Infection, Genetics and Evolution 14 (2013) 223–231

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

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

Genetic characterization and molecular clock analyses of the Crimean-Congo hemorrhagic fever virus from human and ticks in India, 2010–2011

Pragya D. Yadav [a], Sarah S. Cherian [b], Divya Zawar [a], Prasad Kokate [a], Rashmi Gunjikar [a], Santosh Jadhav [b], Akhilesh C. Mishra [a], Devendra T. Mourya [a,*]

[a] Microbial Containment Complex, National Institute of Virology, Sus Road, Pashan, Pune 410021, India
[b] Bioinformatics Group, National Institute of Virology, 20-A Dr Ambedkar Road, Pune 411001, India

ARTICLE INFO

Article history:
Received 11 April 2012
Received in revised form 20 July 2012
Accepted 1 October 2012
Available online 27 November 2012

Keywords:
Crimean Congo hemorrhagic fever virus
India
Sequencin
Phylogeny
Reassortment
Molecular clock

ABSTRACT

A nosocomial outbreak of Crimean Congo hemorrhagic fever (CCHF) was reported among humans in Ahmadabad district, Gujarat, India during January, 2011. In the present study we provide the complete genomic sequences of four CCHFV isolates derived from two human patients and two pools of *Hyalomma anatolicum* ticks during the period of this outbreak and the complete S segment sequence of two retrospective human serum samples, positive for CCHFV in 2010. Sequence-based molecular characterization of the Indian CCHFV showed that they possessed the functional motifs known to occur in the S, M and L gene segment products as in other CCHF viruses. The S segment of the six Indian CCHF viruses showed 99.8% nucleotide identity. Notably both tick isolates shared 100% nucleotide identity with one of the Indian human isolates of 2011. Phylogenetic analysis based on the S segment demonstrated that the Indian CCHFV isolates formed a distinct cluster in the Asian-Middle East group IV of CCHF viruses. The S segment was closest to a Tajikistan strain TADJ/HU8966 of 1990 (98.5% nucleotide identity) and was of South-Asia 2 type while the M segment was of type M2. Both M and L segments were closest to an Afghanistan strain Afg09-2990 of 2009 (93% and 98% nucleotide identity, respectively). The Indian isolates were thus identified as a South-Asia 2/M2 far-east virus combination and the differing parental origin in the S and L/M segments is suggestive that it may be an intra-genotypic reassortant. Molecular clock studies further revealed that the ancestry of the viruses was not very recent and dated back to about 33 years on the basis of the S segment while it was about 15 years based on the M segment. Thus though the 2011 outbreak may not have resulted from a very recent introduction, considering that so far there is no evidence of multiple circulating strains in the country, the possibility of a recent re-introduction of the virus from any of the neighboring countries cannot be ruled out. The study thus warrants the need for continued surveillance and increased sampling of CCHFV in different parts of the country.

1. Introduction

The Crimean Congo hemorrhagic fever virus (CCHFV), is a member of the *Bunyaviridae* family and *Nairovirus* genus. The virus is transmitted by *Ixodid* ticks, typically the *Hyalomma* genus (Ergonul, 2006; Hoogstraal, 1979; Whitehouse, 2004) or infected livestock and causes a hemorrhagic fever with a high case fatality rate (Ergonul, 2006). CCHF outbreaks have been reported from many countries of Asia, the Middle East, the Balkans and Africa since the last 60 years (Hoogstraal, 1979; Ergonul, 2006; Tahmasebi et al., 2010; Olschläger et al., 2011; Altaf et al., 1998; Swanepoel, 1994; Whitehouse et al., 2006; Nichol, 2001) with previously unaffected countries such as Turkey, Iran, Greece, Georgia, etc. reporting cases in the last decade (Ergonul et al., 2004; Chinikar et al., 2004; Tahmasebi et al., 2010; Papa et al., 2010; Zakhashvili et al., 2010). By contrast, in India, though the presence of CCHFV antibodies were reported since long, no documented human case was known (Shanmugam et al., 1976; Rodrigues et al., 1986). On the other hand, the presence of another *Nairovirus* namely the Nairobi sheep disease virus (NSDV) (also referred locally as the Ganjam virus) (Nichol, 2001; Marczinke and Nichol, 2002; Yadav et al., 2011) and antibodies against the same have been recorded in animals and humans in India (Dandawate and Shah, 1969; Ghalsasi et al., 1981; Banerjee, 1996; Joshi et al., 1998; Dandawate et al., 1969).

The CCHFV has a negative-sense, single stranded (ss) RNA genome of length approximately 19.2 kb containing the S (small), M (medium), and L (large) segments. The S segment encodes the nucleoprotein (N), the M segment the glycoprotein precursor (GPC), and the L segment the viral polymerase (L) (Bishop, 1996; Ergonul, 2006; Nichol, 2001). The M segment is translated into a

* Corresponding author. Tel./fax: +91 20 25870640.
E-mail addresses: mouryadt@icmr.org.in, dtmourya@gmail.com (D.T. Mourya).

1567-1348/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2012.10.005

polyprotein rapidly cleaved into two glycoprotein precursors, PreGn (140 kDa) and PreGc (85 kDa) (Sanchez et al., 2002). Proteolytic processing of these glycoprotein precursors results in the structural glycoproteins Gn (37 kDa) and Gc (75 kDa) and possible non-structural proteins (mucin, GP160, GP85, GP38, and NSm) (Altamura et al., 2007; Bergeron et al., 2007; Erickson et al., 2007; Nichol, 2001; Sanchez et al., 2002). The L segment encodes the RNA-dependent RNA polymerase gene (Ergonul, 2006).

Earlier studies on phylogenetic analysis of the S segment had shown that CCHFV strains cluster into 6–7 distinct groups. West-Africa in group I, Central Africa in group II, South-Africa and West Africa in group III, Middle-East and Asia in group IV, Europe in group V and Greece in group VI (Deyde et al., 2006; Hewson et al., 2004a; Carroll et al., 2010; Anagnostou and Papa, 2009). The group IV may be split into two distinct groups, Asia 1 and Asia 2 (Hewson et al., 2004a; Chamberlain et al., 2005). Complete genome analysis of CCHFV isolates revealed not only high genetic diversity but recombination and reassortment (Deyde et al., 2006; Hewson et al., 2004b; Lukashev, 2005). Majorly parallel clustering of the S and L segments have been noted while reassortment involving the M segment seems more frequent. Groups M1, M2 and M3 have been reported in the M segment (Hewson et al., 2004b).

The first reported CCHF outbreak in India having a nosocomial nature occurred during January 2011 in Ahmadabad district, Gujarat State, and included six laboratory confirmed cases and among them five fatal cases (Mishra et al., 2011; Mourya et al., 2012). The outbreak was supported by the involvement of both *Hyalomma anatolicum* ticks and livestock (buffalo, sheep and goats). A large number of animals from the western states of Rajasthan and Gujarat were found seropositive against the CCHFV before and during the outbreak (Mourya et al., 2012). To understand whether the 2011 outbreak was due to the introduction of a novel virus into the country or the evolution of locally circulating viruses in the region, we sequenced the complete genomes of two virus isolates from humans and two from *Hyalomma* ticks during the period of the 2011 outbreak as well as the S genes of two retrospective human serum samples of 2010 from the same geographic region. Phylogenetic and molecular clock analyses were also carried out to determine the time frame for the ancestry of the Indian CCHFV.

2. Materials and methods

2.1. Ethics statement

The clinical samples of the 2011 CCHFV human outbreak were referred to the Institute, in the aftermath of the focal nosocomial outbreak of CCHFV in a tertiary care hospital at Ahmadabad, Gujarat. Being an emergency, as an institutional policy, it was exempted from prior approval of the Institutional Human Ethical Committee (IHEC). However, the committee was notified after the outbreak and approval of IHEC (letter No. 110 (01)/EC-1/866) was obtained. In addition, our study (project No. HCL01/NIV15/2010) is also approved by the Institutional Animal Ethical Committee (IAEC), a national committee called "Committee for the Purpose of Control & Supervision of Experiments on Animals" (CPCSEA) under the Ministry of Environment and Forests, Government of India, permitting the use of infant and adult mice as laboratory animals for isolation of virus and development of antibodies, respectively.

2.2. CCHFV isolation from Human serum samples and ticks and viral RNA extraction

Serum samples were obtained from two patients that were found to be positive for CCHFV by qRT-PCR and both of which were fatal cases of this outbreak. Two pools of *Hyalomma anatolicum* ticks that were positive for CCHFV by qRT-PCR during the period of the 2011 outbreak from the same region were also available (Mishra et al., 2011; Mourya et al., 2012). Two additional retrospective CCHFV qRT-PCR positive human serum samples of 2010 from the same area were also processed for RT-PCR and sequencing of complete S segment. After processing the above mentioned materials by standard laboratory procedures, the individual samples were inoculated in *Vero E6* cells under biosafety level 3 (BSL-3) conditions. The inoculated cell cultures were observed for cytopathic effects (CPE) following which the virus was harvested by repeated freeze–thaw cycles. The virus supernatant was aliquoted and stored at −86 °C until needed. One-hundred microliters of CCHFV infected tissue culture fluid was treated with 1 ml of Trizol LS isolation reagent (Roche Diagnostics, Indianapolis, IN, USA) for 20 min and further addition of 200 μl of isopropanol: chloroform (Merck) lysis buffer. The tubes were incubated in ice and vortexed few times. The aqueous phase was separated after centrifugation at 14,000*g* for 15 min at 4 °C and transferred to a tube and mixed with 70% alcohol. This virus extract was transferred from BSL-3 to BSL-2 laboratory (Microbial Containment Complex, National Institute of Virology, Pune) after surface decontamination and ducked out using Dunk tank. Viral RNA extraction was completed in a biosafety cabinet by following the protocol of QIAamp Viral RNA Mini kit (QIAGEN, Valencia, CA). RNA was eluted in 50 μl of nuclease free water and stored in 10 μl aliquots for RT-PCR amplification.

2.3. RT-PCR and sequencing

Primers were synthesized based on previously described primer sequences (Deyde et al., 2006). Additional primers were designed (sequences are available upon request), based on available CCHFV sequences in GenBank, for filling in the gaps during sequencing of the complete genome. As large fragment amplification did not work, smaller PCR products (1–3 kb) were amplified using a one step RT-PCR with the Superscript III single step RT-PCR system and Platinum Taq High fidelity (Invitrogen, CA). The RT-PCR conditions for the L, M and S segments were 30 min at 50 °C, 5 min at 94 °C and 35 reaction cycles with 15 s at 94 °C, 30 s at 52 °C and 2–3 min extension at 68 °C based on the template size. The amplified products were visualized by ethidium bromide agarose gel staining and extracted from the gels using QIAquick Gel Extraction Kit (Qiagen, Hilden). Cyclic PCR products obtained by 25 reaction cycles with 1 min at 96 °C, 5 s at 45 °C and 4 min at 60 °C, using ABI Big-Dye 3.1 dye chemistry (Applied Biosystems, Foster City, CA) were purified using Dyex 2.0 kit (Qiagen) and the sequencing was performed using the ABI 3100 automated DNA Analyzer.

The GenBank accession numbers of the sequences of the Indian CCHFV isolates are as follows: S segment (JN572087–JN572089, JF922681, JF922679 and JF922674); M segment (JN572083–JN572086) and L segment (JN572090–JN572092 and JN627865) (Table 1).

2.4. Sequence and phylogenetic analysis

The sequences obtained for the S, M and L segments were curated and assembled using KODON v 2.01 software and aligned using CLUSTAL W and BioEdit software version 5.0.6 (North Carolina State University, USA). KODON was also used for translating the nucleotide sequences to amino acid for obtaining the open reading frame (ORF) of the S, M and L segments. Bioinformatics resources at Expert Protein Analysis System (Expasy), Swiss institute of Bioinformatics, ScanProsite and PredictProtein were used for prediction of sequence motifs and conserved patterns in the encoded proteins of the M and L segments, SignalIP was used to predict the signal sequence cleavage site while HMMTOP was used