Infection, Genetics and Evolution 18 (2013) 113–124

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

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

Genetic characterization of E2 region of Chikungunya virus circulating in Odisha, Eastern India from 2010 to 2011

Abhipsa Sahu, Biswadeep Das, Mumani Das, Aparna Patra, Sudarsan Biswal, Santanu K. Kar, Rupenangshu K. Hazra *

Regional Medical Research Centre, Chandrasekharpur, Bhubaneswar, Odisha 751023, India

ARTICLE INFO

Article history:
Received 7 January 2013
Received in revised form 10 April 2013
Accepted 29 April 2013
Available online 16 May 2013

Keywords:
Chikungunya
Genetic analysis
T cell epitope
Positive selection
Transmembrane helix
Structural modeling

ABSTRACT

Chikungunya virus (CHIKV) infection has caught attention yet again as it rages around the globe affecting millions of people. The virus caused epidemic outbreaks affecting more than 15,000 people in Odisha, Eastern India since 2010. In this study, complete genetic characterization of E2 gene of CHIKV circulating in Odisha from 2010 to 2011 was performed by virus isolation, RT-PCR, molecular phylogenetics and bioinformatics methods. Phylogenetic analyses revealed the circulation of Indian Ocean Lineage (IOL) strains of ECSA genotype of CHIKV in Odisha. Several mutations were detected in the E2 gene, viz. E2-R82G, E2-L210Q, E2-I211T, E2-V229I and E2-S375T which had various adaptive roles during the evolution of CHIKV. The CHIKV E2 peptide ^{57}KTDDSHD^{63} was predicted to be the most probable T-cell epitope and peptide ^{84}FVRTSAPCT^{92} predicted to be the common T and B cell epitope having high antigenicity. The amino acid positions 356–379 and 365–385 were predicted to be transmembrane helical domains and indicated E2 protein anchorage in intracellular membranes for effective interaction with the host receptors. Positive selection pressure was observed in five specific sites, 210, 211, 318, 375, and 377 which were observed to be fixed advantageously in most viral isolates. Structural modeling revealed that E2 gene of CHIKV was composed of 3 domains and the major adaptive mutations were detected in domain B, which can modulate binding of CHIKV to host cells, while the transmembrane domain in domain C and the epitopes were located in domain A, which was found to be most conserved. This is the first report from Eastern India demonstrating a predictive approach to the genetic variations, epitopic regions and the transmembrane helices of the E2 region. The results of this study, combined with other published observations, will expand our knowledge about the E2 region of CHIKV which can be exploited to develop control measures against CHIKV.

1. Introduction

Chikungunya virus (CHIKV), an *Alphavirus* is of serious concern in the recent times because of its increased virulence and high infectivity potential to certain mosquito species, like *Aedes albopictus*, which has become one of the most efficient vector of CHIKV in most places of India (Das et al., 2012). CHIKV is known to have emerged from the sylvatic cycle in Africa. In India, it was reported as early as in 1963 (Shah et al., 1964) and later was thought to have disappeared from the sub-continent (Pavri, 1986). Three genotypes of CHIKV reported so far are the Asian; West African; and East, Central and South African (ECSA) genotypes (Kalantri et al., 2006). With the identification of the ECSA genotype of the virus from a field-caught mosquito in 2000 (Mourya et al., 2001) and being responsible for the 2005–2007 epidemics in Asia (Yergolkar et al., 2006), the disease regained attention as a re-emerging infection and has been responsible for unprecedented outbreaks in the Indian Ocean (Schuffenecker et al., 2006) and in India where millions of people have been infected since 2005 (Epstein, 2007). Recent studies based on genetic mutations in CHIKV genome showed the evolution of highly virulent Indian Ocean Lineage (IOL) from ECSA genotype (Volk et al., 2010), which was further subdivided into Indian Ocean and Indian subcontinent groups, the latter being responsible for massive outbreaks of CHIKV in India in recent time (Das et al., 2012).

The adaptation of CHIKV to a new mosquito vector is thought to be a multistep process, involving many amino acid substitutions in the envelope glycoproteins (Tsetsarkin and Weaver, 2011). The majority of the previously described determinants of vector specificity of different *Alphaviruses* are located within the E2 glycoprotein, circumstantially indicating that the process of *Alphavirus* adaptation to new mosquito species occurs via adaptation to a

* Corresponding author. Tel.: +91 674 2301416/9778740543; fax: +91 6742301351.
E-mail address: rupenkh@yahoo.co.in (R.K. Hazra).

1567-1348/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2013.04.037

specific cell surface receptor expressed in the mosquito (Brault et al., 2002; Myles et al., 2003 and Pierro et al., 2008). The *Alphavirus* E2 region constitutes three distinct subdomains; the N-terminal domain A located at the centre of the 3D structure of E2, the domain B positioned at the distal tip of the ectodomain that may interact with a cellular receptor, and the C-terminal domain C which is situated close to the viral membrane. The binding domains on the receptor are conserved and serve as an ideal site for the design of synthetic peptide vaccines (Voss et al., 2010; Li et al., 2010). Therefore complete genetic characterization of E2 gene will aid in developing immunoprophylactic measures against CHIKV in near future.

Many vaccines effective against CHIKV infection has been reported previously (Khan et al., 2012; Edelman et al., 1979; Harrison et al., 1971; Nakao and Hotta, 1973; Tiwari et al., 2009; Muthumani et al., 2008). However, they had their own limitations. With recent developmental surge of T-cell and B-cell epitope based peptide vaccines and its marked efficiency, exploration of potential epitopic domains of CHIKV is very important for designing such vaccines. To successfully design peptide vaccines employing such peptides as T-cell and B-cell epitopes, sequence conservation at these stretches has to be studied. Maximum specificity of a naturally occurring T-cell epitope depends upon its range of length with the minimal glycosylation probability, which might have resulted in a selective advantage for short peptides to become T-cell epitopes (Szabo et al., 2009). Furthermore epitopes that can bind to both T and B cells are likely to act as good vaccine candidates since they can induce both T and B cell immune responses (Singh and Raghava, 2001, 2003). Prediction of transmembrane domains that play a major role in cell adhesion and migration, and more loosely on cell signaling, is a prerequisite, which can be used as drug targets against CHIKV for treatment of the viral infection (Brehin et al., 2008).

Though advances in molecular and bioinformatics tools have expedited in identifying the molecular aspects as well as the evolution and the adaptive ability of the virus, there still lie certain hidden strategies that remains unexplored (Vrati et al., 1988). The E2 region needs to be investigated thoroughly considering its suspected role in virulence and receptor mediated viral transmission (Strauss et al., 1991). The trend of evolution as well as diversifying selection pressure at the E2 region is particularly important in the context of this reemerging infection in a variable ecological milieu. Molecular investigations of the signature residues at the E2 region in these new CHIKV strains would help to understand the changing pattern of this re-emerging virus. This study emphasizes on the genetic variations in E2 gene and possible T-cell and B-cell epitopes that can aid in the formulation of synthetic peptide vaccines against CHIKV as well as probable transmembrane domains which are responsible for receptor binding during the course of the virus cycle.

Odisha recorded a massive CHIKV outbreak in 2010 affecting more than 15000 people, mainly in the coastal districts and revealed the circulation of IOL strains of CHIKV in both patients and *Aedes* mosquitoes, especially *A. albopictus* (Das et al., 2012). According to Health and Family Welfare Department, Govt. of Odisha, large number Chikungunya cases have been continuously reported from many parts of Odisha for the last 3 years. Hence it is important to reveal the genetic characteristics of CHIKV, particularly in E2 region, in order to pave a way for the development of vaccine/drug in future. Thus a reverse genetics study of genotype/lineage specific signature mutations in the E2 region could highlight the probability of persisting IOL CHIKV strains and its adaptation to vectors, especially *A. albopictus* which is the most important vector in this region. The present study was performed to reveal the complete genetic characteristics of the E2 gene of CHIKV circulating in Odisha.

2. Methods

2.1. Virus isolates

CHIKV was isolated from serum samples of suspected CHIKV patients with classical CHIKV symptoms such as fever, skin rashes and polyarthralgia obtained from many places of Odisha in 2010–2011. 85 human sera were collected and tested for anti CHIKV IgM antibodies by ELISA which showed 23 isolates to be positive for CHIKV, which were further subjected to RT-PCR.

2.2. Viral RNA isolation, RT-PCR and sequencing

The viral RNA was isolated from 140 μl serum using QiaAmp viral RNA mini kit (Qiagen, Hilden, Germany) following manufacturer's instruction. The extracted viral RNA was stored at −80 °C till further processing. Reverse transcription was conducted using the QIAGEN LongRange 2Step RT-PCR Kit (Qiagen, Hilden, Germany) for all the samples. Complementary DNA (cDNA) was synthesized in a 25 μl reaction volume with RT mix comprising 1X Long Range RT buffer, 1 mM each of four deoxynucleotide triphosphate (dNTP), 1 μM of oligodT, 0.04 U long range RNase inhibitor, 1 U long range RT enzyme, RNase free water and 50 ng RNA template. The reaction mixture was incubated at 42 °C for 90 min followed by inactivation of RT enzyme at 85 °C for 5 min. Amplification of viral cDNA was carried out in 20 μl reaction volume with PCR mix containing 10X PCR buffer (10 mM Tris–HCl, 50 mM KCl, 0.1% Triton X-100, 1.5 mM $MgCl_2$), 200 μM each of four dNTP, 1.5 U of *Taq* DNA polymerase (Sigma) along with 25 μM of the forward and the reverse primers specific for the E2 genetic region of CHIKV based on the nucleotide sequence of the S27 strain (Schuffenecker et al., 2006). The cycling conditions used were: 95 °C for 1 min followed by 35 cycles of 95 °C for 30 s, T_A °C (annealing temperatute) (Table 1) for 30 s, and 72 °C for 7 min. The PCR products obtained were run in 1.75% agarose gel containing ethidium bromide and bands visualised under UV-transilluminator. To confirm the RT-PCR positive results, isolation of RT-PCR positive CHIKV isolates was attempted in C6/36 *A. albopictus* cells cultured in Eagle's Minimum Essential Medium (EMEM) supplemented with 10% fetal bovine serum and minimum essential amino acids following the standard virus adsorption technique (Gould, 1991). Briefly, preformed monolayer of cells was washed with plain medium prior to infection. 0.5 ml serum of the RT-PCR positive CHIKV isolates was allowed to adsorb to the cells for 1 h at 37 °C. Following adsorption, the inoculum was replenished with 2 ml of maintenance medium (EMEM with 2% FBS) and maintained at 28 °C under sterile conditions. The cells were harvested upon appearance of cytopathic effects or on 6th day post inoculation, whichever was earlier. Identification of the CHIKV obtained from the cells was confirmed by RT-PCR. The CHIKV isolates ($n = 6$) were then sequenced and subjected to comprehensive genetic analysis. For each sequencing reaction, 50 ng of purified PCR product was mixed with a reaction mixture containing 2.5X sequencing buffer, 5X big dye terminator

Table 1
List of primers with annealing temperatures (T_A °C) used in this study.

Gene	Primer	Sequence	T_A (°C)
C	7910F	TCGAAGTCAAGCACGAAGG	
E2	8670R	GTCTGTCGCTTCATTTCTGATG	51.7
E3	8459F	TGCTTGAGGACAACGTCATGAG	
E2	9240R	TTTGTGATTGGTGACCGCG	51.7
E2	9093F	AGTCCGGCAACGTAAAGATCAC	
6K	9861R	AAAGGTTGCTGCTCGTTCCAC	54.8
E2	9648F	AGTTGTGTCAGTGGCCTCGTTC	
E1	10403R	TAAAGGACGCGGAGCTTAGCTG	57.1