

Bioinformatics analysis of polymerase and envelope genes of hepatitis B virus in Egypt

Amal Mahmoud^{1&2} and Medhat H. Hashem³

¹Bioinformatics Dep., Genetic Engineering and Biotechnology Research Institute (GEBRI), University of Sadat City, Egypt

²Alkharizmi Center for Bioinformatics, Egypt

³Animal Biotechnology Dep., Genetic Engineering and Biotechnology Research Institute (GEBRI), University of Sadat City, Egypt

ABSTRACT

Hepatitis B virus (HBV) genotype C is a leading risk factor for hepatocellular carcinoma (HCC). Eight HBV isolates from Sohag governorate, Egypt were confirmed by polymerase chain reaction (PCR), resulted in 1110 nt correspond to partial sequence of polymerase (pol) and envelop (env) genes. Two HBV isolates were associated with HCC (Sohag1 and 7). Phylogenetic analysis of the nucleotide sequence and the amino acid of the pol or env genes revealed that 7 out of the 8 isolates were closely related to genotype D. Only Sohag 1-HCC isolate was clustered with genotype C which is known as risk factor of HCC. Multiple sequence alignment resulted in unique amino acid substitutions (V₃₂₀→L and N₄₆₄→T) and (I₂₈₄→L, T₃₀₀→I and K₃₃₄→R) in pol and env genes respectively. Further prospective studies are needed to confirm the role of these mutations in the development of HCC.

Key word: HBV, PCR, HCC, env gene, pol gene.

INTRODUCTION

Hepatitis or inflammation of the liver can be due to a variety of causes of which viral infection is the most important, and leads to significant morbidity and mortality. Viral hepatitis is caused by infection with one of the five known viruses, which predominantly affect the liver – the hepatitis A, B, C, D and E viruses (Hollinger and Emerson, 2007).

Hepatitis B is caused by the hepatitis B virus (HBV). An estimated 2 billion people have been infected at some point and 350 million people across the world continue to carry chronic (long-term) infection. It is most commonly found in South East Asia, the Middle and Far East, Southern Europe and Africa. It is highly infectious, 50-100 times more so than HIV, It is the 10th leading cause of death worldwide (Lavanchy, 2004). Between 500,000 and 700,000 people die each year (WHO, 2009).

HBV is a leading risk factor for hepatocellular carcinoma (HCC), with over eighty percent of HCC cases occurring in the regions where HBV is endemic (Michielsen and Ho, 2011). Approximately 30-50% of the estimated 320,000 annual HBV-related deaths are due to hepatocellular carcinoma (HCC) (Farazi and DePinho, 2006).

Different HBV genotypes have distinct geographical distributions. Genotype A is found mainly in Northwest Europe, the United States, India, and Sub-Saharan Africa. Genotypes B and C prevail in East Asia, while genotype D is common in the Mediterranean countries. Genotype E is only found in Africa and genotype F is found mainly in Central and South America (Arauz-Ruiz *et al.*, 2002).

Nearly 2-3 million Egyptians are chronic carriers of HBV. Structural and functional differences between genotypes

can influence the severity, course and likelihood of complications.

The HBV genome comprises a partially double stranded 3.2 kb DNA organized into four open-reading frames. The longest open-reading frame encodes the viral polymerase (Pol open-reading frame). The envelope open-reading frame is located within the Pol open-reading frame in a frame-shifted manner. Partially overlapping with the envelope open-reading frame are the core (C) and the X open-reading frames. The covalently closed circular DNA (ccc DNA) is the template that is transcribed to generate four major RNA species: the 3.5-, 2.4-, 2.1-, and 0.7-kb viral RNA transcripts. Expression of these four transcripts is directed by the enhancer II/basal core, large surface antigen (L), major surface antigen (S), and enhancer I/ X gene promoters, respectively. (Ganem *et al.*, 2001).

This study aims to study the sequence variations in polymerase and envelope genes of Hepatitis B Virus in Egypt.

MATERIALS AND METHODS

DNA extraction

DNA was extracted by using QIAamp DNA extraction kit (QIAGEN GmbH, Hilden Germany). Serum samples of eight HBV positive patients from Sohag governorate-Egypt were used where isolates Sohage 1 and 7 were obtained from HBV positive patients with hepatocellular carcinoma (HCC).

PCR procedure and DNA sequencing

Sense 5'-TCACCATATTCTTGGG AACAA-3' and antisense 5'-CGAACC ACTGAACAAATGGC-3' primers were

RESULTS

Eight HBV isolates from Sohag governorate, Egypt were confirmed by

used to amplify 1110 bp of open reading frames of polymerase and S genes of HBV genome.

Briefly, 100 µl of reaction mixture containing 10 µl of extracted DNA, 50 mM potassium chloride, 10 mM Tris hydrochloric acid (pH 8.3), 2 mM magnesium chloride, 200 µM deoxyribonucleosides, 2 U of *AmpliTaq Gold DNA* polymerase (perkin-Elmer, Norwalk, Conn), and 20 pmol each of the oligonucleotide primers. Amplification was performed for one cycle at 95°C for 10 min followed by 35 cycles, each consisting of denaturing for 20 sec. at 94°C, annealing for 20 sec at 55°C, and extension for 1 minute at 72°C.

DNA sequencing was carried out with the BigDye® Terminator v3.1 Cycle Sequencing Kit and an Applied Biosystems 373xl DNA Analyzer.

Sequence analysis

Sequence analysis was done using Blast programs from National Center for Biotechnology Information (NCBI), USA

(<http://www.ncbi.nlm.nih.gov/Blast>).

One hundred sequences of HBV isolates from different types were obtained from GenBank and were used to construct Neighbor-Joining tree between Sohag and GenBank isolates.

Evolutionary relationships

The evolutionary history was inferred using the Neighbor-Joining method (Jaccard, 1908). Phylogenetic analyses were conducted in MEGA4 (Rohlf, 2000).

polymerase chain reaction (PCR), resulted in 1110 nt correspond to partial sequence of polymerase (pol) and envelop (env) genes Phylogenetic analysis of pol and

env genes nucleotide sequence (Fig.1) revealed that 7 out of the 8 isolates (Sohag 2, 3, 4, 5, 6, 7-HCC and 8) were closely related to genotype D. However, isolate Sohag 1-HCC was closely related to genotype C. Similar data was obtained when amino acid of pol and env genes were used (supplementary data). Multiple amino acid sequence alignment showed

specific substitutions distinguished genotype C and Sohag 1-HCC from the other genotypes (Fig. 2 and 3). Thus amino acid substitutions (V₃₂₀→L and N₄₆₄→T) and (I₂₈₄→L, T₃₀₀→I and K₃₃₄→R) were detected in pol and env genes respectively.

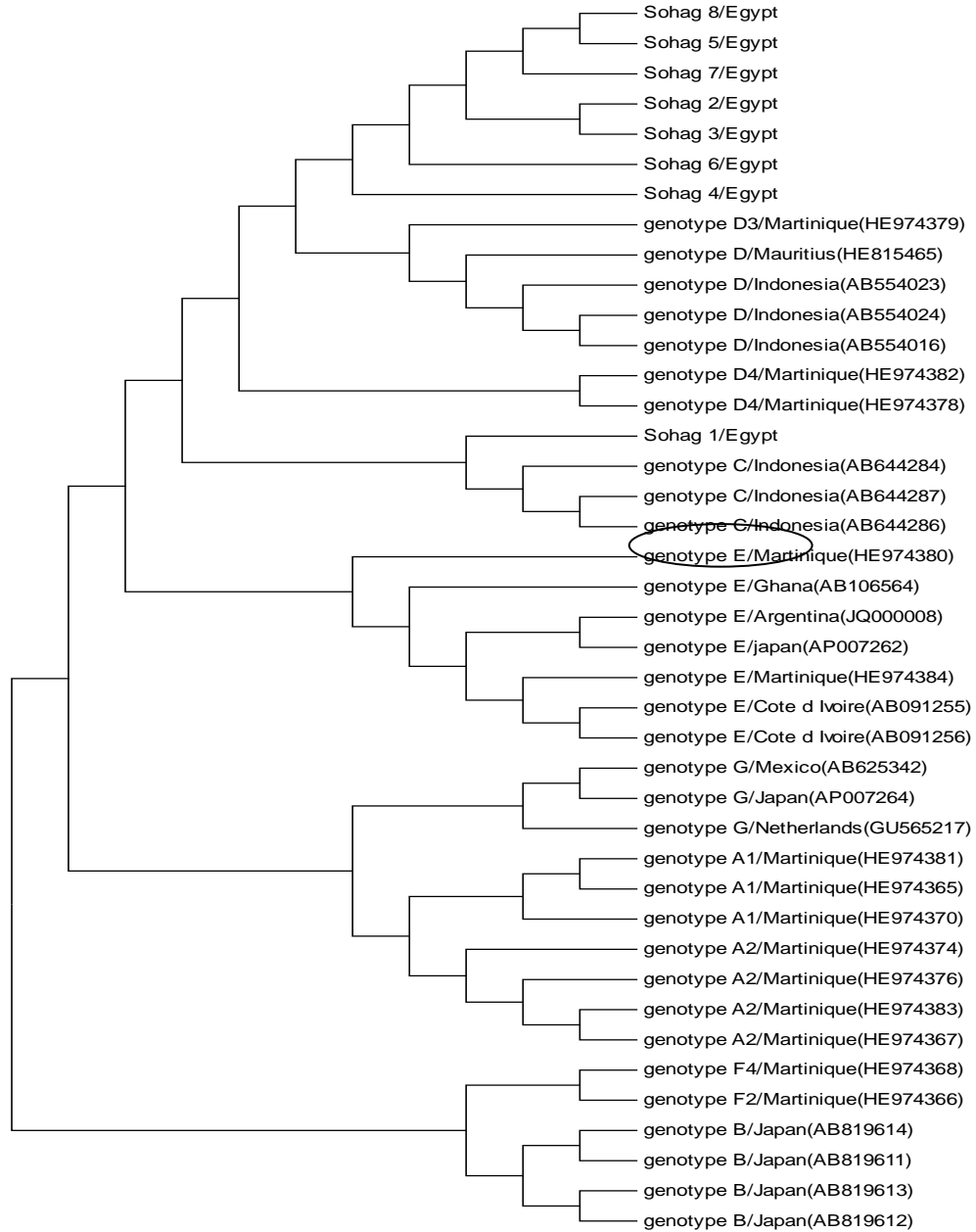


Fig 1. Phylogenetic analysis of HBV nucleotide sequences.

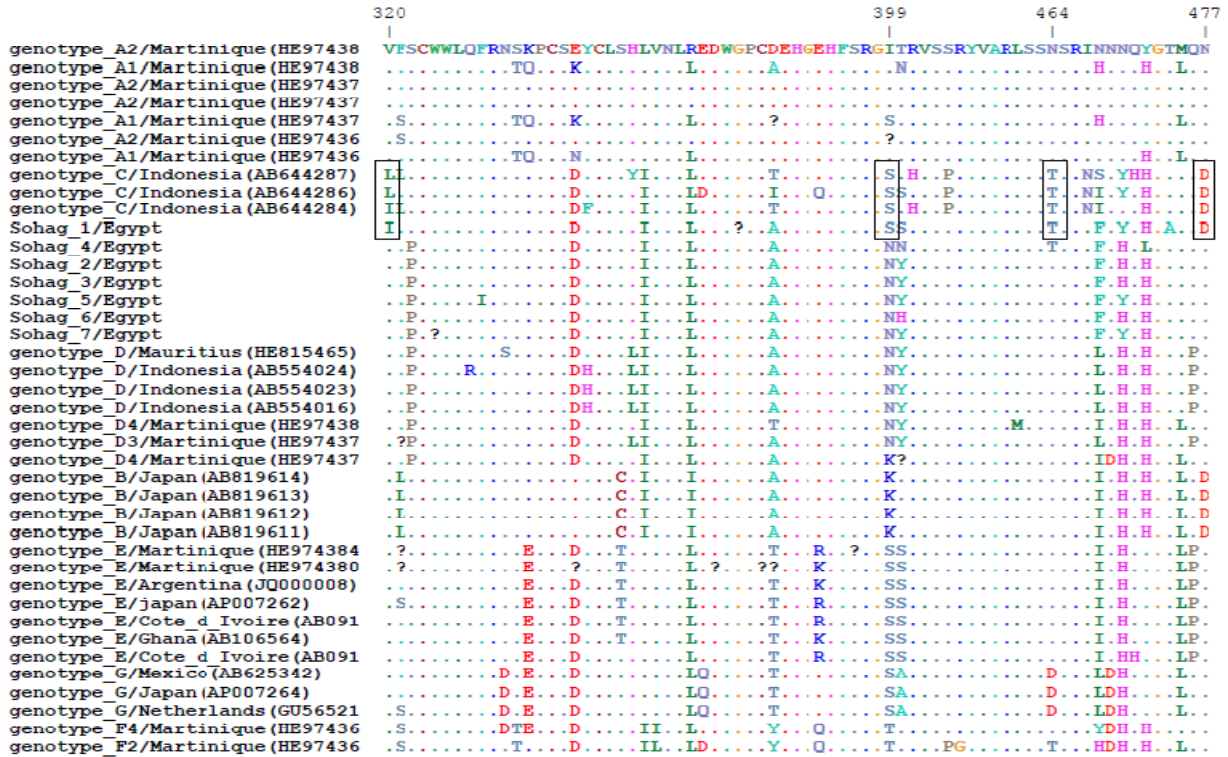


Fig 2. Multiple amino acids sequence alignment of HBV polymerase gene. Amino acids positions correspond to complete genome sequence of genotype C (accession number AB644287).

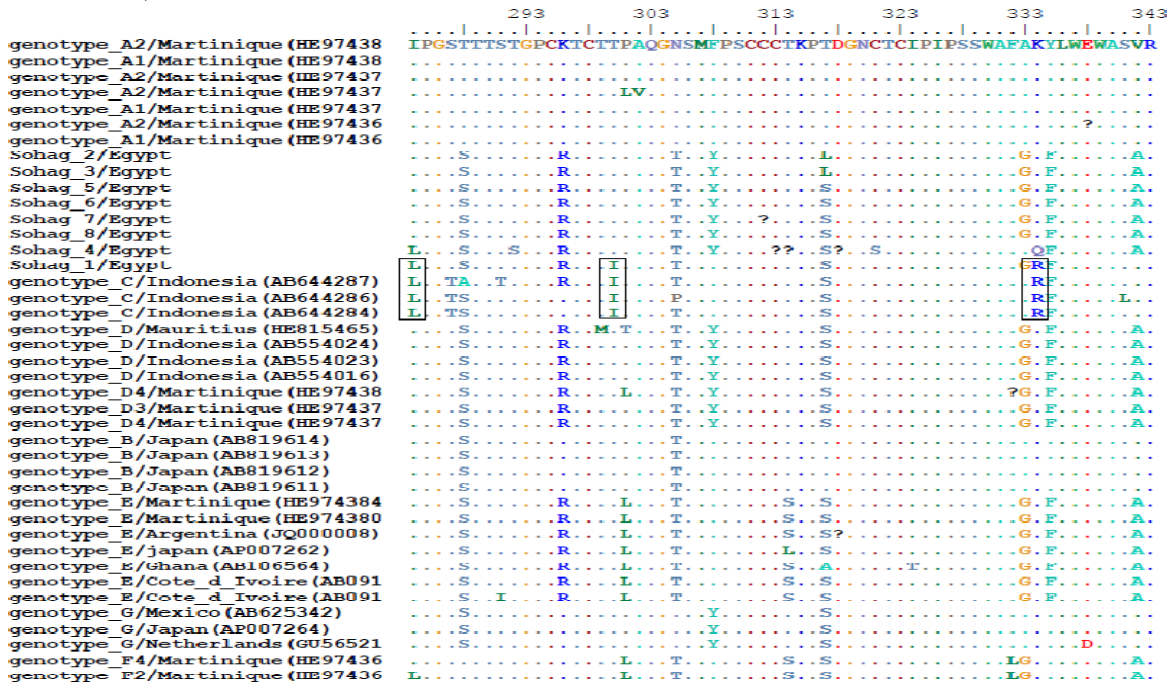


Fig 3. Multiple amino acid sequence alignment of HBV envelope gene. Amino acids positions correspond to complete genome sequence of genotype C (accession number AB644287).

DISCUSSION

HBV is a leading risk factor for hepatocellular carcinoma (HCC), with over eighty percent of HCC cases occurring in the regions where HBV is endemic (Michielsen and Ho, 2011).

In this study, multiple sequence alignment showed unique amino acid substitutions in the envelope genes of Sohag1-HCC and genotype C. These substitutions were V₃₂₀→L and N₄₆₄→T for pol gene and I₂₈₄→L, T₃₀₀→I and K₃₃₄→R for env gene.

Upon Phylogenetic analysis of the nucleotide sequence and the amino acid of pol or env genes, 7 isolates from patients without HCC (Sohag 2, 3, 4, 5 and 6) and 1 isolate from patient with HCC (Sohag 7-HCC) were closely related to genotype D. Only one isolate (Sohag 1-HCC) of the 2 HCC isolates was clustered with genotype C which is known as risk factor of HCC.

Association between HCC and HCV was previously studied (Mahmoud and Hashem 2012). Genotype D is the most widely distributed genotype and has been found universally, with its highest prevalence in a belt stretching from Southern Europe and North Africa. Genotypes B and C prevail in East Asia (Norder, *et al.*, 1994). Wong *et al.*, 2013 reported that genotype C hepatitis B virus is associated with a higher risk of HCC as compared to other major HBV genotypes (A, B and D).

As well as, mutations in S gene sequence were observed in this study that may change the antiviral target site and virus replication. The results of this study indicate that further prospective studies are needed to confirm the role of the mutations in polymerase and envelope genes in the development of HCC.

REFERENCES

- Mahmoud A, and Hashem H M, (2012).** Phylogenetic subtyping of hepatitis C virus 5' UTR isolated in Egypt and the effect of 2 transitions in subdomain III_d on the apical loop structure. *Life Science J.* 9 (1)
- Arauz-Ruiz, Norder P H, Robertson B H, Magnius L O, (2002).** Genotype H: A new Amerindian genotype of hepatitis B virus revealed in Central America. *J Gen Virol.*, 83:2059-2073
- Farazi P A, DePinho R A, (2006).** Hepatocellular carcinoma pathogenesis: From genes to environment. *Nat Rev Cancer*6: 674–687
- Wong G L-H., Chan H L-Y, Yiu K K-L, Lai J W-Y, Chan V K-K, Cheung K K-C, Wong E W-N, Wong V W-S, (2013).** *Aliment Pharmacol Ther* : 37: 517–526
- Ganem D, Schneider R. Hepadnaviridae: The viruses and their replication. In: Knipe D M, Howley P M, eds. *Fields Virology*. Philadelphia: Lippincott-Raven; 2001:2923–2970
- Hollinger F B, Emerson S U, (2007).** Hepatitis A Virus. In *Fields Virology*, 5th edn, pp. 911–940.
- Jaccard P, (1908).** Nouvelle recherches sur la distribution florale. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 44: 223-270.
- Lavanchy D, (2004).** Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. *J Viral Hepat* 11: 97–107.
- Michielsen P, Ho E, (2011).** Viral hepatitis B and hepatocellular carcinoma. *Acta Gastroenterol Belg* 74: 4–8.
- Norder H, Courouce A M, Magnius L O, (1994).** Complete genomes,

phylogenic relatedness and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology*, 198:489-503.

Rohlf F, (2000). On the use of shape spaces to compare morphometric method. *Hystrix, Italian J. Mammology* (n.s.), 11(1): 8-24.

World Health Organization, (2009). Viral hepatitis: Report by the Secretariat. http://apps.who.int/gb/ebwha/pdf_files/A62/A62_22-en.pdf