VirusDis.
https://doi.org/10.1007/s13337-018-0497-6

ORIGINAL ARTICLE

Molecular evolution and genomics of hepatitis B virus subgenotype C2 strain predominant in the chronic patients in Bangladesh

Modhusudon Shaha[1] · Sourav Chakraborty[1] · Md. Saddam Hossain[1] · Abu Hashem[1] · Md. Salimullah[2]

Received: 22 July 2018 / Accepted: 1 October 2018

Abstract Evolution of hepatitis B virus (HBV) is a mystery and caused mainly by genomic mutations as well as recombination. Viral evolution may be responsible for increasing disease severity and render resistance to the existing treatment processes. HBV/C2 strain is associated with chronicity, which may progress to the liver cirrhosis and hepatocellular carcinoma. Furthermore, HBV/C2 strain is highly prevalent in the chronic hepatitis B patients in Bangladesh. Hence, the molecular evolution of that strain and its disease pattern need to be uncovered. Herein, the purpose of this study is to determine the potential mutations of HBV complete genome sequences isolated in Bangladesh and the molecular evolution of HBV/C2 strain. Mutation analysis of the total 57 complete genome sequences of HBV in Bangladesh revealed that 42.11%, 12.28%, 7.02% and 3.51% of the strains were vaccine resistant, HBsAg detection escape, HBV immunoglobulin escape, multi-drug resistant respectively. Furthermore, of the vaccine resistant strains, 16.67% were observed resistant to both vaccine, HBsAg detection and immunoglobulin escape. Bayesian skyline analysis with 462 HBV/C2 strains from 2000 to 2017 revealed the evolution of the strain was in nineteenth century with two rapid sharp increases in the genetic diversity at the end of the twentieth century and then a sudden decrease in the early twenty-first century as observed in C and X gene analysis. This study may help researchers and clinicians to get a better knowledge about the emergence and evolution of HBV/C2 strain that may help to find a proper treatment strategy against hepatitis B.

Keywords Evolution · Hepatitis B virus C2 strain · Bangladesh · Chronicity · Resistance · Mutations

Introduction

Hepatitis B virus (HBV) is still a global health problem despite of the presence of a potential recombinant vaccine [9, 13]. Globally more than 2 billion people have been affected and each year, more than 300 million are chronically infected which in turns may progress to cirrhosis, a life-threatening condition of liver [14]. Being an intermediately prevalent country, Bangladesh bears an increasing rate of chronic HBV infections since several years [10, 15]. From the ten genotypes, genotypes A, C and D are seen to infect people in this region, where subgenotype C2 of HBV is prevalent in the chronic patients in Bangladesh [14, 16, 17].

The molecular evolution of HBV is still a mystery; however, some documentation suggested the existence of HBV is 2000–3000 years old [8, 18]. The evolutionary rates of HBV assessed using the existing sequences are 10^2–10^4 times higher than those derived from archaeological and genetic evidence [7]. As most of the viral genome codes for multiple proteins, a synonymous change in one ORF may result in a non-synonymous mutation in the overlapping ORF. However, host immune response also plays an important role in HBV evolution [11]. Existence of new HBV strains may be related with the increasing of

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13337-018-0497-6) contains supplementary material, which is available to authorized users.

✉ Md. Salimullah
salim2969@gmail.com

[1] Microbial Biotechnology Division, National Institute of Biotechnology, Savar, Dhaka 1349, Bangladesh

[2] Molecular Biotechnology Division, National Institute of Biotechnology, Savar, Dhaka 1349, Bangladesh

Published online: 17 October 2018

pathogenicity and renders the evolution unclear. More specifically, strain HBV/C2 is observed related to the chronic cases and liver abnormalities [20]. Hence, the possible evolution of strain HBV/C2 is selected in this study.

Materials and methods

Processing of HBV full genome

A total of 462 HBV/C2 full genomes sequences were collected from NCBI GenBank. Of these sequences, 57 were documented to be isolated from Bangladesh, where 1 full-genome sequence was assembled and analyzed by our research group and the methods of isolation and assembling were reported previously [15]. The study was approved for ethical concern by National Institute of Biotechnology Ethical Review Committee and the sample was collected after consideration of patient's consent. The collected sequences were edited using MEGA6 [19] to identify them from the similar starting point.

Analysis of HBV genome for genotyping and mutations

The genotypes and subgenotypes of the collected Bangladeshi sequences were collected from the NCBI GenBank with annotations and the unannotated sequences were analyzed by Geno2phenotool (http://hbv.geno2pheno.org/index.php) for subgenotypes. The sequences were then subjected to mutation analysis using several bioinformatics tools, such as HBV Stanford database (https://hivdb.stanford.edu/HBV/HBVseq/development/HBVseq.html), Geno2Pheno tool and MEGA6 [19].

Observation of the molecular evolution of HBV/C2

The evolution of HBV was determined using bioinformatics tools such as Beast 2 software package [2] and Tracer tool (http://tree.bio.ed.ac.uk/software/tracer/). Briefly, about 462 HBV/C2 complete genome sequences deposited in the NCBIGenBank from 2000 to 2017 were collected. The collected sequences were manually double-checked and confirmed for HBV/C2 subgenotype using Geno2Pheno tool. Then the sequences along with our analyzed sequence (NHB17003) were aligned with Emboss Clustal Omega server (https://www.ebi.ac.uk/Tools/msa/clustalo/). The aligned sequences were converted to XML format using BeuTi tool of the Beast 2 software package [2] keeping the Gamma site model of 4 Gamma categories count and GTR subset model. The clock model was set as strict clock with a clock rate of 0.00079 with the initialization panel of 4 dimensional bPopSizes and bGroupSizes and analyzed for Coalescent Bayesian skyline using the Beast tool. The resulted log and tree files were again analyzed for Bayesian Skyline Reconstruction using Tracer tool with 100,000 burn-in.

Results and discussion

Genotyping and mutation profile

HBV molecular analyses still have some gaps to summarize the properties of chronic natures of HBV. Moreover, novel subgenotypes are rising that are related to the chronic cases and bear tremendous medical significance. Genomic analysis of these strains was not carried out before. Hence, the purpose of this study was to determine the molecular evolution pattern of HBV/C2 strain responsible for chronicity.

Genotyping and mutation analysis

Complete genome analysis of the isolated strain by Geno2Pheno and NCBI Genotyping tools revealed the subtype of our sequenced genome is HBV/C2, which seems to be a new strain circulating in Bangladesh. The strain contains an HBsAg escape mutation (I126T) which was reported by our previous study that the patient's plasma was negative for HBsAg test but contained a high viral load [15]. Substitution I126T was documented before to be a mutation causing HBsAg escape [15]. Interestingly, the patient's mother was also infected by HBV few days ago; however, molecular analysis was not performed at that time. Analyzing mutational hotspots of the isolated strain denoted some key mutations, A60V, I126T, R713Q and K130M in Pre-surface, surface, polymerase and X proteins respectively, which are in consistent with other study [12]. Further analysis of the rest of the Bangladeshi sequences reveals that 42.11% (24/57) strains were found to have vaccine resistant, 12.28% (7/57) was HBsAg detection resistant, 7.02% (4/57) HBV immunoglobulin (Ig) resistant, 3.51% (2/57) multi-drug resistant and 3.51% (2/57) contain secondary mutations of unknown medical significance. Of 24 vaccine resistant strains, 16.67% were observed resistant to both vaccine, HBsAg detection and Ig resistant, 4.17% was both vaccine and HBsAg detection resistant, and 4.17% was resistant to both vaccine and multi-drugs (Supplementary Table 1).

Surface gene variations

Some substitutions such as S53L, P62Q and S210N were found in the surface region of the isolated complete

genome that may have clinical significance to escape HBsAg detection and render the strain to cause chronic infection. Analysing all HBV sequences circulating in Bangladesh revealed some common surface gene substitutions such as S207N, S53L, I126T, S210N, T118V in 13, 22, 6, 21 and 15 strains respectively that may be responsible for boosting chronicity of the respective strains. Furthermore, a mutation, 128V was found in 15 strains documented to have escape properties. Some substitutions were observed in 'a' determinant region of surface protein, for example, 143/SL, 145R that may have clinical significance. Moreover, the major hydrophilic region of the surface gene of the strains contains some mutations such as, Y100C, Q101R, L110I, T113S, T115N, T118V, T126S, T127P, A128V, F134Y, T143S, G145R, A159G/V and R160K that may have clinical significance regarding failure of vaccine and B cell activity. The findings are supported by the current HBV resistance status in the world documented by several studies [3, 4, 6, 14].

Molecular evolution of the sequenced HBV/C2

Molecular evolutionary analysis by Bayesian method indicated the single nucleotide conversion is 0.8308 with a highest and lowest rate of AG (2.324) and CG (0.221) substitutions respectively (Table 1). The substation AG with 2.324 showed much more than the other conversions. Bayesian density was observed between 1.25 and $1.25E^{-2}$ analyzing the complete genome, which may signify the substitutions of most sites lie in this region, as shown in Fig. 1. Skyline analysis using HBV/C2 complete genome sequences from 2000 to 2017 revealed the evolution of the strain was in nineteenth century as shown in Fig. 2. There were two rapid sharp increases in the genetic diversity at the end of the twentieth century and then a steady circulation of HBV/C2 strain, which may denote the possible evolutions in these time intervals and gaining stability from the early twenty-first century. Skyline analysis using the S, C and X genes of the above complete sequences denotes the evolution was in the early twentieth century with similar sharp peaks of genetic diversity at the end of the twentieth century as observed in the analysis with complete genome (Fig. 2). The similar tendency of evolution of C, S and X genes of that strain again support the fact of HBV evolution describe above. Furthermore, investigating the history of HBV outbreaks at the end of twentieth century may confirm the possible evolutions of HBV at that time as suggested in this study [1, 5]. Furthermore, Bayesian skyline analysis of C and X genes of HBV revealed a sudden decrease in genetic diversity from early twenty-first century, which may be due to the initiation of mass HBV vaccinations in the world.

In conclusion, as the knowledge of progression of hepatitis B into chronicity and liver abnormalities is still in dark, the involvement of epigenetic factors needs to be known. The genome of HBV is undergoing a rapid

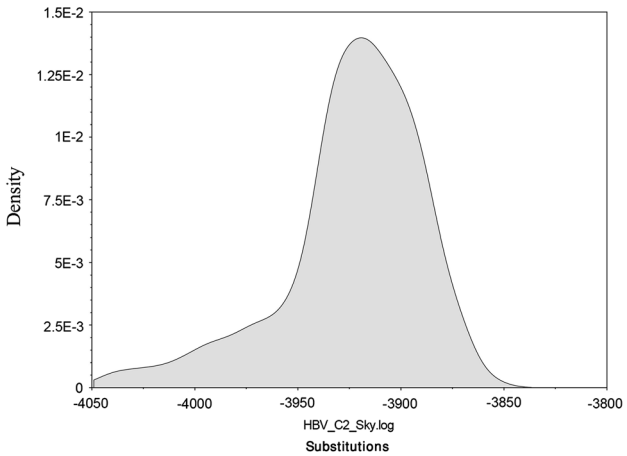

Fig. 1 Analysis of the Bayesian density of the complete genomes HBV/C2 strains, generated by Tracer v1.6

Table 1 Comparative analysis of single nucleotide conversion rate of HBV/C2 complete genome, analyzed using Beast 2 software package

Summary statistics	rateAC	rateAG	rateAT	rateCG	rateGT
Mean	0.7671	2.324	0.4122	0.2213	0.4295
Std. err of mean	8.5442E−3	0.0385	0.01	2.8427E−3	4.8975E−3
Std. dev.	0.0303	0.1281	0.024	0.0138	0.0206
Variance	9.1541E−4	0.0164	5.7795E−4	1.9179E−4	4.2343E−4
Median	0.7677	2.3051	0.4145	0.2189	0.4263
Mode	n/a	n/a	n/a	n/a	n/a
Geometric mean	0.7665	2.3205	0.4115	0.2208	0.4291
95% HPD interval	[0.7156, 0.8138]	[2.1148, 2.5326]	[0.3653, 0.4554]	[0.1987, 0.2459]	[0.3993, 0.4716]
Auto-correlation time (ACT)	71,933.87	81,298.4627	1.5724E5	38,006.6129	51,094.9969
Effective sample size (ESS)	12.5254	11.0826	5.7301	23.7064	17.6338

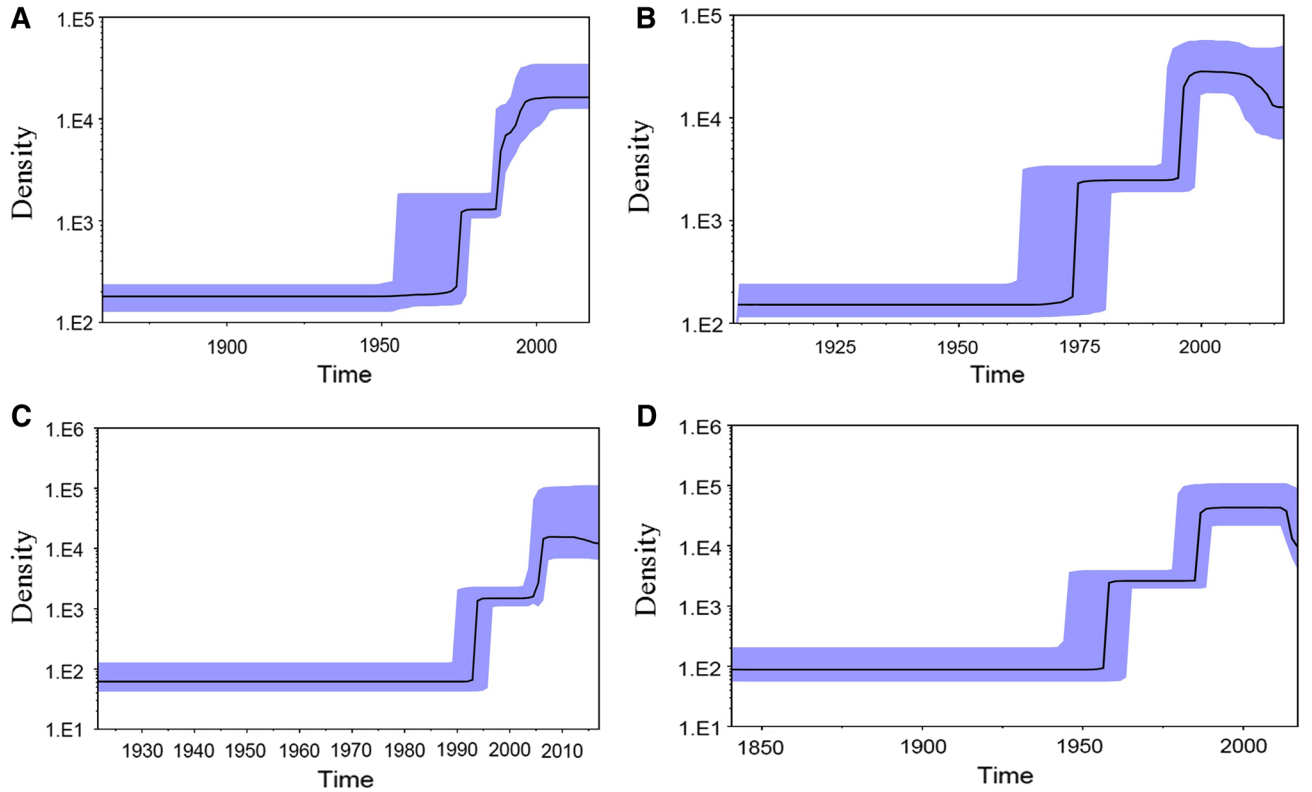

Fig. 2 Skyline analysis of the HBV/C2 strain, generated by Beast 2 software package and Tracer v1.6. **a** Molecular evolution time of HBV/C2 strain using the 462 complete genomes; **b** molecular evolution time of HBV/C2 strain using C gene; **c** molecular evolution time of HBV/C2 strain using S gene; **d** molecular evolution time of HBV/C2 strain using X gene

mutation that is increasing the pathogenicity of the virus. The study is unique for several reasons (1) the mutations responsible for multi-drug resistance, escape to the vaccines, HBsAg detection and Ig resistant were sorted out and (2) the molecular evolution of a newly observed strain in Bangladesh (HBV/C2) was predicted. The findings of this study will help researchers and clinicians to get a depth knowledge about the emergence of HBV/C2 strain and to observe the genomic variations time to time that may help to find a proper treatment strategy against hepatitis B.

Acknowledgements The study was supported by the National Institute of Biotechnology, Ministry of Science and Technology, Bangladesh.

Compliance with ethical standards

Conflict of interest The authors declare that there is no any conflict of interest.

References

1. Allen AM, Irwin GR, Karwacki JJ, Warren DC, Levine RS. Epidemic hepatitis B: a sustained outbreak in a large military population. Am J Epidemiol. 1975;102(6):545–52.
2. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, et al. BEAST 2: a software platform for bayesian evolutionary analysis. PLoS Comput Biol. 2014;10(4):e1003537. https://doi.org/10.1371/journal.pcbi.1003537.
3. de Man RA, Bartholomeusz AI, Niesters HG, Zondervan PE, Locarnini SA. The sequential occurrence of viral mutations in a liver transplant recipient re-infected with hepatitis B: hepatitis B immune globulin escape, famciclovir non-response, followed by Lamivudine resistance resulting in graft loss. J Hepatol. 1998;29(4):669–75.
4. Dos Santos M, Pacheco SR, Stocker A, Schinoni MI, Parana R, Reis MG, et al. Mutations associated with drug resistance and prevalence of vaccine escape mutations in patients with chronic hepatitis B infection. J Med Virol. 2017;89(10):1811–6. https://doi.org/10.1002/jmv.24853.
5. Gerlich WH. Medical virology of hepatitis B: how it began and where we are now. Virol J. 2013;10:239. https://doi.org/10.1186/1743-422x-10-239.
6. Hossain MG, Ueda K. Investigation of a novel hepatitis B virus surface antigen (HBsAg) escape mutant affecting immunogenicity. PLoS ONE. 2017;12(1):e0167871. https://doi.org/10.1371/journal.pone.0167871.
7. Lin YY, Liu C, Chien WH, Wu LL, Tao Y, Wu D, et al. New insights into the evolutionary rate of hepatitis B virus at different biological scales. J Virol. 2015;89(7):3512–22. https://doi.org/10.1128/JVI.03131-14.
8. Littlejohn M, Locarnini S, Yuen L. Origins and evolution of hepatitis B virus and hepatitis D virus. Cold Spring Harbor Perspect Med. 2016;6(1):a021360. https://doi.org/10.1101/cshperspect.a021360.

9. McAleer WJ, Buynak EB, Maigetter RZ, Wampler DE, Miller WJ, Hilleman MR. Human hepatitis B vaccine from recombinant yeast. Nature. 1984;307(5947):178–80.
10. Munshi SU, Tran TTT, Vo TNT, Tabassum S, Sultana N, Nguyen TH, et al. Molecular characterization of hepatitis B virus in Bangladesh reveals a highly recombinant population. PLoS ONE. 2017;12(12):e0188944. https://doi.org/10.1371/journal.pone.0188944.
11. Osiowy C, Giles E, Tanaka Y, Mizokami M, Minuk GY. Molecular evolution of hepatitis B virus over 25 years. J Virol. 2006;80(21):10307–14. https://doi.org/10.1128/JVI.00996-06.
12. Ramezani F, Norouzi M, Sarizade GR, Poortahmasebi V, Kalantar E, Magnius L, et al. Mutation hot spots in hepatitis B surface antigen in chronic carriers from Khoozestan province, southern of Iran. Iran J Allergy Asthma Immunol. 2013;12(3):269–75.
13. Shaha M, Hoque SA, Ahmed MF, Rahman SR. Effects of risk factors on anti-HBs development in hepatitis B vaccinated and nonvaccinated populations. Viral Immunol. 2015;28(4):217–21. https://doi.org/10.1089/vim.2014.0147.
14. Shaha M, Hoque SA, Rahman SR. Molecular epidemiology of hepatitis B virus isolated from Bangladesh. Springerplus. 2016;5(1):1513. https://doi.org/10.1186/s40064-016-3174-5.
15. Shaha M, Das KC, Hossain MS, Jahan M, Hashem A, Rahman SR, et al. Complete genome sequence of a circulating hepatitis B virus genotype C strain isolated from a chronically infected patient identified at an outdoor hospital in Bangladesh. Genome Announc. 2018. https://doi.org/10.1128/genomea.01601-17.
16. Shaha M, Hadisur Rahman M, Jahan M, Dey SK, Das KC, Hashem A, et al. Identification of a novel tri-genotypic recombinant hepatitis B virus in Bangladesh. Virus Res. 2018. https://doi.org/10.1016/j.virusres.2018.07.014.
17. Shaha M, Sarker PK, Hossain MS, Das KC, Jahan M, Dey SK, et al. Analysis of the complete genome of hepatitis B virus subgenotype C2 isolate NHB17965 from a HBV infected patient. F1000Research. 2018;7:1023. https://doi.org/10.12688/f1000research.15090.3.
18. Simmonds P. Reconstructing the origins of human hepatitis viruses. Philos Trans R Soc B Biol Sci. 2001;356(1411):1013–26. https://doi.org/10.1098/rstb.2001.0890.
19. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013;30(12):2725–9. https://doi.org/10.1093/molbev/mst197.
20. Zhang HW, Yin JH, Li YT, Li CZ, Ren H, Gu CY, et al. Risk factors for acute hepatitis B and its progression to chronic hepatitis in Shanghai, China. Gut. 2008;57(12):1713–20. https://doi.org/10.1136/gut.2008.157149.