



## Short Communication

# Quasispecies Analysis Through NGS and Identification of Escape Mutation in a Family Infected with Hepatitis B Virus: A Study from Lahore, Pakistan

Iqra Almas, Samia Afzal\*, Muhammad Idrees, Muhammad Shahid, Iram Amin and Kausar Malik

Centre of Excellence in Molecular Biology, University of the Punjab, Lahore, Pakistan.

## ABSTRACT

Next generation sequencing (NGS) provides an accurate analysis of hepatitis B virus (HBV) as compared to Sanger sequencing. In 2019, samples from HBV infected family of mother and her two sons samples were collected from Lahore, Pakistan. Sequencing of S gene, Geno2pheno (HBV) 2.0 tool and GHOST platform were used to analyze the genotype, mutations (natural and escape) and quasispecies analysis of the HBV. All samples in family were infected with genotype HBV/D1 and had high viral load ( $10^8$  IU). An vaccine escape mutation "C137W" was identified in mother and a son. Mother sample were shown more number of haplotypes and total diversity as compared to sons. The quasispecies analysis showed the mother and sons were infected with HBV having similar genetic makeup, indicating the vertical transfer of infection. This case study provides knowledge about genetic variations that contributes to the improved diagnosis of the infection, treatment therapy, follow-up, and future treatment strategies.

### Article Information

Received 24 August 2022  
Revised 18 October 2022  
Accepted 11 November 2022  
Available online 08 June 2023  
(early access)

### Authors' Contribution

SA and MI planned the study. IA collected samples, planned methodology and wrote the manuscript. IA and MS did the analysis and editing. KM edited the manuscript and gave the approval of study.

### Key words

HBV, Escape Mutation, Quasispecies, S gene

HBV is one of the major liver-related health concerns in Pakistan. According to an estimation by World Health Organization (WHO), HBV is responsible for the cause of deaths of 600,000 population by the acute or chronic infection that leads towards liver cirrhosis or hepatocellular carcinoma (Forbi *et al.*, 2017). Pakistan also has a 2%-4% prevalence rate of HBV due to unhygienic conditions and lack of awareness regarding HBV infection and transmission (Ali *et al.*, 2009).

HBV is one of the smallest DNA infectious viruses with approximately 3.2 kb genome size. HBV is classified into 10 further genotypes based on the genome sequence, (from A to J) with the nucleotide variation exceeding 8%, and sub-genotypes with nucleotide variation ranging from 4%-8% (Kyaw *et al.*, 2020). It is noteworthy that HBV viral genotypes are geographically distributed along a specific

pattern. The HBV genome sequencing may offer significantly valuable information which may be used for viral genotype characterization, resistance to the drug and vaccine, potential transmission networks, and aspects of infection dynamics (McNaughton *et al.*, 2019a).

The "a determinant" region is located at major hydrophilic region (MHR) from amino acid position 120 to 160 in S gene. MHR has a double loop structure that have two pair of cysteine to form disulfide bridge at each loop. Disulfide bridging occur between C124 and C137 and between C139 and C147 to form double loop structure (Purdy, 2007). Mutation in the "a determinant" region can leads towards immune escape mutations. Different vaccine escape mutations has been reported till now. G145R was the first reported vaccine escape mutation along with T116N, P120S/E, I/T126A/N/I/S, Q129H/R, M133L, C137, K141E, P142S, D144A/E, and G145R/A (Dong *et al.*, 2009; Lazarevic *et al.*, 2019).

This case study presents the data of a family of a mother and her two sons infected with HBV. High viral load was quantified with Real-time PCR. Conventional Sanger methods work for giving the consensus sequences (typically for subgenomic fragments) or for dealing with the low clonal coverage only. On the contrary, NGS can investigate the genome for diversity within sample (McNaughton *et al.*, 2019a). It can efficiently deal with

\* Corresponding author: samiaraza@live.com  
0030-9923/2023/0001-0001 \$ 9.00/0



Copyright 2023 by the authors. Licensee Zoological Society of Pakistan.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

the accurate characterization of the viral quasispecies or the study of the virus transmission and its evolutionary dynamics concerning immunological or drug pressure (McNaughton *et al.*, 2019a, b).

Global Hepatitis Outbreak and Surveillance Technology (GHOST) is a cloud-based system that comprises various bioinformatics tools (Longmire *et al.*, 2017). GHOST provides data regarding the genotype, number of haplotypes, total number of reads, total diversity, and maximum frequency of haplotypes of each sample.

The aim of this study is to investigate the mutational analysis and quasispecies identification of HBV of a family (mother and her sons) that are responsible for immune escape and vertical transmission.

#### Materials and methods

In the present study, blood samples of a family of 28 years old mother and her 2 sons (6 and 7 years old) infected with HBV were collected at Genome Centre and Medical Diagnostics (GCMD), Lahore for the detection, quantification and genotyping of hepatitis B (HBV). Well-informed written consents were taken from all the participants at the time of sampling. All the important information such as demographics and clinical data from all the participants was recorded on a data sheet. For the detection of HBeAg in the serum of HBV infected patients, an ETI-EBK PLUS kit was used. MagNA Pure LC Total Nucleic Acid Isolation Kit (Roche, Catalogue Number, 05323738001) was used for the Nucleic Acid Extraction from the serum samples. Primers used for Sanger sequencing 1<sup>st</sup> round PCR were: EXS164F ACATCACATCAGGACTCCTAGGA and EXS656R GAGGCCCACTCCCATAGGTAT whereas primers used for nested PCR second round were INS183F AGGACCCCTTCTCGTGTTACA and INS624R CCAAGATGATGGGATGGGAAT. The S gene was amplified by using Lightcycler 480 instrument (Roche Diagnostics Corporation, Indianapolis, IN) with thermal Profile consists of denaturation at 95°C for 10 min, PCR amplification: 30 cycles at 95°C for 30 sec, 50°C for 30 sec and 72°C for 30 sec and melting curve at 76°C for 10 min. For Sanger sequencing, the S gene PCR product was sequenced in both directions by using forward and reverse primers. The sequencing was performed by using Big Dye-

Terminator version 3.1 cycling methodology (Applied Biosystems, CA). Sequencing was done by an automated sequencer (ABI 3130xl Genetic Analyzer, Applied Biosystems). The obtained sequences were cleaned and aligned by SeqMan (version 10.1.2, DNASTAR, Madison, WI). Further analysis was done by using Geno2pheno (HBV) 2.0 (<https://hbv.geno2pheno.org>). For NGS, 1<sup>st</sup> round was the same as used before in sanger sequencing. 32 specific barcode PCR primers were the nested primers with specific 32 unique 10 bps tags (barcode). 8 bps unique tags were added to the index primer sequences to make 32 unique index primers. The PCR products were purified by using Ampure XP (Agilent). Each product was quantified by using Tape station instrument (Agilent, CA) according to protocol described. The quantified products were normalized and diluted to make an NGS library pool of 10pmol. Sequencing was done by using MiSeq Reagent Kit v3 (600-cycle) in Illumina MiSeq instrument. NGS pooled Library was automatically de-multiplexed by the Miseq Instrument. Further de-multiplexing at paired reads was processed by GHOST system (Longmire *et al.*, 2017). The way to visualize the genetic relatedness among all haplotypes present in each HBV infected specimen is through k-step networks.

#### Results

High viral loads of HBV and HBeAg positive were observed in all three cases of family members. Sanger sequences of all the samples were of genotype D1/HBV, identified by using Geno2pheno (HBV) and further confirmed by GHOST software. Mutational analysis indicated the presence of vaccine Escape mutation C137W in Mother and one of her sons. Whereas another mutation T127L were also identified in family members as shown in supplementary section. NGS data indicated that mother (P42) were shown more number of haplotypes and total diversity as compared to sons (P41 and 43) shown in Table I. K steps represented the Intra-Host variations and Inter Host variations. The circular pattern indicated the acute infection with fewer variations in HBV genetics shown in Figure 1. K step diagram of the 3 samples forming a cluster indicating the similar genetic makeup of HBV species in family indicating the vertical transmission, as shown in Figure 2.

**Table I. Characteristics and GHOST analysis data of HBV infected family.**

ID/ Name	Genotype	nhap	nreads	Total div	Max. freq.	Hbe Ag	Viral load
P41 (Son)	d	11456	19168	0.015004	3540	Positive	> 10 <sup>8</sup>
P42 (Mother)	d	11805	18965	0.015511	3133	Positive	> 10 <sup>8</sup>
P43 (Son)	d	10851	19192	0.014425	4031	Positive	> 10 <sup>8</sup>

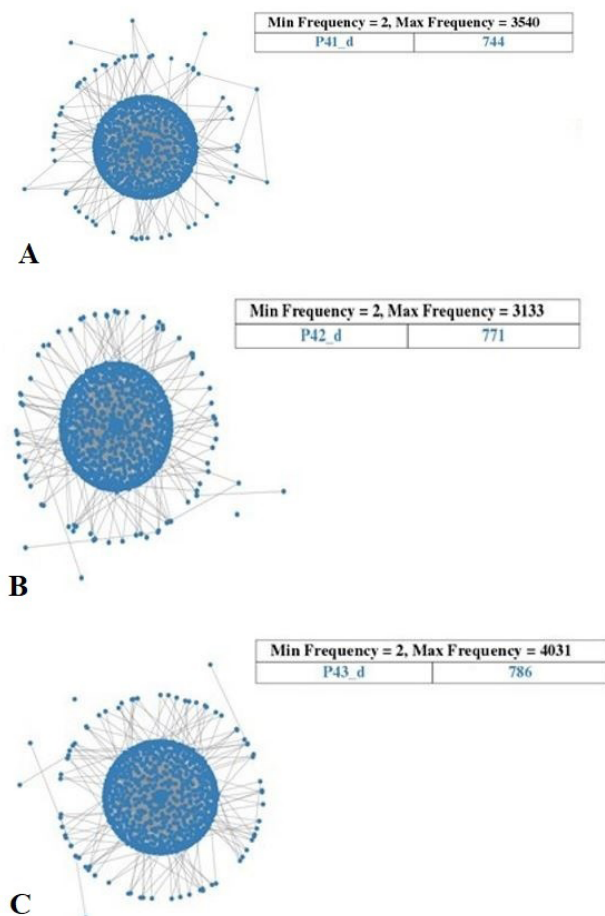


Fig. 1. K-steps diagrams indicating the intra-host variations showing HBV quasispecies of HBV mother (B) and her sons (A and C).

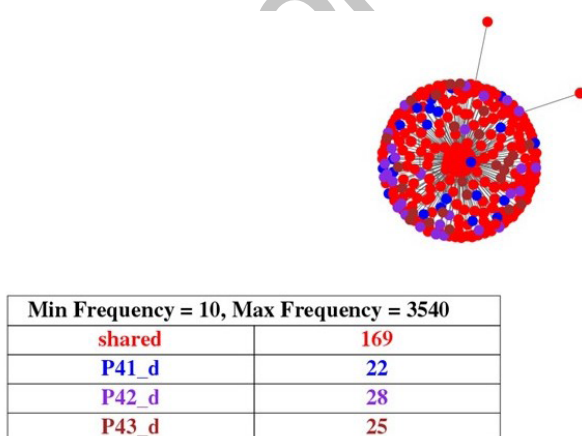


Fig. 2. K step diagram showing the cluster formation of HBV of family, sharing the same genetic makeup. The red colour represents the same genetic sequences present in all samples.

### Discussion

Pakistan has a high prevalence rate of HBV that leads towards a high modality due to liver cirrhosis and hepatocellular carcinoma. Replication of HBV takes place by reverse transcription of an RNA intermediate and the errors in HBV DNA replication occur at a much higher rate than for other DNA viruses due to a lack of RT proofreading function. The nucleotide

substitution estimated rate per year is approximately  $1.4-3.2 \times 10^{-5}$  per site (Okamoto *et al.*, 1987). The evolution of these naturally occurring mutants occurs during the infection period under antiviral pressure of the host immune system or by other factors like specific therapy or immunization (Günther *et al.*, 1999). In this case study, the mother and her sons were infected with HBV/D genotype, the most prevalent genotype in Pakistan. Sequencing identified a mother and one of her sons who share a mutation T127L whereas an escape mutation C137W also be identified in a mother and one of her sons. This vaccine mutation may be responsible for the transmission of HBV from mother to her sons. Escape mutations, C137W along with G145R, was identified by showing false negative result for HbsAg (Gerlich, 2006).

A complex mixture of HBV variants with the related genomic sequence that infects a certain host is known as the viral quasispecies (Domingo *et al.*, 2012). The sequencing technologies have been extensively employed to uncover novel endeavors of HBV and the infection that includes valuable data concerning the identification of the targets for diagnosis and the factors that play their role in the disease progression. For example, a spontaneous mutation in preS/S, pre-C/C and X was noted via Sanger sequencing and a further link was established with the progression of liver disease (Tatsukawa *et al.*, 2011; Yang *et al.*, 2013; Shen and Yan, 2014). However, these research studies pose a limitation that links with the sensitivity of Sanger and endorse the benefit NGS approaches may have contributed, especially with the depth and high quality ensured by the NGS technologies.

In this case study, Sanger sequences and major sequences by NGS showed the same results of all the samples but through NGS, all the minor populations were also obtained. Total diversity defines as the Measurement of the amount of heterogeneity of a sample from a range of 0 to 1, where 0 is more homogenous and 1 is very heterogenous. Mother (P42) were shown more number of haplotypes and total diversity as compared to sons (P41 and 43) indicating more evolution of HBV quasispecies in mother. Intra-host and inter-host variation were observed in the form of K steps by GHOST. All samples indicated acute infections with a similar genetic makeup showing the vertical transfer of HBV infection. In 2002, Pakistan

initiated the HBV Childhood vaccination Program (Jafri *et al.*, 2006). Almost 72% to 86% of children are HBV vaccinated since the launch of the HBV childhood Program (Harris *et al.*, 2017). As the children were supposed to be vaccinated but still, they are infected with HBV also confirms the vertical transmission.

This study contributes to genetic diversity of HBV that will help to improve the diagnosis of the infection, its treatment therapy, follow-up, and future treatment strategies. There is a need to analyse the frequency of specific mutations in HBV genome rather than specific to S gene that are responsible for the vaccination failure and vertical transmission.

#### Acknowledgement

Authors acknowledge Genome Centre for Molecular Diagnostic and Research Lab for providing facility for patients sample collection, and Centre for Disease control and Prevention (CDC), USA for providing facility for analysis of HBV samples.

#### Funding

The funding for this study was provided by Higher Education Commission Pakistan.

#### IRB approval

Institutional Review Board committee, CEMB, Lahore approved the study (CEMB Ref # 03/22, dated 06/06/2022).

#### Ethical statement

A written consent was obtained from all patients before sample collection for this study.

#### Statement of conflict of interest

The authors have declared no conflict of interest.

#### References

- Ali, S.A., Donahue, R.M., Qureshi, H. and Vermund, S.H., 2009. *Int. J. Infect. Dis.*, **13**: 9-19. <https://doi.org/10.1016/j.ijid.2008.06.019>
- Domingo, E., Sheldon, J. and Perales, C., 2012. *Microbiol. Mol. Biol. Rev.*, **76**: 159-216. <https://doi.org/10.1128/MMBR.05023-11>
- Dong, Y., Liu, S.L., Zhai, X.J., Zhu, F.C., Pan, H., Yu, J.X., Chen, Y.Z., Xie, Y.R., Zhang, X.Y. and Zhang, H.M., 2009. *J. med. Virol.*, **81**: 1517-1524. <https://doi.org/10.1002/jmv.21522>
- Forbi, J.C., Dillon, M., Purdy, M.A., Drammeh, B.S., Tejada-Strop, A., McGovern, D., Xia, G.L., Lin, Y., Ganova-Raeva, L.M. and Campo, D.S., 2017. *J. Gen. Virol.*, **98**: 1048-1057. <https://doi.org/10.1099/jgv.0.000776>
- Gerlich, W.H., 2006. *J. clin. Virol.*, **36**: S18-S22. [https://doi.org/10.1016/S1386-6532\(06\)80004-1](https://doi.org/10.1016/S1386-6532(06)80004-1)
- Günther, S., Fischer, L., Pult, I., Sterneck, M. and Will, H., 1999. *Adv. Virus Res.*, **52**: 25-137.
- Harris, B.J., Holzmayer, V., Qureshi, H., Khan, M.A., Khan, S.A., Salamat, N., Jafri, S. and Brennan, C.A., 2017. *PLoS One*, **12**: e0178988. <https://doi.org/10.1371/journal.pone.0178988>
- Jafri, W., Jafri, N., Yakoob, J., Islam, M., Tirmizi, S.F.A., Jafar, T., Akhtar, S., Hamid, S., Shah, H.A. and Nizami, S.Q., 2006. *BMC Infect. Dis.*, **6**: 1-10. <https://doi.org/10.1186/1471-2334-6-101>
- Kyaw, Y.Y., Lwin, A.A., Aye, K.S., Thu, H.M., Htun, M.M., Soe, H.O., Aye, K.T., Thant, K.Z., Hwang, H.J., and Cheong, J., 2020. *BMC Infect. Dis.*, **20**: 1-11. <https://doi.org/10.1186/s12879-020-05269-z>
- Lazarevic, I., Banko, A., Miljanovic, D. and Cupic, M., 2019. *Viruses*, **11**: 778. <https://doi.org/10.3390/v11090778>
- Longmire, A.G., Sims, S., Rytsareva, I., Campo, D.S., Skums, P., Dimitrova, Z., Ramachandran, S., Medrzycki, M., Thai, H. and Ganova-Raeva, L., 2017. *BMC Genomics*, **18**: 21-32. <https://doi.org/10.1186/s12864-017-4268-3>
- McNaughton, A.L., Roberts, H.E., Bonsall, D., de Cesare, M., Mokaya, J., Lumley, S.F., Golubchik, T., Piazza, P., Martin, J.B. and de Lara, C., 2019a. *Sci. Rep.*, **9**: 1-14. <https://doi.org/10.1038/s41598-019-43524-9>
- McNaughton, A.L., D'Arienzo, V., Ansari, M.A., Lumley, S.F., Littlejohn, M., Revill, P., McKeating, J.A. and Matthews, P.C., 2019b. *Gastroenterology*, **156**: 384-399. <https://doi.org/10.1053/j.gastro.2018.07.058>
- Okamoto, H., Imai, M., Tsuda, F., Tanaka, T., Miyakawa, Y. and Mayumi, M. 1987. *J. Virol.*, **61**: 3030-3034.
- Purdy, M.A., 2007. *Asian J. Transfus. Sci.*, **1**: 62. <https://doi.org/10.4103/0973-6247.33445>
- Shen, T. and Yan, X.M., 2014. *World J. Gastroenterol.*, **20**: 5435. <https://doi.org/10.3748/wjg.v20.i18.5435>
- Tatsukawa, M., Takaki, A., Shiraha, H., Koike, K., Iwasaki, Y., Kobashi, H., Fujioka, S.I., Sakaguchi, K. and Yamamoto, K., 2011. *BMC Cancer*, **11**: 1-9. <https://doi.org/10.1186/1471-2407-11-458>
- Yang, J.H., Zhang, H., Chen, X.B., Chen, G. and Wang, X., 2013. *World J. Gastroenterol.*, **19**: 3861.