

RESEARCH ARTICLE

Open Access



# Meta-transcriptomic analysis reveals a new subtype of genotype 3 avian hepatitis E virus in chicken flocks with high mortality in Guangdong, China

Xue-Lian Zhang<sup>1,2</sup>, Wen-Feng Li<sup>2</sup>, Sheng Yuan<sup>2</sup>, Jin-Yue Guo<sup>1,2</sup>, Zhi-Li Li<sup>1,2</sup>, Shi-Hong Chi<sup>2</sup>, Wen-Jing Huang<sup>2</sup>, Xiao-Wen Li<sup>2</sup>, Shu-Jian Huang<sup>1,2\*</sup> and Jian-Wei Shao<sup>1,2,3\*</sup> 

## Abstract

**Background:** Hepatitis E virus (HEV) is one of most important zoonotic viruses, and it can infect a wide range of host species. Avian HEV has been identified as the aetiological agent of big liver and spleen disease or hepatitis-splenomegaly syndrome in chickens. HEV infection is common among chicken flocks in China, and there are currently no practical measures for preventing the spread of the disease. The predominant avian HEV genotype circulating in China have been identified as genotype 3 strains, although some novel genotypes have also been identified from chicken flocks in China.

**Results:** In this study, we used a meta-transcriptomics approach to identify a new subtype of genotype 3 avian HEV in broiler chickens at a poultry farm located in Shenzhen, Guangdong Province, China. The complete genome sequence of the avian HEV, designated CaHEV-GDSZ01, is 6655-nt long, including a 5' UTR of 24 nt and a 3' UTR of 125 nt (excluding the poly(A) tail), and contains three open reading frames (ORFs). Sequence analysis indicated that the complete ORF1 (4599 nt/1532 aa), ORF2 (1821 nt/606 aa) and ORF3 (264 nt/87 aa) of CaHEV-GDSZ01 share the highest nucleotide sequence identity (85.8, 86.7 and 95.8%, respectively) with the corresponding ORFs of genotype 3 avian HEV. Phylogenetic analyses further demonstrated that the avian HEV identified in this study is a new subtype of genotype 3 avian HEV.

**Conclusions:** Our results demonstrate that a new subtype of genotype 3 avian HEV is endemic in Guangdong, China, and could cause high mortality in infected chickens. This study also provides full genomic data for better understanding the evolutionary relationships of avian HEV circulating in China. Altogether, the results presented in this study suggest that more attention should be paid to avian HEV and its potential disease manifestation.

**Keywords:** Avian hepatitis E virus, High mortality, Meta-transcriptomics analysis, New subtype

## Background

Hepatitis E virus (HEV) is the causative agent of hepatitis E, which is an important public health concern in many parts of the world, particularly in developing countries [1, 2]. HEV is a non-enveloped, single-stranded positive-sense RNA virus that belongs to the family Hepeviridae and includes two genera,

*Orthohepevirus* and *Piscihepevirus* [3]. The genome encodes three open reading frames (ORFs) flanked by a capped 5' terminus and a polyadenylated 3' terminus. Among the three ORFs, ORF1 is the longest and is located at the 5' terminus of the genome; this ORF encodes a non-structural polyprotein including a methyltransferase, a papain-like cysteine protease, a viral helicase, and an RNA-dependent RNA polymerase (RdRp) [4]. ORF2 is located at the 3' end of the genome and encodes a capsid protein, which is the major structural protein and functionally binds to host cells, induces

\* Correspondence: [huangshujian@fosu.edu.cn](mailto:huangshujian@fosu.edu.cn); [jwshao1988@163.com](mailto:jwshao1988@163.com)

<sup>1</sup>Key Laboratory for Preventive Research of Emerging Animal Diseases, Foshan University, Foshan 528231, Guangdong, China

Full list of author information is available at the end of the article



neutralizing antibody production, and participates in viral particle assembly [5, 6]. ORF3, which overlaps with ORF2, encodes a cytoskeleton-associated phosphoprotein that interacts with the ORF2 protein and a number of cellular signal transduction pathway proteins [7, 8].

As an important zoonotic virus, HEV can infect a wide range of host species, and its host spectrum is ever expanding. Since the first animal strain of HEV (i.e., swine HEV) was isolated and characterized in 1997 in the United States [9], numerous strains of HEV have been isolated from a number of different animal species and a wide range of geographic locations. Countries that have reported HEV infections include Japan, Taiwan, New Zealand, mainland China, South Korea, Hungary, Australia, the United Arab Emirates, and the Netherlands, among others. The infected animal species include domestic and wild pigs, chickens, rabbits, ruminants, ferrets, minks, bats, rodents, and aquatic birds [3, 10].

Based on sequence analysis, most of the HEVs identified thus far belong to the genus *Orthohepevirus*, which contains four species, designated A, B, C, and D [11, 12]. Avian HEV, belonging to *Orthohepevirus* B, has been identified as the aetiological agent of big liver and spleen disease (BLSD) or hepatitis-splenomegaly syndrome (HS syndrome) in broiler breeder hens and laying hens aged 30 to 72 weeks, via the faecal-oral route [13]. Avian HEV infection can cause increased mortality, a reduction in egg production or subclinical infection, resulting in large economic losses in the poultry industry [14, 15]. The avian HEV genome shares approximately 50% nucleotide sequence identity and some similar antigenic epitopes with mammalian HEVs [13]. Various avian HEVs have been classified into four different genotypes based on full or nearly complete genomes: genotype 1, which is mainly found in Australia and Korea [16, 17], genotype 2, from the USA and Korea [13, 18], genotype 3, from Europe and China [19, 20], and genotype 4, from Hungary and Taiwan [21, 22].

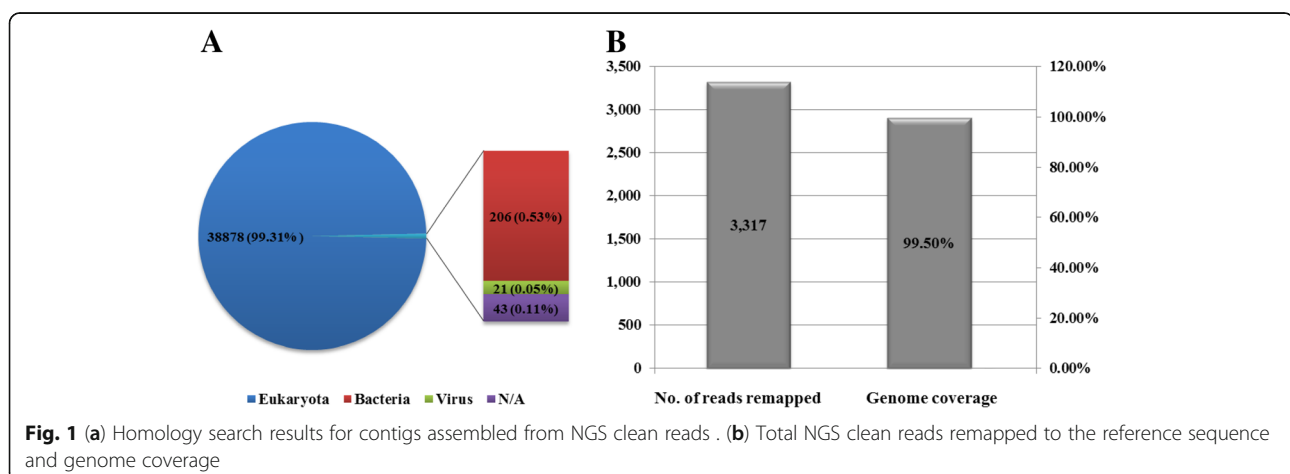
Avian HEV genotype 3 was first characterized in Hungary in 2005 and was later detected in the United Kingdom, Germany, and Austria before 2007 [16]. In China, an Avian HEV genotype 3 strain designated CaHEV was first detected and characterized in Shandong Province in 2010 [20]. Since then, avian HEV strains genotype 3 have been identified in chickens with HS syndrome in many provinces of China [15, 23–26]. In recent years, many previously undescribed genotypes of avian HEV have been found in different regions of China [27, 28], which suggests that there is much greater diversity of avian HEV circulating in chicken flocks in China than previously indicated.

## Results

### Meta-transcriptomics based pathogen discovery

Through an unbiased high-throughput RNA sequencing approach, a total of 42,682,590 paired-end sequencing reads were generated, resulting in 14.5 GB of fastq format sequence data. After default quality control (QC) and de-barcoding steps provided by the Illumina platform, 40,531,484 reads (95%) remained for further analysis.

After de novo assembly using Trinity, a total of 48,891 contigs were generated, which varied from 201 to 9584 nt in length. These contigs were compared to non-redundant nucleotide databases (nt) via nucleotide Blast searches. A total of 38,878 contigs (79.1%), 206 contigs (0.53%), and 21 contigs (0.05%) were annotated as belonging to eukaryotes, bacteria, and viruses, respectively. The remaining 43 contigs (0.11%) were labelled as N/A, showing no taxonomic information (Fig. 1a). All the contigs annotated as eukaryotic were identified as coming from chicken RNA sources, and no fungal pathogens were found. Those contigs identified as bacterial came from species from the genera *Escherichia*, *Acinetobacter*, *Bradyrhizobium*, *Bacillus*, and *Rhizobacter*, which are ubiquitous in the natural environment. Those contigs identified as viral were annotated as avian HEV and avian leukosis virus (ALV).



Strikingly, the assembled contigs covered the nearly complete genome sequence of avian HEV. The sequences were confirmed by reverse transcription polymerase chain reaction (RT-PCR), and 5' and 3' RACE were used to obtain the terminal sequences. The virus sequence obtained in this study was designated CaHEV-GDSZ01 and has been deposited in GenBank under accession number MK050107.

The sequence reads were subsequently mapped to the virus genome to estimate sequencing coverage and depth. Using the complete virus genome of CaHEV-GDSZ01 as the reference sequence, 3317 reads were remapped to the avian HEV sequence, with 99.5% genome coverage (6621/6655 nt) and pairwise identity of 97.1% at a mean depth of 8× (Fig. 1b).

### Sequence comparison

The complete genome of CaHEV-GDSZ01 consisted of 6655 nucleotides (nt) excluding the poly(A) tail at the 3' end, and it contained three major ORFs. The genome comprised a 5' UTR of 24 nt (1–24), ORF-1 of 4599 nt

(25–4623), encoding a non-structural polyprotein of 1532 amino acids, ORF-2 of 1821 nt (4710–6530), encoding a capsid protein of 606 amino acids, ORF-3 of 264 nt (4657–4920), encoding a cytoskeleton-associated phosphoprotein of 87 amino acids, and a 3' UTR of 125 nt (6531–6655).

Multiple sequence comparisons based on individual ORFs showed that the complete ORF1 of CaHEV-GDSZ01 shared 80.8 to 85.8% nucleotide sequence identity and 92.0 to 95.0% amino acid sequence identity with reference strains. The nucleotide sequence identity between the complete ORF2 of CaHEV-GDSZ01 and the reference strains varied from 83.5 to 86.7%, while the amino acid sequence identity between them was 98.0 to 99.3%, which were higher identities than those of the complete ORF1 (Table 1). Additionally, sequence comparison of the complete ORF3 with reference strains revealed that they shared 92.8–95.8% nucleotide sequence identity and 89.7–100% amino acid sequence identity, which was significantly higher than the identities of ORF1 and ORF2. Additionally, the nucleotide sequence identities

**Table 1** Comparison of the nucleotide and amino acid identities of complete ORF1 and ORF2 gene sequences among avian HEV strains

Gene	Strains	1	2	3	4	5	6	7	8	9	10	11	12
ORF1	1 MK050107/China		81.2	80.8	80.8	85.6	85.8	82.6	83.3	81.1	81.5	81.5	81.6
	2 AM943647/ Australia	<b>92.2</b>		86.5	86.3	81.5	81.7	80.2	81.0	81.3	81.5	81.5	81.9
	3 JN597006/South Korea	<b>92.4</b>	<b>95.9</b>		91.5	81.8	81.7	81.0	80.9	81.5	81.6	81.6	81.9
	4 KC454286/South Korea	<b>92.0</b>	<b>96.1</b>	<b>97.8</b>		81.0	81.1	80.2	81.2	81.2	81.5	81.5	81.4
	5 GU954430/China	<b>94.8</b>	<b>93.0</b>	<b>92.6</b>	<b>92.2</b>		98.3	82.5	83.1	81.1	80.6	80.6	81.6
	6 AM943646/Hungary	<b>94.8</b>	<b>93.1</b>	<b>93.0</b>	<b>92.6</b>	<b>98.8</b>		83.0	83.3	81.2	80.9	80.9	81.6
	7 JN997392/Hungary	<b>93.7</b>	<b>92.0</b>	<b>92.2</b>	<b>91.8</b>	<b>94.1</b>	<b>94.3</b>		87.3	80.5	81.1	81.1	81.1
	8 KF511797/Taiwan	<b>95.0</b>	<b>93.1</b>	<b>93.3</b>	<b>92.8</b>	<b>94.8</b>	<b>94.9</b>	<b>95.9</b>		81.6	81.4	81.4	81.6
	9 KM377618/South Korea	<b>92.3</b>	<b>93.3</b>	<b>93.0</b>	<b>92.8</b>	<b>93.0</b>	<b>93.2</b>	<b>92.5</b>	<b>93.1</b>		89.8	89.8	89.3
	10 AY535004/United States	<b>92.6</b>	<b>93.7</b>	<b>93.4</b>	<b>93.3</b>	<b>93.0</b>	<b>93.2</b>	<b>92.7</b>	<b>93.6</b>	<b>96.9</b>		100	89.6
	11 NC023425/ United States	<b>92.6</b>	<b>93.7</b>	<b>93.4</b>	<b>93.3</b>	<b>93.0</b>	<b>93.2</b>	<b>92.7</b>	<b>93.6</b>	<b>96.9</b>	<b>100</b>		89.6
	12 EF206691/ United States	<b>93.2</b>	<b>93.9</b>	<b>93.5</b>	<b>93.5</b>	<b>92.8</b>	<b>93.0</b>	<b>92.6</b>	<b>93.4</b>	<b>96.5</b>	<b>97.4</b>	<b>97.4</b>	
ORF2	1 MK050107/China		84.8	83.5	84.0	86.7	86.5	86.7	84.9	83.9	84.5	84.5	84.7
	2 AM943647/ Australia	<b>98.3</b>		87.9	88.1	84.4	84.1	83.9	84.1	83.6	84.3	84.3	84.5
	3 JN597006/South Korea	<b>98.2</b>	<b>98.8</b>		91.6	85.0	85.0	84.2	84.4	85.0	86.2	86.2	84.6
	4 KC454286/South Korea	<b>98.0</b>	<b>98.3</b>	<b>98.3</b>		84.8	84.6	84.7	84.5	83.8	84.5	84.5	84.2
	5 GU954430/China	<b>99.0</b>	<b>98.8</b>	<b>98.7</b>	<b>98.5</b>		98.5	87.3	85.0	84.0	84.5	84.5	84.1
	6 AM943646/Hungary	<b>98.7</b>	<b>98.7</b>	<b>98.3</b>	<b>98.2</b>	<b>99.7</b>		87.1	84.8	83.6	84.4	84.4	84.0
	7 JN997392/Hungary	<b>99.3</b>	<b>98.7</b>	<b>98.5</b>	<b>98.3</b>	<b>99.7</b>	<b>99.3</b>		87.9	84.0	84.9	84.9	84.6
	8 KF511797/Taiwan	<b>98.5</b>	<b>98.7</b>	<b>98.2</b>	<b>98.0</b>	<b>99.2</b>	<b>99.0</b>	<b>99.0</b>		84.6	84.3	84.3	84.7
	9 KM377618/South Korea	<b>97.9</b>	<b>98.2</b>	<b>98.0</b>	<b>97.9</b>	<b>98.3</b>	<b>98.2</b>	<b>98.2</b>	<b>98.5</b>		91.7	91.7	89.6
	10 AY535004/USA	<b>98.5</b>	<b>98.8</b>	<b>99.0</b>	<b>98.2</b>	<b>99.0</b>	<b>98.7</b>	<b>98.8</b>	<b>98.5</b>	<b>98.7</b>		100	90.7
	11 NC023425/USA	<b>98.5</b>	<b>98.8</b>	<b>99.0</b>	<b>98.2</b>	<b>99.0</b>	<b>98.7</b>	<b>98.8</b>	<b>98.5</b>	<b>98.7</b>	<b>100</b>		90.7
	12 EF206691/USA	<b>98.2</b>	<b>98.5</b>	<b>98.3</b>	<b>97.5</b>	<b>98.3</b>	<b>98.2</b>	<b>98.2</b>	<b>98.5</b>	<b>98.7</b>	<b>99.0</b>	<b>99.0</b>	

The comparison was done using MegAlign ClustalW analysis. Boldface indicates percentage identities of amino acid sequences

of the 5'UTR and 3'UTR of CaHEV-GDSZ01 were 64.0–88.0% and 54.3–83.7%, respectively, compared to other avian HEV strains (Table 2).

**Phylogenetic analysis**

To examine the evolutionary relationships of avian HEV determined in this study with other virus strains, phylogenetic analyses were performed based on the complete nucleotide sequences of ORF1 and ORF2 of HEV. All the phylogenetic trees were estimated using the maximum likelihood method with 1000 bootstrap replicates, and the cutthroat trout virus was used as the outgroup in each case.

Phylogenetic analysis based on the complete ORF1 showed that all known Orthohepeviruses, including the virus newly identified CaHEV-GDSZ01, were divided into four clades. All the viruses isolated in chickens, including CaHEV-GDSZ01, clustered together and constituted the species *Orthohepevirus* B. Furthermore, CaHEV-GDSZ01 was clustered together with two avian HEVs isolated in

China and Hungary with short branch lengths and was allocated to the genotype-3 subclade of the *Orthohepevirus* B species (Fig. 2a). A similar phylogenetic relationship was observed in a maximum-likelihood tree based on the complete sequences of ORF2 (Fig. 2b).

**Selection pressure analysis**

In the selection pressure analyses, as shown in Fig. 3, the mean dN-dS values of ORF1, ORF2, and ORF3 were – 2.04, – 1.61, and – 0.37, respectively. Many negatively selected sites were observed in ORF1, ORF2, and ORF3 (642, 305, and 7, respectively), while no positively selected sites were found, which suggests a lack of positive selected sites in the three ORFs examined.

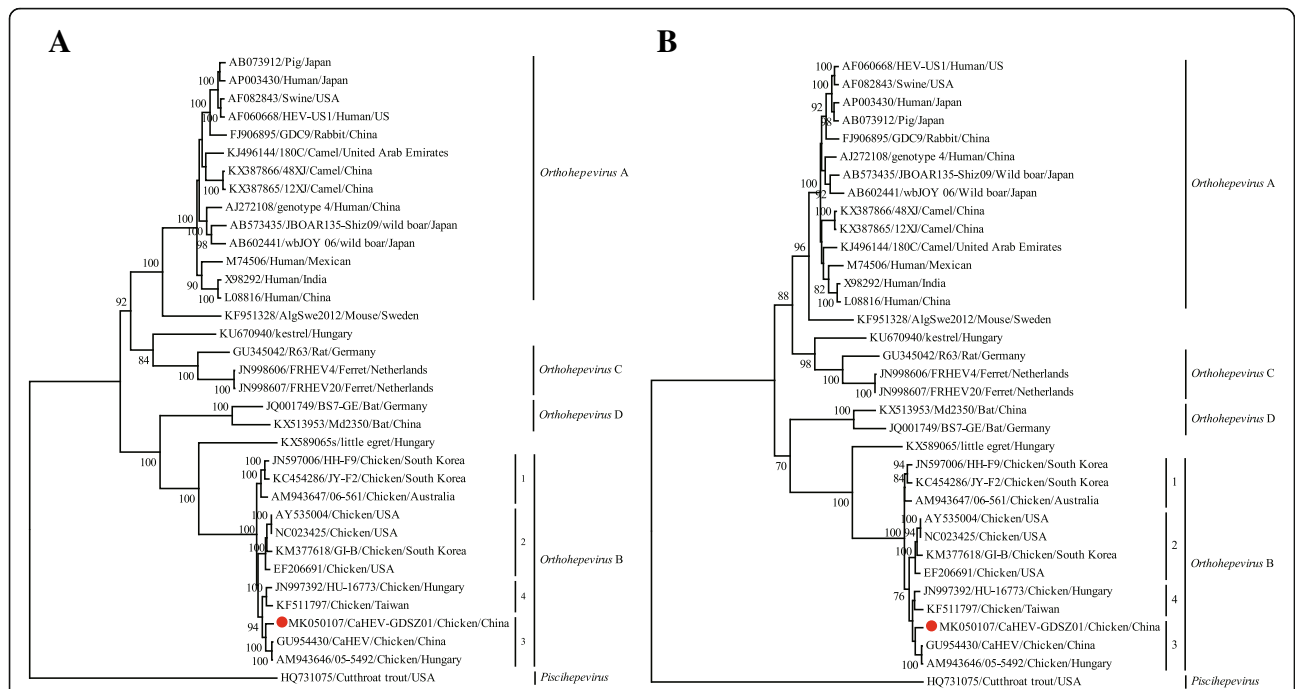
**Discussion**

Avian HEV causes diseases that threaten the healthy development of the poultry industry worldwide. Since the CaHEV strain was first detected and characterized in China in 2010, avian HEV genotype 3 strain has been

**Table 2** Comparison of the nucleotide and/or amino acid sequences identities of complete ORF3, 5'UTR and 3'UTR gene sequences among avian HEV strains

Gene	Strains	1	2	3	4	5	6	7	8	9	10	11	12
ORF3	1 MK050107/China		93.9	92.8	92.8	95.5	95.8	95.8	95.8	93.2	93.2	93.2	93.2
	2 AM943647/ Australia	<b>95.4</b>		96.6	95.8	93.9	93.6	94.3	94.3	95.5	95.5	95.5	95.5
	3 JN597006/South Korea	<b>94.3</b>	<b>96.6</b>		95.8	93.6	93.2	93.9	93.2	95.1	95.1	95.1	95.1
	4 KC454286/South Korea	<b>93.1</b>	<b>95.4</b>	<b>96.6</b>		93.6	93.2	93.9	94.3	93.9	93.2	93.2	93.2
	5 GU954430/China	<b>94.3</b>	<b>92.0</b>	<b>90.8</b>	<b>92.0</b>		98.9	95.8	95.1	93.2	93.9	93.9	93.9
	6 AM943646/Hungary	<b>94.3</b>	<b>92.0</b>	<b>90.8</b>	<b>92.0</b>	<b>97.7</b>		96.2	95.5	92.8	93.6	93.6	93.6
	7 JN997392/Hungary	<b>96.6</b>	<b>94.3</b>	<b>93.1</b>	<b>92.0</b>	<b>93.1</b>	<b>93.1</b>		97.0	92.8	92.8	92.8	94.3
	8 KF511797/Taiwan	<b>100</b>	<b>95.4</b>	<b>94.3</b>	<b>93.1</b>	<b>94.3</b>	<b>94.3</b>	<b>96.6</b>		92.0	92.0	92.0	92.0
	9 KM377618/South Korea	<b>92.0</b>	<b>95.4</b>	<b>92.0</b>	<b>90.8</b>	<b>87.4</b>	<b>87.4</b>	<b>89.7</b>	<b>92.0</b>		97.0	97.0	97.0
	10 AY535004/United States	<b>93.1</b>	<b>96.6</b>	<b>93.1</b>	<b>92.0</b>	<b>88.5</b>	<b>88.5</b>	<b>90.8</b>	<b>93.1</b>	<b>98.9</b>		100	97.0
	11 NC023425/ United States	<b>93.1</b>	<b>96.6</b>	<b>93.1</b>	<b>92.0</b>	<b>88.5</b>	<b>88.5</b>	<b>90.8</b>	<b>93.1</b>	<b>98.9</b>	<b>100</b>		97.0
	12 EF206691/ United States	<b>89.7</b>	<b>94.3</b>	<b>90.8</b>	<b>89.7</b>	<b>88.5</b>	<b>88.5</b>	<b>93.1</b>	<b>89.7</b>	<b>94.3</b>	<b>95.4</b>	<b>95.4</b>	
3'UTR /5'UTR	1 MK050107/China		77.5	76.7	74.4	82.9	83.7	54.3	74.4	78.3	76.0	76.0	81.4
	2 AM943647/ Australia	—		89.1	89.9	79.1	80.6	55.0	76.0	79.1	79.8	79.8	78.3
	3 JN597006/South Korea	84.0	—		96.1	81.4	82.2	53.5	82.9	77.5	77.5	77.5	76.7
	4 KC454286/South Korea	88.0	—	96.0		79.1	79.8	52.7	79.1	78.3	76.7	76.7	76.0
	5 GU954430/China	64.0	—	64.0	68.0		97.7	54.3	83.7	85.3	83.7	83.7	85.3
	6 AM943646/Hungary	—	—	—	—	—	—	55.8	81.4	86.0	82.9	82.9	86.0
	7 JN997392/Hungary	—	—	—	—	—	—	—	51.2	55.8	53.5	53.5	55.8
	8 KF511797/Taiwan	80.0	—	88.0	92.0	60.0	—	—	—	79.8	84.5	84.5	82.2
	9 KM377618/South Korea	88.0	—	96.0	100	68.0	—	—	92.0	—	89.1	89.1	93.0
	10 AY535004/USA	88.0	—	96.0	100	68.0	—	—	92.0	100	—	100.0	91.5
	11 NC023425/USA	88.0	—	96.0	100	68.0	—	—	92.0	100	100	—	91.5
	12 EF206691/USA	88.0	—	96.0	100	68.0	—	—	92.0	100	100	100	—

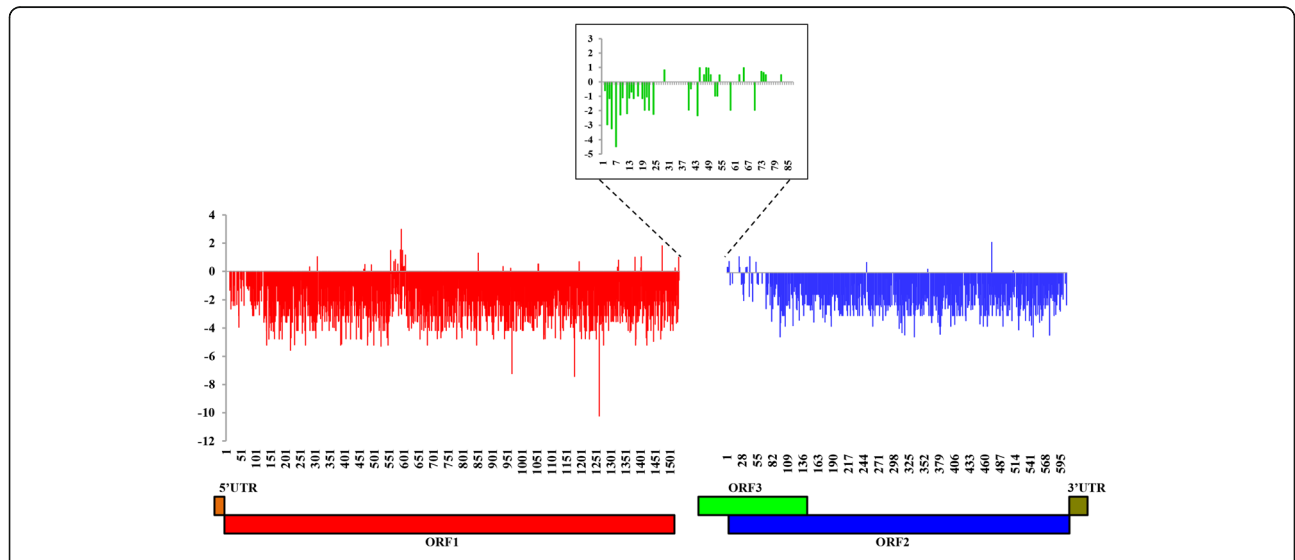
The comparison was done using MegAlign ClustalW analysis. Boldface indicates percentage identities of amino acid sequences. The italic font numbers show the nucleotide sequences identities of 5'UTR among the avian HEV strains. —: Sequence couldn't be obtained from GenBank



**Fig. 2** Phylogenetic analysis based on the complete nucleotide sequences of ORF1 (a) and ORF2 (b) and reference isolates. The trees were constructed based on the maximum likelihood method implemented in PhyML v3.0. Bootstrap values were calculated with 1000 replicates of the alignment. GenBank accession numbers are followed by the name of avian HEV strains. Red dots indicate the avian HEV determined in this study

identified as prevalent in many regions of China [15, 23–26]. In recent years, many new avian HEV strains belonging to a tentatively novel genotype of species *Orthohepevirus* B were detected in chickens in Hebei, Guangdong, Anhui, and Jilin Provinces, China [28]. In this study, a new subtype of genotype 3 avian HEV strain was

determined in broiler chickens with high mortality showing the clinical symptoms of HS syndrome through unbiased high-throughput sequencing. All of these results suggest that there is a larger diversity of avian HEV circulating in China, and avian HEV infection should be monitored as an emerging disease agent on chicken farms.



**Fig. 3** Differences between non-synonymous and synonymous ( $dN-dS$ ) rates plotted for avian HEV ORF1, ORF2 and ORF3.  $dN-dS < 0$  indicates a negatively selected site.  $dN-dS > 0$  indicates a positively selected site

The comparative analysis of individual ORFs among avian HEV strains, including CaHEV-GDSZ01, showed 80.8–85.8%, 83.5–86.7%, and 92.8–95.8% nucleotide sequence identities for ORF1, ORF2, and ORF3, respectively. Comparison of the amino acid sequence showed higher identities: 92.0–95.0% for ORF1, encoding a non-structural polyprotein; 98.0%–99.3 for ORF2, encoding a capsid protein; and 89.7–100% for ORF3, encoding a cytoskeleton-associated phosphoprotein. However, the nucleotide sequence identities of ORF3 among all the avian HEV strains were often higher than the amino acid sequence identities, suggesting that non-synonymous mutations may occur more frequently in this fragment (Table 2).

In the phylogenetic analysis of the complete ORF1 and ORF2, CaHEV-GDSZ01 was grouped with other viruses of the *Orthohepevirus* B species isolated from chickens, which were clustered into the genotype-3 subclade with two avian HEVs isolated in China and Hungary. These results were consistent with the results of sequence comparison analyses of ORF1 and ORF2 indicating that CaHEV-GDSZ01 showed the highest nucleotide sequence identity with the isolate from Hungary (accession no. AM943646; 85.8% identity) and the isolate from China (accession no. GU954430; 86.7% identity). Taken together, these results suggest that CaHEV-GDSZ01 may be a new subtype of avian HEV genotype 3.

In selection pressure analyses, negatively selected sites were observed predominantly in ORF1, ORF2, and ORF3, and no positively selected sites were observed, suggesting that negative selection was predominant in ORF1, ORF2, and ORF3. Additionally, the mean  $dN-dS$  values of the three ORFs were  $-2.04$ ,  $-1.61$ , and  $-0.37$ , respectively, also reflecting the predominance of negative selection. Therefore, the results showed that the micro-evolution of avian HEV seems to be driven by negative selection ( $dN < dS$ ) of all the three ORFs (summarized in Fig. 3). This conclusion is consistent with the expected behaviour of a small-genome virus, in which most components are probably essential for viral viability.

## Conclusion

Through unbiased high-throughput sequencing, we identified a new subtype of genotype 3 avian HEV from a poultry farm experiencing high mortality of broiler chickens showing HS syndrome. In combination with previous studies, our results suggest a larger diversity of avian HEV circulating in China, and much greater efforts should be exerted towards the surveillance of avian HEV.

## Methods

### Case history and clinical sample collection

In May 2018, many 130-day-old broiler chickens on a poultry farm in Shenzhen, Guangdong Province, China,

experienced a sudden mass die-off. The species of these chickens was Qingyuan partridge chicken, one of the three prevalent chicken species in Guangdong Province, and all the chickens were free roaming. All the deceased chickens died naturally, and the clinical symptoms and postmortem lesions presented by the affected chickens were recorded. The deceased chickens were subjected to necropsy, and tissue samples from heart, liver, spleen, lung, kidney, duodenum, brain, pancreas, and bursa of fabricius were aseptically collected and snap-frozen in liquid nitrogen immediately and stored at  $-80^{\circ}\text{C}$  for further use.

### Total RNA extraction, library construction and NGS sequencing

Approximately 100 mg of liver tissue from a chicken with hepatitis-splenomegaly syndrome was homogenized in 600  $\mu\text{L}$  of lysis buffer by using a TissueRuptor instrument (Qiagen). Total RNA was extracted by using an RNeasy Plus Minikit according to the manufacturer's instructions. The quantity and quality of extracted RNA was evaluated with a NanoDrop 2000 (Thermo Fisher Scientific, Waltham, USA).

Ten RNA samples of liver tissues were pooled as one mixed sample in equal-mass amounts. RNA library preparation was conducted following the methodology previously described by Pettersson et al. [29]. Briefly, the host ribosomal RNA (rRNA) was depleted by using a Ribo-Zero-Gold (Epidemiology) kit (Illumina Inc., USA). Subsequently, a library was constructed based on the rRNA-depleted RNA samples using a TruSeq total RNA library preparation kit (Illumina). Library cDNA levels were quantified before, during and after library preparation with Qubit (ThermoFisher Scientific) high-sensitivity RNA/DNA assays, and the fragment sizes were checked with a Bioanalyzer (Agilent Technologies). Paired-end (150-bp) sequencing was then performed on the Illumina HiSeq2500 platform.

### Transcriptome assembly and contig annotation

Transcriptome analyses were conducted following the methodology previously described by Shi et al. [30]. Sequencing reads were demultiplexed and trimmed for quality with Trimmomatic [31] before de novo assembly using Trinity [32]. The resulting contigs were compared against the entire nonredundant protein (nr) database downloaded from GenBank by using BLASTX with an E-value cutoff of  $1\text{E}-5$ . Unmatched sequence reads were assembled using Trinity. To confirm the assembly results, the reads were mapped back to the target contigs with Bowtie2 [33], and any assembly errors were inspected by using Integrated Genomics Viewer (IGV) [34]. The final sequences of the virus genomes were

obtained from the majority consensus of the mapping assembly.

#### Sequence confirmation by RT-PCR and 5'/3'RACE

To confirm the sequences of the virus genomes and fill the gaps generated in the merging of viral contigs with unassembled overlaps by using the SeqMan program implemented in the Lasergene software package v7.1 (DNASTAR), reverse transcriptase-PCR was conducted. All primers were designed based on the transcriptome assembly sequence, and amplicons were sequenced directly by Sanger sequencing. The 5' and 3' ends of the genome of avian HEV determined in this study were obtained by 5' and 3' rapid amplification of cDNA ends (RACE) using a RACE kit (TaKaRa, China). All primer sequences are available upon request. Sequences were assembled and manually edited to produce the final viral genomes using the SeqMan program (DNASTAR, Madison, WI).

#### Sequence comparison and phylogenetic analysis

DNASTAR's Lasergene 12 Core Suite was used for Sanger sequencing assembly and nucleotide sequence translation. Sequence similarity was evaluated via a BLASTn search in GenBank (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The alignment of the sequences obtained in this study and the existing reference sequences in GenBank was carried out using ClustalW (default parameters) as implemented in the MEGA program, version 6.0 [35].

The phylogenetic trees were estimated following the methodology previously described by Lu et al. [36]. The best-fit evolutionary model for all sequence alignments was determined using jModel Test version [37]. The General Time Reversible (GTR) nucleotide substitution model with a gamma ( $\Gamma$ )-distribution model of among-site rate variation and a proportion of invariable sites (i.e. GTR+ $\Gamma$ +I) were found to be the best-fit model for these sequences. Phylogenetic trees were estimated using the maximum-likelihood (ML) method with bootstrap support values calculated from 1000 replicates implemented in PhyML v3.0 [38]. All phylogenetic trees were mid-point rooted for purposes of clarity only.

#### Selection pressure analyses

The selection pressure of code sites was examined site by site across the entire coding region of the genome using the DataMonkey (<http://www.datamonkey.org>) web server and assessed by calculating the difference between non-synonymous (dN) and synonymous (dS) rates (dN-dS) for each ORF. All the analyses of the three ORFs were performed using the SLAC method and the REV nucleotide substitution bias model, and a *P*-value < 0.1 indicated negatively selected sites.

#### Abbreviations

BLSD: Big liver and spleen disease; HEV: Hepatitis E virus; HS syndrome: Hepatitis-splenomegaly syndrome; ORF: Open reading frame; RdRp: RNA-dependent RNA polymerase; rRNA: Ribosomal RNA; RT-PCR: Reverse transcription-polymerase chain reaction; UTR: Untranslated region

#### Acknowledgements

The authors wish to thank the veterinary practitioners and farmers for their help with sample collection and lab members for their help with internal review of the paper.

#### Funding

This study was supported by Key Laboratory for Preventive Research of Emerging Animal Diseases in Foshan University (KLPREAD201801-09; KLPREAD201801-08), Project of Key Laboratory for Preventive Veterinary Medicine, Education Bureau of Guangdong Province (2014KTSPT037) and Guangdong Science and Technology Plan Project (2012A020100001). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

#### Availability of data and materials

The datasets analysed in the present study are available from the corresponding author upon reasonable request. Additional supporting files can be found in the supplementary material section.

#### Authors' contributions

XZ performed experiments, analysed the data and drafted the manuscript; WL, SY, JG, ZL, SC, WH, and XL prepared samples and helped in laboratory analysis. SH contributed to confirming the results and approved the submitted version, JS contributed to the experimental design, supervised the study, and revised the manuscript. All authors read, commented on and approved the final version of the manuscript.

#### Ethics approval and consent to participate

This study was approved by the Research Ethics Committee of the College of Life Science and Engineering, Foshan University. Animal experiments were performed according to the *Guidance for Experimental Animal Welfare and Ethical Treatment* of the Ministry of Science and Technology of China ([www.most.gov.cn/fggw/zfwj/zfwj2006/zf06yw/zf06qt/200612/t20061226\\_39235.htm](http://www.most.gov.cn/fggw/zfwj/zfwj2006/zf06yw/zf06qt/200612/t20061226_39235.htm)). There were no vulnerable populations involved, and no endangered species were used in the experiments. Farm managers gave permission for their animal samples to be used in this study, and sample collection was conducted by local authorities and veterinarians.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Author details

<sup>1</sup>Key Laboratory for Preventive Research of Emerging Animal Diseases, Foshan University, Foshan 528231, Guangdong, China. <sup>2</sup>College of Life Science and Engineering, Foshan University, Foshan 528231, Guangdong, China. <sup>3</sup>Department of Medical Microbiology and Parasitology, Zhejiang University School of Medicine, Hangzhou 310058, Zhejiang, China.

Received: 5 November 2018 Accepted: 25 April 2019

Published online: 06 May 2019

#### References

- Meng XJ. Hepatitis E as a zoonosis. In: Thomas, HC., Zuckermann, A., Lemon, S. (Eds.), *Viral Hepatitis*, 3rd edn. Oxford: Blackwell Publishing Ltd; 2005. p. 611-623.

2. Echevarria JM, Gonzalez JE, Lewis-Ximenez LL, Dos SD, Munne MS, Pinto MA, Pujol FH, Rodriguez-Lay LA. Hepatitis E virus infection in Latin America: a review. *J Med Virol*. 2013;85:1037–45.
3. Smith DB, Simmonds P, Jameel S, Emerson SU, Harrison TJ, Meng XJ, et al. Consensus proposals for classification of the family Hepeviridae. *J. Gen. Virol*. 2014;95:2223–32.
4. Koonin EV, Gorbalenya AE, Purdy MA, Rozanov MN, Reyes GR, Bradley DW. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. *Proc Natl Acad Sci U S A*. 1992;89:8259–63.
5. Tam AW, Smith MM, Guerra ME, Huang CC, Bradley DW, Fry KE, Reyes GR. Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length viral genome. *Virology*. 1991;185:120–31.
6. Purcell RH, Emerson SU. Hepatitis E virus. In: Knipe DM, Howley PM, Griffin DE, Martin MA, Lamb RA, Roizman B, Straus SE, editors. *Fields virology*, fourth ed. Philadelphia: Lippincott Williams and Wilkins; 2001. p. 3051–61.
7. Tyagi S, Korkaya H, Zafullah M, Jameel S, Lal SK. The phosphorylated form of the ORF3 protein of hepatitis E virus interacts with its non-glycosylated form of the major capsid protein, ORF2. *J Biol Chem*. 2002;277:22759–67.
8. Huang YW, Opriessnig T, Halbur PG, Meng XJ. Initiation at the third in-frame AUG codon of open reading frame 3 of the hepatitis E virus is essential for viral infectivity in vivo. *J Virol*. 2007;81:3018–26.
9. Meng XJ, Purcell RH, Halbur PG, Lehman JR, Webb DM, Tsareva TS, Haynes JS, Thacker BJ, Emerson SU. A novel virus in swine is closely related to the human hepatitis E virus. *Proc Natl Acad Sci U S A*. 1997;94:9860–5.
10. Syed SF, Zhao Q, Umer M, Alagawany M, Ujjan IA, Soomro F, Bangulzai N, Baloch AH, Abd EM, Zhou EM, Arain MA. Past, present and future of hepatitis E virus infection: zoonotic perspectives. *Microb Pathog*. 2018;119: 103–8.
11. Sridhar S, Teng J, Chiu TH, Lau S, Woo P. Hepatitis E virus genotypes and evolution: emergence of camel hepatitis E variants. *Int J Mol Sci*. 2017;18. <https://doi.org/10.3390/ijms18040869>.
12. Doceul V, Bagdassarian E, Demange A, Pavio N. Zoonotic hepatitis E virus: classification, animal reservoirs and transmission routes. *Viruses*. 2016;8(10): 270.
13. Haqshenas G, Huang FF, Fenaux M, Guenette DK, Pierson FW, Larsen CT, Shivaprasad HL, Toth TE, Meng XJ. The putative capsid protein of the newly identified avian hepatitis E virus shares antigenic epitopes with that of swine and human hepatitis E viruses and chicken big liver and spleen disease virus. *J. Gen. Virol*. 2002;83:2201–9.
14. Meng XJ. Recent advances in hepatitis E virus. *J Viral Hepat*. 2010;17:153–61.
15. Zhao Q, Liu B, Sun Y, Du T, Chen Y, Wang X, Li H, Nan Y, Zhang G, Zhou EM. Decreased egg production in laying hens associated with infection with genotype 3 avian hepatitis E virus strain from China. *Vet Microbiol*. 2017; 203:174–80.
16. Bilic I, Jaskulska B, Basic A, Morrow CJ, Hess M. Sequence analysis and comparison of avian hepatitis E viruses from Australia and Europe indicate the existence of different genotypes. *J Gen Virol*. 2009;90:863–73.
17. Kwon HM, Sung HW, Meng XJ. Serological prevalence, genetic identification, and characterization of the first strains of avian hepatitis E virus from chickens in Korea. *Virus Genes*. 2012;45:237–45.
18. Moon HW, Lee BW, Sung HW, Yoon BI, Kwon HM. Identification and characterization of avian hepatitis E virus genotype 2 from chickens with hepatitis-splenomegaly syndrome in Korea. *Virus Genes*. 2016;52:738–42.
19. Sprygin AV, Nikonova ZB, Zinyakov NG. Avian hepatitis E virus identified in Russian chicken flocks exhibits high genetic divergence based on the ORF2 capsid gene. *Avian Pathol*. 2012;41:459–63.
20. Zhao Q, Zhou EM, Dong SW, Qiu HK, Zhang L, Hu SB, Zhao FF, Jiang SJ, Sun YN. Analysis of avian hepatitis E virus from chickens, China. *Emerg Infect Dis*. 2010;16:1469–72.
21. Banyai K, Toth AG, Ivanics E, Glavits R, Szentpali-Gavaller K, Dan A. Putative novel genotype of avian hepatitis E virus, Hungary, 2010. *Emerg Infect Dis*. 2012;18:1365–8.
22. Hsu IW, Tsai HJ. Avian hepatitis E virus in chickens, Taiwan, 2013. *Emerg. Infect Dis*. 2014;20:149–51.
23. Liu B, Fan M, Zhang B, Chen Y, Sun Y, Du T, Nan Y, Zhou EM, Zhao Q. Avian hepatitis E virus infection of duck, goose, and rabbit in Northwest China. *Emerg Microbes Infect*. 2018;7:76.
24. Zhao Q, Sun Y, Zhao J, Hu S, Zhao F, Chen F, Clavijo A, Zhou EM, Xiao Y. Development and application of an indirect ELISA for detection of antibodies against avian hepatitis E virus. *J Virol Methods*. 2013;187:32–6.
25. Yang S, Wang L, Sun S. Natural infection with avian hepatitis E virus and Marek's disease virus in Brown layer chickens in China. *Avian Dis*. 2016;60: 698–704.
26. Geng J, Wang L, Wang X, Fu H, Bu Q, Liu P, Zhu Y, Wang M, Sui Y, Zhuang H. Potential risk of zoonotic transmission from young swine to human: seroepidemiological and genetic characterization of hepatitis E virus in human and various animals in Beijing, China. *J Viral Hepat*. 2011;18:e583–90.
27. Sun Y, Du T, Liu B, Syed SF, Chen Y, Li H, Wang X, Zhang G, Zhou EM, Zhao Q. Seroprevalence of avian hepatitis E virus and avian leucosis virus subgroup J in chicken flocks with hepatitis syndrome, China. *BMC Vet Res*. 2016;12:261.
28. Su Q, Li Y, Zhang Y, Zhang Z, Meng F, Cui Z, Chang S, Zhao P. Characterization of the novel genotype avian hepatitis E viruses from outbreaks of hepatic rupture haemorrhage syndrome in different geographical regions of China. *Transbound Emerg Dis*. 2018. <https://doi.org/10.1111/tbed.12987>.
29. Pettersson JHO, Shi M, Bohlin J, Eldholm V, Brynildsrud OB, Paulsen KM, Andreassen Å, Holmes EC. Characterizing the virome of Ixodes ricinus ticks from northern Europe. *Sci Rep*. 2017;7(1):10870.
30. Shi M, Neville P, Nicholson J, Eden JS, Imrie A, Holmes EC. High-resolution metatranscriptomics reveals the ecological dynamics of mosquito-associated RNA viruses in Western Australia. *J Virol*. 2017;91(17):e00680–17.
31. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
32. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29:644–52.
33. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods*. 2012;9:357–9.
34. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*. 2013;14:178–92.
35. Tamura K, Stecher G, Peterson D, Filipi A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 2013;30:2725–9.
36. Lu M, Tian JH, Yu B, Guo WP, Holmes EC, Zhang YZ. Extensive diversity of rickettsiales bacteria in ticks from Wuhan, China. *Ticks Tick Borne Dis*. 2017; 8(4):574–80.
37. Posada D. jModelTest: phylogenetic model averaging. *Mol Biol Evol*. 2008; 25:1253–6.
38. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 2010;59:307–21.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

