

**MOLECULAR CHARACTERIZATION OF THE HEPATITIS B  
VIRUS X GENE**

by

**LESIBANA ANTHONY MALINGA**

RESEARCH DISSERTATION

Submitted in fulfilment of the requirements for the degree of

**MASTER OF MEDICAL SCIENCE**

in

**MEDICAL VIROLOGY**

In the

**FACULTY OF HEALTH SCIENCES**

**(School of Pathology)**

at the

**UNIVERSITY OF LIMPOPO**

**(MEDUNSA CAMPUS)**

**SUPERVISOR: Dr SG Selabe**

**CO-SUPERVISOR: PROF MJ MPHAHLELE**

**2010**

**DECLARATION**

I, **LESIBANA ANTHONY MALINGA** hereby declare that the work done on this dissertation is my own and has not been submitted at this or any other institution for any degree purposes.

This dissertation is being submitted in fulfilment for the requirements of the degree of Master of Medical Science (Medical Virology), in the Department of Virology, School of Pathology, Faculty of Health Sciences, at the University of Limpopo, MEDUNSA campus.

.....

**Initials & Surname (Title)**

**Student Number:.....**

.....

**Date**

## **DEDICATION**

I dedicate this work to Malinga's family, especially my wonderful and supportive parents, my two siblings, Thuli and Tumi and cousin Vincent. Lasty, ACTS branch of MEDUNSA campus for supporting me spiritually in my studies.

## **ACKNOWLEDGEMENTS**

I would like to thank the following:

GOD Almighty for giving me a chance to continue with my studies and for the countless blessings he has given me.

Dr SG Selabe and Prof MJ Mphlahlele for their outstanding supervision, generosity and support that they have given me through my studies.

The Department of Virology, University of Limpopo (MEDUNSA campus) for accepting me to study with them and form a pool of postgraduate students.

To all postgraduate students of Hepatitis and HIV research unit laboratory, Maemu, Lizzy, Nare, Azwi, Khutso for their moral support and technical assistance.

Lastly to the University of Limpopo (MEDUNSA campus), National Research Foundation (NRF) and Poliomyelitis Research Foundation for their financial support.

## **ABSTRACT**

**Introduction:** Hepatitis B virus (HBV) is a serious problem worldwide causing various liver diseases such as chronic hepatitis and hepatocellular carcinoma (HCC). The pathogenesis of HBV related HCC is not well established. Hepatitis B X protein (HBx) plays an important role in the pathogenesis of HCC. HBx coded by HBV X gene enhances several cellular pathways in hepatocytes which may lead to HCC. The genetic variability of other HBV genomic regions plays a significant role in diagnosis, vaccine development and drug resistance. However, the genetic variability of HBV X gene is not well understood. In addition the dual basal core promoter mutations found within the X gene have been implicated in the inhibition of hepatitis B e antigen (HBeAg) expression. Studies focusing on HBV X gene are scarce in South Africa. Consequently HBV X gene variability may reveal interesting mutations and substitutions that are important in chronic liver diseases or HCC. This study aimed at characterising HBV X gene at a molecular level isolated from patients with different serological profiles.

**Methods:** This was an exploratory study which used 20 stored sera (-70°C) collected from adult patients at Dr George Mukhari hospital, Pretoria. The samples were already tested for HBsAg, anti-HBs, anti-HBc and HBeAg serological markers (Elecsys, Roche Diagnostics, Penzburg, Germany). HBV DNA extraction was performed from serum using High Pure Viral Nucleic Acid Assay (Roche Diagnostics, Penzburg, Germany). Nested PCR assay was used for the amplification of 465 nucleotide HBV X gene. Sequencing of PCR positive samples was done using spectruMedix SCE2410 genetic analysis system. Six samples selected, were cloned into the pGEM<sup>®</sup>-T Easy vector system (Promega, Madison, USA). Three clones of each sample were selected and their plasmids purified using Pure Yield<sup>™</sup> Plasmid Miniprep System (Promega, Madison, USA). The plasmid DNA was recovered using optimised nested PCR assay and sequenced. A total of 38 sequences were generated from the study and compared with reference strains retrieved from GenBank. Phylogenetic analysis based on HBV X gene sequences was done using MEGA 4 software to determine different genotype clusters.

**Results:** HBV X gene was successfully detected and amplified in 20 study samples. The sequenced HBV X gene products revealed mutations and insertions. Particularly a six nucleotide insertion, GCATGG between nucleotides 1611 and 1618 which was detected in five samples. In addition, the six cloned samples confirmed the six nucleotide insertion and other mutations associated with inhibition of hepatitis B e antigen (HBeAg) detected in the study. The substitutions within HBx were detected in the N (1-50 amino acids) and C (51-154) terminals by comparing our sequences with archival sequences from GenBank. Important substitutions found within the N and C terminals were S31A, P38S, A42P, F73L, H94Y, P101S, K118T, D119N, I127T/N, K130M and V131I. These substitutions are associated with various biological functions and pathogenesis. Other substitutions with unknown functions detected in the study include A2G, A3G, A4G, C6W, P42S and V116L. Further mutations of T1753M, A1762T and G1764A associated with inhibition of HBeAg expression were detected in most samples and only one sample had C1766T mutation. Phylogenetic analysis resulted in A, C and D HBV genotypes. Five samples and 11 clones clustered with genotype D, two samples and four clones clustered with genotype C and finally 13 samples and 3 clones clustered with genotype A.

**Conclusion:** HBV X gene was successfully characterised using various molecular methods. HBx substitutions detected are involved in various pathogenic effects and may present a risk of HCC for patients infected with HBV. Genotype D samples displayed most mutations/substitutions and this can be regarded as an important genotype with high risk of HCC. The detection of a six nucleotide insertion (GCATGG) in 5 samples may emerge as a new variant of genotype D. Furthermore triple mutations of T1753M/A1762T/G1764A within basal core promoter region were detected mostly in HBeAg negative samples. However further analysis of HBV X gene variability is needed.

## TABLE OF CONTENTS

Page No.

Title page	i
Declaration	ii
Dedication	iii
Acknowledgements	iv
Abstract	v
Table of contents	vii
List of abbreviations	xii
List of tables	xv
List of figures	xvi

### **CHAPTER 1: EXPERIMENTAL PROPOSAL**

1.1	Experimental proposal	1
1.2	Study problem	1
1.3	Aim of the study	2
1.4	Specific Objectives	2
1.5	Expected significance of the study	2

### **CHAPTER 2: LITERATURE REVIEW**

2.1	Introduction	3
2.2	Overview of hepatitis	3
2.2.1	History	3
2.2.2	HBV classification	4
2.2.3	Structure of HBV particles	4
2.3	HBV transmission	6
2.4	Epidemiology of HBV	6
2.4.1	Global prevalence of HBV	6

2.4.2	Distribution of HBV infection in Africa	8
2.4.3	HBV disease burden in South Africa	8
2.5	HBV pathogenesis	8
2.5.1	Disease profile of HBV infection	9
2.5.1.1	Acute hepatitis B	9
2.5.1.2	Chronic hepatitis B	10
2.5.1.3	Chronic hepatitis B phases	11
2.5.1.3.1	Immune tolerance	11
2.5.1.3.2	Immune clearance	12
2.5.1.3.3	Low replicating phase	12
2.5.1.3.4	Reactivation phase	12
2.5.1.4	Fulminant hepatitis B	12
2.6	Laboratory Diagnosis	13
2.6.1	Serological markers	15
2.6.2	Biochemical tests	16
2.6.3	Molecular assays	16
2.7	Prevention, control and management of HBV	17
2.7.1	Prevention and control of HBV	17
2.7.2	Management of HBV infection	18
2.8	Molecular biology of HBV	19
2.8.1	The genome organization of HBV	19
2.8.2	HBV replication	20
2.8.2.1	Entry and uncoating	20
2.8.2.2	Replication of the HBV genome	21
2.8.2.3	Assembly and budding	22

2.8.3	HBV subtypes and genotypes	23
2.8.4	Genetic variability of variability of HBV	25
2.8.4.1	Presurface and surface mutants	25
2.8.4.2	Precore/core mutants	26
2.8.4.3	Polymerase mutants	26
2.8.4.4	X gene mutants	28
2.9	Important roles of HBx in causing liver diseases	29
2.9.1	HBx and regulation of apoptosis	30
2.9.2	HBx and modulation of signalling pathways	31
2.9.3	HBx and transcription modulation	32
2.9.4	HBx and cell cycle	32
2.9.5	HBx and DNA repair	32
2.9.6	HBx and cytokines	32
2.9.7	Integration of X gene in hepatocytes	33
<b>CHAPTER 3: MATERIALS AND METHODS</b>		<b>34</b>
3.1	Ethical considerations	34
3.2	Study design	34
3.3	Study population	34
3.4	Laboratory methods	34
3.4.1	HBV serological testing	34
3.4.2	HBV PCR DNA assay	36
3.4.2.1	HBV DNA extraction	36
3.4.2.2	Amplification of HBV DNA	37
3.4.2.3	Detection of amplified HBV DNA products	39

3.4.2.4	Avoiding PCR contamination	39
3.4.3	HBV DNA cloning	40
3.4.3.1	Ligation of HBV DNA into pGEM <sup>®</sup> -T Easy vector	43
3.4.3.2	Transformation using pGEM <sup>®</sup> -T Easy vector ligation reactions	43
3.4.3.3	Isolation of recombinant plasmid HBV DNA	44
3.4.4	HBV DNA nucleotide sequencing	45
3.4.5	Nucleotide analysis	46
3.4.6	HBV phylogeny	47
<b>CHAPTER 4: RESULTS</b>		<b>48</b>
4.1	Overview of the results	48
4.2	Optimisation of PCR assay for amplification of HBV X gene	49
4.2.1	Detection of PCR products	50
4.3	Identification of HBV positive clones	50
4.4	DNA sequencing analysis	51
4.5	Nucleotide and amino acid sequence analysis of HBV X gene	53
4.6	Nucleotide variations within HBV X gene basal core promoter region	64
4.7	Phylogeny of HBV X gene sequences	66
<b>CHAPTER 5: DISCUSSION</b>		<b>68</b>
5.1	Overview of the results	68
5.2	HBx substitutions associated in its biological function and pathogenesis	69
5.3	Basal core promoter mutations involved in HBeAg expression	73

5.4	Implications of the study	74
		<b>76</b>
	<b>CHAPTER 6: CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS</b>	76
6.1	Conclusions	77
6.2	Limitations of the study	77
6.3	Recommendations	<b>78</b>
	<b>CHAPTER 7: REFERENCES</b>	

## LIST OF ABBREVIATIONS

µl	Microliter
3TC	Lamivudine
ADV	Adefovir dipovoxil
ALT	Alanine aminotransferase
Anti-HBc	Antibody to hepatitis B core antigen
Anti-HBe	Antibody to hepatitis B e antigen
Anti-HBs	Antibody to hepatitis B surface antigen
AP	Activator factor
AST	Aspartate aminotransferases
ATF	Activating transcription factor
Ca <sup>2+</sup>	Calcium
cccDNA	Covalently closed circular DNA
C/EBP	ccAAT/enhancer activity
CREB	cyclic AMP response element binding protein
CTL	Cytotoxic T lymphocytes
DNA	Deoxyribonucleic acid
DR	Direct repeat
Enh	Enhancer
EPI	Expanded Program on Immunisation
ER	Endoplasmic reticulum
ETV	Entacavir
HAV	Hepatitis A virus
HBcAg	Hepatitis B core antigen
HBeAg	Hepatitis B e antigen
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus
HBx	Hepatitis B x protein
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HDV	Hepatitis D virus
HEV	Hepatitis E virus
HLA	Human leukocyte antigen
IL	Interleukin
IgG	Immunoglobulin G
IgM	Immunoglobulin M
kDa	Kilo dalton
LB	Luria bertani
LCR	Ligase chain reaction
Ldt	Telbivudine
LHBs	Large hepatitis B surface antigen
MAPK	Mitogen activate protein kinase
MHBs	Medium hepatitis B surface antigen
MHC	Major histocompatibility complex
MHR	Major hydrophilic region
mRNA	Messenger RNA

NFKB	Nuclear factor kappa B
NPC	Nuclear pore complex
NGFR	Nerve growth factor receptor
°C	Degree celsius
ORF	Open reading frame
PCR	Polymerase chain reaction
PI3-K	Phosphotidylinositol 3-kinase
PkC	Protein kinase C
pgRNA	Pregenomic RNA
Pre C	Pre-core
Pre S	Pre-surface
Rb	Retinoblastoma
RNA	Ribonucleic acid
ROI	Reactive oxygen intermediates
RPB5	RNA polymerase II
ssDNA	Single stranded DNA
STAT	Signal transducer and activator of transcription
TDF	Tenofovir disoproxil fulmarate
TNF $\alpha$	Tumour necrosis-factor alpha
TNFR	Tumour necrosis factor receptor
VDAC3	Voltage dependant anion channel
WHO	World Health Organization

### Nucleotides

A	Adenine
C	Cytosine
G	Guanine
T	Thymine

### Amino acids

A	Alanine
R	Arginine
N	Asparagine
D	Aspartic acid
C	Cysteine
E	Glutamic acid
H	Histidine
I	Isoleucine
L	Leucine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
S	Serine

T	Threonine
W	Tryptophan
Y	Tyrosine
V	Valine

## LIST OF TABLES

<b>Table No.</b>	<b>Legend</b>	<b>Page No.</b>
2.1	Interpretation of HBV serologic tests	14
3.1	Serologic profiles of samples	35
3.2	Sequences of HBV X gene primers	38
4.1	A summary of amino acid substitutions relevant to HBx mediated pathogenesis	62
4.2	A summary of substitutions with unknown functions detected in the study	63
4.3	Correlation of HBeAg positive status with basal core promoter mutations	64
4.4	Correlation of HBeAg negative status with basal core promoter mutations	65

## LIST OF FIGURES

Figure No.	Legend	Page No.
2.1	Electron micrograph of circulating forms of HBV particles in the blood is shown at the top and a schematic drawing of Dane particle	5
2.2	Global prevalence of HBV infection	7
2.3	Acute hepatitis serology and clinical profile	10
2.4	Serologic and clinical profile of chronic hepatitis	11
2.5	HBV genome with a double stranded DNA with four open reading frames (ORFs); namely, polymerase (P), surface (pre S1/pre S2/S), precore (pre C)/core (C) and X	20
2.6	Molecular steps of HBV replication from attachment to translocation of the virion DNA in the nucleus, formation of cccDNA into viral gene products in the nucleus	22
2.7	Worldwide distribution of HBV genotypes and subtypes	24
2.8	Location of the major antiviral drug resistant mutations associated with lamivudine, telbivudine, adefovir, tenofovir and entacavir	27
3.1	The substrate X-Gal turns blue if the gene is intact. White colonies in X-Gal imply the presence of recombinant DNA in the plasmid	41
3.2	pGEM <sup>®</sup> -T Easy vector circle map	42
4.1	A gel electrophoretic diagram showing samples amplified with X gene primers	50
4.2	A gel electrophoretic diagram showing clones amplified with X gene primers	51
4.3	Chromatogram of the X gene visualised by chromas pro version 1.45 showing a six nucleotide insert 5'GCGATGG3' between nucleotide 1612 and 1618	52
4.4.	Alignment of HBV X gene nucleotide sequences compared with corresponding representative references of genotype A to H	59
4.5	The amino acid sequences deduced from HBV X gene nucleotides compared with corresponding representative references of genotypes A to H	61

4.6 A dendrogram obtained by neighbour-joining phylogenetic analysis of HBV X gene sequences rooted on genotype G 67

# CHAPTER 1

## 1. EXPERIMENTAL PROPOSAL

### 1.1. Study problem

Hepatitis B virus (HBV) infection is a serious problem worldwide, causing various liver diseases such as acute and chronic hepatitis, cirrhosis and hepatocellular carcinoma (HCC) and it affects approximately 400 million people in the world (Aliyu et al, 2004). In sub-Saharan Africa, HBV has a high carrier rate ranging between 5 and 20% of the population depending on the region (Kimbi et al, 2004).

The pathogenesis of HBV related to HCC has not yet been well established but evidence suggests that the hepatitis B X protein (HBx) plays a crucial role in the pathogenesis of HCC (Zhang et al, 2005). The HBV X gene codes for a 17 kilodalton (kDa) HBx protein which plays an important role in viral replication and enhances several cellular pathways in hepatocytes which may lead to HCC (Muroyama et al, 2006). The genetic variability of other genomic regions of the HBV have been shown to play a significant role medically as they are involved in diagnosis, immune evasion, vaccine development, immune activation functions and drug resistance (Datta et al, 2008). However, the role of HBx genetic variability is not well understood. The HBV X gene product has been linked to high risk acquisition of HCC due to its role in transactivation of certain cellular genes.

The genetic variability of HBx leads to genotypic-specific variants and mutants that emerge during chronic infection (Lin et al, 2005). In addition, the dual basal core promoter mutations within X gene have been implicated in hepatitis B e antigen (HBeAg) expression, which is a marker for viral replication. Studies focusing only on X gene and its variants are scarce in South Africa. Consequently, X gene variability may reveal interesting and novel mutations, sequence motifs or residues that are important for chronic liver diseases or HCC carcinogenesis.

## **1.2 Aim of the study**

To characterise HBV X gene at a molecular level from HBV isolates with different serological profiles

## **1.3 Specific objectives**

- a) To detect and amplify HBV X gene from HBV infected South African patients
- b) To sequence HBV X gene PCR positive products
- c) To clone and sequence HBV X genes of interest
- d) To analyse and compare generated sequence data set with GenBank HBV X gene sequences
- e) To analyse reported motifs or residues associated with certain biological functions of HBx protein
- f) To analyse novel,if any, motifs or residues found in the generated sequences

## **1.4 Expected significance of the study**

In South Africa, the relevance of X gene in liver diseases has not been studied extensively and this may be important in HBV infection as persistent viral infection can lead to the risk of progression to chronic hepatitis and HCC. Molecular characterisation of X gene may be useful in analysing its biological significance. Moreover, X gene genetic variability can enlighten the differential oncogenic potential of HBV genotypes circulating in the continent.

## **CHAPTER 2**

### **2. LITERATURE REVIEW**

#### **2.1 Introduction**

The word hepatitis is from ancient Greek with “hepa” meaning liver and suffix “-itis” meaning inflammation. Hepatitis results from liver injury characterised by the presence of inflamed hepatocytes. HBV is a global problem causing various liver associated with various liver disease such as HCC (But et al, 2008). The HBV genome has four open reading frames (ORFs); namely surface (S), core (C), polymerase (P) and X gene (Seeger and Mason, 2000). The S, C and P ORFs have high levels of genetic variability which have medical significance such as diagnostics, immune invasion and drug resistance respectively (Datta et al, 2008). The HBx gene is the smallest ORF that codes for HBx which plays an important role in viral replication and enhances several cellular pathways in hepatocytes which may lead to HCC (Muroyama et al, 2006). The medical significance of HBx genetic variability is not well understood. Furthermore, studies on the X gene are scarce in South Africa.

#### **2.2 Overview of hepatitis**

Hepatitis is caused by a range of agents namely; viruses, bacteria, drugs, toxins (i.e. alcohol) and autoimmune diseases. A group of primary viruses such as hepatitis A (HAV; formally called infectious hepatitis) and B (HBV; formerly called serum hepatitis) were the first agents recognised to cause hepatitis and later hepatitis C (HCV), D (HDV), E (HEV) and secondary viruses such as Epstein-Barr virus together with cytomegalovirus were also noted as important causative agents of hepatitis (Ryan and Ray, 2004).

##### **2.2.1 History**

HBV was firstly identified as an Australia antigen mostly isolated in leukaemia patients in 1967. The term Australia antigen was adapted because it was isolated

first in Australian natives, importantly belonging to the aborigine race (Blumberg et al, 1967). In 1970, it was later known as hepatitis because it was found in patients with serum hepatitis (Dane et al, 1970). The antigen was also found in serum of patients with viral hepatitis using immunologic precipitation techniques (Huang, et al 1972).

The partially purified antigen was found to be transmissible to the susceptible individuals and recognised as the antigenic determinant of HBV, which became known as the surface antigen or HBsAg (Huang et al, 1972; Kidd-Ljunggren et al, 2002). It has been shown to be a marker of infection responsible for causing acute and chronic liver diseases (Dudley et al, 1972, Kidd-Ljunggren et al, 2002).

### **2.2.2 HBV classification**

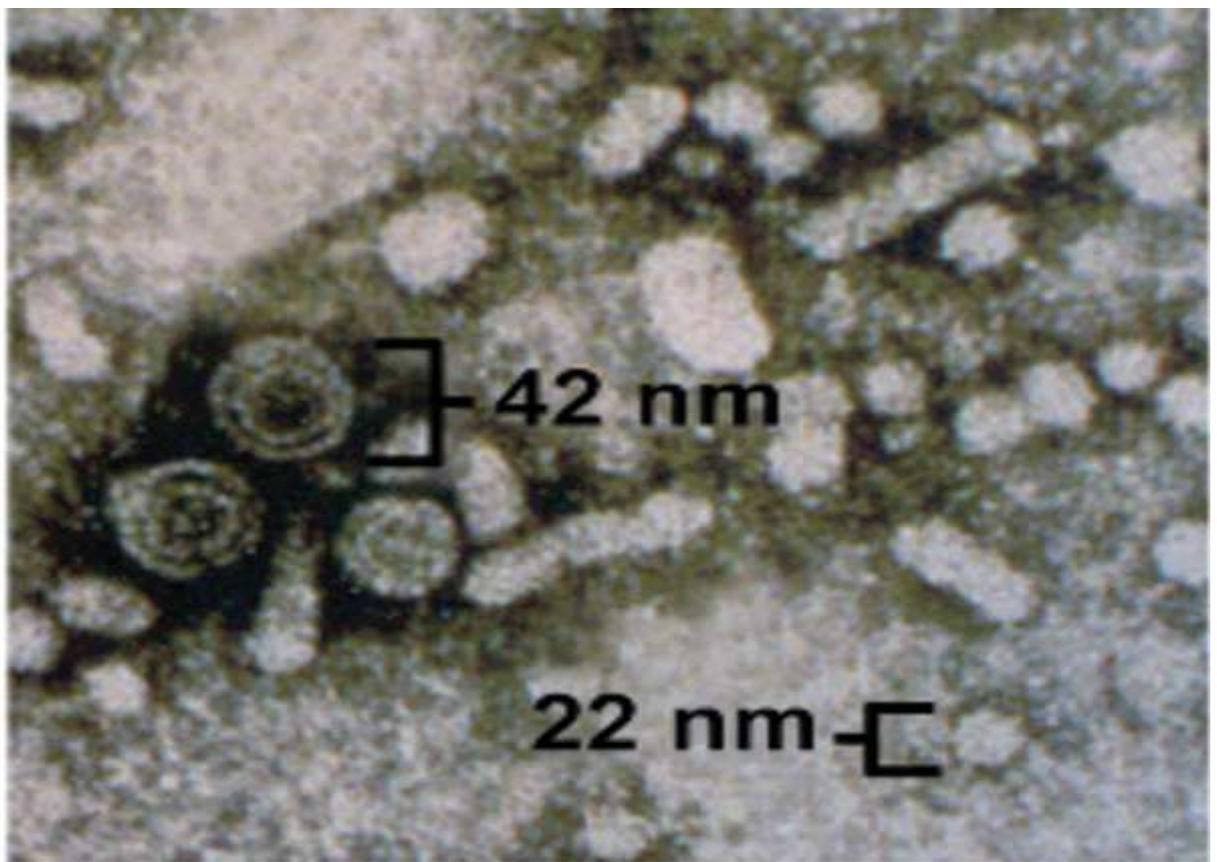
HBV belongs to the *hepadnaviridae* family. Hepadnaviruses are found in both mammals (orthohepadnaviruses) and birds (avihepadnaviruses). Other related hepadnaviruses were discovered in woodchucks (*Marmota monax*), ground squirrel (*Spermophilus beechey*), pekin duck (*Anas domesticus*) and herons (Schaefer, 2007; Liang, 2009). The viruses are small with similar genome organization and have an identical DNA with a sequence homology of about 40% to 70% (Seeger and Mason, 2000; Schaefer, 2007). Finally, hepadnaviruses are hepatotropic and their infection is species-specific or limited to closely related species causing variable degree of pathogenesis (Guo et al, 2005).

### **2.2.3 Structure of HBV particles**

HBV virion (Dane particle) is a spherical, double shelled structure with approximately 42 nm in diameter containing double-stranded relaxed circular DNA (rcDNA) genome. The genome is embedded within a nucleocapsid (core) that is surrounded by a lipid bilayer filled with complexes of viral glycoproteins (Block et al, 2007). The shell of the icosahedral nucleocapsid is formed by multiple copies of single protein, called the hepatitis B core protein (HBcAg), consisting of 153 or 183 amino acid residues depending on the HBV genotype (Bruss, 2007). The HBcAg

has a length of about 27 nm assembled from 240 viral capsid proteins (Block et al, 2007).

The HBV virion is composed of nucleocapsid surrounded by membranous lipoprotein envelope composing of lipids (i.e. triglycerides, phospholipids, cholesterol) derived from infected host (Francois et al, 2001). There are two additional non-infectious particles in the blood of HBV infected individuals: the filamentous forms which have varying lengths and width of 22 nm and spherical structures with a diameter of 20 nm (Liang, 2009) (Figure 2.1).



**Figure 2.1:** Electron micrograph of circulating forms of HBV particles in the blood is shown at the top and a schematic drawing of Dane particle (Adapted from Liang, 2009)

## **2.3 HBV transmission**

HBV is resistant to breakdown outside the body and it is easily transmitted through contact with infected body fluids (Wright, 2006). Asian countries have high endemic rates, with vertical transmission from mother to child regarded as the most common route, whereas in African countries, routes of transmission occur through horizontally between children particularly siblings (Zanetti et al, 2008).

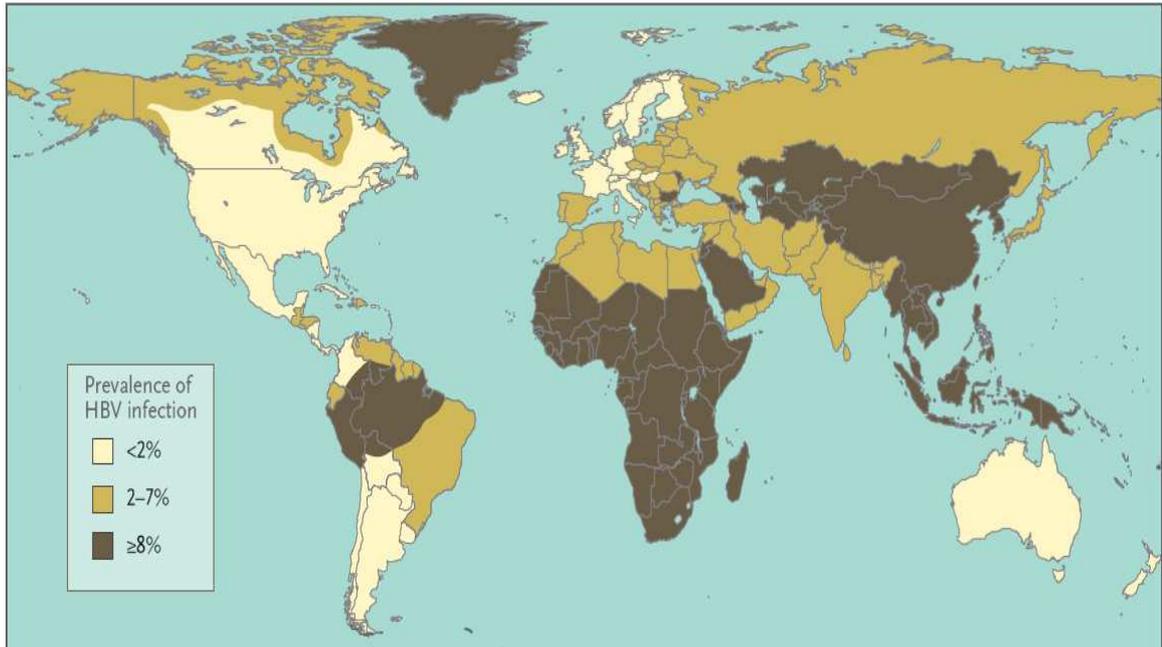
Blood transfusions, dialysis, needle stick accident among healthcare professionals and intravenous drug abuse are regarded primarily as risk factors for HBV infection (Khouri and Dos Santos, 2004). Bites from insects such as mosquitoes and also scarification of the skin by traditional doctors (during ritual initiation ceremonies or when giving treatment) are suspected to play a role in the transmission of the virus because they usually occur in a rural rather than an urban setting (Kew, 1996).

## **2.4 Epidemiology of HBV**

### **2.4.1 Global prevalence of HBV**

HBV infection is common to humans with about 400 million people infected worldwide and prevalence of chronic hepatitis B can be divided into three categories; high (>8%), intermediate (2-7%) and low (<2%) (Violante-Dehesa and Nunez-Nateras, 2007) (Figure 2.2). About 620 000 people die from acute and chronic hepatitis complications. There are around 4.5 million new cases worldwide with a quarter of them likely to progress to liver disease (Zanetti et al, 2008).

Persistent HBV infection contributes to the overall 75% to 80% of global HCC cases (Liu and Kao, 2007). An annual incidence has been estimated to be less than 1% in non-cirrhotic and about 2-3% in cirrhotic patients (Lin and Kao, 2008). In countries with endemic HBV infections such as sub-Saharan Africa and Asia, chronic hepatitis B has shown strong correlation with the development of HCC (Liu and Kao, 2007).



**Figure 2.2:** Global prevalence of HBV infection (Adapted from Diestag, 2008)

Areas of high endemicity include South-east Asia, sub-Saharan Africa, South Pacific, Amazon basin (i.e Colombia, Brazil, Peru and Venezuela). In these areas, the majority of infections are acquired during pregnancy or horizontally in early childhood (Valsakamis, 2007; Zanetti, 2008). Population of Southern parts of Eastern and Central Europe, the Middle East and the Indian subcontinent fall within intermediate HBV infection (Datta, 2008).

Low prevalence (<2%) has been shown in North America, Western and Northern Europe, Latin America and Australia, where transmission route is by sexual or by intravenous drug use (Valsakamis, 2007; Wright, 2006). In the United States, an estimated population of 1.2 million are infected with chronic HBV, but certain areas and populations such as Alaskan natives, Pacific islanders and infants born to mothers from high endemic areas have high HBV prevalence rates (Wright, 2006).

### **2.4.2 Distribution of HBV infection in Africa**

Africa is generally regarded as high endemic areas for HBV infection, but Tunisia and Morocco fall into the intermediate category (<7%) and regions of West and East Africa with high endemicity (>8%) (Andre, 2000). Senegal ranges among the highest with about 90% of the population infected and the rest of Central and Southern Africa fall within the high seroprevalence with an exception to Zambia showing chronic infection of 6.5 to 7.5% (Andre, 2000).

The average incidence through the continent has an average of 10% with other areas having 35% or higher (Kew 1996). About 56% to 98% of the adult population in sub-Saharan Africa, show signs of past or present infection with HBV (Kiire, 1996). Furthermore HBV prevalence rate is higher in black ethnic groups of sub-Saharan Africa (Burnett et al, 2005).

### **2.4.3 HBV disease burden in South Africa**

Previous studies in South Africa indicated HBsAg seroprevalence of 14.6% and 4.6% in adult males and females respectively at the kaNgwane district of Transvaal (Kew, 1996). Chronic HBV carriers in rural areas of former Transkei have higher prevalences of 15.5% as compared to urban areas of Durban and Soweto with 7.4% and 1.3% respectively (Kew, 1996).

This can be contributed by modes of transmission, especially horizontal transmission which is predominant in children but the precise modes are unclear (Kew, 1996). In the White and Indian communities, the carrier rate is estimated to be 0.2%, African-European descendants around 0.4 to 3% and Chinese shown to be 5-3% (Burnett et al, 2005).

## **2.5 HBV pathogenesis**

The HBV on its own is not cytopathic and can co-exist with the hepatocyte in the absence of host auto-immune responses (Hilleman, 2003). The cellular and humoral mechanism of cell-mediated immune system determines the course of the disease and the extent of liver injury (Ocana et al, 2005).

HBV infection is relatively short, with the appearance of neutralising antibodies and the development of immunity to reinfection (Huang et al, 2006). The host's immune

system attack is mediated by cellular response to small epitopes of HBV proteins, especially HBcAg presented on the surface of hepatocytes.

Cytotoxic T lymphocytes (CTL) recognise HBV peptide fragments presented by Human Leukocyte Antigen (HLA) class 1 molecule causing direct cell killing (Lee, 1997). In the process it causes liver injury by an immune response against the virus-infected liver cells although immunosuppression may enhance replication and cause induced cytotoxicity (Huang et al, 2006).

In acute infection, large amounts of interferon are produced which increases expression of HLA's which are associated with viral proteins (Foster et al, 1993). Several studies indicated that CTL response to HBV is polyclonal and multispecific in patients with acute infection but is relatively weak in chronic infected patients except during spontaneous disease flares or interferon therapy (Baumert et al, 2007).

A weak cell-mediated immune response generally represent a non-aggressive host immune responses and this may lead to persistent viral carrier state which can be clinically silent but can result in chronic active liver disease (Hilleman, 2003).

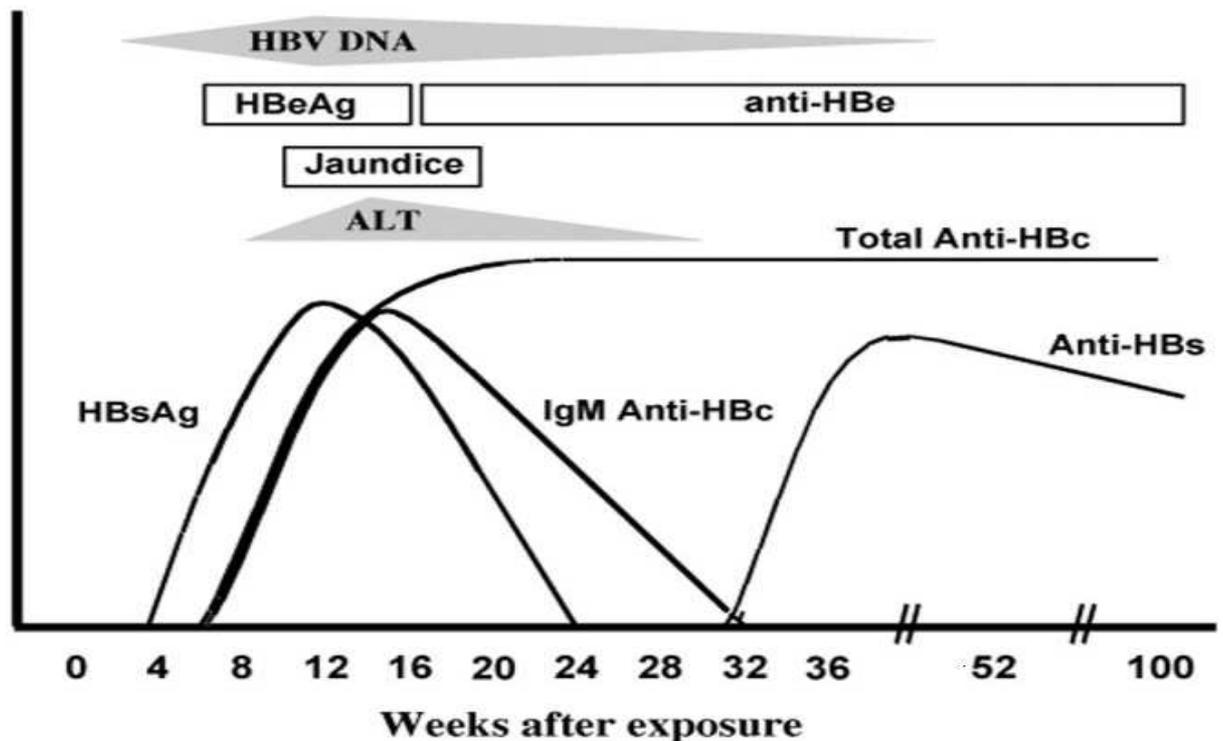
## **2.5.1 Disease profile of HBV infection**

### **2.5.1.1 Acute hepatitis B**

Acute hepatitis has an incubation period of about 2-6 weeks followed by raised levels of serum aminotransferases (Villeneuve, 2005). HBV DNA levels are elevated with a range of 200 million international units per milliliter (IU/ml) to 200 billion IU/ml ( $10^9$ - $10^{12}$  copies per ml) (Pungpapong, 2007). About 30% to 50% of infected adults present with an icteric illness after the incubation period and the outcome of the disease depend on age together with the immune status at the time of infection (Pungpapong, 2007).

It resolves with normalisation of liver function tests but HBsAg persists for few months until cleared from serum but low levels may be present throughout life and can be reactivated by liver transplant (Villeneuve, 2005) (Figure 2.3). There is evidence of strong host antibody response which may result in the complexing of the virus together with HBsAg leading to the loss of all three particles of surface

antigen from the bloodstream. A window period between the loss of HBsAg at the end of acute infection and the appearance of anti-HBs is observed and IgM anti-HBc is tested to reveal acute hepatitis (Valsakamis, 2007).



**Figure 2.3:** Acute hepatitis serology and clinical profile (Liang, 2009)

### 2.5.1.2 Chronic hepatitis B

Chronic hepatitis B is characterised by detection of HBsAg in serum for more than 6 months post-infection, which is accompanied by HBeAg and high HBV DNA consistent with anti-HBc (Figure 2.4) (Elgouhari et al, 2008). HBeAg is initially positive but most patients eventually lose it and seroconvert to anti-HBe (McMahon, 2009). Complications with chronic hepatitis B lead to cirrhosis, liver failure, HCC and extrahepatic diseases (i.e glomerulonephritis) (Villeneuve, 2005). After progression to chronic hepatitis B, the course of the disease is divided into four phases namely, immunotolerance, immuneclearance, low or non- replicating and reactivation (Michielson et al, 2005; Hoofman and Thio, 2007; Fattovich et al 2008).

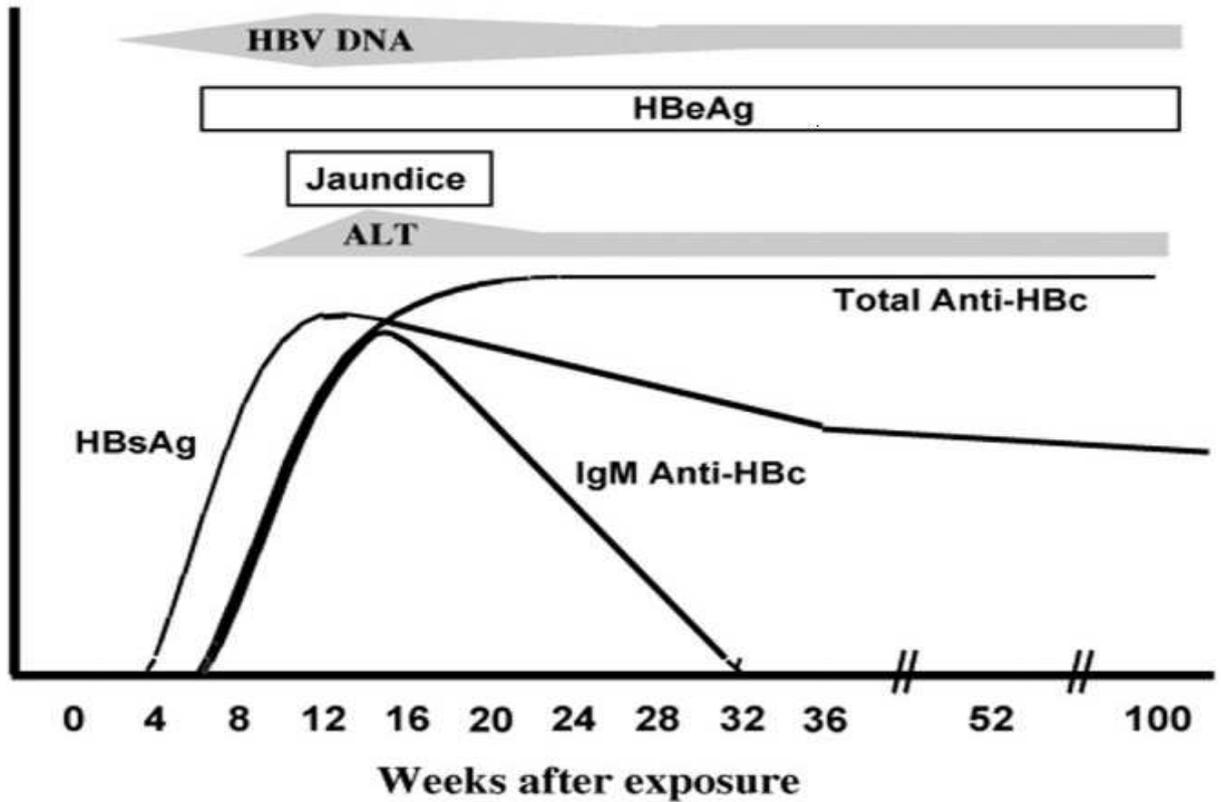


Figure 2.4: Serologic and clinical profile of chronic hepatitis (Liang, 2009)

### 2.5.1.3 Chronic hepatitis B phases

#### 2.5.1.3.1 Immune tolerance

It is characterised by the presence of HBeAg, high serum levels of HBV DNA, normal or minimally elevated serum alanine aminotransferase (ALT) and minimal histological activity with scant fibrosis (Fattovich et al, 2008; Michielson et al, 2005). HBeAg induces immunological tolerance with consistent high levels of viral replication (usually  $\geq 10^8$  copies /ml) and is therefore associated with chronic infection (Francois et al, 2001; Heathcote, 2007). It is primarily observed as a perinatally acquired infection and can last for decades (Valsakamis, 2007). It usually lasts for 2-4 weeks in healthy adults but often several decades in those infected neonatally or in early childhood (Wright, 2006; Pungpapong, 2007).

#### 2.5.1.3.2 Immune clearance

In this phase the HBV DNA levels fluctuate but progressively decrease with elevated ALT and hepatic necroinflammation (Fattovich et al, 2008). Hepatic

inflammation is believed to be associated with immune system and high levels of liver enzymes indicate the severity of liver damage (Hoofman and Thio, 2007). In this phase patients which were asymptomatic may start to show signs and symptoms which are suggestive of hepatitis flares (Lin and Kao, 2008). An important outcome is the seroconversion of HBeAg to anti-HBe but some patients may develop mutations in the precore /core region inhibiting the expression of HBeAg (Hoofman and Thio, 2007).

#### **2.5.1.3.3 Low replicating phase**

The third phase is characterised by HBeAg negativity and anti-HBe positivity, low or undetectable levels of HBV DNA, normal ALT levels and inactive liver histology with minimal fibrosis (Fattovich et al, 2008). It may lead to a decline in HBV replication where HBsAg becomes undetectable (Michielson et al, 2005). HBV carriers in this phase usually confer a favourable prognosis, but in Asian countries the carriers may develop complications even after HBsAg clearance (Lin and Kao, 2008). The immune system minimises active viral replication and thus lowering HBV DNA levels. It has been shown that 50% of patients at this stage clear HBeAg within 5 years of diagnosis and 70% within 10 years (Wright, 2006).

#### **2.5.1.3.4 Reactivation phase**

This phase may be contributed by viral replication due to reactivation of wild type virus reversion to HBeAg positivity or with replication-competent HBV variants that prevent HBeAg expression. HBeAg negativity with the presence of anti-HBe positivity, detectable serum HBV DNA levels ( $10^4$ - $10^8$  copies/ml), ALT elevation and moderate or severe necroinflammation with minimal amounts of liver fibrosis are evident in this phase (Fattovich et al, 2008; Lin and Kao, 2008). About 5% to 10% of inactive carriers undergo HBV reactivation with reversion to HBeAg positive state and in 20% to 30% inactive carriers will reactivate causing HBeAg negativity (Villeneuve, 2005).

#### **2.5.1.4 Fulminant hepatitis B**

Fulminant hepatitis B occurs in about 0.1% to 0.5% of patients acutely infected with HBV and there is lysis of infected hepatocytes by the immune system (Punpapong, 2007). It is marked by sudden appearance of fever, abdominal pain,

vomiting and jaundice accompanied by disorientation, confusion and coma (Liang, 2009). It usually develops after immunosuppression (i.e. chemotherapy), and initially it was believed to be caused by heightened immune response such in the case of HBV Infection with HDV or HCV co-infection (Lee, 1997).

## **2.6 Laboratory Diagnosis**

The routine diagnosis of acute and chronic hepatitis B and prognosis of patients on therapy is a common practice in many laboratories. Markers for HBV infection are normally detected by using antigen-antibody assays and their interpretation is crucial to reveal different clinical phases of the infection (Table 2.1). There are different assays for HBV diagnosis; namely (i) serological (ii) biochemical and (iii) molecular assays

**Table 2.1:** Interpretation of HBV serologic tests (adapted from Shepard et al, 2006).

<b>Tests</b>	<b>Results</b>	<b>Interpretation</b>
HBsAg Anti-HBc Anti-HBs	Negative Negative Negative	Susceptible
HBsAg Anti-HBc Anti-HBs	Negative Negative Positive	Immune due to vaccination
HBsAg Anti-HBc Anti-HBs	Negative Positive Positive	Immune due to natural infection
HBsAg Anti-HBc IgM Anti-HBc Anti-HBs	Positive Positive Positive Negative	Acutely infected
HBsAg Anti-HBc IgM Anti-HBc Anti-HBs	Positive Positive Negative Negative	Chronically infected
HBsAg Anti-HBc Anti-HBs	Negative Positive Negative	Four possible interpretations*

\*1. May be recovering from acute infection

2. May be distantly immune and test is not sensitive enough to detect low levels of anti-HBs in serum

3. May be susceptible with false positive anti-HBc

4. May be chronically infected and have undetectable levels of HBsAg present in the serum

### 2.6.1 Serological Markers

The serological markers are the hallmark of HBV infection. The markers are detected by serological assays which use capture techniques. A target molecule is immobilised on a substrate and allowed to bind to molecules in a solution causing antigen-antibody reaction which result in the formation of a chromogenic product from a colourless background (Greer and Alexander, 1995). Commercial assays usually detect small surface hepatitis B antigen (the 'a' determinant) of HBsAg which is present in acute and chronic hepatitis B. By definition, chronic hepatitis B is when HBsAg is detectable in blood for more than six months.

The previous less sensitive immunoassays resulted in prolonged window period and thus detected acute infection at a later period. Nowadays HBsAg can be detected within 6 to 10 weeks. Newly automated immunoassays are highly sensitive and specific resulting in early detection of acute hepatitis within 9 days (Hatzakis et al, 2006). In about 4 to 6 months, anti-HBs levels increase while HBsAg is decreased (Hatzakis et al, 2006). Commercial assays using antigen-antibody complexes with HBsAg are unreliable in populations with circulating variants due to mutations (Weber et al, 2005).

HBcAg appears as Dane particles in serum and is secreted from infected hepatocytes (Francois et al, 2001). Anti-HBc is mostly detected in chronic hepatitis together with anti-HBe and is unable to neutralize the virus infection (Huang et al, 2006). Most commercially available assays detect both IgG and IgM anti-HBc (Hatzakis et al, 2006). The IgM anti-HBc can also be detected and is important in the diagnosis of acute hepatitis B during the window period.

HBeAg is usually secreted from infected hepatocytes and its detection in serum indicates high levels of viral replication. A decline in detectable HBeAg in chronic hepatitis patients and its seroconversion to anti-HBe is regarded to be beneficial as it leads to low viral replication (Block et al 2007; Francois et al, 2001).

### **2.6.2 Biochemical tests**

Diagnosis of hepatic disease can also be based on monitoring levels of ALT and aspartate aminotransferase (AST) enzymes, which is a sensitive and cost effective assay (Kottilil et al, 2005). ALT level normalization in patients is a good indication of suppression of HBV replication especially before the onset of therapy (Kottilil et al, 2005). In some instances measurement of ALT may be insensitive especially in patients with advanced liver diseases and cirrhosis and thus other laboratory tests such as HBV assay may be useful in making an overall diagnosis (Feld et al, 2009).

### **2.6.3 Molecular assays**

The detection of HBV DNA is one of the criteria used for the diagnosis of hepatitis B virus disease. HBV DNA presence in serum or plasma together with other tests is important to establish diagnosis. The detection of HBV DNA was initially based on solution hybridization and measured in picograms per milliliter (pg/ml) but was less sensitive (Valsakamis, 2007). The latest molecular assays include polymerase chain reaction (PCR) and ligase chain reaction (LCR) which are highly sensitive and specific but are limited to research. The PCR assay has a detection limit of 10 copies of DNA per milliliter and relies on the use of specific primers that attach to each strand of target HBV DNA (Hatzakis et al, 2006). Quantitative PCR for HBV DNA detection is more valuable clinically as it is important for diagnosis and management of liver diseases (Spearman, 1994). The international standard for molecular assays is expressed in IU/ml and this allows direct comparison of HBV DNA levels with disease stages (Hatzakis et al, 2006).

## **2.7 Prevention, control and management of HBV**

### **2.7.1 Prevention and control of HBV**

Effective hepatitis B vaccines containing highly immunogenic viral surface antigens have been around since 1980 (Francois et al, 2001). Immunisation for HBV is done by administering immunoglobulins (passive immunisation), were standard solutions of immunoglobulins are prepared from carriers with high plasma levels of anti-HBs antibodies or vaccines (active immunisation) with purified and heat-activated highly immunogenic HBsAg obtained from plasma of infected persons (Francois et al, 2001).

Recombinant vaccines are produced by genetically manipulated yeast (*Saccharomyces cerevisiae*) with cloned S gene eliciting antibodies (Block et al, 2007). A single epitope within HBsAg called the 'a' determinant is directed against antibodies which produce seroconversion in 95% of recipients (Torresi, 2002). Implementation guidelines by World Health Organization (WHO) for vaccination are universally recognised as one of the most efficient ways of preventing HBV on a global scale (Kramvis et al, 2005). HBV vaccine is the first vaccine to prevent HCC by declining the burden of chronic infections worldwide (Liu and Kao, 2007).

HBV vaccine is given to all children under the age of one year with a dose of 1-5µg through intramuscular injection within 6, 10, 14 weeks of age with regular clinic visits (Kew, 1996). People who are elderly, obese, heavy smokers and immunocompromised including HIV may require higher doses of vaccine and more injections (i.e. at months 0, 1, 2 and 6) to achieve an adequate immune response (Zanetti et al, 2008). The most effective strategy in countries with high (HBsAg ≥8) or intermediate disease endemicity (HBsAg >2% but <8%) is to incorporate the vaccine into Expanded Program on Immunisation (EPI) schedules (Zanetti et al, 2008). In addition HBV vaccine was introduced in the South African EPI programme by 1995 and has shown to be highly effective leading to decreased levels or eradication of HBsAg carriage in vaccinated cohorts of children (Tsebe et al, 2001; Mphahlele et al, 2004).

### 2.7.2 Management of HBV infection

The overall goal of therapy is to suppress HBV DNA to the lowest possible level and ultimately prevent liver sequelae (Delaney and Esoda-Borroto, 2008). This leads to sustained viral replication suppression and HBsAg clearance from hepatocytes (Heathcote, 2007). There are seven drugs that have been licensed for the treatment of HBV infection namely pegylated interferon 2-alpha and nucleoside analogues such as lamivudine, adefovir, entecavir, telbivudine, tenofovir (Diestag, 2008).

Pegylated interferon alpha 2a (40kDa) has two chains of 2kDa polyethylene glycol conjugated to the lysine substrates of the interferon alpha-2a molecule. The drug is incorporated with a polyethylene moiety into the active product which enhances the pharmacokinetic properties of unmodified interferon alpha (Lai et al, 2006).

Lamivudine (LAM) is an orally available L-deoxycytidine analog and appears to be potent *in vitro* against both human immunodeficiency virus (HIV) and HBV (Delaney and Esoda-Borroto, 2008). It inhibits RNA-dependent DNA polymerase of HBV by irreversibly blocking reverse transcription and causing viral replication inhibition (Francois et al, 2001). LAM is the first small antiviral molecule to be approved for treatment of chronic hepatitis B and has the ability to suppress serum HBV DNA levels (Delane and Esoda-Borroto, 2008).

Adefovir dipovoxil (ADV) is an oral prodrug of acyclic phosphonate nucleotide adefovir and has *in vitro* activity against both HIV and HBV (Delaney and Esoda-Borroto, 2008). ADV inhibits priming of reverse transcription by preventing the incorporation of deoxyadenosine triphosphate into the viral primer and elongation of viral minus strand (Keeffe et al, 2008). Treatment with ADV results in suppression of HBV DNA and ALT normalisation (Delaney and Esoda-Borroto, 2008)

Entecavir (ETV) is a potent inhibitor of HBV DNA polymerase approved for treatment of chronic hepatitis B (Delaney and Esoda-Borroto, 2008; Xu and Chen, 2006). Telbivudine (LdT) has been shown to have potent and selective activity against HBV *in vitro* (Delaney and Esoda-Borroto, 2008). LdT targets the synthesis of HBV DNA positive strand (Keeffe et al, 2008).

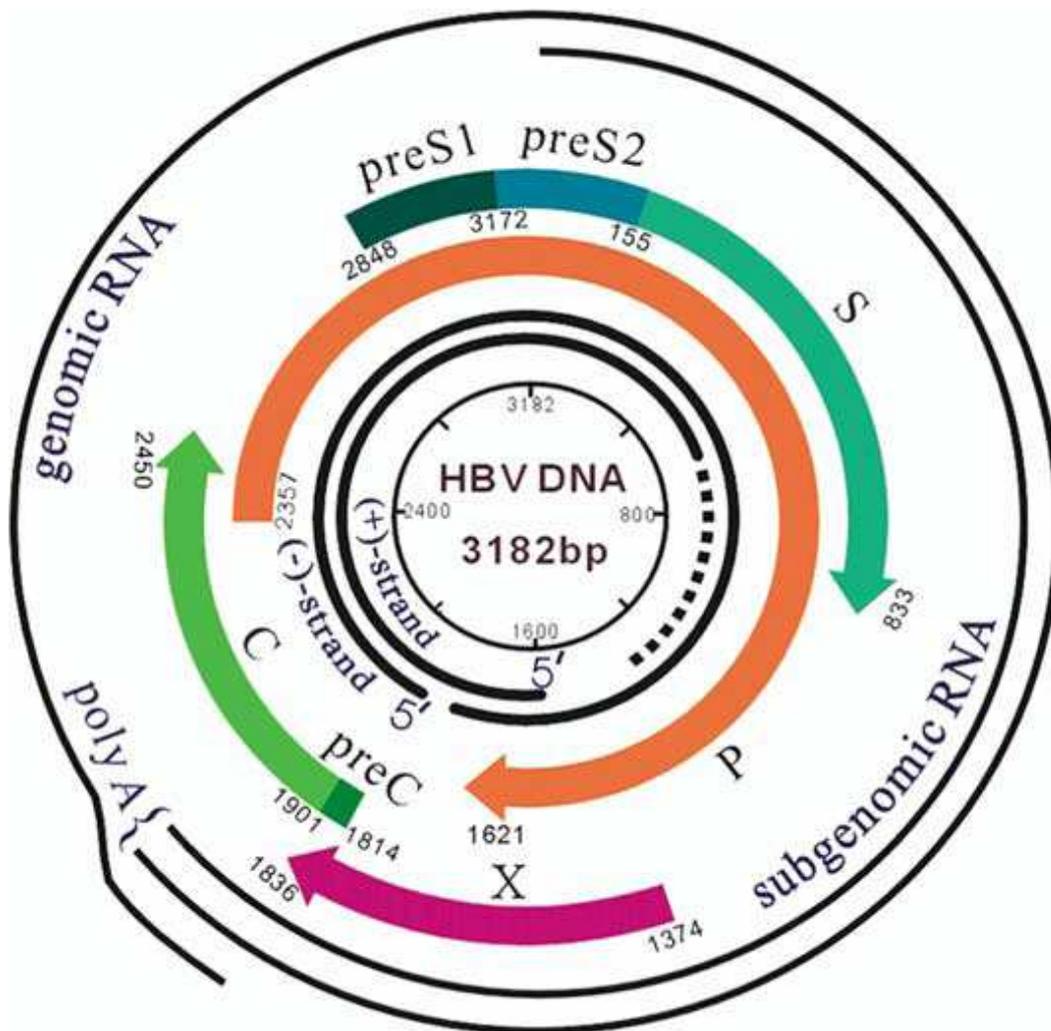
Telbivudine (Ldt) is the L-enantiomer of thymidine and has been shown to have potent and selective activity against HBV *in vitro* (Delaney and Esoda-Borroto, 2008). It targets the synthesis of positive strand HBV DNA which slows the emergence of resistance compared to lamivudine (Keeffe et al, 2008).

Tenofovir disoproxil fumarate (TDF) is an oral prodrug of phosphonate nucleotide tenofovir and has activity against both HIV and HBV. TDF has a high potency against HBV with about 76% of HBeAg positive patients achieving undetectable levels of HBV DNA after 48 weeks of treatment (Delaney and Esoda-Borroto, 2008).

## **2.8 Molecular biology of HBV**

### **2.8.1 The genome organization of HBV**

HBV has a compact genome with partially double stranded DNA of approximately 3200 bases in length and contains four ORFs coding for precore/core (preC/C), polymerase (P), surface (preS1/S2 and S regions) and X genes (Chen et al, 2005) (Figure 2.5). Each DNA molecule has a defined nick in the negative strand with an almost complete positive strand. Adjacent to the area of the nick and gap there are two direct repeat sequences (DR1 and DR2) that facilitate strand transfer during replication of the viral genome. Covalently attached to the 5' end of the negative strand is a viral polymerase protein which is positioned within DR1 (Huang et al, 2006).



**Finger 2.5:** HBV genome with a double stranded DNA with four open reading frames namely polymerase (P), surface antigens (preS1/preS2/S), precore (preC)/core(C) and X. (Adapted from Zhang et al, 2005).

## 2.8.2 HBV Replication

### 2.8.2.1 Entry and uncoating

HBV infection is caused by attachment of preS domain of the surface antigen to host cell membrane receptors such as carboxypeptidase D and triggers endocytosis (Liang, 2009). A low pH and proteolytic cleavage of envelope protein causes uncoating and viral capsids are released into the cytoplasm. In the cytoplasm, the capsid delivers its relaxed circular relaxed circular DNA (rcDNA) from the virus through nuclear pore complex (NPC) into the nucleus (Block et al, 2007). The schematic diagram of replication is shown in figure 2.6 below.

### 2.8.2.2 Replication of the HBV genome

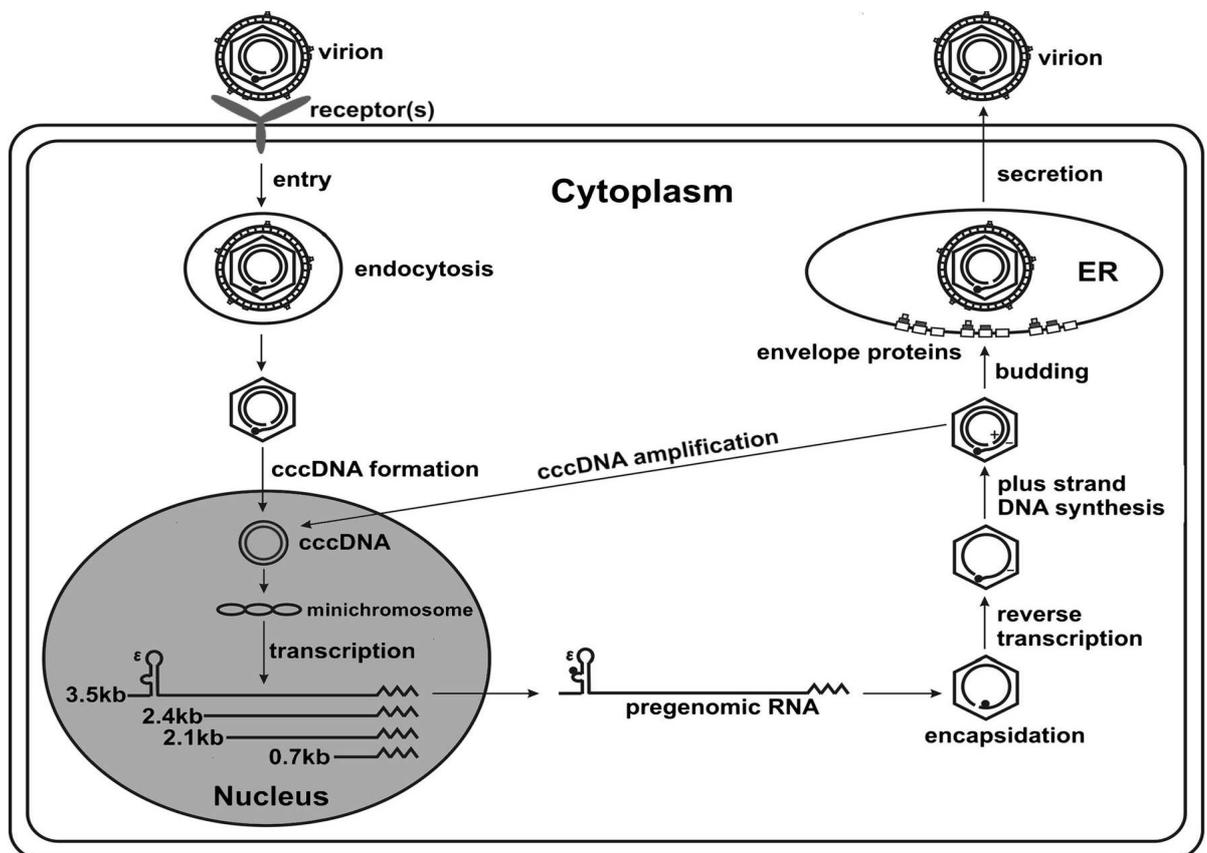
The genome contains diverse modifications (e.g. the polymerase-protein covalently linked to the 5' end of the negative DNA strand) that have to be removed before generation of covalently closed circular DNA (cccDNA) (Beck and Nassal, 2007). Once in the nucleus the partially double stranded is converted into cccDNA, which is the template for transcription by cellular polymerase II (Nassal, 1999; Glebe, 2007).

Four classes of genomic and subgenomic transcripts are synthesised by four promoters and two enhancers (Nassal, 1999). The classes of viral RNA are namely preC mRNA, pregenomic (pg) RNAs with a size of 3.5Kb, 2.4kb mRNA for large (L) envelope protein, 2.1 kb mRNA for middle (M) and major surface (S) proteins and 0.7kb for HBx (Block et al, 2007). These are transcribed by cellular RNA polymerase (the enzyme responsible for cellular mRNA synthesis) using cccDNA as the template. The pgRNAs contain 5'cap structures and are 3' terminally polyadenylated at a common site and serve as mRNA for viral gene products (Beck and Nassal, 2007).

DNA polymerase protein binds to the pgRNA at its 5' epsilon stem-loop structure to initiate nucleocapsid assembly and prime viral negative strand to extend for three nucleotides. Translocation of polymerase and covalently attached nascent DNA to 3' end of DR1 and negative strand DNA synthesis continues by copying pgRNA. At this stage the first DNA nucleotide is covalently linked to P protein extended into complete negative and positive strand DNA resulting into a new molecule of rcDNA which is reverse transcribed (Beck and Nassal, 2007). A crucial event for both processes is the binding of P protein to a 5' proximal stemloop structure on pgRNA called  $\epsilon$  (Nassal, 1999; Beck and Nassal, 2007).

### 2.8.2.3 Assembly and budding

Virions are assembled on and bud from endoplasmic reticulum (ER) membrane through the engulfing of rcDNA containing capsids surrounded by viral envelope proteins (Block et al, 2007). The envelope proteins interact with capsids in the ER to assemble into mature virions, which are then secreted into the extracellular environment (Liang, 2009).



**Figure 2.6:** Molecular steps of HBV replication from attachment to translocation of the virion DNA in the nucleus, formation of cccDNA into viral gene products in the nucleus (adapted from Block et al, 2007).

### 2.8.3 HBV subtypes and genotypes

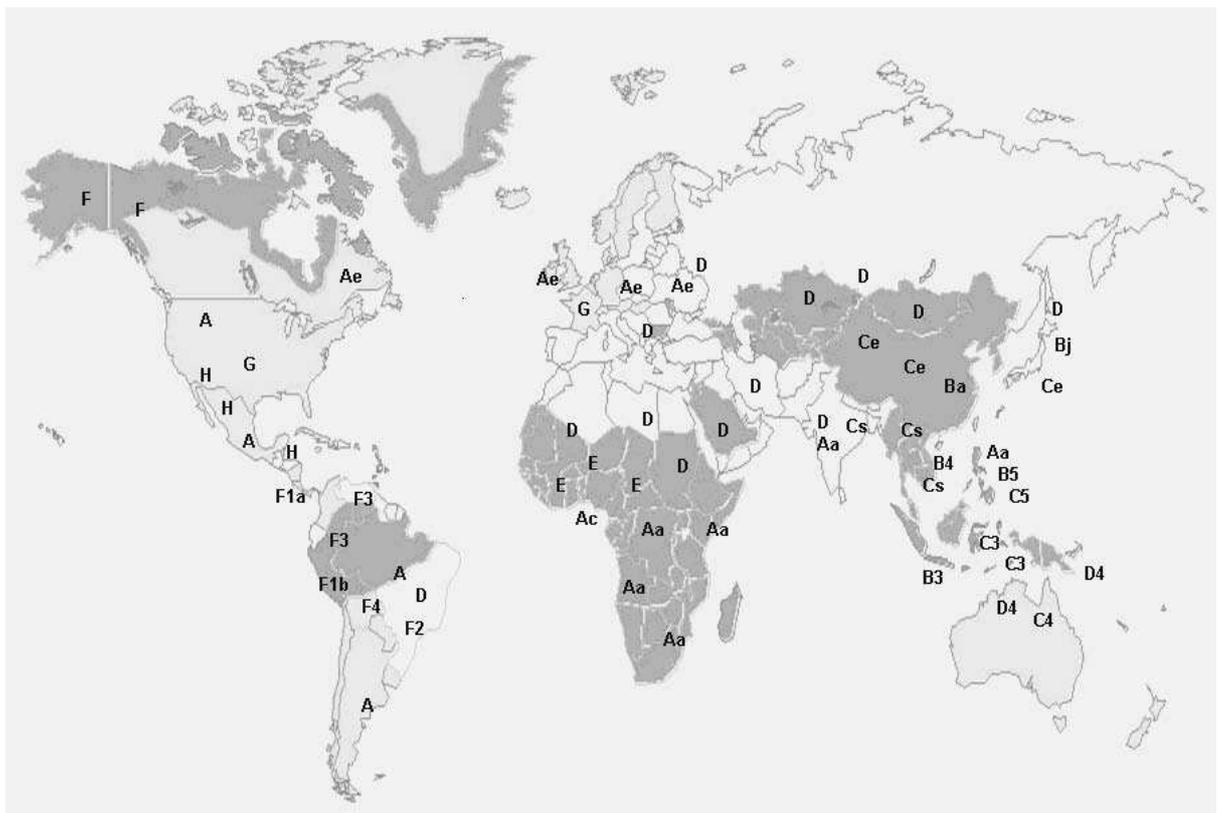
Subtype determinants “d”, “y”, “w” and “r” indicate variability within HBV and these reside within the main antigenic determinant “a” of the surface antigen. HBV strains could be characterised as belonging to either subtype adw, adr, ayw or ayr (Kidd-Ljunggren et al, 2002). The number of serological subtypes resulted to nine, namely: ayw1, ayw2, ayw3, ayw4, ayr, adw2, adwq, adr, adrq (Kramvis et al, 2005). This classification was based on a limited number of amino acid substitution and changes in HBsAg subtype due to nucleotide point mutations on the S gene caused changes in serological subtype determination (Weber, 2005).

A genotype is generally defined as the genetic constitution of an organism (Francois et al, 2001). In 1988 it was suggested that traditional subtypes could be complemented or replaced by classification of different HBV strains into genetics subgroups by comparing the full nucleotide sequences of HBV strains. Kidd-Ljunggren et al (2002) found that these clustered into four groups, A to D with 8% divergence between the groups. Genotype E, F, G and H were identified later and based on their sequences divergence of about 8% and further division of 4% based on S gene lead to subgenotypes (Kramvis et al, 2005; Fattovich et al, 2008). Epidemiological studies show that the genotypes of HBV have distinct geographical or ethnic distribution (Figure 2.7) (Weber, 2005).

HBV genotype A is mainly found in North western areas of Europe, North America, Africa and India (Kidd-Ljunggren et al, 2002; Lin and Kao, 2007). An analysis of genotype A strains found in South Africa demonstrated that a majority clustered into separate groups genotype A, called A1 and their main difference seem to lie mainly in the PreS2 region (Kidd-Ljunggren et al, 2002). Genotype B and C belongs to the indigenous populations of Southeast Asia, but their distribution is fairly intermixed with a tendency of genotype C strains being found in the northern and central Japan (Kidd-Ljunggren et al, 2002). Genotype B is subdivided into two subgenotypes, B1 found in indigenous Japan and B2 in Asia (Kramvis et al, 2005). Most retrospective or case control studies confirmed that the patients with genotype C progress to cirrhosis and HCC more than those with genotype B (Lin and Kao, 2008).

Genotype D is the most widely distributed genotype found worldwide with highest prevalence in belt stretching from southern Europe and North Africa to India, in West and South Africa and among intravenous drug users on all continents (Kidd-Ljunggren, 2002). Genotype E is similar to genotype D genetically and has been interpreted as a subset of genotype D when using X gene for phylogenetic analysis. It is found in West and South Africa and it lacks the 33 nucleotide deletion at the beginning of the PreS1 region which is common to all genotype D strains (Kidd-Ljunggren, 2002; Hoofman and Thio, 2007).

Genotype F is the most divergent genotype found in South and Central America and appears to share some structural features with genotype A strains. A divergence of about 14% is the characteristic feature and may represent the first split from human hepadnaviral ancestors (Kramvis et al, 2005). Genotype G has a high prevalence in the United States, while genotype H is restricted to Central and South America (Weber, 2005).



**Figure 2.7:** Worldwide distribution of HBV genotypes and subgenotypes (Datta, 2008).

## **2.8.4 Genetic variability of HBV**

The viral polymerase enzyme of HBV lacks the proofreading mechanism and uses an RNA intermediate during replication that lead to genetic variability due to random errors as a result of mutations occurring within the entire genome (Datta, 2008). The mutations in the preS/S, preC/C, P and X genes arise during acute or fulminant and chronic HBV infections (Francois et al, 2001; Lin et al, 2005). These are explained in detail in the next section.

### **2.8.4.1 Presurface and surface mutants**

The genomic surface region is divided into the S, Pre S and PreS2 regions with S gene coding for HBsAg made up of 226 amino acids (Francois et al, 2001). Together the Pre S1, PreS 2 and S regions code for large HBV surface protein (LHBs), medium size protein (MHBs) and small HBV surface protein (SHBs). LHBs and MHBs have truncated C terminal region with transcriptional activators (Gurtsevitch, 2008).

The S domain or HBsAg is the most important protein in virions together with spherical and filamentous particles (Francois et al, 2001; Yokosuka and Arai, 2006). The disulphides bonds are essential in stabilising its three dimensional structure. HBsAg is important and contains major neutralising epitopes used in commercial hepatitis B vaccines (Francois et al, 2001). Forming HBsAg is the major hydrophilic region (MHR) with the “a” determinant important for antigenicity. The MHR is divided into five sections, HBs1, HBs2, HBs4 and HBs5 (Francois et al, 2001). The middle part of HBsAg has 55 amino acids and depending on the genotype has additional 119 or 108 amino acids (Yokosuka and Arai, 2006; Bruss, 2007).

Mutations in the S gene are important since they affect the immunogenicity of HBsAg especially the ‘a’ determinant which might lead to failure of antibodies to neutralize HBV (Torresi, 2002). HBsAg mutations include G145R, D144A, P142S, Q129H, I/T126N/A, M133L associated with immune escape and viral persistence (Torresi, 2002; Huang et al, 2006). The G145R found in the HBs4 region is a stable mutation has been reported in several individuals who were infected while

having been vaccinated or received anti-HBs immunoglobulin (Karthigesu et al, 1994; Francois et al, 2001).

#### **2.8.4.2 Precore/core mutants**

Products of preC/C gene are HBcAg and HBeAg which share about 160 amino acids. HBeAg is important in the viral replication cycle and secreted in assembled form while its biological role remains obscure (Conway et al, 1998). It can be considered as a by product of the synthesis of HBcAg and is a polypeptide fragment formed after post translational processing (Francois et al, 2001). Both of the proteins contain 149-residue assembly domain but their termini differ with HBeAg having a 34-residue RNA-binding domain at its C terminus (Conway et al, 1998).

Mutations in the preC/C gene result in the loss of HBeAg and seroconversion to antibodies to HBeAg (anti-HBe) (Huang et al, 2006). The most commonly observed precore/core mutation are G1896A and A1762T/G1764A which causes a stop codon (TAG) at the distal region preventing expression of HBeAg and inhibit transcription of HBeAg respectively (Fattovich et al, 2008).

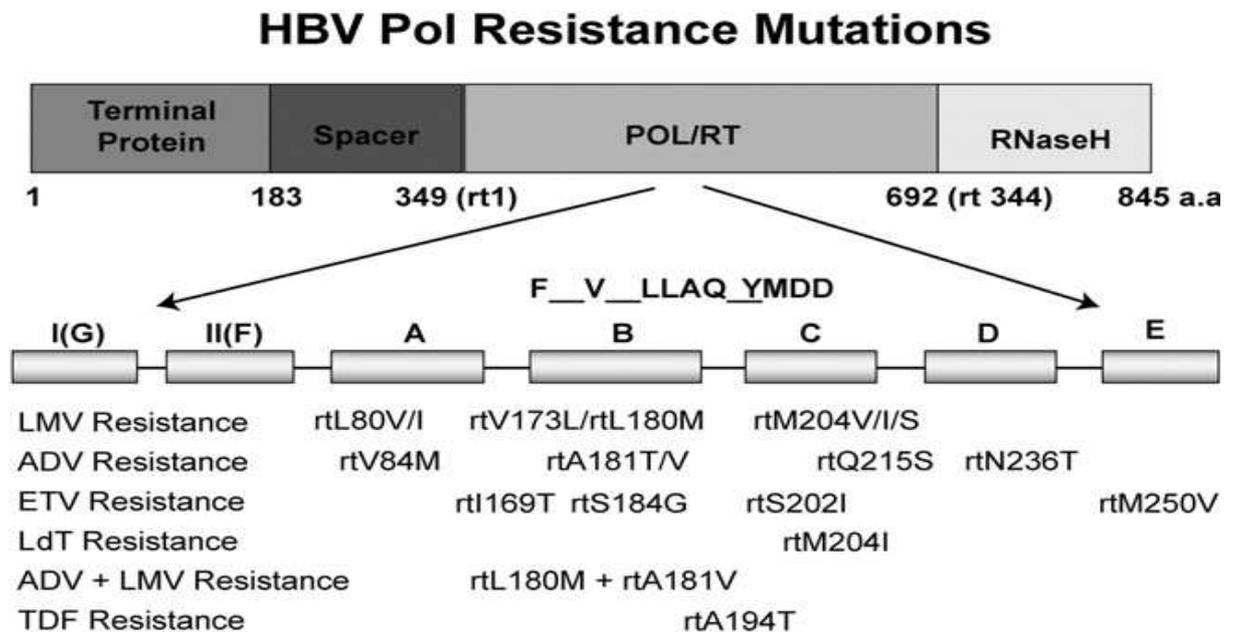
The 1896 mutation which may occur in association with 1899 also leads to HBeAg negativity and may moderately up regulate viral genome replication due to increased transcription of genome RNA (Tong et al, 2005). The T1762/A1764 mutations might contribute to HCC risk because is associated with changes in the overlapping X gene in genotypes B and C infection (Lin and Kao, 2008; Fattovich et al, 2008). The T1653 mutation in the core region is mainly isolated in patients with fulminant hepatitis is associated with the T1762/A1764 double mutation (Gunther et al, 1998).

#### **2.8.4.3 Polymerase mutants**

The polymerase region is covering nearly 80% of the hepadnaviral genome and has the coding capacity of 830 amino acids for HBV (Beck and Nassal, 2007). The P gene has 834-845 codons with four domains namely: N terminal domain, spacer, polymerase and C terminal domain (Yokosuka and Arai, 2006). HBV polymerase can be divided into a number of functional domains designated 'fingers', 'palm' and 'thumb' sub-domains by comparison with the reverse transcriptase of HIV (Torresi,

2002). The five domains (A-E), A and D domains are related to binding to dNTPs corresponding to the finger structure domain, B and E are related to the binding to the template or primer corresponding to the palm structure (Yokosuka, and Arai, 2006; Beck and Nassal, 2007).

Polymerase mutants usually result from viral persistence or resistance due to therapeutic effects of nucleoside analogues such as lamivudine, adefovir and entecavir (Huang et al, 2006, Francois et al, 2001). Lamivudine inhibits viral transcriptase activity and lead to high-level resistance which result in methionine to isoleucine (M204I) or methionine to valine (M204V) mutations (Keeffe, 2008). Mutations in the B and D domain of the viral polymerase of A181V and N236T respectively confer resistance to adefovir drug (Delaney and Esoda-Borroto, 2008). The entecavir drug resistance has been reported to be caused by mutations similar to lamivudine resistance such as I169I and S184G in domain B, S202I in domain C and M250V in domain E (Revil and Locarnini, 2005; Delaney and Esoda-Borroto, 2008). The mutations of the viral polymerase are summarised below.



**Figure 2.8:** Location of the major antiviral drug resistant mutations associated with lamivudine, telbivudine, adefovir, tenofovir and entecavir (adapted from Keeffe et al, 2008).

#### 2.8.4.4 X gene mutants

The X gene is located downstream to enhancer 1 (Enh1) and is partly overlapped by the P ORF at its N terminal and by the preC ORF at its C terminus (Murakami, 1999). The Enh 1 and X promoter complex forms the transcriptional regulation unit (Feitelson and Lee, 2007; Murakami, 1999).

Furthermore Enh 1 is responsive to other extracellular signals such as interleukins and within it, is the X responsive element (XRE) causing it to synergistically act with transactivating HBx. Overlapping the X gene is the direct repeat sequences at the end of the growing linear viral DNA strands and has the tendency to be integrated in hepatocytes (Feitelson and Duan, 1997)

HBx protein is about 154 amino acids long and its viral function is most frequently implicated in oncogenesis. Due to its uncertainty of its function during infection, it is called X protein (Block et al, 2007). The primary structure contains more than 60% of hydrophobic residues and its functional domain has been divided into two domains, namely the N and C terminals (Feitelson et al, 2005; Murakami, 1999). The N terminal domain is responsible for transforming activity, anchorage-independent proliferation activity and can overcome oncogene induced senescence (OIS) (Datta et al, 2008).

It has been shown that the N terminal domain and its variability are involved in contributing to differential pathogenic potentials of HBV genotypes/subtypes (Datta et al, 2008). The C terminal domain exercises a proapoptotic activity to counter balance the proliferative and transforming activities of N terminal domain (Datta et al, 2008).

Mutations in the C terminal of X gene may result in mutations of HBx together with *cis* elements (Murakami, 1999). This usually happens during integration of HBV into host DNA with truncated C terminal, which has been shown to lack antiproliferative or proapoptotic activities in tissue culture cells (Feitelson et al, 2005).

In one of the large study involving a significant number of HCC samples, an insert mutation (insert 204AGGCC) always accompanied a point mutations at position 260 (G-A) and 264 (G/G/T-A) (Chen et al, 2005). The point mutations may be of

diagnostic or clinical significance as they were isolated in about 40% and 70% of tumour tissues and serum samples respectively (Chen et al, 2005).

An eight nucleotide deletion at 3' end of the gene, and within the core promoter region (pos 1770-1770) together with 20 nucleotide deletion at nucleotide 1752-1772 have been described in HBsAg and HBeAg negative patients (Francois et al, 2001). The mutations A1762 plus G1764 causes K130M and V131I, that result in shortened HBx which lacks the C terminus (130-140 amino acids) (Revill and Locarnini, 2005). Some mutations detected in HBV infected patients have been reported to play a significant role in prognosis and among them include K130M and V131I. Mutations at position 1726-1730 showed correlation of decreased HBV DNA levels, indicating a relationship between the X gene and patients prognosis (Zhu et al, 2008).

## **2.9 Important roles of HBx in causing liver diseases**

A fundamental feature in pathogenesis of cancer is the deregulation of cell growth control in cellular proliferation and apoptosis, which are identified as being centrally important in the development of HCC (Feitelson et al, 2005). HBx is selectively over expressed in HCC suggesting that it might play important roles in carcinogenesis (Feitelson and Lee, 2007).

The most well characterised property of HBx is its ability to mediate transactivation of many cellular and viral promoters (Feitelson and Duan, 1997). Cytoplasmic HBx may alter gene expression by promoting the constitutive activation of signalling pathways such as nuclear factor kappa-B (NF-KB), mitogen activate protein kinase (MAPK) and phosphatidylinositol 3-kinase (PI3-K) (Feitelson et al, 2005). The nuclear localisation of the HBx may alter gene expression by binding to a number of transcription factors (Feitelson et al, 2005). Furthermore the transactivation process occurs through multiple *cis*-elements that have been documented to be responsive to HBx such as activation factor 1 and 2 (AP-1, AP-2), CCAAT/enhancer activity (C/EBP) and activating transcription factor (ATF1) (Murakami, 1999).

Persistent HBx expression leads to continual viral replication and activation the immune system therefore causing increased inflammatory responses (Feitelson and Duan, 1997). The HBx has also been shown to interact directly with and modulate the functions mediators of the inflammatory process, including interleukin (IL)-8, intercellular adhesion molecule 1 and major histocompatibility complex (MHC) class 1 (Wang et al, 2002; Block et al, 2007). The increased events of inflammation stimulate liver regeneration in on going chronic hepatitis B infection contributing to pathogenesis of HCC (Feitelson and Duan, 1997).

### **2.9.1 HBx and regulation of apoptosis**

Apoptosis is important for the elimination of redundant, damaged and virtually infected cells. It is regulated by proteins in three ways: as effectors and initiators, inducers and suppressors and also as intermediate proteins of apoptosis (Zhang et al, 2005). The tumour necrosis factor receptor (TNFR) and nerve growth factor receptor (NGFR) receive cell death signals from extracellular environment or internal sensors which triggers the cell to undergo apoptosis (Gupta, 2000; Kim et al, 2003).

The mitochondria play an important role in induction of apoptosis by a receptor independent mechanism (Gupta, 2000). HBx localizes to the alteration of the mitochondrial transmembrane potential and the region was mapped at amino acids 68-117 (Zhang et al, 2005). The voltage dependent anion channel (VDAC3) on the outer membrane is the site where HBx binds which lead to decreased mitochondrial membrane potential (Feitelson and Lee, 2007). VDAC3 can function as a component of mitochondrial permeability transition pore (MPTP) which is modulated by HBx to disrupt movement of calcium ( $Ca^{2+}$ ) into and out of mitochondria, therefore releasing it into the cell (Clippinger and Bouchard, 2008).

This result in the generation of reactive oxygen intermediates (ROI) and potentiates the state of oxidative stress which promotes activation of transcription factors such as NF-KB and the tyrosine phosphorylation and other transcription factors causing a hepatoprotective state (Feitelson et al, 2005). ROI may also contribute to DNA damage in patients with chronic hepatitis B infection and it is likely that integrated template together with mutations may also contribute to ROI

formation (Feitelson and Duan, 1997). In addition, It is hypothesised that oxidative stress and ROI may cause mutations in cancer related genes (Wang et al, 2002).

Depending on the severity of mitochondrial dysfunction HBx may either cause apoptotic cell death or cause oxidative stress due to downregulation of mitochondrial enzymes involved in oxidative phosphorylation or by the activation of NF-KB (Kim et al, 2007). Evidence of mitochondrial aggregation accompanied with mitochondrial dysfunction and/or cell death has been found in cells expressing HBx with increased levels of  $Ca^{2+}$  (Kim et al, 2007).

### **2.9.2 HBx and modulation of signalling pathways**

HBx activation occurs through modulation of a kinase signalling process (Murakami, 1999). Protein kinase C (Pkc) and other kinases such as PI3K, NF-KB, AP-1 activation by HBx suggest that it might act as an adaptor or kinase activator, enhancing the phosphorylation of its associated proteins (Zhang et al, 2005, Gurtsevitch, 2007, Block et al, 2007). A strong NF-KB activation might contribute to blocking the proapoptotic cellular response due to the HBx availability and in turn promote cell growth (Feitelson et al, 2005, Zhang et al, 2005).

The AP-1 and NF-KB are normally stimulated by cytokines and growth factor thus increases their chances of interacting with HBx (Feitelson and Duan, 1997). Furthermore, stimulation of NF-KB by HBx may promote the survival of both infected and mutated cells during chronic hepatitis B infection and thus contribute to HCC (Feitelson and Duan, 1997). The MAPK signal cascade shown to be upregulated by HBx and through RAF-MAPK pathway might accelerate cells into the S phase disrupting the cell cycle (Feitelson et al, 2005).

In stable cells expressing HBx, it has been shown that the tyrosine phosphorylation of signal transducer and activator of transcription (STAT) mainly STAT3 and STAT 5 were upregulated (Zhang et al, 2005). The presence of STAT3 and NF-KB motifs in most cellular genes that promote cell growth and proliferation which lead to DNA replication and repair may suggest that they form a link with HBx and ROI (Feitelson et al, 2005).

### **2.9.3 HBx and transcription modulation**

Many reports suggest that HBx can directly interact with various kinds of transcriptional activators primarily *in vitro* and less *in vivo*. HBx can bind to cAMP response element binding protein, activating transcription factor-2, TATA binding protein, ATF, cyclic AMP response element binding protein (CREB) and p53 (Feitelson and Duan, 1997). The interaction between HBx and RNA polymerase II (RPB5) is the first proof that a transcriptional modulator such as HBx may modulate transcription by directly interacting with RNA polymerase (Murakami, 1999).

### **2.9.4 HBx and cell cycle**

The P53 gene is expressed in cells with damaged DNA and cell division is prevented until DNA repair is finished (Feitelson and Duan, 1997). Association of p53 and HBx was demonstrated by co-immunoprecipitation using anti-p53 and anti-HBx in human liver tissues (Feitelson and Duan, 1997). The p53 binding region is located within its C terminal portion which forms part of the transactivating domain of HBx (Murakami, 1999).

### **2.9.5 HBx and DNA repair**

HBx disrupts the repair mechanisms of excision repair pathway which involves binding and inactivating an ultraviolet-light-induced DNA damage (Feitelson and Duan, 1997). Binding to p53 also suggest that DNA repair mechanisms in chronically infected hepatocytes is compromised (Feitelson and Lee, 2007). Finally binding and inactivating senescence factor, p55sen inactivates retinoblastoma (Rb) gene causing mutations due to faulty DNA repair (Feitelson and Lee, 2007).

### **2.9.6 HBx and cytokines**

The tumour necrosis factor-alpha (TNF- $\alpha$ ) and IL-6 are among major factors that stimulate liver growth in normal conditions and have been shown to be important in early signalling pathways that lead to liver regeneration. The cytokines can delay Fas-mediated apoptosis of liver cells during regeneration and HBx can hypersensitize liver cells to enter the TNF- $\alpha$  mediated apoptosis which deregulates liver cell growth (Kim and Seong, 2003).

### **2.9.7 Integration of X gene in hepatocytes**

Sequence analysis of host-virus junctions in chronic hepatitis B and HCC patients have shown that most viral DNA integrates within the direct repeat sequences located at each end of the viral genome with variable deletion in the 3' end of the X gene (Feitelson and Lee, 2007; Takada and Koike, 1990). The integrated X gene is often truncated and contains mutations that cause the HBx to have diminished co-transactivational activity (Wang et al, 2002).

## **CHAPTER 3**

### **3. MATERIALS AND METHODS**

#### **3.1 Ethical consideration**

The ethical approval to conduct the study was granted by the University of Limpopo, Medunsa campus Research and Ethics Committee (MREC/P/135/2008: PG). HBV sera from infected patients stored at -70°C were selected. The samples formed part of the project MP07/2005 that was previously approved by the Research, Ethics and Publication Committee. There was no need to draw blood from the patients. No information by which samples could be traced back to patients was released or published.

#### **3.2 Study design**

This study was an exploratory study involving stored sera of adult patients at Dr George Mukhari hospital, Pretoria.

#### **3.3 Study population**

The study population consisted of 20 stored sera (-70°C) from HBV DNA positive patients with different serological profiles (Table 2.1).

#### **3.4 Laboratory methods**

##### **3.4.1 HBV serological testing**

As HBV serological markers of 20 selected samples were already available (REPC project number MP07/2005) for HBsAg, anti-HBs, Anti-HBc, and HBeAg (Elecsys version; Roche diagnostics, Penzburg, Germany). There was no need to perform HBV serological tests and anti-HBe was not tested.

**Table 3.1:** Serological profile of samples

<b>SAMPLE</b>	<b>HBsAg</b>	<b>Anti-HBs</b>	<b>Anti-HBc</b>	<b>HBeAg</b>
HX055	Pos	Neg	Neg	Pos
HX149	Neg	Pos	Pos	Neg
HX394	Neg	Neg	Pos	Neg
HX464	Pos	Pos	Pos	Neg
HX558	Pos	Neg	Neg	Neg
HX964	Pos	Neg	Neg	Pos
HX001	Pos	Pos	Neg	Neg
HX056	Neg	Pos	Pos	Neg
HX098	Pos	Neg	Pos	Pos
HX722	Pos	Neg	Pos	Pos
HX061	Pos	Pos	Pos	Pos
HX064	Pos	Neg	Pos	Neg
HX019	Pos	Neg	Pos	ND
HX289	Pos	Pos	Neg	Neg
HX354	Pos	Neg	Neg	Neg
HX036	Neg	Pos	Pos	Neg
HX382	Pos	Pos	Pos	Neg
HX414	Neg	Pos	Pos	Neg
HX362	Neg	Pos	Pos	Neg
HX025	Pos	Neg	Pos	Neg

**Key:** Pos=positive; Neg=negative; ND=not done

**NB:** All samples were not tested for anti-HBe marker

### **3.4.2 HBV PCR DNA assay**

#### **3.4.2.1 HBV DNA extraction**

Viral nucleic acid was extracted from 200µl of serum using the High Pure Viral Nucleic Acid assay (Roche Diagnostics, Penzburg, Germany), following the manufacturers instructions. For each sample, 200µl of serum was aliquoted into 1.5ml microcentrifuge tubes, followed by 200µl of working solution (5µl poly [A] carrier RNA solution and 50µl proteinase K). The 1.5ml microcentrifuge tubes were thoroughly mixed by vortexing for 5 seconds and incubated at 72°C for 10 minutes on a heat block, in order to allow viral lysis to occur. After incubation the tubes were centrifuged to settle down condensation then 100µl of isopropanol was added to the tubes and briefly vortexed. The contents of the tubes were transferred to High Pure filter tubes each attached to collection tube. These tubes were then centrifuged at 8000rpm for 1 minute.

The filtrates in the collection tubes were discarded and the tubes were washed with 450µl wash buffer (10ml wash buffer [20mM NaCl and 2 mM Tris-HCl, pH 7.5] mixed with 20ml ethanol) and centrifuged for 1 minute at 8000rpm. After centrifugation, the filtrates within the collection tubes were discarded and filters were moved to new collection tubes. The wash buffer was added to the filters again and centrifuged for 1 minute at 8000rpm (During the washing steps, DNA would be bound to the matrix of the filter tubes while proteins and other non-nucleic material are washed away).

The pre-warmed elution buffer of 70µl was added onto the filter tubes after the wash step using clean nuclease free 1.5ml tubes as collection tubes. After 5 minutes, the filter tubes inside 1.5ml tubes were centrifuged for 2 minutes at 12000rpm and the filtrate was recovered. The recovered nucleic acid was either used immediately for the detection of HBV-DNA or stored at -20°C for later analysis. For quality control purposes, internal positive and negative controls were included; which involved known un-extracted and extracted positive controls, which were used to determine the precision of PCR and extraction, and negative controls to exclude contamination during processing of the extraction and PCR.

### 3.4.2.2 Amplification of HBV DNA

#### Principle of the PCR assay

PCR is a fast and reliable in vitro technique that uses thermostable enzyme to amplify small quantities of DNA into multiple fragments. The conserved regions on the target DNA serve as binding sites to oligonucleotide primers. The forward and reverse primers bind to 3' ends of the target DNA after denaturation into single strands. Extension of the primers is catalysed by thermostable *Taq* polymerase which results in duplication of DNA strands. The three steps namely, denaturation, annealing and extension are repeated using an automated thermal cycler in a number of cycles into multiple DNA strands.

#### PCR assay for HBV X gene

HBV X gene nested PCR amplification assay was done by using two sets of forward and reverse primers (Table 2.2). In each reaction, positive and negative controls were used for quality control. The primers were able to amplify 465 bp X gene in 20 samples with different serological profiles. The amplification was carried out using the established protocol with modification adapted from our laboratory (Mphahlele et al, 2006.). Briefly, 3µl of extracted DNA was amplified in a 25µl master mix containing the following: 2.5µl of 10X reaction buffer (Biolone, Luckenwalde, Germany), 0.75µl of 25mM MgCl<sub>2</sub> (Biolone, Luckenwalde, Germany), 0.2µl of 25mM dNTPs (Biolone, Luckenwalde, Germany), 0.5µl of 10µM of sense and anti-sense primers (Integrated DNA technologies, Coralville, USA) for first round, 0.2µl of *Taq* polymerase (Invitrogen, Life Technologies, California, USA) and 17.35µl of sterile water (SABAX, Johannesburg, South Africa). The second round PCR was performed using 3µl of the first round product with the second round primers and 17.35µl of sterile water was used to increase the volume to 25µl. The mixture was mixed by gentle flicking. The amplification of HBV DNA was carried out in 0.2 µl thin walled tubes (LasecSA, Cape Town, South Africa) in an automatic thermocycler (BioRad, Canada, USA). The amplification process included the following steps: Initial denaturation of samples at 94°C for 2 minutes, followed by 35 cycles each of the denaturation at 94°C for 10 seconds, annealing at 56°C for 30 seconds, extension at 72°C for 1 minutes and final extension at

72°C for 10 minutes followed by holding temperature of 4°C until the programme was manually terminated. Second round PCR conditions were similar to first round's conditions, except that an annealing temperature of 58 °C was used.

**Table 3.2:** Sequences of HBV X gene primers

Primer	Polarity	Target region	Position	Sequence (5'-3')	Reference
517F	Forward	Outer	1367-1389	ATCGTATCCATGGCTGC TAGGCT	Chen et al, 2005
517R	Reverse	Outer	1863-1883	CACAGCTTGGAGGCTTG AACA	Chen et al, 2005
473F	Forward	Inner	1375-1389	CATGGCTGCTAGGCTGT GCTG	Chen et al, 2005
473R	Reverse	Inner	1822-1847	GAGATGATTAGGCAGAG GTGAAAAAG	Chen et al, 2005

### **3.4.2.3 Detection of amplified HBV X gene PCR products**

Detection of the amplified PCR products was performed using two microliters ( $\mu\text{l}$ ) of the second round PCR product on 2% agarose gel electrophoresis. Briefly 2 $\mu\text{l}$  of a 100bp DNA ladder (Promega, Madison, USA) and 2 $\mu\text{l}$  of the PCR product mixed with 2 $\mu\text{l}$  of blue/orange 6x loading dye containing bromophenol blue, xylene cynol, and orange G (Promega, Madison, USA) were ran on the 2% agarose gel stained with ethidium bromide (1 $\mu\text{l}$ /100ml). The gel was run for 30 minutes in a 1x Tris-acetate-EDTA (TAE) buffer (Promega, Madison, USA) at 100 Voltage. The expected bands were read against 100bp DNA ladder (Promega, Madison, USA) using ultraviolet (UV) transillumination and corresponded to positive controls.

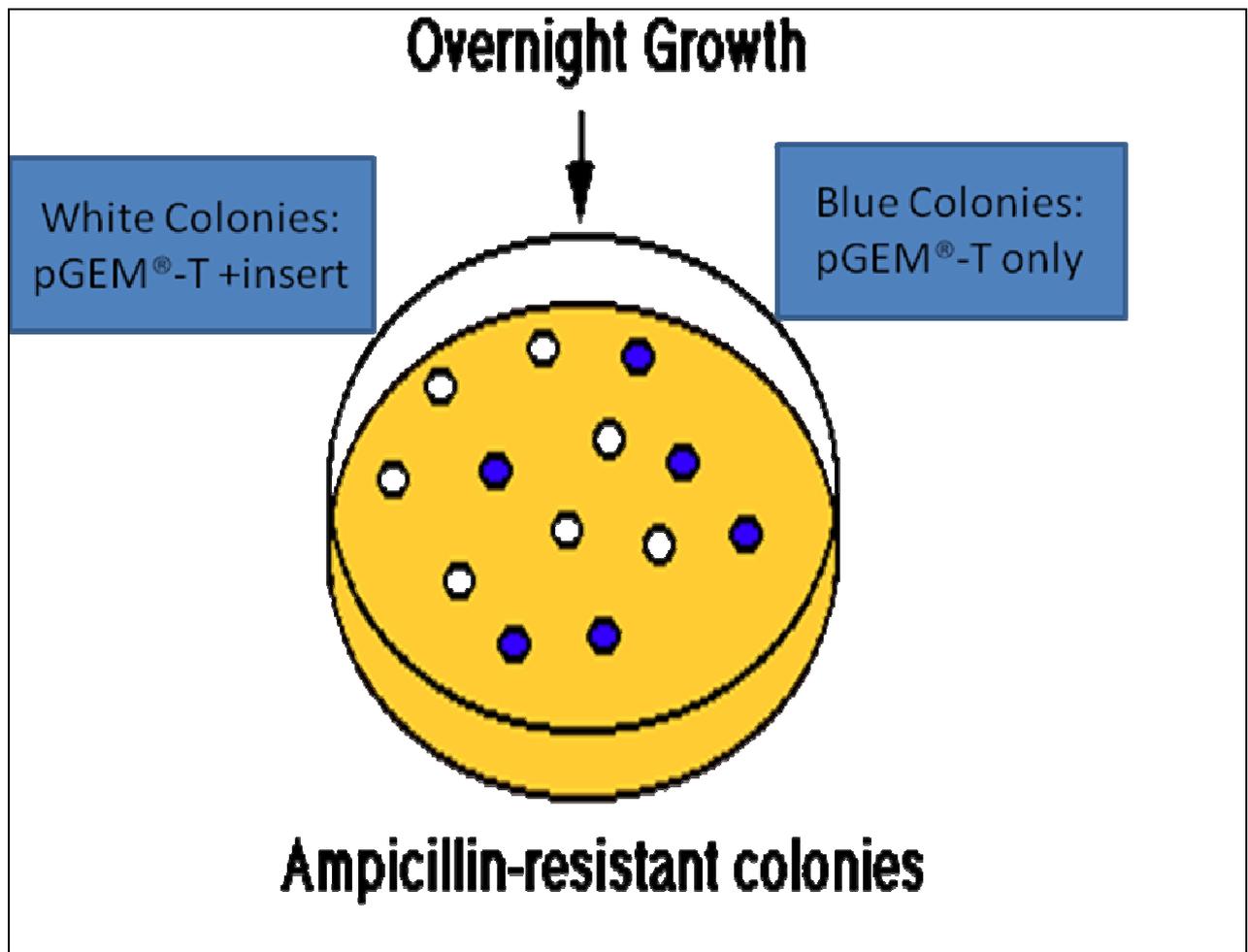
### **3.4.2.4 Avoiding PCR contamination**

To avoid false positive PCR results that might be due to contamination or carry-over of amplified DNA products, the following standard measures were adhered to during the procedures: A UV sterilizing PCR workstation was used which provided an automated process for eliminating contamination by utilizing the microbicidal properties of shortwave 254nm UV light. The UV light was turned on before and after PCR experiment for 30 minutes and also the working area was decontaminated with 70% ethanol in all experiments. Gloves and laboratory coats, strictly for PCR experiments were used and gloves were changed between the procedures. The mastermix was done in a clean room where specimens are not allowed. Nucleic acid extractions and PCR reactions were done in separate room inside air cabinets. In extraction and PCR procedures separate sets of filtered tips and pipetting devices were strictly used. The pipette tips and microcentrifuge tubes were autoclaved before each use. In addition both positive and negative controls were used in each procedure.

### 3.4.3 HBV DNA cloning

#### Principle of the assay

DNA cloning is a technique for reproducing multiple DNA fragments. This allows a copy of a DNA sequence to be produced in an unlimited amount. The HBV DNA was cloned in pGEM<sup>®</sup>-T Easy Vector System with 3'T overhangs at the insertion site (Promega, Madison, USA). Briefly, the target PCR product is ligated into the cut vector with the activity of T4 ligase. The ligation mix is incubated at 4°C overnight. Ligated products are then transformed into the E. coli competent cell by heat shock process and incubated with SOC medium. The transformants are cultured on Luria-Bertani (LB) plates containing Ampicilin, isopropyl β-D-thiogalactopyranoside (IPTG: an inducer of the lac promoter), 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal: a dye that produces a blue colour when hydrolyzed by β-galactosidase) agar plates at 37°C for overnight. The colonies with or without DNA insert are distinguished by white and blue colour respectively (Figure 3.1).

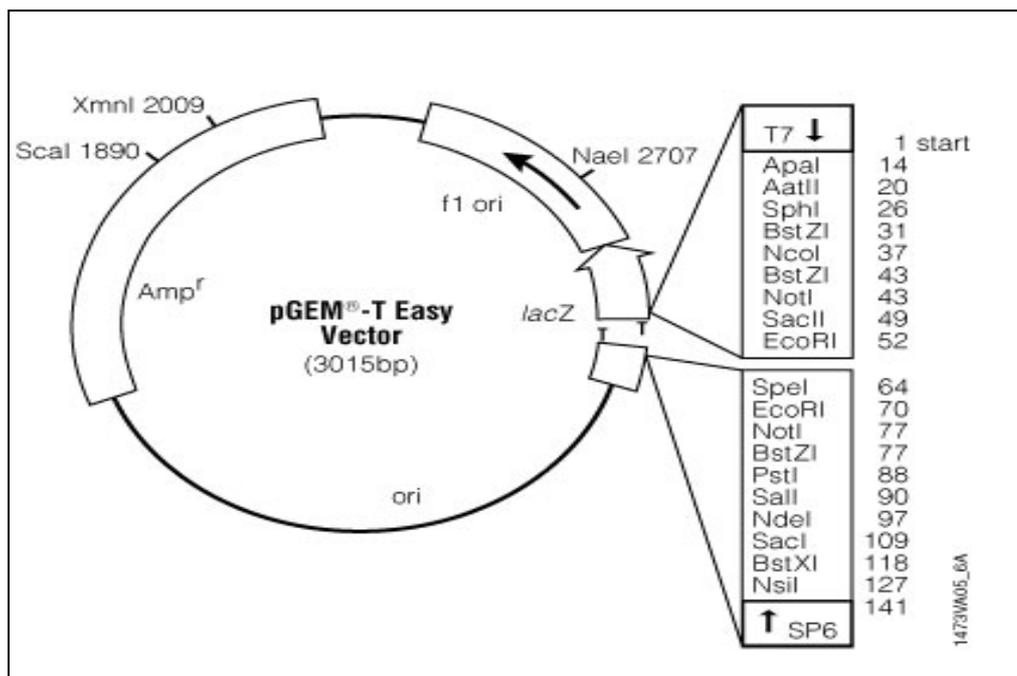


**Figure 3.1:** The substrate X-Gal turns blue if the gene is intact. White colonies in X-Gal imply the presence of recombinant DNA in the plasmid

## Description of pGEM®-T Easy Vector System

The pGEM®-T Easy Vector System is a convenient system for the cloning of PCR products. The vector is prepared by cutting with a blunt-ended restriction endonuclease and adding a 3' terminal thymidine to both ends (Figure 3.2). These single 3' T overhangs at the insertion site greatly improve the efficiency of ligation of a PCR product into the plasmids by preventing recircularization of the vector and provide a compatible overhang for PCR products with 5' A overhangs for efficient ligation with the vector.

In addition the vectors contain T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the coding region for  $\alpha$ -peptide of  $\beta$ -galactosidase. Insertional inactivation of the  $\alpha$ -peptide allows recombinant clones to be directly identified by colour screening on indicator plates containing X-Gal and IPTG (Figure 3.1). The pGEM®-T Easy Vector multiple cloning region is flanked by recognition sites for the restriction enzymes EcoRI, BstZI and NotI, thus providing three single-enzyme digestions for release of the insert (Figure 3.2). The pGEM®-T Vector cloning region is flanked by recognition sites for the enzyme BstZI. Finally the vectors contain the origin of replication of the filamentous phage f1 for the preparation of single-stranded DNA (ssDNA).



**Figure 3.2:** pGEM®-T Easy Vector circle map (Adapted from [www.promega.com](http://www.promega.com))

#### **3.4.3.1 Ligation of HBV DNA into pGEM<sup>®</sup>-T Easy Vector**

The pGEM<sup>®</sup>-T Easy vector and control insert DNA tubes were briefly centrifuged and the 2X rapid ligation buffer was vigorously vortexed before the ligation process. The 0.5ml tubes with low DNA binding affinity were used. The standard ligation reaction in each tube contained: 5µl of ligation buffer, 1 µl of enzyme ligase, 1µl of the vector, 2µl of the DNA template and 1µl of sterilized distilled water which added up to 10µl of the ligation mix. The other two tubes namely positive control (2µl control insert DNA, 5µl of ligation buffer, 1µl of enzyme ligase, 1µl of the vector, 1µl of sterilized distilled water) and background control (5µl of ligation buffer, 1µl of enzyme ligase, 1µl of the vector, 3µl of sterilized distilled water) tubes were used as quality control. The reactions were mixed by pipetting and incubated overnight at 4°C to achieve a maximum number of transformants.

#### **3.4.3.2 Transformation using pGEM<sup>®</sup>-T Easy vector ligation reactions**

For each ligation reaction, two LB/ampicilin/IPTG/X-Gal plates were prepared for each ligation reaction plus other two plates for determining the transformation efficiency. The plates were equilibrated at room temperature prior plating. The tubes containing the ligation reaction were centrifuged to collect the contents at the bottom of the tubes. Two microlitres of each ligation reaction was added to a sterile 1.5ml microcentrifuge tubes on ice. Another tube with 0.1ng of uncut plasmid was set up for the determination of transformation efficiency of the competent cells. Tubes of frozen JM109 High Efficiency Competent Cells were removed from storage (-70°C) and placed in an ice bath for 5 minutes until thawed. The cells were mixed gently by flicking the tubes. Fifty microlitres of cells were transferred carefully into a tube containing the ligation reaction. To determine the transformation efficiency 100µl of cells were used. The tubes were gently flicked to mix the ligation reaction with the cells and placed on ice for 20 minutes. The cells were heat-shocked for 45-50 seconds in a water bath at exactly 42°C and immediately returned to ice for 2 minutes. A room temperature SOC medium of 950µl was added to tubes containing cells transformed with ligation reactions and 900µl of SOC was added to the tube containing cells transformed with uncut plasmid.

The tubes were incubated for 1 hour 30 minutes at 37°C with shaking at approximately 150rpm. After incubation 100µl of each transformation culture was plated onto duplicate LB/ampicillin/IPTG/X-Gal plates. The plates were incubated overnight (16-24 hours) at 37°C to view the results. The plates were stored at 4°C after an overnight incubation until further use.

#### **3.4.3.3 Isolation of recombinant plasmid HBV DNA**

HBV plasmid DNA was isolated and purified prior to sequencing analysis. Several progeny colonies that were grown initially were selected for purification. The Pure Yield™ Plasmid Miniprep System was used to isolate high-quality plasmid DNA (Promega, Madison, USA). The system uses silica–membrane column to purify plasmid DNA rapidly.

A volume of 600µl of bacterial culture grown in LB medium was transferred to a 1.5ml microcentrifuge tube. A solution of cell lysis buffer of 100µl was added and the tube was inverted six times for mixing. Then 350µl of cold neutralization solution was added and also mixed by inverting. The mixture was centrifuged at 12000rpm for 3 minutes. After centrifugation, 900µl of the supernatant was transferred into a Pure Yield™ Minicolumn without disturbing the cell debris pellet. The minicolumn was placed into the collection tubes and centrifuged at 12000rpm for 15 seconds. The flowthrough was discarded and thereafter the minicolumn was placed back into the same collection tube.

A 200µl of Endotoxin Removal Wash (ERB) was added to the minicolumn and centrifuged at 12000rpm for 15 seconds in order to remove protein, RNA, endotoxin contaminants. After centrifugation 400µl of column wash solution was added to the minicolumn and centrifuged at 12000rpm for 30 seconds. The mixture in the minicolumn was transferred to a clean 1.5ml microcentrifuge tube, and then 30µl of elution buffer was added directly onto the minicolumn matrix. And the contents were allowed to stand for 1 minute at room temperature and thereafter centrifuged at 12000rpm for 15 seconds to purify the plasmid DNA. The 1.5µl microcentrifuge tube was capped after centrifugation and the purified DNA was stored at -20°C.

A PCR reaction for the recovery of cloned fragments using second round PCR as explained in section 3.4.2.1 and purified plasmid DNA was used as a template. After the PCR reaction the amplified products were detected as explained in section 3.4.2.2 to determine the presence and size of PCR product insert.

#### **3.4.4 HBV DNA nucleotide sequencing**

HBV X gene (465 bp) positive DNA samples and clones were sequenced at Inqaba Biotechnological Industries (Pty) Ltd, Pretoria, South Africa using SpectruMedix SCE 2410 genetic analysis system (SpectruMedix LLC, State College, PA). The sequencing assay is based on Sanger-Coulson method of sequencing that uses PCR to incorporate chain terminating fluorescent dideoxynucleotides triphosphates (ddNTPs) (Sanger et al, 1977). The PCR reaction mix contains oligonucleotides of different wavelengths each terminating in labelled ddNTPs (ddATP, ddCTP, ddTTP, ddGTP) with each nucleotide designated in the target DNA sequence, T7 DNA polymerase (sequenase, lacks 3→5 exonuclease activity) and inner primers (Table 2.2) which anneals to the single-stranded DNA providing a starting point for complementary strand synthesis. The reactions were terminated by the addition of a stop solution (95% formalide, 20mM EDTA, 0.05% xylene cyanol FF) and denatured by heating.

This is followed by high resolution electrophoretic separation of the oligonucleotides on a polyacrimide gel and laser detection of the characteristic wavelength of fluorescence emitted by each ddNTP which allows computer analysis of the DNA sequence. SpectruMedix SCE 2410 genetic analysis system generated sequenced data through automated sequencing. All sequences were analysed in both forward and reverse directions.

### 3.4.5 Nucleotide analysis

Sequenced data was downloaded from finch server from Inqaba Biotech, Pretoria. Analysis of nucleotide sequences was done using Chromas pro version 1.45 (School of Health Science, Griffith University, Australia). The Chromas software is a free windows application which displays and prints chromatogram files from ABI automated DNA sequencers. Chromas allow base editing and exporting of the sequence to a file or the clipboard. Nucleotide and sequence comparisons and alignments were done using CLUSTAL W (Thompson JD, 1994) and translated into proteins using Bioedit (ClustalW 1.8 software program) (Hall, 1999).

Multiple sequence alignment is the arrangement of several amino acid sequences with postulated gaps so that similar residue is juxtaposed. A positive score is attached to identities, conservative or non-conservative substitutions (the score amplitude measuring the similarity) and a penalty to gaps; an ideal program would maximise the total score, taking account of all possible alignments and allowing for any length gap at any position. Multialign creates a multiple sequence alignment from a group of related sequences using progressive pairwise alignments. Mutations were identified by comparing X gene sequences with corresponding sequences of eight HBV genotype isolates obtained from the National Institutes of health (NIH) and DNA databank of Japan (DDBJ). The eight genotype sequences namely AY233278 (Genotype A), AF121251 (Genotype B), AB014371 (Genotype C), FJ349218 (Genotype D), AP007262 (Genotype E), AB166850 (Genotype F), AP007264 (Genotype G) and EU498228 (Genotype H) were chosen because they were all full genome sequences and have been used in several genotyping studies.

### **3.4.6 HBV phylogeny**

All nucleotide sequences from HBV strains with X gene sequences were aligned using the ClustalW 1.8 software program with a reference panel of reported sequences available in the HBV sequence databases provided by NIH and DDBJ. The 29 reference sequences were full genome and have been used in several genotyping studies. The sequences were of Genotype A (AY738140, EU594395, AB453980, EU85955, AB194951, GQ161837, EU366129, AY934771, AY233278, AF297625, AY233274, AY233287, AY934773, FJ692561), Genotype B (AF121251), Genotype C (DQ683578, EU670263, AB014371), Genotype D (AF121239, EU594415, FJ349218, FJ904420, AY233296, AP007262), Genotype E (DQ060822, GQ161759), Genotype G (AP007264) and Genotype H (EU498228). Pairwise evolutionary distance matrices for the HBV nucleotide sequences were computed using the p-distance algorithm of the MEGA software package (version 4.0, 2007; Pennsylvania State University, University Park, PA) (Tamura et al, 2007). Phylogenetic analysis was performed with the MEGA software package using the p-distance algorithm for distance determination and the neighbor-joining method for tree drawing. The reliability of phylogenetic classification was evaluated by a 1000-replication bootstrap test using the same software package.

## CHAPTER 4

### 4. RESULTS

#### 4.1 Overview of the results

The overall aim of the study was to characterise HBV X gene and the overlapping basal core promoter region at a molecular level in order to determine their genetic variability at nucleotide and amino acid level. The study consisted of 20 HBV positive samples with different serological profiles of which were tested for markers of HBsAg, anti-HBs, anti-HBc and HBeAg (Table 3.1) to establish HBV infection. The HBV X gene has about 465 nucleotides and the direct sequencing of 20 samples revealed spontaneous mutations and inserts in the full length HBV X gene, particularly a six nucleotide insert (GCATGG) between nucleotide 1611 to 1618 of the reading frame detected in 5 samples.

Subsequent cloning of six samples with three clones per sample was done.. The samples were HX055 and HX098 both positive for HBsAg and HBeAg markers, HX025 and HX289 which were HBsAg positive and HBeAg negative. The other two samples (HX019, HX036) were chosen due to their six nucleotide insert (GCATGG) for full characterisation of the X gene. The sequenced clones confirmed a six nucleotide insert (GCATGG) and mutations associated with inhibition of HBeAg expression initially found with direct sequencing from PCR products.

The six nucleotide and three amino acid insert caused a frameshift in the reading frame of nucleotide and amino acid alignment respectively. Mutations within the X gene were observed within the N (1-50 amino acids) and C (51-156) terminals. Important substitutions shown to be relevant to HBx mediated pathogenesis were detected in the two terminals. The serine to alanine substitution at position 31 (S31A), proline to serine at position 38 (P38S), alanine to proline at position 42 (A42P), cystine to arginine at position 69 (C69R), phenylalanine to leucine at position 73 (F73L), histidine to tyrosine at

position 94 (H94Y), proline to serine at position 101 (P101S), methionine to threonine at position 118 (K118T), aspartate to asparagine at position (D119N), isoleucine to threonine or asparagine at position 127 (I127T/N), lysine to methionine at 130 (K130M) and valine to isoleucine at position 131 (V131I) which are involved in biological function of HBx. Other variants unknown functions detected in the study were alanine to glycine at position 2 (A2G), 3 (A3G) and 4 (A4G); cystine to tryptophan at position 6 (C6W), proline to serine at position 42 (P42S) and valine to leucine at position 116 (V116L).

Mutations within basal core promoter region (1756-1776) involved in HBeAg expression were observed. Mainly 1753 C or A (M), 1762T, 1764A were detected in most sequences as compared to 1766T which was found in one isolate.

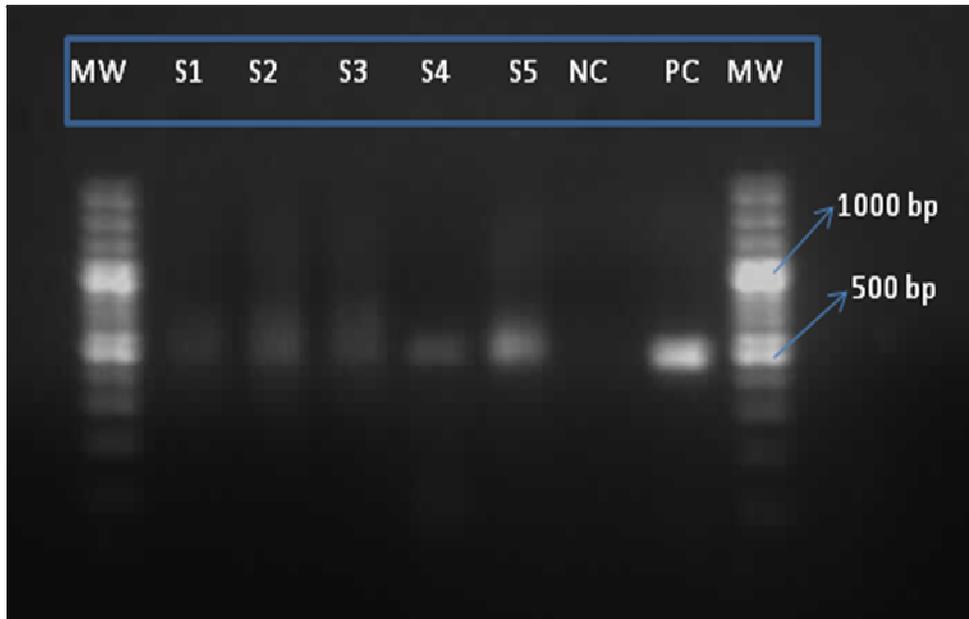
The phylogenetic analysis using HBV X gene sequences with full genome reference sequences of various genotypes from different countries resulted in three clusters: D, C and A. The 38 study sequences were compared with 29 full length references with different genotypes. Five samples and 11 clones clustered with genotype D, two samples and four clones clustered with genotype C and lastly 13 samples and three clones clustered with genotype A.

## **4.2 Optimisation of PCR assay for amplification of HBV X gene**

The study followed a protocol of nested PCR by Chen et al, 2005 which employs two set of primers to successfully amplify a PCR product of 465bp. The annealing temperature of 56°C for first round and 58°C for second round were used during the amplification process. Initially, a volume of 5µl of the extract was used in the first round and 3µl in the second round which produced smears. The volume of the extract was decreased to 3µl with the same amount used in second round and thus yielded results. Furthermore the sensitivity of the assay was found to depend on the annealing temperature. Additionally, the primer concentration of 10mM was effective when using the above mentioned PCR conditions. The study was able to amplify 20 samples with different serological profiles.

### 4.2.1 Detection of PCR products

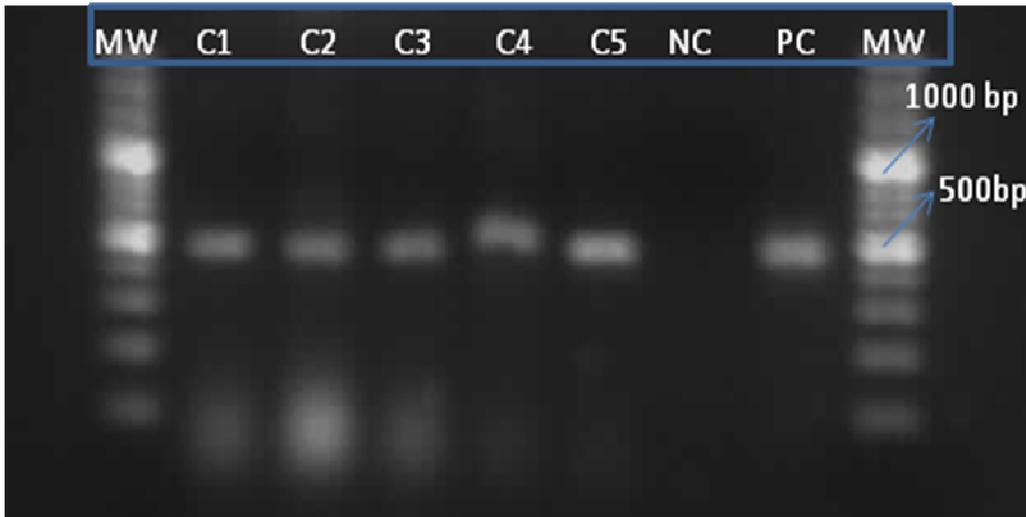
A full length HBV X gene (465 bp) was amplified and the products were run on 2% ethidium bromide stained agarose gel with a positive and negative control and read against 100bp molecular weight marker (Figure 4.1).



**Figure 4.1:** A gel electrophoretic diagram showing samples amplified with X gene primers  
**Key:** MW=molecular weight (100bp) marker, PC=positive control; NC=negative control; S1-S3=X gene weak positive samples; S4-S5=X gene strong positive samples

### 4.3 Identification of HBV positive clones

A total of six HBV positive samples were cloned; two were positive for both HBsAg and HBeAg, the other two negative for HBeAg and positive for HBsAg, and the last two with a six nucleotide insert of GCATGG in their reading frame. The HBV PCR products were successfully cloned with pGEM<sup>®</sup>-T Easy vector and those with X gene insert were identified with a blue colour on LB agar. And from the LB agar, three colonies for each sample were picked for isolation. Following isolation of the recombinant plasmids, a PCR reaction using inner primers targeting HBV X gene was run. After the reaction 18 full lengths of HBV X gene product (465 bp) were detected on 2% ethidium bromide stained agarose gel and the bands were identified against a 100 bp molecular marker (Figure 4.2).

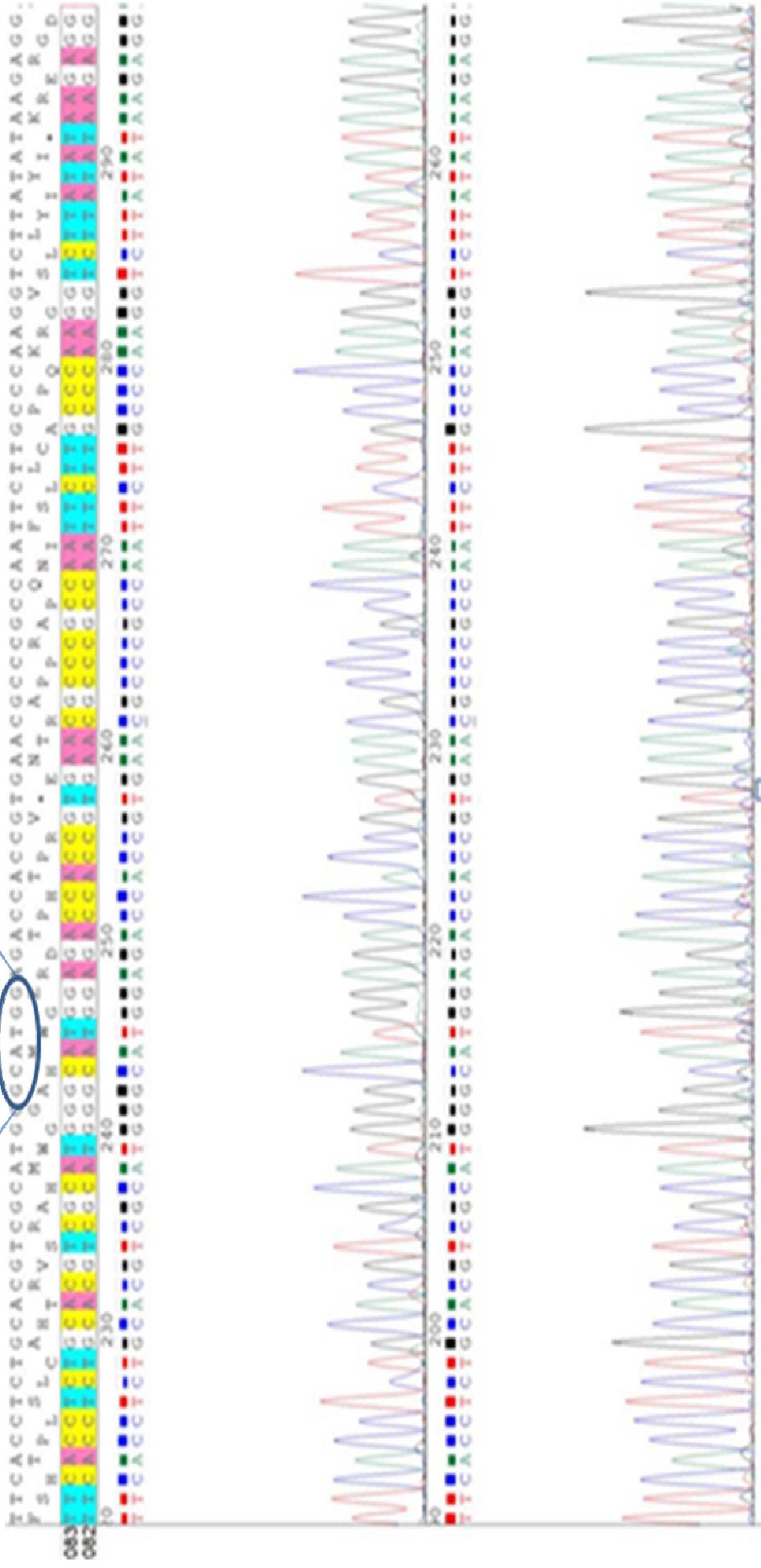


**Figure 4.2:** A gel electrophoretic diagram showing clones amplified with X gene primers  
**Key:** MW=molecular weight (100bp) marker, PC=positive control; NC=negative control; C1-C5=clones

#### 4.4 DNA sequencing analysis

Automated direct sequencing of 20 direct PCR products and their 18 clones revealed a 465 bp HBV X gene product. A six nucleotide insert (GCATGG) between nucleotide 1611 and 1618 was identified in five of the direct PCR products and 11 clones (Figure 4.3).

1611 ← → 1618



**Figure 4.3:** Chromatogram of the X gene visualised by Chromas pro Version 1.45 showing a six nucleotide insert 5'-GCATGG-3' between nucleotide 1611 and 1618.

#### 4.5 Nucleotide and amino acid sequence analysis of HBV X gene

A total of 38 sequences (20 direct PCR products and 18 clones) were compared with eight reference strains from GenBank to determine their sequence variations. The nucleotides were numbered from the start of non-conserved EcoR1 site in the HBV genome according to Okamoto et al, 1994. The nucleotides and amino acids substitutions were identified within regulatory elements and functional domains of the gene which are recognised in patients with severe liver diseases. The N terminal or regulatory domain had mutations of C to G at nucleotide 1378 causing A2G substitution (1 sample), C to G at nucleotide 1381 causing A3G substitution (2 samples and 4 clones), a A to G at nucleotide 1383 causing a A4G substitution(1 sample), C to G at nucleotide 1394 causing C6W substitution (1 sample), T to G at nucleotide causing S31A substitution which was detected in 15 sequences (12 samples and 3 clones) and T to C at nucleotide 1485 causing P38S substitution in 32 sequences (18 samples and 14 clones). Finally A42P substitution caused by a C to T mutation at nucleotide 1497 appearing in 32 sequences (18 samples and 14 clones) and a mutation G to C at nucleotide 1497 appearing in 6 sequences (2 samples and 4 clones).

The C terminal or functional domain displayed most variations. A frame shift due to a six nucleotide insertion at position 1611 to 1618 caused a three amino acid insertion (GME) at position 79 to 83 of the amino acid reading frame in 16 sequences (5 samples and 11 clones) therefore shifting the sequences in the C terminal. The variations include a T to C at position 1578 causing a C69R substitution in 4 sequences (1 sample and 3 clones), T to C at nucleotide 1590 causing F73L substitution in 3 sequences (3 clones), C to T at 1653 nucleotide change causing H94Y substitution in 16 sequences (5 samples and 11 clones), a C to T change at nucleotide 1674 causing P101S substitution appearing in 22 sequences (7 samples and 15 clones), C to T change at nucleotide 1690 causing T106I substitution in 16 sequences (5 samples and 11 clones), C to T change at nucleotide 1719 causing V116L substitution in 16 sequences (5 samples and 11 clones), A to C change at nucleotide 1726 causing K118T and

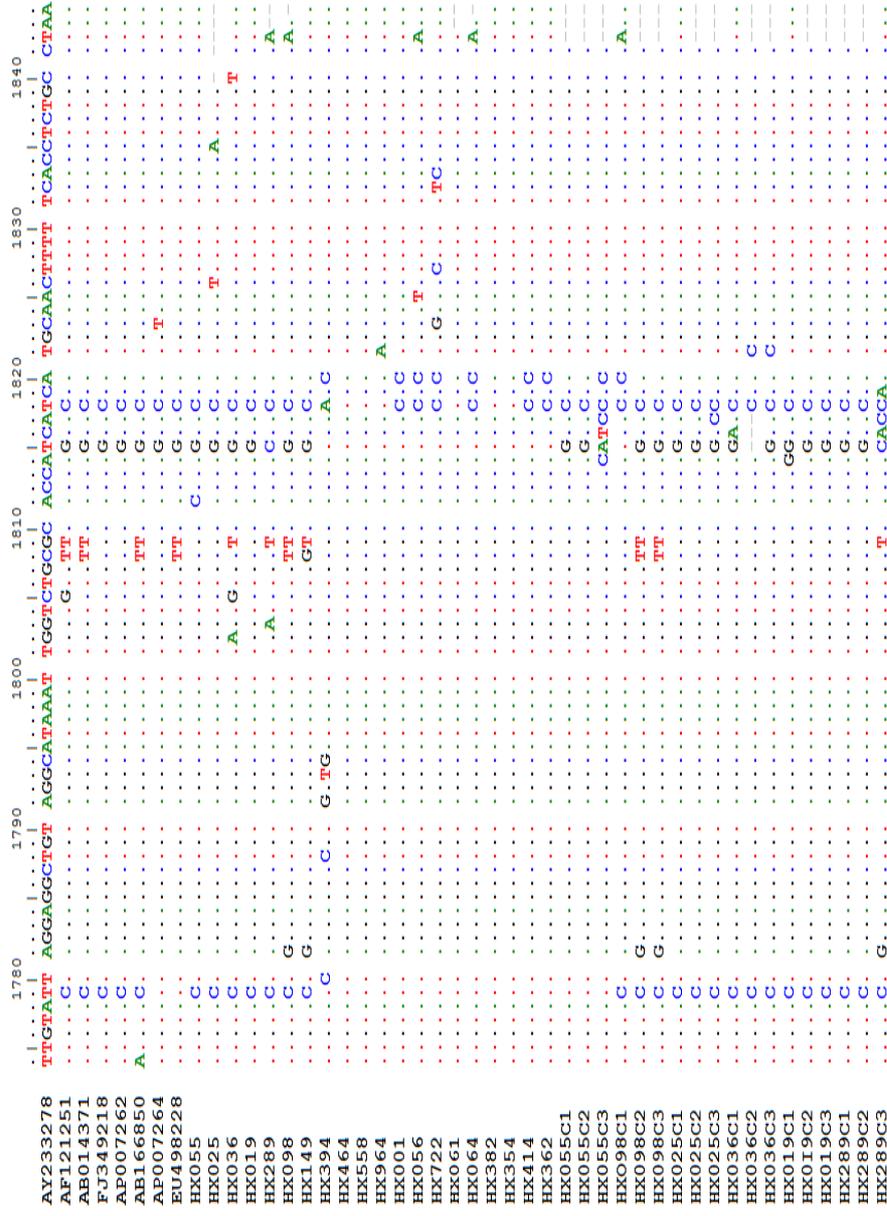
G to A change at nucleotide 1728 causing D119N substitutions in 16 sequences (5 samples and 11 clones). Other mutations within basal core promoter region include a T to C or A nucleotide change at position 1753 causing I127T/N substitution appearing in 32 sequences (14 samples and 18 clones). And lastly a dual mutation of A to T at 1762 and G to A at 1764 nucleotide positions resulting in K130M substitution in 32 sequences (14 samples and 18 clones) and V131I substitution in 33 sequences (15 samples and 18 clones) respectively. Furthermore sequence HX382 had a mutation of C to T at nucleotide 1766 which did not cause an amino acid change within the basal core promoter region. Alignment of nucleotides and amino acids is presented in figures 4.4 and 4.5 respectively. Important mutations are summarised in Tables 4.1 and 4.2.

	1380	1390	1400	1410	1420	1430	1440	1450	1460	1470	
AY233278	ATGGCTG	CTAGGTTGTA	CTGCCAACTG	GATTTCTTCGA	GGGACCTCCT	TGCTTACCT	CCCCTGGCG	CTGAAATCCCG	CGGACGACCC	CTCGCGAGGT	CGC
AF121251	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
AB014371	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
FJ349218	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
AF007262	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
AB166850	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
AP007264	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
EU498228	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX055	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX025	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX036	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX019	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX289	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX098	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX149	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX394	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX464	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX558	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX964	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX001	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX056	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX722	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX061	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX064	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX382	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX354	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX414	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX362	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX055C1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX055C2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX055C3	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX098C1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX098C2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX098C3	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX025C1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX025C2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX025C3	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX036C1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX036C2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX036C3	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX019C1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX019C2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX019C3	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX289C1	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX289C2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
HX289C3	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....

AY233278	1480	1490	1500	1510	1520	1530	1540	1550	1560	1570
AF121251	TTGGGAC	TCTATCGTCC	CCTACTCGGT	CTGGCGTACC	GTCGGACCAC	GTTCTTTACG	CGGTCTCCCC	GGTCTGGCC	GTCTGTGGCC	TACACATCGC
AB014371	.....G	.....C	.....C	.....G	.....A	.....T	.....A	.....A	.....T	.....C
FJ349218	.....C	.....T	.....G	.....T	.....G	.....T	.....A	.....A	.....T	.....T
AP007262	.....GG	.....T	.....T	.....T	.....G	.....T	.....A	.....A	.....T	.....T
AB166850	.....G	.....C	.....T	.....T	.....G	.....T	.....A	.....A	.....T	.....T
AP007264	.....G	.....G	.....C	.....T	.....TG	.....T	.....A	.....A	.....T	.....T
EU498228	.....G	.....GC	.....C	.....T	.....G	.....T	.....A	.....A	.....C	.....T
HX055	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX025	.....T	.....C	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX036	.....C	.....C	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX019	.....C	.....C	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX289	.....C	.....C	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX098	.....C	.....C	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX149	.....T	.....T	.....T	.....T	.....C	.....T	.....A	.....A	.....T	.....T
HX394	.....T	.....T	.....T	.....T	.....C	.....T	.....A	.....A	.....T	.....T
HX464	.....G	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....C
HX558	.....G	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....C
HX964	.....G	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX001	.....T	.....T	.....T	.....T	.....C	.....T	.....A	.....A	.....T	.....T
HX056	.....T	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX722	.....G	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX061	.....T	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX064	.....G	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX382	.....G	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX354	.....G	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX414	.....T	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX362	.....T	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX055C1	.....T	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX055C2	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX055C3	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX098C1	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX098C2	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX098C3	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX025C1	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX025C2	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX025C3	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX036C1	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX036C2	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX036C3	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX019C1	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX019C2	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX019C3	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX289C1	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX289C2	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T
HX289C3	.....C	.....T	.....T	.....T	.....A	.....T	.....A	.....A	.....T	.....T



	1680	1690	1700	1710	1720	1730	1740	1750	1760	1770	
AY233278	GGACTCC	CAGCAATGC	AACGACCGAC	CTTGAGCCCT	ACTTCANAGA	CTGTGTGTTT	AAGACTGGG	AGGAGTTGGG	CGAGGAGATT	AGGTTAAGG	TCT
AF121251	TT	TT	TT	A	A	CT	G		T	T	A
AB014371	T	T	T	A	A	T	G		TC		A
FJ349218	T	T	T	A	A	T			C	A	T
AP007262	T	T	T	A	A	T			CC	A	T
AB166850	TT	G	CG	T	TGG	A	C	A	AA		T
AP007264	GT	TT	TT	A	GG	G	G	AAA	CA	T	A
EU498228	TT	CC	CCG	TGG	A	A	A	A	T	G	A
HX055	T	T	T	A	A	T	A	A	A	A	T
HX025	T	T	T	A	A	T	A	A	A	A	T
HX036	T	T	T	A	A	T	A	A	A	A	T
HX019	T	T	T	A	A	T	A	A	A	A	T
HX289	T	T	T	A	A	T	A	A	A	A	T
HX098	T	T	T	A	A	T	A	A	A	A	T
HX149	T	G		A	A	A	A	A	CC	A	T
HX394	T	G		A	A	A	A	A	CC	A	T
HX464				A	A	A	A	ACA	C	A	T
HX558	A										
HX964											
HX001											
HX056											
HX722											
HX061											
HX064											
HX382											
HX354											
HX414											
HX362											
HX055C1											
HX055C2											
HX055C3											
HX098C1											
HX098C2											
HX098C3											
HX025C1											
HX025C2											
HX025C3											
HX036C1											
HX036C2											
HX036C3											
HX019C1											
HX019C2											
HX019C3											
HX289C1											
HX289C2											
HX289C3											



**Figure 4.4:** Alignment of HBV X gene nucleotides compared with corresponding representative references of genotypes A to H: AY233278, AF121251, AB014371, FJ349218, AP007262, AB166850, AP007264 and EU498228. The samples and clones are HX055-HX362 and HX055C1-HX289C3 respectively. The dots represent similar nucleotides. A six nucleotide insertion, GCATTGG between nucleotide 1611 and 1618 shifted the numbering of the reading frame and the highlighted region represents the basal core promoter region

	10	20	30	40	50	60	70	80	90	100
AY233278	MAARLYCQLD	SSRDVLRRLP	VGAESRGRPL	ARSLGTLSSP	TPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
AF121251	MAARLCCQLD	PARDVLCCLP	VGAESRGRPL	PGPLGALPPA	SPPIVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMX	-XTTVNAHWN	LPKVLHKRTL
AB014371	MAARVCCQLD	PARDVLCCLP	VGAESRGRPV	SGPFGTLSPS	SASAVPANHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQV	LPKVLYKRTL
FJ349218	MAARLCCQLD	PARDVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMX	-XTTVNARQF	LPKVLHKRTL
AP007262	MAARLCCQLD	PARDVLCCLP	VSAESCGRPL	SGPLGGLSSP	SPSAVPADHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMX	-XTTVNAHQI	LPKVLHKRTL
AB166850	MAARMCCQLD	PARDVLCCLP	VGAESRGRPL	SGPLGALPPS	SASAVPADHG	SHLSLRGLPV	CSFSSAGPCA	LRFTSARRMX	-XTTVNAPWS	LPIVLEHRTI
AP007264	MAARLCCQLD	PSRDVLCCLP	VSAESSGRPL	PGPFGALSSP	SPSAVPADHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARYMX	-XTAMNTSHH	LPRQLYKRTL
EU498228	MAARLCCQLD	PARDVLCCLP	VGAESCGRPL	SWSLALPPS	SPPVVPADDG	SHLSLRGLPA	CAFSSAGPCA	LRFTSARRMX	-XTTVNAPWN	LPTTLHKRTL
HX055	MAARLCCQLD	PARDVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX025	MAARLCCQLD	PARDVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX036	MAARLCCQLD	PARDVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX019	MAARLCCQLD	PARGVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLPLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX289	MAARLCCQLD	PARGVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX098	MAARLCCQLD	PARDVLCCLP	VSAESRGRPL	PGPFGTLSPS	SSSAVPTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMX	-XTTVNARQV	LPKVLHKRTL
HX149	MGARLHCHLD	RAGDVLCCLP	VSAESRGRPL	PGPFGTLSPS	SSFAVPPDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMX	-XTTVNARQV	LPKVLHKRTL
HX394	MAARLCCQLD	SSRDVLCCLP	VSAESRGRPL	ARSLGTLSSP	SPSAVPSDHG	AHLSLRGLPV	CVFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX464	MAARLCCQLD	SSRDVLCCLP	VGAESRGRPP	ARPLGALSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX558	MAARLCCQLD	SSRDVLCCLP	VGAESRGRPL	ARPLGALSSP	SPSAVPSNHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX964	MAARLCCQLD	SSRDVLCCLP	VGAESRGRPL	ARPLGALSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX001	MAARLCCQLD	SSRDVLCCLP	VSAESRGRPL	ARSLGTLSSP	SPPAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX056	MAGRLCCQLD	SSRDVLCCLP	VSAESRGRPL	ARSLGTLSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX722	MAARLCCQLD	SSRDVLCCLP	VSAESRGRPL	ARSLGTLSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX061	MAARLCCQLD	SSRDVLCCLP	VGAESRGRPL	ARPLGALSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX064	MAGRLCCQLD	SSRDVLCCLP	VSAESRGRPL	ARSLGTLSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX382	MAARLCCQLD	SSRDVLCCLP	VGAESRGRPL	TGPLGALSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX354	MAARLCCQLD	SSRDVLCCLP	VGAESRGRPL	ARPLGALSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX414	MAARLCCQLD	SSRDVLCCLP	VSAESRGRPL	ARSLGTLSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX362	MAARLCCQLD	SSRDVLCCLP	VSAESRGRPL	ARSLGTLSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX055C1	MAARLCCQLD	PARDVLCCLP	VSAESRGRPL	ARSLGTLSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX055C2	MAARLCCQLD	SSRDVLCCLP	VSAESRGRPL	ARSLGTLSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMX	-XTTVNAHQF	LPKVLHKRTL
HX055C3	MAGGLCCQLD	SSRDVLCCLP	VSAESRGRPL	ARSLGTLSSP	SPSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARCMX	-XTTVNAHQI	LPKVLHKRTL
HX098C1	MAARLCCQLD	PARDVLCCLP	VSAESRGRPL	PGPFGTLSPS	SSSAVPSDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMX	-XTTVNAHQV	LPKVLHKRTL
HX098C2	MAARLCCQLD	PARDVLCCLP	VSAESRGRPL	PGPFGTLSPS	SSSAVPTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMX	-XTTVNARQV	LPKVLHKRTL
HX098C3	MAARLCCQLD	PARDVLCCLP	VSAESRGRPL	PGPFGTLSPS	SSSAVPTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMX	-XTTVNARQV	LPKVLHKRTL
HX025C1	MAARLCCQLD	PARDVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX025C2	MAARLCCQLD	PARDVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX025C3	MAARLCCQLD	PARDVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX036C1	MAARLCCQLD	PARDVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX036C2	MAARLCCQLD	PARGVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLPLRGLPV	CAFSSAGPCA	LRFTSARRMG	MEATVDNAHQF	LPKVLYKRTL
HX036C3	MAARLCCQLD	PARGVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLPLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX019C1	MAGRLCCQLD	PARGVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLPLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX019C2	MAGRLCCQLD	PARDVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX019C3	MAGRLCCQLD	PARGVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLPLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX289C1	MAARLCCQLD	PARGVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	SPKVLKRTL
HX289C2	MAARLCCQLD	PARGVLCCLP	VGAESRGRPF	SGPLGTLSSP	SPSAVSTDHG	AHLPLRGLPV	CAFSSAGPCA	LRFTSARRMG	METTVDNAHQF	LPKVLYKRTL
HX289C3	MAARLCCQLD	PARDVLCCLP	VSAESRGRPL	PGPFGTLSPS	SSSAVPTDHG	AHLSLRGLPV	CAFSSAGPCA	LRFTSARRMX	-XTTVNARQV	LPKVLHKRTL

	110	120	130	140	150	
AY233278	GLPAMSTTDL	EAYFKDCVFK	DWEELGSEIR	LKVFVGGCR	HKLVCAPSSC	NFF TSA*
AF121251	GLSAMSTTDL	EAYFKDCVFT	EWEELGSEFR	LKVFVGGCR	HKLVCSPAPC	NFF TSA*
AB014371	GLSAMSTTDL	EAYFKDCLFK	DWEELGSESR	LMIFVGGCR	HKLVCSPAPC	NFF TSA*
FJ349218	GLSVMSTTDL	EAYFKDCLFK	DWEELGSEIR	LKVFVGGCR	HKLVCAPAPC	NFF TSA*
AP007262	GLSAMSTTDL	EAYFKDCLFK	DWEELGSETR	LMIFVGGCR	HKLVCAPAPC	NFF TSA*
AB166850	GLSGRSMTWI	EEYIKDCVFK	DWEELGSETR	LMIYVGGCR	HKLVCSPAPC	NFF TSA*
AP007264	GLFVMSTTG	EKYFKDCVFA	EWEELGSESR	LMTFVGGCR	HKLVCAPAPC	NFF TSA*
EU498228	GLSPRSTTWI	EEYIKDCVFK	DWEESGSELR	LKVFVGGCR	HKLVCSPAPC	NFF TSA*
HX055	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPC	NFF TSA*
HX025	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPC	NFF TSX-
HX036	GLSVMSTIDL	EAYFKDCLFT	NWEEXXSENR	LMIFVGGCR	HKLVCSPAPC	NFF TSV*
HX019	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPC	NFF TSA*
HX289	VLSVVSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLICSPPPC	NFF TSAX
HX098	GLSAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCSPAPC	NFF TSAX
HX149	GLSAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCSPAPC	NFF TSA*
HX394	GLPAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCSPAPC	NFF TSA*
HX464	GLPAMSTTDL	EAYFKDCVFK	DWEELGSEIR	LKVFVGGCR	HKLVCAPSSC	NFF TSA*
HX558	ELPAMSTTDL	EAYFKDCVFK	DWEELGSEIR	LKVFVGGCR	HKLVCAPSSC	NFF TSA*
HX964	GLPAMSTTDL	EAYFKDCVFK	DWEELGSEIR	LKVFVGGCR	HKLVCAPSSS	NFF TSA*
HX001	GLPAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCAPSPC	NFF TSA*
HX056	GLPAMSTTDL	EAYFKDCAFK	DWEELGSETR	LMIFVGGCR	HKLVCAPSPC	IFF TSAK
HX722	GLPAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCAPSPW	NLF PSA*
HX061	GLPAMSTTDL	EAYFKDCVFK	DWEELGSEIR	LKVFVGGCR	HKLVCAPSSC	NFF TSAX
HX064	GLPAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCAPSPC	NFF TSAX
HX382	GLPAMSTTDL	EAYFKDCVFK	DWEELGSEIR	LKIFVGGCR	HKLVCAPSSC	NFF TSA*
HX354	GLPAMSTTDL	EAYFKDCVFK	DWEELGSEIR	LKVFVGGCR	HKLVCAPSSC	NFF TSA*
HX414	GLPAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCAPSPC	NFF TSA*
HX362	GLPAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCAPSPC	NFF TSA*
HX055C1	GLPAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCAPAPC	NFF TSA-
HX055C2	GLPAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCAPAPC	NFF TSA-
HX055C3	GLPAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCAPIPC	NFF TSA-
HX098C1	GLSAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCAPSPC	NFF TSAK
HX098C2	GLSAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCSPAPC	NFF TSA-
HX098C3	GLSAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCSPAPC	NFF TSA-
HX025C1	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPC	NFF TSA*
HX025C2	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPC	NFF TSA-
HX025C3	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPC	NFF TSA-
HX036C1	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPGPC	NFF TSA*
HX036C2	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAP-PR	NFF TSA-
HX036C3	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPR	NFF TSA-
HX019C1	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPC	NFF TSA*
HX019C2	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPC	NFF TSA-
HX019C3	GLSVMSTIDL	EAYFKDLLT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPC	NFF TSA-
HX289C1	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENR	LMIFVGGCR	HKLVCAPAPC	NFF TSA-
HX289C2	GLSVMSTIDL	EAYFKDCLFT	NWEELGSENG	LMIFVGGCR	HKLVCAPAPC	NFF TSA-
HX289C3	GLSAMSTTDL	EAYFKDCVFK	DWEELGSETR	LMIFVGGCR	HKLVCSPHQC	NFF TSA*

**Figure 4.5:** The amino acid sequences deduced from HBV X gene nucleotides compared with corresponding representative references of genotypes A to H: AY233278, AF121251, AB014371, FJ349218, AP007262, AB166850, AP007264 and EU498228. The samples and clones are HX055-HX362 and HX055C1-HX289C3. A three amino acid insertion, GME between amino acid 79 and 83 shifted the numbering system of the reading frame and the highlighted region represent the basal core promoter region.

**Table 4.1:** A summary of amino acid substitutions relevant to HBx mediated pathogenesis

Nucleotide position	Amino acid Substitution	Number of samples with the mutation (n=20)	Number of clones with the mutation (n=18)	Summary of the reported functions
T1464G	S31A	12	3	Escape from immune surveillance (Yeh et al, 2000).
T1485C	P38S	18	14	Escape from immune surveillance (Utama et al, 2009).
C1497T	S42P	18	14	Phosphorylation motif used for HBx localisation (Noh et al,2004)
T1578C	C69R	1	3	Found within T and B lymphocytes epitopes and may be responsible for immune pressure escape (Shen et al, 2008)
T1590C	F73L	0	3	
C1653T	H94Y	5	11	Responsible for transactivation function, bind to protein kinase C disrupting cell growth (Cong et al, 1997).
C1674T	P101S	7	15	Leads to strong activation of p21, reduction and increased apoptosis (Kwun and Jang,2004)
A1726C/G1728A	K118T/D119N	5	11	Escape of virus from immune surveillance (Zhu et al, 2008)
T1753C/A	I127T/N	14	18	Promotes transactivation and increase antiproliferative activity ( Lin et al, 2005)
A1762T/G1764A	K130M	14	18	Interfere with cell growth control,DNA repair, HBeAg synthesis and inhibitor of p21 (Yuan 2009, Kramvis and Kew 1999, Lin et al, 2005)
	V131I	15	18	

**Table 4.2:** A summary of substitutions with unknown functions detected in the study

<b>Nucleotide substitution</b>	<b>Amino Acid substitution</b>	<b>Number of samples with the mutation (n=20)</b>	<b>Number of clones with the mutation (n=18)</b>
C1378G	A2G	1	0
C1381G	A3G	2	4
A1383G	A4G	1	0
C1497C	P42S	2	4
C1690T	T106I	5	11
G1719T	V116L	5	11

#### 4.6 Nucleotide variations within HBV X gene basal core promoter region

The sequences of 20 HBV strains and their clones were examined for changes within the region spanning the X ORF (Figure 4.1) which may affect HBeAg expression. The changes are summarised in the table below.

**Table 4.3:** Correlation of HBeAg positive status with basal core promoter mutations

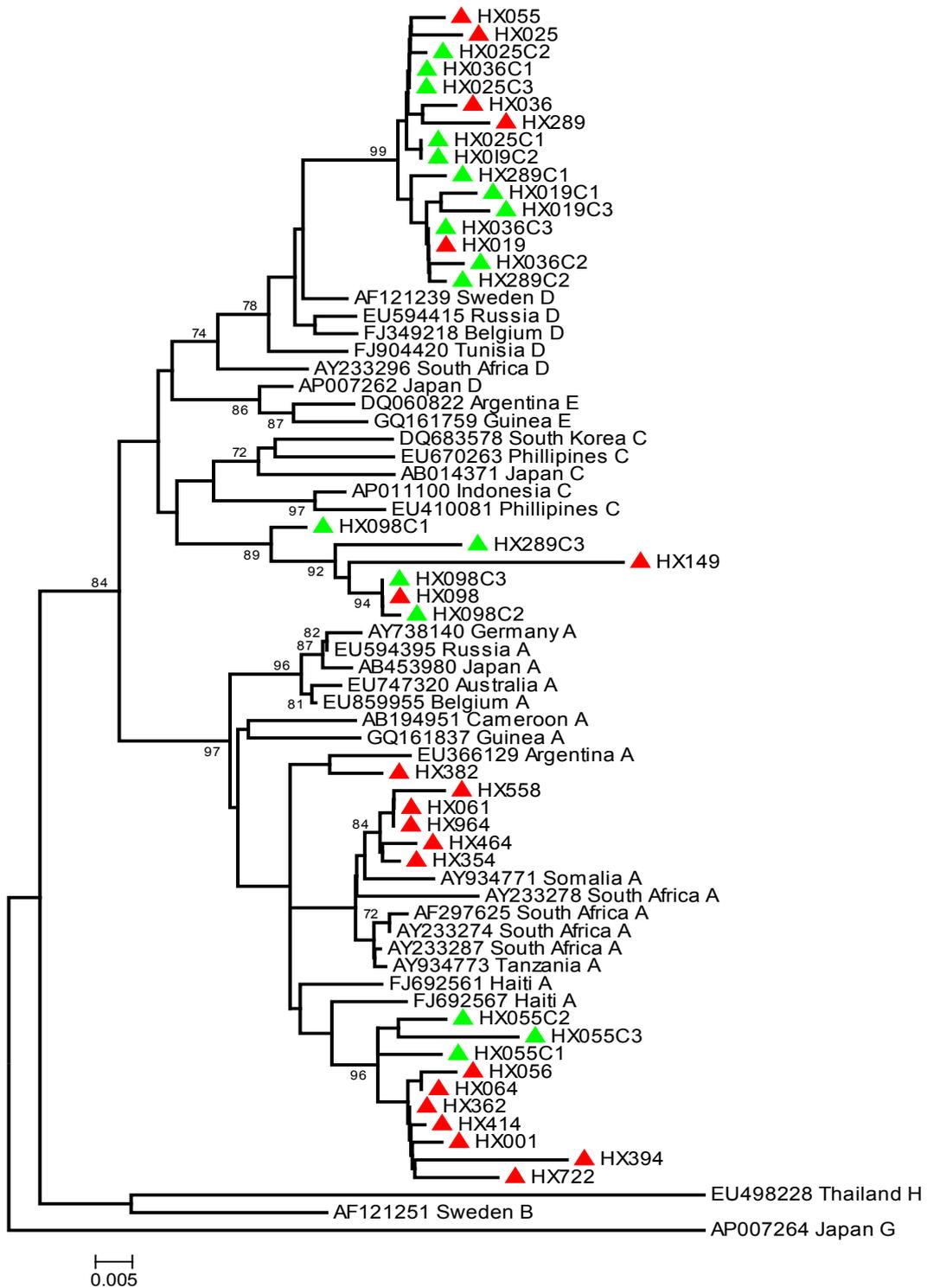
Patients samples with clones	Presence of basal core promoter mutations		
	T1753M	A1762T/G1764A	G1764A/C1766T
HX055	Yes	Yes	No
HX055C1	Yes	Yes	No
HX055C2	Yes	Yes	No
HX055C3	Yes	Yes	No
HX964	No	No	No
HX098	Yes	Yes	No
HX098C1	Yes	Yes	No
HX098C2	Yes	Yes	No
HX098C3	Yes	Yes	No
HX722	Yes	Yes	No
HX061	No	No	No

**Table 4.4:** Correlation of HBeAg negative status with basal core promoter mutations

Patients samples with clones	Presence of basal core promoter mutations		
	T1753M	A1762T/G1764A	G1764A/C1766T
HX149	Yes	Yes	No
HX394	Yes	Yes	No
HX464	No	No	No
HX558	No	No	No
HX001	Yes	Yes	No
HX056	No	No	No
HX064	Yes	Yes	No
HX019	Yes	Yes	No
HX289	Yes	Yes	No
HX289C1	Yes	Yes	No
HX289C2	Yes	Yes	No
HX289C3	Yes	Yes	No
HX354	No	No	No
HX036	Yes	Yes	No
HX036C1	Yes	Yes	No
HX036C2	Yes	Yes	No
HX036C3	Yes	Yes	No
HX382	No	No	Yes
HX414	Yes	Yes	No
HX362	Yes	Yes	No
HX025	Yes	Yes	No
HX025C1	Yes	Yes	No
HX025C2	Yes	Yes	No
HX025C2	Yes	Yes	No

#### 4.7 Phylogeny of HBV X gene sequences

A total of 38 HBV X gene sequences were compared with 29 corresponding full length references from GenBank, a dendrogram was obtained by neighbour-joining phylogenetic analysis (Figure 4.5). The bootstrap statistical analysis was performed using 1000 replicates which supported the nodes of the dendrogram very well as shown by the high percentages (>70%) on the trees. Five samples, namely HX055, HX036, HX289, HX019 and clones (HX025C1-C3, HX019C1-C3, HX036C1-C3 and HX289C1-C2) clustered with genotype D reference strains. The two samples (HX149 and HX098) and two clones (HX098C1-C3 and HX289C3) clustered with genotype C reference strains. Finally, 13 samples (HX382, HX558, HX061, HX964, HX464, HX354, HX056, HX064, HX382, HX414, HX001, HX394 and HX722) and 3 clones (HX055C1-C3) clustered with genotype A. Most of the samples clustered with genotype A which is a major genotype circulating in our region. The hypervariability of the HBV X gene is shown as the samples and their clones did not cluster within the same genotype in most cases.



**Figure 4.6:** A dendrogram obtained by neighbour-joining phylogenetic analysis of HBV X gene sequences rooted on genotype G, showing the relationship of samples and clones with genotypes from various countries. The red and green colours represent samples and clones respectively.

## CHAPTER 5

### 5. DISCUSSION

#### 5.1 Overview of the study

HBV X gene is the smallest ORF of the HBV genome. It consists of 465 nucleotides which code for a 154 amino acid product. HBx is associated with various liver diseases due to its ability to transactivate various cellular genes involved in the regulation and function of cellular processes. Most importantly are the amino acids substitutions which are mediated in the pathogenesis of HBx. The substitutions arise during erroneous replication of HBV genome and are capable of evading immune system and disrupting signal and transcription pathways. Furthermore, basal core promoter region within the X gene responsible for transcription of preC mRNA may develop mutations which can suppress transcription of HBeAg.

In this study, HBV DNA was extracted from 20 samples and the HBx gene was amplified with an optimised PCR assay adapted from Chen et al (2005) protocol. Subsequent sequencing revealed a six nucleotide insertion (GCATGG) within the reading frame of five samples and point mutations in the rest of the samples. Following the use of successful cloning technique, six samples were cloned and three clones were isolated from each sample to further characterise X gene. A total of 38 sequences were generated from the study.

Our samples revealed substitutions within the regulatory and the functional domains of HBx which are involved in certain biological functions of the protein. HBx substitutions associated with its pathogenesis were detected in most samples together with their clones and also within the enhancers II and basal core promoter region. The following sections explain in detail the major finding of the study compared with other studies done elsewhere.

## **5.2 HBx substitutions associated with its biological function and pathogenesis**

HBx has been implicated in several liver diseases such as chronic hepatitis B and HCC (Zhang et al, 2005). This is due to its N and C terminals which act as negative control of proapoptotic activity and promote transactivation respectively (Kidd-Ljunggren et al 1995, Murakami, 1999, Datta et al, 2008). And these functions are mainly performed by residues found within the N and C terminal. The S31A, P38S, P42S substitutions were detected in the N terminal and their functions are well studied.

Yeh et al (2000) reported that S31A was shown to induce low expression of apoptosis which could lead to less regeneration of hepatocytes and was detected in 12 samples and 3 clones of HBV infected patients. Utama et al (2009) reported that S31A was not associated with advanced liver diseases but may result from immune surveillance escape which might contribute to the process of hepatocarcinogenesis. Serine and proline 31 were considered to be the common variants in our study and only one sample HX382, had threonine 31 which was also found by Yeh et al (2000). The S31A substitution was found only in samples and clones which clustered with genotype A and Datta et al (2008) found it to belong to genotype A and suggested that it may be genotype specific.

A change of P38S was detected in 18 samples and 14 clones within the rich serine and glycine region mapped by Kumar et al (1996). Muroyama et al (2006) reported P38S substitution to be frequently associated with HCC than chronic hepatitis B in genotype C. However, the samples and clones with P38S substitution in the present study clustered with genotype A and D and not with genotype C. The P38S was found to have emerged from virus escape immune surveillance because it forms part of B cell epitopes located in positions 29-48 (Park et al, 2000; Utama et al, 2009).

The P42S substitution found within the phosphorylation motif of extracellular-response kinases (ERK) which forms part of the signalling cascades of the cell (Noh et al, 2004). The P42S substitution was detected in samples HX098 and HX149 as well as clones HX098C1-C3 and HX289C3, which may be involved in disturbing the signalling pathways of the cell. However, the S42P substitution detected in 18 samples and 14 clones has been found to be responsible for determining the subcellular localisation of HBx in the nucleus which may be involved in HBV mediated carcinogenesis (Noh et al, 2004). Therefore the P42S might represent a new strategy of the virus in liver disease pathogenesis. The P38S and S42P found within the proline serine rich region formed a positive correlation in samples and clones of genotype A and D in the present study.

The C terminal had most substitutions, namely C69R, F73L, H94Y, P101S, T106I, V116L, K118T, D119N, I127T/N and K130M/V131I. There was also a six nucleotide insertion GCATGG, causing a GME insertion at position 79 to 83 of amino acids in the sequences causing a frameshift in the reading frame. The GME insertion is found within amino acid 77 to 97; which has been shown to be important for localisation of HBx to the mitochondria by forming an amphipathic alpha helix (Boya et al, 2004). The six nucleotide GCATGG insertion found in 5 samples and 11 clones within the X region clustered with genotype D isolates and through the GenBank search it was found earlier by Takahashi et al (1998) and its function is unknown. It was also found in genotype C isolates using full-genome for genotyping. Furthermore, a three amino acid insertion in the C terminal of HBx have been reported before by Kramvis et al (2002) at position 152 to 154 of amino acids.

The C69R and F73L substitutions detected in HX098 sample and clones of HX055C1-C3 found within the T and B lymphocyte epitope and may be responsible for immune pressure escape during viral host interactions (Shen et al, 2008). The C69R mutation was found in genotype C isolate (HX098) and F73L detected in clones of HX055C1-C3 which clustered with genotype A references.

The Enh II and basal core promoter regions found within the X region were shifted by the GME insertion. The H94Y substitution caused by 1653T mutation was detected in 5 samples and 11 clones and is found within the immunodominant domain (amino acid 85 -110) within enhancer II (Shinkai et al 2007; Takahashi, 1999). In addition, codon 94 is essential for transactivation function and is required for binding to protein kinase C-binding protein (XAP3) which is involved in cell growth, differentiation and carcinogenesis (Cong et al 1997). Interestingly sequences with GME insertion had H94Y substitution which is found within the functional domain of HBx and as a result might have the potential to cause multiple harmful effects to the hepatocytes. Furthermore H94Y was shown to be involved in severe liver diseases in Korean patients with genotype C (Kim et al, 2008). However our patients with H94Y together with their clones formed a cluster with genotype D.

The P101S substitution was detected in 7 samples and 15 clones. Furthermore the residue has been shown to be responsible for strong activation of p21 in Hep G2 which eventually disturbs cell cycle control and thus can lead to apoptosis (Kwun and Jang, 2004). Consequently the substitution could cause dysregulation of cell growth which is important for hepatocarcinogenesis. The P101S was found in genotype D and C isolates.

The T106I substitution detected in 5 samples and 11 clones has been reported to be associated with HCC in genotype A (Garmiri et al, 2009). However samples and clones with T106I substitution, clustered with genotype D. The double mutation of A1726C/G1728A which is located within regulatory elements and Enh II has been reported to affect replication of HBcAg *in situ* (Zhu et al, 2008). The K118T and D119N substitutions caused by double mutation of A1726/G1728A detected in 5 samples and 11 clones of genotype D isolates have been associated with liver disease manifestation (Utama et al, 2009). Furthermore it forms part of T helper cell epitope (amino acid 111-135) and T cellular cell epitope (amino acid 115-123) which might affect viral abilities to evade immune clearance (Zhu et al, 2008).

The I127T/N substitution has been shown to promote transactivation and increase antiproliferative activity in cell culture (Lin et al, 2005). Our sequences had I127T/N due to C1753M mutation which appeared in 14 samples and 18 clones. Interestingly I127N only appeared in samples and clones which clustered with genotype D and those with I127T appeared in genotype A and C. It has been speculated that the appearance I127T/N and H94Y substitutions in a patient could be responsible for transactivating of oncogenes responsible for HCC (Liu et al, 2009). Shinkai et al (2007) suggested that these can be due to HBV mechanism to evade the immune system leading to persistent viral replication which is important for hepatocarcinogenesis.

Another interesting mutations detected in the study is the dual mutations of A1762T/G1764A which are nonsynonymous and cause substitutions of K130M and V131I in the X region. The K130M and V131I were detected in 32 (14 samples and 18 clones) and 33 (15 samples and 18 clones) sequences respectively. The substitutions of K130M and V131I has been extensively studied by many groups and has been suggested to interfere with cell growth control, DNA repair, HBeAg synthesis and a strong inhibitor of p21 (Yuan, 2009; Kramvis and Kew,1999; Kwun and jung, 2004; Lin et al, 2005). They have also been associated in several liver diseases such as liver cirrhosis, HCC, asymptomatic carrier state and chronic hepatitis (Utama et al 2009; Baptista et al 1999; Shinkai et al 2007; Kim et al, 2008). Patients with A1762T/G1764A dual mutations can be at risk of developing HCC and cirrhosis as the mutations are involved in multiple pathogenic effects. Furthermore the K130M and V131I substitutions belonged to genotypes A, C and D isolates and might be genotype independent.

In addition, H94Y, P101S, K118T, D119N, V116L, I127T/N, K130M, V131I substitutions were found in combination in samples HX055,HX025, HX036, HX019 and HX289, as well as in clones HX025C1-C3, HX036C1-C3, HX019C1-C3 and HX289C1-C2, which belonged to genotype D.

### 5.3 Basal core promoter mutations involved in HBeAg expression

The basal core promoter region is found between nucleotide 1751 to 1769 (Liu et al, 2009). Our study correlated the HBeAg expression with mutations found within basal core promoter region. From the total of 19 samples tested for HBeAg, 14 were positive and 5 were negative.

Mutations within basal core promoter region have been studied elsewhere and have been found in HBeAg positive (Marrone et al 2003) and negative (Abbas, 2006) patients. Normally liver specific nuclear factor CCAAT/ enhancer binding protein and hepatocyte nuclear factor bind to basal core promoter region and activate transcription of HBeAg (Liu et al, 2009). However, mutational hotspots within basal core promoter region cause less or abolished binding of these factors to the TA rich region and this effect was found to be associated with HBeAg reduction in serum or lead to its negativity (Buckhold et al, 1996; Baptista et al, 1999).

The basal core promoter region mutations, namely; T1753M, A1762T and G1764A were found in both HBeAg positive and negative samples in this study. However most HBeAg negative samples had triple mutations (T1753M, A1762T, G1764A) which were detected in nine samples but were absent in four samples. The HX382 sample which was HBeAg negative had a double mutation of 1764A/1766T and S31T substitution. In contrast, the C1766T usually appears alone and enhance replication by two fold and is also found in the nuclear receptor-binding which therefore suppresses HBeAg (Perekh et al, 2003). In samples positive for HBeAg; triple mutations were present in three (HX055, HX098, HX722) and absent in two (HX964, HX061) samples. The results showed mixed results, but the majority of samples with triple mutation were HBeAg negative. These results are in agreement with Lim and colleagues (2006) where mutations within basal core promoter region did not necessarily mean the absence of HBeAg.

The C1653T mutation has also been reported to affect HBeAg synthesis through its activity with basal core promoter region. Since it is located within the box  $\alpha$  element (nucleotide 1646-1668) which is a strong activator element of basal core promoter, the mutation converts the element into the perfect palindromic sequence 1648-TCTTATATAAGA causing enhanced production of HBeAg (Takahashi et al, 1999; Shinkai et al 2007). And about three and two samples positive and negative for HBeAg respectively were positive for 1653T together with T1753M/A1762T/G1764A triple mutation. And furthermore the above mentioned samples had a six base pair insertion (GCATGC) in the nucleotides of the reading frame.

Moreover the P38S substitution detected in 18 samples has been reported by Muroyama et al (2006) to be associated in HBeAg negative patients. However our findings revealed that serine 38 was found in 13 HBeAg negative patients and to a lesser degree in four HBeAg positive patients. Therefore these finding might be affected by the duration of HBV infection as HBeAg is lost overtime.

HBeAg shares epitopes with HBcAg and has been speculated to cause immunotolerance towards itself or HBcAg or both and in the absence of HBeAg, HBcAg may be targeted by the immune system leading to inflammation of hepatocytes (Kramvis and Kew 1999; Jammeh et al, 2008). In the present study, most samples which were positive for anti-HBc were negative for HBeAg and in other samples where HBeAg was positive, anti-HBc was negative. Therefore, the presence of anti-HBc could lead to HBeAg negativity depending on the immune status of the patient.

#### **5.4 Implications of the study**

The variability within HBV X gene has been established in our study and elsewhere (Datta et al, 2008) to be important for various pathogenenic effects (i.e apoptosis, cell cycle disruption) caused during hepatitis infection by disrupting normal cellular pathways. The oncogenic potential of HBx caused by substitutions can serve as prognostic markers for patients infected with HBV at risk of advanced liver diseases such as HCC. The substitutions are genotype

specific and might depend on geographical regions. Furthermore, triple mutations of T1753M, A1762T, G1764A found in the basal core promoter region within the X gene together with the presence of anti-HBc can lead to HBeAg negativity.

## CHAPTER 6

### 6 Conclusions, limitations and recommendations

#### 6.1 Conclusions

HBV X gene was successfully characterised using various molecular methods and mutations in the X gene of HBV infected patients with different serological profiles were detected. Accumulation of the mutations in the X gene leading to HBx substitutions may present a risk for patients because they are involved in various pathogenic effects.

HBx substitutions which circulate in different genotypes may affect the different pathogenic potential due to complex interactions of the virus with the host. In our study, genotype D isolates displayed most mutations/substitutions in the C terminal of HBx which is responsible for transactivation and patients with this genotype may be at high risk of HCC. Furthermore, the six base pair insertion, GCATGG could emerge as a new variant for genotype D in our region.

The pattern of the mutations formed may serve as prognostic factors as they represent a strategy of HBV to escape the immune system and lead to persistent liver infections which are important for hepatocarcinogenesis. Furthermore, the study represents the first formal investigation of HBV X gene genetic variability in South Africa.

The T1753M, A1762T and G1764A triple basal core promoter mutations mostly detected in HBeAg positive samples and to a lesser degree in HBeAg negatives. Although basal core promoter mutations affect the replication of HBeAg during transcription, other mutations in the core region are important for effective termination of HBeAg.

## **6.2 Limitations of the study**

The patients' sera used in the study lacked clinical information. Therefore, the pattern of mutations/substitutions could not be linked to clinical profiles of the patients. Due to the lack of HCC patients in this study, no correlation of the mutations with the risk of carcinogenesis was made and the study relied on published reports from different countries.

The antibodies to HBeAg were not tested; therefore, no positive correlation was made between HBeAg and basal core promoter mutations in HBeAg negative patients. Furthermore basal core promoter mutations are thought to increase HBV DNA levels; however, no conclusions were made as viral load was not performed on the patients.

## **6.3 Recommendations**

- i. Further characterisation studies on HBV X gene genetic variability are needed to reveal genotype associated mutations/substitutions. Especially prospective studies which might indicate HBx substitutions to be important risk factors for HBV related carcinogenesis.
- ii. Functional studies which involve transfection experiments are needed to determine the pathological significance of HBx substitutions already detected elsewhere and others detected in the present study. The knowledge of specific pathways can lead to development of new therapeutic agents to effectively reduce the period of HBV infection and thus prevent further liver complications.
- iii. Classification of HBV strains on full genome is required as phylogenetic analysis based on X gene needs to be confirmed using classical methods of genotyping and its hypervariability might cause discrepancies.

## CHAPTER 7

Abbas Z, Muzaffar R, Siddiquil A, Naqui SAA, Razzi SAH. Genetic variability in the precore and core promoter regions of Hepatitis B virus strains in Karachi. *BioMed Central Gastroenterology* 2006; 6: 1-6

Aliyu S, Aliyu H, Salinu H, Parma S, Jalal H, Curian M. Rapid detection and quantification of Hepatitis B virus DNA by real-time PCR using new fluorescent (FERT) detection system. *Journal of Clinical Virology* 2004; 30: 191-195

Andre F. Hepatitis B epidemiology in Asia, the Middle East and Africa. *Vaccine* 2000; 18: 20-22

Baptista M, Kramvis A, Kew MC. High prevalence of 1762T1764A mutations in the basic core promoter of hepatitis B virus isolated from Black Africans with hepatocellular carcinoma compared with asymptomatic carriers. *Hepatology* 1999; 29: 946-953

Baumert TF, Thimme R, Von Weizsacker F. Pathogenesis of hepatitis B virus infection. *World Journal of Gastroenterology* 2007; 13: 82-90

Beck J, Nassal M. Hepatitis B virus replication. *World Journal of Gastroenterology* 2007; 13: 48-64

Block TM, Guo H, Guo JT. Molecular virology of hepatitis B virus for clinicians. *Clinics in Liver Diseases* 2007; 11: 685-706

Blumberg BS, Gerstley BJS, Hungerford DA, London WT, Sutnick AJ. A serum antigen (Australian antigen) in Down's Syndrome, leukemia and hepatitis. *Annals of Internal Medicine* 1967; 66: 924-931

Boya P, Pauleau AL, Poncet D, Gonzalez-Polo RA, Zamzam N, Kroemer G. Viral proteins targetting mitochondria: controlling cell death. *Biochemica et Biophysica Acta* 2004; 1659: 178-189

Bruss V. Hepatitis B virus morphogenesis. *World Journal of Gastroentorology* 2007; 13: 65-73

Buckhold VE, Xu Z, Chen M, Benedict YTS, Ou JH. Effects of a naturally occuring mutation in the hepatitis B virus basal core promoter on precore gene expression and viral replication. *Journal of Virology* 1996; 70: 5845-5851

Burnett RJ, Francois G, Kew MC, Leroux-Roels G, Meheus A, Hoosen AA, Mphahlele MJ. Hepatitis B virus and human immunodeficiency virus co-infection in sub-Saharan Africa; a call for further investigation. *Liver International* 2005; 25: 201-213

But DYK, Lai CL, Yeun MF. Natural history of hepatitis-related hepatocellular carcinoma. *World Journal of Gastroentorology* 2008; 14: 1652-1656

Chen GG, Li MY, Ho RLK, Chak ECW, Lau WY, Lai PBS. Identification of hepatitis B virus X gene mutation in Hong Kong patients with hepatocellular carcinoma. *Journal of Clinical Virology* 2005; 34: 7-12

Clippinger AJ, Bouchard MJ. Hepatitis B virus HBx localizes to mitochondria in primary Rat hepatocytes and modulates mitochondrial membrane potential. *Journal of Virology* 2008; 82: 6798-6811

Conway JF, Cheng N, Zlotnick A, Stahl SJ, Wingfield PT, Steren AC. Localization of the N terminus of hepatitis B virus capsid protein by peptide based difference mapping from cytoelectron microscopy. *Biochemistry* 1998; 95: 14622-14627

Cong YS, Yao YL, Yang WM, Kuzhandaivelu N, Seto E. The hepatitis B virus X-associated protein, XAP3, is a protein kinase C-binding protein. *The Journal of Biological Chemistry* 1997; 272: 16482-16489

Dane DJ, Cameron CH, Briggs M. Virus-like particles in serum of patients with Australian-antigen-associated hepatitis. *Lancet* 1970; 1: 695-698

Datta S. An overview of molecular epidemiology of hepatitis B virus (HBV) in India. *Virology journal* 2008; 5: 156

Datta S, Banerjee A, Chandra PK, Biswas A, Panigrahi R, Mahapatra KP, Randa CK, Chakrabanti S, Bhattacharya SK, Chakravarty R. Analysis of hepatitis B virus of X gene phylogeny, genetic variability and its impact on pathogenesis: Implications in Eastern Indian HBV carriers. *Virology* 2008; 5: 1-9

Delaney WE, Esonda-Boroto K. Therapy of chronic hepatitis B: trends and development. *Current Opinion in Pharmacology* 2008; 8: 532-540

Diestag JL. Hepatitis B virus Infection. *The New England Journal of Medicine* 2008; 359: 486-500

Dudley FJ, Sherlock S, Scheufer PJ. Natural history of hepatitis-associated antigen-positive chronic liver disease. *The Lancet* 1972; 2: 1383-1393

Elgouhari HM, Tammi TIA, Carey WD. Hepatitis B virus infection: understanding its epidemiology, course and diagnosis. *Cleveland Clinic Journal of Medicine* 2008; 75: 881-889

Fattovich G, Bortolotti F, Donato F. Natural history chronic hepatitis B: Special emphasis on disease progression and prognostic factors. *Journal of hepatology* 2008; 48: 335-352

Feitelson MA, Duan LX. Hepatitis B virus X antigen in the pathogenesis of chronic infections and the development of hepatocellular carcinoma. *American Journal of Pathology* 1997; 150: 1141-1157

Feitelson MA, Lee J. Hepatitis B virus integration, fragile sites and hepatocarcinogenesis. *Cancer Letters* 2007; 252: 157-170

Feitelson MA, Reis HMGPV, Liu J, Lian Z, Pan J. Hepatitis B virus X antigen and cell cycle control in chronic infection and hepatocarcinogenesis. *Frontiers in Bioscience* 2005; 10: 1558-1572

Feld JJ, Wong DKH, Heathcote J. Endpoints of therapy in chronic hepatitis B. *Hepatology* 2009; 49: 96-102

Foster GR, Thomas HC. Recent advances in the molecular biology of hepatitis B virus: mutant virus and host responses. *Gut* 1993; 34: 1-3

Francois G, Kew M, Van Damme P, Mphahlele MJ, Meheus A. Mutant hepatitis B viruses: a matter of academic interest only or a problem with far reaching implications? *Vaccine* 2001; 19: 3799-3815

Garmiri P, Loua A, Haba N, Candotti D, Allain JP. Deletions and recombinations in the core region of hepatitis B virus genotype E strains from asymptomatic blood donors in Guinea, West Africa. *Journal of General Virology* 2009; 90: 2442-2451

Glebe D. Recent advances in hepatitis B virus research: A German point of view. *World Journal of Gastroenterology* 2007; 13: 8-13

Greer S, Alexandra GJM. Viral serology and detection. *Baillier's Clinical Gastroenterology* 1995; 9: 689-721

Guo H, Mason WS, Aldrich CE, Saputelli JR, Miller DS, Jilbert AR, Newbold JE. Identification and characterisation of Avihepadnaviruses located from exotic Anseroformes maintained in captivity. *Journal of Virology* 2005; 75: 2729-2742

Gunther S, Piwon N, Will H. Wild-type levels of pregenomic RNA and replication synthesis of hepatitis B virus with C (1653) T, A (1762) T and G (1762) and G (1764) A mutation in the core promoter. *Journal of General Virology* 1998; 79: 375-380

Gupta S. Molecular steps of cell suicide: An insight into immune senescence. *Journal of Clinical Immunology* 2000; 20: 229-239

Gurtsevitch VE. Human Oncogenic viruses: Hepatitis B and Hepatitis C viruses and their role in hepatocarcinogenesis. *Biochemistry* 2008; 73; 504-513

Hall TA. A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium seminar* 1999; 42: 95-98.

Hatzakis A, Magiorkinis E, Hida C. HBV virological assessment. *Journal of Hepatology* 2006; 44: 71-76

Heathcote EJ. Treatment of hepatitis B: the next five years. *Clinical Medicine* 2007; 7: 472-477

Hilleman MR. Critical overview and outlook: pathogenesis, prevention and treatment of hepatitis and hepatocellular caused by hepatitis B virus. *Vaccine* 2003; 21: 4626-4649

Hoofman CJ, Thio CL. Clinical implications of HIV and hepatitis B co-infection in Asia and Africa. *Lancet Infectious Diseases* 2007; 7: 402-409

Huang SN, Millman I, O'Connell BA, Aronoff A, Gault H, Blumberg BS. Virus like particles in Australia-Antigen-Associated hepatitis. *American Journal of Pathology* 1972; 67: 453-470

Huang CF, Lin SS, Ho CY, Chen FL, Yuang CC. The immune response induced by hepatitis B virus principal antigens. *Cellular and Molecular Immunology* 2006; 3: 97-106

Jammeh S, Tauner F, Watson R, Thomas HC, Karayiannis P. Effect of basal core promoter and pre-core mutations of hepatitis B virus replication. *Journal of General Virology* 2008; 89: 901-909

Karthigesu VD, Allison LMC, Fortuin M, Mendy M, Whittle HC, Howard CR. A novel hepatitis B virus variant in the sera of immunized children. *Journal of General Virology* 1994; 75: 443-448

Keeffe EB, Dieterich DT, Pawlotsky JM, Benhamou Y. Chronic hepatitis B preventing, detecting and managing viral resistance. *Clinical Gastroenterology and Hepatology* 2008; 6: 268-274

Kew MC. Progress towards the comprehensive control of hepatitis B in Africa: a view from South Africa. *Gut* 1996; 35: 31-36

Khouri ME, dos Santos VA. Hepatitis B: Epidemiological, Immunological and serological considerations emphasizing mutation. *Revista Do Hospital Das Clinicas* 2004; 59: 216-224

Kidd-Ljunggren K, Miyakawa Y, Kidd AH. Genetic variability in hepatitis B viruses. *Journal of General Virology* 2002; 83: 1267-1280

Kidd-Ljunggren K, Oberg M, Kidd AH. The hepatitis B virus X gene: analysis of functional domain variation and gene phylogeny using multiple sequences. *Journal of General Virology* 1995; 76: 2119-2130

Kiire CF. The epidemiology and prophylaxis of hepatitis B in sub-Saharan Africa: a view from tropical and subtropical Africa. *Gut* 1996; 38: 5-12

Kim HJ, Park JH, Jee Y, Lee SA, Kim H, Song BC, Yang S, Lee M, Yoon JH, Kim YJ, Lee HS, Hwang ES, Kook YH, Kim BJ. Hepatitis B virus X mutations occurring naturally associated with clinical severity of liver diseases among Korean patients with chronic genotype C infection. *Journal of Medical Virology* 2008; 80: 1337-1343

Kim KH, Seong BL. Pro-apoptotic function of HBV X protein is mediated by interaction with c-FLIP and enhancement of death-inducing signal. *The European Molecular Biology Organization Journal* 2003; 22: 2104-2116

Kim S, Kim HY, Lee S, Kim SW, Sohn S, Kim K, Cho H. Hepatitis B virus X protein induces perinuclear mitochondrial clustering in microtubule and dynein-dependent manners. *Journal of Virology* 2007; 81: 1714-1726

Kimbi G, Kravis A, Kew M. Distinctive sequence characteristics of subgenotypes A1 isolates of Hepatitis virus from South Africa. *Journal of General Virology* 2004; 85: 1211-1220

Kramvis A, Kew MC. The core promoter of hepatitis B virus. *Journal of Viral hepatology* 1999; 6: 415-427

Kramvis A, Kew M, Francois G. Hepatitis B virus genotypes. *Vaccine* 2005; 23: 2409-2423

Kramvis A, Weitzmann L, Owiredo WKBA, Kew MC. Analysis of subgroup A hepatitis B virus isolates from South Africa. *Journal of General Virology* 2002; 83: 835-839

Kottlil S, Jackson JO, Polis MA. Hepatitis B and Hepatitis C in HIV infection. *Indian Journal of Medical Research* 2005; 121: 474-450

Kumar U, Jarasuryan N, Kumar R. A truncated mutant (residue 58-140) of the hepatitis B virus X protein retains transactivation function. *Medical Sciences* 1996; 93: 5647-5652

Kwun HJ, Jang KL. Natural variants of hepatitis B virus X protein have differential effects on the expression of cyclin-dependent kinase inhibitor p21 gene. *Nucleic Acids Research* 2004; 32: 2202-2213

Lai L, Hui CK, Leung N, Lau G. Pegylated interferon alpha-2a (40kDa) in the treatment of chronic hepatitis B. *International Journal of Nanomedicine* 2006; 3: 255-262

Lee WM. Hepatitis B virus infection. *Medical Progress* 1997; 337: 1733-1745

Liang TJ. Hepatitis B: the virus and disease. *Hepatology* 2009; 49: 13-21

Lim CK, Tan JTM, Khoo JBS, Ravachandran A, Low HM, Chan YC, Ton SH. Correlations of HBV genotypes, mutations affecting HBeAg/anti-HBe status in HBV carriers. *International Journal of Medical Sciences* 2006; 3: 14-20

Liu CJ, Kao JH. Hepatitis B virus-related hepatocellular carcinoma: Epidemiology and pathogenic role of viral factors. *Journal of Chinese Medical Association* 2007; 70: 141-145

Lin CL, Kao JH. Hepatitis B viral factors and clinical outcomes of chronic hepatitis B. *Journal of Biomedical Science* 2008; 15: 137-145

Lin X, Xu X, Huang QL, Liu YQ, Zheng DL, Chen WH, Lin JY. Biological impacts of hot-spot mutations of hepatitis B virus x proteins are genotype B and C differentiated. *World Journal of Gastroenterology* 2005; 11: 4703-4708

Liu S, Zhang H, Gu C, Yin J, He Y, Xie J, Cao G. Association between hepatitis B virus mutations and risk of hepatocellular carcinomas: a meta-analysis. *Journal of the National Cancer Institute* 2009; 101: 1066-1082

Marrone A, Zampiro R, Luongo G, Utilil R, Karagiannis P, Ruggiero G. Low HBeAg serum levels correlate with the double A1762T/1764A core promoter mutation and a positive response to interferon patients with chronic hepatitis B virus infection. *Intervirology* 2003; 46: 222-226

McMahon BJ. The natural history of chronic hepatitis B virus infection. *Hepatology* 2009; 49: 45-55

Michielson PP, Francque SM, van Dongen JL. Viral hepatitis and hepatocellular carcinoma. *World journal of Surgical Oncology* 2005; 3: 1-18

Mphahlele MJ, Burnett RJ, Kyaw T, Schoeman HS, Aspinall S. Immunogenicity and safety of yeast-derived recombinant hepatitis B vaccine (Heberbiovac HB) in South African children. *South African Medical Journal* 2004; 94: 280-281

Mphahlele MJ, Lukhwareni A, Burnett R, Maropeng L, Ngobeni J. High risk occult hepatitis B infection in HIV-positive patients from South Africa. *Journal of Clinical Virology* 2006; 35: 14-20

Murakami S. Hepatitis B virus X protein: structure, function and biology. *Intervirology* 1999; 42; 81-99

Muroyama R, Kato N, Yashida H, Otsuka M, Moriyama M, Wang Y, Shao R-X, Pharei N, Tanaka Y, Ohta M, Tateishi R, ShiinA S, Tatsukawa M, Fukai K, Imazeki F, Yokosuka O, Shiratoni Y, Omata M. Nucleotide change of codon 38 in X gene of hepatitis B virus genotype C is associated with an increased risk of hepatocellular carcinoma. *Journal of hepatology* 2006; 45: 805-812

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

Noh EJ, Jung HJ, Jeong G, Choi KS, Park HJ, Lee JS. Subcellular localization and transcriptional repressor activity of HBx on p21WAF/Cip1 promoter is regulated by ERK-mediated phosphorylation. *Biochemical and Biophysical Research Communications* 2004; 319: 738-745

Ocama P, Opio CK, Lee WM. Hepatitis B virus infection: current status. *American Journal of Medicine* 2005; 118: 1415-1422

Okamoto H, Tsuda F, Akahane Y, Sugai Y, Yoshida M, Moriyama K, Tanaka T, Miyakawa Y, Mayumi M. Hepatitis B virus with mutations in the core promoter for an e antigen-negative phenotype in carriers with antibody to e antigen. *Journal of Virology* 1994; 68: 8102-8110

Parekh S, Zoulim F, Ahn SH, Tsai A, Li J, Kawai S, Khan N, Trepo C, Wands J, Tong S. Genome replication, virion secretion and e antigen expression of naturally occurring hepatitis B virus core promoter mutants. *Journal of Virology* 2003; 77: 6601-6612

Park OY, Jin YH, Lee M, Shin HJ, Kim HI, Cho H, Yun CW, Youn JK, Park S. Characterisation and gene cloning of monoclonal antibody specific for the hepatitis B virus X protein. *Hybridoma* 2000; 19: 73-80

Pungpapong S, Kim WB, Poterucha JJ. Natural history of hepatitis B virus infection: an update for clinicians. *Mayo Clinic Proceedings* 2007; 82: 967-975

Revill P, Locarnini S. HBV variants: clinical significance and public health implications. *Hepatitis B Annual* 2005; 2: 74-92

Ryan KJ, Ray CG. Chapter 37 In *Sherris Medical Microbiology*, 4<sup>th</sup> edition. McGraw-Hill 2004; 541-542

Sanger F, Nickler S, Coulson AR. DNA sequencing with chain termination inhibitors. *Biochemistry* 1977; 74: 5463-5467

Schaefer S. Hepatitis B virus taxonomy and hepatitis B virus genotypes. *World Journal of Gastroenterology* 2007; 13: 14-21

Seeger C, Mason WS. Hepatitis B virus Biology. *Microbiology and Molecular Biology Reviews* 2000; 64: 51-68

Shen T, Yuan XM, Zou YL, Gao JM, Dong H. Virologic characteristics of hepatitis B virus in patients infected via maternal-fetal transmission. *World Journal of Gastroenterology* 2008; 14: 5674-5682

Shephard CW, Simard EP, Finelli L, Fiore AE, Bell BP. Hepatitis B virus infection: Epidemiology and vaccination. *Epidemiology Reviews* 2006; 28: 112-125

Shinkai N, Tadaka Y, Ito K, Mukaide M, Hasegawa I, Asahina Y, Izumi N, Yatsunami H, Orito E, Joh T, Muzikami M. Influence of hepatitis B virus X and core promoter mutations on hepatocellular carcinoma among patients infected with subtype C2. *Journal of Clinical Microbiology* 2007; 45: 3191-3197

Spearman CW. The laboratory diagnosis of acute viral hepatitis. *South African Medical Journal* 1994; 84: 556-559

Takada S, Koike K. Trans-activation function of 3' truncated X gene-cell fusion product from integrated hepatitis B virus DNA in chronic hepatitis tissues. *Medical Sciences* 1990; 87: 5628-5636

Takahashi K, Akahane Y, Hino K, Ohta Y, Mishiro S. Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: comparative analysis of 40 full-length isolates. *Archives of Virology* 1998; 143: 2313-2362

Takahashi K, Ohta Y, Kanai K, Akahane Y, Iwasa Y, Hino K, Ohno N, Yoshizawa H, Mishiro S. Clinical implications of mutations C to T<sup>1653</sup> and T to C/A/G<sup>1753</sup> of hepatitis B virus genotype C genome in chronic liver disease. *Archives of Virology* 1999; 144: 1299-1308

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research* 1994; 22: 4673-4680.

Tong S, Kim KH, Chante C, Wands J, Li J. Hepatitis B virus e antigen variants. *International Journal of Medical Sciences* 2005; 2: 2-7

Torresi J. The virological and clinical significance of mutations in the overlapping envelope and polymerase genes of hepatitis B virus. *Journal of Clinical Virology* 2002; 28: 112-125

Tsebe KV, Burnett RJ, Hlungwani NP, Sibara MM, Venter PA, Mphahlele MJ. The first five years of Universal hepatitis B vaccination in South Africa: evidence for elimination of HBsAg carriage under 5-year olds. *Vaccine* 2001; 19: 3919-3926

Utama A, Purwantoma S, Siburian MD, Dhenni R, Gani RA, Hasan I, Sanityoso A, Miskad UA, Akil F, Yusuf I, Achwan A, Soemohardio S, Lelosutan SAR, Mantamala R, Lukito B, Badihisodo U, Lesmana LA, Salaiman A, Tai S. Hepatitis B virus subgenotypes and basal core promoter mutations in Indonesia. *World Journal of Gastroenterology* 2009; 15: 4028-4236

Valsakamis A. Molecular testing in the diagnosis and management of chronic hepatitis B. *Clinical Microbiology Reviews* 2007; 20: 426-439

Villeneuve JP. The natural history of chronic hepatitis B virus infection. *Journal of Clinical Virology* 2005; 34: 139-142

Violante-Dehesa M, Nunez-Naterans R. Epidemiology of Hepatitis B Virus B and C. *Archives of Medical Research* 2007; 38: 606-611

Wang XW, Hussain SP, Huo TI, Wu CG, Forgues M, Hofseth LJ, Brechot C, Harris CC. Molecular pathogenesis of human hepatocellular carcinoma. *Toxicology* 2002; 181: 43-47

Weber B. Genetic variability of the S gene of hepatitis B virus: Clinical and diagnostic impact. *Journal of Clinical Virology* 2005; 32: 102-112

Wright TL. Introduction to chronic hepatitis B infection. *American Journal of Clinical Virology* 2006; 101: 1-6

Xu XW, Chen YG. Current therapy with nucleoside/nucleotide analogs for patients with chronic hepatitis. *Hepatobiliary Pancreatic Disease International* 2006; 5: 350-359

Yeh CT, Shen CH, Tai DI, Chu CM, Lai WYF. Identification and characterization of a prevalent hepatitis B virus protein X protein mutant in Taiwanese patients with hepatocellular carcinoma. *Oncogenes* 2000; 19: 5213-5220

Yokosuka O, Arai M. Molecular Biology of hepatitis B virus: effect of nucleotide substitutions on the clinical features of chronic hepatitis B. *Medical Molecular Morphology* 2006; 39: 113-120

Yuan JM, Ambinder A, Fan Y, Gao YT, Yu MC, Groopman JD. Prospective evaluation of hepatitis B 1762T/1764A mutations of hepatocellular carcinoma

development in Shanghai, China. *Cancer Epidemiology Biomarkers Preview* 2009; 18: 590-594

Zanetti AR, Van Damme P, Shouval P. The global impact of vaccination against hepatitis B: A historical overview. *Vaccine* 2008; 26: 6266-6273

Zhang X, Zhang H, Ye L. Effects of Hepatitis B virus X protein on the development of liver cancer. *Journal of Laboratory Clinical Methods* 2005; 47: 59-64

Zhu R, Zhang HP, Vu H, Li H, Ling YQ, Hu XQ, Zhu HG. Hepatitis B virus mutations associated with *in situ* expression of hepatitis B core antigen, viral load and prognosis in chronic hepatitis B patients. *Pathology-Research and Practice* 2008; 204: 731-742