Infection, Genetics and Evolution 18 (2013) 42–47

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Analysis of the complete genome sequences of one swine and two human hepatitis E virus genotype 4 strains isolated in Beijing, China

Hongxia Ma [a,b,1], Yansheng Geng [a,c,1], Zhuo Li [d], Tim J. Harrison [e], Weijin Huang [a], Chenyan Zhao [a], Youchun Wang [a,*]

[a] Department of Cell Biology, National Institutes for Food and Drug Control, No. 2, Tiantanxili, Beijing 100050, China
[b] Department of Prevention and Control of Infectious Disease, Henan Centers for Disease and Prevention, No. 105, Nong Ye Dong Lu, Zhengzhou 450016, China
[c] Health Science Center, Hebei University, No. 342, Yuhuadonglu, Baoding 071000, China
[d] Institute of Hepatology, Youan Hospital, Youan Street, Beijing, China
[e] Division of Medicine, University College London Medical School, Cruciform Building, 90 Gower Street, London WC1E 6BT, UK

ARTICLE INFO

Article history:
Received 12 December 2012
Received in revised form 26 April 2013
Accepted 29 April 2013
Available online 16 May 2013

Keywords:
Hepatitis E virus
Swine HEV
Complete genome
Sequence analysis
Zoonotic transmission

ABSTRACT

Full-length sequences were determined and analyzed for two human (MO and W3) and one swine (W2–5) hepatitis E virus (HEV) isolates from Beijing, China. The genomes of the three strains were composed of 7242, 7239, 7239 nucleotides, respectively, excluding the poly (A) tails, and were 84% identical to each other. All were classified into genotype 4. Sequence analysis shows that the 2 human isolates have up to 91–94% nucleotide identity in full length genome with swine strains isolated in China, while the swine isolate share 92% identity with the human strain T1 from Beijing. At the amino acid level, the three strains share 94%, 97% and 89–92% identity in the ORF1, ORF2 and ORF3, proteins respectively. The human strains MO and W3 have the highest identity, 97%, 98–99% and 96–98% in ORFs 1–3, respectively, to swine strains CHN-XJ-SW13 and CHN-XJ-SW33 from Xinjiang, China, while swine strain W2–5 has highest identity with the human strain HE-JA2, 96%, 99% and 91% in ORFs 1–3, respectively. Genotype specific amino acid substitutions were found at a single site in all three ORFs by sequences alignment, and genotype specific short sequences (5–10aa in length) were found in ORF1 and the C-terminus of ORF3. However, no difference was found at any amino acid position that discriminates between human and swine HEVs within genotype 4 for any of the three ORFs. These results indicated that the genotype 4 HEV strains from humans and pigs in China may evolve from the common ancestor.

1. Introduction

Hepatitis E is an important public health problem in many developing countries. Sporadic cases of hepatitis E also have been reported increasingly in developed countries (Meng, 2010; Pavio et al., 2010). The causative agent of hepatitis E is hepatitis E virus (HEV). HEV belongs to the *Hepeviridae* family (Meng et al., 2011). It is a small, non-enveloped virus with a single-stranded, positive-sense RNA genome of approximately 7.2 kb containing a short 5′ untranslated region (UTR), three open reading frames (ORFs), ORF1, ORF2, ORF3 and a short 3′ UTR that is terminated by a poly(A) tract (Tam et al., 1991).

Since swine HEV was first isolated and genetically characterized from pigs in the USA in 1997, new HEV variants have been found in chickens (Haqshenas et al., 2001), rabbits (Zhao et al., 2009), rats (Johne et al., 2010), wild boar (Takahashi et al., 2011) and, most recently, ferrets (Raj et al., 2012) and bats (Drexler et al., 2012). HEV isolates have been divided into at least four major mammalian genotypes (1–4) (Meng et al., 2011): Genotypes 1 and 2 have been identified exclusively in humans and are mainly associated with large waterborne epidemics and sporadic cases of hepatitis E in Asia, Africa, and North America; genotype 3 and 4 have distributed worldwide, which been recognized as a zoonotic pathogen and are mainly responsible for sporadic cases of hepatitis E (Meng, 2010; Pavio et al., 2010).

Accumulating lines of evidence indicate that, in some cases involving HEV genotypes 3 and 4, animal to human transmissions occur (Takahashi et al., 2004; Li et al., 2005; Masuda et al., 2005). Pigs stand out as the major reservoir of genotype 3 and 4 HEV infections in humans, as suggested by the close genetic relationship of swine and human viruses (Meng, 2010; Pavio et al., 2010). Cross-species infections with HEV genotypes 3 and 4 also have been demonstrated experimentally (Arankalle et al., 2006; Feagins et al., 2008; Meng, 2010). However, not all sources of human infections have been identified thus far and in many cases, the origin of HEV infection in humans remains unknown. In some

* Corresponding author. Tel.: +86 10 67095921; fax: +86 10 67053754.
E-mail address: wangyc@nifdc.org.cn (Y. Wang).
[1] These authors are contributed equally to this work.

1567-1348/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2013.04.038

reports, the strains prevalent in the human populations were genetically distinct from those prevalent in swine in same geographic areas, suggesting that cross-species transmission of HEV between human and swine might not occur in these areas (Zhang et al., 2009; Purdy et al., 2012). Therefore, more evidence is needed to determine the risk of swine HEV being transmitted to humans.

In this study, the complete genomic sequences of one swine HEV strain, W2-5, and two human population HEV strains, MO and W3, were amplified, sequenced and compared to the known HEV strains with complete genome sequence representing genotypes1–4, rabbit HEV and avian HEV retrieved from GenBank. Sequence alignments and phylogenetic analysis at the nucleotide and amino acid levels were performed to characterize and identify the relationship between isolates from humans and swine in China.

2. Materials and methods

2.1. Specimens and extraction of RNA

Isolates W3 and MO were amplified from fecal specimens collected from patients with acute hepatitis E in Beijing Youan Hospital, China between May and June, 2006. The study was approved by the ethics committee of Beijing Youan Hospital and the patients gave written informed consent for the use of their samples. Isolate W2–5 was obtained from liver tissue collected in August, 2006 from a domestic pig positive for anti-HEV in a pig farm located in a southern suburb of Beijing. The ground liver tissue and fecal samples were diluted in phosphate buffered saline (PBS; pH 7.4) containing 1% bovine serum albumin (BSA) to make a 10% (wt/vol) suspension and clarified by centrifugation at 3000 rpm at room temperature for 10 min. The supernatants were aliquotted and stored under liquid nitrogen until use. Total RNA was extracted from 140 μL of the 10% supernatant using a QIAamp viral RNA mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

2.2. Reverse transcription and polymerase chain reactions

cDNA was synthesized from 7 μL of RNA by reverse transcription at 50 °C for 50 min with a Superscript III First-Strand Synthesis System for RT-PCR Kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instruction and the first-strand cDNA was used immediately for PCR. A nested or semi-nested polymerase chain reaction (PCR) with 4 sets of specific external and internal primer pairs per strain (Supplementary Table 1) was used to amplify the entire genome, the 5′ and 3′ ends were amplified using the 5′ rapid amplification of cDNA ends (RACE) and a 3′ RACE kit (Invitrogen, Carlsbad, CA, USA). Polymerase chain reactions used in this study were processed as described previously (Zhao et al., 2009).

The expected PCR products were cut out from 2% or 1% agarose gels, purified with a QIAquick Gel Extraction Kit (QIAGEN, Hilden, Germany) and cloned into the pGEM-T easy vector (Promega, Madison, WI). The sequencing of the recombinant plasmids with target PCR fragments was performed commercially by Sangon Biotech Co., Ltd (Shanghai, China) using a BigDye terminator v3.1 kit and ABI-PRISM3730XL DNA sequenceing (Applied Biosystems, Foster City, CA).

2.3. Sequence analysis

The full-length genomes of MO, W3 and W2–5 were acquired by assembling the fragments using DNAstar 5.01 software and the three full-length genomic sequences have been deposited in the GenBank database with accession numbers JQ655733, JQ655735 and JQ655736, respectively. Nucleotide sequences of HEV and amino acid sequences of HEV ORF1-3 were analyzed and compared to the known HEVs from GenBank (Supplementary Table 2) using ClustalW2, MEGA 5.0 and DNAstar 5.0 software.

3. Results

3.1. Genome organization

The full length genomes of the MO, W3 and W2–5 strains were determined to be 7242, 7239, 7239 nucleotides, excluding the 49nt, 18nt, 22nt poly (A) tail (Table 1), respectively. The genomic organization consisted of a 5′ UTR successively covering nucleotides 1–27, 1–25 and 1–26, ORF1 extending from 28 to 5148 (5121nt), 26 to 5149 (5121nt) and 27 to 5147 (5118nt), ORF2 extending from 5187 to 7169 (1983nt), 5188 to 7170 (1983nt) and 5186 to 7168 (1983nt), the small third open reading frame ORF3 extending from 5173 to 5517 (345nt), 5174 to 5518 (345nt) and 5172 to 5516 (345nt), the 3′ UTR extending from 7170 to 7242 (73nt), 7171 to 7239 (69nt), 7169 to 7239 (71nt), respectively, and was followed by a poly (A) tail. BLAST analysis of these three isolates, based on the complete genome, showed 84.2–84.8% identity.

3.2. Phylogenetic analysis of the complete genomes

A phylogenetic tree was constructed, based on the entire genomic sequence of 52 HEV isolates and using an avian HEV as an outgroup (Fig. 1). Sequence analysis indicated that all three isolates were grouped into genotype 4 and MO, W2–5 and W3 fell into subtypes 4a, 4d and 4h, respectively. The human isolate, W3, was clustered with two swine strains swWH09 (GU188851) and CHN-XJ-SW13 (GU119961) and one human strain CHN-NJ-H2011 (JQ740781) as subgenotype 4h, sharing 93.9–94.2% sequence identity with them, the sequence similarity within this subgroup being 93.9–95.5%. The MO isolate was most closely related to the human isolate JYI-ChiSai01C (AB197674) of sub-genotype 4a, sharing a maximum 91.4% nucleotide sequence identity. Four other isolates, two of them from swine, were also classified as subtype 4a and MO had 90.2–91.1% homology with them. The swine W2-5 sequence clustered with six known strains, including five swine strains and one human strain T1 (AJ272180) as subtype 4d, sharing 89.1–92.1% sequence homology with them. Sequence analysis indicated that the swine HEV strain in the present study has higher sequence homology (92.1%) with the T1 human strain isolated from a patient in Beijing than with other isolates from swine.

3.3. Comparison analysis of amino acid sequences of ORF 1

The predicted ORF1 polypeptides of MO, W2–5 and W3 were 1706aa, 1707aa and 1707aa, respectively, in length. In this region,

Table 1
Genomic organization of HEV isolates.

ORFs		W3/JQ655735	W2-5/JQ655736	MO/JQ655733
5′ UTR	Position (nt)	1–25	1–26	1–27
	Length (bp)	25	26	27
ORF 1	Position (nt)	26–5149	27–5147	28–5148
	Length (bp)	5124	5121	5121
ORF 2	Position (nt)	5188–7170	5186–7168	5187–7169
	Length (bp)	1983	1983	1983
ORF 3	Position (nt)	5174–5518	5172–5516	5173–5517
	Length (bp)	345	345	345
3′ UTR	Position (nt)	7171–7239	7169–7239	7170–7242
	Length (bp)	69	71	73
Poly A	Position (nt)	7240–7257	7240–7261	7243–7291
	Length (bp)	18	22	49