Cloning and Mutational Analysis of HBx Gene from the blood of HBV Infected Patients

By

Hijab Javed

NUST201463546MASAB92514F

Atta-ur-Rahman School of Applied Biosciences (ASAB)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

2016

Cloning and Mutational Analysis of HBx Gene from the blood of HBV Infected Patients

A thesis submitted in partial fulfillment of the requirement for the

Degree of Masters of Science

In

Healthcare Biotechnology

By

Hijab Javed

NUST201463546MASAB92514F

Atta-Ur-Rahman School of Applied Biosciences

National University of Sciences & Technology

Islamabad, Pakistan

2016

ACKNOWLEDGEMENTS

Now is the time when, after an intensive period of 1 year, my MS research comes to an end. It has been a tremendous learning experience not only in the scientific arena but also on personal level. I feel obliged to a lot of people whose support was indispensible during this time and I feel compelled to express my gratitude towards them.

First and foremost, I'm grateful to Allah Almighty, the Creator of the universe, without Whose will, nothing can happen. The power that He vested in me was the driving force throughout my degree. All regards to the Holy Prophet Hazrat Muhammad (SAW), the ultimate source of knowledge and righteousness for mankind.

I feel great pleasure to extend my thankfulness to my supervisor, **Dr. Hajra Sadia**, whose towering intelligence and unconditional guidance was something I could always count on. Not only is she dexterous in the competitive field of Biotechnology but also has a very humble nature. It was her competent supervision as well as benevolence that helped me carry out my research.

In addition, I would like to express my gratefulness to **Dr. Peter John**, Principal and Assisstant Professor, ASAB and **Dr. Attia Bhatti**, Assisstant Professor, ASAB, for keeping the environment conducive for research and study.

My acknowledgments would be incomplete without mentioning the unconditional support of my seniors, Wasifa Naushad and Zaira Rehman (PhD scholars) for enlightening my research experience and tolerating me through my never-ending queries and research-related concerns.

Last but not the least; words are not enough to thank my parents and sister for their unlimited love, encouragement, counseling and sympathetic ear. They have always been the wind beneath my wings, the reason why I believe in myself and my ultimate powerhouses.

Thank you all…..

Hijab Javed

Table of Contents

LIST OF ABBREVIATIONS

HCC	Hepatocellular Carcinoma
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
DNA	Deoxy Ribo Nucleic Acid
RNA	Ribo Nucleic Acid
cDNA	Complimentary DNA
NFκB	Nuclear Factor Kappa Beta
AP-1	Activator Protein-1
cAMP	Cyclic Adenosine Monophosphate
CREB	cAMP Responsive Element Binding Protein
ATF-2	Activating Transcription Factor 2
CDK	Cyclin-Dependent Kinase
CK1	Casein Kinase 1
MAPK	Mitogen-activated protein kinases
PI3K	Phosphatidylinositol-4, 5- bisphosphate 3 kinase
PKB	Protein Kinase B
WHO	World Health Organization
ORF	Open Reading Frame
cccDNA	Covalently Closed Circular DNA

SV40	Simian Virus 40
HSV tk	Herpes Simplex Virus thymidine kinase
LTR	Long Terminal Repeat
HIV	Human Immunodeficiency Virus
HTLV	Human T-lymphotropic Virus
MMTV	Mouse Mammary Tumor Virus
RSV	Respiratory Syncytial Virus
MHC	Major Histocompatibility Complex
INF	Interferon
IL	Interleukin
MAPK	Mitogen-Activated Protein Kinase
DISC	Death Inducing Signaling Complex
ELISA	Enzyme-Linked Immunosorbent Assay
TNFa	Tumor Necrosis Factor a
TGF	Tumor Growth Factor
TERT	Telomerase Reverse Transcriptase
GAS2	Growth Arrest Specific 2
HIF-1	Hypoxia Inducible Factor 1
EDTA	Ethylene Diamine Tetra Acetic Acid
$MgCl_2$	Magnesium Chloride
dNTPs	Deoxynucleotide Triphospahate

dATP	Deoxyadenosine 5'-triphosphate
dGTP	Deoxyguanine 5'-triphosphate
dTTP	Deoxythymidine 5'-triphosphate
dCTP	Deoxycytidine 5'-triphosphate
TAE	Tris Acetate EDTA
NCBI	National Centre for Biotechnology Information
BLAST	Basic Local Alignment Search Tool
E. coli	*Escherichia coli*
L.B	Luria Bertani
NaCl	Sodium Chloride
OD	Optical Density
$CaCl_2$	Calcium Chloride
IPTG	Isopropyl β-D-1-thiogalactopyranoside
X-gal	5-bromo-4-chloro-3-indolyl-β-D Galactopyranoside
UV	Ultra-Violet
Tm	Melting Temperature
T_A	Annealing Temperature
×g	Times gravity
Kb	Kilo base pairs
KDa	Kilo Dalton

List of Abbreviations

bp	base pairs
pM	PicoMolar
µL	Micolitre
mL	Millilitre
g	Grams
mg	Milligram
A	Amperes
%	Percentage
°C	Degree Celcius
~	Approximately

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Hepatocellular carcinoma has various etiological agents one of the main one being the chronic persistent infection of Hepatitis B virus. Various studies have speculated the oncogenic potential of its HBx gene and hence its role in development of HCC. The HBx induced carcinogenesis is a result of modulation of various signaling pathways and factors. It has been frequently reported that point mutations and truncations in HBx gene are significantly associated with its oncogenic role and transactivation function and these mutations accumulate in the HBV genome during its natural history of infection. These can be preferentially detected to determine the advancement of disease. This study was designed to investigate the variants of HBx gene amplified through PCR from HBV DNA isolated from 50 patients infected with HBV. The purified products were sequenced and aligned with consensus X sequence to determine the prevalent mutations at nucleotide and protein level. The results indicated higher frequency of mutations at the 3' and 5' ends of the gene whereas the internal regions were found to be relatively conserved. Among other important changes, the most frequently occurring mutations included substitutions at 76T→C/A, 97T→C, 303A→T, 400C→T/G, 136C→T and 70G→A/T which resulted in frameshift mutation in X protein, hence disrupting the dimerization domain of protein. Insertions 68_69insAA and 101_102insA were found to cause a frameshift in the respective X protein, hence likely to disrupt its 14-3-3 binding motif. Another deletion 415_416delCA, observed in 4 sequences, caused a frameshift mutation thereby affecting the p53 binding domain of X protein.

INTRODUCTION

Hepatocellular carcinoma (HCC) is among the primary causes of cancer-linked deaths and accounts for 70-85% of total liver cancer around the world (Yi-Fang *et al*., 2011), hence classifying as a major public health concern. The estimated number is as high as 600,000 deaths per year globally (Tetsuhiro *et al*., 2015). With respect to gender, liver cancer is accounted for the fifth most frequently diagnosed cancer in men whereas the sixth leading cause of cancer-related death in women worldwide (Yi-Fang *et al*., 2001; Wang *et al*., 2013). Maximum incidences of liver cancer are reported in East and South-East Asia and in Central and Western Africa and >80% of the cases occur in developing countries (Yi-Fang *et al*., 2011).

The carcinogenesis in HCC is a multi-factorial and multi-step process. Various etiological factors contribute towards the development of HCC including Hepatitis B virus (HBV), hepatitis C virus (HCV), alcohol consumption, non-alcoholic fatty liver disease (Tetsuhiro *et al*., 2015), ingestion of Aflatoxin B1 and other factors related with chronic inflammatory and hepatic degenerative changes (Hsu *et al.,* 1993). There exist 6 types of hepatitis viruses that lead to liver inflammation: A, B, C, D, E and G among which type A, B and C are the most common. Chronic infection by HBV is the major contributor to majority of HCC cases i.e. >75% (Yi-Fang *et al*., 2011).

Hepatitis B Virus is responsible for 60% of total liver cancer prevailing in developing countries while 23% of it prevailing in developed countries, as opposed to HCV infection which causes 33% and 20% of the total liver cancer in developing and developed countries, respectively (Yi-Fang *et al*., 2011). The geographical distribution of viral cases of Hepatitis (by virus type) occurring in Pakistan is depicted in Figure 1.1. Moreover, HBV infection has been reported to cause about one-third of total liver cirrhosis and greater than three-fourth of total HCC cases around the globe (Yi-Fang *et al*., 2011). It is transmitted through blood and infected body fluid but the precise mechanisms that cause the malignant cellular transformation are still undefined.

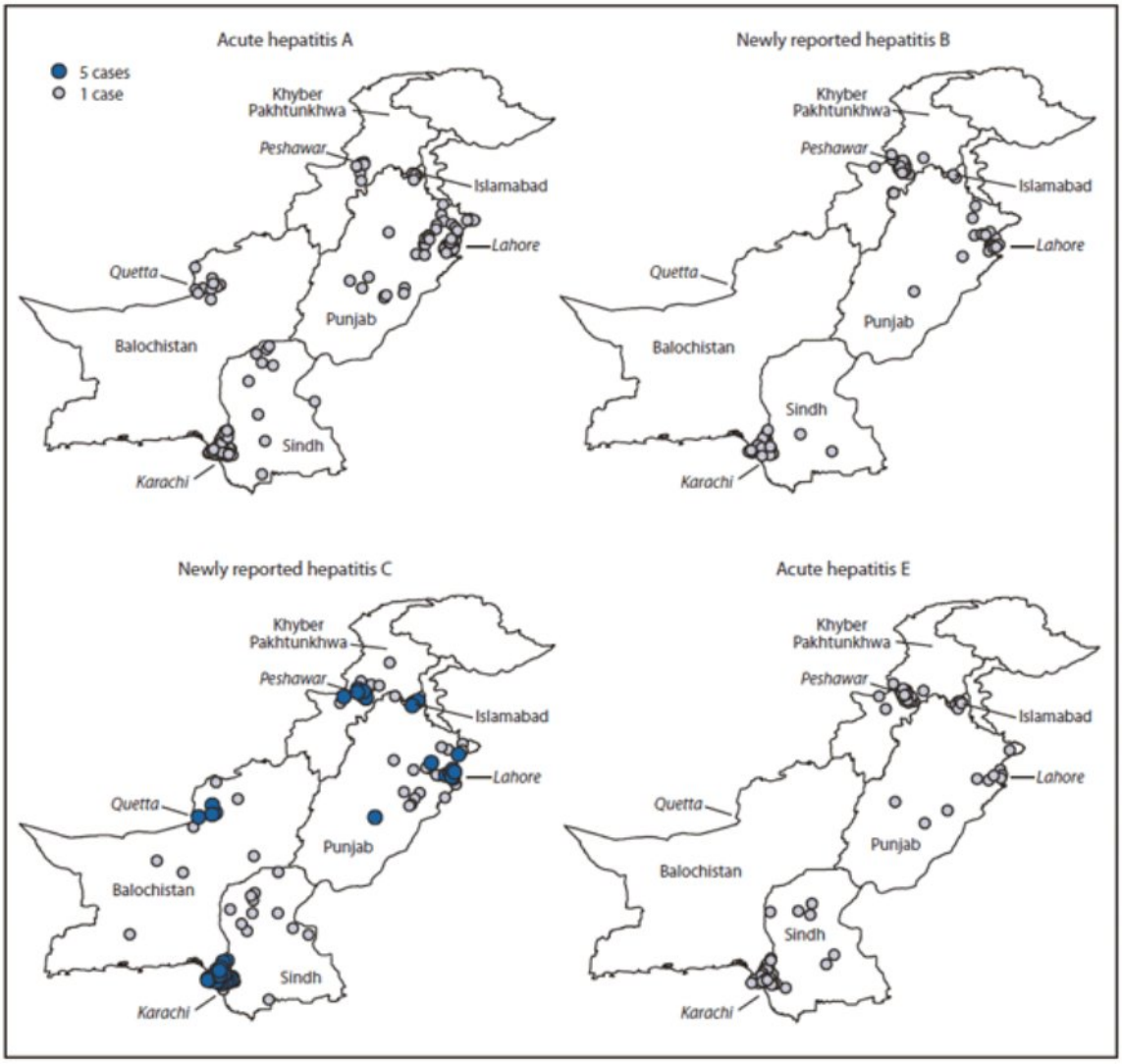


Figure 1.1: Geographical distribution of reported viral hepatitis cases (by virus type) in Pakistan (in major cities) during June 2010-March 2011 (Morbidity and mortality Weekly Report, 2011)

HBV belongs to a group of enveloped hepatotropic viruses that comprise of a circular, partially double stranded DNA genome with a size of 3200 base pairs (3.2 Kb). The replication cycle of HBV genome involves the conversion of DNA into RNA pregenome and then reverse- transcription of this pregenome (Renner M., 1995). HBV has been classified as a key risk factor leading to one of the most malignant tumors globally i.e. HCC (El-Serag, 2004). It translates four proteins out of which X-gene

product (HBx) has been extensively linked with HCC for its oncogenic potential (Koike, 1995). HBx is widely reported to play a role in the malignant transformation of liver cells having chronic HBV infection, which can be attributed to the prevalence of HBx coding sequences in more than 2/3rd of the HBV DNA cloned from genomically integrated sequences of liver carcinoma cells (Renner *et al*., 1995). Most of the time the X sequences are found to be truncated at the 3' end.

Various researches have implicated the contribution of HBx in the infliction of HCC, however, not much is known about the underlying mechanism. HBx gene encodes a transactivating factor/ a multi-functional trans-activator protein with a size of 16.5 kDa and 154 amino acids. It stimulates transcription of various viral as well as cellular genes and tends to perturb the cell proliferation and cell apoptotic pathways. X gene sequences that encode functional HBx are frequently found to be accumulated in serum of chronically infected HBV patients or integrated into the genome of liver carcinoma cells (Renner *et al*., 1995).

It also plays a role in transcription of HBV genome as well as transactivation of a broad spectrum of cellular genes via host-specific transcription factors for instance; HBx interrupts the intra-hepatic cytokine network that ultimately cause liver inflammation (Mi-Ock *et al*., 2002). It has been demonstrated that HBx accelerates the development of HCC through multiple processes that involve apoptosis, cell growth and proliferation, inflammation, angiogenesis, immune responses and multi-drug resistance (Liu *et al*., 2010; Neuveut *et al*., 2010; Ng and Lee, 2011). On the whole, as indicated in Figure 1.2, substantial amount of literature establishes the critical role of HBx in HCC development and progress via modulation of various processes.

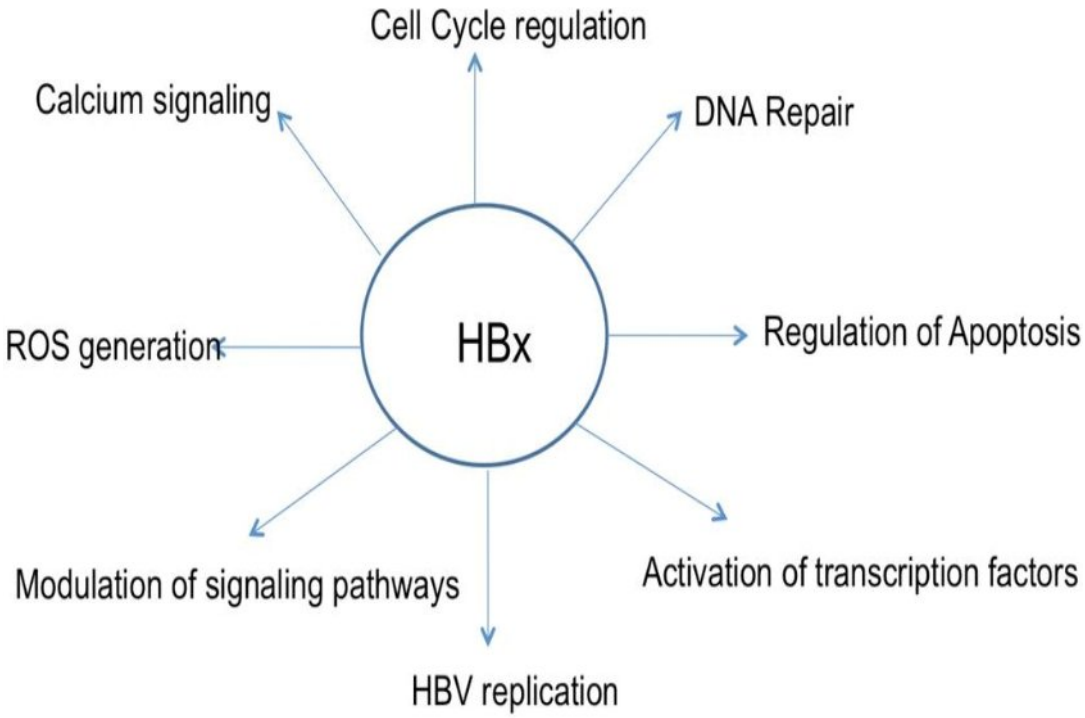


Figure 1.2: Various cellular pathways HBx is known to modulate. Most of these pathways are interconnected hence only few upstream initiating events via HBx expression might invoke a multifaceted response in hepatocytes (Rawat *et al*., 2012).

The HBx induced carcinogenesis is a result of modulation of various signaling pathways and factors, as indicated by previous studies and illustrated in Figure 1.3, such as activation of various transcription factors e.g. NF-κB (Lucito and Schneider, 1992), activator protein 1 (AP-1) (Kekule *et al*., 1993), cAMP-responsive element binding protein/ activating transcription factor 2 (CREB/ATF-2) (Maguire *et al*., 1991), cooperating with cellular oncogenes like Ras and Src (Doria, 1995), regulation of cell apoptosis and cell cycle by interacting with caspases, CDK, CK1 and surviving (Benn and Schneider, 1995; Gottlob *et al*., 1998) and activation of cell signaling pathways such as the Wnt (Ding *et al*., 2005), Ras/MAPK (Benn and Schneider, 1994) and PI3K-Akt/PKB pathway (Lee *et al*., 2001). Recent molecular researches have indicated the occurrence of HBx-induced mutations in proto-oncogenes and tumor suppressor genes such as p53 (Hsu *et al*., 1993).

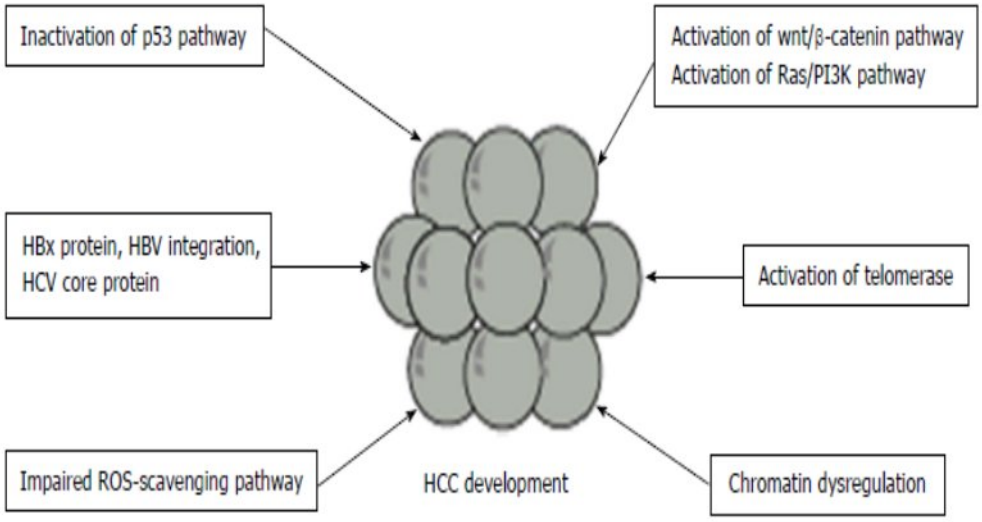


Figure 1.3: Major pathways responsible for the development of hepatocellular carcinoma (Tetsuhiro, 2015)

Early diagnosis of HCC is a rare possibility as patients mostly present late and by that time serious liver cirrhosis has occurred. Only 30% of freshly diagnosed liver cancer cases are detected in earlier stages when it is possible to remove the tumor surgically or go for a liver transplant, with a 5-year survival rate ranging from 50 to 70% (Wang *et al*., 2013). Even in such cases, 60 to 100% of the patients suffer from HCC recurrence and consequently limiting their long-term survival (Deng-Fu *et al*., 2007). Unfortunately, the reported mortality to incidence ratio is close to 1 (Yi-Fang *et al*., 2011; Deng-Fu *et al*., 2007). This figure necessitates the critical identification of some risk factors or prognostic/ diagnostic markers that can be applied clinically in order to predict the advancement of disease in HBV-infected individuals.

Moreover, elucidating the viral characteristics of HBV such as the role of its oncogene HBx in occurrence, recurrence and metastasis of HCC is invaluable to tailor appropriate treatment regimens or novel and guided anti-viral therapies (Tetsuhiro *et al*., 2015). It is reported to acquire various truncations, deletions and point mutations during the course of HBV infection which are likely to be directly associated with tumorogenesis.

Determination of the potential involvement of HBx in the onset of HCC is possible by defining its transactivating domains and identifying its various gene-mutants along with their correlation to its functional properties. Hence, this study aims at analyzing 13 variants of the gene amplified from the HBV genome isolated from HBV infected patients at different stages of chronic infection. To better comprehend the pleiotropic role of HBx, it is essential to determine the various mutations and truncations that prevail in HBx gene and their implications on the various activities that this gene assumes in the host's body.

Moreover, it would be interesting to determine if these mutations possess any diagnostic or clinical value or could potentially serve as significant molecular biomarkers for monitoring the clinical outcomes of patients having chronic HBV infection. Furthermore, if significant mutations are determined, they can be applied for active surveillance of HBV infected subjects to track the probability of developing cancer. The current study is streamlined to cater the very aspect of HBx point mutations which render it using alternative pathways in host's body, to execute its transactivation function.

LITERATURE REVIEW

2.1 HBV as Etiological Agent for Hepatocellular Carcinoma

Hepatocellular carcinoma is a major global health issue and ranks fifth in terms of its incidence among all types of tumors. Moreover, liver cancer is the third leading cause of cancer-mediated deaths per year and is associated with the highest fatality ratio per annum for any tumor. On a global scale, an estimated two billion people are infected with HBV whereas about 350 million are inflicted with chronic infection. Approximately 25% of these chronically infected people tend to develop the hepatocellular carcinoma (WHO, 2008). Various etiological factors are responsible for causing HCC; however, chronic hepatitis B virus infection is the most predominant cause worldwide, accounting for 55% of all global HCC cases (Kew, 2010). The following figure 2.1 illustrates the prevalence of HBV infection globally with reference to the level of chronicity.

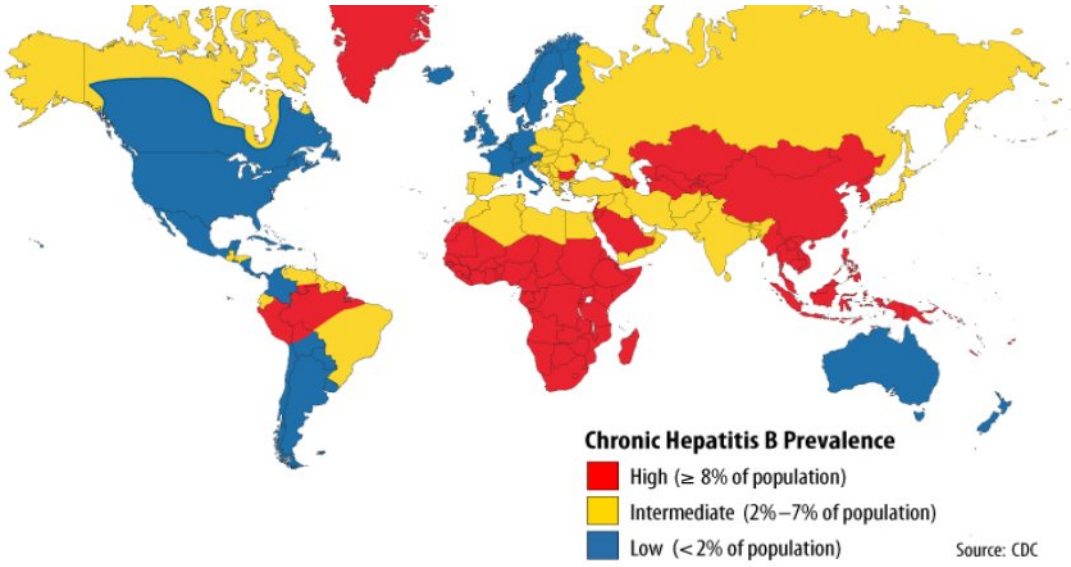


Figure 2.1: World Map showing patterns of chronic HBV infection with countries having highest to intermediate to lowest proportion of population infected (Retrieved from: http://hepbunited.org/hep-b-facts)

Patients suffering from HBV-induced HCC are, in majority of the cases, found to have HBV DNA integrated into the chromosomes of their hepatocytes, the factor which is supposed to play the major role in viral pathogenesis which is a multifactorial process on the whole (Murakami, 2001). Despite having evidence of strong correlation between chronic HBV infection and HCC progression, the exact events are not fully defined. It is suggested that the development of HCC involves the modulation of cell signal transduction pathways by HBV-produced proteins, especially the X protein. One of the most common causes of liver injury include the destruction of HBV-infected liver cells via immune system of the host that further induces repeated cycles of liver regeneration (Gearhart and Bouchard, 2010).

2.2 Characteristics of HBV and its genome

Hepatitis B virus is the prototype member of the *Hepadnaviridae*, a family composed of hepatotropic viruses, and hence it mainly targets the liver and can result in acute or chronic infection (Seeger *et al*., 2007). Viruses of similar type are also found in various species of animals e.g. woodchuks, ground squirrels and ducks which all represent the same family (Kumar *et al*., 1996). Viruses of this nature tend to translate proteins that modify normal cellular activities so as to create a facilitative environment for self-replication. This modulation of cellular signaling pathways can lead to cellular transformation and carcinogenesis (Gearhart and Bouchard, 2010).

Hepatitis B virion is made up of a nucleocapsid surrounded by envelope proteins. Inside the nucleocapsid lie the partially double stranded, circular DNA and a DNA polymerase enzyme (Seegar and Zoulim, 2007). The DNA is composed of four overlapping open reading frames (ORFs) that translate four proteins: viral envelope (S), capsid (C), reverse transcriptase (P) and X, out of which the X-gene product (HBx), expressed from the smallest ORF, has been attributed with oncogenic capacity (Wang *et al*., 2012) and is the major contributor in viral pathogenesis. Moreover, during the entire viral life cycle of HBV shown in Figure 2.2, HBV DNA can assume multiple replicative intermediates such as circulation in patient's blood as virion DNA, cytoplasmic core DNA form and presence in hepatocytes as episomal cccDNA, relaxed circular DNA,

linear DNA or deproteinized dsDNA or integrated linear DNA (Jain *et al*., 2015). The genome of HBV with its various components is shown in Figure 2.3

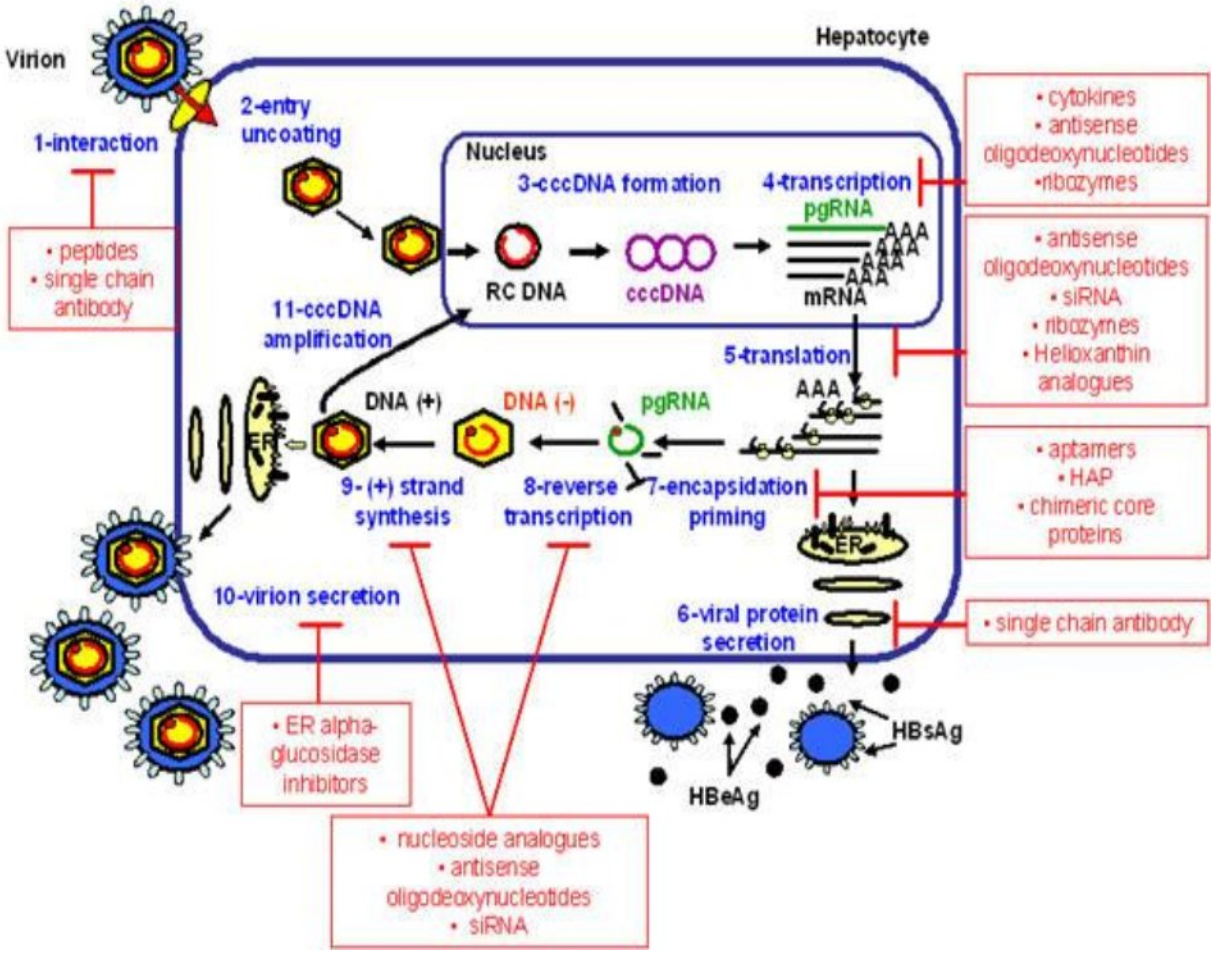


Figure 2.2: Sequence of events taking place in HBV life cycle: Entry in liver cells→ Localization of HBV DNA into nucleus→ replication→ assembling→ excretion of mature virions into extracellular space. The molecular targets of antiviral therapy are also indicated. (**Abbreviations:** *cccDNA, covalently closed circular double-stranded DNA; HAP, heteroaryldihydropyrimidines; HBeAg, hepatitis B e antigen; HBsAg, hepatitis B surface antigen; pgRNA, pre-genomic RNA; siRNA, short interfering RNA)*

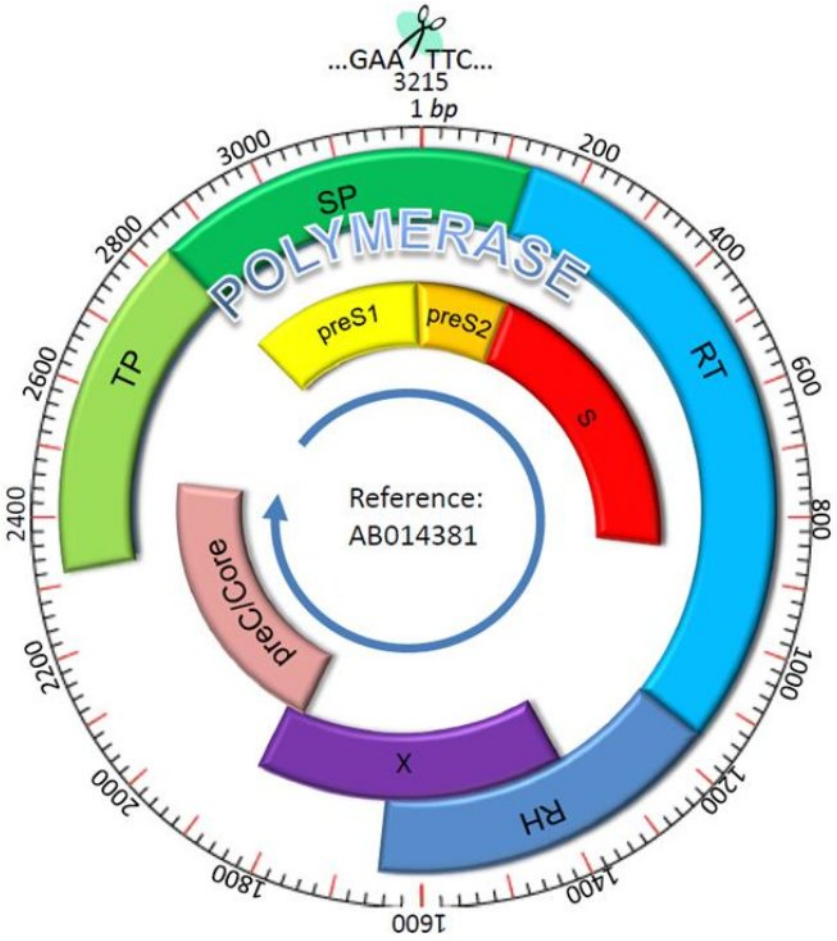


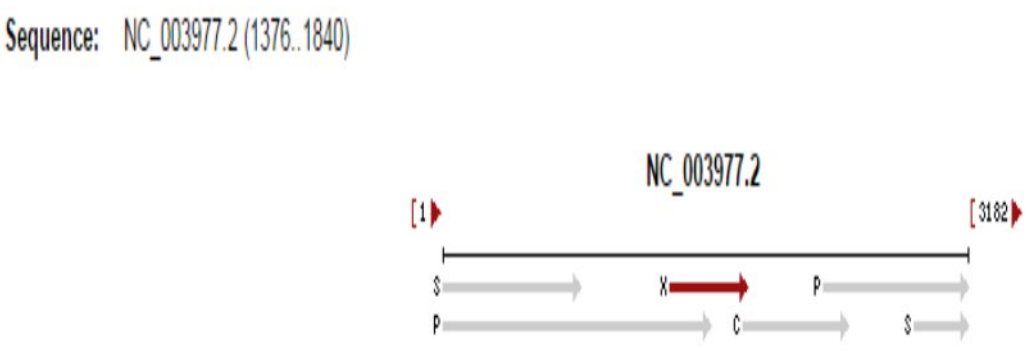
Figure 2.3: HBV genome and location of its four ORFs: (1) Polymerase: 2307..1623 (2). PreS1/S2/S: 2848..3215, 1..835 (3) preS2/S: 3205..3215,1..835 (4) S: 155..835 (5) preC/C: 1814..2452 (6) C: 1901..2452 (Retreived from: http://lifecenter.sgst.cn/biohbv)

There exist almost 10 genotypes of HBV which are responsible for the complexities that arise at the time of diagnosis owing to either different genotypes coinfecting at the same time or recombination events among different genotypes. HBV exhibits a high mutation rate because the reverse transcriptase it uses for viral DNA replication is prone to high error rate which consequently leads to distribution of quasispecies in the infected person (Jain *et al*., 2015).

2.3 Properties of HBx gene/protein

HBx is the smallest gene of the four partially overlapping ORFs of HBV's genome, based on 465 nucleotides that encode a 16.5 kDa protein, composed of 154 amino acids and is said to be an oncogenic protein causing hepatocarcinogenesis upon chronic HBV infection. It is a transactivator molecule with multifunctional activities (Lee *et al*., 2002; Kew, 2011; Wang *et al*., 2013). It is implicated as 'pleitropic' because it can cause the stimulation of cis-elements, not only limited to HBV genome but also a large range of other viral promoters that include SV40 early, HSV-tk, LTRs of HIV type 1, HIV type 2, HTLV type 1, Mouse Mammary Tumor Virus and Respiratory Synctial Virus (Kumar *et al.,* 1996).

The name "X" was designated to this protein because its amino acid sequence exhibited no homology to any known protein at all. It is a highly conserved protein among various HBV subtypes (Feitelson and Lee, 2007). HBx influences pivotal processes like gene transcription, interacellular signaling pathways, cell cycle/ proliferation and cell apoptotic pathways (Qin *et al*., 2000; Lian *et al*., 2001; Qin and Tang, 2002). It mediates complex activities like transcription of the HBV genome and once inside the host, it performs functions such as protein to protein interactions, regulating the phosphorylation, stabilization of mRNA and shuttling of nucleo-cytoplasmic translocation (Diao *et al*., 2001). The locus of HBx gene in the HBV genome is 1376..1840 as depicted in the Figure 2.4.



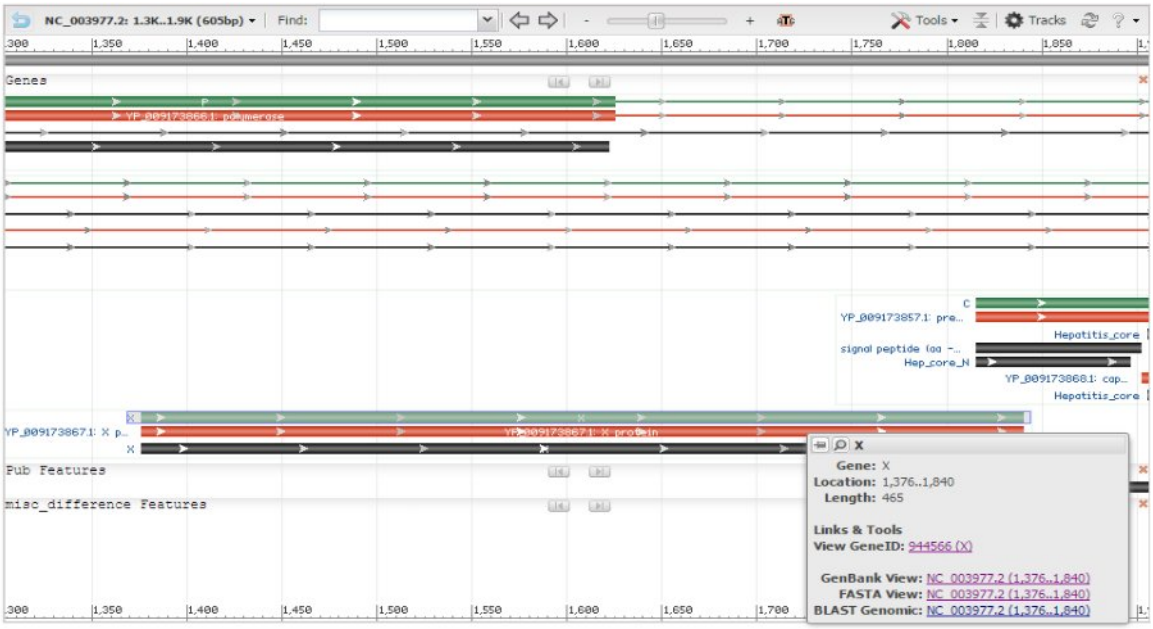


Figure 2.4: Genomic regions, Transcripts and Proteins in the HBV genome (Retrieved from: https://www.ncbi.nlm.nih.gov/gene/944566). The gene of interest 'HBx' resides at the locus from position 1376 to 1840 (465bp)

The role of HBx is indispensable in HBV infection as it promotes viral replication by various mechanisms including: inhibition of anti-HBV proteasomes (Zhang *et al*., 2004), enhancing transcription of virus-encoded RNAs (Gong *et al*., 2013), escalating autophagy pathways (Sir *et al*., 2010) and triggering Ca^{+2} dependent pathways (Bouchard *et al*., 2001). It acts as a promiscuous transactivator of viral and cellular promoters and enhancers (Feitelson and Lee, 2007). Owing to multi-factorial actions of HBx during the development of HCC in humans, it has emerged as a potential therapeutic target (Zhu *et al*., 2015).

The frequency of integration of X gene from the HBV DNA in the hepatocyte chromosome is reported to be the highest in case of HCC and HBxAg is said to always express in the integrated fragment even if no other viral proteins are present in the tumor cells (Wang *et al*., 2013) as it has been indicated by some studies that PCR detection has shown some HCC patients to be positive with HBx that are otherwise negative for HBsAg and HbeAg . Once present inside, HBx protein affects the expression of almost 39 genes (involved in regulation of cell cycle/ cell death and cell adhesion), as revealed

by recent cDNA microarray studies of HBx expression in HepG2 cell lines (Han *et al*., 2000). As represented by Figure 2.4, it is involved in enhancing or alleviating the expression of various genes. Intracellular localization analysis has also depicted that HBx is predominantly found in cytoplasm of tumor cells where it triggers the process of liver cancer (Wang *et al*., 2013). During the integration process, viral sequences are prone to rearrangements and losses, however, HBx is invariably always the part of integrated section in HCC patients and it tends to encode functionally active and viable protein even if it is truncated in the process (Kew, 2011).

Figure 2.5: Pathways and genes upregulated and downregulated by HBx protein during HBV infection (Ayub *et al*., 2013)

Several experiments conducted on transgenic mice carrying the HBx gene indicate the development of liver cancer or present accelerated neoplasmic characteristics (Kim *et al*., 1991). One of the proposed property of HBx is that it leads to accumulation of point mutations in the p53 gene (tumor suppressor gene) thereby rendering it inactive against hepatocyte transformation. In addition, HBxAg is reported to exert anti-apoptotic

effects and also accelerates the cell signaling transduction processes (Wang *et al.*, 2013). Hepatitis B X gene is implicated in the upregulation of expression of certain proto-oncogenes (c-myc, c-fos/c-jun), various cellular genes (Al-antitrypsin, A-fetoprotein, metallothionein, epidermal growth factor receptor, RNA polymerase II and III) and several other immune-system related genes (MHC complex I, MHC complex II, 3 INF and IL-8).

Since HBx reportedly doesn't bind to the DNA, it can be presumed that it mediates the transactivation through protein-protein interactions. No consensus is found among the regulatory regions of the HBx-responsive genes; however, most of them interact with cis-elements for some common trans-factors such as AP-1/2 or NFKB. Interestingly, HBx is capable of mimicking ATPase and protein kinase activities, thereby modulating cellular signal transduction pathways (Kumar *et al*., 1996).

2.4 Oncogenic potential of HBx and its mechanism of action

Despite multiple studies, the question of HBV's direct involvement in the multistep process of hepatocarcinogenesis still needs to be adequately answered. Majority of the investigations have found clonally integrated HBV-sequences in liver tissue, which in most cases, comprises of X-ORF which is maintained and expressed, thereby imparting the cells with a selective growth advantage (Sirma *et al*., 1999). Thus, of the proteins translated by HBV, HBx is speculated to play a critical part in the development and progression of HCC. Various researches tend to have established the oncogenic potential of HBx through putative mechanisms that contribute to HCC, however, the exact mechanism is still not established and there exists dearth of information regarding the precise relationship of HBx with cellular pathways/components that eventually lead to liver cancer. The mapping of X protein (Figure 2.6) indicates its various domains essential for its pleiotropic functions.

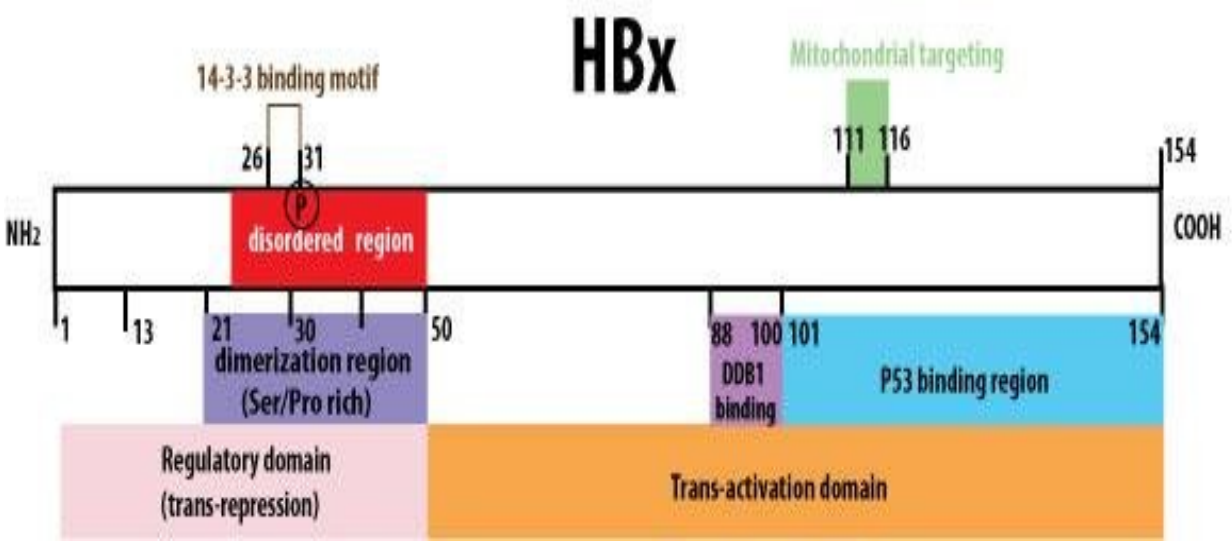


Figure 2.6: Illustration of various domains of HBx protein based on 154 amino acids. Amino acids 1 to 50 constitute the regulatory domain while from 50 to 154 constitute the trans-activation domain responsible for cellular transformation. Mutations affecting this region are note-worthy as they can have implication on the process of carcinogenesis. (Murakami *et al*., 1994; Li *et al*., 2009; Li *et al*., 2008; Diao *et al*., 2008; Elmore *et al*., 1997, Tang *et al*., 2005)

Hepatitis B virus X protein has been found to interact and interfere with a plethora of host proteins and pathways, respectively, so much so that a statement was put forward by Murakami (2001):

"If there is a protein that is not among the list of HBx interactors, then this protein has probably not yet been tested."

Following is the review of different studies that have investigated various mechanisms through which HBx transforms normal cells and contributes to oncogenesis. These may include stimulation of a number of transcription factors e.g. nuclear factor- κB (NF-κB) (Lucito and Schneider, 1992), activation of the activator protein 1 (AP-1) (Kekule *et al.,* 1993), enhancement of cAMP-responsive element binding protein or activating transcription factor 2 (CREB/ATF-2) (Maguire *et al*., 1991) and so on. Moreover, HBx has been reported to interact with cellular oncogenes like Ras and Src (Doria *et al*., 1995; Klein and Schneider, 1997) and regulate cell apoptosis and cell cycle

by affecting the caspases, CDK and CK1 (Gottlob *et al*., 1998; Benn and Schneider, 1995). HBx also stimulates various cell signaling pathways and is reported to perturb Wnt pathway (Dind *et al*., 2005), the Mitogen-activated protein kinase (MAPK) pathway (Benn and Schneider, 1995), Janus family tyrosine kinase (JAK) signal transducers and PI3K-Akt or PKB pathway (Lee *et al*., 2001), as indicated in Figure 2.7.

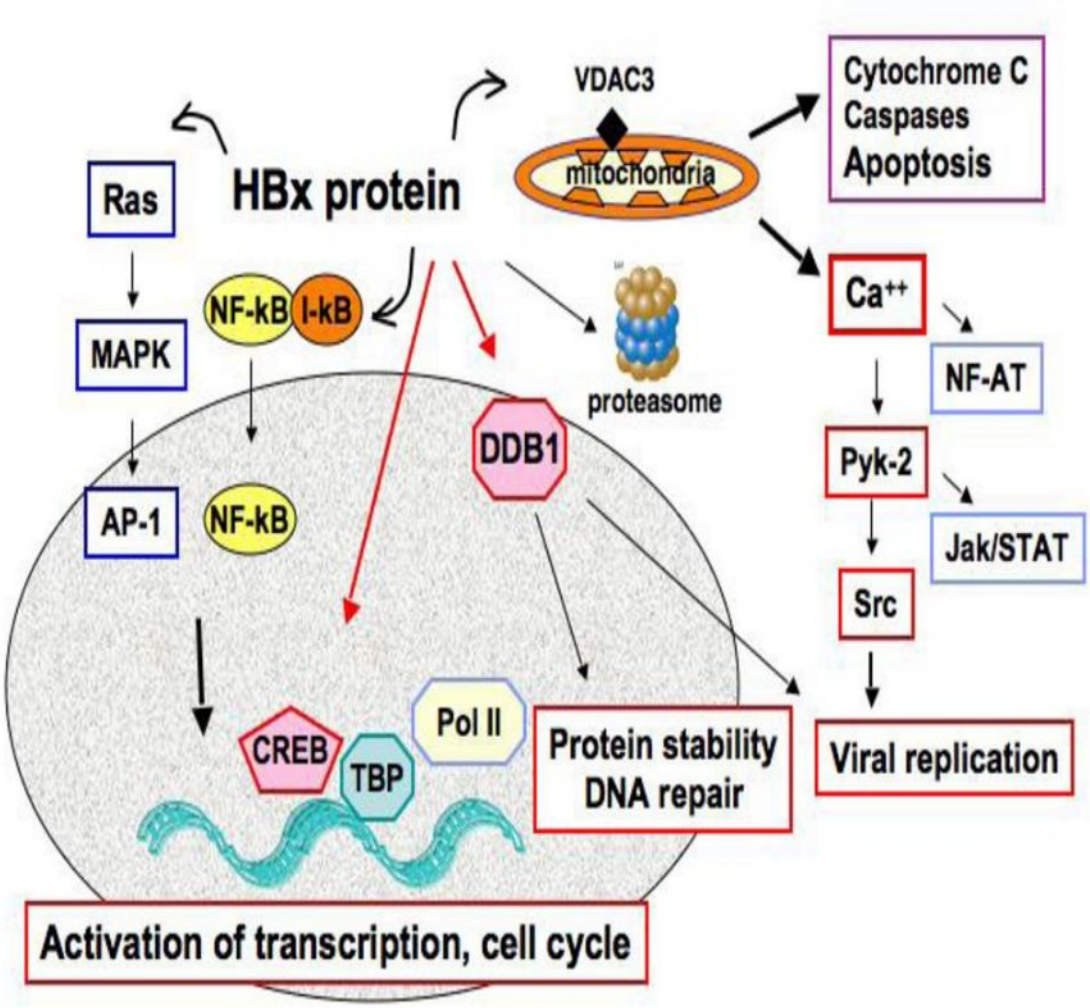


Figure 2.7: Large array of HBx functions that constitute its pleiotropic activities, eventually leading to cellular transformation via trigger of virus replication, modulation of cell cycle, stimulation of signaling pathways, perturbing of apoptosis and interference with DNA repair, while all HBx-interactions taking place in both cytoplasm and nucleus (Benhenda *et al*., 2009).

It has been proposed that HBx tends to perturb the Notch signaling pathway as it upregulates the transcriptional expressions of Notch-1, Jagged-1 and Hes-1 proteins by binding to the intracellular domain of Notch-1, hence enhancing the malignancy. Wang *et al*. (2012) investigated the similar relationship by transfecting HepG2 cells with HBx gene which led to increased biological activities in transfected HepG2 cells as compared to the normal ones. Blocking of Notch signaling pathway led to semi-reversal of some of the HBx-induced abnormalities and attenuated the growth of HBx/HepG2 cells by shortening S phase and inducing their apoptosis. Since Notch signaling pathway is said to be crucial for the modulation of organ development, cell proliferation and differentiation and apoptosis (Artavanis-Tsakonas *et al*., 1999), hence its disturbance by HBx can lead to liver tumorogenesis.

One of the main processes through which the HBx protein induces HCC is inhibition of apoptosis of tumor cells. Disregulation of apoptotic pathways can be through different mechanisms, for instance, Wang *et al.* (2013) demonstrated the effect of HBx on death receptor 'Fas' and its ligand 'FasL' which crosslink to form the 'death inducing signaling complex' (DISC) and cause the apoptotic death of the substrate by release of caspases. They explored the immunohistochemical expression of HBxAg, Fas and FasL in HCC tissues collected from 50 patients as well as the their serum levels using ELISA and concluded that sFas and sFasL both were found in significantly higher levels in HCC and cirrhosis patients as compared to controls ($p<0.01$) but HCC cells tend to become resistant to Fas mediated apoptosis.

Another speculated mechanism through which HBx leads to HCC is via the transcriptional induction of intrahepatic pro-inflammatory cytokine network that lead to acute and chronic liver inflammation and hence injury. These cytokines are also reported to play the role of potential biomarkers for the detection and diagnosis of hepatocellular carcinoma. According to one study, Northern and Western blotting analysis were carried out to check IL-18 expression in HBV integrated liver cancer cell lines and flow cytometry analysis was carried to check expression of FasL and it was suggested that HBx enhances the expression of IL-18 in liver cells which in turn upregulates the

expression of FasL that may be associated with hepatic injury during HBV infection (Lee *et al.,* 2002).

Substantial volume of data has recently supported the role of HBx in induction of pro-inflammatory cytokines. The hepatocytes of patients chronically infected with HBV as well as hepatoma cell lines transfected with either HBV genome or HBx expression cassette, are both reported to produce the tumor necrosis factor-a (TNFa) (Gonzalez-Amaro *et al*., 1994). Similarly, in vitro experiments have depicted the transcriptional activation of IL-6 as well as IL-8 (neutrophil and T-cell chemotactic cytokine) by HBx which are also found in the serum of chronically infected HBV patients and cause liver inflammation. Moreover, Yoo *et al*. (1996) suggested the role of HBx in inducing tumor growth factor-13 (TGF-13) which triggers the production of extracellular matrix proteins in liver cells at the cirrhotic stage.

Apart from this, researches have also elucidated the implications of HBx on proliferative pathways in cells in case of established or immortalized cell lines, expressing the X protein in the absence of HBV replication, however, when cultured primary rat hepatocytes were used as the biological model to investigate HBx effect, it was demonstrated that Hbx modulates the amounts as well as activities of cell-cycle regulatory proteins which consequently cause the normally quiescent hepatocytes to step into the G1 phase of cell cycle and this calcium-requiring action of HBx is in turn needed for HBV replication. Conclusively, HBx is associated with regulation of cell proliferation as well as cytosolic calcium signaling required to stimulate its own replication (Gearhart and Bouchard, 2010).

Recently, a good amount of studies have evidently demonstrated the role of HBx in epigenetic modifications i.e. altering methylation status of host DNA, as one of the mechanisms of HBV-induced malignancy. Hepatitis B virus X gene interacts and activates the transcriptional expression of methyl transferases which leads to hypermethylation of specific regions of DNA or global hypomethylation, consequently leading to silencing of tumor suppressor genes and chromosomal instability, respectively. It has also been suggested that HBx protein enhances the expression of Telomerase

Reverse Transcriptase (TERT) and telomerase action, thereby increasing the life span of liver cells and accelerating the carcinogenesis (Kew *et al*., 2011).

HCC tissues contain predominant expression of HBx variant that has 35 amino acids deleted at the C-terminal (HBx?35). PCR validation and genome wide scanning tests on immortalized hepatocyte lines reveal that this mutant mainly targets the expression of growth arrest specific 2 (GAS2), which is a tumor suppressor gene, by binding to its promoter and attenuating it so as to promote the hepatocyte proliferation. Zhu *et al,*. (2015) demonstrated that the down regulation of GAS2 in HCC samples as compared to normal tissue, rendering the neoplastic hepatocytes with survival advantage.

It is widely known that poorly oxygenated microenvironment (hypoxia) promotes aberrant vessel formation and proliferation and the key mechanism via which it enhances the cancer phenotype is the stimulation of HIF-1 factor. Its expression is reported to increase in various tumor types e.g. bladder, breast, ovarian, liver, pancreatic, prostrate and renal tumors (Harris, 2002; Semenza, 2003). One such research looked into the interaction between HBx and HIF-1a. Wild type HBx has already been demonstrated to activate HIF-1a; however, impact of changes in HBx and their effect on HIF-1a were studied. HBx mutants from HCC 101 tissues were sequenced, cloned and transfected into HCC cell, followed by western blot and luciferase assay analysis of HIF-1a expression and activation. The findings suggested that the dual mutations in HBx i.e. K130M and V131I enhanced its activity as they enhanced the transcription of HIF-1a, indicating that different mutations in HBx cause differential conseqences on the activity of HIF-1a. It was also revealed that C-terminus was significant for the stability and transactivation function of HBx since truncation or deletion in this region weakened its potential to enhance HIF-1a (Liu *et al*., 2014). All these and other multiple interactions of HBx are illustrated in Figure 2.6 below:

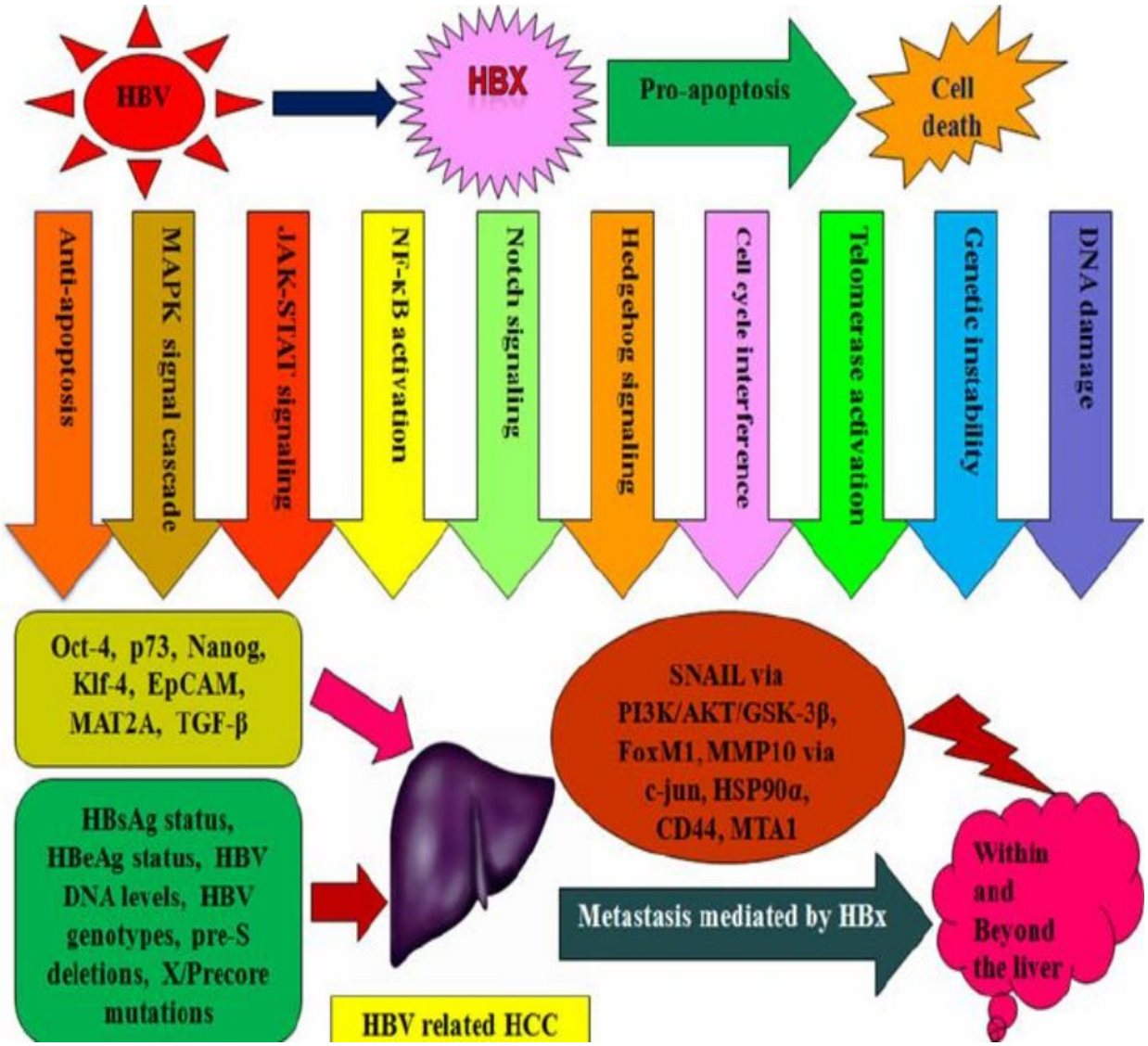


Figure 2.8: A schematic diagram depicting the 'ten-fold oncogenetic cascade' showing functional characterization of HBx in the development and metastasis of HCC.

2.5 Mutational Analysis of HBx Gene

It has also been reported that apart from wild type, different mutants of HBx have been found in specimens of HCC. According to Liu *et al*. (2014), naturally prevalent aberrations of HBx include point mutations (approximately 50% cases), deletions (less than 3% cases) and distal truncations at C-terminal (more than 30% of the cases). Out of

all variants of HBx, the truncated mutants are of special interest owing to their unique functional features. Research has revealed that the carboxyl terminal peptide of HBx is crucial for its transforming potential, anti-proliferative activity and transactivation effects of HBx (Tu *et al.*, 2001; Ma *et al.*, 2008). Various functional domains of HBx gene and their prevalent mutations are given in Figure 2.9.

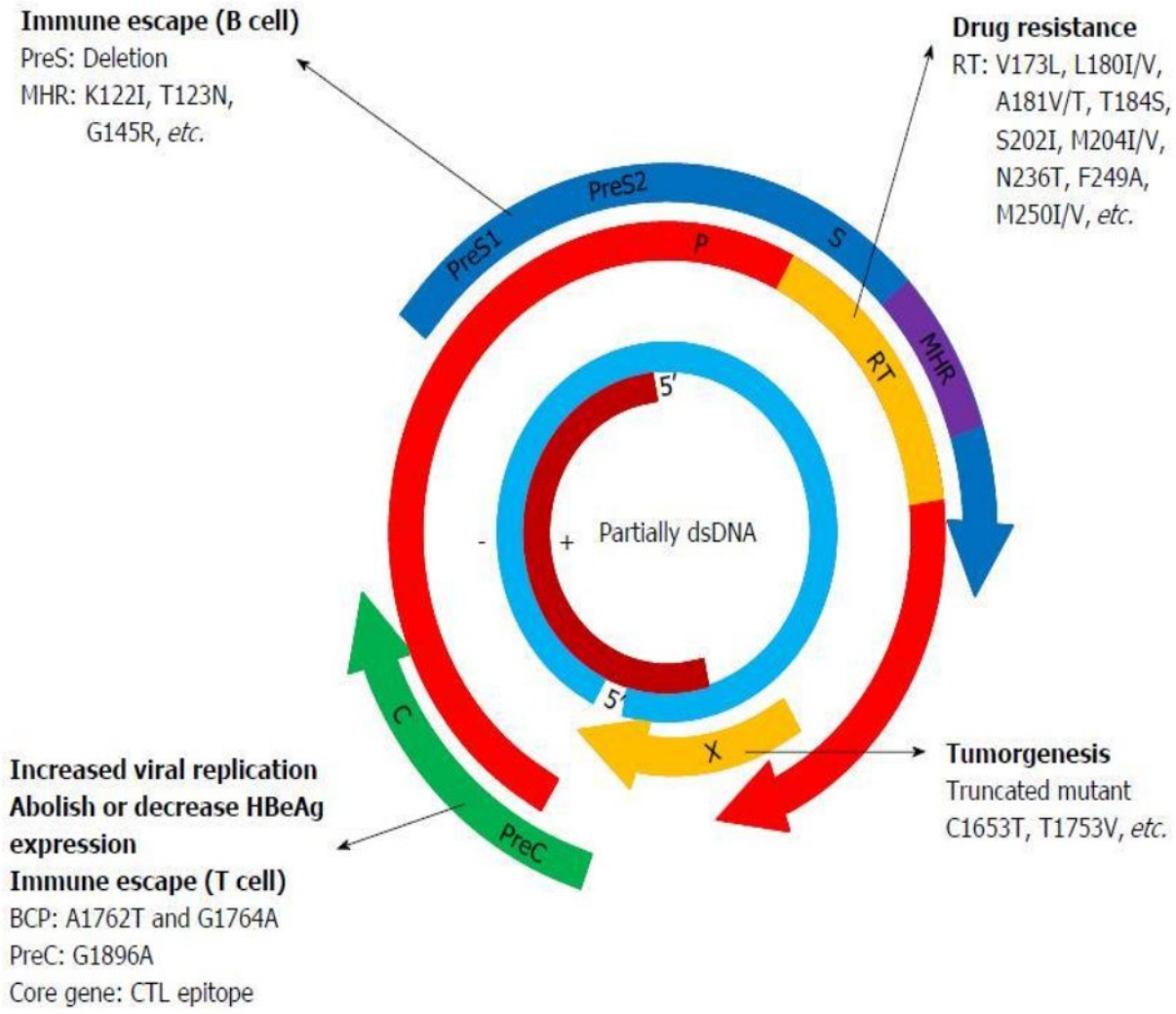


Figure 2.9: Illustration of HBV genome and predominantly occurring mutations arising in its ORFs: the 3.2Kb ds-DNA is organized into 4 ORFs (the polymerase, the envelope, the precore and X). The deletion mutations in the PreS gene and few point mutations in the major hydrophilic region (MHR) of S gene can result in immune escape and occult HBV infection. Mutations in reverse transcriptase (RT) can result in drug resistance e.g. A181T/V mutations in the RT region can cause stop codon mutation. A1762T and

G1764A mutations in the base core promoter or G1896 in PreC can accelerate viral replication and abrogate or alleviate HBeAg expression. Few mutations in the CTL epitope of HBV core gene can result in T cell immune escape whereas few point mutations or truncated mutants in the HBx gene can lead to tumor formation or end-stage liver disease (Zhang *et al*., 2016).

One study of such nature was conducted in China by Zhu *et al*. (2007) where particular HBx gene mutations were investigated in 51 paraffin-embedded HCC tumor tissue samples and 25 serum samples taken from HBV carrier patients in Southern China. Gene Polymorphism analysis via PCR and sequencing depicted that 31.3% of gene fragments taken from tissues showed a unique pattern of having a common deletion at nucleotide 382 to 400 along with 29 point mutations. This resulted in a frame-shift in the HBx open reading frame leading to a C-terminal truncated HBx polypeptide chain. In four of the samples, amplification of full-length HBx gene produced a shorter fragment which upon sequencing showed a 264 bp deletion between 1577 and 1840 nucleotides of the HBV gene. The results implicate the high frequency of HBx mutations in HCC samples or serum of HBV infected patients and the 382 to 400 nucleotide deletion plays a crucial role in the carcinogenesis.

In another study conducted by Renner M. (1995), gene plasmids containing full-length HBx sequences or HBx fragments were constructed and transfected into cells along with a plasmid containing reporter gene. The purpose was to identify the transactivating domain of HBx and it was demonstrated that amino acid 49 to 143 (a 95-amino acid fragment) contains full transactivating potential. Furthermore, a narrower fragment within this larger one i.e. 107 to 130 amino acids comprises of higher proportion of charged amino acids and demonstrated critical role in enhancing gene expression. Selective deletions were made in the fragments to elucidate that stable RNA and proteins were generated if at least codon 1-82 or 70-154 remained intact, indicating it to be the functional domain of HBx.

A similar research using insertion mutation instead of deletion was conducted. Oligonucleotide-directed insertion of 2 codons i.e. Arg and Pro was used to construct a panel of HBx mutants at 10 or 20 amino acid intervals along the full length of gene. They

were transfected in the HepG2 cells to elucidate the transactivation properties of HBx and the findings indicated the presence of two separate internal domains along the gene length, around 68th amino acid residue and the other between 110th and 139th amino acid, that were essential for its activity (Runkel *et al*., 1993).

Researchers have divided the 154 amino acid sequence of HBx into 6 regions (A to F) on the basis of sequence homology with X proteins of other mammalian hepadnaviruses. Among these 6, the highly conserved ones are A, C and E as they comprise of all four conserved cysteines at position C7, C61, C69 and C137. In an investigation carried out by Kumar et al. (1996), significant HBx regions of transactivation were identified by developing a panel of 10 deletion mutants (x5-x14) and 4 single point mutations (x1-x4) and expressions were determined by Western blotting. It was identified that mutation in the most conserved region A as well as non-conserved regions of B and F had no effect on transactivation. However, deletions of regions C and D led to major loss in function (~90%). It was concluded that residues 132 to 140 (conserved 9 amino acids) appeared to be critical for transactivation.

It is reported that HBx mutations tend to prevail and accumulate after prolonged HBV infection and these mutations modify the domains of the protein which cause its interactions with various other proteins and perturb alternative signaling pathways. These point mutations generally occur near the time when HBx sequences tend to integrate in the liver cells leading to carcinogenesis. Sirma *et al*. (1999), in an attempt to elucidate the mechanism of action of HBx sequences, cloned and characterized a set of naturally arising X-mutants derived from HCC patients and studied their impact on cell growth and viability. It was found that X constrains cellular clonal outgrowth and stimulates apoptosis via p53-independant pathway. The most significant finding was that the mutations in the HBx gene arising in HCC abrogated HBx-induced growth arrest and apoptosis. A set of mutants were engineered for the purpose of mapping the growth suppressive effect of HBx to domains reported to have transactivation activity, however, it was proposed that the inhibition of anti-proliferative and apoptotic activities of HBx by naturally occurring mutations may lead to transformation of hepatocytes making them susceptible to proliferation.

Considering the aforementioned background, it seems plausible to amplify X-mutant sequences, analyze their mutations and truncations via sequencing and bioinformatic tools and investigate the relation of these mutations in HBx with various other factors in host's body.

MATERIALS & METHODS

3.1 Sample collection

A total of 50 blood sera samples of HBV patients were collected: 30 from the Diagnostic Lab of Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST, Islamabad and 20 from Shalamar Hospital, Lahore and Ittefaq Hospital, Lahore. These patients were clinically tested to be HBs positive through polymerase chain reaction (PCR) and enzyme linked immunosorbent assay (ELISA). The blood was obtained in Ethylene Diamine Tetra Acetic Acid (EDTA) tubes. Ice packs were used for transporting blood sera samples from Lahore to Islamabad and were stored at 4◦C until further processing. To obtain the sera, the blood was shifted to 1.5 mL microfuges and centrifuged at 2500 ×g for 7 minutes. The separating top layer of serum was pipetted out and shifted to fresh microfuges.

3.2 Viral DNA extraction

Viral DNA was extracted from the sera samples using PureLink Viral RNA/DNA kit (Invitrogen, USA). According to the manufacturer's protocol, 25 µl of Proteinase K was taken in a sterile microfuge tube. 200 µl of cell-free sample (serum) was added to it, followed by addition of 200 µl of Lysis Buffer (containing 5.6 µg Carrier RNA). The tube lid was closed and contents were mixed by vortexing for 15 seconds. The tube was then incubated at 56 ºC for 15 minutes. 250 µl of 100% ethanol was then added to it and again mixed thoroughly by 15 seconds of vortexing. The lysate was incubated at room temperature for 5 minutes. For the purpose of purification, the above lysate was added to Viral Spin Column in a collection tube, which was then centrifuged at 6800× g for 1 minute. The collection tube was discarded and spin column was placed in a new Wash Tube. The column was washed with 500 µl of Wash Buffer (containing ethanol) and again centrifuged at 6800 ×g for 1 minute. The flow through was discarded and another round of washing using 500 µl of Wash Buffer was carried out. The column was then placed in a clean Wash Tube and centrifuged at maximum speed for 1 minute to ensure

the removal of any residual Wash Buffer. Next the spin column was placed in a clean 1.7-ml Recovery Tube and to the center of the cartridge, 50 µl of Elution Buffer (sterile RNase-free water) was added. It was incubated at room temperature for 1 minute, followed by centrifugation at maximum speed for 1 minute, in order to elute nucleic acids. The spin column was discarded and the purified viral DNA collected in Recovery Tube was stored at -20 ºC until desired downstream applications.

3.3 Determination of integrity of purified viral DNA

In order to confirm the presence of purified viral DNA, detection PCR was carried out. It was a Nested PCR based on 2 rounds of amplification and 2 pairs of primers.

3.3.1 HBV Detection Primers

Two pairs of primers were used for detection of HBV DNA (one pair in each round) which were taken from the Diagnostic Lab of Atta ur Rahman School of Biomedical Sciences (ASAB). The sequences of the primers are given in Table 3.1 below:

Table 3.1: The sequences of primers used for HBV detection by Nested PCR

Sr.No.	Primer Code	Sequence 5'→3'	Length
1	MTB5	CATCCTGGTGCTATGCCTCATCT	23
2	MTB6	GGTATGTTGCCCGTTTGTCTTCT	23
3	MTB7	GGCACTAGTAAACTGAGCCA	20
4	MTB8	CGAACCACTGAACAAATGGCACT	23

3.3.2 Nested Polymerase Chain Reaction

The Nested PCR was carried out in order to determine the presence and integrity of viral DNA in the purified samples. It consisted of two rounds, each using a 25 µl PCR reaction mixture containing 1X Taq Polymerase buffer, 2mM $MgCl_2$, 2mM dNTPs (dATP, dGTP, dTTP, dCTP), forward and reverse primers (10pM each) and thermostable Taq Polymerase (1U). The template used in first round was the purified viral DNA

(isolated by the method mentioned above) while for second round of PCR, using the second pair of internal primers, the PCR product from first round was used as the template. Nuclease free water was used to make up the volume to 25 µl. The PCR profiles used are given in Figure 3.1 and 3.2, respectively.

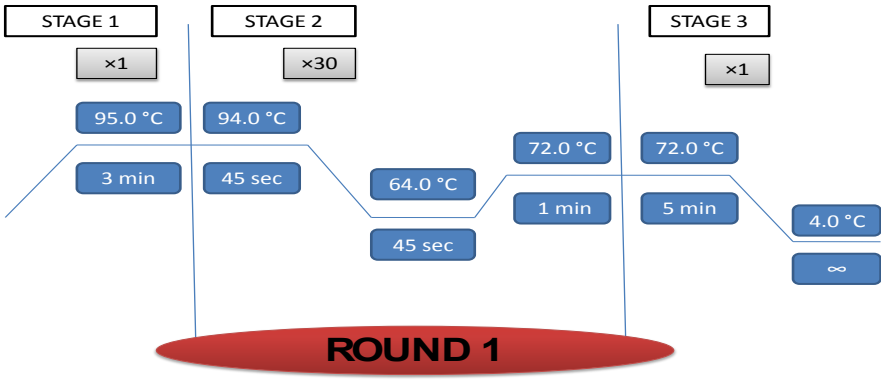


Figure 3.1: PCR profile for Round 1 of Nested PCR for HBV detection

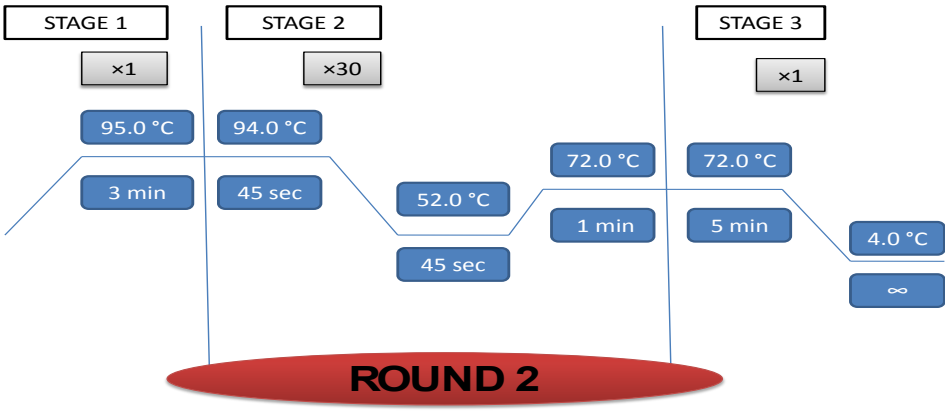


Figure 3.2: PCR profile for Round 2 of Nested PCR for HBV detection

Standard PCR protocol was employed to carry out the PCR reaction in thermocycler (Swift Maxi- Esco Swift MaxPro thermal cycler), in order to prevent contamination. DNA quality was checked by amplification of a ~200 bp fragment of HBV genome.

3.3.3 Detection by Gel Electrophoresis

The PCR product from the second round of Nested PCR was assessed using 2% agarose gel which was prepared in 1X TAE. The stock solution of TAE (50X) was diluted in distilled water in a ratio of 1:49 i.e. 20 µL of stock TAE dissolved in 980 µL of distilled water. 80 µL of this working solution of TAE was taken and 1.6 g of agarose was added to it which was then dissolved by heating in microwave. The gel mixture was then cooled and 5 µL of ethidium bromide (10mg/mL) was added for gel visualization. 5 µl of the product was run on the gel along with 1X loading dye, at a constant current of 80A for 25 minutes. 100bp ladder (ThermoFisher Scientific) was also loaded for determination of the band size and estimation of concentration of the DNA. The gel was visualized in UV-transilluminator and photographed in gel documentation system (Wealtac, Sparks, USA). Until further use, the PCR products were stored in -20 °C.

3.4 Amplification of HBx gene by Polymerase Chain Reaction

The extracted and purified viral DNA present in the samples that gave positive results in the HBV detection Nested PCR, were used to amplify the HBx gene via polymerase chain reaction.

3.4.1 Primer Designing

Primers were designed by obtaining the HBx complete gene sequence published in GenBank of National Centre for Biotechnology Information (NCBI) (NCBI reference number YP_009173867.1). Primers were designed to amplify the HBx region taking cloning considerations into account. For the homology search, Clustal W software and Basic Local Alignment Search Tool (BLAST) of NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was run to determine primer specificity for the gene. Moreover, restriction sites were incorporated in the primers for future purpose of

directional cloning. The forward primer included restriction sites for BbvI, M1uCI, ApoI, EcoRI, FspEI and MspJI restriction enzymes. The reverse primer included restriction sites for AciI, Fnu4HI, BsiEI, CvikI-1, HaeIII, FspEI, EaeI and EagI restriction enzymes. Both primers had 100% sequence identity with the X sequence and Tm, GC content and self-complimentarity was analyzed using online tool OligoCalc. The sequences and properties are tabulated below (Table 3.2):

Table 3.2: Primer sequences for amplification of HBx gene. The Tm and GC content of each primer are also shown whereas the highlighted fragment of the primer sequence indicates the restriction sites.

Sr.No.	**Primer Code**	**Sequence 5'→3'**	**Tm (°C)**	**% GC Content**
1	HBXDF2	CTTGAATTCGTTATGGCTGCTAGGCTGTG CTGCCAA	71.5	54.3
2	HBXDR	GTTGCGGCCGCTTAGGCAGAGGTGAAAA AGTTGCA	72.9	50

3.4.2 Optimization of PCR

The extracted DNA was subjected to gradient PCR, with annealing temperature range of 50 to 60 °C at 2 °C intervals, in order to optimize the annealing temperature (T_A) for amplification of the HBx gene from HBV genome. Once the annealing temperature that gave the sharpest and most clear band of the amplified gene was optimized, it was used for further PCR reactions.

The 25 µL PCR reaction mixture was prepared that comprised of Taq Polymerase buffer (1X), $MgCl_2$ (2 mM), dNTPS (dATP, dGTP, dTTP, dCTP) (0.16mM each), forward and reverse primers (20 pmole each), thermostable Taq polymerase (1U) and viral DNA as template. Nuclease free water was added to make the volume up to 25 µL. PCR mixture was placed in thermocycler (Swift Maxi Thermal Cycler Block, ESCO, Singapore) and Figure 3.2 depict the PCR profile used.

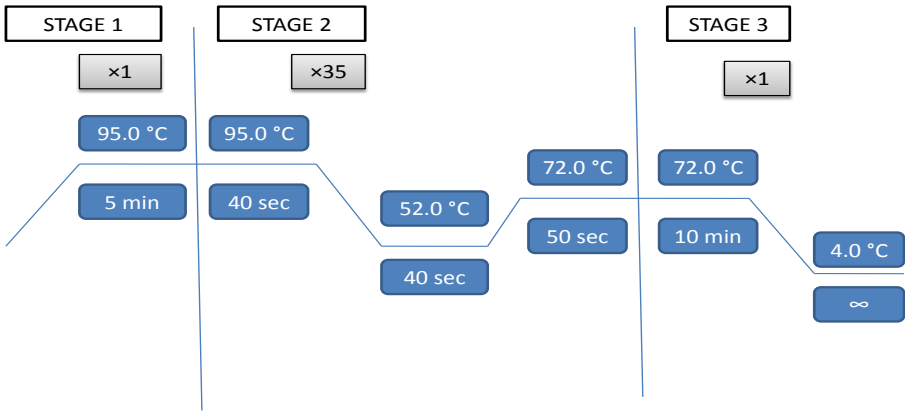


Figure 3.3: PCR profile used for the amplification of HBx gene

3.4.3 Analysis of PCR product by Gel Electrophoresis

To analyze the PCR products, 1.2% agarose gel was prepared using 1X Tris Acetate EDTA (TAE), as per the protocol mentioned in section 3.3.3. 0.96g of agarose was weighted and added to 80 mL of 1X TAE. After solidification of gel, 5 µl of the PCR product was run on the gel along with 1X loading dye. 100 bp ladder (ThermoFisher Scientific) was also loaded for determination of band size.

3.5 Screening of blood samples for HBx gene

Once the PCR for the amplification of HBx gene of HBV was optimized, the viral DNA extracted from the collected 50 serum samples were subjected to PCR in order to screen them for the presence of the gene. The methodology of PCR and product analysis on gel was same as aforementioned.

3.6 Purification of PCR product

For each positive sample, bulk PCR (50 µL) was placed, in order to use the PCR products for purification and prepare them for sequencing applications. GeneJet PCR Purification Kit (Catalogue No. K0701, Thermo Scientific) was used for purifying the PCR product. According to manufacturer's protocol, 100 µL of Binding Buffer was added to 100 µL of PCR product and mixed thoroughly. Occurrence of yellow color indicated optimal pH for DNA binding. This mixture was transferred to purification column and centrifuged for 30-60 seconds at a speed of 13,000 ×g. the flow-through was discarded and 700 µL of Wash Buffer (diluted with ethanol). Next, centrifugation was again carried out for 30-60 seconds at 13,000 ×g, flow-through was discarded and purification column was placed back in the collection tube. In order to ensure the complete removal of any residual buffer, the empty column was spun for an additional 1 minute. Next, the collection tube was discarded and column was shifted to clean 1.5 mL microcentrifuge tubes. To the centre of the purification column membrane, 50 µL of Elution Buffer was added carefully and centrifuged for 1 minute at 13,000 ×g. the column was then discarded and purified PCR product was stored at -20 °C until further use.

3.7 Cloning of HBx gene

One of the potentially important sequences of HBx was cloned in the TA vector for further applications. The procedure followed was as follows:

3.7.1 Ligation in TA vector

In order to ligate the HBx gene in the TA vector, 2 µL of vector, 5 µL of purified PCR product (gene of interest), 1 U of T4 DNA ligase were mixed together and NF water was added to make the mixture volume up to 10 µL. the ligation mixture was incubated at 16 °C overnight in the thermocycler machine and the next day, DH5α strain of *E. coli* was used to transform the ligation mixture.

3.7.2 Preparation of *E. coli* competent cells

A single colony of *E. coli* DH5 alpha cells was picked from streaked LB plates and inoculated in Luria-Bertani (LB) medium to obtain growth. The medium was prepared using 1% trypton, 0.5% yeast extract and 0.5% NaCl in 1L of distilled water. It was sterilized by autoclaving. Post inoculation, the medium was incubated in the shaking incubator at 37 ºC and at an rpm of 180, with constant shaking for 16-24 hours. On the following day, 1mL of this overnight grown starter culture was used to inoculate 50mL of LB medium. It was again incubated at 37 ºC with constant shaking. The concentration of cells was checked every hour by spectrophotometer until the optical density (OD_{600}) of the culture reached 0.6. Bacterial cells were then harvested in a 50mL centrifuge tube through centrifugation at 4000 rpm at 4 ºC (Eppendorf, Germany). The supernatant was discarded while the pellet was re-dissolved in 10mL of ice cold solution 'A' that constituted 100mM $CaCl_2$ and 50mM $MgCl_2$ and placed on ice for 30 minutes. The cells were again harvested by centrifuging at 4000 rpm at 4 ºC for 10 minutes. The supernatant was discarded and pellet was re-suspended in 2mL of solution 'B' that constituted 50mM $MgCl_2$ and 25% glycerol. 100 µL of this cell suspension was aliqouted in microfuge tubes for a single transformation reaction. The entire process was carried out on ice i.e. cold chain maintenance was ensured. The competent cells were stored at -80 ºC until further use.

3.7.3 Transformation

Heat shock method was used to carry out transformation, as described by Woodcock et al. (1989). 5 µL of ligation mixture was added to 100 µL of competent cells, followed by 30 minutes incubation on ice. The cells were then given a heat shock by subjecting them to the temperature of 42 ºC for 2 minutes and immediately snap cooling them on ice for 2-3 minutes. Next, 1 mL of LB medium (without Ampicillin) was added and cells were allowed to grow for 1 hour by incubating at 37 ºC for the expression of antibiotic resistance gene. Cells were then harvested by centrifuging for 2 minutes at 14,000 rpm and were plated on ampicillin containing LB agar plate (200 µg/ mL). Prior to plating, 40 µL of 100 mM and IPTG and 40 µL of X-Gal were spread on the plates in

order to induce the *lacZ* gene and facilitate blue/white selection. To obtain optimal bacterial growth, plates were incubated at 37 ºC overnight.

3.7.4 Screening and Selection of clones

From the plate, 4 different white colonies were picked and inoculated in 3 mL of Ampicillin-containing LB media (at a ratio of 200 µg per mL) and the cells were allowed to grow with constant shaking at 37 ºC for up to 16 hours. These white colonies contained the bacterial cells containing cloned plasmids. The clones were confirmed as follows:

3.7.5 Clone confirmation

3.7.5.1 Restriction digestion analysis

Insert was confirmed by digestion with EcoRI (Fermentas, USA) and BamHI (Fermentas, USA) each, since the pTZ57R/T vector (TA vector) contained restriction sites for these on both sides of the insert. In another round of digestion, HincII was also used to digest the cloned vector. 5 µL of plasmid was treated with 5U of restriction enzyme with a total reaction mixture of 20 µL made up by 1X buffer for each enzyme respectively. The mixture was incubated at 37 °C for 2-3 hours and the digested products were run on 1% agarose gel and visualized under UV transilluminator.

3.7.5.2 Colony PCR

In order to confirm the cloning of our gene of interest in the TA vector, colony PCR was carried out. For this purpose, 4 white colonies were picked with the help of pipette tip and suspended in 15 µL of autoclaved nuclease free water (1 colony in each microfuge tube). The colonies were mixed by gentle pipetting. Ampicillin added LB medium (50 µL in 50 mL media) was then taken and placed at 37 °C for 5 minute for pre-warming. 190 µL of this was further added to the colony and NF water mixture and mixed by vortexing. The tubes were kept in shaking incubator at 37 °C for 2-3 hours at 200 rpm to allow bacterial growth. Next, the colony mixtures were subjected to a denaturation process by keeping it in the thermocycler at a temperature of 95 °C for 15 minutes and then 25 °C for another 15 minutes. This denatured mixture was used as template to carry out PCR amplification of the HBx gene by similar procedure as

mentioned in 3.4.2 and the results were analyzed on gel by similar procedure as mentioned in 3.4.3.

3.8 Plasmid isolation

From the starter culture mixture (of positive colonies) containing picked colonies suspended in NF water and L.B broth, 100 µL was taken and added to 10 ml of L.B media with Ampicillin added. It was incubated at 37 °C overnight for 16 hours, at a constant shaking of 120 rpm. Post-incubation, the inoculated media was centrifuged at maximum speed (× 16,000 g) for 2 minutes. The supernatant was discarded and pellet containing the bacterial cells was used to isolate plasmid by using GenJet Plasmid Miniprep Kit (# K0502, Fermentas, USA). The pellet was resuspended in 250 µL of the 'Resuspension Solution' (containing RNase) and mixed thoroughly by pipetting up and down and vortexing. The mixture was transferred to a fresh microcentrifuge tube and 250 µL of the Lysis Solution was added to it. Thorough mixing was done by inverting the tube 4-6 times until the solution became viscous and slightly clear. Next, 350 µL of 'Neutralization Solution' was added to it and mixed immediately by inverting the tube 4-6 times to prevent localized precipitation of bacterial cell debris, until it attained cloudy appearance. The mixture was then centrifuged at >12000 ×g for 5 minutes to pellet down cell debris and chromosomal DNA. Pellet was discarded and the supernatant was transferred to the supplied GeneJet spin columns by decanting. The column was then centrifuged for 1 min at a speed of > 12,000 ×g and the flow-through was discarded. The column was placed back in the collection tube and 500 µL of 'Wash Solution' (diluted with ethanol) was added to the GeneJet spin column. Centrifugation was again carried out for 30-60 seconds at the same speed and flow-through was discarded. This washing step was again repeated using another 500 µL of Wash Solution. The empty column was spun for an additional 1 minute to remove any residual Wash Solution. Next, the GeneJet spin column was shifted to fresh 1.5 mL microcentrifuge tubes and collection tube was discarded. 50 µL of the Elution Buffer was added to the center of the spin column membrane to elute the plasmid DNA. It was then centrifuged for 2 minutes at 12,000 ×g after 2 minute incubation. The column was discarded and the purified plasmid was stored at -20 °C for further processing.

3.9 Analysis of Plasmid DNA by Gel Electrophoresis

To analyze the purified plasmid, 1% agarose gel was prepared using 0.8g of agarose which was added to 50 mL of 1X Tris Acetate EDTA (TAE) buffer. The gel preparation procedure was same as mentioned in section 3.3.3 above. 3 µL of the purified plasmid was run on gel along with 1X loading dye, at a constant current of 80 A for 20 minutes. The gel was then visualized under UV and photographed in gel documentation system (Wealtac, Sparks, USA).

3.10 Sequencing of Clone

The purified plasmid was sent to Eurofins MWG Operon., Karachi for automated sequencing by forward and reverse universal M13 primers. The sequence was subjected to BLAST homology search to confirm the extent of similarity with published HBx sequence on NCBI at both nucleotide and protein level.

3.11 Mutational Analysis of HBx sequences

The purified PCR products containing HBx variant sequences were sequenced using both forward and reverse gene specific primers and analyzed in order to determine the prevalent mutations. The sequencing results from forward and reverse primers were aligned at CLUSTALW in order to generate a complete X sequence (filling in missing nucleotides in forward sequence from the reverse sequence). These 13 sequences were then aligned with published HBx sequence available at NCBI, after deriving a consensus sequence. In order to determine the changes/ mutations, the sequences were aligned in CLUSTALW and MultiAlin. The complete gene sequences were translated into coresponsing amino acid sequences using the ExPASy Translate Tool which were further aligned on CLUSTALW with consensus HBx protein sequence, retrieved from NCBI, in order to detect the changes at protein level.

RESULTS

4.1 Sample Collection

Blood sera samples were collected from 50 HBV infected patients after their informed consents. The inclusion criteria were patients with chronic HBV infection, preferably having onset of liver disease and high viral titer. Data pertaining to the patients was also recorded, however, no elaborate data was found in the track records of patients from Shalamar Hospital, Lahore. The recorded information about the patients is tabulated as follows (Table 4.1). Among the 50 patients whose sera samples were processed, 30 were male and 20 were female. The samples were collected from Lahore and Islamabad city and the range of age of patients was 17 to 68 years.

Table 4.1: Data pertaining to HBV patients whose sera were collected

Sr.No.	Region	Gender	Age	Viral Titer
1	Lahore	Female	40	<3.0
2	Lahore	Female	38	3421
3	Lahore	Female	54	<3.0
4	Lahore	Male	32	<3.0
5	Lahore	Male	41	234,671
6	Lahore	Male	60	23,524
7	Lahore	Female	35	56,789
8	Lahore	Male	34	<3.0
9	Lahore	Male	17	34,256,192
10	Islamabad	Male	55	562197
11	Islamabad	Male	26	91721
12	Islamabad	Male	33	56190
13	Islamabad	Female	15	103000
14	Islamabad	Male	61	83251
15	Islamabad	Female	60	73007
16	Islamabad	Female	35	84000
17	Islamabad	Male	71	39000
18	Islamabad	Female	55	59300
19	Islamabad	Male	36	89532

20	Islamabad	Male	24	93500
21	Islamabad	Male	20	89300
22	Islamabad	Male	58	79301
23	Islamabad	Male	66	87430
24	Islamabad	Female	20	56009
25	Islamabad	Female	18	73461
26	Islamabad	Male	40	86390
27	Islamabad	Male	25	89409
28	Islamabad	Male	36	873900
29	Islamabad	Female	21	93000
30	Islamabad	Male	57	-
31	Islamabad	Female	35	-
32	Islamabad	Female	32	57910
33	Islamabad	Female	43	31004
34	Islamabad	Male	44	748130
35	Islamabad	Female	40	-
36	Islamabad	Male	22	562197
37	Lahore	Male	29	-
38	Lahore	Male	34	-
39	Lahore	Female	34	-
40	Lahore	Male	62	-
41	Lahore	Male	40	-
42	Lahore	Female	26	-
43	Lahore	Male	41	-
44	Lahore	Female	33	-
45	Lahore	Male	38	-
46	Lahore	Male	68	-
47	Lahore	Male	22	-
48	Lahore	Female	37	-
49	Lahore	Male	68	-
50	Lahore	Female	34	-

4.2 Detection of Viral DNA through Nested Polymerase Chain Reaction

The extracted viral DNA from 50 sera samples was subjected to Nested PCR using 2 pairs of HBV detection primers to confirm the presence and integrity of HBV DNA. The appearance of band size of approximately 200 bp confirmed the presence of HBV genome and these samples were rendered positive and further used to amplify the HBx gene. Results of some of PCR samples are shown in Figure 4.1 and 4.2, respectively.

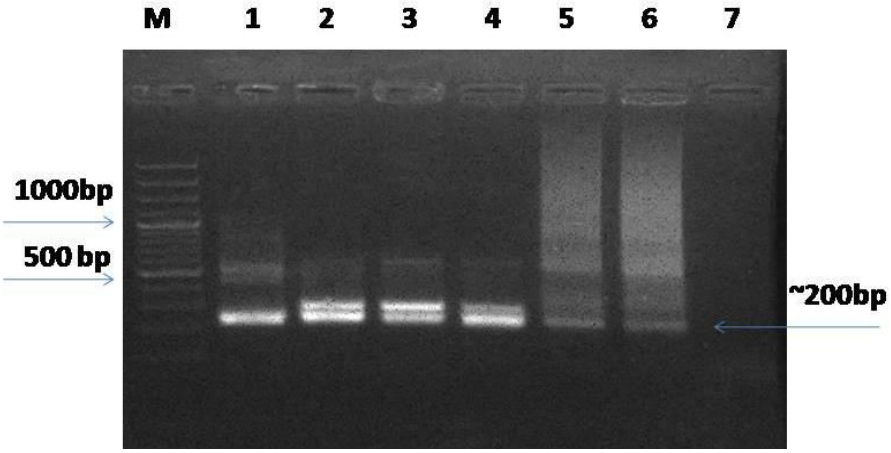


Figure 4.1: Detection of viral DNA through amplification of HBV core region by Nested PCR. Lane M: 100 bp ladder (Fermentas); Lane 1 to 6: PCR products (Lane 1, 2, 3, 4, 5 and 6 indicate positive results for HBV DNA); Lane 7: Negative control

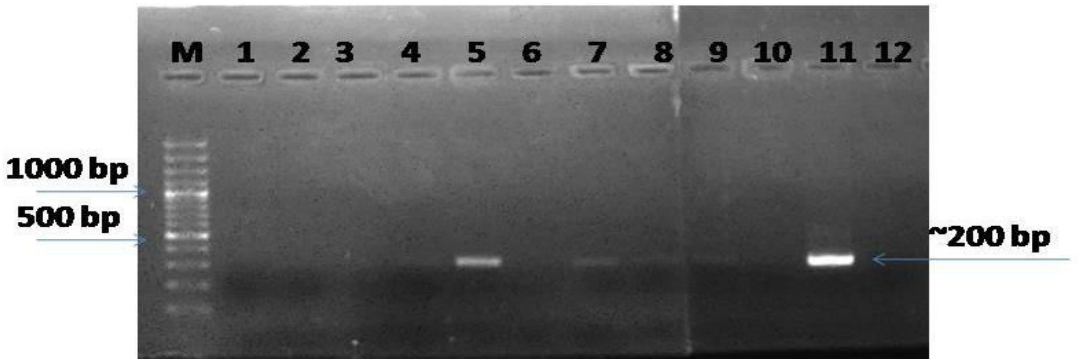


Figure 4.2: Detection of viral DNA through amplification of HBV core region by Nested PCR. Lane M: 100 bp ladder (Fermentas); Lane 1 to 11: PCR amplification from

different patient samples (Lane 5, 7, 8, 9 and 11 are positive for HBV DNA); Lane 12: Negative control

4.3 Amplification of HBx Gene by PCR and Screening of Samples

The annealing temperature for HBx amplification was optimized to be 52 °C. Using this PCR profile, the viral DNA samples that gave positive results for detection PCR were used to amplify the HBx gene by using gene-specific primers. The PCR products were run on gel and the appearance of a band of approximately 465 bp confirmed the amplification of HBx gene in the sample. Figure 4.3 and 4.4 show some of the positive results.

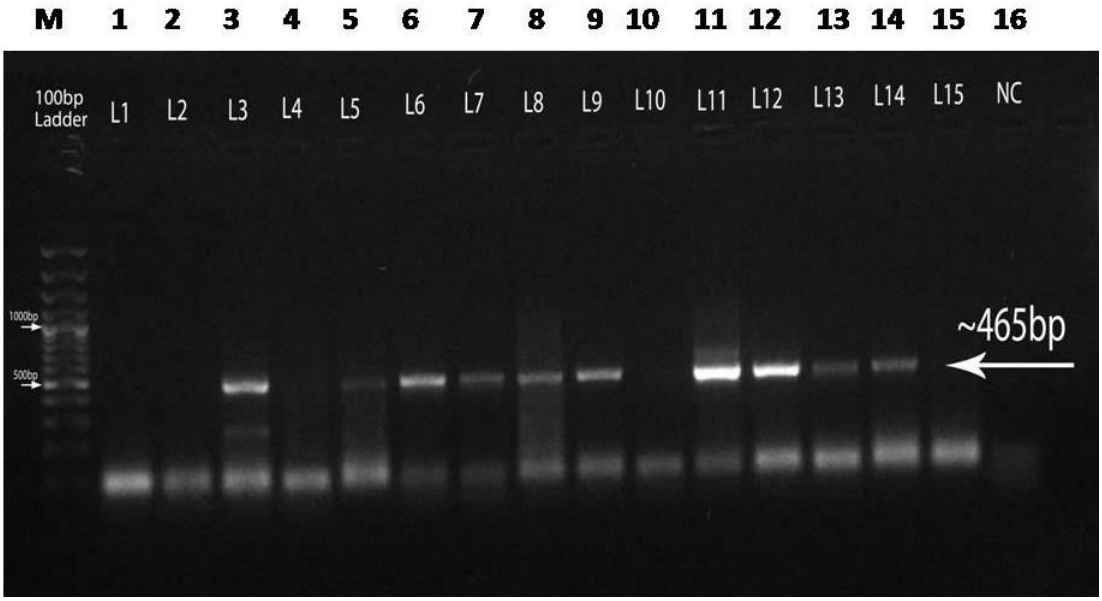


Figure 4.3: Screening of HBx gene from viral DNA isolated from sera samples. Lane M: 100bp ladder (ThermoFisher Scientific, USA); Lane 1-15: PCR products (Sample L3, L5, L6, L7, L8, L9, L11, L12, L13 and L14 are positive for HBx gene); Lane 16: Negative control

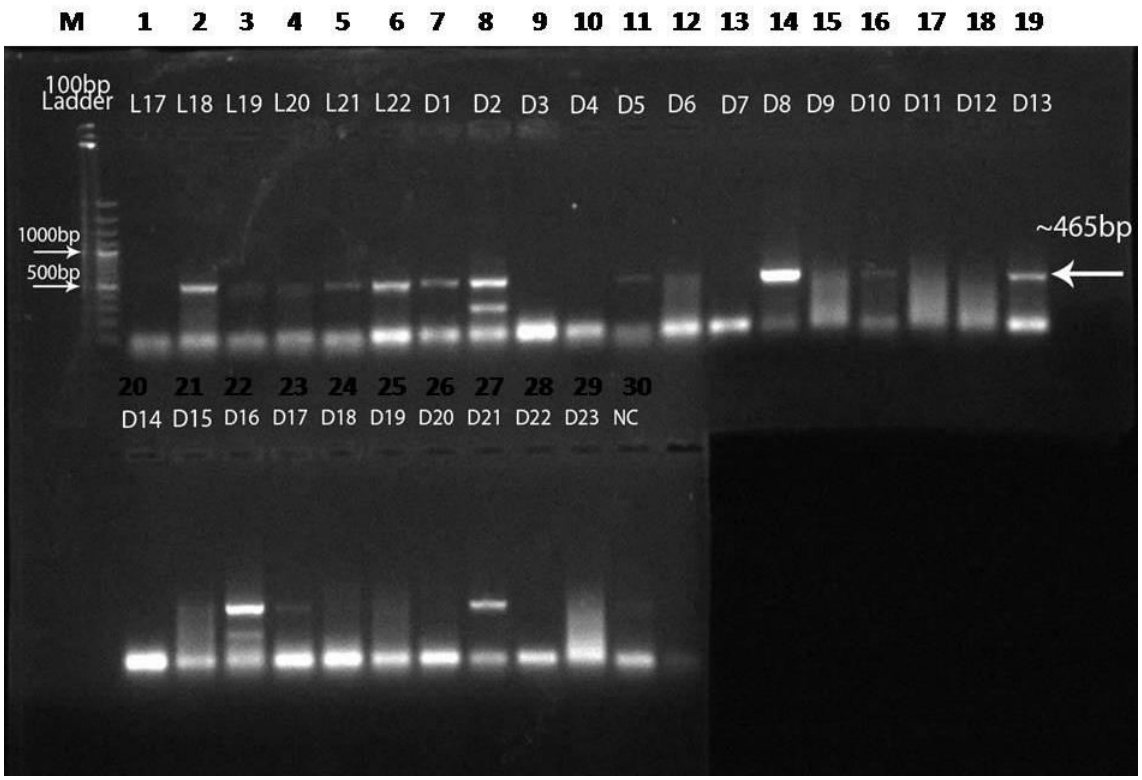


Figure 4.4: Screening of HBx gene from viral DNA isolated from sera samples. Lane M: 100bp ladder (ThermoFisher Scientific, USA); Lane 1-29: PCR products (Sample L18, L19, L20, L21, L22, D1, D2, D5, D8, D10, D13, D16, D17 and D21 are positive for HBx gene); Lane 30: Negative control

4.4 Purification of PCR product

After successful amplification, the PCR products were purified to remove all contaminants (buffer, salts and primer dimers) to avoid any interference with sequencing process. Post-purification, the products were run on gel to check discreteness, sharpness and intensity of bands appropriate for sequencing purposes.

4.5 Cloning of HBx gene

4.5.1 Ligation in TA vector

One of the purified gene products was cloned in the TA vector. Figure 4.5 shows the map of the vector and site of HBx cloning.

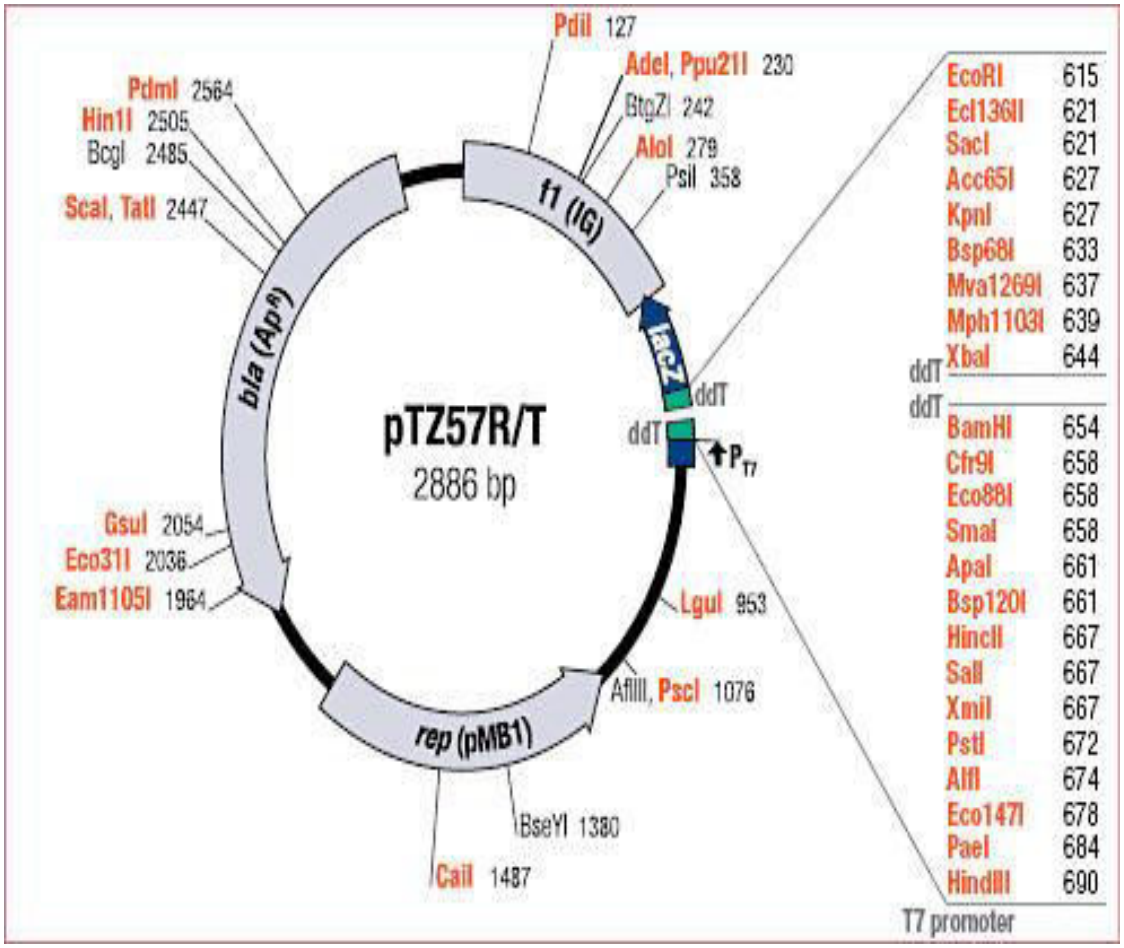


Figure 4.5: Map of pTZ57R/T showing sites for various restriction enzymes, the lacZ gene and the cloning site with T overhangs (dT) where gene of interest with A overhang binds.

4.5.2 Clone confirmation

4.5.2.1 Restriction digestion analysis

The white colonies were picked and inoculated in LB media for further growth. Plasmid was isolated from the positive sample and subjected to restriction digestion analysis for clone confirmation. The clones were subjected to EcoRI digestion that cut out the cloned gene producing a band size of approximately 500 bp, BamHI that produced a band of ~50 bp if the orientation of clone was reverse and HincII that produced a band of about 310 bp, as indicated by Figure 4.6 and 4.7, respectively. The band sizes indicated reverse orientation of the HBx gene.

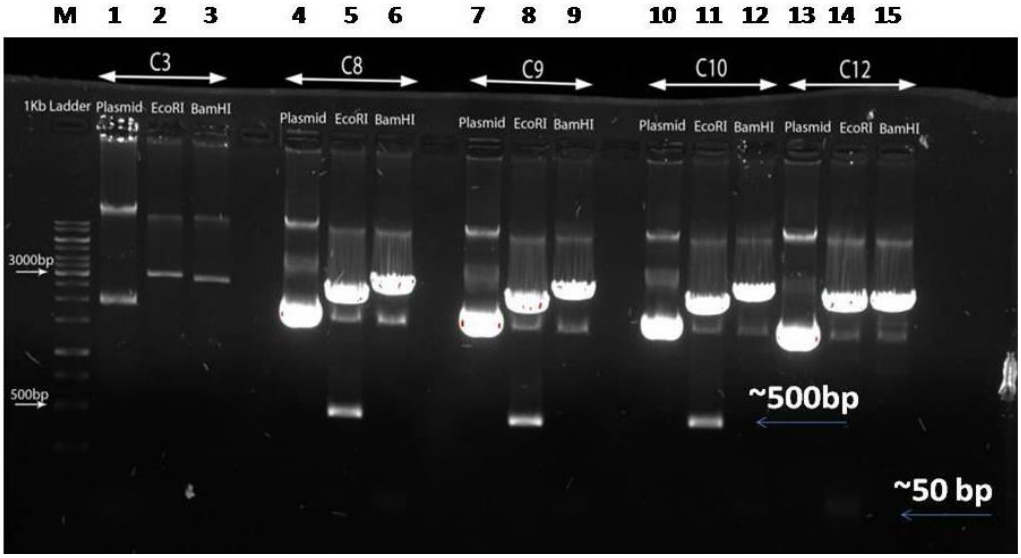


Figure 4.6: Confirmation of presence and orientation of clone through restriction digestion of pTZ57R/T-HBx clone by EcoRI and BamHI. Lane M: 1kb ladder (Fermentas); Lane 1-15: Digestion of each of the five samples (C3, C8, C9, C10 and C12) run in sets of 3 i.e. untreated plasmid, treatment with EcoRI and treatment of BamHI respectively; Appearance of ~500 bp band with EcoRI and ~50bp band with BamHI indicate the successful cloning of gene of interest in reverse orientation

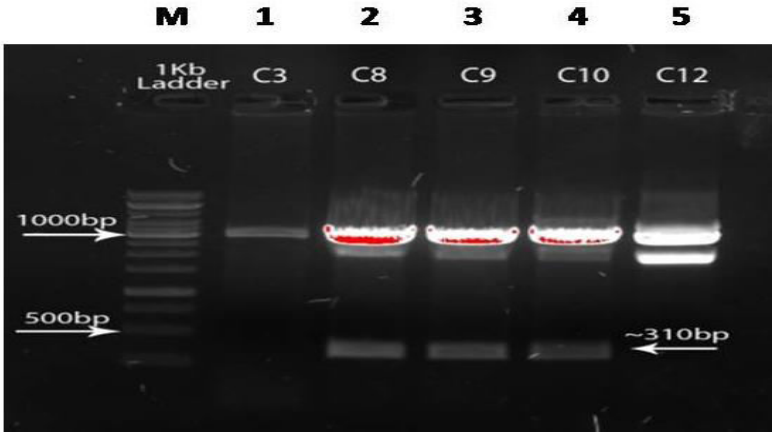


Figure 4.7: Confirmation of presence and orientation of clone through restriction digestion of pTZ57R/T-HBx clone by HincII enzyme. Lane M: 1kb ladder (Fermentas); Lane 1-5: Digestion of each of the five samples C3, C8, C9, C10 and C12; Appearance of ~310 bp band with HincII indicate the successful cloning of gene of interest in reverse orientation in samples C8, C9 and C10

4.5.2.2 Colony PCR

The white colonies expectedly containing the cloned gene of interest were picked and added to LB media. They were then confirmed by colony PCR. Figure 4.8 shows the results:

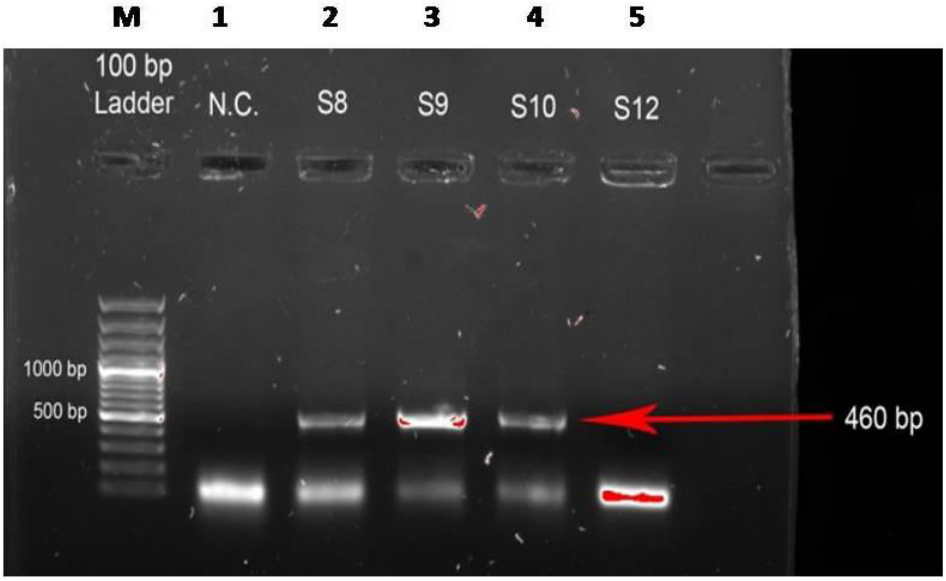


Figure 4.8: Colony PCR results of TA clone of HBx gene. Lane M: 100 bp ladder (ThermoFisher Scientific); Lane 1: Negative control; Lane 2-5: PCR products (S8, S9 and S10 show successful amplification of the HBx sequence rendering the clones positive)

4.6 Plasmid Isolation

For sequencing purpose, the positive clones were isolated and purified and run on gel for analysis. Figure 4.9 shows the purified plasmids containing our gene of interest.

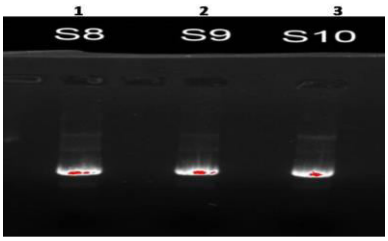


Figure 4.9: Purified plasmids containing gene of interest

4.7 Sequencing of Clone

The TA vector cloned with HBx (our gene of interest) was sent for sequencing from both forward and reverse universal primers (M13) and the results were subjected to BLAST Homology search at NCBI. The results depicted 95% homology of the sequence with published HBx gene sequence. The nucleotide sequence was translated into amino acid sequence using ExPasy Translate tool and subjected to BLAST Homology search for proteins. The results depicted 90% homology with published HBx protein sequence.

4.8 Mutational analysis of X sequences

The purified PCR products containing the amplified X gene were sent for sequencing. The sequencing results from forward and reverse primers were aligned at CLUSTALW in order to generate a complete X sequence (filling in missing nucleotides in forward sequence from the reverse sequence). These 13 sequences were then aligned with published HBx sequence available at NCBI, after deriving a consensus sequence. In order to determine the changes/ mutations, the sequences were aligned in CLUSTALW and MultiAlin. Refer to ***Annexure 1*** for the CLUSTALW sequence alignment results.

The changes detected in the X sequences [point mutations (substitutions), deletions, insertions and truncations at 3' and 5' ends of gene] are reported in the following tables. The most prevalent DNA changes i.e. occurring most frequently in the 13 X variants were considered i.e. any change occurring in at least ≥5 sequences were duly considered and their affect at protein level were studied in detail.

Table 4.2: Insertions detected in HBx sequences

Seq No.	Insertions
L6	90_91insG
L22	68_69insAA 101_102insA
D1	264_265insC

Table 4.3: Deletions detected in HBx sequences

Seq No.	Deletions
L2	36_37delGC
L22	365_369delAGTTG 415_416delCA
D1	28_30delGAT 253delG 351delT 404delG 415_416delCA
D8	29delA 69_70delTG 415_416delCA
D13	415_416delCA
D16	64delG 69delT 142delG 332delA

Table 4.4: Truncation detected at 3' end of HBx sequences

Seq No.	Truncation at 3' end
L22	36 bp truncation from 1st to 36th position (ATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCG)

Table 4.5: Truncation detected at 5' end of HBx sequences

Seq No.	Truncation at 5' end
L22	16 bp truncation from 450th to 465th position (TTTCACCTCTGCCTAA)
D1	46 bp truncation from 420th to 465th position (ATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACC TCTGCCTAA)

D8	46 bp truncation from 420th to 465th position (ATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACC TCTGCCTAA)
D13	19 bp truncation from 447th to 465th position (CTTTTTCACCTCTGCCTAA)

Table 4.6: Most commonly occurring substitutions (Mismatches) detected in X variants

Sr. No.	Type of Mutation	Frequency (out of 13 sequences)	Percentage of X mutants carrying mutation
1.	21C→T/A	6	46%
2.	23A→C/T	7	54%
3.	24A→T	5	39%
4.	25C→T/A	6	46%
5.	35C→A/T	5	39%
6.	43G→T/C	5	39%
7.	54A→C/G/T	6	46%
8.	57T→A/C	6	46%
9.	64G→C/T	5	39%
10.	70G→A	10	77%
11.	76T→C/A	13	100%
12.	89T→A	7	54%
13.	97T→C	13	100%
14.	136C→T	11	85%
15.	262A→T/C	8	62%
16.	264A→C	10	77%
17.	270C→T	7	54%
18.	303A→T	13	100%
19.	354A→G	11	85%
20.	384G→A	9	69%
21.	400C→T	12	92%

4.9 Mutational Analysis of X Protein sequences

The corresponding CLUSTALW sequence alignment results for protein sequences are shown in ***Annexure 2***. The various mutations occurring at the protein level, corresponding directly to the mutations at nucleotide level, were recorded as follows:

Table 4.7: Deletions/ Insertions detected in X protein sequences

Seq No.	Deletions/ Insertions
L12	119delW
D1	109delE
D8	57delG
D16	35_36delGT 90delP 114_115insV

Table 4.8: N-terminal truncations detected in X protein sequences

Seq No.	Truncation at N terminal
L22	12 amino acid truncation from 1st to 12th position (MAARLCCQLDPA)

Table 4.9: C-terminal truncations detected in X protein sequences

Seq No.	Truncation at C terminal
L22	9 amino acid truncation from 145th to 154th position (APCNFFTSA)
D1	17 amino acid truncation from 137th to 154th position (RHKLVCAPAPCNFFTSA)
D8	15 amino acid truncation from 139th to 154th position (KLVCAPAPCNFFTSA)
D13	7 amino acid truncation from 147th to 154th position (CNFFTSA)

DISCUSSION

Chronic infection of HBV has been the leading worldwide cause of liver cancer, comprising of 55% of cases around the globe (Zhang *et al*., 2016). Mounting evidences implicate the role of HBxAg (viral product of X ORF) in the carcinogenesis. It is said to interact with a plethora of cellular proteins leading to deregulation of various signaling pathways pertaining to cell cycle process, cell growth and apoptosis e.g. Src, PI3K/AKT, Wnt/β-catenin, Jak1/STAT, NF-κB etc (Kew, 2011).

It has been reported that mutations tend to accumulate in the genome of HBV during the course of its natural history of infection. Many of these changes occur in the HBx gene and have been preferentially detected in patients with advanced liver disease. The integration of mutated HBx sequences in host's genome mark the onset of carcinogenesis. These alterations can also help in determining the stage of chronic infection as well as its prognosis, for instance, double substitution K130M and V131I have been more frequently detected in HBx sequences derived from the sera of patients presenting cirrhosis or HCC as compared to the HBx sequences derived from sera of patients with HBV infection or mild liver disease (Baptista *et al*., 1999; Venard *et al.,* 2000).

It is a widely known fact that HCC occurs many years post-HBV infection; however, precise events of how HBV leads to cell transformation are not very clearly understood. Owing to the prevalence of many point mutations, insertions and deletion in HBx sequences derived from sera or tissue of chronic patients, we analyzed the HBx polymorphisms to determine whether they are similar to the mutations already reported by other research groups or are there characteristic HBx alterations that strongly disrupt its functional activities and associate with the liver disease. Our study also aimed at determining various mutations that may occur in X sequences and could lead to changes in X protein. The results indicated higher frequency of mutations at the 3' and 5' ends of the gene whereas the internal regions were found to be relatively conserved.

Most frequently occurring mutations identified were substitutions at 76T→C/A, 97T→C, 303A→T, 400C→T/G, 136C→T and 70G→A/T that occurred in all the

examined X sequences. 70G→A/T substitution changed the codon for Valine (24th amino acid in X protein) and resulted in frameshift mutation from 24th to 33 or 39th amino acids, depending on accompanying substitutions and codon redundancy. 76T→C/A substitution changed the codon for Cysteine (26th amino acid in X protein) and led to a frameshift from 26th amino acid onwards. Similarly, substitution 97T→C changed the codon for Serine (33rd amino acid in X protein) to Proline (mostly) or Arginine or Alanine, however, unlike other substitutions, did not lead to a frameshift.

All these aforementioned changes occurred in the regulatory domain of X protein within which is present the Ser/Pro rich dimerization region. It can thus be predicted that these frameshifts could perturb the dimerization of X with other proteins involved in signal transduction, protein decay, cell division or gene transcription. As reported by Reddi *et al.* (2003), this particular region of X has a putative function of signal transduction and also performs the ATPase and GTPase functions in the host cell. Contradictory to this, it is also suggested by researchers that HBx protein with N-terminal truncation was sufficient for dimerization process provided the four cysteine residues at position 7th, 8th, 17th and 26th remained unaffected (Lin and Lo, 1989). Substitutions commonly occurring in our sequences in dimerization region of X protein did not affect any of the 4 important Cysteine residues.

Other noteworthy substitution occurring in all of the analyzed X variants were 303A→T and 400C→T that altered the codons for Serine (position 101) and Lysine (position 134), respectively. Owing to the degeneracy of codons, the substitution 303A→T changed the genetic code from TCA to TCT, both of which code for the same amino acid Serine, hence making no difference to the protein. The other substitution of 400C→T, however, changed the Lysine residue.

According to Chen et al. (2005), one of the most commonly occurring mutation pattern observed in HBx sequences amplified from either liver tissue (40%) or serum samples (70.7%) of patients chronically infected with HBV is the insert mutation at position 204 (insert AGGCCC) which is almost always occuring with point mutations at position 260 (G→A) and 264 (A→ G/C/T). In comparison to our study, 12 out of 13 sequences (i.e. L9, L11, L12, L18, D1, D2, D8, D13, D16, D21) showed similar

substitution at position 264 (A→C) which was always accompanied by substitution at position 262 (A→T) together responsible for amino acid substitution of 88I→F in HBx protein. However, it was not accompanied by any insertions at positions 204, as reported by Chen *et al*. (2005).

Among other commonly reported mutations in the HBx region is the double mutation at 1762/1764 position in HBV genome that makes 386th and 388th nucleotide of the gene (Ma *et al*., 2008) but except sequence L22, D16 (for 386th position) and L2, D8, D16 (for 388th position) none of our sequences significantly showed point mutations at these positions. Moreover, the core promoter overlaps with the X gene and hence common point mutations there A1753C, A1762T and G1764A induce I127T, K130M, V131I and F132Y changes near the X protein's C-terminus (Tong *et al*., 2013). In our results, amino acid substitution at position 131 was also seen in sequences L22, D1, D16, D21 changing Valine to Isoleucine/ Serine or Glutamic Acid, depending on the accompanied point mutation in the very codon.

As per the previous studies, wild-type HBx tends to inhibit cell proliferation and transformation but the occurrence of mutations and truncation at C-terminal region can abrogate the activities of wild-type HBx resulting at the stage of integration of HBV DNA into HCC tissues. In addition, it was found that the wild type HBx enhanced promoter activity for the tumor suppressor p21 which was abrogated by the K130M mutation in HBx, in turn downregulating p21 protein levels (Ma *et al*., 2008; Tu, Bonura and Giannini, 2001; Sirma *et al*., 1999). Similarly, our results in sequences L2, L22, D1, D8, D13 and D16 also indicated change of amino acid Lysine (K) at position 130 into Methionine or Glutamine, as per the substitutions at adjacent positions that changed the code for K accordingly.

Research has also revealed presence of insertions in HBx gene and their potential role in tumorogenesis. According to the findings of a Korean cohort study conducted by Lee *et al*. (2011), four novel types of insertion mutations (PKLL, GM, FFN and tt) were detected in 6 patients. We found novel insertions not previously reported. Insertion 90_91insG in sequence L6 caused substitution of 30th amino acid of X protein from Phenylalanine to Threonine. Another sequence L22 was found to have insertions

(68_69insAA and 101_102insA) which led to frameshift mutation in the respective X protein from 22nd to 34th amino acid, thereby disrupting its 14-3-3 binding motif (ranging from 26 to 35 amino acids). This interaction of 14-3-3 protein with HBx is essential for Raf-1 kinase activation that in turn plays a role in hepatocyte growth regulation (Aoki *et al*., 2000).

An important deletion of 415_416delCA, observed in 4 sequences (D1, D8, D13, L22) in codon 139 caused a frameshift mutation from amino acid 138 to 146 which are present in the p53 binding domain of X protein, hence disrupting the motif. Within the transactivating domain of X protein is present the p53 binding region, however the two functions of HBx are independent of each other. HBx tends to directly or indirectly oppose the functions of p53 e.g. it decreases the repression of alpha-fetoprotein (AFP) via p53, through the p53 binding element on the AFP promoter. Moreover, X binding to p53 can also alleviate certain liver specific transcriptional co-repressors (Murakami, 2001). Hence, it can be concluded that insertion 415_416delCA in X protein can render p53 down-regulated and increase the cancer progression.

Six key mutations were also reported that are located in the core promoter and proximal region of the precore gene (overlapping with X): G1613A, C1653T, T1753V, A1762T, G1764A, A1846T, G1896A and G1899A, however, they were not found to occur significantly in our sequences. Interestingly, a recent study also investigated the post operative value of HBV mutations in HBxAg and identified 8 mutational sites including 1383, 1461, 1485, 1544, 1613, 1653, 1719, and 1753 that could independently predict HCC survival (Zhang *et al*., 2016; Jang *et al*., 2012; Park *et al*., 2014). The first reported mutational site i.e. 1383 (7th nucleotide in X gene) was found consistent with our findings. 5 sequences (i.e. L11, L18, D1, D8, and D13) contained the substitution at position 7 from G to T, thereby causing frameshift in X protein from amino acid 3 to 15 or 18, depending on accompanying substitutions at nucleotide level. This change of a set of amino acids could disrupt the regulatory domain of HBx protein that spans from 1st to 50th amino acid (Li *et al*., 2009).

A very important event occurring in X-protein associated carcinogenesis is the C-terminal truncation of the protein. Various researches have established that disruptions in

HBx that lead to formation of HBx proteins truncated at C-terminal are more prevalent in the tissue or sera of patients with early onset or late stage of liver tumor than in non-tumor patients (Chen *et al.,* 2000; Paterlini *et al*., 2005; Hsai *et al*., 1997). These C-terminally truncated HBx proteins are devoid of their anti-proliferative activity and co-operate with *ras* and *myc* oncogenes in cellular transformation (Tu *et al*., 2001). These frequently occurring C-truncated HBxAg also lose their pro-apoptotic activity that is otherwise functional in wild-type HBx, thereby increasing the transformation process *in vitro* and tumor- promoting activity *in vivo* (Ma *et al*., 2008).

Among the X sequences we analyzed, 7 were found to be truncated at the 5' end (L6, L11, L12, L22, D1, D8, D13), correspondingly translating into HBx proteins with C-terminal truncations. The range of number of amino acids truncated at C-terminal was 2 to 19 amino acids, hence perturbing the transactivation domain (50 to 154 amino acids) within which is present the p53 binding domain (100 to 154 amino acids). Similarly, Ma et al. (2008) found that 3' of X gene was frequently deleted in HCC cells, leading to the X protein truncated at –COOH terminal and that these were sufficient for causing HCC instead of full-length HBx. In addition to that, studies have also revealed that presence of C-terminal truncations in X protein of $\leq$ 24 amino acids (as in case of our X sequences) can be significantly correlated to venous invasion and metastasis (Sze *et al*., 2012).

X protein found in chronically infected patients is almost always truncated at either N-terminal or C-terminal or both. N-terminal of the protein consists of the regulatory domain which is not involved in trans-activation. Our results included 8 out 13 N-terminal truncated X amino acid sequences: L2, L12, L22, D1 D2, D13, D16 and D21. According to the investigation of Gottlob *et al.* (1998), transfection of REV2 cell lines with HBx mutants lacking the NH2 terminal segment showed retention of trans-activation property but were still unable to alter growth characteristics of REV2 cells hence concluding that HBx trans-activation activity (C-terminal domain) is neither essential nor sufficient for tumor progression.

Truncation or frame shift mutations in the N terminal of X protein do not tend to affect its trans-activation function that is responsible for cancer. Various studies have asserted that truncated C-terminal or deletions in the trans-activation domain causes the

loss of p53 dependent transcriptional repression binding site and growth suppressive effect domain, leading to cell proliferation and transformation (Wang *et al*., 2005). According to Gottlob *et al*. (1998), presence of putative p53 binding domain only and not the initial trans-activating domain cannot execute trans-activation functions.

Conclusively, the prevalence of significant mutations in the analyzed HBx sequences manifest important changes in the protein, thereby, altering its functional properties and opting for different mechanisms that lead to increase in cell proliferation, suppression of p53 mediated cell division checks and alleviation of apoptosis, all eventually leading to favorable conditions for cellular transformation. These frequently occurring mutational events can be used as predictive biomarkers to determine the stage of chronic infection and cancer progression. RNA interference technology can be used in future against the X mRNA corresponding to its dimerization or transactivation domains/ p53 binding motif/ 14-3-3 binding region etc, with the target of minimizing HBx-mediated hepatocarcinogenesis. For further studies, a larger sample size may be collected, also including the liver tumor tissue samples and mutational analysis can be carried out using other mutation-specific detection PCRs.

REFERENCES

Aoki H., Yashi J.H., Moriyama M., Arakawa Y. and Hino O. 2000. Hepatitis C virus Core Protein Interacts with 14-3-3 Protein and Activates the Kinase Raf-1. *Journal of Virology*. 74(4): 1736-1741.

Artavanis-Tsakonas S., Rand M.D and Lake R.J. 1999. Notch signaling: Cell fate control and signal integration in development. *Science*. 284 (5415): 770-776.

Baptista M., Kramvis A. and Kew M.C. 1999. High prevalence of 1762(T) 1764(A) mutations in the basic core promoter of hepatitis B virus isolated from black Africans with hepatocellular carcinoma compared with asymptomatic carriers. *Hepatology*. 29: 946−953.

Benhenda S., Cougot D., Neuveut C. and Buendia M.A. 2009. Liver Cell Transformation in Chronic HBV Infection. *Viruses*. 1(3): 630-646.

Benn J. and Schneider R.J. 1994. Hepatitis B virus HBx protein activates Ras-GTP complex formation and establishes a Ras, Raf, MAP kinase signaling cascade. *Proceedings of National Academy of Sciences*. 91: 10350-10354.

Benn J. and Schneider R.J. 1995. Hepatitis B virus HBx protein deregulates cell cycle checkpoint controls. *Proceedings of National Academy of Sciences*. 92: 11215-11219.

Bouchard M.J., Wang L.H. and Schneider R.J. 2001.Calcium signaling by HBxprotein in hepatitis B virus DNA replication. *Science.* 294(5550):2376–2378.

Centers for Disease Control and Prevention. 2011. *Morbidity and Mortality Weekly Report: Establishment of a Viral Hepatitis Surveillance System*. 60(40); 1385-

1390 [ONLINE] Available at: https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6040a3 .htm. [Accessed 24 September 2016].

Chen G.G., Li M.Y., Ho R.L.K, Chak E.C.W., Lau W.Y and Lai P.B.S. 2005. Identification of hepatitis B virus X gene mutation in Hong Kong patients with hepatocellular carcinoma. *Journal of Clinical Virology*. 34: 7−12.

Chen W.N., Oon C.J., Leong A.L., Koh S. and Teng S.W. 2000. Expression of integrated hepatitis B virus X variants in human hepatocellular carcinomas and its significance. *Biochemical and Biophysical Research Communications*. 276: 885−892.

Deng-F.Y., Zhi-Z.D. and Min Y. 2007. Specific molecular markers in hepatocellular carcinoma. *Hepatobiliary & Pancreatic Diseases International.* 6(3): 241-247.

Diao J., Khine A.A. Sarangi F., Hsu E., Iorio C., Tibbles L.A., Woodgett J.R., Penninger J. and Richardson C.D. 2001. X Protein of Hepatitis B Virus Inhibits Fas-mediated Apoptosis and Is Associated with Up-regulation of the SAPK/JNK Pathway. *The Journal of Biological Chemistry*. 276(11): 8328-8340.

Diao J.,Garces R., Richardson C.D. 2001. X protein of hepatitis B virus modulates cytokine and growth factor related signal transduction pathways during the course of viral infections and hepatocarcinogenesis. *Cytokine and Growth Factor Reviews.*12 (2-3):189-205.

Ding Q, Xia W, Liu JC, *et al*: Erk associates with and primes GSK-3beta for its inactivation resulting in upregulation of beta-catenin. Mol Cell 19: 159-170, 2005.

Doria M., Klein N., Lucito R. and Schneider R.J. 1995.The hepatitis B virus HBx protein is a dual specificity cytoplasmic activator of Ras and nuclear activator of transcription factors. *The EMBO Journal*. 14(19): 4747-4757.

Elmore L.W., Hancock A.R., Chang S.F., Wang X.W., Chang S., Callahan C.P., Geller D.A., Will H. and Harris C.C. 1997. Hepatitis B virus X protein and p53 tumor suppressor interactions in the modulation of apoptosis. *Proceedings of National Academy of Sciences*. 94: 14707-14712.

El-Serag H.B. 2004. Hepatocellular carcinoma: recent trends in the United States. *Gastroenterology*. 127: S27-S34.

Fan W., H. Zhou, Y. Yang, X. Xia, Q. Sun, J. Luo and B. Cheng. 2012. Hepatitis B virus X protein promotes the growth of hepatocellular carcinoma by modulation of the Notch signaling pathway. *Oncology Reports*. 27: 1170-1176.

Feitelson M.A. and Lee J. 2007. Hepatitis B virus integration, fragile sites, and hepatocarcinogenesis. *Cancer Letters*. 252(2):157–70.

Gearhart T.L. and Bouchard M.J. 2010.The Hepatitis B Virus X Protein Modulates Hepatocyte Proliferation Pathways To Stimulate Viral Replication. *Journal of Virology*. 84(6): 2675-2686.

Gong D.Y., Chen E.Q., Huang F.J., *et al.* 2013. Role and functional domain of hepatitis B virus X protein in regulating HBV transcription and replication *in vitro* and *in vivo*. *Viruses*. 5(5):1261–1271.

Gonzalez-Amaro R., Garcia-Monzon C., Garcia-Buey L., Moreno-Otero R., Alonso J.L., Yague E., et al. 1994. Induction of tumor necrosis factor alpha production by human hepatocytes in chronic viral hepatitis. *The Journal of Experimental Medicine*. 179(3): 841–848.

Gottlob K., Fulco M., Levrero M. and Graessmann A. 1998.The hepatitis B virus HBx protein inhibits caspase 3 activity. *Journal of Biological Chemistry*. 273(50): 33347-33353.

Guo X., Jin Y., Qian G. and Tu H. 2008. Sequential accumulation of the mutations in core promoter of hepatitis B virus is associated with the development of hepatocellular carcinoma in Qidong, China. *Journal of Hepatology*. 49: 718-725.

Han J., Yoo H.Y., Choi B.H. and Rho H.M. 2000. Selective transcriptional regulations in the human liver cell by hepatitis B viral X protein. *Biochemical and Biophysical Research Communications.* 272(2): 525-530.

Harris A.L. 2002. Hypoxia-a key regulatory factor in tumour growth. *Nature Reviews*. 2: 38–47.

Hsia C.C., Nakashima Y. and Tabor E. 1997. Deletion mutants of the hepatitis B virus X gene in human hepatocellular carcinoma. *Biochemical Biophysical Research Communications*. 241:726−729.

Hsu I.C., Tokiwa1 T., Bennett W., Metcalf R.A., Welsh J.A., Sun T. and Harris C.C. 1993. p53 gene mutation and integrated hepatitis B viral DNA sequences in human liver cancer cell lines. *Cardno Genes Evol Journal*.14(5): 987-992.

Jain S. *et al.* 2015. Comprehensive DNA methylation analysis of hepatitis B virus genome in infected liver tissues. *Scientific Reports.* **5**:10478. doi: 10.1038/srep104783.

Jang J.W., Chun J.Y., Park Y.M., Shin S.K., Yoo W., Kim S.O., Hong S.P. 2012. Mutational complex genotype of the hepatitis B virus X /precore regions as a novel predictive marker for hepatocellular. *Cancer Science.* 103: 296-304.

Kekule A.S., Lauer U., Weiss L., Luber B. and Hofschneider P.H. 1993. Hepatitis B virus transactivator HBx uses a tumor promoter signalling pathway. *Nature.* 361(6414): 742-745.

Kew M.C. 2010. Epidemiology of hepatitis B virus infection, hepatocellular carcinoma, and hepatitis B virus-induced hepatocellular carcinoma. *Pathologie-Biologie (Paris).* 58(4): 273-7.

Kew M.C. 2011. Hepatitis B virus x protein in the pathogenesis of hepatitis B virus-induced hepatocellular carcinoma. *Journal of Gastroenterology and Hepatology.* 26(1): 144-152.

Kim C.M., Koike K., Saito I.,Miyamura T. and Jay G. 1991. HBx gene of hepatitis B virus induces liver cancer in transgenic mice. *Nature.* 351(6324): 317-320.

Klein N.P. and Schneider R.J. 1997. Activation of Src family kinases by hepatitis B virus HBx protein and coupled signaling to Ras. *Molecular and Cellular Biology*. 17(11): 6427-6436.

Koike K. 1995. Hepatitis B virus HBx gene and hepatocarcinogenesis. *Intervirology*. 38: 134-142.

Kumar V., Jayasuryan N. and Kumar R. 1996. A truncated mutant (residues 58-140) of the hepatitis B virus X protein retains transactivtion function. *Proceedings of the National Academy of Sciences*. 93: 5647-5652.

Lee M.O., Choi Y.H., Shin E.C., Kang H.J., Kim Y.M., Jeong S.Y., Seong J.K., Yu D.Y., Cho H., Park J.H. and Kim S.J. 2002. Hepatitis B virus X protein induced expression of interleukin18 (IL-18): a potential mechanism for liver injury caused by hepatitisB virus (HBV) infection. *Journal of Hepatology*. 37(3):380–386.

Lee S.A., Mun H.S., Kim H., Lee H.K., Kim B.J., Hwang E.S., Kook Y.H., Kim B.J. 2011. Naturally occurring hepatitis B virus X deletions and insertions among Korean chronic patients. *Journal of Medical Virology.* 83: 65-70.

Lee Y.I, Kang-Park S. and Do S.I. 2001. The hepatitis B virus-X protein activates a phosphatidylinositol 3-kinase-dependent survival signaling cascade. *The Journal of Biological Chemistry*. 276 (20): 16969-16977.

Lee, F. 2006. Horizons in Cancer Research. *Liver Cancer New Research*. 11th ed. Nova Biomedical Books, pp 1-40. Print.

Li S.K., Ho S.F., Tsui K.W., Fung K.P. and Waye M.Y. 2008. Identification of functionally important amino acid residues in the mitochondria targeting sequence of hepatitis B virus X protein. *Virology*. 381(1): 8-81.

Lian Z.R., Liu J., Pan J.B., Tufan N.L.S., Zhu M.H., Arbuthnot P., Kew M., Clayton M.M. and Feitelson M.A. 2001. A cellular gene up-regulated by hepatitis B virus-encoded X antigen promotes hepatocellular growth and survival. *Hepatology*. 34(1): 146-157.

Lin M.H and Lo S.H. 1989. Dimerization of hepatitis B viral X protein synthesized in a cell-free system. *Biochemical and Biophysical Research Communications*. 164:14-21.

Liu L.P., Hu B.G., Ye C., Ho R.L.K., Chen G.G. and Lai P.B.S. 2014. HBx mutants differentially affect the activation of hypoxia-inducible factor-1a in hepatocellular carcinoma. *British Journal of Cancer*. 110: 1066-1073.

Liu L.P., Hu B.G.,Ye C., *et al.* 2014. HBx mutants differentially affect the activationof hypoxia-inducible factor-1α in hepatocellular carcinoma. *British Journal of Cancer.*110(4):1066–1073.

Liu L.P., Liang H.F., Chen X.P., Zhang W.G., Yang S.L., Xu T., Ren L. 2010. The role of NF-kappa B in Hepatitis b virus X protein-mediated upregulation of VEGF and MMPs. *Cancer Investigation*. 28: 443–451.

Lucito R. and Schneider R.J. 1992. Hepatitis B virus X protein activates transcription factor NF-kappa B without a requirement for protein kinase C. *Journal of Virology*. 66(2): 983-991.

Ma N.F., Lau S.H., Hu L., Xie D., Wu J., Yang J., Wang Y., Wu M.C., Fung J., Bai X., Tzang C.H., Fu L., Yang M., Su Y.A. and Guan X.Y. 2008. COOH-terminal truncated HBV X protein plays key role in hepatocarcinogenesis. *Clinical Cancer Research.* 14: 5061-5068.

Maguire H.F., Hoeffler J.P. and Siddiqui A. 1991.HBV X protein alters the DNA binding specificity of CREB and ATF-2 by protein-protein interactions. *Science*. 252(5007): 842-844.

Mi-O.L., Youn-H.C., Eui-C.H., Hyo-J.K., Young-M.K., Su-Y.J., Je K.S., Dae-Y.Y., Hyeseong C., Jeon H.P. and Se J.K. 2007. Hepatitis B virus X protein induced expression of interleukin 18 (IL-18): a potential mechanism for liver injury caused by hepatitis B virus (HBV) infection. *Journal of Hepatology*. 37: 380–386.

Murakami S. 2001. Hepatitis B virus x protein: a multifunctional viral regulator. *Journal of Gastroenterology*. 36(10): 651–60.

Murakami S., Cheong J. and Kanekol S. 1994. Human Hepatitis Virus X Gene Encodes a Regulatory Domain That Represses Transactivation of X Protein. *The Journal of Biological Chemistry*. 269(21): 15118-15123.

Neuveut C., Wei Y., Buendia M.A. 2010. Mechanisms of HBV-related hepatocarcinogenesis. *Journal of Hepatology*. 52: 594–604.

Ng S.A. and Lee C. 2011. Hepatitis B virus X gene and hepatocarcinogenesis. *Journal of Gastroenterology*. 46: 974–990.

Park Y.M., Jang J.W., Yoo S.H., Kim S.H., Oh I.M., Park S.J., Jang Y.S. and Lee S.J. 2014. Combinations of eight key mutations in the X/preC region and genomic activity of hepatitis B virus are associated with hepatocellular carcinoma. *Journal of Viral Hepatitis*. 21: 171-177.

Paterlini P., Poussin K., Kew M., Franco D. and Brechot C. 1995. Selective accumulation of the X transcript of hepatitis B virus in patients negative for hepatitis B surface antigen with hepatocellular carcinoma. *Hepatology*. 21: 313−321.

Qin L.L., Su J.J., Li Y., Yang C., Ban K.C. and Yian R.Q. 2000. Expression of IGF-II, p53, p21 and HBxAg in precancerous events of hepatocarcinogenesis induced by AFB1 and/or HBV in tree shrews. *World Journal of Gastroenterology*. 6(1): 138-139.

Qin L.X. and Tang Z.Y. 2002. The prognostic molecular markers in hepatocellular carcinoma. *World Journal of Gastroenterology*.8(3): 385-392.

Rawat S., Clippinger A.J. and Bouchard M.J. 2012. Modulation of Apoptotic Signaling by the Hepatitis B Virus X protein. *Viruses*. 4(11): 2945-2972.

Renner M., Haniel A., Btirgelt E., Hofschneider P.H. and Koch W. 1995.Transactivating function and expression of the x gene of Hepatitis B virus.*Journal of Hepatology*. 23: 53-65.

Runkel L., Fischer M. and Schaller H. 1993. Two-Codon Insertion Mutations of the HBx Define Two Separate Regions Necessary for Its Trans-activation Function. *Virology*. 197: 529-536.

Seeger, C., Zoulim F. and Mason W.2007. Hepadnaviruses, p. 2977–3029.*In* D. Knipe and P. Howley (ed.), Fields virology, vol. 2. Lippincott Williams & Wilkins, Philadelphia, PA.

Semenza G.L. 2003. Targeting HIF-1 for cancer therapy. *Nature Reviews*. 3: 721–732.

Shilu M, Ashraf A, Hany AH, Kaneez F, Mohd S,Govindaraju A, Nargis B, Syed J, Muhammad I, Adeel G.A.C, Mohammad AQ, Salem MB and Ishtiaq Q. Biomarkers for virus-induced hepatocellular carcinoma (HCC)Infection, Genetics and Evolution 26: 327–339, 2014.

Sir D., Tian Y., Chen W.L., *et al.* 2010.The early autophagic pathway isactivated by hepatitis B virus and required for viral DNA replication.*Proceedings of National Academy of Sciences of USA*.107(9):4383–4388.

Sirma H., Giannini C., Poussin K., Paterlini P., Kremsdorf D. and Brechot C. 1999. Hepatitis B virus X mutants, present in hepatocellular carcinoma tissue abrogate both the antiproliferative and transactivation effects of HBx. *Oncogene*. 18: 4848-4859.

Smith S.S., Michael M.Z., Graham M.W. 1989. Quantitative evaluation of Escherichia coli host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Research*. 17(9): 3469-78.

Tang H., Delgermaa L., Huang F., Oishi N., Liu L. He F., Zhao L. and Murakami S. 2005. The Transcriptional Transactivation Fuction of HBx Protein Is Important for Its Augmentation Role in Hepatitis B Virus Replication. *Journal of Virology*. 79(9): 5548-5556.

Tetsuhiro C, E. Suzuki, T. Saito, S. Ogasawara, Y. Ooka, A. Tawada, A. Iwama, O. Yokosuka. Biological features and biomarkers in hepatocellular carcinoma. *World Journal of Hepatology*. 7(16): 2020-2028, 2015.

Tong S., Li J., Wands J.R. and Wen Y.M. 2013. Hepatitis B virus genetic variants: biological properties and clinical implications. *Emerging Microbes and Infections*. e10; doi:10.1038/emi2013.10.

Tu H., Bonura C., Giannini C., Mouly H., Soussan P., Kew M., Paterlini- Brechot P. *et al.* 2001. Biological impact of natural COOH-terminal deletions of hepatitis B virus X protein in hepatocellular carcinoma tissues. *Cancer Research*. 61: 7803−7810.

Venard V., Corsaro D., Kajzer C., Bronowicki J.P. and Le Faou A. 2000. Hepatitis B virus X gene variability in French-born patients with chronic hepatitis and hepatocellular carcinoma. *Journal of Medical Virology*. 62: 177−184.

Wang F., Zhuo H., Yang Y., Xia X., Sun Q., Luo J. and Cheng B. 2012. Hepatitis B virus X protein promotes the growth of hepatocellular carcinoma by modulation of the Notch signaling pathway. *Oncology Reports*. 27(4):1170-1176.

Wang W.W., Ang S.F., Kumar R., Heah C., Utama A. 2013. Identification of Serum Monocyte Chemoattractant Protein-1 and Prolactin as Potential Tumor Markers in Hepatocellular Carcinoma. *PLoS ONE*. 8(7): e68904, doi:10.1371/journal.pone.0068904.

Wang X.Z., Chen X.C., Chen Y.X., Zhang L.J., Li D., Chen F.L., Chen Z.X., Chen H.Y. and Tao Q.M. 2003. Overexpression of HBxAg in hepatocellular carcinoma and its relationship with Fas/FasL system. *World Journal of Gastroenterology*. 9(12):2671-2675.

Woodcock D.M., Crowther P.J., Doherty J., Jefferson S., DeCruz E., Noyer-Weidner M.,

World Health Organization, (2008). Hepatitis B. Available at: http://www.who.int/mediacentre/factsheets/fs204/en/.

Xiao-Z.W., Xiao-C.C., Yun-X.C., Li-JZ., Dan L., Feng-L.C., Zhi-X.C. 2003. Overexpression of HBxAg in hepatocellular carcinoma and its relationship with Fas/FasL system. *Journal of Gastroenterology*. 9(12):2671-2675.

Yi-Fang H., Zhao J., Ma L.Y., Yin J.H, Chang W.J, Zhang H.W and Cao J.W. 2011. Factors predicting occurrence and prognosis of hepatitis-B-virus-related hepatocellular carcinoma. *World Journal of Gastroenterology*. 17(38): 4258-4270, 2011.

Yoo Y.D., Ueda H., Park K., Flanders K.C., Lee Y.I., Jay G., *et al.* 1996. Regulation of transforming growth factor-beta 1 expression by the hepatitis B virus (HBV) X transactivator- Role in HBV pathogenesis. *The Journal of Clinical Investigation*. 97(2):388–395.

Zhang Z., Protzer U., Hu Z., *et al.* 2004. Inhibition of cellular proteasome activities enhances hepadnavirus replication in an HBX-dependentmanner. *Journal of Virology*. 78(9): 4566–4572.

Zhang Z.H., Wu C.C., Chen X.W., Li X., Li J. and Lu M.J. 2016. Genetic variation of hepatitis B virus and its significance for pathogenesis. *World Journal of Gastroenterology*. 22(1): 126-144.

Zhu P., Tan D., Peng Z., Liu F. and Song L. 2007. Polymorphism Analyses of Hepatitis B Virus X Gene in Hepatocellular Carcinoma Patients from Southern China. *Acta Biochimica et Biophysica Sinica*. 39(4): 265-272.

Zhu R., Mok M.T.S., Kang W., Lau S.S.K. *et al*. 2015. Truncated HBx-dependent silencing of GAS2 promotes hepatocarcinogenesis through deregulation of cell cycle, senescence and p53-mediated apoptosis. *Journal of Pathology*.**237:** 38–49.

ANNEXURE 1

clustalw.aln

CLUSTAL 2.1 multiple sequence alignment

```
L22             -------------------------TGCTATTGATCTTGCTCGGCCTACANGGCCGCAA
D13             TATGCCCTGGGAACTTTTGTTACCTCTGGATTCTAACCGGGACGTCCTTTGTTTCCGCAA
HBx             -ATGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCC
D1              --TACAATGTAAATTTTTTTTACCTCTG---CCTAAGCGGCACCTCTTTTGTTTACGTCC
L6              -ATGGCTGCTAGGCTGGGTTACCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCC
D21             ---AGCTGGTAGGCTGTGCTGACAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGCCC
L12             --------ATAGGCTGTGCTGCCAAACGGATCCTGCGCTGGACGTCCTTTNTTCCAGTCC
L9              -ATGGCTGATAGGCTGTGTCGCCTACAGGATCCTGCGCGGGACGTCCTTTGTTTACGTCC
D2              --TGGCTGCTAGGCTGTGCTGCCAACTGGATCCTGCGCGGGACGTCCTTTGTTTACGTCC
L18             -GTAATTTTTTGTTACGTCTGCCTAAGTGGCCGCTAACGGGACGTCCTTTGTTTACGTCC
L2              -----------GAATTGGTGTTTCCTTGTGGCTTTGAGGGATTTTTTTTGTTGCGTACA
D16             ---TACTGCTAGGGTGCGCTGTACAGTGGATCCTGCGCCCCTNTTCCTTAGTTTTCGACC
L11             TATGGGTTCTAGGCTGTGCGCCCATATGGAGCCNGCGCGGGACGTCCTTTGTTTACGTCC
D8              -ATGGCAACTTGGAGATTTTTGTTACGTTTGTCTTAGTTTGCCCCTACATTTGTTCGTCC
                                                          :  .     ..

L22             TANGTCT---------CACGGACGACCCTTC-TCGGGGCCGCAATGGGACTTTCTCGTCC
D13             CAGATCCGCATAATCCCAGGGACGACCCTAT-TCGGGGCCGCTG-GGGACTTTCTCGACC
HBx             CGTCGGCGCTGAATCCTGCGGACGACCCTTC-TCGGGGTCGCTT-GGGACTCTCTCGTCC
D1              CGTCTGCGCTTTTTCCCGCGGACCCCCCTAT-TCGGGGGCGCTT-GGGACTCTCTCGTCC
L6              CGCCGGCGCTAAATCCCGCGGACGACCCTACGTCGGGGCCGCTT-GGGACTCTCTCGTCC
D21             CGTCGGCGCTGAATCCCGCGGACGACCCTTC-TCGGGGCCGCTT-GGAACTCTCTCGTCC
L12             CGTCGGCGCTAAATCCCGCGGACGACCCTTC-TCGGGGCCGCTT-GGGACTCTCTCGTCC
L9              CGTCGGCGATAAATCCCGCGGACGACCCTAC-TCGGGGCCGCTT-GGGACTCTCTCGTCC
D2              CGTCGGCGCTAAATCCCGCGGACGACCCTTC-TCGGGGCCGCTT-GGGACTCTCTCGTCC
L18             CTTCCGCGCTAATCCCAGCGGACCACCCTAC-TCGGGGCCGCTT-GGGACTCTCTCGTCC
L2              GGGCCGCGACAAATCCCGCGGACGACCCTAC-TCGGGGCCGCTT-GGGACTCTCTCGTCC
D16             CGTCG--CGCTGAATCAGCGGACGATTTTTC-TCGGTGGCGCTN-GGGACTCTCTCCCCC
L11             CGTAGGCGCTGAATCCCACGGACGACCCTAC-TCGGGGCCGCTT-GGGACTCTCTCGTCC
D8              CGNCGGCGCACAATCCCACGGACGACCCTAC-TCTTTGCCGATT-GGCACTCTCTCGTCC
                          . **** .  *: **  * **.: ** *** ****  **
```

```
L22             CCTTCTCCGTCTGCCGTTTCGCCCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCC
D13             CCTTCTCCGTCTGCCGTTTCTTCCGTCTGTGGGGGGCAGCTCTGTTTATGCGGACTCCCC
HBx             CCTTCTCCGTCTGCCGTTCCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCC
D1              CTTTCTCCCTCTGCCGTTTCGACCGACCATGGGGGGCACCTCTCTTTACCCGGACTCCCC
L6              CCTTCTCCGTCTGCCGTTTCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCC
D21             CCTTCTCCGTCTGCCGTTTCGACCGACCACGGGGCCCACCTCTCTTTACNCGGACTCCCC
L12             CCTTCTCCGTCTGCCGTTTCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCC
L9              CCTTCTCCGTCTGCCGTTTCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCC
D2              CCTTCTCCGTCTGCCGTTTCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCC
L18             CCTTCTCCGTCTGCCGTTTCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCC
L2              CCTTCTCCGTCTGCCGTTTCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCC
D16             CCCTCTCCGGNTGCCCCTCCGACC-ACCGCGAGGCGCACCTCTCTCGACGNTGACTGGCA
L11             CCTTCTCCGTCTGCCGTTTCGACCGACCACGGGGCGCACCTCTCTTTACGCGGACTCCCC
D8              CCTTCTCCGTCTGCCGTGCCACCGGAACTCGGGGCGCACCTCTCTTTATGCTGACTCCCC
                *  *****  ****   * * :.  *.** ** **** *  *   **** *.

L22             GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATG
D13             TTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCATGTCGCATG
HBx             GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATG
D1              GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCCCTTCACCTCTGCTTGTCTCATG
L6              GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATG
D21             GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCCCTTCACCTCTGCACGTCGCATG
L12             GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATG
L9              GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATG
D2              GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATG
L18             GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATG
L2              GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCCCATG
D16             GTCTGGGCCTTTTCATCTGTAGGACCGTGTGCAGTTCGGTTCNTTTCGGCACGTCTCAGG
L11             GTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCACCTCTGCACGTCGCATG
D8              GTCTGTGCCTTCTCTTCCCCTGGACCGTGTGCACTTCGCTTCCTCTCTGCACGTCGCATG
                 **** ***** **:**   ********** ***  ***   ** **: *** ** *
```

```
L22     GAGACCACCGTGAACGCCCACCAATTAGACCCCAAGGTCTTACATAAGAGGACTCTTGNA
D13     GAGACCACCGTGAACGCCCACCAAATCTTGCCTAAGGTCTTACATAAGAGGACTCTTGGA
HBx     GAGACCACCGTGAACGCCCACCAAATATTGCCCAAGGTCTTACATAAGAGGACTCTTGGA
D1      GAGACCTCCGTGATCCCCACCAATTCCTTGCCTAAGGTCTCACATAAGAGGACTCTTGGT
L6      GAGACCACCGTGAACGCCCACCAATTCTTGCCTAAGGTCTTACATAAGAGGACTCTTGGA
D21     GAGACCACCGTGAACCCCCACCAATTCTTGCCTAAGGTCTTACATGAGAGGACTCTTGGA
L12     GAGACCACCGTGAACGCCCACCAATTCTTGCCTAAGGTCTTACATAAGAGGACTCTTGGA
L9      GAGACCACCGTGAACGCCCACCAATTCTTGCCTAAGGTCTTACATAAGAGGACTCTTGGA
D2      GAGACCACCGTGAACGCCCACCAATTCTTGCCTAAGGTCTTACATAAGAGGACTCTTGGA
L18     GAGACCACCGTGAACGCCCACCAATTCTTGCCTAAGGTCTTAATAAGGAGGACTCTTGGA
L2      GAGACCACCGTGAACGCCCACCAATTCTTGCCTAAGGTCTTACATAAGAGGACTCTTGGA
D16     GAGACCGCCGTGAACCCCCATTAGCTCTTGTATAAGGTCTTAGATAAGAGGACTCTTGGA
L11     GAGACCACCGTGAACGCCCACCAATTCTTGCCTAAGGTCTTACATAAGAGGACTCTTGGA
D8      GACACCACCGAGAACGCCCACCAATTCTTGCCTAAGGTCTTACTTAACAGGACTCTTGGA
        ** *** ***:**:* **.. *  .:  . ******* * :::. *********.:

L22     CTCTCGGTAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAACAA
D13     CTCTCTGTAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAGGAC
HBx     CTCTCAGCAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAAGAC
D1      CTCTCTGTCATGTCTTCGACCGACCTTTAGGCTTACTTCAATGACTGTTTGTTAAGGACT
L6      CTCTCTGTAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAGGAC
D21     CTCTCTGTAATGTCACCGACAGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAGGAC
L12     CTCTCTGTAATGTCAACGACCGACCTTGNGGCATACTTCAAAGACTGTTNGTTTAAGGAC
L9      CTCTCTGTAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAGGAC
D2      CTCTCTGTAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAGGAC
L18     CTCTCTGTAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAGGAC
L2      CTCTCTGTAATGTCAACGACCGACCTTGAAGCATACTTCAAAGACTGTTTGTTTAAGGAC
D16     CTTTCTGTAATGTCAGTGAACGACGTTGACGCATTTTTAAAG-ACTGTTTGTTCAATGAC
L11     CTCTCTGTAATGTCAACGACCGACCTTGAGGCATACTTCAAAGACTGTTTGTTTAAGGAC
D8      CTCTCTGTAATGTCCACGACCGACCTTGAGGCATACTTCCAAGACTGTTTGTTTAAGGAC
        ** ** * .*****  **.:*** **  **:*: **.:*  ******.*** *.  .
```


L22 TGGGAG-----GAGGGAGCCGGAGGGAAAAAATGCTCTTGAATTGGG--GATGGAAACTA
D13 TGGGAGGAGTTGGGGGATCAGGTTAGGTTTATGGTTCTTGAACTAG---TCTGTANNTTT
HBx TGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGTACTAGGAGGCTGTAGGCAT
D1 GGGGGGCAGTTGGGCGAGGAGATTTGATTAAAGGTTCTTGAATTAG----CTATAGGTGT
L6 TGGGAGGAGTTGGGGGAGGAGATTAGATTAAAGGTCTTTGTATTAGGAAGCTGTAGGCAT
D21 TGGGAGGAGTTGGGGGAGGAGATTAGATTAAAGGCCTTTGTATTAGGAGGCTGTAGGNAT
L12 TAGGAGGAGTTGGGGGAGGAGATTAGATTAAAGGTCTTTGTATTAGGAGGCTGTAGGCAT
L9 TGGGAGGAGTTGGGGGAGGAGATTAGATTAAAGGTCTTTGTATTAGGAGGCTGTACGCAT
D2 TGGGAGGAGTTGGGGGAGGAGATTAGATTAAAGGTCTTTGTATTAGGAGGCTGTAGGCAT
L18 TGGGAGGAGTTGGGGGAGGAGATTAGATTAAAGGTCTTTGTATTAAGAGGCTGTAGGCAT
L2 TGGGAGGAGTTGGCGGAGGAGATTAGATTAGAGGTCTTTGTATTAGGAGGCTGTAGGCAT
D16 TGACTGGAGGTGGGGGAGGAGATTAGATCNGAGGNCTTTCTATTAGGAGGCTGGAGGCAT
L11 TGGGAGGAGTTGGGGGAGGAGATTAGATTAAAGGTCTTTGTATTAGGAGGCTGTAGNCAT
D8 TGGGAGGAGTTGGGGGAGGAGATTAGATTACAGGTCTCCCTATTAGGAGGATTCATCTAT
GGTTGGCTCCGAAAACGAAAATCCAAAAACT----------
TTATTGGTCTGACACCAGGCNGCCCNACGA----------
AAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAA------
TAG----------
AAATTGGTCTGCACACCAGCACCATGCAACTTTTTCACCTCTGCCT--------
AAATTGGTCTGCGTCCCAGCACCATGGAACTTTTTCACCTCTGCCGCGGCCGTA
AAATTGGTCTGCGCACCATCACCATGCAACTTTTTCACCTCTGCCTA-------
AAATTGGTCTGCGCACCAGCACCACGCAACTTTTTCACCTCTGCAAAA------
AAATTGGTCTGCGCACCANCACCATGCAACTTTTTCACCTCTGCCTAA------
AAATTGGTCTGCGCACCAGCACCATGCAACCTTTTCACCTCTGCCTAA------
AAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCCTAA------
AAATTGGTTTGCTCACCATTACCCTGCAACGCTTTCCACTCTCCCGAA------
AAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCC---------

Annexure 2

clustalw.aln

```
CLUSTAL 2.1 multiple sequence alignment


L22             -----------CYSCSAYXAAIXLTDDPSRGRNGTFSSPSPSAVSPDHGAHLSLRGLPV
D13             -MPWELLLPLDSNRDVLCFRNRSASQGRPYSGPLGTFSTPSPSAVSSVCGGQLCLCGLPF
L12             ---RLCCQTDPALDVLXSSPVGAKSRGRPFSGPLGTLSSPSPSAVSTDHGAHLSLRGLPV
D2              -AARLCCQLDPARDVLCLRPVGAKSRGRPFSGPLGTLSSPSPSAVSTDHGAHLSLRGLPV
Consensus       MAARLCCQLDPARDVLCLRPVGAESCGRPFSGSLGTLSSPSPSAVPTDHGAHLSLRGLPV
D21             --AGRLCQLDPARDVLCLRPVGAESRGRPFSGPLGTLSSPSPSAVSTDHGAHLSLXGLPV
L9              MADRLCRLQDPARDVLCLRPVGDKSRGRPYSGPLGTLSSPSPSAVSTDHGAHLSLRGLPV
L11             MGSRLCAHMEXARDVLCLRPVGAESHGRPYSGPLGTLSSPSPSAVSTDHGAHLSLRGLPV
L2              ----NWCFLVALRDFFCCVQGRDKSRGRPYSGPLGTLSSPSPSAVSTDHGAHLSLRGLPV
L18             -VIFCYVCLSGRRDVLCLRPFRANPSGPPYSGPLGTLSSPSPSAVSTDHGAHLSLRGLPV
L6              WLLGWVTNWILRGTSFVYVPPALNPADDPTSGPLGTLSSPSPSAVSTDHGAHLSLRGLPV
D8              -MATWRFLLRLSFAPTFVRPXGAQSHGRPYSLPIGTLSSPSPSAVPPELGAHLSLC-LPV
D1              --QCKFFLPLPKRHLFCLRPVCAFSRGPPYSGALGTLSSLSPSAVSTDHGGHLSLPGLPV
D16             -TARVRCTVDPAPXFLSFRPVALNQRTIFLGGAX--DSLPPLSXCPSDHREAHLSRXLAV
                                                 *  . *  ..         *..

L22             CAFSSAGPCALRFTSARRMETTVNAHQLDPKVLHKRTLXLSVMSTTDLEAYFKD-CLFKQ
D13             CAFSSAGPCALRFTSACRMETTVNAHQILPKVLHKRTLGLSVMSTTDLEAYFKD-CLFKD
L12             CAFSSAGPCALRFTSARRMETTVNAHQFLPKVLHKRTLGLSVMSTTDLXAYFKD-CXFKD
D2              CAFSSAGPCALRFTSARRMETTVNAHQFLPKVLHKRTLGLSVMSTTDLEAYFKD-CLFKD
Consensus       CAFSSAGPCALRFTSARRMETTVNAHQILPKVLHKRTLGLSAMSTTDLEAYFKD-CLFKD
D21             CAFSSAGPCALPFTSARRMETTVNPHQFLPKVLHERTLGLSVMSPTDLEAYFKD-CLFKD
L9              CAFSSAGPCALRFTSARRMETTVNAHQFLPKVLHKRTLGLSVMSTTDLEAYFKD-CLFKD
L11             CAFSSAGPCALRFTSARRMETTVNAHQFLPKVLHKRTLGLSVMSTTDLEAYFKD-CLFKD
L2              CAFSSAGPCALRFTSARPMETTVNAHQFLPKVLHKRTLGLSVMSTTDLEAYFKD-CLFKD
L18             CAFSSAGPCALRFTSARRMETTVNAHQFLPKVLIRRTLGLSVMSTTDLEAYFKD-CLFKD
L6              CAFSSAGPCALRFTSARRMETTVNAHQFLPKVLHKRTLGLSVMSTTDLEAYFKD-CLFKD
D8              CAFSSPGPCALRFLSARRMDTTENAHQFLPKVLLNRTLGLSVMSTTDLEAYFQD-CLFKD
D1              CAFSSAGPCALPFTSACLMETSVIPTNSLPKVSHKRTLGLSVMSSTDL-AYFND-CLLRT
D16             WAFSSVGPCAVRFXSARLRETAVNPHLLY-KVLDKRTLGLSVMSVNDVDAFLKTVCSMTD
                 **** ****: * **   :*:  .     **  .*** **.** .*: *:::  * :
```

```
L22             WEEGAGGKKCSIGDGNVGSENENPKT------------
D13             WEELGDQVRFMVLELVCXFLLVHQXAXR----------
L12             -EELGEEIRLKVFVLGGCRHKLVCAPSPCNFFTSA---
D2              WEELGEEIRLKVFVLGGCRHKLVCAPXPCNFFTSA---
Consensus       WEELGEEIRLKVFVLGGCRHKLVCAPAPCNFFTSA---
D21             WEELGEEIRLKAFVLGGCRXKLVCVPAPWNFFTSAAAV
L9              WEELGEEIRLKVFVLGGCTHKLVCAPAPRNFFTSAK--
L11             WEELGEEIRLKVFVLGGCXHKLVCAPAPCNFFTSA---
L2              WEELAEEIRLEVFVLGGCRHKLVCAPAPCNFFTSA---
L18             WEELGEEIRLKVFVLRGCRHKLVCAPAPCNLFTSA---
L6              WEELGEEIRLKVFVLGSCRHKLVCTPAPCNFFTSA---
D8              WEELGEEIRLQVSLLGGFIY------------------
D1              GGQLGEEILKVLELAIGV--------------------
D16             WRWGRRLDXRXFYEAGGINWFAHHYPATLSTLP-----
```

---

clustalw.dnd