Infection, Genetics and Evolution 55 (2017) 277–280

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Short communication

Complete genome sequences and phylogenetic analysis of encephalomyocarditis virus strains isolated from pigs and rats origin

Huimin Liu[a], Yongtao Li[b], Guozhi Zhang[a], Siyao Sang[a], Chuanqing Wang[b], Hongtao Chang[b,*]

[a] College of Life Science, Henan Agricultural University, Zhengzhou, Henan Province, China
[b] College of Animal Science and Veterinary Medicine, Henan Agricultural University, Zhengzhou, Henan Province, China

ARTICLE INFO

Keywords:
Encephalomyocarditis virus
Phylogenetic analysis
Comparative analysis

ABSTRACT

In order to evaluate the genetic variability of encephalomyocarditis virus (EMCV), the whole genomes of six EMCV field isolates originating from pigs and rats origin in different regions of central China, were phylogenetically and comparatively analyzed. Phylogenetic analysis of whole genome sequences, open reading frame (ORF), capsid coding region (CCR) and VP3/VP1 using neighbor-joining analysis revealed that these isolates belonged to lineage 1. Nucleotide sequences of six isolates showed more than 99% pairwise identity rates, and the sequences of isolates from pig and boar in the same region were completely identical with each other, without any genetic deletion or insertion. From comparative analysis of variability of each EMCV protein coding region, 3D and VP3 regions showed that the highest average identity rates, and was confirmed as highly conserved. In contrast, the protein coding regions 3A and 3B was confirmed to be highly variable region with the lowest average identity rate. Our data confirmed that the EMCV strains isolated from pigs and rats origin had high homology with each other, which implied rats may play an important role in EMCV transmission between domestic pigs and wild boars.

1. Introduction, methods and results

Encephalomyocarditis virus (EMCV; family *Picornaviridae*, genus *Cardioviruses*), is a positive sense single-stranded RNA virus, of approximately 7.8 kb in length (Palmenberg et al., 1984). The genome of EMCV is composed of an open reading frame that encodes the four structural proteins, composing the capsid coding region (CCR) and eight nonstructural proteins. The coding region flanked by two untranslated regions (UTR). The 5′ UTR is between 800 and 1200 nucleotides (nt) long and contains the internal ribosome entry site (IRES) and poly(C) tract, while the 3′ UTR is about 120 nt long and composed of short stem-loop structures followed by a poly(A) tail of variable length (20 to 70 nucleotides) (Carocci and Bakkali-Kassimi, 2012; Duke et al., 1992; Jang et al., 1989).

Although it is considered a rodent virus, EMCV has a wide host-range among domestic and wild animals, and its infection is distributed worldwide (Feng et al., 2015; LaRue et al., 2003; Liu et al., 2016a; Luo et al., 2017; Maurice et al., 2016; Oberste et al., 2009). Moreover, the evidence regarding human EMCV infection implies that the virus poses a potential risk to public health (Feng et al., 2015; Oberste et al., 2009). Till now, approximately 40 strains of EMCV have been isolated worldwide from many different animals, and phylogenetic analysis revealed the origin of EMCV outbreaks has been assumed to be local rodent populations. To extend these findings and gain a better understanding of the genetic relationships and evolution of EMCV, the complete genomic sequences of six EMCV strains from pigs and rats origin were determined and the phylogenetic relationships were analyzed.

The tissues of brain, heart, spleen, lung and kidney were collected from 62 dead pigs (domestic pigs, domestic and semi-captive wild boars) that exhibited with acute myocarditis from different pig farms in central China between 2012 and 2015. Meanwhile, the rats around pig farms were also collected in spite of no apparent clinical symptoms. Viral genomic RNA and DNA were separately extracted from the homogenized tissues with the PureLink® Viral RNA/DNA Extraction Kit according to the manufacture's protocol (Thermo Fisher Scientific, America). DNA and RNA were subjected to PCR or RT-PCR for the detection of a panel of potential pathogens including EMCV, porcine parvovirus (PPV), pseudorabies virus (PRV), classical swine fever virus (CSFV), porcine circovirus (PCV2) and porcine reproductive and respiratory syndrome virus (PRRS). The results of PCR indicated that the tested specimens were positive for EMCV but negative for other potential pathogens tested, moreover, no pathogenic bacteria was isolated (data not shown). The purified PCR product was sequenced by a

* Corresponding author.
E-mail address: ndcht@163.com (H. Chang).

http://dx.doi.org/10.1016/j.meegid.2017.09.032
Received 28 August 2017; Received in revised form 7 September 2017; Accepted 27 September 2017
Available online 28 September 2017

Table 1
EMCV isolates from Henan province between 2012 and 2015.

Virus designation	Geographic origin	Gene accession No.	Species	Isolation/ submission year	Length (bp)	5′-UTR (bp)	Poly(C)	ORF (bp)	3′-UTR (bp)	VP1 aa62	VP1 aa231	VP2 aa156	League
JZ1201	Jiaozuo	KF836386	Pig	2012	7735	724	$C_7TCTC_3TC_{10}$	6879	132	N	K	R	1
JZ1202	Jiaozuo	KF836387	Boar	2012	7735	724	$C_7TCTC_3TC_{10}$	6879	132	N	K	R	1
JZ1203	Jiaozuo	KF836388	Rat	2012	7725	708	C_7	6879	138	N	K	R	1
YM13	Yima	KF836389	Pig	2013	7729	713	C_{12}	6879	137	N	K	R	1
YY13	Yuanyang	KF836390	Pig	2013	7724	714	C_{10}	6879	131	N	K	R	1
HN13	Xinyang	KF771002	wild boar	2013	7725	709	C_7	6879	137	N	K	R	1

commercial corporation (Sangon, Shanghai, China), and sequence alignment indicated that the isolates matched other EMCV isolates in GenBank. Thus, we formally named the six EMCV isolates as JZ1201, JZ1202, JZ1203, YM13, YY13, HN13 (Table 1), and cultured on baby hamster kidney 21 (BHK-21) cells. After three passages on BHK-21 cells, a distinct CPE was observed by cell rounding, pyknosis, and degeneration of the cell monolayer.

Viral genomic RNA of the EMCV isolates were extracted from the cultural supernatant of infected BHK-21 cells, and the whole genomes were amplified and sequenced. The whole ORF of EMCV isolates were amplified and sequenced using a set of specific primers (Liu et al., 2016a). Both 5′- and 3′-UTRs were amplified by the 5′ and 3′ Full RACE Core sets (TaKaRa Biotechnology Company, Dalian, China). DNA fragments corresponding to the whole ORF of EMCV strains were amplified and cloned into the pMD19-T vector. Amplicons were sequenced by a commercial corporation (Sangon, Shanghai, China), and the sequences assembly were carried out using the SeqMan program of the DNASTAR software. Full length of the whole genome sequences of six EMCV strains with a 5′-UTR, a 3′-UTR and a large ORF were amplified and have been submitted to GenBank (Table 1).

Nucleotide and amino acid sequences of each other of the 11 protein coding regions of the 32 virus isolates (Table 2) were compared and pairwise identity rates of them were calculated (Table 3). Pairwise identity rates of these 32 virus isolates ranged from 79.9% to 99.9%

Table 2
EMCV isolates used in the study.

Virus designation	Gene accession No.	Geographic origin	Species	Isolation/ submission year
Mengo-M	L22089	Uganda	Monkey	1948/1993
K3	EU780148	Korea	Swine	1990/2008
K11	EU780149	Korea	Swine	1990/2008
CBNU	DQ517424	Korea	Swine	2006/2006
BEL-2887A/91	AF356822	Belgium	Swine	1991/2001
PV2	X87335	Germany	Swine	1985/1995
PV21	X74312	Germany	Swine	1993/1993
RD1338	JX257003	Germany	Wood mouse	2005/2012
EMCV-30	AY296731	USA	Swine	1987/2003
D variant	M37588	USA	Swine	1980/1988
EMCV-B	M22457	USA	Swine	1980/1989
EMCV-D	M22458	USA	Swine	1980/1989
EMCV-R	M81861	USA	Chimpanzee	1945/2000
pEC9	DQ288856	USA	Mice	1995/2005
BJC3	DQ464062	China	Swine	2005/2006
HB1	DQ464063	China	Swine	2005/2006
GX0602	FJ604853	China	Mice	2006/2009
GX0601	FJ604852	China	Swine	2006/2009
NJ08	HM641897	China	Swine	2008/2010
GXLC	FJ897755	China	Swine	2008/2009
HB10	JQ864080	China	Swine	2010/2012
FJ13	KF293299	China	Tigers	2013/2013
BD2	KF709977	China	Swine	2013/2013
Sing-M105-02	KC310738	Singapore	Orangutan	2002/2012
Sing-M100-02	KC310737	Singapore	Orangutan	2002/2012
GS01	KJ524643	China	Swine	2014/2014

Table 3
Pairwise identity rates of each protein coding region of the six virus isolates and other EMCV reference isolates.

Genome region[a]		No. of positions aligned	Average identity rate (%)
ORF	nt	6678	95.32 (79.9–99.9)
	aa	2226	98.33 (93.5–99.9)
VP4	nt	210	95.71 (77.1–100)
	aa	70	98.86 (95.7–100)
VP2	nt	768	95.02 (79.0–99.9)
	aa	256	99.36 (97.7–100)
VP3	nt	693	96.11 (79.7–100)
	aa	231	99.23 (97.4–100)
VP1	nt	831	95.76 (80.5–100)
	aa	277	98.60 (96.8–99.6)
2A	nt	471	94.21 (70.7–99.6)
	aa	143	98.86 (97.1–100)
2B	nt	408	98.52 (90.0–100)
	aa	136	99.21 (97.1–100)
2C	nt	975	95.39 (81.0–99.6)
	aa	325	98.96 (96.0–99.7)
3A	nt	258	95.33 (72.5–100)
	aa	86	97.73 (90.7–100)
3B	nt	66	94.48 (66.7–100)
	aa	22	96.76 (86.4–100)
3C	nt	615	95.52 (75.3–99.7)
	aa	205	97.91 (88.8–100)
3D	nt	1383	96.74 (83.7–100)
	aa	461	99.44 (94.8–99.8)

[a] nt, nucleotide; aa, amino acid.

(average 95.32%). For these 11 proteins, average identities of amino acid sequences of 3D were the highest, 96.74%. Therefore, among these 11 proteins, 3D was confirmed to be the most highly conserved region. The 3D gene is a RNA-dependent RNA polymerase and plays an important role in virus replication. In contrast, the protein coding regions 3A and 3B were confirmed to be highly variable region with the lowest average identity rate.

We further analyzed the sequences of 5′ UTR in the six isolates and other reference strains. Multiple sequence alignment based on 5′ UTR was completed by using Clustalx software. Sequence difference among 5′ UTRs was mainly in poly(C) tract. JZ1201 and JZ1202 strains from the same pig farm contained a poly(C) tract with a $C_7TCTC_3TC_{10}$ motif in the 5′ UTR, but JZ1203 isolated from rats contains C7 motif (Table 1). However, the 3′ UTR also contained a secondary structure, a pseudoknot, which controls viral RNA synthesis and has been identified in other picornaviruses, and the mean length of 3′ UTR in the six isolates varied ranging from 130 to 138 nt (Table 1). The length of poly(C) tract in Chinese isolates was between 6 nt and 30 nt, and JZ1201 and JZ1202 have the longest poly(C) tract of 24 nt. Previous studies reveal that the length of poly(C) tract in 5′ UTR correlated with the pathogenicity of EMCV, shorter poly(C) exhibited lower pathogenicity. However, EMCV strains isolated by Larue has short poly(C) (7 nt to 10 nt) but high pathogenicity to mice, swine and rhesus (LaRue et al., 2003). The relationship between poly(C) tract and virus pathogenicity is still to be investigated.

Phylogenetic analysis was performed separately on the whole