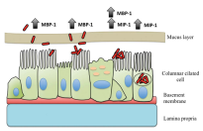
Virulence

Virulence



ISSN: 2150-5594 (Print) 2150-5608 (Online) Journal homepage: http://www.tandfonline.com/loi/kvir20

Genotype-specific acquisition, evolution and adaptation of characteristic mutations in hepatitis E virus

Aqsa Ikram, Mohamad S. Hakim, Jian-hua Zhou, Wenshi Wang, Maikel P. Peppelenbosch & Qiuwei Pan

To cite this article: Aqsa Ikram, Mohamad S. Hakim, Jian-hua Zhou, Wenshi Wang, Maikel P. Peppelenbosch & Qiuwei Pan (2017): Genotype-specific acquisition, evolution and adaptation of characteristic mutations in hepatitis E virus, Virulence, DOI: 10.1080/21505594.2017.1358349

To link to this article: http://dx.doi.org/10.1080/21505594.2017.1358349

VIRULENCE
2017, VOL. 0, NO. 0, 1–12
https://doi.org/10.1080/21505594.2017.1358349

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

Genotype-specific acquisition, evolution and adaptation of characteristic mutations in hepatitis E virus

Aqsa Ikram[a,b], Mohamad S. Hakim[a,c], Jian-hua Zhou[a,d], Wenshi Wang[a], Maikel P. Peppelenbosch[a], and Qiuwei Pan[a]

[a]Department of Gastroenterology and Hepatology, Erasmus MC-University Medical Center, Rotterdam, the Netherlands; [b]Atta Ur Rahman School of Applied Biosciences, National University of Science and Technology, Islamabad, Pakistan; [c]Department of Microbiology, Faculty of Medicine, Universitas Gadjah Mada, Yogyakarta, Indonesia; [d]State Key Laboratory of Veterinary Etiological Biology, National Foot-and-Mouth Disease Reference Laboratory, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, Gansu, P.R. China

ABSTRACT

Hepatitis E virus (HEV) infection is a major cause of acute hepatitis but also provokes chronic infection in immunocompromised patients. Although the pathogenesis and treatment outcome involve complex interplay between the virus and host, the nature of adaptive responses of HEV to the host immune system remain obscure at best. In this study, we used large-scale proteomic bioinformatics to profile characteristic mutations in human HEV isolates associated to ribavirin treatment failure, chronic hepatitis, hepatic failure or altered immunoreactivity. The prevalence of specific mutations was examined in a large number of protein sequences of ORF1 and ORF2 regions of the 3 major human-derived HEV genotypes (1, 3 and 4). By analyzing potential B, CD4+ and CD8+ T cell epitopes, we found that many of these mutations overlap with the predicted epitopes and are frequently present among the 3 HEV genotypes. These overlapping mutations mediate reduced antigenicity. Finally, by delineation of diversification and evolution of the underlying epitopes, we observe that most of these variants apparently evolved earlier in genotype 1 when compared with genotypes 3 and 4. These results indicate that HEV is under substantial evolutionary pressure to develop mutations enabling evasion of the host immune response and resistance to antiviral treatment. This indicates the existence of an ongoing evolutionary arms race between human immunity, antiviral medication and HEV.

ARTICLE HISTORY
Received 3 April 2017
Revised 4 July 2017
Accepted 17 July 2017

KEYWORDS
B and T cells; epitope; evolution; hepatitis E virus; mutation

Introduction

Hepatitis E virus (HEV) is a non-enveloped, single-stranded positive sense RNA virus which mainly infects the liver.[1] It causes over 3 million acute cases and 57,000 deaths every year.[2] Although 8 HEV genotypes are now recognized, there are 4 well-defined genotypes infecting humans, including genotype 1, 2, 3 and 4. Genotypes 1 and 2 are found only in humans and are responsible for most cases of infection in the developing countries. Genotypes 3 and 4 circulate in several animals (e.g. pigs, wild boars and deer) and are the main cause of sporadic infection in the developed countries.[3,4] Although no FDA-approved treatment is available, pegylated interferon-α (PEG-IFN-α) or ribavirin has been used as off-label treatment of some cases of HEV infection.[5,6]

As an RNA virus, HEV possesses a high mutation rate. It was indirectly estimated from clinical isolates that the mutation rate of HEV was ~1.5 base substitutions per site per year, which is quite similar to that reported for hepatitis C virus (HCV).[7] The viral RNA-dependent RNA polymerase (RdRp), which lacks the proof-reading capacity, is an important factor contributing to the high rate of mutations in the HEV genome.[8] Furthermore, the selection pressure imposed by the host immune responses may also contribute to the variability of HEV genome. Recently, studies have hinted at the acquisition by the HEV genome of certain mutations associated with ribavirin treatment failure, chronic hepatitis, hepatic failure and reduced immunoreactivity.[9] The observation that different mutations change the HEV proteome with respect to not only the therapeutic response but also the

CONTACT Qiuwei Pan, PhD q.pan@erasmusmc.nl Department of Gastroenterology and Hepatology, Erasmus MC-University Medical Center Room Na-617, 's-Gravendijkwal 230 NL-3015 CE Rotterdam, the Netherlands.
Supplemental data for this article can be accessed on the publisher's website.

immunoreactivity highlights the importance of studying HEV mutations and their effects on the host immune responses.

Only limited knowledge is available on the contribution of HEV genome variants toward susceptibility, pathogenesis and therapeutic responses. To elucidate these processes, we have used a large-scale proteomic bioinformatics to profile human HEV characteristic mutations. In the present study, we have comprehensively investigated the proteomic variation of the open reading frame 1 (ORF1) and ORF2 regions among the major HEV genotypes 1, 3 and 4, by retrieving a large data set of HEV sequences. We first mapped the presence and abundance of the mutations related to ribavirin treatment failure, chronic hepatitis, hepatic failure or reduced immunoreactivity. Furthermore, we examined the overlap of these mutations with predicted B cell, $CD4^+$ and $CD8^+$ T cell epitopes, and assessed their antigenicity. Interestingly, we have observed that these overlapped mutations evolved earlier in genotype 1 when compared with genotype 3 and 4. Thus, the acquisition and evolution of these characteristic mutations may help the virus to evade host immune response, develop resistance to antiviral treatment, and facilitate its adaptation in human population.

Results

Identification of HEV characteristic mutations

Several HEV variants were previously reported both *in vitro* and *in vivo*. Using a combination of phrases/keywords in PubMed (Table S1), we explored clinical and *in vitro* data reporting mutations within the ORF1 and ORF2 of HEV. In our analysis, 20 mutations (Table S2) were identified in these regions, 17 in ORF1 and 3 in ORF2 (Fig. 1). Among these mutations, one was related to chronic hepatitis, 9 with hepatic failure, 8 with ribavirin treatment failure, and 2 with altered immunoreactivity (Table S2).

Frequency of mutations within ORF1 and ORF2

To evaluate the relevance of these reported characteristic mutations, we first evaluated their presence in the circulating strains based on the HEV sequences deposited in the GenBank. We have searched the 4 main genotypes identified from the human host. However, limited sequences are available for genotype 2. Thus, we retrieved 57 ORF1, 51 ORF2 (genotype 1), 131 ORF1, 131 ORF2 (genotype 3) and 99 ORF1, 96 ORF2

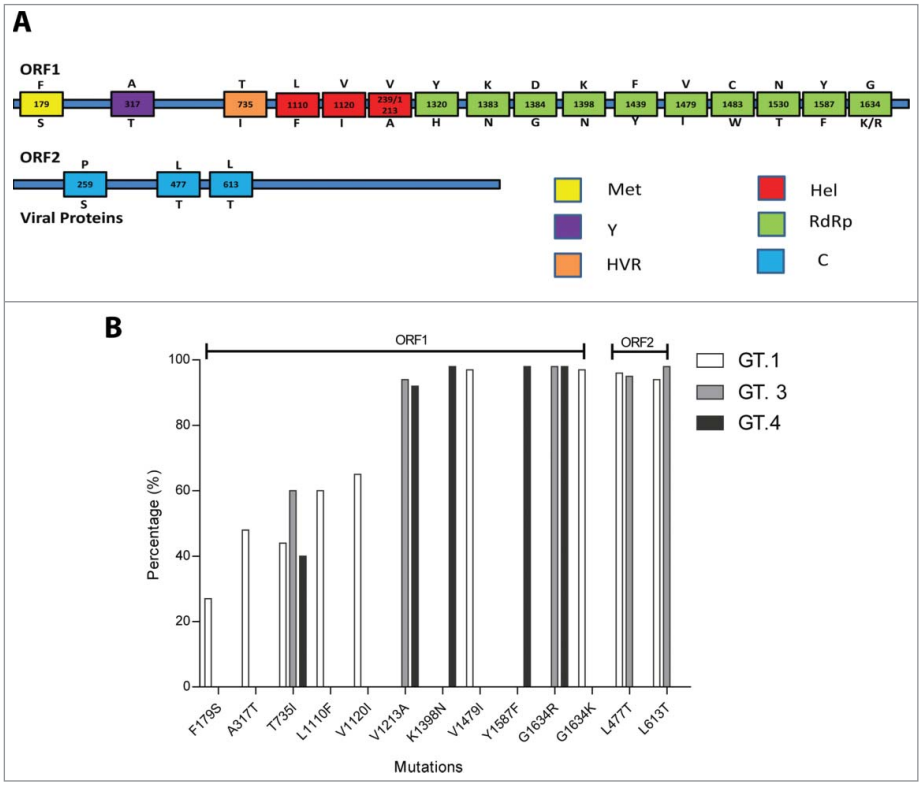


Figure 1. (A). Identified mutations of ORF1 and ORF2 regions of HEV. The number within the box represents the amino acid position; the letter(s) above the box refer to the wild type amino acid, and the letter below the box are relevant mutations reported in previous studies. (Met: methyltransferase; Y: Y-domain; HVR: hypervariable regions; Hel: RNA helicase; RdRP: RNA-dependent RNA polymerase; C: capsid protein) (B). The prevalence of mutations within HEV genotype 1 (n = 81), 3 (n = 182) and 4 (n = 143). Amino acid diversity was measured as the proportion of sequences that varies from the consensus sequence.

(genotype 4) of HEV sequences (retrieved in January 2017). The selected sequences represent all major genotypes (1, 3 and 4). The full-length sequences used in this study based on each ORF and genotype are provided in Table S3.

After removal of closely related and redundant sequences, 406 full length ORF1 and ORF2 sequences were finally selected for further analysis. Table 1, Table S2 and Fig. 1 show the number of all possible and experimentally confirmed amino acid variants and frequency of their variation for each HEV genotype. The number of variants found in ORF1 region was higher (17 variants) than that of ORF2 region (3 variants). Out of 17 variants in ORF1, 8 were related to hepatic failure; one with chronic hepatitis; and 8 with ribavirin treatment failure. While in ORF2, one variation was associated with hepatic failure and 2 with immunoreactivity. When analyzing HEV genotypes, it was revealed that genotype 1 possesses a higher number of mutations (8) with considerable frequency when compared with genotype 3 and 4 which possesses only 6 and 3 mutations, respectively (Table 1 and Fig. 1B). This observation further suggests the high level of polymorphism in HEV genotype 1.

Mutations within genotype 1, 3 and 4

In genotype 1, 9 out of 17 reported mutations (V1213A, Y1320H, K1383N, D1384G, K1398N, C1483W, N1530T, Y1587F, G1634R) in ORF1 and one (P259S) out of 3 in ORF2, were found to be conserved (i.e. not present in our analyzed sequences), while others have shown considerable variations (Table 1). Among these variable mutations, T735I and G1634R/K were observed to be frequent among all selected genotype (1, 3 and 4). Another mutation (A317T) was found in 2 genotypes (1 and 3). Four mutations (ORF1 = 2; ORF2 = 2) reached the frequency of > 90% (Table 1, Fig. 1B). Among these, 2 mutations were related to ribavirin treatment failure [G1634K/R (ORF1; genotype 1, 3, 4) and Y1587F (ORF1; genotype 4)] and 2 were associated with altered immunoreactivity [L477T (ORF1; genotype 1, 3) and L613T (ORF1; genotype 1, 3)].

In genotype 3, we found that the reported mutation sites [13 in ORF1 (F179S, L1110F, V1120I, Y1320H, K1383N, D1384G, K1398N, F1439Y, V1479I, C1483W, N1530T, Y1587F, G1634K) and one (P259S) in ORF2] were conserved (i.e., not present in our analyzed sequences) (Table 1, Fig. 1B). Mutations of ORF1, i.e., V1213A and G1634R, were found with a considerable frequency in genotypes 3 and 4. The most important variants were A317T and V1213A which showed a high prevalence in genotype 3. Four mutations approached a frequency of > 90%, where one was related to chronic hepatitis (V1213A), one with ribavirin treatment failure (G1634R) and 2 with immunoreactivity (L477T, L613T) (Table 1, Fig. 1B).

In genotype 4, 13 sites in ORF1 (F179S, A317T, L1110F, V1120I, V1213A, Y1320H, K1383N, K1398N, D1384G, F1439Y, V1479I, C1483W, N1530T) and 3 in ORF2 (P259S, L477T, L613T) were found to be conserved. Mutation Y1587F was frequent in genotype 4. Two mutations reached the frequency of > 90% (Y1587F, G1634R) (Table 1, Fig. 1B). Both mutations have been reported to be related to ribavirin treatment failure (Y1587F, G1634R).

Table 1. All possible and experimentally confirmed (*in vivo* and *in vitro*) mutations and their prevalence in the major HEV genotypes (1, 3, 4). Different colors indicate the mutations related to ribavirin treatment failure (yellow), chronic hepatitis (purple), hepatic failure (red) or altered immunoreactivity (blue).

		Percentage		
	Mutations	GT1 ORF1 (57)	GT3 ORF1 (131)	GT4 ORF1 (99)
ORF1	F179S	16 (28%)	0	0
	A317T	28 (49%)	128 (98%)	0
	T735I	26 (46%)	60 (46%)	43 (43%)
	L1110F	35 (61%)	0	0
	V1120I	38 (67%)	0	0
	V1213A	0	125 (95%)	0
	Y1320H	0	0	0
	K1383N	0	0	0
	D1384G	0	0	0
	K1398N	0	0	0
	F1439Y	9 (16%)	0	0
	V1479I	54 (95%)	0	0
	C1483W	0	0	0
	N1530T	0	0	0
	Y1587F	0	0	96 (97%)
	G1634R	0	128 (98%)	96 (97%)
	G1634K	54 (95%)	0	0
		GT1 ORF2 (51)	GT3 ORF2 (131)	GT4 ORF2 (96)
ORF2	P259S	0	0	0
	L477T	46 (90%)	127 (97%)	0
	L613T	46 (90%)	125 (95%)	0

Mutations as a missense SNPs and their effects on protein structural stability

In our analysis, 4 in silico SNP prediction algorithms were used to predict the selected mutations as neutral or deleterious. According to predicted results, most of the selected mutations with high prevalence in HEV infected papulation are deleterious (Table S5). Next, most of these mutations are predicted to cause destabilization of the protein as calculated using 3 web servers (Table S6), suggesting the importance of these mutations. Because the crystal structure of HEV ORF2 (2ZTN-amino acid 129–606) is available (Fig. 3A). We have modeled the effect of 2 mutations that are located

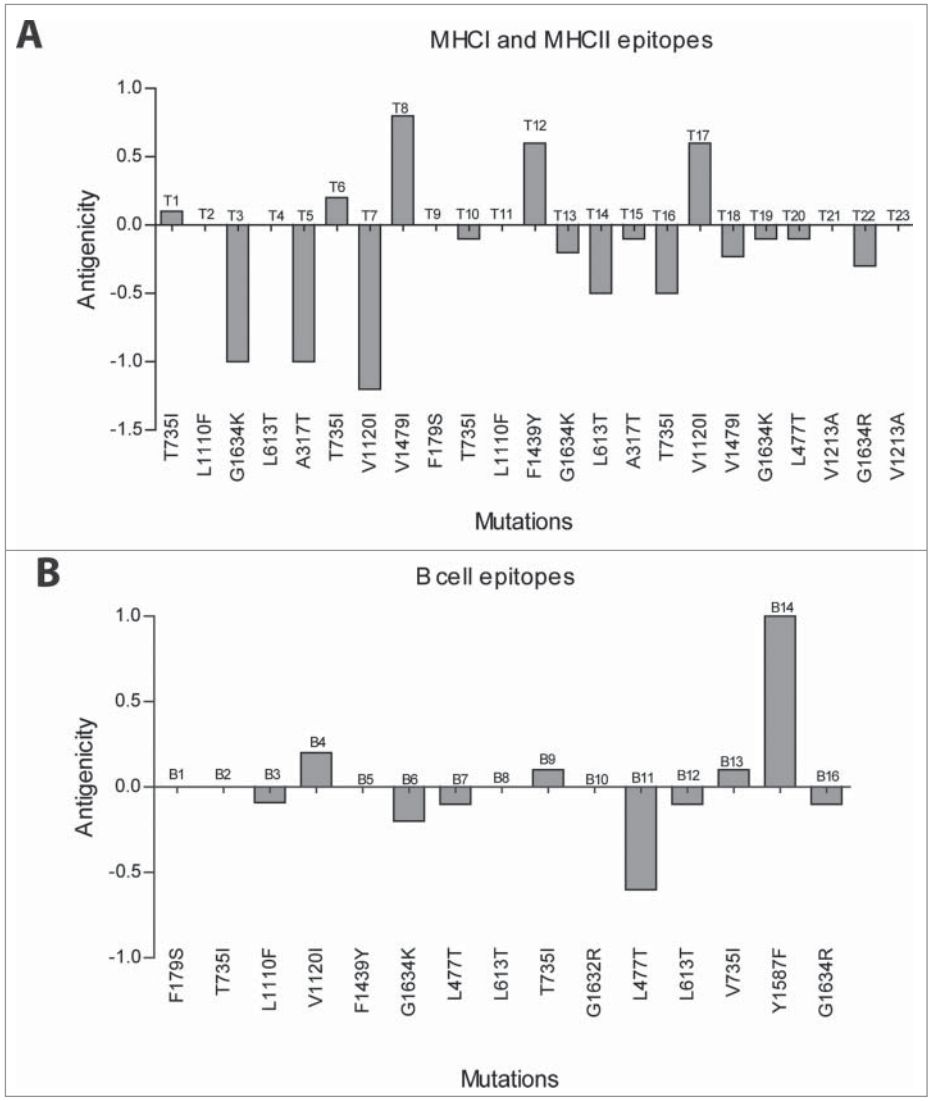


Figure 2. Antigenicity difference between wild-type and mutated T cell (MHCI and MHCII) (A) and B cell epitopes (B).

within this region on structural stability of the protein. We found that both P259S (Fig. 3B) and L477T (Fig. 3C) are predicted to cause destabilization of ORF2 protein.

Overlap of characteristic mutations within the predicted B cell, CD4+ and CD8+ T cell epitopes

The consensus sequence of ORF1 and ORF2 of genotypes 1, 3 and 4 (respectively) was used to predict B cells, major histocompatibility complex class I (MHCI) and class II (MHCII) T cell epitopes using IEDB, ProPred-1 and ProPred, respectively.[10-12] In case of MHCI and II, epitope's binding to maximum number of alleles and binding capacity of < 500 mM were selected. The predicted epitopes were further confirmed by BLASTp[13] to avoid considering the epitopes that have a homology with human proteins.

The predicted epitopes were then evaluated for the presence or absence of the reported mutations (Table 1). We found that many of these characteristic mutations reported in ORF1 and ORF2 regions overlap with the predicted epitopes and were also frequently observed among selected HEV genotypes (1, 3 and 4) (Table 2 and Table 3). In our analysis, all epitopes possessing mutations that are present in our analyzed HEV sequences were considered. Table 2 and Table 3 show mutations in the predicted epitopes of ORF1 and ORF2 (consensus genotype 1, 3 and 4) against B cells, MHC-I and MHC-II T cells. It was observed that ORF1 and ORF2 of genotype 1 possess more epitopes with reported mutations when compared with genotype 3 and 4 (Table 2 and 3). Most of these overlapped mutations were related to hepatic failure (HF), followed by ribavirin treatment failure (RTF), altered immunoreactivity (AI) and chronic hepatitis (CH) (HF>RTF>AI>CH).

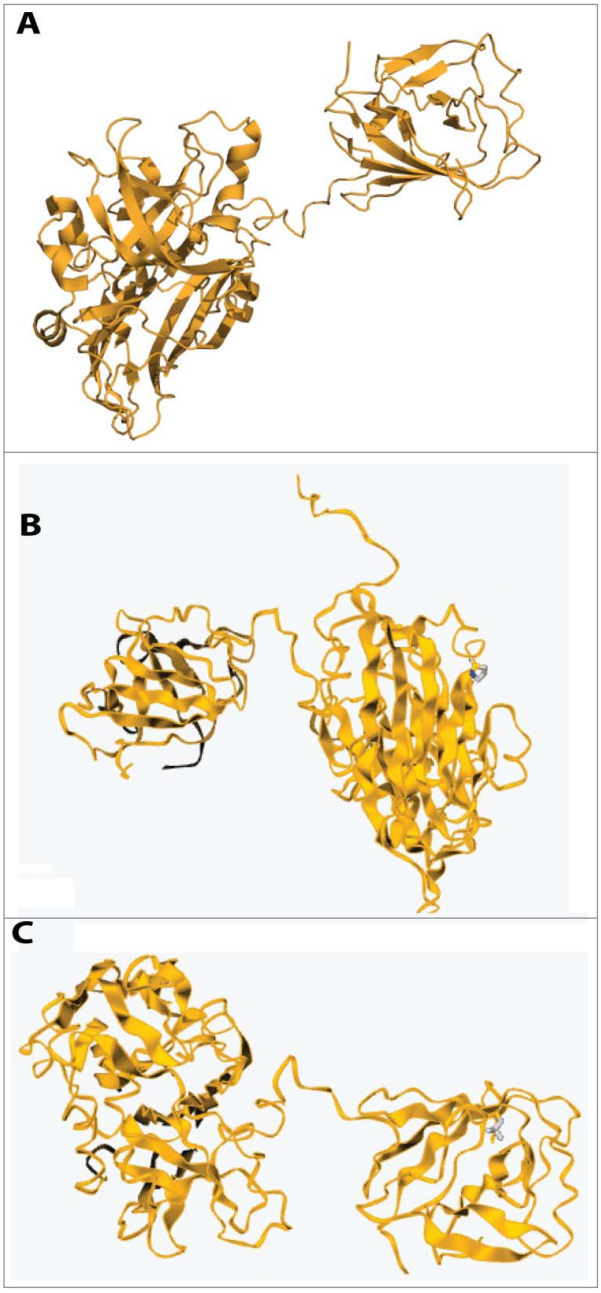


Figure 3. The effects of mutations (P259S, L477T) on the structural stability of ORF2 region of HEV predicted by DUET web server. Mutations are shown in blue color ribbon. (A) Crystal struture without mutation of ORF2 (2ZTN-amino acid 129–606). (B). P259S; feature: distabilizing; secondary structure: loop or irregular. (C). L477T; feature: distabilizing; secondary structure: extended β-strand.

Alteration of epitope antigenicity by characteristic mutations

We hypothesize that HEV may alter its epitopes by acquiring mutations to evade the immune recognition by both B and T cells.[14] The online tool VaxiJen[15] was used to detect the effect of each mutation on the antigenicity of the epitopes (wild type and mutation containing epitopes) (Table 2 and Table 3). Interestingly, in many cases, mutated epitopes have a reduced antigenicity when compared with the wild-type epitope (Table 2, Table 3, and Fig. 2). Some mutated epitopes have a sustained antigenicity, while only few mutated epitopes have an increased antigenicity. These findings suggest that most of the reported mutations within predicted epitopes have a reduced antigenicity. Consequently, they are less recognized by both B and T cells and thus, decreasing the effective roles of the adaptive immune cells to clear the infections.

Evolution of the characteristic mutations

To visualize the mutation evolution within the predicted epitopes of each genotype, a heat map was generated for each reported mutation (Fig. 4). The concept of evolving mutations among the HEV genotypes within antigenic area will be helpful in determining their role in both immune and therapeutic responses. Different trends of evolving mutations have been observed in our analysis. The mutations A317T, T735I, L1110F, V1120I, V1479I and G1634K which were frequently observed in genotype 1 ORF1 have an evolving period of 3 y (1998–2001), 6 y (1998–2003), 8 y (1992–1999) 7 y (1992–1998), 5 y (1983–1987) and 4 y (1983–1986), respectively (Fig. 4 and Fig. 5). A commonly observed variation in all selected genotypes (1, 3 and 4), i.e., T735I, was found to be evolved earlier in ORF1 of genotype 1 (1998–2003) when compared with genotype 3 and 4. Another common variation G1634K/R evolved earlier in ORF1 of genotype 1 (1983–1986) as compared with genotype 3 (1993–1997) and genotype 4 (1995–1998) (Fig. 4 and Fig. 5). Similarly, the mutations in ORF2 (L477T and L613T) appear to evolve earlier in genotype 1 compared with genotype 3 (Fig. 4 and Fig. 5). These data collectively indicate an earlier acquisition of several characteristic mutations in genotype 1 as compared with genotype 3 and 4.

Discussion

Ribavirin monotherapy has been widely used for treatment of chronic hepatitis E.[5] It is in general very effective. However, for a subset of patients, ribavirin treatment fails.[16,17] Mutations in the viral polymerase have been noted before or during therapy in these patients.[18] Furthermore, many mutations related to hepatic failure, chronic hepatitis and immunoreactivity have also been reported, conferring the importance of these mutations in HEV pathogenesis and treatment responses. In support of this, several studies have shown that proteomic variations in certain epitopes that are

Table 2. Antigenicity evaluation of wild-type and mutated epitopes. The effect of mutations on the antigenicity (threshold level = 0.4) of T cell predicted epitopes (MHCI and MHCII). Bold letters (one-letter amino acid code) represent mutations within the predicted epitopes.

Region/ Genotype	Epitopes name	Position	Wild-type epitope	Antigenicity	Mutated epitope	Antigenicity
			MHCI			
ORF1/GT1	T1	F179S	MFRHGMTRL	0.1	M**F/S**RHGMTRL	0.2
	T2	T735I	ATPTPAAPL	0.2	ATP**T/I**PAAPL	0.2
	T3	L1110F	TTSRVLRSL	0.4	TTSRV**L/F**RSL	−0.6
	T4	F1439Y	FYGDAFDDT	0.2	FYGDA**F/Y**DDT	0.2
	T5	G1634K	AVSDFLRGL	0.4	AVSDFLR**G/K**L	−0.6
ORF2/GT1	T6	L613T	SALALLEDL	0.3	SALALLED**L/T**	0.5
			MHCII			
ORF1/GT1	T7	A317T	FHAVPAHIW	0.5	FH**A**VP**A/T**HIW	−0.7
	T8	T735I	IPSRAATPT	0.2	IPSRAATP**T/I**	1
	T9	V1120I	FWGEPAVGQ	0.2	FWGEPA**V/I**GQ	0.2
	T10	V1479I	LGLECAVME	0.8	LGLECA**V/I**ME	0.7
	T11	G1634K	LRGLTNVAQ	0.7	LR**G/K**LTNVAQ	0.7
ORF2/GT1	T12	L477T	WLSLLAAEY	1.1	WLSL**L/T**AAEY	1.7
			MHCI			
ORF1/GT3	T13	V1213A	VIVNNFFLV	0.1	VIVNNFFL**V/A**	−0.1
	T14	G1634R	LGLAVCDFL	0.2	L**G/R**LAVCDFL	−0.3
			MHCII			
ORF1/GT3	T15	V1213A	LVGGEVGHH	1.1	L**V/A**GGEVGHH	1
	T16	G1634R	LGLAVCDFL	0.2	L**G/R**LAVCDFL	−0.3
ORF2/GT3	T17	L477T	WLSLLAAEY	1.1	WLSL**L/T**AAEY	1.7
	T18	L613T	VLEDLIDYP	0.4	VLED**L/T**IDYP	−0.17
			MHCI			
ORF1/GT4	T19	T735I	AEADTPVAV	0.2	AEAD**T/I**PVAV	0.1
	T20	V1213A	REVGISDAI	0.8	REVGISD**V/A**I	0.7
	T21	G1634R	SERAEQLRL	0.3	SE**G/R**AEQLRL	0.3
			MHCII			
ORF1/GT4	T22	T735I	VATDVPPPA	0.4	VA**T/I**DVPPPA	0.1
	T23	V1213A	VIVNNFFLS	0.03	**V/A**IVNNFFLS	0.02

associated with these mutations can critically influence the outcome of the immune responses.[19] These antigenic variations have been observed among HEV strains using genotype- and strain-specific monoclonal antibodies.[20] Even though several studies have reported frequencies of mutations in the HEV genome, the global prevalence of these characteristic mutations has not been comprehensively studied. Thus, this study has evaluated the HEV ORF1 and ORF2 variability in major genotypes (1, 3 and 4) by a systematic retrieval of human-derived HEV sequences. Furthermore, important co-occurrence of B and T cell ($CD4^+$ and $CD8^+$) epitope mutations was revealed, suggesting adaptation of viruses to escape immune surveillance.

By exploring the intra-genotypic diversity from representative human HEV genotypes, we have demonstrated that genotype 1 produces the largest number of intra-genotypic variants (Table 1). These variants were found to be evolved with a period of on average ~4 y but varied from 3 to 5 y in the human population (Fig. 4 and Fig. 5). Surprisingly, some reported mutations, including Y1320H, K1383N, D1384G and

Table 3. Antigenicity evaluation of wild-type and mutated epitopes. The effect of mutations on the antigenicity (threshold level = 0.4) of predicted B cells epitopes. Bold letters (one-letter amino acid code) represent mutations within the predicted epitopes.

Region/Genotype	Epitopes name	Position	Wild-type B cell epitopes	Antigenicity	Mutated B cell epitopes	Antigenicity
	B1	F179S	FRHGMTRLY	0.1	**F/S**RHGMTRLY	0.1
ORF1/GT1	B2	T735I	TPAAPLPPP	0.1	**T/I**PAAPLPPP	0.1
	B3	L1110F	TTSRVLRSL	0.1	TTSRV**L/F**RSL	0.01
	B4	V1120I	WGEPAVGQK	0.8	WGEPA**V/I**GQK	1
	B5	F1439Y	FYGDAFDDT	0.2	FYGDA**F/Y**DDT	0.2
	B6	G1634K	SDFLRKLTN	0.1	SDFLR**G/K**LTN	−0.1
ORF2/GT1	B7	L477T	SLTAAEYDQ	1.1	SLL**/T**AAEYDQ	1.0
	B8	L613T	LLDYPARAH	0.4	**L/T**LDYPARAH	0.4
ORF1/GT3	B9	T735I	LPHPTPPVS	0.5	LPHP**T/I**PPVS	0.6
	B10	G1632R	ERAEQLRLA	0.1	ERAEQL**G/R**LA	0.1
ORF2/GT3	B11	L477T	SLTAAEYDQ	1.1	SLL**/T**AAEYDQ	0.5
	B12	L613T	LIDYPARAH	0.5	**L/T**IDYPARAH	0.4
ORF1/GT4	B13	V735I	TPVAVDVPP	0.5	TPVA**V/I**DVPP	0.6
	B14	Y1587F	CGLKLKVDY	0.7	CGLKLKVDY**/F**	1.7
	B15	G1634R	KNWGPGSEG	1.8	KNWGPGSE**G/R**	1.7

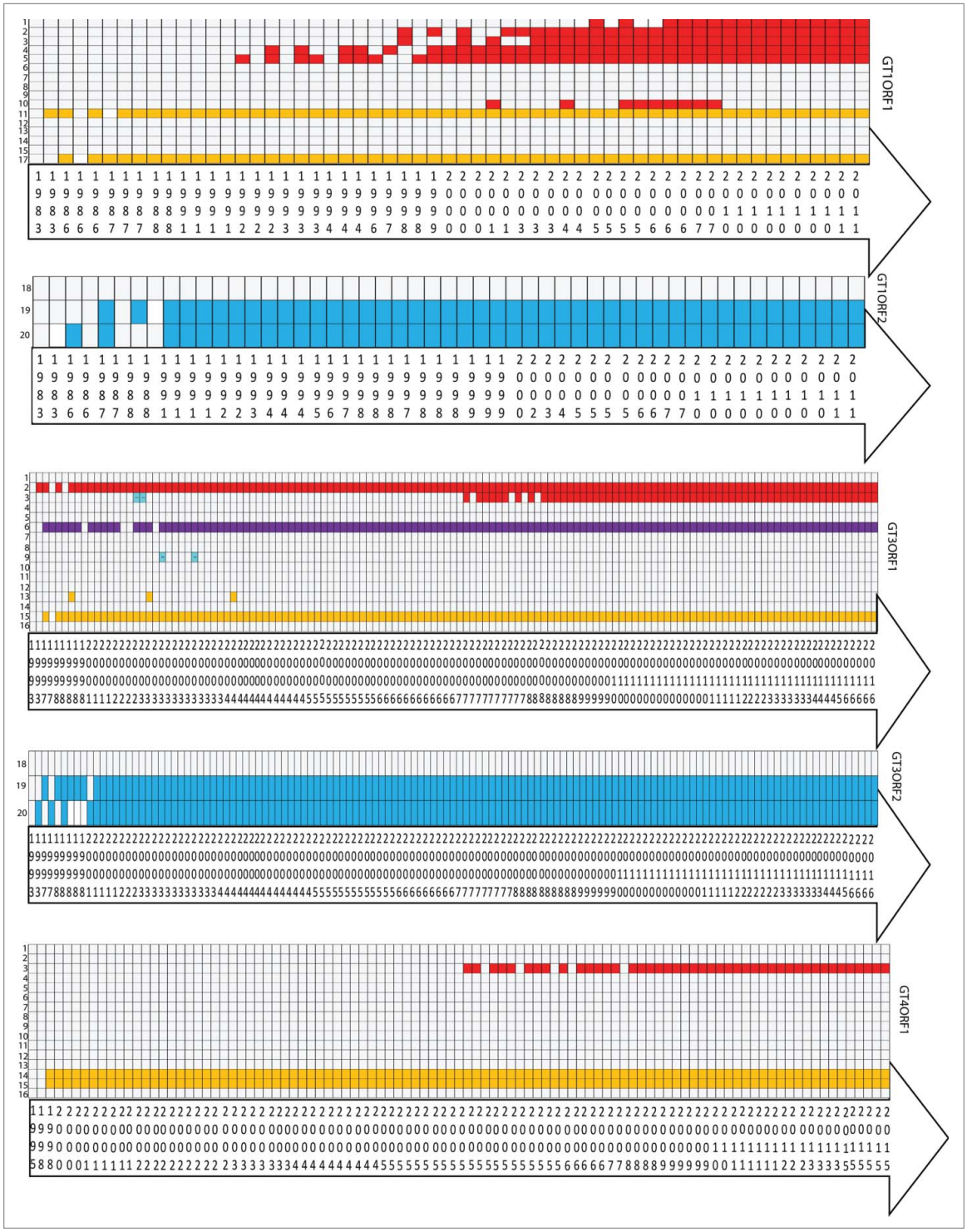


Figure 4. A heat map showing the evolvment of mutations with years. Mutations 1 = F179S, 2 = A317T, 3 = T735I, 4 = L1110F, 5 = V1120I, 6 = V1213A, 7 = Y1320H, 8 = K1383N, 9 = D1384G, 10 = K1398N, 11 = F1439Y, 12 = V1479I, 13 = C1483W, 14 = N1530T, 15 = Y1587F, 16 = G1634R, 17 = G1634K, 18 = P259S, 19 = L477T, 20 = L613T. Each box represents the sequence of ORF 1 or ORF2, from genotype 1, 3 or 4 in a particular year. Red colored boxes represent the mutations related to hepatic failure mutations; purple to chronic hepatitis; yellow to ribavirin treatment failure; and blue to altered immunoreactivity. The year of deposition of each sequence is mentioned in the arrow below the heat map.

K1398N (related to ribavirin treatment failure) and C1483W, N1530T and P259S (related to hepatic failure), were hardly present in our retrieved large set of sequences, suggesting their insignificance. In contrast, 2 mutations T735I (related to hepatic failure) and G1634R (related to ribavirin treatment failure) were found in all selected genotypes (1, 3 and 4) with a considerably high frequency. Among these mutations,

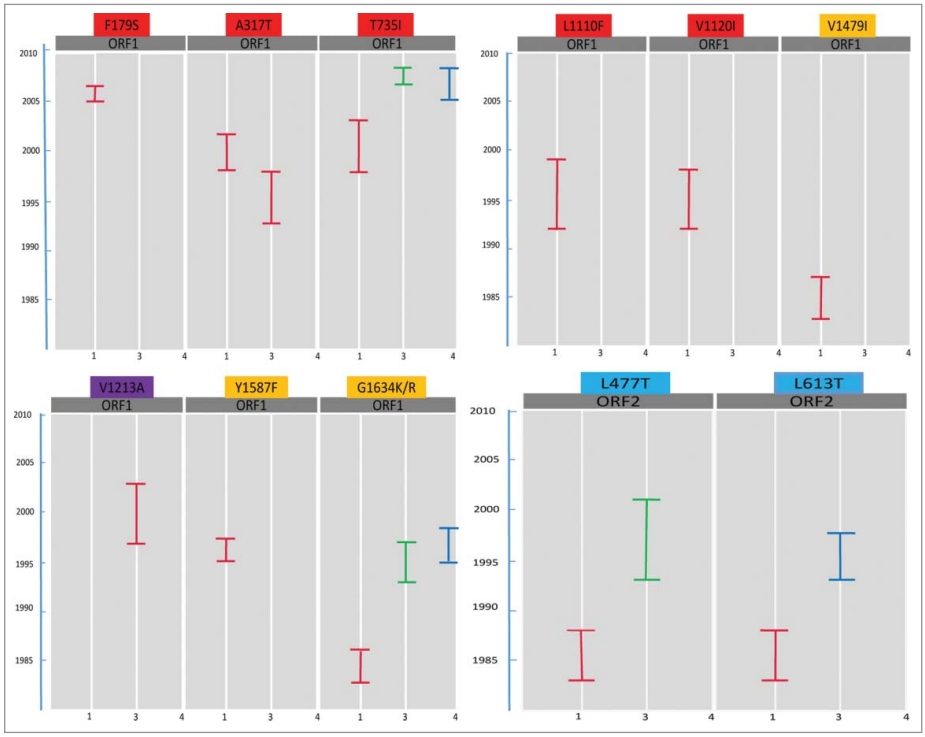


Figure 5. Evolution period of overlapped mutations. Each bar represents the evolving period (years) of each mutations.

G1634R was identified in patients as a baseline mutation that effects ribavirin treatment response.[21] This mutation was also observed in a patient with chronic hepatitis E, experiencing ribavirin treatment failure with a completely resistant phenotype.[16] However, the exact clinical relevance of such mutations in ribavirin treatment failure is uncertain since findings from *in vitro* studies show that some of these mutations facilitate HEV replication but paradoxically seem to increase ribavirin sensitivity,[16] thus requiring further investigation. Furthermore, effective new antiviral therapy is needed for HEV patients with ribavirin treatment failure.[22] Two immunoreactivity-related mutations (L477T and L613T) were mainly found in genotype 1 and 3 highlighting the role of such mutations in affecting host immune response (Table 1).

To investigate the implication of these characteristic mutations in the host immune response, we have profiled their overlap with predicted B and T cell epitopes. We found that the hotspot sites where mutations and predicted epitopes overlap are frequently present among many HEV genotypes (Table 1 and Fig. 2). The subsequent analysis of these mutations showed that this overlap mostly decreases and in few cases, sustains or enhances the antigenicity of the mutated epitopes (Table 2, Table 3 and Fig. 2). Experimental studies have demonstrated that only antibodies recognizing conformational epitopes are neutralizing, and the aa residues Leu477 and Leu613 in the capsid protein are important in forming a neutralization-sensitive epitope representing the importance of these mutations in epitope deformation.[23] In patients, HEV-specific T cells target relatively conserved HEV peptides, and those are predominantly located in the ORF2 capsid protein. The T-cell responses persist over years after resolution of HEV infection, suggesting the role in both clearance of primary infection and protective immune response against secondary infection.[24] Based on our study, we propose a model of immune evasion by HEV (Fig. 6). In this model, mutated epitopes will result in escape recognition by B and T cells. Furthermore, we have mapped the evolution of these mutations in all selected genotypes. We found that most of the overlapped mutations are more abundantly present in genotype 1 as compared with genotype 3 and 4. Interestingly, the common mutations evolved earlier in genotype 1 than genotype 3 and 4.

In summary, our study represents a comprehensive analysis of the characteristic mutations in the major HEV genotypes. Some of these mutations overlap with predicted B and T cell epitopes that are expected to affect the antigenicity. We further revealed the evolution of these mutations among the 3 major genotypes. These results indicate that HEV is under substantial evolutionary pressure to develop mutations enabling evasion of the host immune response and resistance to antiviral treatment. This indicates the existence of an ongoing evolutionary arms race between human immunity, antiviral medication and the HEV.

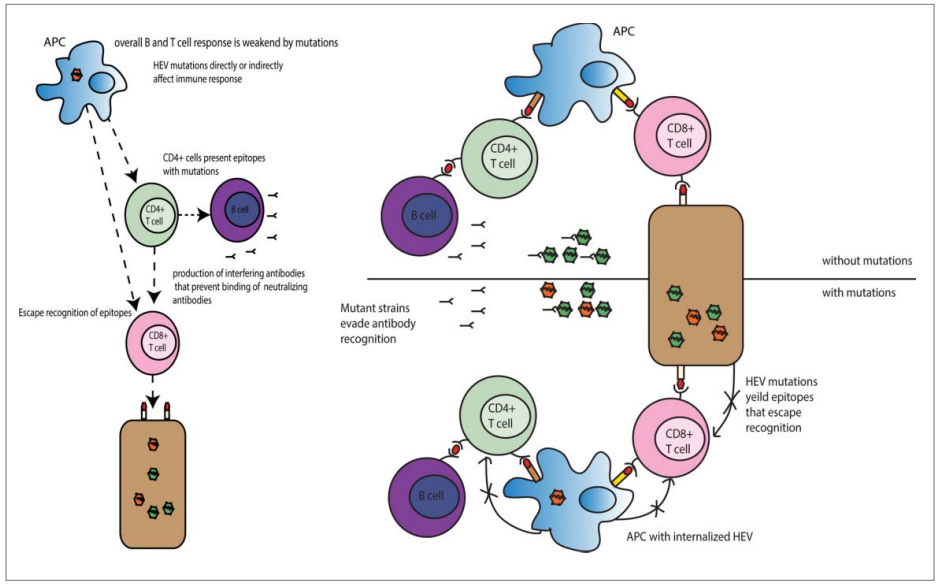


Figure 6. Possible mechanisms of immune evasion by hepatitis E virus. Main routes by which HEV mutations may result in evasion of the host immune responses. Mutated epitopes presented by antigen presenting cells, B and T cells will results in escape recognition of the epitopes.

Material and methods

Data collection

A database on reported mutations within the ORF1 and ORF2 regions of HEV is currently not available. We reviewed earlier studies (through January 2017), using a combination of the keywords that are listed in Table S1, and evaluated mutations in HEV-infected individuals, as well as in HEV replicons. These studies were searched from the PubMed (ncbi.nlm.nih.gov/pubmed), EMBASE, and Cochrane Library databases.

Retrieval of sequences

We retrieved 57 ORF1, 51 ORF2 (genotype 1), 131 ORF1, 131 ORF2 (genotype 3) and 99 ORF1, 96 ORF2 (genotype 4) of HEV protein sequences from the GenBank[25] (accessed on January 2017). The selected protein sequences represent all major genotypes (1, 3 and 4). The GenBank accession numbers of HEV protein sequences used in this study against each ORF and genotype are provided in Table S3. The sequences were trimmed manually and analyzed using the reference protein sequences of selected genotypes [AF185822 (genotype 1), AB291960 (genotype 3), AB200239 (genotype 4)]. Different quality control measures were performed for the sequences, and many sequences were disqualified for further analysis based on the following 2 conditions: (a) if the sequence derived from a non-human host; and (b) if the sequence was a clonal sequence from the same patient. Sequences were also annotated by the year of sampling. In some cases where the source did not provide the sampling year, we used the submission date to the GenBank as the sampling year.

Consensus sequence and mutations analysis

The HEV ORF1 and ORF2 sequences were aligned using ClustalW, BioEdit and CLC Workbench 7 (http://www.clcbio.com). A consensus analysis for each HEV genotype was performed to observe the presence or absence of mutation at each site. These mutations were selected on the basis of published data reporting all experimentally proven mutations (*in vitro* and *in vivo*). The prevalence of each mutation was then measured within the selected regions (ORF1 and ORF2) of each genotype (1, 3 and 4).

Analysis of mutations as a missense SNPs and their effects on structural stability

To validate selected mutations as missense SNPs in ORF1 and ORF2 regions of HEV genotype 1, 3 and 4, computational analysis was performed using 4 tools PROVEAN (Protein Variation Effect Analyzer),[26] nsSNP analyzer,[27] SNPs & GO and PMUT.[28] These tools describe missense SNPs as damaging or neutral to function and structure. To predict the change in protein stability due to these SNPs, DUET,[29] I-Mutant version 2.0,[30] and STRUM[31] web servers were used. As an input in I-Mutant and STRUM servers, FASTA sequences of ORF1, ORF2 regions of selected genotypes were used. PDB files of 3D structures (2ZTN) of ORF2 were used as an input in DUET web server.

Prediction of the B and T cell epitopes and comparison with the host proteome

Nine-mer B cell epitopes were predicted against HEV genotype (1, 3 and 4) from their consensus sequence by using online tool, the Immune Epitope Database (IEDB).[32] Similarly, 9-mer T cell epitopes (MHC class I and II), following the same criteria as B cell epitopes, were predicted by online T cell epitope prediction tools, including ProPred-I (MHCI) and ProPred (MHCII).[33,34] ProPred-I and ProPred identify and predict 47 types of MHCI and 57 types of MHCII allele specific binding peptides in a provided protein, respectively.[33,34] The predicted T cell epitopes were also confirmed by IEDB tool.[35] The selected epitopes were analyzed for comparisons with the human proteome using the Protein Blast program (BLASTp).[13] This was performed to validate that these epitopes will not trigger an autoimmune response.

Overlapping sites of mutations and predicted epitopes

A comprehensive exploration was performed to find out any reported mutations positioned in a predicted epitope. Predicted B cells and T cells epitopes were mapped to ORF1 and ORF2 regions that contain most of the reported mutation sites. Finally, the incidence of these overlapped mutations was determined among the selected genotypes (1, 3 and 4) by using percentage formula.

Prediction of antigenicity of epitopes

To investigate the antigenic properties of epitopes before and after mutations, VaxiJen tool was used.[15] The analysis was performed to find out whether these mutations lead to decreased, enhanced or sustained antigenicity of the specific epitopes.

Evolutionary analysis of mutations

To visualize the evolution of amino acid mutation related to chronic hepatitis, hepatic failure, ribavirin treatment failure or immunoreactivity within each genotype, a heat map was developed by GraphPad Prism version 7 (GraphPad Software, La Jolla California, USA) to calculate the number of amino acid and the isolation year differences between 3 individual genotypes. Isolation years were extracted from the strain-annotated information. The difference values were added into a matrix where the y-axis represents the isolation year differences and the x-axis represents the amino acid differences.

Abbreviations

AI	Altered immunoreactivity
CH	Chronic hepatitis
HCV	Hepatitis C virus (HCV)
HEV	Hepatitis E virus
HF	Hepatic failure
MHCI	Major histocompatibility complex class I
MHCII	Major histocompatibility complex class II
ORF	Open reading frame
PEG-IFN-α	Pegylated interferon-α
RdRp	RNA-dependent RNA polymerase
RNA	Ribonucleic acid
RTF	Ribavirin treatment failure

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Funding

The authors thank to the Higher Education Commission of Pakistan for financial support to A. Ikram; the Indonesia Endowment Fund for Education (LPDP) for PhD scholarship to Mohamad S. Hakim; the China Scholarship Council for funding PhD fellowship to W. Wang (201303250056); the Dutch Digestive Foundation (MLDS) for a career development grant (No. CDG 1304), the Daniel den Hoed Foundation for a Centennial Award fellowship and the Erasmus MC Mrace grant to Q. Pan.

References

[1] Kamar N, Bendall R, Legrand-Abravanel F, Xia NS, Ijaz S, Izopet J, Dalton HR. Hepatitis E. Lancet. 2012; 379:2477-88. https://doi.org/10.1016/S0140-6736(11)61849-7. PMID:22549046

[2] Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380:2095-128. https://doi.org/10.1016/S0140-6736(12)61728-0. PMID:23245604

[3] Teshale EH, Hu DJ. Hepatitis E: Epidemiology and prevention. World J Hepatol. 2011;3:285-91. https://doi.org/10.4254/wjh.v3.i12.285. PMID:22216368

[4] Hakim MS, Wang W, Bramer WM, Geng J, Huang F, de Man RA, Peppelenbosch MP, Pan Q. The global burden of hepatitis E outbreaks: A systematic review. Liver Int. 2017;37:19-31. https://doi.org/10.1111/liv.13237. PMID:27542764

[5] Kamar N, Mallet V, Izopet J. Ribavirin for chronic hepatitis E virus infection. N Engl J Med. 2014;370:2447-8. https://doi.org/10.1056/NEJMoa1215246. PMID:24941183

[6] Wedemeyer H, Pischke S, Manns MP. Pathogenesis and treatment of hepatitis e virus infection. Gastroenterology.

2012;142:1388-97 e1. https://doi.org/10.1053/j.gastro.2012.02.014. PMID:22537448

[7] Takahashi K, Toyota J, Karino Y, Kang JH, Maekubo H, Abe N, Mishiro S. Estimation of the mutation rate of hepatitis E virus based on a set of closely related 7.5-year-apart isolates from Sapporo, Japan. Hepatol Res. 2004;29:212-5. https://doi.org/10.1016/j.hepres.2004.04.004. PMID:15288013

[8] Lhomme S, Garrouste C, Kamar N, Saune K, Abravanel F, Mansuy JM, Dubois M, Rostaing L, Izopet J. Influence of polyproline region and macro domain genetic heterogeneity on HEV persistence in immunocompromised patients. J Infect Dis. 2014;209:300-3. https://doi.org/10.1093/infdis/jit438. PMID:23964111

[9] van Tong H, Hoan NX, Wang B, Wedemeyer H, Bock CT, Velavan TP. Hepatitis E virus mutations: Functional and clinical relevance. EBioMedicine. 2016;11:31-42. https://doi.org/10.1016/j.ebiom.2016.07.039. PMID:27528267

[10] Singh R, Kaul R, Kaul A, Khan K. A comparative review of HLA associations with hepatitis B and C viral infections across global populations. World J Gastroenterol. 2007;13:1770-87. https://doi.org/10.3748/wjg.v13.i12.1770. PMID:17465466

[11] Sturniolo T, Bono E, Ding J, Raddrizzani L, Tuereci O, Sahin U, Braxenthaler M, Gallazzi F, Protti MP, Sinigaglia F, et al. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. Nat Biotechnol. 1999;17:555-61. https://doi.org/10.1038/9858. PMID:10385319

[12] Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, Wheeler DK, Gabbard JL, Hix D, Sette A, et al. The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 2015;43:D405-12. https://doi.org/10.1093/nar/gku938. PMID:25300482

[13] Lavigne R, Seto D, Mahadevan P, Ackermann HW, Kropinski AM. Unifying classical and molecular taxonomic classification: Analysis of the Podoviridae using BLASTP-based tools. Res Microbiol. 2008;159:406-14. https://doi.org/10.1016/j.resmic.2008.03.005. PMID:18555669

[14] Gu Y, Tang X, Zhang X, Song C, Zheng M, Wang K, Zhang J, Ng MH, Hew CL, Li S, et al. Structural basis for the neutralization of hepatitis E virus by a cross-genotype antibody. Cell Res. 2015;25:604-20. https://doi.org/10.1038/cr.2015.34. PMID:25793314

[15] Doytchinova IA, Flower DR. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007;8:4. https://doi.org/10.1186/1471-2105-8-4. PMID:17207271

[16] Debing Y, Ramiere C, Dallmeier K, Piorkowski G, Trabaud MA, Lebosse F, Scholtès C, Roche M, Legras-Lachuer C, de Lamballerie X, et al. Hepatitis E virus mutations associated with ribavirin treatment failure result in altered viral fitness and ribavirin sensitivity. J Hepatol. 2016;65:499-508. https://doi.org/10.1016/j.jhep.2016.05.002. PMID:27174035

[17] Debing Y, Gisa A, Dallmeier K, Pischke S, Bremer B, Manns M, Wedemeyer H, Suneetha PV, Neyts J. A mutation in the hepatitis E virus RNA polymerase promotes its replication and associates with ribavirin treatment failure in organ transplant recipients. Gastroenterology. 2014;147:1008-11 e7. quiz e15-6. https://doi.org/10.1053/j.gastro.2014.08.040. PMID:25181691

[18] Todt D, Walter S, Brown RJ, Steinmann E. Mutagenic effects of ribavirin on hepatitis E virus-viral extinction versus selection of fitness-enhancing mutations. Viruses. 2016; 8(10): 283-297. https://doi.org/10.3390/v8100283. PMID:27754363.

[19] Liang JH, Dai X, Dong C, Meng JH. A single amino acid substitution changes antigenicity of ORF2-encoded proteins of hepatitis E virus. Int J Mol Sci. 2010;11:2962-75. https://doi.org/10.3390/ijms11082962. PMID:21152284

[20] Liang J-H, Dai X, Dong C, Meng J-H. A single amino acid substitution changes antigenicity of ORF2-encoded proteins of hepatitis E virus. Int J Mol Sci. 2010;11:2962-75. https://doi.org/10.3390/ijms11082962. PMID:21152284

[21] Lhomme S, Kamar N, Nicot F, Ducos J, Bismuth M, Garrigue V, Petitjean-Lecherbonnier J, Ollivier I, Alessandri-Gradt E, Goria O, et al. Mutation in the hepatitis E virus polymerase and outcome of ribavirin therapy. Antimicrob Agents Chemother. 2016;60:1608-14. https://doi.org/10.1128/AAC.02496-15.

[22] Kamar N, Wang W, Dalton HR, Pan Q. Direct-acting antiviral therapy for hepatitis E virus? Lancet Gastroenterol Hepatol. 2017;2:154-5. https://doi.org/10.1016/S2468-1253(16)30242-4. PMID:28404127

[23] Zhang H, Dai X, Shan X, Meng J. The Leu477 and Leu613 of ORF2-encoded protein are critical in forming neutralization antigenic epitope of hepatitis E virus genotype 4. Cell Mol Immunol. 2008;5:447-56. https://doi.org/10.1038/cmi.2008.56. PMID:19118511

[24] Brown A, Halliday JS, Swadling L, Madden RG, Bendall R, Hunter JG, Maggs J, Simmonds P, Smith DB, Vine L, et al. Characterization of the specificity, functionality, and durability of host T-Cell responses against the full-length hepatitis E virus. Hepatology. 2016;64:1934-50. https://doi.org/10.1002/hep.28819. PMID:27631819

[25] Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2013;41:D36-42. https://doi.org/10.1093/nar/gks1195. PMID:23193287.

[26] Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics. 2015;31:2745-7. https://doi.org/10.1093/bioinformatics/btv195. PMID:25851949

[27] Bao L, Zhou M, Cui Y. nsSNPAnalyzer: Identifying disease-associated nonsynonymous single nucleotide polymorphisms. Nucleic Acids Res. 2005;33:W480-W2. https://doi.org/10.1093/nar/gki372. PMID:15980516

[28] Ferrer-Costa C, Gelpí JL, Zamakola L, Parraga I, De La Cruz X, Orozco M. PMUT: A web-based tool for the annotation of pathological mutations on proteins. Bioinformatics. 2005;21:3176-8. https://doi.org/10.1093/bioinformatics/bti486. PMID:15879453

[29] Pires DE, Ascher DB, Blundell TL. DUET: A server for predicting effects of mutations on protein stability using an integrated computational approach. Nucleic Acids Res. 2014;42:W314-9. https://doi.org/10.1093/nar/gku411. PMID:24829462

[30] Capriotti E, Fariselli P, Casadio R. I-Mutant2. 0: Predicting stability changes upon mutation from the

protein sequence or structure. Nucleic Acids Res. 2005;33:W306-W10. https://doi.org/10.1093/nar/gki375. PMID:15980478

[31] Quan L, Lv Q, Zhang Y. STRUM: Structure-based prediction of protein stability changes upon single-point mutation. Bioinformatics. 2016;32:2936-46. https://doi.org/10.1093/bioinformatics/btw361. PMID:27318206

[32] Vita R, Zarebski L, Greenbaum JA, Emami H, Hoof I, Salimi N, Damle R, Sette A, Peters B. The immune epitope database 2.0. Nucleic Acids Res. 2010;38:D854-62. https://doi.org/10.1093/nar/gkp1004. PMID:19906713

[33] Singh H, Raghava GP. ProPred: Prediction of HLA-DR binding sites. Bioinformatics. 2001;17:1236-7. https://doi.org/10.1093/bioinformatics/17.12.1236. PMID:11751237

[34] Singh H, Raghava GP. ProPred1: Prediction of promiscuous MHC Class-I binding sites. Bioinformatics. 2003;19:1009-14. https://doi.org/10.1093/bioinformatics/btg108. PMID:12761064

[35] Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, Lundegaard C, Sette A, Lund O, Bourne PE, et al. Immune epitope database analysis resource. Nucleic Acids Res. 2012;40:W525-30. https://doi.org/10.1093/nar/gks438. PMID:22610854