Acta virologica **62**: 196 – 201, 2018 doi:10.4149/av_2018_212

Prevalence and genetic features of hepatitis E virus in swine, in Gansu, China

J.-H. ZHOU[1,2#], Y.-N. WANG[1#], X.-R. LI[1], Y.-S. LIU[1], Q. PAN[2], X. LAN[1*]

[1]State Key Laboratory of Veterinary Etiological Biology, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, 730046, Gansu, P. R. China; [2]Department of Gastroenterology and Hepatology, Erasmus MC-University Medical Center, Rotterdam, The Netherlands

Received September 11, 2017; revised October 17, 2017; accepted March 27, 2018

Summary. – The prevalence of hepatitis E virus (HEV) in domestic swine threats the public health, due to the risk of zoonotic transmission. In this study, we investigated the prevalence and genetic features of HEV in swine, in Gansu, which is geographically an important province in China. 377 fecal samples were collected from pig farms in Gansu province of China and tested for HEV RNA by RT-PCR. The prevalence rate is about 23% in pig farms of Gansu province. By RT-nPCR, 5' and 3' RACE methods, a whole genome with 7,284 nt in length, termed as swCH189, was obtained and investigated by nucleotide, codon and amino acid usage analyses. Phylogenetic tree analysis classified swCH189 strain into genotype 4e. Although this subtype has never been reported in the local population, genotype 4 is known as zoonotic and more pathogenic than other genotypes. According to the synonymous codon usage patterns of the three open reading frames (ORFs) of swCH189, compositional constraint mainly influences usage patterns of synonymous codons with A-end, while natural selections dominate in usage patterns of synonymous codons with G, C and U-ends. Genetic diversities of each ORFs, in respect to codon and amino acid usage patterns, are closely related to other members of genotype 4 in general, but with distinct features. Thus, the prevalence rate and the genetic features of HEV determined by this study are important for the prevention of zoonotic transmission of HEV from swine to human in this region as well as in China.

Keywords: hepatitis E virus; swine; nucleotide usage; synonymous codon usage; genetic diversity

Introduction

Hepatitis E virus (HEV) is an emerging zoonotic pathogen threatening the public health. It is classified into eight genotypes (HEV 1-8), but the zoonotic potential varies among different genotypes (Meng, 2016; Smith *et al*., 2016; Woo *et al*., 2016). The genome of HEV is a single-stranded positive-sense RNA with about 7.2 kb in length. It encodes three open reading frames (ORF 1-3), a non-structural polyprotein, the capsid and a small phosphoprotein with a multifunctional C-terminal region, respectively (Holla *et al*., 2013).

Currently, it has been widely recognized that many animal species could serve as viral reservoirs, including domestic swine, wild boar, cattle, camel, rabbit, goat and dog (Huang *et al*., 2016; Long *et al*., 2017; Park *et al*., 2016). However, which type of animal serves as the primary reservoir probably depends on the geographical region. Domestic swine is the predominant breed of livestock in China. It is widely distributed across the country, and generally used for meat production. Because HEV infection in swine often fails to produce in notable clinical symptoms, it thus represents a silent risk of transmitting to human via direct contact or consuming meat products. Therefore, assessing the prevalence and understanding the genetic features of HEV in swine shall contribute to establish preventive measures of zoonotic transmission.

Gansu province is located in the northwest of China. It has a population of 26 million and covers an area of 425,800

*Corresponding author: E-mail: lanxi@caas.cn; phone: +86 931 8342771. #These authors contributed equally to this work.

Abbreviations: HEV = hepatitis E virus; COA = correspondence analysis; ORF = open reading frame

square kilometers. Historically and geographically, it is an important region lying between the Tibetan and Loess plateaus and the Northern Silk road passed through the province. In this study, we have collected 377 fecal samples from domestic swine in Gansu province. We determined the prevalence of HEV by RT-PCR of the viral genome. Importantly, we have succeeded in amplifying the full-genome of a particular strain and its genetic features were characterized.

Materials and Methods

Sampling. Based on the collection requirements that the age of pigs with health ranges from 30 to 90 days, a total of 377 fecal samples (one sample for each pig) were collected in Gansu province from nine commercial pig farms where more than 1000 pigs were raised. Each sample was suspended in 10% (w/v) calcium-and magnesium-free phosphate-buffered saline. After centrifugation, these supernatants were collected, and total RNA was extracted by using TRIzol reagent (Invitrogen, USA). RNA samples served as templates to generate full-length cDNA by reverse transcription PCR (Qiagen, Germany).

Amplification and sequencing of the full-length genome of HEV. Seven sets of primers were designed (Table S1), based on a multiple sequence alignment of several whole genome sequences of HEV in GenBank (AY594199, DQ279091, DQ450072 and AB108537). The first round of PCR was carried out using 5 μl of the synthesized cDNA and a set of external primers with Ex Taq DNA polymerase (TakaRa, Japan). Nested PCR was carried out with internal primers and 2 μl of the first round of PCR product. The thermal profile for all PCRs was 94°C for 2 min, followed by 35 cycles of 94°C for 45 s, 50–60°C (depending on the specific PCR reaction) for 45 s, 72°C for 3 min, and finally, 72°C for 10 min. The 5' and 3' RACE were carried out with the SMART™ RACE cDNA amplification kit (Clontech, USA) following the manufacture's instructions. Briefly, 5' terminal first-strand cDNA was synthesized in a 10 μl reaction using 1 μl 5' RT external antisense primer which was specific to the swine HEV (HE1EA), 3 μl total RNA, 1 μl SMART II A oligo and 1 μl MMLV reverse transcriptase. After transcription, first-strand cDNA reaction product was diluted with tricine-EDTA buffer and incubated for 7 min at 72°C. First round of PCR was carried out in 5 cycles of 94°C for 30 s, 72°C for 3 min, followed by 5 cycles of 94°C for 30 s, 70°C for 30 s, 72°C for 3 min, and 25 cycles of 94°C for 30 s, 68°C for 30 s, 72°C for 3 min using 10-fold tricine-EDTA buffer diluted cDNA as template, primers of HE1EA and 5'- and 3'-RACE PCR primer UPM (provided by SMART™ RACE cDNA amplification kit). Nested PCR was carried out at 94°C for 2 min followed by 25 cycles of 94°C for 45 s, 50°C for 45 s, 72°C for 1 min, and finally, 72°C for 10 min using 2 μl 100-fold diluted first round of PCR product, 1 μl internal antisense primer HE1IA and NUP. To amplify the 3' end, first-strand cDNA was synthesized with 1 μl 3'-CDS primer A in a 10 μl reaction. First-strand cDNA (10 μl) was used in a first-round PCR with primers HE7ES and UPM, and 5 μl first-round of PCR product was used in a nested PCR with primers HE7IS and NUP (provided by SMART™ RACE cDNA amplification kit). RT-PCR products were purified using a Gel extraction mini kit (TaKaRa) and inserted into the pGEM-T Easy vector (Promega, USA). The recombinant plasmid was transformed into DH5α competent *E. coli* cells according to the manufacturer's instructions. Positive plasmids containing the inserted fragment were identified by PCR. Three of the positive clones were sequenced by the Sangon Biotech (Shanghai) Co., Ltd.

Phylogenetic analysis for full-length sequence of HEV. According to the previous reported HEV full genome sequences (HEV1-8) (Smith *et al*., 2016; Woo *et al*., 2016) (Table S2), a phylogenetic tree was constructed by maximum composite likelihood distances with 1000 bootstrap replications using MEGA5 software to investigate genetic diversity.

The nucleotide and codon usage patterns of HEV ORFs. To better investigate the physical structure of HEV ORFs, the following compositional properties were calculated: (1) the overall nucleotide usage patterns (A%, U%, G% and C%); (2) nucleotide usage patterns at the different codon positions (A1%, U1%, G1%, C1%, A2%, U2%, G2%, C2%, A3%, U3%, G3% and C3%). In addition, 59 synonymous codon usage patterns were represented by the relative synonymous codon usage values (RSCU) (Sharp *et al*., 1986).

Analyses of genetic diversity of codon and amino acid usages. To better understand genetic diversity in respect to codon/amino acid usages, correspondence analysis (COA), which is a multivariate statistical method which reduces data dimensionality by performing a covariance analysis between factors, was performed by CodonW1.4 version.

Results

HEV prevalence and complete sequence of the swCH189 strain

Detection of HEV RNA from fecal samples identified that 88 of the 377 pigs (23%) were positive. Importantly, we have succeeded in amplification and sequencing of a whole genome of HEV, termed as swCH189 strain (Acc. No. of GenBank: FJ610232). The complete genome of swCH189 was 7,241 nucleotides (nt) long, excluding the 3' poly (A) tail. The genomic organization consists of 5' untranslated region (5'UTR) of 26 nt (1–26), ORF1 of 5124 nt (27–5150), ORF2 of 2025 nt (5147–7171), ORF3 of 345 nt (5175–5519) and 3'UTR of 70 nt (7172–7241), followed by a poly (A) tail of 43 residues. Excluding the poly(A) tail of this strain, the overall nucleotide usage patterns were A% 18.49%, U% 27.23%, G% 25.92% and C% 28.36%. Compared with the overall nucleotide usage patterns of the reference genomes of HEV strains (Table S2), the genome organization of the swCH189 strain was similar to those of HEV strains.

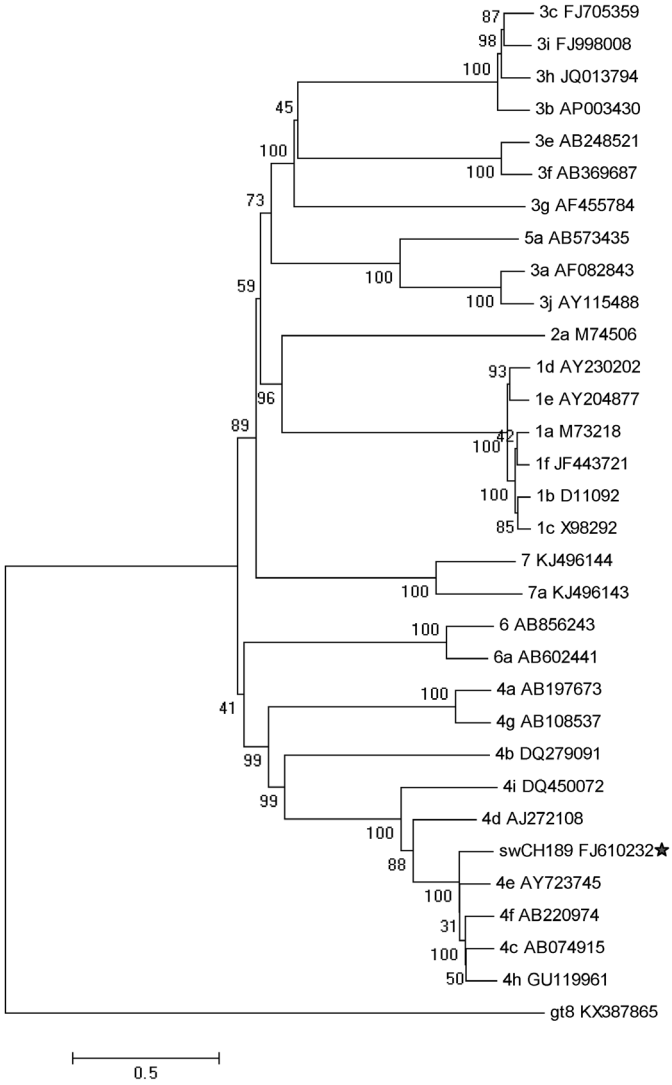

Fig. 1

Phylogenetic analyses of the whole genome of HEV using the neighbor-joining tree of maximum-likelihood distances

Phylogenetic analysis

Phylogenetic tree classified swCH189 strain into genotype 4 and tended to cluster with the India swine HEV isolate (AY723745, genotype 4e) (Fig. 1). According to geographic factor, it is close to the swine HEV isolate (GU119961, genotype 4h) from Xinjiang Uygur autonomous region; whereas distinct from the recently discovered camel HEV isolate (KX387865, genotype 8) from this region (Fig. 1).

Nucleotide and codon usage patterns of the swCH189 strain

The overall nucleotide usage patterns are similar among the three ORFs of the swCH189 strain. However, it has different patterns at the different code positions in the three ORFs and the whole genome (Table S2). At the third code position, U3 was strongly selected by ORF 1 and 2, and C3 was strongly selected by ORF3 (Table S3). This may imply that optimal codons might tend to select those with U-end in ORF 1 and 2, and those with C-end in ORF3. Since the nucleotide A usage was the lowest, no optimal codons selected A-end and many rare codons were with A-end (Table 1). Although the nucleotide U usage in ORF 1 and 2 and nucleotide C usage were highest, codons with U-end in ORF 1 and 2 and those with C-end in ORF3 had no tendency to be selected in high frequency (Table 1). These results suggested that other evolutional factors, which are involved in gene transcription/translation and gene functions, likely

Table 1. Synonymous codon usage patterns for ORFs of the swCH189 strain

Codon	ORF1	ORF2	ORF3
UUU(F)	1.22	1.14	0.80
UUC(F)	0.78	0.86	1.20
UUA(L)	**0.37**	**0.48**	**0.00**
UUG(L)	0.84	1.06	**0.40**
CUU(L)	1.94	2.23	1.20
CUC(L)	1.46	1.16	2.00
CUA(L)	**0.48**	**0.39**	**0.40**
CUG(L)	0.91	0.68	2.00
AUU(I)	1.64	1.88	1.00
AUC(I)	0.66	0.63	1.00
AUA(I)	0.70	**0.50**	1.00
GUU(V)	1.56	1.69	0.89
GUC(V)	0.81	1.07	1.78
GUA(V)	**0.44**	**0.18**	**0.00**
GUG(V)	1.19	1.07	1.33
UCU(S)	1.95	2.69	**0.55**
UCC(S)	0.92	1.24	1.09
UCA(S)	1.03	0.62	**0.00**
UCG(S)	0.65	0.72	1.64
AGU(S)	1.03	**0.52**	**0.55**
AGC(S)	**0.43**	**0.21**	2.18
CCU(P)	1.50	1.51	**0.57**
CCC(P)	1.09	1.06	0.95
CCA(P)	0.77	0.68	0.95
CCG(P)	0.64	0.75	1.52
ACU(T)	1.08	1.88	**0.00**
ACC(T)	1.46	1.09	3.00
ACA(T)	0.88	**0.55**	1.00
ACG(T)	**0.58**	**0.48**	**0.00**
GCU(A)	1.11	2.00	**0.33**
GCC(A)	1.49	1.27	1.67
GCA(A)	0.82	**0.36**	0.67
GCG(A)	**0.58**	**0.36**	1.33
UAU(Y)	1.00	1.36	**0.00**
UAC(Y)	1.00	0.64	**0.00**
CAU(H)	1.00	0.83	0.67
CAC(H)	1.00	1.17	1.33
CAA(Q)	**0.43**	**0.40**	0.80
CAG(Q)	1.57	1.60	1.20
AAU(N)	1.11	1.41	0.00
AAC(N)	0.89	**0.59**	**0.00**
AAA(K)	**0.51**	1.09	**0.00**
AAG(K)	1.49	0.91	**0.00**
GAU(D)	1.35	1.10	**0.00**
GAC(D)	0.65	0.90	2.00
GAA(E)	**0.32**	**0.35**	**0.00**
GAG(E)	1.68	1.65	2.00
UGU(C)	1.22	0.67	**0.57**
UGC(C)	0.78	1.33	1.43
CGU(R)	1.46	2.38	1.20
CGC(R)	1.62	2.38	3.60
CGA(R)	**0.52**	**0.50**	**0.00**
CGG(R)	1.57	0.75	1.20
AGA(R)	**0.10**	**0.00**	**0.00**
AGG(R)	0.73	**0.00**	**0.00**
GGU(G)	1.18	1.41	**0.00**
GGC(G)	1.54	1.26	1.78
GGA(G)	**0.16**	**0.52**	**0.00**
GGG(G)	1.11	0.81	2.22

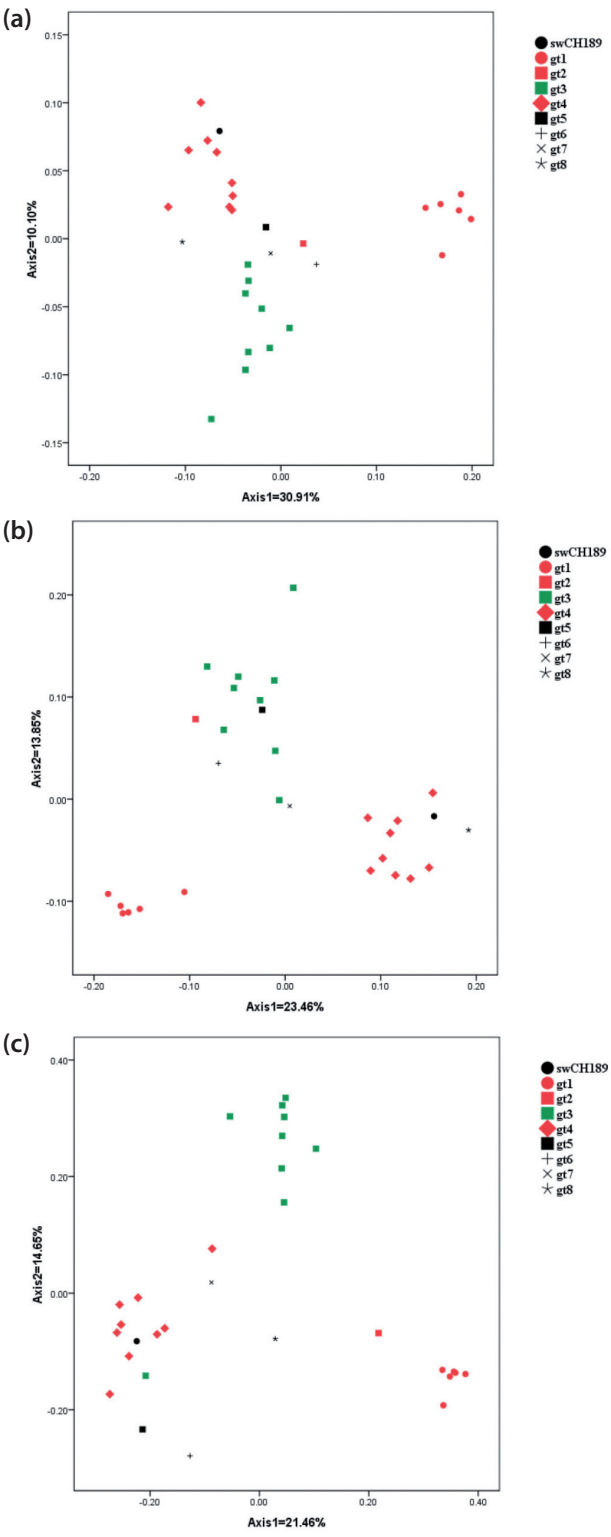

Fig. 2

Genetic diversity of HEV strains reflected by principal axes 1 and 2 generated from COA method based on synonymous codon usage patterns of the corresponding strain
(**a**) HEV ORF1; (**b**) HEV ORF2; (**c**) HEV ORF3.

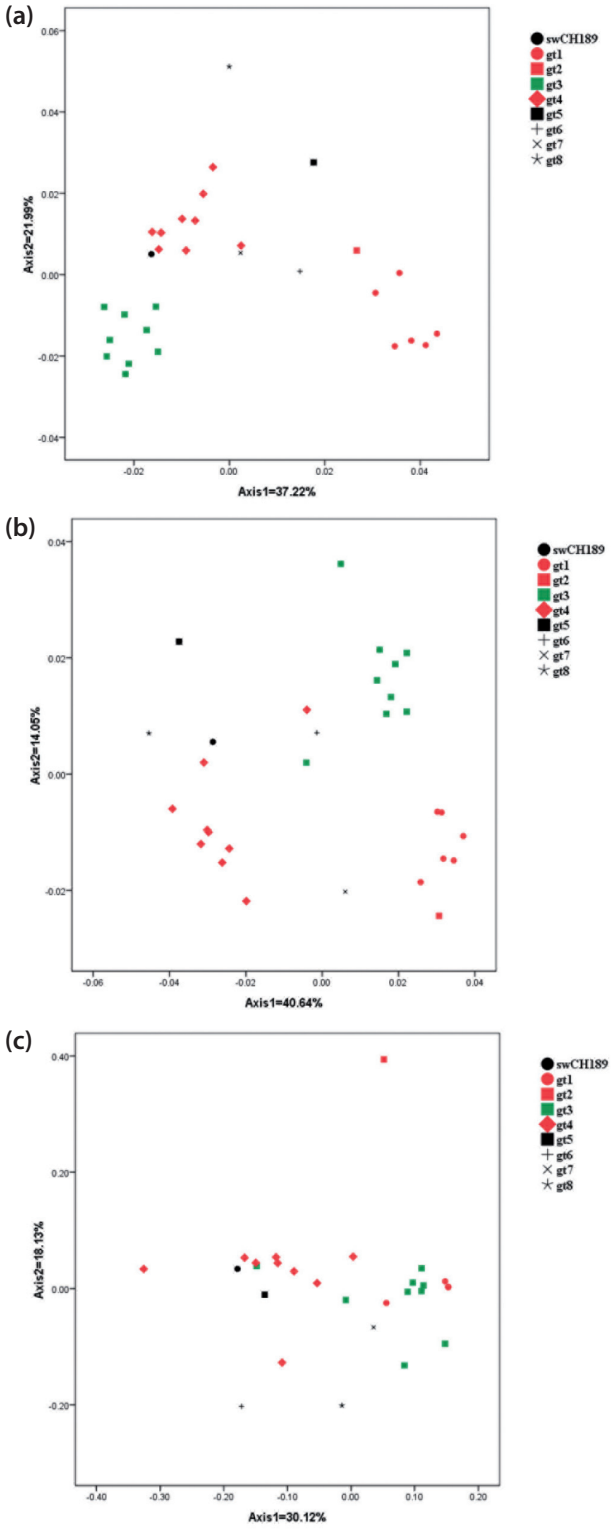

Fig. 3

Genetic diversity of HEV strains reflected by principal axes 1 and 2 generated from COA method based on amino acid usage patterns of the corresponding strains
(**a**) HEV ORF1; (**b**) HEV ORF2; (**c**) HEV ORF3.

influenced the formation of nucleotide composition and synonymous codon usage.

Genetic diversities for codon and amino acid usages

Projection of synonymous codon usages of each ORF on the two-dimensional map represented the genetic classification of the swCH189 strain into the HEV variants (Fig. 2). Generally, the pattern of synonymous codon usages for each ORF of this strain is closely related to the members of the same genotype. As for genetic diversity of ORF1 and ORF2, both are close to genotype 4; whereas ORF2 is more close to genotype 8. Similarly, projection of amino acid usage of each ORF revealed that ORF 1 and 3 are more closely related to the member of genotype 4 than the ORF2 of this strain (Fig. 3).

Discussion

In China, the meat and related products from domestic swine are the important sources of food to the general public. Exposure to domestic swine and the related products has been recognized as the primary cause of HEV infection (Hsieh *et al*., 1999; Meng *et al*., 2002). In this study, we have revealed a prevalence rate of 23% in domestic swine in Gansu province of China, by RT-PCR of HEV genome from fecal samples. This rate is comparable to that in Beijing (33%) and Shandong (30%), but lower than that in Henan (92%) (Xia *et al*., 2015).

Among these samples, a full-length HEV genome (swCH189) was obtained by means of 5' and 3' RACE methods. In comparison with reference genome of different genotypes, this strain belongs to genotype 4e. It is different from the genotypes identified in other provinces, including genotype 4h in Henan and genotype 4d in Beijing and Shandong (Xia *et al*., 2015), indicating a region specificity of the prevalence of subtypes. Genotype 4d has been detected in patients in Shandong and Jiangsu of China (Liu *et al*., 2012; Xia *et al*., 2015). Genotype 4f has been found in patients from Gunsu province (Ma *et al*., 2010). Up to date, no report about genotype 4e has been found in patients in Gansu, but it clearly bears a high risk of zoonotic transmission of this subtype from pigs to human being.

We have characterized the genetic features of the swCH189 strain, in respect to the nucleotide, synonymous codon and amino acid usage patterns. Overall, it is closely clustered to the other members of genotype 4. These genetic features implicated that mutation pressure from compositional constraint was one of evolutional forces in the formation of synonymous codon usage patterns in the three ORFs of swCH189 strain and natural selections related to viral fitness may also take part in this formation.

In conclusion, we have revealed a prevalence rate of 23% of HEV infection in swine in Gansu, China. We obtained and characterized a full-length strain, belonging to genotype 4e, which has never been reported in this region. These results are important for the prevention of zoonotic transmission of HEV from swine to human in this region as well as in China.

Acknowledgments. This work was supported by National Natural Science foundation of China (No. 31302100).

Supplementary information is available in the online version of the paper.

References

Holla RP, Ahmad I, Ahmad Z, Jameel S (2013): Molecular virology of hepatitis E virus. Semin. Liver Dis. 33, 3–14. https://doi.org/10.1055/s-0033-1338110

Hsieh SY, Meng XJ, Wu YH, Liu ST, Tam AW, Lin DY, Liaw YF (1999): Identity of a novel swine hepatitis E virus in Taiwan forming a monophyletic group with Taiwan isolates of human hepatitis E virus. J. Clin. Microbiol. 37, 3828–3834.

Huang F, Li Y, Yu W, Jing S, Wang J, Long F, He Z, Yang C, Bi Y, Cao W, Liu C, Hua X, Pan Q (2016): Excretion of infectious hepatitis E virus into milk in cows imposes high risks of zoonosis. Hepatology 64, 350–359. https://doi.org/10.1002/hep.28668

Liu P, Li L, Wang L, Bu Q, Fu H, Han J, Zhu Y, Lu F, Zhuang H (2012): Phylogenetic analysis of 626 hepatitis E virus (HEV) isolates from humans and animals in China (1986-2011) showing genotype diversity and zoonotic transmission. Infect. Genet. Evol. 12, 428–434. https://doi.org/10.1016/j.meegid.2012.01.017

Long F, Yu W, Yang C, Wang J, Li Y, Li Y, Huang F (2017): High prevalence of Hepatitis E virus infection in goats. J. Med. Virol. 89, 1981–1987. https://doi.org/10.1002/jmv.24843

Ma Z, Feng R, Zhao C, Harrison TJ, Li M, Qiao Z, Feng Y, Wang Y (2010): Seroprevalence and distribution of hepatitis E virus in various ethnic groups in Gansu province, China. Infect. Genet. Evol. 10, 614–619. https://doi.org/10.1016/j.meegid.2010.04.003

Meng XJ, (2016): Expanding host range and cross-species infection of hepatitis E virus. PLoS Pathog. 12, e1005695. https://doi.org/10.1371/journal.ppat.1005695

Meng XJ, Wiseman B, Elvinger F, Guenette DK, Toth TE, Engle RE, Emerson SU, Purcell RH (2002): Prevalence of antibodies to hepatitis E virus in veterinarians working with swine and in normal blood donors in the United States and other countries. J. Clin. Microbiol. 40, 117–122. https://doi.org/10.1128/JCM.40.1.117-122.2002

Park WJ, Park BJ, Ahn HS, Lee JB, Park SY, Song CS, Lee SW, Yoo HS, Choi IS (2016): Hepatitis E virus as an emerging zoonotic pathogen. J. Vet. Sci. 17, 1–11. https://doi.org/10.4142/jvs.2016.17.1.1

Sharp PM, Tuohy TM, Mosurski KR (1986): Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. Nucleic. Acids. Res. 14, 5125–5143. https://doi.org/10.1093/nar/14.13.5125

Smith DB, Simmonds P, Izopet J, Oliveira-Filho EF, Ulrich RG, Johne R, Koenig M, Jameel S, Harrison TJ, Meng XJ, Okamoto H, Van der Poel WH, Purdy MA (2016): Proposed reference sequences for hepatitis E virus subtypes. J. Gen. Virol. 97, 537–542. https://doi.org/10.1099/jgv.0.000393

Woo PC, Lau SK, Teng JL, Cao KY, Wernery U, Schountz T, Chiu TH, Tsang AK, Wong PC, Wong EY, Yuen KY (2016): New Hepatitis E Virus Genotype in Bactrian Camels, Xinjiang, China, 2013. Emerg. Infect. Dis. 22, 2219–2221. https://doi.org/10.3201/eid2212.160979

Xia J, Zeng H, Liu L, Zhang Y, Liu P, Geng J, Wang L, Wang L, Zhuang H (2015): Swine and rabbits are the main reservoirs of hepatitis E virus in China: detection of HEV RNA in feces of farmed and wild animals. Arch Virol 160, 2791–2798. https://doi.org/10.1007/s00705-015-2574-0

Supplementary information

Prevalence and genetic features of hepatitis E virus in swine, in Gansu, China

J.-H. ZHOU[1,2#], Y.-N. WANG[1#], X.-R. LI[1], Y.-S. LIU[1], Q. PAN[2], X. LAN[1*]

[1]State Key Laboratory of Veterinary Etiological Biology, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, 730046, Gansu, P.R. China; [2]Department of Gastroenterology and Hepatology, Erasmus MC-University Medical Center, Rotterdam, The Netherlands

Received September 11, 2017; revised October 17, 2017; accepted March 27, 2018

Table S1. Primers used for swCH189 complete genome sequence amplification

Primer name	Nucleotide position*	Nucleotide sequence (5'-3')	Length (bp)
HE1IA	389~366	AGCAGTGTACCACCGCTGAACATC	389
HE1EA	523~498	AGTGAGTARAGGGCAAYCCCCGTCTC	498
HE2IS	110~128	GGCGAATGCTGTGGTGGTT	1067
HE2IA	1176~1154	GCTGRTGGCARATRGTMAGATAG	
HE2ES	52~74	CTCCTGGCRTYACTACTGCYATT	1279
HE2EA	1330~1311	CGGCACTGGGCRTAAAACTG	
HE3IS	1096~1116	TTGCYAAYGAGGGYTGGAATG	2394
HE3IA	3489~3469	GRACYGTAATYGCACCAGGGT	
HE3ES	1028~1048	GTTTTGCTGYTCRMGGCTAAT	2567
HE3EA	3594~3574	GRGTVAGKGCMACWATAGCRT	
HE4IS	3206~3225	CTCBTCRGTCCATCTCCTTG	994
HE4IA	4199~4180	MCGSGAMACATCACGGYTGC	
HE4ES	3168~3185	CTYCCACCRCAYYTGCTG	1042
HE4EA	4209~4186	AGAATGTKATACGSGAMACATCAC	
HE5IS	4173~4193	GAYCTATGYAACCGTGATGTG	1396
HE5IA	5568~5549	GCGRGARTCAACATCAGGBA	
HE5ES	3881~3901	CTACMTGCCBCAGGARCTBAC	1823
HE5EA	5703~5683	ATGAGTRTTRGTGCCRTCYTG	
HE6IS	5341~5362	CTCCCCTATATTCATCCAACCA	733
HE6IA	6073~6049	GGGCACCAGTATAAGGTGTATTAGT	
HE6ES	5333~5351	CCTTCGCCCTCCCCTATAT	892
HE6EA	6224~6204	GTAGCMGCMGTRGTRGTHAGC	
HE7ES	5954~5971	GCTGGCGTTCGGTTGAGA	
HE7IS	6004~6024	GGCCTTGTYATGCTYTGYATC	1281

*Primers nucleotide positions is related to HEV strain swCH31 (GenBank Acc. No. DQ450072). The primers HE1IA and HE1EA were used for 5´RACE reaction and the primers HE7ES and HE7IS were used for 3´RACE reaction.

Table S2. Reference HEV strains with different genotypes 1–8

Acc. No.	Host	Strain	Genotype
M73218	*Homo sapiens*	B1/Myanmar	1a
D11092	*Homo sapiens*	Xinjiang/China/1988	1b
X98292	*Homo sapiens*	hev037/India	1c
AY230202	*Homo sapiens*	Morocco	1d
AY204877	*Homo sapiens*	T3/Chad	1e
JF443721	*Homo sapiens*	IND-HEV-AVH5-2010/India/2010	1f
M74506	*Homo sapiens*	M1/Mexico	2a
AF082843	Swine	Meng	3a
AP003430	*Homo sapiens*	JRA1/Japan	3b
FJ705359	Wild boar	wbGER27/Germany/2006	3c
AB248521	Swine	swJ8-5/Japan	3e
AB369687	*Homo sapiens*	E116-YKH98C/Japan/1998	3f
AF455784	Swine	Osh 205/Japan	3g
JQ013794	*Homo sapiens*	TR19 / France / 2007	3h
FJ998008	Wild boar	BB02/Germany/2007	3i
AY115488	Swine	Arkell/Canada	3j
AB197673	*Homo sapiens*	JKO-ChiSai98C/China/1998	4a
DQ279091	Swine	swDQ/China	4b
AB074915	*Homo sapiens*	JAK-Sai/Japan	4c
AJ272108	*Homo sapiens*	T1/China	4d
AY723745	Swine	IND-SW-00-01/India	4e
AB220974	*Homo sapiens*	HE-JA2/Japan	4f
AB108537	*Homo sapiens*	CCC220/China/2000	4g
GU119961	Swine	CHN-XJ-SW13/China/2009	4h
DQ450072	Swine	swCH31/China/2006	4i
AB573435	Wild boar	JBOAR135-Shiz09/Japan/2009	5a
AB602441	Wild boar	wbJOY_06/Japan/2006	6a
KJ496143	Camel	178C/ United Arab Emirates/2013	7a
KX387865	Camel	12XJ/China	8

Table S3. Nucleotide usage patterns for ORFs of the swCH189 strain

	ORF1	ORF2	ORF3
U%	26.5	28.2	20.9
C%	27.8	29.7	40.0
A%	18.8	18.3	11.0
G%	26.9	23.8	28.1
U1%	18	19	17
C1%	27.2	27.3	41.7
A1%	19.6	21.5	12.2
G1%	34.9	32.4	28.7
U2%	27	24	30
C2%	28.9	35.0	37.4
A2%	24.2	22.8	9.6
G2%	20.0	18.2	23.5
U3%	34	42	16
C3%	27.3	27.0	40.9
A3%	12.5	10.7	11.3
G3%	25.9	20.6	32.2