How to Cite:
Almahdawi, M. A., Saleh, M. K., & Al-Waysi, S. A. A. (2022). Genetic variability of the S gene of chronic hepatitis B virus in Baghdad. *International Journal of Health Sciences*, *6*(S3), 3699–3714. https://doi.org/10.53730/ijhs.v6nS3.6157

Genetic variability of the S gene of chronic hepatitis B virus in Baghdad

Musaab A. Almahdawi
Department of Biology, College of Education for Pure Sciences, Tikrit University, Iraq

Mahmood K. Saleh
Department of Biology, College of Education for Pure Sciences, Tikrit University, Iraq

Safaa A. A. Al-Waysi
Director of Gastroenterology Hepatology Hospital Laboratory, Ministry of Health, Iraq

***Abstract*---**A total of 240 patients with chronic hepatitis B virus were included in the study, including 124 men and 116 females ranging in age from 6 to 78 years, PCR was used to amplify the Hepatitis B viral gene segment. There were only 26 positive samples for PCR amplification of a gene HBV S with strong positive bands submitted for sequencing. The NCBI BLASTn engine showed about 98% to 99% sequences similarities between the sequenced samples and the intended reference target sequences. By comparing the observed nucleic acid sequences of these investigated samples with the retrieved nucleic acid sequences (GenBank acc. MK840532.1), the details of its sequences were highlighted, and the total length of the amplified amplicons was also determined, the alignment results of the 548 bp samples revealed the presence of ten nucleic acid variations represented by ten nucleic acid substitutions in the analyzed samples in comparison with the most similar referring reference nucleic acid sequences (GenBank acc. no. MK840532.1), Our results indicated the presence of ten nucleic acid variants observed in the investigated samples, namely 163C>T, 184T>C, 210T>C, 217C>A, 274A>G, 310T>C, 336G>A, 346G>A, 372A>T, and 405A>C. Two variants of these identified substitutions (336G>A and 346G>A) were found in the majority of the investigated clinical samples, Amino acid alignment of these amino acid sequences with their references showed that six variants exhibited silent effects on the surface protein within the size of the amplified loci, These synonymous (silent) variants were detected in various samples and exemplified in the entire surface protein sequences. Meanwhile, four non-synonymous (missense) variants

were detected in the identified variants. The observed variations were respectively deposited in the NCBI-bank database under the accession numbers (OM972674 to OM972699) to represent S1, to S26 samples respectively, The investigated samples were clustered into seven phylogenetic clades within the Hepatitis B virus sequences. These genotypes are genotypes A, B, C, D, E, and F. Within the genotype D, all investigated samples (S1 to S26) were incorporated to constitute one large clade, clade-D that was made of 36 samples, our samples slightly deviated into two closely associated positions.

***Keywords*---**CHB, S gene, phylogenetic analysis, genotype.

Introduction

Two billion individuals worldwide have serological indicators of current or prior Hepatitis B virus (HBV) infection, with 257 million chronically infected. If HBV is not cleared after an acute infection, it might become dormant or active in the long run. (Munshi et al., 2017). 80% to 90% of newborns and toddlers under the age of one year, as well as 25–30 percent of children infected between the ages of 1 and 6 years, will get a chronic infection, in which Hepatitis B replicate in a liver for the rest of their lives (Lok AS,2002). The majority of infections in infants and children are asymptomatic, Adults have a 30% chance of getting symptomatic acute hepatitis B. Low, medium, high, and endemicity are the four categories of HBV prevalence described as (HBsAg) positive. (Chang,2000).

HBV is spread by direct contact with contaminated blood or sperm. It is usually transmitted perinatally from infected mothers to new borns in regions with high endemicity. In low-endemic areas, sexual transmission is the most widely used methodology of infection. People who have had a lot of sexual partners, males who have sexually with men, and those who have had other sexually transmitted infections are more like to get Aids.. (Trépo et al., 2014). HBV is generally categorized into 10 genotypes based on complete genome diversity of more than 8% at the nucleotides (nt) level (Sugauchi et al,2001) and phylogenetic studies of linked full genome sequences labelled (A–J) (Tatematsu et al.,2009). In the Mid-East, HBV is common, with a high prevalence of genotype D and a lack of hepatitis B envelope antigen (HBeAg) (Teriaky&Al-Judaibi,2013).

Large "L", medium "M", Small "S", are the three surfaces of protein that are associated with that makeup virus envelop and are found in a(4:1:1) ratio in whole viral particles. The S protein gene is found at the 3′ end of the open reading frame of the envelope gene, which is separated into three regions: (Pre-S1, Pre-S2, and S). Pre-S2 and S sections encode protein M, while Pre-S1, Pre-S2, and S encode the protein L. (Vierling & Manns,2003). HBV is diagnosed using a variety of serological markers, including HBsAg, HBcAg, HBeAg, HBsAb, HBV DNA, too. The presence of HBV DNA in the blood is a sign that HBV replication is active, assesses antiviral medication response, and detects resistance development (Teriaky&Al-Judaibi, 2013).

HBV DNA is a direct indicator of viral load, indicating how effectively virus replicate. Real-time PCR methods with a measurement device of 10–20 IU/mL and a broad detection range of up to 10 M are used in the majority of HBV DNA testing. With chronic HBV infection, serum HBV levels can range from detectable to more than 109 IU/L. (Trépo et al, 2014). To detect and quantify HBV DNA, two approaches are used: signals amplification, including such hybrids capturing and branching DNA techniques, and target amplification, such as polymerase chain reaction (PCR). (Datta et al, 2014). (Caliendo et al,2011). The lower range of viral load (10 to 15 IU/mL) and the highest viral load (107 to 108 IU/mL) can be detected using real-time PCR. As a result, in clinical settings, it has become the gold standard for identifying and measuring HBV DNA. It's also entirely automated and doesn't produce contamination that spreads (Bustin,2005).

HBV infection is a worldwide health problem , HBV infection is quite common in China (Ott JJ et al.,2017). HBV is an infection in the Hepadnaviridae species with a (3.2) kilobyte genomic of partly double-stranded circular DNA. PreS1, PreS2, and S, which are code by the genome's (PreS1, PreS2) , and S OPF , respectively, make up HBV surface proteins. The PreS1 and PreS2 proteins are involved in viral hepatocyte invasion, In the body, viruses assemble, replicate, and activate the immune system . In terms of facilitating HBV attachment to hepatocytes, The most important component of the protein is the PreS1 amino acid residues (2147). HBV can enter hepatocytes by pinocytosis because (PreS2) contains the PHSA receptors (polymerized human serum albumin) (Xiang et al., 2017). (Pollicino et al.,2012).

Immuno resistance with occult Hepatitis B infection has previously been associated with PreS/S gene variations (Zhang et al.,2016) (Chen et al., 2009) Serological identification of PreS/S gene variants shows that the presence of both HBV surface antigen and HBV surface antibody, and the absence of HBsAg. Previous research has discovered that high HBV reverse transcriptase spontaneous error rates have an influence on HBsAg expression, with viral variants constantly developing throughout infection and sensitive to both internal and external selection pressure (Cento et al.,2013)(Kim et al.,2015). HBsAg expression can be affected by PreS (PreS1 and PreS2) gene mutations (Wang et al., 2018).

Materials and Methods

This study was conducted for patients and patients diagnosed with HBV to the Medical City / Gastrointestinal Hospital / Dubai Laboratories and Sama Dubai Specialized Medical Center on Al Maghreb Street. A total of 240 patients with chronic hepatitis B virus were included in the study, including 124 men and 116 females ranging in age from 6 to 78 years. The QIA amp DNA Blood Mini Kit (250) Systems is used to extract viral DNA from serum samples. This system provides a quick and easy way to prepare pure DNA. A binding column in a micro centrifuge tube is used to purify samples, and (200 µl) of serum may be processed each purification. The DNA extracted is of excellent quality. Real-time PCR was used to assess viral load (Rotor-Gene Q, QIAGEN). RT-PCR reaction components and mixing amounts are indicated according to the manufacturer's instructions.

These primers were supplied in a lyophilized form. Lyophilized primers were dissolved in a nuclease free water to give a final concentration of 100 pmol/µl as a stock solution. A working solution of these primers was prepared by adding 10µl of primer stock solution (stored at freezer -20 C) to 90µl of nuclease free water to obtain a working primer solution of 10pmol/µl. PCR was used to amplify the Hepatitis B viral gene segment. The sense and antisense primers, as well as the PCR software for PCR, are listed in Tables (1) (2).

Table 1
S-gene Primer Sequence in HBV (Gandhe *et al.*, 2003)

Primer Name	Primer sequence	nt	Product length
Sense	5 -ACC CCT GCT CGT GTT ACA GGC-3	184–204	548
Anti-sense	5-AAA GCC AGA CAG TGG GGG AAA-3	711–731	

Table 2
Amplification of the S gene using PCR (Gandhe *et al.*, 2003)

Round	◦C	Time	Cycle
Initial Denaturation	95	5 min	1
Denaturation	95	30 sec	40
Annealing	55	30 sec	40
Extension	72	30 sec	40
Final extension	72	7 min	1
Hold	10	10 min	1

Polymerase chain reaction was used to amplify a portion of the S gene (PCR). In this study, Master Mix was prepared in a total volume of 25 microliters, and (12.5) microliters of GoTaq Green Master Mix were added to every sample in a PCR tube as per manufacturer instructions, followed by (1) microliter (10 pmol) of forward primer and (1) microliter of reverse primer, the DNA template (9) microliters, and (1.5) microliters of ddH2O, that is (16) microliters of master mix per tube and add (9) microliters of template. After PCR amplification, agarose gel electrophoresis was adopted to confirm the presence of amplification. PCR was completely dependable on the extracted DNA criteria.

The PCR product was sent to Macrogen Corporation – Korea to Sanger sequencing using ABI3730XL, automated DNA Sequencer The results were received by email then analyzed using Genious software. The sequencing results of the PCR products of the targeted samples were edited, aligned, and analyzed as long as with the respective sequences in the reference database using Bio Edit Sequence Alignment Editor Software Version 7.1 (DNASTAR, Madison, WI, USA). The observed variations in each sequenced sample were numbered in PCR amplicons as well as in its corresponding position within the referring genome. The observed nucleic acids were numbered in PCR amplicons as well as in their corresponding positions within the referring genome. Each detected variant within the viral sequences was annotated by SnapGene Viewer ver. 4.0.4 (https://www.snapg

ene.com). The observed variations were respectively deposited in the NCBI-bank kit database under a unique accession number for each analyzed sample.

Results and Discussion

After HBV-DNA extraction followed by purification of PCR products, the genomes including the S gene were amplified using its primer (product length 548 bp) , the amplification results of this gene are shown in Figure (1) after it was migrated by the electrophoresis, and there were only 26 positive samples for PCR amplification of a gene HBV S with strong positive bands submitted for sequencing.

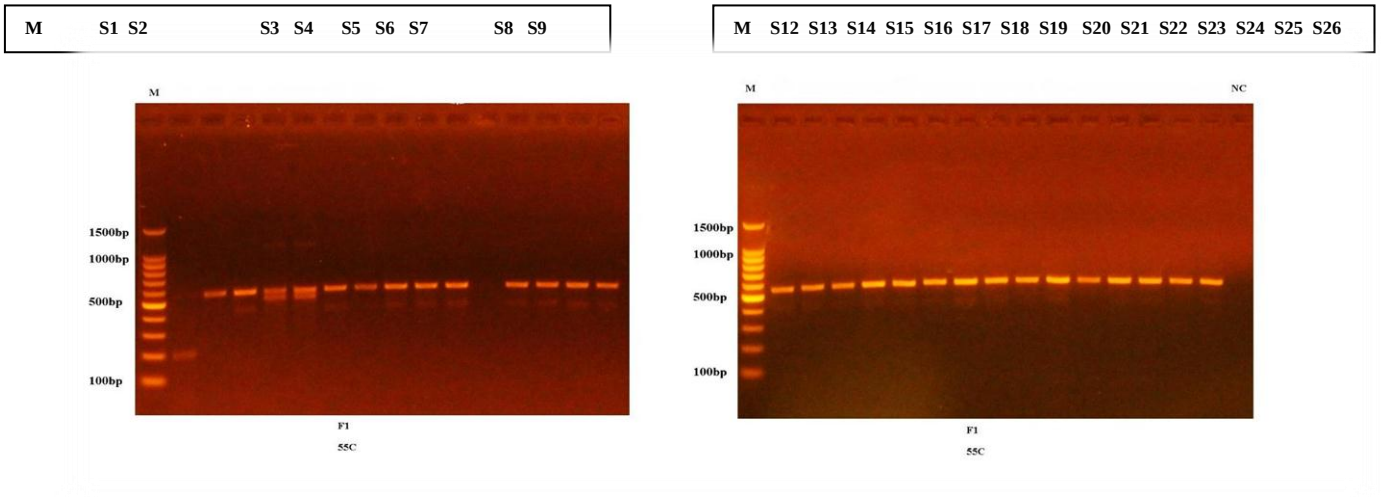

Fig 1. The amplification results of the Hepatitis B primers fractionated on 1.5% agarose gel electrophoresis stained with Eth.Br. M:100bp ladder marker

Twenty-six samples were included in the present study. These samples were screened to amplify S gene sequences of the Hepatitis B virus. Thus, the variation of the S gene can be used for Hepatitis B virus genotyping due to its possible ability to adapt to variable genetic diversity as it was seen in different viral types. The sequencing reactions indicated the exact identity after performing NCBI blastn for these PCR amplicons (Zhang *et al.*, 2000). Concerning the 548 bp amplicons, the NCBI BLASTn engine showed about 98% to 99% sequences similarities between the sequenced samples and the intended reference target sequences. By comparing the observed nucleic acid sequences of these investigated samples with the retrieved nucleic acid sequences (GenBank acc. MK840532.1), the accurate positions and other details of the retrieved PCR fragments were identified. The total length of the targeted locus was localized in the NCBI server, and the positions of the start and end of the targeted locus were also confirmed within the most homologous viral target, Fig (2).

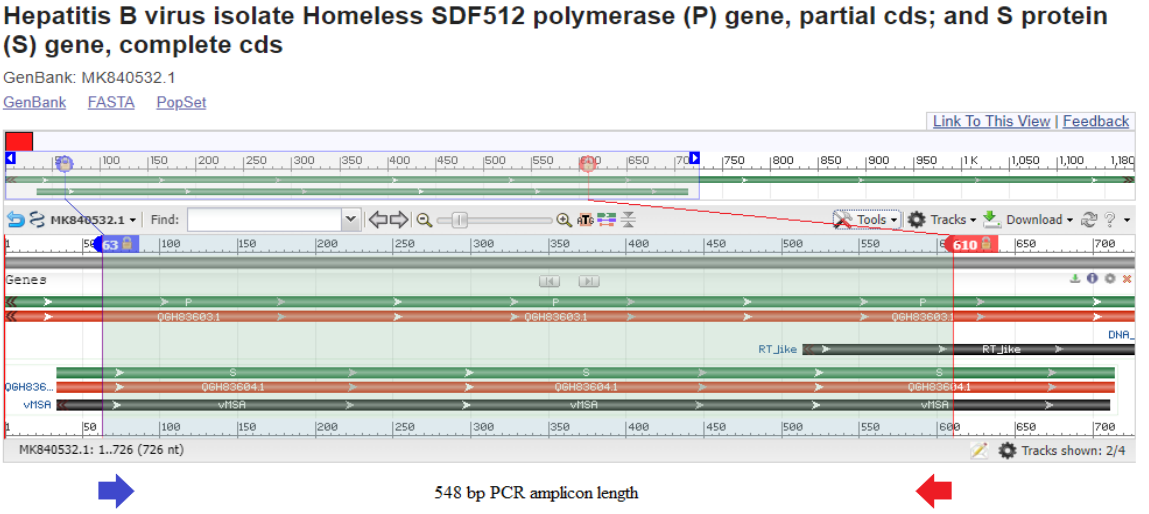

Fig 2. The exact position of the retrieved 548 bp amplicon partially covered the majority of the coding portions of the S gene within Hepatitis B virus genomic sequences (GenBank acc. no. MK840532.1). The blue arrow refers to the starting point of this amplicon while the red arrow refers to its endpoint

After positioning the 548 bp amplicons' sequences within the genomic sequences of the Hepatitis B virus, the details of its sequences were highlighted, and the total length of the amplified amplicons was also determined, Table (3).

Table 3
The position and length of the 548 bp PCR amplicons that used to amplify the majority of the coding portions of the S gene within Hepatitis B virus genomic sequences (GenBank acc. no. MK840532.1)

Amplicon	Reference locus sequences (5′ - 3′)	length
S gene Nucleic acid sequences of human-infecting viral S gene	*ACCCCTGCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAAT CCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAA TTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTCGCAGTC CCCAACCTCCAATCACTCACCAACCTCCTGTCCTCCAACTTGTCCT GGTTATCGCTGGATGTGTCTGCGGCGTTTTATCATCTTCCTCTTCA TCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCA AGGTATGTTGCCCGTTTGTCCTCTAATTCCAGGATCTTCAACCACC AGCACGGGACCATGCAGAACCTGCACGACTCCTGCTCAAGGAAC CTCTATGTATCCCTCCTGTTGCTGTACCAAACCTTCGGACGGAAAT TGCACCTGTATTCCCATCCCATCATCCTGGGCTTTCGGAAAATTCC TATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAG TGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGC TTT**	548 bp

*refers to forward primer placed in the forward direction
**refers to reverse primer placed in the reverse complement direction

Interestingly, the alignment results of the 548 bp samples revealed the presence of ten nucleic acid variations represented by ten nucleic acid substitutions in the analyzed samples in comparison with the most similar referring reference nucleic acid sequences (GenBank acc. no. MK840532.1) , Fig (3).

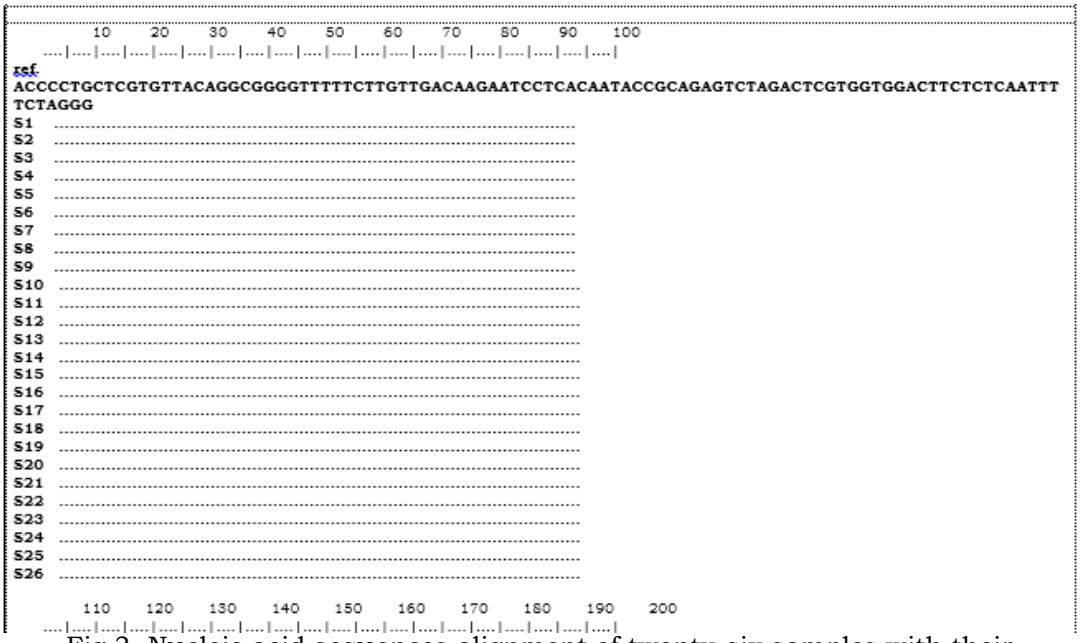

Fig 3. Nucleic acid sequences alignment of twenty-six samples with their corresponding reference sequences of the 548 bp amplicons of the S genetic sequences. The symbol “ref” refers to the NCBI referring sequence (GenBank acc. no. MK840532.1), letters “S”, followed by a number refers to the sample number

Our results indicated the presence of ten nucleic acid variants observed in the investigated samples, namely 163C>T, 184T>C, 210T>C, 217C>A, 274A>G, 310T>C, 336G>A, 346G>A, 372A>T, and 405A>C. Two variants of these identified substitutions (336G>A and 346G>A) were found in the majority of the investigated clinical samples. Whereas the other variations were only detected in a variable intensity in the investigated sample. These differences observed in the currently observed nucleic acid sequences in the analyzed samples were not found in the corresponding reference sequences (GenBank acc. no. MK840532.1), which refers to the presence of a possible novelty of such variations. To confirm these extremely high number of variations, the sequencing chromatograms of the investigated samples, as well as their detailed annotations, were verified and documented, and the chromatograms of their sequences were shown according to their positions in the PCR amplicons , Fig (4).

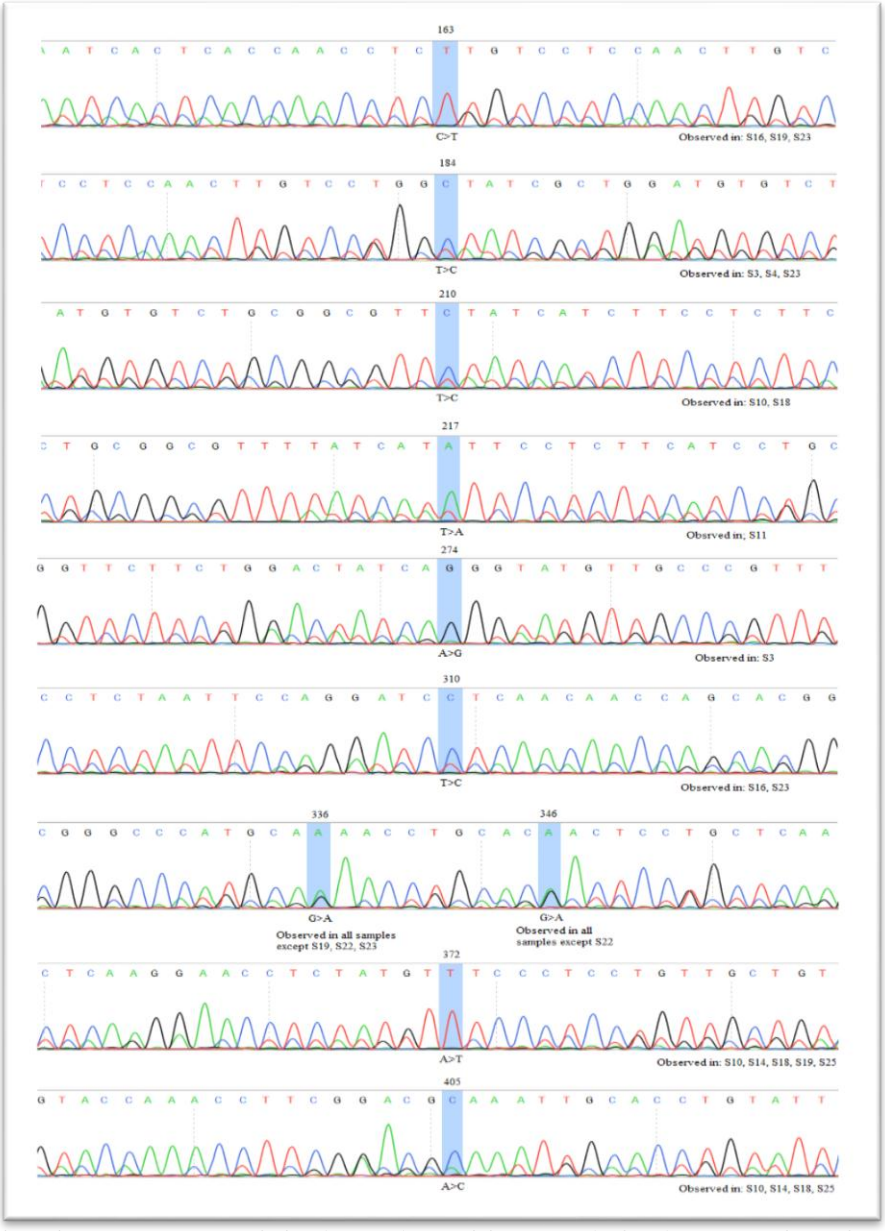

Fig 4. The chromatogram of the investigated human-infecting samples of Hepatitis B virus. The letter "S" refers to the code of the investigated samples in this study. The identified variations were respectively arranged according to the names of the investigated samples (S1 to S26). The cyan highlighted regions represent the site of the observed variation

Nucleic acid variations

The observed nucleic acid variations were further analyzed to identify whether such substitutions induce possible alteration in their corresponding positions in the viral S protein. All nucleic acid sequences of S1 to S26 were translated to their

corresponding amino acid sequences using the Expasy translate suite. Amino acid alignment of these amino acid sequences with their references showed that six variants exhibited silent effects on the surface protein within the size of the amplified loci , Fig(5a). These synonymous (silent) variants were detected in various samples and exemplified in the entire surface protein sequences, namely p.65C=, p.71G=, p.82I=, p.101Q=, p.113S=, and p.125T=. Meanwhile, four non synonymous (missense) variants were detected in the identified variants. These variants are p.80F>S, p.122R>K, p.134Y>F and p.145G>A , Fig(5b).

A) sequences of amino acid residues within PCR amplicons

```
          10        20        30        40        50        60        70        80        90       100
    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ref.
PLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNHSPTSCPPTCPG
YRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLI
S1  ..................................................................................................
S2  ..................................................................................................
S3  ..................................................................................................
S4  ..................................................................................................
S5  ..................................................................................................
S6  ..................................................................................................
S7  ..................................................................................................
S8  ..................................................................................................
S9  ..................................................................................................
S10 ..........................................................S.....................................
S11 ..................................................................................................
S12 ..................................................................................................
S13 ..................................................................................................
S14 ..................................................................................................
S15 ..................................................................................................
S16 ..................................................................................................
S17 ..................................................................................................
S18 ..........................................................S.....................................
S19 ..................................................................................................
S20 ..................................................................................................
S21 ..................................................................................................
S22 ..................................................................................................
S23 ..................................................................................................
S24 ..................................................................................................
S25 ..................................................................................................
S26 ..................................................................................................

          110       120       130       140       150       160       170       180
    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|..
ref.
PGSSTTSTGPCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASAR
FSWLSLLVPFVQWFVGLSPTVWL
S1  ..........K.....................................................................
S2  ..........K.....................................................................
```

S3 ……K……
S4 ……K……
S5 ……K……
S6 ……K……
S7 ……K……
S8 ……K……
S9 ……K……
S10 ……K……F……A……
S11 ……K……
S12 ……K……
S13 ……K……A……
S14 ……K……F……
S15 ……K……
S16 ……K……
S17 ……K……A……
S18 ……K……F……
S19 ……F……
S20 ……K……
S21 ……K……
S22 ……
S23 ……
S24 ……K……
S25 ……K……F……A……
S26 ……K……

B) sequences of amino acid residues within the entire protein

The entire structure of the surface protein [Hepatitis B virus] (GenBank QGH83604.1)

MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPTSNH
SPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRT
CTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSWLSLLVPFVQW
FVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI

Fig 5. (a,b)Amino acid residues alignment of the detected variations of the S-encoded surface protein within the investigated human-infecting Hepatitis B virus samples

A) The amino acid substitutions are highlighted according to their corresponding positions within the amplified 548 bp locus. B) The amino acid substitutions are highlighted according to their corresponding positions within the entire protein. The grey highlights refer to the amplified region of the S-encoded protein. The cyan and yellow colors refer respectively to the silent and missense amino acid substitution in the alignment chart.

However, these amino acid alterations may be developed by the invading viruses as an adaptation to drugs that are directed toward their targeted surface protein (Legnardi et al., 2020). The observed variations were respectively deposited in the NCBI-bank it database under the accession numbers (OM972674 to OM972699) to represent S1, to S26 samples respectively , To summarize all the results obtained from the sequenced 548 bp fragments, the exact positions of the observed mutation were described in Table (4).

Table 4
The pattern of the observed mutation in the 548 bp of the S gene amplicons in comparison with the NCBI referring sequences (GenBank acc. MK840532.1). The identified variations were respectively arranged according to the names of the investigated samples (S1 to S26). The letter "p." refer to the "protein" position in which the variation was detected

No.	Sample	Native	Allele	Position in the PCR fragment	Position in the reference amino acid sequences	Type of variant	Variant summary
1.	S16, S19, S23	C	T	163	65Cys	Silent	p.65C=
2.	S3, S4, S23	T	C	184	71Gly	Silent	p.71G=
3.	S10, S18	T	C	210	80Phe	Missense	p.80F>S
4.	S11	C	A	217	82Ile	Silent	p.82I=
5.	S3	A	G	274	101Gln	Silent	p.101Q=
6.	S16, S23	T	C	310	113Ser	Silent	p.113S=
7.	S19, S22, S23	G	A	336	122Arg	Missense	p.122R>K
8.	S1-S18, S20-S21, S24-S26	G	A	346	125Thr	Silent	p.125T=
9.	S10, S14, S18, S19, S25	A	T	372	134Tyr	Missense	p.134Y>F
10.	S10, S14, S18, S25	A	C	405	145Gly	Missense	p.145G>A

Phylogenetic tree

To give a phylogenetic understanding of the actual distances between our investigated samples and the most relative reference strains of Hepatitis B virus, a comprehensive phylogenetic tree was generated in the present study according to nucleic acid variations observed in the amplified 548 bp of the S gene amplicons. This phylogenetic tree was contained S1 to S26 samples alongside other relative nucleic acid sequences of Hepatitis B virus sequences. Within this tree, our investigated samples were incorporated alongside relative sequences to constitute the majority of the incorporated sequences within the cladogram. The total number of the aligned nucleic acid sequences in this comprehensive tree was 91. In the constructed cladogram, the investigated samples were clustered into seven phylogenetic clades within the Hepatitis B virus sequences. The most interesting fact observed in our investigated viral isolates is correlated with the positioning of our investigated samples and their neighbour sequences into these

related important genotypes within the Hepatitis B virus. These genotypes are genotype A, B, C, D, E, and F. Within the genotype D, all investigated samples (S1 to S26) were incorporated to constitute one large clade, clade-D that was made of 36 samples (Fig 6). Within this major clade, our samples slightly deviated into two closely associated positions. This sort of slight diversity was reflected by slight potential evolutionary effects of the observed nucleic acid substitutions in inducing such tilts in the incorporated tree. However, these nucleic acid alterations were only minor deviations within the same clade of genotype-D. Furthermore, the aggregation of all investigated viral samples with each other may refer to the presence of only close patterns of the phylogenetic distribution of these sequences. The current observation of this tree has confirmed the sequencing reactions since it explained the actual neighbour-joining-based positioning in these observed variations. Despite the narrow biological diversity observed in our investigated samples with respect to the clade-D, highly distinct phylogenetic distances were observed between the samples of genotype-D and the nearest genotype of Hepatitis B viruses. Thus, the currently observed nucleic acid variations were only a minor tilt within the same viral genotype without taking and noticeable evolutionary effect in altering the current positioning of the investigated S1 – S26 samples. This finding was strongly suggested that these (S1 – S26) samples may represent known genetic variations of the genotype-D within the Hepatitis B virus sequences.

Next to the genotype-D, HBV samples of genotype-F were suited. This observation indicated the HBV genotype-F represents the closest genotype to the samples of genotype-D based on the S gene-based sequencing. Besides the genotype-F, the genotype-E was suited with distinct phylogenetic positions. In the vicinity of the genotype-E, the genotype-C, genotype-B, genotype-G, and genotype-A were respectively located in distinct phylogenetic distances. However, the S gene sequences showed closer phylogenetic positions between these five genotypes (genotype-E, genotype-C, genotype-B, genotype-G, and genotype-A) than that found between the other incorporated genotypes (genotype-D and genotype-F). This tree showed that all the incorporated clades were represented by ten reference sequences, with one exception of genotype-G. Within this genotype, a lower number of HBV sequences were incorporated due to the scarce sequences of the S gene-based genotype-D compared with the other incorporated genotype.

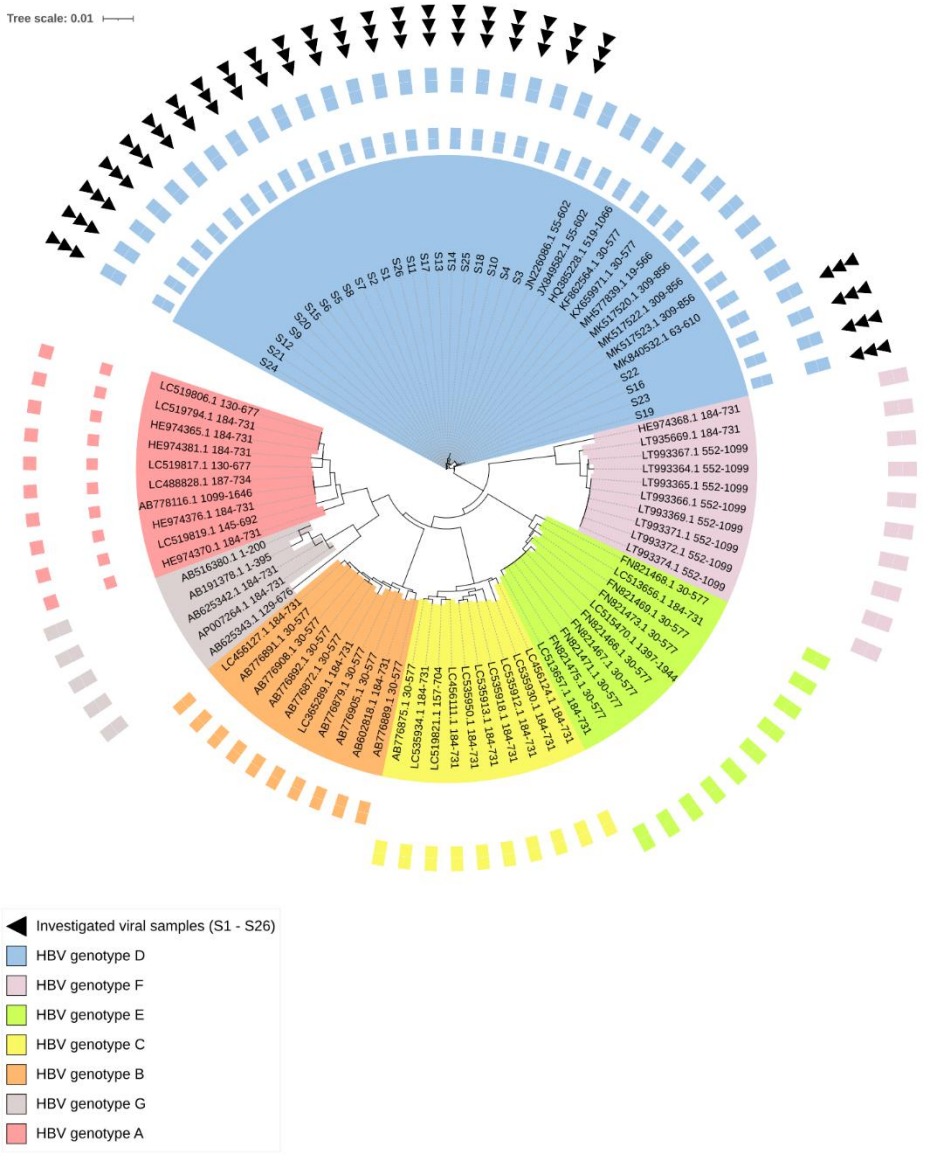

Fig 6. The comprehensive cladogram phylogenetic tree of genetic variants of the S gene fragment of ten human-infecting Hepatitis B viral samples. The black-colored triangle refers to the analyzed viral variants. All the mentioned numbers referred to GenBank accession number of each referring species. The number "0.01" at the top portion of the tree refers to the degree of scale range among the comprehensive tree categorized organisms. The letter "S#" refers to the code of the investigated samples.

Noteworthy, the utilization of the S gene sequences in this study has given further indication for the presence of the precise identification of the actual genotype of this viral organism. It was found that the investigated viral samples were occupied distinct positions within the major clade-D of the cladogram with few phylogenetic

distances among them. This indicated the presence of high genetic homology among the viral sequences of these strains within the major clade. The other related clades, which consist of genotypes F, E, C, B, G, and A were found to occupy vast phylogenetic positions away from major clade-D sequences. This observation showed the presence of obvious phylogenetic distances among these genotypes and the genotype of our investigated samples. This data may have shown no potential phylogenetic effect of pathological status on inducing any noticeable nucleic acid alteration in the investigated viral strains. However, this S gene-based comprehensive tree has provided an inclusive tool about the high ability of such genetic fragments to efficiently identify viral genotypes using this genetic fragment. This, in turn, gives a further indication of the ability of the currently utilized S gene-specific primers to describe the investigated Hepatitis B viruses and their accurate phylogenetic positions. Thus, according to the observed nucleic acid substitutions, as shown in the constructed phylogenetic tree, all these samples were positioned within variable closely associated phylogenetic positions. These observations suggested no potential participation of the detected mutations in inducing remarkable evolutionary alterations of the currently investigated viral samples in the Hepatitis B virus sequences. However, the high ability of such genetic fragments to efficiently identify viral sequences was proved by using these S gene-based fragments. This, in turn, gives a further indication of the ability of the currently utilized S gene-specific primers to describe the currently investigated Hepatitis B viruses and their accurate phylogenetic positions.

Conclusion

Through the current study, it was found that the genotype S is the type that causes human infection within the hepatitis B virus, and it is considered good from a diagnostic point of view as well as distinguishing polymorphisms within the species, and that the genotype D is predominant in patients with hepatitis B type in Iraq.

References

Bustin SA, Benes V, Nolan T, et al. Quantitative real-time RT-PCR--a perspective. J Mol Endocrinol 2005;34:597-601.

Caliendo AM, Valsamakis A, Bremer JW, et al. Multilaboratory evaluation of real-time PCR tests for hepatitis B virus DNA quantification. J Clin Microbiol 2011;49:2854-8.

Cento V, Van Hemert F, Neumann–Fraune M, Mirabelli C, Di Maio VC, Salpini R, Bertoli A, Micheli V, Gubertini G, Romano S, *et al*: Anti–HBV treatment induces novel reverse transcriptase mutations with reflective effect on HBV S antigen. J Infect 67: 303–312, 2013.

Chang MH. Natural history of hepatitis B virus infection in children. J Gastroenterol Hepatol. 2000;15(Suppl):E16–9.

Chen CH, Changchien CS, Lee CM, Tung WC, Hung CH, Hu TH, Wang JH, Wang JC and Lu SN: A study on sequence variations in pre–S/surface, X and enhancer II/core promoter/precore regions of occult hepatitis B virus in

non–B, non–C hepatocellular carcinoma patients in Taiwan. Int J Cancer 125: 621–629, 2009.

Datta S, Chatterjee S, Veer V. Recent advances in molecular diagnostics of hepatitis B virus. World J Gastroenterol 2014;20:14615-25.

Gandhe, S. S., Chadha, M. S., & Arankalle, V. A. (2003). Hepatitis B virus genotypes and serotypes in western India: lack of clinical significance. *Journal of medical virology*, *69*(3), 324-330.

Kim H, Lee SA, Won YS, Lee H and Kim BJ: Occult infection related hepatitis B surface antigen variants showing lowered secretion capacity. World J Gastroenterol 21: 1794–1803, 2015.

Legnardi M, Tucciarone CM, Franzo G, Cecchinato M. Hepatitis B Virus Evolution, Diagnosis and Control. Vet Sci. 2020 Jun 22;7(2):79. doi: 10.3390/vetsci7020079.

Letunic I, Bork, P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 2; 47(W1): W256-W259.

Lok AS. Chronic hepatitis B. N Engl J Med. 2002;346(22):1682–3.

Munshi, S. U., Tran, T. T. T., Vo, T. N. T., Tabassum, S., Sultana, N., Nguyen, T. H., ... & Rahman, M. (2017). Molecular characterization of hepatitis B virus in Bangladesh reveals a highly recombinant population. PloS one, 12(12), e0188944.

Ott JJ, Horn J, Krause G and Mikolajczyk RT: Time trends of chronic HBV infection over prior decades–A global analysis. J Hepatol 66: 48–54, 2017.

Pollicino T, Amaddeo G, Restuccia A, Raffa G, Alibrandi A, Cutroneo G, Favaloro A, Maimone S, Squadrito G and Raimondo G: Impact of hepatitis B virus (HBV) preS/S genomic variability on HBV surface antigen and HBV DNA serum levels. Hepatology 56: 434–443, 2012.

Sarhan, R. S, Hashim, H.O., Al-Shuhaib, M.B.S. 2019. The Gly152Val mutation possibly confers resistance to beta-lactam antibiotics in ovine Staphylococcus aureus isolates. Open Veterinary Journal 9(4): 339-348. https://doi.org/10.4314/ovj.v9i4.12.

Sugauchi F, Mizokami M, Orito E, et al. A novel variant genotype C of hepatitisB virus identified in isolates from Australian Aborigines: complete genomesequence and phylogenetic relatedness. J Gen Virol 2001;82:883–92.

Tatematsu K, Tanaka Y, Kurbanov F, et al. A genetic variant of hepatitisB virus divergent from known human and ape genotypes isolated froma Japanese patient and provisionally assigned to new genotype J. J Virol 2009;83:10538–47.

Teriaky, A., & Al-Judaibi, B. (2013). Correlation between HBsAg quantitation and HBV DNA in HBeAg-Negative HBV/D patients. Saudi journal of gastroenterology: official journal of the Saudi Gastroenterology Association, 19(6), 243.

Trépo, C., Chan, H. L., & Lok, A. (2014). Hepatitis B virus infection. The Lancet, 384(9959), 2053-2063.

Vierling, J. M., & Manns, M. P. (2003). Liver immunology. M. E. Gershwin (Ed.). Philadelphia, Pa, USA: Hanley & Belfus.

Wang, T., Dai, Y., Zhang, M., Cui, D., Xu, X., Sun, C., & Cheng, J. (2018). Sequence analysis of the Pre-S gene in chronic asymptomatic HBV carriers

with low-level HBsAg. International journal of molecular medicine, 42(5), 2689-2699.

Xiang KH, Michailidis E, Ding H, Peng YQ, Su MZ, Li Y, Liu XE, Dao Thi VL, Wu XF, Schneider WM, *et al*: Effects of amino acid substitutions in hepatitis B virus surface protein on virion secretion, antigenicity, HBsAg and viral DNA. J Hepatol 66: 288–296, 2017.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T (2012) Primer-BLAST: A tool to design target-specific primers for surface protein chain reaction. BMC Bioinformatics 13:134.

Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. J Comput Biol. 7(1-2):203-14.

Zhang ZH, Wu CC , Chen XW, Li X, Li J and Lu MJ: Genetic variation of hepatitis B virus and its significance for pathogenesis. World J Gastroenterol 22: 126–144, 2016.