Pak. J. Biotechnol. Vol 15 (3) 793-801 (2018) ISSN Print: 1812-1837
www.pjbt.org ISSN Online: 2312-7791

BIOINFORMATICS ANALYSES OF SURFACE PROTEIN GENE OF HBV GENOME OCCURRED IN BLOOD SAMPLES FROM DIFFERENT REGIONS OF KSA

Al Harthi J.H.[a], Al-Yami M.R.[b] and Sadik A.S.[a,c*]

[a] Department of Biology, Faculty of Science, Taif University, Taif, Al-Haweiah, P.O. Box 888, Zip code 21974, Taif, KSA
[b] Child Hospital in Taif, Shehar, El-Shafa Road, Zip code 26514, Ministry of Health, KSA. [c] Department of Agricultural Microbiology, Faculty of Agriculture, Ain Shams University, P.O. Box 68, Hadayek Shobra 11241, Cairo, Egypt

Article received 27.7.2018, Revised 19.9.2018, Accepted 23.9.2018

ABSTRACT

The major surface antigen (MSA) genes of hepatitis B virus (HBV) were PCR-amplified from different DNA templates of HBV-infected blood samples. On determination of nucleotide sequences these genes were subjected to some bioinformatics analyses at the levels of DNA, amino acids and putative domains. The lengths of partial nucleotide sequences and deduced amino acids in open reading frames (ORFs) of the HBV strains under investigation were 752 & 217, 838 & 257, 893 & 231, 665 & 148 and 840 & 183 nts for HBV strains LC152750.1, LC152754.1, LC152755.1, LC152756.1 and LC152758.1, with percent identities ranged from 94-99% & 86-99%, respectively, compared to HBV overseas related strains in GenBank. The ORFs and their amino acids of the partial sequences of HBV strains confirmed the dependence of these genes to MSA. Presences of general primers of MSA genes, and primers of each of PCR, sequencing and mutagenic were detected in the PCR-amplified sequences of HBV strains. The putative conserved domains of the five HBV strains of MSA genes were belonging to vMSA (Accession ≠: pfam00695) domain which described as major surface antigen from hepadnavirus of the superfamily cl02933. Phylogenetic tree of the HBV P, S genes, partial sequence, and strains of this study confirmed that these genes were belonging to the MSA from hepadnavirus compared to the most related genes. Bioinformatics comparisons between the five HBV strains showed that they were closely related to each other, with 90-94%, nucleotides similarity.

Keywords: HBV, Surface antigen genes, vMSA domain, Similarity, Bioinformatics.

INTRODUCTION

The prevalence of chronic hepatitis B infection varies widely in different parts of the world (Sharavanan et al., 2014). Hepatitis B virus (HBV) infections are still a major health issue, with approximately 350 million people chronically infected with HBV worldwide. Therefore, HBV infection is one of the most important community health problems and also one of the most common infectious diseases in the world (Li et al., 2007; Liang et al. 2013). HBV infection is endemic in Asia and causes major public health problems worldwide (Xu et al., 2008). HBV is responsible for 50%-80% of Hepatocellular carcinoma cases worldwide (Anaedobe et al., 2015). HBV infection is an important public health problem that requires high priority efforts towards prevention and control. Active immunization is the single most important and effective preventive measure against HBV infection (Dassah et al., 2015). Suresh et al., (2016) reported that blood serves as a vehicle for transmission of blood-borne pathogens including, HBV and Hepatitis C virus (HCV). Serological and molecular data showed evidence of the circulation of a virus similar to hepatitis B virus in swine (Vieira et al., 2015). Information about the minimum copy number of HBV genomes required for infection would be useful as a reference for drug and vaccine development; for monitoring HBV patients during treatment; for screening of blood, organ, and tissue donors; and for regulating nucleic acid amplification assays for HBV (Hsia et al., 2006). A simple and effective test to identify viral genotypes would greatly aid efforts to under- stand and control the spread of this disease. The preS/S and core regions were amplified by multiplex-polymerase chain reaction (PCR) and delivered to 12 wells containing genotype-specific Invader probes. The association between mutations in the hepatitis B surface antigen (HBsAg) gene and the occurrence of Occult HBV (OBI) in patients has not been studied adequately to determine if the two are correlated (Arababadi et al., 2011). Different HBV genotypes have characteristic geographical distribution, which is important epidemiologically. HBV strains have been classified into eight different genotypes (A to H) based on >8% differences in the entire genomic sequence. Each genotype has a distinct geographic distribution. The viral genome encodes four over-lapping open reading frames (ORFs: S, C, P, and X) (Ganem et al., 2001). Genotypes A and D are predominant in Europe, Africa, and the USA, genotypes B and C are restricted to East Asia, genotype E is found in Africa, and genotype F is found in indigenous populations in Central and South America (Eroglu et al., 2004). This study aimed to carry out the bioinformatics analyzes of the nucle-

otide sequence of S protein genes of HBV amplified from DNA-blood samples obtained from different regions of Saudi Arabia, in a trail to detect any mutations at the level of DNA or protein.

MATERIALS AND METHODS

Source of serum blood samples: A total number of 10 serum blood samples infected with HBV were collected from Taif (AJMS-01-2016, AJMS-02-2016 and AJMS-03-2016); Abha (AJMS-04-2016 and AJMS-05-2016); Jeddah (AJMS-06-2016 and AJMS-07-2016); Hafr Al Batin (AJMS-08-2016) and Riyadh (AJMS-10-2016) (Table 1). Bas-ed on the gender four (AJMS-01-2016, AJMS-06-2016, AJMS-07-2016 and AJMS-09-2016 and 6 (AJMS-02-2016, AJMS-03-2016, AJMS-04-2016, AJMS-05-2016, AJMS-08-2016 and AJMS-10-2016) blood samples belonging to females and males were collected, respectively. The ages of the collected blood serum samples were ranged from 23 to 56 years.

Table 1: Source of blood samples infected with HBV.

Sample codes	Age (Year)	Gender	Regions	Source of samples
AJMS-05-2016	23	M	Abha	General Abha Hospital
AJMS-04-2016	27	M	Abha	Asir Central Hospital
AJMS-08-2016	29	M	Hafr Al Batin	General King Khalid Hospital
AJMS-07-2016	48	F	Jeddah	General King Fahd Hospital
AJMS-06-2016	56	F	Jeddah	General King Fahd Hospital
AJMS-10-2016	34	M	Riyadh	Riyadh El Shamasy Hospital
AJMS-09-2016	46	F	Riyadh	Riyadh El Shamasy Hospital
AJMS-02-2016	35	M	Taif	Central laboratory
AJMS-03-2016	41	M	Taif	King Abdul Aziz Hospital
AJMS-01-2016	52	F	Taif	Central Laboratory

HBV serological testing: BIO ELISA (BIO-RAD, la-Coquette, France) kit was used to assess HBsAg in collected blood samples. The Monolisa™ HBs Ag ULTRA kit for the detection of the surface anti- gen of the hepatitis B in human serum or plasma by the enzyme immunoassay technique was used.

PCR confirmation of HBV: Cobas *Taq*Screen MPX test was used as described by the manufacturer to confirm the presence of HBV-DNA by PCR in the collected blood samples. HBV-DNAs were amplified and detected using automated, real time PCR on the COBAS® TaqMan® Analyzer. Detec-tion of PCR products occurred *via* detection of probes which were specific for HBV. Real time detection of PCR products was accomplished by measuring the fluorescence of released reporter dyes representing the viral targets using the following primer pairs.

Primer pairs	Sequences (5'……..3')	5' Position
Pair 1	Sense: GGATGTGTCTGCGGCGTTT	5'377
	Antisense: ACCCCATCTTTTTGTTTTGTTAGGTA	5'840
Pair 2	Sense: AGACTCGTGGTGGACTTCTCT	5'252
	Antisense: CAAAAGAAAATTGGTAACAGCGGTA	5'794

Amplification reaction and PCR program: The PCR was conducted in a total volume of 25 µL tot-al which consists of Go*Taq* Green Master Mix, 2X (12.5 µL); Sense Primer, 20 µM (1.25 µL); Anti-sense Primer, 20 µM (1.25 µL); HBV-DNA Temp-late (2.50 µL); Nuclease-Free water (7.50 µL). The PCR program was started with one cycle of 95°C for five minutes followed by 35 cycles each of 95 °C, 45-55°C and 72°C for minute for each. The fin-al cycle at 72°C was extended for 10 minutes.

Sequencing of PCR products: PCR product from round 2 was gel purified and sequenced using auto-mated DNA sequencing system (ABI 3100) and BigDye® Terminator v3.1 cycle sequencing Kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions and using three primers: Sense: 5'GGA TGT GTC TGC GGC GTT T3'; Sense: 5'AGA CTC GTG GTG GAC TTC TCT3' and Antisense: 5'CAA AAG AAA ATT GGT AAC AGC GGTA3'.

Bioinformatics analysis: DNA sequence was then analyzed using BLASTN 2.2.23+ software (http:// www.ncbi.nlm.nih.gov/blast/) against all sequences in the database for genotyping of each sample. The deduced amino acids *i.e.*, open reading frame (OR-F), types of domains and their superfamilies were also analyzed using the 2.2.23 + software (http://www.ncbi.nlm.nih.gov/ blast/). The sequence that showed the lowest e-value and maximum identity was taken as the genotype of the sample analyzed. On sequencing, the general primers of S large pro-tein region, and primers of each of PCR, sequencing and mutagenic were detected in the obtained partial sequences of HBV strains as shown below.

Primer names	Sequences (5'……..3')	5' Position
	S Large protein region	
F1	CCTGCTGGTGGCTCCAGTTC	5'56
F2	YTGGCCWAAATTCGCAGTCCC	5'298
S2	GTCACCATATTCTTGGGAAC	5'2816
S3	TATTGGGGGCCAAGTCTGTA	5'752
P3	CATCCTCAGGCCATGCAGTGG	5'3193
TPR1	TCGGGAAAGAATCCCAGAGGATTGG	5'2933
	PCR products	
Pair 1	Sense: GGATGTGTCTGCGGCGTTT	5'377
	Antisense: ACCCCATCTTTTTGTTTTGTTAGGTA	5'840
Pair 2	Sense: AGACTCGTGGTGGACTTCTCT	5'252
	Antisense: CAAAAGAAAATTGGTAACAGCGGTA	5'794
	Sequencing primers	
Sense 5'252	AGACTCGTGGTGGACTTCTCT	5'252
Sense 5'377	GGATGTGTCTGCGGCGTTT	5'377
Antisense 5'794	CAAAAGAAAATTGGTAACAGCGGTA	5'794
	Mutagenic primers	
M741VAYWF	GGCTTTCAGTTATGTGGATGATGTGGTATTGGG	5'741
L669MAYWF	CAGCCCGTTTCTCATGGCTCAGTTTACTAGTGCC	5'669

RESULTS AND DISCUSSIONS

HBV causes B, a serious and common infectious disease of the liver (Almeida et al., 2012; Dai et al., 2015). Chronic infection is associated with an increased risk to develop severe liver diseases, including liver cirrhosis, and hepatocellular carcinoma (HCC), one of the most common forms of human cancer (Anaedobe et al., 2015). In this study *S* protein gene belonging to major surface antigen of HBV was PCR-amplified using the DNA templates extracted from 10 HBV-infected blood samples representing different regions (Taif, Jeddah, Riaydh, Hafr Al Batin and Abha) of KSA. This was followed by bioinformatics analysis at differ-rent levels *i.e.*, DNA, deduced amino acids (ORFs) and putative domains compared to related HBV strains in GenBank. The experimental results showed that the lengths of partial nucleotide sequences of the HBV strains under investigation were 752, 838, 893, 665 and 840 nts for HBV strains AJMS-01-2016 (LC152750.1), AJMS-05-2016 (LC1527-54.1), AJMS-06-2016 (LC152755.1), AJMS-07-2016 (LC152756.1) and AJMS-09-2016 (LC15 27-58.1), respectively, with percent identities ranged from 94 to 99% compared to HBV overseas related strains in GenBank. These sequences were identified as small surface envelope protein gene (LC1-52750.1), preS protein gene (LC1527 54.1 and LC152755.1) and large S protein gene (LC1527-56.1 and LC152758.1) of HBV as shown in Table 2. The HBV genome encodes four overlapping open reading frames (ORFs) that are translated into viral core protein, surface proteins, polymerase/ reverse transcriptase (RT), and heaptitis B virus X (HBx). PCR products were directly sequenced from both directions and analyzed as described earlier (Datta et al., 2006). The HBV genotypes are geographically distributed. Genotype A is predominant in North America, Northern Europe, India, and Africa. Genotypes B and C are predominant in Asia (Echevarría et al., 2005).

Table 2: Characters of nucleotide sequences, ORFs and putative domains of the S protein genes of HBV strains LC152750.1, LC152754.1, LC152755.1, LC152756.1 and LC152758.1.

Levels	Characters	HBV strains belonging to S protein genes				
		LC152750.1	LC152754.1	LC152755.1	LC152756.1	LC152758.1
DNA	Length (nts)	752	838	893	665	840
	Identities (%)	98-99	96	94	96	95-96
Protein	ORF#-Frame	1-1	1-3	1-1	1-3	1-3
	ORF Length (nts)	654	771	696	447	552
	Eaa	217	257	231	148	183
	Stop codon	+	-	+	+	+
	Gene	Small S envelope protein	pre-S protein	pre-S protein	S protein	preS protein
	Identities (%)	93-99	89-91	86	95-96	89-91

Eaa: Encoding amino acids. +: Present. -: Absent.

The lengths of ORFs of HBV strains (LC1-52750.1, LC152754.1, (LC152755.1, LC15-2756.1 and LC152758.1) were 654, 771, 696, 447 and 552 nts and deducing 217 (with stop codon), 257, 231 (with stop codon), 148 (with stop codon) and 183 (with stop codon) amino acids, respectively. The identities (%) between these strains and HBV overseas related strains in GenBank were ranged from 86%, for LC152755.1, to 99 % for LC1527-50.1. Also, the ORFs and their amino acids of the partial sequences of HBV strains confirmed the dependence of these genes to major surface antigen (Table 3).

Table 3: Genes identification of five partial sequences of HBV strains (LC152750.1, LC152754.1, (LC152755.1, LC152756.1 and LC152758.1) based on DNA and ORFs and maximum identities compared to those related HBV strains in GenBank.

HBV isolates	Genes based DNA		Genes based ORFs (Protein)	
	Genes	Identity (%)	Genes	Identity (%)
LC152750.1	polymerase (P)	99	small surface envelope protein	99
LC152754.1	preS protein	96	putative preS protein	89
LC152755.1	preS protein	94	putative pre-S protein	86
LC152756.1	large S protein	96	S protein	95
LC152758.1	large S protein	96	putative preS protein	89

It is well known that the largest ORF in the HBV genome encodes for the hepatitis B polymerase protein (HBp) (Toh et al., 1983). Based on the nucleotide sequences and deduced amino acids (ORFs) differences of 04 & 03, 36 & 15, 52 & 29, 26 & 08 and 31 & 15, respectively, were recorded between the partial sequences of HBV strains: LC152750.1, LC152754.1, LC152755.1, LC152-75-6.1 and LC152758.1 and most related overseas HBV strains in GenBank (Table 4). Phylogenetic tree of the HBV P, S genes, partial sequence, and strains: LC152750.1, LC152754.1, LC152755.1, LC152756.1 and LC152758.1 of this study compared the most related genes confirmed that these genes are belonging to the major surface antigen from hepadnavirus (Fig. 1). Based on the complete genome sequence the Iranian HBV isolates have grouped into eight genotypes, A to H. The complete genome sequence (around 3185 bp) organization and phylogenetic analysis of the five-partial sequence of HBV strains, were obtained from Iranian chronic infected patients were studied (Amini-Bavil-Olyaee et al., 2005). Phylogenetic analysis based on the precore/core gene sequence revealed that all strains were of genotype D, sub-genotype D1 with bootstrap value 99%. The S gene encoded Arg122, Pro127, and Lys160 corresponding to sub-type ayw2. All strains had a nucleotide length of 3,182 bp except IR-P4 strain, with a 3,185 bp in length and with a unique Phe89 insertion in the X gene (Amini-Bavil-Olyaee et al., 2005).

Table 4: Total differences between the nucleotide sequences and deduced amino acids (ORFs) of the five partial sequences of HBV strains (LC152750.1, LC152754.1, (LC152755.1, LC152756.1 and LC152758.1) compared to the most related overseas strains in GenBank.

HBV isolates	DNA		ORFs	
	MROS	Total differences	MROS	Total differences
LC152750.1	KP712860.1	04	ABC00823.1	03
LC152754.1	FJ715414.1	36	ABV21658.1	15
LC152755.1	FJ715412.1	52	ABV02797.1	29
LC152756.1	KM108606.1	25	ALG40620.1	08
LC152758.1	JX125373.1	31	ABV21658.1	15

MROS: Most related overseas strain.

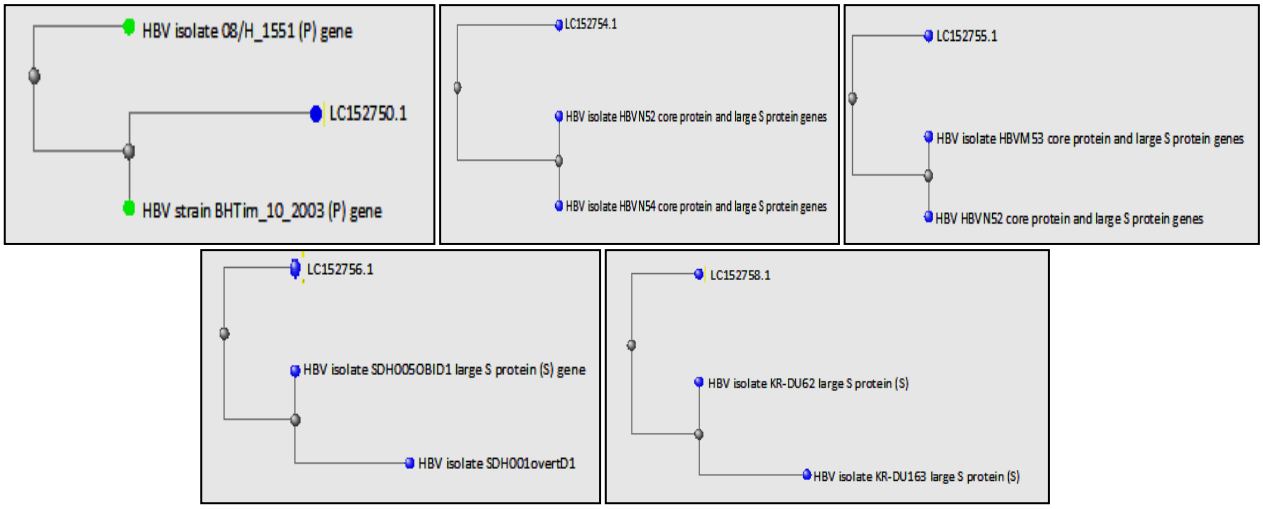

Figure 1. Phylogenetic tree of surface protein region genes of HBV strains LC152750.1, LC152754.1, LC152755.1, LC152756.1 and LC152758.1 and most related overseas strains in GenBank.

Presences of general primers of S large protein region genes, and primers of each of PCR, seq-uencing and mutagenic were detected in the PCR-amplified sequences of HBV strains LC152-750.1, LC152754.1, LC152755.1, LC152756.1 and LC15-2758.1 (Table 5). The results showed that there was a difference in the presence of primers in the PCR-amplified DNAs of the five HBV strains under investigation, where the total number of 7, 3, 2, 3 and 3 primers were recorded in the HBV strains LC152750.1, LC152754.1, LC1-52755.1, LC1527- 56.1 and LC152758.1, respectively.

Table 5: Presences of general primers of S large protein region genes, and primers of each of PCR, sequencing and mutagenic recorded in the PCR-amplified sequences of HBV strains LC152750.1, LC152754.1, LC152755.1, LC152756.1 and LC152758.1.

Primers names	LC152750.1	LC152754.1	LC152755.1	LC152756.1	LC152758.1
F1	-	+	-	-	+
F2	+	-	-	-	-
S2	-	+	+	-	+
S3	+	-	-	-	-
P3	-	+	+	-	-
TPR1	-	-	-	-	-
F58	+	-	-	+	+
R58	+	-	-	-	-
F252	+	-	-	-	-
M741VAYWF	+	-	-	+	-
L669MAYWF	+	-	-	+	-
Total	7	3	2	3	3

The primers of the surface protein region genes were repeated more frequently than other primers. Results are illustrated in Fig. 2. Data showed that two differences in the nucleotide sequences of two primers (F2-0298: 5'YTG GCC AAA ATT CGC AGT CCC3' and P3-3193: 5'CAT CCT CAG GCC ACG CAG TGG3' of major surface antigen region were recorded in the PCR-amplified sequence genes of HBV strains: LC152750.1 and LC152755.1. Bioinformatics comparisons between the five HBV strains under investigation showed that they were closely related to each other, with 90-94%, nucleotides similarity. Bartholomeusz et al., (1998) showed that sequence of the BCP/pre-core region revealed no double mutations at nucleotides 1762/1764 positions of basal core promoter or at nucleotide 1896 of the precore region, nucleotide at 1858 was T. The putative conserved domains of the five HBV strains (LC152750.1, LC-152754.1, (LC15-2755.1, LC152756.1 and LC152-758.1) of surface protein region genes under inves-tigation were belonging to vMSA (Accession num-ber: pfam-00695) domain (Fig. 3) which described as major surface antigen from hepadnavirus of the superfamily cl02933 with Query interval of 10-212, 82-153, 99-231, 1-148 and 82-153 as well as Query E-value of 4.19e-85, 3.17e-08, 7.67e-29, 7.64e-51 and 3.17e-08, respectively.

TCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACAGGCGGTGGTTTTTCTTGTTGACAATAATCCTCACTATACCGCATAGTCTAGACTCGTGGTGGACTT
CTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCCAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCTTGTCCTCCAACTTGTCCTGGTTA
TCGCTGGATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTT
TGTCCTCTAATTCCAGGATCCTCAACAACCAGCACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGTTGCTGTACCA
AACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTT
ACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTG
AGTCCCTTTTTACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCTCTAAATTTTATGGGTTA
TGTCATTGGATGTTATGGGTCCTTGCCACAA

Forward 58-AGACTCGTGGTGGACTTCTCT-38
F2:sense, nt 298-320, YTGGCCWAAATTCGCAGTCCC
Forward 58-GGATGTGTCTGCGGCGTTT-38
L669MAYWF(58-CAGCCCGTTTCTCATGGCTCAGTTTACTAGTGCC-38
M741VAYWF(58-GGCTTTCAGTTATGTGGATGATGTGGTATTGGG-38
S3 (Forward, nt752-771): 5'-TATTGGGGGCCAAGTCTGTA-3'
Reverse 58-ACCCCATCTTTTTGTTTTGTTAGGTA-38 840

LC152750.1

AAGGCATTAAACCGTATTATCCTGAATATCCAGTTAATCATTACTTCAAAACTAGGCATTATTAACATACTCTGTGGAAGGCTGGCATTCTATATAAGAGAGA
ATCTACACGCAGCGCCTCATTATGTGGGTCACCATATTCTTGGGAACAAGAGCTACAGCATGGGAGGTTGGTCTTCCAAACCTCGATAAGGAATTTGGACGAG
TCTTaCTGTTCCCAATCGTGTGGGATTATTaCCAGATCACCAGTTGGACCCTGCGTTCGGAGCCAACTCAAACAATTCAGACAGGGACTTCAACTCCAACAAG
GATCACCGACCAGAGGCAAATCAGGTAGGAGCGGGAGCTTTTGGGCCAGGGTTCACCCCACCACACGGCGGTCTTGTGGGGTGGAGCCATTAGGCTCAGGGCG
TATTGACAACAGAGCCAGCCGGACATCCTCGTGCCTCCGCCAATCGGCAGTCAGGAAGACAGCCTACTCCCATCTCTCCACCTCTAAGAGACAGTCATCCTCA
GGCCATGCAGTGGAACTCCACAACATTCCACCAAGCTATGCTAGATCCCAGAGTGAGGGGCCTATATTTTCCTGCTGGTGGCTCCAGTTCCGGAACAGTAAAC
CCTGTTCCGACTACTGCCTCACCCATATCGTCAATCTTCTCGAGGACTGGGGACCCTGCACCGAACATGGAGAACACAACATCAGGATTCCTAGGACCCCTGC
TCGTGTTACAGGCGGGGTCTTTCTTGTTGACAAGAATCCTCACAATACCACAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTATCTAGGGGGAGCACCCAC
GTGTCCTGGCCAAATT

S2: Forward, nt2816-2835) GTCACCATATTCTTGGGAAC
P3: Forward, nt3193-3213) CATCCTCAGGCCATGCAGTGG
F1: sense, nt 56-76, CCTGCTGGTGGCTCCAGTTC

LC152754.1

ATATTATTAATAGATGTCAACAATATGTGGGCCCTCTTACAGTTAATGAAAAAAGGAGATTAAATTTAATTATGCCTGCTAGGTTTAATCCTAACCTTACCAA
ATATTTGCCCTTGGACAAAGGCATTAAACCGTATTATCCTGAATATGCAGTTAATCATTACTTCAAATCTAGGCATTATTAACATACTCTGTGGAAGGCTGGC
ATTCTATATAAGAGAGAGACTACACGCAGCGCCTCATTTTGTGGGTCACCATATTCTTGGGAACAAGAGCTACAGCATGGGAGGTTGGTCTTCCAAACCTCGA
CAAGGCATGGGGACGAATCTTTCTGTTCCCAATCCTCTGGGATTCTTTCCCGATCACCAGTTGGACCCTGCGTTCGGAGCCAACTCAAACAATCCAGATTGGG
ACTTCAACCCCAAAAAGGATCACTGGCCAGAGGCAAATCAGGTAGGAGCGGGAGCATCCAGACCAAGGTTCACCCCACTACAAGGCAGTCTTTTGGGGTGGAG
CCCTCGGGCTCCGGGCACATTGACCACAGTGCCGGCAGCACCCCCTCCGGCCTCCACAAATCTTCTGGCAGGAAGAACGCCTACTCCAATATCTCCACCTGTA
CGAAACACTCATCCTCAGGCCACGCAGTGGAACTCCACGACAGTCCACCCAGCTCTGCTAGAGCACAGAGTGAGGGGCCGATACTTTCCTGCTGGTGGCTCCG
GTTCCGGAAGGGTAAATCCTGTTCCGACTACTGCCTCACCCACATCGTAGGTCCTCTCGAGGACTGGGGACCCTGCACCGAACATGGAGAACACAACATCAGG
ATTCCTAGGACCCCTGCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCATCCCAATACCACA

S2: (Forward, nt2816-2835) GTCACCATATTCTTGGGAAC
P3: (Forward, nt3193-3213) CATCCTCAGGCCATGCAGTGG

LC152755.1

CCTCCTGTCCTCCAACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTAATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCT
ACTGGACTATCAAGGTATGTTGCCGGTTAGTCCTCTAATTCCAGGATCTTCAACCACCAGCACGGGACCATGCAGGACCTGCACGACTCCTGCTCAAGGAACC
TCTTTGTATCCCTCCTGTAGCTGTACCAAACCTTCGGACGGAAATTGCACGGTTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCCTATGGGAGTGGG
CCTCAGCCCGTTTGTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTA
TTGGGGGCCAAGTCTGTACAGCATCGTGAGTCCCTGATTACCGCTGTTACGCAATATTCCTACTTGTACTTTGGGTATACATTTAAACCTCTAACAAAATTAA
ATGATGGGGTTACTCTTTACATTTCATGGGCTATGTCATTGGATGTTATGGGTCATTGCGACAAATTCACATCATACAGAAATTCAAAGAATGTTTAAGGAAA
CTTCCTGTTAACAGGCCTATTGATTGGAAAGTCTGTCAACGTATTGT

Forward: 58-GGATGTGTCTGCGGCGTTT-38
L669MAYWF: Sense(58-CAGCCCGTTTCTCATGGCTCAGTTTACTAGTGCC-38)
M741VAYWF: Sense(58-GGCTTTCAGTTATGTGGATGATGTGGTATTGGG-38)

LC152756.1

CATTAAACCTTATTATCCGGAACACGCAGTTAATCATTATTTCAAAACTAGGCATTATTAACATACTCTGTGGAAGGCTGGCATTCTATATAAGAGAGAATCC
ACACGCAGCGCCTCATTTAGTGGGTCACCATATTCTTGGGAACAAGAGCTACAGCATGGGAGGTTGGTCTTCCAAACCTCGACAAGGCATGGGGACGAATCTT
TCTGTTCCCAATCCTCTGGGATTCTTTCCCGATCACCAGTTGGACCCTGCGTTCGGAGCCAACTCAAACAATCCAGATTGGGACTTCAACCCCAACAAGGATC
ACTGGCCAGAGGCAAATCAGGTAGGAGCGGGAGCATTCAGACCAGGGTTCACCCCACTACAAGGCGGTCTTTTGGGGTGGAGCCCTCGGAGCTCCGGGCACAT
TGACAACAGATGCCAGCAGACCCCCCACCTGCGTCCACAAATCGGCAGTCAGGTAGACAGCCTACTCCAATCTCTCCACCTCTACCAAACACTCATCCTCACG
CCATGCAATGGAACTCCACAACAGTCGACCAAGCTCTGCTAGAGCACAAGTGAGGGGCCTATATTTTCCTGCTGGTGGCTCCAGTTCCGGAACAGTAAACCCT
GTTCCGACTACTGCCTCACCCATATCGTCAATCTTCTCGAGGACTGGGGACCCTGCACCGAACATGGAGAACACAACATCAGGATTCCTAGGACCCCTGCTCG
GTTACAGGCGGGGTTTTACTTGTTGACAAGAATCCTCACAATACCACAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGT
CCTGGCCTAAATTCGC
S2: (Forward, nt2816-2835) GTCACCATATTCTTGGGAAC
F1: (sense, nt 56-76, CCTGCTGGTGGCTCCAGTTC
Forward: 58-AGACTCGTGGTGGACTTCTCT-38

LC152758.1

Figure 2. Primers of surface protein genes recorded in the partial sequences of HBV strains LC152750.1, LC1527-54.1, LC152755.1, LC152756.1 and LC152758.1 belonging to small envelope protein, prepreS, prepreS, large S and large S, protein genes.

Figure 3. Putative conserved domains of the ORF number 1 of the partial sequences of HBV strains LC152750.1, LC152754.1, LC152755.1, LC152756.1 and LC152758.1 belonging to small envelope protein, prepreS, prepreS, large S and large S, protein genes.

The differences in amino acids between the putative conserved domains of the five partial sequences of HBV strains (LC152750.1, LC152754.1, (LC152755.1, LC152756.1 and LC152758.1) belonging to the vMSA (pfam00695) domain and the superfamily cl02933 were 1, 3, 11, 1, and 3, respectively (Fig. 4). Carman (1997) showed that following vaccination HBV variants with mutations in the surface genes of HBV have now been described. As a conclusion, some HBV-infected blood samples were collected from different regions of KSA, and the presence of HBV was confirmed by using ELISA and Rt-PCR techniques. Primers specific to HBV-S protein gene were successfully used to flank S gene from the DNA extracted from blood samples as templates among PCR. The nucleotide sequences of S protein gene, was subjected to bioinformatics analyses based on the level of types of domains, its accession, and superfamily compared to those similar in GenBank, and was described as major surface antigen from hepadnavirus of the superfamily cl02933.

```
gi 138789      30 .[163].FFLLTRILTIPQSLDSWWTSLNFLGGSPVCLGQNSQSPTSNHSPTSCPPICPGYRWMCLRRFIIFLFILLLCL 265
LC152750.1      1 .[  9].FFLLTIILTIPHSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCL  82
gi 138787      59 .[164].YFLWTKILTIAQSLDWWWTSLSFPGGIPECTGQNLQFQTCKHLPTSCPPTCNGFRWMYLRRFIIYLLVLLLFL 295
gi 172045835    1 .[181].FFLLIKILEILRRLDWWWISLSSPKGKMQCAFQDTGAQISPHYVGSCPWGCPGFLWTYLRLFIIFLLILLVAA 254
gi 138788       1 .[184].FFLLTKILEILRKLDWWWISLSSPKEKMLCAFQNTGAQTSPHYVGSCPWGCPGFLWTYLRLFIIFLLLLLVAA 257

gi 138789     266 IFLLVLLD.[54].SSWAFAKYLWEWASVRFSWLSLLVPFVQWF.[1].GLSPTVWLSAIWMMWYWGPSLYSIVSPFIPL 389
LC152750.1     83 IFLLVLLD.[54].SSWAFGKFLWEWASARFSWLSLLVPFVQWF.[1].GLSPTVWLSVIWMMWYWGPSLYSILSPFLPL 206
gi 138787     296 TFLLVLLD.[52].SSWALGSYLWEWALARFSWLSLLVPLLQWL.[1].GISLTVWLLLIWMIWFWGPVLMSILPPFIPI 417
gi 172045835  255 GLLYLTDN      MSIILGKLQWESVSALFSSISSLLPSDQKS     LVALIFGLLLIWMTSSSATQTLVTLTQLATL 323
gi 138788     258 GLLFLTEN      KSTIFEKLQWESVSALSSSIYSLLPSEPKS     LVALTFGLFLIWTTSSSVTQVLVTLTQLATL 326

gi 138789     390 LPIFFC 395
LC152750.1    207 LPIFFC 212
gi 138787     418 FALFFL 423
gi 172045835  324 SALFYK 329
gi 138788     327 SALFFK 332

gi 138789      30 LDPAFGANSNNPDWDFNPVKDDWPAANQVG.[1].GAF.[1].PRLTPPH.[2].ILGWS.[15].PAST     NRQSG 102
LC152754.1     82 LDPAFGANSNNPDWDFNPNKDHWPEANQVG.[1].GAF.[1].PGFTPLQ.[2].LLGWS      PRSS.[1].HIDNR 140
gi 138787      59 ILMTRYKEIDWDNWQGFPVNQRLPVSNNNP     PSG     QRAETFE     IKSRP.[15].PQTP.[1].NRDQR 128
gi 172045835    1 MGQQPAKSMDVRRIEGGELLLNQLAGRMIP.[1].GTV.[4].KFPTIDH     LLDHV.[15].PAGA.[1].RRLGL 75
gi 138788       1 MGHTQAKSTTDRRVEGGELLLQHLAGRMIP.[4].GPI.[4].KFPTIQH     VMDHI.[15].PEGT.[1].RRLGL 78

gi 138789     103 RQPTP.[1].SPPLRDS.[280]. 395
LC152754.1    141 CQQTP.[1].LRPQIGS.[ 31]. 184
gi 138787     129 RKPTP.[1].TPPLRDT.[282]. 423
gi 172045835   76 TNPTP.[1].ETPQPQW.[241]. 329
gi 138788      79 DQPRP     TPPPITW.[242]. 332

gi 138789      30 LDPAFGANSNNPDWDFNPVKDDWPAANQVG.[1].GAF.[1].PRLTPPH.[2].ILGWSPQAQGILTTVSTIPPPAST 97
```

Figure 4. Alignment between the superfamily (cl02933) of vMSA domain (pfam00695) of ORFs of the partial sequen- ces of HBV strains LC152750.1, LC152754.1, LC152755.1, LC152756.1 and LC152758.1 which belongs to the major surface antigen from hepadnavirus superfamily.

Acknowledgments

The authors would like to thank King Abdulaziz City for Science and Technology (KACST) represented by the General Directorate for Research Grants through Research Proposal # AT1518-37 for their financial support to accomplish this study. Thanks also due to Prof. Sonya Mohamed, ARC, Giza, Egypt for her sincere help to accomplish the bioinformatics analyses.

REFERENCES

Almeida, M.K.C., de Santos, K.N., dos Fecury, A.A., Oliveira, C.S.F., de Freitas, A.S., Quaresma, J.A.S., Fuzii, H.T. and L.C. Martins, Prevalence of viral hepatitis B and C in riverside communities of the Tucurui Dam, Para, Brazil. Journal of Medical Virology 84: 1907-1912 (2012).

Amini-Bavil-Olyaee, S., Sarrami-Forooshani, R., Mahboudi, F., Sabahi, F., Adeli, A., Noorinayer, B., Azizi, M. and M.R. Zali, Genotype characterization and phylogenetic analysis of hepatitis B virus isolates from Iranian patients. J. Med. Virol. 75: 227-234 (2005).

Anaedobe, C., Fowotade, A., Omoruyi, C. and R. Bakare, Prevalence, sociodemographic features and risk factors of Hepatitis B virus infection among pregnant women in Southwestern Nige-ria. Pan African Medical Journal 20: 406 (2015).

Arababadi, M.K., Pourfathollah, A.A., Jafarzadeh, A., Hassanshahi, G., Salehi, M., Ahmadabadi, B.N. and D. Kennedy, Hepatitis B virus genotype, HBsAg mutations and co-infection with HCV in occult HBV infection. Clinics and Research in Hepatology and Gastroenterology 35: 554-559 (2011).

Bartholomeusz, A., Schinazi, R. and S. Locarnini, Significance of mutations in the hepatitis B virus polymerase selected by nucleoside analogues and implications for controlling chronic disease. Viral Hepatitis Rev. 4:67-187 (1998).

Carman, W.F., The clinical significance of surface antigen variants of hepatitis B virus. J. Viral. Hepat. 4: 11-20 (1997).

Dai, J.J., HuaMing, T., QingXi, M. and Z. Quan Hong, Anti-hepatitis B virus activities of friedelolactones from Viola diffusa Ging. Phytomedicine 22: 724-729 (2015).

Dassah, S., Sakyi, S.A., Frempong, M.T., Luuse, A.T., Ephraim, R.K.D. and E.O.A. Anto

Oduro, Serocon-version of hepatitis B vaccine in young children in the Kassena Nankana district of Ghana: A Cross-sectional Study. PLoS ONE 10: 1-10 (2015).

Datta, S., Banerjee, A., Chandra, P.K., Chowdhury, A. and R. Chakravarty, Genotype, phylogenetic analysis and transmission pattern of occult hepatitis B virus (HBV) infection in families of asymptomatic HBsAg carriers. J. Med. Virol. 78: 53-59 (2006).

Echevarrìa, J.M., Avellòn A. and L.O. Magnius, Molecular epidemiology of hepatitis B virus in Spain: identification of viral genotypes and prediction of antigenic subtypes by limited sequencing. J. Med. Virol. 76: 176–184 (2005).

Eroglu, C., Leblebicioglu, H., Gunaydin, M., Turan, D., Sunbul, M., Esen, S. and A. Sanic, Distinguishing hepatitis B virus (HBV) genotype D from non-D by a simple PCR. Journal of Virological Methods 119: 183-187 (2004).

Ganem, D. and R.J. Schneider, Hepadnaviridae and their replication. In: D.M. Knipe, P.M. How-ley, D.E. Griffin, M.A. Martin, R.A. Lamb, B. Roizman, et al., editors. Fields Virology. 4. Philadelphia, PA: Lippincott-Raven Publishers (2001).

Garson, J.A., Grant, P.R., Ayliffe, U., Ferns, R.B. and R.S. Tedder, Real-time PCR quantitation of hepatitis B virus DNA using automated sample preparation and murine cytomegalovirus internal control. Journal of Virological Methods 126: 207-213 (2005).

Hsia, C.C., Purcell, R.H., Farshid, M., Lachenbruch, P.A. and M.W. Yu, Quantification of hepatitis B virus genomes and infectivity in human serum samples. Transfusion 46: 1829-1835 (2006).

Li, W., She, R., Wei, H., Zhao, J., Liu, L., Li, X., Wang, Y. and D. Wang, Detection of swine hepatitis B virus and E viruses in the liver and serum in pigs in China. Veterinary Science in China 37: 921-925 (2007).

Liang, X., Bi, S., Yang, W., Wang, L., Cui, G., Cui, F., Zhang, Y., Liu, J., Gong, X., et al., Reprint of: epidemiological serosurvey of hepatitis B in China-declining HBV prevalence due to hepatitis B vaccination. Vaccine 31: J21 -J28 (2013).

Sharavanan, T.K.V., Premalatha, E. and N. Dinakaran, Seroprevalence of hepatitis B surface antigen among rural pregnant women attending a tertiary care hospital. Scholars Journal of Applied Medical Sciences 2: 1351-1354 (2014).

Suresh, B., Babu, K.V.S., Mouli, P.C. and D.S. Jothibai, A study on patterns of co-infections among blood donors at the blood bank of a tertiary care referral teaching hospital in South India. Journal of Clinical and Scientific Rese-arch 5: 105-111 (2016).

Toh, H., Hayashida, H. and T. Miyata, Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus. Nature 305: 827-829 (1983).

Vieira, Y.R., Silva, M.F.M., Santos, D.R.L., Vieira, A.A., Ciacci-Zanella, J.R., Barquero, G., Lago, B.V., do Gomes, S.A., Pinto, M.A., Paula and V.S. de, Serological and molecular evidence of hepadnavirus infection in swine. Annals of Agricultural and Environmental Medicine 22: 11-16 (2015).

Xu, J., Wang, J., Deng, F., Hu, Z. and H. Wang, Green tea extract and its major component epi-gallocatechin gallate inhibits hepatitis B virus *in vitro*. Antiviral Research 78: 242-249 (2008).