

https://dspace.org/

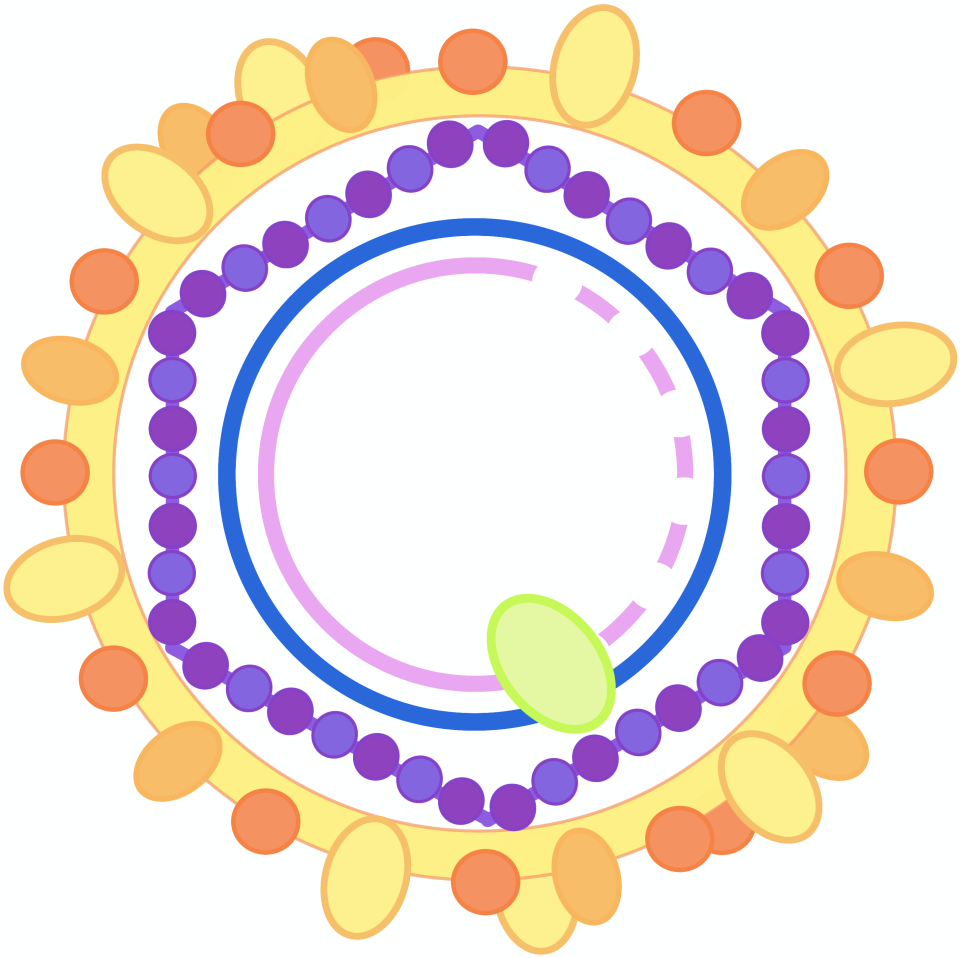
Functional characterisation of pre S1/pre S2 deletion mutants of hepatitis B virus isolated from Southern Africans

Wolhuter, Suzanne

2016
http://hdl.handle.net/10539/22277

Downloaded from a DSpace repository. https://dspace.org/

Functional Characterisation of PreS1/PreS2 Deletion Mutants of Hepatitis B Virus Isolated from southern Africans



Suzanne Wolhuter

Degree of Master in Science of Medicine
A dissertation submitted to the Faculty of Health Sciences at the University of the Witwatersrand, Johannesburg, South Africa in fulfillment of the requirements for the degree of Master in Science of Medicine

Johannesburg, 2016

Declaration

I, Suzanne Wolhuter, declare that this dissertation: Functional Characterisation of preS1/preS2 Deletion Mutants of Hepatitis B Virus Isolated from southern Africans, is my own original work and that all sources that I have used have been indicated by complete references. This thesis is being submitted for the degree of Masters of Science in Medicine at the University of the Witwatersrand, Johannesburg, South Africa. It has not been submitted before for any degree at this or any other University.

Signature of Candidate

Date

In memory of my late grandfather

Pieter Johan Gerard Kieser

29 June 1922 – 31 October 2007



` *Science begins with myth, and with the criticism of myths.* '
Karl Popper

Oupa, you taught me the power of the mind, and the power of thinking critically.

Publications and Presentations Arising from this Thesis

Presentations

Two presentations by the same title (Functional Characterisation of preS1/preS2 Deletion Mutants of Hepatitis B Virus Isolated from southern Africans) were given in an oral format at the International Hepatitis B Virus Conference in Bad Nauheim, Germany (October 2015) and at the Virology Africa Conference in Cape Town, South Africa (November 2015).

Abstract

Both HBV and HIV are hyperendemic in sub-Saharan Africa, and there is a correspondingly high incidence of hepatocellular carcinoma (HCC) in this region, contributing to a high burden of disease. In regions where HBV genotypes B and C prevail, a strong relationship exists between the rapid and more likely development of HCC and infection with HBV containing deletions in the preS region. Similar preS deletion mutants have been detected in southern African and Indian HCC patients infected with subgenotype A1. Disturbingly in a cohort study conducted in Mpumalanga, South Africa by our team, equivalent deletion-mutants were detected in 5 treatment-naïve HBV-HIV coinfected patients. Thus, the aim of this study was to characterize the quasispecies of HBV in these patients and to construct plasmids containing preS deletion-mutants in a subgenotype A1 backbone, in order to functionally characterize them *in vitro*. Such studies may determine how these preS deletions contribute to the working model of hepatocarcinogenesis and explain the high hepatocarcinogenic potential of subgenotype A1.

The quasispecies populations of HBV deletion mutants isolated from four HIV-positive patients were analysed phylogenetically, using both Neighbor-Joining and Bayesian methods. It was found that the preS deletion mutants represented the majority population, as 70% or more of clones in each of the four patients were sequentially highly similar to the respective parental strains. In addition to the major populations in each patient, minor populations were identified in the quasispecies.

The overlength subgenotype A1 wild-type replication competent plasmid, was successfully altered to remove an unwanted *Xba*I site. This altered construct served as a positive control to test whether the removal of this restriction site would affect replication and viral protein expression of this construct compared to the original plasmid. This plasmid was then successfully used as the backbone to construct three overlength deletion mutant constructs. Using a new strategy a 784 bp fragment flanking preS deletions from each of 4 patients, was inserted into the subgenotype A1 backbone.

Three fragments were derived from isolates with preS deletion mutants and one fragment, without the deletion, was isolated from a patient with occult HBV infection.

The resulting plasmids together with the appropriate control plasmids were used to transfect Huh7 cells in culture. Viral replication was followed at days 1, 3 and 5 using enzyme linked immunosorbent assay (ELISA) for hepatitis B surface antigen (HBsAg) and Hepatitis B e antigen (HBeAg). All deletion mutants were shown to express HBsAg and HBeAg at levels comparable to the controls, with the highest levels on day 3. There were no significant differences in the expression of HBeAg in Huh7 transfected cells between the deletion-mutant constructs and the positive controls. The HBV viral loads (VL) were measured at the same time points, using real-time quantitative PCR (qPCR). Both the extracellular (supernatant) and intracellular (lysate) compartment of the Huh7 cells were tested. Furthermore, to clarify whether encapsidated virus was being secreted, an immunocapture technique on the supernatant was employed prior to DNA extraction and qPCR. In a similar manner to the ELISA experiments, qPCR measurement of VL showed that despite the replacement of the 784 bp fragment with a fragment containing deletion mutations, all constructs were producing detectable amounts of HBV DNA suggesting that viral replication was occurring. Analogous to the HBsAg results, the highest VL was seen in all constructs on day 3 post transfection in both the supernatant and lysate experiments. Cell-associated (lysate compartment) VL was not particularly increased when compared to the VL measured in the supernatant, which does not indicate an inability to secrete the mature virions, nor an accumulation of viral DNA within the cell.

Negligible HBsAg expression was observed on days 1, 3 and 5 following transfection with the overlength construct that contained the 784 bp fragment derived from a patient with occult HBV infection. The VL following transfection with this construct was higher than both positive controls at all three time points, in both the supernatant and lysate compartment. This is again consistent with the clinical characteristics of the patient, which also had high viral loads. This finding is novel, since the *in vitro* experiments mimicked the phenotype of occult infection seen in the patient *in vivo*.

As far as we are aware, we are the first group to have constructed plasmids with deletion mutants and an occult mutant in a subgenotype A1 backbone and to show that they express viral proteins and viral DNA following transfection into HuH7 cells. These constructs will be important resource for further research into the high hepatocarcinogenic potential of subgenotype A1.

Acknowledgements

Firstly, I would like to thank my family, Hanneke Kieser, Gerda Nicholson, William Nicholson, Mark Wolhuter and Nick Gardener. Without your encouragement, support and constant love, I really would not have been able to produce this thesis, nor would I have survived the marathon that is a Masters in Science. You guys made all the difference, and I have so much love for you.

I would like to extend immense thanks to Anna Kramvis, my mentor, my supervisor and a role model of particular integrity. Prof you have not only guided me as a supervisor should, through the ups and downs of my Masters degree, but you have inspired, encouraged, and sometimes rescued me. Thank you for finding that amazing balance of being a supervisor, a friend as well as a person I constantly look up to and admire.

To my laboratory colleagues at the HVDRU, with special mention of Dr Trevor Bell, Dr Mukhlid Yousif and Dr Nimisha Bhoola, I want to say thank you so much. Particularly in terms of lending your expertise, for being willing to help build capacity and teach the baby Masters student, and for your endless support. Our unit has a culture of building people up, and supporting them as well as of course a social cohesion and some good kidding about that made doing a Masters with the HVDRU team a very special experience. A special mention also, to Roshni Desai, for all her help and assistance with the admin and financial aspects required around this project.

From the AGTRU, I'd like to extend thanks to Professor Patrick Arbuthnot and Dr Abdullah Ely. Thank you for allowing me time in your laboratory to carry out my Real-Time PCR experiments, and for being kind enough to lend advice and training too.

I would like to extend my deepest gratitude to the CANSA Research Division and the Poliomyelitis Research Fund, who were the generous benefactors that made this research possible. My thanks goes also to the Hepatitis B Foundation and WhiteSci South Africa for the substantial travel grants, without which I would not have been able to attend the 2015 International Molecular Biology of Hepatitis B Viruses meeting in Germany.

Table of Contents

		Page Number
1.	Chapter One: Introduction	1
1.1	The Hepatitis B Virus: Let's Get to Know Each Other	2
1.1.1	Classification: What's in a Name?	2
1.1.2	HBV Genome, Structure and Biology:	7
1.1.3	HBV Viral Proteins: Their Translation and Functions	12
	Surface Proteins	12
	The Hepatitis B X Protein	15
	The Enigmatic HBeAg and Core Proteins	16
	The Polymerase Enzyme	19
1.1.4	The Replication Cycle of HBV	20
1.1.5	HBV Transmission in an African Setting	22
1.1.6	Epidemiology of HBV – Worldwide and Southern Africa	24
1.1.7	Phylogeny and the Genotypes and Subgenotypes of HBV	26
	A Focus on Genotype A: Prevalent in Africa	28
1.1.8	HBV and HIV Coinfection: A Dangerous Liaison	29
	Occult HBV Infection	31
	Clinical Outcomes of HBV HIV Coinfection	32
1.1.9	HBV Mutants and PreS and S Deletions: What is Missing?	35
1.1.10	The Cell Culture Systems Available for Hepatitis B Virus Study	39
1.1.11	Study Rationale: Why Carry Out Functional Characterisation on S Deletion Mutants	41
2.	Chapter 2: Materials and Methods	45
	Overview and Workflow of the Study	45
2.1	List of Materials and Equipment	46
2.1.1	Complete List of Materials Used for this Study	46

2.1.2	Complete List of Equipment Used for this Study	48
2.1.3	Complete List of Bioinformatics Tools, Suites and Software Used for this Study	50
2.1.4	List of Plasmids Generated and Utilised During this Study	52
2.1.5	List of Bacterial Strains and Cell Culture Lines Utilised in this Study	53
2.2	Overview of Statistics Employed in this Study	53
2.3	Study Participants and Ethical Clearance	53
2.4	Quasispecies Analysis – General Recombinant Nucleic Acid Manipulation, Cloning and Phylogenetic Analysis	55
2.4.1	DNA Extraction	55
2.4.2	PCR Amplification	56
2.4.3	Cloning and Sequencing	57
2.4.4	Phylogenetic Analysis	59
2.5	Generation of Overlength Clones Containing PreS1 Mutations	60
2.5.1	Cloning Strategy Design	60
2.5.2	Replication Capacity of Clones	64
3.	Chapter 3: Results	67
	Overview and Workflow of the Study	67
3.1	Quasispecies Analysis - General Recombinant Nucleic Acid Manipulation, Cloning and Phylogenetic Analysis	68
3.1.1	PCR Amplification of the S Region and Sequencing	68
3.1.2	Cloning Amplicons and Sequencing Clones	70
3.1.3	Phylogenetic Analysis	71
3.2	Generation of Overlength Replication Competent Constructs with preS Mutations	79
3.2.1	Initial Restriction Mapping of the A12C15_OL A1WTSA1.281MER Construct	79

3.2.2	Alteration of the A12C15_OL A1WTSA1.281MER	81
3.2.3	Production of Final Overlength Constructs	84
3.2.4	Sequence Analysis of Overlength Clones	92
3.3	Determination of the Replication Competence of Overlength Clones	97
3.3.1	Expression of HBsAg and HBeAg	97
3.3.2	Viral Load Quantification of Overlength Constructs	102
4.	Chapter 4: Discussion	109
4.1	Quasispecies and Phylogenetic Analysis	110
4.2	Generation of Overlength Constructs for Transfection Experiments	114
4.3	Extracellular Protein Expression After Transfection	115
4.4	HBV Viral Loads Following Transfection	117
4.5	The Way Forward	119
4.6	Conclusion	120
5	Appendices	121
5.1	Appendix A – Detailed Protocols	121
5.2	Appendix B – Composition of Reagents and Solutions	147
5.3	Appendix C – Complete Plasmid Maps	151
5.4	Appendix D – Ethics Clearance Certificate	159
5.5	Appendix E – Molecular Weight Markers	160
5.6	Appendix F – Protocols for Phylogenetic Tree Production	161
5.7	Appendix G - Complete List of Sequences Used for Phylogenetic Trees	164
6.	References	165

List of Figures

Figure		Page Number
	Chapter One: Introduction	
1.1	Relative Size of Hepatitis B Virus to Other Viruses	8
1.2	Representation of the HBV Genome, Transcriptome and Subsequent Viral Proteins	10
1.3	Representation of the Viral Proteins in the Mature Virion	11
1.4	Important Functional Domains of the LHBs	14
1.5	Graphic Representation of the 'a'-determinant Region	16
1.6	Structure of the HBV Polymerase	19
1.7	Stages of the Replication Cycle	20
1.8	Facts and Figures of the Global HBV Epidemic	24
1.9	Global HBsAg Endemicity (1957-2013)	25
1.10	Overview of Geographical Distribution of HBV Genotypes	27
1.11	The Prevalence of HBV and the HIV Globally	30
1.12	Shongwe Cohort Outcomes on Coinfection and Occult HBV	41
1.13	Graphical Representation of Sequences Selected from the Shongwe Study	43
	Chapter Two: Materials and Methods	
2.1	Algorithm for DNA Extraction	55
2.2	Nested PCR Strategy, Primers and Amplicons	56
2.3	The PCR Cycling Conditions for Amplification of the Entire S Region	57
2.4	Graphical Representation of Sequencing Primers Used to for Determining DNA Sequence of Quasispecies.	58
2.5	Layout of the Overlength A12C15_OL A1WTSA1.281MER Backbone and the ORFs of HBV as Well as Identified Restriction Sites	60

2.6	Algorithm to Alter the Backbone of the Plasmid for Subcloning Purposes	61
2.7	Algorithm for 784 bp Fragment Generation	62
2.8	Identification of Suitable Restriction Sites and Cloning Workflow	63
2.9	The GFP Containing Plasmid Used to Determine Transfection Efficiency	65
	Chapter Three: Results	
3.1	High Fidelity PCR of the S Region	68
3.2	Alignment of PCR Product Against Parent Sequence	69
3.3	Screening for Clones Using Restriction by *Xba*I	70
3.4	Neighbor Joining Phylogenetic Tree	74
3.5	Neighbor Joining Phylogenetic Tree With All the Genotypes	76
3.6	Maximum Clade Credibility Tree with all the Genotypes	78
3.7	Plasmid Map of the A12C15_OL A1WTSA1.281MER Vector	79
3.8	Restriction Mapping of A12C15_OL A1WTSA1.281MER Vector	80
3.9	Restriction of A12C15_OL A1WTSA1.281MER with *Apa*I and *Not*I	81
3.10	Plasmid Map of the A12C15_OL A1WTSA1.281MER Altered 9.3 Vector	82
3.11	Restriction Mapping of A12C15_OL A1WTSA1.281MER Altered 9.3	83
3.12	Plasmid Map of the OL_DELMUT SHH011A- F Construct	84
3.13	Restriction Mapping of the OL_DELMUT SHH011A- F Construct	85
3.14	Plasmid Map of the OL_DELMUT SHH045A- F Construct	86
3.15	Restriction Mapping of the OL_DELMUT SHH045A- F Construct	87
3.16	Plasmid Map of the OL_DELMUT SHH167A- F Construct	88
3.17	Restriction Mapping of the OL_DELMUT SHH167A- F Construct	89
3.18	Plasmid Map of the OL_FLCC SHH193A- F Construct	90
3.19	Restriction Mapping of the OL_FLCC SHH193A- F Construct	91
3.20	Translated aa Sequence of the preS1 Region of All Constructs	93

3.21	Translated aa Sequence of the preS1 Region of All Constructs	94
3.22	Translated aa Sequence of the preS2 Region of All Constructs	95
3.23	Detection of Mean HBsAg Expression in Culture Supernatant of Transfected Huh7 Cells	98
3.24	Normalised HBsAg Expression in Culture Supernatant of Transfected Huh7 Cells	98
3.25	Detection of HBeAg Expression in Culture Supernatant of Transfected Huh7 Cells	100
3.26	Normalised HBeAg Expression in Culture Supernatant of Transfected Huh7 Cells	100
3.27	Expression of HBsAg relative to HBeAg	101
3.28	Box and Whisker plot of HBV DNA Expression in Culture Supernatant of Huh7 Cells	103
3.29	Normalised HBV DNA Expression in Culture Supernatant of Huh7 Cells	103
3.30	Box and Whisker Plot of Immunocaptured Supernatant HBV DNA Expression of Huh7 Cells	105
3.31	Normalised Secreted Virion HBV DNA Expression in Immunocaptured Supernatant of Huh7 Cells	105
3.32	Box and Whisker plot of HBV DNA Expression in Culture Lysates of Huh7 Cells	107
3.33	Normalised HBV DNA Expression in Culture Lysates of Huh7 Cells	107

List of Tables

Table		Page Number
	Chapter One: Introduction	
1.1	Virus Classification Systems Summary as Related to HBV	3
1.2	The Two Major Genera of the Hepadnaviridae Family	5
1.3	Summary of All Hepatitis Viruses that Affect the Liver	7
1.4	Different Length Transcripts Result in Three Surface Proteins	13
1.5	Different Length Transcripts of the pre-Core and Core Genes	18
1.6	Stages of the HBV Replication Cycle	21
1.7	HBV Routes of Transmission	22
1.8	The Effect of Coinfection on Disease Outcome for bot HBV and HIV	33
1.9	HBV Variants and Their Relevance	36
1.10	Comparison of Cell Lines used for HBV	40
1.11	Summary of Patient Sequences Selected from the Shongwe Study	42
	Chapter Two: Materials and Methods	
2.1	Constructs Generated and Utilised During this Study	52
2.2	Summary of the Bacterial Strains and Cell Lines used for this Study	53
2.3	Key Clinical Data for Selected Shongwe Patients	54
	Chapter Three: Results	
3.1	Inter- and Intra- Group Divergence within quasispecies and their parents	71
3.2	Molecular characteristics of clones relative to the parental strain	72
3.3	Summary of restriction enzymes utilised and expected band sizes	80
3.4	Summary of restriction enzymes utilised and expected band	83

	sizes	
3.5	Summary of restriction enzymes utilised and expected band sizes	85
3.6	Summary of restriction enzymes utilised and expected band sizes	87
3.7	Summary of restriction enzymes utilised and expected band sizes	89
3.8	Summary of restriction enzymes utilised and expected band sizes	91
3.9	Functional domains pertaining to Translation of the LHBs	92
3.10	Summary of Mutations Observed in Final Constructs preS1 Region	96
3.11	Significant Differences in HBsAg Expression	99
3.12	Significant Differences in Supernatant Viral Loads	104
3.13	Significant Differences in Immunocapture Viral Loads	106
3.14	Significant Differences in Lysate Viral Load	108

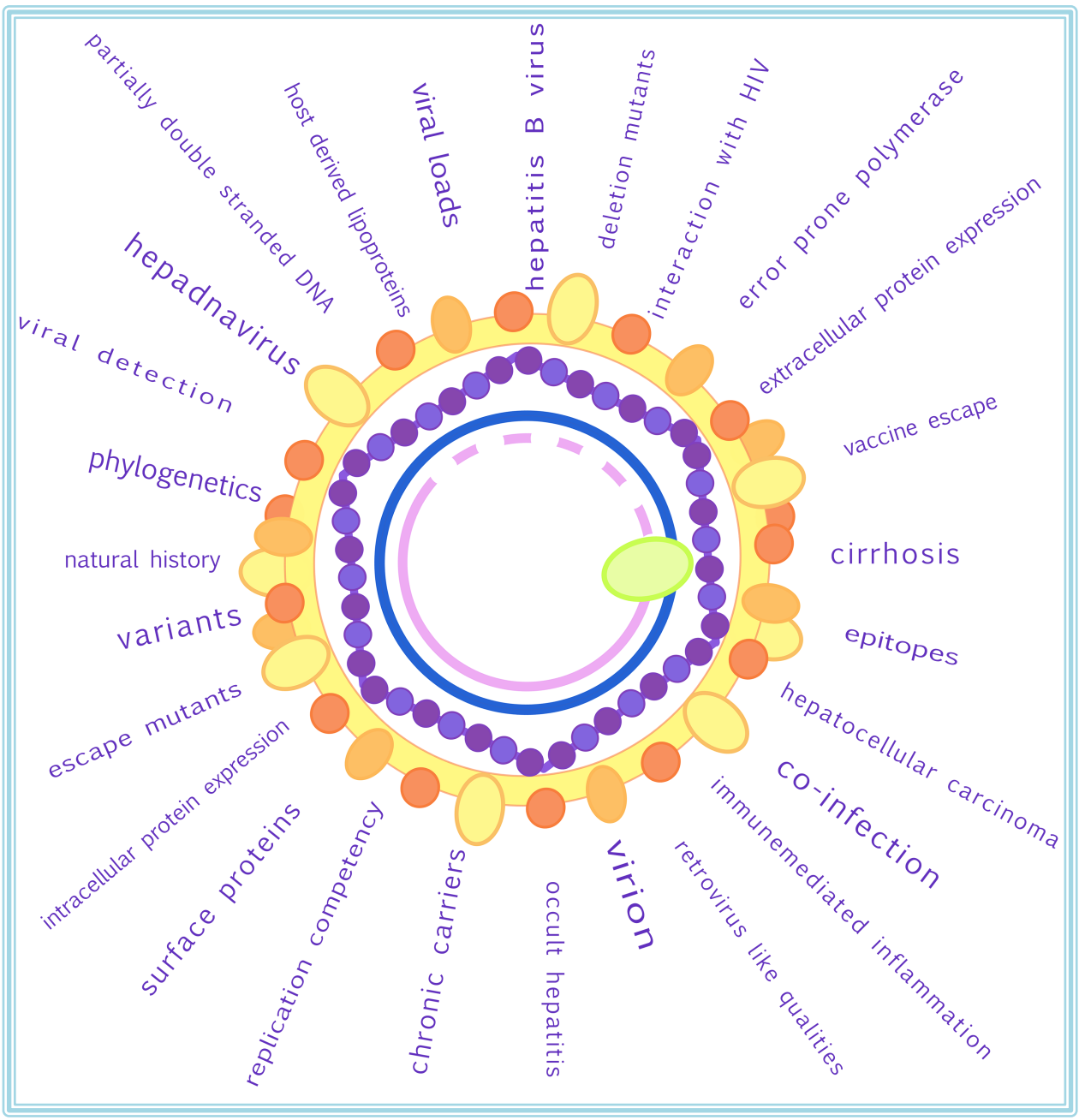
List of Abbreviations and Acronyms:

+DNA	plus strand DNA (as in positive sense DNA)
-DNA	minus strand DNA (as in negative sense DNA)
+RNA	plus strand RNA (as in positive sense RNA)
-RNA	minus strand RNA (as in negative strand RNA)
+ssDNA	plus single strand DNA
-ssDNA	minus single strand DNA
+dsDNA	plus double strand DNA
-dsDNA	minus double strand DNA
aa/s	amino acid/s
AGSHBV	Arctic Ground Squirrel hepatitis B virus
AGSHV	Ashy Headed Sheldgoose HBV
AIDS	autoimmune deficiency syndrome
anti-HBe	anti-hepatitis B e antibodies
anti-HBs	anti-hepatitis B s antibodies
ART	anti-retroviral treatment/therapy
bp	base pairs
BtHBV	Bat HBV
BQW	best-quality water
cccDNA	covalently closed circular DNA
CHB	chronic HBV in infection
CHBV	Crane hepatitis B virus
ChHBV	Chimpanzee hepatitis B virus
COX-2	cyclooxygenase-2
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
ELISA	enzyme-linked immunosorbent assay

ER	endoplasmic reticulum
FCS	Foetal Calf Serum
GFP	Green Fluorescent Protein
GGH	Ground Glass Hepatocytes
GHBV	Gorilla hepatitis B virus
GSHV	Ground Squirrel hepatitis B virus
GTR	generalised time-reversible
HAART	highly active anti-retroviral therapy
HBIG	hepatitis B immunoglobulin
HBeAg	hepatitis B e antigen
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HBx	Hepatitis B X protein
HHBV	Heron hepatitis B virus
HIV	human immunodeficiency virus
HSPG	heparan-sulfate proteoglycans
hTERT	human telomerase reverse transcriptase
IDT	Integrated DNA Technologies
ICTV	International Committee on Taxonomy of Viruses
IRIS	immune reconstitution inflammatory syndrome
LHBs	large hepatitis B surface protein
MHBs	middle hepatitis B surface protein
MAPK	Mitogen-activated protein kinases
mRNA	messenger Ribonucleic acid
mt	mutant
mTOR	mammalian target of rapamycin
NAT	nucleic acid testing
nt	nucleotide/s
NTCP	sodium taurocholate cotransporting polypeptide
NF-KB	nuclear factor kappa-light-chain-enhancer of activated B cells

ORF/s	open reading frame/s
OuHV	Orang-utan hepatitis B virus
PCR	polymerase chain reaction
PEI	Polyethylenimine
PIC	proteinase inhibitor cocktail
PHBV	Parrot hepatitis B virus
pgRNA	pregenomic RNA
PLHIV	people living with HIV
qPCR	quantitative Real-Time PCR
RB	retinoblastoma
RB/HBHBV	Old World Bats hepatitis B virus
RCT	randomized control trials
rcDNA	relaxed circular DNA
RGHBV	Ross Goose hepatitis B virus
SCID	severe combined immunodeficiency
SGHBV	Snow Goose hepatitis B virus
SHBs	small hepatitis B surface protein
STHBV	Stork hepatitis B virus
TBHBV	New World Bats hepatitis B virus
TSHBV	Tree Shrew (*Tupaia Belangeri*) hepatitis B virus
VL	Viral Load
WHV	Woodchuck hepatitis virus
WMHV	Wooly Monkey hepatitis B virus
WT	wild-type

1 Chapter One: Introduction



Hepatitis B Virus (HBV), the smallest DNA virus known to infect humans, has successfully infected a third of the world's population [1-3]. This literature review will attempt to look at the classification, structure and molecular biology of this virus; how this virus impacts its human host; and what happens when the genome of the virus is altered by naturally occurring deletions. Finally, the rationale for the study will be presented.

1.1 The Hepatitis B Virus: Let's Get to Know Each Other

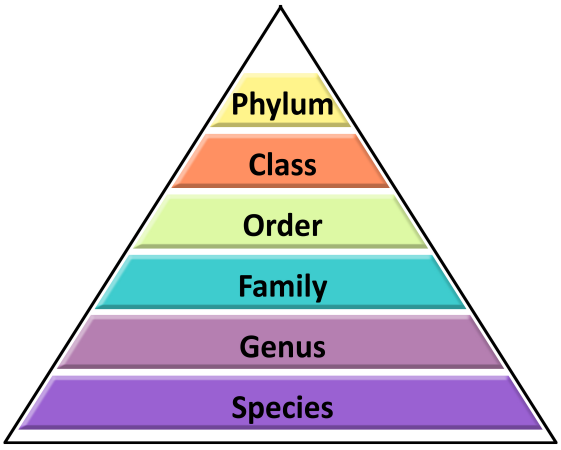
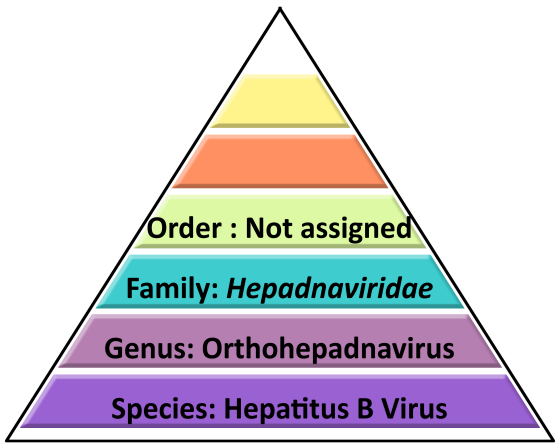
1.1.1 Classification: What's in a Name?

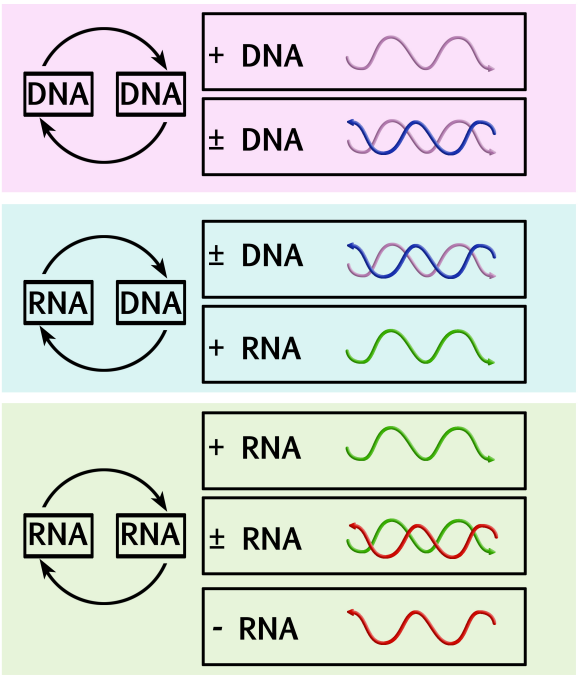
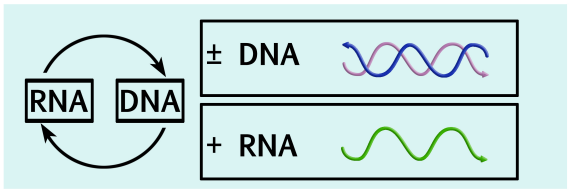
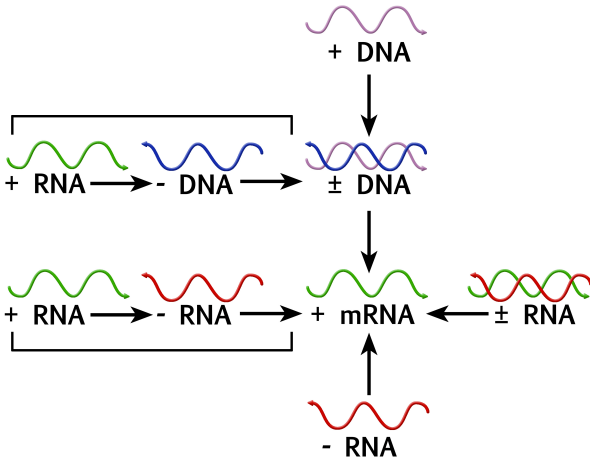
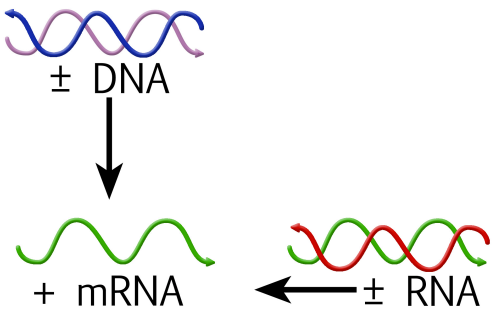
While the virologist V. R. Racaniello neatly summarises that all viruses are "obligate intracellular parasites", and this describes the nature of viruses perfectly, it does not give us enough resolution when we are looking at one particular virus [4]. The classification of viruses is often confusing or contradictory; one of the reasons for this is the difficulty in defining what a virus is. Most leaders in the field of virology seem to be satisfied with defining viruses as non-living entities, falling within a certain parameter of size that have some of the characteristics of living organisms [4]. However, this definition is complicated because it suggests that viruses are then not organisms or microbes (as they are not living), and they cannot all be called pathogens, as they do not always cause illness or disease [4]. Therefore, we should consider other important factors such as which hosts or species the virus infects and affects, which system or organ (if any) is targeted by the virus, the viral morphology, and the organisation of the viral genome [4]. Nassal *et al.,* (2005) neatly summarised "*the fundamental properties of a functional virus are the ability to replicate the genome, to form infectious virions, and to cope with the host defense in order to establish and maintain infection; a virus meeting all these criteria is biologically fit*" [5].

Another source of confusion is that there are a few different systems of classification for viruses [4]. The classical system, which is still in use today for the classification of animal viruses, is the Linnaean Hierarchical System. The Hierarchy consists of phylum, class, order, family, genus and finally species. This system also takes into account the kind of nucleic acid present, the symmetry of the capsid or protein shell as well as its dimensions and finally the presence or absence of a lipid membrane or envelope [4]. It is still the primary system used by the International Committee on Taxonomy of Viruses (ICTV), which was established in 1966 and released its latest Virus Taxonomy report in 2011 [4, 6]. Classification by Genome Type is the second most commonly used system. Here viruses are known by their classical term as well as being grouped into one of seven genome types that represent all the virus families [4, 6]. Thirdly, the most modern system of classification is the Baltimore System. Here viruses are described by the relationship of the viral genome to its messenger Ribonucleic Acid (mRNA). Thus the

Baltimore System take into account the Francis and Crick central dogma of the flow of nucleic acid to protein synthesis in the traditional sense of DNA transcribed into RNA and then translated into protein [4]. Table 1.1 summarises the three classification systems and illustrates where HBV fits in.

Table 1.1: Virus Classification Systems Summary as Related to HBV

Viral Classification Systems and Hepatitis B Virus		
	General	Hepatitis B Virus
Linnaean Hierarchy	Classical system of viral taxonomy by assigning a hierarchy based on: 1. Nature of nucleic acid 2. Symmetry of Protein Shell 3. The presence of a lipid envelope 4. The measurements and symmetry of the viral capsid 	Order - Not assigned Family - *Hepadnaviridae* Genus - Orthohepadnavirus Species - Hepatitis B Virus 

Genome Type	The viral genome carries the entire blueprint of viral propagation and thus survival. All 22 families of viruses are grouped into seven genome types: 	HBV falls into the seventh genome type as the virus has gapped double-stranded DNA, and the virus must replicate via an RNA intermediate. 
Baltimore Classification	All viruses have to generate mRNA to be translated into proteins by host ribosomes in order to complete their life cycle. The pathways to generate these mRNAs are dictated by the viral genome type. This classification system points out the obligatory relationship between viral genomes and the mRNA it must produce. 	According to the Baltimore scheme, HBV has a partially double stranded DNA genome, that at one point uses reverse transcriptase to replicate this genome via an RNA intermediate, and is thus similar to retroviruses in animals and pararetroviruses in plants. 

The table was constructed drawing on multiple references [4, 6-9].

Therefore, when we refer to hepatitis B virus as the prototype of the family *Hepadnaviridae*, we mean that it is the most well-known and understood member of this family of viruses, or a model/exemplar of this family. The word *Hepadnaviridae* is constructed from *Hepa*, from classic Greek meaning liver, and *dna* referring to the type of nucleic acid [3]. In the *Hepadnaviridae* family, there are two major groups or genera, namely *Orthohepadnaviridae,* which infect mammals and *Avihepadnaviridae,* which infect birds [3, 9-12]. This classification is based on host specificity, similarities in organ tropism, ability to cause both acute and persistent infection as well as replication strategy. See Table 1.2 [3, 9, 13].

Table 1.2: The Two Major Genera of the *Hepadnaviridae* Family

Hepadnaviridae Family	
Orthohepadnaviruses (Infect mammals)	*Avihepadnaviruses* (Infect birds)
Human hepatitis B Virus (HBV)*	Duck hepatitis B virus (DHBV)*
Non-human primate HBV: Chimpanzee hepatitis B virus (ChHBV) Wooly monkey hepatitis B virus (WMHV) Gorilla hepatitis B virus (GHBV) Orang-utan hepatitis B virus (OuHV)	Miscellaneous bird HBV: Parrot hepatitis B virus (PHBV) Heron hepatitis B virus (HHBV) Stork hepatitis B virus (STHBV) Crane hepatitis B virus (CHBV)
Rodent HBV: Woodchuck hepatitis virus (WHV) Arctic ground squirrel hepatitis B virus (AGSHBV) Tree shrew (*Tupaia Belangeri*) hepatitis B virus (TSHBV) Ground squirrel hepatitis B virus (GSHV)	Geese HBV: Snow goose hepatitis B virus (SGHBV) Ashy headed sheldgoose HBV (AGSHV) Ross goose hepatitis B virus (RGHBV)
Bat HBV (BtHBV): New world bats hepatitis B virus (TBHBV) Old world bats hepatitis B virus (RB/HBHBV)	

This table was compiled by drawing on a number of literature resources [2, 3, 6, 8, 10, 14-17] to illustrate the host range of the *Hepadnaviridae* family of viruses. The prototype member of each genus is indicated with an asterisk (*) [17, 18]. While all *Hepadnaviridae* show very narrow host specificity, some genera have a narrower host range than others, for example; Woodchuck hepatitis virus (WHV) occurs naturally in Marmots, but cannot infect Ground Squirrels [3, 17]. Of note: Human HBV can experimentally infect Chimpanzees and Chacma Baboons, and only transiently infect primary hepatocytes of the Woodchuck, Duck and Tree Shrew [16, 18]. A stable animal model for infection includes transgenic animals such as severe combined immunodeficiency (SCID) mice, but only where their livers have been humanised [19]. Further, a putative third genus of hepadnaviruses, white sucker hepatitis B virus (WSHBV), was recently discovered in the genetic material of a White Sucker fish (*Catostomus commersonii*). This however, remains to be confirmed [20]. Endogenous integrated hepadnavirus DNA has also been detected in the Zebra finch (*Taeniopygia guttata*); these kinds of integrated HBV sequences found in host genomes are sometimes referred to as "fossilised" viral sequences as they can give clues to the origin and deep evolution of the virus [10, 21].

HBV has been shown to primarily infect liver cells. However, HBV has also been found in other tissues like lymphoid cells, such as bone marrow cells and peripheral blood mononuclear cells (PMBCs) as well as spleen and kidney cells [11]. While it has been shown that HBV has no disease development in those extrahepatic tissues, viral presence and a reservoir of cccDNA in these sites may be important for establishing viral persistence [11]. Furthermore, HBV is often grouped together with other viruses that also cause liver disease, or have what is referred to as liver tropism, as these viruses infect mostly liver cells. These include Hepatitis A, C, D and E the main features of which are briefly summarised in Table 1.3 below.

Table 1.3: Summary of All Hepatitis Viruses that Affect the Liver

Virus	Virology	Clinical Manifestation and Route of Transmission
Hepatitis A virus	Order: *Picornavirales* Family: *Picornaviridae* Genus: *Hepatovirus* Species: *Hepatitis A virus* Structure: 27 nm non-enveloped spherical particle Genome: +ssRNA	Acute hepatitis, self-limiting Water-borne, enteral or oro-faecal
Hepatitis B virus	Order: none Family: *Hepadnaviridae* Genus: *Orthohepadnavirus* Species: Hepatitis B virus Structure: 42 nm Dane particle, enveloped Genome: partially dsDNA	Asymptomatic acute phase, followed by possible protracted asymptomatic carrier status. Subsequent probable development of inflammatory liver disease including cirrhosis and liver cancer. Complicated by co-infection with HIV Blood-borne, parenteral
Hepatitis C virus	Order: none Family: *Flaviviridae* Genus: *Hepacivirus* Species: *Hepatitis C virus* Structure: 55-65 nm in diameter, enveloped Genome: +ssRNA	Acute phase can be asymptomatic or include jaundice. Chronic infection can lead to cirrhosis, liver cancer and immune disorders. Complicated by co-infection with HIV Blood-borne, parenteral
Hepatitis D virus	Order: none Family: none Genus: *Deltavirus* Species: *Hepatitis delta virus* Structure: 36 nm, coated with HBsAg Often called a viroid/satellite virus Genome: circular –ssRNA, smallest viral genome known to infect humans	Only causes disease in individuals co-infected with Hepatitis B virus, thus often referred to as a satellite virus. Requires to be coated in HBsAg for replication to occur. Co-infection of HBV and HDV increases the likelihood of fulminant hepatitis. Blood-borne, parenteral
Hepatitis E virus	Order: none Family: *Hepeviridae* Genus: *Hepevirus* Species: *Hepatitis E virus* Structure: 27-34 nm, non-enveloped spherical particle Genome: +ssRNA	Acute hepatitis, self-limiting Water-borne, enteral or oro-faecal

Several resources, including textbooks, lecture notes and reports were used to construct the following table [2, 4, 22, 23]. (Note size refers to viral diameter, ssRNA stands for single-stranded Ribonucleic acid, whereas dsDNA stands for double-stranded ribonucleic acid, the + and – denote plus and minus strand)

1.1.2 HBV Genome, Structure and Biology:

All viral classification systems place a lot of importance on the genome; this is because the nature of the viral genome predicts how the virus will replicate, which proteins it will produce and thus ultimately, how it will interact with its host/s [4]. The HBV genome is a very compact genome of ~3200 nucleotides in length and consists of partially double-stranded DNA [2, 6]. Taken together with the fact that each virion is only 42-47 nm in diameter, HBV is the smallest known DNA pathogen to infect humans [1-3, 18, 24-28]. Moreover the fact that the genome consists of overlapping reading frames means that the HBV has a fairly unique method of replication, which will be discussed below [3]. The virions, or infectious particles themselves are named Dane particles after the primary author on the paper of HBV's debut visualisation under the electron microscope [29].

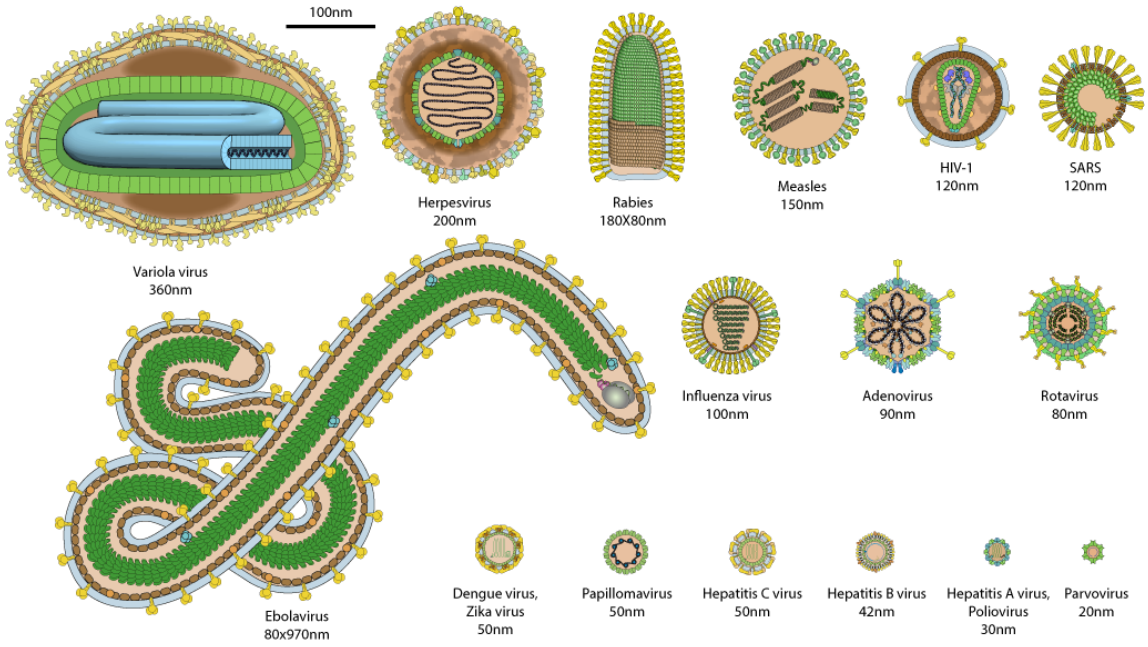


Figure 1.1: Relative Size of Hepatitis B Virus to Other Viruses

This picture originates from the ExPASy Viral Zone web-resource [25] and clearly demonstrates the relative size of Hepatitis B virus in relation to other well characterised viruses.

This circular genome is organised into four overlapping reading frames (ORFs), located on the minus strand of the DNA genome, as the plus strand of the genome DNA is roughly only two thirds of the full genome in length, and has a variable size on the 3' end [3, 8, 30]. These ORFs code for the seven viral proteins that are translated from five mRNA transcripts, of varying length (Figure 1.3) [9, 13, 26]. The pregenomic RNA (pgRNA), consists of an overlength mRNA (3.5 kb) with a polyA-tail and a 5' cap and

serves as the primer for the plus strand during DNA synthesis as well as the transcript for the polymerase and core protein [9, 13, 26, 31]. This overlength mRNA is often referred to as the major core gene transcript or just the major transcript [13]. A second overlength mRNA, called the pre-core/core mRNA, is translated into a precursor molecule, which is subsequently post-translationally modified into the enigmatic HBV antigen (HBeAg) [9, 13, 26, 31, 32]. Subgenomic mRNAs, which are shorter than overlength mRNAs, are the transcripts for the three surface proteins (2.4 kb and 2.1 kb) and the X protein (0.7 kb) [9, 13, 26]. Thus the seven viral proteins translated consist of the three surface proteins, the X protein, the Hepatitis B e antigen (HBeAg), HBV core antigen or nucleocapsid protein (HBcAg) and the polymerase enzyme (see Figures 1.2 and 1.3) [9, 26]. The gene encoding the polymerase (P) overlaps with all three of the other ORFs (covering ≈80% of the entire genome), with complete overlap of the entire preS1, preS2 and Surface (S) ORF, and partial overlap of the Core (C) and Hepatitis B X (HBx) ORFs [9, 26, 33].

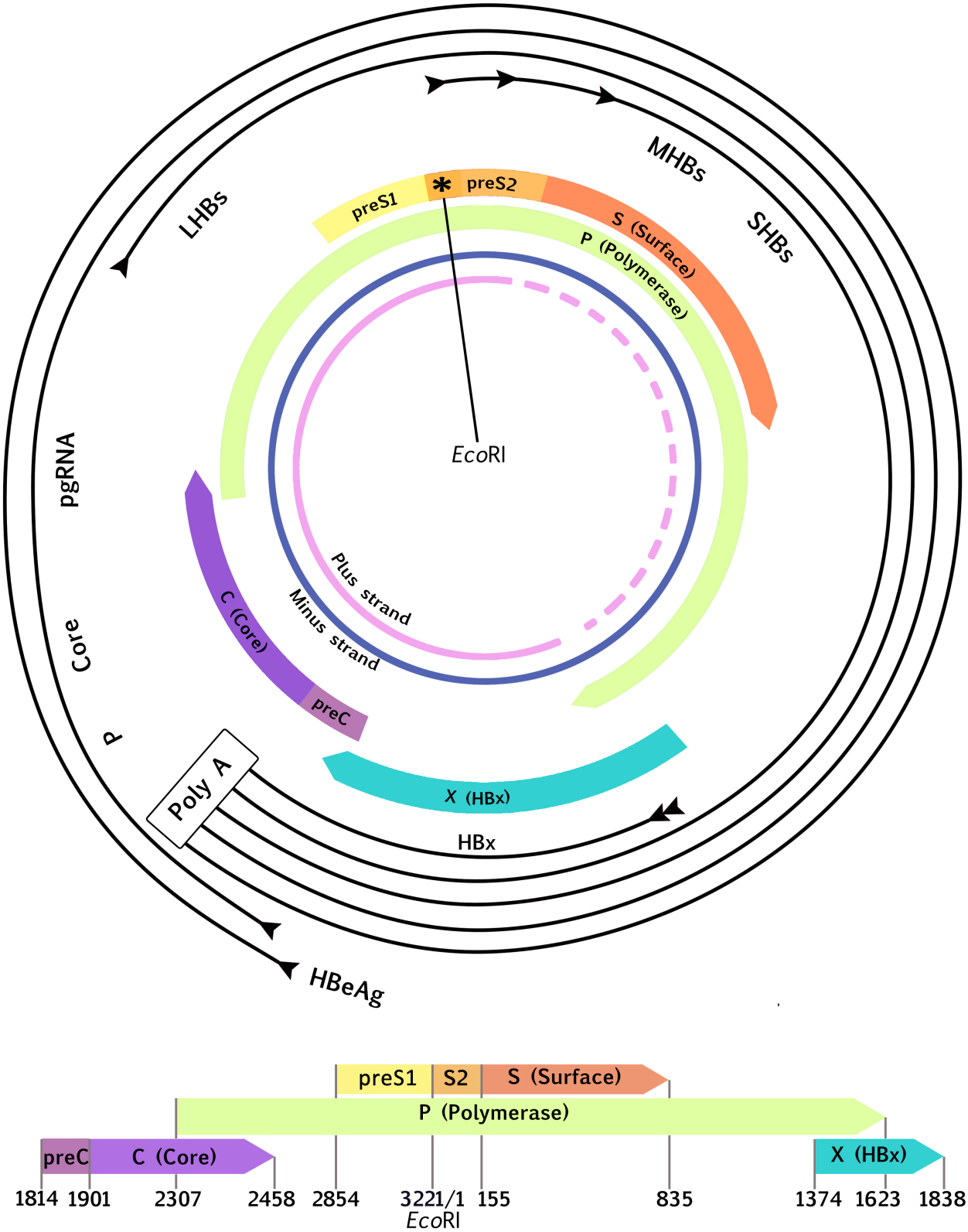


Figure 1.2: Representation of the HBV Genome, Transcriptome and Subsequent Viral Proteins

This diagram shows a circular (top) and the linear (bottom) representation of the HBV genome, its ORFs and the RNA transcripts. The various mRNA transcripts are from where HBV genes are translated, all of which are 'capped' with a poly A tail [13, 26]. The linear representation includes genome numbering from the *Eco*RI site; here the numbering is based on Genotype A. For a complete reference of numbering for all genotypes, please see Kramvis *et al.,* 2005 [34].

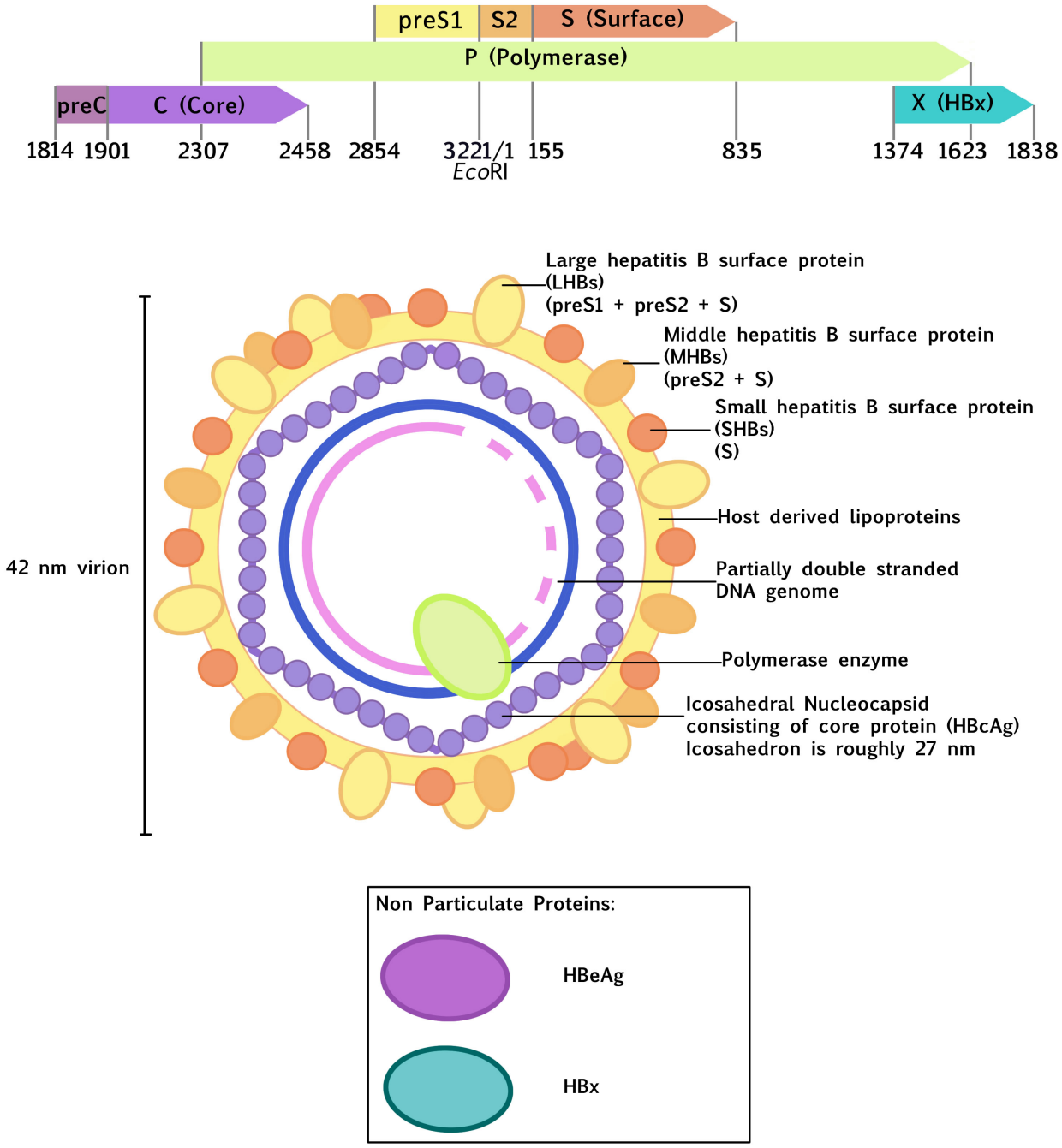


Figure 1.3: Representation of the Viral Proteins in the Mature Virion

The flat HBV genome layout with numbering relative to the *Eco*RI site is repeated in the top diagram and below is the virion structure with each protein represented in the matching colour.

1.1.3 HBV Viral Proteins: Their Translation and Functions

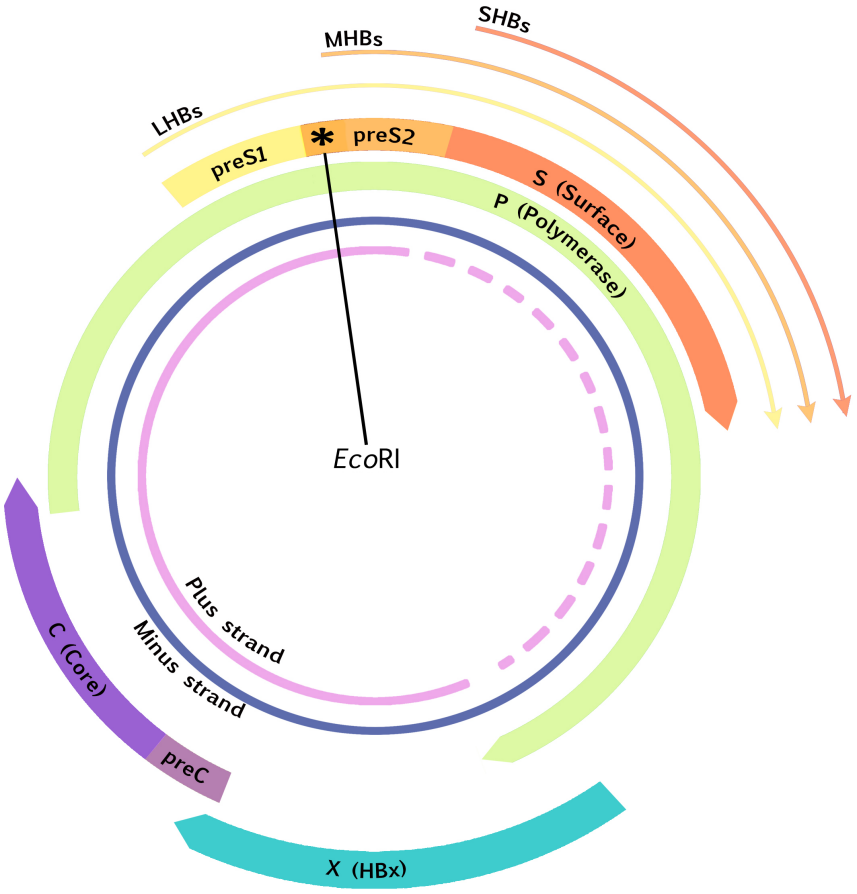
Surface Proteins

The overlapping genetic organisation of the HBV genome means it is possible for HBV to make use of alternative translation initiation and/or termination sites, allowing for the production of multiple proteins from one ORF (Figure 1.3). For example, the multiple initiation sites of preS1, preS2 and Surface genes permit the translation of three different surface proteins, the Large, Middle and Small hepatitis surface proteins (LHBs, MHBs and SHBs respectively) [2, 9, 26, 35, 36] (Table 1.4). Collectively the LHBs, MHBs and SHBs are known as the hepatitis B surface antigen (HBsAg) [18]. While the MHBs is dispensable, and not required for virion assembly, the SHBs and LHBs are essential to the HBV lifecycle [24, 37-39]. The immature virion acquires host-derived lipoproteins from either the endoplasmic reticulum (ER) or the Golgi apparatus via the secretory pathways during budding [3]. Thus, the budding process allows the viral particle to gain lipoprotein envelope encrusted with a mosaic of surface proteins of the virus [36]. The SHBs is also known as the major peptide because it is produced in such major excess compared to the MHBs and LHBs [35, 40]. A large excess of the HBsAg is also produced, which then form spherical (16-25 nm diameter) and filamentous (width of ~20 nm) particles, which are visible under the electron microscope [2, 17, 26, 36]. These non-infectious empty subviral particles are produced 1000:1 relative to the infectious Dane particle (up to 300 µg/ml) and accumulate in the blood [13, 17, 26, 36]. They display the T and B cell epitopes and thus contribute to antigenemia [13, 40]. This ratio between SHBs and LHBs is important for normal viral secretion, and overexpression of the LHBS may lead to retention of HBsAg in the ER [28, 41, 42].

The primary function of the LHBs is binding to the cellular entry receptor for HBV, the bile acid transporter designated sodium taurocholate cotransporting polypeptide (NTCP) [11, 40, 43]. While it was known that HBV binds to heparan sulfate for attachment to human cells, it was unknown which receptor HBV was binding in order to gain entry into hepatocytes (amongst other cells) up until 2012 [43, 44]. Now, however, it has been shown that a 75 aa-long domain on the N-terminus the LHBs, which is encoded by the PreS1, binds to NTCP with high specificity [3, 18, 43, 45, 46].

Table 1.4: Different Length Transcripts Result in Three Surface Proteins

Sub Genomic Transcripts for Surface Proteins

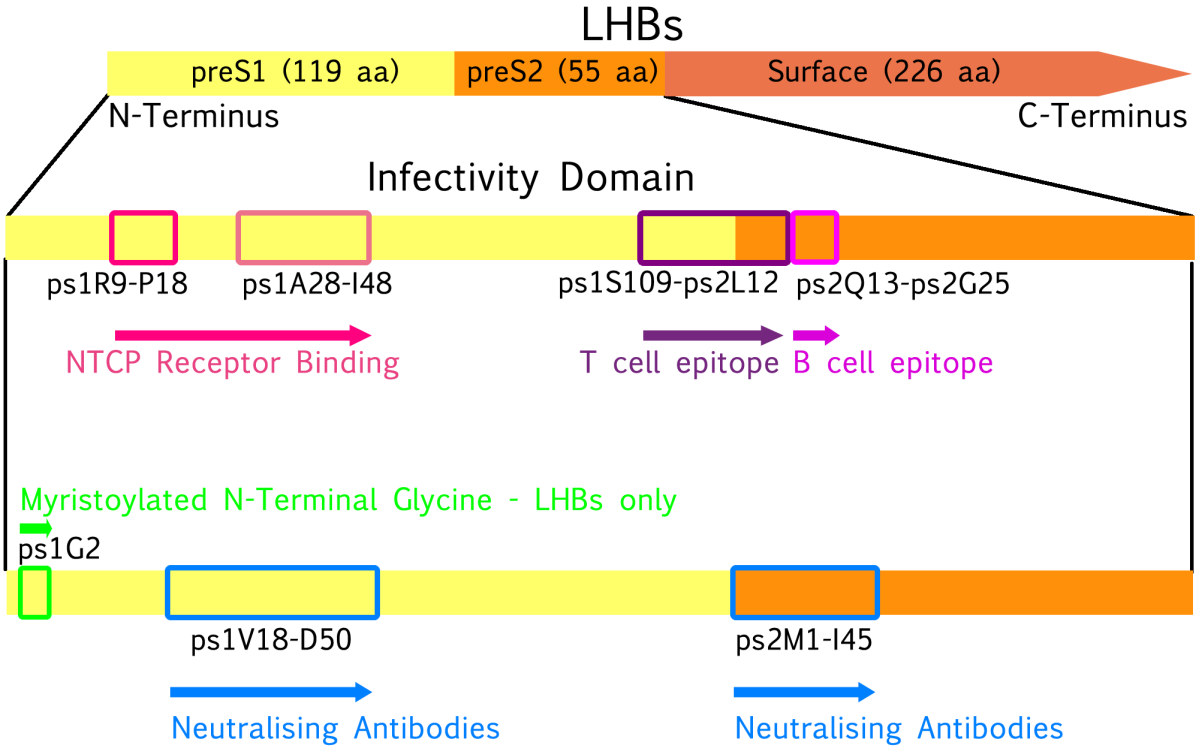


Surface Protein	mRNA transcripts	Protein Components	Size in Amino Acids (aa)*
LHBs	2.4 kb preS1 mRNA	preS1 + preS2 + S preS1 preS2 Surface	400
MHBs	2.1 kb preS2 mRNA	preS2 + S preS2 Surface	281
SHBs	2.1 kb preS2 mRNA with alternate initiation site	S Surface	226

Amino acid (aa) size is based on subgenotype A1; variations in the size of these proteins are dependent on genotype as well as whether deletions or insertions are present. More exact length of the surface proteins for each genotype is reviewed in Kramvis *et al.*, 2005 [26, 34, 35].

This is indicated as the NTCP receptor binding domain in Figure 1.4 below, along with other important functional domains. The LHBs is required for the encapsidation of core particles, a vital step for budding and secretion of viral particles [13]. It has also been shown that this region on the preS1 region has to have a myristoylated residue on the second glycine for viral entry to occur, and thus is essential for infectivity [13, 43, 47].

This domain is referred to as the key domain of preS1 and is also crucial in orienting the LHB towards the ER or the cytoplasm for budding and nucleocapsid envelopment [24]. As seen in Figure 1.4, the T cell epitope is found in the region overlapping the end of the preS1 and beginning of the preS2. Mutations in this region are often responsible for immune escape and viral breakthrough by reducing T cell reactivity [48]. It is thought that vaccine escape mutants may contribute to a rise in horizontal transmission of the virus [48].



No.	Functional Domain	Region	Size in aa	Amino Acid Sequence
1.	Myristylated N-Terminal Glycine	preS1	1	G
2.	NTCP Reception binding site	preS1	39	RKGMGTNLSVPNPLGFFPDHQLDPAFGANSNNPDWDFNPI
3.	T Cell Epitope	preS1/2	23	SPPLRDSHPQAMQWNSTAFHQAL
4.	B Cell Epitope	preS2	13	QDPRVRGLYFPAG
5.	Neutralising Antibody Target 1	preS1	33	VPNPLGFFPDHQLDPAFGANSNNPDWDFNPIKD
6.	Neutralising Antibody Target 2	preS2	45	MQWNSTAFHQALQDPRVRGLYFPAGGSSSGTLNPVPNTASHISPI

Figure 1.4: Important Functional Domains of the LHBs

The LHBs of the surface protein of HBV contains a few important functional domains, numbering here is from the beginning of the preS1 region for ease of reference. For genotype A, the preS1 region encodes 119 aas and the preS2 55 aas. Together these two regions make up what is known as the infectivity domain as it contains the areas required for hepatocyte entry (NTCP receptor). The NTCP receptor has an essential component from aa 9-18 highlighted in bright pink, as well an accessory domain from aa 28-48 highlighted here in a dull pink. Myristoylation of the N-terminal Glycine on aa 2 is essential for viral entry. Further highlighted in this diagram are immune related functional domains including both B and T cell epitopes as well as the areas to which neutralising antibodies are targeted [13]. The accompanying table outlines the positions and aa sequences of all the above functional domains, and lists aa numbering both relative to the start of the LHBs as well as to the preS1 and preS2 regions.

A region of the surface protein known as the 'a'-determinant within the antigenic loops (AGL), consists of three central hydrophilic loops, stretching from residues 124-147 relative to the EcoRI site [13, 49-51]. The exact three-dimensional conformation of the "a"-determinant's loops are as yet unknown, as it consists of 8 conserved cysteines connected by disulphide linkages, however the exact linkages have not been discovered [13]. It is however proposed that the first loop extends form residues 124-138 and the second loop from 139-148 [48]. Most of the immune systems neutralising antibodies are targeted at these loops [11, 13, 26, 35, 49, 51]. Thus, these epitopes on the immunodominant 'a'-determinant region of the surface proteins provide the major target that ultimately stimulates the production of anti-HBs antibodies and thus underlies the efficacy of the Hepatitis B virus vaccine [11, 13, 35, 51]. Finally, these same loops are also the target for detection of HBsAg during serological tests [11, 35]. It follows therefore, that mutations in this area result in detection and immune escape mutants, and importantly, affects the antigenicity of the virus [48, 52]. Mutations may either change the conformation of the loops' structure, or by changing the polarity of the aa side chains as a result of substitutions, the ability antibodies to bind the antigenic site with high avidity is diminished or even abolished [48, 53].

The Hepatitis B X Protein

The last of the viral proteins, HBx is encoded for by the X gene, the viruses' smallest gene [11]. Being a non-particulate or accessory protein, HBx has no role in the structure of the Dane particle [18]. With an anti-apoptotic domain on its C-terminus and a pro-apoptotic domain on the N-terminus, it functions as a transcriptional transactivator or a viral regulatory protein, pivotal in the transcription of cccDNA and influencing host as well as viral gene expression [9, 11, 18, 24]. It is thought to play a role in hepatocarcinogenesis because of its ability to modulate cellular functions that activate immune responses [11]. While its exact functions and mechanism of action remain poorly understood, we know that X gene mutations have frequently been reported when isolated from HCC cases [11].

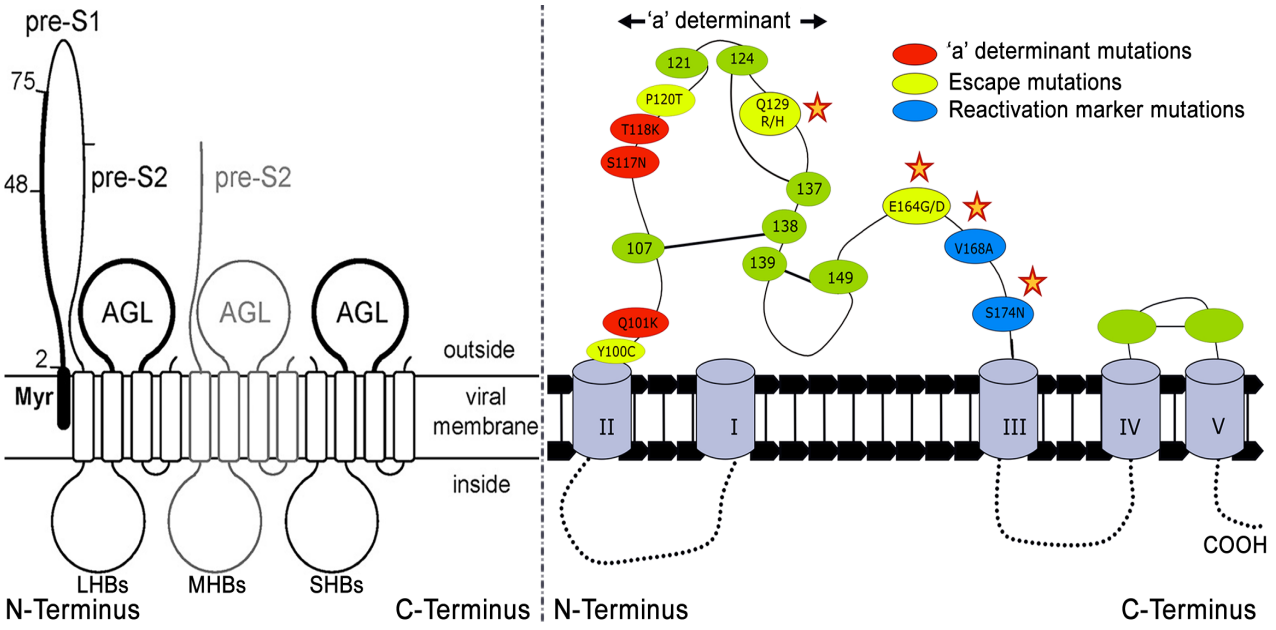


Figure 1.5: Graphic Representation of the 'a'-determinant Region

This is a graphical representation of the 'a'-determinant within its context in the surface protein as a whole. The as the exact structure is not known this diagram is hypothetical in nature, but serves to illustrate that mutations arising, particularly in the S gene portion, of the surface proteins can lead to various escape mutants arising [13]. This is discussed in more depth under the section on HBV mutants and S deletions. This figure was reproduced with permission from Copyright © 2009, American Society for Microbiology from publications: Le Duff et al., (2009) The Pre-S1 and Antigenic Loop Infectivity Determinants of the Hepatitis B Virus Envelope Proteins Are Functionally Independent, Journal of Virology 83(23): 12443–12451 (http://jvi.asm.org/content/83/23/12443.full.pdf+html) and Copyright © 2012, PLOS One Makondo et al., (2012), Genotyping and Molecular Characterisation of Hepatitis B Virus from Human Immunodeficiency Virus-Infected Individuals in Southern Africa, PLOS One 7(9):e46345 (http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0046345#pone-0046345-g003)

The Enigmatic HBeAg and Core Proteins

The pre-Core, Core gene consists of the pre-Core region of 29 amino acid (aa) codons and the core section of 181 aa codons [54]. Alternate initiation and termination sites, together with post-translational modification, allows the pre-Core (pre-C) and Core (C) genes to result in the expression of two proteins, HBeAg and HBcAg [54, 55]. These two proteins are translated from separate mRNAS, the pgRNA (the major gene transcript) and the pre-Core mRNA, respectively (Table 1.5) [54, 55].

The pre-Core mRNA stretches from upstream of pre-Core ATG initiation (1814 from the *Eco*RI site) at its 5' end to the all the way to the Poly A tail, differing from the Core gene by only the 29 aa codons on the N terminal end [56]. However, a termination site at the end of the core genes 3' end (2458 from *Eco*RI) means the immature molecule expresses a longer precursor than the pre-core/core gene [54]. This translated precursor, which is subsequently post-translationally cleaved at position 19 on the amino end and

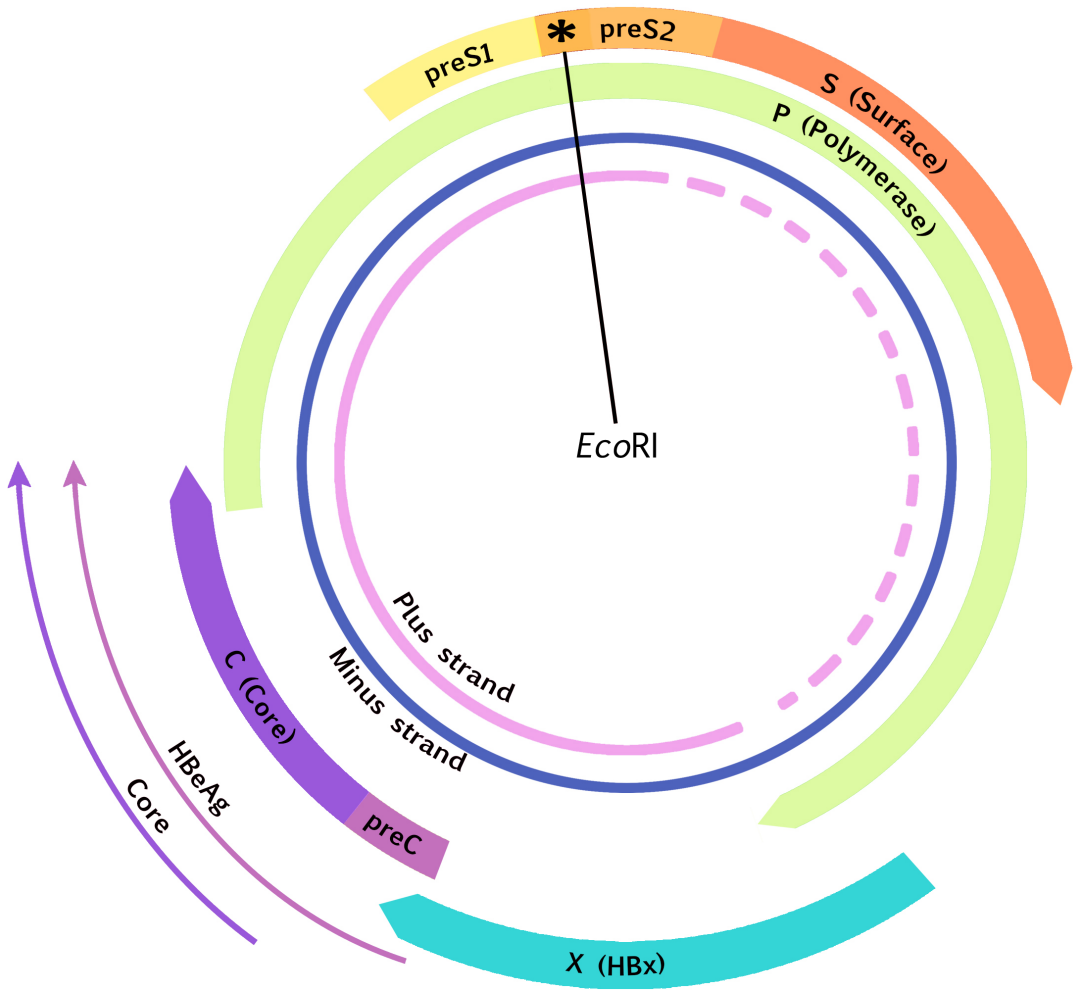
a variable site on the carboxyl end, results in the mature HBeAg protein [54]. After the two proteolytic cleavage events, the mature HBeAg differs from the core protein by ten additional residues on the amino or N-terminal end, as well as the lack of a DNA-binding sequence on the carboxyl or C-terminal end [26, 54]. While the exact role of the HBeAg in the viral life cycle is unclear, it appears to have conflicting roles in regulating the immune response. It seems to function as an immunogen or target for inflammatory responses as well as a viral decoy, or even as a tolerogen; summarily the protein either encourages viral persistence or delays the antiviral immune response in some way [13, 18, 26, 31, 53, 54, 56-59]. Importantly, as it is an extracellular or accessory protein, HBeAg is often used as a marker for active replication of HBV and infectivity [9, 56]. This is because high viral loads correlate strongly with high HBeAg expression (and identify individuals that pose a higher risk of transmitting the virus), and conversely as anti-hepatitis B e antibodies (anti-HBe) appear, viremia is significantly reduced [56, 60]. Furthermore, while HBeAg is not essential for viral replication or viral assembly, it is often observed that with strong immune pressure from anti-HBe over time, expression may be reduced or lost altogether [26, 53, 54, 56, 60, 61].

On the other hand, the 181 aa core protein is translated from the pgRNA gene transcript, starting at the core ATG initiation (position 1901 from the *Eco*RI site) and ending at the gene termination site (2458 from *Eco*RI) [54, 55]. It should be noted that in some HBV genotypes, including subgenotype A1, the core protein is 183 aa long. This is due to a 6-nt insertion and is the reason for which the subgenotype A1 genome is 3221 bp in length [54, 55]. This core protein is composed of an assembly domain as well as a nucleic acid binding domain, from aa's 1-149 and 150-183 respectively [55]. The translated core protein is essential for the formation of the nucleocapsid [3, 54, 55]. Genetic economy dictates that this core protein is repeatedly used to form the protective icosahedral protein shell or packaging around the viral genome and polymerase that is seen in Dane particles [3, 4, 62, 63]. Thus, the 21 kDa viral core protein forms homodimers, which in turn self-assemble into the icosahedral capsids [62]. The icosahedral symmetry allows for simultaneous stability of the virion and maximal internal volume [4]. The HBV virion's capsid is icosahedral instead of helical in shape because the core proteins are not symmetrical [4]. The core proteins are not always the same size, resulting in different sized capsids, ranging from 25nm (small) to 28-30 nm (large). The smaller capsids are said to have a T3 icosahedral symmetry and consist of

180 copies of the core protein, whereas the larger capsids, resulting in T4 symmetry, contains 240 copies [13, 55, 62, 63]. The viability of the T3 virions have however not been established [13]. The core proteins will specifically encapsidate the pgRNA as well as the viral Polymerase [55].

Table 1.5: Different Length Transcripts of the pre-Core and Core Genes

Sub Genomic Transcripts for Core and HBeAg Proteins



Protein	mRNA transcripts	Protein Components	Size in aa
Core	pgRNA	Core Protein Core	181
HBeAg	pre-Core mRNA	Immature HBeAg preCore Core	210

The Polymerase Enzyme

Another manner in which *Hepadnaviridae*, and particularly HBV, is unusual is that typically smaller viruses do not encode their own polymerase enzymes but use host polymerases to replicate their genetic material. Furthermore, the HBV polymerase is atypical in that; it is a single enzyme with error-prone RNA-dependent DNA polymerase activity as well as reverse transcriptase capability (Figure 1.4). Thus, the replication and life cycle of HBV is similar to retroviruses in animals and pararetroviruses in plants [3, 4, 6, 9, 18].

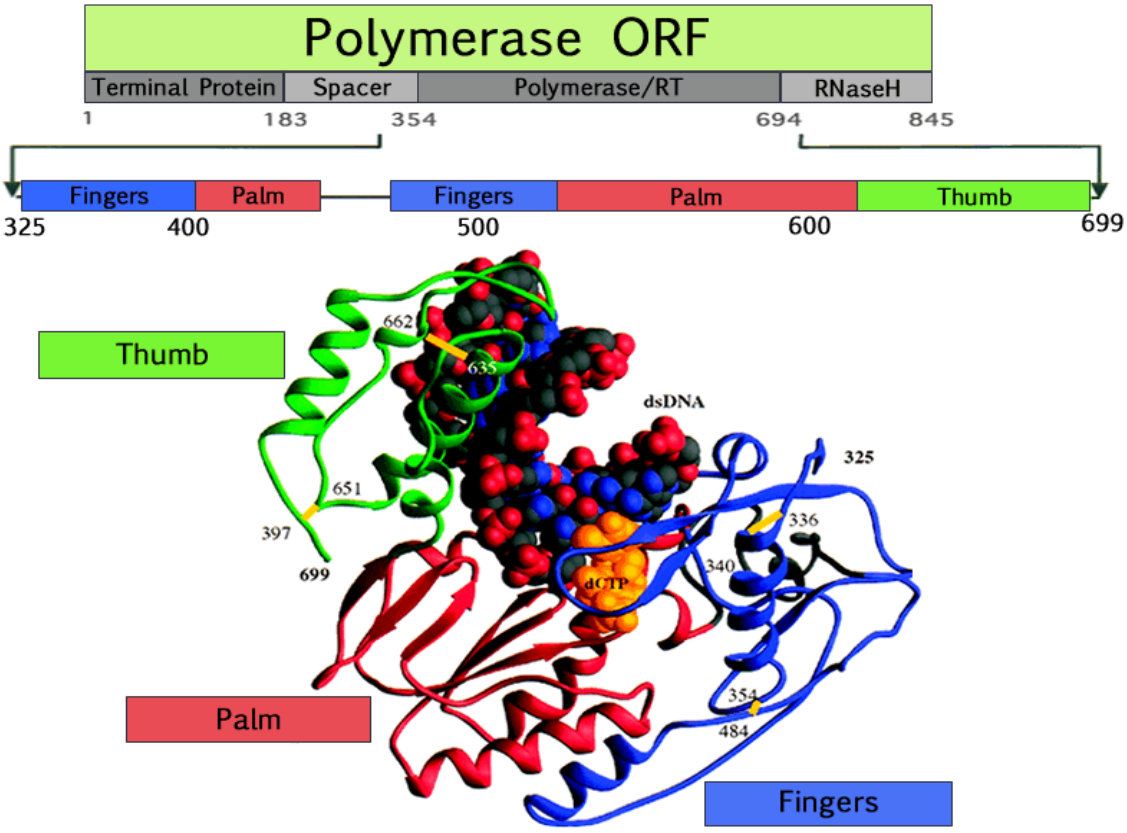


Figure 1.6: Structure of the HBV Polymerase

This image is adapted from the study by Das *et al.,* 2001 where they carried out crystal structure analysis to achieve this schematic of the enzyme [64]. Copyright was obtained from the journal. The HBV endogenous polymerase is comprised of three functional domains (also known as catalytic subdomains) [3, 13, 18]:

Polymerase Terminal Protein – a domain located at the N-terminal end of the enzyme; this acts a primer for the synthesis of the negative DNA strand.

A Spacer Region – section in the middle of the enzyme that separates the C and N termini, and allows for movement of the enzyme.

The Polymerase/Reverse Transcriptase (RT) and RNase H Domain – the RT portion is placed after the spacer towards the C-terminal end of the enzyme; this section allows the reverse transcription of mRNA into negative strand DNA. Finally the RNase H section - immediately adjacent to the RT portion, on the extreme end of the C-terminal part of the enzyme; RNase H is a ribonuclease that hydrolytically cleaves RNA allowing the removal of the RNA primer and completion of the nascent DNA synthesis. As indicated in the image the double-stranded DNA (dsDNA) fits between the 'thumb' and the 'fingers' of the RT portion of the enzyme.

1.1.4 The Replication Cycle of HBV

Despite a good understanding of the endogenous polymerase, the minutia of the HBV viral replication cycle has not yet been fully elucidated [3]. This is partly due to the fact that the cellular receptor for HBV, NTCP, was only recently discovered [18, 43]. It is also partly because finding cell lines that are permissive for HBV infection or suitable animal models has been difficult [3, 18]. Figure 1.5 and Table 1.6 below explores the major steps of the HBV replication cycle.

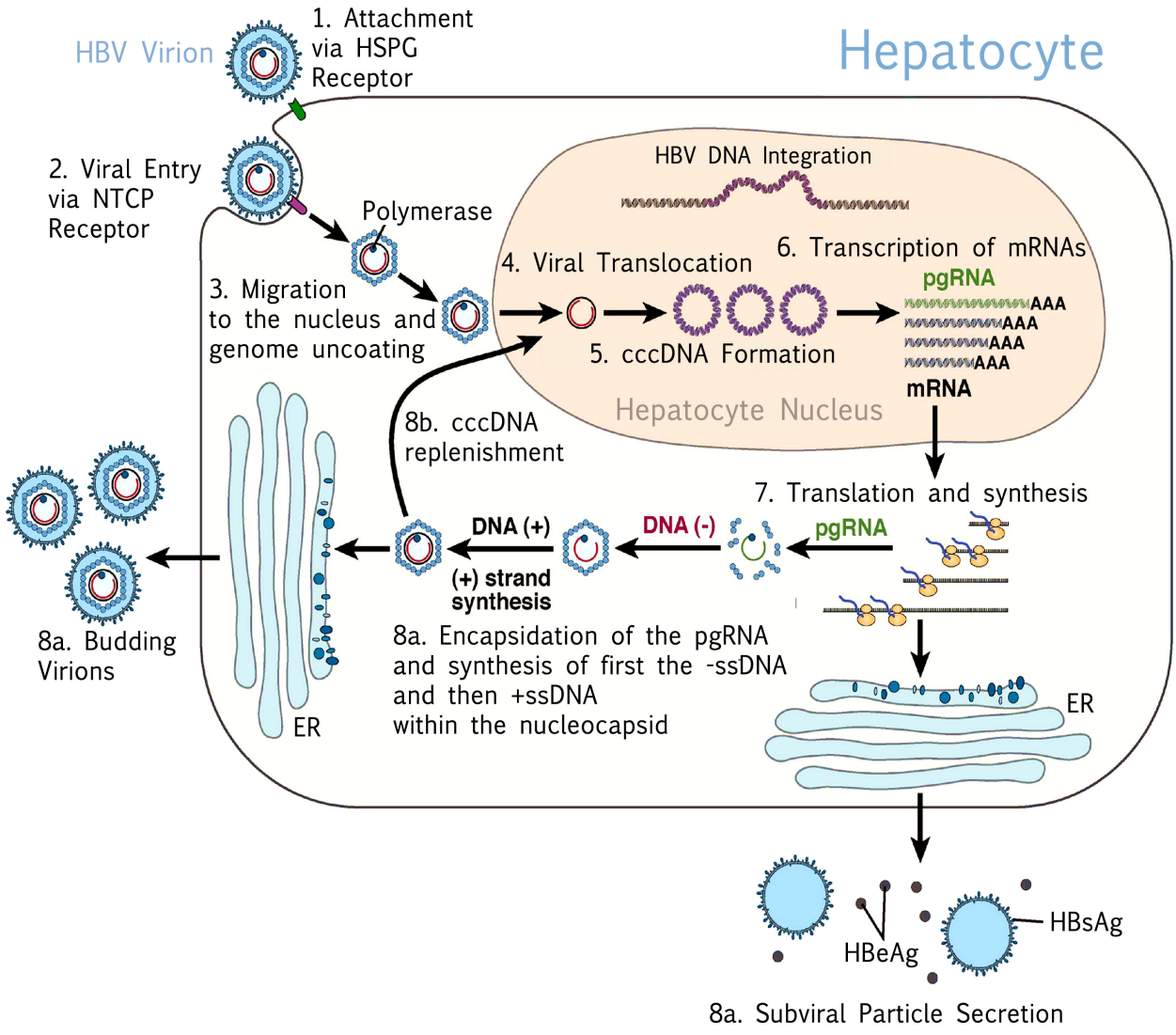


Figure 1.7: Stages of the Replication Cycle

The numbers on the diagram co-ordinate with the steps of the replication cycle as indicated in Table 1.6 below. Reprinted from Gastroenterology, 137, F. Zoulim and S. Locarnini, Hepatitis B Virus to Nucleos(t)ide Analogues, 1593-1608, Copyright 2009, with permission from Elsevier [65].

Table 1.6: Stages of the HBV Replication Cycle

<table>
<tr><th colspan="4">Life Cycle of HBV in Hepatocytes</th></tr>
<tr><td colspan="4">1. HBV binds to heparan sulfate proteoglycans (HSPG) to achieve viral attachment.</td></tr>
<tr><td colspan="4">2. The key domain of the preS1 (on the LHBs) binds to the NTCP to gain entry into the cell via endocytosis. It has been shown that the first 75 N-terminal amino acids of the preS1 domain, as well as an attachment on the on the N-terminal end of a myristic acid at the 2nd glycine of preS1, are essential for viral entry to occur. However, the endosomal compartment required for endocytosis has not yet been characterised.</td></tr>
<tr><td colspan="4">3. The Dane particle is transported/migrated to the nucleus. It is assumed the nucleocapsid is disassembled or removed, uncoating the genome, at the nuclear pore, but this has not been clearly shown.</td></tr>
<tr><td colspan="4">4. The viral genome is discharged into the nucleus, referred to as translocation of the viral genome.</td></tr>
<tr><td colspan="4">5. The relaxed circular viral DNA (rcDNA) genome is converted into covalently closed circular DNA (cccDNA), which is essential to the replication cycle as it is not possible to transcribe mRNAs from the gapped rcDNA. All viruses must produce mRNA to replicate. The mechanism of this conversion and whether host or viral enzymes carry out this process is not well understood, however it is probable that three things must occur:
• A polymerase (cellular?) or reverse transcriptase (viral?) is required to close the gap in the plus strand of the rcDNA
• Endonucleases should remove the RNA primer on both minus and plus strands of DNA as well as remove the short terminally redundant sequence on the rcDNA minus strand that was created during DNA synthesis
• Ligation of both strands to form circular strands is required
Finally, the cccDNA is complexed with cellular histones forming a minichromosome, also referred to as a minicistron, which structurally resembles a plasmid or episomal DNA construct. This structure serves as the template for transcription of all the viral genes. Because this minichromosome is also similar to the host chromosome in structure, it is possible for epigenetic modification to occur.
At some stage during this step, viral DNA may be incorporated into the host genome as well.</td></tr>
<tr><td colspan="4">6. Transcription of the pgRNA and subgenomic mRNAs is carried out by cellular RNA polymerase II (the RNA intermediate step of the replication cycle), followed by the transcripts being translocated from the nucleus to the cytosol.</td></tr>
<tr><td colspan="4">7. Translation and synthesis of the viral proteins occur at the ER; thus, the core particles are ready for assembly and encapsidation.</td></tr>
<tr><td>Afferent Arm</td><td>8a. Some pgRNA is encapsidated together with the newly synthesised viral polymerase, in immature virions. The viral particle matures as the viral reverse transcriptase converts the pgRNA into -ssDNA inside the nucleocapsid. The +ssDNA strand is then synthesised using the –ssDNA as a template, and the rcDNA is restored. Thus, the encapsidated pgRNA regenerates the partially dsDNA in budding virions after reverse transcription. The RNase H activity of the polymerase gradually degrades the pgRNA. As the genome matures, the virion buds via the ER, the Golgi apparatus and finally the cellular membrane (as a part of multivesicular particles). As this mature virion buds, it gains surface proteins and the host-derived portion of the envelope along the way. The subviral particles are similarly secreted but via a constitutively active secretory pathway.</td><td>Efferent Arm</td><td>8b. Some pgRNA will be encapsidated, but then, some mature genomes and core particles (from step 8a.) are transported back to the nucleus, and these are used to replenish the cccDNA pool to continue driving the replication cycle.</td></tr>
</table>

This table was compiled using several resources [3, 4, 7, 9, 11, 13, 17, 18, 24, 26, 43, 47, 55, 66] to best explain the replication cycle of HBV.

With the nature of the replication mechanism and the lack of proofreading ability of the viral polymerase sequence heterogeneity and quasispecies as well as variants and mutations of the virus arise [9, 11, 13, 56, 65]. These will be discussed in detail later, but it is important to note that these altered viral genomes arise naturally and frequently, and may or may not have functional consequences for the virus regarding fitness, replication rates, pathogenicity and immune/drug evasion and/or susceptibility.

1.1.5 HBV Transmission in an African Setting

Being a blood-borne virus, transmission of HBV is parenteral, meaning percutaneous or mucosal exposure to infected blood or body fluids can result in infection [9, 27, 67]. Body fluids that can transmit HBV include blood, mucous and saliva, as well as menstrual, vaginal and seminal fluids [1, 27]. Further, transmission can either be vertical or horizontal (Table 1.7).

Table 1.7: HBV Routes of Transmission

Routes of Transmission	
Vertical Transmission	Horizontal Transmission
Vertical transmission of HBV includes: Perinatal: HBV being passed from mother to child during birth or in utero (very rarely) Nonsexual transmission: Long-term household exposure	Horizontal transmission includes: Iatrogenic or healthcare-related exposure: Blood transfusions, needle stick injuries, kidney dialysis, organ transplantation and dental procedures Body fluid transmission: Sexual transmission, sharing of contaminated needles (as is common with intravenous drug users), contact with a bleeding wound, tattooing, scarification, acupuncture, traditional circumcision, piercing, sharing of toothbrushes or razors, exposure to contaminated saliva and rough play amongst children, bites and scratches Indirect transmission: Via contaminated objects or surfaces

Table compiled from resources [2, 9, 26, 27, 67-72].

In areas where HBV is endemic, the vast majority of persons infected with HBV acquired it either at birth or during their early childhood, which is in contrast to low endemicity areas where risky sexual behaviour and intravenous drug use are the principal sources of infection [1, 26, 61, 70, 73, 74]. Mothers with a high viral load, and thus a high

HBeAg seropositivity, will transmit HBV to their offspring 90% of the time [61, 69, 75]. In Africa however, the low HBeAg seropositivity amongst pregnant women results in horizontal transmission in early childhood being more common than vertical transmission [26, 70, 74, 76-78]. Unfortunately, this early stage infection most often results in the individual becoming a chronic carrier of HBV, which increases the risk of serious clinical sequelae [67, 73, 74]. For a long time, it was thought that there should be a zoonotic reservoir to explain high rates of horizontal childhood infection in Africa [1, 79-81]. Evidence exists that bats, rodents, a few primates and some species of birds do carry hepadnaviruses (orthohepadnaviruses and avihepadnaviruses) that are closely related to HBV. Nevertheless, the very narrow host range of these viruses and their limited ability to infect human hepatocytes means they may pose minimal threat [4, 15, 19]. While tropism studies have shown that bats can be infected with human HBV and that BtHBV has the potential to infect human hepatocytes, it remains unclear to what extent bats spread HBV in human populations [15, 43, 82]. By the same token, there is as yet no convincing evidence to show there is an insect or parasitic vector carrying HBV. Although many blood-sucking insects test positive for HBsAg and HBeAg, viral replication within these arthropods and actual transmission has not been clearly demonstrated [79-81]. Thus, maybe traditional practices such as scarification, traditional circumcision, close contact and rough play as well as most certainly the substandard clinical practices in resource-limited facilities are a more likely source of higher transmission rates in the African setting [67, 76, 83]. That said, the failure of maternal health, child health and vaccination programmes is probably the biggest contributors to ongoing high transmission rates [84]. Unsafe injections have been shown in particular to contribute to high rates of HBV transmission in developing countries [72]. It is interesting to note that HBV is much more infectious than Human Immunodeficiency Virus (HIV) [72, 85-88]. The minimum infectious dose for HBV is as little as 4-10 virions per millilitre whereas HIV requires exposure to body fluids containing more than 400 virions per millilitre to establish infection [26, 67, 72, 86-90]. One microlitre of HBV-infected blood can contain as many as 10^6 to 10^9 virions, and these HBV virions are stable in dried blood for up to one week [67, 87, 90]. Additionally, body fluids other than blood and serum can serve as reservoirs for transmission; fluids such as semen and saliva have also been shown to be infectious [67].

1.1.6 Epidemiology of HBV – Worldwide and Southern Africa

The worldwide epidemiology and prevalence are summarised in Figure 1.6. Two billion people show evidence in serology of either past or present infection with HBV, that is nearly a third of the world's population [1]. Prevalence rates, measured by the seroprevalence of HBsAg in serum, vary extensively across the globe as well as in different age categories. Rates are lower than 2% in Central Latin America, North America and Western Europe, and are highest (8% and above) in the North Western part of Africa [1, 78]. HBV infection is considered chronic if HBsAg has been present for six months or more. It is often accompanied by the presence of Ground Glass Hepatocytes (GGH) in the liver; and chronic HBV in infection (CHB) is associated with a 100-fold increased relative risk of developing HCC [1, 65, 91-93].

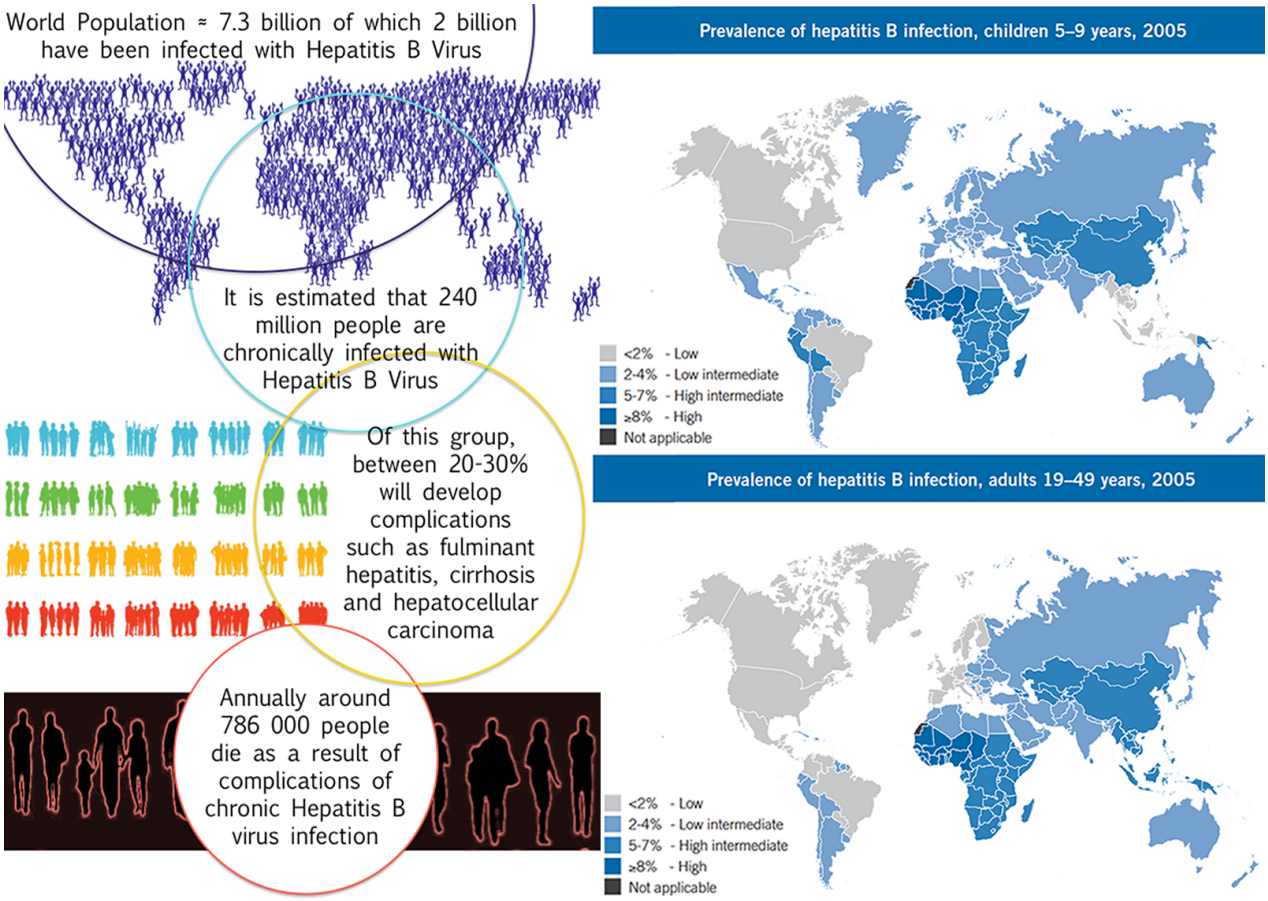


Figure 1.8: Facts and Figures of the Global HBV Epidemic

On the left, this infographic was constructed using the statistics reported in a recent World Health Organization Report, and the prevalence images on the right have been presented from the same report: *Guidelines for the prevention, care and treatment of persons with chronic hepatitis B infection* 2015 [94] as well as a report compiled by Lozano *et al.,* summarising recent global and regional mortality patterns [95]. The infographic depicts the major epidemiological facts of HBV concerning burden of disease and morbidity and mortality on a global scale.

Underreporting, inefficient data collection resources and general lack of health care infrastructure are huge barriers to obtaining accurate prevalence data in Africa [22, 95]. However, a paper published in 2007 considered data reported in 111 studies, and concluded that 65 million chronically infected individuals reside in Africa, which translates to 18% of the global burden of HBV disease [96]. Moreover, it is estimated that HBV is the cause of roughly 250 000 deaths per annum in Africa, and we know that these deaths result from liver cancer and cirrhosis decades after initial infection [22, 96]. By measuring population antibodies against hepatitis B core antigen (anti-HBcAg), one can show the exposure rate to HBV. In Africa, this ranges from 1.8-98% [76, 96]. Here the lower prevalence is associated with blood donors and children, whereas the high exposure rates are associated with patients who have chronic liver disease or hepatocellular carcinoma (HCC) [76, 96].

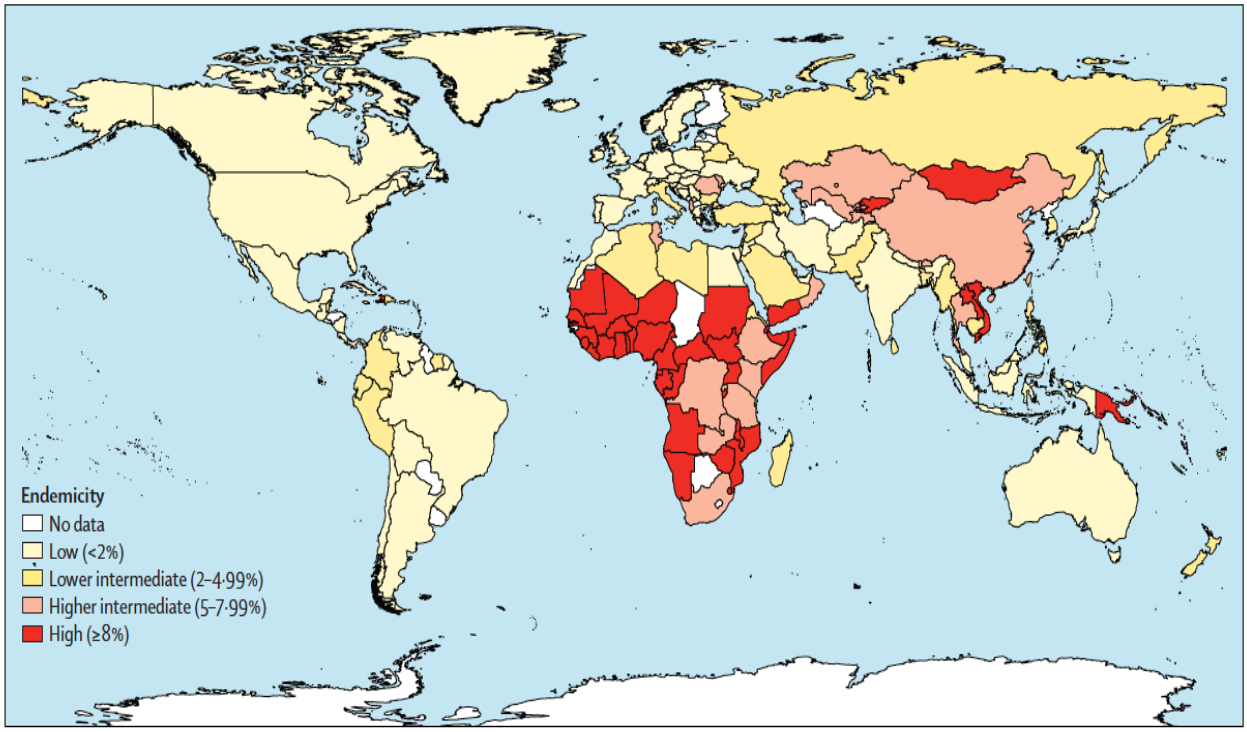


Figure 1.9: Global HBsAg Endemicity (1957-2013)

This figure clearly shows that Africa has the highest endemicity of HBV drawn from a systematic review compiled from studies spanning 56 years. It can be argued that Africa is now the only hyperendemic region [78]. Reprinted from the Lancet, 386, Schweitzer *et al.*, Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013, 1546–55. Copyright (2015), with permission from Elsevier.

In sub-Saharan Africa, HBsAg seroprevalence rates suggest that the virus is hyperendemic in this region, as 8% of the population or more test positive in the serum [96]. There are pockets of sub-Saharan Africa however, that have intermediate endemicity

of 2-8%, including Cote d'Ivoire, Kenya, Liberia, Senegal, Sierra Leone and Zambia [96]. In most countries in the world, HBsAg carrier status is higher in men than in women, and sub-Saharan Africa is no exception to this pattern [97]. Older studies showed that while the prevalence of HBsAg carriers ranges from 9-20% in sub-Saharan Africa, in South Africa the rate is around 9.6%, so HBV is also considered to be hyperendemic in this country [97, 98]. There is a marked increase in prevalence in rural areas as compared to urban areas; rural inhabitants are twice as likely to be HBsAg positive as their urban counterparts [70, 96]. Thus, roughly 10-15% of rural black children were carriers, as opposed to the 1% of urban children who were carriers [70, 77, 97-101]. In short, like so many other communicable diseases, HBV hits lower-income and more resource-limited settings harder than their socioeconomically more privileged counterparts [1, 70, 99].

1.1.7 Phylogeny and the Genotypes and Subgenotypes of HBV

Phylogenetic analysis and its ability to explain evolutionary theory is a relatively new science. As Baldauf states, phylogenetics is the science of estimating the evolutionary past based on the comparison of DNA or protein sequences [102]. Modern phylogeny thus depends on the numerical calculation of trees using quantitative methods [102-106]. Phylogeny represents a potent tool for making previously overwhelming data sets comprehensible and allowing the interpretation of trends and relationships [102].

The system of genotype and subgenotype designation refers to the relative genetic relatedness of HBV genome sequences within the family *Hepadnaviridae* or different genera of the family [33]. The description of genotypes and subgenotypes relies on sequence alignment, sequence divergence calculations and phylogenetic comparison [6]. Considering HBV has such a compact genome with overlapping ORFs, this limits the amount of mutation possible without affecting viral viability. Genotypes are identified where there is 7.5% or more intergenomic sequence divergence between strains, and subgenotypes where there is more than 4% intergenomic sequence divergence [33, 71, 107, 108]. Phylogenetic analysis of sequences from all over the world have shown that there are none confirmed genotypes, designated A-I, and one putative genotype 'J', as well as 35 or more associated subgenotypes which are linked with all the genotypes except for E and G [6, 11, 35, 71, 108, 109]. The distinct geographical distribution of

these genotypes (Figure 1.8), and in some cases of subgenotypes, has enabled epidemiological and transmission-related studies, furthered understanding of the differing prognosis and clinical manifestation of different genotypes, as well as varying response to treatment and has contributed the knowledge around the age of HBV infection in humans as well as tracing human migrations [10, 11, 35, 71, 76, 108, 110, 111]. Furthermore, the study of the geographical distribution of HBV genotypes and subgenotypes has cast some light on the long-term evolutionary history, viral mutation rate and to some extent the origin of the virus [10, 112]. While it appears that the origin of the virus is ancient (potentially as long as 80 million years), the earliest known human HBV sequence appears to be between 3000 and 100 000 years old [10, 110]. Evidently there are gaps in our knowledge, and while the study of genotypes and their evolution has aided in this regard, we still have no consensus on a viral mutation rate and many unanswered questions as to the origin and actual age of HBV [10, 112-114].

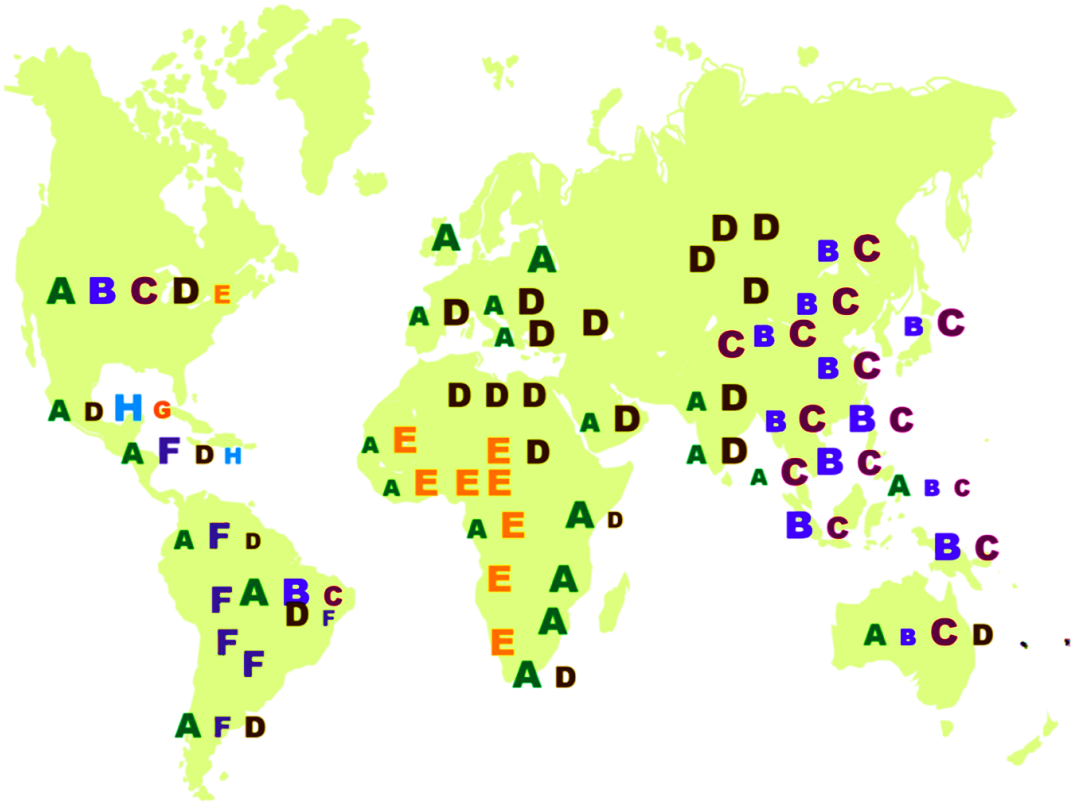


Figure 1.10: Overview of Geographical Distribution of HBV Genotypes

The genotypes of HBV have a distinct geographic distribution [108]. The most predominant strains in southern Africa are subgenotype A1 (75%) and genotype D (25%) [108]. Picture used here with permission from A. Kramvis (from publication title The Clinical Implications of Hepatitis B Virus Genotypes in Paediatrics currently submitted to Reviews of Medical Virology).

From a more clinical perspective, it has been established that the genotypes and some subgenotypes behave differently in terms of clinical manifestation of infectivity, susceptibility to antiviral therapy, prognosis, viral antigen expression, relative likelihood of developing resistance or escape mutants and relative risk of developing advanced liver disease [33, 35, 76, 108, 115-120]. Of course, this may all be complicated in cases of infection with recombinants strains or superinfection with more than one genotype [6, 76].

A focus on Genotype A: Prevalent in Africa

Nucleotide divergence shows us that the oldest endemic genotype on the African continent is genotype A, and currently still dominates the southern, eastern and central parts of Africa [76]. In northern Africa, genotype D prevails and genotype E is most prevalent in the central and western parts of Africa [6, 76, 108]. The subgenotypes circulating in South Africa are predominantly A1 and D3; however, subgenotype A1 remains the most widespread at 75% as well as the most studied in South Africa [76]. This is largely because of the severe clinical outcomes associated with this subgenotype. Genotype A is often coupled with a higher viral load, which results in increased infectivity and transmission [35]. The characterisation that has been carried out on subgenotype A1 has revealed that it has five very distinct features compared to subgenotype A2, which is the subgenotype that predominates outside Africa. Namely, subgenotype A1 is associated with low viral loads, early HBeAg seroconversion, high transmissibility, which is mostly horizontal and greater hepatocarcinogenic potential [86, 96, 121-123]. Of these possibly the most important feature is that subgenotype A1 is associated with an increased risk of developing HCC in Black Africans, which means progression to HCC is accelerated and there may be a different mechanism by which HCC develops in the African setting [26, 115, 122, 124, 125].

1.1.8 HBV and HIV Co-Infection: A Dangerous Liaison

Co-infection occurs when a single individual is concurrently infected with both HBV and HIV [1]. These two blood-borne viruses share many common routes of transmission as well as sharing similar risk factors. Add to this the reality that Africa has both the highest HIV prevalence per capita and high HBV endemicity, it seems evident that high rates of co-infection will occur in the African setting [76, 84, 97, 126, 127]. Furthermore, while we know there is a significant proportion of co-infection in Africa, it is also likely that the number of co-infected individuals is underestimated due to the dependence on serological-based screening [84, 97, 128-131]. Again poor systematic surveillance, underreporting and limited resources mean we do not have accurate statistics; however from examining available data a disturbing pattern emerges [78].

Of the 34 million HIV-infected individuals worldwide, 5-15% are also chronically infected with HBV, (Figure 1.9) [1, 129, 130, 132]. A recent report from the WHO on the state of Viral Hepatitis in Africa estimates that 5-20% of HIV-positive individuals are also chronically infected with HBV, meaning that of the 24.7 million people living with HIV in Africa, between 1.2 and 4.9 million are also HBV-positive [22, 127, 129, 132]. Small cohort studies carried out all over the continent, but mostly in South Africa, have reported anything from a 1.6% to 83.6% co-infection rate [97, 100, 127, 130, 133-141]. A big multinational cohort study, which included African countries, reported a 5.5% co-infection prevalence [132]. Interestingly, the most common genotype amongst co-infected Africans and southern Africans was genotype A and then D, which is reflective of the predominant genotypes especially in South Africa [127, 132, 142]. The great variation of this co-infection prevalence can be attributed to the fact that these statistics were not drawn from population-based studies, and depended heavily on the subpopulation that was chosen for the study, and the sample size of the study. It is also worth noting that data on complete genome sequences in South Africa (in a predominantly subgenotype A1 background) are severely limited, as only a handful of studies have been conducted to date [52, 143, 144].

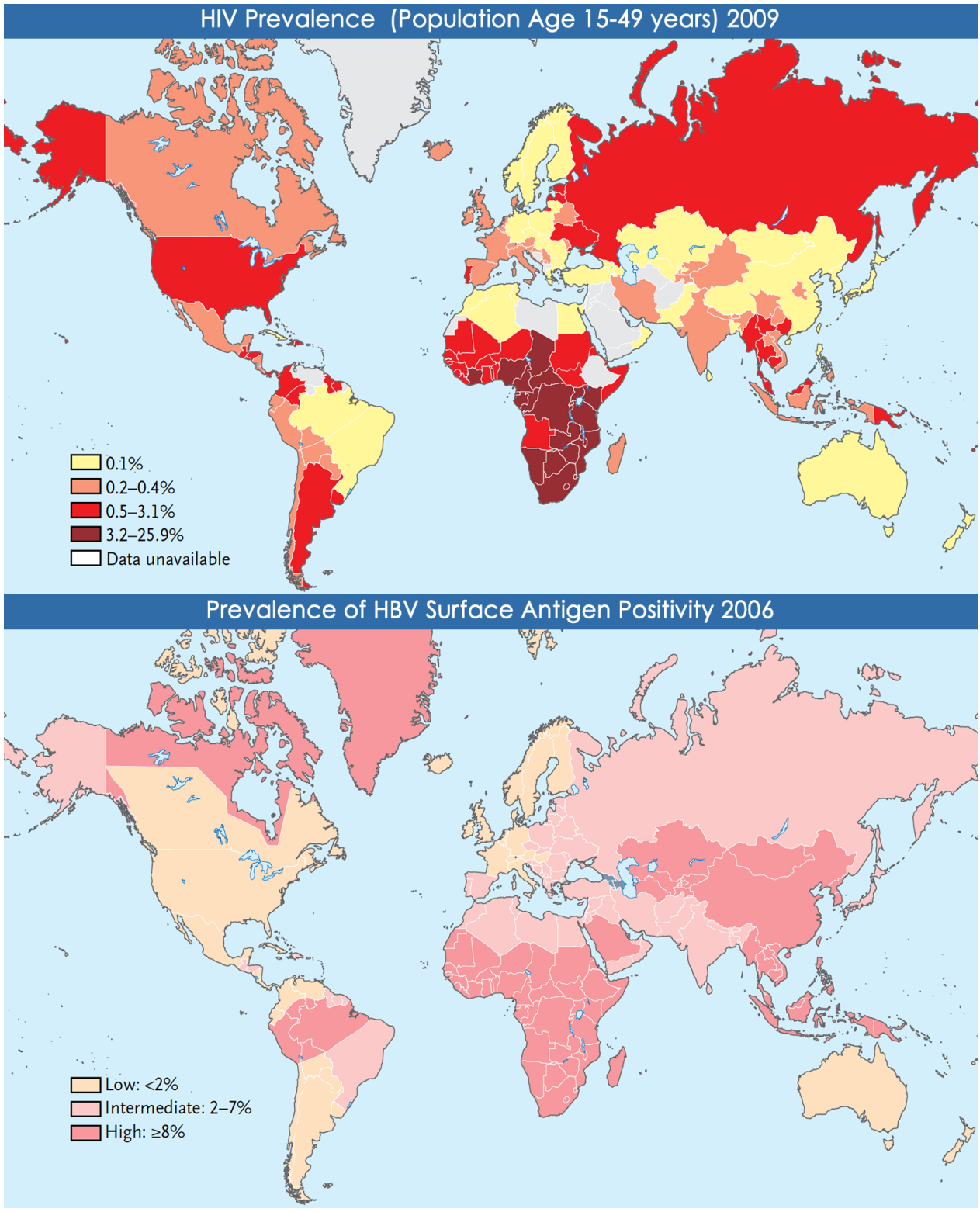


Figure 1.11: The Prevalence of HBV and the HIV Globally

Note the very high rates of prevalence of both diseases especially in Africa and Asia. The following diagram was adapted and reproduced with permission, copyright Massachusetts Medical Society [84].

Occult HBV Infection

At this point, it is important to introduce the concept of occult HBV infection. While the term was first coined in the late 1970s, prior to the Taormina expert meeting in March of 2008 the definition and relevance of occult hepatitis was widely debated [145]. The new agreed definition of occult HBV is the presence of HBV DNA in the liver where HBsAg is negative, or when HBV DNA is detectable but under 200 IU/ml [145]. False occult HBV cases can occur when variants of HBV that affect the S gene, resulting in either escape mutants or mutants where a modified and thus unrecognisable HBsAg is produced; therefore these cases are not always serologically positive for HBsAg [39, 145]. In these false occult cases, the HBV viral load is above 200 IU/ml, and often comparable to HBsAg positive cases [145]. Conversely, true occult cases are defined by instances in which HBsAg is not detectable in serum (detection escape mutants) and additionally HBV viral load is below the 200 IU/ml [145]. Other than the challenge of many cases of HBV being missed as they are false serological negatives, clinically, occult HBV is important in that the seronegative nature of most occult cases means there is a larger risk of HBV transmission resulting from blood transfusions and organ transplantations [39, 146, 147]. Further, occult HBV infection is strongly associated with HCC development as well as re-emergence of HBV viraemia following immune suppression [39, 148]. While extensive studies are required to confirm, occult HBV is particularly common in HIV-co-infected patients [97, 128, 138, 146, 147, 149-151].

Most prevalence studies are based on serological detection of HBV. In the absence of nucleic acid testing (NAT), occult HBV would not be detected by these assays, and this may result in an underestimate of HBV prevalence in the presence or absence of HIV. This means that occult HBV infection itself can be a confounding factor in those studies looking at co-infection prevalence [145].

Clinical Outcomes of HBV-HIV Co-infection

Less obvious than the epidemiology of co-infection is exactly how these two viruses interact, and what co-infection means for the individuals who are infected with both pathogens. To reiterate, in Europe, North America and Australia, HIV and HBV are usually acquired in either adolescence or adult life as a result of sexual transmission or intravenous drug use [78, 84, 127]. Conversely, in the African setting HBV is usually acquired at a young age, and thus normally precedes infection with HIV [84, 127, 132]. Most co-infected individuals in the African setting will, therefore, have an already established chronic HBV infection by the time they acquire HIV, which may have important bearings on their disease progression and outcome [132]. Adults who acquire HBV after HIV are more likely to become chronic HBV carriers compared to HBV mono-infected adults where the infection is usually acute [97]. A third scenario also occurs where individuals who had recovered from or had an effective immune response to an HBV infection or even occult HBV infection, then become HIV-infected, the subsequent immunosuppression may result in reactivation of the HBV infection [97]. Additionally, it should be noted that immune reconstitution inflammatory syndrome (IRIS) or immune recovery is likely (in up to 40% of cases) when a patient is placed on anti-retroviral treatment (ART) [126, 149, 152]. These individuals will also be more susceptible to contracting HBV or having reactivated HBV disease. Thus, the appearance of HBV after IRIS is often referred to as either reactivation or unmasking of occult infection [126]. There is also increased risk of mother-to-child or perinatal transmission of both pathogens, especially since HIV enhances HBV replication and co-infected mothers are more likely to be HBeAg-positive and thus more infectious [127]. All these scenarios make it clear that acquiring HIV and HBV co-infection is not straightforward and can have multiple outcomes; some of these clinical outcomes and complications are highlighted in Table 1.8 below.

Table 1.8: The Effect of Co-Infection on Disease Outcome for Both HBV and HIV

Effect of Co-Infection on HBV Disease	Effect of Co-Infection on HIV Disease
Exacerbates liver disease, increased inflammation of the liver during IRIS, elevated liver enzyme levels, increased jaundice, increased hepatotoxicity (3- to 5- fold) due to ART and faster progression to cirrhosis and HCC. Finally, higher increased cases of liver-related mortality.	Where ART has been implemented liver failure and end-stage liver disease due to HBV have become the leading causes of death in co-infected individuals.
Increased proportions of chronicity after acute HBV infection has been noted and faster progression to chronic disease than in mono-infection (up to 5 times faster).	Hyperbilirubinemia, mitochondrial toxicity, systemic reactions and hypersensitivity reactions are common in combination with certain ART medications.
Less spontaneous clearance of HBV, and lower rates of HBeAg clearance, as well as increased cases of HBV recrudescence subsequent to immune control.	CD4 T-cell counts were consistently shown to be lower in co-infected individuals.
Increased number of occult HBV infection cases.	Lower body mass index (BMI) was reported amongst co-infected individuals.
Higher replication rates of HBV or higher viral loads/viremia have been observed.	No differences in HIV viral loads have been reported.
Reduced response to treatment especially interferons. What's more, immunosuppression can cause an accelerated loss of anti-HBs antibodies, blunting the immune response increasing the risk of HBV reactivation episodes.	Higher chances of treatment interruptions and changes in order to accommodate greater hepatotoxicity and viral resistance of the two pathogens as well as blunted immune responses or greater immune deficiency due to active HBV replication.
The development of HBV resistance is common, particularly when lamivudine or emtricitabine is prescribed in the absence of tenofovir.	Individuals with chronic HBV had a higher possibility of experiencing either an AIDS-defining or death event. Progression to AIDS-related outcomes, concerning morbidity and mortality, is also more likely and more rapid.

This table was populated with data from multiple studies and reports to show how co-infection affects each disease singularly in terms of progression and outcome [1, 84, 127-132, 140, 149, 153-155]. The effect of HBV on HIV disease is less apparent than the relationship of HIV on HBV disease. However, it is thought that HBV may blunt the immune system and cause a greater predisposition towards liver toxicity following ART administration than in mono-infected patients [130]. It is important to note that the pathogenic and molecular mechanisms for the increase in liver disease as a result of co-infection are poorly understood [131].

Presently neither HBV nor HIV can be eradicated with the therapies we have available to us [84]. In the case of HBV, new antiviral therapies have been slow in discovery, particularly In comparison to HIV or even hepatitis C virus treatment [156]. Furthermore, the challenge of treatment that can remove cccDNA in the nuclei of hepatocytes has not yet been overcome [156]. New curative therapies are desperately needed for the treatment of HBV if we ever hope to prevent the more than 1 million deaths that occur due to end-stage liver disease and HCC associated with HBV [156]. In affluent settings, where HBV-HIV co-infection studies were carried out, it was found that simultaneous

infection profoundly affected every characteristic of the natural history of HBV and that liver disease emerged as the leading cause of death in co-infected patients after the initiation of ART [1, 84, 127, 131, 154]. In the African setting, this is not as clearly delineated, and it would be wise to assume that other major infectious diseases still play a bigger role in morbidity and mortality in HIV-co-infected patients than HBV [1, 84, 128, 131]. Although ideally all HIV-positive individuals should be screened for HBV and occult HBV, the infrastructure and resources to do so are unfortunately severely limited [127, 130].

There are few important factors that that need to be highlighted at this point:

- In resource-limited settings adequate screening for HBV in HIV patients is scarce
- The occurrence of occult HBV means that even if screening is carried out there will be cases that are missed
- Co-infected individuals' lifespans are prolonged by placing them on ART
- Those that are co-infected are not necessarily placed on HIV treatment regimens that are also effective against HBV (especially regimens including Tenofovir, which has been shown to be effective treatment of HBV with minimal occurrence of resistance)
- Co-infection with HIV accelerates advancement of serious HBV-related liver disease

Taken together it becomes evident that in the African setting this situation is creating greater opportunities for chronic liver disease, especially HCC, to take hold in co-infected patients [97, 130, 131, 138, 151, 153]. Finally, it imperative to note that despite a general significant increase in numbers and types of cancers related to HIV infection and autoimmune deficiency syndrome (AIDS), liver cancer, and HCC in particular, remains the fourth most common cancer in Africa [96, 126, 157].

1.1.9 HBV Mutants and preS and S Deletions: What is Missing?

As mentioned before, variants of HBV arise naturally due to the error-prone nature of the polymerase enzyme [9, 13, 24, 31, 114, 158, 159]. As Bouckaert *et. al,* point out the variation that is introduced by these errors is opposed by a conflicting force, namely the extreme compactness of the overlapping genome, limiting the degree of variability possible [114]. With all HBV mutations, it is important to remember that because of the overlapping genome, a mutation can affect more than one gene at a time. Thus, even a single point mutation can affect many aspects of the virus [9, 11, 31, 65]. Particularly, for the purposes of this study, its important to remember the S ORF is completely overlapped by the Polymerase ORF and thus mutants that have drug resistance and/or altered replication fitness can simultaneously be vaccine-escape or HBsAg mutants [24]. Due to the high replication rate of HBV, and the polymerase being error-prone, it is reasonable to assume that each person infected with HBV harbours a virus population of quasispecies [9, 11, 24, 31, 56]. In virology, when we refer to quasispecies, we mean "*a population of viruses that share a common origin but which have distinct genomic sequences as a result of mutation, drift and the impact of selection*" (Smith *et al*., 1997) [160]. Consequently, any infected individual may be infected with several variants at any one time, and not all of these variants will be viable [9, 159]. Nassal *et al.,* 2005 above, summarise the requirements for viability of the virus. While viable variants will replicate autonomously, partially defective virus strains can be 'rescued' by trans-complementation with a wild-type (WT) or viable strain [28, 37, 159, 161-165]. Some of the mutants may arise within an infected individual as de novo mutations, but there are variants that have been found to be transmissible [24, 91]. Thus, these transmissible variants circulate as the predominant viral population [166]. There are several kinds of variants that arise in HBV, and they may arise from immune or drug pressure as well as the error-prone polymerase. All are shown in Table 1.9. While it is not yet clear, another selection pressure exerted on the HBV genome may be co-infection with HIV, as a relatively high incidence of variants have been seen in co-infected patients [52, 97, 138, 151], but the mechanism by which this occurs is unclear.

Table 1.9: HBV Variants and Their Relevance

HBV Variants			
Type of Variant	Cause of Mutation	Location of Mutation	Result
Vaccine Escape Mutants	Host immune pressure	The 'a'-determinant of HBsAg, where B and T cell epitopes are found	The immune system is primed by the vaccines, but the circulating HBV does not display the same epitopes, thus, HBs antibodies will be ineffective
Immune Escape Mutants	Host immune pressure/ vaccine pressure	The 'a'-determinant of HBsAg, where B and T cell epitopes are found	The immune system no longer recognises the virus, as epitopes are either deleted or obscured, lower HBsAg titers are common
Surface Protein Deletion-mutants	Not explicitly known	The preS1 and preS2 regions	These deletions are usually in-frame and are common in CHB and are associated with necro-inflammation, fulminant hepatitis, cirrhosis, progressive liver disease and increased HCC development
	Host immune pressure	The 'a'-determinant of HBsAg	Occult infection; HBsAg is not serologically detectable due to the deletion or decreased HBsAg secretion Reduced binding of HBsAg to anti-HBs, thus immune escape as well as reduced virion secretion
Drug Resistance Mutants	Drug pressure	The polymerase gene	Resistance to Nucleos(t)ide Analogues
Pre-Core Mutant	Host immune pressure	The HBeAg protein	Higher rates of fulminant hepatitis and exacerbation of clinical outcomes of CHB, stop codon mutations can lead to the complete blockage of HBeAg production
Core Promoter Mutations	Error-prone polymerase, shown to accumulate over time	In the promoter section of the pre-core	The most common of the naturally occurring variants Decreased HBeAg production and increased seroconversion, escape of anti-HBe antibody mediated immunity, increased severity of disease prognosis as well as HCC development
X Gene Mutants	Error-prone polymerase	Usually point mutations, often close to the C-terminus of the HBx protein	HCC development

This table was populated using numerous resources [11, 26, 31, 35, 37, 40, 51, 56, 65, 166-170].

Owing to the research focus here being deletion-mutants in the S gene, its important to remember that alterations of this gene may affect infectivity, virion binding to and secretion from hepatocytes, virus transmission, response to vaccine-primed immunity as well as pathogenesis of liver disease [31, 65]. It is still not clear exactly how HBV causes HCC, as the virus is not cytopathic [171]. It is thought to be largely due to the interplay of the viral replication and the host immune system that causes continuous liver injury and regeneration and ultimately leads to tumorigenesis [11, 26, 31, 73, 166]. While HBV DNA integration was previously thought to be a possible culprit, it has since

been shown that integration is both random and does not have a *cis*-effect on flanking genes, so it is unlikely to be an HCC-causative mechanism [91]. However, some studies have shown that when integrated into specific sites of the host genome, the S and X genes can possibly contribute to genomic instability and mutagenesis [170]. Certain genotypes have a stronger link with HCC development than others; subgenotype A1, predominant in sub-Saharan Africa, for example, is known to be the cause of most cases of HCC in that region [26]. There is a growing body of work showing an association between variants that contain preS deletions or produce truncated envelope proteins and the development of advanced liver disease, including cirrhosis and HCC [24, 26, 40, 51, 119, 166-168, 170, 172-178]. Of particular note, a meta-analysis that scrutinised nearly 12 000 patients involved in 43 different studies found that preS deletion-mutants were associated with a 3.77-fold increased risk of developing HCC [167]. While the S region of the genome is moderately well conserved, the preS1 and preS2 are much more variable and tend towards mutation, and we now know that preS variants frequently arise in patients with chronic HBV [24, 35, 170, 179]. Furthermore, preS deletion-mutants become more likely the older the HBV-infected individual becomes [24, 167, 172]. In-frame preS deletions are most commonly deletions in the C-terminal end of the preS1, central deletions in the preS2, the start codon of preS2 or miscellaneous point mutations [24, 91]. Not surprisingly, because the MHBs protein is not essential for virion assembly, infectivity or viral secretion, mutants where MHBs is severely truncated or not synthesised at all are frequently selected [24, 91]. This said, balanced production of the surface proteins seems to be important for normal viral secretion [24, 164]. Where preS1 mutants contain deletions towards their C-terminal ends and preS2 mutants contain deletions towards their N-terminal ends the likelihood of defective or no MHBs production rises, and correspondingly these mutants are more commonly seen in HCC [24]. Moreover, the preS2 region overlaps with the spacer domain of the HBV polymerase; thus large deletions in this region are tolerated as they do not always negatively affect replication fitness [24, 180]. Infection with preS/S mutants is associated with low circulating and supernatant HBsAg levels, retention and accumulation of surface proteins in the ER of GGH cells, impaired secretion of virions and increased amounts of cccDNA building up in hepatocyte nuclei [24, 91, 93, 166, 170].

The hypotheses or viral model for hepatocarcinogenesis where S deletion-mutants are concerned include the following cellular occurrences:

- It has been shown by several studies that preS deletion-mutants are incompetent for secretion, because the LHBs is truncated and the transmembranous section required for cytosolic directionality and budding is missing or defective [24, 166]. The preS1 deletion-mutants tend to aggregate more centrally causing the appearance of 'classic' or inclusion-like type I GGH cells, whereas preS2 deletion-mutants will accumulate in the periphery or margin causing the appearance of type II GGH cells [24, 91, 166, 170, 181-183]. This is important because we can infer that the deletion-mutants produce two different kinds of truncated surface proteins and may even have differing oncogenic mechanisms [91]. For this reason, subsequent identification of the GGH type could be a diagnostic and prognostic tool for detection of precursor lesions of HCC [24, 91].
- A strong ER stress response occurs as a result of the cellular defence response triggered by the accumulation of unfolded proteins, leading to increased recruitment of cellular chaperones like Grp78 and Grp94 [91, 182, 183]. These chaperones will prevent apoptosis and increase tolerance to hypoxia thus creating a favourable environment for tumour growth [91, 182]. ER stress also leads to an accumulation of reactive oxygen species, which will damage host DNA; in due course this leads to destabilisation of the host genome in hepatocytes and mutagenesis [24, 91, 182].
- Several cellular pathways crucial for normal growth are affected by abnormal S protein accumulation and ER stress, and together they lead to hepatocyte hyperproliferation. The first of these is the augmented production of cyclin A protein; by increasing centrosome duplication this increased cyclin A leads to cell aneuploidy and may exacerbate ER stress [24, 91, 170, 182]. The second process is the induction of the cyclooxygenase-2 (COX-2) enzyme; ordinarily important for anchorage-independent growth, here COX-2 allows cells to proliferate abnormally [91]. Similarly, retinoblastoma (RB) protein becomes hyperphosphorylated, resulting in abnormally increased cell cycling conditions [24, 91]. In analogy, over-expression of the mammalian target of the rapamycin (mTOR) signal cascade, the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) protein, and mitogen-activated protein kinases (MAPK) pathways similarly lead to abnormal cell cycling and hyperproliferation [24, 91, 170, 182]. Importantly NF-κB also upregulates COX-2 [182].

- Mutants affecting the MHBs specifically have been shown to upregulate human telomerase reverse transcriptase (hTERT); where it normally regulates telomerase activity, overexpression promotes carcinogenesis [24].
- In mouse models, it has been shown that where very high viral loads are reached the HBV can become directly cytopathic and can burst hepatocytes after the extreme accumulation of LHBs [24, 91, 166, 170, 182, 184].

Ultimately the retention of truncated surface proteins leads to ER stress resulting in oxidative damage to host DNA, genome instability and mutagenesis and the induction of cell cycle disturbances and hepatocyte hyperproliferation [24, 91, 170]. It is thus very important to keep studying the incidence of these S deletion-mutants as they can be strong prognostic tools; but more importantly, it is possible with a better understanding of the carcinogenesis model of HBV that we would be able to identify novel drug targets [91, 182].

1.1.10 The Cell Culture Systems Available for HBV Study

It is important to remember in a study like this where the aim is to functionally characterise mutant strains of Hepatitis B virus, there are two major limiting factors. Firstly HBV has a very narrow range of species that it can readily infect; and secondly because of this; there is only a handful of *in vitro* systems available in which to carry out functional studies. The four core *in vitro* model systems that can be employed when studying the Hepatitis B Virus and their main advantages and disadvantages are summarised in Table 1.10. Transient transfection of human hepatic cell lines provides a readily available and relatively cost-effective model that adequately simulates the liver microenvironment [5, 60, 185-187]. As the cell lines can be transiently transfected with overlength clones, they can be used to study the replication competence of these HBV clones as well as carry out other experiments towards the functional characterisation and prediction of the clinical phenotype of such clones [5, 60, 65, 187-189]. Of course transient transfection systems have shortcomings too, including inter-experimental variability where a stably transfected cell line is not employed, variable transfection efficiencies are frequent and the peak of viral expression being reached prematurely due to incomplete transfection of all cells [5]. Establishing a stably transfected cell line is not always feasible if the time frame in which to do the study is limited.

Table 1.10: Comparison of Cell Lines Used for HBV

Primary Hepatocytes	
Primary Human Hepatocytes (PHH) Origin: derived from surgically resected liver pieces	
Advantages: Most closely represent average human hepatocytes, naturally susceptible to infection and thus can be used for infection and viral entry studies These cells allow for the study of the entire replication cycle	Disadvantages: Differences arising from donor heterogeneity mean infection efficiencies vary; cells are only viable for a few days post-infection as they lose liver-specific function and de-differentiate back to fibroblast-like cells. It is expensive and difficult to acquire these cells, ethics clearance is required, donors may not be suitable due to liver disease, requires surgery on a live person
Primary *Tupaia Belangeri* Hepatocytes (PTH) Origin: derived from surgically resected liver pieces of the Northern Tree shrew	
Advantages: These cells are immediately susceptible to infection and infection has been shown to be specific, entire replication cycle of the virus is possible. As Tree shrews can be kept in captivity, these cells are easier to obtain than human hepatocytes, and often the cells obtained are more homogenous for the same reason	Disadvantages: Only transient infection can be achieved Will also lose hepatic phenotype if not maintained by chemical induction
Human Hepatic Cell Lines	
HepaRG Origin: A hepatoma cell line isolated from a female patient with both HCC and hepatitis C infection	
Advantages: Unlike the other hepatoma cell lines, HepaRG is susceptible to HBV infection, and thus, the entire replication cycle, viral attachment, entry, drug metabolism as well as viral uncoating can be studied in this cell line	Disadvantages: Requires constant supplementation of DMSO and hydrocortisone to keep cells from de-differentiating and losing NTCP expression Requires both DMSO and polyethylene glycol for efficient infection;ß despite supplements it is still likely that only a subpopulation of the cells remains susceptible to infection Low HBV replication is seen in this cell line
Huh7 and HepG2 Origin of Huh7: A hepatoma cell line obtained from a 57-year-old Japanese male with HCC Origin of HepG2: A human hepatoblastoma cell line isolated from a 15-year-old Caucasian American male with HCC	
Advantages: Can be transfected with overlength genome plasmids, which then support viral replication, viral gene and protein expression, viral assembly and production of infectious virions, which can be stably maintained for extended periods of continuous study and can be propagated in a purely chemically defined media. HBV replication is high in this cell line, continuous study lasting up to nine months in passaged transfected cells is possible, they retain functional hepatocyte activity to a large degree If constitutive expression of NTCP can be achieved both these cell lines are susceptible to HBV infection Can be propagated into a stable cell line, which can overcome inter-experimental variability	Disadvantages: Not susceptible to direct HBV infection Susceptible to mycoplasma infection Low drug metabolizing capability The mutated p53 gene in Huh7 cells results in an increased half-life as well as increased expression of p53 in these cells Both cell lines do not allow for study of the oncogenic potential of HBV as they are already derived from cancerous cell lines Transfection efficiencies may vary

This table was constructed using a number of resources to summarise the main features of several popular cell culture systems. [5, 16, 45, 60, 185, 186, 188-191]

1.1.11 Study Rationale: Why Carry Out Functional Characterisation on S Deletion-mutants

It is apparent that preS deletion-mutants play an integral role in the rapid development of HCC [91, 124, 170]. HIV co-infection is not just another risk factor for HCC development, but may lend an opportunity for HBV occult infection going unnoticed and probably untreated [126-132, 138, 157]. While much work has been done in genotypes other than genotype A, limited work on molecular characterisation has been undertaken and as yet no functional characterisation of preS deletion-mutants in subgenotype A1 has been performed to date in the African context [39, 100, 117, 133, 134, 137, 138, 140, 142, 147, 153, 179].

The cohort from which this current study draws, enrolled 298 treatment-naïve and HIV-positive patients from Mpumalanga in South Africa, of which 23.8% were co-infected with HBV (Figure 1.10 below) [153]. It was found that out of a pool of 71 HIV-co-infected patients, 43 (63%) of those had occult HBV infection (Figure 1.10) [153].

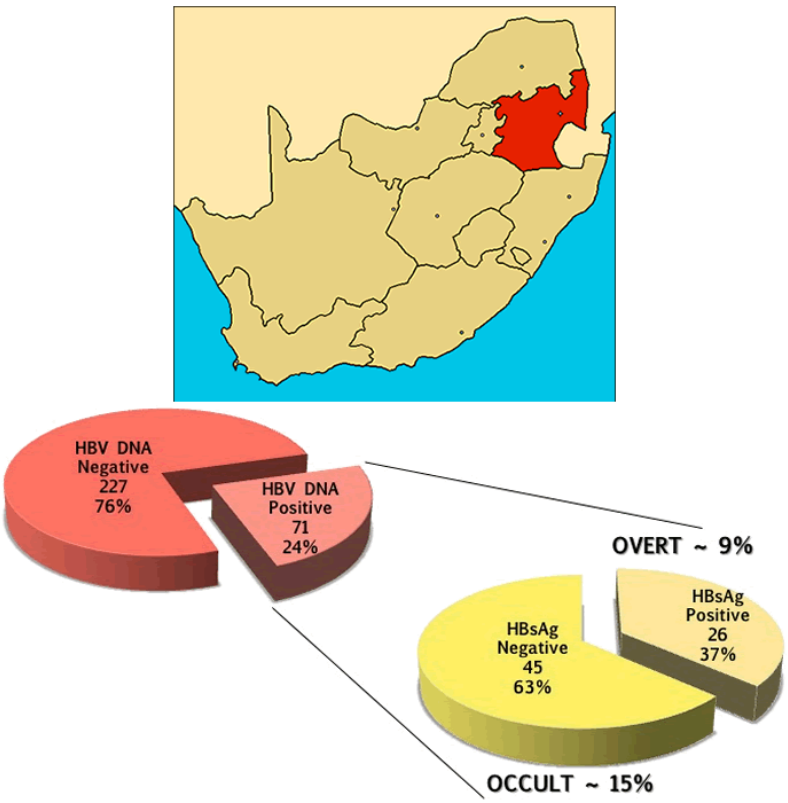


Figure 1.12 Shongwe Cohort Outcomes on Co-Infection and Occult HBV

The numbers reported here are from the Bell *et. al.,* 2012 Shongwe cohort study, which was carried out in Mpumalanga South Africa, highlighted in red on the map [153]. The red pie chart represents the number of individuals who were HIV-positive, and either mono- (76%) or co-infected (24%). The number appearing above the percentage represents actual participants, the total of which was 298. The yellow pie chart shows the proportions of participants from the co-infected group who were classified as either overt or occult infection cases.

When sequencing of the S region of 29 of these patients was carried out, isolates from five patients were found to have deletions in either the preS1, preS2 or both [52]. Four of these isolates were from HBsAg-positive participants, but one was a case of occult HBV infection [52]. The deletions were in-frame and ranged from 30-66 nt deletions which are mapped in Table 1.11 and depicted in Figure 1.13 [52]. One additional patient sequence, SHH193A was chosen as it was a full-length sequence and could thus be employed as a full-length cloning control. Notably, this patient was confirmed to have an occult infection, as this patient had undetectable HBsAg in serum and had a viral load above 200 IU/ml [52, 145].

Importantly, in other studies carried out in southern Africa and India it was found that a very high percentage of patients, infected with subgenotype A1 strains and having developed HCC, harboured HBV with the same deletions in the preS region [118, 192].

Table 1.11: Summary of Patient Sequences Selected from the Shongwe Study

No.	Patient	Subgenotype	Deletion		
			Region	No of nucleotides	Positions from the *Eco*RI site
1.	SHH011A	A1	preS1 and preS2	30 nt and 66 nt	2900 to 2929 and 3211 to 3255
2.	SHH045A	A1	preS2	33 nt	23 to 55
3.	SHH167A	A1	preS2	45 nt	9 to 54
4.	SHH274A	A1	preS2	54 nt	2 to 55
5.	SHH300A	A1	preS2	54 nt	3 to 56
			Point Mutations (Positions from the *Eco*RI site).		
6.	SHH193A	A1	Full-length		

It should be noted that while the sequence from patient SHH193A was a full-length sequence, and a WT relative to the deletion-mutants. This patient sequence was the most similar to the HBV backbone into which it would be cloned, and could thus be used as a cloning control. This patient was confirmed to have a false occult infection [52]. Thus HBsAg was undetectable in serum however viral load was above 200 IU/ml at baseline [52, 145].

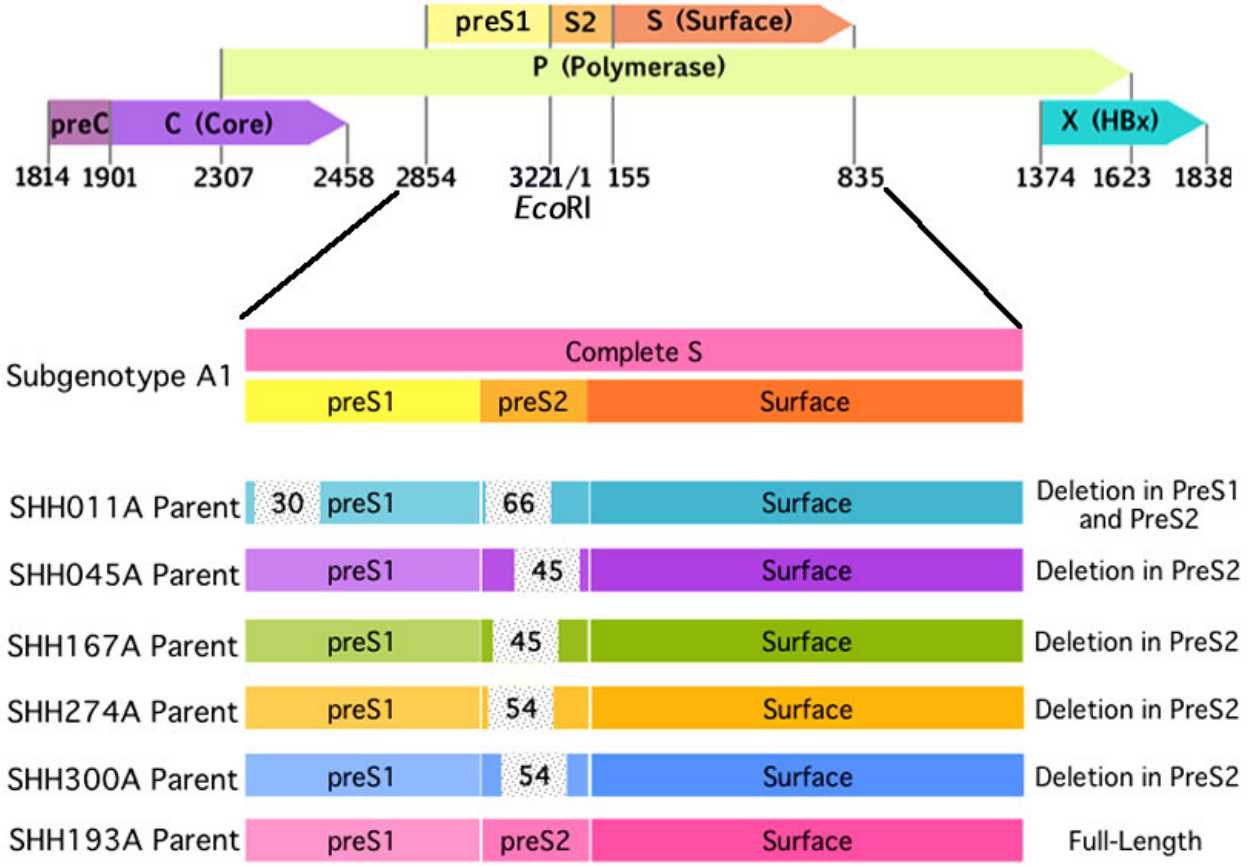


Figure 1.13: Graphical Representation of Sequences Selected from the Shongwe Study

Each participant was identified by the SHH denoting the Shongwe study as well as a number indicating the order in which they entered the cohort study. In each case, the deletion size and relative positions are indicated on the grey dotted area. The deletions are outlined in this figure relative to the WT subgenotype A1 S gene of the A12C15_OL 1.28 MER construct and the full-length SHH193A patient sequence.

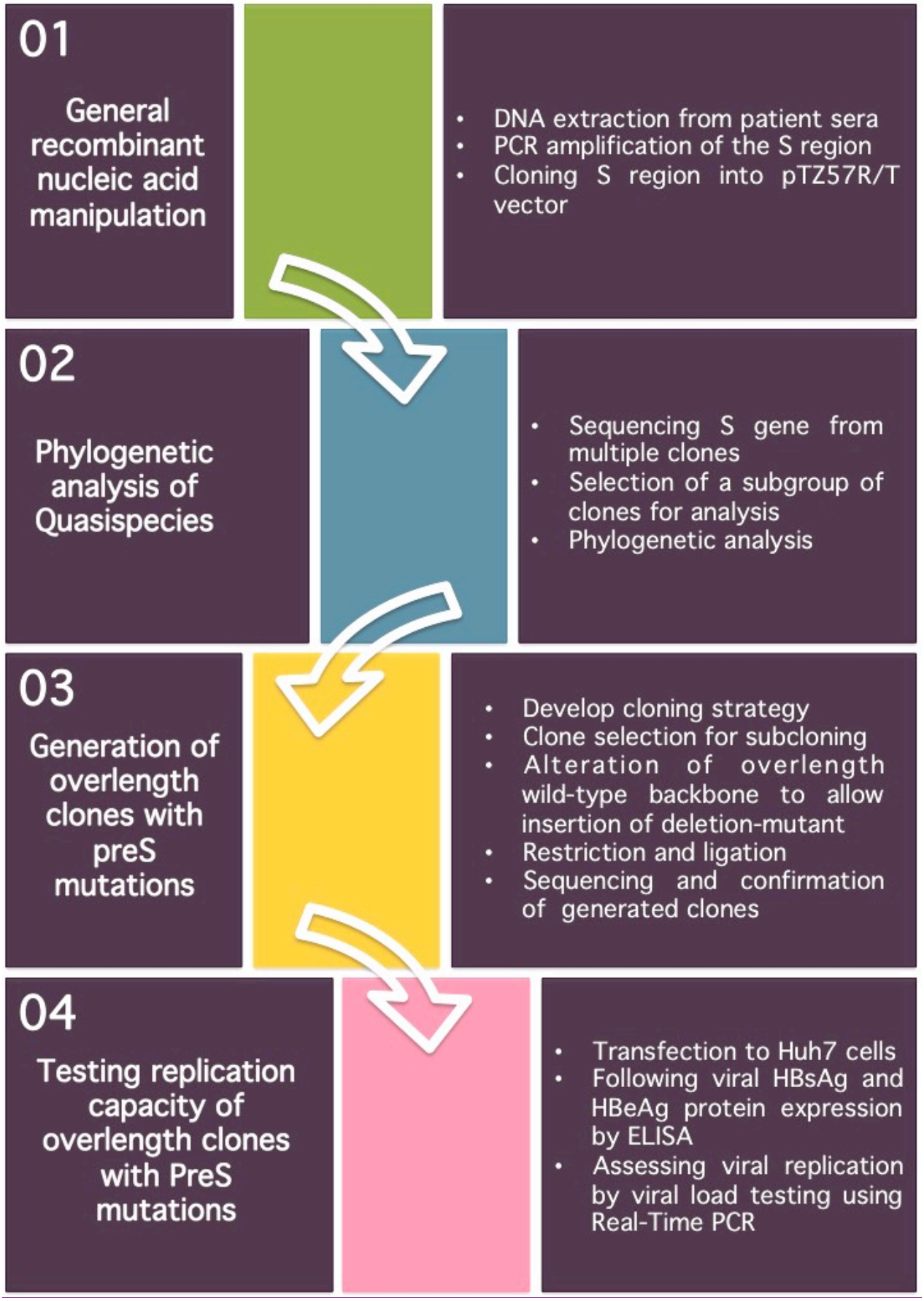
The primary aim of this study is to construct a plasmid with the preS1 and preS2 deletion-mutants discovered in the Shongwe study in a subgenotype A1 backbone and to establish whether these deletion-containing constructs are replication-competent. The objectives for this study include:

- Amplification of the complete S region from DNA extracted from Shongwe patient samples
- Cloning of the S region into pTZ57R/T plasmid for a quasispecies analysis of the viral population in each of these patients.
- Subcloning deletion sequences closest to parent sequence isolated from patients into an overlength, replication-competent, WT subgenotype A1 backbone (originally constructed by Nimisha Bhoola [158, 193]) using restriction and ligation enzymes.

- Transfecting the WT replication-competent clone (as a control and basis of comparison) as well as the newly constructed deletion-mutant and occult clones into Huh7 cells, which simulate the liver microenvironment.
- To establish replication competence by following viral replication using the expression of viral proteins using enzyme-linked immunosorbent assay (ELISA) as well as viral load (VL) quantification using quantitative Real-Time PCR (qPCR) [51].

2 Chapter Two: Materials and Methods

Overview and Workflow of the Study:



2.1 List of Materials and Equipment

All experimental procedures were performed according to standardised protocols either set out by manufacturers or published methods as described [194]. For detailed protocols, and the composition of reagents and solutions, electrophoresis gels, as well as other related compounds, please refer to Appendices A and B respectively.

2.1.1 Complete List of Materials Used for This Study:

Primers and synthesised DNA:

Integrated DNA Technologies (IDT) of Whitehead Scientific (Cape Town, South Africa) was the preferred supplier for the following:

1. Forward and reverse primers for the complete S-gene (refer to Figure 2.2 and Table 5.1 in Appendix A)

2. HBV Primers for Real-Time PCR (refer to Table 5.10 in Appendix A)

3. Two synthesised genes, within pIDTBlue plasmids (Figures 5.6 and 5.7 in Appendix C).

Enzymes: Some Mung Bean Nuclease, DNaseI and some restriction enzymes and their respective buffers, as well as 100x BSA buffer, were obtained from New England Biolabs (Ipswich, USA). The large majority of restriction enzymes and T4 DNA ligase was acquired from Thermo Fisher Scientific Inc. (Renfrew, Scotland). HotStarTaq® Polymerase was supplied by Qiagen GmbH (Hilden, Germany), and Roche (Mannheim, Germany) supplied the Taq/Tgo Polymerase mix in the High Fidelity Expand Kit. The PCR Terminator® End Repair Enzyme was acquired from Lucigen (Middleton, USA). Remaining DNaseI recombinant, RNase-free was obtained from Roche (Mannheim, Germany). SsoFast™ EvaGreen® Supermix for Real-Time PCR was obtained from BioRad (California, USA).

Molecular Weight Markers: Ladders for nucleic acids, such as the O'GeneRuler™ 1kb DNA Ladder, GeneRuler™ 1kb DNA Ladder, 1kb DNA Plus GeneRuler™ Ladder and O'GeneRuler DNA Ladder Mix were acquired from Thermo Fisher Scientific Inc. (Renfrew, Scotland). The 1 kb DNA Ladder was supplied by Promega (Madison, USA).

Gel Electrophoresis: Molecular Grade Agarose was supplied by White Scientific (Cape Town, South Africa). Ethidium Bromide was from Bio Basic Inc. (Ontario, Canada). Novel Juice Gel Staining Reagent was obtained from GeneDireX (Gueishan, Taiwan).

Kits: The NucleoBond® PC 500 Kit, NucleoBond® XtraMidi EF, NucleoSpin® Plasmid QuickPure and NucleoSpin® Gel & PCR Cleanup Kit were supplied by Machenerey-Nagel GmbH & Co (Düren, Germany). Qiagen GmbH (Hilden, Germany) was the preferred supplier of the following: HotStarTaq® Mastermix PCR Kit, QIAprep® Spin Miniprep Kit, QIAamp®DNA Blood Mini Kit. ZymoResearch (California, USA) supplied the DNA Clean&Concentrator Kit. The High Fidelity Expand PCR Kit was procured from Roche (Manheim, Germany). PCR Terminator® End Repair Kit was acquired from Lucigen (Middleton, USA). The Murex® HBsAg Version 3 ELISA and ETI-EBK PLUS HBeAg Kits were obtained from DiaSorin (Sallugia, Italy). Monolisa™ HBsAg ULTRA Kit for immunocapture was obtained from BioRad (California, USA).

Antibiotics and Bacterial Strains: Kanamycin and Ampicillin were acquired from Melford Laboratories (Ipswich, UK). For the bacterial strains, DH5α *Escherichia coli* was obtained from Invitrogen Life Technologies (Paisley, Scotland), and the SCS110 *Escherichia coli* was obtained from Stratagene-Agilent Technologies (California, USA).

Dry Chemicals: Peptone/Tryptone, Yeast Extract and Piperazine-1,2-bis(2-ethanesulfonic acid) (PIPES) were obtained from Merck (Darmstadt, Germany). Merck (Wadeville, South Africa), supplied Bacteriological Agar, Sodium Chloride and Tris(hydroxymethyl)aminomethane. Both Magnesium Chloride and Calcium Chloride were supplied by Saarchem (Krugersdorp, South Africa).

Liquids: Alcohols such as Absolute Ethanol, Methanol, Propan-2-ol and Glycerol we supplied by Merck (Darmstadt, Germany), which was also the preferred supplier of the cleaning agent Extran. Acids such as Hydrochloric Acid were obtained from Saarchem Merck (Krugersdorp, South Africa). Sterile saline water (used as best-quality water in PCR, was supplied by Adcock Ingram (Johannesburg, South Africa), and ULTRAPure Water was supplied by GibCo by Life Tech Technologies (Paisley, Scotland).

Plasticware and other consumables: Lab coats were obtained from Supreme Workwear (Johannesburg, South Africa). All nitrile and latex gloves were acquired from Lasec (Cape Town, South Africa). 5, 10, 20 and 50 ml pipettes were acquired from Techno Plastic

Products® (Trasadingen, Switzerland). 15 ml falcon tubes and 50ml tubes were obtained from Greiner Bio-One (GmbH, Austria). Minisart 0.2 μm Filter units were obtained from Sartorius (Goettingen, Germany). All 1,5 and 2,0 ml tubes were supplied by Eppendorf (Barkhausenweg, Hamburg). Thin wall PCR* tubes were acquired from QSP Hoffman-LaRoche (Basel, Switzerland). Petri dishes 47 mm for bacterial work were supplied by Merck Millipore (Darmstadt, Germany). MicroAmp™ Optical 96-Well reaction plates for Real-Time PCR were supplied by Applied Biosystems (Singapore, Republic of Singapore). Microseal® 'B' film to seal 96-well plates was obtained from BioRad (California, USA).

Tissue Culture: Hyclone Foetal Bovine Serum was acquired from Thermo Scientific (Northumberland, UK). Gibco™ Phosphate Buffered Saline tablets, Gibco™ 0.25% Trypsin EDTA (10x), DMEM GlutaMAX™ media and 100x Pen Strep Glutamine were supplied by GibCo™ by Life Tech Technologies (Paisley, Scotland). 75 cm^2 cell culture flasks were obtained from Sigma (St. Louis, USA). Tissue culture 12-well plates, as well as 5, 10, 20 and 50 ml pipettes, were acquired from Techno Plastic Products® (Trasadingen, Switzerland). Minisart 0.2 μm filter units were obtained from Sartorius (Goettingen, Germany). TransIT® Transfection reagent was obtained from Mirus Bio (Madison, USA). Polyethylenimine linear (PEI) was supplied by Polysciences Inc. (Pennsylvania, USA). Protease Inhibitor Cocktail tablets cOmplete Mini was obtained from Roche (Manheim, Germany).

2.1.2 Complete List of Equipment Used for This Study:

Pipettes either from Gilson Inc. (Middleton, USA), or from Finnpipette made by Thermo Fisher Scientific Inc. (Renfrew, Scotland). The Eppendorf Easypet 4221 was supplied by Eppendorf (Barkhausenweg, Hamburg). Other Eppendorf acquired equipment included centrifuges S810R, 5424, 5417C and 5453, the Mastercycler gradient 5331 PCR machine as well as the Thermomixer Comfort with 1.5ml adaptor (Barkhausenweg, Hamburg). The Hettich Universal II centrifuge in the tissue culture area was obtained from DJB Labcare Ltd. (Buckinghamshire, UK). Transilluminators used to visualise non-ethidium bromide electrophoresis results included the Kodak Electrophoresis Documentation and Analysis System supplied by Panavision (Midrand, South Africa) and the DarkReader Transilluminator DR-45M from Clare Chemical Research (Dolores, US). The gel electrophoresis workstations including the Sub-Cell® GT electrophoresis tanks and the PowerPac™ Basic power packs were supplied by BioRad (California, USA). For visualising

and recording gel electrophoresis results the BioRad Geld Doc system and the accompanying software; Quantity-One v4.6.9 from BioRad (California, USA) was used. Also from BioRad: the Microplate Reader Model 680, the Plate Washer PW40, the T100™ Thermal Cycler PCR machine, as well as the CFX96™ Real-Time System with C100™ Thermal Cycler (California, USA). The NanoDrop ND 1000 spectrophotometer was from NanoDrop Technologies-Thermo Fisher Scientific Inc. (Waltham, USA). The heating block used was the Accublock Digital Dry bath and vortex machine used was the VX10 as well as the Prism microcentrifuges were all supplied by Labnet International Inc. (Edison, USA). The Heidolph MR1 magnetic stirrer was obtained from Heidolph (Schwabach, Germany). The -70˚C chest freezer, fume hood without UV light, the water baths and the 37˚C shaking incubator were all acquired from Labcon Air and Vacuum Technologies (Midrand, South Africa). The Esco Ascent ductless fume hood was obtained from Esco® (Pennsylvania, USA). The Biosan DNA/RNA UV cleaner hood for PCR and extraction was procured from BioSan Medical-Biological Research and Technologies (Riga, Latvia). The Labotec Bio-Flow laminar flow fume hood in the tissue culture room was supplied by Labotec (Midrand, South Africa). The stationary 37˚C incubator, Jouan EB 53, was obtained from Labexchange® (Burladingen, Germany). The tissue culture incubators included the Heal Force HF90 CO2 Incubator from Heal Force (Shanghai, China) and the Forma™ Scientific™ Water Jacketed CO_2 Incubator supplied by Thermo Fisher Scientific Inc. (Renfrew, Scotland). A number of upright and chest fridges and freezers were obtained from Defy Appliances (Pty) Ltd (Durban, South Africa). The Orion Star™ A111 pH benchtop meter was acquired from Thermo Fisher Scientific Inc. (Renfrew, Scotland). The Olympus CK2 inverted microscope was obtained from Olympus Optical Co., Ltd. (Tokyo, Japan). Two autoclaves were used; the Erma Optical Works autoclave was acquired from Optical Works Inc. (Tokyo, Japan) and the SteriLClav-28 from Raypa® (Barcelona, Spain). The UltraSonic Cleaner – Memory Quick was obtained from EUMAX (Shenzhen, China). Refrigerated Microfuge SIGMA 1-14K (Osterode am Harz, Germany).

2.1.3 Complete List of Bioinformatics Tools, Suites and Software Used for This Study:

- CLC Main Workbench v6.6 and CLC Sequence Viewer v6.6 from CLC Bio (Aarhus, Denmark http://www.clcbio.com)
- Unipro UGENE v1.10.1 from UniPro (Novosibirsk, Russia http://ugene.unipro.ru)
- Finch v1.4.0 from Geospiza Inc. (Seattle, USA www.geospiza.com/finchtv)
- MEGA v5.05, Molecular Evolutionary Genetics Analysis version 5 from the Institute of Molecular Evolutionary Genetics (Pennsylvania, USA http://www.megasoftware.net/) [195, 196]
- EnzymeX v3.1 by Nucleobytes.com Software (Amsterdam, The Netherlands http://nucleobytes.com/index.php/enzymex)
- Fragment Merger Tool by Trevor Graham Bell (Johannesburg, South Africa http://hvdr.bioinf.wits.ac.za/fmt/) [197]
- Pipeline: TreeMail by Trevor Graham Bell (Johannesburg, South Africa http://hvdr.bioinf.wits.ac.za/treemail/) {unpublished data}
- Babylon Tool by Trevor Graham Bell (Johannesburg, South Africa http://hvdr.bioinf.wits.ac.za/SmallGenomeTools/) [198]
- Mutation Reporter Tool by Dr Trevor Graham Bell (Johannesburg, South Africa http://hvdr.bioinf.wits.ac.za/mrt/) [199]
- Quantity One 1-D Analysis Software from BioRad (California, USA http://www.bio-rad.com/en-us/product/quantity-one-1-d-analysis-software)
- GeneDoc: Analysis and Visualisation of Genetic Variation by Karl B. Nicholas, Hugh B. Nicholas Jr. and David W. Deerfield II. (San Francisco, USA http://www.nrbsc.org/gfx/genedoc/ebinet.htm) [200]
- RANDOM.ORG - True Random Number Service by Dr Mads Haahr of the School of Computer Science and Statistics at Trinity College (Dublin, Ireland http://www.random.org/sequences/)
- MEGA 5: Molecular Evolutionary Genetics Analysis by Sudhir Kumar, Koichiro Tamura, Daniel Peterson and Glen Stecher (Tokyo, Japan http://www.megasoftware.net/) [196]
- jModelTest v2.1.4 by D. Darriba, G.L. Taboada, R. Doallo and D. Posada of the Department of Biochemistry, Genetics and Immunology at the University of Vigo and the Department of Electronics and Systems at the University of A Coruna (Vigo, Spain https://code.google.com/p/jmodeltest2/) [105]

- PAUP* v4.0a142 by David L. Swofford at Duke University (Durham, USA http://paup.csit.fsu.edu/) [201]
- GARLI: Genetic Algorithm for Rapid Likelihood Inference by Derrick J. Zwickl (Austin, USA http://www.bio.utexas.edu/faculty/antisense/garli/garli.html) [202]
- BEAUti: Bayesian Evolutionary Analysis Utility v1.8.0 by Alexei J. Drummond, Andrew Rambaut, Marc A. Suchard and Walter Xie of the Institute of Evolutionary Biology, University of Edinburgh (Edinburgh, Scotland http://beast.bio.ed.ac.uk/) [104]
- BEAST: Bayesian Evolutionary Analysis Sampling Trees v1.8.0 by David Geffen and Andrew Rambaut of the Institute of Evolutionary Biology, University of Edinburgh (Edinburgh, Scotland http://beast.bio.ed.ac.uk/) [104]
- Tracer v1.6.0 by Andrew Rambaut, Marc A. Suchard, Walter Xie and Alexei J. Drummond of the Institute of Evolutionary Biology, University of Edinburgh (Edinburgh, Scotland http://beast.bio.ed.ac.uk/) [103, 104]
- Tree Annotator v1.8.0 by Andrew Rambaut and Alexei J. Drummond of the Institute of Evolutionary Biology, University of Edinburgh. (Edinburgh, Scotland http://beast.bio.ed.ac.uk/treeannotator) [104]
- FigTree: Tree Figure Drawing Tool v1.4.0 by Andrew Rambaut of the Institute of Evolutionary Biology, University of Edinburgh. (Edinburgh, Scotland http://tree.bio.ed.ac.uk/) [104]
- R Studio – open-source commercial statistics package by R Studio. (Boston, USA https://www.rstudio.com/products/rstudio/)
- Bio-Rad CFX Manager v2.1.1 Real-Time Software System by developed by BioRad Laboratories (California, USA https://bior-rad.com/pcrsoftware)

2.1.4 List of Plasmids Generated and Utilised During This Study:

Table 2.1: Constructs Generated and Utilised During This Study

Plasmid Name:	Description:	Source/Reference
pIDTBlue	A high-copy vector of 2951 bp in size. Based on the pBluescript II SK (+) phagemid; originally designed for cloning, sequencing and *in vitro* transcription. Used here as a backbone for synthesised gene products	Integrated DNA Technologies (IDT) Gene Products (Alting-Mees, 1989 [203])
pIDTBlue_S011A-1	A 3710 bp plasmid containing a 709 bp synthesised segment of SHH011A plus flanking *Xba*I sites with Ampicillin resistance gene	Generated for this work by IDT Gene Products
pIDTBlue_S045A-1	A 3773 bp plasmid containing a 772 bp synthesised segment of SHH045A plus flanking *Xba*I sites with Ampicillin resistance gene	Generated for this work by IDT Gene Products
pTZ57R/T	TA cloning vector that allows for blue-white screening and has an ampicillin resistance gene	Thermo Scientific/ Fermentas
pEGFP-C3	A mammalian expression vector containing a fluorescent tag (Green Fluorescent Protein) under the control of the CMV promoter. Contains a Kanamycin resistance gene. Used here to calculate transfection efficiency	Addgene/Clonetech
pcDNA™4/TO Vector	pcDNA™4/TO is a 5.1 kb expression vector available from Invitrogen. The vector allows expression of the gene of interest in transfected mammalian host cells	Invitrogen Life Technologies
A12C15_OL A1WTSA1.281MER	1.28 overlength replication-competent Hepatitis B Virus genome of subgenotype A1 within pcDNA4_TO backbone	Bhoola, N.H., *et. al.*, 2014 [193]
A12C15_OL A1WTSA1.281MER Altered 9.3	1.28 overlength replication-competent Hepatitis B Virus genome of subgenotype A1 within pcDNA4_TO backbone, where the third *Xba*I site present in the backbone was removed and is thus 32 bp shorter than the parent plasmid	Present study
OL _DELMUT SHH011A-F	The 1.28 overlength replication-competent subgenotype A1 backbone containing the SHH011A deletion-mutant in the correct orientation	Present study
OL _DELMUT SHH045A-F	The 1.28 overlength replication-competent subgenotype A1 backbone containing the SHH045A deletion-mutant in the correct orientation	Present study
OL _DELMUT SHH167A-F	The 1.28 overlength replication-competent subgenotype A1 backbone containing the SHH167A deletion-mutant in the correct orientation	Present study
OL _OCFL SHH193A-F	The 1.28 overlength replication-competent subgenotype A1 backbone containing the S region of patient SHH193A, which is a occult full-length, sequence in the correct direction	Present study

For complete plasmid maps of all plasmids used in this study, please refer to Appendix C.

2.1.5 List of Bacterial Strains and Cell Culture Lines Utilised in This Study:

Table 2.2: Summary of the Bacterial Strains and Cell Lines Used for This Study

Bacterial Strain or Cell Line:	Description:	Supplier and Reference:
Escherichia coli DH5α	This bacterial strain is used for routine cloning	Invitrogen Life Sciences [204]
Escherichia coli SCS110	This is a Dam and Dcm methylation-free strain used for cloning when DNA needs to be restricted with methylation-sensitive enzymes	Stratagene-Agilent Technologies [205]
Escherichia coli XLBlue	Host strain for routine cloning, allows for blue-white screening	Stratagene-Agilent Technologies [205]
Huh7 human hepatic cell line	A human hepatic cell line generated from the explant of the hepatocellular carcinoma of a male patient in Japan in 1982	Dr Abdullah Ely from the AGTRU

2.2 Overview of Statistics Employed in This Study

Where appropriate statistical analysis was done using the student T-test to establish significant relationships in the data. This was done in a pairwise fashion using the R Studio software. Only where the p-values were equal to or below 0.05 was it considered significant.

2.3 Study Participants and Ethical Clearance

A cohort was established the Shongwe Hospital, in Malelane, a rural area in the Mpumalanga province of South Africa [52]. All patient sera were anonymised using a number referring to the order in which they were enrolled in the cohort from July 2009 to December 2010 [133, 153]. Biomedical and clinical data were matched to the number in a database. During a previously completed Masters project, it was found that five HIV-infected patients were co-infected with HBV with deletions in the PreS1 and PreS2 regions [52]. Isolates from these five patients were chosen for the present study. Furthermore, an isolate from a sixth patient, SHH193A, was chosen as it had no deletions (i.e. full-length) however the patient was classified as having false occult HBV infection [52, 153]. The quasispecies population of HBV isolated from the patients, at baseline, before the initiation of antiretroviral therapy, was analysed. Table 2.3 below outlines some key clinical data for the selected patients.

Table 2.3: Key Clinical Data for Selected Shongwe Patients

Patient	Gender	Age	BMI	HBsAg	HBeAg	Viral Load copies/ml	Viral Load IU/ml
SHH011A	F	25	27.1	+	-	6.53e x 03	1.39e + 03
SHH045A	F	18	26.3	-	-	2.98e + 04	6.33e + 03
SHH167A	M	38	16.7	+	-	1.35e + 05	2.86e + 04
SHH274A	F	52	21.2	+	+	5.16e + 04	1.10e + 04
SHH300A	M	33	20.9	+	-	1.55e + 03	3.30e + 02
SHH193A	F	23	18.7	-	-	2.06e + 04	4.93e + 03

The medical arm of the Human Research Ethics Committee of the University of the Witwatersrand was consulted for ethical clearance for this study. The study received unconditional approval from the committee on the 11th of May in 2012, certificate number M1204106 (Appendix D).

2.4 Quasispecies Analysis – General Recombinant Nucleic Acid Manipulation, Cloning and Phylogenetic Analysis

2.4.1 DNA Extraction

To obtain purified DNA from patient sera, the QIAamp® DNA Blood Mini Kit was used (full protocol refer to Appendix A). 200 µl of serum from patients, as well as from controls was used (Figure 2.1). The HBV isolated from SHH193A had no deletions in the S gene region and thus served as a full-length cloning control. DNA was eluted into 75 µl of best-quality water (BQW) and then used as template for amplification by PCR.

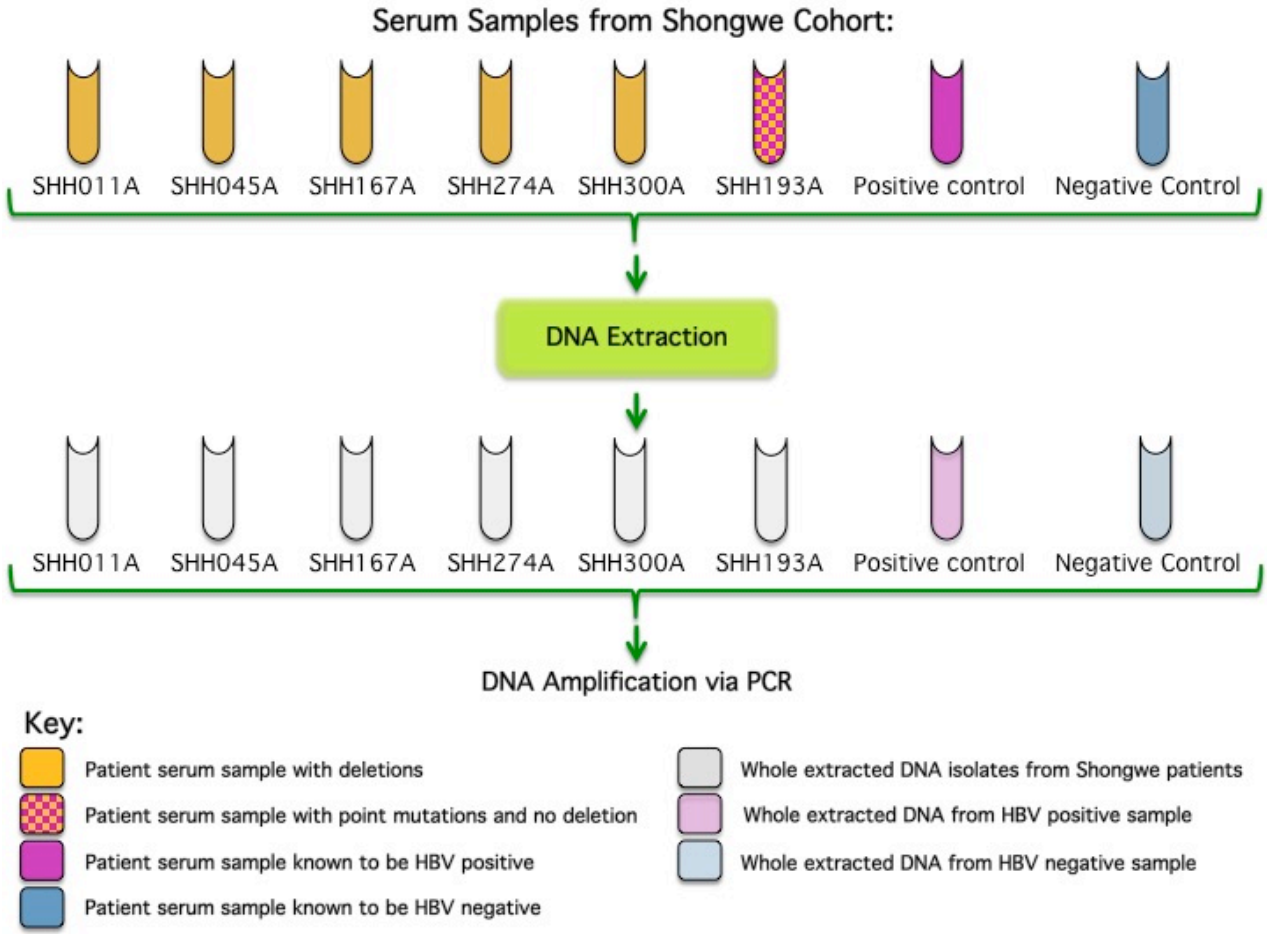


Figure 2.1: Algorithm for DNA Extraction

Each round of extraction required 200 µl of serum; DNA was eluted into 75 µl of BQW.

2.4.2 PCR Amplification

The complete S region was amplified using a nested PCR with specific primers [86], (Figure 2.2).

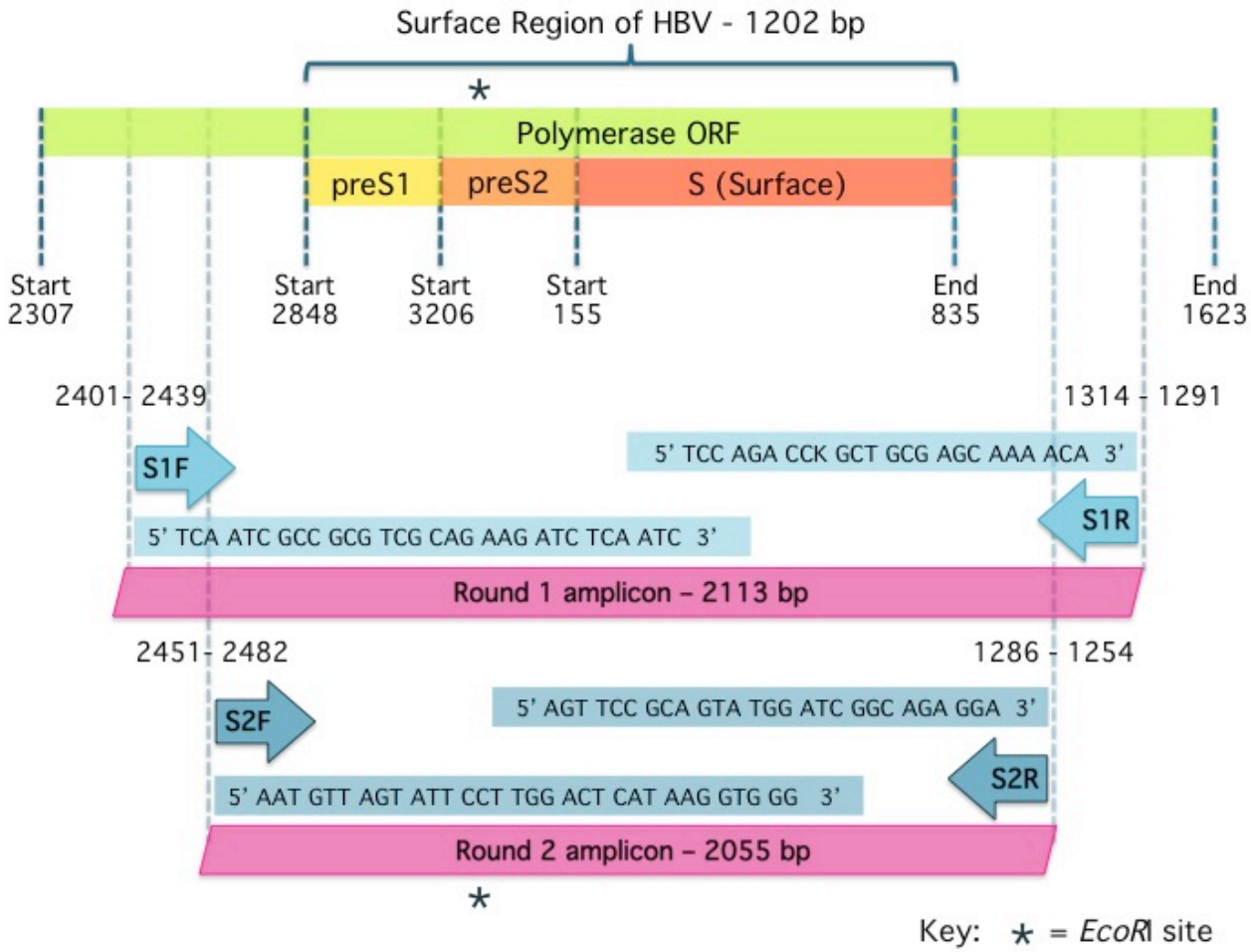


Figure 2.2: Nested PCR Strategy, Primers and Amplicons

Nested PCR, strategy, where S1F and S1R are the primers for the first round and S2F and S2R for the second round. Primer sequences are represented in light blue and amplicons in pink. Note the position of the *EcoR*I site, which is important in the numbering of the HBV genome.

Initially, a Hot Start Taq polymerase was used, just to confirm that DNA extraction was successful, and the desired amplicons were being produced (HotStarTaq® MasterMix PCR kit). Subsequently, the Roche Expand High Fidelity PCR System was employed as it provides TA overhangs, which assists in easy cloning into a vector. Also, this system contains an enzyme blend of both thermostable *Taq* Polymerase and *Tgo* Polymerase. The *Tgo* enzyme possesses proofreading ability and thus provides high specificity of PCR products. The final dNTP concentration was 200 μM of each dNTP in both PCR protocols, (see content detail in Table 5.2, Appendix A). In both PCR protocols, the positive control consisted of DNA extracted from an HBV-positive serum sample and the negative

controls included DNA extracted from an HBV-negative serum samples, as well as BQW only where no DNA was added.

Both PCR processes were optimised using the Taguchi method [206] and the cycling conditions are detailed (Figure 2.3). To confirm the size of the PCR products, 5 µl of the PCR product was run on a 1% agarose gel using a current to separate products by size. The BioRad GelDoc system was used to visualise and capture the images of the PCR products on the gel for analysis and confirmation of PCR success.

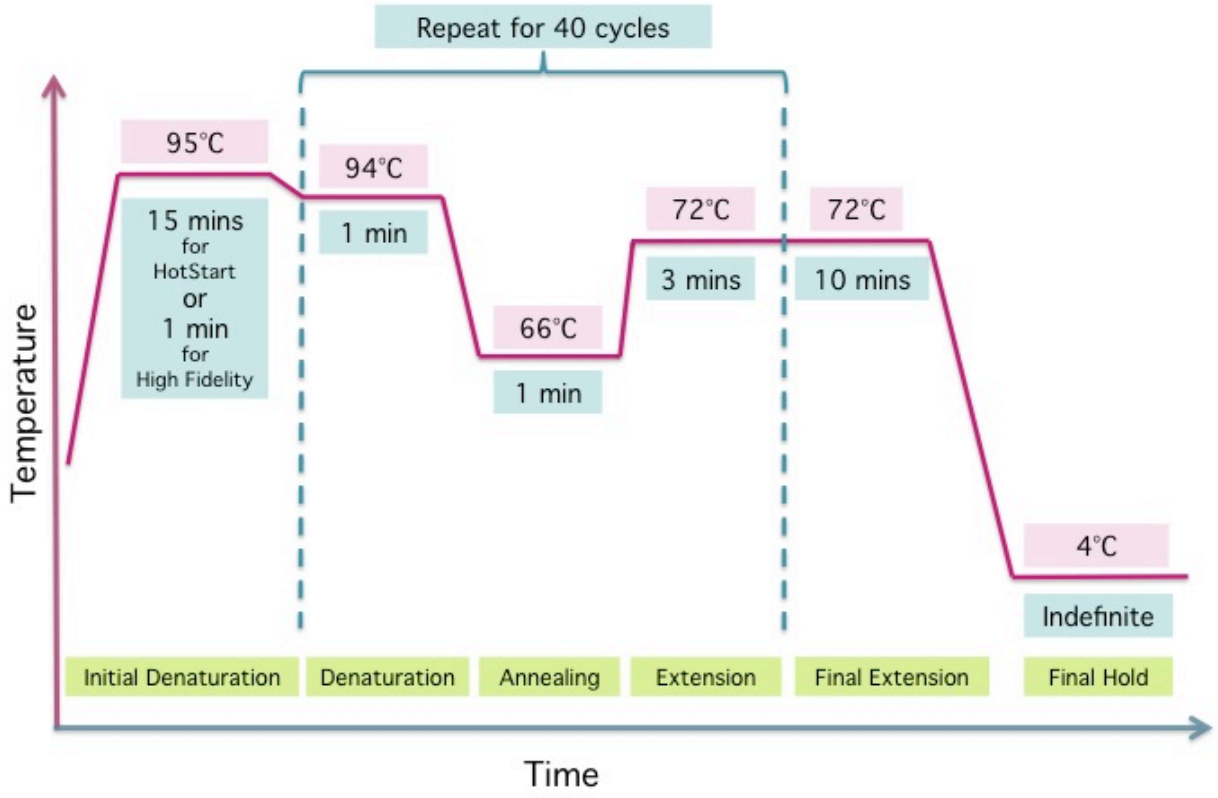


Figure 2.3: The PCR Cycling Conditions for Amplification of the Entire S Region
The cycling conditions for the first round and the second round of the two PCR reactions are identical, except for the difference in time during the initial denaturation step. The HotStart protocol requires a 10-minute initial denaturation, whereas the High Fidelity protocol only needs a 1-minute initial denaturation.

2.4.3 Cloning and Sequencing

The PCR products that were the correct size were carefully excised from the gel. The amplified DNA was extracted and purified from the gel slice using the Zymo Research DNA Clean&Concentrator Kit (Appendix A 5.1.2). Where it was intended to use DNA for downstream cloning, gels were stained with Novel Juice instead of ethidium bromide so as to avoid damage to the DNA. T and A overhangs were added by Taq, thus the purified DNA segments were directly ligated into a cloning vector. In this case, the

InsTAclone™ PCR Cloning Kit with the pTZ57R/T cloning vector was employed. Once the PCR products were ligated into the pTZ57R/T, these circular constructs were transformed into DH5α competent cells (Appendix A; 5.1.4 - 5.1.8 for details). White colonies were then screened for clones using the cracking method and restriction digestion with *Xba*I (see Appendix A; 5.1.5 and 5.1.9). After mini-prep (Appendix A 5.1.6), DNA from selected clones was sent to the Stellenbosch Sequencing Unit for Sanger sequencing. Figure 2.4 below shows the details of the sequencing primers. After chromatograms had been carefully checked and any wobble bases disambiguated, the DNA Fragment Merger Tool [197] was used to merge the three overlapping sequences into the complete S sequence fragment (Figure 2.4). These sequences were used to identify clones that were suitable for subcloning into the overlength subgenotype A1 backbone, as well as being used for analysis of the quasispecies via phylogenetic analysis.

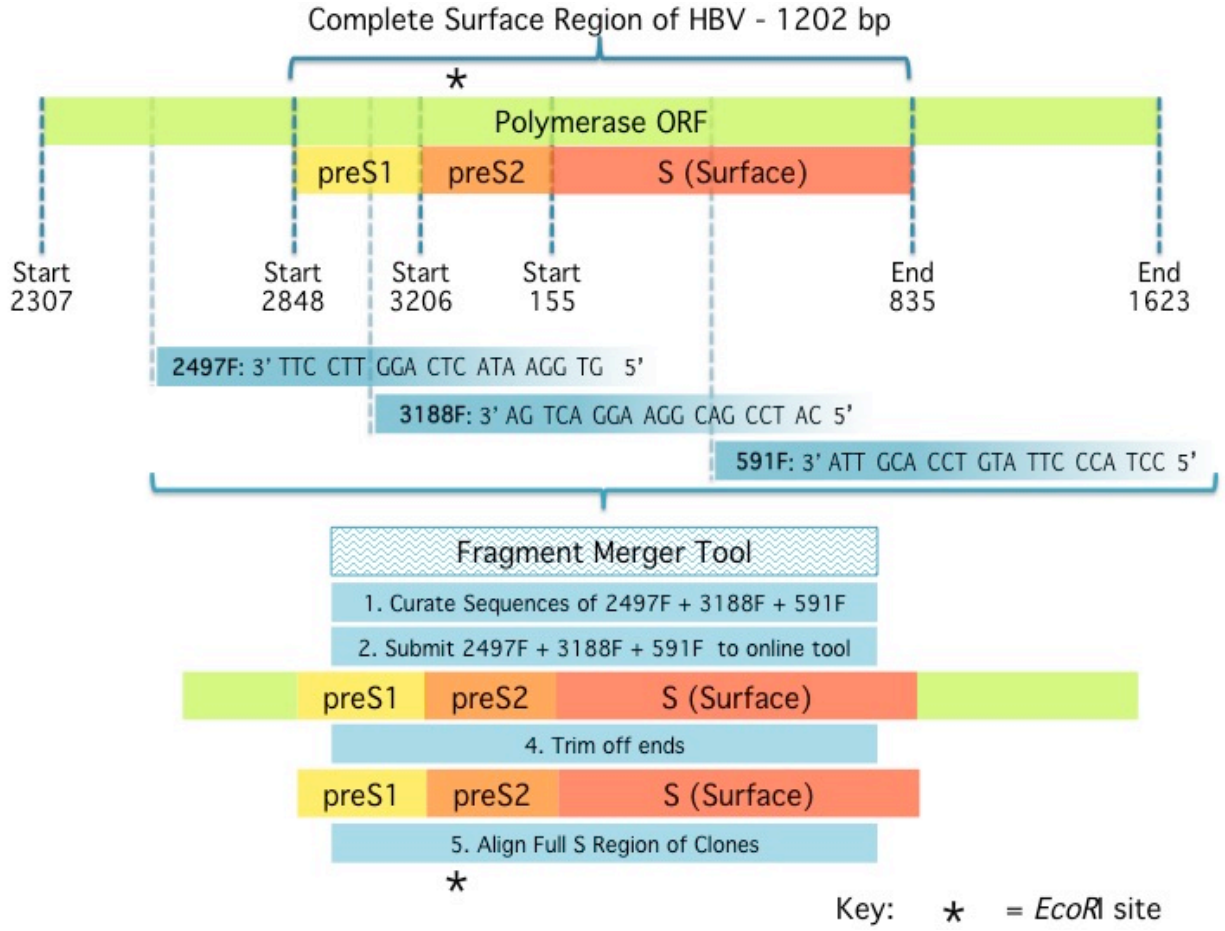


Figure 2.4: Graphical Representation of Sequencing Primers Used for Determining DNA Sequence of Quasispecies.

Each primer is likely to give a read of 800 bp of good quality sequence; the overlapping fragments were merged in the correct order using the Fragment Merger Tool. These merged sequences could then be aligned and used for phylogenetic analysis.

2.4.4 Phylogenetic Analysis

For phylogenetic analysis 30 HBV clones from each patient were chosen randomly out of the total pool of clones for phylogenetic analysis using the free online random number generator:

RANDOM.ORG (http://www.random.org/sequences/)

The random selection process ensured that the quasispecies analysis was not skewed or biased by manual selection of clones. Using ClustalW alignment in CLC Main Workbench, the sequences of the selected clones were aligned with the preS1/S2/Surface region of 6 subgenotype A1 sequences and a genotype F sequence as the root, from GenBank, [102]. For purposes of comparing the preS1/S2/Surface region of representatives from each genotype of HBV, 100 sequences were selected from Genbank. These are listed in Appendix G (Table 5.12). Three phylogenetic trees were constructed:

1. Neighbor-Joining Phylogenetic tree with bootstrapping of 1000 data sets – the multiple sequence alignment was submitted as a PHYLIP file to Pipeline: TreeMail (http://hvdr.bioinf.wits.ac.za/treemail/). The resulting tree was annotated in FigTree (http://tree.bio.ed.ac.uk/).
2. Maximum likelihood tree with bootstrapping of 1000 data sets – the multiple sequence alignment was submitted as a Nexus file (http://www.megasoftware.net/) to Garli (http://www.bio.utexas.edu/faculty/antisense/garli/garli.html). The resulting tree was annotated in FigTree (http://tree.bio.ed.ac.uk/).
3. Maximum Clade Credibility tree using BEAUti and BEAST programs (http://beast.bio.ed.ac.uk/). Model testing was carried out using the programmes jModelTest v2.1.4 (https://code.google.com/p/jmodeltest2/) and PAUP* (http://paup.csit.fsu.edu/). The tree was then analysed using Tracer v1.6.0 (http://beast.bio.ed.ac.uk/) and finally annotated in FigTree (http://tree.bio.ed.ac.uk/).

The complete algorithms for producing all trees are included in Appendix F.

2.5 Generation of Overlength Clones Containing PreS1 Mutations

2.5.1 Cloning Strategy Design

The A12C15_OL A1WTSA1.281MER was chosen as the backbone for cloning the deletion-mutants. This construct is a 1.28 overlength replication-competent HBV genome of subgenotype A1 within pcDNA4_TO backbone that was constructed using a novel method described by Bhoola *et al.* [193]. *Xba*I restriction sites were identified in the S region of HBV that flanked the preS deletions, and these sites were present in the same genomic position in the HBV overlength genome inserted into pcDNA4_TO vector (Figures 2.5 and 2.6).

Thus, the genomic region containing the mutations could be restricted and ligated into the overlength backbone in the equivalent position, to replace the WT. However, in addition to the two suitable *Xba*I sites in the HBV, the vector contained a third *Xba*I site. The latter *Xba*I site was removed by restriction with *Apa*I and *Not*I (Figure 2.6). The ends were blunted using the Lucigen End Repair Kit; the modified plasmid was re-ligated to become circular again and transformed to SCS110-competent cells. Full restriction mapping was carried out to confirm the integrity of this resulting plasmid.

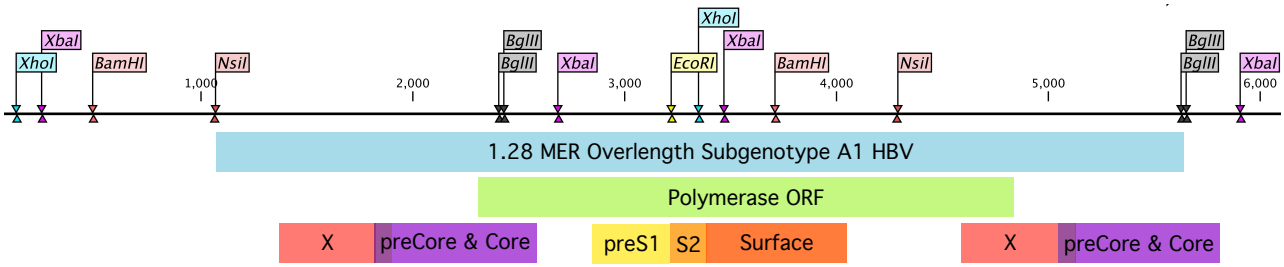


Figure 2.5: Layout of the Overlength A12C15_OL A1WTSA1.281MER Backbone and the ORFs of HBV Including Identified Restriction Sites

This image shows a scale representation of how the overlength genome of HBV is oriented. The black line above represents a two times the full genome and all the relevant restriction sites. Note the position of the two central *Xba*I sites, which were the sites that flank the Pres1 and PreS2 regions where deletions were identified. These were the *Xba*I sites that were used for inserting fragments containing deletions into the overlength backbone.

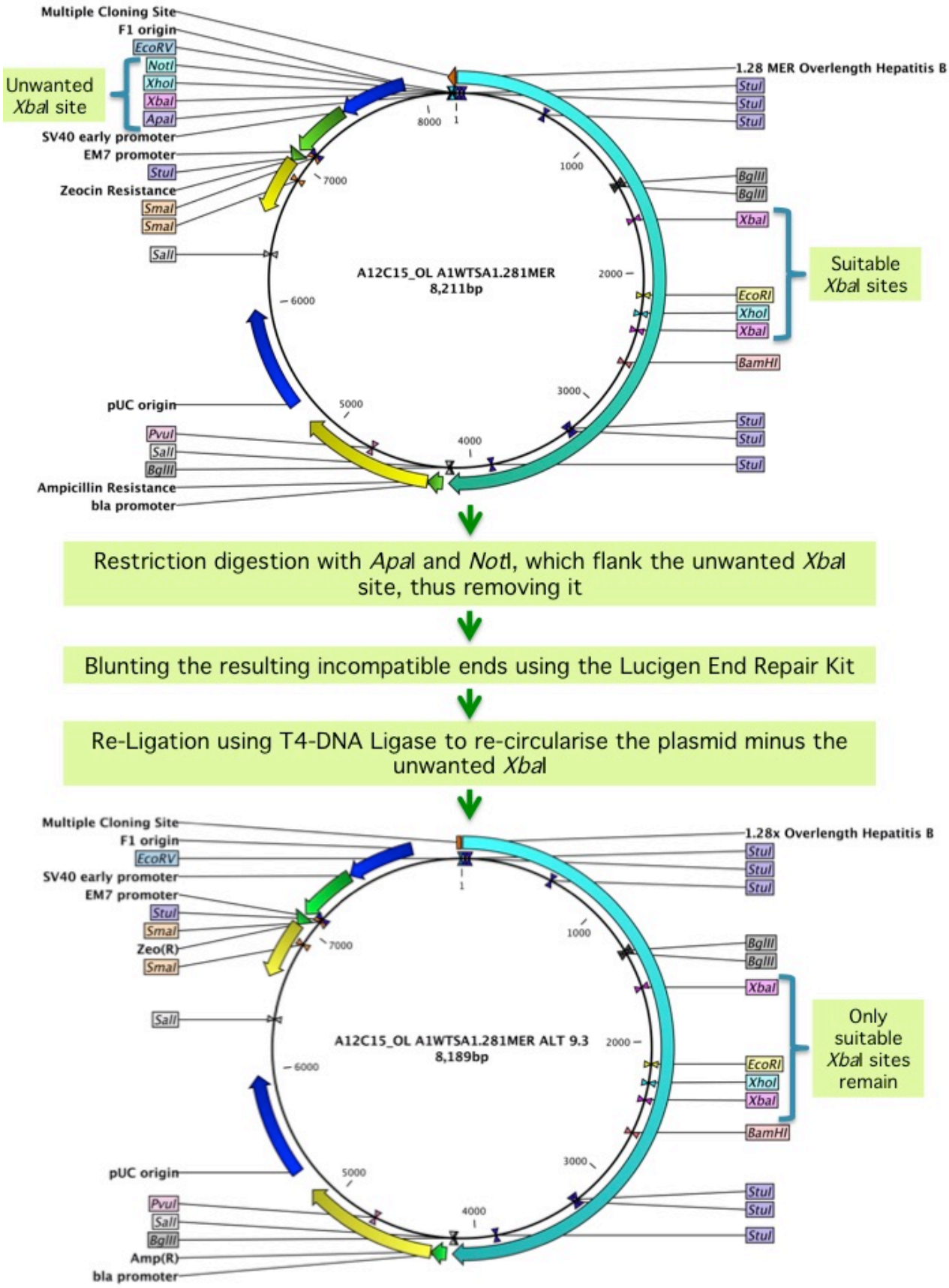


Figure 2.6: Algorithm to Alter the Backbone of the Plasmid for Subcloning Purposes

Clones of the deletion-mutants that were identical to those of the major population of HBV isolated from each patient (i.e. homologous to the parent strain) were chosen for subcloning into the modified plasmid using the identified *Xba*I sites (Figure 2.7 and 2.8). The resultant clones were transformed into SCS110-competent cells (See Appendix A 5.14). Colonies were screened by cracking and restriction for clones that contained the insert in the correct orientation (see Appendix A 5.1.5 and 5.1.9 respectively). Complete restriction mapping and sequencing were used to confirm correct insertion and orientation.

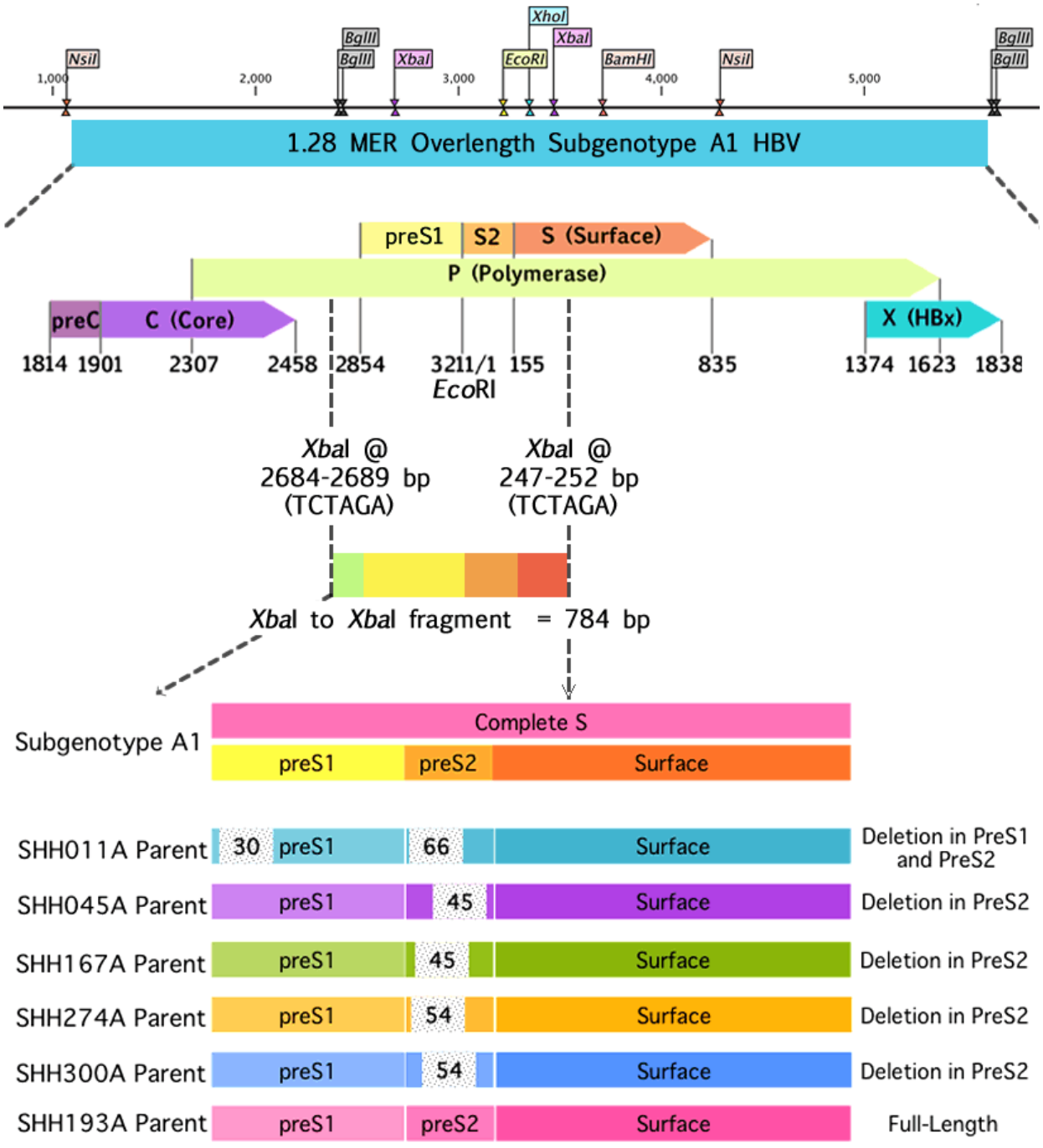
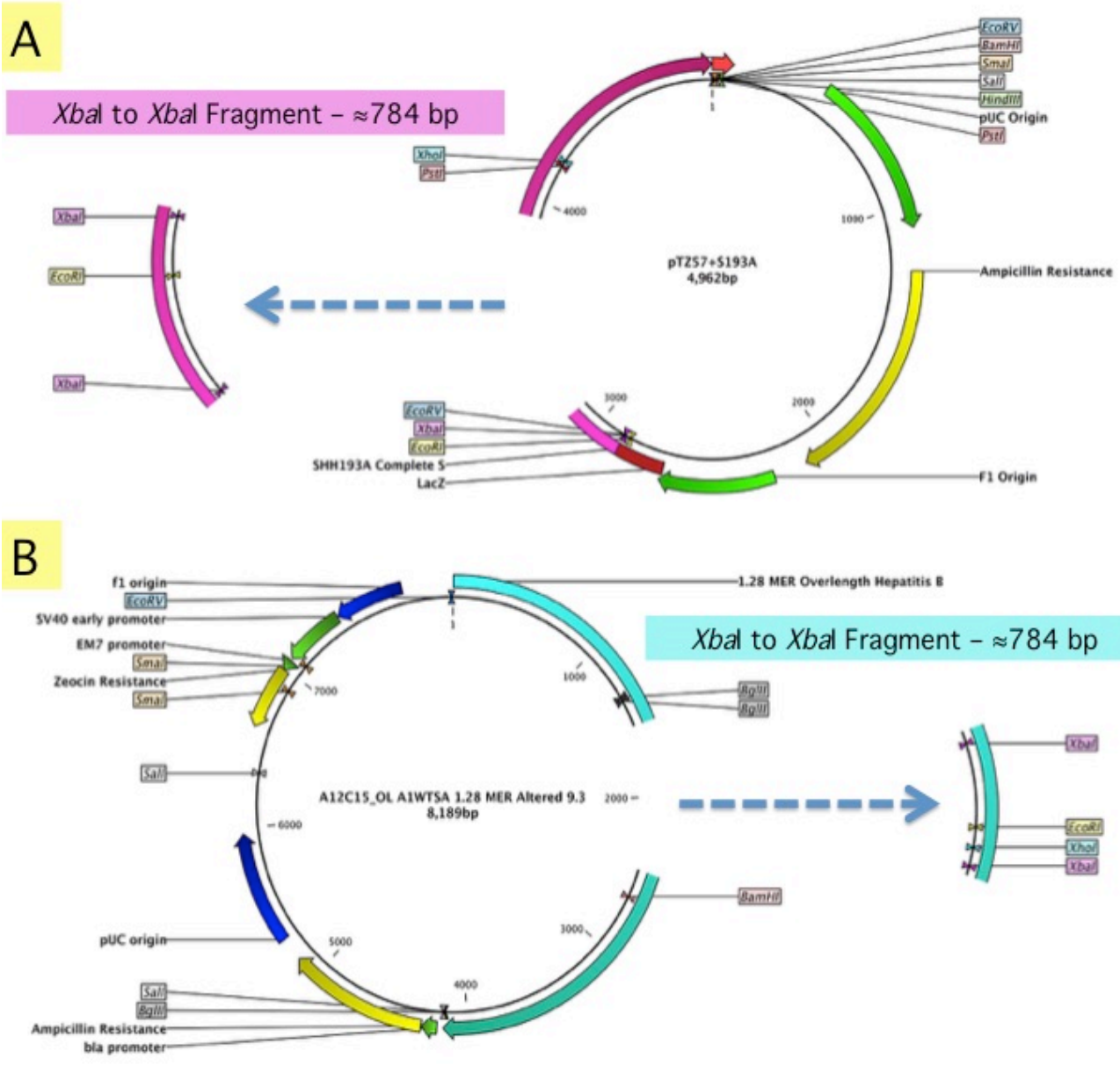


Figure 2.7: Algorithm for 784 bp Fragment Generation


A
XbaI to XbaI Fragment – ≈784 bp
EcoRV
BamHI
SmaI
SalI
HindIII
pUC Origin
PstI
XhoI
PstI
4000
1000
pT257+S193A
4,962bp
Ampicillin Resistance
XbaI
EcoRI
XbaI
3000
2000
EcoRV
XbaI
EcoRI
SHH193A Complete S
LacZ
F1 Origin
B
f1 origin
EcoRV
SV40 early promoter
EM7 promoter
SmaI
Zeocin Resistance
SmaI
1.28 MER Overlength Hepatitis B
1000
BglII
BglII
7000
SalI
A12C15_OL A1WTSA 1.28 MER Altered 9.3
8,189bp
2000
6000
XbaI
EcoRI
XhoI
XbaI
BamHI
3000
pUC origin
5000
4000
SalI
BglII
Ampicillin Resistance
bla promoter

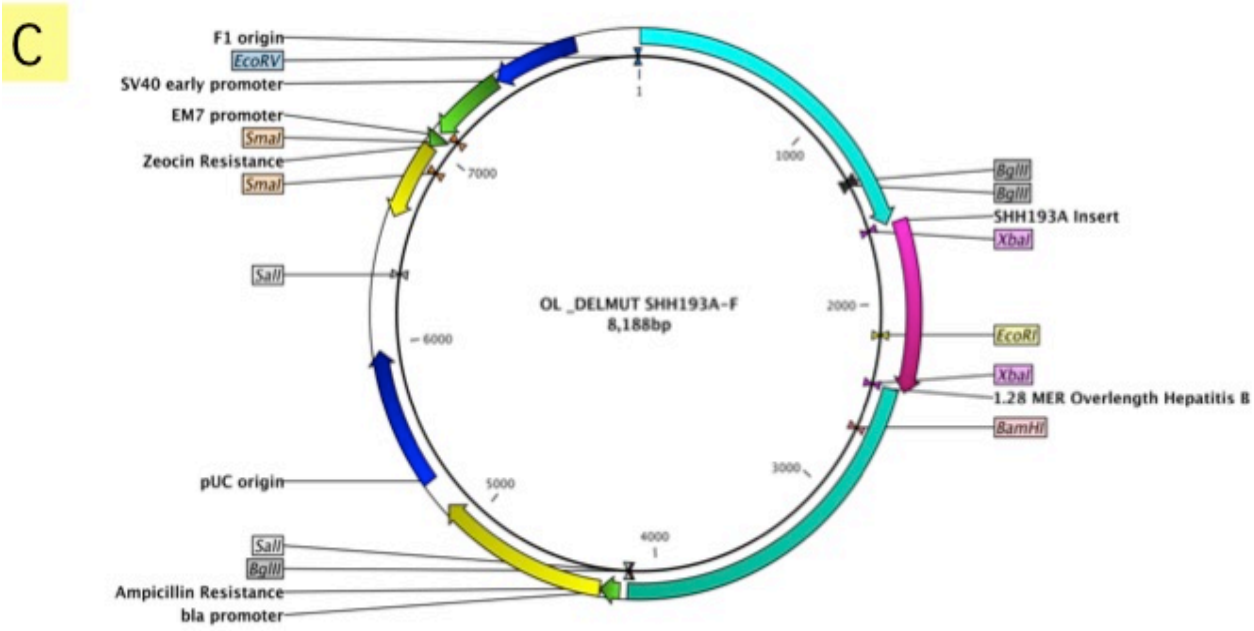


Figure 2.8: Identification of Suitable Restriction Sites and Cloning Workflow

Initially, a multiple sequence alignment was used to check whether all PCR amplified fragments contained the *Xba*I sites in the same genomic positions as the plasmid overlength backbone. Once it was confirmed that the *Xba*I sites were a suitable enzyme to use for cloning purposes, a cloning workflow was established. First PCR amplicons were cloned into pTZ57R/t plasmid, and once confirmed, the *XbaI* to *Xba*I fragment was restricted out of the appropriate clone (A), and similarly the fragment was removed from the A12C15_OL A1WTSA1.281MER Alt 9.3 backbone (B). Finally, the fragment identified in a patient isolate was ligated into the A12C15_OL A1WTSA1.281MER Alt 9.3 backbone (C), to produce an overlength clone containing the mutant fragment.

2.5.2 Replication Capacity of Clones

The nucleic acid sequences of the cloned final constructs were translated using GeneDoc and the Babylon tool [198], to determine whether the start and stop codons were still in-frame and to ensure correct expression of the viral proteins and to prevent the transfection of clones with artefactual mutations. Those constructs that were homologous to the parent strains were transiently transfected into an Huh7 cell line. These cells were grown in complete Dulbecco's Modified Eagle's Medium (DMEM) which contains 10% Foetal Calf Serum (FCS), L-Glutamine (2.92 mg/ml) with antibiotics Streptomycin (1 mg/ml) and Penicillin (1000 units/ml). Huh7 cells were grown in 75 cm^2 culture flasks until they were at 90% confluent, then they were either frozen for stocks, split/passaged or seeded into 12-well plates for transfection (detailed protocols in Appendix A 5.1.11, details for solutions Appendix B). Transfections were carried out in 12-well plates in triplicate, when cells were 90% confluent, with 2 μg of DNA and 5 μl of Polyethylenimine (PEI) transfection reagent. Transfection efficiency was calculated by

determining the percentage of cells successfully transfected with the pEGFP-c3 plasmid in a separate well. This plasmid contains Green Fluorescent Protein, which is self-fluorescent under 395 nm light, emitting light at a 509 nm wavelength (Figure 2.8). The percentage of cells transfected was calculated by counting the number of fluorescent cells, relative to the total number of cells.

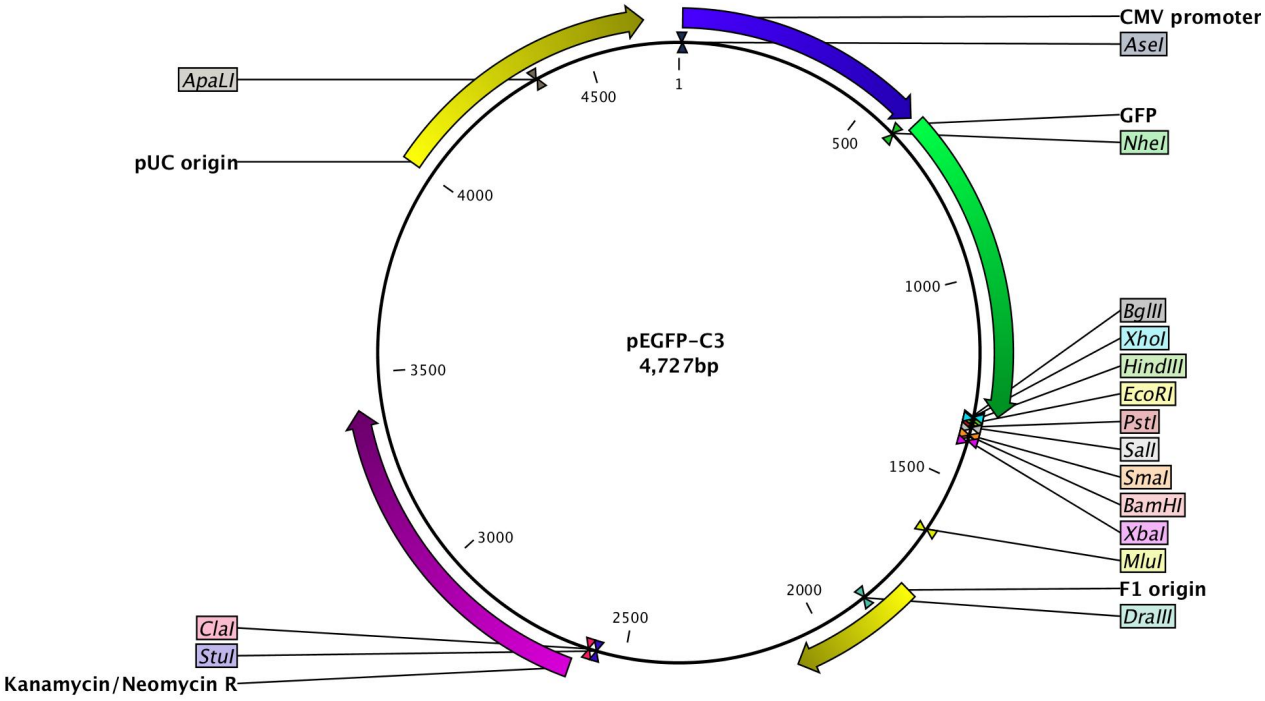


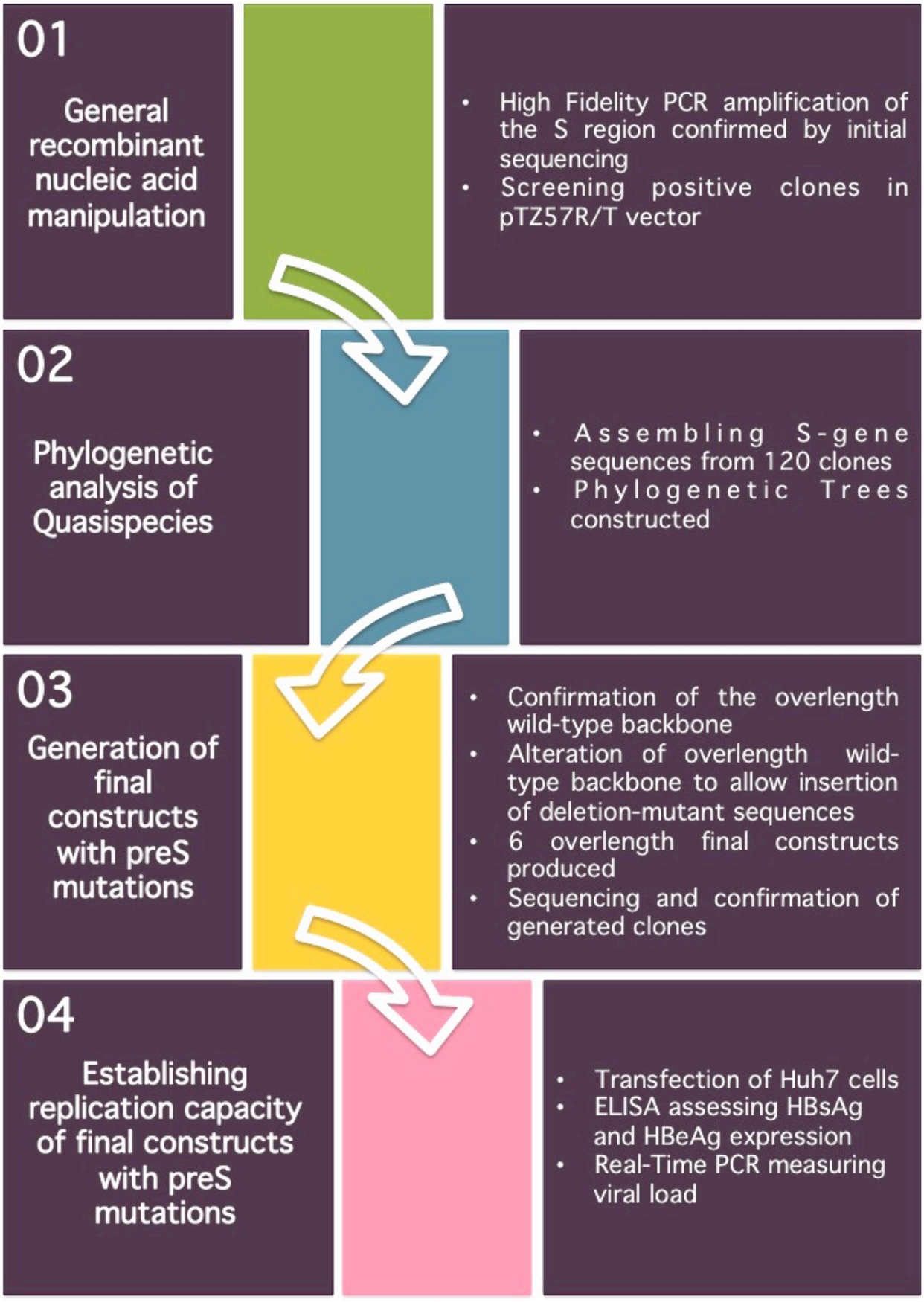
Figure 2.9: The GFP-Containing Plasmid Used to Determine Transfection Efficiency

Supernatant as well as the cells were collected on days 1, 3 and 5 post-transfection. For supernatants, a proteinase inhibitor cocktail (PIC) was added to preserve proteins and viral particles and then frozen at -70˚C. For harvesting of cells, a protocol was followed that allowed for the disruption of cell membranes and the addition of PIC for storage at -70˚C. The frozen samples were then thawed and used for ELISA and DNA extraction to be followed by qPCR quantification of VL. For DNA extraction, both supernatants and cell lysates were treated with DNaseI and RNaseA to remove any plasmid DNA, as this would falsely inflate VL detection. (Appendix A 5.1.11 - 5.1.17). ELISA was used to measure HBsAg and HBeAg expression after transfection. A plate was prepared where 75-100 µl of a 1:20 dilution of the supernatant was loaded into each well, in each case two biological replicates were added to the plate. The plate was only considered valid if it passed all the internal quality control specifications, and ODs were read at 620 and 450 nm. Samples were considered negative if the OD measure was

below the cut-off values, which was calculated by adding 0.05 to the mean of the kit provided negative control replicates. To analyse the data, the 620 nm reading was subtracted from the 450nm reading to get a 'blanked' result, and the average readings were normalised to the WT control. Furthermore, VL testing was done on the DNA extracted from both the supernatants and cell lysates via a SYBR Green qPCR strategy using primers designed for a conserved area of the BCP region of HBV. A serial dilution of plasmid pCH-9/3091 DNA was prepared from 10^9 copies decreasing to 10^3 copies to establish a standard curve. Only values with an R^2 value of 0.92 and above were considered significant, and efficiency of the run should be between 95% and 105%. For qPCR a sample was considered negative if the Sq values were lower than the threshold and/or the water negative control. Each qPCR plate was prepared with three biological replicates from three independent transfections as well as three technical replicates for each sample. Finally, an immunocapture technique using a Monolisa™ HBsAg ULTRA ELISA plate to capture secreted virions from the supernatant on day 3 post-transfection and then extracting DNA for qPCR was employed. Thus, this technique developed by Samal *et al.*, allows a method for the rapid quantification of VL of encapsidated DNA present in secreted virions only [207].

3 Chapter Three: Results

Overview and Workflow of the Study:



3.1 Quasispecies Analysis - General Recombinant Nucleic Acid Manipulation, Cloning and Phylogenetic Analysis

3.1.1 PCR Amplification of the Envelope (preS1/pre2/S) Region and Sequencing

Specific S region primers were used to amplify the complete S region from extracted DNA. Once amplification was successful with a Hot Start Taq protocol, the Roche Expand High Fidelity PCR System was used to obtain amplicons of high fidelity. These amplicons were used for downstream cloning (Figure 3.1). SHH011A and SHH045A failed to amplify despite multiple attempts using both the Hot Start and High Fidelity protocols, and serum for these samples was depleted. Thus, it was decided to use synthetic nucleotide fragments, with the deletions, for downstream cloning (refer to Figure 5.6 and 5.7 in Appendix C).

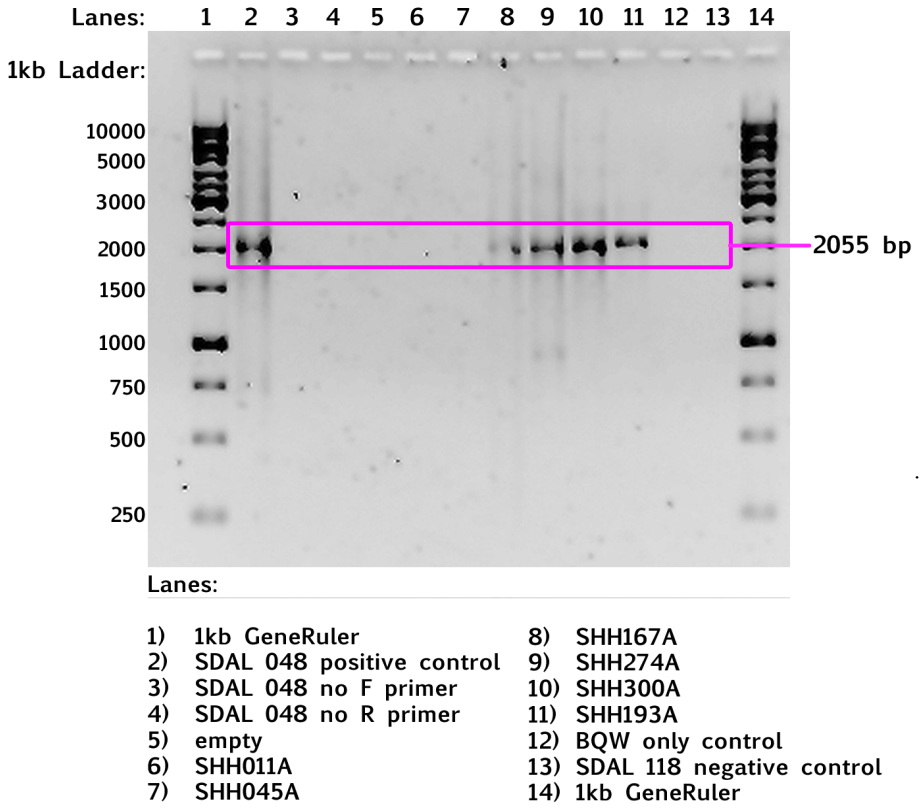


Figure 3.1: High Fidelity PCR of the S Region

PCR results of the second round of nested PCR with Roche Expand High Fidelity PCR System. The figure shows a 1% Agarose gel stained with ethidium bromide. Here templates from the first round of PCR were diluted 1 in 100 and expected band size is 2055 bp. Primers were S2F 3' AAT GTT AGT ATT CCT TGG ACT CAT AAG GTG GG 5' and S2R 3' AGT TCC GCA GTA TGG ATC GGC AGA GGA 5'. All controls were as expected. SHH011A and SHH045A were not amplified. This gel is representative of the High Fidelity PCR experiment, which was repeated a number of times to attain enough DNA for sequencing and cloning.

Before proceeding with purification and cloning, the amplicons were sent for Sanger sequencing to confirm the correct sequences for cloning (Figure 3.2).

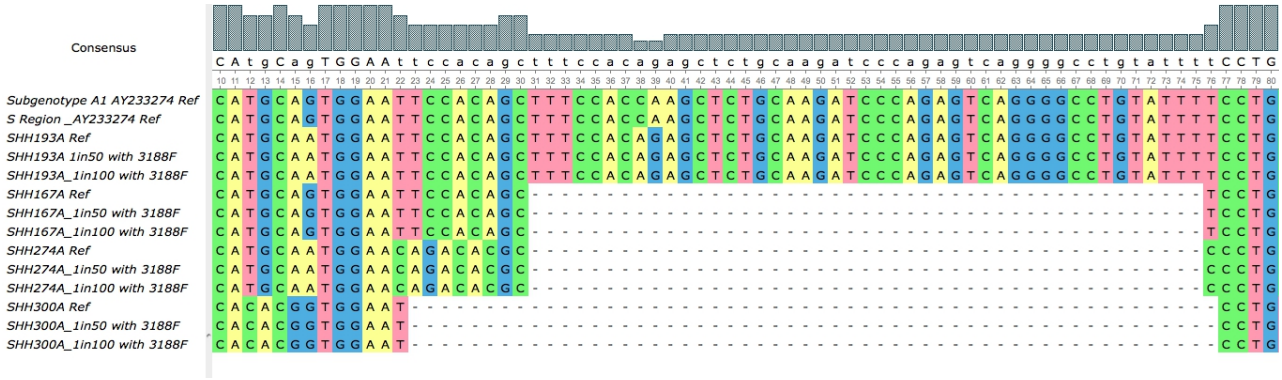


Figure 3.2: Alignment of PCR product against parent sequence

This small segment of the overall alignment shows that the Roche High Fidelity PCR has reproduced the sequences identically, and thus these sequences were taken forward to cloning. This section shows the results obtained when the 3188F sequencing primer was used and similar results were obtained with both the 2497F and 519F primers. Note the figure shows both a 1 in 50 and a 1 in 100 dilution product of the second round of the nested PCR.

3.1.2 Cloning Amplicons and Sequencing Clones

Following transformation with pTZ57R/T containing the amplicons, white colonies were picked randomly from plates and screened. Cracking and restriction digest were used to identify positive clones (see Figure 3.3). In total 257 clones were produced 75 for SHH167A, 78 for SHH193A, 43 for SHH274A and 61 for SHH300A. All of these clones represent the pool from which quasispecies were analysed by phylogenetic comparison.

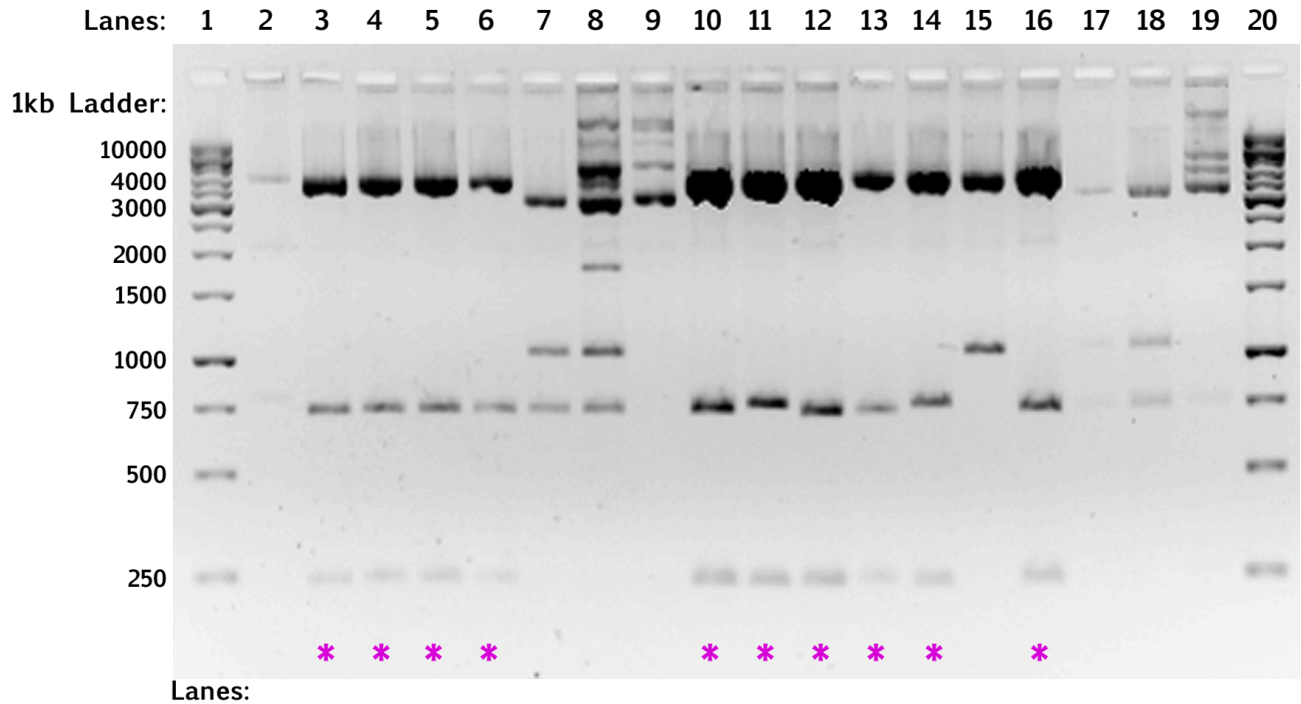


Figure 3.3: Screening for clones using restriction by *Xba*I

This gel is representative of screening by restriction digestion with *Xba*I. For pTZ57R/T+SHH193A the bands indicating a clone are 3935 bp, 783 bp and 244 bp; for pTZ57R/T+SHH274A the bands indicating a clone are 3916 bp, 739 bp and 244 bp; for pTZ57R/T+SHH300A the bands indicating a clone are 3916 bp, 730 bp and 244 bp. Not shown here, bands for pTZ57R/T+SHH167A would be 3916 bp, 739 bp and 244 bp. On this gel, all lanes marked with an asterisk show the correct bands and thus were counted as clones. This figure depicts a 1% Agarose gel, stained with ethidium bromide.

3.1.3 Phylogenetic Analysis:

Using a random number generator (https://www.random.org/), 30 clones of HBV from each patient were sequenced. Fragments produced by the three sequencing primers were curated and then merged using the Fragment Merger tool [197]. A multiple sequence alignment was constructed containing the 4 parent sequences, sequences of the 120 (4 x 30) clones. HBV subgenotype F was used as an outlier to root for the tree, and 19 subgenotype A1 sequences for comparison. Both inter- and intra-group nucleotide divergences were calculated for the clones and their parental sequences (Table 3.1). A summary of molecular characteristics of the cloned sequences from the multiple sequence alignment is detailed in Table 3.2. While the parental sequence was found to represent the majority population within the quasispecies sequences, subpopulations of different sequences were identified.

Table 3.1: Inter- and Intra- Group Divergence Within Quasispecies and the Parental Strains

	SHH167A	SHH193A	SHH274A	SHH300A
SHH167A	**11±0.07**			
SHH193A	11±0.05	**4±0.03**		
SHH274A	12±0.05	10±0.01	**2±0.02**	
SHH300A	12±0.05	10±0.01	10±0.01	**4±0.01**

Mean nucleotide divergence in percentage was calculated using DAMBE (Data Analysis in Molecular Biology and Evolution) software [208]. Sequence divergence is reported as percentages of the mean divergence ± the standard deviation where intra-group divergence is shaded in yellow. The smaller the divergence, the more similar sequences are to each other.

Table 3.2: Molecular Characteristics of Clones Relative to the Parental Strain

Molecular Characteristics of Quasispecies				
Sequence	Sequence characteristic	Position of Indel	Region	Frequency no., (%) in 30 clones
Parent SHH167A	45 nt deletion	9 - 54	PreS2	N/A
SHH167A 1.70	30 nt & 66 nt deletions overlapping the 45 nt deletion	2899 – 2929 & 3212 -56	PreS1	1 (3.3%)
SHH167A 1.77	63 nt deletion overlapping the 45 nt deletion	2 - 64	PreS1	1 (3.3%)
SHH167A 2.134	19 nt deletion instead of 45 nt deletion	3201 - 3219	PreS1-PreS2	1 (3.3%)
SHH167A 1.121	183 nt deletion instead of 45 nt deletion	3021 - 3209	PreS1	1 (3.3%)
SHH167A 1.12	224 nt deletion instead of 45 nt deletion	127 - 350	Surface	1 (3.3%)
SHH167A 1.3	134 nt deletion instead of 45 nt deletion	3020 - 3154	PreS1	1 (3.3%)
SHH167A 1.133 SHH167A 2.22 SHH167A 1.69	No 45 nt deletion	-	-	3 (10%)
Remaining Clones	45 nt deletion	9 - 54	PreS2	21 (70,2%)
Parent SHH193A	Full-length	-	-	N/A
SHH193A 1.147 SHH193A 2.143	1 nt deletion	608 2911	Surface PreS1	1 (3.3%)
SHH193A 2.141 SHH193A 1.121 SHH193A 1.150 SHH193A 2.146 SHH193A 2.144 SHH193A 2.122	54 nt deletion	3 - 56	Surface	6 (20%)
Remaining Clones	Full-length	-	-	23 (76.7%)
Parent SHH274A	54 nt deletion	2 - 55	PreS2	N/A
SHH274A 2.21	96 nt deletion	3120 - 3215	PreS1	1 (3.3%)
SHH274A 1.11	101 nt deletion overlapping the 54 nt deletion	3167 - 46	PreS1-PreS2	1 (3.3%)
SHH274A 1.25	1 nt deletion	322	Surface	1 (3.3%)
Remaining Clones	54 nt deletion	2 - 55	PreS2	27 (90%)
Parent SHH300A	54 nt deletion	3 - 56	PreS2	96.6%
SHH300A 2.14	21 nt deletion instead of the 54 nt deletion	35 - 56	PreS2	1 (3.3%)
SHH300A 2.109	1 nt deletion	733	Surface	1 (3.3%)
SHH300A 1.125	1 nt deletion	286	Surface	1 (3.3%)
Remaining Clones	54 nt deletion	3 - 56	PreS2	27 (90%)

Initially, the sequences of the clones were compared using a Neighbor-Joining tree, and a few reference sequences to establish how the major and minor populations would cluster (Figure 3.4). Reference sequences were obtained from GenBank, and were trimmed and aligned such that only the PreS1/PreS2/Surface section is shown.

The grouping and relative close relatedness of these sequences suggest that the deletion-mutants represent the major population of viral quasispecies circulating in each of the patients. Only one of the quasispecies, SHH167A 2.134 showed no deletion whereas the parental strain showed a 45 nt deletion. However, it also contained a separate 19 nt deletion in a different location. Interestingly, six clones derived from patient SHH193A (with the full-length parental strain) showed a novel deletion of 54 nt in a similar position as the parental strains from patients SHH274A and SHH300A.

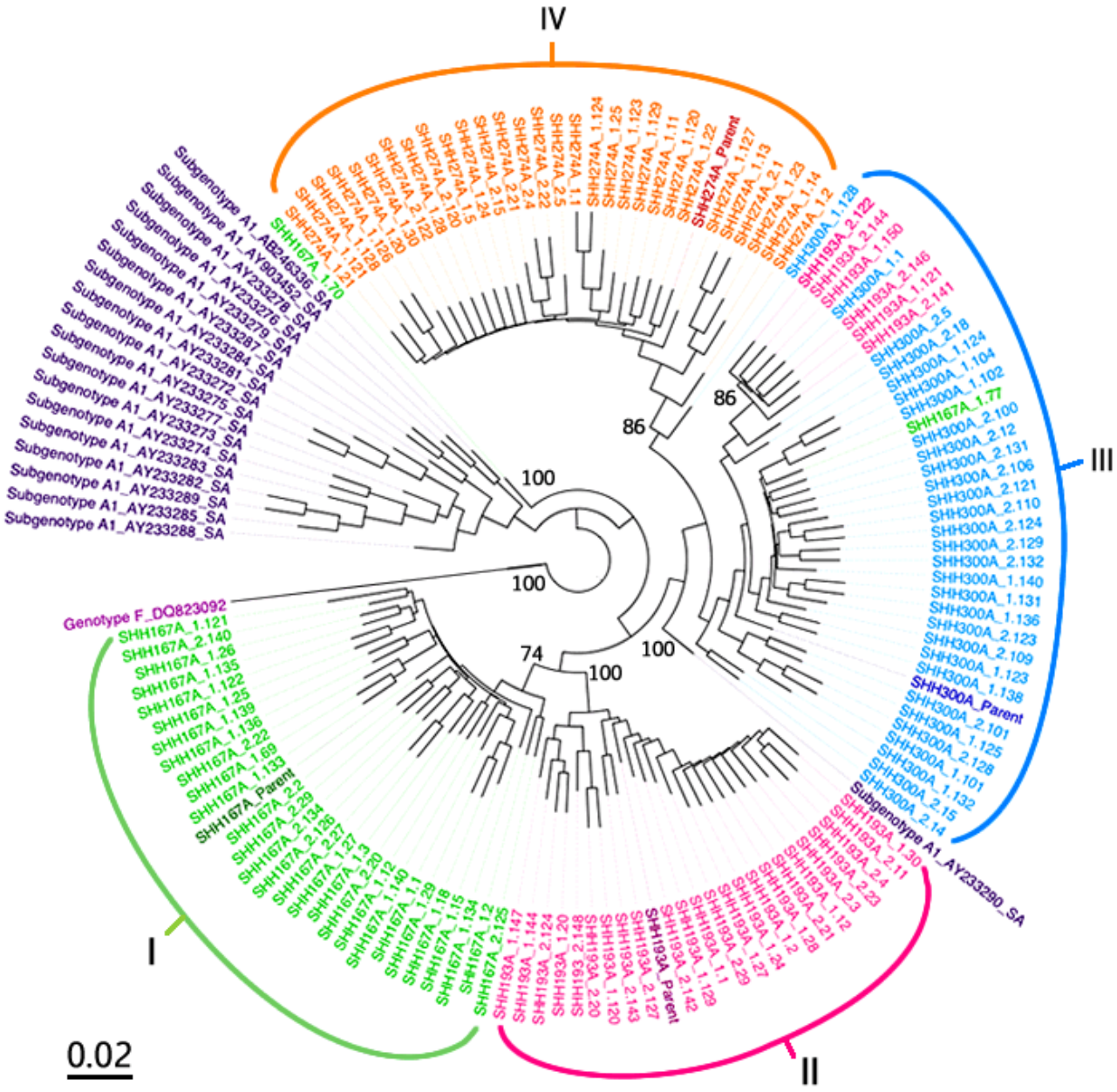


Figure 3.4: Neighbor-Joining Phylogenetic Tree

This tree demonstrates the phylogenetic similarity of complete PreS1/PreS2/Surface sequences of cloned quasispecies from four patients, established using the Neighbor-Joining method. Bootstrap statistical analysis was included to 1000 replicates, values above 65% are considered significant, some values are indicated as a percentage. All sequences obtained from GenBank are designated by their genotype, accession number and an acronym for the country of origin. Genotype F was used as the outlier to root the tree (light purple), and 19 Subgenotype A1 sequences from South Africa (dark purple) were included for comparison (detailed in Table 5.12, Appendix F). Branch lengths represent the amount of genetic change over time to scale of substitutions per site. Most sequences from patient SHH167A cluster together in group I, including four clones (SHH167A 2.134; SHH167A 1.133; SHH167A 2.22; SHH167A 1.69) that have no deletion. The SHH167A_1.77 displayed a 63 nt deletion which more closely resembled the deletion pattern of the SHH300A parent, and thus clustered with group III. SHH167A_1.70 contained extra novel deletions of 30 nt and 66 nt, which resulted in this sequence not clustering with any of the others. Most of the WT group from patient SHH193A clustered together in group II except 6 clones that had a novel 54 nt deletion which cluster with SHH300A patient sequences in group III. Refer back to Table 3.2 for the complete list of novel mutations.

Once it was established how the quasispecies relate to each other and the reference PreS1/PreS2/Surface sequences, it was decided to carry out additional phylogenetic analysis of the quasispecies to the PreS1/PreS2/Surface sequences representative of all 9 HBV genotypes. This tree was again constructed using Neighbor-Joining statistical methods and thus depicts evolutionary relatedness of the PreS1/PreS2/Surface sequences by minimum evolution, or maximum parsimony (Figure 3.5) [209].

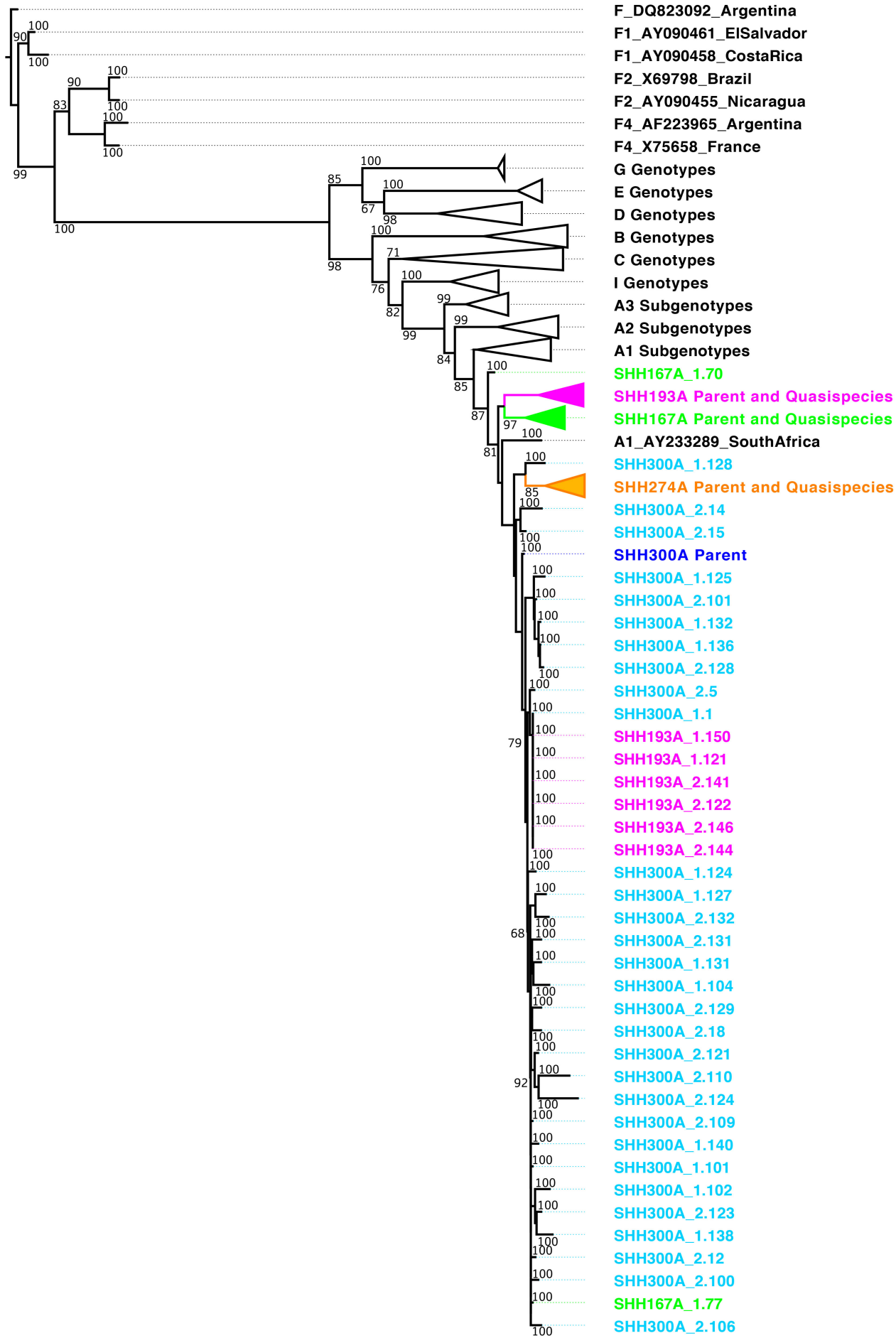

F_DQ823092_Argentina
F1_AY090461_ElSalvador
F1_AY090458_CostaRica
F2_X69798_Brazil
F2_AY090455_Nicaragua
F4_AF223965_Argentina
F4_X75658_France
G Genotypes
E Genotypes
D Genotypes
B Genotypes
C Genotypes
I Genotypes
A3 Subgenotypes
A2 Subgenotypes
A1 Subgenotypes
SHH167A_1.70
SHH193A Parent and Quasispecies
SHH167A Parent and Quasispecies
A1_AY233289_SouthAfrica
SHH300A_1.128
SHH274A Parent and Quasispecies
SHH300A_2.14
SHH300A_2.15
SHH300A Parent
SHH300A_1.125
SHH300A_2.101
SHH300A_1.132
SHH300A_1.136
SHH300A_2.128
SHH300A_2.5
SHH300A_1.1
SHH193A_1.150
SHH193A_1.121
SHH193A_2.141
SHH193A_2.122
SHH193A_2.146
SHH193A_2.144
SHH300A_1.124
SHH300A_1.127
SHH300A_2.132
SHH300A_2.131
SHH300A_1.131
SHH300A_1.104
SHH300A_2.129
SHH300A_2.18
SHH300A_2.121
SHH300A_2.110
SHH300A_2.124
SHH300A_2.109
SHH300A_1.140
SHH300A_1.101
SHH300A_1.102
SHH300A_2.123
SHH300A_1.138
SHH300A_2.12
SHH300A_2.100
SHH167A_1.77
SHH300A_2.106
0.02

Figure 3.5: Neighbor-Joining Phylogenetic Tree with all the Genotypes

Phylogenetic analysis of complete PreS1/PreS2/Surface sequences of cloned quasispecies from four patients to 105 complete PreS1/PreS2/Surface sequences representative of all 9 HBV genotypes. This analysis was established using the Neighbor-Joining method. Bootstrap statistical analysis was included to 1000 replicates, values above 65% are considered significant, these are indicated on the branches as a percentage. All sequences obtained from GenBank are designated by their genotype, accession number and country of origin, a full list of the sequences used are available in Table 5.12, Appendix F. Genotype F_DQ823092 from Argentina was used as the outlier to root the tree; the tree is presented in descending order. This tree was generated using the Pipeline Treemail tool, the tree was viewed in FigTree [104].

While the Neighbor-Joining method is known for its efficiency in obtaining a correct tree, it also assumes the smallest amount of total evolutionary change and constant rate of nucleotide substitution [209]. For comparison, a Maximum Clade Credibility Tree using Bayesian Markov Chain Monte Carlo (MCMC) inference method was also constructed, to see whether using this statistical model would show a different evolutionary relationship between the PreS1/PreS2/Surface sequences (Figure 3.6) [210]. This method uses prior knowledge (phylogeny distribution, statistical models and substitution rates) to evaluate the likelihood of a given tree [210]. Therefore, this method infers whether a tree is likely, based on the observations made [210]. The posterior probability is calculated by measuring the proportion of times that a tree is visited by the inferential chain [210]. Thus the closer to 1 the posterior probability, the better the evidence is for the placement of each node [210].



Figure 3.6: Maximum Clade Credibility Tree with all the Genotypes

Bayesian Markov Chain Monte Carlo (MCMC) inference method, comparing complete PreS1/PreS2/Surface sequences of cloned quasispecies from four patients to 105 complete PreS1/PreS2/Surface sequences representative of all 9 HBV genotypes. The genotype, accession number and country of origin for all designated sequences was obtained from GenBank, a full list of the sequences used are available in Table 5.12, Appendix F. A relaxed molecular clock was used for analysis and the generalised time-reversible (GTR) substitution model and a gamma distribution model was employed. The MCMC analysis was run for 100 000 000 states, the Effective Sample Size (ESS) was 190, and their uncertainties were represented in the 95% Highest Posterior Density intervals. Branch lengths represent a 2% divergence as per the scale. The tree is presented in descending order. The tree was generated using BEAUti and BEAST; and analysis was conducted in Tracer and FigTree [103-106, 201, 210].

3.2 Generation of Overlength Replication-Competent Constructs with PreS Mutations

3.2.1 Initial Restriction Mapping of the A12C15_OL A1WTSA1.281MER Construct

This construct was originally generated by Dr. N. Bhoola [193]. To confirm the integrity of this plasmid, full restriction mapping was carried out (Figures 3.7 and 3.8 plus Table 3.3)

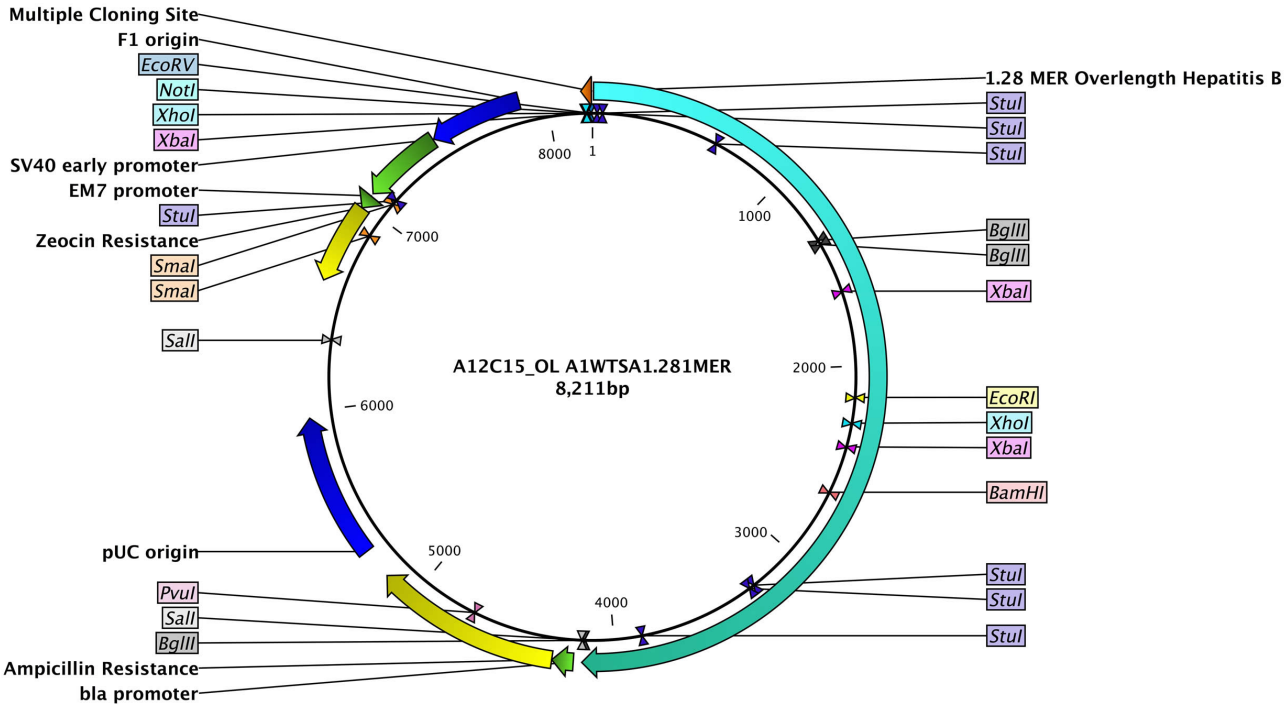
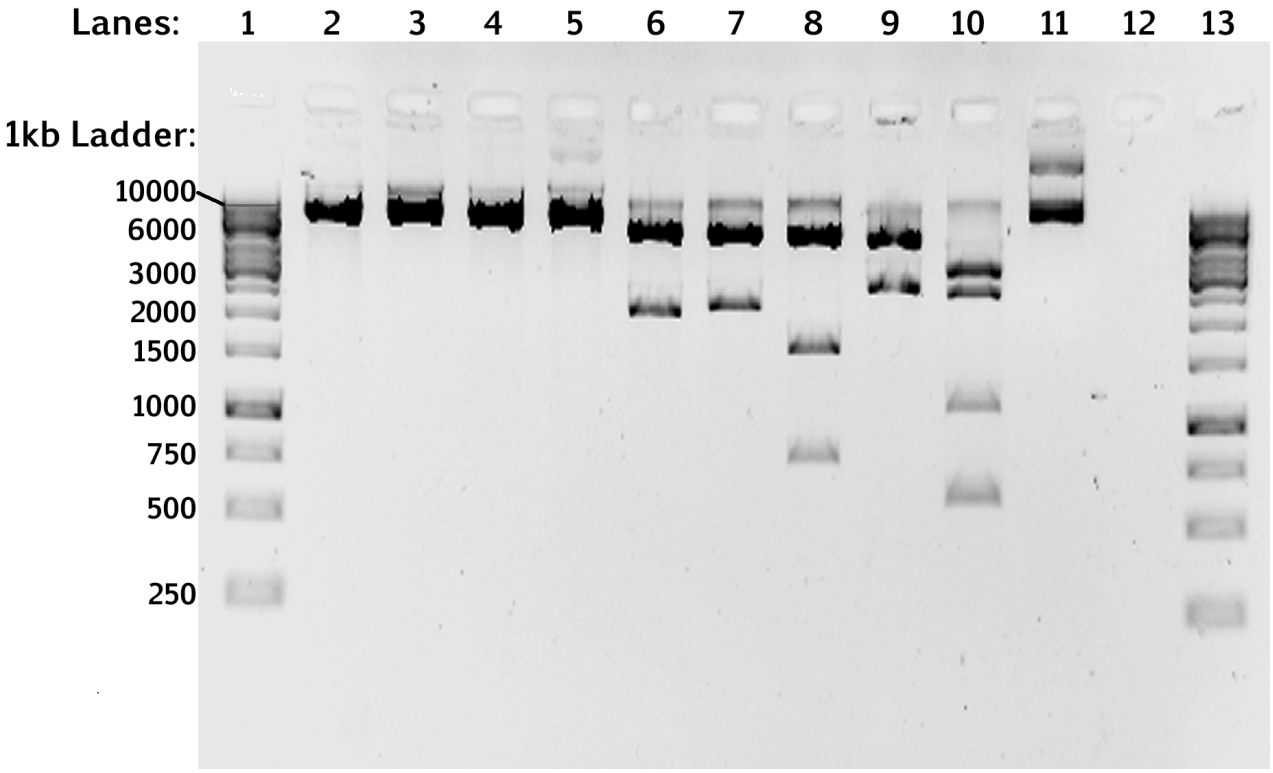


Figure 3.7: Plasmid Map of the A12C15_OL A1WTSA1.281MER Vector

This plasmid map shows the organisation and relevant restriction sites in this construct.

Table 3.3: Summary of Restriction Enzymes Utilised and Expected Band Sizes

Enzyme	No of cuts	Size of expected fragments for correct orientation
*Pvu*I	1	8211 bp linear
*Not*I	1	8211 bp linear
*BamH*I	1	8211 bp linear
*EcoR*I	1	8211 bp linear
*Sal*I	2	6026 bp and 2185 bp
*Xho*I	2	5895 bp and 2316 bp
*Xba*I	3	5769 bp, 1658 bp and 784 bp
*Bgl*II	3	5411 bp, 2776 bp and 24 bp
*Stu*I	7	3257 bp, 2607 bp, 1119 bp, 587bp, 587 bp. 27 bp and 27 bp



Lanes:

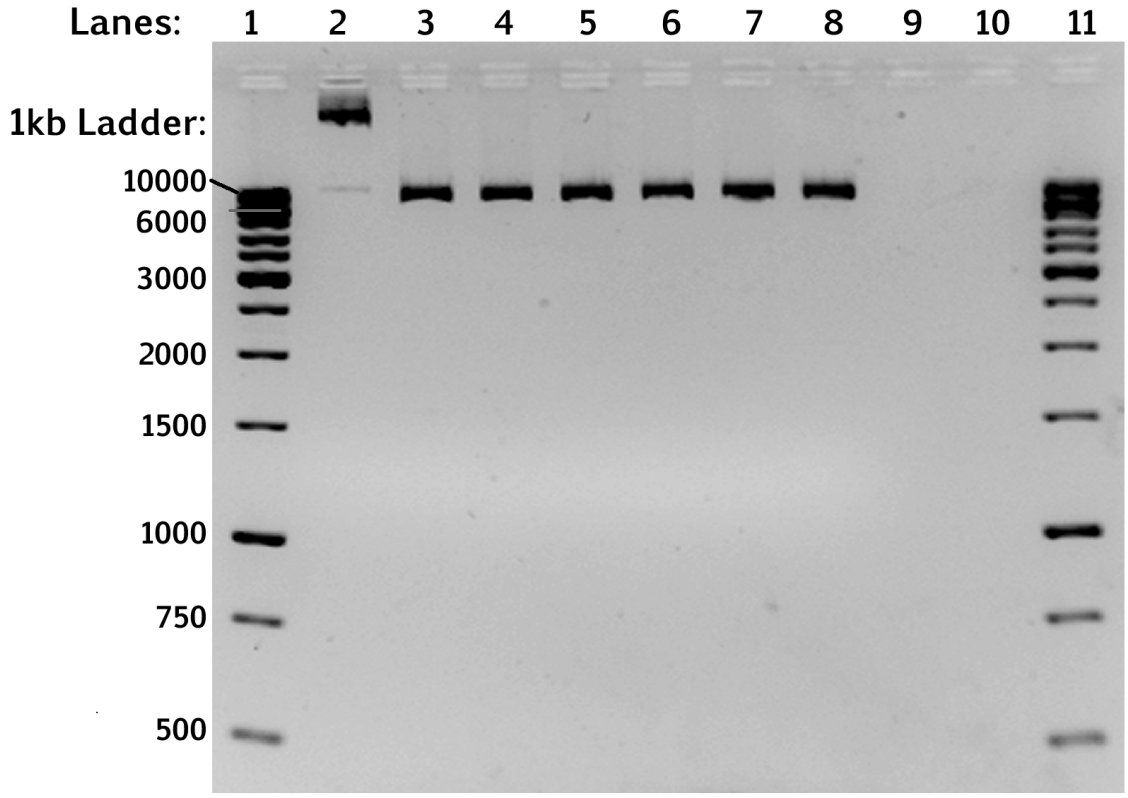
1) GeneRuler™ 1kb DNA Ladder
2) A12C15_OL A1WTSA 1.28 MER + *Pvu*I
3) A12C15_OL A1WTSA 1.28 MER + *Not*I
4) A12C15_OL A1WTSA 1.28 MER + *BamH*I
5) A12C15_OL A1WTSA 1.28 MER + *EcoR*I
6) A12C15_OL A1WTSA 1.28 MER + *Sal*I
7) A12C15_OL A1WTSA 1.28 MER + *Xho*I
8) A12C15_OL A1WTSA 1.28 MER + *Xba*I
9) A12C15_OL A1WTSA 1.28 MER + *Bgl*II
10) A12C15_OL A1WTSA 1.28 MER + *Stu*I
11) Uncut A12C15_OL A1WTSA 1.28 MER
12) No DNA + *Xba*I
13) GeneRuler™ 1kb DNA Ladder

Figure 3.8: Restriction Mapping of A12C15_OL A1WTSA1.281MER Vector

The resulting restriction map is considered successful as all expected bands were present and were also confirmed to be the correct size using a standard curve graph. Distance travelled was plotted against the log of the molecular weight; band distances from the well were substituted back into the line equation and compared to actual sizes. The uncut plasmid and a no DNA control were included to show enzyme activity was as expected. This figure represents a 1% Agarose gel, stained with ethidium bromide.

3.2.2 Alteration of the A12C15_OL A1WTSA1.281MER

The A12C15_OL A1WTSA1.281MER was successfully altered for the purposes of cloning. After removing the third unwanted *Xba*I site with *Apa*I and *Not*I, the ends were blunted using the Lucigen End Repair Kit (Figure 3.9).



Lanes:

1)	GeneRuler™ 1kb DNA Ladder
2)	Uncut A12C15_OL A1WTSA 1.28 MER
3) - 6)	A12C15_OL A1WTSA 1.28 MER #9 + *Apa*I and *Not*I
4)	A12C15_OL A1WTSA 1.28 MER #9 + *Apa*I only
5)	A12C15_OL A1WTSA 1.28 MER #9 + *Not*I only
6)	Empty
10)	No DNA control (water + *Apa*I and *Not*I)
11)	GeneRuler™ 1kb DNA Ladder

Figure 3.9: Restriction of A12C15_OL A1WTSA1.281MER with *Apa*I and *Not*I

Restriction with these enzymes removed the undesired *Xba*I site from the vector backbone. After restriction, the total length is 8189 bp as opposed to the 8211 bp it was before the restriction, as 22 bp were removed around the undesired *Xba*I site. This restriction also eliminated a second *Xho*I site. The uncut plasmid and a no DNA control were included to show enzyme activity was as expected. This figure represents a 1% Agarose gel, stained with Novel Juice.

After blunting the restricted ends and recircularising the altered plasmid, full restriction mapping was carried out on the resulting plasmid (Figures 3.10 and 3.11 as well as Table 3.4).

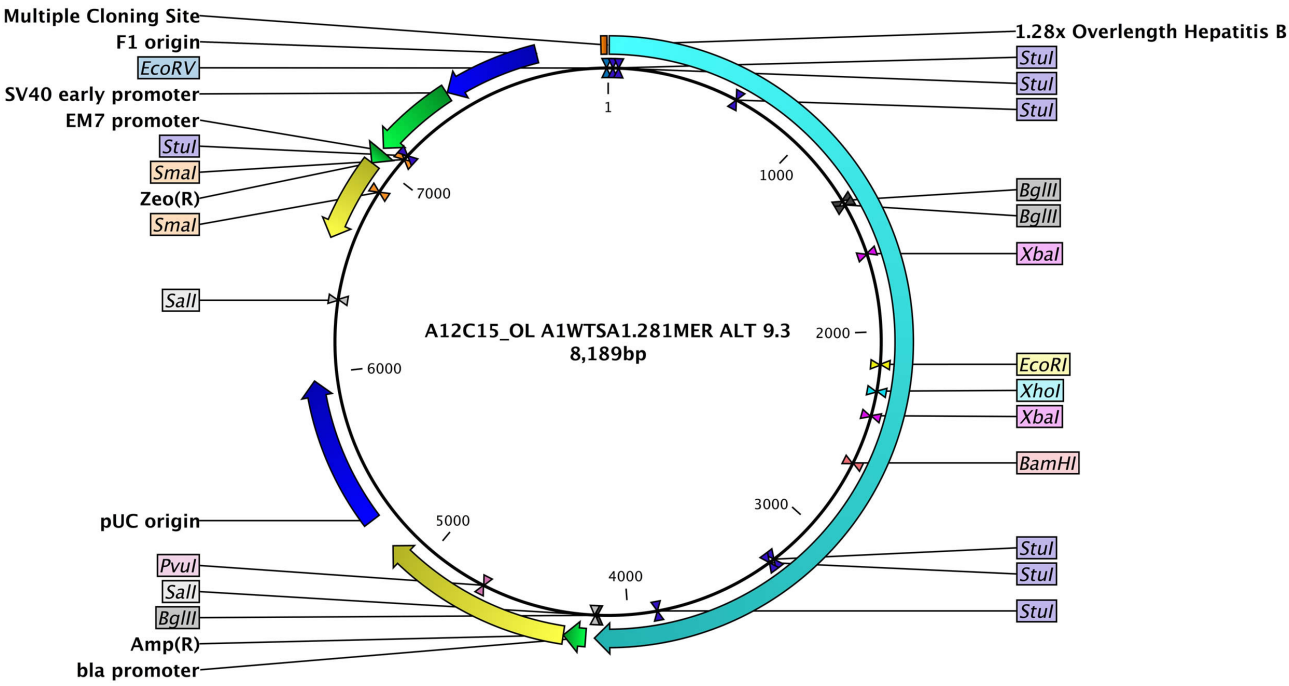
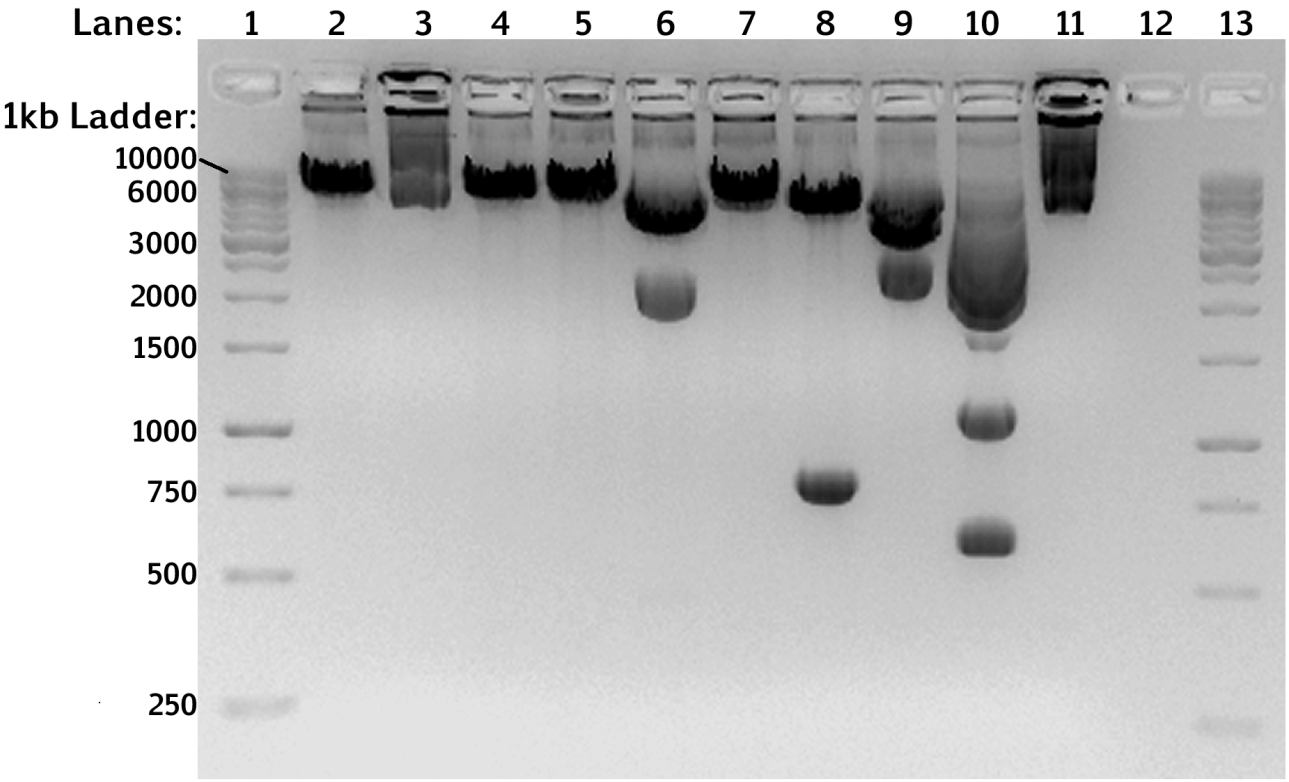


Figure 3.10: Plasmid Map of A12C15_OL A1WTSA1.281MER Altered 9.3 Vector

This plasmid map shows the organisation and relevant restriction sites of the now altered construct. Thus, in this construct, there is one less *Xba*I site and one less *Xho*I site as result of the restriction with *Apa*I and *Not*I. Also as a result of the restriction with *Apa*I and *Not*I, this construct is 22 bp shorter than the A12C15_OL A1WTSA1.281MER. The two remaining *Xba*I sites were used for subcloning.

Table 3.4: Summary of Restriction Enzymes Utilised and Expected Band Sizes

Enzyme	No of cuts	Size of expected fragments for correct orientation
*Pvu*I	1	8189 bp linear
*Not*I	0	8189 bp circular (*Not*I site was removed with alteration)
*EcoR*I	1	8189 bp linear
*BamH*I	1	8189 bp linear
*Sal*I	2	6004 bp and 2185 bp
*Xho*I	1	8189 bp
*Xba*I	2	7427 bp, 784 bp
*Bgl*II	3	5389 bp, 2776 bp and 24 bp
*Stu*I	7	3257 bp, 2607 bp, 1097 bp, 587bp, 587 bp. 27 bp and 27 bp



Lanes:

1) GeneRuler™ 1kb DNA Ladder
2) A12C15_OL A1WTSA 1.28 MER Alt 9.2 + *Pvu*I
3) A12C15_OL A1WTSA 1.28 MER Alt 9.3 + *Not*I
4) A12C15_OL A1WTSA 1.28 MER Alt 9.3 + *BamH*I
5) A12C15_OL A1WTSA 1.28 MER Alt 9.3 + *EcoR*I
6) A12C15_OL A1WTSA 1.28 MER Alt 9.3 + *Sal*I
7) A12C15_OL A1WTSA 1.28 MER Alt 9.3 + *Xho*I
8) A12C15_OL A1WTSA 1.28 MER Alt 9.3 + *Xba*I
9) A12C15_OL A1WTSA 1.28 MER Alt 9.3 + *Bgl*II
10) A12C15_OL A1WTSA 1.28 MER Alt 9.3 + *Stu*I
11) Uncut A12C15_OL A1WTSA 1.28 MER Alt 9.3
12) No DNA + *Xba*I
13) GeneRuler™ 1kb DNA Ladder

Figure 3.11: Restriction Mapping of A12C15_OL A1WTSA1.281MER Altered 9.3

The resulting restriction map is considered successful as all expected bands were present and were also confirmed to be the correct size using a standard curve graph. Distance travelled was plotted against the log of the molecular weight; band distances from the well were substituted back into the line equation and compared to actual sizes. The uncut plasmid and a no DNA control were included to show enzyme activity was as expected. This figure represents a 1% Agarose gel, stained with ethidium bromide.

3.2.3 Production of Final Overlength Constructs

A small subset of pTZ57R/T clones were identified for subcloning; this was done on the basis that the sequences spanning the two *Xba*I sites were most homologous to the parental sequence. Thus, the fragment containing the mutations could be ligated into the A12C15_OL A1WTSA1.28 MER Altered 9.3 overlength backbone for functional analysis. All final constructs had to be screened carefully to ensure the section fragment containing the deletion was oriented in the forward direction. The sequences of the final constructs were confirmed by complete restriction mapping and by Sanger Sequencing (Figures 3.12 – 3.19 and Tables 3.5 – 3.8).

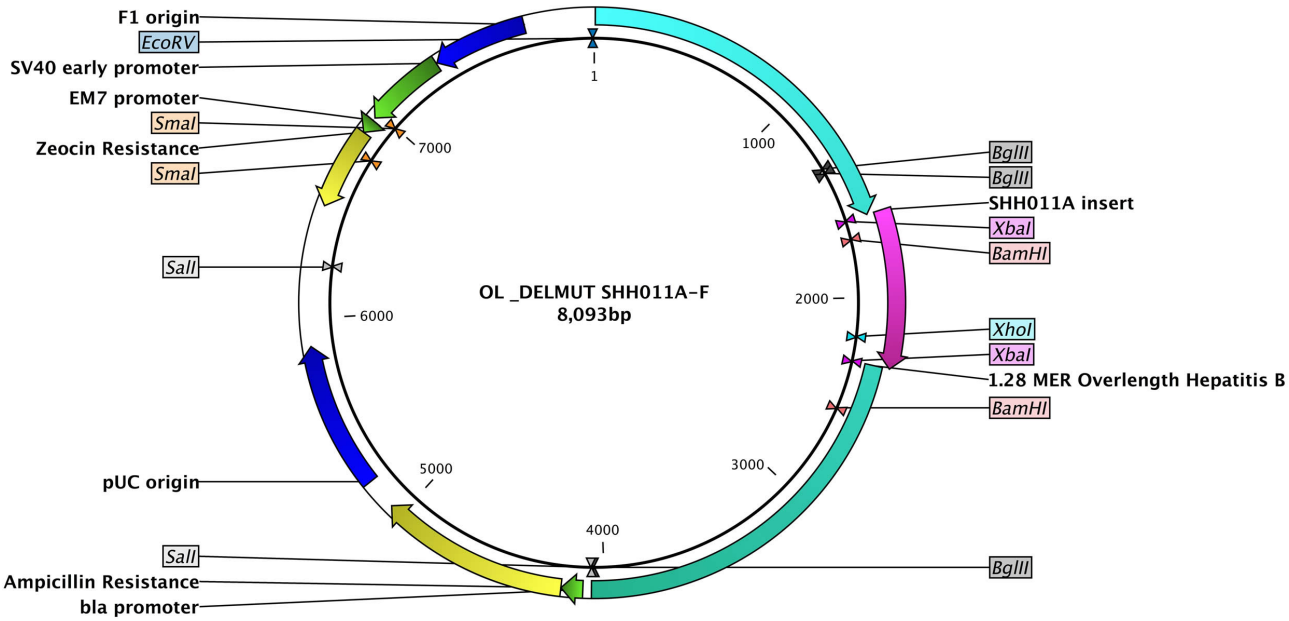
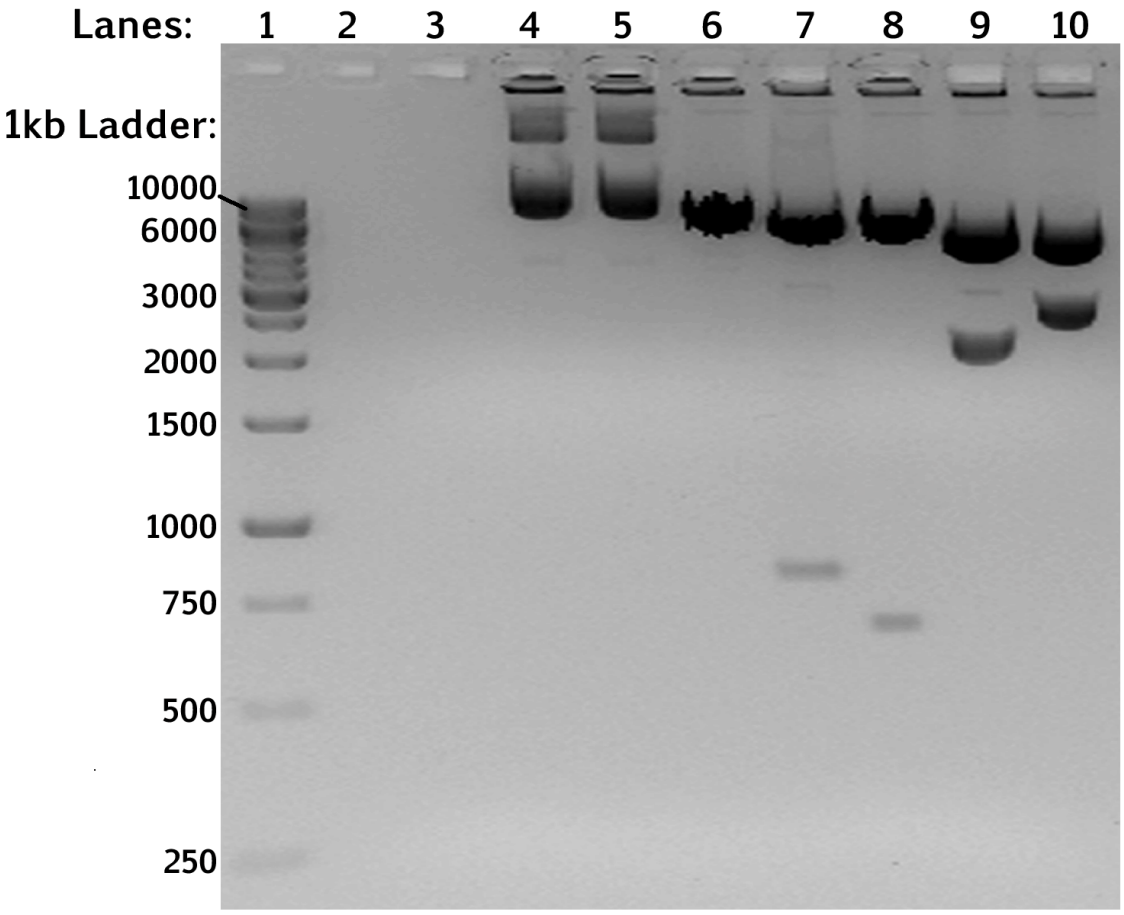


Figure 3.12: Plasmid Map of the OL_DELMUT SHH011A- F Construct

This plasmid map shows the organisation of the deletion-mutant, amplified from the patient SHH011A, inserted in the correct orientation within the OL backbone and the relevant restriction sites.

Table 3.5: Summary of Restriction Enzymes Utilised and Expected Band Sizes

Enzyme	No of cuts	Size of expected fragments for correct orientation
*Pst*I	0	8093 bp circular
*EcoR*I	0	8093 bp circular
*Xho*I	1	8093 bp linear
*BamH*I	2	7256 bp and 837 bp
*Xba*I	2	7405 bp and 688 bp
*Sal*I	2	5908 bp and 2185 bp
*Bgl*II	3	5389 bp, 2680 and 24 bp



Lanes:

1)	GeneRuler™ 1kb DNA Ladder	6)	OL_DELMUT SHH011A-F + *Xho*I
2)	No DNA + *Xba*I	7)	OL_DELMUT SHH011A-F + *BamH*I
3)	Empty	8)	OL_DELMUT SHH011A-F + *Xba*I
4)	OL_DELMUT SHH011A-F + *Pst*I	9)	OL_DELMUT SHH011A-F + *Sal*I
5)	OL_DELMUT SHH011A-F + *EcoR*I	10)	OL_DELMUT SHH011A-F + *Bgl*II

Figure 3.13: Restriction Mapping of the OL_DELMUT SHH011A- F Construct

The resulting restriction map is considered successful as all expected bands were present and were also confirmed to be the correct size using a standard curve graph. Distance travelled was plotted against the log of the molecular weight; band distances from the well were substituted back into the line equation and compared to actual sizes. The uncut plasmid and a no DNA control were included to show enzyme activity was as expected. This figure represents a 1% Agarose gel, stained with ethidium bromide.

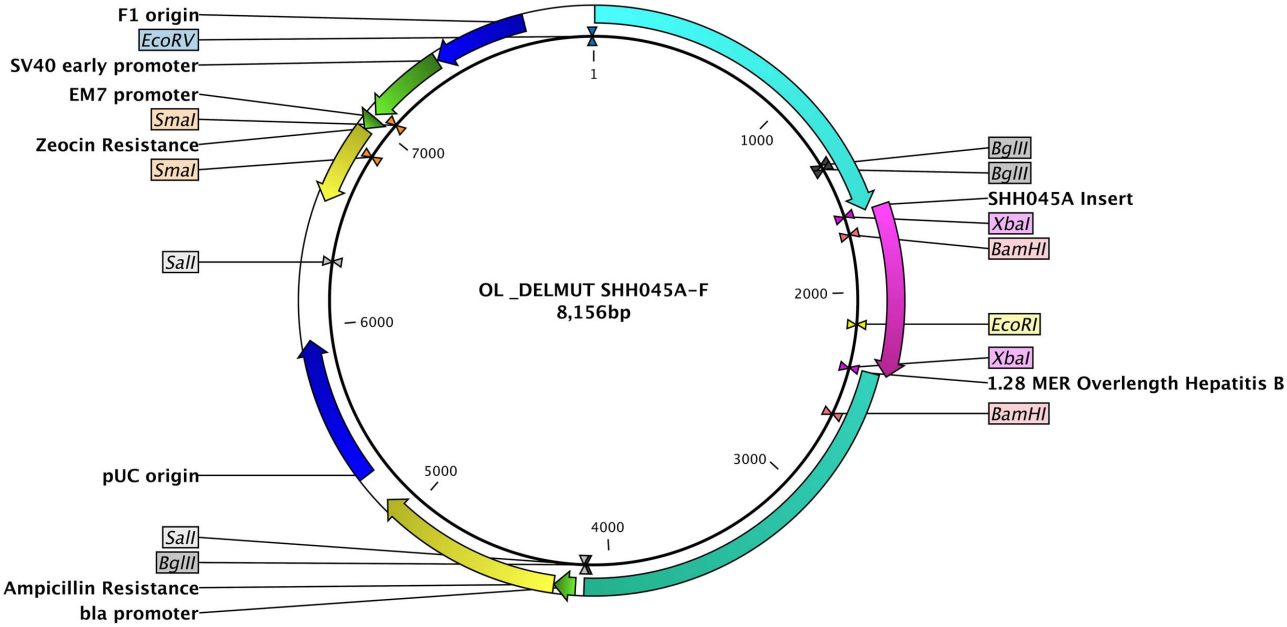
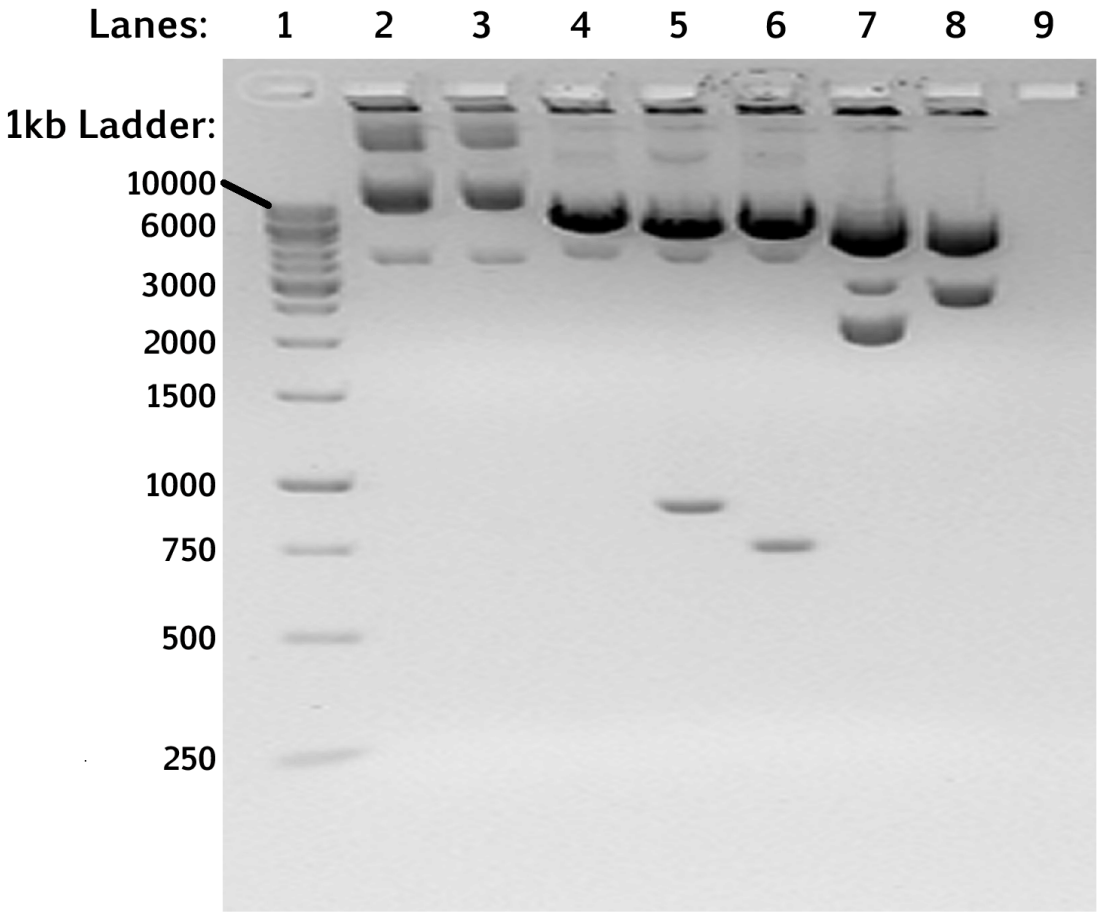


Figure 3.14: Plasmid Map of the OL_DELMUT SHH045A- F Construct

This plasmid map shows the organisation the deletion-mutant, amplified from the patient SHH045A, inserted in the correct orientation within the OL backbone and the relevant restriction sites.

Table 3.6: Summary of Restriction Enzymes Utilised and Expected Band Sizes

Enzyme	No of cuts	Size of expected fragments for correct orientation
*Pst*I	0	8156 bp circular
*Xho*I	0	8156 bp circular
*EcoR*I	1	8156 bp linear
*BamH*I	2	7256 bp and 751 bp
*Xba*I	2	7405 bp and 688 bp
*Sal*I	2	6001 bp and 2185 bp
*Bgl*II	3	5419 bp, 2743 and 24 bp



Lanes:

1) GeneRuler™ 1kb DNA Ladder
2) OL_DELMUT SHH045A-F + *Pst*I
3) OL_DELMUT SHH045A-F + *Xho*I
4) OL_DELMUT SHH045A-F + *EcoR*I
5) OL_DELMUT SHH045A-F + *BamH*I
6) OL_DELMUT SHH045A-F + *Xba*I
7) OL_DELMUT SHH045A-F + *Sal*I
8) OL_DELMUT SHH045A-F + *Bgl*II
9) No DNA + *Xba*I

Figure 3.15: Restriction Mapping of the OL_DELMUT SHH045A- F Construct

The resulting restriction map is considered successful as all expected bands were present and were also confirmed to be the correct size using a standard curve graph. Distance travelled was plotted against the log of the molecular weight; band distances from the well were substituted back into the line equation and compared to actual sizes. The uncut plasmid and a no DNA control were included to show enzyme activity was as expected. This figure represents a 1% Agarose gel, stained with ethidium bromide.

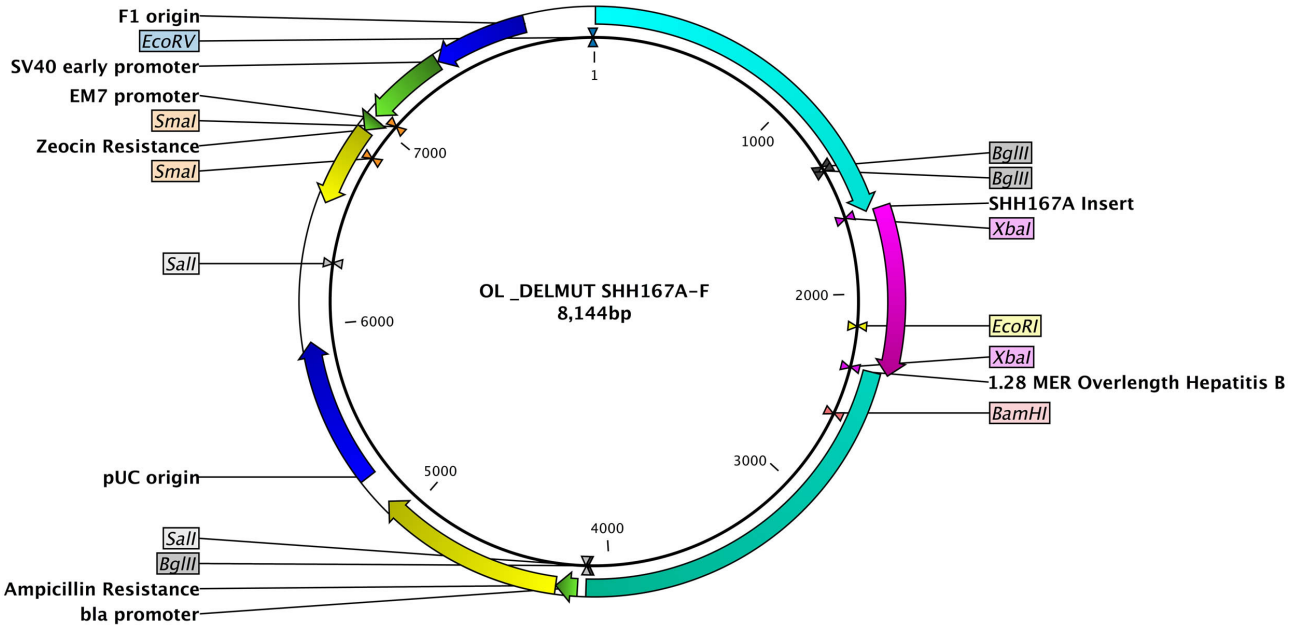
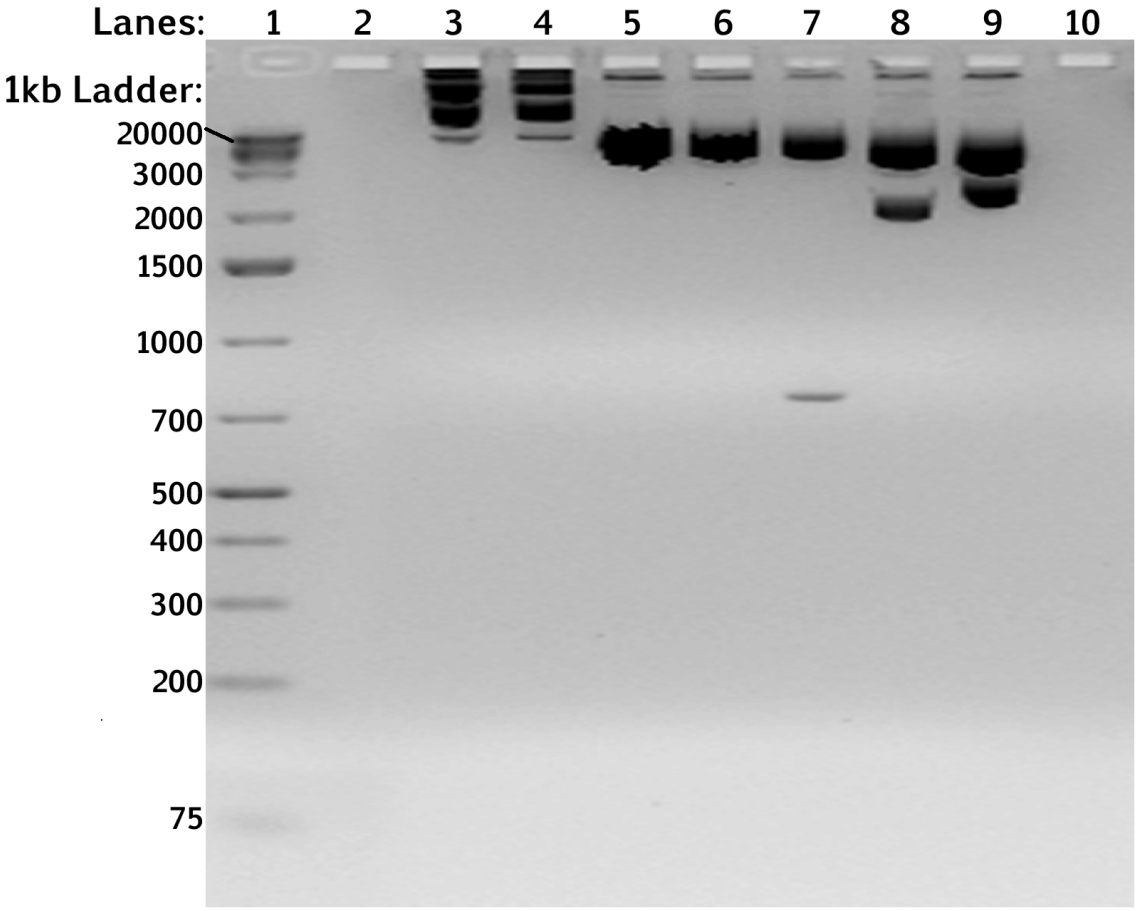


Figure 3.16: Plasmid Map of the OL_DELMUT SHH167A- F Construct

This plasmid map shows the organisation of the deletion-mutant, amplified from the patient SHH167A, inserted in the correct orientation within the OL backbone and the relevant restriction sites.

Table 3.7: Summary of Restriction Enzymes Utilised and Expected Band Sizes

Enzyme	No of cuts	Size of expected fragments for correct orientation
*Pst*I	0	8144 bp circular
*Xho*I	0	8144 bp circular
*EcoR*I	1	8144 bp linear
*BamH*I	1	8144 bp linear
*Xba*I	2	7405 bp and 739 bp
*Sal*I	2	5959 bp and 2185 bp
*Bgl*II	3	5389 bp, 2731 and 24 bp



Lanes:

1) 1kb Plus GeneRuler™ DNA Ladder
2) Empty
3) OL_DELMUT SHH167A-F + *Pst*I
4) OL_DELMUT SHH167A-F + *Xho*I
5) OL_DELMUT SHH167A-F + *EcoR*I
6) OL_DELMUT SHH167A-F + *BamH*I
7) OL_DELMUT SHH167A-F + *Xba*I
8) OL_DELMUT SHH167A-F + *Sal*I
9) OL_DELMUT SHH167A-F + *Bgl*II
10) No DNA + *Xba*I

Figure 3.17: Restriction Mapping of the OL_DELMUT SHH167A- F Construct

The resulting restriction map is considered successful as all expected bands were present and were also confirmed to be the correct size using a standard curve graph. Distance travelled was plotted against the log of the molecular weight; band distances from the well were substituted back into the line equation and compared to actual sizes. The uncut plasmid and a no DNA control were included to show enzyme activity was as expected. This figure represents a 1% Agarose gel, stained with ethidium bromide.

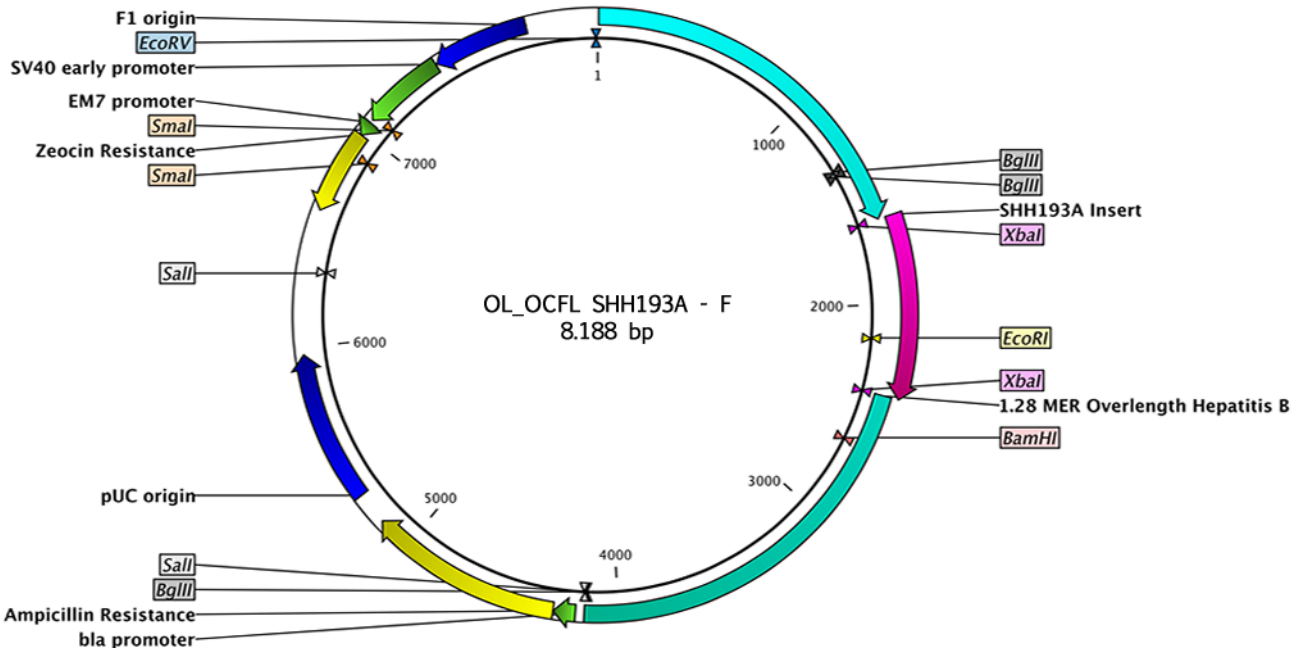
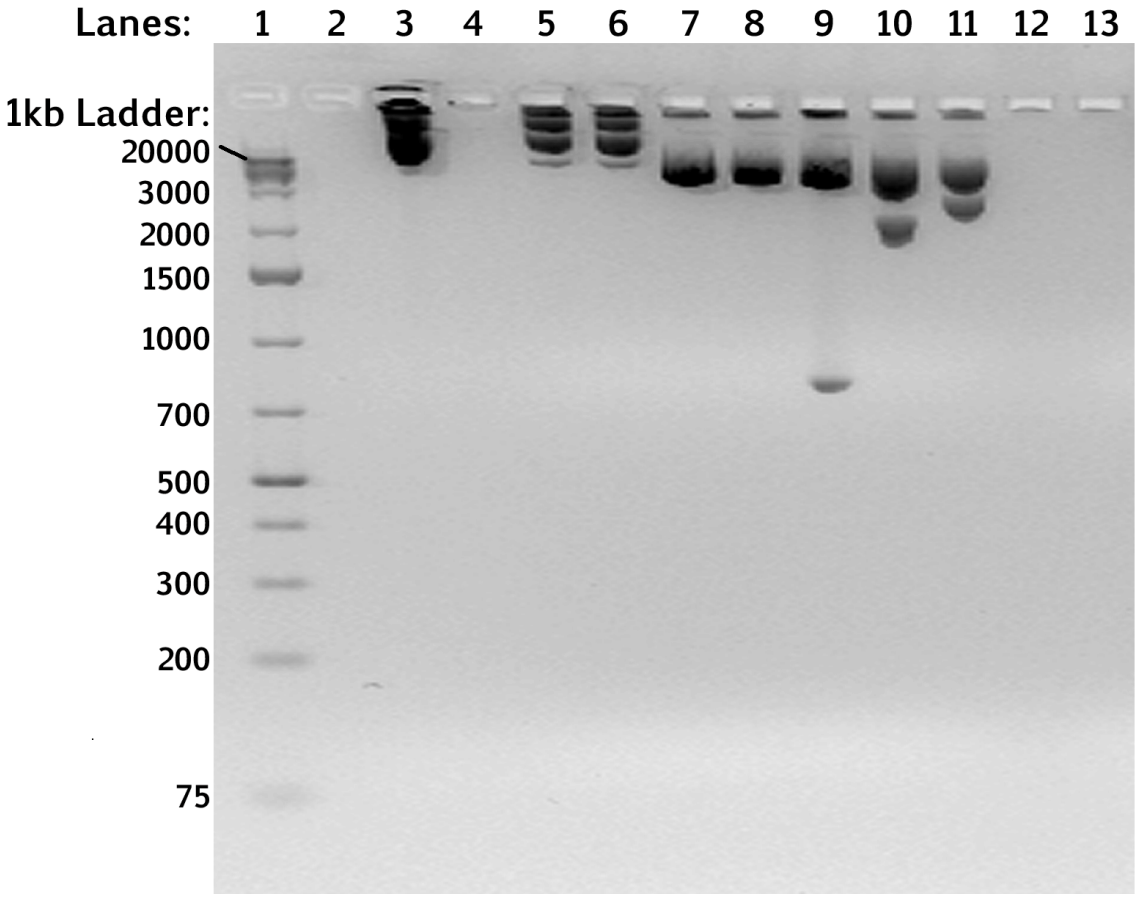


Figure 3.18: Plasmid Map of the OL_OCFL SHH193A- F Construct

This plasmid map shows the organisation of full-length cloning control, amplified from patient SHH193A, inserted in the correct orientation within the OL backbone and the relevant restriction sites.

Table 3.8: Summary of Restriction Enzymes Utilised and Expected Band Sizes

Enzyme	No of cuts	Size of expected fragments for correct orientation
*Pst*I	0	8188 bp circular
*Xho*I	0	8188 bp circular
*EcoR*I	1	8188 bp linear
*BamH*I	1	8188 bp linear
*Xba*I	2	7405 bp and 783 bp
*Sal*I	2	6003 bp and 2185 bp
*Bgl*II	3	5389 bp, 2775 and 24 bp



Lanes:

1) GeneRuler™ 1kb DNA Ladder
2) Empty
3) Uncut OL_OCFL SHH193A-F
4) Empty
5) OL_OCFL SHH193A-F + *Pst*I
6) OL_OCFL SHH193A-F + *Xho*I
7) OL_OCFL SHH193A-F + *EcoR*I
8) OL_OCFL SHH193A-F + *BamH*I
9) OL_OCFL SHH193A-F + *Xba*I
10) OL_OCFL SHH193A-F + *Sal*I
11) OL_OCFL SHH193A-F + *Bgl*II
12) No DNA + *Xba*I
13) Empty

Figure 3.19: Restriction Mapping of the OL_OCFL SHH193A- F Construct

The resulting restriction map is considered successful as all expected bands were present and were also confirmed to be the correct size using a standard curve graph. Distance travelled was plotted against the log of the molecular weight; band distances from the well were substituted back into the line equation and compared to actual sizes. The uncut plasmid and a no DNA control were included to show enzyme activity was as expected. This figure represents a 1% Agarose gel, stained with ethidium bromide.

3.2.4 Sequence Analysis of Overlength Clones

Sequences of the final overlength constructs were aligned and translated using the Babylon Tool [198]. The three deletion mutant and one full-length occult constructs, that were the most homologous to the parental strains were selected for transfection [52]. The sequences of these four selected constructs, their parental strains and the two WT control constructs, A12C15_OL A1WTSA1.281MER and A12C15_OL A1WTSA1.281MER ALT 9.3, were aligned against a WT subgenotype A1 reference (GenBank number AY233274) for comparison. Each reading frame was checked to see whether the translated sequence was in-frame, and whether any unwanted start and/or stop codons were introduced as a result of ligating the 784 bp fragment into the A12C15_OL A1WTSA1.281MER ALT 9.3 backbone. Moreover the functional domains were inspected to determine whether they were conserved (Table 3.9). All amino acid changes reported here include those present the parental strains, as well as any novel changes observed in the final constructs [52]. This same analyses was also carried out on the Polymerase gene, however, this is not included as there were no amino acid changes in the final constructs.

Table 3.9: Functional Domains of the LHBs

Functional Domain	Indicator	Region	Amino Acid Position (Number of aa)	Amino Acid Sequence
Myristoylated N-Terminal Glycine	→	PreS1	ps1G2 (1)	G
NTCP Receptor Binding Site	→	PreS1	ps1R9-I48 (39)	RKGMGTNLSVPNPLGFFPDHQLDP AFGANSNNPDWDFNPI
T Cell Epitope	→	PreS1/2	ps1S109-ps2L12 (23)	SPPLRDSHPQAMQWNSTAFHQAL
B Cell Epitope	→	PreS2	ps1Q132-ps2G25 (13)	QDPRVRGLYFPAG
Neutralising Antibody Target 1	→	PreS1	ps1V18-ps1D50 (33)	VPNPLGFFPDHQLDPAFGANSNNP DWDFNPIKD
Neutralising Antibody Target 2	→	PreS2	ps2M1-ps2I45 (45)	MQWNSTAFHQALQDPRVRGLYFP AGGSSSGTLNPVPNTASHISPI

The colour of the indicator arrow correlates with each functional domain as represented Figures 3.20-3.22.

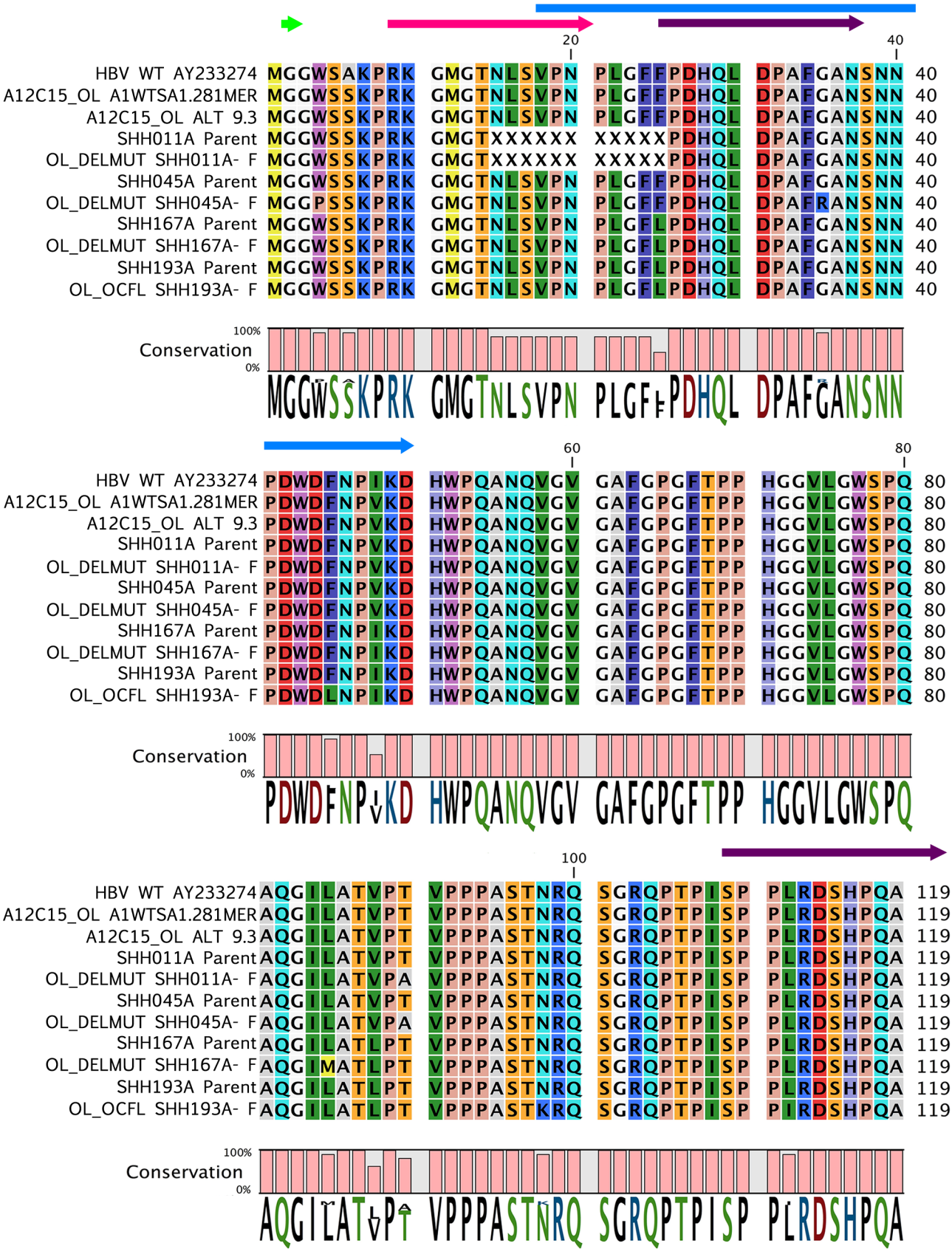


Figure 3.20: Translated aa Sequence of the preS1 Region

This alignment depicts the translated sequences as compared to a WT subgenotype A1 reference sequence (GenBank Accession AY233274). Functional domains are indicated by coloured arrows that corresponding to Table 3.9 above, and deletions are indicated by plain X letters. All mutations observed are detailed below. A base change compared to the WT is not reported as a mutation if it is present in all the constructs.

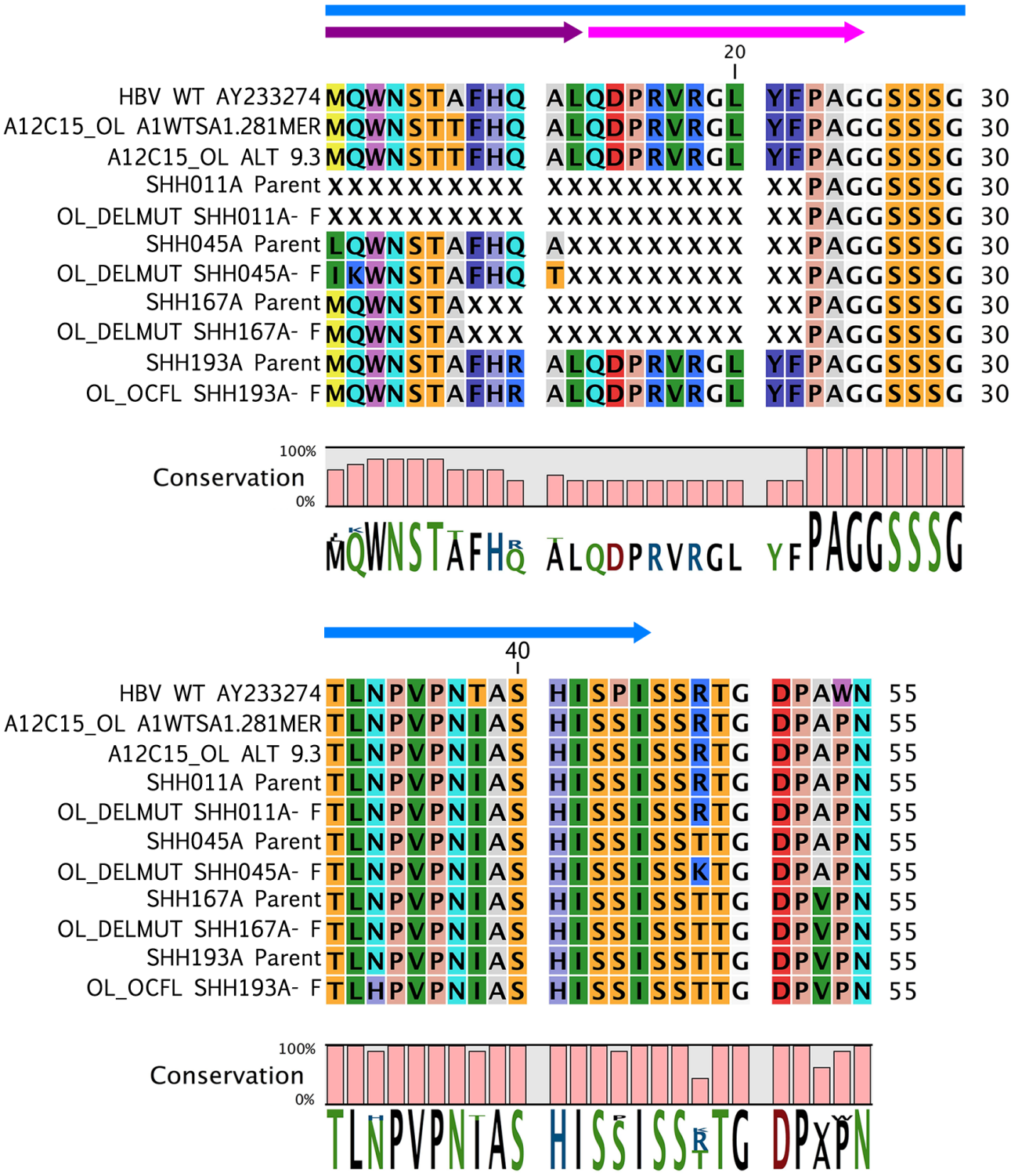


Figure 3.21: Translated aa Sequence of the preS2 Region

This alignment depicts the translated sequences as compared to a WT subgenotype A1 reference sequence (GenBank Accession AY233274). Functional domains are indicated by coloured arrows that corresponding to Table 3.9 above, and deletions are indicated by plain X letters. All mutations observed are detailed below. A base change compared to the WT is not reported as a mutation if it is present in all the constructs.

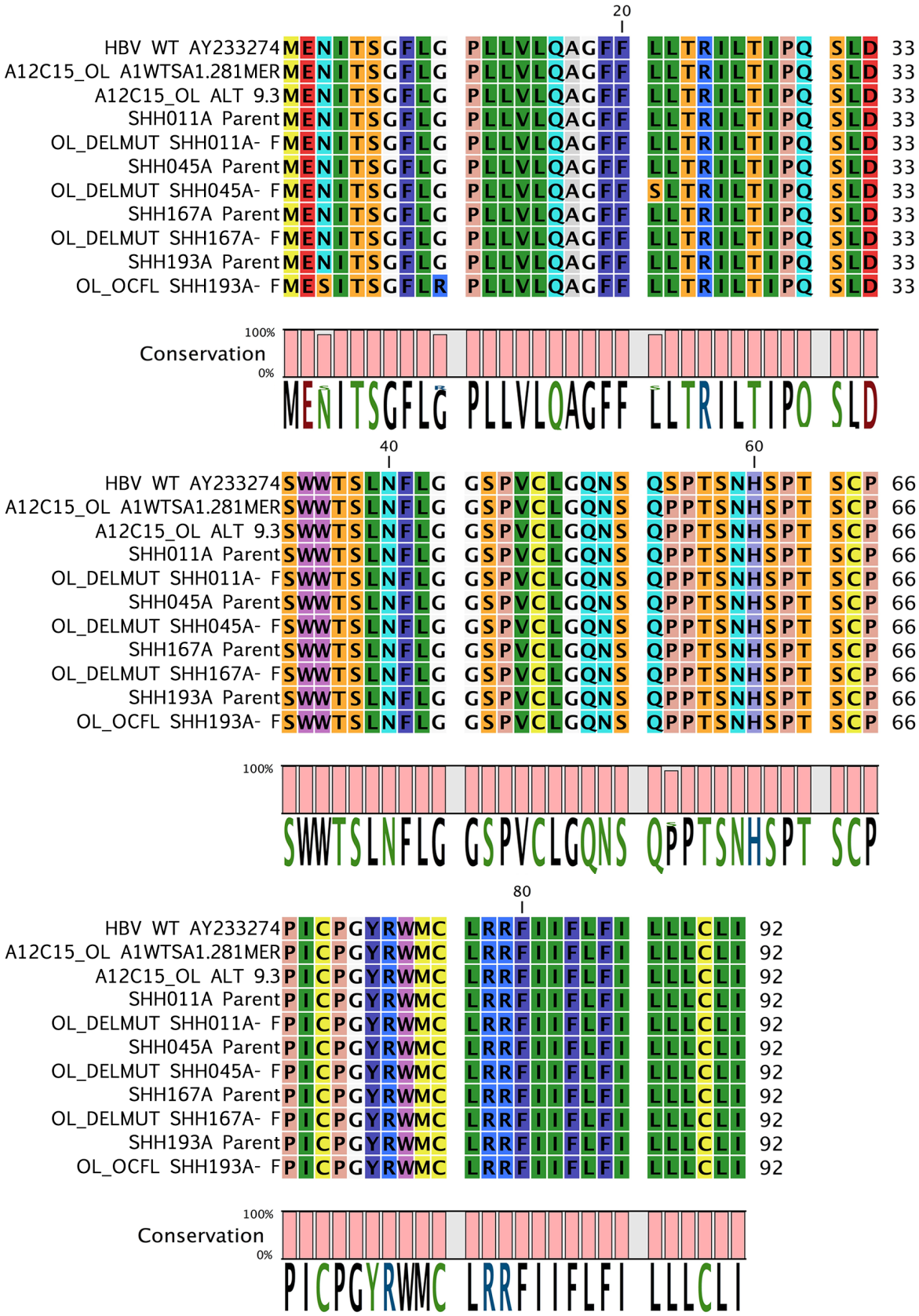


Figure 3.22: Translated aa Sequence of the Surface Region of All Constructs

This alignment depicts the translated sequences as compared to a WT subgenotype A1 reference sequence (GenBank Accession AY233274). Functional domains are indicated by coloured arrows that corresponding to Table 3.9 above, and deletions are indicated by plain X letters. All mutations observed are detailed below. A base change compared to the WT is not reported as a mutation if it is present in all the constructs.

Only 784 bp fragment, from position 2684 bp to position 252 bp from the *Eco*RI siteof the A12C15_OL A1WTSA1.281MER ALT 9.3 backbone was replaced. This fragment only extends 92 aa into the Surface region, and only the first the first 23 aa contained deletions (Figure 3.22).

Table 3.10: Missense Mutations in the preS1, preS2 and S Regions of the Final Overlength Constructs:

Construct	Region in Which aa Changes Were Observed		
	preS1	preS2	S
A12C15_OL A1WTSA1.281MER	pS1I48V	ps2A7T	
A12C15_OL A1WTSA1.281MER ALT 9.3	pS1I48V	ps2A7T	
OL_DELMUT SHH011A - F	pS1del15-25, pS1I48V pS1T90A	ps2del1-22, ps2T48K	
OL_DELMUT SHH045A - F	pS1I48V pS1W4P, pS1G35R, pS1T90A	ps2del12-22, ps2L1I, ps2Q2K, ps2A11T	sL21S
OL_DELMUT SHH167A - F	pS1F25L, pS1V88L pS1L85M	ps2del8-22, ps2R48T	
OL_OCFL SHH193A - F	pS1F25L, pS1V88L pS1F45L, pS1N98K, pS1L112I	ps2Q10R, psR48T, ps2A53V ps2N33H, ps2R48T	sN3S, sG10R

All translated sequences were compared to a WT subgenotype A1 reference sequence (GenBank accession number AY233274). The underlined mutations represent the aa changes already in the parental strain [52]. In plain text are the inadvertent aa changes not found in the parental strain but only in the final construct.

3.3 Determination of the Replication Competence of Overlength Clones

The A12C15_OL A1WTSA1.281MER, which is replication-competent [193] was used as the backbone for the generation of the deletion-mutant constructs and as a positive control.

In order to insert the subgenomic fragment, containing the deletions, two *Xba*I sites spanning the 784 bp fragment were utilized. However, because the vector contained three *Xba*I sites, the redundant site had to be removed. This removal resulted in a backbone, which was 22 bp shorter than A12C15_OL A1WTSA1.281MER. The shorter plasmid was named A12C15 ALT 9.3. The restricted 22 bp fragment was in the multiple cloning site of the pcDNA™ 4/TO plasmid vector, which is not functionally important. The fragments containing the deletion-mutants were all inserted into this A12C15 ALT 9.3 backbone. Thus transfection experiments with A12C15 ALT 9.3 were used for the normalisation and comparison of the replication of the deletion mutant constructs.

An additional control was constructed in which the S region of HBV, without deletions, was amplified from the serum derived from patient SHH193A and inserted into A12C15 ALT 9.3. This resulted in the full-length cloning control OL_OCFL SHH193A.

The negative controls included mock-transfected Huh7 cells, with and without the transfection agent PEI as well as cells transfected with the eGFP-C3 plasmid, a mammalian expression vector, containing a GFP fluorescent tag under the control of the CMV promoter and used here to measure transfection efficiency.

3.3.1 Expression of HBsAg and HBeAg

HBsAg and HBeAg are both expressed extracellularly, thus were harvested at day 1 (24 hours), day 3 (72 hours) and day 5 (120 hours) after transfection. The quantitative measurements were repeated 3 times with 2 biological replicates each. The expression of these proteins indicates that a construct is producing virus in cell culture. Samples were considered negative if the OD measure was below the cut off values, which was calculated by adding 0.05 to the mean of the kit provided negative control replicates. The results of HBsAg expression in transfected Huh7 cells are shown Figures 3.23 – 3.27 and the significant differences when measuring HBsAg expression is summarized in Table 3.11.

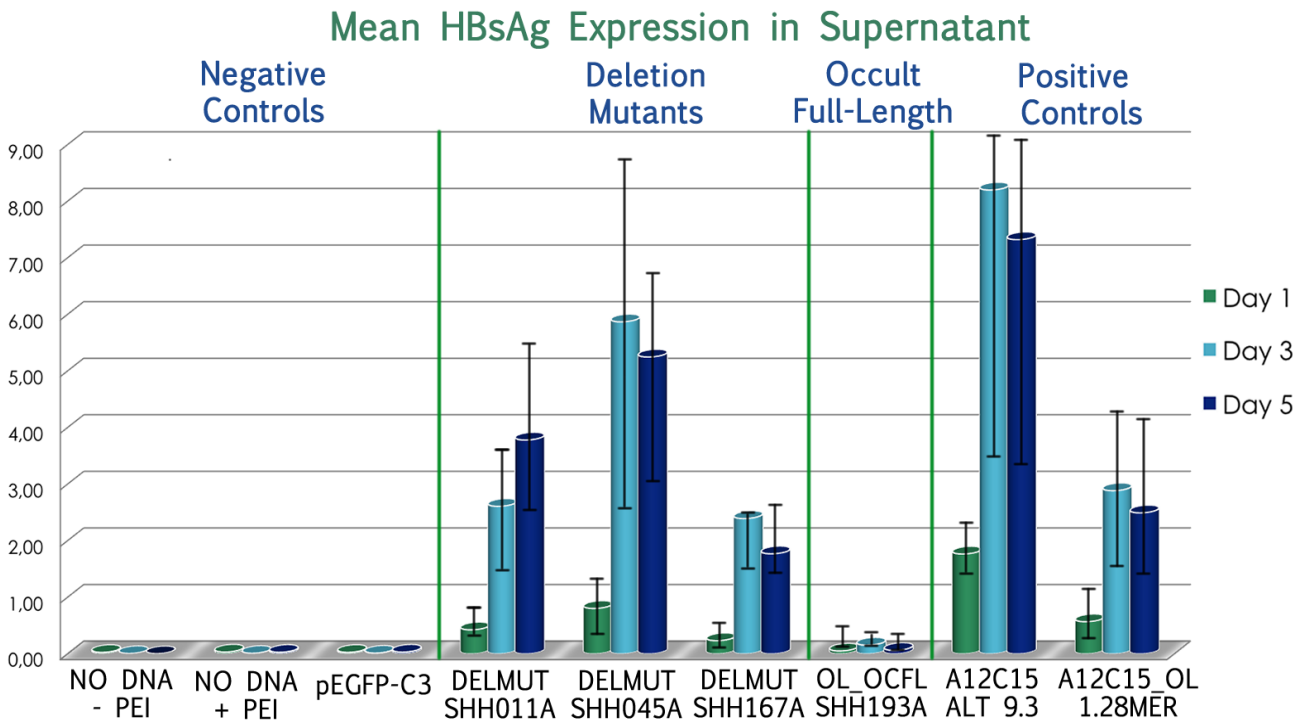


Figure 3.23: Detection of Mean HBsAg Expression in Culture Supernatant of Transfected Huh7 Cells

This bar graph presents the observed HBsAg expression in supernatant of transfected Huh7 cells as measured by ELISA. Error bars represent the standard deviation of the mean. Expression was tested at a dilution 1:20 of supernatant, unless OD measures were out of the linear scale. In which case those samples were re-measured at 1:80 dilution, and multiplied by 4.

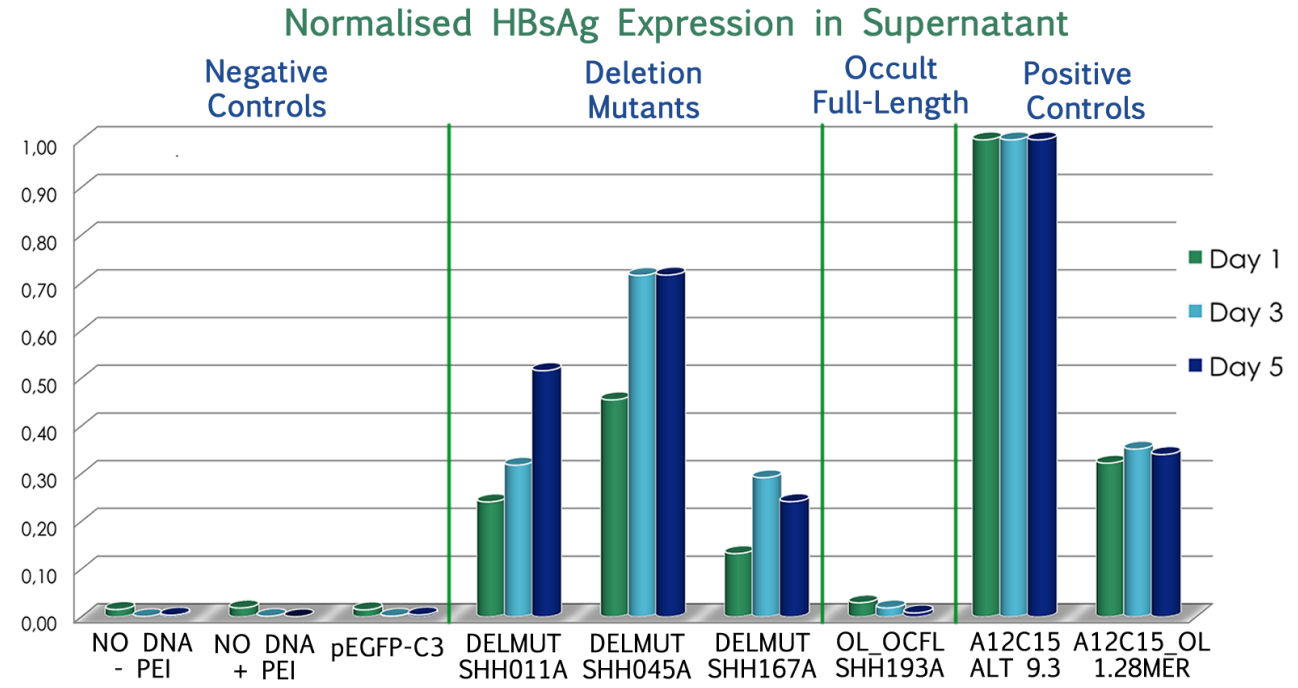


Figure 3.24: Normalised HBsAg Expression in Culture Supernatant of Transfected Huh7 Cells

The observed HBsAg Expression in supernatant of transfected Huh7 cells was normalized against the A12C15 ALT 9.3 construct to show relative expression.

Table 3.11: Significant Differences in HBsAg Expression

	A12C15_OL 1.28MER DAY 1 DAY 3 DAY 5	A12C15 ALT 9.3 DAY 1 DAY 3 DAY 5	OL_OCFL SHH193A DAY 1 DAY 3 DAY 5	DELMUT SHH011A DAY 1 DAY 3 DAY 5	DELMUT SHH045A DAY 1 DAY 3 DAY 5	DELMUT SHH167A DAY 1 DAY 3 DAY 5
DAY 1						
A12C15_OL 1.28MER						
A12C15 ALT 9.3	↑ (p = 0.014)		↑ (p = 0.001)	↑ (p = 0.007)	↑ (p = 0.040)	↑ (p = 0.002)
OL_OCFL SHH193A					↓ (p = 0.040)	
DELMUT SHH011A				↓ (p = 0.039) ↓ (p = 0.015)		
DELMUT SHH045A					↓ (p = 0.025)	
DELMUT SHH167A						↓ (p = 0.010)
DAY 3						
A12C15_OL 1.28MER						
A12C15 ALT 9.3						
OL_OCFL SHH193A	↓ (p = 0.029)	↓ (p = 0.048)		↓ (p = 0.028)	↓ (p = 0.039)	↓ (p = 0.007)
DELMUT SHH011A						
DELMUT SHH045A						
DELMUT SHH167A						↓ (p = 0.013)
DAY 5						
A12C15_OL 1.28MER						
A12C15 ALT 9.3						
OL_OCFL SHH193A	↓ (p = 0.033)	↓ ns: (p = 0.075)		↓ (p = 0.010)	↓ (p = 0.014)	↓ (p = 0.008)
DELMUT SHH011A						
DELMUT SHH045A						
DELMUT SHH167A						

Arrows indicate a significant relationship where experimental samples on the left column are compared relative to controls and experimental samples in the top row. Significance was determined to be any p-value below 0.05 by employing a student *t*-test. Any blank cell indicates there is no significant difference.

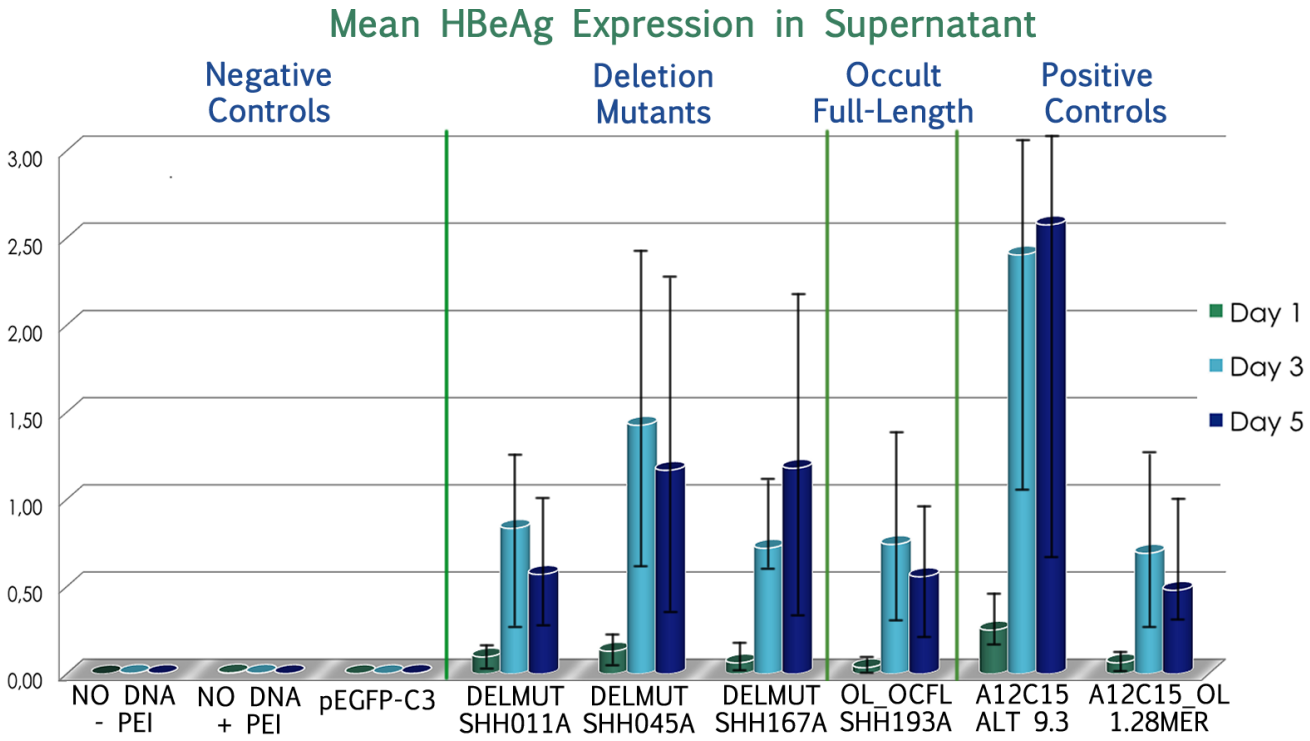


Figure 3.25: Detection of HBeAg Expression in Culture Supernatant of Transfected Huh7 Cells

This bar graph presents the observed HBsAg expression in supernatant of transfected Huh7 cells as measured by ELISA. Error bars represent the standard deviation of the mean. Expression was tested at a dilution 1:20 of supernatant, unless OD measures were out of the linear scale. In which case those samples were re-measured at 1:80 dilution, and then the OD was multiplied by 4.

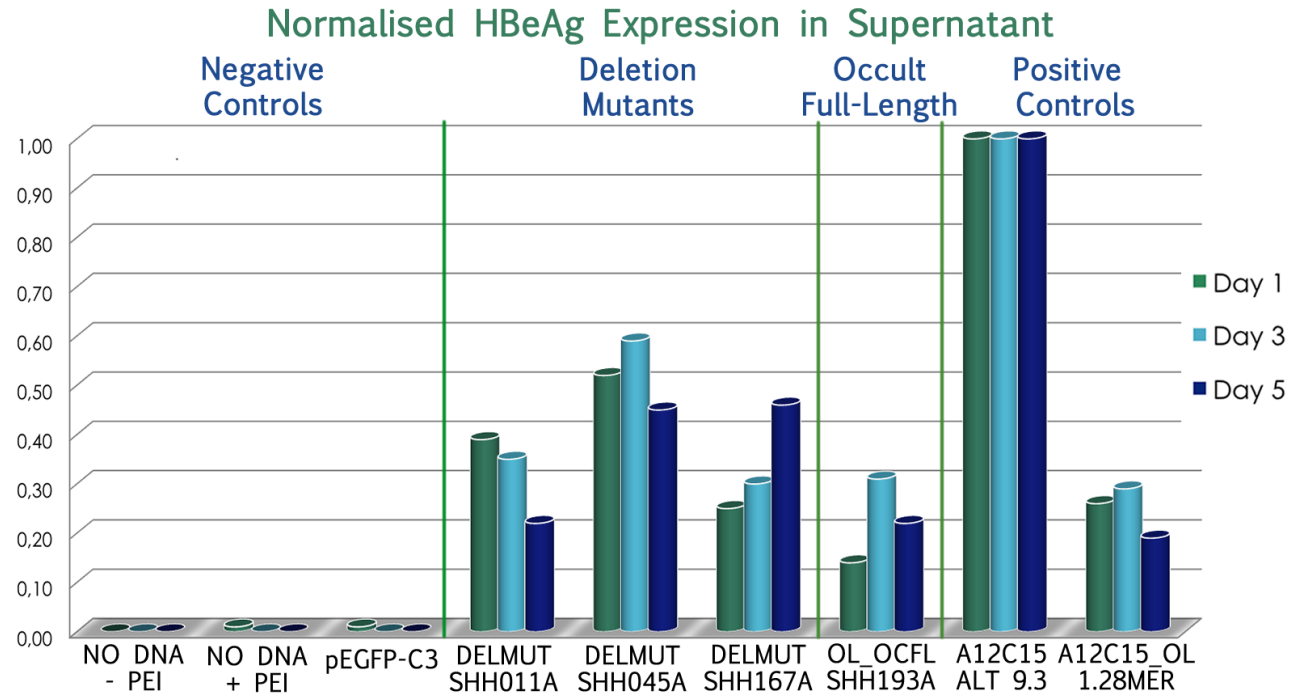


Figure 3.26: Normalised HBeAg Expression in Culture Supernatant of Transfected Huh7 Cells

HBeAg expression in supernatant of transfected Huh7 cells was normalized against the A12C15 ALT 9.3 construct to show relative expression.

There were no significant differences in the expression of HBeAg in Huh7 transfected cells between the deletion-mutant constructs and the positive controls. Furthermore, there was no significant difference in HBeAg expression between days 1, 3 and 5.

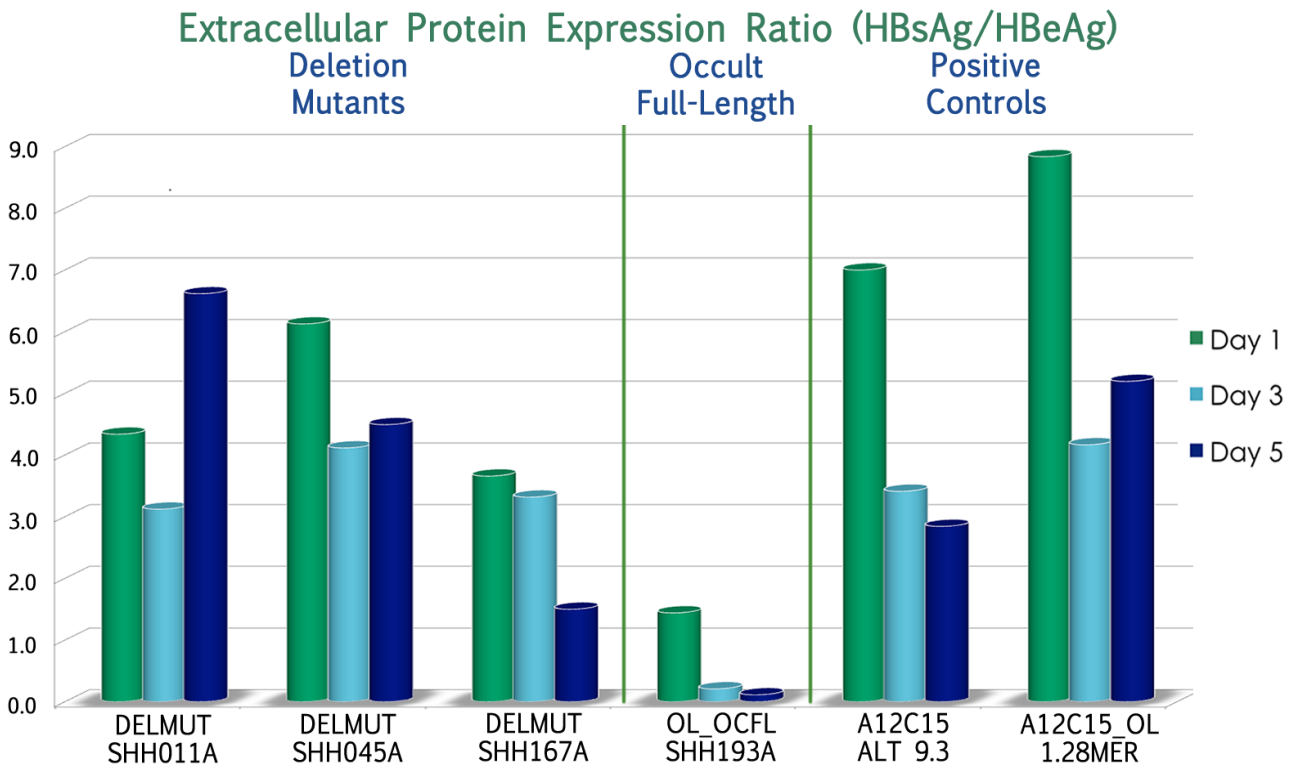


Figure 3.27: Expression of HBsAg Relative to HBeAg

This figure shows the mean values of HBsAg were divided by mean values of HBeAg to show extracellular protein expression as a ratio.

3.3.2 Viral Load Quantification of Overlength Constructs

Samples for qPCR were collected on day 1 (24 hours), day 3 (72 hours) and day 5 (120 hours) after transfection. Supernatant, lysate and immunocaptured virions were collected in order to analyse extracellular, intracellular (cell-associated) and secreted/encapsidated virion DNA levels, respectively. Immunocapture was carried out only on day 3 (72 hours) after transfection. Secreted hepatitis B virions were estimated in undiluted supernatants of transfected cells by using an immunocapture ELISA strategy before isolating the DNA. Three biological replicates were employed for the Real-Time plates from three independent transfections as well as three technical replicates for each sample. For qPCR, a sample was considered negative if the Sq values were lower than the threshold and the negative water control. qPCR results are shown in Figures 3.28 to 3.33.

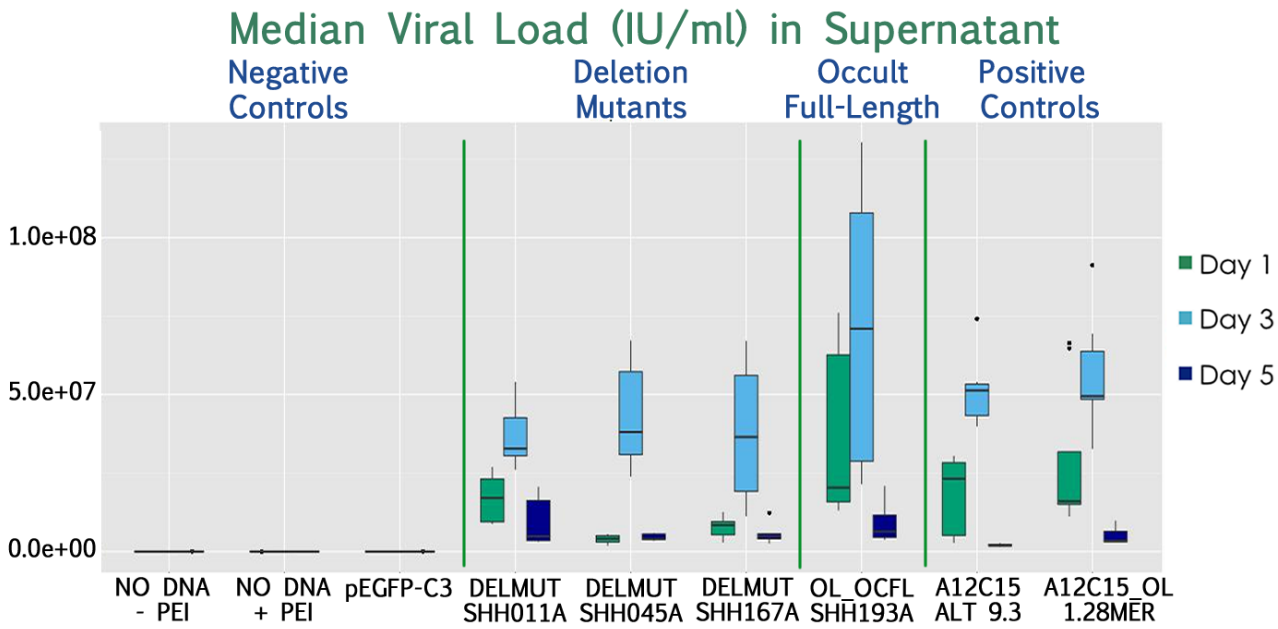


Figure 3.28: Box and Whisker plot of HBV DNA Expression in Culture Supernatant of Huh7 Cells

The central value or median is shown by the black horizontal line through the bar. The vertical lines on the upper and lower extremes of each bar represent the upper and lower quartiles, meaning that the coloured bar itself represents the interquartile range.

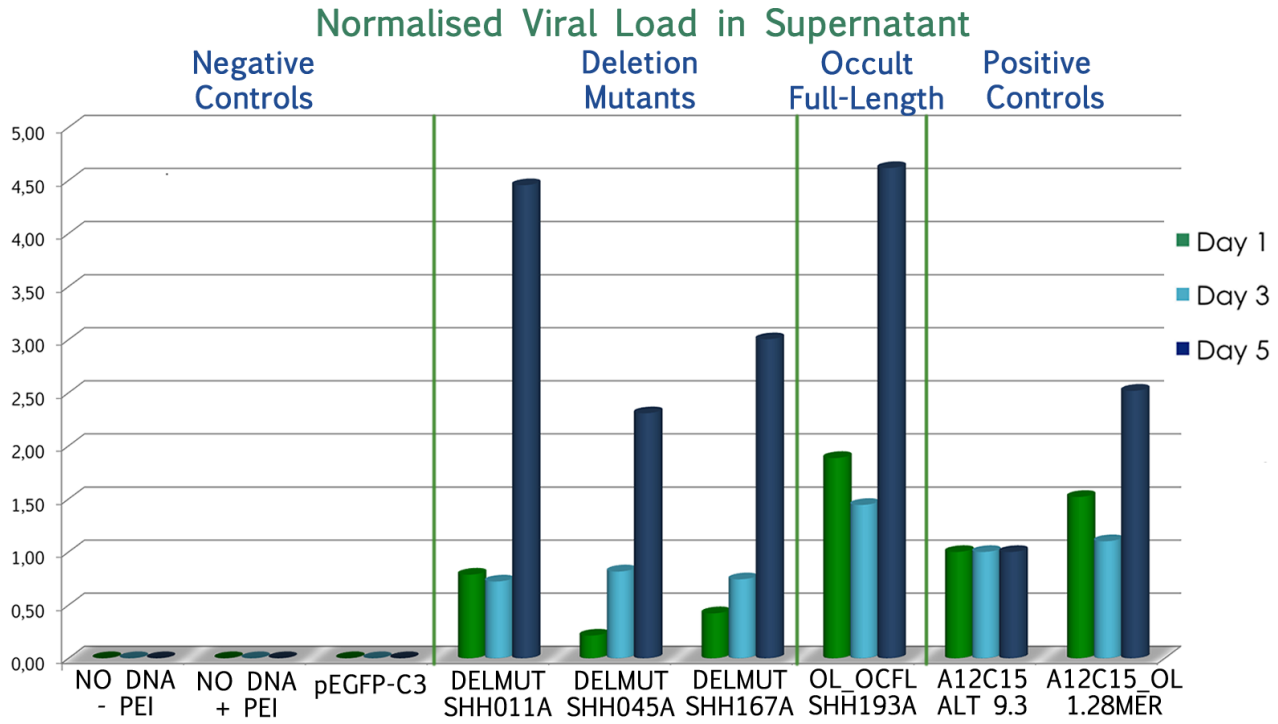


Figure 3.29: Normalised HBV DNA Expression in Culture Supernatant of Huh7 Cells

Represented here is the normalised data for quantified VL in the supernatant of transfected Huh7 cells. Normalisation was carried out against the A12C15 ALT 9.3 construct.

Table 3.15 summarises all the significant differences when measuring HBV DNA in the supernatant of Huh7 transfected cells.

Table 3.12: Significant Differences in Supernatant Viral Loads

	A12C15_OL 1.28MER DAY 1 DAY 3 DAY 5	A12C15 ALT 9.3 DAY 1 DAY 3 DAY 5	OL_OCFL SHH193A DAY 1 DAY 3 DAY 5	DELMUT SHH011A DAY 1 DAY 3 DAY 5	DELMUT SHH045A DAY 1 DAY 3 DAY 5	DELMUT SHH167A DAY 1 DAY 3 DAY 5
DAY 1						
A12C15_OL 1.28MER	↓ (p = 0.008) ↑ (p = 0.006)				↑ (p = 0.005)	↑ (p = 0.015)
A12C15 ALT 9.3		↓ (p = 0.000) ↑ (p = 0.000)			↑ (p = 0.003)	↑ (p = 0.021)
OL_OCFL SHH193A			↑ (p = 0.011)	↑ (p = 0.039)	↑ (p = 0.003)	↑ (p = 0.007)
DELMUT SHH011A				↓ (p = 0.001)	↑ (p = 0.000)	↑ (p = 0.022)
DELMUT SHH045A					↓ (p = 0.000)	↓ (p = 0.003)
DELMUT SHH167A						↓ (p = 0.020)
DAY 3						
A12C15_OL 1.28MER	↑ (p = 0.000)			↑ (p = 0.029)		
A12C15 ALT 9.3		↑ (p = 0.000)		↑ (p = 0.032)		
OL_OCFL SHH193A			↑ (p = 0.001)			
DELMUT SHH011A				↑ (p = 0.000)		
DELMUT SHH045A					↑ (p = 0.000)	
DELMUT SHH167A						↑ (p = 0.001)
DAY 5						
A12C15_OL 1.28MER						
A12C15 ALT 9.3	↓ (p = 0.006)		↓ (p = 0.003)	↓ (p = 0.012)	↓ (p = 0.000)	↓ (p = 0.005)
OL_OCFL SHH193A					↑ (p = 0.044)	
DELMUT SHH011A						
DELMUT SHH045A						
DELMUT SHH167A						

Arrows indicate a significant relationship where experimental samples on the left column are compared relative to controls and experimental samples in the top row. Significance was determined to be any p-value below 0.05 by employing a student *t*-test. Any blank cell indicates there is no significant difference.

Immunocapture VL were measured only on day 3 post transfection.

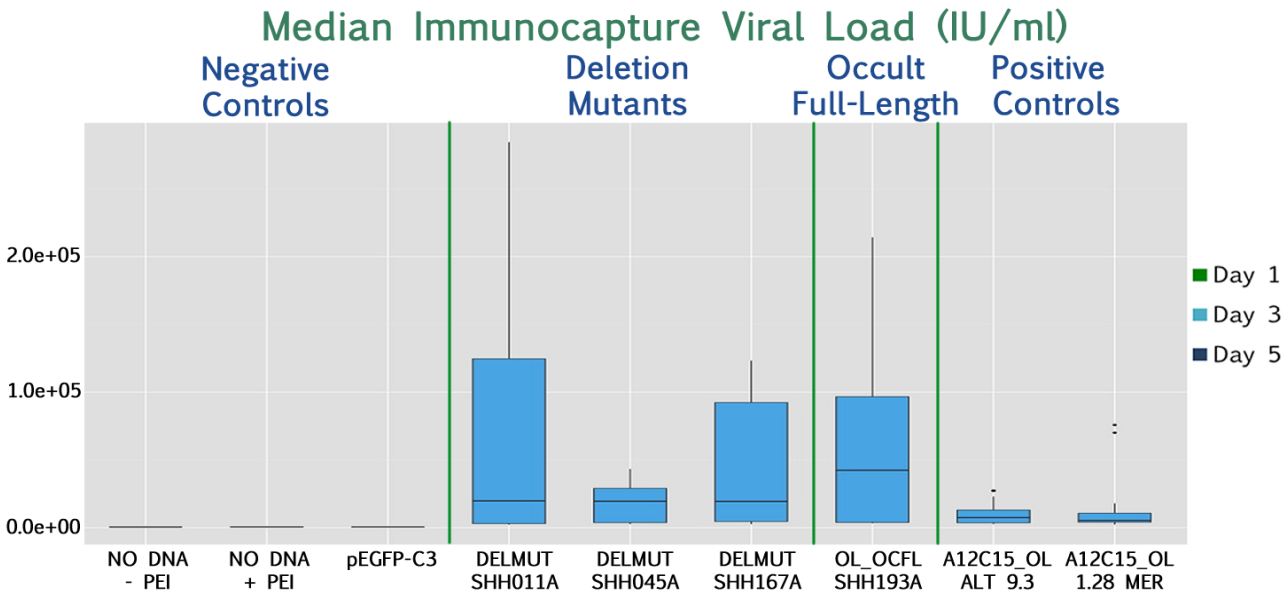


Figure 3.30: Box and Whisker Plot of Immunocaptured Supernatant HBV DNA Expression of Huh7 Cells

This figure represents the distribution of data for quantified VL following immunocapture. The central value or median represented by the black horizontal line through the bar. The vertical lines on the upper and lower extremes of each bar represent the upper and lower quartiles, meaning that the coloured bar itself represents the interquartile range.

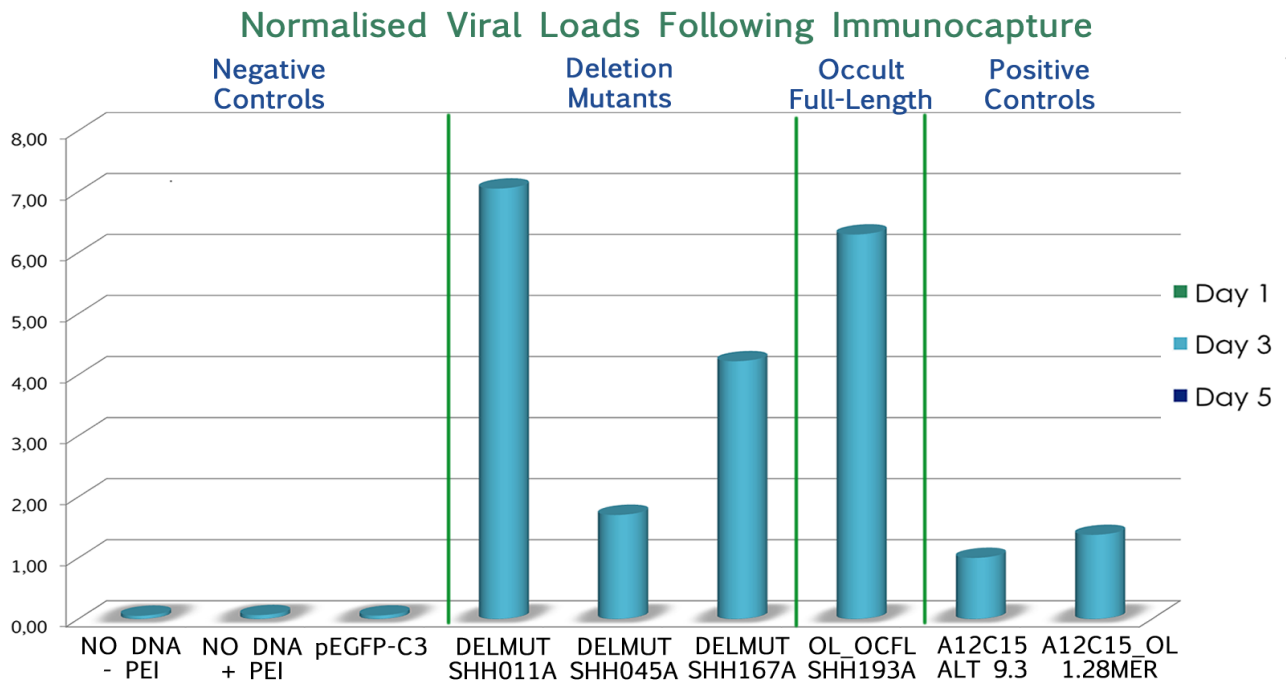


Figure 3.31: Normalised Secreted Virion HBV DNA Expression in Immunocaptured Supernatant of Huh7 Cells

Depicted here is the normalised data for quantified VL in immunocaptured virions found in supernatant of transfected Huh7 cells. Normalisation was carried out against the A12C15 ALT 9.3 construct to show relative HBV DNA measurement.

Table 3.13: Significant Differences in Immunocapture Viral Loads

	A12C15_OL 1.28MER DAY 3	A12C15 ALT 9.3 DAY 3	OL_OCFL SHH193A DAY 3	DELMUT SHH011A DAY 3	DELMUT SHH045A DAY 3	DELMUT SHH167A DAY 3
DAY 3						
A12C15_OL 1.28MER			↓ (p = 0.000)	↓ (p = 0.029)	↓ (p = 0.000)	↓ (p = 0.000)
A12C15 ALT 9.3			↓ (p = 0.000)	↓ (p = 0.001)	↓ (p = 0.001)	↓ (p = 0.000)
OL_OCFL SHH193A				↓ (p = 0.000)	↑ (p = 0.000)	↑ (p = 0.000)
DELMUT SHH011A					↑ (p = 0.000)	↑ (p = 0.000)
DELMUT SHH045A						↓ (p = 0.000)
DELMUT SHH167A						

Arrows indicate a significant relationship where experimental samples on the left column are compared relative to controls and experimental samples in the top row. Significance was determined to be any p-value below 0.05 by employing a student *t*-test. Any blank cell indicates there is no significant difference.

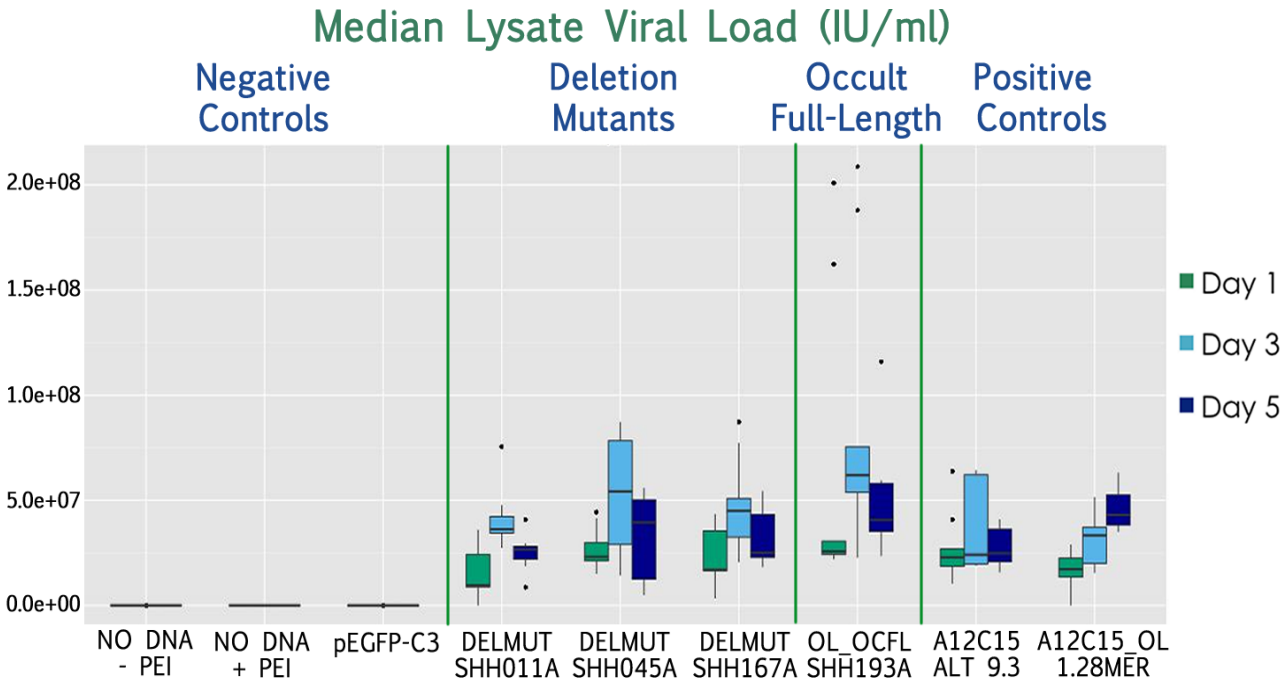


Figure 3.32: Box and Whisker Plot of HBV DNA Expression in Culture Lysates of Huh7 Cells

Quantification of the VL in the lysates, and thus cell-associated virus. The central value or median represented by the black horizontal line through the bar. The vertical lines on the upper and lower extremes of each bar represent the upper and lower quartiles, meaning that the coloured bar itself represents the interquartile range.

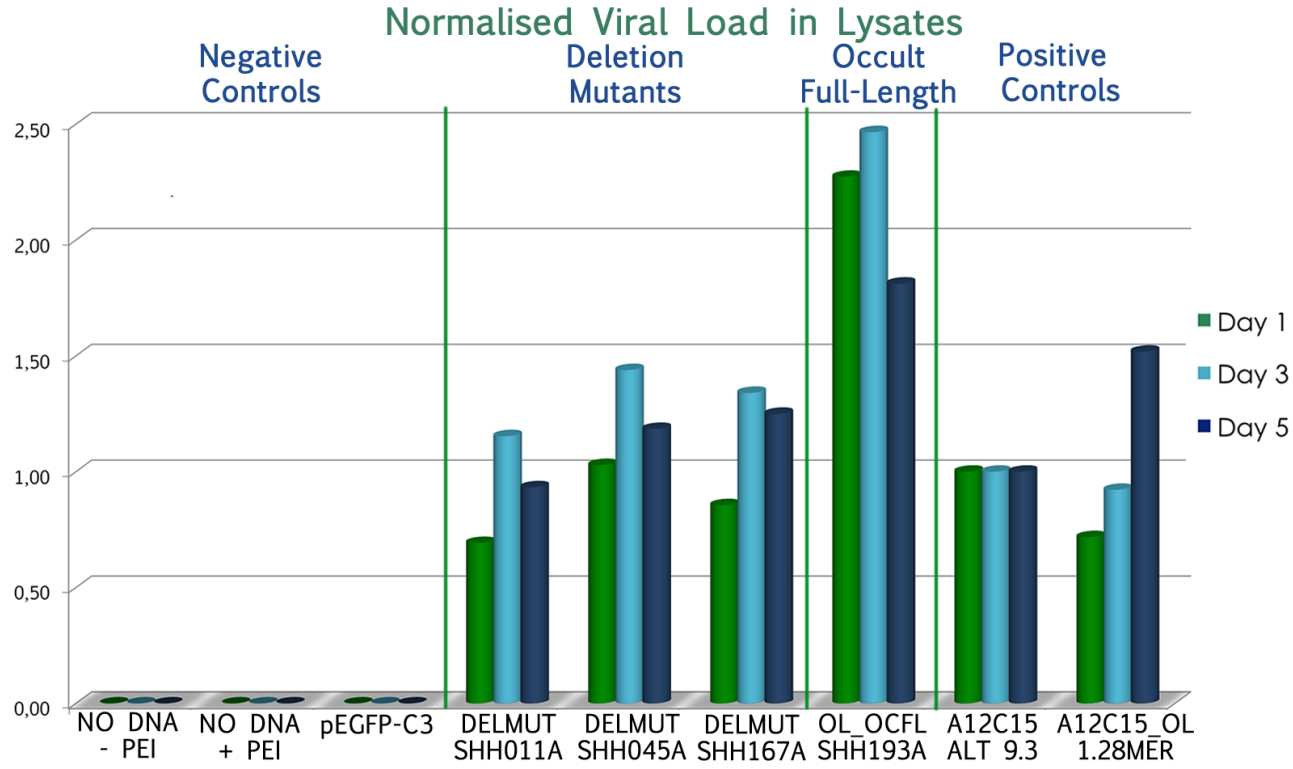


Figure 3.33: Normalised HBV DNA Expression in Culture Lysates of Huh7 Cells

This figure represents the normalised data for quantified VL in lysates, thus cell-associated virus. Normalisation was carried out against the A12C15 ALT 9.3 construct to show relative HBV DNA measurement.

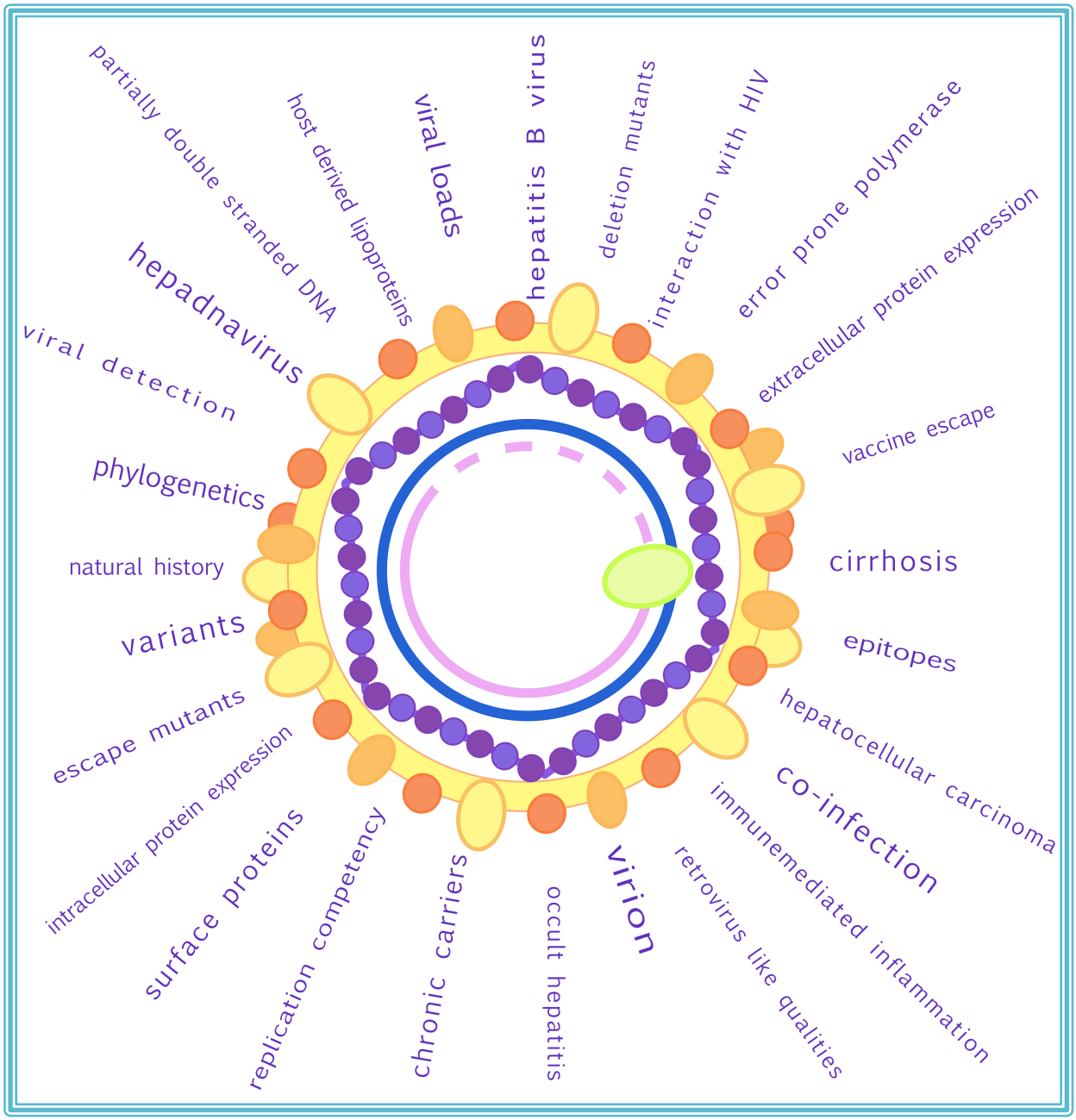
All significant differences for the lysate or cell-associated qPCR results are highlighted in Table 3.17.

Table 3.14: Significant Differences in Lysate Viral Load

	A12C15_OL 1.28MER DAY 1 DAY 3 DAY 5	A12C15 ALT 9.3 DAY 1 DAY 3 DAY 5	OL_OCFL SHH193A DAY 1 DAY 3 DAY 5	DELMUT SHH011A DAY 1 DAY 3 DAY 5	DELMUT SHH045A DAY 1 DAY 3 DAY 5	DELMUT SHH167A DAY 1 DAY 3 DAY 5
DAY 1						
A12C15_OL 1.28MER	↓ (p = 0.016) ↓ (p = 0.000)					
A12C15 ALT 9.3						
OL_OCFL SHH193A						
DELMUT SHH011A				↓ (p = 0.003)		
DELMUT SHH045A					↓ (p = 0.030)	
DELMUT SHH167A						↓ (p = 0.011)
DAY 3						
A12C15_OL 1.28MER	↓ (p = 0.028)		↓ (p = 0.026)			
A12C15 ALT 9.3			↓ (p = 0.038)			
OL_OCFL SHH193A						
DELMUT SHH011A				↑ (p = 0.013)		
DELMUT SHH045A						
DELMUT SHH167A						
DAY 5						
A12C15_OL 1.28MER				↑ (p = 0.002)		↑ (p = 0.050)
A12C15 ALT 9.3	↓ (p = 0.000)		↓ (p = 0.043)			
OL_OCFL SHH193A				↑ (p = 0.030)		
DELMUT SHH011A						
DELMUT SHH045A						
DELMUT SHH167A						

Arrows indicate a significant relationship where experimental samples on the left column are compared relative to controls and experimental samples in the top row. Significance was determined to be any p-value below 0.05 by employing a student *t*-test. Any blank cell indicates there is no significant difference.

4 Chapter Four: Discussion



Both HBV and HIV are hyperendemic in sub-Saharan Africa, and there is a correspondingly high incidence of HCC in this region [1, 22, 78, 97, 151, 157], contributing to a considerable burden of disease in this region. In regions where HBV genotypes B, C and D prevail, a strong relationship exists between the rapid and more likely development of HCC and infection with preS deletion mutants of HBV [24, 35, 37, 39, 40, 51, 91, 117, 118, 159, 166-168, 172, 178, 183, 211, 212]. Similar preS deletion mutants have been detected in southern African and Indian HCC patients infected with subgenotype A1 [52, 118, 192]. Disturbingly, in a cohort study conducted in Mpumalanga, South Africa by our team, equivalent deletion-mutants were detected in 5 treatment-naïve HBV-HIV coinfected patients [52]. Thus, the aim of this study was to characterise the quasispecies of HBV in these patients and to construct plasmids containing preS deletion-mutants in a subgenotype A1 backbone, in order to functionally characterize them *in vitro*. Such studies may determine how these preS deletions contribute to the working model of hepatocarcinogenesis and explain the high hepatocarcinogenic potential of subgenotype A1 [115, 192].

4.1 Quasispecies and Phylogenetic Analysis

Sequence variability is a feature of HBV because of the error-prone nature of the HBV polymerase. Thus, in infected individuals, HBV exists as a quasispecies population [171, 213], which can be shaped by pressures exerted by host immunity, co-infection with other pathogens, vaccination or antiviral therapies [9, 155, 171]. The deletion mutants detected in the 5 treatment-naïve HBV-HIV coinfected patients by direct sequencing [52] represent the major population because only sequences comprising at least 20% of quasispecies can be detected by direct sequencing [24, 159, 171, 213]. In order to determine whether the parental sequences are the only strains in the quasispecies population, or whether they co-circulate with other variants or WT HBV, 30 clones of the complete S region of HBV isolated from each of 4 patients (SHH167A, SHH193A, SHH274A and SHH300A) were analysed.

Both inter- and intra-group nucleotide divergence were calculated in order to determine the diversity of quasispecies and to compare how the clones differed from the parental sequence, obtained by direct sequencing (see Table 3.1) [52]. The clones of HBV from

patient SHH167A showed the highest intra-group divergence of 11% (Table 3.1). A total of 9 of the 30 clones (26,7%) were distinct from the parent sequence, including 5 deletion patterns that were different from that of the parental strain (Table 3.2). These additional deletion mutants are the reason for the high intra-group divergence. On the other hand, the intra-group divergence of the complete S region of HBV isolated from the 4 remaining patients ranged from 2% to 4%. Studies characterising quasispecies in genotypes other than subgenotype A1, report similar intra-group divergences [39, 177, 213, 214] and also found that the preS2 is the most variable region of the complete S [37, 119, 213].

Sequencing of the clones of the S region amplicons from patients SHH167A, SHH193A, SHH274A and SHH300A, revealed that the majority of the clones (70% or more) were similar to their respective parental strains (Table 3.2). This corresponds with other studies, which found that the mutant population isolated from the patient sera was the predominating species [37, 39, 166]. For patient SHH167A, three clones (10%) had sequences with WT sequence (with no deletion). This is similar to the findings of others who investigated preS deletion-mutants of genotype B and C in patients with advanced liver disease [40]. Similarly, Melegari *et. al.*, found a majority population of preS deletion-mutants with a small proportion of a WT population being simultaneously present [162]. However, this was the case in a patient that was infected with genotype D following interferon therapy and followed-up at several time points, thus drug pressure was present which not the case in our study [162]. It is possible these WT strains are trans-complementing the deletion mutants, thus rescuing the major population, if replication and/or protein expression are compromised by the deletion/s [28, 37, 161-165].

In all four patients, minority populations, which differed from the major population, were detected. Five clones of HBV from SHH167A had deletion patterns, which differed from that of the parental strain. Three percent of the HBV clones from patients SHH274A and SHH300A showed deletion patterns different to the deletion pattern in the parental strain. The reverse was observed in the HBV isolates from patient SHH193A. The majority population of HBV clones had preS regions without deletions (i.e. full-length) and were homologous to the parental strain. However, one clone had a 1 nt deletion, and 5 clones had a 54 nt deletion that was not detected by direct sequencing. These results

demonstrate that deletion-mutants can co-circulate with WT or full-length virus in various proportions, representing either a major or minor population of the quasispecies.

The opposing forces of the error-prone polymerase and the compactness of the overlapping HBV genome can be factors that influence the dynamic nature of the quasispecies population within a patient [9, 114, 171]. However, the human immune system and drug pressure can act to positively select for mutations that allow evasion of detection and elimination [9, 171, 215]. As all patients selected for this study were HBV-HIV co-infected and treatment naïve, it is likely they were immunosuppressed at the time of serum collection. This may mean the immune system was not exerting selective pressure on the quasispecies population [9], which could explain why so little variation is seen in the quasispecies sequences presented here [97, 149, 152]. This is supported by the fact that CD4 counts at the time of serum collection from these patients were low (SHH167A CD4 = 49; SHH193A CD4 = 12; SHH274A CD4 = 182; SHH300A = 98) [52]. Moreover, in the African context, we know that HBV transmission often occurs at a young age, with a competent immune system at the time of infection [26, 70, 74, 77, 78, 107]. Assuming that the HIV infection occurred later than the HBV infection, and as it is known that HIV alters the natural history of HBV, immunosuppression as a result of HIV co-infection may exert the little or no selection pressures on the quasispecies, explaining the minimal variation [9, 84, 127, 128, 130, 132, 140, 149, 153-155]. Only 2 of the 4 patients, SHH167A and SHH193A, had mixed populations of deletions and full-length strains, which contrasts with a study where a very high proportion of patients had a quasispecies population comprised of various deletion mutants and WT strains, all belonging to genotype C [40]. In another study all five patients, infected with genotype C, had mixed populations [212]. However, in another this study where limited clones were sequenced (only 5 to 10 clones per patient), the deletion mutants were found to be the major population [37, 39]. Our findings have clinical relevance, as the severity of disease has been shown to be increased in patients infected with a quasispecies where escape mutants represent the major population [9].

Phylogeny allows the comparison of sequences in such a way that evolutionary past, relatedness, and trends can be established [102]. Thus, we used Neighbor-Joining and Bayesian inference phylogenetic methods to analyse the relationship of the S regions of clones to each other. The Neighbor-Joining phylogenetic method is based on the least

squares estimation of branch length and assumes a constant rate of nucleotide substitution [209]. Thus, this method assumes minimum evolution or maximum parsimony, minimizing the sum of the branch lengths at each clustering of an operational taxonomic unit, or sequence [209]. The resulting 'true' tree is made up of the most pairs of neighbors with the shortest branch length [209]. Although the trees constructed using this method may not be the most accurately representative of evolutionary pathways, this method is highly efficient in producing a tree that groups the sequences with highest similarity to each other [209].

As expected, the tree comparing the 4 groups of 30 clones, the clustering of parental sequences and subgenotype A1 references sequences, reflects the sequence variation seen in the nucleotide divergence calculations (Tables 3.1 and 3.2). The major populations clustered together in four main groups, each set of clones grouping with its parental sequence, except where significant sequence variation was present (Figure 3.4). For example, SHH167A 1.70, clustered separately from any of the clones in the Neighbor-Joining tree. This can be attributed to the deletion pattern present that was unlike any other sequence in the group, namely a 33 nt as well as a 66 nt deletion overlapping the region where the 45 nt deletion was present in the parental strain. Furthermore, clone SHH167A 1.77 contained a 63 nt deletion similar to the deletion pattern seen in SHH274A, and thus this clone clustered with the SHH274A group. The 5 clones of SHH193A, with the 54 nt deletion distinct from the parent sequence, clustered with SHH300A group, which also has a 54 nt deletion in the same position. The relative clustering of the strains from SHH167A, SHH193A, SHH274A and SHH300A, to each other was the same as seen in our previous analysis of the sequences obtained by direct sequencing [52].

In addition to the Neighbor-Joining method, the sequences were analysed phylogenetically using Bayesian inference. The Bayesian inference approach has become progressively popular as it estimates maximum posterior probability allowing more accurate estimations of evolutionary pathways compared to Neighbor-Joining and Maximum Likelihood trees [103, 113]. It is a hierarchical approach that requires specification of priors; such as evolutionary model, molecular substitution rate, selection of a molecular clock, accurate sampling dates, to infer the most likely tree by coalescent algorithms [103, 104, 113, 114]. In this study, we found there were essentially

no differences between the two trees (Fig 3.5 and 3.6). The clustering pattern was the same for both trees and the four groups of clones clustered in the corresponding pattern as in the previous Neighbor-Joining tree where only subgenotype A1 sequences were included as references. Upon investigation of the literature in this regard, it was found that there are some challenges with Bayesian inference analysis of data. The first of which being that Bayesian analysis is time-consuming compared to Neighbor-Joining methods, and while it has the power to elucidate virus evolution accurately, all priors have to be known for an accurate outcome [103, 112, 113, 215-218]. In the case of HBV, these priors are not always available. HBV demonstrates rapid evolutionary dynamics, and there is controversy over the HBV nucleotide substitution rate [112-114, 217]. This is because it is affected by the overlapping viral genome (where different ORFs may evolve at different rates), substitutions can occur in both directions, recombination events are not accounted for and finally calculation of the substitution rate affects the selection of the molecular clock [112, 113, 216]. Furthermore, the model requires accurate sampling dates and times before it can precisely extrapolate the evolution of viruses like HBV [112, 113, 216-218]. The optimum use of the Bayesian inference method in the present study was limited. Accurate sampling date and time data were not available for all the reference sequences, which were obtained from GenBank. As we only sequenced HBV isolated at baseline, we did not have multiple longitudinal time points to follow up for each patient over a long period of time. Finally, Bayesian analysis is most accurate when the sample group is large, and we only had 30 clones sequences for the four patients, which did not represent a powerful enough sample size [114, 215, 216]. This meant that an evolutionary rate or time of origin estimation of quasispecies could not be determined using Bayesian inference on this data set in the present study.

4.2 Generation of Overlength Constructs for Transfection Experiments

The most cost-effective method to establish whether the deletion mutants would be replication competent and how the deletions would affect virus protein expression and viral loads, is to construct overlength clones that can be transfected into a cell-line like Huh7 for functional characterisation. The replacement of the WT 784 bp region

corresponding to the fragment flanking the preS deletions, with those sequences from patients with variations in this region, minimised the confounding effects of any other variations in other regions of the genome. Thus, once the constructs were generated, it was important to confirm that the deletions were present and in the correct genomic context, that there was high sequence similarity to the respective parental strain, that the backbone sequence was not altered and finally that no unwanted start or stop codons were introduced in the HBV overlength sequence.

After carrying out restriction mapping, it was confirmed that the constructs were correct regarding the position and size of deletions (Figures 3.7 – 3.19 and Tables 3.3 – 3.8). Sequencing of a number of clones allowed for the selection of the sequences with the least nucleotide divergence compared to the parental strain. Bioinformatic analyses of both the nucleotide and amino acid sequences revealed that no unwanted start and stop codons, nor lethal missense mutations were introduced by the introduction of the 784 bp fragment into the subgenotype A1 replication competent clone. Some inadvertent aa changes were introduced. However, none of these mutations were in functionally important domains in the 'a'-determinant region or polymerase ORF, where such mutations could impact on antigen presentation and/or replication, respectively (Figures 3.20 – 3.22 and Tables 3.9 - 3.10). The removal of the *Xba*I restriction site and 22 nucleotides from the vector in the replication competent clone did not compromise viral replication and protein expression significantly.

As far as we are aware, we are the only group to have adopted this cloning strategy for the functional characterisation of preS deletion mutants. Studies that have carried out functional characterisation in other genotypes and subgenotypes have employed overlength clones constructed from the entire original patient sequence [24, 28, 37, 39, 117, 164, 166, 183, 219, 220]. However, the strategy employed here may provide more specificity in exploring the effect of the deletions themselves as the chance of introducing additional changes in the remainder of the HBV genome sequence is minimised.

4.3 Extracellular Protein Expression After Transfection

Expression of extracellular proteins is an important marker of viral replication and function. The ELISA used for HBsAg detection employs numerous monoclonal mouse

antibodies with slightly different antigen binding sites to ensure high specificity when binding the antigenic loops of the 'a'-determinant region [11, 35]. However, it should be noted that while ELISA is a quantitative assay for HBsAg, it cannot distinguish between virion-associated or subviral particle HBsAg [36]. This said, detection of HBsAg *in vitro* indicates that subviral particles are expressed, and/or virions are being assembled, and thus suggests that viral replication is occurring [36]. Moreover, the cloning strategy used here, where the 784 bp fragment flanking the mutations was replaced, did not affect the conformation of the antigenic loop structure, and thus should not affect HBsAg detection.

While HBeAg is not essential for viral replication, it is used as an indicator of viral replication [56]. Again, as the cloning strategy employed in the present study only replaced a 784 bp fragment flanking the mutations in the preS region, the *pre-core* ORF was unaffected and therefore HBeAg expression should not be disturbed at the transcriptional and translational levels [56, 60]. However, if the preS deletions interfere with the expression of HBsAg in the endoplasmic reticulum/endoplasmic reticulum Golgi intermediate compartment (ER/ERGIC), this may interfere with the secretion of HBeAg at that level. The ratio of SHBs to LHBs, where SHBs vastly outnumbers SHBs is important for viral assembly. Thus, if this ratio is disrupted by the preS deletions, it may negatively affect HBsAg secretion, increase retention of HBsAg in the ER and decrease virion secretion [28, 37, 41, 42, 162]. This, in turn, may block HBeAg secretion from the ER and cause considerably higher amounts of cccDNA accumulating within the cell and its nucleus [37]. It has been shown that accumulation of cccDNA can lead to HBV being directly cytopathic [184].

The ELISA experiment it was found that the deletion mutant constructs; DELMUT SHH011A, DELMUT SHH045A and DELMUT SHH167A expressed HBsAg at levels comparable to the positive controls (Figures 3.23 and 3.24 and Table 3.11). HBeAg expression was not affected by the introduction of deletions in the preS region (Figures 3.25 and 3.26), and no significantly different expression levels of HBeAg were noted across all the constructs; meaning that HBeAg expression was comparable to the positive controls for all constructs at all time points. Functional studies carried out on other genotypes of HBV, concurred with the findings of the present study that deletion mutants expressed both HBsAg and HBeAg in tissue culture [28, 37, 39, 117, 162]. The

relatively lower levels of HBeAg expressed by the deletion mutants compared to the WT constructs, was also observed in several other studies investigating genotypes other than subgenotype A1 [28, 37, 39, 117, 162].

The negligible HBsAg expression by the full-length cloning control OL_OCFL SHH193A was initially surprising, because it had been assumed that the full-length construct, which was chosen because of its sequence similarity to wild-type HBV, would behave much like the WT controls. After subsequent closer examination of the clinical data for patient SHH193A, it was discovered that the HBV-HIV co-infected patient had occult HBV infection [145]. Thus, the finding, that the subgenotype A1 occult OL_OCFL SHH193A construct produced very low HBsAg levels at all three time points (Figure 3.23), is novel, since the *in vitro* experiments mimicked the phenotype of occult infection seen in the patient *in vivo* [145, 153]. This result also indicates that the 784 bp fragment from patient SHH193A that was replaced in the WT A12C15 ALT 9.3 backbone was sufficient to drastically reduce HBsAg expression, and to produce the occult phenotype.

When HBsAg was analysed as a ratio of HBeAg, it was found that HBsAg is generally expressed earlier and HBeAg later (Figure 3.27). HBsAg expression was retarded on day 1 post transfection for the DELMUT SHH011A construct, but then increased amply on day 5. This retardation may be because this construct has an additional deletion as compared to the other two deletion mutant constructs. The additional deletion is in the preS1 region as opposed to the preS2 region, which encodes for MHBs, which is dispensable for HBV replication whereas SHBs is necessary for HBsAg secretion [24, 37-39, 162]. The negative effect of this additional preS1 deletion on HBsAg expression was also observed in another study, which characterised a preS1 deletion mutant in a genotype D overlength construct [162].

4.4 HBV Viral Loads Following Transfection

Real-Time PCR used to measure VL is a quantitative assay for the amount of HBV DNA present. To establish whether replication occurs and whether viral particles were being secreted, both the extracellular (supernatant) and intracellular (lysate or cell-associated) compartment of the Huh7 cells were tested. Furthermore, to clarify whether the encapsidated virus was being secreted, an immunocapture technique was employed on

the supernatant prior to DNA extraction and qPCR. Thus, qPCR was employed in all three compartments to establish whether the full-length cloning control and the deletion-mutant constructs would replicate once transfected into an Huh7 tissue culture model.

In a similar manner to the ELISA experiments, qPCR measurement of VL of showed that despite the replacement of the 784 bp fragment with a fragment containing deletion mutations, all constructs were producing detectable amounts of HBV DNA suggesting that viral replication was occurring (Figures 3.28, 3.29, 3.32 and 3.33 and Tables 3.12 and 3.14). Analogous to the HBsAg results, the highest VL was seen in all constructs on day 3 post transfection in both the supernatant and lysate experiments. Most notably, the VL of the occult construct, OL_OCFL SHH193A, was higher than both positive controls at all three time points and in both the supernatant and lysate compartment. This is again consistent with the clinical characteristics of false occult HBV infection where viral loads are higher than 200 IU/ml [145, 153]. Cell-associated (lysate compartment) VL was not particularly increased when compared to the VL measured in the supernatant, which indicates the ability to secrete the mature virions with no accumulation of viral DNA within the cell.

The immunocapture VL method measures the HBV DNA that is encapsidated in virions and have been secreted into the supernatant, and not any free DNA [207]. Incongruously, where the deletion mutants expressed lower levels of HBsAg and lower VL (in the supernatants or lysates) compared to the positive controls, a significantly higher VL was observed as compared to the positive controls following immunocapture (Figures 3.30 and 3.31 and Table 3.13). The occult construct (OL_OCFL SHH193A) also showed a significantly higher VL than the positive controls. These results would indicate that the changes made by replacing the 784 bp fragment in the S region were sufficient to alter the encapsidation and secretion of virions. This effect was more pronounced when the double deletion mutant construct DELMUT SHH011A was transfected. The reason for the increased VL in encapsidated virions needs further investigation.

4.5 The Way Forward

In terms of assessing whether the mutations studied here are directly linked to the development of HCC and other advanced liver disease, it would be necessary to

determine whether the mutations precede the manifestation of disease. The only way to establish this convincingly would be to implement a long-term longitudinal prospective study [91, 169, 172, 221] and to collect additional clinical data from liver biopsies and/or examination using a Fibroscan. The prospects of such studies are limited because of both financial and human resource constraints in the South African setting, where these are already overburdened by the largest ARV roll-out programme in the world. Moreover, many patients are lost to follow-up either due to their inability to attend the clinic or succumbing to illness. A shorter follow-up study is currently underway where 40 patients from the original Shongwe cohort are being followed-up at 3, 6, 12 and 18 months after the commencement of ARV treatment. One of the objectives of the follow-up study is to analyse the quasispecies population at follow-up time points using next generation sequencing, which should also yield a big enough sample group for more accurate analysis by Bayesian inference methods and follow the evolution of the deletion mutants following the initiation of ARV treatment.

The functional characterization of the preS deletion mutants is being continued. This will include Southern hybridization to characterize the replication intermediates and to convincingly show viral replication in the Huh7 cells following transfection with the deletion mutants. The expression of the envelope proteins will be studied by Western blotting as well as by immunohistochemistry combined with confocal microscopy [24, 37-39, 117, 166, 219, 220]. These experiments will shed light on whether HBV DNA and/or surface proteins of the deletion-mutants accumulate inside Huh7 cells, whether they cause apoptosis, and thus whether they are likely to cause ER stress which has been implicated in the development of HCC [24, 91, 166, 170, 181-183]. Further, it may prove useful to carry out qPCR with primers specifically to amplify up cccDNA in the lysate/cell-associated compartment of transfected cells to confirm whether there is a significant increase in the HBV DNA pool within the hepatocyte nucleus [37, 184]. It may be prudent to repeat the transfection experiments in a different cell line, such as HepG2, to establish whether the present findings are reproducible in alternative cell lines [37].

4.6 Conclusion

The quasispecies circulating in treatment naïve HBV-HIV co-infected patients infected with HBV preS deletion mutants, have been molecularly characterised. It was found that the preS deletion mutants represented the majority population, as 70% or more of clones in each of the four patients were sequentially highly similar to the respective parental strains. In addition to the major populations in each patient, minor populations were identified in the quasispecies. It appears we are the first group to characterise the quasispecies in patients who are co-infected with HIV and HBV, where the patients are infected with subgenotype A1.

The overlength subgenotype A1 WT replication competent plasmid [193], was successfully altered to remove an unwanted *Xba*I site. This altered construct served as a positive control to test whether the removal of this restriction site would affect replication and viral protein expression of this construct compared to the original plasmid. This plasmid was then successfully used as the backbone to construct three overlength deletion mutant constructs, which were shown to be replication competent. Finally, the altered backbone was used to successfully construct an overlength construct that contains the 784 bp fragment derived from a patient with occult HBV infection and transfection with this construct mimicked the phenotype seen *in vivo.* As far as we are aware, we are the first group to have constructed plasmids with deletion mutants and an occult mutant in a subgenotype A1 backbone and to show that they express viral proteins and viral DNA following transfection into Huh7 cells. These constructs will be an important resource for further research into the high hepatocarcinogenic potential of subgenotype A1, the genotype A strain predominating in Africa.

5 Appendices

5.1 Appendix A – Detailed Protocols

5.1.1 DNA Extraction Protocol

- *Care was taken to follow the manufacturers' instructions. When carried out correctly, this kit should yield purified total DNA (genomic + viral + any other DNA present in the sample), which does not contain proteins, nucleases or other contaminants. There are elements of serum that can act as inhibitors to the amplification process of PCR; this kit should also strip away such elements*
- *Aliquot out appropriate volumes of each solution. Equilibrate all necessary solutions to room temperature and set heating block to 56°C*
- *Pipet 20 µl Qiagen Proteinase K into the bottom of a 1.5 ml microcentrifuge tube*
- *Add 200 µl serum sample to the microcentrifuge tube, if you want to rid the sample of RNA, add 4 µl of RNase A stock solution (100 mg/ml) to the sample before adding buffer AL*
- *Add 200 µl buffer AL to the sample, mix contents thoroughly using a pulse vortex for 15 seconds (Do not add Qiagen Proteinase K directly to buffer AL)*
- *Incubate the mixture at 56°C for 10 minutes and then briefly centrifuge the 1.5 ml tube to remove drops from the inside of the lid*
- *Add 200 µl Ethanol (96-100%) to the sample and mix again by pulse-vortexing for 15 seconds. Again briefly centrifuge the 1.5 ml tube to remove drops from inside the lid*
- *Carefully apply the mixture from step 5 to the QIAamp spin column (in a 2 ml collection tube) without wetting the rim, close the cap and centrifuge at 6 000 x g or 8 000 RPM for 1 minute. Discard the collection tube containing the filtrate and place column in a new collection tube*
- *Carefully open the column and add 500 µl of buffer AW1 without wetting the rim, close the cap and spin for 1 minute at 6 000 x g or 8 000 RPM. Discard the collection tube containing the filtrate and place column in a new collection tube*
- *Carefully open the column and add 500 µl of buffer AW2 without wetting the rim, close the cap and spin at full speed (20 000 x g or 14 000 RPM) for 3 minutes, replace column into a new collection tube and spin again at full speed for 1 minute (to prevent any carryover of buffer AW2)*
- *Place the QIAamp spin column in a new sterile 1.5 ml microcentrifuge tube. Carefully open the column and add 75 µl of buffer AE, warmed in the water bath. Incubate at room temperature for 1-5 minutes and the centrifuge at 6 000 x g or 8 000 RPM for 1 minute. Measure DNA concentration by NanoDrop and then store at -20°C*

1.1.1. PCR Amplification

- *For all PCR reactions the hood was first prepared by wiping the surfaces with bleach and then 70% ethanol, then all remaining DNA was destroyed using the UV lamp in the hood for 15 minutes*
- *All required PCR tubes were autoclaved beforehand and any necessary racks were cleaned in a diluted bleach solution*
- *For the Hot Start PCR reactions, the Qiagen HotStarTaq® Mastermix was used, thus there was no need to add dNTPs separately. However this Mastermix ensures that the final dNTP concentration is 200 µM of each dNTP. In both PCR protocols the positive control consisted of DNA extracted from a sera sample that was known to be HBV positive. Whereas, the negative controls would include the DNA extraction from a known HBV negative serum sample, as well as a best quality water only control where no DNA whatsoever was added. In both cases they were added instead of template DNA.*
- *A 20 µmol/l primer solution was prepared by adding 5 µl of the 200 µmol/l stock to 45 µl of Best Quality Water*
- *For the Roche High Fidelity Expand kit the 50 µl aliquot of 10mM dNTP solution had to be prepared as follows:*
    - *In a 1.5ml (Eppendorf) tube, mix the following to get 2.5x dNTP stock:*
    - *250ul 100 mM dATP*
    - *250ul 100 mM dTTP*
    - *250ul 100 mM dGTP*
    - *250ul 100 mM dCTP.*
    - *Aliquot 500 µl into two 1.5 ml tubes.*
    - *Add 750 ul ddH_2O to each of the two tubes to dilute to 10mM.*
    - *Aliquot 10 mM dNTP into 50 tubes, 50 µl per tube. (Note: Small aliquots are made in order to reduce the number of times that dNTP mixes are thawed and frozen, which degrade the nucleotides.)*
    - *Store @ -20°C.*
- *The Roche High Fidelity Expand kit's buffer includes $MgCl_2$ in 25 mM concentration, thus final concentration of $MgCl_2$ in PCR reactions was 1.5 mM*
- *See Table 5.1 below for details on primer sequences, which were originally designed by Vermeulen et al., [86]. See Table 5.2 for detailed content of PCR reactions for both Hot Start and High Fidelity Expand protocols*

Table 5.1: Complete S Gene Primer Sequences

	Primer	Primer sequence (3'-5')	Region in HBV genome	Amplicon size
1st round	S1F	TCA ATC GCC GCG TCG CAG AAG ATC TCA ATC	2401 -2439	2113 bp
	S1R	TCC AGA CCK GCT GCG AGC AAA ACA	1314 – 1291	2113 bp
2nd round	S2F	AAT GTT AGT ATT CCT TGG ACT CAT AAG GTG GG	2451 – 2482	2055 bp
	S2R	AGT TCC GCA GTA TGG ATC GGC AGA GGA	1286 - 1254	2055 bp

The region in the genome refers to the relative position of the primers to the EcoRI site within the HBV genome. These primers were originally designed for the purposes of amplifying the complete S region by Vermeulen et al., [86]

Table 5.2: Master Mixes used for Qiagen HotStart PCR reactions and subsequent Roche High Fidelity PCR reactions for amplification of the entire S region

Qiagen HotStart PCR			
First Round PCR Reaction		Second Round PCR Reaction	
Sample Content	Quantity in µl	Sample Content	Quantity in µl
Forward primer (S1F)	0.375	Forward primer (S2F)	0.75
Reverse primer (S1R)	0.375	Reverse primer (S2R)	0.75
Qiagen Master Mix	12.5	Qiagen Master Mix	25
Template DNA	2.5	Template DNA	5
Best Quality Water	9.25	Best Quality Water	18.5
Total Volume	**25**	**Total Volume**	**50**
Roche Expand High Fidelity PCR System			
First Round PCR Reaction		Second Round PCR Reaction	
Forward primer (S1F)	0.375	Forward primer (S2F)	0.75
Reverse primer (S1R)	0.375	Reverse primer (S2R)	0.75
Roche Expand Enzyme	0.125	Roche Expand Enzyme	0.25
Roche Expand Buffer	2.5	Roche Expand Buffer	5.0
dNTPs	0.5	dNTPs	1.0
Template DNA	2.5	Template DNA	5.0
Best Quality Water	18.625	Best Quality Water	37.25
Total Volume	**25**	**Total Volume**	**50**

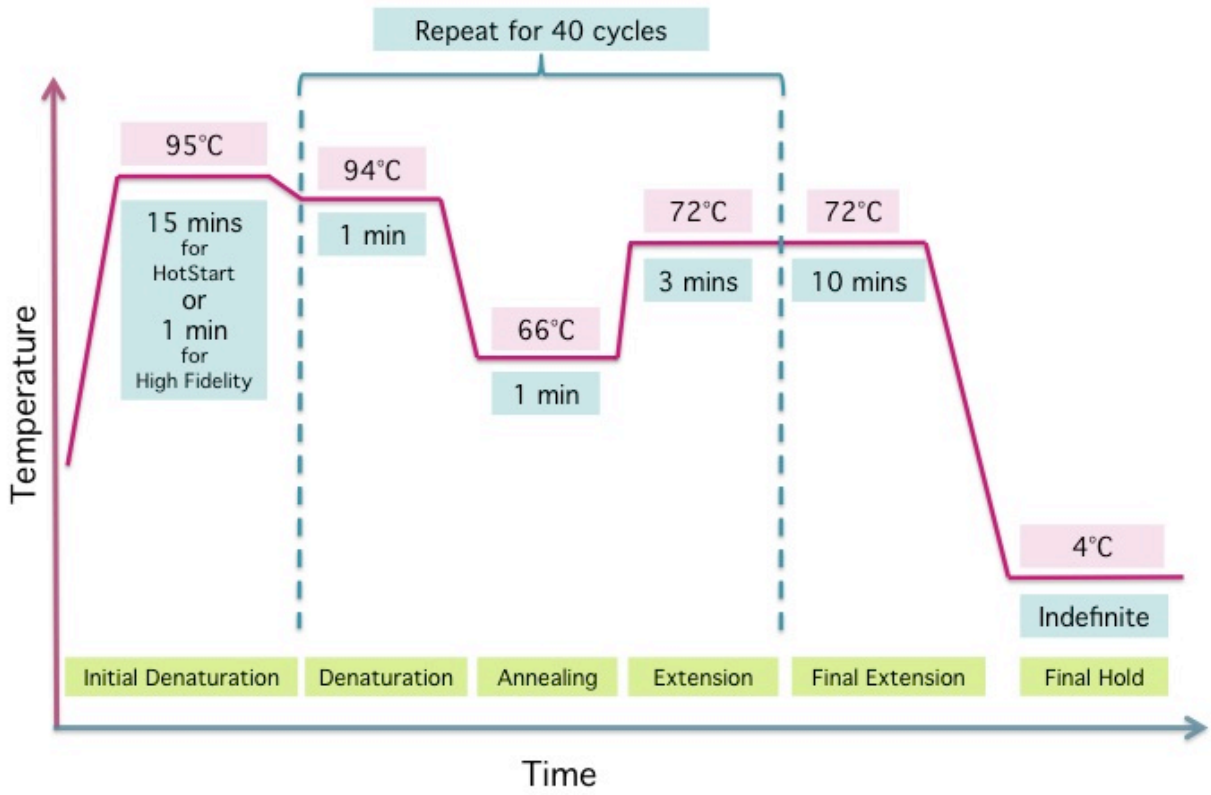


Figure 5.1: PCR ThermocyclingCconditions

All PCR product were visualised by running them on a 1% (w/v) agarose gel for 1.5 hours at 75 V. 5 µl 1 kb DNA Ladder was run alongside. They were then visualised and captured on the Gel Doc system.

5.1.2 Gel Electrophoresis and DNA Amplicon Purification from Gel

- *Agarose gels were usually prepared to be 1 % Agarose, they were run at 70-80V for 2 hours, depending on whether the DNA had travelled far enough the length of the gel. Mostly gels were stained with ethidium bromide. However when DNA was intended to be used for downstream cloning, gels were stained with Novel Juice instead of ethidium bromide as to avoid damage to the DNA. This is because ethidium bromide intercalates with DNA, and is not easily removed upon purification.*
- *To Prepare a 1 % Agarose Gel*
    - *Add 1 g of Agarose for every 100ml of 1x TAE Buffer, heat in microwave until Agarose is completely dissolved. Allow to cool. Prepare gel tray and select desired comb, place comb in the tray. Add 3 µl of ethidium bromide to every 100 ml of Agarose solution. Pour into tray and allow to set. Similarly a 2% gel would be 2 g Agarose for every 100ml of 1x TAE Buffer.*

- *Gel Purification Protocol*
    - *Excise the DNA fragment from the agarose gel using razor blade or scalpel and transfer it into a sterile Eppendorf tube. Note: The amount of agarose excised from the gel should be as small as possible and should not exceed 150 µl (150 mg) per tube.*
    - *Add 3 volumes of ADB to each volume of agarose excised from the gel (e.g. for 100 µl (mg) of agarose gel slice add 300 µl of ADB). Not For DNA fragments <8 kb, following the incubation step, add one additional volume (equal to that of the gel slice) of water to the mixture for better DNA recovery (e.g. 100 µl agarose, 300 µl ADB and 100 µl water).*
    - *Incubate at 37-55 °C for 5-10 minutes until the gel slice is completely dissolved.*
    - *Transfer the melted agarose solutions to the wells of the Zymo-Spin™ Column and place within the collection tube*
    - *Centrifuge for 1 minute at 10 000 RPM until the sample mixtures have been completely filtered. Discard the flow-through*
    - *Add 200 µl of DNA Wash Buffer to each tube. Centrifuge for 1 minute at 10 000 RPM. Repeat the wash step, but centrifuge for 1.5 minutes.*
    - *Add ≥15 µl preheated DNA Elution Buffer or water directly to the column matrix in each well. Transfer the Zymo-Spin column into a fresh Eppendorf tube, incubate at 60°C for 1 minute and centrifuge for 3 minutes to elute the DNA.*
    - *Ultra-pure DNA is now ready for use.*

5.1.3 Preparing Chemically competent cells

It is important to choose a strain of Escherichia coli (E. coli) that will lend itself to the needs of your cloning protocols.

After doing a comparison of the strains of E. coli according to genotype, it was decided to use SCS110 this project as it is a strain which is has dam (DNA adenine methylase) and dcm (DNA cytosine methylase) mutations which allows cleavage with methylation sensitive enzymes. This was important during this project because XbaI was the enzyme primarily employed for cloning as well as screening during the cloning processes employed during this project, and XbaI is a methylation sensitive enzyme.

This strain also allows blue/white screening due to the N-terminal deletion (lacZΔM15), which permits the α-complementation segment present on pUC-based plasmids, to make a functional lacZ protein. Thus empty vectors will appear to have a blue colour, whereas those containing a cloned insert will be white.

Finally SCS110 cells are easily transformed by electrochemical methods, and will produce relatively high transformation efficiency even with large plasmids.

Table 5.3 below outlines genotypes of commercially available used E. coli strains that were considered for this study.

Table 5.3: Choosing Escherichia coli strain by Genotype

Strain Name	*Stockist*	*Genotype*
XL1-BLue	*Stratagene*	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIqZΔM15 Tn10 (Tetr)]*
XL1-Blue MR	*Stratagene*	*Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac*
XL1-Blue MRF	*Stratagene*	*Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac {F´ proAB lacIqZΔM15 Tn10 (Tetr)}*
XL2-Blue	*Stratagene*	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac {F´ proAB lacIqZΔM15 Tn10 (Tetr) Amy Camr}**
SCS110	*Stratagene*	*rpsL (Strr) thr leu endA thi-1 lacY galK galT ara tonA tsx dam dcm supE44 Δ(lac-proAB) {F´ traD36 proAB lacIqZΔM15}*
JM110	*Stratagene*	*rpsL (Strr) thr leu thi-1 lacY galK galT ara tonA tsx dam dcm supE44 Δ(lac-proAB) [F´ traD36 proAB lacIqZΔM15]*
Top10	*Invitrogen*	*F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG*
Top10F	*Invitrogen*	*F´{lacIq Tn10 (TetR)} mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL endA1 nupG*
DH5α	*Invitrogen*	*F- φ80lacZΔM15 Δ(lacZYA-argF)U169 recA1 endA1 hsdR17(rk-, mk+) phoA supE44 thi-1 gyrA96 relA1 tonA (confers resistance to phage T1)*
SURE	*Stratagene*	*e14–(McrA–) Δ(mcrCB-hsdSMR-mrr)171 endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC::Tn5 (Kanr) uvrC [F´ proAB lacIqZΔM15 Tn10 (Tetr)]*

PIPES Protocol for Preparing Chemically Competent E. coli:

- *Inoculate 10 ml of LB medium with 100 µl of untransformed DH5α/SCS110/XLBlue E. coli stock, or previously prepared competent bacteria of the necessary strain and shake overnight in a bacterial incubator at 37°C (120 RPM). OR spread stock the required strain on an appropriate LB plate, grow at 37°C overnight, and inoculate 10ml of media with one colony*
- *Split the overnight culture into two 50ml tubes and make up to 50ml with fresh LB medium, culture bacteria for 90 minutes (@ 37°C, 120 RPM) until bacteria are in mid log phase (OD_{600} of 0.3-0.5) OR Innoculate 50 ml of SOB with 50 µl of the overnight culture, grow at 37°C, shaking at 120 RPM till cells are in mid-log phase (OD_{600} of 0.3-0.5), thus ±5hours*
- *Pellet the bacteria (4x25 ml per 50 ml falcon tube) by centrifuging at 4500 RPM for 15 minutes at 4°C and remove supernatant*
- *Gently resuspend the bacterial pellet from each 25 ml tube in 2.5 ml PIPES buffer and place on ice for 15-30 minutes. Note: the bacteria in PIPES buffer are extremely fragile and should be treated gently making sure not to keep them on ice for too long*
- *Pellet the bacteria (4x25 ml per 50 ml falcon tube) by centrifuging at 2500 RPM for 10 minutes at 4°C and remove supernatant*
- *Gently resuspend each bacterial pellet in 1ml PIPES buffer, aliquot 100 µl of the now competent bacteria into 2 ml capped sterile tubes. The PIPES buffer contains 15% glycerol thus enabling storage at -70 to -80°C for up to 6 months*
- *Test transformation efficiency using a standard high replication plasmid such as pTZ57R*

5.1.4 Transformations

Using the InsTAclone Kit: Transformation Protocol from Overnight Bacterial Culture

This protocol is sufficient for 2 transformations; if more are required the protocol should be adjusted accordingly.

- *The day prior to the transformation, seed an overnight culture (of the required bacterial cell line either SCS110 of DH5α) by inoculating 2 ml of C-medium with a single bacterial colony. Make sure to use freshly streaked bacterial colonies (not older than 10 days). Incubate the culture overnight at 37°C in a shaker.*
- *2 ml of overnight culture is sufficient for 26 transformations. The culture can be kept at 4°C for one week and used for preparation of competent cells.*
- *The day of transformation ensure culture tubes are pre-warmed and contain the required amount of C-medium (1.5 ml for each 2 transformations) at 37°C for at least 20 min. Pre-warm LB agar plates in a 37°C incubator for at least 20 min before plating. These should be supplemented with ampicillin, X-Gal and IPTG.*
- *Prepare the T-solution by thawing T-solution (A) and T-solution (B), and mixing contents thoroughly. Combine 250 µl of T-solution (A) and 250 µl of T-solution (B) in a separate tube and keep on ice.*
- *Add 150 µl of the overnight bacterial culture to 1.5 ml of pre-warmed C-medium and incubate with shaking for 20 min at 37°C.*
- *Pellet bacterial cells from this overnight culture by centrifugation for 1 min at 5000 RPM, discard the supernatant.*
- *Resuspend the pelleted cells in 300 µl of T-solution. Incubate on ice for 5 min.*
- *Centrifuge for 1 min at 6000 RPM in a microcentrifuge, discard the supernatant.*
- *Resuspend pelleted cells in 120 µl of T-solution. Incubate 5 min on ice.*
- *The ligation mixture is now added ina volume of 2.5 µl (containing 14 ng vector DNA) or 1 µl of supercoiled Control DNA (10-100 pg) into new microcentrifuge tubes. Chill on ice for 2 min.*
- *Add 50 µl of the prepared cells that are on ice to each tube containing DNA, mix and incubate again on ice for 5 min.*
- *Plate immediately on pre-warmed LB-ampicillin X-Gal/IPTG agar plates. Incubate overnight at 37°C.*
- *Document the plates and controls by doing colony counts and taking photographs on the GeneDoc*

Transformation Using Chemically Competent Cells

- *Pre-chill Eppendorf tubes on ice. (Some tubes are for the experimental transformations and one tube is for the pTZ57R positive control). Preheat SOC medium to 42°C.*
- *Thaw the chemically competent cells on ice. When thawed, gently mix and aliquot 100 µl of cells into each of the two pre-chilled tubes. Add 1.7 µl of β-mercaptoethanol to each aliquot of cells. Gently swirl the tubes to mix. Incubate the cells on ice for 10 minutes, swirling gently every 2 minutes.*
- *Add 0.1–50 ng of the experimental DNA (see Quantity and Volume of DNA, reverse page, for guidelines) to one aliquot of cells and add 1 µl of the pUC18 control DNA to the other aliquot. Swirl the tubes gently.*
- *Incubate the tubes on ice for 30 minutes.*

- *Heat-pulse the tubes in a 42°C water bath for 45 seconds. The duration of the heat pulse is* **critical** *for maximum efficiency.*
- *Incubate the tubes on ice for 2 minutes.*
- *If the antibiotic resistance gene is anything other than ampicillin, add 0.9 ml of preheated SOC medium and incubate the tubes at 37°C for 1 hour with shaking at 225–250 rpm, to allow time for antibiotic expression.*
- *Plate ≤200 µl of the transformation mixture on LB agar plates containing the appropriate antibiotic‡ (and containing IPTG and X-gal if colour screening is desired). For the pTZ57R control transformation, plate 200 µl of the transformation on LB–ampicillin agar plates.*

Note: Cells may be concentrated by centrifuging at 1000 RPM for 10 minutes. Resuspend the pellet in 200 µl of SOC medium. If plating <100 µl of cells, pipet the cells into a 200 µl pool of SOC medium and then spread the mixture with a sterile spreader. If plating ≥100 µl, the cells can be spread on the plates directly. Tilt and tap the spreader to remove the last drop of cells. Some β-galactosidase fusion proteins are toxic to the host bacteria. If an insert is suspected to be toxic, plate the cells on media without X-gal and IPTG. Color screening will be eliminated, but lower levels of the potentially toxic protein will be expressed.

- *Incubate the plates at 37°C overnight (at least 17 hours for blue-white color screening). Colonies containing plasmids with inserts will remain white, while colonies containing plasmids without inserts will be blue. The blue color can be enhanced by incubating the plates for two hours at 4°C following the overnight incubation at 37°C.*
- *For the pUC18 control, expect 100 colonies (≥5 °— 106 CFU/µg pUC18 DNA). For the experimental DNA, the number of colonies will vary according to the size and form of the transforming DNA, with larger and non-supercoiled DNA producing fewer colonies.*

5.1.5 Cracking

- *An aliquot of 500 µl of the bacterial culture was placed into a sterile 1.5 ml Eppendorf tube, and centrifuged at 12 000 RPM for 1 minute, then the supernatant was discarded.*
- *The pellet was resuspended in 100 µl of Cracking Buffer (Appendix B) and vortexed. The resuspended bacterial cells were centrifuged at 12 000 RPM for 20 minutes.*
- *The bands resulting from the cracking reaction were visualised on a 1% agarose gel by loading 30 µl of the supernatant. 5 µl 1 kb DNA Ladder was run alongside. Importantly samples had to be loaded onto a "dry" 1% (w/v) agarose gel, meaning the running buffer in the tank was lower than the top edge of the agarose gel but not covering the wells, to prevent the sample from floating out of the well. After the samples had run into the gel, the entire gel was submerged in running buffer, and electrophoresis for 1.5 hours at 75 V.*

5.1.6 Plasmid Mini Preparation

Sigma Solution Mini Prep with no columns.

Procedure:

- *8ml of cells were grown overnight with shaking at 120 RPM and 37°C, transfer 2ml of culture into Cryo tube for glycerol stocks*
- *Pellet the remaining 6ml of culture by microfuging for 5 min at 8 000 RPM, discard supernatant*
- *Resuspend the pellet in 200 µl Solution I*
- *Add 200 µl Solution II, invert gently to mix and leave at room temperature for 5min*
- *Add 350 µl Solution III, shake to mix and incubate on ice for 10 min*
- *Microfuge to pellet the cell debris at room temperature for 10 min at 13 000 RPM*
- *Decant the supernatant in to a sterile Eppendorf with a pipette*
- *Add 600 µl Isopropanol and microfuge at room temperature for 30 min (13 000 RPM) to precipitate the DNA, pour off the supernatant*
- *Wash with 1 ml 70 % ethanol, microfuge at room temperature for 15 minutes (13 000 RPM), pour off ethanol supernatant*
- *Dry isolates overnight at 37°C*
- *Resuspend in 40µl of sterile water*
- *Do NanoDrop spectrophotometry on a number of the samples to check DNA concentration*
- *Store at -20°C until further use*

5.1.7 Ethanol Precipitation of DNA

- *To 180 µl of resuspended plasmid isolate add 18 µl of 3M Sodium Acetate (pH 5.2) and mix*
- *Add 600 µl of -20ºC 100% ethanol, mix and centrifuge at 13 000 RPM for 20 minutes, pour off supernatant*
- *Add 1 ml of 70% ethanol and centrifuge at 13 000 RPM for 10 mins, pour off supernatant*
- *Dry pellet in the stationary 37°C Incubator and then resuspend in 50µl*
- *DNA concentration was estimated using the NanoDrop spectrophotometer and running some linearised sample on a 1% Agarose gel*

5.1.8 Ligation

Using the InsTAclone Kit:

The InsTAclone Kit recommends that for DNA amplicons of 2000 bp in length 350 ng of DNA should be added to the ligation mixture. Thus the appropriate amount of DNA would be calculated from the concentration established by NanoDrop spectrophotometry. Table 5.4 below shows the components of a ligation reaction. All reactions were carefully mixed and then placed first on the bench at room temperature for 30 minutes and then in the thermocycler at 4°C overnight. Controls always included:

1. *Positive control (vector with a control PCR fragment) – to test efficiency of the ligation reaction*
2. *Second positive control (supercoiled circular DNA vector)*
3. *Empty vector control (where the open vector was added with no insert) - to test whether the enzyme was working*
4. *No DNA control – which would show no contamination of competent cells once they are transformed*

Note: pTZ57R – closed circular supercoiled vector

pTZ57R/T - linearized cloning vector with single 3'-ddT overhangs

Table 5.4: Components of Ligation Reaction using the InsTAclone Kit

Component:	Quantity (volume in µl):					
	Ligation Reaction	Positive Control	pTZ57R	pTZ57R/T (Empty Vector control)	No T4-DNA Ligase	No DNA Control
Vector pTZ57R/T	3	7	7	3	7	0
10x Ligation Buffer	3	3	3	3	3	3
Purified PCR Product	1-15 depending on concentration	0	0	0	0	0
Control PCR Fragment	0	4	0	0	4	0
Best Quality Water (Nuclease Free)	8-23 depending on concentration	15	19	23	15	27
T4 DNA Ligase	1	1	1	1	0	1
Total Volume	30	30	30	30	30	30

Using T4-DNA Ligase:

The DNA concentration of all DNA to be ligated was determined first by NanoDrop Spectrophotometry. For ligations the optimal insert/vector ratio is 3:1, however to be safe, a 1:1 ratio was included as well.

Controls included:

- *A pTZ57R (circular) positive control*
- *A cut and ligate control (vector only control)*
- *A no T4-DNA Ligase control*
- *A no DNA negative control*

Once reactions were set up contents were carefully mixed and then placed at 4°C in the thermocycler overnight.

Table 5.5: Typical Contents of a Ligation Reaction

Component:	Quantity (volume in µl)				
	pTZ57R positive control	Experimental Ligation	Cut +Ligate Control	No T4-DNA Ligase control	No DNA Control
Insert	0 µl	1-20 µl depending on concentration	0 µl	10 µl	0 µl
Vector	6 µl	1 µl	1 µl	1 µl	0 µl
Best Quality Water	18 µl	1-10 µl **depending on concentration**	23 µl	16 µl	24 µl
10x ligation buffer	3 µl	3 µl	3 µl	3 µl	3 µl
T4-DNA Ligase	3 µl	3 µl	3 µl	0 µl	3 µl
Total Volume	30 µl	30 µl	30 µl	30 µl	30 µl

5.1.9 Restriction Digestion

Before setting up any Restriction reaction the concentration of the DNA was determines by NanoDrop spectrophotometry. This is because 1 µl of Enzyme solution is sufficient to digest up to 1 µg of lambda DNA. Thus the between 800 ng – 1 µg of DNA was added to each restriction digestion reaction.

Table 5.6: Typical Contents of a Restriction Digest Reaction

Content:	*Experimental Digest*	*No DNA Control*
Plasmid DNA	*1 – 8 µl depending on concentration (always aim to digest 800-1000 ng/ µl)*	*0 µl*
Appropriate Buffer	*1 µl*	*1 µl*
Enzyme	*1 µl*	*1 µl*
Best Quality Water	*1 – 7 µl depending on concentration*	*8 µl*
Total Volume	*10 µl*	*10 µl*

All restriction digestions were mixed and then incubated at 37°C for one hour unless the Enzyme used required a different temperature. After restriction loading dye was added (typically 5 µl) to stop the reaction, and together were loaded onto a 1% (w/v) Agarose Gel and ran at 75V for 1.5 hours. 5 µl 1 kb DNA Ladder was run alongside. The results were visualised and images captured on the GelDoc System.

5.1.10 Plasmid Maxi Preparation

Qiagen Protocol for DNA Maxi Prep from Bacterial Culture for Cloning Purposes:

- *Cells were harvested from culture by centrifuging at 5 500 RPM for 15 minutes at 4^{0}C*
- *Cells were then resuspended thoroughly by aspirating with a pipette in 10 ml of Buffer P1 (check that RNase has been added)*
- *Cells were then lysed by adding 10 ml of Buffer P2 incubate at room temp for 5 min and then add 10 ml of Buffer P3, incubate on ice for 15min*
- *The lysate was then cleared by centrifuging at 11 000 RPM for 30 minutes at 4^{0}C, transfer supernatant to a fresh Nunc tube and centrifuging again at 11 000 RPM for 15 minutes at 4^{0}C*
- *A Qiagen column was then equilibrated by adding 4 ml of Buffer QBT, the flow through is discarded*
- *The cleared lysate was then loaded onto the column with the aid of an inoculation loop to prevent the flocculant from blocking the column, flow through was discarded*
- *The column was then washed twice with 30 ml of Buffer QC and the flow through was discarded*
- *The plasmid DNA was then eluted into a fresh tube using 15 ml of buffer QF*
- *Then 10.5 ml of isopropanol was added to the tube to precipitate the DNA, this was then carefully mixed and then centrifuged at 4^{0}C for 30 minutes at a speed of 12 000 RPM, after which the supernatant was carefully poured off*

- *Added 5 ml of 70% ethanol to each tube, mixed carefully and centrifuged for 15 minutes at 12 000 RPM at room temperature, again supernatant was carefully poured off*
- *The pellet was dried until pellet was translucent*
- *The pellet was resuspended in 50-100 µl of sterile preheated TE buffer*
- *Each tube was tapped vigourously, incubated at 60^{0}C for a few minutes, then again at 37^{0}C for a few minutes and then spun briefly*
- *The DNA concentration was estimated by NanoDrop, and DNA was stored at -20°C*

Qiagen Protocol for DNA Endo Free Maxi Prep from Bacterial Culture for Transfection Purposes:

- *Cells were harvested from culture by centrifuging at 5 500 RPM for 15 minutes at 4 °C*
- *Cells were then completely resuspended by aspirating with a pipette in 10 ml of Buffer P1 (check that RNase has been added)*
- *Cells were then lysed by adding 10 ml of Buffer P2 incubate at room temp for 5 min and then add 10 ml of Buffer P3, incubate on ice for 15 minutes*
- *During the incubation, prepare QIAfilter cartridge*
- *Pour the lysate into the QIAfilter Cartridge. Incubate at room temperature for 10 minutes, do not insert the plunger yet. Remove the cap from the QIAfilter Cartridge outlet nozzle. Gently insert the plunger into the QIAfilter Cartridge and filter the cell lysate into a 50 ml tube.*
- *Add 2.5 ml Buffer ER to the filtered lysate, mix by inverting the tube approximately 10 times, and incubate on ice for 30 minutes.*
- *Equilibrate a QIAGEN-tip 500 by applying 10 ml Buffer QBT, and allow the column to empty by gravity flow.*
- *Apply the filtered lysate from step 7 to the QIAGEN-tip column and allow it to flow through the tip. Flow through was discarded.*
- *The column was then washed twice with 30 ml of Buffer QC and the flow through was discarded*
- *The plasmid DNA was then eluted into a fresh tube using 15 ml of buffer QF*
- *Then 10.5 ml of isopropanol was added to the tube to precipitate the DNA, this was then carefully mixed and then centrifuged at 4 °C for 30 minutes at a speed of 12 000 RPM, after which the supernatant was carefully poured off*
- *Added 5 ml of 70% ethanol to each tube, mixed carefully and centrifuged for 15 minutes at 12 000 RPM at room temperature, again supernatant was carefully poured off*
- *The pellet was dried until pellet was translucent*
- *The pellet was resuspended in 50-100 µl of sterile preheated TE buffer*
- *Each tube was tapped vigorously, incubated at 60 °C for a few minutes, then again at 37 °C for a few minutes and then spun briefly*
- *The DNA concentration was estimated by NanoDrop*

5.1.11 Cell Culture Techniques

The UV lights (in hood and on tissue culture room ceiling) were switched on for 30 mins prior to starting any work in the tissue culture room, and again after completing work. The water bath was switched on to 37°C. All elements that go into the hood were sprayed with 70% ethanol before placing them inside the hood.

Thawing Cells:

Cells should be thawed rapidly, i.e. not on ice, and then slowly (dropwise) thawed in 10 ml of warm medium. Add cells to a 75 ml culture flask to avoid osmotic shock.

Passaging/Splitting Cells:

After 2 days of growth (or more depending on confluency), remove media with a pipette.

Wash cells twice with 5-7 ml of Sodium Chloride or 1% EDTA Sodium Chloride Solution, incubating at room temperature for 5 mins each time. Remove the solution after each wash.

Add 1,5 ml of Trypsin to the cells, and incubate at 37°C for 2 mins (or less). Neutralise the Trypsin by adding 4-5 ml of Complete Medium, resuspend the cells gently and transfer to a 15 ml Falcon tube. Spin cells down (at 50% for 5 minutes). Resuspend in the desired amount and split into as many flasks as is required usually 2 x 75 cm^2 flasks).

Incubate at 37°C with 5% CO_2 for as 2 days before changing medium or splitting again.

Freezing Cells:

It is recommended to freeze rapidly growing cells.

Wash cells twice with 5-7 ml of Sodium Chloride or 10% EDTA Sodium Chloride Solution, incubating at room temperature for 5 mins each time. Remove the solution after each wash.

Add 1,5 ml of Trypsin to the cells, and incubate at 37°C for 2 mins (or less).

Neutralise the Trypsin by adding 4-5 ml of Complete Medium, resuspend the cells gently and transfer to a 15 ml Falcon tube. Spin cells down (at 50% for 5 minutes).

Resuspend them in Freezing Solution 1 (F1) (20% Serum) to half the desired final volume (i.e. if you want to freeze 6 mls of cells resuspend in 3 ml), place the tubes in an ice bucket and allow cells to cool.

Once cells are cool add Freezing Solution 2 (F2) (20% Serum, 20% DMSO) dropwise (same volume as F1). Chill the cells on ice for exactly 30 mins, then transfer to pre chilled vials for freezing (rack should be chilled as well), use 1 ml per vial.

Place vials in a well insulate box and place in the -70°C freezer overnight for slow cooling. Place samples in liquid nitrogen and store at -70°C indefinitely.

Seeding Protocol:

Grow Huh7 cells to the desired confluency in a 75 cm flask, here we use 100% confluency. First trypsinise cells and spin them down, resuspend in 10ml of complete media.

For seeding at 80%

100% confluent

1 well is 3.6 cm

seed 10 wells

1 x 75 cm flask

(100 x 75) / (80 x3.6) = 7500 / 288 = 26.04

10 / 26.04 = 0.38 %

Therefore 380 µl per well.

Thus add 3.8 ml of resuspended cells to 6.2 ml of complete media, mix thoroughly and seed 1 ml per well. Always rock the plate gently to distribute cells evenly.

Similarly:

For seeding at 70%

(100 x 75) / (70 x3.6) = 7500 / 252 = 29.76

10 / 29.76 = 0.33 %

Therefore 330 µl per well.

Thus add 3.3 ml of resuspended cells to 6.7 ml of complete media, mix thoroughly and seed 1 ml per well.

For seeding at 60%

(100 x 75) / (60 x3.6) = 7500 / 216 = 34.7

10 / 34.7 = 0.28 %

Therefore 280 µl per well.

Thus add 2.8 ml of resuspended cells to 7.2 ml of complete media, mix thoroughly and seed 1 ml per well.

For seeding at 50%

(100 x 75) / (50 x 3.6) = 7500 / 180 = 42.6

10 / 42.6 = 0.22 %

Therefore 230 µl per well.

Thus add 2.3 ml of resuspended cells to 7.7 ml of complete media, mix thoroughly and seed 1 ml per well.

For seeding at 40%

(100 x 75) / (40 x 3.6) = 7500 / 144 = 52.08

10 / 52.08 = 0.11 %

Therefore 110 µl per well.

Thus add 1.1 ml of resuspended cells to 8.9 ml of complete media, mix thoroughly and seed 1 ml per well.

Cell Transfection:

Day 1

Wash cells twice with 5-7 ml of Sodium Chloride or 10% EDTA Sodium Chloride Solution, incubating at room temperature for 5 mins each time. Remove the solution after each wash.

Add 1,5 ml of Trypsin to the cells, and incubate at 37°C for 2 mins (or less). Neutralise the Trypsin by adding 4-5 ml of Complete Medium, resuspend the cells gently and transfer to a 15 ml Falcon tube. Spin cells down (at 50% for 5 minutes).

Resuspend in 10ml (preferably in a 50 ml Nunc Tube)

Use the seeding protocol described above to seed at 80%, thus cells should be at 90% confluency on the day or transfection.

Day 2

Reagents and equipment necessary for PEI transfection include:

- *10 × PEI*
- *300 mM NaCl*
- *1.5 mL sterile microfuge tubes*
- *1 mL, 200 µL and 10 µL pipette and tips*
- *Ultrapure Sabax distilled water*
- *Vortex mixer*
- *Incubator at 37°C with 5% CO_2*

To prepare 10 × PEI mix (1 mg/mL) dissolve 50 mg of PEI (Polyethyeneimine unbranched) in 50 mL of water and heating the mixture to 80°C for a few hours with periodic mixing. Filter sterilise, aliquot and store at -20°C.

PEI Transfection Procedure:

- *Seed cells (one day before transfection or on day of transfection and wait 4 hours). Check that cells have grown to 90% confluency*
- *Remove and replace the complete medium on cells before transfection.*
- *Bring all reagents to room temperature.*
- *Prepare DNA solution. Dilute DNA to be 2,5 µg in 5 µl. Prepare PEI solution. Mix DNA and PEI solutions. Add the PEI solution to the DNA solution (not the other way around). See table below for ratios. Transfection mix (PEI mix + DNA in 150 mM NaCl) should be 10% of medium ie 100 µl transfection mix in 900 µl of medium.*
- *Incubate 10 minutes at room temperature (not longer).*
- *Add transfection mix to cells (drop wise).*
- *Change transfection medium after 8-12 hours (or next morning).*

Format	15 cm	10 cm	6 cm	6-well	12-well	24-well
Volume (medium)	18 mL	9 mL	4.5 mL	1.8 mL	900 µL	450 µL
Volume (transfection mix)	2 mL	1 mL	500 µL	200 µL	100 µL	50 µL
DNA (mass) **300 mM NaCl (volume)** **H_2O (volume)**	50 µg 500 µL To 1 mL	25 µg 250 µL To 500 µL	12.5 µg 125 µL To 250 µL	5 µg 50 µL To 100 µL	2.5 µg 25 µL To 50 µL	1.25 µg 12.5 µL To 25 µL
10 × PEI (volume) **300 mM NaCl (volume)** **H_2O (volume)**	100 µL 500 µL 400 µL	50 µL 250 µL 200 µL	25 µL 125 µL 100 µL	10 µL 50 µL 40 µL	5 µL 25 µL 20 µL	2.5 µL 12.5 µL 10 µL
PEI:DNA Ratio (2:1)*	100 µg : 50 µg	50 µg : 25 µg	25 µg : 12.5 µg	10 µg: 5 µg	5 µg: 2.5 µg	2.5 µg: 1.25 µg

*Higher PEI:DNA ratios can be used.

Figure 5.2: DNA and PEI ratios for Transfection

Table 5.7: Optimised Transfection Contents

Plate Format	*12 Well Plate (Content Per Well)*
Volume of Medium	*900 µl*
Volume of Transfection Mix	*100 µl*
Total Volume	*1 ml*
DNA Mix: *DNA (mass)* *Volume of 300nM NaCl* *Volume of Best Quality Water*	 *2 µg* *25 µl* *To 50 µl*
PEI Mix: *Volume of 10 x PEI* *Volume of 300nM NaCl* *Volume of Best Quality Water*	 *5 µl* *25 µl* *20 µl*
PEI : DNA Ratio	*5 µl : 2 µg*

Day 3

Check Transfection Efficiency of the well containing the eGFP plasmid under the fluorescent microscope. Count total amount of cells in your field, then count how many are fluorescent (x/y x100 = % efficiency).

Harvest supernatant and lysates on days 1, 3 and 5 post transfection.

Add Protease Inhibitor Cocktail Tablets cOmplete mini solution (Roche) to a final concentration of 50X to the serum and stored at -70°C until further use.

5.1.12 ELISA for HBsAg

The Murex® HBsAg Version 3 (DiaSorin) ELISA is designed specifically to determine presence of HBsAg with high sensitivity. The 96 well ELISA plate is coated in mouse monoclonals that are specific for different epitopes on the 'a' determinant of HBsAg. Therefore this kit should be sensitive even for some mutant strains of HBV.

- *Supernatant that was stored at -70°C was thawed on ice until fluid. Supernatant was diluted 1:20 with neat cell culture media (DMEM GlutaMAX™ media without foetal bovine serum)*
- *The plates were taken from the fridge, and allowed to reach room temperature, and then only removed from protective bag. Also allow all reagents to come to room temperature.*
- *Determine which samples go on the plate, their positioning, and the positioning of the controls. Only use the wells required for your number of samples and controls. For all experimental samples, three biological replicates and two technical replicates were loaded onto the plate.*

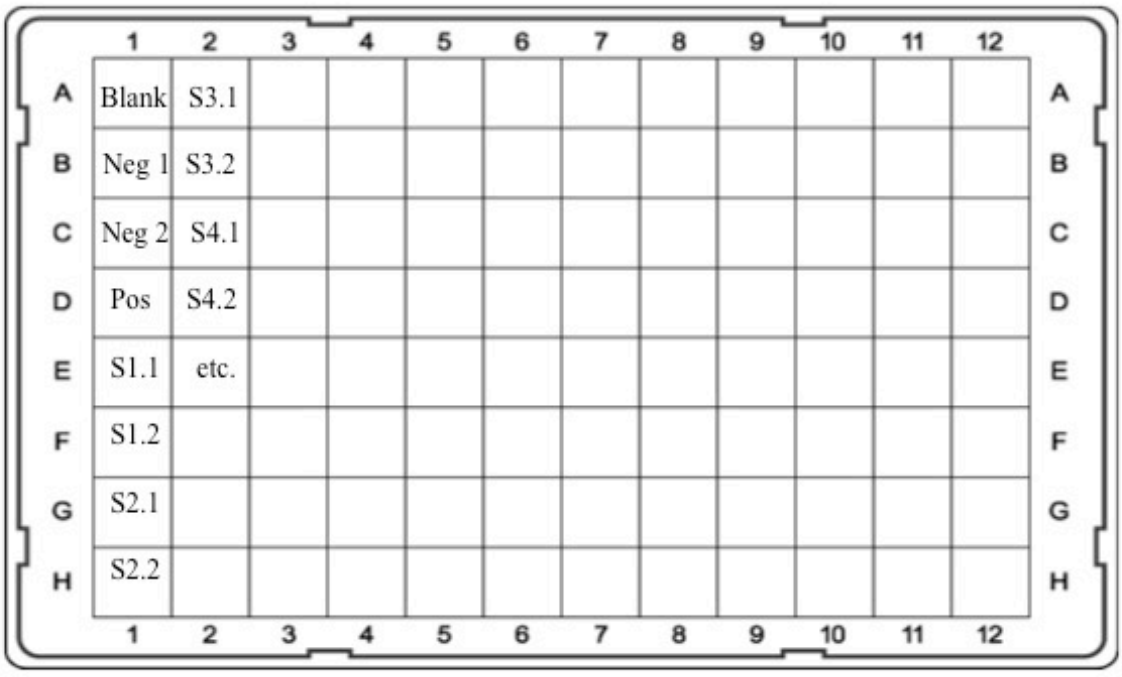


Figure 5.3: Plate Layout for HBsAg ELISA

The blank contains no incubation buffer, and no serum ie substrate only. The Negative control (Neg) is human serum/plasma non-reactive for HBsAg. The Positive control (Pos) is deactivated human serum that is reactive/positive for HBsAg.

- *Prepare the Substrate and Wash Solutions*
    - *To the Substrate Solution add a volume of colourless Substrate Diluent to an equal volume of pink Substrate Concentrate in either a clean glass or plastic vessel. Prepare this solution just prior to use. It is important that this order of addition is followed and that any pipettes and glassware used to prepare Substrate Solution are clean. Alternatively, the Substrate Solution may be made by pouring the entire contents of the bottle of Substrate Diluent into the bottle of Substrate Concentrate. One bottle of Substrate Solution provides sufficient reagent for at least five plates. The Substrate Solution should be pink; if it is purple before being used, it should be discarded and fresh Substrate Solution prepared.*
    - *For the Wash Solution - Add one volume of Concentrated Wash Fluid (Glycine/Borate wit 0.2% Bronidox) to 19 volumes of distilled or deionised water (best quality water) to give the required volume.*
- *To each well add, first 25 µl of Sample Diluent (mix by shaking before use) and then 75 µl of sample/control. Do not add any Sample Diluent or sample to the blank well. Controls and samples were aliquoted in duplicate (i.e. two wells each) onto the plate. Add samples first and then controls, which include:*
    - *Positive Control – included in Kit, contains human sera positive for HBsAg*
    - *Negative Control – included in Kit, contains human sera that is non-reactive for HBsAg*
- *Cover the plate with provided lid and allow to incubate at 37°C for 1 hour*
- *Add 50 µl of Conjugate (invert bottle to mix before use) to each well except the blank well, add the lid, and shake the plate on a shaker for 10 seconds*
- *Incubate for 30 minutes at 37°C*
- *Wash the plate 5 times by filling (±500 µl) all active wells with the Wash Solution, aspirate the liquid and allow to soak (±30 seconds). After the last round of washing invert the plate onto absorbent paper (blot), taking particular care on the last wash to tap out any residual Wash Solution.*
- *Immediately add 100 µl Substrate Solution to each active well, cover plate with lid and incubate at 37°C for 30 minutes while colour develops. Reactive wells should have a purple colour.*
- *After the incubation add 50 µl of Stop Solution to each well, reactive wells should have a yellow colour.*
- *Blank the plate reader on air (no plate in cartridge). Within 15 minutes of adding the Stop Solution, place the plate in the plate reader, and measure the absorbance of each active well at 450 nm. Use 620 nm to 690 nm as a reference wavelength.*

CALCULATION OF RESULTS

Each plate must be considered separately when calculating and interpreting results of the assay. Approved software may be used for calculation and interpretation of results.

Quality Control

The absorbance of the negative control (or average of the negative controls) must be less than 0.2. The absorbance of the positive controls must be above the average of the negative controls plus 0.8 (i.e. the absorbance of the positive control is more than 0.8 above the absorbance of the negative control)

The assay is only valid if these criteria are met.

Negative Control

Calculate the mean absorbance of the replicates of the Negative Control.

If one of the Negative Control wells has an absorbance more than 0.03 above the other discard the higher value.

Cut-off Value

Calculate the cut-off Value by adding 0.05 to the mean of the Negative

Control replicates.

Example

Negative Control absorbance: well 1 = 0.071, well 2 = 0.075

Mean Negative Control = (0.071 + 0.075)/2 = 0.073

Cut-off Value = 0.073 + 0.05 = 0.123

Note:

The term biological replicates describes three separate samples of the same construct that were treated with the exact same experimental conditions but independently of one another. Whereas, technical replicates are when the exact same sample, collected after all experimental techniques are complete is analysed multiple times.

5.1.13 ELISA for HBeAg

The ETI-EBK PLUS (HBeAg) (DiaSorin) ELISA kit is designed specifically to determine presence of HBeAg with high sensitivity. The 96 well ELISA plate is coated in mouse monoclonals that are specific for the HBeAg.

- *Supernatant that was stored at -70°C was thawed on ice until fluid. Supernatant was diluted 1:20 with neat cell culture media (DMEM GlutaMAX™ media without foetal bovine serum)*
- *The plates were taken from the fridge, and allowed to reach room temperature, and then only removed from protective bag. Also allow all reagents to come to room temperature.*
- *Determine which samples go on the plate, their positioning, and the positioning of the controls. Only use the wells required for your number of samples and controls. For all experimental samples, three biological replicates and two technical replicates were loaded onto the plate.*

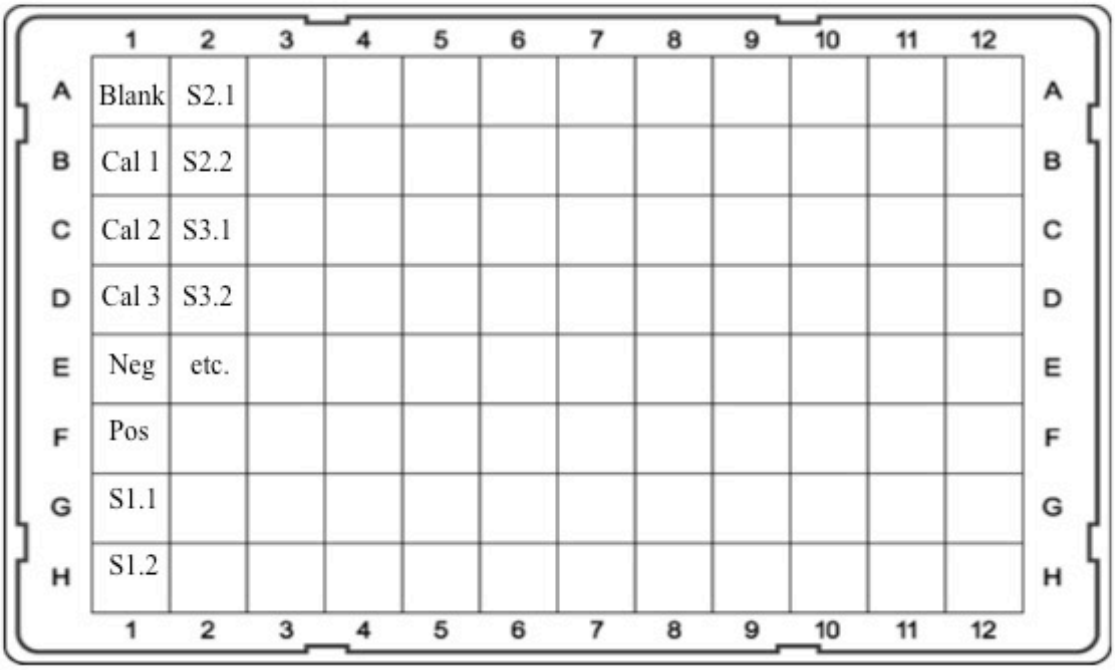


Figure 5.4: Plate Layout for HBeAg ELISA

The blank contains no incubation buffer, and no serum ie substrate only. The three calibrators (Cal 1, 2 and 3) are three replicates of a human serum/plasma non-reactive for HBeAg. The Negative control (Neg) is human serum/plasma non-reactive for HBeAg. The Positive control (Pos) is recombinant HBeAg expressed in E. coli.

- *To each well add, first 50 µl of Incubation Buffer, except for the blank well (mix by shaking before use) and then add 100 µl of sample/calibrator/control. The Positive and Negative Controls are loaded in singlicate. Calibrators were loaded in triplicate and samples were aliquoted in duplicate (i.e. two wells each) onto the plate. Add samples first and then controls and calibrators. Incubation buffer should turn green upon adding the sample/calibrator/control.*
- *Cover the plate with provided cardboard sealer, tap the plate gently to get rid of air bubbles and mix (can put on shaker for 30 seconds), and allow to incubate at 37°C for 2 hours.*
- *Just before the end of this incubation step prepare the Enzyme Tracer Solution (or Conjugate), prepare by diluting 1:50 with tracer diluent (i.e. add about 100 µl conjugate + 4.9 ml tracer diluent).*
- *Remove the cardboard sealer and discard it, wash the plate 5 times by filling (±300 µl) all active wells with the Wash Solution, aspirate the liquid and allow to soak (±30 seconds). After the last round of washing invert the plate onto absorbent paper (blot), taking particular care on the last wash to tap out any residual Wash Solution.*
- *Add 100 µl of Enzyme Tracer Solution to each well except the blank well, add a new the cardboard sealing lid, and shake the plate on a shaker for 10 seconds. Incubate for 1 hour at 37°C*
- *Remove the cardboard sealer and discard it, wash the plate 5 times by filling (±300 µl) all active wells with the Wash Solution, aspirate the liquid and allow to soak (±30 seconds). After the last round of washing invert the plate onto absorbent paper (blot), taking particular care on the last wash to tap out any residual Wash Solution.*
- *Add 100 µl chromogen/substrate solution into all active wells, and incubate for 30 minutes at room temperature (but protect from light.*

- *Add 100 µl Stop Solution to each well*
- *Measure absorbance immediately after adding Stop Solution at 450/630 nm, subtract the 630 nm reading from the 450 nm reading.*

CALCULATION OF RESULTS

Each plate must be considered separately when calculating and interpreting results of the assay. Approved software may be used for calculation and interpretation of results.

Quality Control

- *Blank absorbance must be between 0.0000 and 0.150*
- *The Mean absorbance for the calibrator must be less than 0.120 and greater than -0.020 nm*
- *The absorbance of the negative control must be less than 0.120 and greater than -0.020 nm*
- *The absorbance of the positive control must be less than 2.500 and greater than 0.500 nm*
- *The difference between the positive control absorbance and negative control absorbance must be greater than 0.450 nm (PC – NC > 0.450). If not the run is invalid and should be repeated.*

Cut-off Value

Calculate the cut-off Value by adding 0.060 to the mean of the calibrator replicates.

Cut-off Value = Mean (Cal) + 0.060

Example

Calibrators: Cal 1 = 0.071, Cal 2 = 0.075, Cal 3 = 0.069

Mean Negative Control = (0.071 + 0.075 + 0.069)/3 = 0.072

Cut-off Value = 0.072 + 0.05 = 0.122

Note:

The term biological replicates describes three separate samples of the same construct that were treated with the exact same experimental conditions but independently of one another. Whereas, technical replicates are when the exact same sample, collected after all experimental techniques are complete is analysed multiple times.

5.1.14 Supernatant and Cell Lysate Harvesting Protocol

For Supernatants:

Collect the supernatant, add to an equal volume of proteinase inhibitor buffer (which was made up as 1 tablet per 25 ml ultra pure water, i.e. final concentration is 1X PIC Buffer) in a Eppendorf tube. Transfer to -70°C for storage until DNA Removal Protocol was executed

For Cell Lysates:

Pre-cool the centrifuge to 4°C. Add 500 µl 1% EDTA Sodium Chloride Solution and 75 µl Trypsin EDTA to each of the 12 wells on the transfection plates, after supernatant has been removed. Incubate at 37°C for 5 minutes, then neutralise the trypsin by adding 600 µl of complete DMEM media. Transfer to a clean Eppendorf tube. Pellet the cells by centrifuging for 5 mins 800 RPM at 4°C. Wash 3 times with 1 ml of ice-cold Phosphate Buffered Saline (PBS). Add 200 µl of chilled RIPA buffer with 200 µl of PIC Buffer, incubate on ice for mins. Sonicate for 1 min the samples to break cells in rounds of 10 seconds, keep the samples on ice during sonication. Centrifuge at 12 000 RPM for 12 mins at 4° to pellet cell debris, transfer supernatant to a fresh Eppendorf tube without disturbing the pellet. Transfer to -70°C for storage until DNA Removal Protocol was executed

5.1.15 Pre Real-Time PCR Plasmid DNA Removal Protocol

Thaw supernatants or cell lysates on ice for 30 minutes. Centrifuge at 22 000g or 12 000 RPM for 5 minutes at 4°C, resuspend in best quality water (130 µl). Treat the supernatant with DNase I recombinant RNase-free Nuclease to remove double stranded plasmid DNA, and Mung Bean Nuclease to degrade single stranded DNA as indicated in Tables 5.8 and 5.9.

Tables 5.8 and 5.9: Components of Nuclease Digestion Reactions

Component	*Quantity in µl*
Supernatant	*130*
10x Incubation Buffer	*7,5*
10U/ul DNase I	*1,5*
10x Mung Bean Buffer	*7,5*
80U/µl RNase A (Mung Bean Nuclease)	*1,5*
Nuclease Free Water (Best Quality Water)	*2*
Total Volume in µl	***150***

Alternatively when using NEB Mung Bean Nuclease and DNase I:

Component	*Quantity in µl*
Supernatant	*175*
NEB Mung Bean Nuclease Buffer	*10*
NEB Mung Bean Nuclease	*4 (1 µl for every 1 µg of DNA)*
10x DNase I Reaction Buffer	*10*
NEB DNase I	*1 (2 units for every 5 µg of DNA)*
Nuclease Free Water (Best Quality Water)	*0*
Total Volume	***200***

Mix thoroughly and incubate at 37°C for 20 minutes. Stop the reaction by adding % 5 µl of 0.2 M EDTA pH 8.0 (to a final concentration of 5mM) and incubate at 75°C for 10 minutes. Proceed with the Qiagen NucleoSpin® Tissue Extraction protocol (see above in section 1.1.1)

5.1.16 Imunnocapture Protocol to collect Secreted Viral DNA

This protocol was developed by Samal et al., for the quantitation of HBV DNA in secreted virions in cell culture methods [207]. Supernatants collected on the third day post transfection were collected as outlined in 1.1.15. The supernatants were subjected to DNAseI and RNaseA for plasmid DNA removal as outlined above in 5.1.16.

Add 100 µl of supernatant to the Monolisa™ HBsAg ULTRA plate, along with 50 µl of the conjugate solution (which is made up of kit solutions R6 and R7 and mixed thoroughly to homogenise the conjugate which is lyophilised).

To allow binding this mixture was incubated at 37°C for 90 mins, and then washed 5 times with PBS.

Treat the bound material with 10 µl Proteinase K and 100 µl of Lysis Buffer from the Qiagen NucleoSpin® Tissue Extraction Kit.

Proceed with the Qiagen NucleoSpin® Tissue Extraction protocol (see above in section 5.1.1)

5.1.17 Real-Time PCR for Determining Viral Load

Real-Time PCR Quantitation was carried out using a SYBR Green method.

Table 5.10: Real-Time Probes and Primers

Primer:	*Sequence:*
HBVs F	*5'- tgcacctgtattcccatc -3'*
HBVs R	*5'- ctgaaagccaaacagtgg -3'*

The standard curve was prepared using the pCH-9/3091 plasmid, which using 3 916 154.16 g/mol as the molecular weight, 6.503 ng equates to 10^9 copies. Thus diluted the plasmid to 6.503 ng/µl and made serial dilutions to get 10^8, 10^7, 10^6, 10^5, 10^4, 10^3, 10^2, 10^1 and 10^0. Then diluted these 1:5 to get these copy numbers per 5 µl [222].

First both the PCR hood and the DNA hoods are treated with bleach, then distilled water and then 70% ethanol, followed by at least 30 minutes of UV light, in order to decontaminate the hoods.

All components for were defrosted on ice, thawed primers, SsoFast™ EvaGreen® Supermix, extracted DNA and plasmid serial dilutions were kept on ice.

The Master Mix was prepared as outlined in Table 5.11

In each case, the controls included

- *A positive control – Eurohep standard, a 10^6 HBV U/ml standard from the National Standards International Bank*
- *A negative control – extraction negative samples*
- *A reagent blank – no DNA template*

Table 5.11: Real-Time PCR Master Mix Contents

Reagent	Initial Concentration	Volume Per reaction (in µl per 20 µl reaction)	Final Concentration
SsoFast™ EvaGreen® Supermix	2x	10	1x
HBVs F and R	10 µM	1	0.5 µM
Best Quality Water		0	
DNA		9	
Total		20	
For Standards			
DNA		5	
Best Quality Water		4	
Total		20	

After adding all the contents of the Master Mix to a 2ml Eppendorf tube, the mixture was thoroughly mixed, and then 11 µl was aliquoted into the 96 well plate. Standards were loaded onto the plate in triplicate, 5 µl into each well. Extracted DNA samples were loaded in duplicate, 9 µl into each well. For all experimental samples, three biological replicates and three technical replicates were loaded onto the plate.

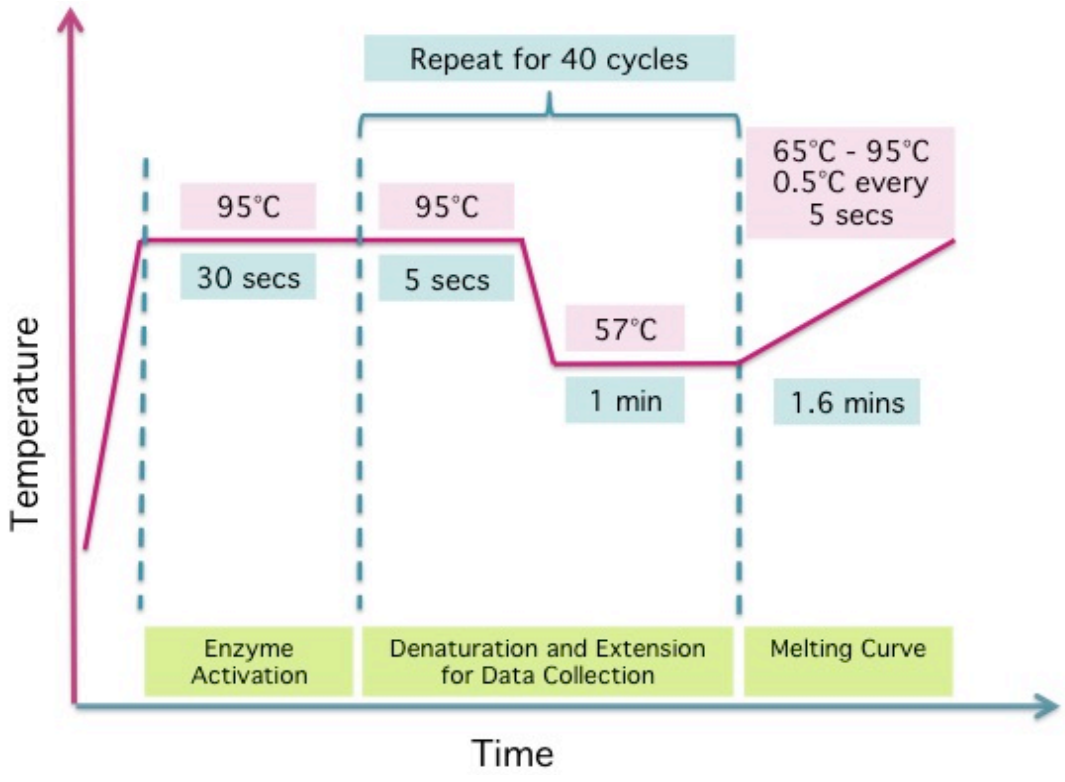


Figure 5.5: Real-Time PCR Cycling Conditions

Note that a Melting Curve was added at the end of the run to confirm that primer binding was specific. Reactions were placed in the CFX96™ Real-Time System with C100™ Thermal Cycler with the aforementioned conditions. Data was analysed on the CFX 96 BioRad Real-Time System, and then exported into an Excel document for further statistical analysis. Sample results are reported in HBV U/ml, only values with an R^2 value of 0.92 and above were considered significant. Also it was ensure that only runs with efficiency between 95 and 105% were considered valid.

Note:

Biological replicates describes three separate samples of the same construct that were treated with the exact same experimental conditions but independently of one another. Whereas, technical replicates are when the exact same sample, collected after all experimental techniques are complete is analysed multiple times.

5.2 Appendix B – Composition of Reagents and Solutions

100x TAE Buffer (for gel electrophoresis)

Dissolved 484 g of Tris buffer in 100 ml in ddH_2O. Added 500 ml 0.5 M Na_2 EDTA (pH 8.0) and 57.1 ml glacial acetic acid. Adjusted volume to 1 L with ddH_2O. Autoclaved at 121°C for 30 minutes, and allowed to cool. Stored at room temperature. Diluted to a 1× working solution for gel electrophoresis by adding 10 ml of 100x TAE to 1 litres of ddH_2O.

PIPES Transformation Buffer:

Add 1,47 g of CaCl2 (final concentration should be 100 nmol/l), 0.3g of PIPES.HCL power, 15 ml of 100% glycerol to a 200 ml Schotts Glass. Add 50 ml of best quality distilled water, and pH to 7.0 using 1 M NaOH solution. Make up to 100 ml with best quality distilled water. Autoclave at 121 °C for 30 minutes, allow to cool. Store at - 20°C in the dark until use.

NaOH Solutions

For a 1 M solution add 4 g NaOH in 100 ml best quality water

For a 10 M solution add 40 g NaOH in 100 ml best quality water

Can be stored at room temperature

Cracking buffer:

Add 1% (w/v) SDS (2.5 g), 2 mM EDTA Disodium Salt (0.19 g), 400 mM Sucrose (34.22 g), 50 mM Tris-HCl, pH 6.8 (125 ml), 100 µg/ml RNase A, DNase- and Protease-Free (2.5 µl), 0.01% (w/v) Bromophenol Blue Sodium Salt for Electrophoresis (0.025 g), 0.01% (w/v) Xylene Cyanol FF (0.025 g).

Use Best quality distilled water, add to make up to a final volume of 250 ml, after which it was stored at room temperature until use.

Ampicillin (100 mg/ml)

Dissolve 1g ampicillin powder in 10ml of best quality distilled water

Filter sterilize into aliquots of 1 ml and store at –20°C

Kanamycin (100 mg/ml)

Dissolve 1g kanamycin powder in 10 ml of best quality distilled water

Filter sterilise into aliquots of 1 ml and store at -20°C

0.2 M EDTA pH 8.0 Solution

Add 7.44 g EDTA-2Na + 80 ml dH_20

Adjusted to pH 8.0 using 1M NaOH (10 ml of NaOH)

Bring volume to 100 ml with dH_2O

*Stir vigorously on a magnetic stirrer**

Sterilize by autoclaving

Store at room temperature

**The disodium salt of EDTA will not dissolve until the pH of the solution is adjusted to 8.0 by the addition of NaOH*

10% (0.5M) EDTA Solution

Add 186.1 g of disodium EDTA (Na2EDTA) to 800 mL of dH_2O

Adjust the pH to 8.0 with 1M NaOH (~50 ml of NaOH)

Bring volume to 1 L with dH_2O

*Stir vigorously on a magnetic stirrer**

Sterilize by autoclaving

Store at room temperature

**The disodium salt of EDTA will not dissolve until the pH of the solution is adjusted to 8.0 by the addition of NaOH*

10% EDTA Sodium Chloride

1 ml of 10% EDTA to 1L of Sodium Chloride Solution

DMEM

6.69 g of DMEM Powder

1.85 g of Sodium Hydrogen Carbonate

500 ml of SABAX or sdH_20

Mix all ingredients thoroughly and then filter sterilize

Store at 4°C until use

Foetal Calf Serum (FCS)

Filter sterilize as needed (8 ml at a time usually works)

Complete Medium (10% Serum, 2x Streptomycin)

For 50 ml of medium:

44 ml of DMEM

5 ml FCS (filter sterilized)

1 ml 100x Streptomycin

Huh7 Freezing Solution 1 (F1) (20% Serum)

8 ml DMEM

2 ml FCS

Huh7 Freezing Solution 2 (F2) (20% Serum, 20% DMSO)

6 ml DMEM

2 ml FCS

2 ml DMSO (sterile)

100x cOmplete PIC solution

Dissolve 1 tablet in 25 ml of sdH_20, vortex till tablet is completely dissolved.

Store at 4°C

1x PBS Buffer

For 1000 ml

1.42 g of 10 mM Na_2HPO_4

0.22 g of 1.8 mM NaH_2PO_4

8.19 g of 140 mM NaCl

Adjust to pH 7.4

Make up to 1000 ml

Ripa Buffer

For 1000 ml:

50 ml of 50 mM Tris-HCl, pH 7.4

8.76 g of 150 mM NaCl

10 ml of 1% Triton X-100 or NP-40

5 g of 0.5% Sodium deoxylcholate

1 g of 0.1% SDS

2 ml of 1 mM EDTA (0.5M stock)

0.42 g of 10mM NaF

Make up to 1000 ml

5.3 Appendix C – Complete Plasmid Maps

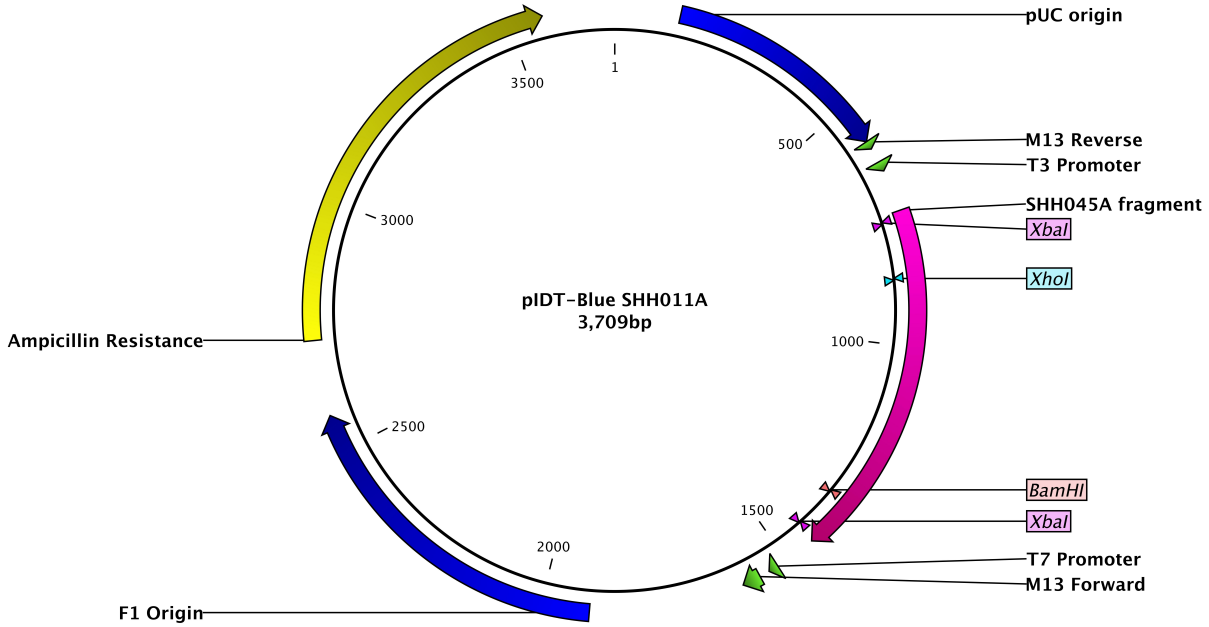


Figure 5.6: The pIDT-Blue Plasmid with Synthesized SHH011A Fragment

A 709 bp fragment was synthesized by IDT and placed within the pIDT-Blue backbone for subcloning purposes. This is after multiple attempts to amplify the S region from this patient serum failed. Note that the fragment is flanked by two XbaI sites for ease of subcloning.

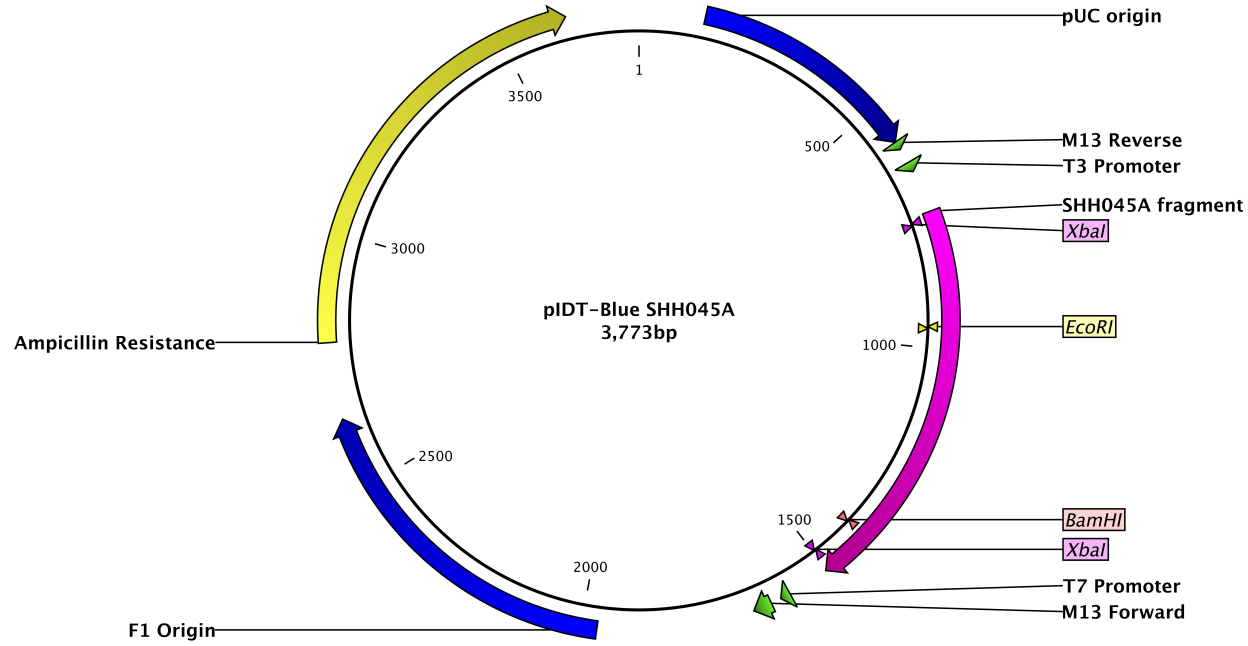


Figure 5.7: The pIDT-Blue Plasmid with Synthesized SHH045A Fragment

A 772 bp fragment was synthesized by IDT and placed within the pIDT-Blue backbone for subcloning purposes. This is after multiple attempts to amplify the S region from this patient serum failed. Note that the fragment is flanked by two XbaI sites for ease of subcloning.

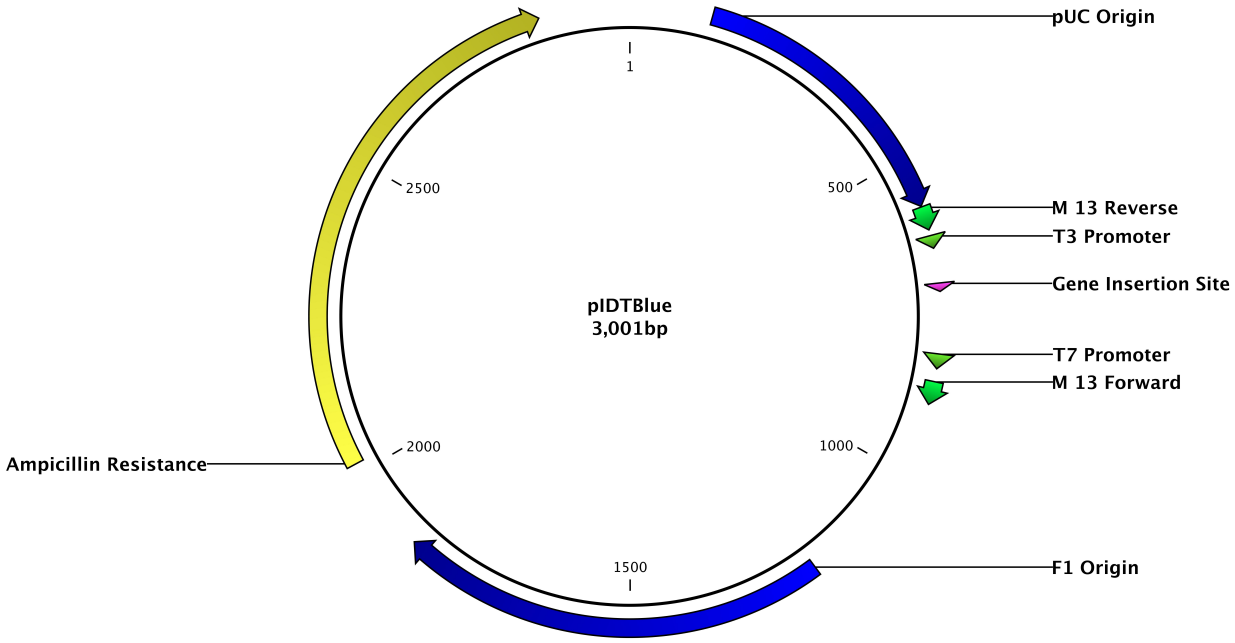


Figure 5.8: The pIDT-Blue Plasmid Without Insert

This is the plasmid that was used as the backbone for the synthesized genes. It is a high copy vector of 2951 bp in size, with an ampicillin resistance gene.

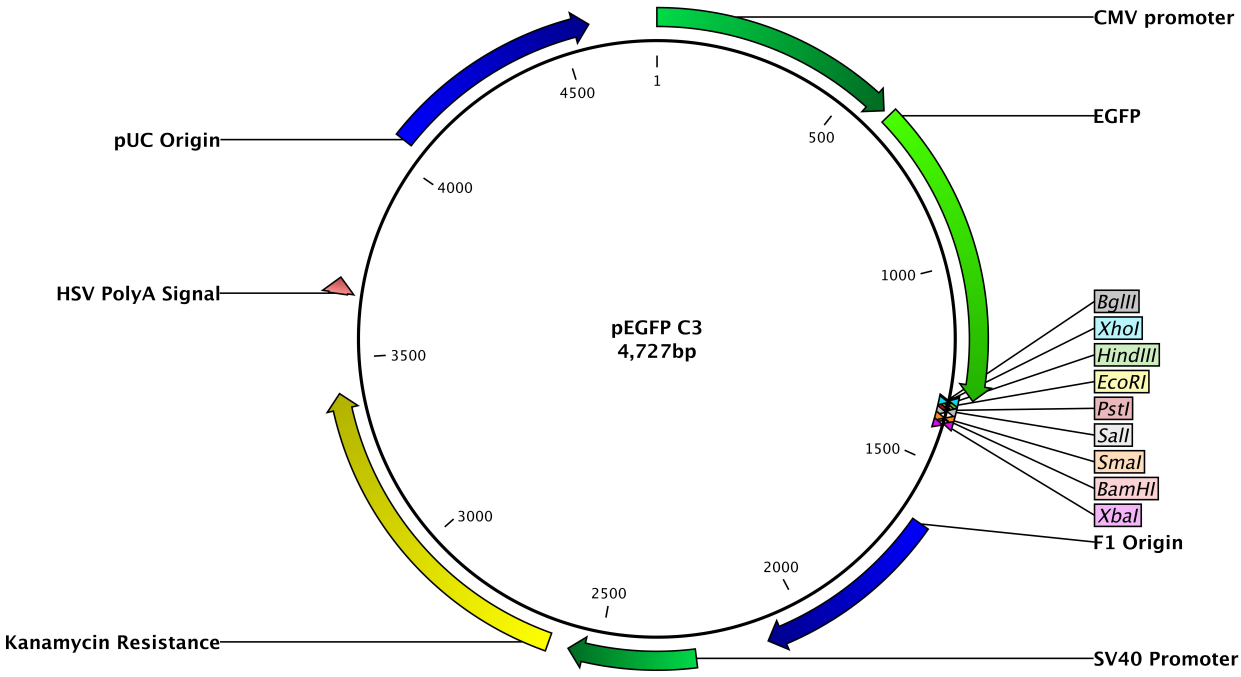


Figure 5.9: The pEGFP-C3 Plasmid with Green Fluorescent Protein Gene

Green Fluorescent Protein is and inherently fluorescent protein that is used here as a reporter in order to calculate transfection efficiency. Huh7 cells successfully transfected with this protein can be visualised directly using a fluorescent microscope. Transfection efficiency is then calculated, depending on the amount of GFP fluorescence.

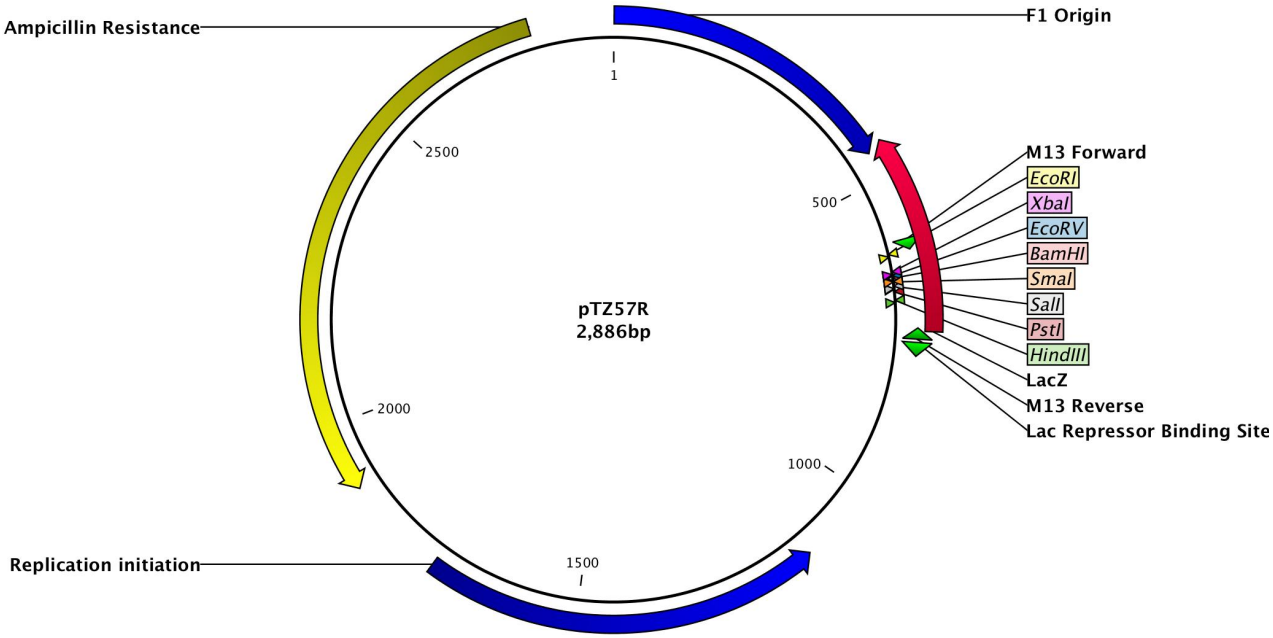


Figure 5.10: The pTZ57RT Cloning Vector

Shown here as a closed circular vector, which is how it was used and a positive transformation control. However, in the InsTAclone kit, it is also available as a linearised vector (pTZ57R/T) which is used for direct insertion of a PCR amplicon with TA ends.

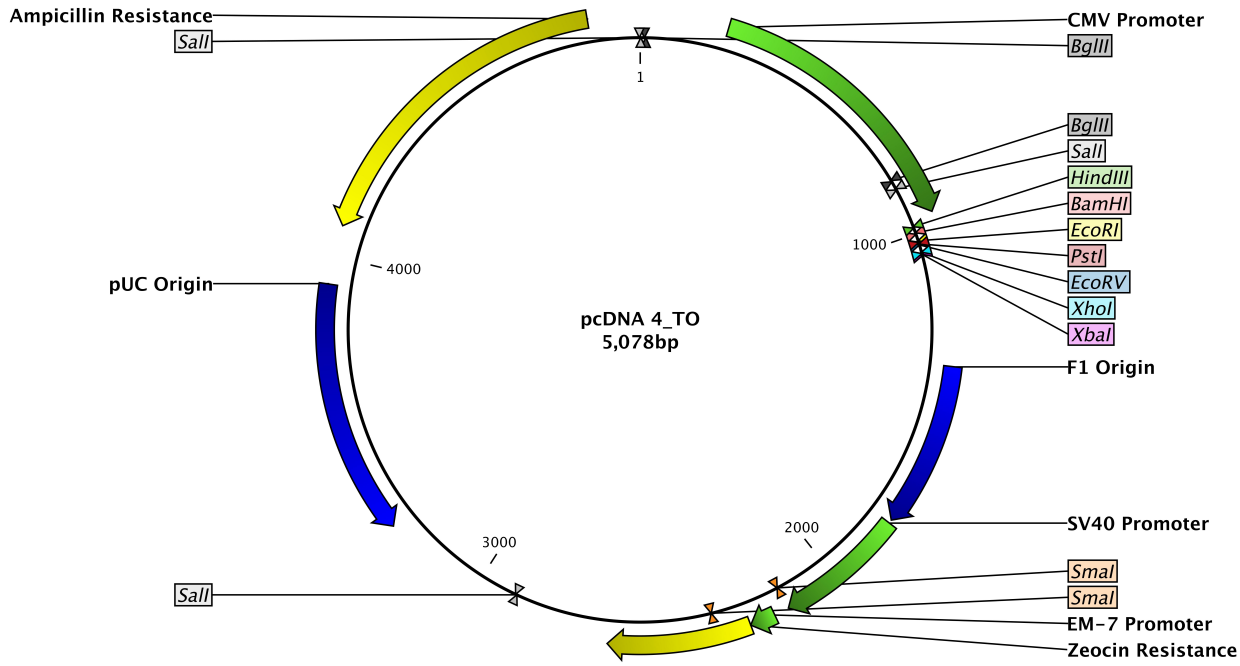


Figure 5.11: The pcDNA™ 4/TO Cloning Vector

This plasmid was used as the backbone for the overlength clone constructs.

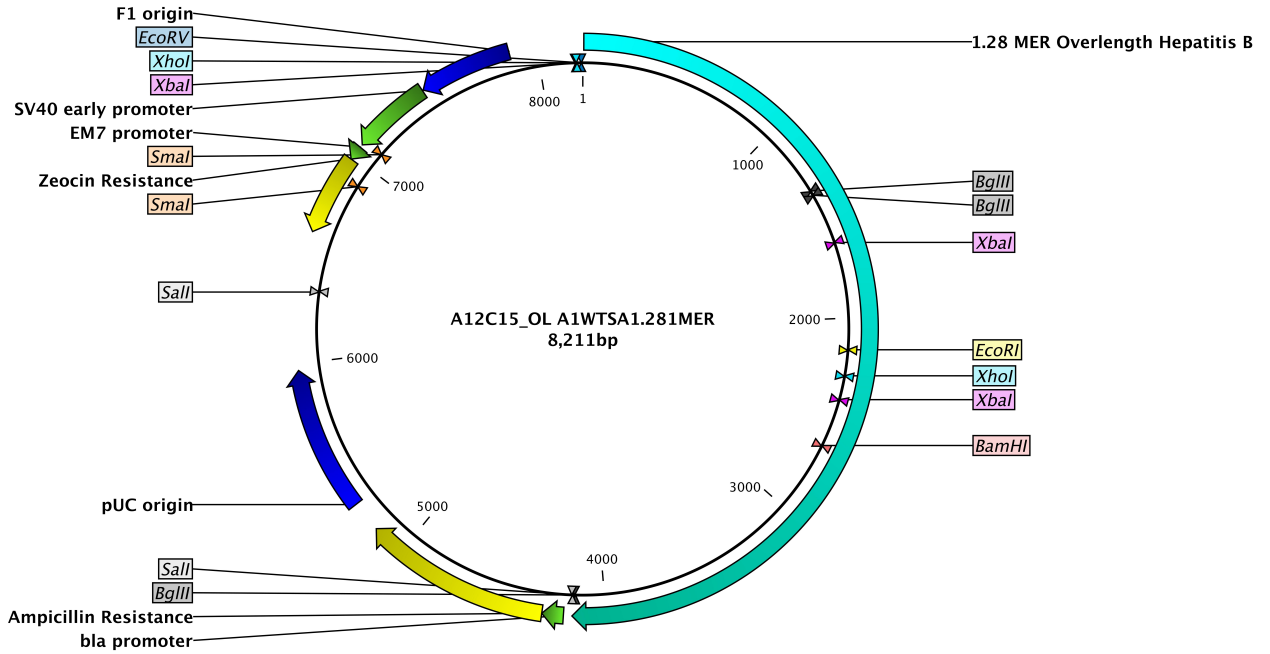


Figure 5.12: The A12C15_OL A1WTSA 1.28 MER Overlength Clone

This clone has a 1.28 MER overlength Subgenotype A1 wild type HBV within the pcDNA™ 4/TO backbone. The 1.28 MER HBV segment contains a full Polymerase, full S gene, as well as two times PreCore, Core and X Open Reading Frames (ORFs). Please note there are three XbaI sites.

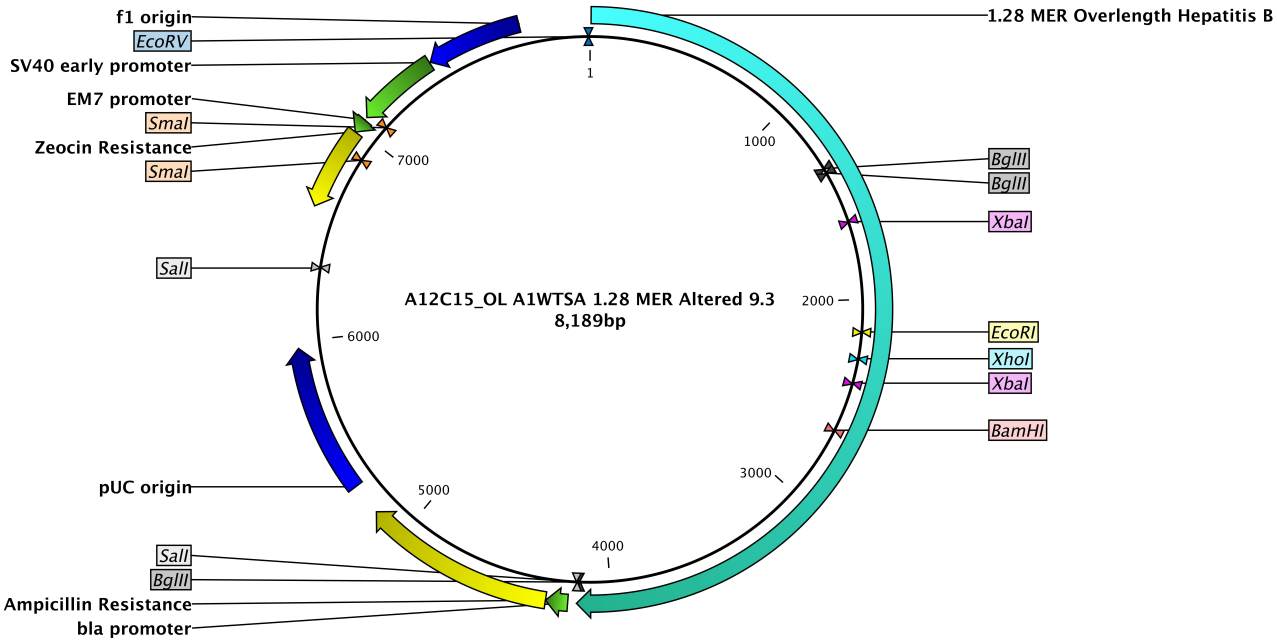


Figure 5.13: The A12C15_OL A1WTSA 1.28 MER Altered 9.3 Clone

This clone still contains an unaltered 1.28 MER overlength Subgenotype A1 wild type HBV within the pcDNA™ 4/TO backbone. However the backbone was restricted with ApaI and NotI, and the resulting ends were blunted. After re-ligating this construct the undesired XbaI site in the vector was removed.

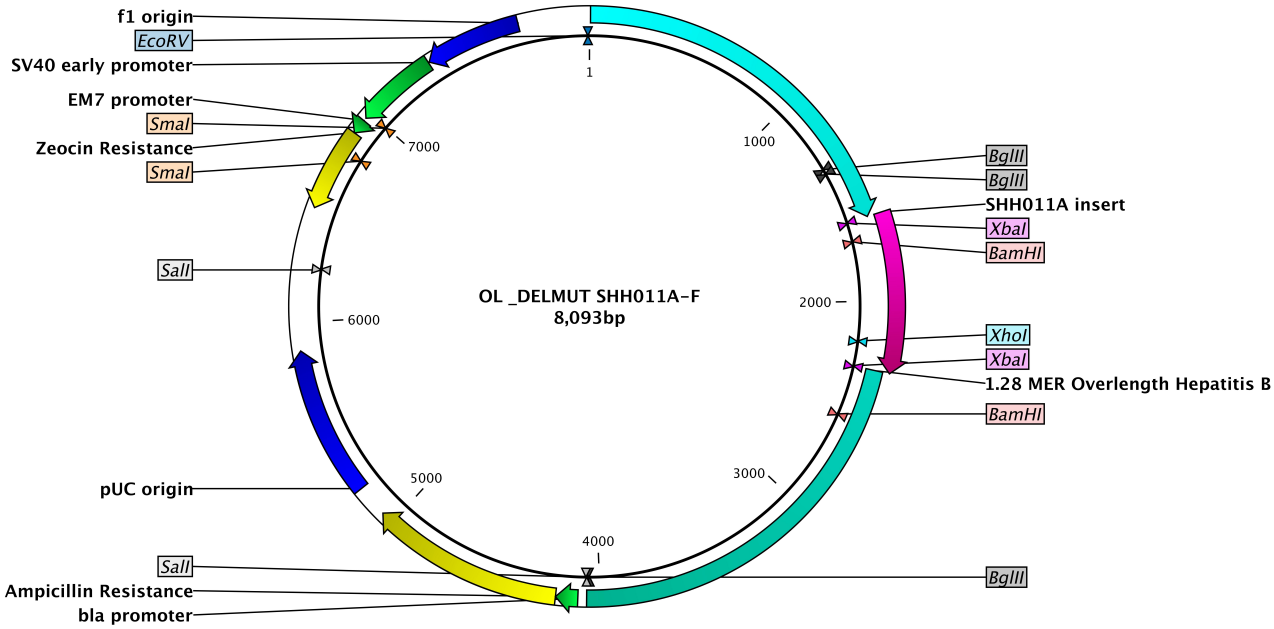


Figure 5.14: OL_DELMUT SHH011A-F Clone

In this construct a portion of the SHH011A S gene (pink arrow) was inserted using the two remaining XbaI sites. Here SHH011A is in the correct orientation.

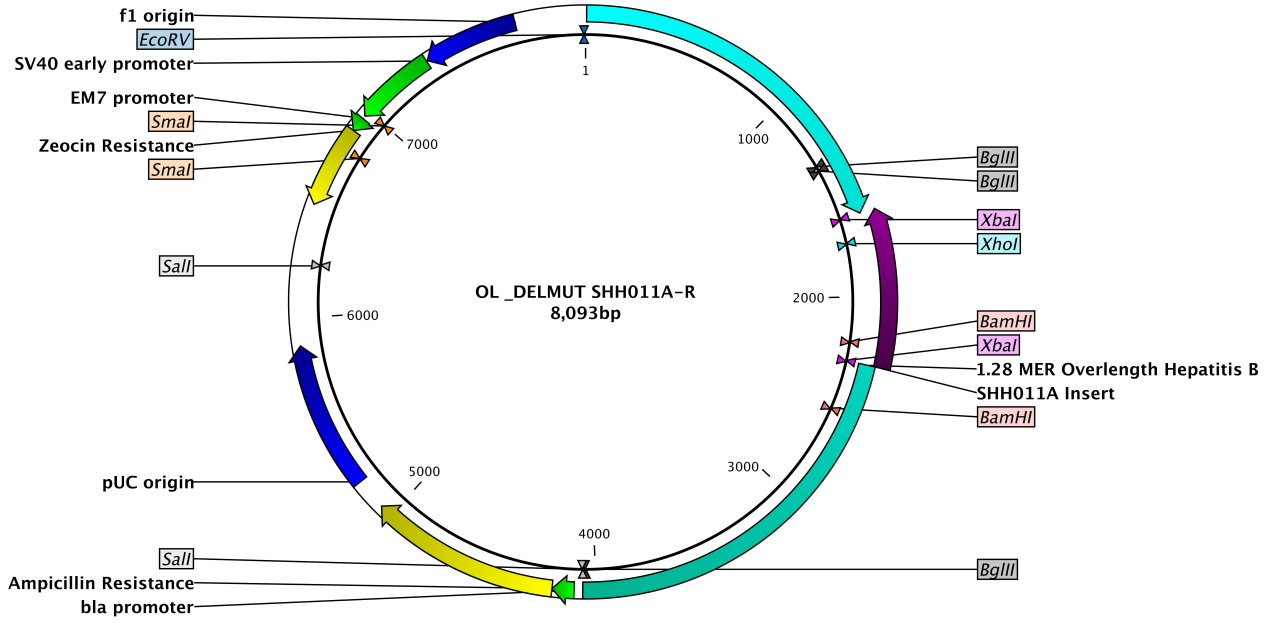


Figure 5.15: OL_DELMUT SHH011A-R Clone

In this construct a portion of the SHH011A S gene (purple arrow) was inserted using the two remaining XbaI sites. Here SHH011A is in the incorrect orientation.

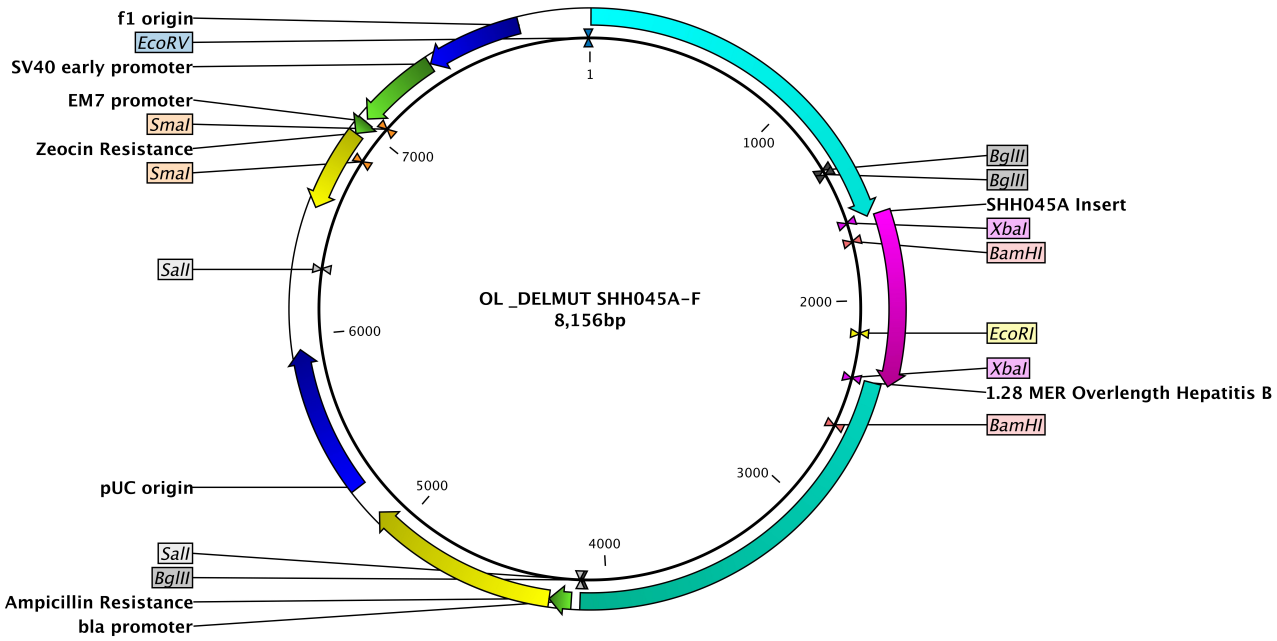


Figure 5.16: OL_DELMUT SHH045A-F Clone

In this construct a portion of the SHH045A S gene (purple arrow) was inserted using the two remaining XbaI sites. Here SHH045A is in the correct orientation.

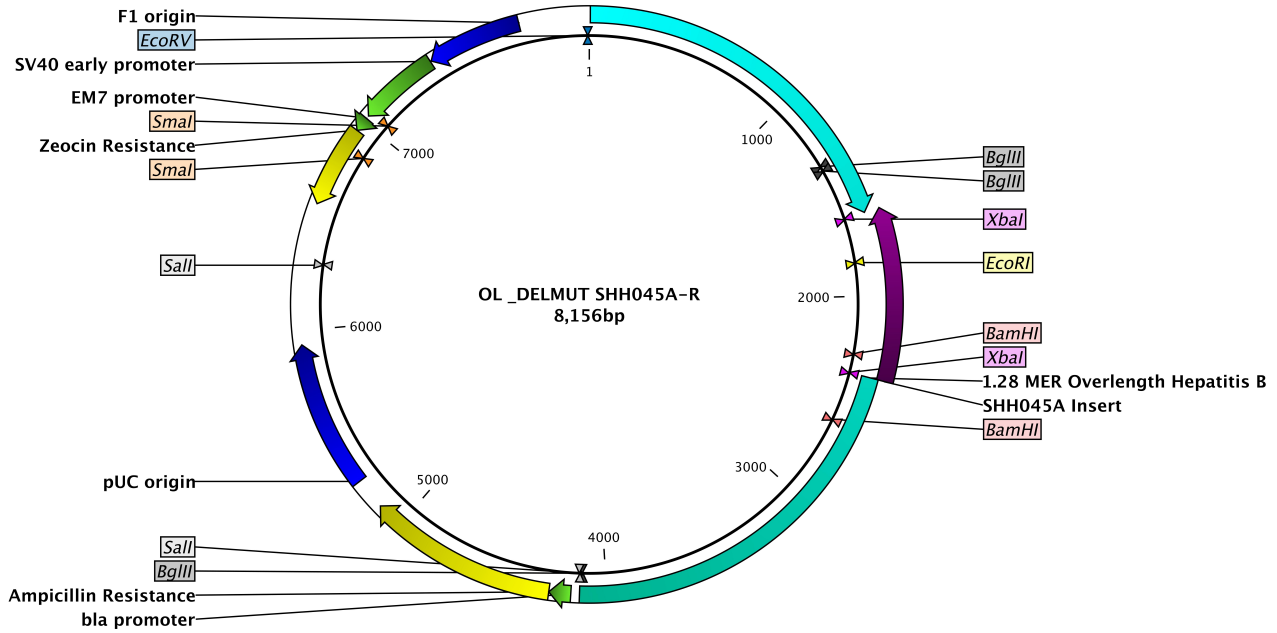


Figure 5.17: OL_DELMUT SHH045A-R Clone

In this construct a portion of the SHH045A S gene (purple arrow) was inserted using the two remaining XbaI sites. Here SHH045A is in the incorrect orientation.

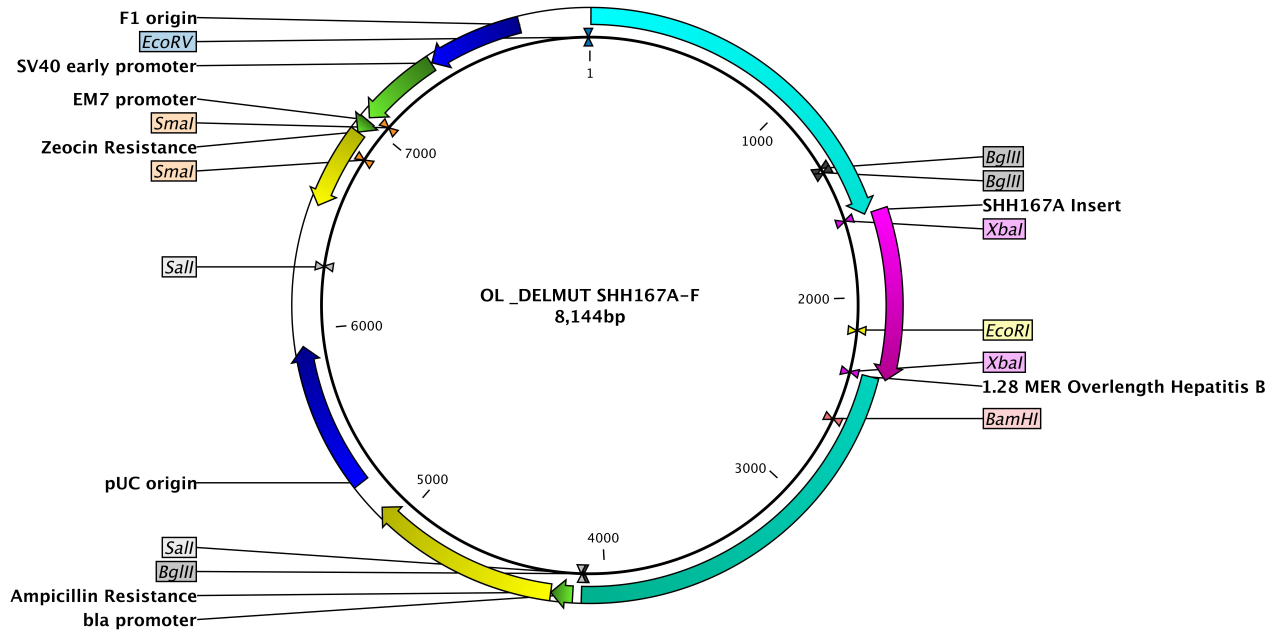


Figure 5.18: OL_DELMUT SHH167A-F Clone

In this construct a portion of the SHH167A S gene (pink arrow) was inserted using the two remaining XbaI sites. Here SHH167A is in the correct orientation.

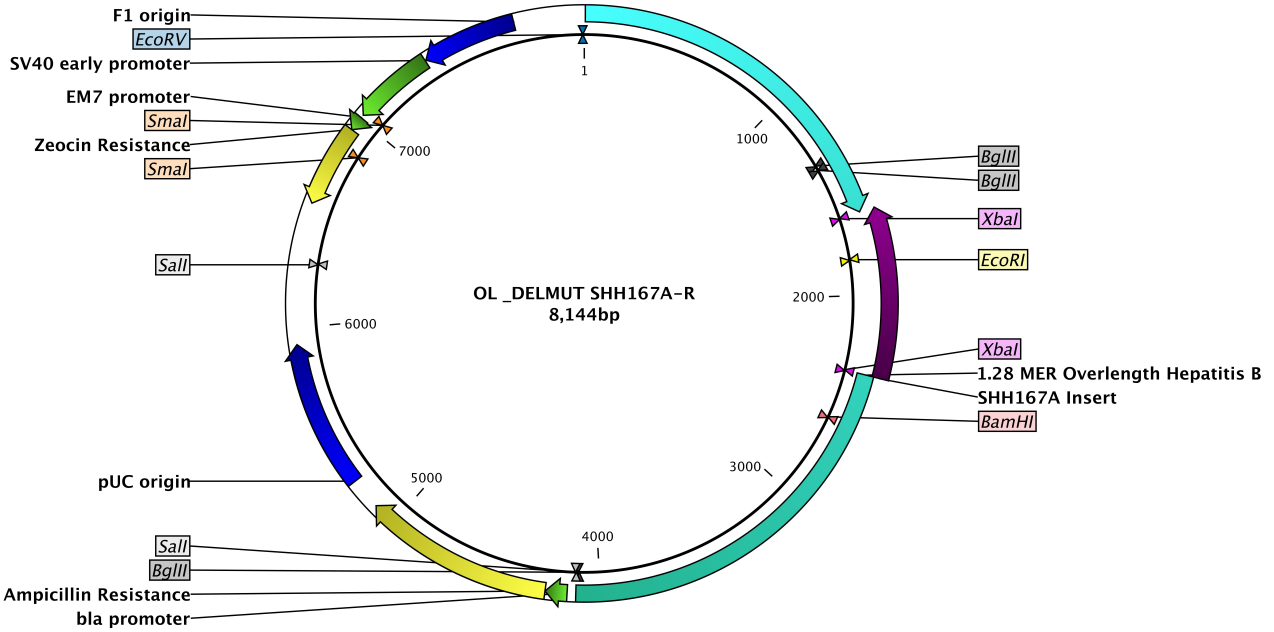


Figure 5.19: OL_DELMUT SHH167A-R Clone

In this construct a portion of the SHH167A S gene (purple arrow) was inserted using the two remaining XbaI sites. Here SHH167A is in the incorrect orientation.

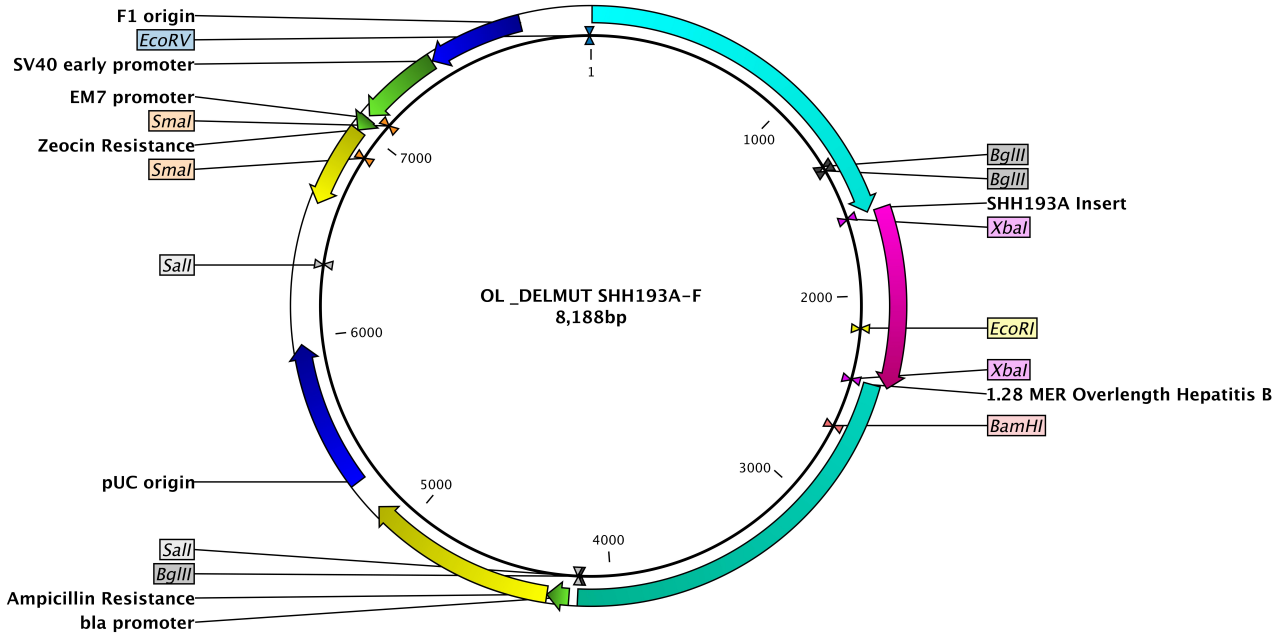


Figure 5.20: OL_DELMUT SHH193A-F Clone

In this construct a portion of the SHH193A S gene (pink arrow) was inserted using the two remaining XbaI sites. Here SHH193A is in the correct orientation. SHH193A contained no deletions and was thus used as a wild type control.

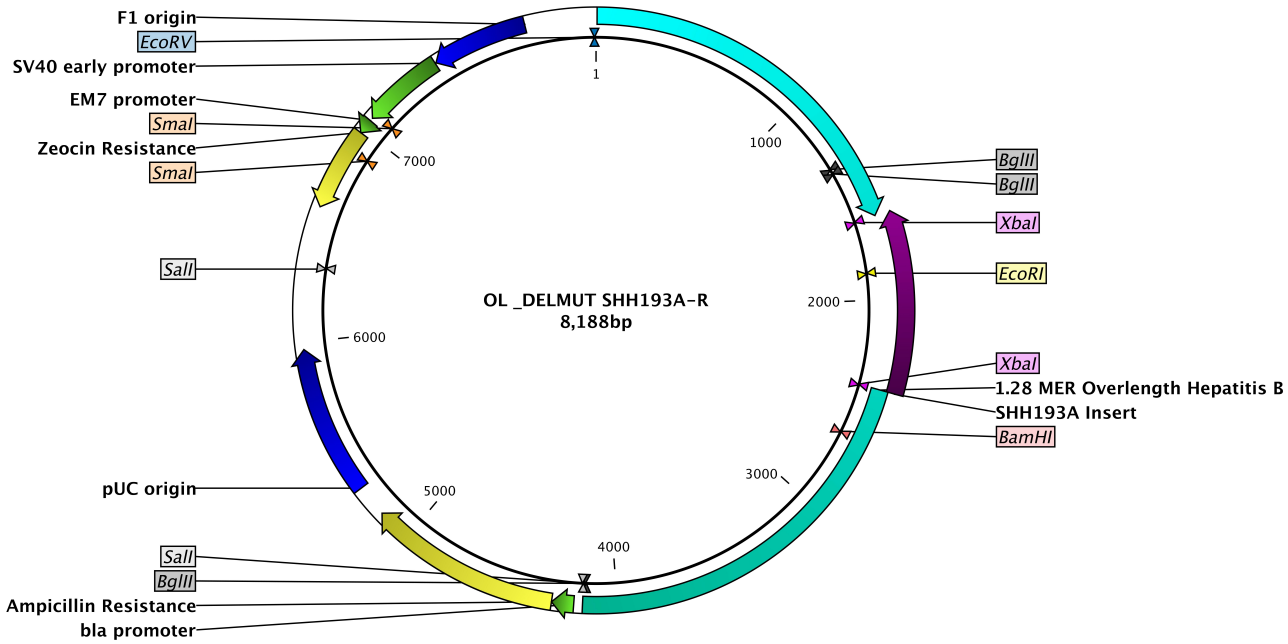


Figure 5.21: OL_DELMUT SHH193A-R Clone

In this construct a portion of the SHH193A S gene (purple arrow) was inserted using the two remaining XbaI sites. Here SHH193A is in the incorrect orientation. SHH193A contained no deletions and was thus used as a wild type control.

5.4 Appendix D – Ethics Clearance Certificate

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Mrs Suzaane Nicholson

CLEARANCE CERTIFICATE	**M1204106**
PROJECT	Functional Charaterisation of preS1/preS2 Deletion Mutants of Hepatitis B Virus Isolated from Southern Africans
INVESTIGATORS	Mrs Suzaane Nicholson.
DEPARTMENT	Internal Medicine/Heptatitis Virus Diversity Res.
DATE CONSIDERED	Ad hoc
+DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 11/05/2012 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor : Prof A Kramvis

DECLARATION OF INVESTIGATOR(S)
To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

5.5 Appendix E – Molecular Weight Markers

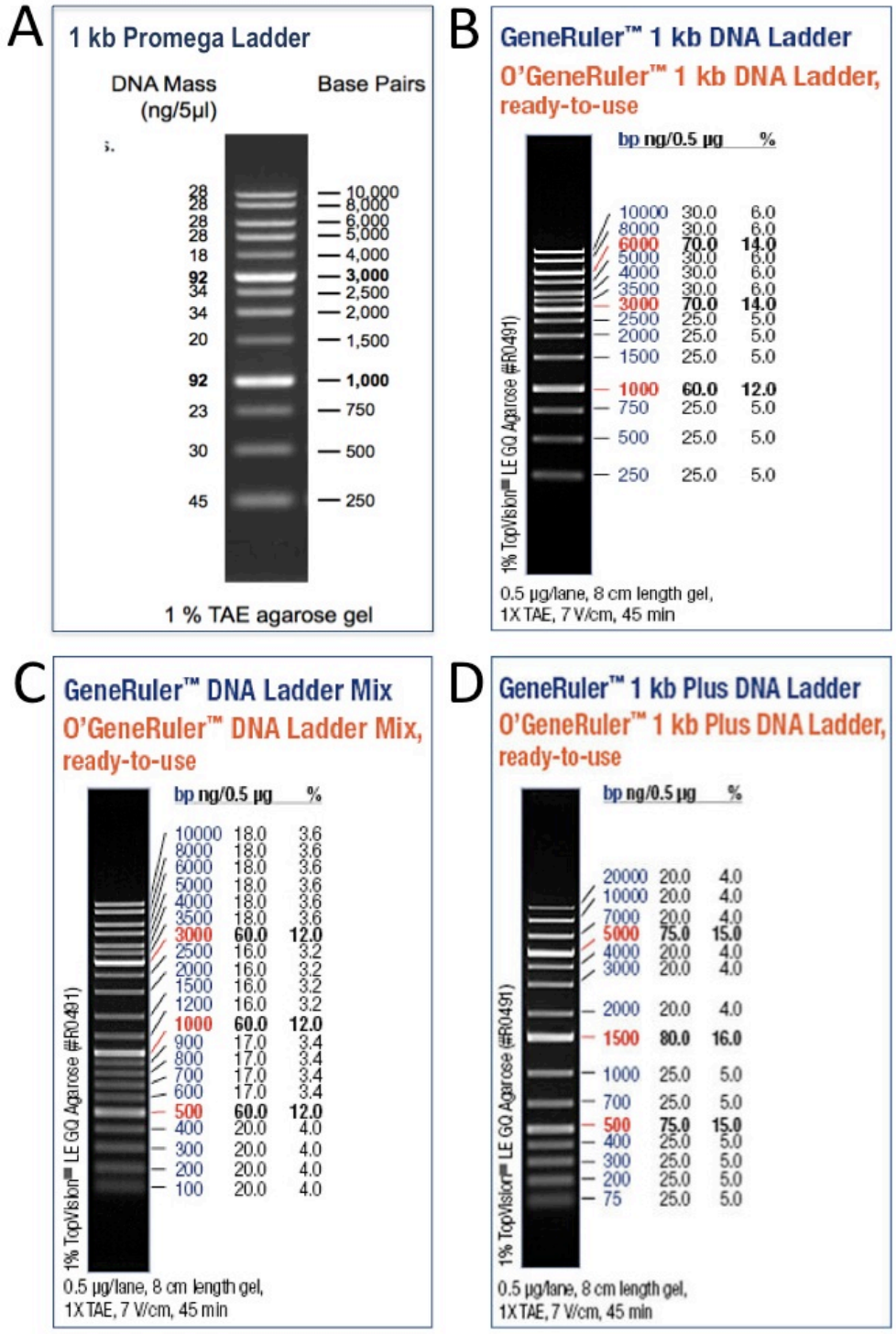


Figure 5.22: DNA Molecular Weight Markers Used for this Study

The 1kb Promega Ladder (A) was used for to determine the size of the products of the initial HotStart and some high fidelity PCR reactions. The GeneRuler™ 1 kb DNA Ladder (B) was used for most high fidelity PCR reactions, all cracking, digestion reactions and ligations, as well as for most restriction mapping. The GeneRuler™ DNA Ladder Mix (C) was used to determine concentration after Gel purification and after restriction mapping (prior to ligations) in addition to NanoDrop spectrophotometry. The 1kb Plus DNA GeneRuler™ (A) was used for some of the restriction maps.

5.6 Appendix F – Protocols for Phylogenetic Tree Production

Standard Operating Procedures for Phylogenetic Analysis using Neighbor-Joining Methods with Bootstrapping

1.Multiple Sequence Alignment (MSA)

Prepare a multiple sequence alignment of all sequences for analysis using a program of your choice (ClustalW, Muscle, MEGA, BioEdit, CLC Main Workbench). Check the alignment by visualising it in a multiple sequence alignment viewer (AliView, SeaView, GeneDoc, BioEdit, MEGA, CLC Main Workbench), and curate sequences as necessary (verify INDELS and wobble bases).

2.Create a Phylip file

Use either GeneDoc, Mega or CLC Main Workbench to export your MSA as a phylip file with the extension .phy

3.Run the Tree through Pipeline: TreeMail

Submit the phylip file you just created to Pipeline: TreeMail online at http://hvdr.bioinf.wits.ac.za/treemail/). Select Neighbor-Joining 1000 bootstrapping replicates and press Run TreeMail. The result is emailed to you.

4.Creating a Publication Ready Tree

Open the final tree (output from Pipeline: TreeMail, a .tre file) in FigTree (or TreeView). For correct display of your tree select Node and under Node Labels, set display to posterior, and make sure the scale bar reads 0.02 (2% divergence), Your Order must be decreasing.

Standard Operating Procedures for Bayesian Analysis using BEAUti and BEAST – Procedure for producing a Maximum Clade Credibility Tree:

1.Multiple Sequence Alignment (MSA)

Prepare a multiple sequence alignment of all sequences for analysis using a program of your choice (ClustalW, Muscle, MEGA, BioEdit, CLC Main Workbench). Check the alignment by visualising it in a multiple sequence alignment viewer (AliView, SeaView, GeneDoc, BioEdit, MEGA, CLC Main Workbench), and curate sequences as necessary (verify INDELS and wobble bases).

2.Quick Check using Maximum Likelihood Tree

Run the MSA (FASTA, Phylip or Nexus/PAUP) through Garli to check the tree. Garli produces a Maximum Likelihood (ML) tree. Open this tree in FigTree or TreeView. Garli does not accept Nexus files from CLC Main Workbench. All lines in a Nexus file, which contain square brackets, must be deleted. This should be done in a standard text editor,

such as Notepad (Windows) or TextEdit (Mac). Check that the ML tree looks "reasonable" and that all sequences, which should appear on this tree, are included.

Note: Garli does not accept long sequence names, thus edit sequences names to be Genotype, Accession number and country only eg. A1_AY233090_SA

Note: On Mac to Run Garli...

Set up your Garli bin folder which should contain your FASTA/Nexus file of your MSA, make sure the file name and output name is correct in your config file

Open Terminal and direct to the correct cd to run your Garli

3.Create a Nexus File

Open the MSA FASTA file in MEGA, export it as a Nexus/PAUP file (.nxs or .nex or .nexus). CLC Main Workbench does not export a valid Nexus file, so use MEGA instead. Open the Nexus file in Notepad (Windows) or TextEdit (Mac) and delete all lines containing square brackets.

4.Model Test

At this stage you can run a Model Test using either a FASTA or Nexus file of your alignment, a Model Test allows one to carry out statistical selection of best-fit models of nucleotide substitution JModelTest or PAUP are two of the software packages that can be used for this purpose. Import your file, select likelihood calculations, and then run/execute the model test. This will tell you which statistical model is best applied to your data set. However it has been shown in the past that the generalised time-reversible (GTR) model is best suited to HBV.*

5.Create a BEAUti file

This is the most important step, as this dictates the Bayesian analysis parameters, such as the evolutionary model, the log normal and exponent settings. Open BEAUti, go to File ➔ Import Data ➔ Select your Nexus file.

Choose appropriate settings:

Under Sites, select the GTR evolutionary model, for Base Frequencies select Estimated, and under Site Heterogeneity Model select Gamma + Invariant Sites

(Thus the model is GTR + I + G)

Under MCMC select length of chains (10 000 000 is usually acceptable)

You will run BEAUti twice, 1st to establish your log normal .xml file (Select Clocks and under Model select log normal relaxed clock) and the 2nd time to establish your exponential .xml file (Select Clocks, and under model select exponential relaxed clock)

At this stage you can save this BEAUti file, and it will save your settings, so you can save one for log normal and one for exponential if you wish. It may also be useful to save the log normal files into one folder and exponential files into another folder.

6.Run BEAST

This programme will create many possible trees. Later, Tree Annotator is used to select the best tree (most supported tree). You will run BEAST twice, once with your log normal file and then again with your exponential file.

Open BEAST ➔ Select XML file ➔ Run. The input for BEAST is the XML file created by BEAUti. BEAST takes around 5.5 hours to process 10 million chains for 225 samples.

Copy or rename the "log" and "tree" files from the first BEAST run, and then run BEAST again with the second XML file. (Or put your log and exponential files in separate folders)

7.Use Tracer to Identify the best Model

Import the .log files of both the log normal and the exponential runs from BEAST. Select the posterior statistic.

Under the Marginal Probability Distribution tab, make sure your distribution is a normal distribution (i.e. not skewed). Under the trace tab, make sure your trace/hairy caterpillar is relatively level (not uneven in height or density). The burn in should appear as grey.

Under the Estimates tab, check the effective sample size (ESS), an acceptable ESS is > 100 however an ESS ≥ 200 is desirable.

Compare the two files' AICM values by selecting both, click on analysis, and select Model Comparison [103], this may take some time

Ultimately you should choose the file with the option (log normal or exponential) with the highest ESS and the lowest AICM to be the file you work on for your final tree. You should also identify you HPD percentage and include this in the legend under your tree.

8.Find the 'Best' Tree or Single Target Tree

Open Tree Annotator ➔ Select Burn-in as 10 (%) i.e. 1000 trees and set posterior probability as 0.5%. Choose the input file (which will be the BEST result, either the log normal or exponential .trees file, depending on which one scored best in step 7). Create a name for your output file, and select a destination for the result. Press Run

9.Creating an Publication Ready Tree

Open the final tree (out put from Tree Annotator) in FigTree (or TreeView).

For correct display of your tree select Node and under Node Labels, set display to posterior, and make sure the scale bar reads 0.02 (2% divergence), Your Order must be decreasing.

5.7 Appendix G – Complete List of Sequences Used for Phylogenetic Trees

Other than the sequences or the quasispecies and their parental strains, the PreS1/PreS2/Surface section of sequences from all 9 genotypes were used as reference sequences.

Table 5.12: Reference and Genotype Sequences Used for Phylogenetic Trees

	GenBank Accession Number	Country of Origin
Subgenotype A1	AY233288	South Africa
	AY233284	South Africa
	AY233281	South Africa
	AY233277	South Africa
	AY233279	South Africa
	AY233276	South Africa
	AY233287	South Africa
	AY233278	South Africa
	AY233289	South Africa
	AY233283	South Africa
	AY233282	South Africa
	AY233285	South Africa
	AY233288	South Africa
	AY233290	South Africa
	AY233275	South Africa
	AY233274	South Africa
	AB246336	South Africa
	AY903452	South Africa
	AB116085	Bangladesh
	JQ023661	Colombia
	KJ854687	Brazil
	AB453988	Japan
	AF043560	Argentina
	AY934771	Somalia
	AB116094	Phillipines
	AB076678	Malawi
	FJ692592	Haiti
	FM199974	Rwanda
Subgenotype A2	AY233286	South Africa
	Z35717	Poland
	Z72478	Germany
	AF537371	USA
	AJ344115	France
Subgenotype A3	FJ692554	Nigeria
	GQ161813	Guinea
	AM180624	Cameroon
Genotype B	AB073831	Thailand
	AF282918	China
	AB073836	Taiwan
	AB115551	Cambodia
	AB073835	Vietnam
	AB033555	Sumatra
	M54923	Indonesia
	AB241117	Phillippines
	AB010291	Japan
	DQ463797	Canada
	DQ463801	Canada
	DQ463799	Canada

	GenBank Accession Number	Country of Origin
Genotype C	JN315779	Korean Mummy
	AB014376	Japan
	AY066028	China
	AF458664	China
	AY057947	Tibet
	AB117758	Cambodia
	AB112066	Myanmar
	AY167099	Taiwan
	AB112065	Vietnam
	AB112472	Thailand
	AP011107	Indonesia
	HM011493	Indonesia
	AP011108	Indonesia
	X75656	Polynesia
	X75665	New Caledonia
	AB493842	Indonesia
	AB554020	Indonesia
	AB554018	Indonesia
	AB644287	Indonesia
	AB540583	Indonesia
	AB241109	Philippines
	AB048704	Australia
Genotype D	FJ386590	China
	FJ904415	Tunisia
	JN642129	Lebanon
	GU456684	Iran
	JF754629	Turkey
	AY233296	South Africa
	X85254	Italy
	GQ922000	Canada
	AB048701	Australia
	FJ692532	Haiti
	FJ904438	Tunisia
	FJ904394	Tunisia
	AB03358	Japan
	GQ205378	India
Genotype E	X75664	Senegal
	AB091255	Cote d'Ivoire
Genotype F	DQ823092	Argentina
	X69798	Brazil
	AY090455	Nicaragua
	AF223965	Argentina
	X75658	France
	AY090461	El Salvador
	AY090458	Costa Rica
Genotype I	EU833891	Canada
	FJ023666	Laos
	AF241407	Vietnam
	EU835241	India

References:

1. WHO, *World Health Statistics 2015*, in *World Health Statistics*, WHO, Editor. 2015: Luxembourg.
2. Lai, C.L. and S. Locarnini, *Human Virus Guides: Hepatitis B Virus*. Vol. 1. 2002: International Medical Press.
3. Lusebrink, J., Schildgen, V., Schildgen, O., *The human hepatitis B virus-Classification, biology, life cycle, in vitro and in vivo models.* Hepatology-A Clinical Text Book. 2nd ed. The Johns Hopkins University, 2009: p. 37-57.
4. Flint, S., et al., *Principles of Virology: Vol. 1. Molecular Biology*. 2009: ASM Press, Washington, DC.
5. Nassal, M., et al., *Phenotyping hepatitis B virus variants: from transfection towards a small animal in vivo infection model.* Journal of clinical virology, 2005. **34**: p. S89-S95.
6. Schaefer, S., *Hepatitis B virus taxonomy and hepatitis B virus genotypes.* World Journal of Gastroenterology, 2007. **13**(1): p. 14.
7. Summers, J. and W.S. Mason, *Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate.* Cell, 1982. **29**(2): p. 403-415.
8. Summers, J., A. O'Connell, and I. Millman, *Genome of hepatitis B virus: restriction enzyme cleavage and structure of DNA extracted from Dane particles.* Proceedings of the National Academy of Sciences, 1975. **72**(11): p. 4597-4601.
9. Locarnini, S. *Molecular virology of hepatitis B virus*. in *Seminars in liver disease*. 2004.
10. Littlejohn, M., S. Locarnini, and L. Yuen, *Origins and Evolution of Hepatitis B Virus and Hepatitis D Virus.* Cold Spring Harbor perspectives in medicine, 2016. **6**(1): p. a021360.
11. Gao, S., Z.-P. Duan, and C.S. Coffin, *Clinical relevance of hepatitis B virus variants.* World J Hepatol, 2015. **7**(8): p. 1086-1096.
12. King, A.M., M.J. Adams, and E.J. Lefkowitz, *Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses*. Vol. 9. 2011: Elsevier.
13. Glebe, D. and C.M. Bremer, *The Molecular Virology of Hepatitis B Virus.* Semin Liver Dis, 2013. **33**(02): p. 103-112.
14. Tang, H. and A. McLachlan, *Avian and Mammalian hepadnaviruses have distinct transcription factor requirements for viral replication.* Journal of virology, 2002. **76**(15): p. 7468-7472.
15. Drexler, J.F., et al., *Bats carry pathogenic hepadnaviruses antigenically related to hepatitis B virus and capable of infecting human hepatocytes.* Proceedings of the National Academy of Sciences, 2013. **110**(40): p. 16151-16156.
16. Walter, E., et al., *Hepatitis B virus infection of tupaia hepatocytes in vitro and in vivo.* Hepatology, 1996. **24**(1): p. 1-5.
17. Fauquet, C.M., et al., *Virus taxonomy: VIIIth report of the International Committee on Taxonomy of Viruses*. 2005: Academic Press.
18. Seeger, C. and W.S. Mason, *Molecular biology of hepatitis B virus infection.* Virology, 2015. **479–480**(0): p. 672-686.
19. Dickens, C., et al., *Occult hepatitis B virus infection in chacma baboons, South Africa.* Emerging infectious diseases, 2013. **19**(4): p. 598.
20. Hahn, C.M., et al., *Characterization of a Novel Hepadnavirus in the White Sucker (Catostomus commersonii) from the Great Lakes Region of the United States.* Journal of virology, 2015. **89**(23): p. 11801-11811.
21. Gilbert, C. and C. Feschotte, *Genomic fossils calibrate the long-term evolution of hepadnaviruses.* PLoS biology, 2010. **8**(9): p. 2061.

22. WHO, A., *Report from the WHO Regional Comittee for Africa; by theSecretariat of WHO Region Comittee for Africa entitled Viral Hepatitis: Situation Analysis and Perspectives in the African Region*, in *VIRAL HEPATITIS: SITUATION ANALYSIS AND PERSPECTIVES IN THE AFRICAN REGION*. 2014, World Health Organisation: Cotonou, Republic of Benin. p. 6.
23. Stannard, P.L. *Lecture notes: Hepatitis*. Hepatitis [Lecture notes] 2011 2012 [cited 2015 18 September 2015]; Lecture notes on viral hepatitis]. Available from: http://www.virology.uct.ac.za/teachhepatitis.html.
24. Pollicino, T., et al., *Hepatitis B virus PreS/S gene variants: pathobiology and clinical implications.* Journal of hepatology, 2014. **61**(2): p. 408-417.
25. P. Le Mercier, et al. *Human Virus Relative Size*. ViralZone [Web-resource] 2015 April 2015 [cited 2015 27 July 2015]; 2011 Jan; 39(Database issue):D576-82. :[ViralZone is a SIB Swiss Institute of Bioinformatics web-resource for all viral genus and families, providing general molecular and epidemiological information, along with virion and genome figures. Each virus or family page gives an easy access to UniProtKB/Swiss-Prot viral protein entries.]. Available from: http://viralzone.expasy.org/all_by_species/5216.html.
26. Tong, S., et al., *Hepatitis B virus genetic variants: biological properties and clinical implications.* Emerging Microbes & Infections, 2013. **2**(3): p. e10.
27. Van Damme, P., et al., *Strategies for global prevention of hepatitis B virus infection*, in *Hot Topics in Infection and Immunity in Children VI*. 2010, Springer. p. 175-188.
28. Bock, C., et al., *The Pre‐S region determines the intracellular localization and appearance of hepatitis B virus.* Hepatology, 1999. **30**(2): p. 517-525.
29. Dane, D., C. Cameron, and M. Briggs, *Virus-like particles in serum of patients with Australia-antigen-associated hepatitis.* The lancet, 1970. **295**(7649): p. 695-698.
30. Lutwick, L.I. and W.S. Robinson, *DNA synthesized in the hepatitis B Dane particle DNA polymerase reaction.* Journal of virology, 1977. **21**(1): p. 96-104.
31. Chotiyaputta, W. and A.S. Lok, *Hepatitis B virus variants.* Nature Reviews Gastroenterology and Hepatology, 2009. **6**(8): p. 453-462.
32. Gerlich, W.H., *Medical virology of hepatitis B: how it began and where we are now.* Virol J, 2013. **10**(1): p. 239-263.
33. Norder, H., et al., *Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes.* Intervirology, 2004. **47**(6): p. 289-309.
34. Kramvis, A., M. Kew, and G. Francois, *Hepatitis B virus genotypes.* Vaccine, 2005. **23**(19): p. 2409-2423.
35. Cao, G.-W., *Clinical relevance and public health significance of hepatitis B virus genomic variations.* World journal of gastroenterology: WJG, 2009. **15**(46): p. 5761.
36. Locarnini, S. and S. Bowden, *Hepatitis B surface antigen quantification: not what it seems on the surface.* Hepatology, 2012. **56**(2): p. 411-414.
37. Pollicino, T., et al., *Impact of hepatitis B virus (HBV) preS/S genomic variability on HBV surface antigen and HBV DNA serum levels.* Hepatology, 2012. **56**(2): p. 434-443.
38. Ni, Y., et al., *The pre-s2 domain of the hepatitis B virus is dispensable for infectivity but serves a spacer function for L-protein-connected virus assembly.* Journal of virology, 2010. **84**(8): p. 3879-3888.
39. Pollicino, T., et al., *Molecular and functional analysis of occult hepatitis B virus isolates from patients with hepatocellular carcinoma.* Hepatology, 2007. **45**(2): p. 277-285.
40. Chen, C.H., et al., *Pre-S Deletion and Complex Mutations of Hepatitis B Virus Related to Advanced Liver Disease in HBeAg-Negative Patients.* Gastroenterology, 2007. **133**(5): p. 1466-1474.

41. Lu, C.-C., et al., *Key role of a CCAAT element in regulating hepatitis B virus surface protein expression.* Virology, 1995. **206**(2): p. 1155-1158.
42. Xu, Z., G. Jensen, and T. Yen, *Activation of hepatitis B virus S promoter by the viral large surface protein via induction of stress in the endoplasmic reticulum.* Journal of virology, 1997. **71**(10): p. 7387-7392.
43. Yan, H., et al., *Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus.* elife, 2012. **1**: p. e00049.
44. Schulze, A., P. Gripon, and S. Urban, *Hepatitis B virus infection initiates with a large surface protein–dependent binding to heparan sulfate proteoglycans.* Hepatology, 2007. **46**(6): p. 1759-1768.
45. Gripon, P., et al., *Infection of a human hepatoma cell line by hepatitis B virus.* Proceedings of the National Academy of Sciences, 2002. **99**(24): p. 15655-15660.
46. Blanchet, M. and C. Sureau, *Infectivity determinants of the hepatitis B virus pre-S domain are confined to the N-terminal 75 amino acid residues.* Journal of virology, 2007. **81**(11): p. 5841-5849.
47. Slijepcevic, D., et al., *Impaired uptake of conjugated bile acids and Hepatitis B Virus preS1-binding in Na+-taurocholate cotransporting polypeptide knockout mice.* Hepatology, 2015: p. n/a-n/a.
48. Howard C. Thomas, A.S.F.L., Stephen A. Locarnini, Arie J. Zuckerman, *Viral Hepatitis, Fourth Edition*. Vol. Fourth Edition. 2013, New York, USA: John Wiley and Sons, Ltd.
49. Le Duff, Y., M. Blanchet, and C. Sureau, *The pre-S1 and antigenic loop infectivity determinants of the hepatitis B virus envelope proteins are functionally independent.* Journal of virology, 2009. **83**(23): p. 12443-12451.
50. Glebe, D., et al., *Pre-s1 antigen-dependent infection of Tupaia hepatocyte cultures with human hepatitis B virus.* Journal of virology, 2003. **77**(17): p. 9511-9521.
51. Zuckerman, J.N. and A.J. Zuckerman, *Mutations of the surface protein of hepatitis B virus.* Antiviral research, 2003. **60**(2): p. 75-78.
52. Makondo, E., T.G. Bell, and A. Kramvis, *Genotyping and molecular characterization of hepatitis B virus from human immunodeficiency virus-infected individuals in southern Africa.* PloS one, 2012. **7**(9): p. e46345.
53. Chen, M.T., et al., *A function of the hepatitis B virus precore protein is to regulate the immune response to the core antigen.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(41): p. 14913-14918.
54. Parekh, S., et al., *Genome replication, virion secretion, and e antigen expression of naturally occurring hepatitis B virus core promoter mutants.* Journal of virology, 2003. **77**(12): p. 6601-6612.
55. Porterfield, J.Z., et al., *Full-length hepatitis B virus core protein packages viral and heterologous RNA with similarly high levels of cooperativity.* Journal of virology, 2010. **84**(14): p. 7174-7184.
56. Locarnini, S. and M. Omata, *Molecular virology of hepatitis B virus and the development of antiviral drug resistance.* Liver International, 2006. **26**(S2): p. 11-22.
57. Milich, D. and T.J. Liang, *Exploring the biological basis of hepatitis B e antigen in hepatitis B virus infection.* Hepatology, 2003. **38**(5): p. 1075-1086.
58. Milich, D.R., et al., *The secreted hepatitis B precore antigen can modulate the immune response to the nucleocapsid: a mechanism for persistence.* The Journal of Immunology, 1998. **160**(4): p. 2013-2021.
59. Milich, D.R., et al., *Is a function of the secreted hepatitis B e antigen to induce immunologic tolerance in utero?* Proceedings of the National Academy of Sciences, 1990. **87**(17): p. 6599-6603.
60. Tong, S., et al., *Hepatitis B virus e antigen variants.* Int J Med Sci, 2005. **2**(1): p. 2-7.

61. Magnius, L.O. and J.Å. Espmark, *New specificities in Australia antigen positive sera distinct from the Le Bouvier determinants.* The Journal of Immunology, 1972. **109**(5): p. 1017-1021.
62. Pairan, A. and V. Bruss, *Functional surfaces of the hepatitis B virus capsid.* Journal of virology, 2009. **83**(22): p. 11616-11623.
63. Newman, M., et al., *Stability and morphology comparisons of self-assembled virus-like particles from wild-type and mutant human hepatitis B virus capsid proteins.* Journal of virology, 2003. **77**(24): p. 12950-12960.
64. Das, K., et al., *Molecular modeling and biochemical characterization reveal the 19 mechanism of hepatitis B virus polymerase resistance to lamivudine (3TC) and 20 emtricitabine (FTC).* J Virol, 2001. **75**: p. 4771-9.
65. Zoulim, F. and S. Locarnini, *Hepatitis B virus resistance to Nucleos(t)ide analogues.* Gastroenterology, 2009. **137**: p. 1593 - 1608.
66. Gish, R.G., et al., *Chronic hepatitis B: virology, natural history, current management and a glimpse at future opportunities.* Antiviral research, 2015. **121**: p. 47-58.
67. Shepard, C.W., et al., *Hepatitis B virus infection: epidemiology and vaccination.* Epidemiologic reviews, 2006. **28**(1): p. 112-125.
68. Mauss, S., et al., *Hepatology-A clinical textbook*. 2014: flying publisher.
69. Liaw, Y.F., Chu, C.M., *Hepatitis B virus infection.* The Lancet, 2009. **373**(9663): p. 582-592.
70. Kew, M.C., *Hepatitis B virus infection: the burden of disease in South Africa.* Southern African Journal of Epidemiology and Infection, 2008. **23**(1): p. 4-8.
71. Pourkarim, M.R., et al., *Molecular identification of hepatitis B virus genotypes/subgenotypes: revised classification hurdles and updated resolutions.* World J Gastroenterol, 2014. **20**(23): p. 7152-68.
72. Simonsen, L., et al., *In Focus-Unsafe injections in the developing world and transmission of bloodborne pathogens: A review.* Bulletin of the World Health Organization, 1999. **77**(10): p. 789-800.
73. Shi, Y.-H. and C.-H. Shi, *Molecular characteristics and stages of chronic hepatitis B virus infection.* World journal of gastroenterology: WJG, 2009. **15**(25): p. 3099.
74. François, G., et al., *Hepatitis B vaccination in Africa: mission accomplished?* Southern African Journal of Epidemiology and Infection, 2008. **23**(1): p. 24-28.
75. Supran, E.M., et al., *Enzyme-linked immunosorbent assay (ELISA) for the detection of hepatitis Be antigen and antibody: report of a field trial.* Journal of clinical pathology, 1983. **36**(5): p. 581-585.
76. Kramvis, A., Arbuthnot, P., Swaby, L.A., Ukomadu, C., *Hepatitis B virus prevalence, natural history, and treatment in Africans in Africa and the United States.* Curr Hepatitis B Rep, 2008. **2**: p. pp. 143–149.
77. Botha, J., et al., *Hepatitis B virus carrier state in black children in Ovamboland: role of perinatal and horizontal infection.* The Lancet, 1984. **323**(8388): p. 1210-1212.
78. Schweitzer, A., et al., *Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013.* The Lancet, 2015. **386**(10003): p. 1546-1555.
79. Prince, A.M., et al., *Hepatitis B Antigen in Wild-Caught Mosquitoes in Africa.* The Lancet. **300**(7771): p. 247-250.
80. Brotman, B., A. Prince, and H. Godfrey, *Role of arthropods in transmission of hepatitis-B virus in the tropics.* The Lancet, 1973. **301**(7815): p. 1305-1308.
81. Jupp, P., et al., *Infection of the common bedbug (Cimex lectularius L.) with hepatitis B virus in South Africa.* S Afr Med J, 1978. **53**(15): p. 598-600.
82. Ni, Y., et al., *Hepatitis B and D viruses exploit sodium taurocholate co-transporting polypeptide for species-specific entry into hepatocytes.* Gastroenterology, 2014. **146**(4): p. 1070-1083. e6.

83. Kramvis, A. and C.J. Clements, *Implementing a birth dose of hepatitis B vaccine for home deliveries in Africa—Too soon?* Vaccine, 2010. **28**(39): p. 6408-6410.
84. Kourtis, A.P., et al., *HIV–HBV Coinfection — A Global Challenge.* New England Journal of Medicine, 2012. **366**(19): p. 1749-1752.
85. Fultz, P.N., et al., *Vaccine protection of chimpanzees against challenge with HIV-1-infected peripheral blood mononuclear cells.* Science, 1992. **256**(5064): p. 1687-1690.
86. Vermeulen, M., et al., *Hepatitis B virus transmission by blood transfusion during 4 years of individual‐donation nucleic acid testing in South Africa: estimated and observed window period risk.* Transfusion, 2012. **52**(4): p. 880-892.
87. Komiya, Y., et al., *Minimum infectious dose of hepatitis B virus in chimpanzees and difference in the dynamics of viremia between genotype A and genotype C.* Transfusion, 2008. **48**(2): p. 286-294.
88. Attia, S., et al., *Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis.* Aids, 2009. **23**(11): p. 1397-1404.
89. Tabuchi, A., et al., *Titration of hepatitis B virus infectivity in the sera of pre‐acute and late acute phases of HBV infection: Transmission experiments to chimeric mice with human liver repopulated hepatocytes.* Journal of medical virology, 2008. **80**(12): p. 2064-2068.
90. Bond, W.W., et al., *Inactivation of hepatitis B virus by intermediate-to-high-level disinfectant chemicals.* Journal of Clinical Microbiology, 1983. **18**(3): p. 535-538.
91. Wang, H.C., et al., *Hepatitis B virus pre‐S mutants, endoplasmic reticulum stress and hepatocarcinogenesis.* Cancer science, 2006. **97**(8): p. 683-688.
92. Beasley, R.P., *Hepatitis B virus. The major etiology of hepatocellular carcinoma.* Cancer, 1988. **61**(10): p. 1942-1956.
93. Hadziyannis, S., et al., *Cytoplasmic hepatitis B antigen in" ground-glass" hepatocytes of carriers.* Arch. Pathol., 1973. **96**(5): p. 327-30.
94. Organization, W.H., *Guidelines for the prevention, care and treatment of persons with chronic hepatitis B infection.* 2015.
95. Lozano, R., et al., *Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010.* The Lancet, 2013. **380**(9859): p. 2095-2128.
96. Kramvis, A. and M.C. Kew, *Epidemiology of hepatitis B virus in Africa, its genotypes and clinical associations of genotypes.* Hepatology research, 2007. **37**(s1): p. S9-S19.
97. Burnett, R., et al., *Hepatitis B virus and human immunodeficiency virus co‐infection in sub‐Saharan Africa: a call for further investigation.* Liver international, 2005. **25**(2): p. 201-213.
98. Kiire, C., *The epidemiology and prophylaxis of hepatitis B in sub-Saharan Africa: a view from tropical and subtropical Africa.* Gut, 1996. **38**(Suppl 2): p. S5-12.
99. Dibisceglie, A.M., et al., *Prevalence of hepatitis B virus infection among black children in Soweto.* British Medical Journal (Clinical research ed.), 1986. **292**(6533): p. 1440-1442.
100. Firnhaber, C., et al., *The prevalence of hepatitis B co-infection in a South African urban government HIV clinic.* SAMJ: South African Medical Journal, 2008. **98**(7): p. 541-544.
101. Schoub, B., et al., *Universal immunization of infants with low doses of a low-cost, plasma-derived hepatitis B vaccine in South Africa.* Bulletin of the World Health Organization, 2002. **80**(4): p. 277-281.
102. Baldauf, S.L., *Phylogeny for the faint of heart: a tutorial.* TRENDS in Genetics, 2003. **19**(6): p. 345-351.

103. Baele, G., et al., *Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty.* Molecular biology and evolution, 2012. **29**(9): p. 2157-2167.
104. Drummond, A.J., et al., *Bayesian Phylogenetics with BEAUti and the BEAST 1.7.* Molecular Biology and Evolution, 2012. **29**(8): p. 1969-1973.
105. Darriba, D., et al., *jModelTest 2: more models, new heuristics and parallel computing.* Nature methods, 2012. **9**(8): p. 772-772.
106. Posada, D. and T.R. Buckley, *Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests.* Syst Biol, 2004. **53**(5): p. 793-808.
107. Kramvis, A., et al., *Relationship of serological subtype, basic core promoter and precore mutations to genotypes/subgenotypes of hepatitis B virus.* J Med Virol, 2008. **80**: p. 27 - 46.
108. Kramvis, A., *Genotypes and genetic variability of hepatitis B virus.* Intervirology, 2014. **57**(3-4): p. 141-150.
109. Arauz-Ruiz, P., et al., *Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America.* J Gen Virol, 2002. **83**: p. 2059 - 2073.
110. Kahila Bar‐Gal, G., et al., *Tracing hepatitis B virus to the 16th century in a Korean mummy.* Hepatology, 2012. **56**(5): p. 1671-1680.
111. Kramvis, A. and D. Paraskevis, *Subgenotype A1 of HBV–tracing human migrations in and out of Africa.* Antivir Ther, 2013. **18**(3 Pt B): p. 513-521.
112. Locarnini, S., et al. *Possible origins and evolution of the hepatitis B virus (HBV)*. in *Seminars in cancer biology*. 2013. Elsevier.
113. Zhou, Y. and E.C. Holmes, *Bayesian estimates of the evolutionary rate and age of hepatitis B virus.* Journal of molecular evolution, 2007. **65**(2): p. 197-205.
114. Bouckaert, R., M.V. Alvarado-Mora, and J.R. Pinho, *Evolutionary rates and HBV: issues of rate estimation with Bayesian molecular methods.* Antivir Ther, 2013. **18**: p. 497-503.
115. Kew, M.C., et al., *Increased hepatocarcinogenic potential of hepatitis B virus genotype A in Bantu‐speaking sub‐saharan Africans.* Journal of medical virology, 2005. **75**(4): p. 513-521.
116. Sugiyama, M., et al., *Influence of hepatitis B virus genotypes on the intra‐and extracellular expression of viral DNA and antigens.* Hepatology, 2006. **44**(4): p. 915-924.
117. Lin, C.-M., et al., *Functional analysis of hepatitis B virus pre-s deletion variants associated with hepatocellular carcinoma.* J Biomed Sci, 2012. **19**(1): p. 17-25.
118. Ochwoto, M., et al., *Genotyping and molecular characterization of hepatitis B virus in liver disease patients in Kenya.* Infection, Genetics and Evolution, 2013. **20**: p. 103-110.
119. Huy, T.T.-T., et al., *High prevalence of hepatitis B virus pre-s mutant in countries where it is endemic and its relationship with genotype and chronicity.* Journal of clinical microbiology, 2003. **41**(12): p. 5449-5455.
120. McMahon, B.J., *The influence of hepatitis B virus genotype and subgenotype on the natural history of chronic hepatitis B.* Hepatology international, 2009. **3**(2): p. 334-342.
121. Tanaka, Y., et al., *A case control study for differences among hepatitis B virus infections of genotypes A (subtypes Aa and Ae) and D.* Hepatology, 2004. **40**(3): p. 747-755.
122. Kramvis, A. and M.C. Kew, *Molecular characterization of subgenotype A1 (subgroup Aa) of hepatitis B virus.* Hepatology Research, 2007. **37**(s1): p. S27-S32.

123. Kimbi, G.C., A. Kramvis, and M.C. Kew, *Distinctive sequence characteristics of subgenotype A1 isolates of hepatitis B virus from South Africa.* Journal of general virology, 2004. **85**(5): p. 1211-1220.
124. Liu, S., et al., *Associations between hepatitis B virus mutations and the risk of hepatocellular carcinoma: a meta-analysis.* Journal of the National Cancer Institute, 2009.
125. Skelton, M., M.C. Kew, and A. Kramvis, *Distinct mutant hepatitis B virus genomes, with alterations in all four open reading frames, in a single South African hepatocellular carcinoma patient.* Virus research, 2012. **163**(1): p. 59-65.
126. Lucas, S. and A.M. Nelson, *HIV and the spectrum of human disease.* The Journal of pathology, 2015. **235**(2): p. 229-241.
127. Sonderup, M.W. and C.W. Spearman, *HIV/HBV and HIV/HCV Co-infection in Sub-Saharan Africa: Transmission, Disease Outcomes, and Treatment Options.*
128. Hoffmann, C.J. and C.L. Thio, *Clinical implications of HIV and hepatitis B co-infection in Asia and Africa.* The Lancet infectious diseases, 2007. **7**(6): p. 402-409.
129. Thio, C.L., *Hepatitis B and human immunodeficiency virus coinfection.* Hepatology, 2009. **49**(S5): p. S138-S145.
130. Stabinski, L., et al., *Prevalence of HIV and hepatitis B virus co-infection in sub-Saharan Africa and the potential impact and program feasibility of hepatitis B surface antigen screening in resource-limited settings.* JAIDS Journal of Acquired Immune Deficiency Syndromes, 2015. **68**: p. S274-S285.
131. Thomas, D.L., *Growing importance of liver disease in HIV‐infected persons.* Hepatology, 2006. **43**(S1): p. S221-S229.
132. Thio, C.L., et al., *Characterization of HIV-HBV co-infection in a multi-national HIV-infected cohort.* AIDS (London, England), 2013. **27**(2): p. 191.
133. Bell, T.G., *Clinical and molecular characterisation of Hepatitis B Virus Infection in Human Immunodeficienfy Virus-Positive Southern African Adults, Facilitated by newly developed Bioinformatic Tools* in *Department of Internal Medicine*. 2013, University of the Witwatersrand: Johannesburg. p. 202.
134. Mphahlele, M.J., et al., *High risk of occult hepatitis B virus infection in HIV-positive patients from South Africa.* Journal of Clinical Virology, 2006. **35**(1): p. 14-20.
135. Burnett, R., et al., *Increased exposure to hepatitis B virus infection in HIV-positive South African antenatal women.* International journal of STD & AIDS, 2007. **18**(3): p. 152-156.
136. Hoffmann, C.J., et al., *Hepatitis B virus infection and response to antiretroviral therapy (ART) in a South African ART program.* Clinical Infectious Diseases, 2008. **47**(11): p. 1479-1485.
137. Lukhwareni, A., et al., *Increased detection of HBV DNA in HBsAg‐positive and HBsAg‐negative South African HIV/AIDS patients enrolling for highly active antiretroviral therapy at a Tertiary Hospital.* Journal of medical virology, 2009. **81**(3): p. 406-412.
138. Barth, R.E., et al., *Presence of occult HBV, but near absence of active HBV and HCV infections in people infected with HIV in rural South Africa.* Journal of medical virology, 2011. **83**(6): p. 929-934.
139. Boyles, T.H. and K. Cohen, *The prevalence of hepatitis B infection in a rural South African HIV clinic.* SAMJ: South African Medical Journal, 2011. **101**(7): p. 470-471.
140. Matthews, P.C., et al., *Characterizing the epidemiology and interaction between HIV-1 and HBV co-infection in South Africa.* BMC Infectious Diseases, 2014. **14**(Suppl 2): p. P18.
141. Matthews, G.V., et al., *Impact of lamivudine on HIV and hepatitis B virus-related outcomes in HIV/hepatitis B virus individuals in a randomized clinical trial of antiretroviral therapy in southern Africa.* Aids, 2011. **25**(14): p. 1727-1735.

142. Andersson, M., et al., *The epidemiology of hepatitis B virus infection in HIV-infected and HIV-uninfected pregnant women in the Western Cape, South Africa.* Vaccine, 2013. **31**(47): p. 5579-5584.
143. Mayaphi, S.H., et al., *Variability of the preC/C region of hepatitis B virus genotype A from a South African cohort predominantly infected with HIV.* Journal of medical virology, 2013. **85**(11): p. 1883-1892.
144. Gededzha, M.P., et al., *Complete genome analysis of hepatitis B virus in human immunodeficiency virus infected and uninfected South Africans.* Journal of medical virology, 2016.
145. Raimondo, G., et al., *Statements from the Taormina expert meeting on occult hepatitis B virus infection.* Journal of Hepatology, 2008. **49**(4): p. 652-657.
146. Raimondo, G., et al., *Occult hepatitis B virus infection.* Journal of hepatology, 2007. **46**(1): p. 160-170.
147. Powell, E.A., et al., *Mutations associated with occult hepatitis B in HIV‐positive South Africans.* Journal of medical virology, 2015. **87**(3): p. 388-400.
148. Huang, X. and F. Hollinger, *Occult hepatitis B virus infection and hepatocellular carcinoma: a systematic review.* Journal of viral hepatitis, 2014. **21**(3): p. 153-162.
149. Sagnelli, E., et al., *Clinical impact of occult hepatitis B virus infection in immunosuppressed patients.* World journal of hepatology, 2014. **6**(6): p. 384.
150. Jardim, R.N.C.M., et al., *Occult hepatitis B virus infection in immunocompromised patients.* Brazilian Journal of Infectious Diseases, 2008. **12**(4): p. 300-305.
151. Barth, R.E., et al., *Hepatitis B/C and HIV in sub-Saharan Africa: an association between highly prevalent infectious diseases. A systematic review and meta-analysis.* International Journal of Infectious Diseases, 2010. **14**(12): p. e1024-e1031.
152. Chang, C., et al., *Immune reconstitution disorders in patients with HIV infection: from pathogenesis to prevention and treatment.* Current HIV/AIDS Reports, 2014. **11**(3): p. 223-232.
153. Bell, T.G., et al., *Hepatitis B virus infection in human immunodeficiency virus infected southern African adults: occult or overt–that is the question.* 2012.
154. Chun, H.M., et al., *Hepatitis B virus coinfection negatively impacts HIV outcomes in HIV seroconverters.* Journal of Infectious Diseases, 2012. **205**(2): p. 185-193.
155. Oliveira, M.P., et al., *Overt and occult hepatitis B virus infection among treatment‐naïve HIV‐infected patients in Brazil.* Journal of Medical Virology, 2016.
156. Revill, P., et al., *Global strategies are required to cure and eliminate HBV infection.* Nature Reviews Gastroenterology and Hepatology, 2016.
157. Parkin, D.M., et al., *Part I: Cancer in Indigenous Africans—burden, distribution, and trends.* The lancet oncology, 2008. **9**(7): p. 683-692.
158. Günther, S., et al., *A novel method for efficient amplification of whole hepatitis B virus genomes permits rapid functional analysis and reveals deletion mutants in immunosuppressed patients.* Journal of virology, 1996. **69**(9): p. 5437-5444.
159. Sterneck, M. and H. Will, *Naturally occurring variants of hepatitis B virus.* Adv Virus Res, 1999. **52**(25): p. 60298-5.
160. Smith, D.B., et al., *Virus' quasispecies': making a mountain out of a molehill?* Journal of General Virology, 1997. **78**(7): p. 1511-1519.
161. Samal, J., M. Kandpal, and P. Vivekanandan, *Hepatitis B "e" antigen-mediated inhibition of HBV replication fitness and transcription efficiency in vitro.* Virology, 2015. **484**: p. 234-240.
162. Melegari, M., P.P. Scaglioni, and J.R. Wands, *The small envelope protein is required for secretion of a naturally occurring hepatitis B virus mutant with pre-S1 deleted.* Journal of virology, 1997. **71**(7): p. 5449-5454.

163. Kalinina, T., et al., *Selection of a secretion-incompetent mutant in the serum of a patient with severe hepatitis B.* Gastroenterology, 2003. **125**(4): p. 1077-1084.
164. Garcia, T., et al., *Drastic reduction in the production of subviral particles does not impair hepatitis B virus virion secretion.* Journal of virology, 2009. **83**(21): p. 11152-11165.
165. Guo, H., et al., *Characterization of the intracellular deproteinized relaxed circular DNA of hepatitis B virus: an intermediate of covalently closed circular DNA formation.* Journal of virology, 2007. **81**(22): p. 12472-12484.
166. Xu, Z. and T. Yen, *Intracellular retention of surface protein by a hepatitis B virus mutant that releases virion particles.* Journal of virology, 1996. **70**(1): p. 133-140.
167. Yin, J., et al., *Significant association of different preS mutations with hepatitis B-related cirrhosis or hepatocellular carcinoma.* Journal of Gastroenterology, 2010. **45**(10): p. 1063-1071.
168. Kao, J.H., et al., *Fine mapping of hepatitis B virus pre-S deletion and its association with hepatocellular carcinoma.* Liver International, 2012. **32**(9): p. 1373-1381.
169. Chen, B.Ä., et al., *High prevalence and mapping of pre-S deletion in hepatitis B virus carriers with progressive liver diseases.* Gastroenterology, 2006. **130**(4): p. 1153-1168.
170. Sukowati, C.H., et al., *Significance of hepatitis virus infection in the oncogenic initiation of hepatocellular carcinoma.* World journal of gastroenterology, 2016. **22**(4): p. 1497.
171. Locarnini, S., J. McMillan, and A. Bartholomeusz. *The hepatitis B virus and common mutants*. in *Seminars in liver disease*. 2003.
172. Sinn, D.H., et al., *Pre-s mutation is a significant risk factor for hepatocellular carcinoma development: a long-term retrospective cohort study.* Digestive diseases and sciences, 2013. **58**(3): p. 751-758.
173. Bock, C.T., et al., *A preS mutation isolated from a patient with chronic hepatitis B infection leads to virus retention and misassembly.* Gastroenterology, 1997. **113**(6): p. 1976-1982.
174. Qu, L.S., et al., *Combined pre‐S deletion and core promoter mutations related to hepatocellular carcinoma: A nested case‐control study in China.* Hepatology Research, 2011. **41**(1): p. 54-63.
175. Lee, S.-A., et al., *Nucleotide change of codon 182 in the surface gene of hepatitis B virus genotype C leading to truncated surface protein is associated with progression of liver diseases.* Journal of hepatology, 2012. **56**(1): p. 63-69.
176. Yang, J.D. and L.R. Roberts, *Hepatocellular carcinoma: a global view.* Nature Reviews Gastroenterology and Hepatology, 2010. **7**(8): p. 448-458.
177. Sugauchi, F., et al., *Genotype, serotype, and phylogenetic characterization of the complete genome sequence of hepatitis B virus isolates from Malawian chronic carriers of the virus.* Journal of medical virology, 2003. **69**(1): p. 33-40.
178. Takahashi, K., et al., *Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: comparative analysis of 40 full-length isolates.* Archives of virology, 1998. **143**(12): p. 2313-2326.
179. Nurainy, N., et al., *Genetic study of hepatitis B virus in Indonesia reveals a new subgenotype of genotype B in east Nusa Tenggara.* Archives of virology, 2008. **153**(6): p. 1057-1065.
180. Stuyver, L.J., et al., *Nomenclature for antiviral‐resistant human hepatitis B virus mutations in the polymerase region.* Hepatology, 2001. **33**(3): p. 751-757.
181. Hsieh, Y.H., et al., *Hepatitis B virus pre‐S2 mutant large surface protein inhibits DNA double‐strand break repair and leads to genome instability in hepatocarcinogenesis.* The Journal of pathology, 2015. **236**(3): p. 337-347.

182. Su, I.-J., et al., *The emerging role of hepatitis B virus pre-S2 deletion mutant proteins in HBV tumorigenesis.* J Biomed Sci, 2014. **21**: p. 98.
183. Hsieh, Y.-H., et al., *Pre-S mutant surface antigens in chronic hepatitis B virus infection induce oxidative stress and DNA damage.* Carcinogenesis, 2004. **25**(10): p. 2023-2032.
184. Chisari, F.V., et al., *Molecular pathogenesis of hepatocellular carcinoma in hepatitis B virus transgenic mice.* Cell, 1989. **59**(6): p. 1145-1156.
185. Urban, S., et al., *Strategies to inhibit entry of HBV and HDV into hepatocytes.* Gastroenterology, 2014. **147**(1): p. 48-64.
186. Nakabayashi, H., et al., *Growth of human hepatoma cell lines with differentiated functions in chemically defined medium.* Cancer research, 1982. **42**(9): p. 3858-3863.
187. Sureau, C., et al., *Production of hepatitis B virus by a differentiated human hepatoma cell line after transfection with cloned circular HBV DNA.* Cell, 1986. **47**(1): p. 37-47.
188. Cavallone, D., et al., *Optimization of in vitro HBV replication and HBsAg production in HuH7 cell line.* Journal of virological methods, 2013. **189**(1): p. 110-117.
189. Krelle, A.C., A.S. Okoli, and G.L. Mendz, *Huh-7 Human Liver Cancer Cells: A Model System to Understand Hepatocellular Carcinoma and Therapy.* 2013.
190. Ochiya, T., et al., *An in vitro system for infection with hepatitis B virus that uses primary human fetal hepatocytes.* Proceedings of the National Academy of Sciences, 1989. **86**(6): p. 1875-1879.
191. Sells, M.A., M.L. Chen, and G. Acs, *Production of hepatitis B virus particles in Hep G2 cells transfected with cloned hepatitis B virus DNA.* Proceedings of the National Academy of Sciences, 1987. **84**(4): p. 1005-1009.
192. Gopalakrishnan, D., et al., *Hepatitis B virus subgenotype A1 predominates in liver disease patients from Kerala, India.* World J Gastroenterol, 2013. **19**(48): p. 9294-9306.
193. Bhoola, N.H., et al., *Construction of replication competent plasmids of hepatitis B virus subgenotypes A1, A2 and D3 with authentic endogenous promoters.* Journal of virological methods, 2014. **203**: p. 54-64.
194. Sambrook, J., D.W. Russell, and D.W. Russell, *Molecular cloning: a laboratory manual (3-volume set)*. Vol. 999. 2001: Cold spring harbor laboratory press Cold Spring Harbor, New York:.
195. Kumar, S., K. Tamura, and M. Nei, *MEGA: molecular evolutionary genetics analysis software for microcomputers.* Computer applications in the biosciences: CABIOS, 1994. **10**(2): p. 189-191.
196. Tamura, K., et al., *MEGA6: molecular evolutionary genetics analysis version 6.0.* Molecular biology and evolution, 2013. **30**(12): p. 2725-2729.
197. Bell, T.G. and A. Kramvis, *Fragment Merger: An Online Tool to Merge Overlapping Long Sequence Fragments.* Viruses, 2013. **5**(3): p. 824-833.
198. Bell, T. and A. Kramvis, *Bioinformatics Tools for Small Genomes, Such as Hepatitis B Virus.* Viruses, 2015. **7**(2): p. 781-797.
199. Bell, T. and A. Kramvis, *Mutation Reporter Tool: An online tool to interrogate loci of interest, with its utility demonstrated using hepatitis B virus.* Virology Journal, 2013. **10**(1): p. 62.
200. Nicholas, K.B., H.B.J. Nicholas, and D.W. Deerfield, *GeneDoc: analysis and visualization of genetic variation.* Embnew. news, 1997. **4**(1).
201. Posada, D., *Using MODELTEST and PAUP* to select a model of nucleotide substitution.* Curr Protoc Bioinformatics, 2003. **Chapter 6**: p. Unit 6.5.
202. Gutell, R.R. and R.K. Jansen, *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion.* 2006.

203. Alting-Mees, M.A. and J.M. Short, *pBluescript II: gene mapping vectors.* Nucleic acids research, 1989. **17**(22): p. 9494.
204. Chung, C.T., S.L. Niemela, and R.H. Miller, *One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution.* Proceedings of the National Academy of Sciences, 1989. **86**(7): p. 2172-2175.
205. Jerpseth, B. and B.L. Kretz, *SCS110: dam, dcm, endA Epicurian coli competent cells.* Strategies, 1993. **6**(22): p. 15.
206. Cobb, B.D. and J.M. Ciarkson, *A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods.* Nucleic Acids Research, 1994. **22**(18): p. 3801-3805.
207. Samal, J., M. Kandpal, and P. Vivekanandan, *A simple and rapid method for the quantitation of secreted hepatitis B virions in cell culture models.* Indian journal of medical microbiology, 2015. **33**(2): p. 290.
208. Xia, X. and Z. Xie, *DAMBE: software package for data analysis in molecular biology and evolution.* Journal of heredity, 2001. **92**(4): p. 371-373.
209. Saitou, N. and M. Nei, *The neighbor-joining method: a new method for reconstructing phylogenetic trees.* Molecular biology and evolution, 1987. **4**(4): p. 406-425.
210. Yang, Z. and B. Rannala, *Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method.* Molecular biology and evolution, 1997. **14**(7): p. 717-724.
211. Tong, S., *Impact of viral genotypes and naturally occurring mutations on biological properties of hepatitis B virus.* Hepatology Res, 2007. **37**: p. S3 - S8.
212. Lee, S.-A., et al., *Male-specific W4P/R mutation in the pre-S1 region of hepatitis B virus, increasing the risk of progression of liver diseases in chronic patients.* Journal of clinical microbiology, 2013. **51**(12): p. 3928-3936.
213. Gous, N., et al., *Retrospective characterization of the S open reading frame of HBV isolated from children with membranous nephropathy treated with interferon-alpha2b.* Antivir Ther, 2010. **15**(1): p. 61-69.
214. Lago, B.V., et al., *Hepatitis B Virus Subgenotype A1: Evolutionary Relationships between Brazilian, African and Asian Isolates.* PLoS ONE, 2014. **9**(8): p. e105317.
215. He, J.-D., et al., *Evolutionary perspective on hepatitis B virus with an expanded sampling strategy.* Virus research, 2013. **178**(2): p. 525-529.
216. Iles, J.C., et al., *Phylogeography and epidemic history of hepatitis C virus genotype 4 in Africa.* Virology, 2014. **464**: p. 233-243.
217. Holmes, E.C., *Evolutionary history and phylogeography of human viruses.* Annu. Rev. Microbiol., 2008. **62**: p. 307-328.
218. Mora, M.V.A., et al., *Molecular characterization of the Hepatitis B virus genotypes in Colombia: a Bayesian inference on the genotype F.* Infection, genetics and Evolution, 2011. **11**(1): p. 103-108.
219. Khan, N., et al., *Modulation of hepatitis B virus secretion by naturally occurring mutations in the S gene.* Journal of virology, 2004. **78**(7): p. 3262-3270.
220. Melegari, M., S. Bruno, and J.R. Wands, *Properties of hepatitis B virus pre-S1 deletion mutants.* Virology, 1994. **199**(2): p. 292-300.
221. Zhang, A.-Y., et al., *Evolutionary Changes of Hepatitis B Virus Pre-S Mutations Prior to Development of Hepatocellular Carcinoma.* PLoS ONE, 2015. **10**(9): p. e0139478.
222. Nassal, M., *The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly.* Journal of Virology, 1992. **66**(7): p. 4106-16.