

## RESEARCH ARTICLE

# A tightly clustered hepatitis E virus genotype 1a is associated with endemic and outbreak infections in Bangladesh

Trang Nguyen Hoa<sup>1</sup>, Saif Ullah Munshi<sup>2</sup>, Khanh Nguyen Ngoc<sup>1</sup>, Chau Le Ngoc<sup>1</sup>, Thanh Tran Thi Thanh<sup>1</sup>, Tahmina Akther<sup>2</sup>, Shahina Tabassum<sup>2</sup>, Nilufa Parvin<sup>3</sup>, Stephen Baker<sup>4</sup>, Motiur Rahman<sup>1,5\*</sup>

**1** Oxford University Clinical Research Unit, Wellcome Asia Programme, The Hospital for Tropical Diseases, Ho Chi Minh City, Vietnam, **2** Bangabandhu Sheikh Mujib Medical University, Shahbag, Dhaka, Bangladesh, **3** Sir Salimullah Medical College and Hospital (SSMCH), Dhaka, Bangladesh, **4** Cambridge Institute of Therapeutic Immunology & Infectious Disease, Cambridge University, Cambridge, England, **5** Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, Oxford University, Oxford, United Kingdom

\* [mrahman@oucru.org](mailto:mrahman@oucru.org)



## OPEN ACCESS

**Citation:** Hoa TN, Munshi SU, Ngoc KN, Ngoc CL, Thanh TTT, Akther T, et al. (2021) A tightly clustered hepatitis E virus genotype 1a is associated with endemic and outbreak infections in Bangladesh. PLoS ONE 16(7): e0255054. <https://doi.org/10.1371/journal.pone.0255054>

**Editor:** Pierre Roques, CEA, FRANCE

**Received:** December 16, 2020

**Accepted:** July 8, 2021

**Published:** July 22, 2021

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0255054>

**Copyright:** © 2021 Hoa et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All sequences were submitted to GenBank. Accession no: (MK005535 to MK005551) for the ORF2-3 region, and (MH991993 to MH992013) for the whole genome sequences.

## Abstract

### Background

Hepatitis E virus (HEV) infection is endemic in Bangladesh and there are occasional outbreaks. The molecular characteristics and pathogenesis of endemic and outbreak HEV strains are poorly understood. We compared the genetic relatedness and virulence associated mutations of endemic HEV strains with outbreak strains.

### Methods

We analyzed systematically collected serum samples from HEV immunoglobulin M (IgM) positive patients attended at Bangabandhu Sheikh Mujib Medical University, Dhaka from August 2013 to June 2015. HEV RNA positive samples were subjected to whole genome sequencing. Genotype and subtype of the strains were determined by phylogenetic analysis. Virulence associated mutations e.g. acute viral hepatitis (AVH), fulminant hepatic failure (FHF), chronic hepatitis, ribavirin treatment failure (RTF), B and T cell neutralization epitopes were determined.

### Results

92 HEV immunoglobulin M (IgM) antibody positive plasma samples (43 in 2013–2014 and 49 in 2014–2015) were studied. 77.1% (70/92) of the samples were HEV RNA positive. A 279 bp open reading frame (ORF) 2 and ORF 3 sequence was obtained from 54.2% (38/70) of the strains. Of these 38 strains, whole genome sequence (WGS) was obtained from 21 strains. In phylogenetic analysis of 38 (279 bp) sequence all HEV sequences belonged to genotype 1 and subtype 1a. Further phylogenetic analysis of 21 HEV WGS, Bangladeshi HEV sequences clustered with genotype 1a sequences from neighboring countries. Within genotype 1a cluster, Bangladesh HEV strains formed a separate cluster with the 2010 HEV

**Funding:** This study was supported by the Oxford University Clinical Research Unit, Vietnam.

**Competing interests:** The authors have declared that no competing interests exist.

outbreak strains from northern Bangladesh. 80.9 to 100% of the strains had A317T, T735I, L1120I, L1110F, P259S, V1479I, G1634K mutations associates AVH, FHF and RTF. Mutations in T cell recognition epitope T3, T5, T7 was observed in 76.1%, 100% and 100% of the strains respectively.

## Conclusion

Strains of HEV genotype 1a are dominant in Bangladesh and are associated with endemic and outbreak of HEV infection. HEV isolates in Bangladesh have high prevalence of virulence associated mutations and mutation which alters antigenicity to B and T cell epitopes.

## Introduction

Hepatitis E virus (HEV) infection account for 20 million infections globally per year, 3.3 million of these are symptomatic, and associated with 57,000 deaths [1,2]. The principal burden of HEV disease is in low-middle income countries (LMICs); an estimated 2 billion people live in regions where HEV is endemic [3]. The clinical outcome of HEV infection is varied and largely dependent the genotype of the infecting virus and host factors. The overwhelming majority of HEV infections are self-limiting, but infections in pregnant women are associated with high mortality [4].

HEV is a non-enveloped, single-stranded, positive-sense RNA virus, and a member of the family *Hepeviridae* [5]. The HEV genome is 7.2 kb and is comprised of a short 5' untranslated region (UTR), three open reading frames (ORF1, ORF2, and ORF3), and a 3' UTR, followed by a poly A tail [6]. ORF1, ORF2, and ORF3 encode the nonstructural protein, the viral capsid protein, and a phosphorylated protein, respectively [7,8]. HEV from human, animals, and primates are currently divided into eight genotypes, which includes four genotypes known to cause disease in humans (genotype 1, 2, 3 and 4). These four genotypes are further subdivided into 24 subtypes (genotype 1 (n = 5), 2 (n = 2), 3 (n = 10), and 4 (n = 7)) and a single serotype [5,9].

The various HEV genotypes have a distinct geographic distribution, clinical presentation, and mode of transmission. Genotypes 1 and 2 are found only in humans and are responsible for most HEV cases in Asian LMICs (Myanmar, Pakistan, Bangladesh, India, and Nepal). Genotype 2 is more commonly identified in Latin America and Africa [10]. These genotypes are transmitted fecal-orally, are associated with contaminated water, and typically cause acute infection. Alternatively, genotypes 3 and 4 circulate in various animals (pigs, wild boar, and deer) and are the main cause of HEV infection in high-income countries [10–12]. Genotype 3 is chiefly identified in Europe; genotype 4 is mainly found in more affluent Asian countries including China, Taiwan, Vietnam, and Japan. These genotypes are associated with infection via cross-species transmission often through the consumption of contaminated food [12–14].

The pathogenesis of HEV is poorly understood, although a number of genetic loci have been found to be associated with disease progression. There are several mutations in the HEV genome that have been found to be associated with fulminant hepatic failure (FHF), acute viral hepatitis (AVH), and chronic hepatitis. Additionally, further mutations correlating with ribavirin treatment failure, and B and T cell epitope neutralization have been reported [15–17]. It has been further suggested that these B and T cell epitope mutations are deleterious and decrease the antigenicity of HEV; the mutated epitopes result in variants that evade host B and T recognition [8,18,19].

Bangladesh has a high HEV burden, where small outbreaks are common. The prevalence of HEV antibody among the general population is 22%-60% [20–22]; the incidence of HEV sero-conversion is 40 per 1,000 people per year in rural settings [21]. Drinking municipal tap water, outdoor employment, recent urban travel, and recent contact with a patient with jaundice have been identified as HEV infection risk factors [21,23]. In 2008–2009 a HEV outbreak associated with 4,521 infections and 17 deaths was reported near the capital Dhaka [23]. In 2010 a similar outbreak with 200 hospitalization has been reported in northeast city Rajshahi [22]. Despite the significance of HEV in Bangladesh there has been limited characterization of the HEV circulating in this highly endemic country. Here, we conducted a longitudinal study to identify patients infected with HEV in Bangladesh to characterize the infecting viruses and assess the composition of mutations that may associate with a more severe disease progression.

## Materials and methods

### Study design and population

A prospective cross-sectional study was conducted from August 2013 to June 2015 at Bangabandhu Sheikh Mujib Medical University (BSMMU) in Dhaka. Patients attending at the virology department of BSMMU with a HEV IgM test request and tested positive were eligible for enrollment. Systematically selected (first four HEV IgM ELISA (Beijing Wantai Biological Pharmacy Enterprise Co., Ltd., Beijing, China) positive patients each month) leftover samples were stored at -86°C. Plasma samples were unlinked with patient data by replacing the patient ID with a study ID. Samples were shipped to Oxford University Clinical Research Unit (OUCRU), Vietnam for further analysis. The study was approved by the Bangabandhu Sheikh Mujib Medical University (BSMMU) ethical review committee (Approval No. BSMMU/2013/3027).

### Amplification and sequencing of HEV

HEV IgM positive samples were screened for HEV RNA using quantitative PCR (qPCR) [24]. Total viral RNA was extracted from 140ul of plasma using a MagNa pure viral RNA extraction kit (Roche Diagnostics, Basel). RNA was reverse transcribed into cDNA using superscript III reverse transcriptase. cDNA from the qPCR positive samples were used for amplification of the HEV genome by nested PCR. Degenerate oligonucleotide primers were designed to amplify 450–600 bp overlapping fragments to cover the entire HEV genome (S1 File). Additionally, where necessary the 5' and 3' cDNA ends were amplified using RACE PCR as previously described [25]. When required a primer walking strategy was implemented to close gaps in the genome sequences. PCR amplicons were visualized by agarose gel electrophoresis and purified using the QIAamp PCR purification kit (QIAGEN, Germany). Purified amplicons were sequenced in both directions using an Applied Biosystem 3130XL Genetic Analyzer. The entire genome sequences (amplified by primer set 1 to 20 (S1 File)) and a 279bp region (position 5972–6319 of GenBank accession no M73218) of ORF2 and ORF3 (part of the 440 bp fragment amplified by primer set 20 (S1 File)) were used for final analysis [26].

### Genome assembly and phylogenetics

To construct the HEV whole genome sequences the overlapping sequences from the PCR amplicons were aligned and a consensus sequence was generated. Additionally, 109 HEV genome sequences representing all genotypes and subtypes including 20 HEV genotype 1 sequences from Asia were accessed from GenBank (S2 File). Sequences generated during this study and the reference sequences were subjected to phylogenetic analysis using Geneious

8.0.5 software [27]. The entire genome sequences and a 279bp fragment of ORF2 and ORF3 were aligned using MUSCLE alignment program within Geneious [28]. Sequence alignments were subjected to Jmodel test to identify the best model for phylogenetic analysis [29]. The suggested nucleotide substitution model (GTR+G+I) was used for phylogenetic analysis in RAxML v7.2.8 (available in Geneious package). To confirm the reliability of phylogenetic tree, bootstrap resampling and reconstruction were repeated on 500 occasions.

Phylogenetic analysis of the whole genome sequences was performed in a two-step process. Firstly, 21 whole genome sequences from the present study and 109 existing genome sequences (including two outbreak sequences from Bangladesh) were analyzed. Secondly, 20 genotype 1a genome sequences from Asia (Bangladesh, India, Nepal, Pakistan, and Myanmar) and 21 genotype 1a HEV genome sequences from the present study were analyzed. All sequences were submitted to GenBank. Accession no; (MK005535 to MK005551) for the ORF2-3 region, and (MH991993 to MH992013) for the whole genome sequences.

### Virulence associated mutation analysis

Known mutations associated with i) FHF, ii) AVR, iii) chronic hepatitis, iv) ribavirin treatment failure, and v) B and T (MHC I and II) cell epitope neutralization were identified in 21 HEV WGS. These mutations were: F179S, A317T, T735I, L1110F, V1120I, F1439Y, C1483W, and N1530T in ORF1 and P259S in ORF2 (FHF and AVH); v1213A in ORF1 (chronic hepatitis); Y1320H, G1634R/K, K1383N, D1384G, K1398R, V1479I, and Y1587F in ORF1 (ribavirin treatment failure); and A317T, T735I, L1110F, V1120I, and G1634K in ORF1 and L477T and L613 in ORF2 (B cell and T cell neutralization). The potential effect of T (MHCI and MHC II) cell and B cell epitope mutations on antigenicity were inferred with published amino acid sequences of wild-type T cell and B cell epitopes [15]; this included epitope T2-T5 in MHCI, T7-T12 in MHCII in T cell epitope, and epitope B2-B8 in B cell epitope.

### Statistical analysis

All data collected in this study were cleaned and entered into an SPSS database. Analysis was performed using Statistical Package for Social Science (SPSS) software (IBM SPSS Statistics 23, NY USA). Categorical variables were compared with Fisher's exact test and continuous variables by Student's t-test and Mann-Whitney U-test as appropriate. A  $p$ -value  $<0.05$  was considered to be statistically significant.

## Results

### Study population

From August 2013 to June 2015, 92 HEV IgM antibody positive plasma samples were collected (44 (47.8%) in 2013–2014 and 48 (52.2%) in 2014–2015). 53 (57.6%) samples were from female patients, including 24 (45.3%) from pregnant women. The median age of the patients was 25 (IQR; 21–30 years). Of these 92 patients 70 (76.1%) patients were HEV RNA positive (33 men and 37 women (14 pregnant and 23 non-pregnant women)). 84.6% (33/39) of the samples from men were HEV RNA positive compared to 69.8% (37/53) of samples from women ( $p = 0.08$ ; chi squared test). Among the women patients, 79.3% (23/29) non-pregnant women and 58.3% (14/24) pregnant women were HEV RNA positive. There was no difference in age, sex, HEV RNA positivity among patients recruited during 2013–2014 and 2014–2015 (S3 File).

## Sequencing of HEV

HEV RNA positive plasma samples were subjected to sequencing. A schematic diagram of the DNA amplicons sequenced for each sample is shown in [S4 File](#). At least one DNA fragment could be sequenced in all of the plasma samples. Ultimately, a 279bp fragment of ORF 2–3 (used for genotyping) or an entire genome sequence (when generated) were subjected to phylogenetic analysis. The final data set included 38 (17 from incomplete genome sequence and 21 from complete genome sequence) ORF2–3 sequences and 21 whole genome sequences (1 from pregnant women, 6 from non-pregnant women, and 14 from men).

## Genome assembly and phylogenetics

The average length of novel HEV genomes was 7,215bp (range; 7,202–7,221bp) excluding the 3' poly (A) tail. Each genome had a 17–28bp untranslated region at 5' end, a 66bp untranslated region at 3' end, and a poly A tail of variable length (8–33bp). The complete HEV genome contains 3 open reading frames (ORF); ORF1 was from 28 to 5,109 bp; ORF2 was from 5,147bp–7,129bp and; ORF3 was from 5,106bp–5,477bp. ORF2 starts after 41 bases of ORF1. ORF3 overlapped ORF1 by four bases and ORF2 by 333bp.

The predicted translated protein of ORF1 was 1,694 amino acids. The NTP-binding domains GVPGSGKS and DEAP in the putative helicase region and the GDD site was found in RDRP regions of all isolates. The ORF1 sequences were highly conserved, with 98.1% identity (min-max; 97.05–97.91%) at the DNA level and 99.5% (min-max; 98.75–100%) identity at the amino acid level. The predicted translated product of ORF2 was 660 amino acids. Identity within ORF2 was 98.6% (min-max; 98.44–97.81%) at a nucleotide level, and 99.5% (min-max, 98.34–100%) at the amino acid level. The translated product of ORF3 product was 124 amino acids. The nucleotide similarity of ORF3 was 99.38% (min-max 98.39–99.19%) and 99.23% (min-max, 97.56–100%) identical at amino acid level.

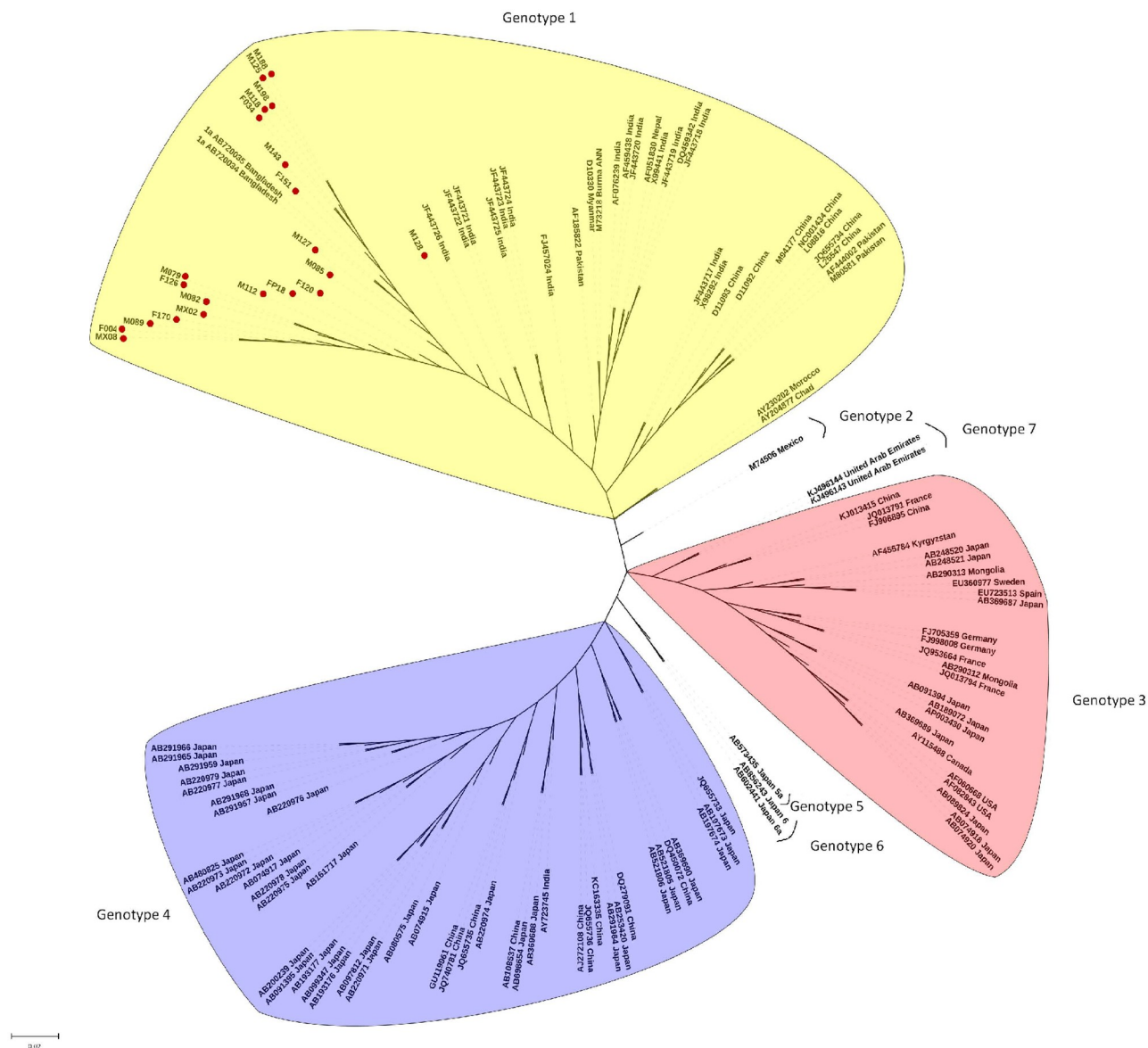
In a phylogenetic analysis of the 279 bp fragment of ORF2–3 of all 38 Bangladeshi HEV sequences (17 from incomplete genome sequence and 21 from complete genome sequence) from this study clustered in proximity to HEV genotype 1a sequences ([S5 File](#)). In a further phylogenetic analysis, the 21 whole genome sequences from this study clustered alongside genotype 1a HEV from India, Burma, and Nepal ([Fig 1](#)) and two near complete WGS from a 2010 outbreak in Rajshahi. Notably, the genotype 1 HEV sequences from China, Pakistan, and various African countries formed an independent cluster.

A subsequent phylogenetic analysis with whole genome sequences (n = 20) from counties neighboring Bangladesh found that genotype 1a sequences fell into two main clusters. The first cluster contained HEV from India, Nepal, and Myanmar, the second cluster included the 21 HEV from the present study and two HEV sequence from a previous outbreak in northern Bangladesh ([Fig 2](#)); these HEV DNA sequences were >99% identical across the genome.

## Virulence associated mutation analysis

Mutations associated with severe disease in 21 HEV WGS from this study are shown in [Table 1](#). All HEV sequences harbored the A317T, T735I, and L1120I mutations in ORF1; 80.9% had an L1110F mutation in ORF1 and a P259S mutation in ORF2; these are known to be associated with FHF and AVH. None of sequences had a V1213A mutation associated with chronic infection. The ribavirin resistance-associated mutations V1479I and G1634K were also detected in all sequences. Conversely, none of the sequences had the Y1320H, K1383N,

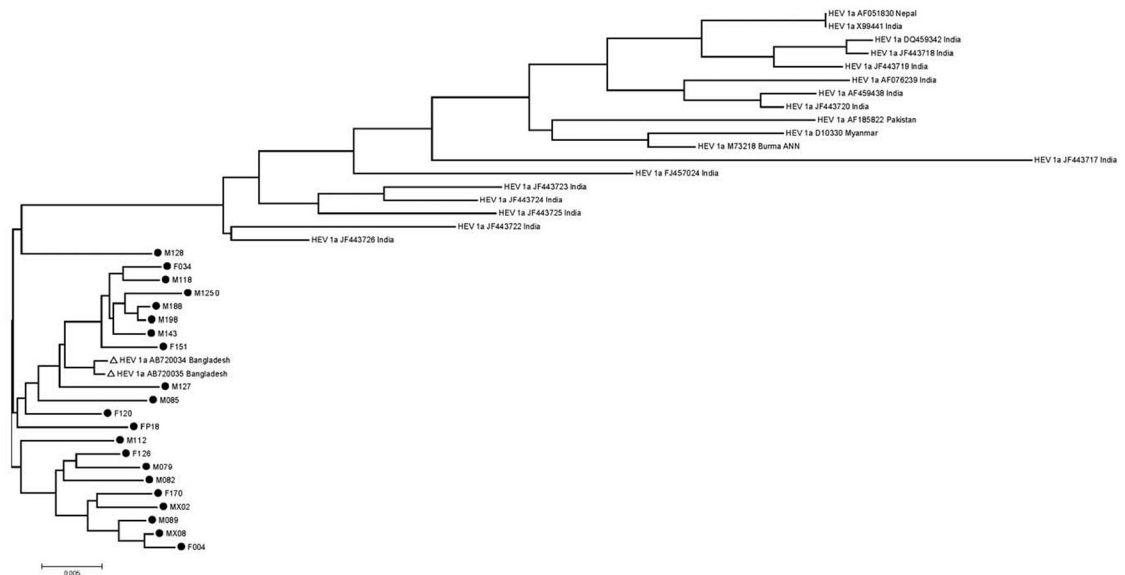




**Fig 1. Phylogenetic analysis of 109 reference HEV genome sequence and 21 Bangladeshi endemic HEV genome sequences from present study.** A midpoint rooted tree showing the relationship between the 21 Bangladeshi HEV genome sequences with 109 reference sequences representing all genotype and subtypes. The tree was constructed using RAxML v7.2.8 available in Geneious software using GTR+G+I nucleotide substitution model with 500 bootstrapping replicates. The Bangladeshi HEV strains from this study is presented as followed by strain number, and the reference genomes are presented as genotype, subtype followed by GenBank accession number and the country of origin. The scale bar indicates the number of nucleotide substitution.

<https://doi.org/10.1371/journal.pone.0255054.g001>

D1384G, K1398N, and Y1587F mutations. With respect to immune escape mutations, all sequences had the L477T mutation associated with the B and T cell immune reactivity (Table 2). Additionally, all viruses had mutations associated with altered T cell antigenicity (MCHI and MCHII) and B cell epitopes, exceptions were the T3 and B3 epitopes where 76.1% (16/21) of the HEV had mutations in these loci (Table 2). Mutations in T cell recognition epitope (T3, T5, and T7) with the highest reduction antigenicity were observed in 76.1%, 100%, and 100% of the HEV sequences, respectively (Table 2).



**Fig 2. Phylogenetic analysis of 20 reference HEV genome sequence from Bangladesh and neighboring countries and 21 Bangladeshi endemic HEV genome sequences from present study.** A midpoint rooted tree showing the relationship between the 21 Bangladeshi HEV genome sequences with 20 genotype 1a reference sequences from neighboring countries. The tree was constructed using RAxML v7.2.8 available in Geneious software using GTR+G+I nucleotide substitution model with 500 bootstrapping replicates. The Bangladeshi HEV strains from this study are presented as followed by strain number, and reference genomes are presented as genotype, subtype followed by GenBank accession number and the country of origin. The scale bar indicates the number of nucleotide substitution.

<https://doi.org/10.1371/journal.pone.0255054.g002>

## Discussion

HEV outbreaks are commonly undetected in LMICs due to a lack of disease surveillance and/or diagnostic capacity. Here, we analyzed samples collected systematically over a two year period to characterize endemic HEV genotypes in Bangladesh and compare with outbreak HEV genotypes. To our knowledge, this is the first report regarding the molecular characterization and WGS of endemic HEV strains from Bangladesh. Although there was no outbreak reported during the study period, the possibility of unidentified and/or unreported outbreak cannot be excluded. Smaller outbreaks often remain unreported due to lack of active surveillance in Bangladesh [30]; therefore, it is possible that viral strains characterized in this study may include outbreak strains.

Among the HEV IgM positive patients, 70% were HEV RNA positive. This discrepancy may be due to several factors including i) low viral load, ii) poor storage condition of the sample, or iii) the sensitivity of the qPCR assays for HEV isolates from Bangladesh. The prevalence of RNA positivity was higher in male patients compared to female (however, the difference was not significant) and was lowest in pregnant women. It is possible that male patients visited the hospital during active infection. In Bangladesh, men are main earning member in the family, and they seek healthcare earlier to minimize the wage loss due to sickness. Pregnant women in Bangladesh usually seek health care from “Maternal and Child Health (MCH)” clinics and referred to tertiary facilities when conditions worsen. This delay may result in lower RNA positivity in pregnant women.

All viral sequence in the present study were genotype 1a. When compared with genotype 1a strains from India, Nepal, Pakistan, and Myanmar, the Bangladeshi HEV strains formed a separate cluster within the genotype 1a cluster. HEV strains from the present study (and from the 2010 outbreak from northern Bangladesh) had high amino acid and DNA sequence similarity.

Table 1. Mutations associated with FHF, chronic hepatitis, ribavirin treatment failure and B and T cell epitope neutralization in 21 HEV WGS isolated in this study.

Isolate No	Fulminant hepatic failure										Chronic hepatitis	Ribavirin treatment failure										Epitope neutralization		
	ORF1											ORF2	ORF1										ORF2	
	MT	Y-Dom	HVR	Helicase	RDRP			RDRP					Helicase	RDRP					RDRP					
F179S	A317T	T735I	L1110F	V1120I	F1439Y	C1483W	N1530T	P259S	V1213A	ORF1	ORF2	Y1320H	K1383N	D1384G	K1398N	V1479I	Y1587F	G1634K	L477T	L613T				
% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)	% (n)			
0 (0)	100 (21)	100 (21)	81.8 (17)	100 (21)	0 (0)	0 (0)	0 (0)	100 (21)	0 (0)	0 (0)	S	0 (0)	0 (0)	0 (0)	0 (0)	100 (21)	0 (0)	100 (21)	100 (21)	0 (0)				
BDHEV_F004	T	I	I	I				S	NA	ORF1	S					I		K	T					
BDHEV_F034	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_F120	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_F126	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_F151	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_F170	T	I	I	I	I			S	NA	ORF1	S					I		K	T					
BDHEV_P018	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M079	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M082	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M085	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M089	T	I	I	I	I			S	NA	ORF1	S					I		K	T					
BDHEV_M112	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M118	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M125	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M127	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M128	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M143	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M188	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_M198	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_MX02	T	I	I	F	I			S	NA	ORF1	S					I		K	T					
BDHEV_MX08	T	I	I	I	I			S	NA	ORF1	S					I		K	T					

MT: Methyl transferases; Y-Dom: Y domain; HVR: Hyper variable region; RDRP: RNA dependent RNA polymerase; NA: Not applicable.

<https://doi.org/10.1371/journal.pone.0255054.t001>



Table 2. Antigenicity evaluation of wild-type and mutated epitopes of 21 HEV strains from this study.

No of isolate % (N)	Region	Epitope number	Position	Predicted epitope aa	Antigenicity	Actual epitopes aa	Antigenicity
<b>MHC I</b>							
100 (21)	ORF1	T2	T735I	ATPTPAAPL	0.2	ATPT/IPAAPL	0.2
76.1 (16)	ORF1	T3	L1110F	TTSRVLRSL	0.4	TTSRVL/FRSL	-0.6
100 (21)	ORF1	T5	G1634K	AVSDFLRGL	0.4	AVSDFLRG/KL	-0.6
<b>MHC II</b>							
100 (21)	ORF1	T7	A317T	FHAVPAHIW	0.5	FHAVPA/THIW	-0.7
100 (21)		T8	T735I	IPSRAATPT	0.2	IPSRAATPT/I	1
100 (21)		T9	V1120I	FWGEPAVGQ	0.2	FWGEPAV/IGQ	0.2
100 (21)		T10	V1479I	LGLECAVME	0.8	LGLECAV/IME	0.7
100 (21)		T11	G1634K	LRGLTNVAQ	0.7	LRG/KLTNVAQ	0.7
100 (21)	ORF2	T12	L477T	WLSLLAAEY	1.1	WLSLL/TAAEY	1.7
<b>B cell epitopes</b>							
100 (21)	ORF1	B2	T735I	TPAAPLPSP	0.1	T/IPAAPLPSP	0.1
76.1 (16)		B3	L1110F	TTSRVLRSL	0.1	TTSRVL/FRSL	0.01
100 (21)		B4	V1120I	WGEPAVGQK	0.8	WGEPAV/IGQK	1
100 (21)		B6	G1634K	SDFLRKLTN	0.1	SDFLRG/KLTN	-0.1
100 (21)	ORF2	B7	L477T	SLTAAEYDQ	1.1	SLL/TAAEYDQ	1.0
100 (21)		B8	L613T	LLDYPARAH	0.4	L/TLDYPARAH	0.4

The effect of mutation and antigenicity was adopted from Ikram et al., Bolt letters indicate mutations in the predicted epitopes.

<https://doi.org/10.1371/journal.pone.0255054.t002>

We suspect that these HEV strains have entered endemic circulation and associated with sustained infections and HEV outbreaks. The exact etiology of outbreak or epidemic is beyond our investigation, but could be associated with various host or environmental factors. Most of the HEV sequences in our study harbored FHF and AVH associated mutations. It is possible that the majority of patients had severe infections as severely ill patients are generally referred to tertiary hospitals. These mutations have also been reported in HEV strains from FHF patients in India [16]. The clustering of Bangladeshi HEV strains among genotype 1a might be due possible evolution and host genetic adaptation of these viruses in this population [5].

The predicted B and T cells epitopes overlap in HEV, and these mutations might alter (increase or decrease) the antigenicity of the HEV. Mutations with highest antigenicity alteration in MHC class I (T3 and T5) and MCH II (T7) were present in most of the HEV in our study. It is possible that this is a strategy adopted by HEV to escape the host immune response. For MHC I, two immune reactivity related mutations (L477T and L613T) overlap with B and T cell epitopes. Generally, antibody recognizing conformational epitopes are neutralizing, and the amino acid residues L477 and L613 in the capsid protein might play a role in epitope deformation [31]. HEV specific T cell targets are generally conserved and located in the capsid protein in ORF2. As T cell response persists for a longer period, these mutations can influence both clearance of virus after primary infection and protective immune response against secondary infection [32]. The presence of G1634K mutation affects the ribavirin treatment response and a V1479I mutation is associated with ribavirin treatment failure. All HEV in our

study had a G1634K and a V1479I mutation. Therefore, it is possible that HEV patients in Bangladesh treated with ribavirin might respond poorly [17,18,33].

This study has several limitations; the samples were collected from patients attending a single tertiary care facility in Dhaka and may not be representative of the general population of Bangladesh. Patients with mild or moderate infection usually seek care at the local hospitals or private practitioners. As the samples were unlinked, clinical condition of the patients and disease outcome is unknown. We collected only 4 samples per months, increasing the sample size might have captured more variations present in the HEV strains in Bangladesh. Although there was no reported outbreak during the study period, the possibility of unreported outbreaks cannot be excluded.

## Conclusions

Taken together, our data show that strains of HEV genotype 1a are dominant in Bangladesh and are associated with endemic and outbreak of HEV infection. HEV isolates in Bangladesh have high prevalence of virulence associated and ribavirin resistant mutations which may influence treatment with Ribavirin. HEV isolates in our study contains mutation which alters antigenicity to B and T cell epitopes and facilitates evading host immune response.

## Supporting information

**S1 File. List of primers used for whole genome amplification in this study.**  
(DOCX)

**S2 File. Accession number, genotype and subtype, and country of origin for 109 Hepatitis E virus used as representative reference strain.**  
(DOCX)

**S3 File. Sociodemographic characteristics, HEV RNA positivity and HEV WGS information 92 patient samples analyzed in this study.**  
(DOCX)

**S4 File. Schematic presentation of the sequencing of 92 HEV IgM ELISA positive samples.** Among 92 samples, 70 were qPCR positive (column 3), qPCR negative samples are presented in transparent rows. A total of 825 fragments (column 28) could be amplified and sequenced (dark block) (range 1–24). Whole genome sequence could be obtained for 21 HEV viral strain. Fragment 20 (nucleotide sequence 5822 to 6211) contains the ORF 2–3 overlapping section could be sequence in 38 (21+ 17) strain.  
(DOCX)

**S5 File. Phylogenetic analysis of 109 reference HEV genome sequence and 279bp fragment (ORF 2 and ORF) of 38 Bangladeshi endemic HEV sequence from present study.** A mid-point rooted tree showing the relationship between the 38 Bangladeshi HEV sequences with 109 reference sequences representing all genotype and subtypes. The tree was constructed using RAxML v7.2.8 available in Geneious software using GTR+G+I nucleotide substitution model with 500 bootstrapping replicates. The Bangladeshi HEV strains from this study is presented as ● followed by strain number, and the reference genomes are presented as genotype, subtype followed by GenBank accession number and the country of origin. The scale bar indicates the number of nucleotide substitution.  
(DOCX)

## Acknowledgments

We thank OUCRU laboratory management staff for supporting the sequencing of HEV strains.

## Author Contributions

**Conceptualization:** Saif Ullah Munshi, Tahmina Akther, Shahina Tabassum, Nilufa Parvin, Stephen Baker, Motiur Rahman.

**Data curation:** Trang Nguyen Hoa, Chau Le Ngoc, Thanh Tran Thi Thanh.

**Formal analysis:** Trang Nguyen Hoa, Khanh Nguyen Ngoc, Thanh Tran Thi Thanh.

**Funding acquisition:** Motiur Rahman.

**Investigation:** Trang Nguyen Hoa, Saif Ullah Munshi, Khanh Nguyen Ngoc, Chau Le Ngoc, Nilufa Parvin.

**Methodology:** Trang Nguyen Hoa, Saif Ullah Munshi, Khanh Nguyen Ngoc, Chau Le Ngoc, Thanh Tran Thi Thanh.

**Project administration:** Trang Nguyen Hoa, Saif Ullah Munshi, Tahmina Akther, Shahina Tabassum, Motiur Rahman.

**Resources:** Thanh Tran Thi Thanh, Shahina Tabassum, Nilufa Parvin, Stephen Baker, Motiur Rahman.

**Supervision:** Thanh Tran Thi Thanh, Tahmina Akther, Nilufa Parvin, Stephen Baker.

**Validation:** Khanh Nguyen Ngoc, Chau Le Ngoc, Shahina Tabassum.

**Writing – original draft:** Stephen Baker, Motiur Rahman.

**Writing – review & editing:** Trang Nguyen Hoa, Saif Ullah Munshi, Khanh Nguyen Ngoc, Chau Le Ngoc, Thanh Tran Thi Thanh, Tahmina Akther, Shahina Tabassum, Nilufa Parvin, Stephen Baker, Motiur Rahman.

## References

1. Rein DB, Stevens GA, Theaker J, Wittenborn JS, Wiersma ST: The global burden of hepatitis E virus genotypes 1 and 2 in 2005. *Hepatology* 2012, 55(4):988–997. <https://doi.org/10.1002/hep.25505> PMID: 22121109
2. Webb GW, Dalton HR: Hepatitis E: an underestimated emerging threat. *Ther Adv Infect Dis* 2019, 6:2049936119837162. <https://doi.org/10.1177/2049936119837162> PMID: 30984394
3. Global Hepatitis Report 2017. In: World Health Organization; 2017.
4. Kmush BL, Labrique A, Li W, Klein SL, Schulze K, Shaikh S, et al: The Association of Cytokines and Micronutrients with Hepatitis E Virus Infection During Pregnancy and the Postpartum Period in Rural Bangladesh. *Am J Trop Med Hyg* 2016, 94(1):203–211. <https://doi.org/10.4269/ajtmh.15-0238> PMID: 26621563
5. Smith DB, Purdy MA, Simmonds P: Genetic variability and the classification of hepatitis E virus. *J Virol* 2013, 87(8):4161–4169. <https://doi.org/10.1128/JVI.02762-12> PMID: 23388713
6. Ahmad I, Holla RP, Jameel S: Molecular virology of hepatitis E virus. *Virus Res* 2011, 161(1):47–58. <https://doi.org/10.1016/j.virusres.2011.02.011> PMID: 21345356
7. Holla RP, Ahmad I, Ahmad Z, Jameel S: Molecular virology of hepatitis E virus. *Semin Liver Dis* 2013, 33(1):3–14. <https://doi.org/10.1055/s-0033-1338110> PMID: 23564385
8. 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

9. Huang FF, Sun ZF, Emerson SU, Purcell RH, Shivaprasad HL, Pierson FW, et al: Determination and analysis of the complete genomic sequence of avian hepatitis E virus (avian HEV) and attempts to infect rhesus monkeys with avian HEV. *J Gen Virol* 2004, 85(Pt 6):1609–1618. <https://doi.org/10.1099/vir.0.79841-0> PMID: 15166445
10. Hakim MS, Wang W, Bramer WM, Geng J, Huang F, de Man RA, et al: The global burden of hepatitis E outbreaks: a systematic review. *Liver Int* 2017, 37(1):19–31. <https://doi.org/10.1111/liv.13237> PMID: 27542764
11. Boxall E, Herborn A, Kochethu G, Pratt G, Adams D, Ijaz S, et al: Transfusion-transmitted hepatitis E in a 'nonhyperendemic' country. *Transfus Med* 2006, 16(2):79–83. <https://doi.org/10.1111/j.1365-3148.2006.00652.x> PMID: 16623913
12. Cooper K, Huang FF, Batista L, Rayo CD, Bezanilla JC, Toth TE, et al: Identification of genotype 3 hepatitis E virus (HEV) in serum and fecal samples from pigs in Thailand and Mexico, where genotype 1 and 2 HEV strains are prevalent in the respective human populations. *J Clin Microbiol* 2005, 43(4):1684–1688. <https://doi.org/10.1128/JCM.43.4.1684-1688.2005> PMID: 15814985
13. Huang F, Li Y, Yu W, Jing S, Wang J, Long F, et al: Excretion of infectious hepatitis E virus into milk in cows imposes high risks of zoonosis. *Hepatology* 2016, 64(2):350–359. <https://doi.org/10.1002/hep.28668> PMID: 27286751
14. Khuroo MS, Khuroo MS, Khuroo NS: Transmission of Hepatitis E Virus in Developing Countries. *Viruses* 2016, 8(9). <https://doi.org/10.3390/v8090253> PMID: 27657112
15. Ikram A, Hakim MS, Zhou JH, Wang W, Peppelenbosch MP, Pan Q: Genotype-specific acquisition, evolution and adaptation of characteristic mutations in hepatitis E virus. *Virulence* 2018, 9(1):121–132. <https://doi.org/10.1080/21505594.2017.1358349> PMID: 28727933
16. Mishra N, Walimbe AM, Arankalle VA: Hepatitis E virus from India exhibits significant amino acid mutations in fulminant hepatic failure patients. *Virus Genes* 2013, 46(1):47–53. <https://doi.org/10.1007/s11262-012-0833-7> PMID: 23054435
17. Todt D, Gisa A, Radonic A, Nitsche A, Behrendt P, Suneetha PV, et al: In vivo evidence for ribavirin-induced mutagenesis of the hepatitis E virus genome. *Gut* 2016, 65(10):1733–1743. <https://doi.org/10.1136/gutjnl-2015-311000> PMID: 27222534
18. Debing Y, Ramiere C, Dallmeier K, Piorkowski G, Traubad MA, Lebosse F, et al: Hepatitis E virus mutations associated with ribavirin treatment failure result in altered viral fitness and ribavirin sensitivity. *J Hepatol* 2016, 65(3):499–508. <https://doi.org/10.1016/j.jhep.2016.05.002> PMID: 27174035
19. Talukdar PK, Rahman M, Rahman M, Nabi A, Islam Z, Hoque MM, et al: Antimicrobial resistance, virulence factors and genetic diversity of *Escherichia coli* isolates from household water supply in Dhaka, Bangladesh. *PLoS One* 2013, 8(4):e61090. <https://doi.org/10.1371/journal.pone.0061090> PMID: 23573295
20. Ayesha Begum S T, Md. Nazrul Islam, Afzalunnessa: Seroprevalence of Hepatitis E Virus (HEV) Infection Among Patients Attending BSMMU, Dhaka. *Bangladesh J Med Microbiol* 2007, 01(02):52–55.
21. Labrique AB, Zaman K, Hossain Z, Saha P, Yunus M, Hossain A, et al: An exploratory case control study of risk factors for hepatitis E in rural Bangladesh. *PLoS One* 2013, 8(5):e61351. <https://doi.org/10.1371/journal.pone.0061351> PMID: 23675410
22. Sheikh A, Sugitani M, Kinukawa N, Moriyama M, Arakawa Y, Komiyama K, et al: Hepatitis e virus infection in fulminant hepatitis patients and an apparently healthy population in Bangladesh. *Am J Trop Med Hyg* 2002, 66(6):721–724. <https://doi.org/10.4269/ajtmh.2002.66.721> PMID: 12224580
23. Gurley ES, Hossain MJ, Paul RC, Sazzad HM, Islam MS, Parveen S, et al: Outbreak of hepatitis E in urban Bangladesh resulting in maternal and perinatal mortality. *Clin Infect Dis* 2014, 59(5):658–665. <https://doi.org/10.1093/cid/ciu383> PMID: 24855146
24. Jothikumar N, Cromeans TL, Robertson BH, Meng XJ, Hill VR: A broadly reactive one-step real-time RT-PCR assay for rapid and sensitive detection of hepatitis E virus. *J Virol Methods* 2006, 131(1):65–71. <https://doi.org/10.1016/j.jviromet.2005.07.004> PMID: 16125257
25. Harun-Or-Rashid M, Akbar SM, Takahashi K, Al-Mahtab M, Khan MS, Alim MA, et al: Epidemiological and molecular analyses of a non-seasonal outbreak of acute icteric hepatitis E in Bangladesh. *J Med Virol* 2013, 85(8):1369–1376. <https://doi.org/10.1002/jmv.23601> PMID: 23703666
26. Lu L, Li C, Hagedorn CH: Phylogenetic analysis of global hepatitis E virus sequences: genetic diversity, subtypes and zoonosis. *Rev Med Virol* 2006, 16(1):5–36. <https://doi.org/10.1002/rmv.482> PMID: 16175650
27. Smith DB, Simmonds P, Izopet J, Oliveira-Filho EF, Ulrich RG, John R, et al: Proposed reference sequences for hepatitis E virus subtypes. *J Gen Virol* 2016, 97(3):537–542. <https://doi.org/10.1099/jgv.0.000393> PMID: 26743685
28. Smith DB, Bukh J, Kuiken C, Muerhoff AS, Rice CM, Stapleton JT, et al: Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource. *Hepatology* 2014, 59(1):318–327. <https://doi.org/10.1002/hep.26744> PMID: 24115039

29. Posada D: jModelTest: phylogenetic model averaging. *Molecular biology and evolution* 2008, 25(7):1253–1256. <https://doi.org/10.1093/molbev/msn083> PMID: 18397919
30. Sazzad HM, Labrique AB, Teo CG, Luby SP, Gurley ES: Surveillance at Private Laboratories Identifies Small Outbreaks of Hepatitis E in Urban Bangladesh. *Am J Trop Med Hyg* 2017, 96(2):395–399. <https://doi.org/10.4269/ajtmh.16-0411> PMID: 27994104
31. 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(6):447–456. <https://doi.org/10.1038/cmi.2008.56> PMID: 19118511
32. Brown A, Halliday JS, Swadling L, Madden RG, Bendall R, Hunter JG, et al: Characterization of the Specificity, Functionality, and Durability of Host T-Cell Responses Against the Full-Length Hepatitis E Virus. *Hepatology* 2016, 64(6):1934–1950.
33. Debing Y, Gisa A, Dallmeier K, Pischke S, Bremer B, Manns M, et al: 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(5):1008–1011 e1007; quiz e1015–1006. <https://doi.org/10.1053/j.gastro.2014.08.040> PMID: 25181691