbit anti-GFP primary antibody (Thermo Fisher Scientific), diluted; 1:200 Mouse anti-Flag primary antibody (Sigma) and /or diluted; 1:2000 Rabbit anti-HA primary antibody (Sigma) to each coverslip of the cells and incubated for 1 hour at room temperature or overnight at 4°C. However, only a Blocking buffer was added to the controls. After removing the primary antibody solution, the cells were washed three times with 500 µl of 1x PBS (+). Then, 50-100 µl of the diluted: 1:500 Donkey anti-mouse Alexa Fluor 594 secondary antibodies (Thermo Fisher Scientific) and/or the diluted: 1:500 Donkey anti-rabbit Alexa Fluor 488 secondary antibodies (Thermo Fisher Scientific) along with adding DAPI (diluted 1 in 500 µl 10% FBS in 1x PBS). The secondary antibodies were incubated for 45 minutes at room temperature and protected from light. Cells were washed 3 times with 500 µl of 1x PBS (+). The mounting step was done by adding two drops of mounting medium (ProLong™ Gold Antifade Mountant, Thermo Fisher Scientific) to the slide (one sample coverslip and one control coverslip). The coverslips (protected from light) were air-dried at room temperature overnight and visualize under Fluorescence microscopy or stored at 4°C until use.

2.2.12 Detection of plasmid Expression by Western Blotting

Cells were seeded and transfected in 6-well plates and harvested at 48h post-transfection. 200 µl RIPA lysis buffer (Sigma) with 1:1000 dilution of the Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific) was used to extract protein. The lysate was spun at 13,000 rpm, 4°C for 10min to separate protein and cell debris. 5x loading buffer and 1:20 dilution of Dichlorodiphenyltrichloroethane (DDT, SIGMA) were added to the supernatant and incubated at 95°C for 5 min. 15 µl of sample and 5 µl of the Chameleon Duo Pre-stained

Protein Ladder (LI-COR, USA) were loaded in the 4–15% Mini-PROTEAN TGX Precast Protein Gels (BIO-RAD). The gel was run at 50V for 15 min and then at 200V for 25 min. The proteins in the gel were transferred to a nitrocellulose membrane at 100V for 1h. The membrane was washed with PBS once and blocked in 5% milk for 1h. Then the membrane was washed with PBS for 5 min 3 times, incubated in 2% milk with 1:5000 dilution of the Rabbit anti-GFP primary antibody (Thermo Fisher Scientific) and 1:1000 dilution of the Mouse monoclonal anti-GAPDH antibody (Abcam) overnight. The next day, the membrane was washed with PBS for 5 min 3 times and incubated in 2% milk with 1: 20,000 dilution of the donkey anti-mouse secondary antibody conjugated with IRDye680LT (LI-COR) and 1: 15,000 dilution of the goat anti-rabbit secondary antibody conjugated with IRDye800CW (LI-COR). Then the membrane was washed with PBS for 5 min 3 times and dried with filter papers before being analyzed. It was visualized using the Odyssey CLx Imaging System (LI-COR).

2.2.13 RNA Extraction

Total RNA from HepG2, HepG2.2.15, and HEK293 cells was extracted using Qiagen RNeasy Plus Mini Kit (Cat No./ID: 74134). RLT buffer was prepared by adding 10 µl β-mercaptoethanol (β-ME) per 1 ml of Buffer. 350 µl of Buffer RLT Plus was added to each well of cells in a 6-well plate and incubated for 1-2 minutes, and then the lysate was then transferred to an Eppendorf tube and put on ice. The homogenized lysate was transferred to a gDNA Eliminator spin column placed in a 2 ml collection tube and centrifuged for 30 s at ≥8000 x g (≥10,000 rpm). The column was discarded, and the flow-through was saved. One volume (350µl) of 70% ethanol was added to the flow-through and mixed well by pipetting. Immediately after, up to 700 µl of the sample, including any precipitate, was transferred to a RNeasy spin column placed in a 2 ml collection tube and centrifuged for 15 seconds at ≥8000x g. The flow-through was discarded. 700 µl Buffer RW1 was added to the RNeasy Mini spin

column (in a 2 ml collection tube) and centrifuged for 15 seconds at ≥8000 x g. The flow-through was discarded. 500 µl Buffer RPE was added to the RNeasy spin column and centrifuged for 15 seconds at ≥8000 x g. The flow-through was discarded. Another 500 µl Buffer RPE was added to the RNeasy spin column and centrifuged for 2 min at ≥8000 x g (≥10,000 rpm). Then, the RNeasy spin column was placed in a new 2 ml collection tube and centrifuged at full speed for 1 min to dry the membrane further. The RNeasy spin column was placed in a new 1.5 ml collection tube, and 30µl RNase-free water was directly added to the spin column membrane. Then, the tubes were centrifuged for 1 min at ≥8000 x g to elute the RNA. Then, the RNeasy spin column was discarded, and the 1.5 ml collection tube containing the eluted RNA was put on ice for subsequent use or DNase treatment.

Possible genomic DNA contamination was removed by treating the purified RNA samples with DNA-free Turbo DNase (Ambion). This was done by adding 0.5 µl of DNA-free Turbo DNase (2 units/µl), 2 µl of 10x DNase buffer and 8 µg of RNA and made up to a final reaction volume of 20µl total with nuclease-free water. These samples were incubated at 37˚C for 30 minutes. Then, 1µl of kit inactivation reagent was added to each sample to inactivate and remove the Turbo DNase, and the samples were incubated at room temperature for 5 minutes. Next, the samples were centrifuged at 10,200 rpm (ca.10,000 xg) for 1.5 minutes, and each supernatant was transferred to a new 1.5 ml Eppendorf tube. Finally, the RNA concentrations of the DNase-treated RNA samples were measured using a Nanodrop 1000 spectrophotometer. The RNA samples were stored at -80ºC.

2.2.14 cDNA Reverse Transcription

cDNAs were synthesized by using a High-Capacity cDNA Reverse Transcription Kit (Thermofisher Scientific, Catalog number: 4368814). For each reaction, 2x reverse transcription master mix was prepared as follows: 10 µl of the reagent master mix (2 µl of 10x Reverse Transcription buffer, 0.8 µl of 25x dNTP Mix (100 mM), 2 µl 10x RT Random

Primers, 1 µl MultiScribeTM Reverse Transcriptase and 4.2 µl nuclease-free H_2O). 2 µg of DNase treated RNA was added to the reverse transcription reaction. The tube was briefly centrifuged to spin down the contents and eliminate air bubbles. Reverse transcription was performed in a thermal cycler (10 minutes at 25°C, 120 minutes at 37°C, 5 minutes at 85°C and then at 4°C). The cDNA samples were stored at -20 °C.

2.2.15 Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR)

Quantitative Polymerase Chain Reaction was carried out using Applied Biosystems ViiA™ 7 Real-Time PCR System; 2 µl (20 ng/µl) of diluted cDNA was mixed with 1 µl of reverse primer (300 nM), 1 µl of forward primer (300 nM) and 10 µl PowerUp™ SYBR™ Green Master Mix (Invitrogen, A25741) and 8 µl of nuclease-free water. RT-qPCR runs were performed using GAPDH as a reference gene. Pre-designed qRT PCR primers were used and blasted to check their specificity to the target genes using NCBI primer blast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/), after which primers annealing temperature were calculated and optimized by qRT-PCR, then used to measure the gene expression of target genes. Then the results from qRT-PCR were analyzed as relative gene expression for the housekeeping gene GAPDH using the comparative CT method (Delta Delta CT) to analyze real-time PCR data (Livak and Schmittgen 2001).

Delta Delta CT equation:

ΔCt = Ct (Target A -treated) – Ct (Ref B-treated)

ΔCt = Ct (Target A-control) – Ct (Ref B-control)

$\Delta\Delta$ Ct = Δ Ct (treated) – Ct (control)

Fold change = $2(-\Delta\Delta Ct)$

GraphPad Prism version 6 was used to analyze the qPCR data.

2.2.16 Extraction of cytoplasmic Core-associated HBV DNA

HepG2.2.15 cells were grown in 6-well tissue culture plates until 60-80% confluent. 3 wells were transfected with APOBEC3 plasmids for 72-96 hours. Then, the cells in each well were washed with 2 ml PBS, trypsinized, transferred to Eppendorf tubes, and centrifuged at 1500 rpm for 5 minutes. The supernatant in each Eppendorf tube was discarded, and these tubes containing the pellets were either stored at -80˚C or the extraction was continued.

The pellets were re-suspended in 1 ml PBS each. For the extraction, 800µl of this suspension was utilized. Next, the cells were centrifuged at 1500 rpm for 5 minutes to form pellets. These cell pellets were lysed with 400 µl of the HBV lysis buffer (100 mM Tris-HCl, pH 8, and 0.2% (v/v) NP-40) for one minute at room temperature. Then, the cell lysates were centrifuged at 14000 rpm for 1 minute and the supernatants were transferred to a new Eppendorf tube. After that, 2.7 µl of 1 M Magnesium Acetate, 5 µl of 20 µg/µl DNase I and 5 µl of 10 µg/µl RNase A were added to each one, then mixed and incubated for 2 hours at 37°C.

Then, the mixtures were centrifuged at 14000 rpm for one minute, and the supernatants were transferred to new Eppendorf tubes. Next, 10 µl of 0.5 M EDTA and 20 µl of 20% SDS were added to each Eppendorf tube containing the transferred supernatant and mixed by pipetting. Then, 10 µl of 5 M NaCl and 7.5 µl 20mg/ ml Proteinase K were added to each one and mixed, then incubated at 55°C for one hour. After this, a 460 µl of a mixture of phenol: chloroform: isoamyl alcohol (125:24:1) was added to each tube and mixed by vortexing (until the formation of white emulsion) and centrifuged at 13000 rpm for 5 minutes at room temperature. After that, 400 µl of the upper aqueous phase was collected in a new Eppendorf tube, and 40 µl of Sodium Acetate, 1 µl of 20 µg/µl glycogen and 800 µl of ice-cold 100% ethanol were added, and the tube then was incubated at -20°C for two hours.

The tube was centrifuged at 14000 rpm for 20 minutes at 4°C. The supernatant was removed, and each pellet was washed with 1000 µl of ice-cold 70% ethanol and dried using a vacuum

desiccator. The pellet was then re-suspended in 20 µl of sterile deionized water. Lastly, 1 µl of 2 µg/µL RNase A was added to each sample, and the samples were incubated at 37°C for 30 minutes. Then, core-associated HBV DNA extracts were stored at -20°C.

2.2.17 Extraction of supernatant virion-associated HBV DNA

HepG2.2.15 cells were grown in 6-well tissue culture plates until they reached 60-80% confluency; then, the cells were transfected with APOBEC3 plasmids. Eight- or twenty-four-hours post-transfection, the medium was removed and replaced with a fresh complete culture medium. Four days post-transfection, the culture medium was collected and centrifuged at 1500 rpm for 5 minutes to remove cellular debris. A 400 µl of the clarified supernatant was transferred to an Eppendorf tube. Then, 1.6 µl of 10 µg/µl DNase I and 4 µl of 1 M $MgCl_2$ were added to the clarified supernatant and incubated for 2 hours at 37°C. 25 µl of 20 mg/ ml Proteinase K was mixed thoroughly. After that, 100 µl of 5x HBV Buffer was added and incubated at 56°C for 30 minutes. The tubes were left to cool down to room temperature. Under the fume hood, 530 µl of a mixture of phenol: chloroform: isoamyl alcohol (125:24:1) was added to each tube and mixed well by vortexing (until a white emulsion was formed) and centrifuged at 13000 rpm for 5 minutes at room temperature. About 400 µl of the upper aqueous phase was transferred to a new Eppendorf tube, and 40 µl of 3 M Sodium Acetate, 1 µl of 20 µg/µl glycogen and 800 µl of ice-cold 100% ethanol was added to it and incubated at -20°C for two hours. After that, the mixture was centrifuged at 13000 rpm for 30 minutes at 4 °C. The supernatant was removed, and the pellet was washed twice with 1000 µl of ice-cold 70% ethanol. Then, the tubes were dried using a vacuum desiccator, and the pellet was resuspended in 20 µl of sterile deionized water. Finally, 1 µl of 2 µg/µl RNase A was added and incubated at 37°C for 30 minutes. The HBV DNA was stored at -20°C.

2.2.18 Quantification of HBV DNA by Real-Time PCR

Quantitative Polymerase Chain Reaction (qPCR) was carried out to quantify the extracted HBV DNA to elucidate the effect of APOBEC3 proteins on HBV replication. Quantitative Polymerase Chain Reaction (qPCR) was done using a TaqMan™ Universal PCR Master Mix (Thermo Fisher, Catalog number: 4304437). TaqMan qPCR has high sensitivity and high sequence specificity in comparison to SYBR Green qPCR, but TaqMan qPCR is more expensive.

HBV DNA standards were prepared by quantifying the HBV plasmids copy number using the number copies using the formula:

Number of copies = (amount of dsDNA in nanograms * $6.022x10^{23}$) / (length of dsDNA in base pairs * $1x10^{9}$ * 650)

Then, in small tubes with caps, 2E9 HBV plasmid copies were used to make 1:10 dilutions of the standard in nuclease-free water (10 µl standard in 90 µl nuclease-free water) to create standards that range from 1E9 to 1E1 or 1E0 HBV copies/reaction. The HBV standards were stored at -20°C. Also, TaqMan Assay 20x (forward primer [0.4 µM], reverse primer [0.4 µM] and probe [0.2 µM]) was prepared and put at -20°C.

HBV DNA extracted from cytoplasm/supernatant was amplified and quantified using ViiA™ 7 real-time PCR instrument (Applied Biosystems, CA, USA). Plate Preparation: Each plate should contain triplicate HBV standard and NTC (no template control) wells and the "unknown" samples. 15 µl of the master mix was added to each well of the 96-well plate, followed by 5 µl of respective HBV standards, NTC or unknown samples. An Optical Adhesive cover was placed over the plate and sealed. The plate was spun for 1 minute at 1500 rpm and inserted into the instrument. An initial two minutes of incubation at 50°C to activate uracil-DNA-glycosylase (UNG) and destroy potential carry-over amplicons was followed by 10 min

at 95ºC to activate the thermostable DNA polymerase—subsequently, 40 cycles of 95˚C for 15 seconds and 60˚C for 1 minute.

2.2.19 Differential DNA Denaturation Polymerase chain reaction (3D-PCR)

Differential DNA Denaturation PCR (3D-PCR) was used to identify hypermutated HBV genomes. This 3D-PCR method was carried out in a two-round procedure (nested PCR) (Suspene, Guetard *et al.* 2005, Suspène, Henry *et al.* 2005).

2.2.19.1 Primers Design

Fragments of the HBV genome (Precore, X and S regions) were amplified. A hot start PCR was performed to increase the sensitivity and specificity of this method. The first-round primers were External Forward (ayw) and External Reverse (ayw) for each region (Precore, X and S) of the extracted HBV DNA from HepG2.2.15 cell. The second-round primers were Internal forward (ayw) and Internal Reverse (ayw) for each region (Precore, X and S) of the extracted HBV DNA from HepG2.2.15 cells (Appendix A.8).

2.2.19.2 Taq polymerase-based 3D-PCR

2.2.19.2.1 Nested PCR reaction

The first reaction of the nested PCR involved standard amplification; the reaction conditions were 95°C for 5 min, followed by 35 cycles (95˚C for 30 s, 55˚C for 30 s, and 72˚C for 30 s), and finally for 10 min at 72°C for the first round. The cycling conditions for the second round were 90°C or 89°C for 5 min, followed by 35 cycles (90°C or 89°C for 60 s, 45°C for 30 s, and 72°C for 30 s), and finally 10 min at 72°C. The input amount used for the differential amplification that occurred in the second round was 0.5 µl of the first-round PCR. Immediately after, the amplified fragments of the HBV genome (Precore, X and S regions) were capture-cloned into the linearized shuttle vector pCR2.1 (TA cloning kit, Invitrogen).

2.2.19.2.2 Capture-Cloning into Transfer Vector pCR2.1

The amplified fragment of the HBV genome (Precore, X and S regions) was capture-cloned into the linearized transfer vector pCR2.1 (TA cloning kit, Invitrogen, Cat No. K2020-20). Around (1.5:1) insert to vector ratio was used whereby 3 µl of the fresh PCR product was mixed with 1 µl of 10x ligation buffer, 2 µl of pCR2.1 vector (25 ng/µl), 1 µl of ExpressLink™ T4 DNA Ligase, and 3 µl of distilled water in an Eppendorf tube, mixed gently and incubated at 25°C overnight for two hours. The ligation products were used to transform *E. coli* (Top10), and the transformation reaction was used to prepare 1:10 and 1:100 serial dilutions. These dilutions were plated on 50 µg/mL kanamycin plates containing x-Gal and IPTG and were incubated overnight at 37°C. This method of using plates containing x-Gal and IPTG allows distinguishing between the colonies containing the pCR2.1 plasmid and insert from a circularised plasmid containing colonies by blue/white screening. The next day, 10 white colonies of each reaction were picked up and prepared for Miniprep extraction. The transfer vector was extracted using a QIAprep Spin Miniprep Kit (Qiagen, Cat No./ID: 27104) and sent for sequencing.

2.2.19.3 Phusion U polymerase-based 3D-PCR

2.2.19.3.1 Nested PCR reaction

The first reaction of the nested PCR involved standard amplification; the reaction conditions were 95°C for 5 min, followed by 35 cycles (98°C for 30 s, 61°C for 30 s, and 72°C for 30 s), and finally for 10 min at 72°C for the first round. The cycling conditions for the second round were 90°C or 89°C for 5 min, followed by 35 cycles (98°C or 89°C for 60 s, 61°C for 30 s, and 72°C for 30 s), and finally 10 min at 72°C. The input amount used for the differential amplification that occurred in the second round was 0.5 µl of the first-round PCR. Immediately after, the amplified fragments of the HBV genome (Precore, X and S regions) were capture-

cloned into the linearized shuttle pCR-Blunt (Zero Blunt PCR Cloning Kit, Invitrogen, Cat No. K2700-40).

2.2.19.3.2 Capture-Cloning into Transfer Vector pCR-Blunt

The amplified fragment of the HBV genome (Precore, X and S regions) was capture-cloned into the linearized transfer vector pCR-Blunt (Zero Blunt PCR Cloning Kit, Invitrogen, Cat No. K2700-40). Around (1.5:1) insert to vector ratio was used whereby 3 µl of fresh Blunt PCR product was mixed with 1 µl of 10x ligation buffer, 2 µl of the pCR-Blunt vector (25 ng/µl), 1 µl of ExpressLink™ T4 DNA Ligase, and 3 µl of distilled water in an Eppendorf tube, mixed gently and incubated at 25˚C overnight for two hours. The ligation products were used to transform *E. coli* (Top10), and the transformation reaction was used to prepare 1:10 and 1:100 serial dilutions. These dilutions were plated on 50 µg/mL kanamycin plates containing x-Gal and IPTG and were incubated overnight at 37˚C. This method of using plates containing x-Gal and IPTG allows distinguishing between the colonies containing the pCR-Blunt plasmid and insert from a circularised plasmid containing colonies by blue/white screening. The next day, 10 white colonies of each reaction were picked up and prepared for Miniprep extraction. The transfer vector was then extracted using a QIAprep Spin Miniprep Kit (Qiagen, Cat No./ID: 27104) and sent for sequencing.

2.2.19.4 Clonal Sequencing

Sequencing the amplified fragment of the HBV genome (Precore, X and S regions) was carried out commercially by sending the transfer plasmids to GENEWIZ sequencing company, London, UK using the M13 Reverse sequencing primer. In a 0.2 ml tube, 200-500 ng of plasmid DNA and 5 µl of 5 pmol/ µl (µM) of sequencing primer were mixed with distilled water to bring the final volume to 10 µl. Sequence Analysis: the obtained sequences were aligned, and parameters of hypermutation were evaluated with Hypermut Program Package22 (http://www.hiv.lanl.gov/HYPERMUT/hypermut.html) (Noguchi, Ishino *et al.* 2005) .

2.2.20 Next-Generation Sequencing (NGS)

Next-Generation sequencing (NGS) was carried out to quantify the frequency and extent of APOBEC3G mutations in HBV pre-core, X, and S regions in the extracted HBV DNA to elucidate the effect of APOBEC3 proteins on HBV replication.

2.2.20.1 Primers ID Design

The PCR forward primers (primers ID) used were designed to consist of universal tag sequences at their 5′ ends for a later amplification step, random barcode (ID) of the form NNNNTNNNNTNNNN and sequence-specific to the HBV pre-core, X, S regions (Appendix A.9):-

HBVpre-core:(5'-TGTCCAGCACGCTTCAGGCTNNNNTNNNNTNNNNTAAATTGGTCTGCGCACCAGCA-3'),

HBVX:(5'-TGTCCAGCACGCTTCAGGCTNNNNTNNNNTNNNNATGGCTGCTAGGCTGTGCTGCCAA-3'),

HBVS:(5'-TGTCCAGCACGCTTCAGGCTNNNNTNNNNTNNNNTCTTCATCCTGCTGCTATGCCTCAT-3').

2.2.20.2 Nested PCR reaction

Nested PCR rounds were performed with Phusion U Hot Start PCR Master Mix (Thermo Scientific, F533S). The first round was to amplify single copies of primer ID amplicons. TaqMan™ Universal PCR Master Mix (Thermo Fisher, Catalog number: 4304437) qPCR was used to calculate the number of HBV DNA molecules to use for each PCR reaction. 20,000-40,000 molecules were used for each for the first round of the Nested PCR reaction. The cycling conditions for the first round were as follows; one cycle of 98°C for the 30s and

then two cycles of (98°C for 10s, 61°C for 120 s, 72 °C for 10 s) and an extension of 72 °C for 10 s. The first-round PCR products were purified with Qiaquick PCR purification Kit (Qiagen, Cat No./ID: 28104); the Qiagen PCR purification protocol was followed except for the elution, which was in 25 µl water. The second-round PCR were 25 cycles in a 50µl reaction: 30 µl Master mix using Primer ID out and reverse primer of each HBV region (pre-core, X, S) plus 20 µl from the first round PCR. The cycling conditions for the second round were as follows; one cycle of 98°C for the 30s, 25 cycles of (98°C for 10s, 61°C for 120 s, 72 °C for 10 s) and an extension of 72 °C for 15 s. The second PCR products were confirmed by running 1/6 of each sample in 1% Agarose gel (60v for 70 minutes). The first-round PCR products were purified by gel extraction of 5/6 of each sample using a Qiaquick gel extraction kit (Qiagen, Cat No./ID: 28704) according to manufacturer protocol. Samples were stored at -20°C.

2.2.20.3 NGS Library preparation

NEBNext® Ultra™ II DNA Library Prep Kit (Illumina, E7645) was used to prepare the library for the Next Generation Sequencing. First, the NEBNext End Prep step was done by adding 50 µl of the gel-purified sample, 3 µl NEBNext Ultra II End Prep Enzyme Mix (green), 7 µl NEBNext Ultra II End Prep Reaction Buffer (green) to a sterile nuclease-free tube. A 100 µl or 200 µl pipette was set to 50 µl, and the entire volume was gently pipetted up and down at least 10 times to mix the mixture thoroughly. A quick spin was performed to collect all liquid from the sides of the tube. The sample was placed in a thermocycler, with the heated lid set to ≥ 75°C, and the following program was run: 30 minutes at 20°C, 30 minutes at 65°C and held at 4°C. Second, the Adaptor Ligation step was done by adding the following components 30 µl NEBNext Ultra II Ligation Master Mix (red), 1 µl NEBNext Ligation Enhancer (red) and 2.5 µl NEBNext Adaptor (Multiplex Oligos Set1: NEB #E7335 and Multiplex Oligos Set2: #E7500) for Illumina (red) to the 60 µl End Prep Reaction Mixture. The Ultra II Ligation Master Mix was mixed by pipetting up and down several times before adding to the reaction.

A 100 µl or 200 µl pipette was set to 80 µl, the entire volume was gently pipetted up and down at least 10 times to mix the mixture thoroughly. A quick spin was performed to collect all liquid from the sides of the tube. The Ligation mixture was incubated at 20°C for 15 minutes in a thermocycler with the heated lid off. Then, 3 µl of USER Enzyme (red) was added to the ligation mixture, mixed well and incubated at 37°C for 15 minutes with the heated lid set to ≥ 47°C. Samples were stored overnight at -20°C. Third, the samples were purified with Qiaquick PCR purification Kit (Qiagen, Cat No./ID: 28104); the Qiagen PCR purification protocol was followed except for the elution in 18 µl water. Fourth, the PCR Enrichment step of Adaptor ligated DNA was done by adding the 15 µl Adaptor Ligated DNA fragments, 25 µl NEBNext Ultra II Q5 Master Mix (blue), 5 µl Index Primer/i7 Primer (blue), 5 µl Universal PCR Primer/i5 Primer (blue) to a sterile strip tube. A 100 µl or 200 µl pipette was set to 40 µl, and then the entire volume was gently pipetted up and down at least 10 times to mix the mixture thoroughly. A quick spin was performed to collect all liquid from the sides of the tube. The sample was placed in a thermocycler and performed PCR amplification using the following PCR cycling conditions: one cycle of the initial denaturation at 98°C for 30 seconds, 15 cycles of denaturation annealing extension (98°C at 65°C for 10 seconds, 75 seconds and one cycle of final extension 65°C 5 minutes) and held at 4°C. Fifth, the samples were purified with Qiaquick PCR purification Kit (Qiagen, Cat No./ID: 28104); the Qiagen PCR purification protocol was followed except for the elution in 18 µl water. Sixth, the concentration of samples was measured with The Qubit fluorometer (Invitrogen, Catalog number: Q33238). Seventh, an equal number of samples (copy number) were pooled in one tube. In the final step, 4 µl of the pooled samples were run in a 2% agarose gel. According to manufacturer protocol, the library band was extracted using a Qiaquick gel extraction kit (Qiagen, Cat No./ID: 28704). The purified pool sample was stored at -20°C.

2.2.20.4 Next-Generation Sequencing and analysis

Samples from the previous step were sequenced, producing 150 bp paired-end reads on an Illumina MiSeq sequencer (next-generation sequencing) at Molecular Diagnostic Unit (MDU), Imperial College London (St Mary's Campus, Norfolk Place, London W2 1PG). Sequencing data for the samples were processed and analysed by Dr Daniel Goldhill (Barcley's Group) using custom scripts in Python and R. the generated consensus sequences for each barcode had more than three reads, with the consensus taken as the majority of the reads. These generated consensus sequences were mapped and compared to the reference sequences of HBV pre-core, X, and S regions.

3 Chapter Three: Results

The overall goal of this project was to study the effects of the APOBEC3 proteins on the hepatitis B virus (HBV), carried out using the following experimental plan:

The first step consisted of constructing GFP-tagged APOBEC3 expressing plasmid and characterising their optimal expression using Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR) and their subcellular localisation using Immunofluorescence staining techniques.

The second step was to investigate the impact of individual and synergistic effects of APOBEC3 family members on the levels of cytoplasmic and extracellular virion-associated HBV DNA in the HBV-producing cell line HepG2.2.15. Their impact was measured using Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR).

The third step was to use the experimental system to analyse the frequency and extent of mutations in HBV DNA caused by APOBEC3 proteins detected using 3D-PCR sequencing and Next-generation sequencing.

The fourth step was to evaluate the effect of the Human Upstream stimulatory factor 1 (USF1) on APOBEC3 protein expression levels and its impact on HBV replication using Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR).

3.1 Construction of APOBEC3 Mammalian Expression Vectors

APOBEC3 overexpression will determine the individual and synergistic effect of the APOBEC3 family members on the cytoplasmic and extracellular virion-associated HBV DNA levels. The APOBEC3 overexpression system was developed by constructing APOBEC3DE, and APOBEC3G GFP tagged destination vectors (Appendix A.1), in which the GFP epitope in the destination vector is fused with APOBEC3 proteins. This C-terminal-tagged GFP epitope allows the detection and visualisation of the APOBEC3 proteins under a fluorescence microscope. APOBEC3DE and APOBEC3G DNA from shuttle entry vectors were digested with two rare-cutting restriction endonucleases (*Sgf*I and *Mlu*I), gel extracted, purified, and then ligated with correspondingly double-digested GFP-tagged destination vectors (Figure 9). Four colonies resistant to ampicillin were picked, and APOBEC3DE and APOBEC3G containing GFP-tagged destination plasmids were subsequently confirmed by sequencing. The junction between the plasmid and the insert (in the pCMV6-GFP, pCMV6-A3A-GFP, pCMV6-A3B-GFP, pCMV6-A3F-GFP pCMV6-A3DE-GFP and pCMV6-A3G-GFP) was confirmed by sequencing using T7 sequencing primers. (Figure 10).

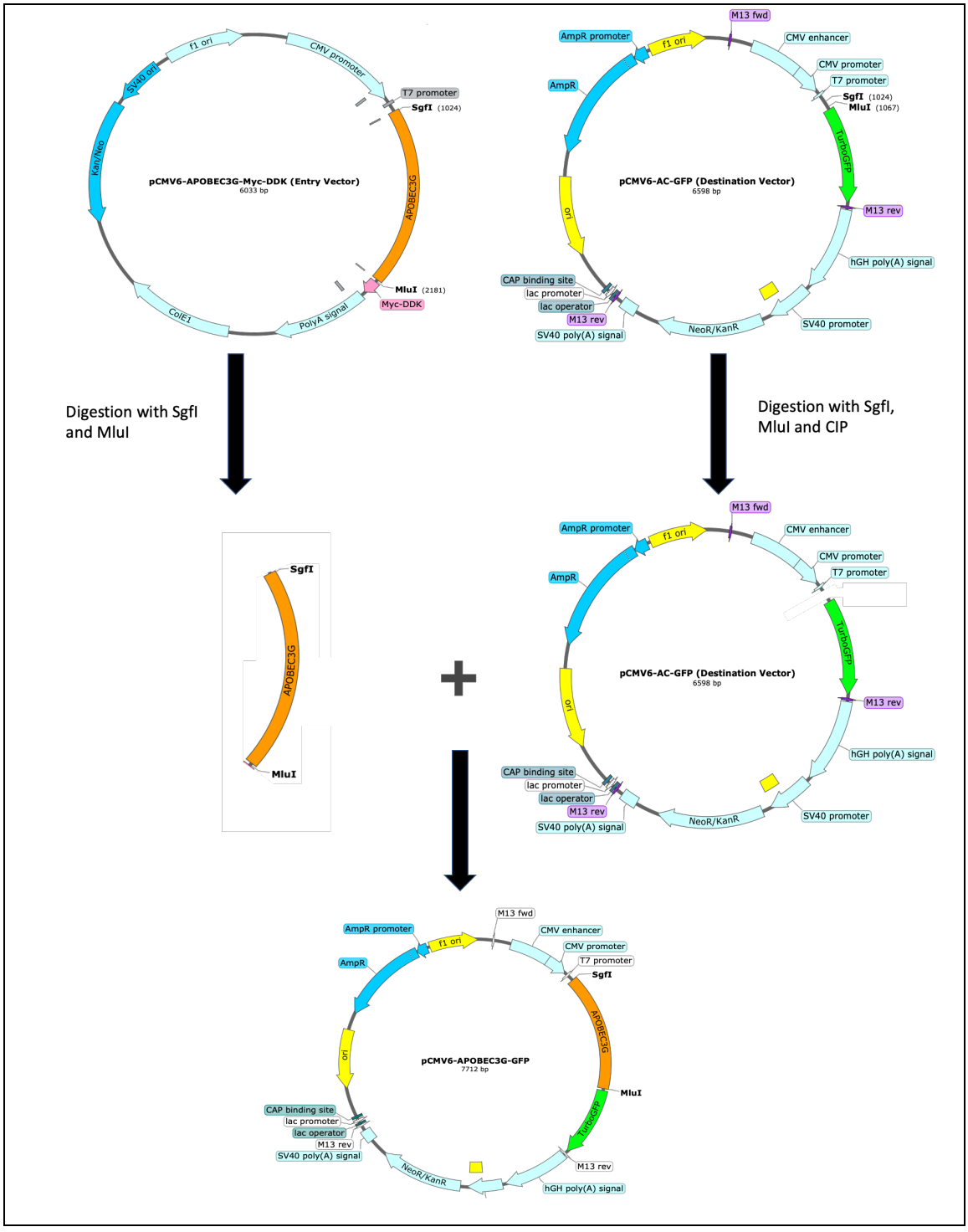


Figure 9. Structure of the Precision Shuttle subcloning procedure of A3DE and A3G.

The APOBEC3DE and APOBEC3G entry vectors and GFP tagged destination vectors were digested with two rare-cutting restriction endonucleases (*Sgf*I and *Mlu*I), gel extracted and purified and then ligated with correspondingly double-digested GFP tagged destination vectors. After a ligation reaction, the resulting clones are grown on ampicillin-containing plates to select for successful subcloning of the ORF into the destination vector. This diagram was adapted from the Origene company website.

pCMV6-GFP

```
CLUSTAL O(1.2.4) multiple sequence alignment


pCMV6-GFP           TACGCGGCCGCTCGAGATGGAGAGCGACGAGAGCGGCCTGCCCGCCATGGAGATCGAGTG
GFP                 ----------------ATGGAGAGCGACGAGAGCGGCCTGCCCGCCATGGAGATCGAGTG
                                    ****************************************

pCMV6-GFP           CCGCATCACCGGCACCCTGAACGGCGTGGAGTTCGAGCTGGTGGGCGGCGGAGAGGGCAC
GFP                 CCGCATCACCGGCACCCTGAACGGCGTGGAGTTCGAGCTGGTGGGCGGCGGAGAGGGCAC
                    ************************************************************

pCMV6-GFP           CCCCGAGCAGGGCCGCATGACCAACAAGATGAAGAGCACCAAAGGCGCCCTGACCTTCAG
GFP                 CCCCGAGCAGGGCCGCATGACCAACAAGATGAAGAGCACCAAAGGCGCCCTGACCTTCAG
                    ************************************************************
```

pCMV6-APOBEC3A-GFP

```
CLUSTAL O(1.2.4) multiple sequence alignment


pCMV6-APOBEC3A-GFP     NNNNNNNTTNNNANNNNATNGGTANNAGGANATCTGCCGCCGCGATCGCCATGGAAGCCA
APOBEC3A               ---------------------------------------------------ATGGAAGCCA
                                                                           **********

pCMV6-APOBEC3A-GFP     GCCCAGCATCCGGGCCCAGACACTTGATGGATCCACACATATTCACTTCCAACTTTAACA
APOBEC3A               GCCCAGCATCCGGGCCCAGACACTTGATGGATCCACACATATTCACTTCCAACTTTAACA
                       ************************************************************

pCMV6-APOBEC3A-GFP     ATGGCATTGGAAGGCATAAGACCTACCTGTGCTACGAAGTGGAGCGCCTGGACAATGGCA
APOBEC3A               ATGGCATTGGAAGGCATAAGACCTACCTGTGCTACGAAGTGGAGCGCCTGGACAATGGCA
                       ************************************************************
```

pCMV6-APOBEC3B-GFP

```
CLUSTAL O(1.2.4) multiple sequence alignment


pCMV6-APOBEC3B-GFP     NNNNNNNNNNNAANNGNTCGGTANNAGGANATCTGCCGCCGCGATCGCCATGAATCCACA
APOBEC3B               ---------------------------------------------------ATGAATCCACA
                                                                           ***********

pCMV6-APOBEC3B-GFP     GATCAGAAATCCGATGGAGCGGATGTATCGAGACACATTCTACGACAACTTTGAAAACGA
APOBEC3B               GATCAGAAATCCGATGGAGCGGATGTATCGAGACACATTCTACGACAACTTTGAAAACGA
                       ************************************************************

pCMV6-APOBEC3B-GFP     ACCCATCCTCTATGGTCGGAGCTACACTTGGCTGTGCTATGAAGTGAAAATAAAGAGGGG
APOBEC3B               ACCCATCCTCTATGGTCGGAGCTACACTTGGCTGTGCTATGAAGTGAAAATAAAGAGGGG
                       ************************************************************
```

pCMV6-APOBEC3F-GFP

```
CLUSTAL O(1.2.4) multiple sequence alignment


pCMV6-APOBEC3F-GFP     NNNNNTNNNNANNGNTCGGTANNAGGANATCTGCCGCCGCGATCGCCATGAAGCCTCACT
APOBEC3F               ---------------------------------------------ATGAAGCCTCACT
                                                                     *************

pCMV6-APOBEC3F-GFP     TCAGAAACACAGTGGAGCGAATGTATCGAGACACATTCTCCTACAACTTTTATAATAGAC
APOBEC3F               TCAGAAACACAGTGGAGCGAATGTATCGAGACACATTCTCCTACAACTTTTATAATAGAC
                       ************************************************************

pCMV6-APOBEC3F-GFP     CCATCCTTTCTCGTCGGAATACCGTCTGGCTGTGCTACGAAGTGAAAACAAAGGGTCCCT
APOBEC3F               CCATCCTTTCTCGTCGGAATACCGTCTGGCTGTGCTACGAAGTGAAAACAAAGGGTCCCT
                       ************************************************************
```



Figure 10. APOBEC3-GFP plasmids were confirmed by sequencing.

Sequencing the GFP tagged APOBEC3 expressing plasmids was carried out commercially by sending the plasmids to the GENEWIZ sequencing company in London, UK, using the T7 sequencing primers. In a 0.2 ml tube, 200-500 ng of plasmid DNA and 5 µl of 5 pmol/ µl (µM) of sequencing primer were mixed with distilled water to bring the final volume to 10 µl.

HEK293T cells were seeded in 10cm dishes and transfected with 10 µg of pCMV6-AC-GFP or pCMV6-A3DE-GFP or pCMV6-A3G-GFP plasmids using calcium phosphate transfection. The expression of constructed pCMV6-A3DE-GFP and pCMV6-A3G-GFP was examined and detected by visualising the C-terminal-tagged GFP epitope in the transfected HEK293T cells under the fluorescence microscope. After forty-eight hours, the GFP was checked under a fluorescence microscope at 200x magnification. GFP was visualised in the cytoplasm of the pCMV6-A3DE-GFP or pCMV6-A3G-GFP transfected HEK293T cells, while GFP appears to be present throughout the cell in the pCMV6-AC-GFP transfected HEK293T cells (Figure 11).

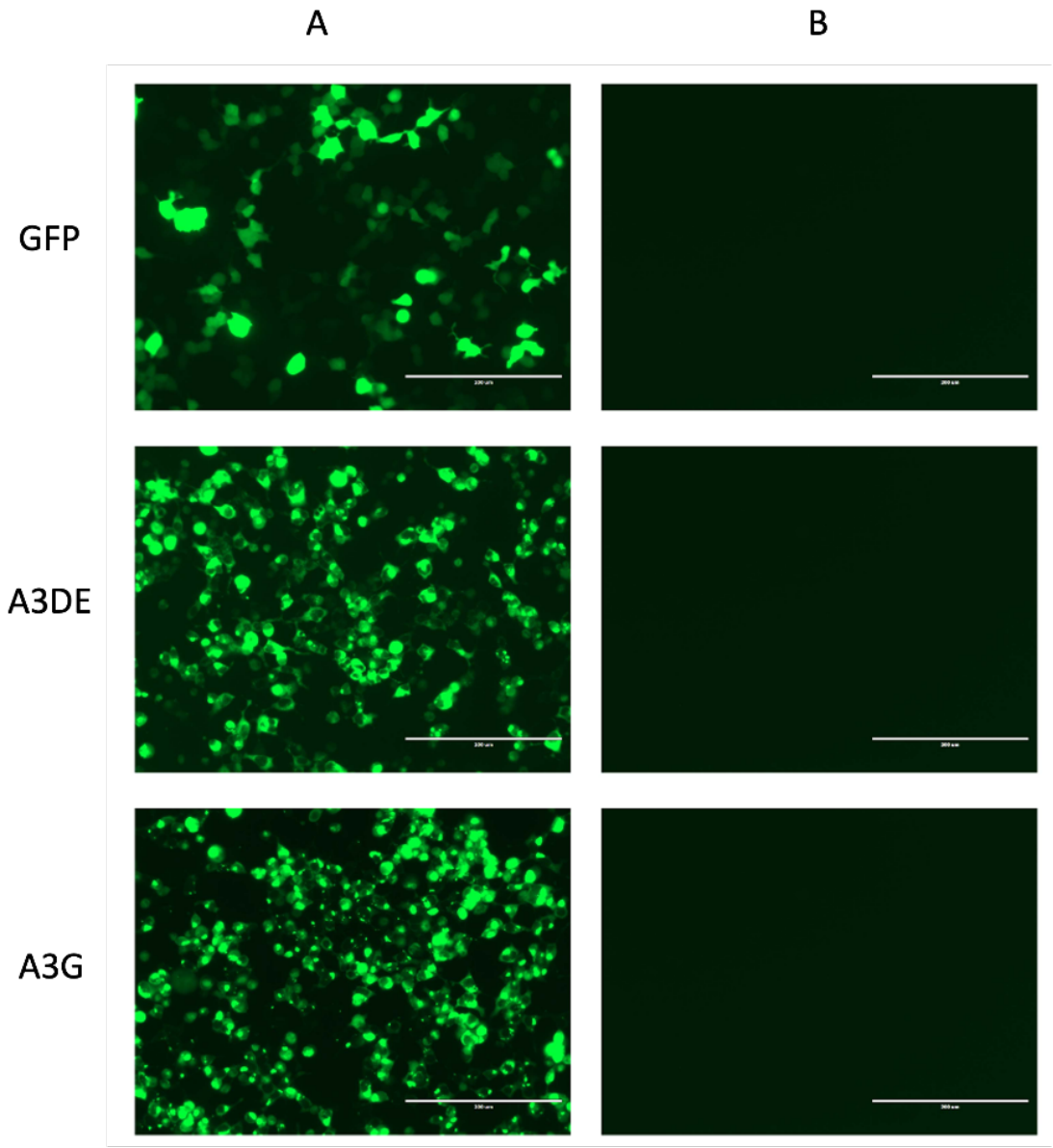


Figure 11. Exogenous expression of APOBEC3DE and APOBEC3G proteins in HEK293T cells.

HEK293T cells were transfected with 10 µg of pCMV6-A3DE-GFP and pCMV6-A3G-GFP, and after 48 hours, the cells were visualised by fluorescence microscopy. A: HEK293T cells transfected with pCMV6-GFP or pCMV6-A3DE-GFP or pCMV6-A3G-GFP plasmids, B: HEK293T cells transfected with pCMV6 empty plasmid (Mock transfection). GFP was visualised in the cytoplasm of the pCMV6-A3DE-GFP or pCMV6-A3G-GFP transfected HEK293T cells, while GFP appears to be present throughout the cell in the pCMV6-GFP transfected HEK293T cells (control).

HEK293T cells were transfected with 10 µg of pCMV6-GFP, pCMV6-A3A-GFP, pCMV6-A3DE-GFP and pCMV6-A3G-GFP, and after 48 hours. The proteins were detected using Rabbit anti-GFP primary antibody and Mouse monoclonal anti-GAPDH antibody, then adding donkey anti-mouse secondary antibody conjugated with IRDye680LT (LI-COR) and 1: 15,000 dilution of the goat anti-rabbit secondary antibody conjugated with IRDye800CW (LI-COR). Then, GAPDH as reference protein (red) and GFP tagged APOBEC3 proteins (green) were visualized using the Odyssey CLx Imaging System (LI-COR). The expected proteins sizes; GFP (26 kDa), A3A (26 kDa), A3DE (46 kDa), A3G (46 kDa). The expected sizes APOBEC3-GFP proteins (APOBEC3 plus GFP): A3A-GFP (49 kDa), A3DE-GFP (72 kDa), and A3G-GFP (72 kDa). (Figure 12).

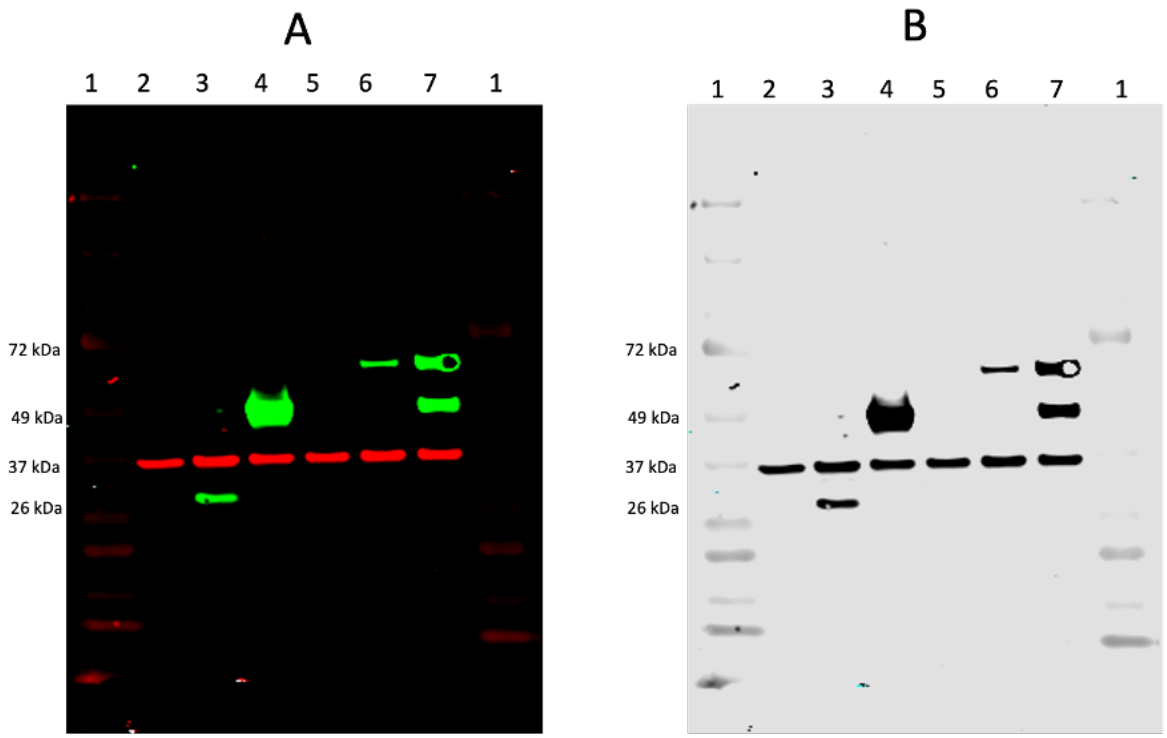


Figure 12. The western blot image shows the protein expression of Empty-GFP, APOBEC3A, APOBEC3DE and APOBEC3G proteins in HEK293T cells.

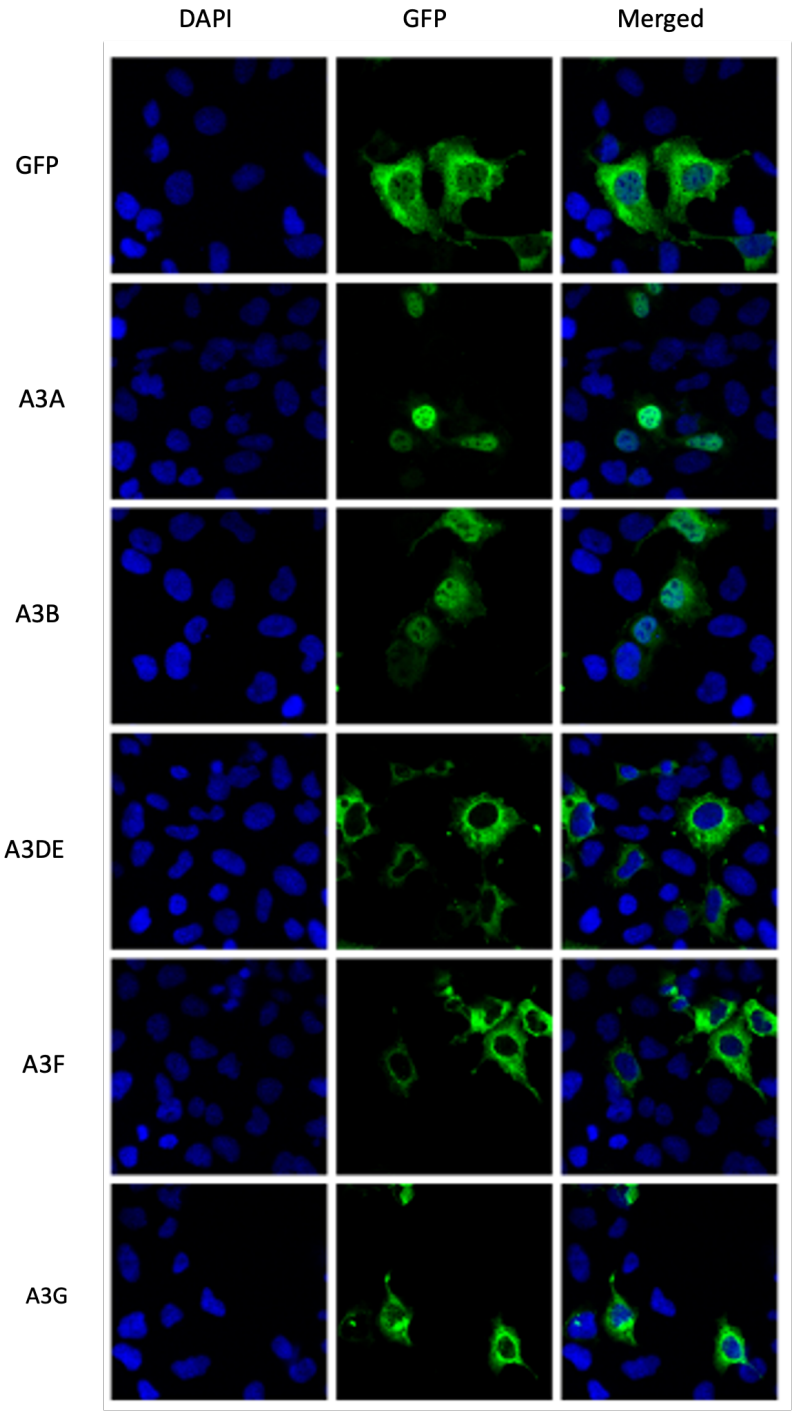
HEK293T cells were transfected with 10 µg of pCMV6-GFP, pCMV6-A3A-GFP, pCMV6-A3DE-GFP and pCMV6-A3G-GFP, and after 48 hours. Lane 1: ladder, 2: Mock transfected HEK293T cells, 3: Empty-GFP vector transfected HEK293T cells, 4: A3A-GFP transfected HEK293T cells, 5: Mock transfected HEK293T cells 6: A3DE-GFP transfected HEK293T cells 7: A3G-GFP transfected HEK293T cells. (A) GAPDH as reference protein (red) and GFP tagged APOBEC3 proteins (green) were visualized using the Odyssey CLx Imaging System (LI-COR). (B) is the white and black image of (A).

3.2 The Intracellular localisation of APOBEC3 proteins and the colocalisation of HBV polymerase protein with APOBEC3G

3.2.1 The Intracellular localisation of APOBEC3 proteins

The localisation of APOBEC3 proteins determines their effects on specific parts of the hepatitis B virus life cycle. Immunofluorescence staining was performed to define where these deaminases are localised in the cell. HEK293T cells were seeded on coverslips, transfected with GFP-tagged APOBEC3 plasmids (APOBEC3A, APOBEC3B, APOBEC3DE, APOBEC3F and APOBEC3G) and stained with Rabbit anti-GFP antibody (green staining). An immunofluorescence confocal microscope was used to take images. APOBEC3DE, APOBEC3F and APOBEC3G proteins expression were restricted to the cytoplasm where HBV reverse transcription occurs. In contrast, APOBEC3A and APOBEC3B were observed to have mainly nuclear localisation. A plasmid expressing GFP was used as a control, and its expression was distributed throughout the cell (Figure 13).

A



B

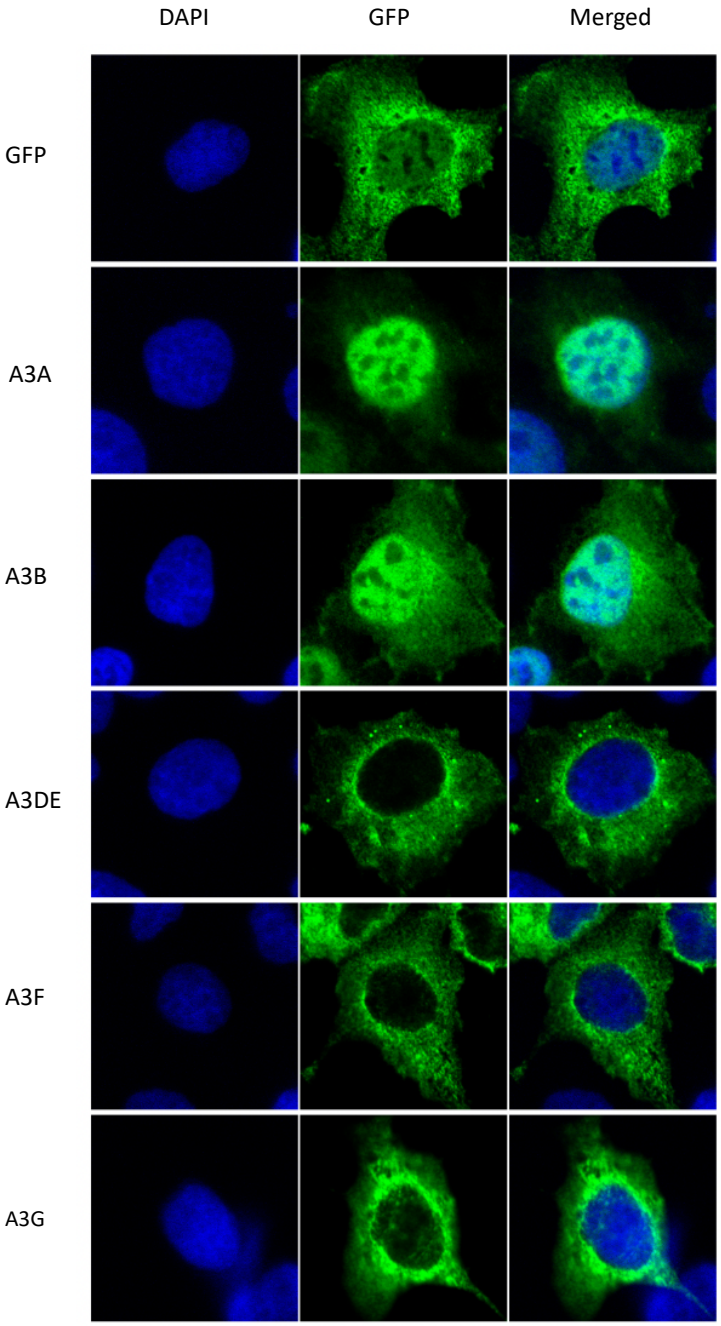


Figure 13. The subcellular localisation of APOBEC3 Proteins in HEK293T cells.

APOBEC3 proteins staining of APOBEC3 transfected HEK293T cells. A) Confocal microscopy images showing localization of APOBEC3 proteins. B) Higher magnification of images in A. The APOBEC3 proteins were stained using a Rabbit anti-GFP antibody (green staining). Cell nuclei were also stained with DAPI (blue staining). APOBEC3A and APOBEC3B have mainly nuclear localisation, while APOBEC3DE, APOBEC3F and APOBEC3G have a cytoplasmic localisation. APOBEC3 proteins staining in the transfected cells was compared to GFP transfected cells. Images were captured using an immunofluorescence confocal microscope at a magnification of 60x.

3.2.2 The colocalisation of HBV polymerase protein with APOBEC3G protein

HBV polymerase protein localisation determines its effect on specific parts of the hepatitis B virus life cycle. HBV polymerase protein co-localisation with APOBEC3G protein was examined by immunofluorescence staining, which was performed to define where these proteins are co-localised in the cell. HEK293T cells were seeded on coverslips, transfected with GFP-tagged APOBEC3G (Appendix A.1) and Flag-tagged HBV polymerase plasmids (Appendix A.3), then stained with Rabbit anti-GFP primary antibody (green staining) and Mouse anti-Flag antibody (red staining). Immunofluorescence confocal microscope was used to take images. APOBEC3G proteins and HBV polymerase proteins expression was mainly in the cytoplasm. There are areas where the APOBEC3G proteins and HBV polymerase proteins colocalise, as shown in the merged image (yellow color). Confocal microscopy images of a number of different HEK293T cells are representative to show co-localization of APOBEC3G and HBV polymerase (Figure 14) and Figure 15 shows Confocal microscopy images with Higher magnification of the image number 4 in the Figure14.

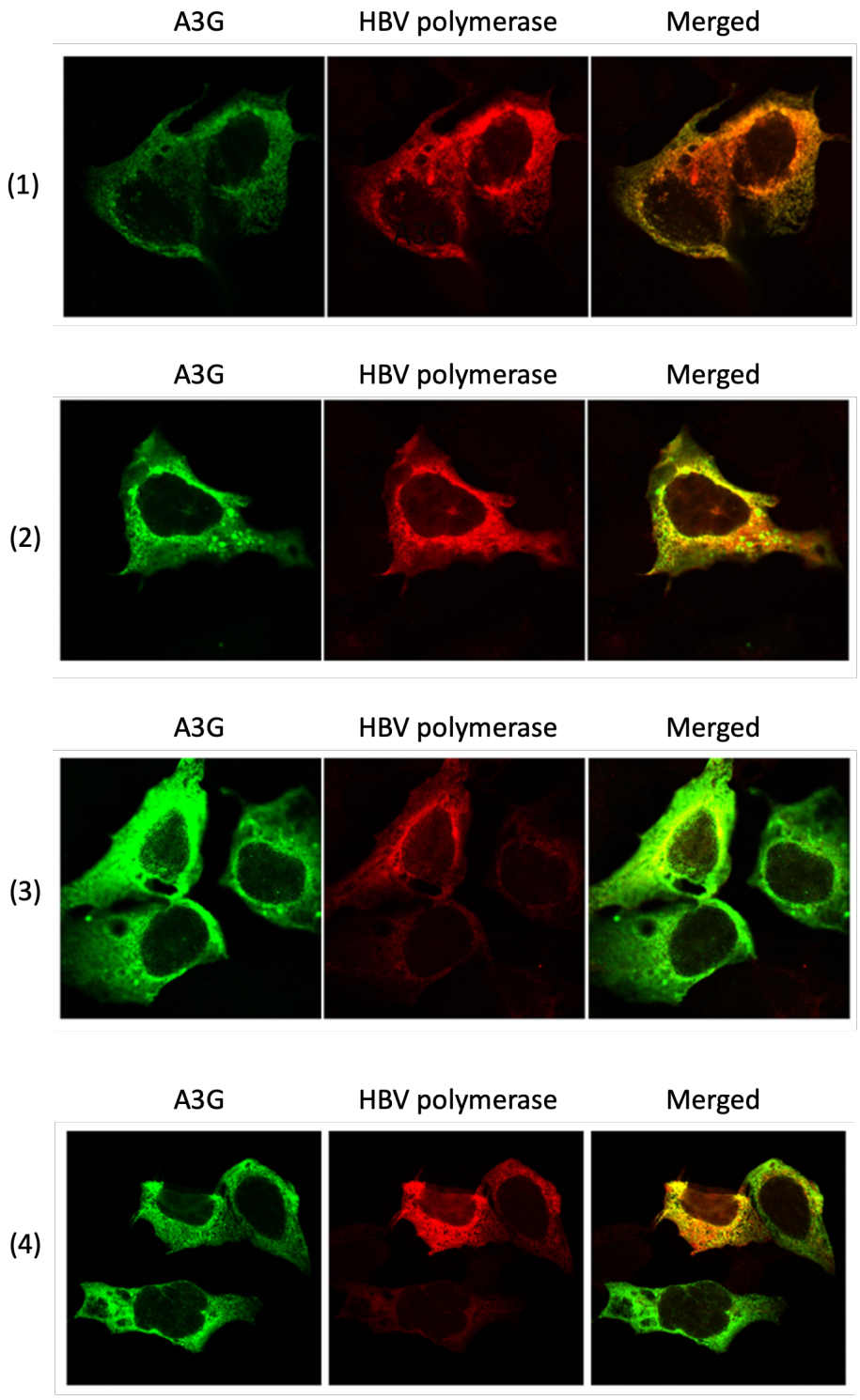


Figure 14. Localization and co-localization of HBV polymerase and APOBEC3G.

Confocal microscopy images of a number of different HEK293T cells showing co-localization of APOBEC3G and HBV polymerase. (1), (2), (3), and (4). APOBEC3G proteins expression was mainly in the cytoplasm (green) and HBV polymerase proteins expression was mainly in the cytoplasm (red). The colocalisation of APOBEC3G proteins and HBV polymerase proteins in areas of the cytoplasm as shown in the merged image (yellow color).

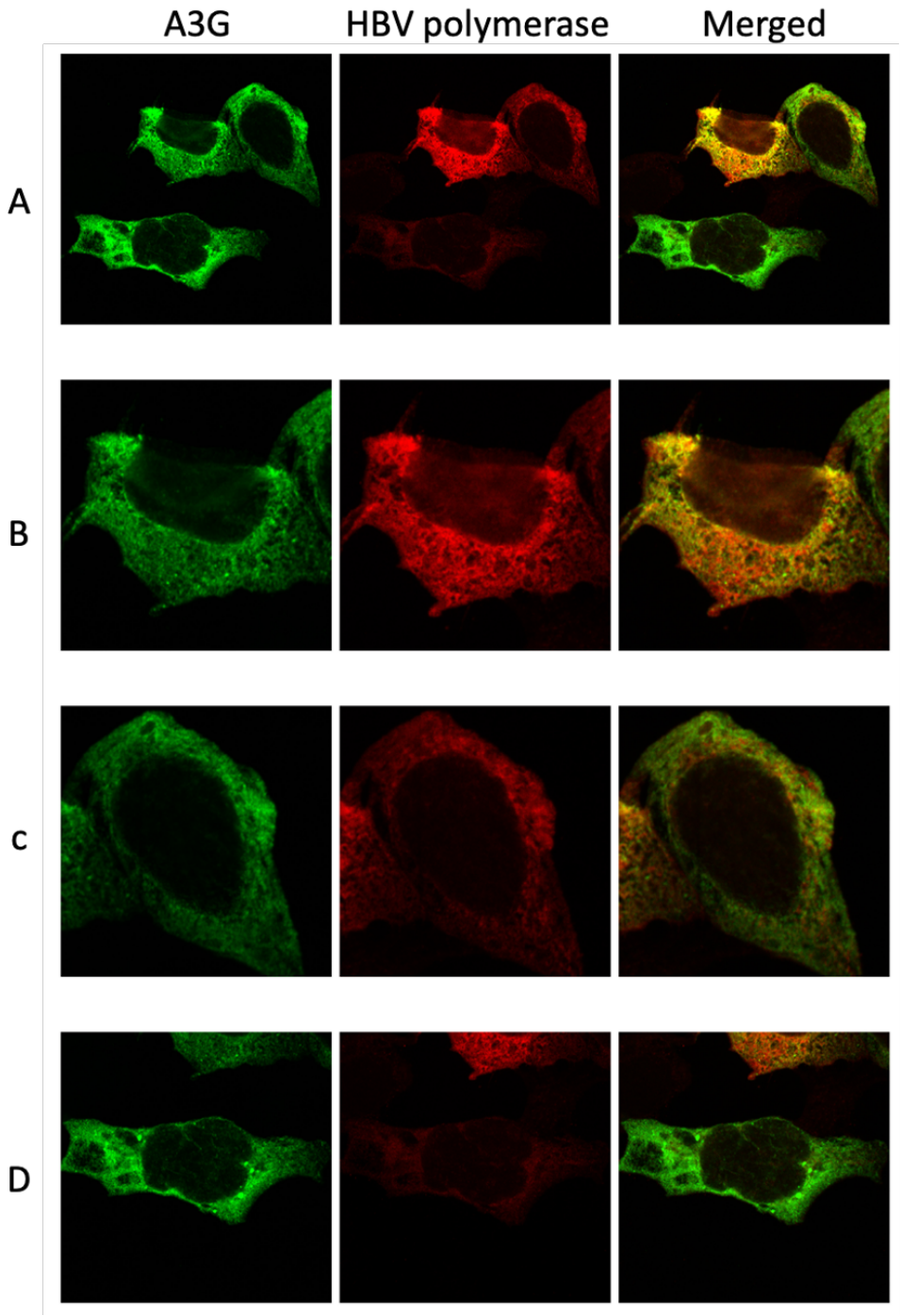


Figure 15. Localization and co-localization of HBV polymerase and APOBEC3G.

Confocal microscopy images with Higher magnification of image number (4) in the (figure 14) showing co-localization of APOBEC3G and HBV polymerase. B, C and D) Higher magnification of ones in A.

3.2.3 The colocalisation of HBV polymerase protein with DDX3

DDX3 was used as positive control for experiment examining the possible co-localisation of HBV polymerase with APOBEC3G. Immunofluorescence staining examined HBV polymerase protein co-localisation with DDX3 protein, which was performed to define where these proteins are co-localised in the cell. HEK293T cells were seeded on coverslips, transfected with HA-tagged DDX3 (Appendix A.4) and Flag-tagged HBV polymerase plasmids (Appendix A.3), and stained with Rabbit anti-HA primary antibody (green staining) and Mouse anti-Flag antibody (red staining). Immunofluorescence confocal microscope was used to take images. DDX3 proteins and HBV polymerase proteins expression was mainly in the cytoplasm. There are areas where the DDX3 proteins and HBV polymerase proteins colocalise, as shown in the merged image (yellow color). Also, it was observed that in these areas where DDX3 proteins and HBV polymerase proteins colocalise, the DDX3 protein expression was less prominent compared to areas where HBV polymerase proteins and DDX3 proteins do not colocalise (Figure 16).

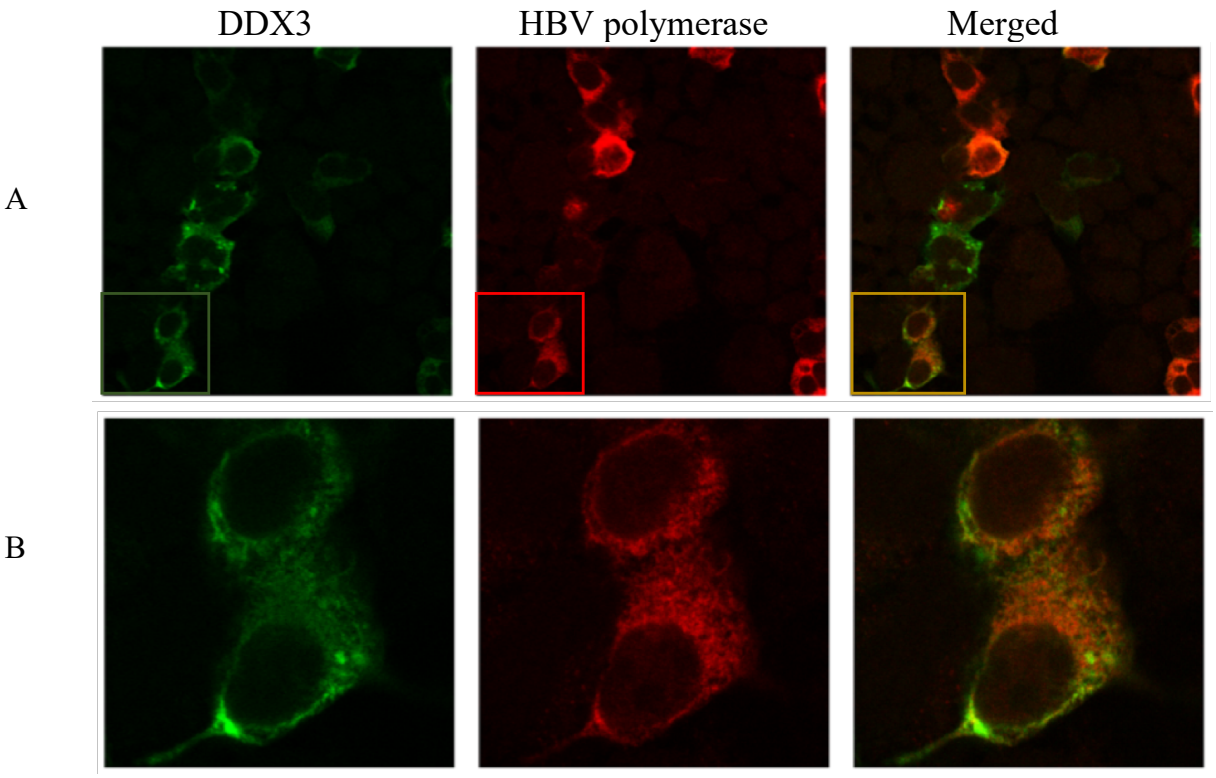


Figure 16. Localization and co-localization of HBV polymerase and DDX3.

A) Confocal microscopy images showing co-localization of DDX3 and HBV polymerase. B) Higher magnification of ones enclosed in Boxes in A.

3.3 APOBEC3 proteins inhibit HBV replication

The experimental approach to evaluate the effect of APOBEC3 proteins on HBV DNA replication involves exogenous overexpression of APOBEC3 proteins in HBV producing HepG2.2.15. Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR) was used to investigate the impact of individual and synergistic APOBEC3 family members on the levels of cytoplasmic and extracellular virion-associated HBV DNA (Figure 17).

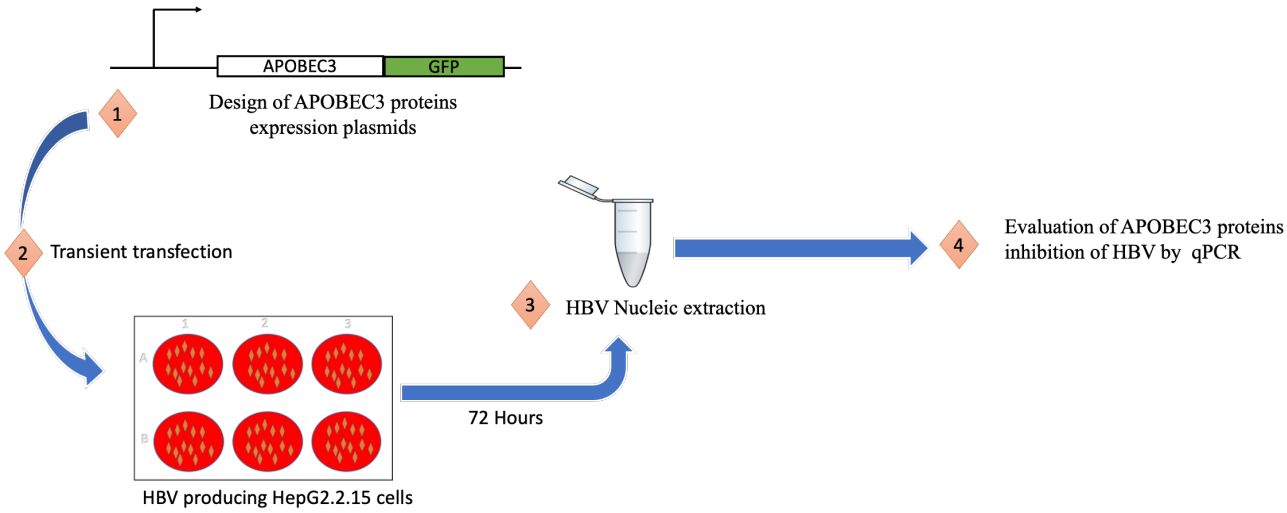


Figure 17. The experimental approach to evaluate the effect of APOBEC3 proteins on HBV DNA replication.

1) APOBEC3 expression plasmids were constructed. 2) HBV producing HepG2.2.15 cells were transfected with GFP-tagged APOBEC3 expression plasmids. 3) 72 hours post transfection; the cytoplasmic capsid-associated HBV DNA and the extracellular virion-associated HBV DNA was extracted from HepG2.2.15 cells 4) The evaluation of APOBEC3 proteins inhibition of HBV DNA levels were determined by using Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR).

3.3.1 APOBEC3G protein expression in HepG2 cells over different time-points

The optimal time point of APOBEC3G expression in HepG2 cells was examined by transfecting HepG2 cells with an APOBEC3G expression plasmid for 24 hours, 48 hours, 72 hours, and 96 hours. Post-transfection, total RNA was extracted, and the overexpression of APOBEC3G was measured using quantitative Real-Time Polymerase Chain Reaction (RT-qPCR). The expression levels of APOBEC3G ranged from 1.8×10^4 fold change at 24-hours to 4.3×10^4 fold change at 72-hours post-transfection (Figure 18).

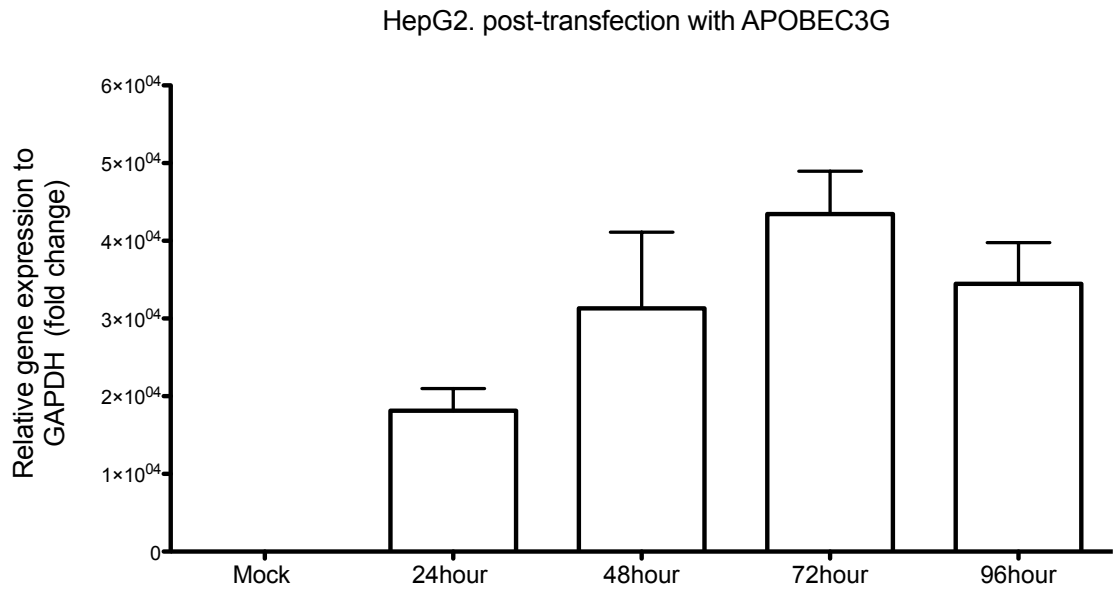


Figure 18. APOBEC3G overexpression in HepG2 cells at different time points.

HepG2 cells were transfected with APOBEC3G expression plasmid, Total RNA was extracted 24 hours, 48 hours, 72 hours, and 96 hours post-transfection and the fold change of APOBEC3G was quantified with real-time qPCR. The fold change of APOBEC3G gene expression levels is shown in comparison to the un-transfected sample (Mock), and the data represent the mean (+/-SD), N=9. Thus, the data are representative of two independent experiments.

3.3.2 APOBEC3 proteins inhibit cytoplasmic HBV DNA

The effect of APOBEC3 proteins on the cytoplasmic HBV DNA levels was determined by using Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR). HBV producing HepG2.2.15 cells were transfected with GFP-tagged APOBEC3 expression plasmids (APOBEC3A, APOBEC3B, APOBEC3DE, APOBEC3F or APOBEC3G). Seventy-two hours post-transfection, the cytoplasmic capsid-associated HBV DNA was extracted from cells using NP40 lysis buffer that only lyses the plasma membrane realising the contents of the cytoplasm but leaves the nuclear membrane intact (Galvis, Fisher et al. 2014, Galvis, Fisher et al. 2017). The effects of APOBEC3 protein on cytoplasmic HBV DNA were compared to the control; GFP transfected HBV producing HepG2.2.15 cells (Figure 19). APOBEC3F and APOBEC3G caused the highest inhibition of the cytoplasmic HBV DNA by 50.8 and 65.8 %, respectively. While the percentages of the cytoplasmic HBV DNA inhibition by APOBEC3A and APOBEC3B were 37.8 % and 57.2 %, APOBEC3DE inhibited the cytoplasmic HBV DNA by 39.7 % (Figure 20 and Table 5). Thus, all five APOBEC3 proteins show variable inhibitory effects on the cytoplasmic HBV DNA, with the highest inhibition being by APOBEC3F and APOBEC3G and the lowest inhibition by APOBEC3A and APOBEC3DE.

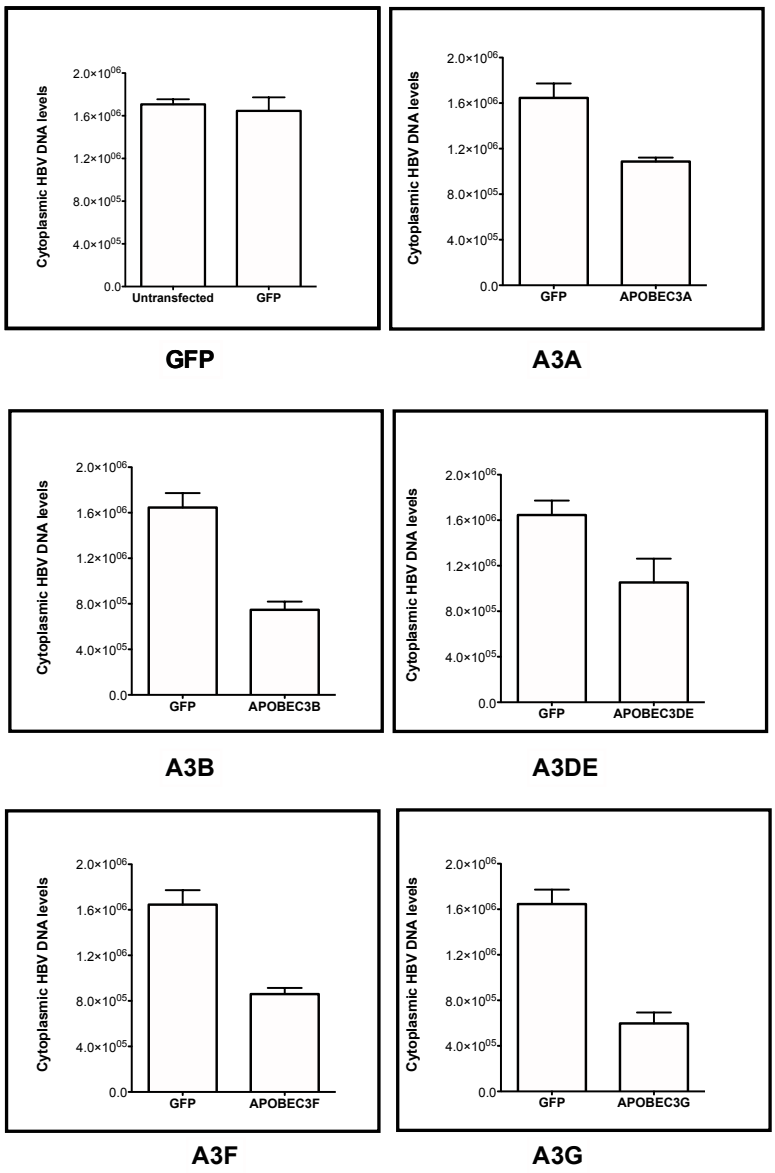


Figure 19. The effect of individual APOBEC3 proteins on the cytoplasmic HBV DNA levels (cytoplasmic HBV DNA copy number).

HepG2.2.15 cells were transfected with individual APOBEC3 vectors (2 μg), and 72 hours post-transfection, cytoplasmic HBV DNA was extracted and quantified with real-time qPCR using TaqMan Assay (forward primer, reverse primer, and probe). The cytoplasmic HBV DNA inhibition levels are shown as compared to the control GFP vector transfected and un-transfected samples, and the data represent the mean (+/-SD), N=9. Thus, the data are representative of three independent experiments.

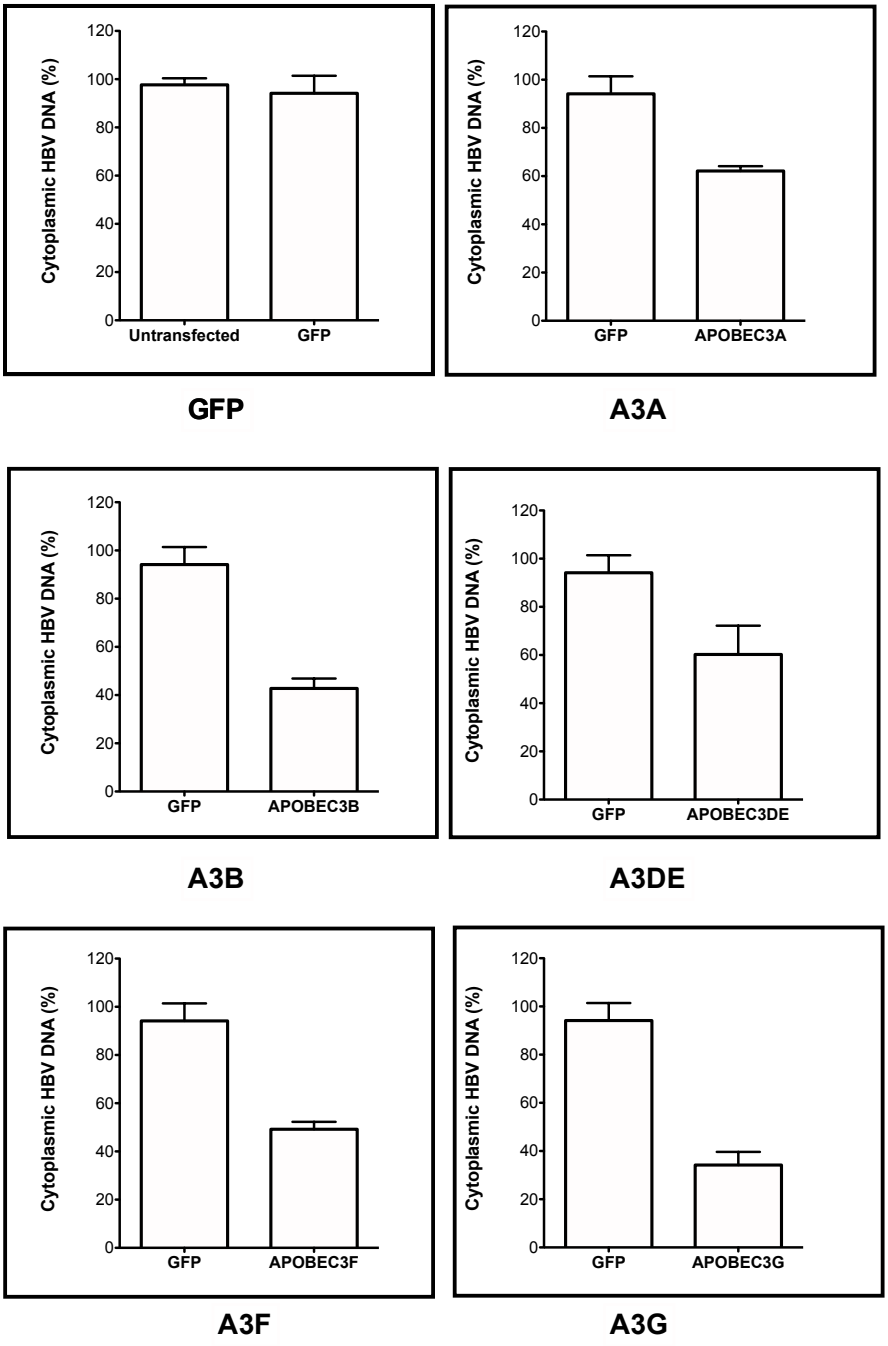


Figure 20. The effect of individual APOBEC3 proteins on the cytoplasmic HBV DNA levels (percentages of cytoplasmic HBV DNA inhibition).

This figure shows the same experimental data as Figure 19 represented as percentages of cytoplasmic HBV DNA inhibition.

HepG2.2.15	*GFP*	*A3A*	*A3B*	*A3DE*	*A3F*	*A3G*
(%) of cytoplasmic HBV DNA degradation	5.85	37.84	57.25	39.76	50.81	65.82

Table 5. The inhibition of cytoplasmic HBV DNA levels by individual APOBEC3 proteins.

The results show the percentages of cytoplasmic HBV DNA degradation in the APOBEC3 transfected cells versus GFP transfected cells.

3.3.3 APOBEC3 proteins inhibit extracellular capsid-associated HBV DNA

Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR) was used to measure the effects of APOBEC3 proteins on the levels of the extracellular virion-associated HBV DNA in HBV producing HepG2.2.15 cells. The extracellular virion-associated HBV DNA was extracted from the supernatant of HBV producing HepG2.2.15 cells after 72-hours post-transfection with GFP-tagged APOBEC3 expression plasmids (APOBEC3A, APOBEC3B, APOBEC3DE, APOBEC3F or APOBEC3G). The levels of APOBEC3 proteins inhibition of the extracellular virion-associated HBV DNA were compared to the control; GFP transfected HBV producing HepG2.2.15 cells (Figure 21). APOBEC3F and APOBEC3G caused the highest inhibition of the extracellular virion-associated HBV DNA by 55.4 % and 48.7 %, respectively. While the percentages of the extracellular virion-associated HBV DNA inhibition by APOBEC3A and APOBEC3B was 32.6 % and 33.9 %, APOBEC3DE inhibited the extracellular virion-associated HBV DNA by 16 % (Figure 22 and Table 6). All five APOBEC3 proteins show variable inhibitory effects on the extracellular virion-associated HBV DNA, with the highest inhibition was by APOBEC3F and APOBEC3G and the lowest inhibition by APOBEC3DE.

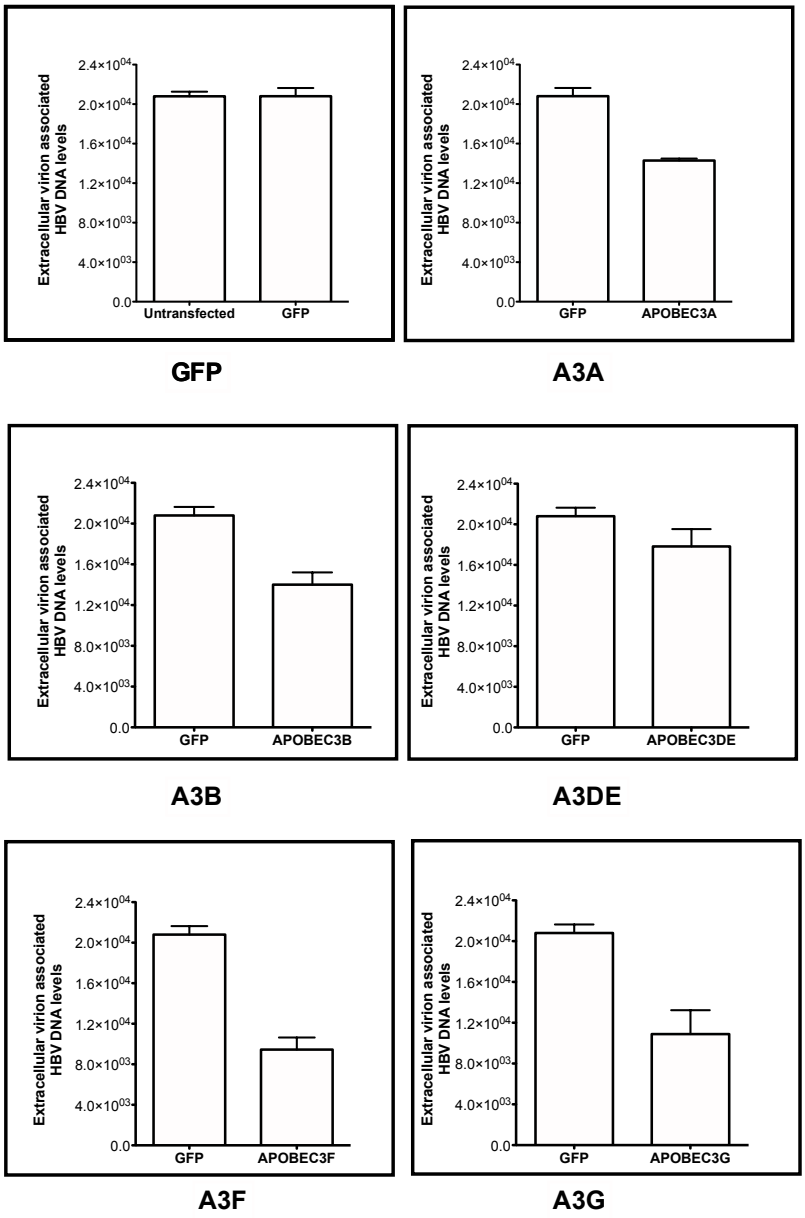


Figure 21. The effect of individual APOBEC3 proteins on the extracellular virion-associated HBV DNA levels (extracellular virion-associated HBV DNA copy number).

HepG2.2.15 cells were transfected with individual APOBEC3 vectors (2 μg), and 72 hours post-transfection, extracellular virion-associated HBV DNA was extracted and quantified with real-time qPCR using TaqMan Assay (forward primer, reverse primer, and probe). The extracellular virion-associated HBV DNA inhibition levels are shown as compared to the control GFP vector transfected and un-transfected samples, and the data represent the mean (+/-SD), N=9. Thus, the data are representative of three independent experiments.

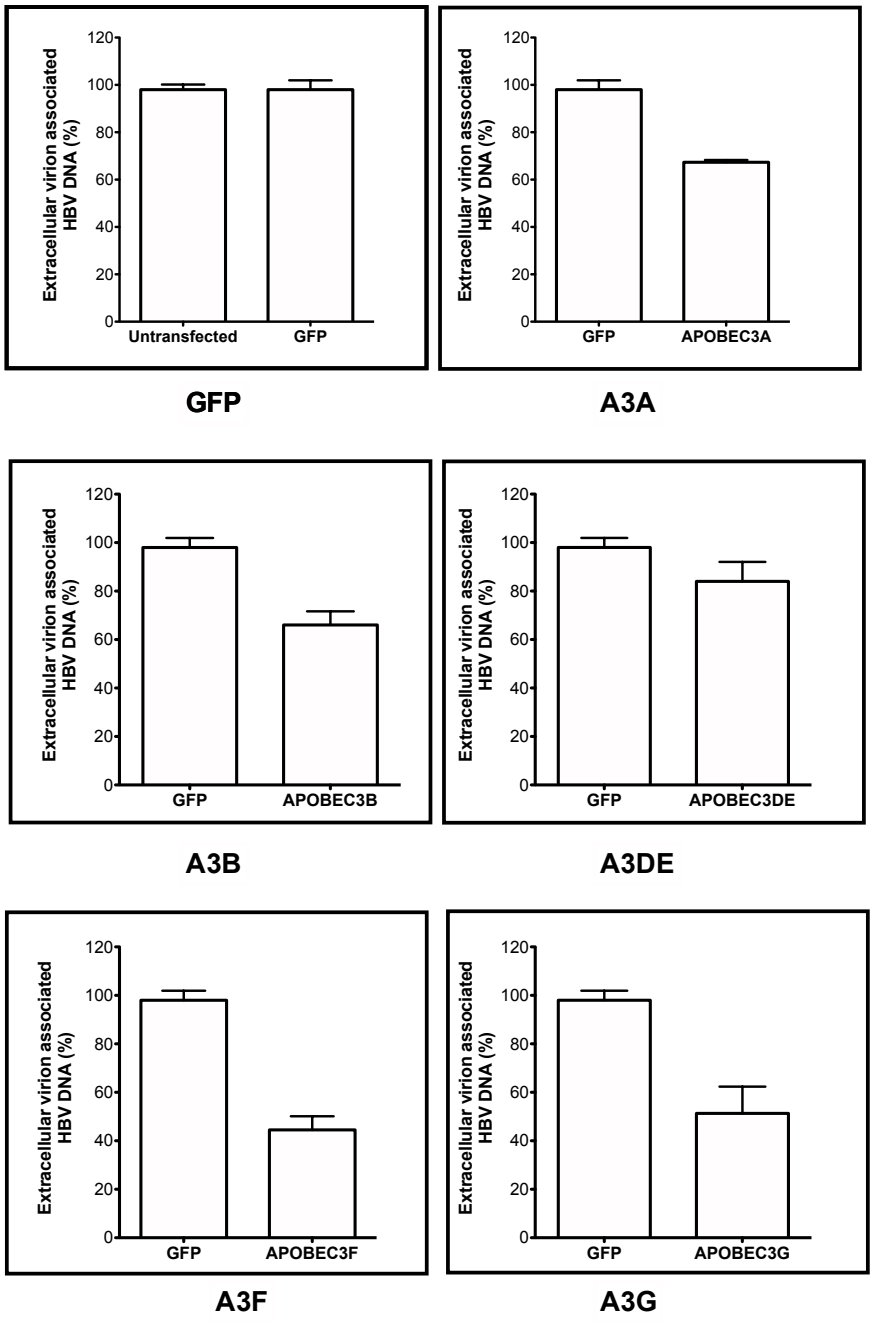


Figure 22. The effect of individual APOBEC3 proteins on the extracellular virion-associated HBV DNA levels (percentages of extracellular virion-associated HBV DNA inhibition).

This figure shows the same experimental data as Figure 21 represented as percentages of the extracellular virion-associated HBV DNA inhibition.

HepG2.2.15	*GFP*	*A3A*	*A3B*	*A3DE*	*A3F*	*A3G*
(%) of Extracellular virion-associated HBV DNA degradation	2.00	32.68	33.99	16.02	55.48	48.70

Table 6. The inhibition of extracellular virion-associated HBV DNA levels by individual APOBEC3 proteins.

The results show the percentages of extracellular virion-associated HBV DNA degradation in the APOBEC3 transfected cells versus GFP transfected cells.

3.3.4 APOBEC3 proteins synergistically inhibit cytoplasmic HBV DNA

The synergistic effect of GFP-tagged APOBEC3A, APOBEC3F and APOBEC3G proteins on the cytoplasmic HBV DNA levels in the cells' cytoplasm was determined using Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR). Huh7 cells were co-transfected with HBV DNA plasmid alongside GFP-tagged APOBEC3A, APOBEC3F and APOBEC3G proteins. Seventy-two hours post-transfection, the cytoplasmic HBV DNA was extracted from cells using NP40 lysis buffer that only liberates the cell's cytoplasmic contents and leaves the nuclear membrane intact. The synergistic effect of APOBEC3 proteins on cytoplasmic HBV DNA were compared to the control; GFP transfected Huh7 cells (Figure 23). Together APOBEC3A, APOBEC3F, and APOBEC3G proteins inhibited the cytoplasmic HBV DNA by 67.2 % (Figure 24 and Table 7). In addition, the APOBEC3 proteins' synergistic inhibitory effect on the cytoplasmic HBV DNA was higher in comparison to the inhibition by individual APOBEC3A, APOBEC3F or APOBEC3G.

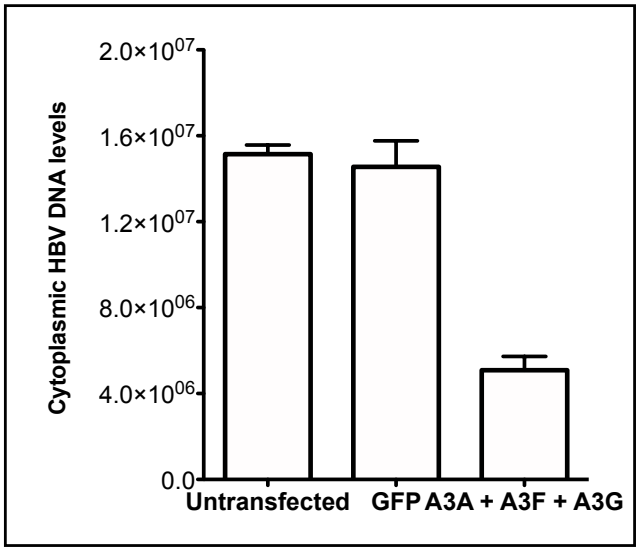


Figure 23. The synergistic effect of APOBEC3 proteins on the cytoplasmic HBV DNA levels (cytoplasmic HBV DNA copy number).

Real-time qPCR using TaqMan Assay (forward primer, reverse primer, and probe) was used to determine the synergistic effect of APOBEC3 proteins on HBV replication. Huh7 cells were co-transfected with 2 μg in total, 0.5 μg of each HBV DNA plasmid together with APOBEC3A, APOBEC3F, and APOBEC3G vectors. At 72 hours post-transfection, cytoplasmic HBV DNA was extracted and quantified with real-time qPCR using TaqMan Assay (forward primer, reverse primer, and probe). The cytoplasmic HBV DNA inhibition levels are shown as compared to the control GFP vector transfected and un-transfected samples, and the data represent the mean (+/-SD), N=9. The data are representative of three independent experiments.

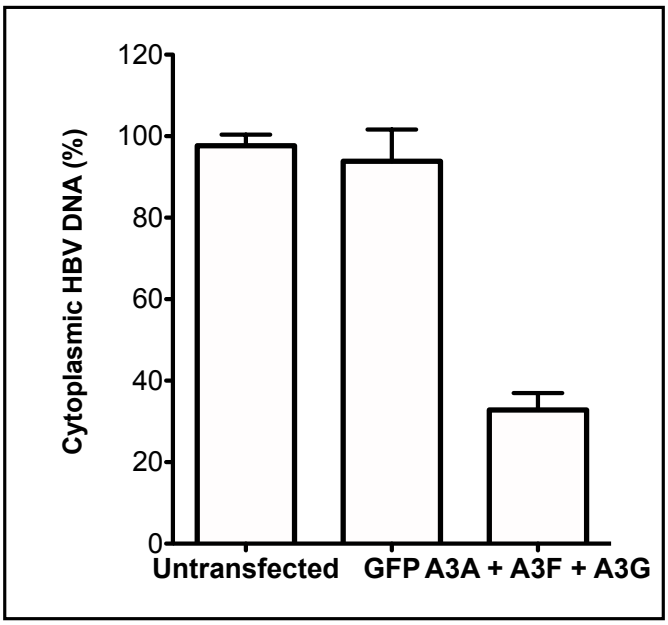


Figure 24. The synergistic effect of APOBEC3 proteins on the cytoplasmic HBV DNA levels (percentages of cytoplasmic HBV DNA inhibition).

This figure shows the same experimental data as Figure 23 represented as percentages of cytoplasmic HBV DNA inhibition.

HepG2.2.15	***GFP***	***A3A+A3F+A3G***
(%) cytoplasmic HBV DNA degradation	6.15	67.20

Table 7. The synergistic inhibition of cytoplasmic HBV DNA levels by APOBEC3 proteins.

The results show the percentages of cytoplasmic HBV DNA degradation in the APOBEC3 transfected cells versus GFP transfected cells.

3.4 APOBEC3 proteins edit the HBV genome

APOBEC3 proteins exert their inhibition effect on hepatitis B virus replication through editing and non-editing mechanisms. HBV has GC-rich regions in its genome (Appendix A.4)(Zhang, Li et al. 2013). These GC-rich regions are targeted by the cytidine to uridine deamination effect of APOBEC3 proteins (Figure 25). The editing of GC-rich regions in the HBV DNA genome by APOBEC3 proteins leads to C to U mutations in the negative strand and G to A mutations in the positive strand. This editing results in two possible outcomes. Either viruses survive the mutations in their HBV DNA genomes and thus have AT-rich regions instead of GC-rich regions in their HBV DNA genomes (Figure 26.1 -26.3), or the editing results in viruses that do not survive the mutations in their HBV DNA genomes and are degraded (Figure 26.4 -26.6). This project focused on understanding how APOBEC3 proteins exert their editing inhibition effect on hepatitis B virus replication. Despite several studies conducted to understand the frequency and extent of APOBEC3 protein mutations in the HBV DNA genome, there are opposing claims. Using two approaches, this project examined the frequency and extent of APOBEC3 protein mutations in the HBV DNA genome. The first approach determined the mutations in the HBV DNA genome by Clonal Sanger sequencing using the Differential DNA Denature PCR (3D-PCR) method. The second approach used a primer-ID Next Generation Sequencing (NGS) method.

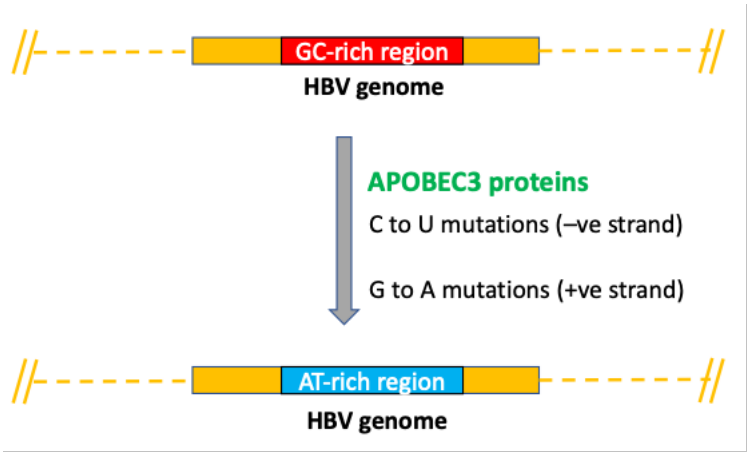


Figure 25. The diagram illustrates the APOBEC3 proteins edition of GC-rich regions in the HBV genomic DNA.

This editing leads to C to U mutations in the negative strand and G to A mutations in the positive strand, resulting in HBV genomic DNA with AT-rich regions instead of GC-rich regions.

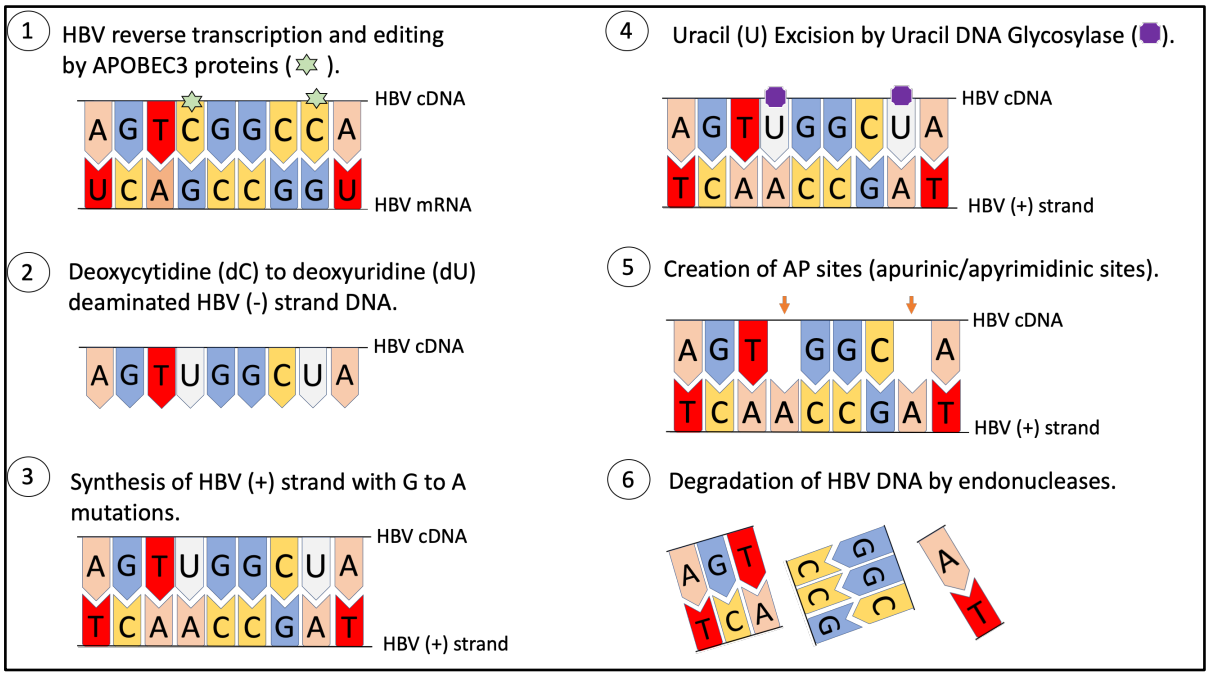


Figure 26. Diagram explains how APOBEC3 proteins can edit single-strand DNA during reverse transcription.

APOBEC3 proteins edit HBV during reverse transcription, which leads to HBV degradation. 1) HBV reverse transcription and editing by APOBEC3 proteins 2) Deoxycytidine (dC) to deoxyuridine (dU) deaminated HBV (-) strand DNA. 3) Synthesis of HBV (+) strand with G to A mutations. 4) Uracil (U) Excision by Uracil DNA Glycosylase. 5) Creation of AP sites (apurinic/apyrimidinic sites). 6) Degradation of HBV DNA by endonucleases.

3.4.1 3D-PCR analysis of the frequency and extent of APOBEC3 proteins mutations in HBV

To examine the frequency and extent of mutations caused by APOBEC3 proteins in the HBV DNA genome in the surviving viruses, a standard (i.e. one round) PCR method was used by Turelli *et al. (*Turelli, Mangeat *et al. 2004)* could not detect significant G to A hypermutations in the HBV genome. Subsequently, Suspène *et al. (*Suspène, Henry *et al. 2005)* developed a two-round PCR method to selectively amplify the AT-rich regions in the virus genome by lowering the denaturing temperature of the second-round PCR. This method was named Differential DNA Denaturation PCR (3D-PCR). In this method, PCR was carried out using Taq polymerase, which is not a high-fidelity polymerase, as high-fidelity polymerases do not amplify uracil-containing DNA. This project has introduced a new approach using the 3D-PCR method by using Phusion U polymerase, a high-fidelity polymerase that can amplify uracil-containing DNA. Using a high-fidelity polymerase is important to eliminate the errors caused by Taq polymerase in the amplified HBV DNA sequences and amplify the uracil in the edited DNA. It was previously shown in Huh 7 co-transfected with an HBV plasmid and APOBEC3G plasmids that 3D-PCR products were recovered from reactions carried out at as low as 90ºC or 88 ºC (Suspene, Guetard *et al.* 2005). The HBV-producing HepG2.2.15 cell line in this project was transfected with APOBEC3 expression plasmids. cytoplasmic HBV DNA was recovered from the HepG2.2.15 cells 72 hours post-transfection with APOBEC3 expression plasmids. 3D-PCR was used to analyse the frequency and extent of APOBEC3 deamination of the HBV genome. A 167-nucleotide segment of the X gene of the HBV genome was amplified at 88 ºC using the 3D-PCR method and cloned and sequenced (Figure 27). Ten clones of the products of the 3D-PCR second-round amplification were cloned and sequenced. These ten clones represent an independent individual molecular event of the mutated HBV genome DNA by APOBEC3 proteins.

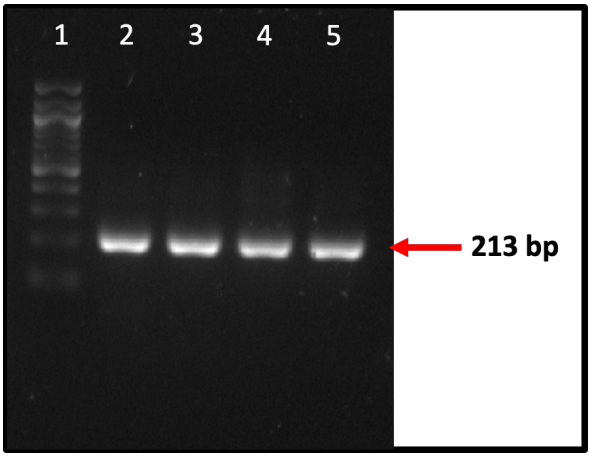


Figure 27. The agarose gel image shows the 3D-PCR products of the X region of HBV extracted from the HepG2.2.15 cell line.

HBV-producing HepG2.2.15 cells were transfected with APOBEC3 expression plasmids, and the cytoplasmic HBV DNA was extracted 72 hours post-transfection. The 3D-PCR technique was employed to amplify a 167 bp segment 5' to the X gene of the HBV genome using specific first outside primers and second inside primers. These products shown are from the second round of 3D-PCR amplification at 88°C; these products were then capture-cloned into transfer vector pCR2.1, and ten clones were sequenced using the M13 Reverse sequencing primers. Lane 1:100bp ladder and HBV DNA from 2: Mock transfected HepG2.2.15 cells, 3: Empty vector transfected HepG2.2.15 cells, 4: A3DE transfected HepG2.2.15 cells, 5: A3G transfected HepG2.2.15 cells.

Table 8 shows the lengths of the HBV (ayw) X region sequenced and the number of Guanines in these regions of the HBV genomic DNA. It also shows the numbers of the other nucleotides in these sequences. This information was used to help determine the number of guanine nucleotides that may transition to adenine. It also allows the determination of the percentage of G>A mutations in the sequences amplified by 3D-PCR.

Table 8. shows the length of the HBV (ayw) X sequence and the number of Adenine, Guanine, Cytosine and Thymine in the HBV (ayw) X sequence.

It also shows the number of Guanines that potentially may undergo Guanine to Adenine (G>A) transitions due to HBV editing by APOBEC3 family members.

	Length	*Adenine*	*Guanine*	*Cytosine*	*Thymine*
HBV (ayw) X sequence	*167*	*14 (8.4%)*	*44 (26.35%)*	*65 (38.9%)*	*44 (26.35%)*

3.4.1.1 Taq polymerase-based 3D PCR analysis of APOBEC3-induced mutations in HBV DNA

Cells were either Mock transfected, transfected with an empty vector, or transfected with APOBEC3DE or APOBEC3G for 72 hours, after which the intracellular HBV DNA was extracted and amplified by Taq polymerase-based 3D PCR technique. Ten clones from each transfection were sequenced, and their sequences were aligned to wild type HBV X region sequence. This HBV X region sequence is 167 nucleotides long and contains 58 guanines. G to A mutations in the X region of the extracted intracellular HBV DNA genomic are shown in empty transfected (Figure 28), APOBEC3G transfected (Figure 29), and APOBEC3DE transfected (Figure 30).

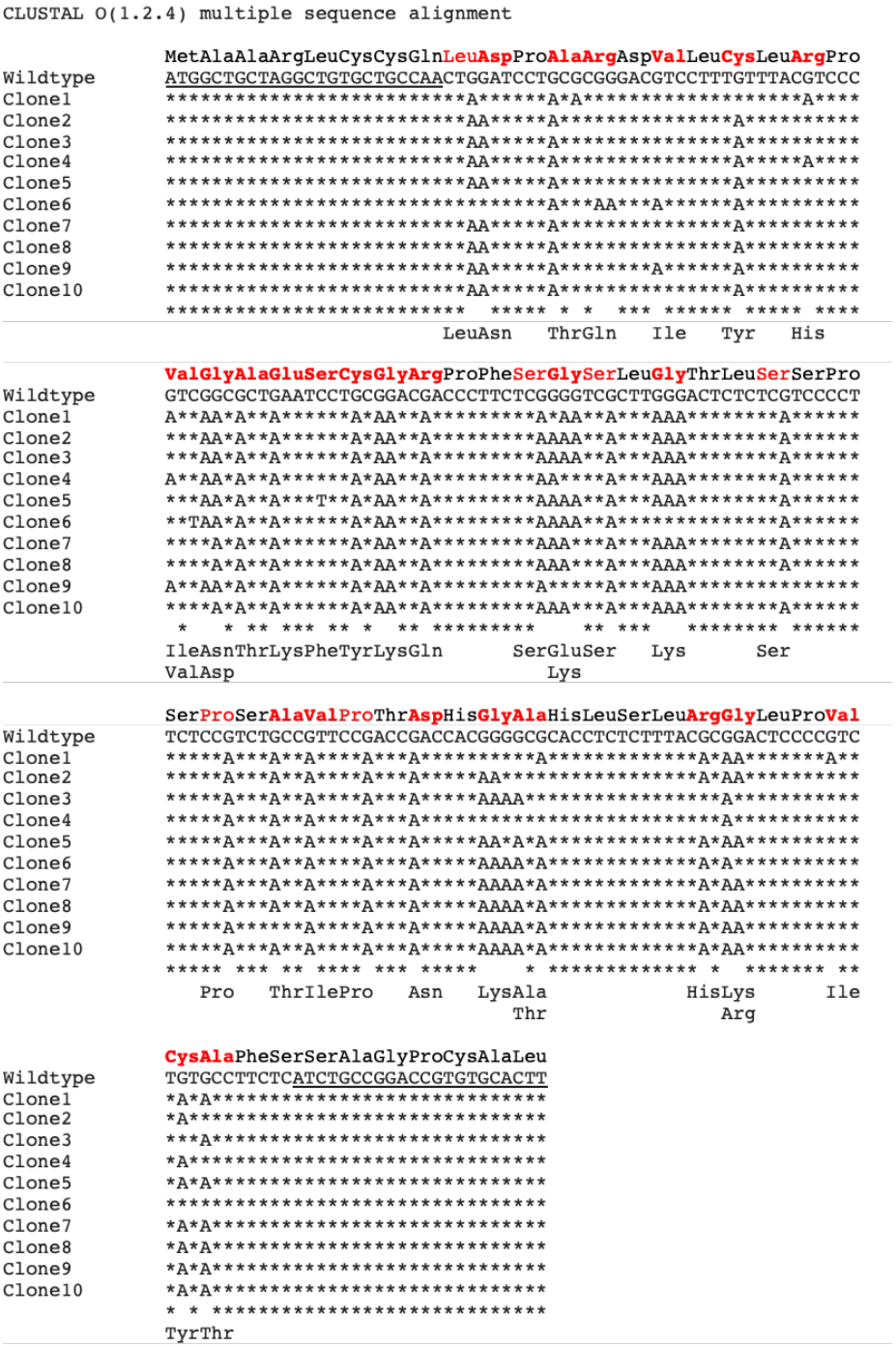


Figure 28. G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the empty vector transfected HepG2.2.15 cell line.

The wild type (HBV ayw) sequence (top line) was used as a reference sequence for the aligned sequences of the ten clones amplified by Taq polymerase-based 3D-PCR. Underlined regions are PCR primer sequences. (*) Asterix means that there is no nucleotide difference between the clones' sequences and the reference sequence. Bold Red amino acids are changed due to G to A base changes. Red amino acids are not changed, but the codon is changed.

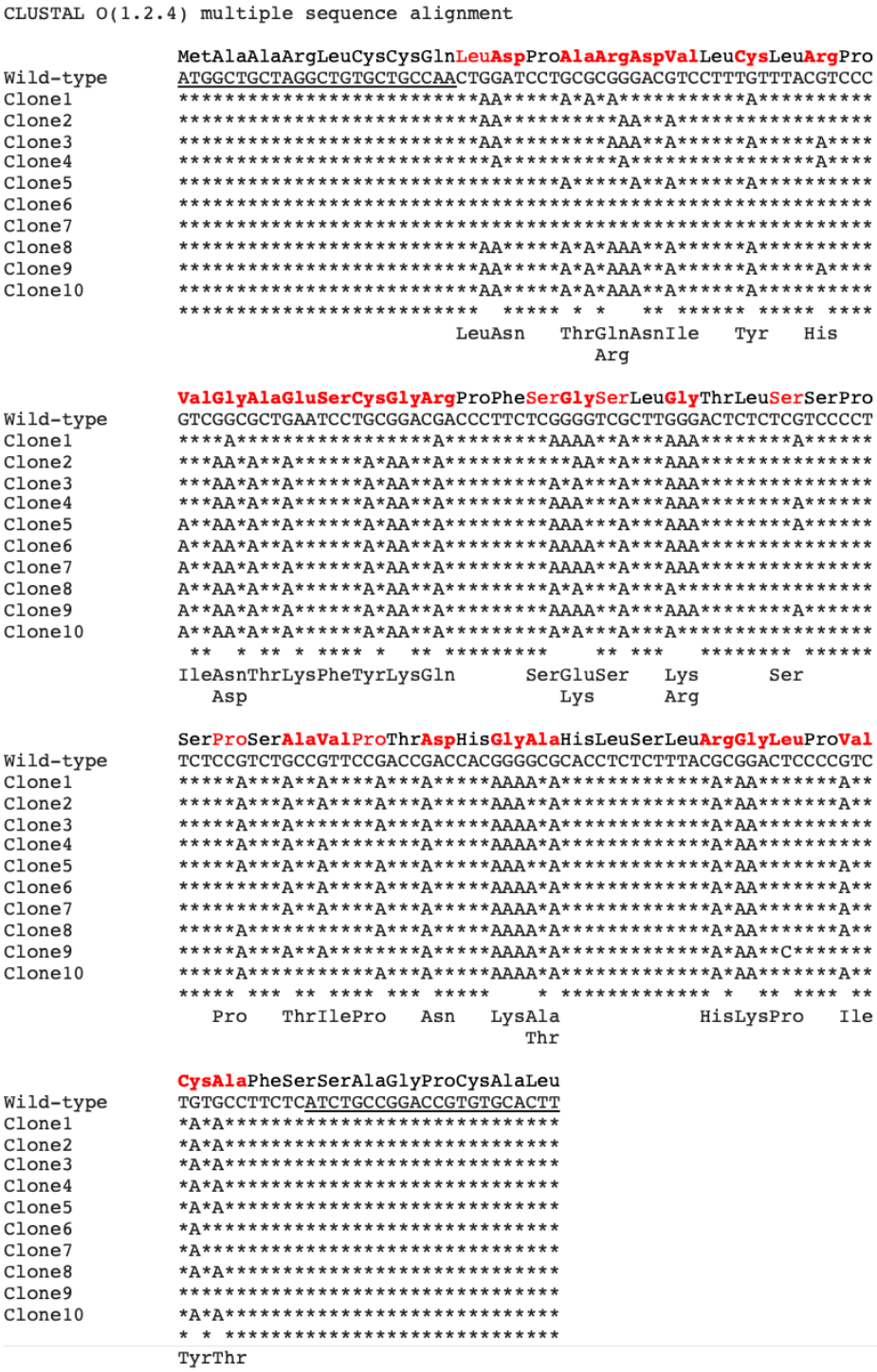


Figure 29. G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the APOBEC3G transfected HepG2.2.15 cell line.

The wild type (HBV ayw) sequence (top line) was used as a reference sequence for the aligned sequences of the ten clones amplified by Taq polymerase-based 3D-PCR. Underlined regions are PCR primer sequences. (*) Asterix means that there is no nucleotide difference between the clones' sequences and the reference sequence. Bold Red amino acids are changed due to G to A base changes. Red amino acids are not changed, but the codon is changed.

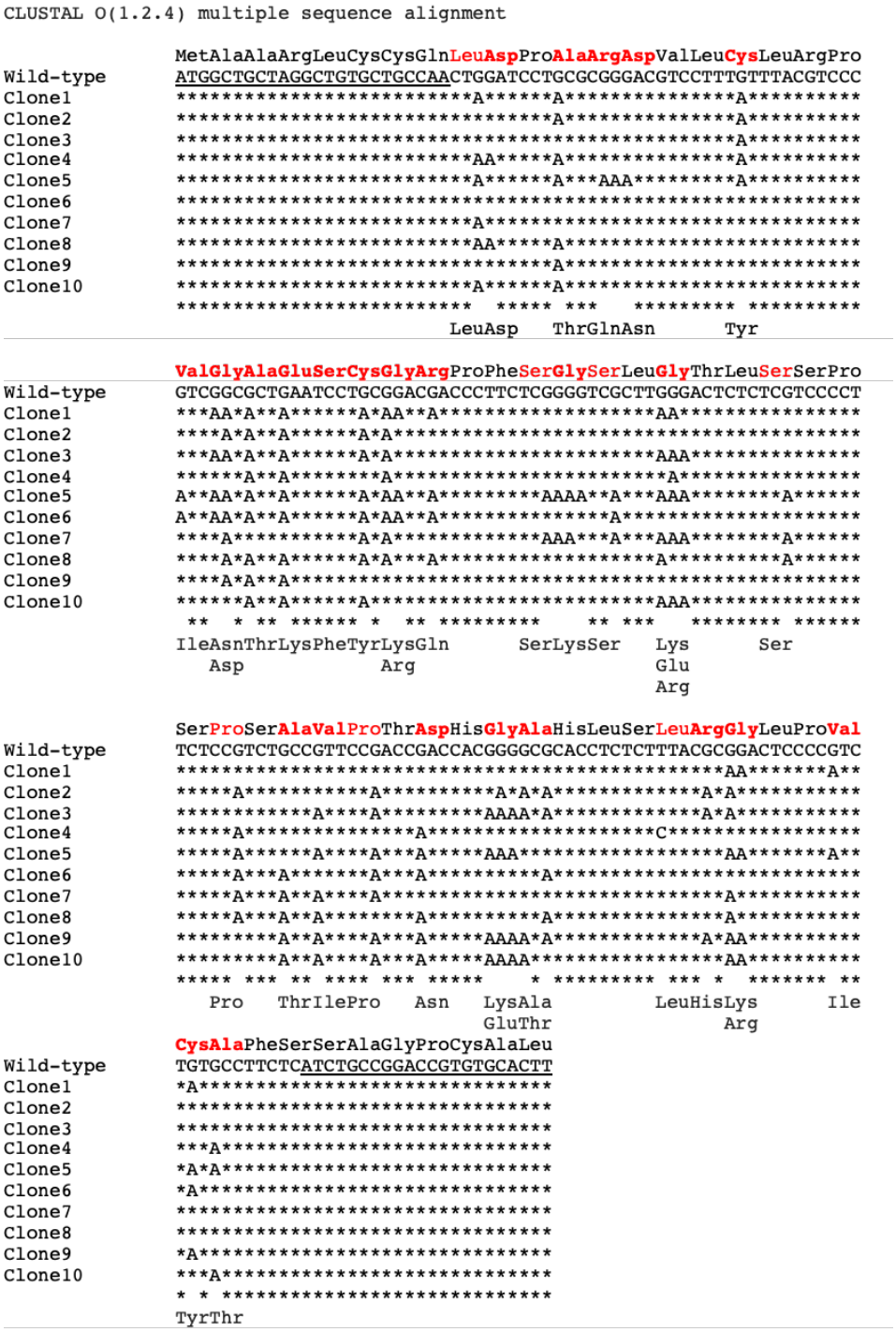


Figure 30 G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the APOBEC3DE transfected HepG2.2.15 cell line.

The wild type (HBV ayw) sequence (top line) was used as a reference sequence for the aligned sequences of the ten clones amplified by Taq polymerase-based 3D-PCR. Underlined regions are PCR primer sequences. (*) Asterix means that there is no nucleotide difference between the clones' sequences and the reference sequence. Bold Red amino acids are changed due to G to A base changes. Red amino acids are not changed, but the codon is changed.

Table 9 represents the number of G to A mutations as well as the percentages of the G to A mutations in the 167 nucleotides of the X region of HBV that was extracted from HepG2.2.15 cell line and was amplified by Taq polymerase-based 3D-PCR.10 clones were sequenced. The Hypermute package programme was used to analyse the numbers and percentages of G to A mutations in the X region in comparison to the reference sequence, which is wild type hepatitis B virus (ayw). In the A3G transfected HepG2.2.15 cell lines, the number of substitutions of G to A mutations ranges from 31 to 40 Guanines out of 44 Guanines in the X region, and the percentage ranges from 70.45% to 68.97% of Guanines were substituted by Adenines whereas the A3DE transfected HepG2.2.15 cell lines, the number of substitutions of G to A mutations ranges from 11 to 36 Guanines out of 44 Guanines in the X region and the percentage ranges from 25% to 81.82% of Guanines were substituted by Adenines. It was also noted that the empty vector transfected HepG2.2.15 cell lines show Guanines into Adenines substitutions, the number of substitutions of G to A mutations ranges from 28 to 35 Guanines out of 44 Guanines in the X region, and the percentage ranges from 63.64% to 79.55% of Guanines were substituted by Adenines (Table 9).

Table 9. Nucleotide Substitutions of Clones Amplified by Taq polymerase-based 3D-PCR in HepG2.2.15 cells.

HepG2.2.15 cells were transfected with the APOBEC3G, APOBEC3DE or empty vector. The number of G to A substitutions and the percentage of G to A transitions per clone were shown.

APOBEC3G transfected HepG2.2.15 Cell Lines. The sequenced clones of the HBX region amplified by Taq polymerase-based 3D-PCR.

	No. of G>A substitutions	% of G>A transitions
Clone1	33	75
Clone 2	34	77.27
Clone 3	36	81.82
Clone 4	33	75
Clone 5	36	81.82
Clone 6	31	70.45
Clone 7	31	70.45
Clone 8	36	81.82
Clone 9	40	90.91
Clone 10	36	81.82

APOBEC3DE transfected HepG2.2.15 Cell Lines. The sequenced clones of the HBX region amplified by Taq polymerase-based 3D-PCR.

	No. of G>A substitutions	% of G>A transitions
Clone1	17	38.64
Clone 2	14	31.82
Clone 3	19	43.18
Clone 4	11	25
Clone 5	36	81.82
Clone 6	16	36.36
Clone 7	17	38.64
Clone 8	17	38.64
Clone 9	17	38.64
Clone 10	19	43.18

Empty vector transfected HepG2.2.15 Cell Lines. The sequenced clones of the HBX region amplified by Taq polymerase-based 3D-PCR.

	No. of G>A substitutions	% of G>A transitions
Clone1	33	75
Clone 2	32	72.73
Clone 3	32	72.73
Clone 4	28	63.64
Clone 5	35	79.55
Clone 6	31	70.45
Clone 7	34	77.27
Clone 8	34	77.27
Clone 9	33	75
Clone 10	34	77.27

3.4.1.2 Phusion U polymerase-based 3D PCR analysis of APOBEC3 induced mutations in HBV DNA

Cells were either Mock transfected, transfected with an empty vector, or transfected with APOBEC3DE or APOBEC3G for 72 hours, after which the intracellular HBV DNA was extracted and amplified by Phusion U polymerase-based 3D PCR technique. Ten clones from each transfection were sequenced, and their sequences were aligned to wild type HBV X region sequence. This HBV X region sequence is 167 nucleotides long and contains 44 guanines. G to A mutations in the X region of the extracted intracellular HBV DNA genomic are shown in empty transfected (Figure 31), APOBEC3G transfected (Figure 32), and APOBEC3DE transfected (Figure 33).

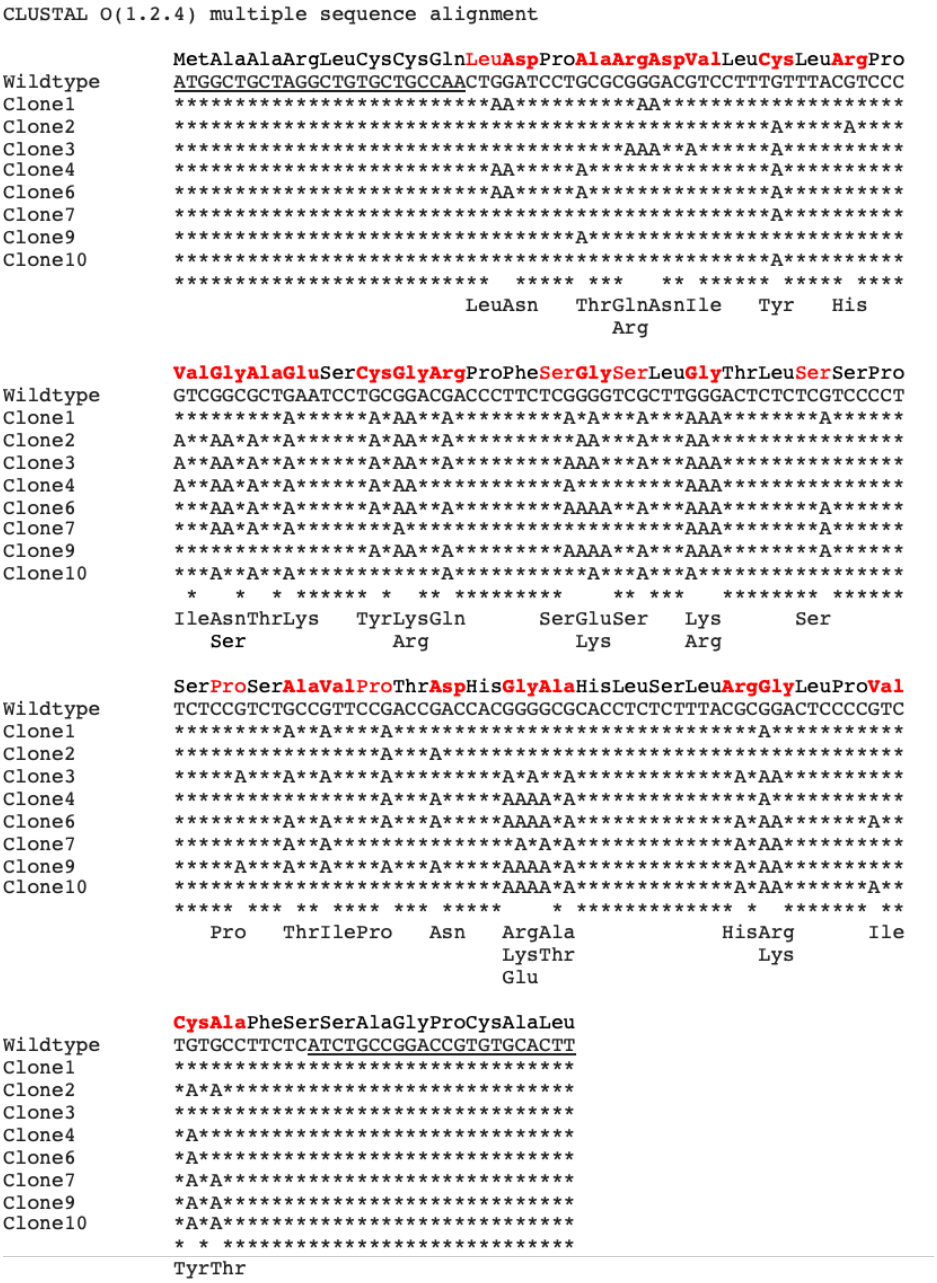


Figure 31. G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the empty vector transfected HepG2.2.15 cells.

The wild type (HBV ayw) sequence (top line) was used as a reference sequence for the aligned sequences of the ten clones amplified by Phusion U polymerase-based 3D-PCR. Underlined regions are PCR primer sequences. (*) Asterix means that there is no nucleotide difference between the clones' sequences and the reference sequence. Bold Red amino acids are changed due to G to A base changes. Red amino acids are not changed, but the codon is changed.

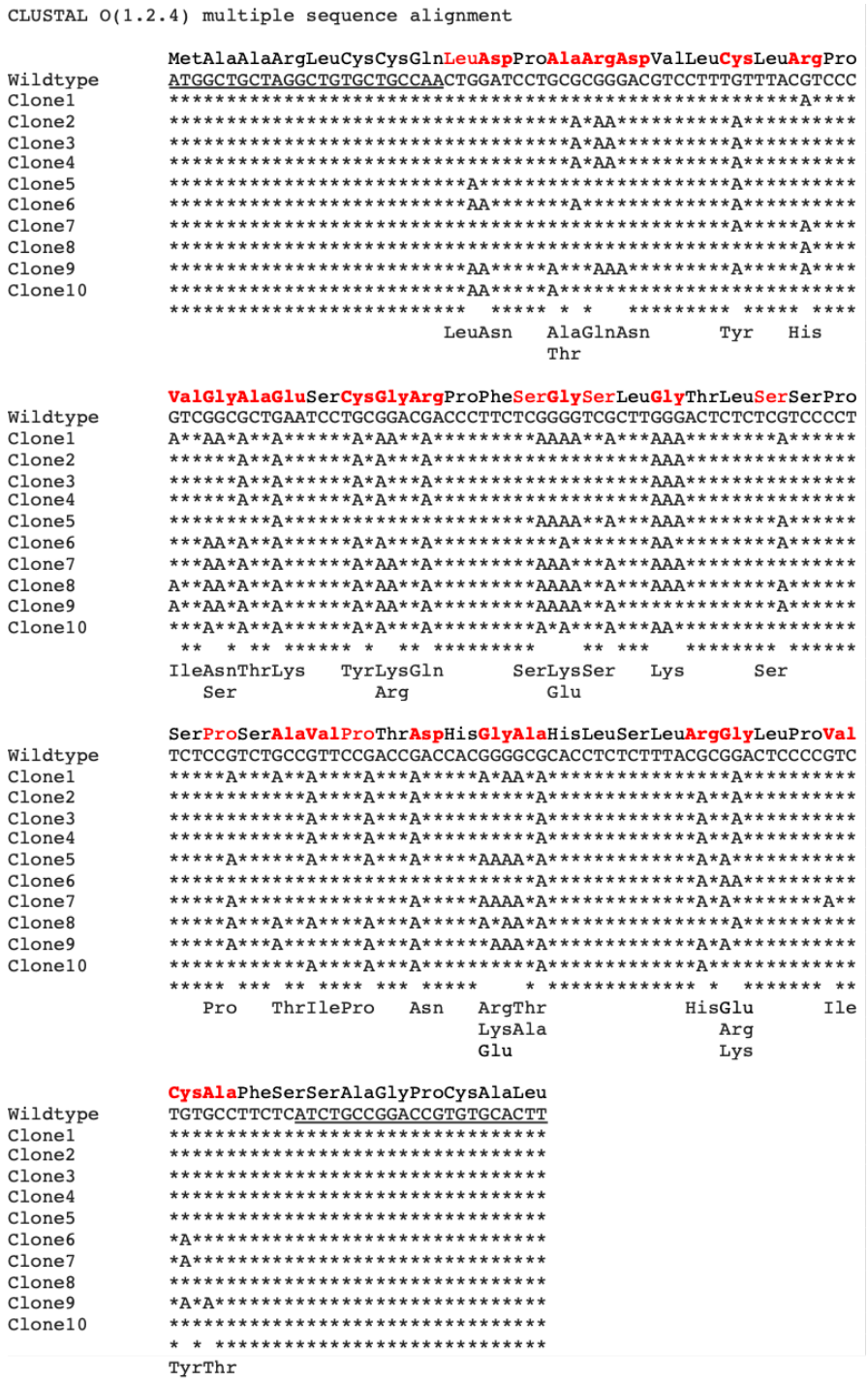


Figure 32. G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the APOBEC3G transfected HepG2.2.15 cells.

The wild type (HBV ayw) sequence (top line) was used as a reference sequence for the aligned sequences of the ten clones amplified by Phusion U polymerase-based 3D-PCR. Underlined regions are PCR primer sequences. (*) Asterix means that there is no nucleotide difference between the clones' sequences and the reference sequence. Bold Red amino acids are changed due to G to A base changes. Red amino acids are not changed, but the codon is changed.

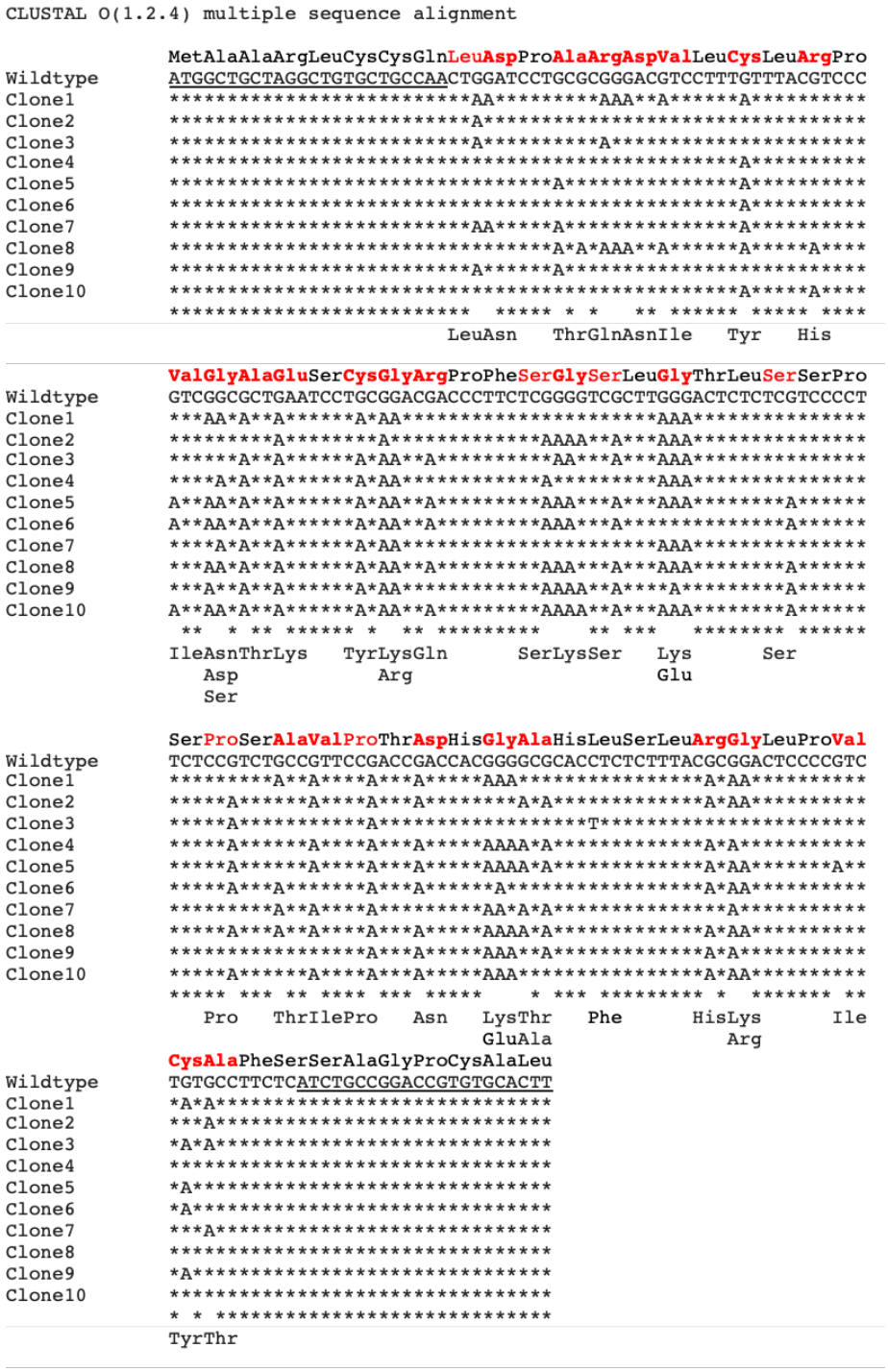


Figure 33. G to A mutations were observed in sequences (X-region) of the cytoplasmic HBV DNA extracted from the APOBEC3DE transfected HepG2.2.15 cells.

The wild type (HBV ayw) sequence (top line) was used as a reference sequence for the aligned sequences of the ten clones amplified by Phusion U polymerase-based 3D-PCR. Underlined regions are PCR primer sequences. (*) Asterix means that there is no nucleotide difference between the clones' sequences and the reference sequence. Bold Red amino acids are changed due to G to A base changes. Red amino acids are not changed, but the codon is changed.

Table 10 represents the number of G to A mutations, and the percentages of the G to A mutations in the 167 nucleotides of the X region of HBV extracted from HepG2.2.15 cell line and amplified by Phusion U polymerase-based 3D-PCR.10 clones were sequenced. The Hypermute package programme was used to analyse the numbers and percentages of G to A mutations in the X region compared to the reference sequence, which is wild-type hepatitis B virus (ayw). In the A3G transfected HepG2.2.15 cell lines, the number of G to A substitutions ranges from 18 to 35 Guanines out of 44 Guanines in the X region. The percentage ranges from 40.91% to 79.55% of Guanines were substituted by Adenines, whereas the A3DE transfected HepG2.2.15 cell lines, the number of substitutions of G to A mutations ranges from 15 to 37 Guanines out of 44 Guanines in the X region and the percentage ranges from 34.09% to 84.09% of Guanines were substituted by Adenines. It was also noted that the empty vector transfected HepG2.2.15 cell lines show Guanines into Adenines substitutions, the number of substitutions of G to A mutations ranges from 11 to 31 Guanines out of 44 Guanines in the X region, and the percentage ranges from 24.99% to 70.45% of Guanines were substituted by Adenines (Table 10).

Table 10. Nucleotide Substitutions of Clones Amplified by Phusion U polymerase-based 3D-PCR in HepG2.2.15 cells.

HepG2.2.15 cell**s** were transfected with the APOBEC3G, APOBEC3DE or empty vector. The number of G to A substitutions and the percentage of G to A transitions per clone were shown.

APOBEC3G transfected HepG2.2.15 Cell Lines. The sequenced clones of the HBX region amplified by Phusion U polymerase-based 3D-PCR.

	No. of G>A substitutions	% of G>A transitions
Clone1	29	65.91
Clone 2	18	40.91
Clone 3	18	40.91
Clone 4	18	40.91
Clone 5	23	52.27
Clone 6	20	45.45
Clone 7	28	63.64
Clone 8	29	65.91
Clone 9	35	79.55
Clone 10	19	43.18

APOBEC3DE transfected HepG2.2.15 Cell Lines. The sequenced clones of the HBX region amplified by Phusion U polymerase-based 3D-PCR.

	No. of G>A substitutions	% of G>A transitions
Clone1	29	65.91
Clone 2	21	47.73
Clone 3	15	34.09
Clone 4	22	50
Clone 5	33	75
Clone 6	24	54.55
Clone 7	22	50
Clone 8	37	84.09
Clone 9	24	54.55
Clone 10	30	68.18

Empty vector transfected HepG2.2.15 Cell Lines. The sequenced clones of the HBX region amplified by Phusion U polymerase-based 3D-PCR.

	No. of G>A substitutions	% of G>A transitions
Clone1	17	38.63
Clone 2	20	45.45
Clone 3	31	70.45
Clone 4	25	56.82
Clone 5	11	24.99
Clone 6	22	49.99
Clone 7	20	45.45
Clone 8	11	24.99
Clone 9	29	65.91
Clone 10	19	43.18

Table 11 shows a summarised comparison of the number and percentages of G to A mutations in the 167 nucleotides of the X region of HBV amplified by Phusion U polymerase-based 3D-PCR compared to those amplified by Taq polymerase-based 3D-PCR. Ten clones of each were sequenced. The Hypermute package programme was used to analyse the numbers and percentages of G to A mutations in the 10 (X region) clones from each amplification compared to the wild type of hepatitis B virus (ayw) reference sequence. The total number and percentages of G to A mutations shown by Taq polymerase-based 3D-PCR were 346 (78.63%), 183 (41.59%), and 326 (74.09%) of G to A mutations in the A3G, A3DE and empty vector transfected HepG2.2.15, respectively. While the total number and percentages of G to A mutations shown by Phusion U polymerase-based, 3D-PCR were 237 (53.86%), 257 (58.41%) and 205 (46.58%) of G to A mutations in the A3G, A3DE and empty vector transfected HepG2.2.15, respectively (Table 11).

Table 11. a summarised comparison of the number and percentages of G to A mutations in the X region of the extracted cytoplasmic HBV genomic DNA.

The total number and percentages of Guanine to Adenine (G>A) substitutions in the sequenced ten clones of the 167 nucleotides of the X region of the extracted cytoplasmic HBV genomic DNA amplified by Phusion U polymerase-based 3D-PCR were compared to those amplified by Taq polymerase-based 3D-PCR.

	Taq polymerase-based 3D-PCR (Total No. of G>A substitutions) *	Phusion U polymerase-based 3D-PCR (Total No. of G>A substitutions) *
A3G	346 (78.63%)	237 (53.86%)
A3DE	183 (41.59%)	257 (58.41%)
Empty vector	326 (74.09%)	205 (46.58%)

*The total number of substitutions in ten clones compared to the reference sequence.

3.4.2 Next-Generation Sequencing approach to the analysis of the frequency and extent of APOBEC3 proteins mutations in HBV DNA

Further investigation of the frequency and extent of APOBEC3DE and APOBEC3G mutations in HBV pre-core, X, and S regions were explored by primer-ID Next-Generation sequencing (NGS) to elucidate the effect of APOBEC3 proteins on HBV replication. Primer-ID Next-Generation sequencing (NGS) uses barcodes (primer-ID) that allow the amplification and identification of every amplified and sequenced segment as an independent individual molecular event compared to conventional NGS. The HBV-producing HepG2.2.15 cells were transfected with APOBEC3 expression plasmids for 72 hours. The cytoplasmic HBV DNA was recovered and used to determine the frequency and extent of APOBEC3DE and APOBEC3G mutations in HBV pre-core, X, and S regions in the extracted HBV genome DNA. To determine the frequency and extent of APOBEC3DE and APOBEC3G mutations in the extracted HBV DNA, HBV pre-core, X, and S regions were amplified with nested Phusion U Hot Start PCR, and the library for the Next Generation Sequencing was prepared using NEBNext® Ultra™ II DNA Library Prep Kit (Figure 34 and 35).

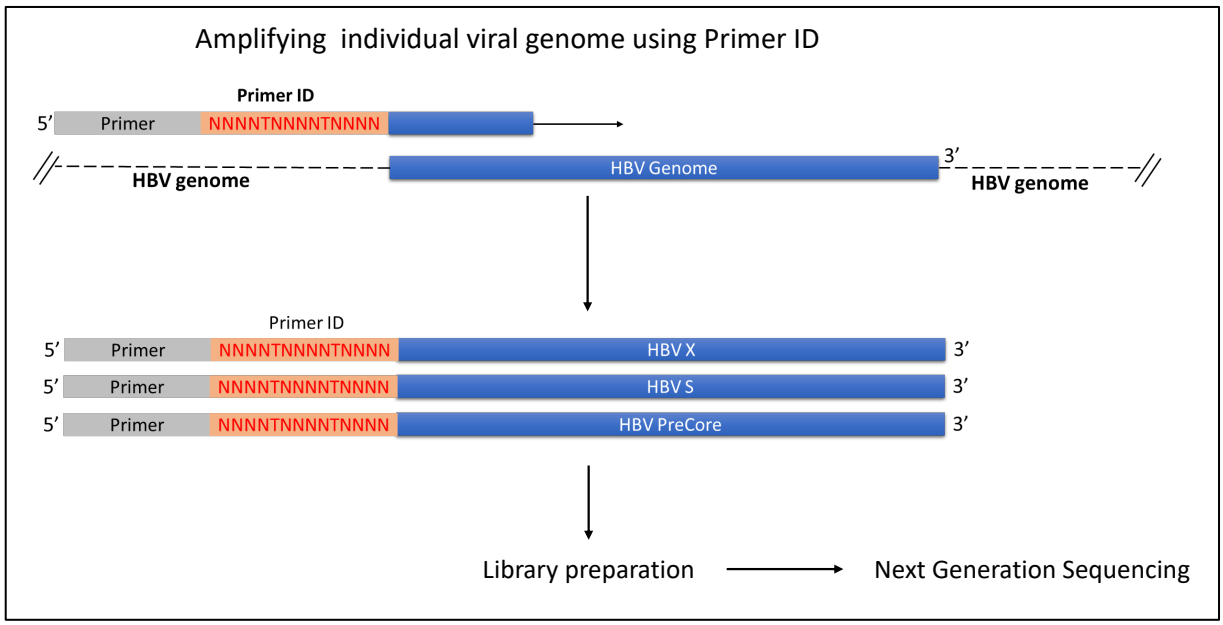


Figure 34. Next-Generation Sequencing approach to analysing the frequency and extent of APOBEC3 protein mutations in HBV DNA in Hep2.2.15 cells.

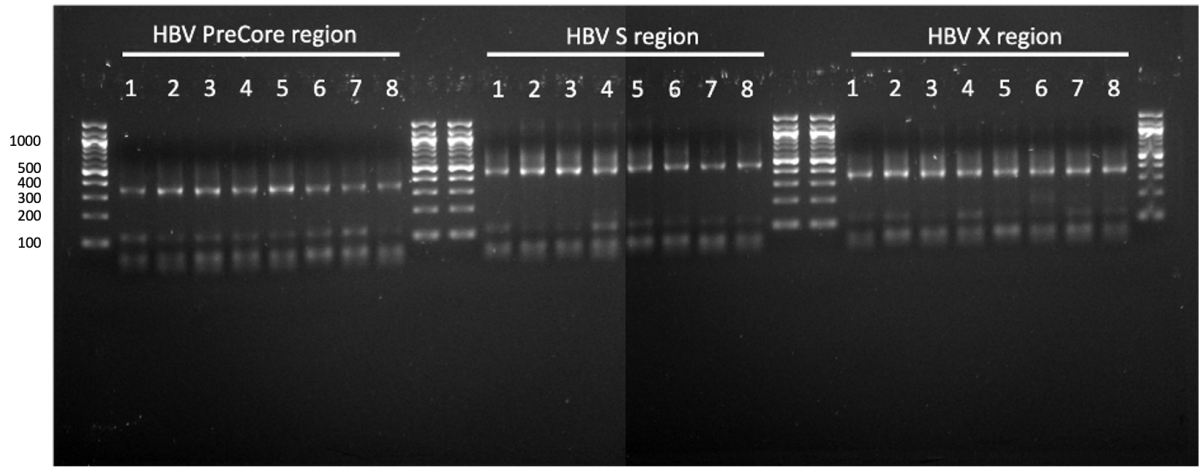


Figure 35. The agarose gel image shows the NGS library products of the HBV PreCore, S, and X regions extracted from APOBEC3 transfected HepG2.2.15 cells.

HepG2.2.15 cells were transfected with APOBEC3 expression plasmids, and the cytoplasmic HBV DNA was extracted 72 hours post-transfection. Lanes 1 and 4: Mock transfected HepG2.2.15 cells, 2 and 6: Empty vector transfected HepG2.2.15 cells, 3 and 7: A3DE transfected HepG2.2.15 cells, 4 and 8: A3G transfected HepG2.2.15 cells.

The number of G to A mutations and C to T mutations caused by APOBEC3G transfected HepG2.2.15 cells in the X region of HBV DNA were 980 G to A mutations and 98 C to T mutations. In contrast, the number of G to A mutations and C to T mutations caused by APOBEC3DE transfected HepG2.2.15 cells in the X region of HBV DNA were 886 G to A mutations and 88 C to T mutations. The number of G to A mutations and C to T mutations observed in the Empty vector transfected HepG2.2.15 cells in the X region of HBV DNA were 604 G to A mutations and 68 C to T mutations. The number of G to A mutations and C to T mutations observed in the Mock transfected HepG2.2.15 cells in the X region of HBV DNA were 633 G to A mutations and 49 C to T mutations (Table 12 and 15).

HBV X region

MOCK

A

→	A	C	T	G
A	0	4	5	8
C	31	0	49	8
T	11	27	0	6
G	633	7	31	0

Empty Vector

B

→	A	C	T	G
A	0	1	8	11
C	32	0	68	3
T	19	14	0	4
G	604	8	43	0

APOBEC3DE

C

→	A	C	T	G
A	0	0	8	6
C	48	0	88	8
T	19	16	0	7
G	886	7	37	0

APOBEC3G

D

→	A	C	T	G
A	0	9	7	12
C	44	0	98	2
T	24	23	0	9
G	980	4	50	0

Table 12. Nucleotide substitutions of the HBV X region sequence amplified by Next Generation Sequencing.

The HBV X region sequence was amplified by Next Generation Sequencing in HepG2.2.15 cells that were Mock, empty vector, APOBEC3DE, or APOBEC3G transfected. The number of G to A mutations and the number of C to T substitutions were shown.

The number of G to A mutations and C to T mutations caused by APOBEC3G transfected HepG2.2.15 cells in the S region of HBV DNA were 4048 G to A mutations and 160 C to T mutations. In contrast, the number of G to A mutations and C to T mutations caused by APOBEC3DE transfected HepG2.2.15 cells in the S region of HBV DNA were 3466 G to A mutations and 186 C to T mutations. The number of G to A mutations and C to T mutations observed in the Empty vector transfected HepG2.2.15 cells in the S region of HBV DNA were 2770 G to A mutations and 144 C to T mutations. The number of G to A mutations and C to T

mutations observed in the Mock transfected HepG2.2.15 cells in the S region of HBV DNA were 1960 G to A mutations and 70 C to T mutations (Table 13 and 15).

HBV S region

MOCK

A

→	A	C	T	G
A	0	10	4	17
C	11	0	70	2
T	13	38	0	6
G	1960	1	18	0

Empty Vector

B

→	A	C	T	G
A	0	24	5	29
C	47	0	144	10
T	10	61	0	14
G	2770	7	41	0

APOBEC3DE

C

→	A	C	T	G
A	0	32	6	36
C	71	0	186	7
T	13	56	0	20
G	3466	5	49	0

APOBEC3G

D

→	A	C	T	G
A	0	24	7	22
C	62	0	160	11
T	16	63	0	11
G	4084	4	40	0

Table 13. Nucleotide substitutions of the HBV S region sequence amplified by Next Generation Sequencing.

The HBV S region sequence was amplified by Next Generation Sequencing in HepG2.2.15 cells that were APOBEC3G, APOBEC3DE, empty vector or Mock transfected. The number of G to A mutations and the number of C to T substitutions were shown.

The number of G to A mutations and the number of C to T mutations caused by APOBEC3G transfected HepG2.2.15 cells in the PreCore region of HBV DNA were 189 G to A mutations and 62 C to T mutations. In contrast, the number of G to A mutations and the number of C to T mutations caused by APOBEC3DE transfected HepG2.2.15 cells in the PreCore region of HBV DNA were 60 G to A mutations and 35 C to T mutations. The number of G to A mutations and C to T mutations observed in the Empty vector transfected HepG2.2.15 cells in the PreCore

region of HBV DNA were 85 G to A mutations and 32 C to T mutations. The number of G to A mutations and C to T mutations observed in the Mock transfected HepG2.2.15 cells in the PreCore region of HBV DNA were 289 G to A mutations and 72 C to T mutations (Table 14 and 15).

HBV PreCore region

A

MOCK

→	A	C	T	G
A	0	14	39	35
C	42	0	72	17
T	21	54	0	15
G	289	4	30	0

B

Empty Vector

→	A	C	T	G
A	0	9	18	5
C	17	0	32	8
T	8	25	0	4
G	85	1	22	0

C

APOBEC3DE

→	A	C	T	G
A	0	10	19	14
C	17	0	35	6
T	6	22	0	8
G	60	5	14	0

D

APOBEC3G

→	A	C	T	G
A	0	15	41	23
C	23	0	62	13
T	17	35	0	10
G	189	3	18	0

Table 14. Nucleotide substitutions of the HBV PreCore region sequence amplified by Next Generation Sequencing.

The HBV PreCore region sequence was amplified by Next Generation Sequencing in HepG2.2.15 cell line that was APOBEC3G, APOBEC3DE, empty vector or Mock transfected. The number of G to A mutations and the number of C to T substitutions were shown.

Table 15. a summarised comparison of the number of G to A and the number of C to T substitutions in the HBV X, S, and PreCore regions sequences amplified by Next Generation Sequencing.

HepG2.2.15 cell line that was APOBEC3G, APOBEC3DE, empty vector or Mock transfected. The HBV X, S, and PreCore regions sequences were amplified by Next Generation Sequencing. The number of G to A mutations and the number of C to T substitutions were shown.

No. of G>A substitutions

	HBV X region	HBV S region	HBV PreCore region
A3G	980	4084	189
A3DE	886	3466	60
Empty vector	604	2770	85
Mock	633	1960	289

No. of C>T substitutions

	HBV X region	HBV S region	HBV PreCore region
A3G	98	160	62
A3DE	88	186	35
Empty vector	68	144	32
Mock	49	70	72

3.5 Human Upstream stimulatory factor 1(USF1) effect on APOBEC3 proteins expression levels and the effect on HBV replication.

USF1 can take part in basal transcription regulation of the human A3G gene in hepatocytes by binding to identified E-box of the A3G promoter (Zeng, Li *et al.* 2016).

3.5.1 Human Upstream stimulatory factor 1(USF1) upregulates APOBEC3DE and APOBEC3G proteins

The effect of USF1 on APOBEC3 family members was examined by transfecting HepG2.2.15 cells with a USF1 expression plasmid for 48 hours. Post-transfection, Total RNA was extracted, and the overexpression of USF1 was measured using quantitative Real-Time Polymerase Chain Reaction (RT-qPCR). The expression of USF1 was 2376-fold in USF1-transfected HepG2.2.15 cells compared to the untransfected HepG2.2.15 cells (Figure 36). The APOBEC3DE and APOBEC3G gene expression increased by 4.8-fold change and 4-fold change, respectively, while APOBEC3B and APOBEC3F gene expression increased by 1.4-fold change and 1.9-fold change. APOBEC3DE and APOBEC3G showed higher gene expression than the other APOBEC3 family members (Figure 37).

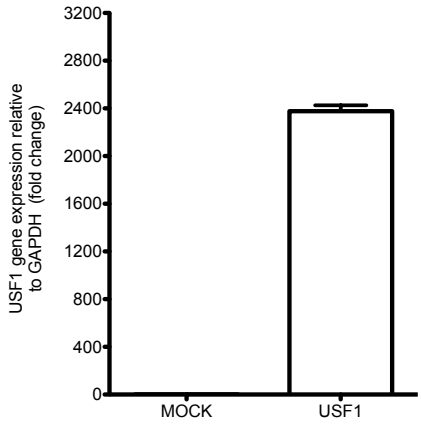


Figure 36. Human Upstream stimulatory factor 1(USF1) overexpression in HepG2.2.15 cells.

HepG2.2.15 cells were transfected with a USF1 expression vector, and 48 hours post-transfection, Total RNA was extracted, and the fold change of USF1 was quantified with real-time qPCR. The fold change of USF1 gene expression level is compared to the untransfected samples, and the data represent the mean (+/-SD), N=9. Thus, the data are representative of three independent experiments.

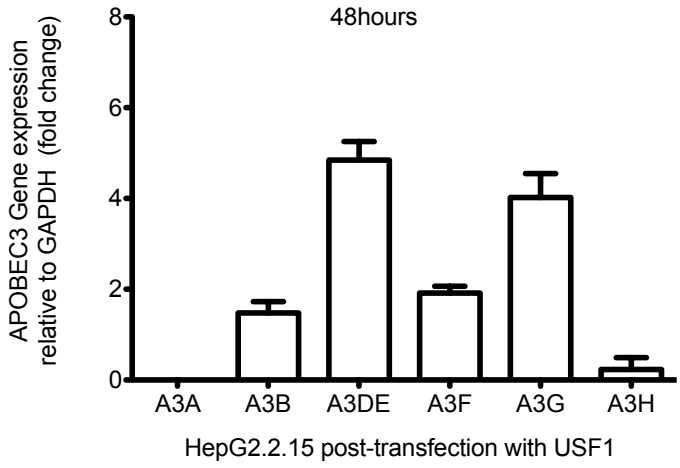


Figure 37. Human Upstream stimulatory factor 1 (USF1) upregulates APOBEC3DE AND APOBEC3G in HepG2.2.15 cells.

HepG2.2.15 cells were transfected with a USF1 expression vector, and 48 hours post-transfection, Total RNA was extracted, and the fold change of APOBEC3 family members was quantified with real-time qPCR. The fold change of APOBEC3 family members' gene expression levels is shown in comparison to the untransfected samples with the USF1 expression vector, and the data represent the mean (+/-SD), N=9. Thus, the data are representative of three independent experiments.

3.5.2 Human Upstream stimulatory factor 1(USF1) enhance HBV replication

The Effect of USF1 on HBV replication was determined by transfecting HepG2.2.15 cell with a USF1 expression plasmid for 48 hours. Post-transfection, the extracellular virion-associated HBV DNA was extracted from the supernatant of USF1 transfected HBV-producing HepG2.2.15 cells. Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR) was used to measure the effects of USF1 on the extracellular virion-associated HBV DNA levels in HBV producing HepG2.2.15 cells (Figure 38). USF1 enhanced extracellular HBV virion replication by 25 % (Figure 39).

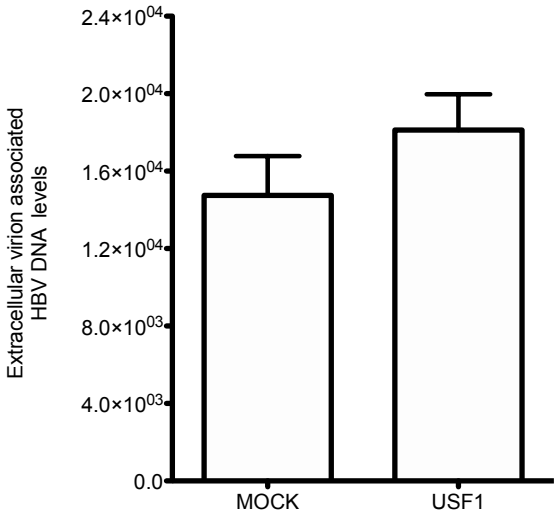


Figure 38. Human Upstream stimulatory factor 1(USF1) enhance HBV replication in HepG2.2.15 cells.

HepG2.2.15 cells were transfected with USF1 expression vector, and 48-hours post-transfection, extracellular virion-associated HBV DNA was extracted and quantified with real-time qPCR using TaqMan Assay (forward primer, reverse primer, and probe). The extracellular virion-associated HBV DNA levels are shown as compared to the untransfected samples, and the data represent the mean (+/-SD), N=9. Thus, the data are representative of three independent experiments.

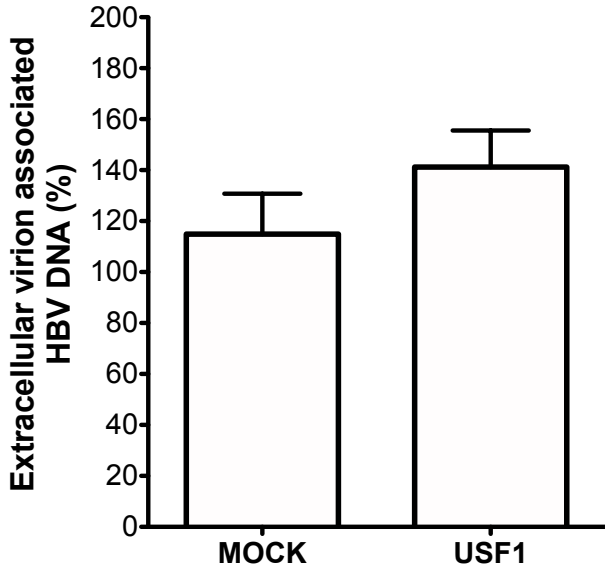


Figure 39. Human Upstream stimulatory factor 1(USF1) enhance HBV replication in HepG2.2.15 cells.

HepG2.2.15 cells were transfected with USF1 expression vector, and 48-hours post-transfection, extracellular virion-associated HBV DNA was extracted and quantified with real-time qPCR using TaqMan Assay (forward primer, reverse primer, and probe). The extracellular virion-associated HBV DNA levels are shown as percentages of the untransfected samples, and the data represent the mean (+/-SD), N=9. Thus, the data are representative of three independent experiments.

4 Chapter Four: Discussion

This project investigated the effect of APOBEC3 proteins HBV replication. The APOBEC3 family members are described as mutator proteins and have been shown to reduce the levels of HBV production in infected HBV culture systems; it needs to be clarified to what extent APOBEC3 proteins can target replicating HBV genomes in the cytoplasm. The objectives, findings, discussions, and conclusions of this study were:

1) To construct GFP-tagged APOBEC3 expressing plasmids, characterising their optimal expression and the intracellular localisation of APOBEC3 proteins.
2) To investigate the impact of APOBEC3 family members on the cytoplasmic and extracellular virion-associated HBV DNA levels using Quantitative Polymerase Chain Reaction (qPCR).
3) The analysis of the frequency and extent of APOBEC3 mutations in HBV DNA the cytoplasmic HBV DNA using 3D-PCR, Modified 3D-PCR and prime ID NGS.
4) An investigation of the effects of the Human Upstream stimulatory factor 1(USF1) on APOBEC3 family members and HBV replication

APOBEC3 proteins exert their deamination effect on single-strand DNA during the reverse transcription step of the HBV replication cycle and lead to the formation of uracil-containing HBV DNA. Subsequently, this leads to G to A mutations in the positive strand of HBV DNA and C to T mutations in the negative strand. The fate of the APOBEC3 mutated virus genomes is that either they are degraded or remain intact with mutations and are packaged into virus capsids. This study measured the percentages of the degraded HBV genomes due to APOBEC3 proteins inhibition of HBV replication, using Quantitative Polymerase Chain Reaction (qPCR) and measured the extent and frequency of APOBEC3 mutations in the surviving Hepatitis B viruses by using 3D-PCR, modified 3D-PCR and primer ID NGS. The critical question was

determining the effect of APOBEC3 family members on degraded, mutated, and packaged HBV genomes.

Previous studies examined the effect of APOBEC3 proteins on HBV replication using transient co-transfection of APOBEC3 and HBV expression constructs in hepatoma Huh7 cells. However, this study used an HBV-producing cell line (HepG2.2.15) as a tool to study the effect of APOBEC3 family members on HBV replication because this study focused on studying the effect on the reverse transcription step of the HBV replication cycle in the cytoplasm and on evaluating the effect in this tool in comparison to the results of previous studies that used APOBEC3 and HBV transient co-transfection in hepatoma Huh7 cells to examine the effect of APOBEC3 protein on HBV replication.

HepG2.2.15 cells have integrated HBV genomes that produce an intermediate pregenomic RNA (pgRNA), which is then transported to the cytoplasm and packaged into the capsid. Subsequently, the pgRNA is reverse-transcribed to negative-strand HBV DNA, and then a positive-strand is synthesized. After that, HBV particles are assembled and secreted from the cells (Sells, Chen *et al.* 1987, Sells, Zelent *et al.* 1988).

There are two ways to examine the effect of these mutators, either by upregulating them with IFN or overexpressing them using APOBEC3-expressing plasmids. The *in vitro* overexpression of APOBEC3 proteins in the hepatoma cell line (HepG2.2.15) method was selected because it singles out the effect of APOBEC3 proteins on HBV replication from the other effects using IFN to upregulate APOBEC3 protein expression. IFN can inhibit HBV replication without the effect of APOBEC3 proteins (Jost, Turelli *et al.* 2007, Proto, Taylor *et al.* 2008, Turelli, Liagre-Quazzola *et al.* 2008). One potential drawback of overexpressing

APOBEC3 proteins using expression construct is that their expression levels may exceed those of APOBEC3 proteins expressed when IFN upregulates them, which may show a more prominent effect of these proteins on HBV replication that does not mimic the actual effect *in vivo*.

This study uses GFP-tagged APOBEC3 expressing plasmids. One advantage of using GFP-tagged APOBEC3 expressing plasmids is that they can be used for the localisation and co-localisation of APOBEC3 proteins. The localisation of APOBEC3 proteins provides an insight into which part of the HBV replication cycle will be targeted by different APOBEC3 proteins. This study showed that APOBEC3DE, APOBEC3F and APOBEC3G proteins expression was restricted to the cytoplasm where HBV reverse transcription occurs. In contrast, APOBEC3A and APOBEC3B were observed to have mainly nuclear localisation. These results indicate that APOBEC3DE, APOBEC3F and APOBEC3G may have their effect on cytoplasmic parts of the HBV replication cycle, where reverse transcription occurs. While APOBEC3A and APOBEC3B may have their effect mainly on nuclear parts of HBV replication cycle. However, these two APOBEC3 proteins were synthesised in the cytoplasm. Due to this, they may also exert an effect on HBV replication in the cytoplasm before their nuclear localisation. The number of APOBEC3 proteins that would enter in a single HBV capsid is not well understood.

Using immunofluorescence to examine the possible co-localisation of HBV polymerase with APOBEC3G, this study showed that APOBEC3G and HBV polymerase proteins colocalise in some areas of the cytoplasm. This may suggest an interaction between APOBEC3G and HBV polymerase proteins. DDX3 was used as positive control for experiment examining the possible co-localisation of HBV polymerase with APOBEC3G. DDX3 is an RNA helicase which unwinds RNA in an ATPase-dependent manner. Using immunofluorescence to examine

the possible co-localisation of HBV polymerase with DDX3, this study showed that DDX3 and HBV polymerase proteins colocalise in some areas of the cytoplasm. Surprisingly, DDX3 was less prominent in these areas where it co-localises with HBV polymerase compared to the area without co-localisation. An explanation for this observation is that DDX3 may be incorporated into the HBV capsid through HBV polymerase interaction, as suggested by Wang *et al.* (Wang, Kim *et al.* 2009) and therefore not detectable by antibody staining. One of the interesting future experiments would be to use immunofluorescence to examine the possible co-localisation of HBV polymerase with APOBEC3A and APOBEC3B (nuclear APOBEC3 proteins).

4.1 APOBEC3 proteins inhibit HBV replication

This study used an experimental system that overexpressed GFP-tagged APOBEC3 plasmids in the HBV-producing HepG2.2.15 cell line to study the effect of APOBEC3 proteins on HBV replication. Another benefit of using fluorescence microscopy to visualise real-time GFP-tagged APOBEC3 proteins expression is that it allows the estimation of real-time transfection efficiency of GFP-tagged APOBEC3 plasmids in HBV-producing HepG2.2.15 cell line. Therefore, experiments with a transfection efficiency of 30% or more of GFP-tagged APOBEC3 plasmids were subsequently used to evaluate the effect of APOBEC3 proteins on HBV replication using methods such as Quantitative Polymerase Chain Reaction (qPCR).

In addition to the visual estimation of APOBEC3 expression using GFP, this study used quantitative Real-Time Polymerase Chain Reaction (RT-qPCR) to quantify the optimal time point of APOBEC3 gene expression in HepG2 cells and found that the highest gene expression of APOBEC3G in HepG2 cells was 72-hour post-transfection, suggesting that this would be the optimal time point to examine the APOBEC3 protein inhibition of HBV replication. This experiment's results face two challenges that should acknowledged, the first one that one reference gene (GAPDH) was used. GAPDH sometimes shows variations in its levels from

cell to cell, the solution to this would be using two reference genes to minimise the possibility of this cell-to-cell variations. The second challenge is the result of qPCR showing the optimal time point of APOBEC3 gene expression in HepG2 cells is 72-hour post-transfection only shows the levels of APOBEC3G gene expression but not the APOBEC3G proteins levels, therefore, showing the protein levels using western blotting would validate the result of qPCR showing this the optimal time point of APOBEC3 gene expression in HepG2 cells is 72-hour post-transfection.

HBV producing HepG2.2.15 cells were transfected with GFP-tagged APOBEC3 plasmids (APOBEC3A, APOBEC3B, APOBEC3DE, APOBEC3F or APOBEC3G) for 72-hours. Then, the effect of APOBEC3 proteins on the cytoplasmic HBV DNA and the extracellular virion-associated HBV DNA levels were evaluated using Quantitative Real-Time Polymerase Chain Reaction (RT-qPCR).

Previously, Turelli *et al.* (Turelli, Mangeat *et al.* 2004) showed a 50 times reduction of the intracellular HBV levels in Huh7 that were transfected with HBV plasmids and co-transfected with APOBEC3G plasmids. Similarly, this study showed a 65% inhibition of the cytoplasmic HBV DNA levels in HepG2.2.15 cells transfected with a GFP-tagged APOBEC3G plasmid. Interestingly, the levels of cytoplasmic HBV DNA in the HepG2.2.15 cell line transfected with GFP-tagged APOBEC3 plasmids were reduced by 57% by APOBEC3B and 50% by APOBEC3F, this result indicated that nuclear localised APOBEC3B exerted a similar inhibition effect on HBV levels to that of the APOBEC3F localised in the cytoplasm. APOBEC3A also showed inhibition of the HBV levels in the cytoplasm but much lower than that of APOBEC3B. One explanation is that these nuclear APOBEC3 proteins may be able to exert their cytoplasmic inhibitory effects on HBV levels after ribosomes synthesize them in the

cytoplasm and before they are translocated to the cell nucleus. Chen *et al.* (Chen, Hu *et al.* 2018) reported similar results to this study finding that APOBEC3B exerts an inhibitory effect on HBV levels during reverse transcription in the cytoplasm—an interesting study by Lucifora *et al.* (Lucifora, Xia *et al.* 2014) used lymphotoxin-β-receptor agonists to upregulate APOBEC3B and IFN-α to upregulate APOBEC3A, leading to deamination and degradation of HBV cccDNA in the nucleus. This indicated that APOBEC3A and APOBEC3B might act as antiviral factors in the HBV replication cycle's nuclear and cytoplasmic parts. Overall, APOBEC3 proteins showed lower inhibition percentages of the extracellular virion-associated HBV DNA levels compared to the inhibition of the cytoplasmic HBV DNA in HepG2.2.15 cells transfected with GFP-tagged APOBEC3 plasmids. However, APOBEC3F showed similar inhibition percentages of the extracellular virion-associated HBV DNA levels by 55% and the cytoplasmic HBV DNA by 50% in HepG2.2.15 cells line transfected with GFP-tagged APOBEC3F plasmid.

Bouzidi *et al.* showed that APOBEC3DE could antagonise APOBEC3F and APOBEC3G restriction of HBV; therefore, APOBEC3DE could enhance extracellular HBV DNA levels. Bouzidi *et al.* argued that this was due to the catalytically inactive state of APOBEC3DE due to a Tyr320Cys substitution in the C-terminal domain (Bouzidi, Caval *et al.* 2016). However, Bouzidi *et al.* did not show the effect of APOBEC3DE on cytoplasmic HBV DNA. In comparison to Bouzidi *et al.* This study has shown that APOBEC3DE could inhibit the extracellular virion-associated HBV DNA levels by 16% and the cytoplasmic HBV DNA by 39% in HepG2.2.15 cell line transfected with GFP-tagged APOBEC3DE plasmid. This raised the question of whether APOBEC3DE exerts its inhibitory effect through editing. Another finding of the current study is that APOBEC3 inhibitory effects may be enhanced when more than one APOBEC3 is used. This study showed that APOBEC3A, APOBEC3F, and

APOBEC3G proteins synergistically inhibited the cytoplasmic HBV DNA by 67.20 %. This suggests that APOBEC3 proteins may act synergistically to inhibit HBV replication.

4.2 APOBEC3 proteins edit the HBV genome

APOBEC3 proteins are cytidine deaminases, which are hydrolases that use water in their reaction to remove the amino group from cytidines and convert them to uridines. This leads to the formation of uracil-containing DNA. This leads to G to A mutations in the plus strand and C to T mutations in the negative strand. The overall effect leads to the formation of AT-rich regions in the GC-rich islands of the HBV DNA.

APOBEC3 proteins can exert their inhibitory effect on HBV replication through non-deaminating and deaminating mechanisms. The deaminating mechanism inhibits HBV by inducing G to A mutations in the plus strand and C to T mutations in the negative strand during HBV reverse transcription of the pgRNA. Estimating these HBV mutations due to APOBEC3 proteins is crucial in determining the actual effect of these cytidine deaminases on HBV replication. To explore this, several researchers tried to amplify HBV sequences after APOBEC3 protein expression to determine the levels of mutation. This study builds on their efforts and proposes alternative approaches to authentic determination of the frequency and extent of HBV DNA mutations due to APOBEC3 proteins.

Turelli *et al.* could not detect G to A mutations using sequences from conventional polymerase chain reaction amplification of core-associated HBV DNA extracted from the Huh7 hepatoma cell line co-transfected with APOBEC3G and HBV expressing plasmids (Turelli, Mangeat *et al.* 2004).

Suspene *et al.* detected G to A mutations using a two-round nested PCR (3D-PCR) that could selectively amplify G to A mutations by lowering the DNA denaturation temperature in the second round. This two-round PCR (3D-PCR) used non-high-fidelity Taq polymerase to amplify the sequences of core-associated HBV DNA extracted from Huh7 hepatoma cells co-transfected with APOBEC3 family members and HBV-expressing plasmids Field (102). This method uses external primers for the first PCR round and internal primers for the second round. Suspene used this method to selectively amplify the HBV variants that had AT-rich regions using 3D-PCR.

However, Suspene *et al.* succeeded in detecting G to A mutations using 3D-PCR; they faced challenges such as: 1) 3D-PCR uses Taq polymerase (non-high-fidelity) to amplify the sequences of core-associated HBV DNA. This Taq polymerase is prone to errors and makes the estimation of G to A mutations questionable. 2) 3D-PCR selectively amplifies HBV molecules that have G to A mutation by lowering the DNA denaturation temperature in the second round. This excludes the other part of HBV variants and magnifies the number of G to A mutated HBV molecules compared to the whole HBV population.

The rationale for using Taq polymerase (non-high-fidelity) in the 3D-PCR method, by Suspene *et al.* , comes from the issues arising when using the high-fidelity polymerase to amplify uracil containing DNA that resulted from the APOBEC3 proteins editing of HBV genome DNA (Suspene, Guetard *et al.* 2005) (Vartanian, Henry *et al.* 2010). These issues are 1) In the case when one strand of HBV DNA is edited to contain uracil, high-fidelity polymerase preferentially amplifies the other DNA strand that does not contain uracil because the high-fidelity polymerase halts at an uracil nucleotide and cannot complete the amplification of uracil containing DNA strands. 2) In the other case when the two strands of HBV DNA are edited to

contain uracil, high-fidelity polymerase halts at a uracil nucleotide and cannot complete the amplification of uracil-containing DNA strands (Arbeithuber, Makova et al. 2016). Another way to validate the 3D-PCR technique would be to use the 3D-PCR technique to amplify DNA of known AGCT composition and to amplify a cloned HBV X region and examine if 3D-PCR did not introduce mutations.

These two scenarios led Suspene *et al.* and others to refrain from using high-fidelity polymerase in the 3D-PCR method to detect and quantify the G to A mutations in HBV genome DNA caused by APOBEC3 proteins.

This study addressed the inability to use high-fidelity polymerase by modifying the 3D-PCR method by using the Phusion U high-fidelity polymerase. This Phusion U high-fidelity polymerase-based 3D-PCR uses Phusion U high-fidelity polymerase that 1) can amplify uracil-containing DNA and 2) has high-fidelity that eliminates the DNA amplification errors that may occur during using Taq polymerase (non-high-fidelity). Phusion U high-fidelity polymerase-based 3D-PCR results were compared to Taq polymerase-based 3D-PCR.

This study uses the same primers that were used by Suspene *et al.* and his colleagues to sequence the HBV (ayw) X region (with a length of 167 nucleotides) in the 3D-PCR estimation of G to A mutations in the edited HBV sequence by APOBEC3 proteins (Suspene, Guetard *et al.* 2005). In the 167-nucleotide long HBV (ayw) X region sequenced, 44 guanines are potential targets for editing by APOBEC3 proteins.

In this study, the Taq polymerase-based 3D-PCR (original) and the Phusion U high-fidelity polymerase-based 3D-PCR (Modified) techniques show that the result of the number of G to

A mutations varied from one clone to another in the 10 clones sequenced suggesting the likelihood of that APOBEC3 proteins may randomly targeting cytosines in this HBV X region. However, when looking at these results, overall, it is apparent that the numbers and percentages of G to A mutations in the HBV X region detected by Phusion U polymerase-based 3D-PCR was lower in empty vector and APOBEC3G transfected HepG2.2.15 cells in comparison to the numbers and percentages of G to A mutations in the HBV X region detected by Taq polymerase-based 3D-PCR in empty vector and APOBEC3G transfected HepG2.2.15 cells. However, in the case of APOBEC3DE transfected HepG2.2.15 cells, the numbers and percentages of G to A mutations in the HBV X region by Phusion U polymerase-based 3D-PCR were higher than those detected by Taq polymerase-based 3D-PCR. These findings suggest that using Phusion U high-fidelity polymerase that can amplify uracil-containing DNA shows a different picture of the accurate estimation of the numbers and percentages of G to A mutations in the HBV X region. Surprisingly, both Taq polymerase-based 3D-PCR and the Phusion U high-fidelity polymerase-based 3D-PCR show G to A mutations in the HBV X region in the empty vector-transfected HepG2.2.15 cells. This raises the question of whether these G to A mutations already exist in the HBV X region of HBV-producing HepG2.2.15. To answer this, 3D-PCR results of the numbers and percentages of G to A mutations in the HBV X region in the HBV that was extracted from HepG2.2.15 cells show that these mutations do already exist in the HBV genome DNA extracted from HepG2.2.15 cells. One interpretation of this finding is that endogenous APOBEC3 proteins may exert their effect on the HBV genome DNA and as 3D-PCR selectively amplify HBV molecules that have G to A mutation by lowering the DNA denature temperature. This excludes the other part of HBV variants and magnifies the number of G to A mutations in HBV molecules compared to the whole HBV population. Seager *et al.* claimed that the APOBEC3-mediated HBV G to A mutations already exist in the HBV input for 3D-PCR, and the problem lies in the 3D-PCR technique. Seager *et*

al. raised another critical issue: a resulting artifact may be due to re-annealing and extension of incomplete PCR products during the denaturation step of the 3D-PCR (Seeger and Sohn 2016). Another Study by Nair *et al.* suggested that endogenous APOBEC3 proteins may edit between 10 to 25% of the HBV genomic DNA in HepAD38 and HepG2.2.15 cell lines (Nair and Zlotnick 2018). The results of the current study validate Nair's results that G to A mutations may already exist but refute the suggestion that G to A mutations is an artifact produced by re-annealing and extension of incomplete PCR products during the denaturation step of 3D-PCR as suggested by Seager *et al.* because this study used the modified 3D-PCR which uses Phusion U high-fidelity polymerase that can amplify uracil containing DNA and still shows that there are G to A mutations in the HBV X region in the HBV DNA that was extracted from HepG2.2.15 cells.

Next-Generation sequencing (NGS) or second-generation sequencing is a technique that uses a technology that can perform the sequencing of millions of DNA fragments in parallel (Behjati and Tarpey 2013). This technique is superior to First-Generation Sequencing (Sanger Sequencing) by providing greater depth of sequence information over the traditional Sanger sequencing method. It allows greater precision in determining mutations' true depth and actual bias in the sequenced samples. PCR amplification of the viral segment is needed before proceeding to Next-Generation sequencing (NGS). This PCR step may lead to errors such as PCR-mediated recombination, nucleotide misincorporation and PCR-mediated recombination. These PCR errors may affect the result of mutations detected Next-Generation sequencing (Meyerhans, Vartanian *et al.* 1990, Wyles 2019). To solve this issue Next-Generation sequencing with barcoding primers (Primer ID next) was developed (Jabara, Jones *et al.* 2011, Zhou, Jones *et al.* 2015). This Primer ID Next-Generation sequencing uses an identifier sequence (primer ID barcode) added to the primers used in the PCR step preceding Next-

Generation sequencing to identify each individual viral molecule. This allows the ability to know when the PCR resamples the original template sequenced. These PCR resampled sequences can be pooled together to form a consensus sequence that can then be aligned to its original template sequence. Primer ID Next-Generation sequencing was used to study the antiviral resistance mutations in several viruses, including HIV, Influenza and Coronaviruses (Zhou, Jones *et al.* 2015, Goldhill, Langat *et al.* 2019, Zhou, Hill *et al.* 2021). For example, the Goldhill *et al.* study used Primer ID Next-Generation sequencing by adding a barcoding sequence to each cDNA molecule amplified during reverse transcription. This allowed them to detect the accurate favipiravir-induced mutations in the Influenza virus and quantify and determine more mutations than that could be determined by conventional Sanger sequencing (Goldhill, Langat *et al.* 2019). The current study adopted their protocol by tagging each individual HBV DNA molecule of HBV pre-core, X, S regions of the extracted intracellular HBV DNA genome from HepG2.2.15 cells with barcoding sequences during the first round of nested PCR using Phusion U high-fidelity polymerase. This was to quantify the frequency and extent of APOBEC3G mutations in HBV pre-core, X, and S regions in the extracted HBV DNA to elucidate the effect of APOBEC3 proteins on HBV replication. The use of primer ID Next Generation Sequencing also addressed the issues of using the 3D-PCR (which uses Sanger sequencing) that may magnify the number of G to A mutated HBV molecules caused by APOBEC3 proteins. This magnification of the number of G to A mutated HBV genomes may be because 3D-PCR selectively amplifies HBV molecules that have G to A mutation by lowering the DNA denaturation temperature in the second round and excluding the other part of HBV variants. Using primer ID Next Generation Sequencing (NGS) can give a more accurate measure of the frequency of mutations in HBV DNA caused by APOBEC3 proteins. The greater depth of NGS sequence information over traditional 3D-PCR sequencing methods allowed greater precision in determining the bias of 3D-PCR in detecting G to A and C to U

mutations caused by APOBEC3 proteins the extracted intracellular HBV DNA genome from HBV-producing HepG2.2.15 cell line.

The first observation of the results of the current study using primer ID Next Generation Sequencing showed that there are existing G to A and C to U mutations in HBV pre-core, X, and S regions of the extracted intracellular HBV DNA genome from HepG2.2.15 cells that have not been transfected with APOBEC3 plasmids. This observation further confirms the results of the 3D-PCR carried out in the current study and validates the claims by Seager *et al.* and Nair *et al.* that the APOBEC3 mediated HBV G to A mutations already exist in the HBV DNA input for 3D-PCR and that 10 to 25% of the HBV genome DNA population in HepAD38 and HepG2.2.15 cell lines have G to A mutations. One suggested explanation is that these mutations in the HBV genome may be due to the deamination effect of endogenous APOBEC3 proteins. Another surprising finding of the current study using primer ID Next Generation Sequencing was that the S region of the HBV genome showed a higher number of G to A and C to U mutations than that of the HBV X region. The HBV pre-core region had the lowest number of G to A and C to U mutations. This may suggest that the region of the HBV genome where the surface protein and polymerase coding regions overlap was more targeted by APOBEC3 than other parts of the HBV genome. The current study results using primer ID Next Generation Sequencing also showed other mutations that 3D-PCR could not detect. These findings may indicate that using primer ID Next Generation Sequencing may help to demonstrate more accurately the effects of APOBEC3 proteins on HBV *in vitro* and in vivo. This would be important if the effects of APOBEC3 proteins on HBV DNA are tested in a future clinical setting to evaluate the possible role of APOBEC3 proteins in treatments for HBV infection.

Overall, there are two parts to the story of the APOBEC3 proteins on HBV replication. One part is where virus genomes survive the mutations caused by APOBEC3 proteins, and the other part is where the genomes do not survive the mutations and are subsequently degraded. Studying the mutations induced by APOBEC3 proteins in the surviving genomes is important to determine the deamination effect of APOBEC3 protein on HBV genomic DNA. However, this does not show the mutations induced by APOBEC3 proteins in the degraded genomes. This is called survival bias which was first proposed by Wald *et al. ,* an American mathematician who was studying the damaged areas in bomber warplanes that survived their missions and returned (Wald 1980). That study proposed that the areas that were intact in the warplanes that were damaged and returned safely (survived) are those that were damaged in the warplanes that did not return (that did not survive). Therefore, the study suggested that the areas which were intact areas in the warplanes that were damaged and returned safely (survived) are the areas of the planes' bodies that should be strengthened to avoid losing them. This approach increased the percentage of surviving (returned safely) warplanes. The same analogy could be used regarding APOBEC3, where the more damaging mutations are caused by APOBEC3 proteins that lead to the degradation of HBV genomic DNA. These mutations could be studied using cell-free APOBEC3 deamination of the intact but important regions of the survived HBV where the possible lethal APOBEC3 mutations lie. The other important possibility is that the surviving genomes may undergo DNA repair by cellular DNA repair mechanisms, which may lead to viral diversity and may subsequently increase viral replication and infectivity by counteracting the effects of APOBEC3 proteins on the HBV genomic DNA (Kitamura, Wang *et al.* 2013).

4.3 Human Upstream stimulatory factor 1 (USF1) upregulate APOBEC3 family members and enhances HBV replication

APOBEC3 gene expression can be upregulated by type I (IFN-α) (Bonvin, Achermann *et al.* 2006) and type II interferons (IFN-γ) (Proto, Taylor *et al.* 2008), and lymphotoxin-β-receptor agonist (Lucifora, Xia *et al.* 2014). The upregulation of APOBEC3 proteins was examined by Zeng *et al.* , who showed that USF1 binds to the E-box in the promotor of the APOBEC3G gene, which subsequently leads to the APOBEC3G upregulation (Zeng, Li *et al.* 2016). This study further examined the effect of USF1 on the gene expression of APOBEC3 family members. This was carried out by overexpressing USF1 in HepG2.2.15 cells and evaluating the level of gene expression of APOBEC3 family members. The current study's findings showed similar results of APOBEC3G upregulation to those of Zeng *et al.* In addition, it showed that the gene expression of APOBEC3DE, APOBEC3B and APOBEC3F were upregulated upon UF1 overexpression in HepG2.2.15 cells. This suggests that USF1 can upregulate APOBEC3 family members and may subsequently have an inhibitory effect on HBV replication. However, USF1 also enhanced HBV replication despite its effect on the gene expression of APOBEC3 family members. Several studies showed that USF1 increased the genome transcription of a number of viruses such; as HIV (Maekawa, Sudo *et al.* 1991, Sirito, Walker *et al.* 1992), varicella-zoster virus (Meier, Luo *et al.* 1994), EB virus and adenovirus (Sirito, Walker *et al.* 1992, Corre and Galibert 2005). This may explain why USF1 enhanced HBV replication, as it may exert this effect through activating HBV transcription. The results of this study showed that transcription factor (USF1) plays a role in the APOBEC3, where it upregulates APOBEC3 family members and enhances HBV replication simultaneously.

4.4 Conclusions

This study has investigated the effects of APOBEC3 proteins in Hepatitis B virus replication and provided novel approaches that may help to measure more accurately the effects of APOBEC3 proteins on the HBV virus. This study concluded that:

APOBEC3DE and APOBEC3G localise to the cytoplasm, suggesting that these proteins exert their effect on HBV replication in the cytoplasm, it was also observed that APOBEC3G proteins colocalise with HBV polymerase proteins in areas of the cytoplasm.

The transcription factor (USF1) could upregulate APOBEC3 gene expression and USF1 could enhance HBV replication.

The levels of extracellular virion-associated HBV DNA and cytoplasmic HBV DNA could be inhibited by APOBEC3DE and APOBEC3G.

The level of G to A mutations in the HBV X region detected by Phusion U polymerase-based 3D-PCR were lower than those detected by Taq polymerase-based 3D-PCR in the empty vector and APOBEC3G vector transfected HepG2.2.15 cells. However, the level of G to A mutations in the HBV X region detected by Phusion U polymerase-based 3D-PCR were higher than those detected by Taq polymerase-based 3D-PCR in APOBEC3DE transfected HepG2.2.15 cells. The primer ID Next Generation Sequencing (NGS) demonstrated a higher number of G to A and C to U mutations in the HBV S region compared to the HBV X region. The primer ID NGS also showed that G to A and C to U mutations already existed in HBV pre-core, X, and S regions of the extracted intracellular HBV genomic DNA from HepG2.2.15 cells that were not transfected with APOBEC3 plasmids.

4.5 Possible Future Work

Following the results of this study, further investigations could expand and extend the approach that was used. These might include:

1) Using immunofluorescence to examine the possible co-localisation of HBV polymerase with APOBEC3A and APOBEC3B (nuclear APOBEC3 proteins).
2) Inactivate APOBEC3 genes so no protein is made, and then analyse to see if any G to A mutations are present in the HBV genome.
3) Exploring the *In vitro* APOBEC3 deamination of HBV genome DNA to detect possible lethal mutations in the intact areas of the surviving viruses.
4) Transfect HepG2.2.15 cells with APOBEC3 plasmids, and then isolate HBV particles and assess live virus by infecting cells and compare to virus isolated from HepG2.2.15 cell line. To examine if virus particles less infectious.
5) Using Modified 3D-PCR (Phusion U polymerase-based 3D-PCR) to examine the *In vivo* APOBEC3 mutations in HBV genome DNA and compare it to conventional 3D-PCR (Taq polymerase-based 3D-PCR).
6) Using primer ID Next Generation Sequencing to examine the *In vivo* APOBEC3 mutations in HBV genome DNA.
7) As USF1 can upregulate the basal endogenous APOBEC3 gene expression, it would be interesting to know if these upregulated APOBEC3 members cause G to A mutation in HBV genome DNA even though USF1 enhance HBV replication.

5 References:

Albin, C. and W. S. Robinson (1980). "Protein kinase activity in hepatitis B virus." J Virol **34**(1): 297-302.

Allweiss, L. and M. Dandri (2017). "The Role of cccDNA in HBV Maintenance." Viruses **9**(6).

Anant, S. and N. O. Davidson (2001). "Molecular mechanisms of apolipoprotein B mRNA editing." Curr Opin Lipidol **12**(2): 159-165.

Arbeithuber, B., K. D. Makova and I. Tiemann-Boege (2016). "Artifactual mutations resulting from DNA lesions limit detection levels in ultrasensitive sequencing applications." DNA Res **23**(6): 547-559.

Arias, J. F., T. Koyama, M. Kinomoto and K. Tokunaga (2012). "Retroelements versus APOBEC3 family members: No great escape from the magnificent seven." Frontiers in microbiology **3**: 275.

Armitage, A. E., K. Deforche, J. J. Welch, K. Van Laethem, R. Camacho, A. Rambaut and A. K. Iversen (2014). "Possible footprints of APOBEC3F and/or other APOBEC3 deaminases, but not APOBEC3G, on HIV-1 from patients with acute/early and chronic infections." J Virol **88**(21): 12882-12894.

Beck, J. and M. Nassal (2007). "Hepatitis B virus replication." World J Gastroenterol **13**(1): 48-64.

Behjati, S. and P. S. Tarpey (2013). "What is next generation sequencing?" Arch Dis Child Educ Pract Ed **98**(6): 236-238.

Belloni, L., L. Allweiss, F. Guerrieri, N. Pediconi, T. Volz, T. Pollicino, J. Petersen, G. Raimondo, M. Dandri and M. Levrero (2012). "IFN-alpha inhibits HBV transcription and replication in cell culture and in humanized mice by targeting the epigenetic regulation of the nuclear cccDNA minichromosome." J Clin Invest **122**(2): 529-537.

Blumberg, B. S. (1964). " Polymorphisms of the serum proteins and the development of iso-preciptins in transfused patients " Bull N Y Acad Med **40**: 377-386.

Blumberg, B. S., A. I. Sutnick and W. T. London (1968). "Hepatitis and leukemia: their relation to Australia antigen." Bull N Y Acad Med **44**(12): 1566-1586.

Bock, C. T., S. Schwinn, S. Locarnini, J. Fyfe, M. P. Manns, C. Trautwein and H. Zentgraf (2001). Structural organization of the hepatitis B virus minichromosome. J Mol Biol. England, 2001 Academic Press. **307:** 183-196.

Bonvin, M., F. Achermann, I. Greeve, D. Stroka, A. Keogh, D. Inderbitzin, D. Candinas, P. Sommer, S. Wain-Hobson, J. P. Vartanian and J. Greeve (2006). "Interferon-inducible expression of APOBEC3 editing enzymes in human hepatocytes and inhibition of hepatitis B virus replication." Hepatology **43**(6): 1364-1374.

Bonvin, M. and J. Greeve (2007). Effects of point mutations in the cytidine deaminase domains of APOBEC3B on replication and hypermutation of hepatitis B virus in vitro. J Gen Virol. England. **88:** 3270-3274.

Bottcher, B., S. A. Wynne and R. A. Crowther (1997). "Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy." Nature **386**(6620): 88-91.

Bouzidi, M. S., V. Caval, R. Suspène, C. Hallez, P. Pineau, S. Wain-Hobson and J. P. Vartanian (2016). "APOBEC3DE Antagonizes Hepatitis B Virus Restriction Factors APOBEC3F and APOBEC3G." J Mol Biol **428**(17): 3514-3528.

Bransteitter, R., P. Pham, M. D. Scharff and M. F. Goodman (2003). Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. Proc Natl Acad Sci U S A. United States. **100:** 4102-4107.

Brar, S. S., M. Watson and M. Diaz (2004). Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. J Biol Chem. United States. **279:** 26395-26401.

Bruss, V. (2007). "Hepatitis B virus morphogenesis." World J Gastroenterol **13**(1): 65-73.

Burns, M. B., L. Lackey, M. A. Carpenter, A. Rathore, A. M. Land, B. Leonard, E. W. Refsland, D. Kotandeniya, N. Tretyakova and J. B. Nikas (2013). "APOBEC3B is an enzymatic source of mutation in breast cancer." Nature **494**(7437): 366-370.

Burns, M. B., N. A. Temiz and R. S. Harris (2013). "Evidence for APOBEC3B mutagenesis in multiple human cancers." Nature genetics **45**(9): 977-983.

Centers for Disease Control and Prevention. (2019). "Hepatitis B." Yellow Book Retrieved July, 2019, from https://wwwnc.cdc.gov/travel/yellowbook/2020/travel-related-infectious-diseases/hepatitis-b#map404.

Chen, Y., J. Hu, X. Cai, Y. Huang, X. Zhou, Z. Tu, J. Hu, J. E. Tavis, N. Tang, A. Huang and Y. Hu (2018). "APOBEC3B edits HBV DNA and inhibits HBV replication during reverse transcription." Antiviral Res **149**: 16-25.

Chisari, F. V., W. S. Mason and C. Seeger (2014). "Comment on "Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA"."

Conticello, S. G., R. S. Harris and M. S. Neuberger (2003). "The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G." Curr Biol **13**(22): 2009-2013.

Corre, S. and M. D. Galibert (2005). "Upstream stimulating factors: highly versatile stress-responsive transcription factors." Pigment cell research **18**(5): 337-348.

Crowther, R. A., N. A. Kiselev, B. Bottcher, J. A. Berriman, G. P. Borisova, V. Ose and P. Pumpens (1994). "Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy." Cell **77**(6): 943-950.

Dane, D. S., C. H. Cameron and M. Briggs (1970). "Virus-like particles in serum of patients with Australia-antigen-associated hepatitis." Lancet **1**(7649): 695-698.

Daub, H., S. Blencke, P. Habenberger, A. Kurtenbach, J. Dennenmoser, J. Wissing, A. Ullrich and M. Cotten (2002). "Identification of SRPK1 and SRPK2 as the major cellular protein kinases phosphorylating hepatitis B virus core protein." J Virol **76**(16): 8124-8137.

Desimmie, B. A., K. A. Delviks-Frankenberrry, R. C. Burdick, D. Qi, T. Izumi and V. K. Pathak (2014). "Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all." J Mol Biol **426**(6): 1220-1245.

Dickerson, S. K., E. Market, E. Besmer and F. N. Papavasiliou (2003). AID Mediates Hypermutation by Deaminating Single Stranded DNA. J Exp Med. **197:** 1291-1296.

Dryden, K. A., S. F. Wieland, C. Whitten-Bauer, J. L. Gerin, F. V. Chisari and M. Yeager (2006). "Native hepatitis B virions and capsids visualized by electron cryomicroscopy." Mol Cell **22**(6): 843-850.

European Association For The Study Of The Liver (2017). "EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection." Journal of hepatology **67**(2): 370-398.

Fields, B. N., D. M. Knipe and P. M. Howley (2013). Fields virology, Volume 2 Volume 2. Philadelphia, Wolters Kluwer Health/Lippincott Williams & Wilkins.

Galibert, F., E. Mandart, F. Fitoussi, P. Tiollais and P. Charnay (1979). "Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli." Nature **281**(5733): 646-650.

Galvis, A. E., H. E. Fisher and D. Camerini (2017). "NP-40 Fractionation and Nucleic Acid Extraction in Mammalian Cells." Bio Protoc **7**(20): e2584.

Galvis, A. E., H. E. Fisher, T. Nitta, H. Fan and D. Camerini (2014). "Impairment of HIV-1 cDNA synthesis by DBR1 knockdown." J Virol **88**(12): 7054-7069.

Gao, B., Z. Duan, W. Xu and S. Xiong (2009). "Tripartite motif-containing 22 inhibits the activity of hepatitis B virus core promoter, which is dependent on nuclear-located RING domain." Hepatology **50**(2): 424-433.

Glebe, D. and S. Urban (2007). "Viral and cellular determinants involved in hepadnaviral entry." World J Gastroenterol **13**(1): 22-38.

Goldhill, D. H., P. Langat, H. Xie, M. Galiano, S. Miah, P. Kellam, M. Zambon, A. Lackenby and W. S. Barclay (2019). "Determining the Mutation Bias of Favipiravir in Influenza Virus Using Next-Generation Sequencing." J Virol **93**(2).

Graham, F. L., J. Smiley, W. C. Russell and R. Nairn (1977). "Characteristics of a human cell line transformed by DNA from human adenovirus type 5." J Gen Virol **36**(1): 59-74.

Gripon, P., S. Rumin, S. Urban, J. Le Seyec, D. Glaise, I. Cannie, C. Guyomard, J. Lucas, C. Trepo and C. Guguen-Guillouzo (2002). Infection of a human hepatoma cell line by hepatitis B virus. Proc Natl Acad Sci U S A. United States. **99:** 15655-15660.

Guo, J. T. and H. Guo (2015). "Metabolism and function of hepatitis B virus cccDNA: Implications for the development of cccDNA-targeting antiviral therapeutics." Antiviral Res **122**: 91-100.

Harris, R. S., S. K. Petersen-Mahrt and M. S. Neuberger (2002). RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. Mol Cell. United States. **10:** 1247-1253.

Hohn, T. (2013). "Plant pararetroviruses: interactions of cauliflower mosaic virus with plants and insects." Curr Opin Virol **3**(6): 629-638.

Hu, J. and C. Seeger (2015). "Hepadnavirus Genome Replication and Persistence." Cold Spring Harb Perspect Med **5**(7): a021386.

Hu, X., M. A. Zimmerman, K. Bardhan, D. Yang, J. L. Waller, G. B. Liles, J. R. Lee, R. Pollock, D. Lev, C. F. Ware, E. Garber, V. Bailly, J. L. Browning and K. Liu (2013). Lymphotoxin beta receptor mediates caspase-dependent tumor cell apoptosis in vitro and tumor suppression in vivo despite induction of NF-kappaB activation. Carcinogenesis. England. **34:** 1105-1114.

Hultquist, J. F., J. A. Lengyel, E. W. Refsland, R. S. LaRue, L. Lackey, W. L. Brown and R. S. Harris (2011). "Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1." J Virol **85**(21): 11220-11234.

Isaacs, A. and J. Lindenmann (1957). "Virus interference. I. The interferon." Proc R Soc Lond B Biol Sci **147**(927): 258-267.

Ito, S., H. Nagaoka, R. Shinkura, N. Begum, M. Muramatsu, M. Nakata and T. Honjo (2004). Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. Proc Natl Acad Sci U S A. United States. **101:** 1975-1980.

Jabara, C. B., C. D. Jones, J. Roach, J. A. Anderson and R. Swanstrom (2011). "Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID." Proc Natl Acad Sci U S A **108**(50): 20166-20171.

Janahi, E. M. and M. J. McGarvey (2013). "The inhibition of hepatitis B virus by APOBEC cytidine deaminases." J Viral Hepat **20**(12): 821-828.

Jarmuz, A., A. Chester, J. Bayliss, J. Gisbourne, I. Dunham, J. Scott and N. Navaratnam (2002). An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. Genomics. United States. **79:** 285-296.

Jost, S., P. Turelli, B. Mangeat, U. Protzer and D. Trono (2007). Induction of antiviral cytidine deaminases does not explain the inhibition of hepatitis B virus replication by interferons. J Virol. United States. **81:** 10588-10596.

Kann, M., B. Sodeik, A. Vlachou, W. H. Gerlich and A. Helenius (1999). "Phosphorylation-dependent Binding of Hepatitis B Virus Core Particles to the Nuclear Pore Complex." J Cell Biol **145**(1): 45-55.

Kao, S., M. A. Khan, E. Miyagi, R. Plishka, A. Buckler-White and K. Strebel (2003). "The human immunodeficiency virus type 1 Vif protein reduces intracellular expression and inhibits packaging of APOBEC3G (CEM15), a cellular inhibitor of virus infectivity." J Virol **77**(21): 11398-11407.

Kaplan, P. M., R. L. Greenman, J. L. Gerin, R. H. Purcell and W. S. Robinson (1973). "DNA polymerase associated with human hepatitis B antigen." J Virol **12**(5): 995-1005.

Kitamura, K., Z. Wang, S. Chowdhury, M. Simadu, M. Koura and M. Muramatsu (2013). "Uracil DNA glycosylase counteracts APOBEC3G-induced hypermutation of hepatitis B viral genomes: excision repair of covalently closed circular DNA." PLoS Pathog **9**(5): e1003361.

Knisbacher, B. A., D. Gerber and E. Y. Levanon (2016). "DNA editing by APOBECs: a genomic preserver and transformer." Trends in Genetics **32**(1): 16-28.

Kock, J., M. Nassal, S. MacNelly, T. F. Baumert, H. E. Blum and F. von Weizsacker (2001). "Efficient infection of primary tupaia hepatocytes with purified human and woolly monkey hepatitis B virus." J Virol **75**(11): 5084-5089.

Kramvis, A., M. Kew and G. Francois (2005). "Hepatitis B virus genotypes." Vaccine **23**(19): 2409-2423.

Kuroki, K., F. Eng, T. Ishikawa, C. Turck, F. Harada and D. Ganem (1995). "gp180, a host cell glycoprotein that binds duck hepatitis B virus particles, is encoded by a member of the carboxypeptidase gene family." J Biol Chem **270**(25): 15022-15028.

Lada, A. G., C. F. Krick, S. G. Kozmin, V. I. Mayorov, T. S. Karpova, I. B. Rogozin and Y. I. Pavlov (2011). Mutator effects and mutation signatures of editing deaminases produced in bacteria and yeast. Biochemistry (Mosc). United States. **76:** 131-146.

Lamontagne, R. J., S. Bagga and M. J. Bouchard (2016). "Hepatitis B virus molecular biology and pathogenesis." Hepatoma Res **2**: 163-186.

Lau, P. P., W. J. Xiong, H. J. Zhu, S. H. Chen and L. Chan (1991). "Apolipoprotein B mRNA editing is an intranuclear event that occurs posttranscriptionally coincident with splicing and polyadenylation." J Biol Chem **266**(30): 20550-20554.

Lei, Y. C., Y. J. Tian, H. H. Ding, B. J. Wang, Y. Yang, Y. H. Hao, X. P. Zhao, M. J. Lu, F. L. Gong and D. L. Yang (2006). "N-terminal and C-terminal cytosine deaminase domain of APOBEC3G inhibit hepatitis B virus replication." World J Gastroenterol **12**(46): 7488-7496.

Leistner, C. M., S. Gruen-Bernhard and D. Glebe (2008). Role of glycosaminoglycans for binding and infection of hepatitis B virus. Cell Microbiol. England. **10:** 122-133.

Li, D., J. Liu, F. Kang, W. Guan, X. Gao, Y. Wang and D. Sun (2011). Core-APOBEC3C chimerical protein inhibits hepatitis B virus replication. J Biochem. England. **150:** 371-374.

Li, J., S. Lin, Q. Chen, L. Peng, J. Zhai, Y. Liu and Z. Yuan (2010). Inhibition of hepatitis B virus replication by MyD88 involves accelerated degradation of pregenomic RNA and nuclear retention of pre-S/S RNAs. J Virol. United States. **84:** 6387-6399.

Li, N., L. Zhang, L. Chen, W. Feng, Y. Xu, F. Chen, X. Liu, Z. Chen and W. Liu (2012). "MxA inhibits hepatitis B virus replication by interaction with hepatitis B core antigen." Hepatology **56**(3): 803-811.

Liao, W., S. H. Hong, B. H. Chan, F. B. Rudolph, S. C. Clark and L. Chan (1999). APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. Biochem Biophys Res Commun. United States, 1999 Academic Press. **260:** 398-404.

Liu, F., M. Campagna, Y. Qi, X. Zhao, F. Guo, C. Xu, S. Li, W. Li, T. M. Block, J. Chang and J. T. Guo (2013). "Alpha-interferon suppresses hepadnavirus transcription by altering epigenetic modification of cccDNA minichromosomes." PLoS Pathog **9**(9): e1003613.

Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-408.

Lok, A. S., F. Zoulim, G. Dusheiko and M. G. Ghany (2017). "Hepatitis B cure: From discovery to regulatory approval." J Hepatol **67**(4): 847-861.

Luangsay, S., C. R. C. o. L. C. Lyon University, INSERM, Lyon, France, F. Zoulim and L. Hospices Civils de Lyon, France (2013). "Structure and Molecular Virology." 63-80.

Lucifora, J., Y. Xia, F. Reisinger, K. Zhang, D. Stadler, X. Cheng, M. F. Sprinzl, H. Koppensteiner, Z. Makowska, T. Volz, C. Remouchamps, W. M. Chou, W. E. Thasler, N. Huser, D. Durantel, T. J. Liang, C. Munk, M. H. Heim, J. L. Browning, E. Dejardin, M. Dandri, M. Schindler, M. Heikenwalder and U. Protzer (2014). "Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA." Science **343**(6176): 1221-1228.

MacLachlan, J. H. and B. C. Cowie (2015). "Hepatitis B virus epidemiology." Cold Spring Harb Perspect Med **5**(5): a021410.

Maekawa, T., T. Sudo, M. Kurimoto and S. Ishii (1991). "USF-related transcription factor, HIV-TF1, stimulates transcription of human immunodeficiency virus-1." Nucleic acids research **19**(17): 4689-4694.

Magnius, L. O. and H. Norder (1995). "Subtypes, genotypes and molecular epidemiology of the hepatitis B virus as reflected by sequence variability of the S-gene." Intervirology **38**(1-2): 24-34.

Malim, M. H. and P. D. Bieniasz (2012). "HIV Restriction Factors and Mechanisms of Evasion." Cold Spring Harb Perspect Med **2**(5): a006940.

Mao, R., J. Zhang, D. Jiang, D. Cai, J. M. Levy, A. Cuconati, T. M. Block, J. T. Guo and H. Guo (2011). Indoleamine 2,3-dioxygenase mediates the antiviral effect of gamma interferon against hepatitis B virus in human hepatocyte-derived cells. J Virol. United States. **85:** 1048-1057.

Marin, M., K. M. Rose, S. L. Kozak and D. Kabat (2003). "HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation." Nat Med **9**(11): 1398-1403.

McLean, A. A., M. R. Hilleman, W. J. McAleer and E. B. Buynak (1983). "Summary of worldwide clinical experience with H-B-Vax (B, MSD)." J Infect **7 Suppl 1**: 95-104.

Meier, J. L., X. Luo, M. Sawadogo and S. E. Straus (1994). "The cellular transcription factor USF cooperates with varicella-zoster virus immediate-early protein 62 to symmetrically activate a bidirectional viral promoter." Molecular and cellular biology **14**(10): 6896-6906.

Meyerhans, A., J. P. Vartanian and S. Wain-Hobson (1990). "DNA recombination during PCR." Nucleic Acids Res **18**(7): 1687-1691.

Mogensen, K. E., M. Lewerenz, J. Reboul, G. Lutfalla and G. Uze (1999). "The type I interferon receptor: structure, function, and evolution of a family business." J Interferon Cytokine Res **19**(10): 1069-1098.

Muhlemann, B., T. C. Jones, P. B. Damgaard, M. E. Allentoft, I. Shevnina, A. Logvin, E. Usmanova, I. P. Panyushkina, B. Boldgiv, T. Bazartseren, K. Tashbaeva, V. Merz, N. Lau, V. Smrcka, D. Voyakin, E. Kitov, A. Epimakhov, D. Pokutta, M. Vicze, T. D. Price, V. Moiseyev, A. J. Hansen, L. Orlando, S. Rasmussen, M. Sikora, L. Vinner, A. Osterhaus, D. J. Smith, D. Glebe, R. A. M. Fouchier, C. Drosten, K. G. Sjogren, K. Kristiansen and E. Willerslev (2018). "Ancient hepatitis B viruses from the Bronze Age to the Medieval period." Nature **557**(7705): 418-423.

Muramatsu, M., K. Kinoshita, S. Fagarasan, S. Yamada, Y. Shinkai and T. Honjo (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell. United States. **102:** 553-563.

Muramatsu, M., V. S. Sankaranand, S. Anant, M. Sugai, K. Kinoshita, N. O. Davidson and T. Honjo (1999). "Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells." J Biol Chem **274**(26): 18470-18476.

Nair, S. and A. Zlotnick (2018). "Asymmetric Modification of Hepatitis B Virus (HBV) Genomes by an Endogenous Cytidine Deaminase inside HBV Cores Informs a Model of Reverse Transcription." J Virol **92**(10).

Nassal, M. (1999). Hepatitis B virus replication: novel roles for virus-host interactions. Intervirology. Switzerland. **42:** 100-116.

Nassal, M. (2008). "Hepatitis B viruses: reverse transcription a different way." Virus Res **134**(1-2): 235-249.

Nguyen, D. H., S. Gummuluru and J. Hu (2007). Deamination-independent inhibition of hepatitis B virus reverse transcription by APOBEC3G. J Virol. United States. **81:** 4465-4472.

Nguyen, D. H. and J. Hu (2008). Reverse transcriptase- and RNA packaging signal-dependent incorporation of APOBEC3G into hepatitis B virus nucleocapsids. J Virol. United States. **82:** 6852-6861.

Nik-Zainal, S., L. B. Alexandrov, D. C. Wedge, P. Van Loo, C. D. Greenman, K. Raine, D. Jones, J. Hinton, J. Marshall and L. A. Stebbings (2012). "Mutational processes molding the genomes of 21 breast cancers." Cell **149**(5): 979-993.

Nik-Zainal, S., D. C. Wedge, L. B. Alexandrov, M. Petljak, A. P. Butler, N. Bolli, H. R. Davies, S. Knappskog, S. Martin and E. Papaemmanuil (2014). "Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer." Nature genetics **46**(5): 487-491.

Noguchi, C., H. Ishino, M. Tsuge, Y. Fujimoto, M. Imamura, S. Takahashi and K. Chayama (2005). "G to A hypermutation of hepatitis B virus." Hepatology **41**(3): 626-633.

Patient, R., C. Hourioux and P. Roingeard (2009). "Morphogenesis of hepatitis B virus and its subviral envelope particles." Cell Microbiol **11**(11): 1561-1570.

Patterson Ross, Z., J. Klunk, G. Fornaciari, V. Giuffra, S. Duchene, A. T. Duggan, D. Poinar, M. W. Douglas, J. S. Eden, E. C. Holmes and H. N. Poinar (2018). "The paradox of HBV evolution as revealed from a 16th century mummy." PLoS Pathog **14**(1): e1006750.

Pei, R. J., X. W. Chen and M. J. Lu (2014). "Control of hepatitis B virus replication by interferons and Toll-like receptor signaling pathways." World J Gastroenterol **20**(33): 11618-11629.

Phillips, S., R. Jagatia and S. Chokshi (2022). "Novel therapeutic strategies for chronic hepatitis B." Virulence **13**(1): 1111-1132.

Pollicino, T., L. Belloni, G. Raffa, N. Pediconi, G. Squadrito, G. Raimondo and M. Levrero (2006). "Hepatitis B virus replication is regulated by the acetylation status of hepatitis B virus cccDNA-bound H3 and H4 histones." Gastroenterology **130**(3): 823-837.

Popp, C., W. Dean, S. Feng, S. J. Cokus, S. Andrews, M. Pellegrini, S. E. Jacobsen and W. Reik (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature. England. **463:** 1101-1105.

Prasad, B. V. and M. F. Schmid (2012). "Principles of virus structural organization." Adv Exp Med Biol **726**: 17-47.

Proto, S., J. A. Taylor, S. Chokshi, N. Navaratnam and N. V. Naoumov (2008). APOBEC and iNOS are not the main intracellular effectors of IFN-gamma-mediated inactivation of Hepatitis B virus replication. Antiviral Res. Netherlands. **78:** 260-267.

Proto, S., J. A. Taylor, S. Chokshi, N. Navaratnam and N. V. Naoumov (2008). "APOBEC and iNOS are not the main intracellular effectors of IFN-gamma-mediated inactivation of Hepatitis B virus replication." Antiviral Res **78**(3): 260-267.

Purcell, R. H. (1993). "The discovery of the hepatitis viruses." Gastroenterology **104**(4): 955-963.

Rada, C., J. M. Jarvis and C. Milstein (2002). AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. Proc Natl Acad Sci U S A. United States. **99:** 7003-7008.

Robinson, W. S., D. A. Clayton and R. L. Greenman (1974). "DNA of a human hepatitis B virus candidate." J Virol **14**(2): 384-391.

Rogozin, I. B., M. K. Basu, I. K. Jordan, Y. I. Pavlov and E. V. Koonin (2005). APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases predicted by computational analysis. Cell Cycle. United States. **4:** 1281-1285.

Rosler, C., J. Kock, M. Kann, M. H. Malim, H. E. Blum, T. F. Baumert and F. von Weizsacker (2005). "APOBEC-mediated interference with hepadnavirus production." Hepatology **42**(2): 301-309.

Rösler, C., J. Köck, M. H. Malim, H. E. Blum and F. v. Weizsäcker (2004). "Comment on "Inhibition of Hepatitis B Virus Replication by APOBEC3G"."

Samuel, C. E. (2001). "Antiviral actions of interferons." Clin Microbiol Rev **14**(4): 778-809, table of contents.

Sato, Y., H. C. Probst, R. Tatsumi, Y. Ikeuchi, M. S. Neuberger and C. Rada (2010). Deficiency in APOBEC2 leads to a shift in muscle fiber type, diminished body mass, and myopathy. J Biol Chem. United States. **285:** 7111-7118.

Schmitz, A., A. Schwarz, M. Foss, L. Zhou, B. Rabe, J. Hoellenriegel, M. Stoeber, N. Pante and M. Kann (2010). "Nucleoporin 153 arrests the nuclear import of hepatitis B virus capsids in the nuclear basket." PLoS Pathog **6**(1): e1000741.

Schulze, A., P. Gripon and S. Urban (2007). "Hepatitis B virus infection initiates with a large surface protein-dependent binding to heparan sulfate proteoglycans." Hepatology **46**(6): 1759-1768.

Seeger, C. (2018). "Control of viral transcripts as a concept for future HBV therapies." Curr Opin Virol **30**: 18-23.

Seeger, C. and W. S. Mason (2000). Hepatitis B Virus Biology. Microbiol Mol Biol Rev. **64:** 51-68.

Seeger, C. and W. S. Mason (2013). Sodium-dependent taurocholic cotransporting polypeptide: a candidate receptor for human hepatitis B virus. Gut. England. **62:** 1093-1095.

Seeger, C. and W. S. Mason (2015). "Molecular biology of hepatitis B virus infection." Virology **479-480**: 672-686.

Seeger, C. and J. A. Sohn (2016). "Complete Spectrum of CRISPR/Cas9-induced Mutations on HBV cccDNA." Mol Ther **24**(7): 1258-1266.

Sells, M. A., M. L. Chen and G. Acs (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**(4): 1005-1009.

Sells, M. A., A. Z. Zelent, M. Shvartsman and G. Acs (1988). "Replicative intermediates of hepatitis B virus in HepG2 cells that produce infectious virions." J Virol **62**(8): 2836-2844.

Sirito, M., S. Walker, Q. Lin, M. T. Kozlowski, W. H. Klein and M. Sawadogo (1992). "Members of the USF family of helix-loop-helix proteins bind DNA as homo-as well as heterodimers." Gene Expression The Journal of Liver Research **2**(3): 231-240.

Sohn, J. A., S. Litwin and C. Seeger (2009). "Mechanism for CCC DNA synthesis in hepadnaviruses." PLoS One **4**(11): e8093.

Sowden, M. P., N. Ballatori, K. L. Jensen, L. H. Reed and H. C. Smith (2002). "The editosome for cytidine to uridine mRNA editing has a native complexity of 27S: identification of intracellular domains containing active and inactive editing factors." J Cell Sci **115**(Pt 5): 1027-1039.

Stoeckl, L., A. Funk, A. Kopitzki, B. Brandenburg, S. Oess, H. Will, H. Sirma and E. Hildt (2006). Identification of a structural motif crucial for infectivity of hepatitis B viruses. Proc Natl Acad Sci U S A. **103:** 6730-6734.

Suspene, R., D. Guetard, M. Henry, P. Sommer, S. Wain-Hobson and J. P. Vartanian (2005). Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. Proc Natl Acad Sci U S A. United States. **102:** 8321-8326.

Suspène, R., M. Henry, S. Guillot, S. Wain-Hobson and J. P. Vartanian (2005). "Recovery of APOBEC3-edited human immunodeficiency virus G->A hypermutants by differential DNA denaturation PCR." J Gen Virol **86**(Pt 1): 125-129.

Tanaka, Y., H. Marusawa, H. Seno, Y. Matsumoto, Y. Ueda, Y. Kodama, Y. Endo, J. Yamauchi, T. Matsumoto, A. Takaori-Kondo, I. Ikai and T. Chiba (2006). Anti-viral protein APOBEC3G is induced by interferon-alpha stimulation in human hepatocytes. Biochem Biophys Res Commun. United States. **341:** 314-319.

Teng, B., C. F. Burant and N. O. Davidson (1993). "Molecular cloning of an apolipoprotein B messenger RNA editing protein." Science **260**(5115): 1816-1819.

Tumanov, A. V., E. P. Koroleva, P. A. Christiansen, M. A. Khan, M. J. Ruddy, B. Burnette, S. Papa, G. Franzoso, S. A. Nedospasov, Y. X. Fu and R. A. Anders (2009). T cell-derived lymphotoxin regulates liver regeneration. Gastroenterology. United States. **136:** 694-704 e694.

Turelli, P., A. Liagre-Quazzola, B. Mangeat, S. Verp, S. Jost and D. Trono (2008). "APOBEC3-Independent Interferon-Induced Viral Clearance in Hepatitis B Virus Transgenic Mice."

Turelli, P., B. Mangeat, S. Jost, S. Vianin and D. Trono (2004). "Inhibition of hepatitis B virus replication by APOBEC3G." Science **303**(5665): 1829.

Valsamakis, A. (2007). Molecular testing in the diagnosis and management of chronic hepatitis B. Clin Microbiol Rev. United States. **20:** 426-439, table of contents.

Vartanian, J. P., M. Henry, A. Marchio, R. Suspene, M. M. Aynaud, D. Guetard, M. Cervantes-Gonzalez, C. Battiston, V. Mazzaferro, P. Pineau, A. Dejean and S. Wain-Hobson (2010). "Massive APOBEC3 editing of hepatitis B viral DNA in cirrhosis." PLoS Pathog **6**(5): e1000928.

Vasudevan, A. A., S. H. Smits, A. Hoppner, D. Haussinger, B. W. Koenig and C. Munk (2013). "Structural features of antiviral DNA cytidine deaminases." Biol Chem **394**(11): 1357-1370.

Vieira, V. C. and M. A. Soares (2013). "The role of cytidine deaminases on innate immune responses against human viral infections." Biomed Res Int **2013**: 683095.

Wald, A. (1980). "A method of estimating plane vulnerability based on damage of survivors, CRC 432, July 1980." Center for Naval Analyses.

Wang, H., S. Kim and W. S. Ryu (2009). "DDX3 DEAD-Box RNA helicase inhibits hepatitis B virus reverse transcription by incorporation into nucleocapsids." J Virol **83**(11): 5815-5824.

Watanabe, T., E. M. Sorensen, A. Naito, M. Schott, S. Kim and P. Ahlquist (2007). Involvement of host cellular multivesicular body functions in hepatitis B virus budding. Proc Natl Acad Sci U S A. United States. **104:** 10205-10210.

Wedekind, J. E., G. S. Dance, M. P. Sowden and H. C. Smith (2003). Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. Trends Genet. England. **19:** 207-216.

Will, H., R. Cattaneo, H. G. Koch, G. Darai, H. Schaller, H. Schellekens, P. M. van Eerd and F. Deinhardt (1982). "Cloned HBV DNA causes hepatitis in chimpanzees." Nature **299**(5885): 740-742.

World Health Organization. (2022). "Hepatitis B fact sheet." Retrieved June, 2022, from https://www.who.int/news-room/fact-sheets/detail/hepatitis-b.

Wyles, D. L. (2019). "Antiretroviral Effects on HBV/HIV Co-infection and the Natural History of Liver Disease." Clin Liver Dis **23**(3): 473-486.

Yan, H., G. Zhong, G. Xu, W. He, Z. Jing, Z. Gao, Y. Huang, Y. Qi, B. Peng, H. Wang, L. Fu, M. Song, P. Chen, W. Gao, B. Ren, Y. Sun, T. Cai, X. Feng, J. Sui and W. Li (2012). Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. eLife. **1**.

Yang, Y. and H. C. Smith (1997). "Multiple protein domains determine the cell type-specific nuclear distribution of the catalytic subunit required for apolipoprotein B mRNA editing." Proc Natl Acad Sci U S A **94**(24): 13075-13080.

Yuen, M. F., D. S. Chen, G. M. Dusheiko, H. L. A. Janssen, D. T. Y. Lau, S. A. Locarnini, M. G. Peters and C. L. Lai (2018). "Hepatitis B virus infection." Nat Rev Dis Primers **4**: 18035.

Zeng, Y., H. Li, X. Zhang, J. Shang and Y. Kang (2016). "Basal transcription of APOBEC3G is regulated by USF1 gene in hepatocyte." Biochem Biophys Res Commun **470**(1): 54-60.

Zhang, W., X. Zhang, C. Tian, T. Wang, P. T. Sarkis, Y. Fang, S. Zheng, X. F. Yu and R. Xu (2008). Cytidine deaminase APOBEC3B interacts with heterogeneous nuclear ribonucleoprotein K and suppresses hepatitis B virus expression. Cell Microbiol. England. **10:** 112-121.

Zhang, Y., C. Li, Y. Zhang, H. Zhu, Y. Kang, H. Liu, J. Wang, Y. Qin, R. Mao, Y. Xie, Y. Huang and J. Zhang (2013). "Comparative analysis of CpG islands among HBV genotypes." PLoS One **8**(2): e56711.

Zhao, D., X. Wang, G. Lou, G. Peng, J. Li, H. Zhu, F. Chen, S. Li, D. Liu, Z. Chen and Z. Yang (2010). APOBEC3G directly binds Hepatitis B virus core protein in cell and cell free systems. Virus Res. Netherlands, 2010 Elsevier B.V. **151:** 213-219.

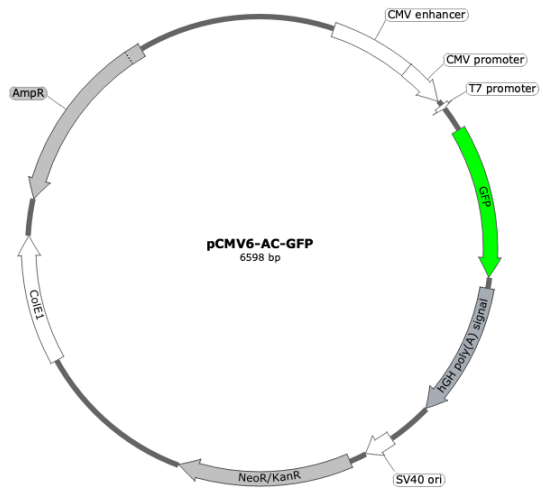
Zhou, S., C. S. Hill, M. U. Clark, T. P. Sheahan, R. Baric and R. Swanstrom (2021). "Primer ID Next-Generation Sequencing for the Analysis of a Broad Spectrum Antiviral Induced Transition Mutations and Errors Rates in a Coronavirus Genome." Bio Protoc **11**(5): e3938.

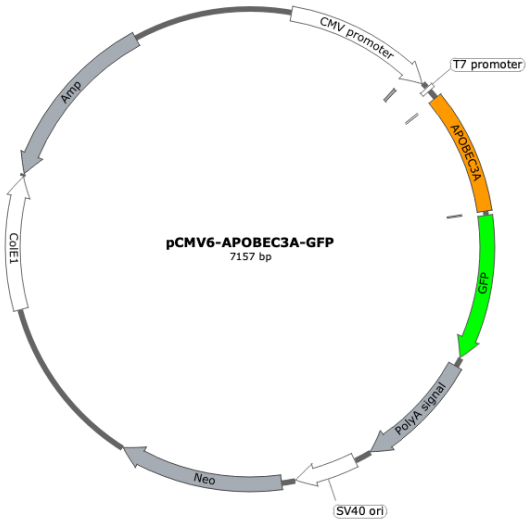
Zhou, S., C. Jones, P. Mieczkowski and R. Swanstrom (2015). "Primer ID Validates Template Sampling Depth and Greatly Reduces the Error Rate of Next-Generation Sequencing of HIV-1 Genomic RNA Populations." J Virol **89**(16): 8540-8555.

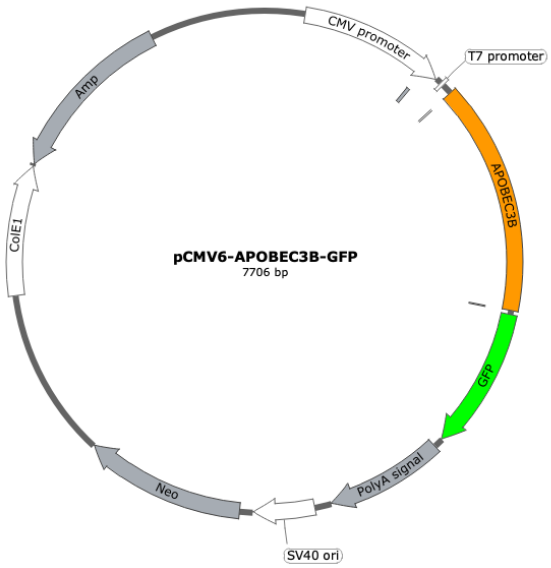
Zoulim, F. and C. Seeger (1994). "Reverse transcription in hepatitis B viruses is primed by a tyrosine residue of the polymerase." J Virol **68**(1): 6-13.

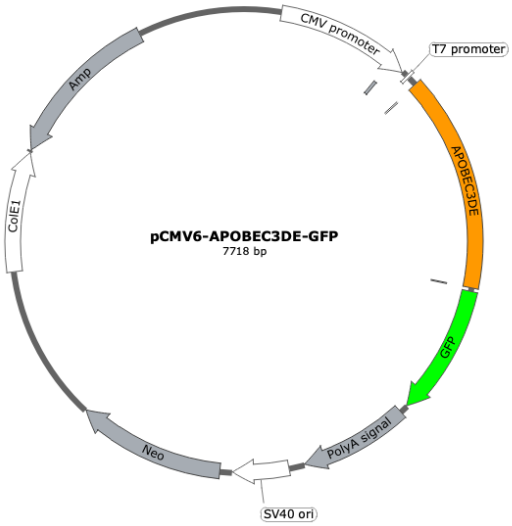
6 Appendix

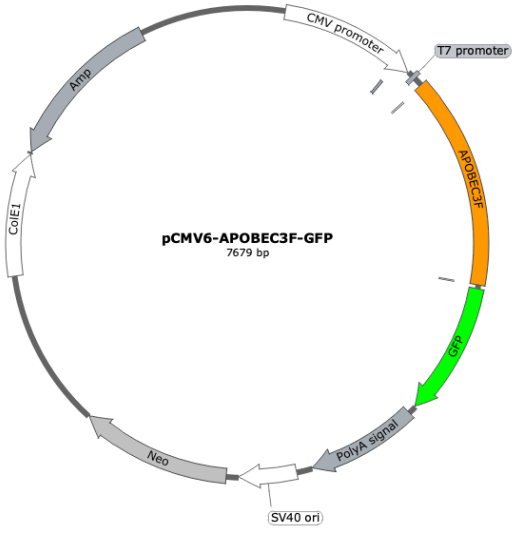
A.1 APOBEC3 plasmids maps

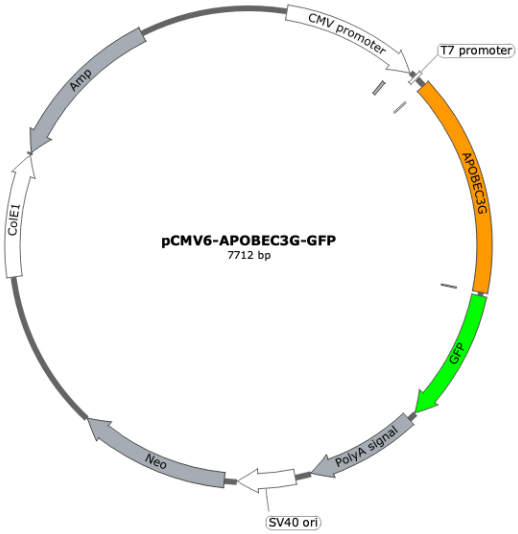




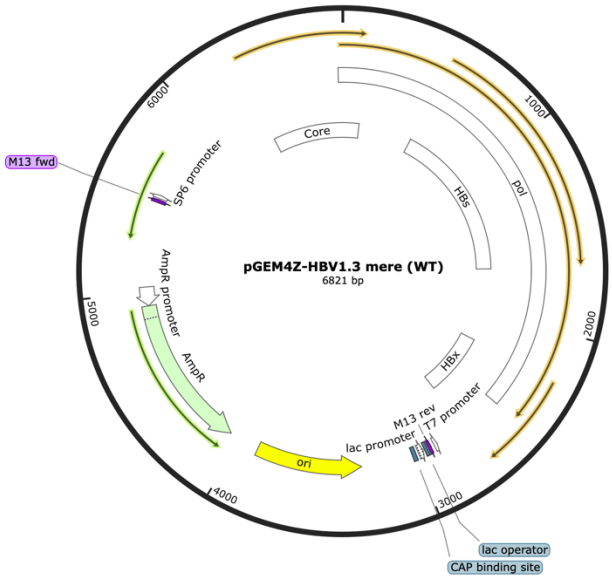




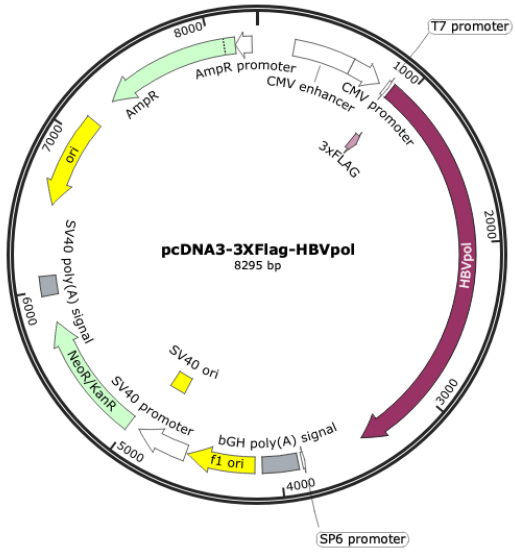




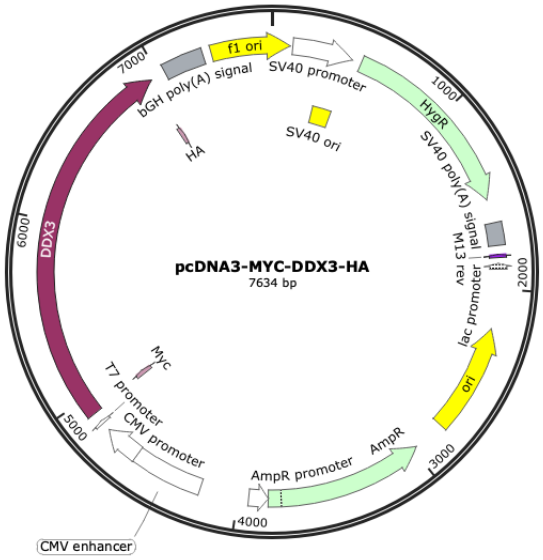
A.2 HBV1.3 mere (wildtype) plasmid map



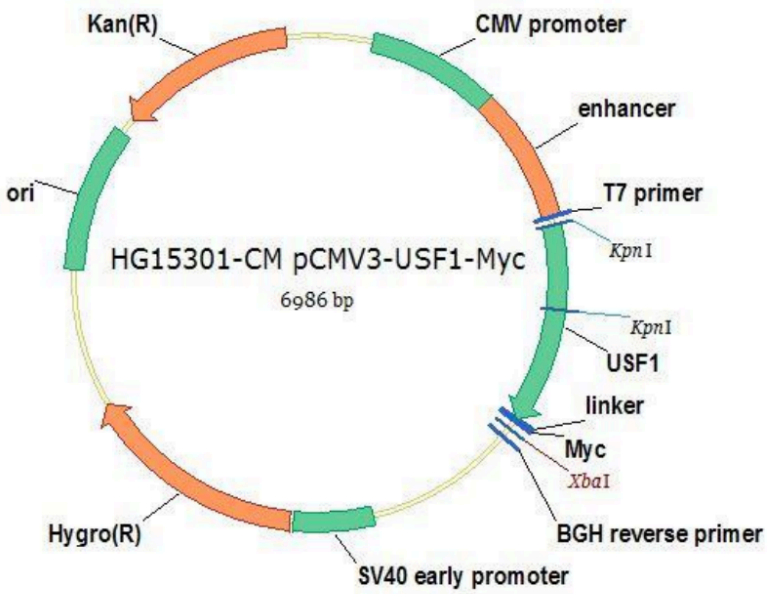
A.3 HBV polymerase plasmid map



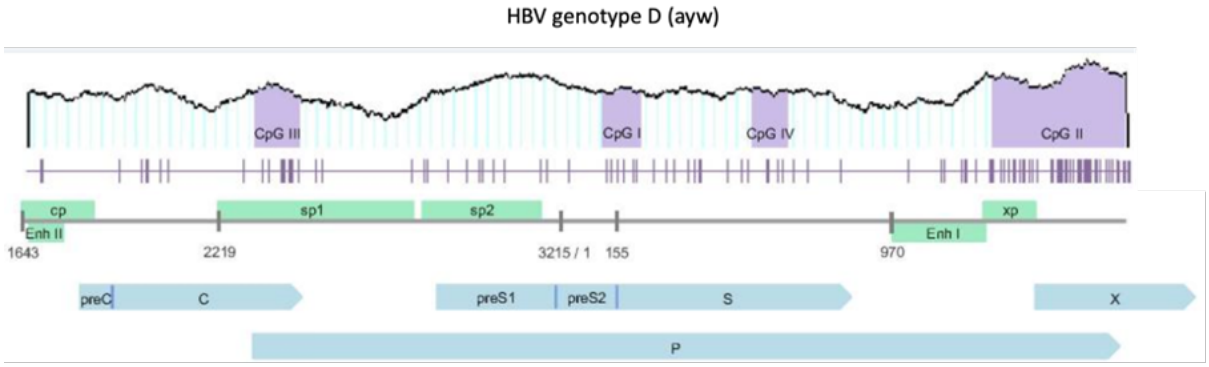
A.4 DDX3 plasmid map



A.5 USF1 plasmid map



A.6 The four GC regions (CpG islands) in the HBV genome (ayw). This figure adapted from (Zhang, Li et al. 2013).



The four GC regions (CpG islands) in the HBV genome (ayw). The polymerase, pre-core/core, S, and X genes (blue). The four promoters, enhancers I and II, (green). The GC regions (CpG islands I to VI) (purple areas) within the HBV genome and single CpG dinucleotide are shown as vertical purple line under the HBV genome. This figure adapted from (Zhang, Li et al. 2013).

A.7 The complete sequence of HBV ayw strain sequence

```
   1 AATTCCACAA CCTTCCACCA AACTCTGCAA GATCCCAGAG TGAGAGGCCT GTATTTCCCT
  61 GCTGGTGGCT CCAGTTCAGG AACAGTAAAC CCTGTTCTGA CTACTGCCTC TCCCTTATCG
 121 TCAATCTTCT CGAGGATTGG GGACCCTGCG CTGAACATGG AGAACATCAC ATCAGGATTC
 181 CTAGGACCCC TTCTCGTGTT ACAGGCGGGG TTTTTCTTGT TGACAAGAAT CCTCACAATA
 241 CCGCAGAGTC TAGACTCGTG GTGGACTTCT CTCAATTTTC TAGGGGGAAC TACCGTGTGT
 301 CTTGGCCAAA ATTCGCAGTC CCCAACCTCC AATCACTCAC CAACCTCTTG TCCTCCAACT
 361 TGTCCTGGTT ATCGCTGGAT GTGTCTGCGG CGTTTTATCA TCTTCCTCTT CATCCTGCTG
 421 CTATGCCTCA TCTTCTTGTT GGTTCTTCTG GACTATCAAG GTATGTTGCC CGTTTGTCCT
 481 CTAATTCCAG GATCCTCAAC AACCAGCACG GGACCATGCC GGACCTGCAT GACTACTGCT
 541 CAAGGAACCT CTATGTATCC CTCCTGTTGC TGTACCAAAC CTTCGGACGG AAATTGCACC
 601 TGTATTCCCA TCCCATCATC CTGGGCTTTC GGAAAATTCC TATGGGAGTG GGCCTCAGCC
 661 CGTTTCTCCT GGCTCAGTTT ACTAGTGCCA TTTGTTCAGT GGTTCGTAGG GCTTTCCCCC
 721 ACTGTTTGGC TTTCAGTTAT ATGGATGATG TGGTATTGGG GGCCAAGTCT GTACAGCATC
 781 TTGAGTCCCT TTTTACCGCT GTTACCAATT TTCTTTTGTC TTTGGGTATA CATTTAAACC
 841 CTAACAAAAC AAAGAGATGG GGTTACTCTC TAAATTTTAT GGGTTATGTC ATTGGATGTT
 901 ATGGGTCCTT GCCACAAGAA CACATCATAC AAAAAATCAA AGAATGTTTT AGAAAACTTC
 961 CTATTAACAG GCCTATTGAT TGGAAAGTAT GTCAACGAAT TGTGGGTCTT TTGGGTTTTG
1021 CTGCCCCTTT TACACAATGT GGTTATCCTG CGTTGATGCC TTTGTATGCA TGTATTCAAT
1081 CTAAGCAGGC TTTCACTTTC TCGCCAACTT ACAAGGCCTT TCTGTGTAAA CAATACCTGA
1141 ACCTTTACCC CGTTGCCCGG CAACGGCCAG GTCTGTGCCA AGTGTTTGCT GACGCAACCC
1201 CCACTGGCTG GGGCTTGGTC ATGGGCCATC AGCGCATGCG TGGAACCTTT TCGGCTCCTC
1261 TGCCGATCCA TACTGCGGAA CTCCTAGCCG CTTGTTTTGC TCGCAGCAGG TCTGGAGCAA
1321 ACATTATCGG GACTGATAAC TCTGTTGTCC TATCCCGCAA ATATACATCG TTTCCATGGC
1381 TGCTAGGCTG TGCTGCCAAC TGGATCCTGC GCGGGACGTC CTTTGTTTAC GTCCCGTCGG
1441 CGCTGAATCC TGCGGACGAC CCTTCTCGGG GTCGCTTGGG ACTCTCTCGT CCCCTTCTCC
1501 GTCTGCCGTT CCGACCGACC ACGGGGCGCA CCTCTCTTTA CGCGGACTCC CCGTCTGTGC
1561 CTTCTCATCT GCCGGACCGT GTGCACTTCG CTTCACCTCT GCACGTCGCA TGGAGACCAC
1621 CGTGAACGCC CACCAAATAT TGCCCAAGGT CTTACATAAG AGGACTCTTG GACTCTCAGC
1681 AATGTCAACG ACCGACCTTG AGGCATACTT CAAAGACTGT TTGTTTAAAG ACTGGGAGGA
1741 GTTGGGGGAG GAGATTAGGT TAAAGGTCTT TGTACTAGGA GGCTGTAGGC ATAAATTGGT
1801 CTGCGCACCA GCACCATGCA ACTTTTTCAC CTCTGCCTAA TCATCTCTTG TTCATGTCCT
1861 ACTGTTCAAG CCTCCAAGCT GTGCCTTGGG TGGCTTTGGG GCATGGACAT CGACCCTTAT
1921 AAAGAATTTG GAGCTACTGT GGAGTTACTC TCGTTTTTGC CTTCTGACTT CTTTCCTTCA
1981 GTACGAGATC TTCTAGATAC CGCCTCAGCT CTGTATCGGG AAGCCTTAGA GTCTCCTGAG
2041 CATTGTTCAC CTCACCATAC TGCACTCAGG CAAGCAATTC TTTGCTGGGG GGAACTAATG
2101 ACTCTAGCTA CCTGGGTGGG TGTTAATTTG GAAGATCCAG CGTCTAGAGA CCTAGTAGTC
2161 AGTTATGTCA ACACTAATAT GGGCCTAAAG TTCAGGCAAC TCTTGTGGTT TCACATTTCT
2221 TGTCTCACTT TTGGAAGAGA AACAGTTATA GAGTATTTGG TGTCTTTCGG AGTGTGGATT
2281 CGCACTCCTC CAGCTTATAG ACCACCAAAT GCCCCTATCC TATCAACACT TCCGGAGACT
2341 ACTGTTGTTA GACGACGAGG CAGGTCCCCT AGAAGAAGAA CTCCCTCGCC TCGCAGACGA
2401 AGGTCTCAAT CGCCGCGTCG CAGAAGATCT CAATCTCGGG AATCTCAATG TTAGTATTCC
2461 TTGGACTCAT AAGGTGGGGA ACTTTACTGG GCTTTATTCT TCTACTGTAC CTGTCTTTAA
2521 TCCTCATTGG AAAACACCAT CTTTTCCTAA TATACATTTA CACCAAGACA TTATCAAAAA
2581 ATGTGAACAG TTTGTAGGCC CACTCACAGT TAATGAGAAA AGAAGATTGC AATTGATTAT
2641 GCCTGCCAGG TTTTATCCAA AGGTTACCAA ATATTTACCA TTGGATAAGG GTATTAAACC
2701 TTATTATCCA GAACATCTAG TTAATCATTA CTTCCAAACT AGACACTATT TACACACTCT
2761 ATGGAAGGCG GGTATATTAT ATAAGAGAGA AACAACACAT AGCGCCTCAT TTTGTGGGTC
2821 ACCATATTCT TGGGAACAAG ATCTACAGCA TGGGGCAGAA TCTTTCCACC AGCAATCCTC
2881 TGGGATTCTT TCCCGACCAC CAGTTGGATC CAGCCTTCAG AGCAAACACC GCAAATCCAG
2941 ATTGGGACTT CAATCCCAAC AAGGACACCT GGCCAGACGC CAACAAGGTA GGAGCTGGAG
3001 CATTCGGGCT GGGTTTCACC CCACCGCACG GAGGCCTTTT GGGGTGGAGC CCTCAGGCTC
3061 AGGGCATACT ACAAACTTTG CCAGCAAATC CGCCTCCTGC CTCCACCAAT CGCCAGTCAG
3121 GAAGGCAGCC TACCCCGCTG TCTCCACCTT TGAGAAACAC TCATCCTCAG GCCATGCAGT
3181 GG
```

A.8 The actual DNA sequences of HBV pre-core, X, and S regions, its translated protein sequences, and the 3D-PCR primers sequences showing the size of PCR products.

HBV X region:

```
Protein     MetAlaAlaArgLeuCysCysGlnLeuAspProAlaArgAspValLeuCysLeuArgPro
Wildtype    ATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCCC
Primer      ATGGCTGCTAGGCTGTGCTGCCAA>

Protein     ValGlyAlaGluSerCysGlyArgProPheSerGlySerLeuGlyThrLeuSerSerPro
Wildtype    GTCGGCGCTGAATCCTGCGGACGACCCTTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCT

Protein     SerProSerAlaValProThrAspHisGlyAlaHisLeuSerLeuArgGlyLeuProVal
Wildtype    TCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCCGTC

Protein     CysAlaPheSerSerAlaGlyProCysAlaLeu
Wildtype    TGTGCCTTCTCATCTGCCGGACCGTGTGCACTT
Primer                <TAGACGGCCTGGCACACGTGAA
```

HBV preC region:

```
Protein                               MetGlnLeuPheHisLeuCysLeuIleIleSerCys
Wildtype    TAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAATCATCTCTTGT
Primer      TAAATTGGTCTGCGCACCAGCA>

Protein     SerCysProThrValGlnAlaSerLysLeuCysLeuGlyTrpLeuTrpGlyMetAspIle
Wildtype    TCATGTCCTACTGTTCAAGCCTCCAAGCTGTGCCTTGGGTGGCTTTGGGGCATGGACATC

Protein     AspProTyrLysGluPheGlyAlaThrValGluLeuLeuSerPheLeuProSerAspPhe
Wildtype    GACCCTTATAAAGAATTTGGAGCTACTGTGGAGTTACTCTCGTTTTTGCCTTCTGACTTC
Primer                                                              <AAG

Protein     PheProSerValArgAsp
Wildtype    TTTCCTTCAGTACGAGATC
Primer      AAAGGAAGTCATGCTCTAG
```

HBV S region:

```
Protein       LeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTyrGlnGlyMet
Wildtype    TCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATG
Primer      TCTTCATCCTGCTGCTATGCCTCAT>

Protein     LeuProValCysProLeuIleProGlySerSerThrThrSerThrGlyProCysArgThr
Wildtype    TTGCCCGTTTGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACC

Protein     CysMetThrThrAlaGlnGlyThrSerMetTyrProSerCysCysCysThrLysProSer
Wildtype    TGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCG

Protein     AspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheGlyLysPheLeuTrp
Wildtype    GACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGG

Protein     GluTrpAlaSerAerArgPheSerTrpLeuSerLeuLeuValProPheValGlnTrpPhe
Wildtype    GAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTC
Primer                                                <TAAACAAGTCACCAAG

Protein     ValGlyLeu
Wildtype    GTAGGGCTTT
Primer      CATCCCGAAA
```

A.9 The actual DNA sequences of HBV pre-core, X, and S regions, its translated protein sequences, and the Next-Generation Sequencing (NGS) primers ID sequences showing the size of PCR products.

HBV X region:

```
Protein                                          MetAlaAlaArgLeuCysCysGlnLeu
Wildtype                                         ATGGCTGCTAGGCTGTGCTGCCAACTG
Primer      TGTCCAGCACGCTTCAGGCTNNNNTNNNNTNNNNATGGCTGCTAGGCTGTGCTGCCAA>

Protein     AspProAlaArgAspValLeuCysLeuArgProValGlyAlaGluSerCysGlyArgPro
Wildtype    GATCCTGCGCGGGACGTCCTTTGTTTACGTCCCGTCGGCGCTGAATCCTGCGGACGACCC

Protein     PheSerGlySerLeuGlyThrLeuSerSerProSerProSerAlaValProThrAspHis
Wildtype    TTCTCGGGGTCGCTTGGGACTCTCTCGTCCCCTTCTCCGTCTGCCGTTCCGACCGACCAC

Protein     GlyAlaHisLeuSerLeuArgGlyLeuProValCysAlaPheSerSerAlaGlyProCys
Wildtype    GGGGCGCACCTCTCTTTACGCGGACTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGT
Primer                                                 <TAGACGGCCTGGCACA

Protein     AlaLeu
Wildtype    GCACTT
Primer      CGTGAA
```

HBV preC region:

```
Protein                                                          Met
Wildtype                                        TAAATTGGTCTGCGCACCAGCACCATG
Primer     TGTCCAGCACGCTTCAGGCTNNNNTNNNNTNNNNTAAATTGGTCTGCGCACCAGCA>

Protein     GlnLeuPheHisLeuCysLeuIleIleSerCysSerCysProThrValGlnAlaSerLys
Wildtype    CAACTTTTTCACCTCTGCCTAATCATCTCTTGTTCATGTCCTACTGTTCAAGCCTCCAAG

Protein     LeuCysLeuGlyTrpLeuTrpGlyMetAspIleAspProTyrLysGluPheGlyAlaThr
Wildtype    CTGTGCCTTGGGTGGCTTTGGGGCATGGACATCGACCCTTATAAAGAATTTGGAGCTACT

Protein     ValGluLeuLeuSerPheLeuProSerAspPhePheProSerValArgAsp
Wildtype    GTGGAGTTACTCTCGTTTTTGCCTTCTGACTTCTTTCCTTCAGTACGAGATC
Primer                                  <AAGAAAGGAAGTCATGCTCTAG
```

HBV S region:

```
Protein                                            LeuPheIleLeuLeuLeuCysLeu
Wildtype                                          TCCTCTTCATCCTGCTGCTATGCCTC
Primer      TGTCCAGCACGCTTCAGGCTNNNNTNNNNTNNNNTCTTCATCCTGCTGCTATGCCTCAT>

Protein     IlePheLeuLeuValLeuLeuAspTyrGlnGlyMetLeuProValCysProLeuIlePro
Wildtype    ATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTAATTCCA

Protein     GlySerSerThrThrSerThrGlyProCysArgThrCysMetThrThrAlaGlnGlyThr
Wildtype    GGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACC

Protein     SerMetTyrProSerCysCysCysThrLysProSerAspGlyAsnCysThrCysIlePro
Wildtype    TCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCC

Protein     IleProSerSerTrpAlaPheGlyLysPheLeuTrpGluTrpAlaSerAerArgPheSer
Wildtype    ATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCC

Protein     TrpLeuSerLeuLeuValProPheValGlnTrpPheValGlyLeu
Wildtype    TGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTT
Primer                          <TAAACAAGTCACCAAGCATCCCGAAA
```