

## RESEARCH ARTICLE

## Genetic diversity in enhancer II region of HBV genotype D and its association with advanced liver diseases

Majid Khan<sup>1</sup>, Sanaullah Khan<sup>1\*</sup>, Mehrunnisa Fatima Gondal<sup>2</sup>, Safia Bibi<sup>3</sup>, Bakht Tarin Khan<sup>1,4</sup>, Abdul Majid<sup>2</sup>, Ayesha Khattak<sup>1</sup>, Muhammad Nasir Khabir<sup>1</sup>, Muhammad Anwar<sup>1</sup>, Aisha Gul<sup>1</sup>, Maryam Naseem<sup>1</sup>, Sobia Attaullah<sup>5</sup>

**1** Department of Zoology, University of Peshawar, Peshawar, Khyber Pakhtunkhwa, Pakistan, **2** Rural Health Centre Barki, District Lahore, Pakistan, **3** Department of Zoology, Kohat University of Science and Technology Kohat, Kohat, Khyber Pakhtunkhwa, Pakistan, **4** Department of Zoology, University of Buner, Buner, Khyber Pakhtunkhwa, Pakistan, **5** Department of Zoology, Islamia College Peshawar University, Peshawar, Khyber Pakhtunkhwa, Pakistan

\* [sanaullahkhan@uop.edu.pk](mailto:sanaullahkhan@uop.edu.pk)



## OPEN ACCESS

**Citation:** Khan M, Khan S, Gondal MF, Bibi S, Khan BT, Majid A, et al. (2022) Genetic diversity in enhancer II region of HBV genotype D and its association with advanced liver diseases. PLoS ONE 17(1): e0261721. <https://doi.org/10.1371/journal.pone.0261721>

**Editor:** Shah Jahan, University of Health Sciences Lahore, PAKISTAN

**Received:** July 5, 2021

**Accepted:** December 7, 2021

**Published:** January 4, 2022

**Copyright:** © 2022 Khan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and the minimum data set were submitted to GenBank with accession numbers, MZ666825, MZ676992, MZ676993, MZ676994, MZ676995, MZ711089, MZ711090, MZ711091, MZ711092.

**Funding:** The author received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

## Background

Hepatitis B Virus (HBV) is one of the most common human infectious agents, and the mutations in its genome may cause chronic hepatitis (CH), liver cirrhosis (LC), and hepatocellular carcinoma (HCC). This study was designed to characterize the enhancer-II (Enh-II) region of X gene in HBV positive patients to assess the association of such mutations with CH, LC, and HCC.

## Methods

HBV positive samples (N = 200) with patients' demographic and clinical data were collected from different regions of Khyber Pakhtunkhwa (KP), Pakistan. The Enh-II region of the HBx gene was sequenced and analyzed for polymorphism associated with advanced liver disease. Univariate and logistic regression analyses were performed to evaluate potent mutations associated with a risk for LC and HCC.

## Results

HBV Enh-II region sequences analysis revealed 25 different mutations. The highest frequency of mutations S101F (62.2%), A102V/R/G/I (56.25%), M103L/A (68.75%) were found in HCC, followed in LC and CH patients as 57.1%, 42.8%, 28.52% 16%, 15.2% and 18.4% respectively. H94 deletion in the  $\alpha$ -box of the Enh-II region, associated with a high risk of HCC was found in half of the HCC patients. This deletion was present in 28.5% of LC and 6.5% of CH patients. Importantly, the high frequency of some notable mutations such as E109A/Y, A110S/K, Y111D/E, and F112L was first time reported in the entire study population. The frequencies of these mutations were high in HCC (43.75%, 37.5%, 50% and

43.75% respectively) as compared to LC (14.28%, 14.28%, 28.2% and 42.8%) and CH patients (12.8%, 15.2%, 16.8% and 16% respectively).

## Conclusion

Mutations associated with LC and HCC are prevalent in the Enh-II region in Pakistani HBV isolates. The mutations found are alarming in CH patients as these may progress to LC and HCC in a large number of patients.

## Introduction

HBV has been proved to be the most important virological factor for the development of advanced liver disease as LC and HCC. The replication capabilities of the virus and the accumulation of potent mutations in the HBV genome, are common events for HCC development, and thus considered a life-threatening pathogen leading to significant rates of mortality worldwide [1, 2]. According to WHO, 257 million people are living with HBV infection with an estimated number of 887,000 deaths in 2015 attributed to HBV complications [3].

The infectious virion has an internal capsid (core particle) that protects the partially double-stranded DNA genome. HBV genome consists of four overlapping reading frames (ORFs) corresponding to four major genes—polymerase (P), surface (S), core (C), and regulatory (X). Hepatitis B X (HBx) is the smallest of HBV genes, encoding for 154 amino acids that constitute 17.5 kilodalton protein. HBx is involved in transcription regulation and interacts with a variety of cellular pathways responsible for oncogenesis [4]. The HBV genome contains two transcriptional enhancer (Enh) regions i.e. Enh-I and Enh-II, which are important for viral replication. The Enh-II, located in nt 1636–1741, overlaps the HBx coding region and stimulates the transcriptional activity of the surface, core, and HBx gene promoters which shows high hepatocyte-specific activity [5]. Enh-II is further composed of two regions i.e. box- $\alpha$  (nt 1644–1666) and box- $\beta$  (nt 1702–1713) [6]. Enh-II region of the HBx plays a key role in controlling the transcription of all genes in a hepatocyte-specific manner. Moreover, Enh-II plays an important role in the viral life cycle by regulating the formation of the 3.5 kb pre-genomic RNA which is then translated to produce the viral core and polymerase proteins and HBeAg [7].

Enh-II stimulates the transcription of surface protein-I (SP1), surface protein-II, and X protein in a position and orientation-independent manner [8]. On the other hand, this element regulates the basal core proteins (BCP) region in a position and orientation-dependent manner, functioning as a core upstream regulatory sequence, which overlaps the sequence of Enh-II [9]. Hepatocyte-enriched transcription factors, such as hepatocyte nuclear factor1 (HNF1) [10], HNF3 [11, 12], HNF4 [13, 14], CCAAT/enhancer-binding protein (C/EBP) [15–17], and fetoprotein transcription factor (FTF) [18], which are known to be responsible for the hepatocyte-specific Enh-II activity, and regulate its function [19]. Different mutations in the Enh-II region (as T1653 mutations in box- $\alpha$ ) have been identified which are responsible for the development of HCC. Most of the studies primarily focused on certain single point nucleotide mutations, however, it is unclear whether a single or specific combination of these mutations is associated with the development of HCC [20].

Pakistan is highly endemic for HBV with approximately nine million people infected and the infection rate is on a steady rise [21, 22]. Even though in Pakistan data related to HBV and its genotype prevalence are available but studies that cover information of HBV mutations especially in the Enh-II region and its clinical association with the development of HCC are

scarce. This study is therefore designed to assess the diversity of mutations associated with LC and HCC in Enh-II regions in HBV positive patients of Khyber Pakhtunkhwa (KP) province, Pakistan.

## Materials and methods

### Study population

Samples were collected randomly from clinically confirmed HBV-positive patients from five districts (Peshawar, Mardan, Dir Lower, Kohat, Bannu) of KP, Pakistan.

### Samples and data collection

Blood (3mL) was collected in a sterile tube from HBs Ag-positive patient. Clinical data were obtained through a predesigned questionnaire. The samples were transported to the Laboratory of Virology and Immunology, Department of Zoology University of Peshawar for further process. Proper consent was obtained from each participant before including in the study and the study protocol was approved by the Research Ethical Committee, University of Peshawar, Pakistan.

### Inclusion/Exclusion criteria

Patients positive for HBsAg through Immunochromatography (ICT) test was included while those with HIV and HCV antibody positivity and whose HBV DNA was undetectable by PCR were excluded from the study.

### DNA extraction

HBV DNA was extracted by using DNA isolation kits (TRIzol™ Reagent, ThermoFisher scientific) as per the manufacturer's instructions.

### HBV DNA and genotype/s detection

HBV DNA was detected through the protocol of Kyaw et al [23]. HBV genotype/s (A-F), were identified using the protocol of Naito et al [24]. Briefly, the first round of PCR was performed for the HBV S gene using, 10 pM of primers (P1 and S1-2) and green Master Mix having a final volume of 20 µL. The thermocycler conditions were adjusted to 10 minutes of denaturation at 95°C followed by amplification for 35 cycles (94°C for 20 sec, 55°C for 20 sec, and 72°C for 1 minute), and a final extension at 72°C for 7 minutes. Second-round PCR primers were mixed in two separate mixes as Mix-A and Mix-B depending on their targets and amplification size differences.

### HBV X gene amplification and detection

HBx gene was amplified through nested PCR using the protocol of Al-Qahtani et al [25]. Briefly, first-round PCR consisted of 10 µL of extracted DNA, 20 pM of each primer F1: 5' ATTGATTGGAAAGTMTGTM 3' and R1: 5' TCCACAGTAGCTCCAAATTCCTT 3' primer, and Green Master Mix (20 µL). The thermocycler conditions were initial 5 minutes of denaturation at 94°C followed by amplification for 30 cycles (94°C for 45 sec, 55°C for 30 sec, and 72°C for 1 minute), and a final extension at 72°C for 10 minutes. The second round of PCR primers were F2: 5' CGCTTGTTTTGCTCGCAGC 3' and R2: 5' GGCACAGCTTGAGGCTTG 3' and the reaction and PCR conditions were the same as for the previous round.

## Gel electrophoresis

PCR amplified product was electrophoresed on 2% agarose gel, stained with ethidium bromide, and compared with 100bp DNA ladder (ThermoFisher USA). The amplified product was visualized under an ultraviolet (UV) trans-illuminator.

## Sequencing of DNA

Amplified HBV X gene was sequenced through Sanger method from commercially available services. The obtained sequences were submitted to GenBank with accession numbers, MZ666825, MZ676992, MZ676993, MZ676994, MZ676995, MZ711089, MZ711090, MZ711091, MZ711092.

## Trimming of enhancer II region

Redundant sequences present upstream and downstream of the Enh-II region were trimmed using BioEdit software and final sequences were obtained.

## Sequence analysis

Sequences were analysed for potent mutations using online sequence alignment tools like BLAST (<https://blast.ncbi.nlm.nih.gov/BlastAlign.cgi>). BioEdit software was used for the alignment of query sequences with the subject and different genotypes (A-J) sequences to compare similarities and differences.

## Phylogenetic analysis

Phylogenetic analysis was performed by comparing the obtained gene sequences with HBV reference sequences for each genotype available in Genbank. A phylogenetic tree was constructed in MEGA software (version 6.1) using the neighbor-joining method. Tamura Nei Model method was used to estimate the distances. Bootstrap statistical analysis was performed using 1000 datasets.

## Data analysis

Statistical software SPSS version 11.5, package (SPSS Inc, Chicago, IL, USA) was used for statistical analysis of the data. The data were compared by applying the student's t-test and statistical significance was given to the groups having p-values of <0.05.

## Results

### Demographics of the study population

Different baseline characteristics such as age, sex, HBeAg status of these positive samples are summarized in [Table 1](#). Phylogenetic analysis of HBV showed that all patients (100%) had genotype D ([Fig 1](#)). Among 200 HBsAg positive HBV patients, 148 had PCR confirmed HBV infection while 52 were PCR negative. Out of these 148 PCR positive samples, 125 (84.45%) had CH, 16 (10.81%) had HCC and 7 (4.72%) had LC. Gender wise study of HBV infection revealed a higher frequency of HCC in males (63.2%) patients as compared to females (36.8%) and was most prevalent in age group  $\leq 40$  years (60.8%), 40–49 years (30.4%), 50–59 years (17.6%) and  $>60$  years (9.6%). The p-value was significant for all three groups (CH, LC, and HCC).

**Table 1. General characteristics of HBV patients.**

Basic characteristics	CH N (%)	HCC N (%)	LC N (%)	p-value
<b>Age in years</b>				
≤40	64 (51.2)	8 (50.0)	3 (42.8)	0.0238
40–49	31 (24.8)	5 (31.2)	2 (28.6)	
50–59	19 (15.2)	2 (12.5)	1 (14.3)	
≥60	11 (8.8)	1 (6.25)	1 (14.2)	
<b>Sex</b>				
Male	79 (63.2)	10 (62.5)	5 (71.4)	0.9041
Female	46 (36.8)	6 (37.5)	2 (28.6)	
<b>HBeAg</b>				
Positive	84 (67.2)	9 (56.2)	6 (85.7)	0.3785
Negative	41 (32.8)	7 (43.7)	1 (14.3)	

<https://doi.org/10.1371/journal.pone.0261721.t001>

### Mutations in the Enh-II region

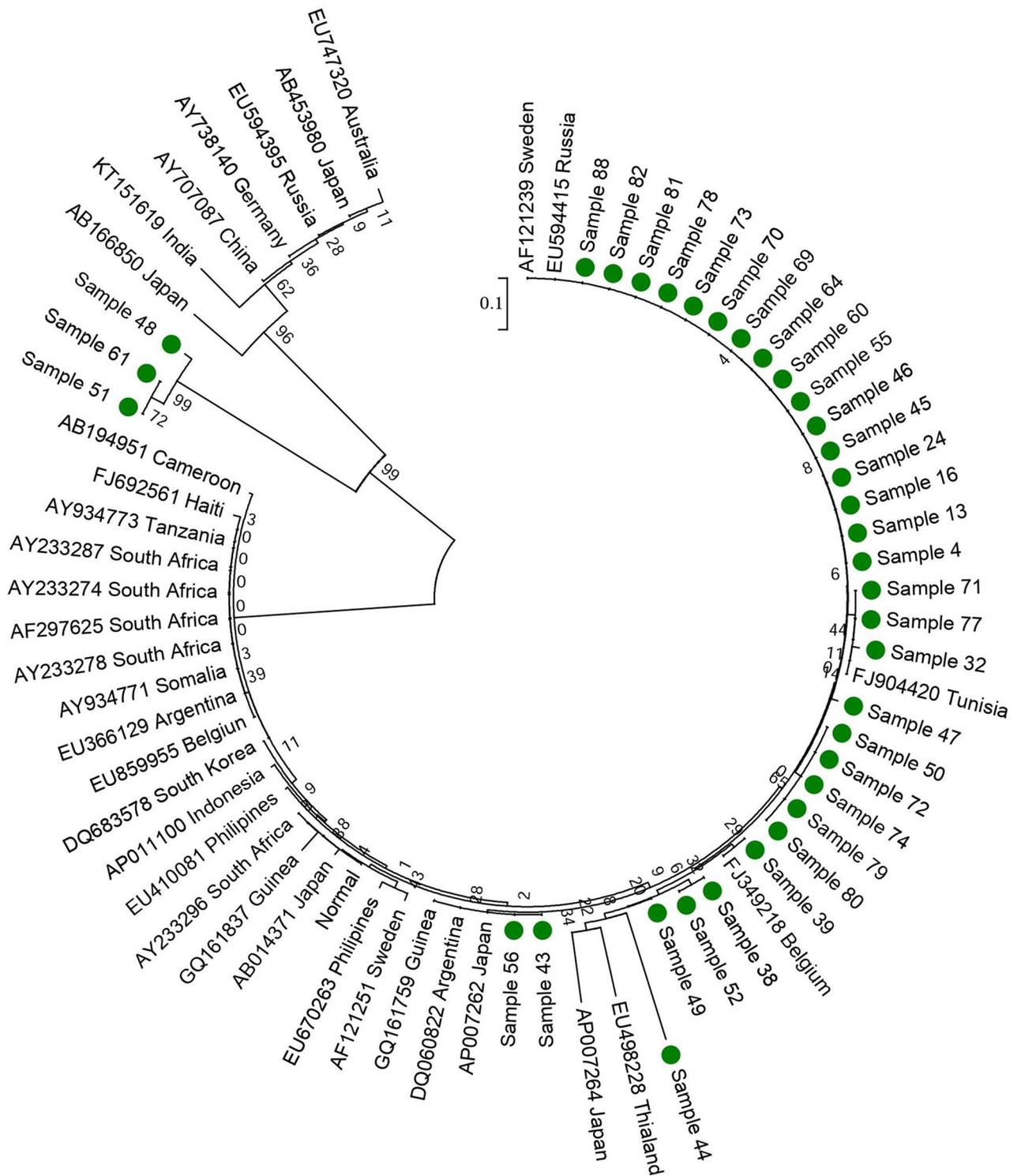
Twenty-five different substitution and deletion mutations were observed in HBV-positive individuals in this study which are illustrated in Table 2. Sequencing and data analysis showed the highest frequency of S101F (62.2%), A102V/R/G/I (56.25%), and M103L/A (68.75%) mutation in HCC and 57.1%, 42.8%, and 28.52% respectively in LC as compared to 16%, 15.2%, and 18.4% respectively in CH patients. H94 deletion in the  $\alpha$ -box of the Enh-II region has been reported to increase the risk of HCC and its frequency was higher in HCC (50%) as compared to and LC (28.5%) and CH group (6.5%) in this study. Deletion mutations of various sizes were also reported in this study. Moreover, some non-reported mutations such as E109A/Y (43.75%), A110S/K (37.5%), Y111D/E (50%), and F112L (43.75%) were also found in high frequency in HCC and LC (14.28%, 14.28%, 28.2% and 42.8% respectively) patients as compared to CHB (12.8%, 15.2%, 16.8% and 16% respectively) (Figs 2 and 3).

### Mutation in transcription factors controlling Enh-II activity

In this study mutations in hepatocyte-enriched transcription factors HNF1, HNF3, HNF4, FTF (hB1F), C/EBP,  $\alpha$ -box, and  $\beta$ -box were also found which are shown in Fig 4. The highest frequency of mutation was found in  $\alpha$ -box followed by the HNF4 site 1 while the lowest frequency of mutation was observed in HNF1 site (Fig 5).

### Single, double and triple nucleotide(s) mutations causing amino acid substitution in CH, LC, and HCC patients

The significance of mutations related to HCC progression was evaluated by Univariate and multivariate logistic regression analysis among CH, LC, and HCC groups. Nine single nucleotide mutations were found in the Enh-II region, including the highest incidence rate of mutation L108D, Y111D, K113E in HCC as compared to other groups. Similarly, 9 double nucleotide mutations were also detected in different regions of the Enh-II region having A102G, M103A, and K118I in HCC compared to CH and LC. Fourteen triple nucleotide mutations were found among which the more frequent mutation in LC and HCC groups were A110S/K, Y111E, and F112L while K91W/R and R96C were found in the lowest frequency (Figs 6–8).



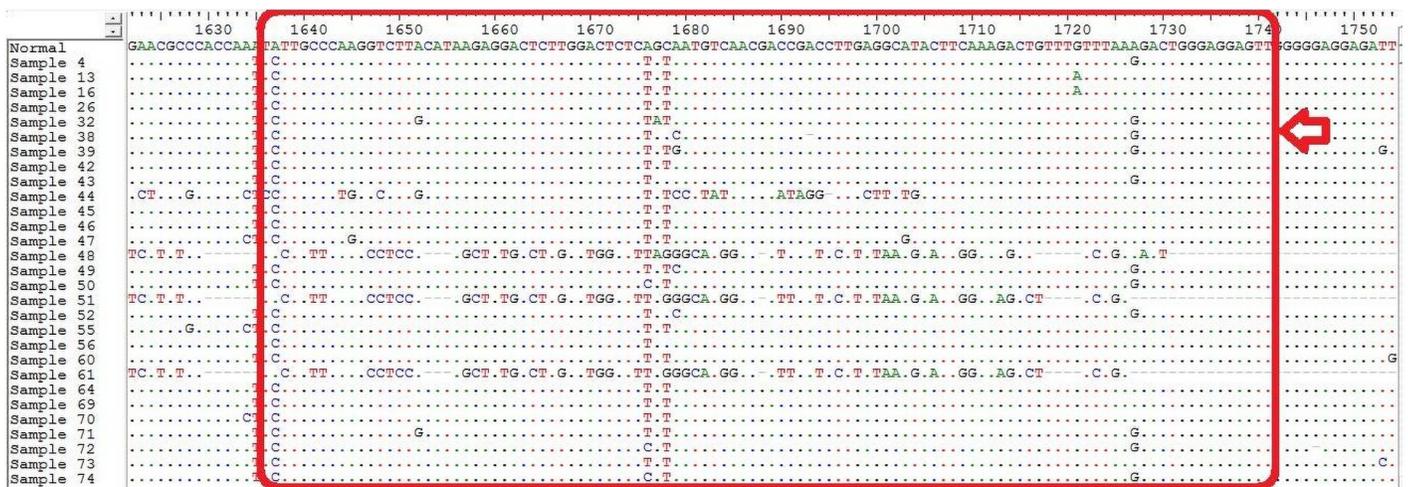
**Fig 1. A rooted phylogenetic tree of 38 enhancer II region sequences obtained from Pakistani HBV patients aligned with reference sequences using neighbour joining method.** Bootstrap statistical analysis was performed using 1000 datasets. The sequences are labelled by their lab number and country.

<https://doi.org/10.1371/journal.pone.0261721.g001>

**Table 2. Mutations in the Enh-II region (1636–1744).**

Nucleotide	Amino Acid	CH N (%)	HCC N (%)	LC N (%)	P value	95% Conf. Interval	
						Lower	Upper
CCG1640TTT	P90F	12 (9.6)	4 (25)	1 (14.28)	0.0247	2.0911	19.869
AAA1644TGG/CGT	K91W/R	7 (5.6)	3 (18.75)	2 (28.57)	0.054	-0.2736	21.914
GTG1647GCG/CCG	V92A/P	11 (8.8)	5 (31.25)	1 (14.28)	0.0399	0.8127	22.964
T1650C	L93P	10 (8)	4 (25)	1 (14.28)	0.0307	1.4327	19.327
CAT1653Deletion	H94-	7 (6.5)	8 (50)	2 (28.5)	0.0758	-2.5629	36.563
AA1657GC	K95S	6 (4.8)	3 (18.75)	1 (14.28)	0.038	0.6521	15.291
CGT1659TGC	R96C	9 (7.2)	5 (31.25)	3 (42.85)	0.0597	-0.9819	33.749
ACC1662GCG	T97A	5 (4)	2 (12.5)	1 (14.28)	0.0363	0.6105	12.316
CT1670GG	L100G	16 (12.8)	8 (50)	2 (28.52)	0.0399	1.3328	37.774
AGC1674TTT	S101F	20 (16)	10 (62.2)	4 (57.1)	0.0396	1.9843	54.449
GCG1677GTG/ATT/CGT/GGC	A102V/R/G/I	19 (15.2)	9 (56.25)	3 (42.8)	0.0357	2.3877	46.029
ATG1680/CTG/GCG	M103L/A	23 (18.4)	11(68.75)	2 (28.52)	0.0446	0.8982	49.658
AGC1683ATT/TGG	S104I/W	20 (16)	4 (25)	1 (14.28)	0.0165	3.6732	23.087
CC1689AC/TT	T106N/I	17 (13.6)	7 (43.75)	2 (28.52)	0.0305	2.6013	34.689
GAT1692CGT	D107R	15 (12)	6 (37.5)	1 (14.28)	0.0386	1.1071	27.486
TG1695CG	L108P	14 (11.2)	5 (31.25)	2 (28.52)	0.0267	2.6423	28.014
GAA1699GCG/TAT	E109A/Y	16 (12.8)	7 (43.75)	1 (14.28)	0.0469	0.3171	31.293
GCG1701AGC/AAA	A110S/K	19 (15.2)	6 (37.5)	1 (14.28)	0.03	2.2361	28.757
TAT1704GAT/GAA	Y111D/E	21 (16.8)	8 (50)	2 (28.2)	0.0291	3.1875	38.812
TTT1707CTG	F112L	20 (16)	7 (43.75)	3 (42.8)	0.0271	3.7241	40.459
A1710G	K113E	17 (13.6)	6 (37.5)	2 (28.52)	0.0249	3.2839	31.589
AT1714GC	D114G	18 (14.4)	7 (43.7)	1 (14.28)	0.0413	0.9599	31.833
TTT1722AGC	F117S	21 (16.8)	5 (31.25)	2 (28.52)	0.0162	4.8376	30.019
AAA1725GAA/ATT	K118E/I	15 (12)	6 (37)	1 (14.28)	0.0375	1.2172	27.209
GAT1728AAC	D119N	22 (17.6)	5 (31.25)	2 (28.28)	0.0154	5.0792	30.297

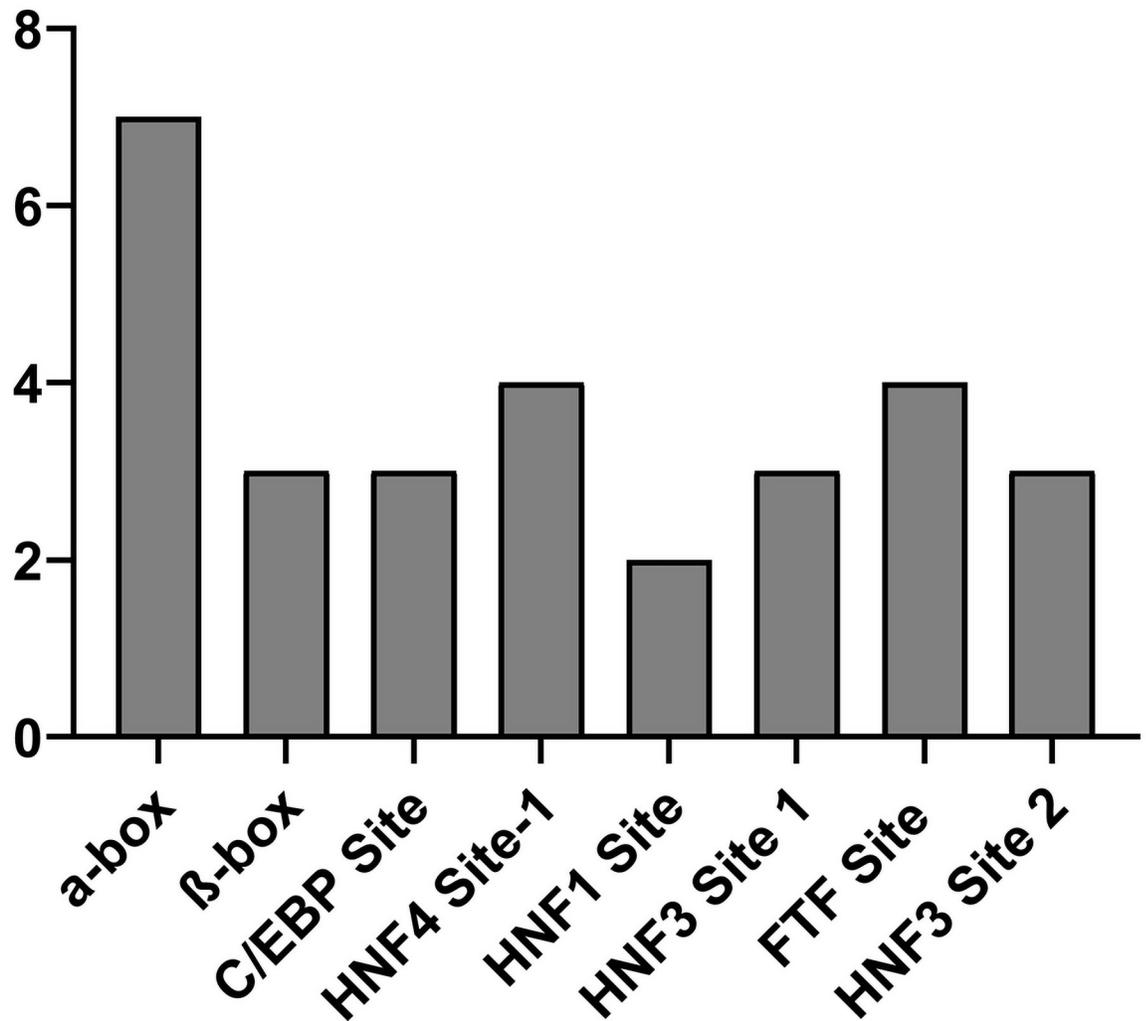
<https://doi.org/10.1371/journal.pone.0261721.t002>



**Fig 2. Enhancer II region nucleotide alignment showing comparison with the representative sequence (Normal). The dots represent similar nucleotides.**

<https://doi.org/10.1371/journal.pone.0261721.g002>





**Fig 4. Frequency of mutation in different transcriptional factors controlling enhancer II activity.**

<https://doi.org/10.1371/journal.pone.0261721.g004>

### Deletion or insertion mutations in Enh-II region

We found 17 different patterns of deletion. The deletion was observed in different nucleotides of the Enh-II region such as 1687, 1688, 1693, and 1695. Similarly, nt deletions 1718→1721 (23.64%) and 1630→1635 (18.24%) were found as the most frequent deletion mutations. A nucleotide deletion of 5 bp was detected in 16.86% (1610→1614) and 3.37% (1717→1721) cases while 6 nt deletion was found at 1630→1635 in 18.24% cases with a remarkable pattern. Two different nt insertions in the X gene were CCTCCA (6 bp) and AGGGCA (6 bp) at nt 1647 and nt.1677 were also found. We also found truncation in the Enh-II region from 1726→end in 10.81% cases and 1731→end in 14.84% cases (Table 3).

### Non-reported mutations

We also found some non-reported mutations. Non-reported mutations such as E109A/Y (43.75%), A110S/K (37.5%), Y111D/E (50%), F112L (43.75%) were found in high frequency in the HCC group as compared to other groups (Table 2). Overall frequencies of non-reported mutations are given in Table 4.



Fig 5. Mutations in different transcriptional factors affecting enhancer II activity.

<https://doi.org/10.1371/journal.pone.0261721.g005>

### Discussion

HBV infection is one of the leading causes of the advanced form of liver disease and continues to remain a medical concern in Pakistan despite prophylaxes. Most of the studies in Pakistan have been focussing on HBV’s prevalence, its epidemiology, and genotyping [22]. Studies are scarce on HBV’s mutations and their association with advanced forms of liver disease.

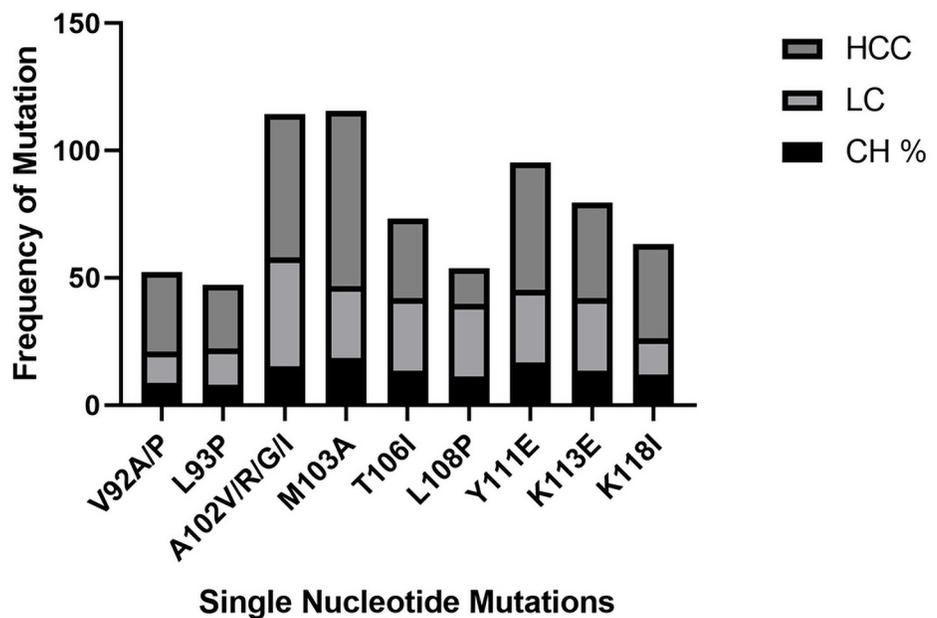


Fig 6. Frequency of single nucleotide mutations in CH, LC, and HCC groups.

<https://doi.org/10.1371/journal.pone.0261721.g006>

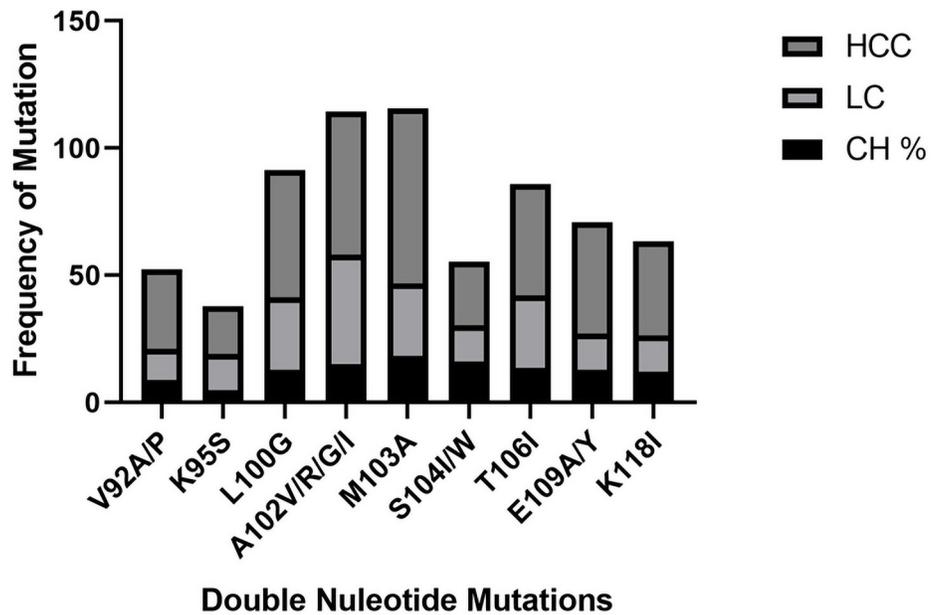


Fig 7. Frequency of double nucleotide mutations in CH, LC, and HCC groups.

<https://doi.org/10.1371/journal.pone.0261721.g007>

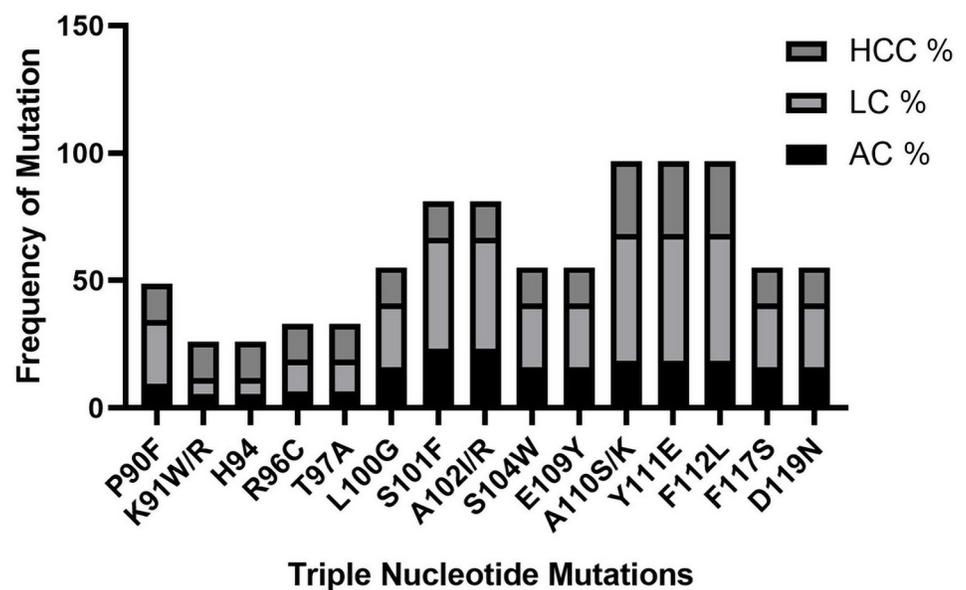


Fig 8. Frequency of triple nucleotide mutations in CH, LC, and HCC groups.

<https://doi.org/10.1371/journal.pone.0261721.g008>

Previously, the association of mutations in Enh-II region with the development of advanced liver disease were not investigated in this area. Various studies around the globe suggested that certain mutations in the HBV genome may be associated with HCC development [26, 27]. It is considered that mutations in HBV genome may responsible for escape from the host immune system and may affect the oncogenic potential in chronic HBV diseases [28]. Sequences in this study revealed nucleotide substitutions, deletions, and insertions within the Enh-II region,

Table 3. Deletion mutation in the Enh-II region of HBV.

Length of Fragment	Nucleotide Deletion	Frequency	Percentage (%)
1 bp	1687,	8	5.40
	1688,	4	2.70
	1693	11	7.43
	1695	17	11.48
	1746	4	2.70
	1757	17	11.48
3bp	1653→1655,	23	15.54
4 bp	1718→1721	35	23.64
5 bp	1610→1614,	22	16.86
	1717→1721	5	3.37
6 bp	1630→1635	27	18.24
Truncated	1731→end	16	10.81
	1726→end	22	14.84

<https://doi.org/10.1371/journal.pone.0261721.t003>

Table 4. Non-reported mutations.

Nucleotide Mutation	Amino Acid Mutation	Frequency (N)	Percentage (%)	95% Conf. Interval		
				P values	Lower	Upper
T1650C	L93P	15	10.13	0.1218	-31.84	75.4
CAT1653 Deletion	H94-	17	11.48	0.1218	-31.84	75.4
CGT1659TGC	R96C	17	11.48	0.1218	-29.386	69.596
AGC1683ATT/TGG	S104I/W	25	16.89	0.1217	-36.681	86.951
GAA1699GCG/TAT	E109A/Y	24	16.21	0.1219	-20.829	49.309
GCG1701AGC/AAA	A110S/K	26	17.56	0.1219	-18.375	43.505
TAT1704GAT/GAA	Y111D/E	31	20.94	0.1219	-20.829	49.309
TTT1707CTG	F112L	30	20.27	0.1217	-30.579	72.469
AT1714GC	D114G	26	17.56	0.1218	-37.942	89.882

<https://doi.org/10.1371/journal.pone.0261721.t004>

which are shown associated with certain biological functions of the protein that may lead to cancer.

Deletion mutation was found in this study at nt 1653 (H94-) in 6.5% of CH, 28.5% of LC, and 50% of HCC patients which showed an increasing trend from CH to HCC. These mutations might be a potential biomarker to predict the prognosis for patients with an advanced form of liver diseases such as LC and HCC. Our findings are consistent with other studies [20, 29] that reported high frequency of HCC due to substitution mutation at nt 1653 in the Enh-II region of the HBx gene. But in our study at this point, it was a deletion mutation. This region in Enh-II is located in the box  $\alpha$  which is a strong activation element of the Enh-II/core promoter that can enhance the box  $\alpha$  binding affinity and Enh-II/core promoter activity [15, 29]. Many trans-regulated nuclear factors bind HBV at the 1653 site, and such mutation can alter the binding affinity of the nuclear factors [30, 31]. Mutation at this site was reported to be a predictive factor for HCC in genotype C [20] and genotype D [32]. A meta-analysis data also reported that C1653T is associated with an increased risk of HCC [33].

Substitution mutation S101F was detected in a considerable number of samples (N = 37) and found highest in HCC patients (62%) followed by LC patients (57%). It is shown that mutation at this site is responsible for the disturbance of the cell cycle leading to apoptosis caused by the increased expression of p21 in Hep G2 and may lead to the advanced form of

liver disease [4]. The level of HBV transcription varies with the extent of transcription factor binding to HBV DNA, which can be altered by mutations in the HBV genome. When a new mutation allows a transcription factor to bind a new region in HBV DNA, HBV replication is stimulated resulting in the progression of the disease [34]. The substitution mutation T106N/I associated with HCC [35] was found in 43.75% of HCC and 28.52% of LC patients of this study. The Enh-II region from nt 1640 to 1663 containing box- $\alpha$  and C/EBP site is an independent sequence and sufficient for the enhancer activity. In an experimental study, mutation in this sequence results in a great decrease in luciferase gene expression under the control of the minimal promoter [10]. In the present study eight different (P90F, K91W/R, V92A/P, L93P, H94-, K95S, R96C, T97A) substitutions and deletion mutations were found in this region that may affect the viral replication and cell cycle and might play a role in the development of the advanced form of liver disease.

Hepatocyte enriched transcription factors have various binding motif and regulate the activities of Enh-II region [10–17]. Of these transcriptions factors, HNFs regulate cccDNA transcription by directly interacting with HBx, which enhances the DNA-binding activity of HNFs. HNF1 up-regulates Enh-II activity by interacting with either the hB1F or B element in Enh-II. In the absence of HNF1 $\alpha$ , the concentration of HBV pgRNA is decreased, resulting in decreased genome replication [31]. In this study, two mutations were found in the HNF1 Site (K118E/I and D119N) which were more prevalent in HCC (37%) patients as compared to LC (14.2%) and CH (12%). It is reported that due to mutations new binding sequences of HNF1 may emerge that activate increased viral replication. Moreover, double combinational mutation (K118E/I, D119N) within regulatory elements of Enh-II has been shown to affect the replication of HBcAg in situ and can manifest liver disease [36, 37]. Additionally, this double mutation was studied in T cellular cell epitope (amino acid 115–123) and T helper cell epitopes (amino acid 111–135) which plays a role in the evasion of the HBV from host immune clearance [38]. In the current study, 4 bp deletion was found in the SP1 binding site of the C-terminal region (nt 1717–1721). Deletion in this region causes carcinogenesis which might be due to the deletion of the Sp1 binding site that controls the expression of the X gene negatively. The deletion of the Sp1 binding site is increased with the expression of the X gene [39].

Likewise, three mutations T97A, L100G, and S101F were reported in the HNF4 site which were more prevalent in HCC groups (12.5%, 50%, and 62% respectively) as compared to other two groups. Such deletion may increase HNF4 expression and HNF3 down-regulation in CH patients [35].

Overall, twelve non-reported mutations were also identified (P90F, V92W/R, R96C, T97A, S104I/W, E109A/Y, D107R, A110S/K, Y111E, F112L, and D114G). Mutations in the Enh-II region of the amino-terminal domain of core protein may be responsible for the nucleocapsid assembly and stability [40]. The incidence and impact of these mutations are unclear; however, the presence of these mutations in high frequency in HCC and LC patients is alarming as these mutations might be responsible for the advanced form of liver disease.

## Conclusion

In conclusion, in this age, sex, and HBeAg status matched study, most of the mutations were associated with HCC development in infected patients of KP with HBV/D which are also confirmed by other studies. The occurrence of multiple mutations represents a strategy by which HBV can escape immune surveillance and thus contribute to hepatocarcinogenesis. These mutations cause structural changes and conformation alteration of X protein and hence alter the regulatory and transactivation functions. Some unreported mutations were also detected in our study which may provide new insights into the mechanism by which multi-site

mutations in the Enh-II region may contribute to the development of advanced liver disease. However, further experiments need to be conducted to confirm the inference. In short, our study confirms certain mutations and their association with an advanced form of liver disease in Pakistani patients.

## Author Contributions

**Conceptualization:** Majid Khan, Sanaullah Khan, Muhammad Nasir Khabir.

**Data curation:** Majid Khan, Safia Bibi, Abdul Majid, Muhammad Nasir Khabir, Aisha Gul.

**Formal analysis:** Safia Bibi, Bakht Tarin Khan, Abdul Majid, Ayesha Khattak, Aisha Gul.

**Methodology:** Bakht Tarin Khan, Ayesha Khattak, Muhammad Nasir Khabir, Muhammad Anwar, Maryam Naseem.

**Resources:** Bakht Tarin Khan, Maryam Naseem.

**Supervision:** Sanaullah Khan.

**Writing – original draft:** Majid Khan, Safia Bibi, Abdul Majid.

**Writing – review & editing:** Mehrunnisa Fatima Gondal, Abdul Majid, Muhammad Anwar, Maryam Naseem, Sobia Attaullah.

## References

1. Wang J, Shen T, Huang X, Kumar GR, Chen X, Zeng Z, et al. Serum hepatitis B virus RNA is encapsidated pregenome RNA that may be associated with persistence of viral infection and rebound. *Journal of Hepatology*. 2016; 65(4):700–10. <https://doi.org/10.1016/j.jhep.2016.05.029> PMID: 27245431
2. Zhao J, Wu J, Cai H, Wang D, Yu L, Zhang W-H. E3 Ubiquitin ligase Siah-1 is Down-regulated and fails to target natural HBx truncates for degradation in hepatocellular carcinoma. *Journal of Cancer*. 2016; 7(4):418. <https://doi.org/10.7150/jca.13019> PMID: 26918055
3. World Health Organization. Hepatitis B: Fact sheets. Accessed on 1 January 2019; [www.who.int/news-room/fact-sheets/detail/hepatitis-b](http://www.who.int/news-room/fact-sheets/detail/hepatitis-b).
4. Gededzha MP, Sondlane TH, Malinga LA, Burnett RJ, Lebelo RL, Blackard JT, et al. Molecular characterization of hepatitis B virus X gene in HIV-positive South Africans. *Virus genes*. 2018; 54(2):190–198. <https://doi.org/10.1007/s11262-018-1536-5> PMID: 29411271
5. Tong S, Revill P. Overview of hepatitis B viral replication and genetic variability. *Journal of hepatology*. 2016; 64(1):S4–S16. <https://doi.org/10.1016/j.jhep.2016.01.027> PMID: 27084035
6. De Mitri M, Bernardi M. Predicting the development of hepatocellular cancer in hepatitis B carriers. *Gut*. 2008; 57(1):12–5. <https://doi.org/10.1136/gut.2007.128710> PMID: 18094200
7. Qu LS, Zhu J, Liu TT, Shen XZ, Chen TY, Ni ZP, et al. Effect of combined mutations in the enhancer II and basal core promoter of hepatitis B virus on development of hepatocellular carcinoma in Qidong, China. *Hepatology Research*. 2014; 44(12):1186–95. <https://doi.org/10.1111/hepr.12291> PMID: 24341484
8. Yuh C, Ting L. The genome of hepatitis B virus contains a second enhancer: cooperation of two elements within this enhancer is required for its function. *Journal of virology*. 1990; 64(9):4281–7. <https://doi.org/10.1128/JVI.64.9.4281-4287.1990> PMID: 2166817
9. Ishida H, Ueda K, Ohkawa K, Kanazawa Y, Hosui A, Nakanishi F, et al. Identification of multiple transcription factors, HLF, FTF, and E4BP4, controlling hepatitis B virus enhancer II. *Journal of virology*. 2000; 74(3):1241–51. <https://doi.org/10.1128/jvi.74.3.1241-1251.2000> PMID: 10627534
10. Wang W, Li M, Wu X, Wang Y, Li Z. HNF1 is critical for the liver-specific function of HBV enhancer II. *Research in virology*. 1998; 149(2):99–108. [https://doi.org/10.1016/s0923-2516\(98\)80085-x](https://doi.org/10.1016/s0923-2516(98)80085-x) PMID: 9602504
11. Johnson JL, Raney AK, McLachlan A. Characterization of a functional hepatocyte nuclear factor 3 binding site in the hepatitis B virus nucleocapsid promoter. *Virology*. 1995; 208:147–58. <https://doi.org/10.1006/viro.1995.1138> PMID: 11831695
12. Li M, Xie Y, Wu X, Kong Y, Wang Y. HNF3 binds and activates the second enhancer, ENII, of hepatitis B virus. *Virology*. 1995; 214(2):371–8. <https://doi.org/10.1006/viro.1995.0046> PMID: 8553537

13. Guo W, Chen M, Yen T, Ou J-H. Hepatocyte-specific expression of the hepatitis B virus core promoter depends on both positive and negative regulation. *Molecular and Cellular Biology*. 1993; 13(1):443–8. <https://doi.org/10.1128/mcb.13.1.443-448.1993> PMID: 8417343
14. Raney AK, Johnson JL, Palmer C, McLachlan A. Members of the nuclear receptor superfamily regulate transcription from the hepatitis B virus nucleocapsid promoter. *Journal of virology*. 1997; 71(2):1058–71. <https://doi.org/10.1128/JVI.71.2.1058-1071.1997> PMID: 8995626
15. López-Cabrera M, Letovsky J, Hu K-Q, Siddiqui A. Multiple liver-specific factors bind to the hepatitis B virus core/pregenomic promoter: trans-activation and repression by CCAAT/enhancer binding protein. *Proceedings of the National Academy of Sciences*. 1990; 87(13):5069–73. PMID: 2367525
16. López-Cabrera M, Letovsky J, Hu K-Q, Siddiqui A. Transcriptional factor C/EBP binds to and transactivates the enhancer element II of the hepatitis B virus. *Virology*. 1991; 183(2):825–9. [https://doi.org/10.1016/0042-6822\(91\)91019-d](https://doi.org/10.1016/0042-6822(91)91019-d) PMID: 1853580
17. Yuh C, Ting L. C/EBP-like proteins binding to the functional box-alpha and box-beta of the second enhancer of hepatitis B virus. *Molecular and Cellular Biology*. 1991; 11(10):5044–52. <https://doi.org/10.1128/mcb.11.10.5044-5052.1991> PMID: 1922032
18. Li M, Xie Y-H, Kong Y-Y, Wu X, Zhu L, Wang Y. Cloning and characterization of a novel human hepatocyte transcription factor, hB1F, which binds and activates enhancer II of hepatitis B virus. *Journal of Biological Chemistry*. 1998; 273(44):29022–31. <https://doi.org/10.1074/jbc.273.44.29022> PMID: 9786908
19. Zhang P, Raney AK, McLachlan A. Characterization of functional Sp1 transcription factor binding sites in the hepatitis B virus nucleocapsid promoter. *Journal of Virology*. 1993; 67(3):1472–81. <https://doi.org/10.1128/JVI.67.3.1472-1481.1993> PMID: 8437225
20. Kim H, Lee S-A, Kim B-J. X region mutations of hepatitis B virus related to clinical severity. *World Journal of Gastroenterology*. 2016; 22(24):5467–78. <https://doi.org/10.3748/wjg.v22.i24.5467> PMID: 27350725
21. Noorali S, Hakim ST, McLean D, Kazmi SU, Bagasra O. Prevalence of Hepatitis B virus genotype D in females in Karachi, Pakistan. *The Journal of Infection in Developing Countries*. 2008; 2(05):373–8. <https://doi.org/10.3855/jidc.200> PMID: 19745506
22. Hakim S, Kazmi S, Bagasra O. Seroprevalence of hepatitis B and C genotypes among young apparently healthy females of Karachi-Pakistan. *Libyan Journal of Medicine*. 2008; 3(2):66–70.
23. Kyaw Y.Y., Lwin A.A., Aye K.S. et al. Distribution of hepatitis B virus genotypes in the general population of Myanmar via nationwide study. *BMC Infectious Diseases* 20, 552 (2020). <https://doi.org/10.1186/s12879-020-05269-z> PMID: 32727389
24. Naito H, Hayashi S, Abe K. Rapid and specific genotyping system for hepatitis B virus corresponding to six major genotypes by PCR using type-specific primers. *J Clin Microbiol*. 2001 Jan; 39(1):362–4. <https://doi.org/10.1128/JCM.39.1.362-364.2001> PMID: 11136801
25. Al-Qahtani A. A., Al-Anazi M. R., Nazir N., Ghai R., Abdo A. A., Sanai F. M., et al. (2017). Hepatitis B virus (HBV) X gene mutations and their association with liver disease progression in HBV-infected patients. *Oncotarget*, 8(62), 105115–105125. <https://doi.org/10.18632/oncotarget.22428> PMID: 29285238
26. Shi H-P, Zhang J, Shang X-C, Xie X-Y. Hepatitis B virus gene C1653T polymorphism mutation and hepatocellular carcinoma risk: an updated meta-analysis. *Asian Pacific Journal of Cancer Prevention*. 2013; 14(2):1043–7. <https://doi.org/10.7314/apjcp.2013.14.2.1043> PMID: 23621183
27. Liao Y, Hu X, Chen J, Cai B, Tang J, Ying B, et al. Precore mutation of hepatitis B virus may contribute to hepatocellular carcinoma risk: evidence from an updated meta-analysis. *PLoS One*. 2012; 7(6) 1–9. <https://doi.org/10.1371/journal.pone.0038394> PMID: 22675557
28. Lee J-H, Han K-H, Lee JM, Park JH, Kim H-S. Impact of hepatitis B virus (HBV) x gene mutations on hepatocellular carcinoma development in chronic HBV infection. *Clin Vaccine Immunol*. 2011; 18(6):914–21. Epub 04/13. <https://doi.org/10.1128/CVI.00474-10> PMID: 21490166
29. Yuh C, Chang Y, Ting L. Transcriptional regulation of precore and pregenomic RNAs of hepatitis B virus. *Journal of virology*. 1992; 66(7):4073–84. <https://doi.org/10.1128/JVI.66.7.4073-4084.1992> PMID: 1602534
30. Shinkai N, Tanaka Y, Ito K, Mukaide M, Hasegawa I, Asahina Y, et al. Influence of hepatitis B virus X and core promoter mutations on hepatocellular carcinoma among patients infected with subgenotype C2. *Journal of clinical microbiology*. 2007; 45(10):3191–7. <https://doi.org/10.1128/JCM.00411-07> PMID: 17652471
31. Tanaka Y, Mukaide M, Orito E, Yuen M-F, Ito K, Kurbanov F, et al. Specific mutations in enhancer II/core promoter of hepatitis B virus subgenotypes C1/C2 increase the risk of hepatocellular carcinoma. *Journal of hepatology*. 2006; 45(5):646–53. <https://doi.org/10.1016/j.jhep.2006.06.018> PMID: 16935384

32. Hamid A, Muzamil & Salama Abosufyan. X gene variability and genotyping of hepatitis b virus (hbv) in sudanese patients with liver diseases, khartoum, sudan. *World Journal of Pharmaceutical Research*. 2018; 7.
33. Liu S, Zhang H, Gu C, Yin J, He Y, Xie J, et al. Associations between hepatitis B virus mutations and the risk of hepatocellular carcinoma: a meta-analysis. *Journal of the National Cancer Institute*. 2009; 101(15):1066–82. <https://doi.org/10.1093/jnci/djp180> PMID: 19574418
34. Pult I, Chouard T, Wieland S, Klemenz R, Yaniv M, Blum HE. A hepatitis B virus mutant with a new hepatocyte nuclear factor 1 binding site emerging in transplant-transmitted fulminant hepatitis B. *Hepatology*. 1997; 25(6):1507–15. <https://doi.org/10.1002/hep.510250633> PMID: 9185776
35. Garmiri P, Loua A, Haba N, Candotti D, Allain JP. Deletions and recombinations in the core region of hepatitis B virus genotype E strains from asymptomatic blood donors in Guinea, West Africa. *Journal of General Virology* 2009; 90: 2442–2451
36. Long Y, Chen E, Liu C, Huang F, Zhou T, He F, et al. The correlation of hepatocyte nuclear factor 4 alpha and 3 beta with hepatitis B virus replication in the liver of chronic hepatitis B patients. *J Viral Hepat*. 2009; 16(8):537–46. <https://doi.org/10.1111/j.1365-2893.2009.01089.x> PMID: 19302415
37. Tuteja A, Siddiqui AB, Madan K, Goyal R, Shalimar, Sreenivas V, et al. Mutation profiling of the hepatitis B virus strains circulating in North Indian population. *PloS one*. 2014; 9(3):e91150–e. <https://doi.org/10.1371/journal.pone.0091150> PMID: 24637457
38. Zhu R, Zhang HP, Vu H, Li H, Ling YQ, Hu XQ, et al. Hepatitis B virus mutations associated with in situ expression of hepatitis B core antigen, viral load and prognosis in chronic hepatitis B patients. *Pathology-Research and Practice* 2008; 204: 731–742. <https://doi.org/10.1016/j.prp.2008.05.001> PMID: 18597951
39. Utama A, Purwantoma S, Siburian MD, Dhenni R, Gani RA, Hasan I, et al. Hepatitis B virus subgenotypes and basal core promoter mutations in Indonesia. *World Journal of Gastroenterology* 2009; 15: 4028–4236 PMID: 19705499
40. Salarnia F., Besharat S., Zhand S., Javid N., Khodabakhshi B., Moradi A. Mutations in Hepatitis-B X-Gene Region: Chronic Hepatitis-B versus Cirrhosis. *Journal of clinical and diagnostic research: JCDR*, 2017. 11(3), 31–34. <https://doi.org/10.7860/JCDR/2017/22570.9498> PMID: 28511432