Sanjay Gupta et al. / Journal of Pharmacy Research 2012,5(7),3860-3863

Research Article
ISSN: 0974-6943

Available online through
www.jpronline.info

Molecular cloning and characterization of chikungunya virus genes from indian isolate of 2006 Outbreak

Sanjay Gupta[1*], Namrata Dudha[1], Mohan Babu Appaiahgari[2], Kaushik Bharati[3], Dinesh Gupta[4], Yask Gupta[4], Kapila Kumar[1], Reema Gabrani[1], Sanjeev K. Sharma[1], Amita Gupta[5], Vijay K. Chaudhary[6], Sudhanshu Vrati[2,3*],

[1]*Department of Biotechnology, Jaypee Institute of Information Technology, Noida 201307, India.*
[2]*Vaccine & Infectious Disease Research Centre, Translational Health Science and Technology Institute, Gurgaon 122016, India.*
[3]*Virology Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110067, India.*
[4]*Bioinformatics laboratory, Structural and Computational Biology Group, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi 110067, India.*
[5]*Department of Microbiology, University of Delhi South Campus, New Delhi 110021, India.*
[6]*Department of Biochemistry, University of Delhi South Campus, New Delhi 110021, India.*

Received on:07-04-2012; Revised on: 12-05-2012; Accepted on:16-06-2012

ABSTRACT

ABSTRACT: Re-emergence of Chikungunya virus (CHIKV) and its disease severity in the South-East Asian region necessitates the study of virus strains circulating in this part of the world. Cloning, sequencing and phylogenetic analyses of a CHIKV isolate (IND-06-Guj) from Gujarat 2006 outbreak, India is reported here. The amino acid sequence of IND-06-Guj was found to be similar to those of the 41 CHIKV isolates reported from South Asian region and it differed at 57 positions from CHIKV-S27, the prototype virus belonging to ECSA genotype. Most of the data in the literature pertaining to CHIKV biology has been generated using S27 strain. Considering the observed sequence differences between South Asian and CHIKV-S27 strain, the ready availability of cloned genes representing ECSA genotype shall facilitate studies of CHIKV disease biology in South Asian context.

KEYWORDS. *Chikungunya* virus; CHIKV; South Asia; sequence and phylogenetic analysis

INTRODUCTION:

There have been massive outbreaks of Chikungunya fever in recent years in the South-East Asian region [1]. The frequency and severity of these outbreaks highlight the need to characterize and understand the biology of Chikungunya virus CHIKV isolates. CHIKV is an alphavirus and belongs to the *Togaviridae* family of viruses. The genome of CHIKV is ~12 kb long positive sense single stranded RNA molecule. The viral genome has two large open reading frames encoding four nonstructural (nsP1, nsP2, nsP3 and nsP4) and five structural (capsid, E3, E2, 6K and E1) proteins, and is flanked by 5' and 3' NTR sequences [2]. CHIKV isolates have been categorized into three genotypes worldwide on the basis of phylogenetic analyses of cDNA sequences encoding E1 protein: East/Central/South African (ECSA), West African and Asian [3-5]. Asian genotypes have high degree of conservation whereas African genotypes exhibit wide range of diversity [3, 5]. In the resurgent Chikungunya epidemic that affected more than two million people around the world in last seven years, especially, in the South Asian region, a unique clade of ECSA genotype was found responsible [6]. Studies indicate that CHIKV strains currently circulating in this region also belong to ECSA genotype, and have incorporated various molecular features into their genome making them significantly different from the African prototype virus [5, 7, 8]. Given the widespread nature of CHIKV epidemic in South Asia and attention of scientific community towards sequencing of CHIKV (almost half of the full length CHIKV sequence submissions being from South Asia), there is an urgent need of studying a local strain in South Asia.

***Corresponding author.**
Dr. Sanjay Gupta
Centre for Emerging Diseases,Department of Biotechnology,Jaypee Institute of Information Technology,A-10, Sector-62, NOIDA, U.P. 201307,India
&
Dr. Sudhanshu Vrati
Virology Laboratory,National Institute of Immunology,Aruna Asaf Ali Marg, JNU Complex,New Delhi 110067,India

MATERIAL AND METHODS:

Virus isolation and propagation:

The isolate was obtained from a male patient, about 60 year old belonging to Chandlodia area in Ahmedabad, Gujarat. The patient, during sample collection, was suffering from severe chikungunya-related symptoms. For virus isolation, the blood sample collected from the patient was processed to obtain the serum sample. In order to obtain high titered virus, the serum sample, diluted 1:10 in 1X PBS (pH 7.2), was injected intracranial into 2-3 day old Balb/c mice. The brain tissues from the infected mice showing retarded movements were harvested and the clarified brain lysate was used to infect Vero cell monolayers. The Vero cell cultured virus was subjected to two rounds of plaque purification and then propagated through scale-up Vero cell cultures to prepare the stock virus (IND-06-Guj) for further use.

RNA isolation, RT-PCR and sequencing:

Total RNA (1µg) extracted from IND-06-Guj infected Vero cells was subjected to RT-PCR using Superscript III-One step RT-PCR system and *Pfu* DNA polymerase according to manufacturer's instructions (Invitrogen, Life Technologies, CA and Promega Corp., WI, respectively). Primers were designed to amplify cDNA sequences encoding the viral proteins. The RT-PCR products corresponding to nsP1, nsP2, nsP3, nsp4, capsid, E3, E2, 6K and E1 were cloned into pCR 2.1 TOPO TA vector (Invitrogen) according to the manufacturer's protocol. Clones were sequenced using the Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems, Life Technologies, CA) and the universal M13 forward/reverse primers or the cDNA-specific primers.

Sequence comparison and phylogenetic analyses:

The nucleotide sequences obtained for each gene of IND-06-Guj were assembled in contig using ORF finder for generating nonstructural (GenBank Accession no. for cDNA of IND-06-Guj: nsP1 - JF272473, nsP2 - JF272474,

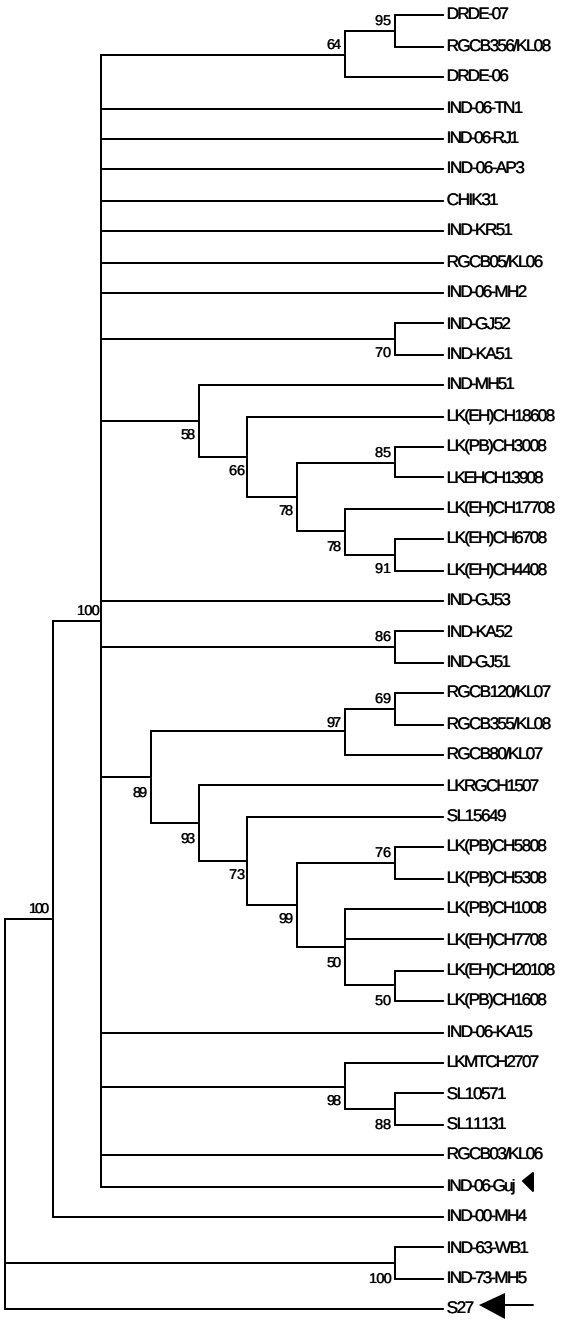

Fig. 1a

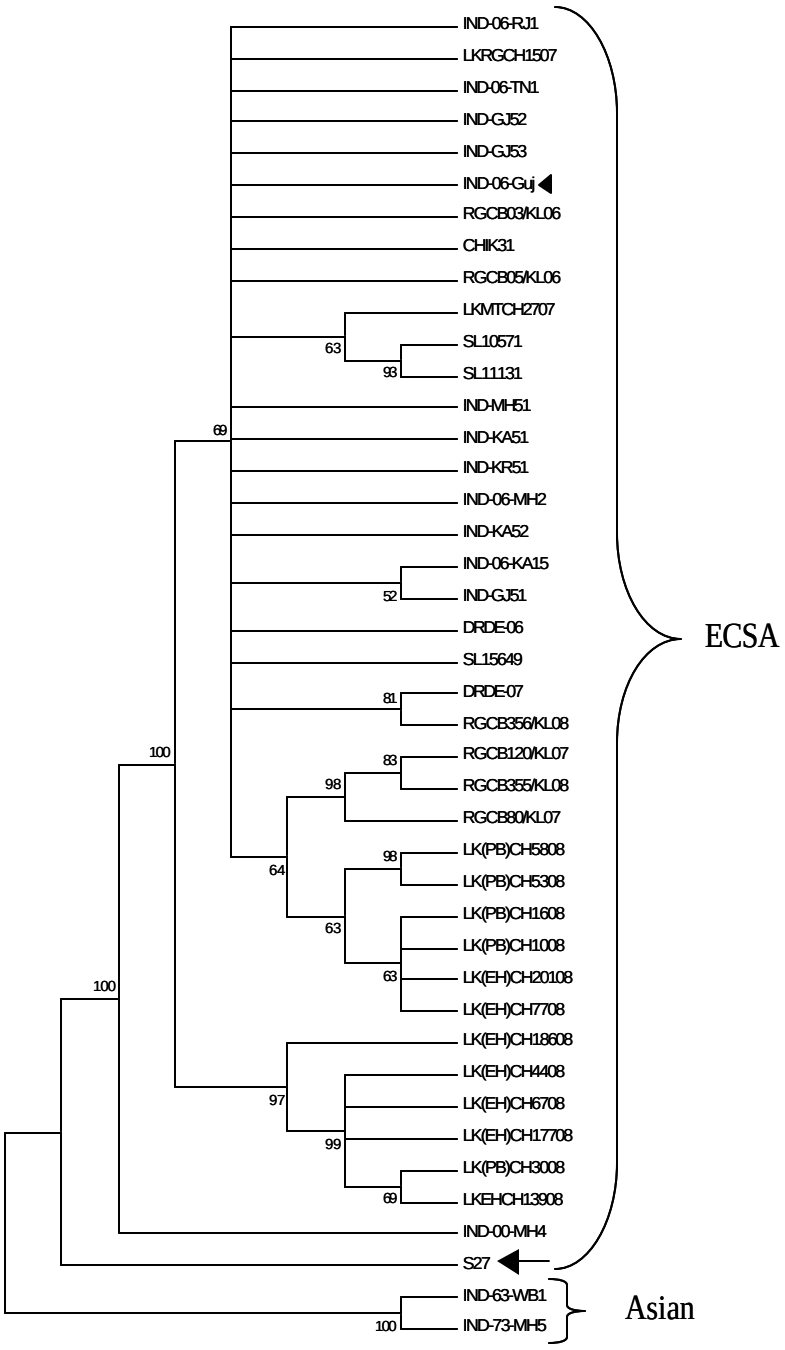

Fig. 1b

Fig. 1 Neighbor-joining tree based on nonstructural protein nucleotide sequences (1a), and structural nucleotide sequences of 43 CHIKV isolates (1b). The bootstrap supports (%), are indicated at the branching points. The details of the isolates are given in Table 2. Arrowhead and arrow represent IND-06-Guj and S27 strain, respectively

nsP3 - JF272475, nsP4 - JF272476) and structural (GenBank Accession no. for cDNA of IND-06-Guj: Capsid - JF272477, E3 - JF272478, E2 - JF272479, 6K - JF272481 and E1 - JF272480) ORFs. The sequence identity of these cloned genes was independently verified with the full length genome sequence reported earlier for the clinical isolate obtained from the patient serum (GenBank accession no. JF274082). The assembled nonstructural and structural polyprotein sequences of IND-06-Guj were compared with the respective polyprotein sequences of the CHIKV-S27 using the NCBI BLAST suite (http://www.ncbi.nlm.nih.gov/BLAST). Sequence alignment of structural polyprotein of IND-06-Guj with the 41 CHIKV isolates reported from South Asia (17 from Sri Lanka and 24 from India) (Table 1) and the CHIKV-S27 using ClustalW 2.0.12 (www.ebi.ac.uk/Tools/clustalW2) and MUSCLE

Table 1: CHIKV isolates used in multiple sequence alignment and phylogenetic analysis. 37 of the 42 strains listed, except S27, are isolated from different regions/outbreaks of South Asia since year 2006

CHIKV Isolate	GenBank Accession No.	Year of Isolation	Origin of isolate	Genotype	Identity score of nonstructural polyprotein with IND-06-Guj	Identity score of structural polyprotein with IND-06-Guj
IND-63-WB1	EF027140	1963	Kolkota	Asian	97	97
IND-73-MH5	EF027141	1973	Barsi, Maharashtra	Asian	97	96
IND-06-AP3	EF027134	2006	Andhra Pradesh	ECSA	99	99
IND-06-KA15	EF027135	2006	Karnataka	ECSA	99	99
IND-06-MH2	EF027136	2006	Maharashtra	ECSA	99	100
IND-06-RJ1	EF027137	2006	Rajasthan	ECSA	99	100
IND-06-TN1	EF027138	2006	Tamilnadu	ECSA	99	99
IND-00-MH4	EF027139	2000	Maharashtra	ECSA	99	98
CHIK31	EU564335	2006	Rajasthan	ECSA	99	100
DRDE-06	EF210157	2006	Hyderabad	ECSA	99	100
DRDE-07	EU372006	2007	Kerala	ECSA	99	99
IND-GJ52	FJ000062	2006	Gujarat	ECSA	99	99
IND-KA52	FJ000063	2006	Karnataka	ECSA	99	99
IND-GJ51	FJ000064	2006	Gujarat	ECSA	99	99
IND-GJ53	FJ000065	2006	Gujarat	ECSA	99	100
IND-KR51	FJ000066	2006	Kerala	ECSA	99	99
IND-MH51	FJ000067	2006	Maharashtra	ECSA	99	99
IND-KA51	FJ000068	2006	Karnataka	ECSA	99	100
RGCB03/KL06	GQ428210	2006	Kerala	ECSA	99	99
RGCB05/KL06	GQ428211	2006	Kerala	ECSA	99	99
RGCB80/KL07	GQ428212	2007	Kerala	ECSA	99	99
RGCB120/KL07	GQ428213	2007	Kerala	ECSA	99	99
RGCB355/KL08	GQ428214	2008	Kerala	ECSA	99	99
RGCB356/KL08	GQ428215	2008	Kerala	ECSA	99	99
LK(EH)CH20108	FJ513679	2008	Sri Lanka	ECSA	99	99
LK(EH)CH18608	FJ513675	2008	Sri Lanka	ECSA	99	99
LK(EH)CH17708	FJ513673	2008	Sri Lanka	ECSA	99	99
LK(EH)CH7708	FJ513657	2008	Sri Lanka	ECSA	99	99
LK(EH)CH6708	FJ513654	2008	Sri Lanka	ECSA	99	99
LK(EH)CH4408	FJ513645	2008	Sri Lanka	ECSA	99	99
LK(PB)CH5808	FJ513637	2008	Sri Lanka	ECSA	99	99
LK(PB)CH5308	FJ513635	2008	Sri Lanka	ECSA	99	99
LK(PB)CH3008	FJ513632	2008	Sri Lanka	ECSA	99	99
LK(PB)CH1608	FJ513629	2008	Sri Lanka	ECSA	99	99
LK(PB)CH1008	FJ513628	2008	Sri Lanka	ECSA	99	99
LKRGCH1507	FJ445428	2007	Sri Lanka	ECSA	99	99
LKMTCH2707	FJ445427	2007	Sri Lanka	ECSA	99	100
LKEHCH13908	FJ445426	2008	Sri Lanka	ECSA	99	99
SL15649	GU189061	2006	Sri Lanka	ECSA	99	99
SL10571	AB455494	2006	Sri Lanka	ECSA	99	99
SL11131	AB455493	2006	Sri Lanka	ECSA	99	99
S27	NC_004162	2001	Africa	ECSA	98	98

[9, 10] was carried out. Phylogenetic relationship of IND-06-Guj with South Asian isolates was assessed by constructing neighbor-joining trees based on distance estimations by the Kimura 2- parameter model of nucleotide substitution [11]. The reliability of the tree branching was tested by bootstrap consensus tree, inferred from 1000 replicates to represent the evolution of the sequences [12]. Branches corresponding to supported partitions represented in less than 50% bootstrap replicates were collapsed. Further, since several interior branches of the tree were having low statistical support, we produced a multifurcating tree by assuming that all the interior branches have a branch length equal to zero (0). Hence branch lengths are not proportional to the number of nucleotide substitutions. Phylogenetic analyses and tree plotting was performed using MEGA4 [13].

RESULTS AND DISCUSSION:

It has been suggested that almost all CHIKV isolates from the subcontinent have undergone genomic microevolution, especially the virus strains belonging to ECSA genotype, leading to enhanced transmissibility and infectivity[5, 7, 14, 15]. Thus, the deduced amino acid sequences of both the structural and nonstructural polyproteins of IND-06-Guj were compared with isolates reported from the South Asian region and also with CHIKV-S27 to understand its similarity with other isolates from the region and also to study its evolution from CHIKV-S27, the prototype strain of ECSA genotype. The comparison of deduced amino acid sequences of both the structural and nonstructural polyproteins of IND-06-Guj and CHIKV-S27 strains by BLAST showed 98% amino acid sequence similarity with 57 amino acid substitutions across the two polyproteins (Table 2). Sequence alignment of structural polyprotein of IND-06-Guj with the 41 CHIKV isolates reported from South Asia and the CHIKV-S27 demonstrated that seven isolates (EF027136, EF027137, EF027140, EF027141, FJ000065, FJ000068 and FJ445427; 6 Indian and 1 Sri Lankan) had 100% sequence identity with IND-06-Guj, while 31 isolates (16 Sri Lankan and 15 Indian) displayed 99% sequence identity. However, IND-06-Guj was 97% and 96% identical to the two isolates belonging to the Asian genotype, EF027140 and EF027141, respectively (Table 1). Further, 50 out of 57 mutations identified between

Table 2: The amino acid substitutions of IND-06-Guj in comparison to CHIKV- S27 isolate

Protein	Position in the Protein	Position in the polyprotein	Amino Acid in S27	Amino Acid in IND-06-Guj
nsP1	128	128	T	K
	172	172	L	V
	234	234	E	K
	376	376	T	M
	383	383	M	L
	384	384	I	L
	481	481	T	I
	488	488	Q	R
	507	507	L	R
	512	512	P	S
nsP2	54	589	S	N
	374	909	H	Y
	641	1176	G	D
	642	1177	C	Y
	643	1178	S	N
nsP3	175	1508	V	I
	217	1550	Y	H
	326	1659	P	S
	331	1664	V	A
	337	1670	T	I
	352	1685	K	E
	376	1709	I	T
	382	1715	A	T
	419	1742	N	I
	461	1794	L	P
	462	1795	S	N
	471	1804	P	S
	534	1857	R	X
nsP4	75	1938	T	A
	254	2117	T	A
	500	2363	Q	L
	514	2377	I	T
	555	2418	V	I
	604	2467	V	I
Capsid	23	23	P	S
	27	27	V	I
	63	63	K	R
E3	23	284	I	T
E2	57	382	G	K
	74	399	I	M
	79	404	G	E
	160	485	N	T
	164	489	A	T
	181	506	L	M
	204	519	S	G
	211	536	I	T
	267	592	M	R
	299	624	S	N
	312	637	T	M
	344	669	A	T
	375	700	S	T
	386	711	V	A
6K	8	756	V	I
	54	802	I	V
E1	269	1078	M	V
	284	1093	D	E
	322	1131	V	A

S27 and IND-06-Guj were also present in other South Asian strains except IND-63-WB1, IND-73-MH5 and IND-00-MH4. Similarly, sequence alignment of nonstructural polyprotein of IND-06-Guj with those of 42 CHIKV isolates (including CHIKV-S27) demonstrated 99% sequence identity with 39 (17 Sri Lankan and 22 Indian) South Asian isolates belonging to ECSA genotype, whereas, two isolates that belonged to Asian genotype (IND-63-WB1 and IND-73-MH5) showed 97% identity.

In this analysis, various amino acid substitutions were observed in the structural as well as the nonstructural polyprotein. Only three amino acid substitutions (nsP1-P512S, nsP2-G1176D and nsP3-N1742I) observed in the nonstructural region were unique to IND-06-Guj. However, two amino acid substitutions in the nonstructural polyprotein, one each in nsP1 (Q488R) and nsP3 (V1664A), found in at least 85% of the South Asian isolates were also found in our cloned sequences from IND-06-Guj. An opal stop codon, which is maintained in all the South Asian isolates except in IND-00-MH4, was also present at position 524 in the nsP3 protein. It has been proposed that this change in nsP3 protein regulates the expression of RNA polymerase by read through mechanism [5]. Interestingly, the amino acid change D284E (Asp to Glu) in E1 gene (IND-06-Guj) has been observed only in CHIKV strains isolated since 2006 outbreak but not in the CHIKV strains (S27, IND-63-WB1, IND-73-MH5 and IND-00-MH4) before 2006. The substitution glutamic acid instead of aspartic acid may introduce a slight distortion in the icosahedral T=4 symmetry of the virion which may affect assembly or disassembly of the virus particle in infected cells [5].

Phylogenetic analysis placed the two Indian isolates (EF027140 and EF027141; belonging to year 1963 and 1973 outbreaks, respectively) in the Asian genotype. Remaining South Asian isolates, including IND-06-Guj, were clustered together into the ECSA genotype along with CHIKV-S27. However, S27 formed a separate branch in this genotype in both the phylogenetic trees. Further it was observed that most isolates from Sri Lanka and India formed separate clusters indicative primarily of their year and region of isolation (fig. 1a and 1b). Interestingly, in the phylogenetic tree constructed based on the nonstructural nucleotide sequences, IND-06-Guj branches out of the clade formed by the South Asian isolates reported since the 2006 outbreak (fig. 1a), whereas it stands out separately with 8 South Asian isolates in the phylogenetic tree generated for structural nucleotide sequences, probably indicative of a common lineage among them (fig. 1b).

Our study suggests that readily available genes of South Asian strain such as IND-06-Guj, in comparison to CHIKV-S27, may help study the biology of CHIKV in the context of South Asian region, where the severity of the disease is more pronounced compared to other parts of the world. Further, expression and purification of proteins encoded by cloned CHIKV cDNAs would provide the much needed reagents to facilitate studies addressing molecular aspects of Chikungunya pathogenesis and would also help to develop tools for the prevention and/or treatment of Chikungunya.

ACKNOWLEDGEMENT:

This work was funded by Department of Biotechnology, Govt. of India. (Grant No. BT/PR11162/Med/29/97/2008). DG acknowledges the DBT support to Bioinformatics Infrastructure Facility at ICGEB.

REFERENCES:

1. Pialoux G, Gaüzère BA, Jauréguiberry S, Strobel M, Chikungunya, an epidemic arbovirosis, Lancet Infect. Dis., 7, 2007, 319-327.
2. Khan AH, Morita K, Parquet Md Mdel C, Hasebe F, Mathenge EG, Igarashi A, Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site, J. Gen. Virol., 83, 2002, 3075-3084.
3. Powers AM, Brault AC, Tesh RB, Weaver SC, Re-emergence of chikungunya and o'nyong-nyong viruses: evidence for distinct geographical lineages and distant evolutionary relationships, J. Gen. Virol., 81, 2000, 471- 479.
4. Powers AM, Logue CH, Changing patterns of chikungunya virus: re-emergence of a zoonotic arbovirus, J. Gen. Virol., 88, 2007, 2363-2377.
5. Schuffenecker I, Iteman I, Michault A, Murri S, Frangeul L, Vaney MC, Lavenir R, Pardigon N, Reynes JM, Pettinelli F, Biscornet L, Diancourt L, Michel S, Duquerroy S, Guigon G, Frenkiel MP, Bréhin AC, Cubito N, Desprès P, Kunst F, Rey FA, Zeller H, Brisse S, Genome microevolution of chikungunya viruses causing the Indian Ocean outbreak, PLoS Med., 2006, 3 e263.
6. Arankalle VA, Shrivastava S, Cherian S, Gunjikar RS, Walimbe AM, Jadhav SM, Sudeep AB, Mishra AC, Genetic divergence of Chikungunya viruses in India (1963-2006) with special reference to the 2005-2006 explosive epidemic, J. Gen. Virol., 88, 2007,1967-1976.
7. Sreekumar E, Issac A, Nair S, Hariharan R, Janki MB, Arathy DS, Regu R, Mathew T, Anoop M, Niyas KP, Pillai MR, Genetic characterization of 2006-2008 isolates of Chikungunya virus from Kerala, South India, by whole genome sequence analysis, Virus Genes, 40, 2010, 14-27.
8. Tsetsarkin KA, Vanlandingham DL, McGee CE, Higgs S, A single mutation in chikungunya virus affects vector specificity and epidemic potential, PLoS Pathog., 2007, 3 e201.
9. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG, Clustal W and Clustal X version 2.0, Bioinformatics, 23, 2007, 2947-2948.
10. Edgar RC, MUSCLE: a multiple sequence alignment method with reduced time and space complexity, BMC Bioinformatics, 5, 2004,113.
11. Kimura M, A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences, J. Mol. Evol., 16, 1980, 111-120.
12. Felsenstein J, Confidence limits on phylogenies: An approach using bootstrap, Evol., 39, 1985, 783-791.
13. Tamura K, Dudley J, Nei M, Kumar S, MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) softare version 4.0, Mol. Biol. Evol., 24, 2007, 1596-1599.
14. Tsetsarkin KA, McGee CE, Volk SM, Vanlandingham DL, Weaver SC, Higgs S, Epistatic roles of E2 glycoprotein mutations in adaption of chikungunya virus to *Aedes albopictus* and *Ae. aegypti* mosquitoes, PLoS One 4, 2009, e6835.
15. Hapuarachchi HC, Bandara KB, Sumanadasa SD, Hapugoda MD, Lai YL, Lee KS, Tan LK, Lin RT, Ng LF, Bucht G, Abeyewickreme W, Ng LC, Re-emergence of Chikungunya virus in South-east Asia: virological evidence from Sri Lanka and Singapore, J. Gen. Virol., 91, 2010,1067-1076.

Source of support: DBT, Conflict of interest: None Declared