Virus Research 144 (2009) 290–293

Contents lists available at ScienceDirect

Virus Research

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

Short communication

Full genomic sequence analysis of swine genotype 3 hepatitis E virus isolated from Shanghai

Fu-sheng Si[a,c], Yu-min Zhu[a,b], Shi-juan Dong[a,b], Shui-sheng Yu[a,b], Rui-song Yu[a,b], Shi-yuan Shen[a,b], Qian Yang[c], Zhen Li[a,b,*]

[a] Institute of Animal Science and Veterinary Medicine, Shanghai Academy of Agricultural Sciences, Shanghai 201106, China
[b] Shanghai Key Laboratory of Agricultural Genetics and Breeding, Shanghai 201106, China
[c] Veterinary College, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China

ARTICLE INFO

Article history:
Received 27 January 2009
Received in revised form 11 April 2009
Accepted 14 April 2009
Available online 23 April 2009

Keywords:
Hepatitis E virus
Genotype 3
Full genomic sequence
Phylogenetic analysis
Swine viral isolates

ABSTRACT

The full genomic nucleotide sequence of a previously identified genotype 3 hepatitis E virus (HEV), strain SAAS-JDY5, was obtained using RT-PCR and rapid amplification of cDNA ends (RACE). The genome consisted of 7225 nucleotides, excluding a poly-A tail at the 3′ terminus, and contained three open reading frames (ORFs), ORF-1, ORF-2 and ORF-3, encoding 1702, 660 and 113 amino acids, respectively. Phylogenetic analysis confirmed that SAAS-JDY5 belonged to genotype 3 HEV and was most closely related to the Japanese isolate wbJYG1 (AB222184). SAAS-JDY5 shared approximately 87% nucleotide similarity to human and swine strains from the United States, compared with 74–75% similarity to Asian (genotype 4) and Mexican strains (genotype 2). Alignment of the SAAS-JDY5 genomic sequence with reference sequences of the same genotype revealed one nucleotide substitution and one deletion at positions 5145 and 7189 (3′ UTR), respectively. Moreover, SAAS-JDY5 contained two additional nucleotides (AC) at the very end of the 3′-terminus preceding the poly-A tail of the genome. Comparison of the putative amino acid sequence encoded by the SAAS-JDY5 genome with sequences of other genotype 3 isolates revealed 15 unique amino acid substitutions and one deletion in ORF-1, and three substitutions in ORF-2.

Hepatitis E is an important public health problem in many developing countries of Asia and Africa, and also occurs sporadically in some industrialized countries. The disease mainly affects young adults and has a relatively high mortality (up to 25%) in affected pregnant women (Emerson and Purcell, 2003). Hepatitis E is transmitted primarily by the fecal–oral route, often through contaminated water (Purcell, 1996; Bradley, 1992).

Hepatitis E virus (HEV), a member of the genus Hepevirus, is a single-stranded, positive-sense RNA virus without an envelope (Purcell and Emerson, 2001). The genome of HEV is approximately 7.2 kb, and contains a short 5′ untranslated region (UTR), three open reading frames (ORFs), and a short 3′ UTR terminated by a poly(A) tract (Reyes et al., 1990). ORF-1 is located at the 5′ end of the genome and encodes non-structural proteins including a methyltransferase, a papain-like cysteine protease, RNA helicase, and RNA-dependent RNA polymerase (RdRp). ORF-2 maps to the 3′ terminus and encodes for a major structural protein (Tam et al., 1991; Purcell and Emerson, 2001). ORF-3, which overlaps with ORF-2 (Graff et al., 2006; Huang et al., 2007), encodes a smaller protein which interacts with the ORF2 protein (Tyagi et al., 2002) and a number of cellular signal transduction pathway proteins (Korkaya et al., 2001; Tyagi et al., 2004).

The first swine strain of HEV was isolated and characterized in the US (Meng et al., 1997), followed by similar findings in Japan (Okamoto et al., 2001; Nishizawa et al., 2003), Taiwan (Hsieh et al., 1999; Wu et al., 2002), New Zealand (Garkavenko et al., 2001), and mainland China (Wang et al., 2002). Other animals such as goats, cattle and rodents are also reported to be infected with HEV (Usmanov et al., 1994; Favorov et al., 1996, 2000; Maneerat et al., 1996; Tsarev et al., 1998). These studies have gradually confirmed that HEV is a zoonotic pathogen, and that pigs are a major reservoir (Zheng et al., 2006).

Based on sequence analysis, HEVs have been classified into four major genotypes (1–4). Genotype 1 is the main cause of hepatitis E in developing countries in Asia and Africa. Genotype 2 has been documented in Mexico and a number of African countries. Genotype 3 strains originated mainly from North America, Europe, New Zealand and some Asian countries including Japan, South Korea, Taiwan and Thailand, and genotype 4 has mainly been associated with Asian

* Corresponding author at: Institute of Animal Science and Veterinary Medicine, Shanghai Academy of Agricultural Sciences, Shanghai 201106, China.
Tel.: +86 21 62206391; fax: +86 21 62207858.
E-mail address: zhenli60@sh163.net (Z. Li).

0168-1702/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2009.04.009

countries including China, Japan, India and Vietnam (Schlauder and Mushahwar, 2001).

Although previous epidemiological research on HEV had reported that strains circulating in China belonged to genotypes 1 and 4 (Wang et al., 1999; Wei et al., 2006), our more recent studies have demonstrated the widespread existence of genotype 3 HEV strains in Shanghai pig farms (Ning et al., 2007; Ning et al., 2008). This finding has generated considerable interest in the origins, genomic characteristics and zoonotic nature of genotype 3 HEV in Shanghai. We now report the first full-length sequence of a genotype 3 swine HEV strain, SAAS-JDY5, from China.

The strain of HEV, designated SAAS-JDY5, was isolated from a fecal specimen collected from a pig farm located in a Shanghai suburb. The strain was identified as genotype 3 HEV by RT-PCR and sequencing of the ORF-2 150 bp fragment (Li et al., 2009). Viral RNA was extracted with TRIZOL reagent (Invitrogen, USA) in accordance with the manufacturer's protocol. The extracted RNA was dissolved in 10 μl RNase-free pure water for use in determining the full genomic sequence of the HEV strain.

Different primer sets were used for the amplification of the entire SAAS-JDY5 genomic sequence. First strand cDNA of the whole genomic RNA was synthesized with the SuperScript™ III First-strand Synthesis System (Invitrogen, USA). First round PCR was carried out using 5 μl of the synthesized cDNA and an external set of forward and reverse primers with Ex *Taq* DNA polymerase (TakaRa, Japan). Nested PCR was carried out with an internal primer set and 5 μl of the first PCR product. The resulting PCR products were excised from agarose gels, purified using the Axyprep DNA Gel Extraction Kit (Axygen, USA), and cloned into PMD18-T vector (TakaRa, Japan) using T4 DNA ligase (TakaRa, Japan). Recombinant plasmid DNA was transformed into *Escherichia coli* DH5α cells (TakaRa, Japan), and plasmids containing the insert fragment were identified by PCR. At least three of the positive clones were sequenced.

Amplification of the extreme 5′ end sequence was carried out using the SMART™ RACE cDNA Amplification Kit (Clontech Laboratories, Japan) according to the manufacturer's protocol. Briefly, extracted RNA was treated with 1 μl SMART ™ A Oligonucleotide (supplied with the kit) and used as the template to synthesize cDNA. First-strand cDNA was synthesized by reverse transcription using 5′ RACE CDS primer A (supplied with the kit) and the SuperScript™ III First-strand Synthesis System. This cDNA was then amplified by two rounds of PCR with Ex *Taq* DNA polymerase (TakaRa, Japan) using 10× Universal Primer A Mix and Nested Universal Primer A (supplied with the kit) as forward primers, and specific genotype 3 HEV primers 5′ EA and 5′ IA as reverse primers, respectively. Purified PCR products were TA-cloned and sequenced.

Amplification of the extreme 3′ end was also carried out with SMART™ RACE cDNA Amplification Kit (Clontech Laboratories, Japan). cDNA was synthesized by incubating 10 μl HEV RNA template solution with the SuperScript™ III First-strand Synthesis System at 50 °C for 50 min. Reactions were terminated by exposure to 85 °C for 5 min. The external reverse primer, 3′ RACE CDS primer A (supplied with the kit), which has a poly(T) tract, was used to prime the cDNA synthesis. The resultant cDNA was then amplified by two rounds of PCR using the 10× Universal Primer A Mix and Nested Universal Primer A (supplied with the kit) as reverse primers, and the specific HEV primers 3′ ES and 3′ IS as forward primers, respectively. The PCR reaction mixture and amplification conditions were the same as for 5′ RACE.

Sequence assembly was accomplished and percent identity was calculated using Lasergene (version 7.10; DNAstar). Sequence alignments were generated by CLUSTAL-W (version 1.8). Genetic distances between pairs of viral isolates were calculated with MEGA software (version 4.0) and the phylogenetic tree was constructed using the neighbor-joining method and an avian HEV strain (AY535004) (Huang et al., 2004) as the outgroup.

Nine RT-PCR fragments of SAAS-JDY5 genomic RNA were obtained, and the overlapping cDNA sequences were cloned and sequenced to construct the full-length genome. The full-length genome of SAAS-JDY5 consisted of 7225 nucleotides (nt), excluding the poly(A) tail at the 3′ terminus, and contained three major ORFs. The genome comprised a 5′ UTR of 26 nt (1–26), ORF-1 of 5109 nt (27–5135), ORF-2 of 1983 nt (5170–7152), ORF-3 of 342 nt (5159–5500) and a 3′ UTR of 73 nt (7153–7225), followed by a poly(A) tail of 30 residues. ORF-2 and ORF-3 were defined according to Graff et al. (2006) and Huang et al. (2007).

Phylogenetic analysis confirmed the classification of SAAS-JDY5 into genotype 3 HEV (Fig. 1). Similar topologies were obtained with ORF-1, 2- and 3-based phylogenetic analyses (data not shown). The phylogenetic tree revealed that SAAS-JDY5 was most closely related (90.8% sequence similarity) to a Japanese strain, wbJYG1 (AB222184), isolated from a wild boar. The ORF-1, ORF-2 and ORF-3 sequences of the two strains were 90.3%, 92.0% and 97.0% similar, respectively. The similarity between SAAS-JDY5 and the USA HEV strain Meng (AF082843), isolated from swine, was 87.0%. The percent nucleotide identity (PNI) of SAAS-JDY5 with isolates of genotype 1 ranged between 74.1% and 75.3%, between 75.5% and 76.3% for genotype 4 isolates, and was 74.6% in the case of a genotype 2 isolate from Mexico.

The 26 nt of the 5′ UTR and the 73 nt of the 3′ UTR were obtained by 5′ and 3′ RACE, respectively. The 5′ UTR sequence was ACGCGGGGGCAGACCACGTATGTGGTCGATGCC and the 3′ UTR sequence was TTAATTCCTCCTGTGCCCCCTTCGTAGTCTTCTTTGCTTTATTTCTCTTTTCTGCTTTCCGCGCTCCCCGGAC. Alignment of the SAAS-JDY5 5′ UTR with those of other available genotype 3 HEV strains revealed a high level of conservation. However, alignment of the 3′ UTR of SAAS-JDY5 with six other genotype 3 HEV strains revealed the existence of several mutations in this part of the genome. SAAS-JDY5 resembled other genotype 3 isolates in having an ORF-2 with a TAA termination codon, but also had two additional nucleotides (AC) at the very end of the 3′-terminal preceding the poly-A tail. Furthermore, SAAS-JDY5 had a nucleotide deletion within this region at nt position 7189 whereas a G or C was present in other genotype 3 sequences.

ORF-1 consisted of 5109 nucleotides (nt 27–5135) capable of encoding a protein of 1702 amino acids. This was four residues longer than the ORF-1 of HEV-US1 (genotype 3) but shorter than other genotype 3 swine and human isolates. SAAS-JDY5 exhibited only 76.4–84.5% identity with other genotype 3 isolates at the hyper-variable region of ORF-1 (date no shown). Percent nucleotide and amino acid identities of SAAS-JDY5 ORF-1 with ORF-1 sequences of other human genotype 3 isolates ranged between 86.1–88.7% and 96.3–96.9%, respectively. When compared independently with the Japanese strain, wbJYG1, percent nucleotide identity and amino acid similarity values were 90.3% and 96.5%, respectively. ORF-1 motifs, i.e. the NTP-binding domains GVPGSGKS and DEAP in the putative helicase region, and the GDD site found in the RNA-dependent RNA polymerase, were conserved. When compared to other genotype 3 isolates, 15 unique amino acid substitutions and one deletion were observed. Substitutions were observed at the following amino acid positions: 465 (K to R), 598 (R to H), 739 (D to E), 766 (L to P), 870 (V to A), 929 (A to V), 930 (A to T), 961 (A to T), 969 (A to T), 1017 (R to W), 1298 (T to P), 1443 (F to L), 1558 (K to R), 1693 (L to P) and 1696 (S to P), and the Leu residue at position 1048 was deleted.

ORF-2 of SAAS-JDY5 potentially encoded a protein of 660 amino acids of identical size to the corresponding protein of other human and swine genotype 3 HEVs. The N-terminal (residues 1–113) and C-terminal (residues 481–660) regions of the protein exhibited more variability than the central (residues 114–480) region. Substitutions