Genotyping of virus involved in the 2006 Chikungunya outbreak in South India (Kerala and Puducherry)

N. Pradeep Kumar[1,*], M. Madhu Mitha[2], N. Krishnamoorthy[1], T. Kamaraj[3], Rajan Joseph[3] and P. Jambulingam[1]

[1]Vector Control Research Centre (ICMR), Indira Nagar, Puducherry 605 006, India
[2]VIT University, Vellore 632 014, India
[3]Kerala State Institute of Virology and Infectious Diseases, Alappuzha 688 009, India

An epidemic outbreak of Chikungunya virus occurred during 2006 in 15 States/Union Territories of India. The magnitude of the outbreak was unprecedented. We carried out a preliminary investigation to characterize the genotype of the virus involved in the outbreak in the State of Kerala and the Union Territory of Puducherry. We also looked into mutations in the gene sequences of the *E1* region, the most informative gene of the virus on phylogeny, so as to understand the evolutionary trends of the species. The results indicate that the outbreak in Kerala and Puducherry was caused by the East Central South African (ECSA) strain of the virus. Owing to the absence of the A226V mutation in the samples analysed, it is likely that the outbreak was caused by the introduction of the ancestral stock of the Reunion clade to India, either by a human host or by the vector population. The magnitude of the outbreak might be due to mutations other than the already proposed one, i.e. A226V. Interestingly, one of the samples (Puducherry) showed a mutation K211E in the *E1* gene, specific only to the Asian strain. In-depth studies are required in order to have a thorough understanding of the phylogenetic trends of the virus in India.

Keywords: Chikungunya virus, genotyping, outbreak, phylogeny.

THE Chikungunya virus (CHIKV) belonging to the genus *Alphavirus*, family Togaviridae, is transmitted by *Aedes* species of mosquitoes. This virus caused a pandemic outbreak in the islands of Indian Ocean during 2005–06, affecting more than 70% of the population[1]. The higher magnitude of the outbreak was attributed to a mutation A226V in the *E1* gene, which offers cholesterol independence to the virus, in infecting the vector mosquito. Recently, an outbreak of this disease occurred in India affecting about 15 States/Union Territories[2]. The number of suspected cases was estimated to be 1,392,027 in the country, with 70,731 and 542 cases in Kerala and Puducherry respectively[2]. The magnitude of the outbreak was higher than the Chikungunya epidemic reported[3] in India during 1971. While the local authorities of the Government of Kerala reported mortality of 74 cases in Alappuzha District, these were attributed to other infections/ disease conditions of the patients by Health authorities of the Government of India[2].

A recent study on the CHIKV outbreak in Andhra Pradesh (AP), Karnataka and Maharashtra[4] indicated the involvement of the East African strain of the virus. We isolated the virus from a few serum samples collected from suspected Chikungunya cases in Kerala and Puducherry and amplified the gene for non-structural protein 1 (NSP1), reported to be of diagnostic use for viral infection. The viral gene for the membrane fusion glycoprotein envelope 1 (*E1*) gene, reported to be of phylogenetic significance was amplified to elucidate the strain of the parasite involved in the viral outbreak. Since the *E1* gene is reported to be the most phylogenetically informative one[5], a larger portion of the gene was also amplified to elucidate the evolutionary trends of the parasite.

About 500 µl of blood samples from eight suspected or clinically diagnosed Chikungunya cases attending Primary Health Centres in Alappuzha District, were collected into RNAase-free eppendorff tubes, which were immediately transferred to ice packs (4°C). These were transported to the Kerala State Institute of Virology and Infectious Diseases, Alappuzha and subsequently to the Vector Control Research Centre, Puducherry. These samples were centrifuged at 6000 rpm for about 15 min, and the sera separated into fresh RNAase-free tubes. Four serum samples of suspected cases from Puducherry collected within 3–4 days of onset of fever and one sample collected from the nearby district of Cuddalore, Tamil Nadu (TN) were also used for virus isolation. A negative control (blood of the first author) was maintained. Viral RNA was isolated from the serum sample using QIAGEN Viral RNA Kit (QIAGEN, Germany). The RNA was reverse-transcribed to cDNA, and *NSP1* and *E1* gene sequences were amplified using Roche One Step RT–PCR kit (Roche, USA). The DNA primers used for RT–PCR were those already reported[5]. The amplified fragments were custom-sequenced (Microsynth, Switzerland). The larger portion of the *E1* gene (~1000 bp) was amplified using the 'FG20' DNA primers[1]. The cDNA sequences were aligned using ClustalW and analysed using phylogenetic software, viz. MEGA 3.1 (Kumar *et al.*[6]) and DNASP (Rozas *et al.*[7]).

Five samples among eight collected from Kerala and all the five samples from Puducherry/TN were found positive for CHIKV infection. The case details of the patients and respective GenBank accession numbers for the cDNA sequences amplified are given in Table 1. Custom sequencing of one of the *E1* sequences from Kerala did not yield good results, and hence was not included for analysis. The cDNA sequences of the *NSP1* gene amplified ranged from 322 to 359 bp.

*For correspondence. (e-mail: kumarnp@yahoo.com)

Table 1. Details of Chikungunya patients from Kerala, Puducherry and Tamil Nadu

Patient ID no.	Date of collection	Age	Sex	State	Region	Symptoms when sampled	GenBank accession no.	
							NSP1	*E1*
CH1	20–10–06	25	M	Kerala	Alappuzha	Fever		Negative
CH2	21–10–06	12	F	Kerala	Alappuzha	Fever		Negative
CH3	20–10–06	7	M	Kerala	Alappuzha	Fever*	EF674318	EF674326
CH4	22–10–06	2	M	Kerala	Alappuzha	Fever*	EF674319	EF674327
CH5	22–10–06	6	F	Kerala	Mararikulam	Fever*	EF674320	EF674328
CH6	23–10–06	4	F	Kerala	Alappuzha	Fever*		Negative
CH7	23–10–06	6	F	Kerala	Alappuzha	Fever*	EF674321	EF674329
CH8	24–10–06	13	F	Kerala	Kuttanad	Fever*	EF674322	EF674330
PCH1	12–10–06	60	F	Puducherry	Puducherry	Fever*	EF113093	EF113095
PCH2	12–10–06	14	M	Puducherry	Puducherry	Fever*	EF113094	EF113096
PCH3	02–11–06	24	F	Puducherry	Puducherry	Fever*	EF674323	EF674331
PCH4	03–11–06	49	M	Puducherry	Puducherry	Control		Negative
PCH5	03–11–06	18	F	Tamil Nadu	Cuddalore	Fever*	EF674324	EF674332
PCH6	06–11–06	16	M	Puducherry	Puducherry	Fever*	EF674325	–

*Other symptoms included arthralgia, myalgia headache, reddishness of eyes and rashes.

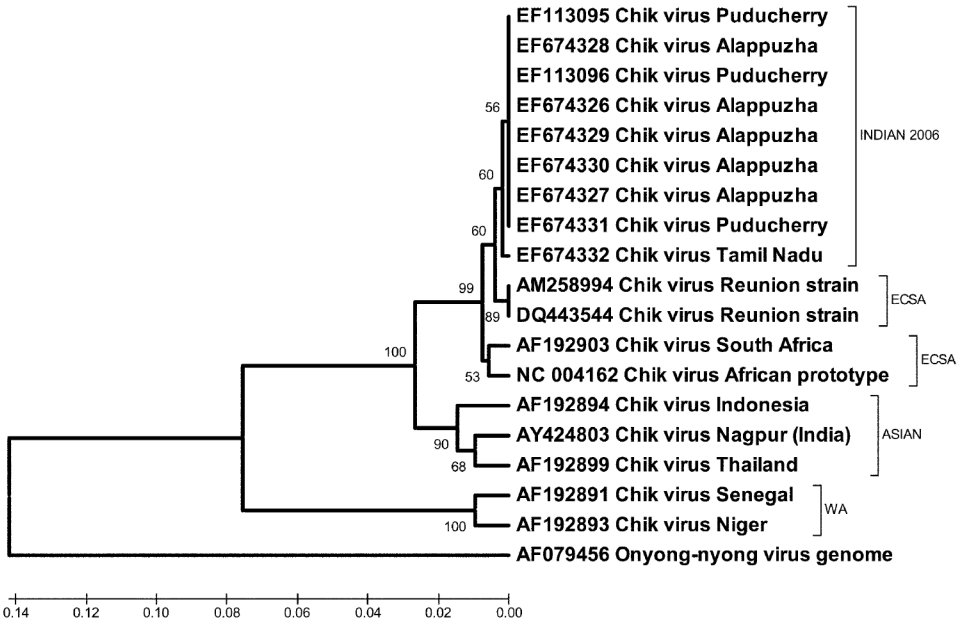

Figure 1. Phylogenetic analysis of the *E1* gene sequences of Chikungunya virus (bootstrap consensus tree).

The *E1* gene amplified ranged from 247 to 286 bp. The mean genetic distance among the nine sequences was negligible (Kimura two-parameter = 0.0008). In the phylogenetic tree these sequences grouped with the ECSA strain when analysed along with the representatives of the three strains of CHIKV[2] (Figure 1), and were found to be genetically similar to the Reunion (Indian Ocean islands) isolates. The genetic distance between these samples and the ECSA strain was only 0.0110, compared to the Asian strain (0.0522). These results indicate that the current outbreak was caused by the Reunion clade of the virus, which may have been transported by either human host or vector to the highly receptive South Indian region. An earlier study of the outbreak in AP, Karnataka and Maharashtra also implicated the cause of the outbreak to the ECSA strain[4].

The larger region of five *E1* sequences amplified included three samples from Puducherry and two from Alappuzha region. On sequencing, the DNA sequences of this fragment ranged from 933 to 957 bp (GenBank accession no. EF555196–EF555200).

Two mutations, viz. T/C at 10593 and A/G at 10624 were found in the 933 bp of the *E1* gene in the isolates from Puducherry region, corresponding to 10182–11114 bp of Chikungunya genome (DQ443544). While the former was a synonymous mutation, the latter was non-synonymous

Table 2. Alignment of the deduced amino acids of *E1* gene sequences of Chikungunya virus isolates (non-synonymous mutations K211E and A226V are highlighted)

```
#MEGA
!Title Chik E1 sequences;
!Format
   DataType=Protein
   NSeqs=14 NSites=283
   Identical=. Missing=? Indel=-;

!Domain=Data;
#EF555196_Chik_virus_E1_Alappuzha      GAYCFCDAEN TQLSEAHVEK SESCKTEFAS AYRAHTASAS AKLRVLYQGN NITVTAYANG
#EF555197_Chik_virus_E1_Alappuzha      .......... .......... .......... .......... .......... ..........
#EF555198_Chik_virus_E1_Puducherry     .......... .......... .......... .......... .......... ..........
#EF555199_Chik_virus_E1_Puducherry     .......... .......... .......... .......... .......... ..........
#EF555200_Chik_virus_E1_Puducherry     .......... .......... .......... .......... .......... ..........
#AM258994_Chik_virus_Reunion           .......... .......... .......... .......... .......... ..........
#DQ443544_Chik_virus_Reunion           .......... .......... .......... .......... .......... ..........
#AF192903_Chik_virus_South_Africa      .......... .......... .......... .......... .......... ..........
#NC_004162Chik_virus_African_prototype .......... .......... .......... .......... .......... ..........
#AY424803_Chik_virus_E1_Nagpur         .......T.. .......... .......... .......... G......... .V..S.....
#AF192899_Chik_virus_E1_Thailand       .......T.. .......... .......... .......... .......... SV..S.....
#AF192894_Chik_virus_E1_Indonesia      .......T.. .......... .......... .......... .......... .V..S.....
#AF192891_Chik_virus_E1_Senegal        .......... .......... .......... .......... .......... ....A.....
#AF192893_Chik_virus_E1_Niger          .......... .......... .......... .......... .......... ....A.....

#EF555196_Chik_virus_E1_Alappuzha      DHAVTVKDAK FIVGPMSSAW TPFDNKIVVY KGDVYNMDYP PFGAGRPGQF GDIQSRTPES
#EF555197_Chik_virus_E1_Alappuzha      .......... .......... .......... .......... .......... ..........
#EF555198_Chik_virus_E1_Puducherry     .......... .......... .......... .......... .......... ..........
#EF555199_Chik_virus_E1_Puducherry     .......... .......... .......... .......... .......... ..........
#EF555200_Chik_virus_E1_Puducherry     .......... .......... .......... .......... .......... ..........
#AM258994_Chik_virus_Reunion           .......... .......... .......... .......... .......... ..........
#DQ443544_Chik_virus_Reunion           .......... .......... .......... .......... .......... ..........
#AF192903_Chik_virus_South_Africa      .......... .......... .......... ..G....... .......... ..........
#NC_004162Chik_virus_African_prototype .......... .......... .......... .......... .......... ..........
#AY424803_Chik_virus_E1_Nagpur         .......... .......... .......... .......... .......... ..........
#AF192899_Chik_virus_E1_Thailand       .......... .......... .......... .......... .......... ..........
#AF192894_Chik_virus_E1_Indonesia      .......... .......... .......... .......... .......... ..........
#AF192891_Chik_virus_E1_Senegal        .......... .V........ .......... .......... .......... ..........
#AF192893_Chik_virus_E1_Niger          .......... .V........ .......... .......... .......... ..........
```

(*Contd.*)

Table 2. (*Contd.*)

```
#EF555196_Chik_virus_E1_Alappuzha      KDVYANTQLV LQRPAAGTVH VPYSQAPSGF KYWLKERGAS LQHTAPFGCQ IATNPVRAVN
#EF555197_Chik_virus_E1_Alappuzha      .......... .......... .......... .......... .......... ..........
#EF555198_Chik_virus_E1_Puducherry      .......... .......... .......... .......... .......... ..........
#EF555199_Chik_virus_E1_Puducherry      .......... .......... .......... .......... .......... ..........
#EF555200_Chik_virus_E1_Puducherry      E......... .......... .......... .......... .......... ..........
#AM258994_Chik_virus_Reunion            .......... .....V.... .......... .......... .......... ..........
#DQ443544_Chik_virus_Reunion            .......... .....V.... .......... .......... .......... ..........
#AF192903_Chik_virus_South_Africa       .......... .......... .......... .......... .......... ........M.
#NC_004162Chik_virus_African_prototype  .......... .......... .......... .......... .......... ........M.
#AY424803_Chik_virus_E1_Nagpur          E......... ....S..... .......... .......... .......... ........M.
#AF192899_Chik_virus_E1_Thailand        E......... ....S..... .......... .......... .......... ........M.
#AF192894_Chik_virus_E1_Indonesia       E.I....... ....S..... .......... .......... .......... ........M.
#AF192891_Chik_virus_E1_Senegal         .......... .......... .......... .......... .......... ..........
#AF192893_Chik_virus_E1_Niger           .......... .......... .......... .......... .......... ..........

#EF555196_Chik_virus_E1_Alappuzha      CAVGNMPISI DIPEAAFTRV VDAPSLTDMS CEVPACTHSS DFGGVAIIKY AASKKGKCAV
#EF555197_Chik_virus_E1_Alappuzha      .......... .......... .......... .......... .......... ..........
#EF555198_Chik_virus_E1_Puducherry      .......... .......... .......... .......... .......... ..........
#EF555199_Chik_virus_E1_Puducherry      .......... .......... .......... .......... .......... ..........
#EF555200_Chik_virus_E1_Puducherry      .......... .......... .......... .......... .......... ..........
#AM258994_Chik_virus_Reunion            .......... .......... .......... .......... .......... ..........
#DQ443544_Chik_virus_Reunion            .......... .......... .......... .......... .......... ..........
#AF192903_Chik_virus_South_Africa       .......... ...D...... .......... .......... .......... ..........
#NC_004162Chik_virus_African_prototype  .......... ...D...... .......... .......... .......... .V........
#AY424803_Chik_virus_E1_Nagpur          .......... ...D...... .......... .......... ....A..... ..........
#AF192899_Chik_virus_E1_Thailand        .......... ...D...... .......... .......... .......... ..........
#AF192894_Chik_virus_E1_Indonesia       .......... ...D...... .......... .......... .......... ..........
#AF192891_Chik_virus_E1_Senegal         .....I.... ...D...... .....V.... .......... .......... T.........
#AF192893_Chik_virus_E1_Niger           .....I.... ...D...... .....V.... .......... .......... T.........

#EF555196_Chik_virus_E1_Alappuzha      HSMTNAVTIR EAEIEVEGNS QLQISFSTAL ASAEFRVQVC STQ
#EF555197_Chik_virus_E1_Alappuzha      .......... .......... .......... .......... ...
#EF555198_Chik_virus_E1_Puducherry      .......... .......... .......... .......... ...
#EF555199_Chik_virus_E1_Puducherry      .......... .......... .......... .......... ...
#EF555200_Chik_virus_E1_Puducherry      .......... .......... .......... .......... ...
#AM258994_Chik_virus_Reunion            .......... .......... .......... .......... ...
#DQ443544_Chik_virus_Reunion            .......... .......... .......... .......... ...
#AF192903_Chik_virus_South_Africa       .......... .......... .......... .......... ...
#NC_004162Chik_virus_African_prototype  .......... .......... .......... .......... ...
#AY424803_Chik_virus_E1_Nagpur          .......... .......... .......... .......... ...
#AF192899_Chik_virus_E1_Thailand        .......... .......... .......... .......... ...
#AF192894_Chik_virus_E1_Indonesia       .......... .......... .......... .......... ...
#AF192891_Chik_virus_E1_Senegal         ........T. ..DV...... .......... .......... ...
#AF192893_Chik_virus_E1_Niger           .......... ..DV*.*... ........S. .....H.... ...
```

(K211E). Such mutations were also recorded in the Asian strains (AY424803-Nagpur, India, AF192899-Thailand and AF192894-Indonesia). However, other non-synonymous mutations that were recorded in the Asian strain were totally absent in the Puducherry isolate. Besides these, two mutations were recorded among the sequences of South Indian isolates compared to the Reunion isolates, viz. C/T at 10670 and A/G at 10743. The former was the non-synonymous mutation, which led to A226V in Reunion strains, while the latter was a synonymous mutation. The alignment of the common region of the sequences (deduced amino acids) with the representatives of all the three strains of CHIKV is given in Table 2 (region 10264–11112 corresponding to DQ443544).

During the initial period of the CHIKV outbreak in the Indian Ocean islands (March–June 2005), the isolates did not show prevalence of A226V mutations. However, during the explosive epidemic peak of infection from September to December 2005, and subsequently till March 2006, the mutation frequency of A226V in the population increased tremendously to reach about 90% of the population[1]. Since this mutation offers cholesterol independence to the virus, as described in the closely related Selmiki Forest virus, it was postulated that this function would have provided the virus a chance to survive and multiply with better efficiency in mosquito vectors, which in turn would have contributed to its rapid spread[1]. However, the Indian isolates that we studied did not have this mutation. It may be concluded that the current Indian outbreak might be due to the introduction of ancestral virus strain, which caused the first period of outbreak in the Reunion Islands. The absence of this mutation indicates the probable contribution of other mutations in the genome towards the comparatively higher magnitude of the outbreak in India. Sequencing and analysing the entire genome of this virus, may throw light on this aspect. The findings of this preliminary study on the phylogeny of CHIKV infection in Kerala and Puducherry warrant a large-scale investigation involving more samples, so as to have a thorough understanding of the evolutionary trends of the virus.

1. Schuffenecher, I. *et al.*, Genome microevolution of Chikungunya viruses causing the Indian Ocean outbreak. *PloS. Med.*, 2006, **3**, 1058–1070.
2. National Vector Borne Disease Control Programme, Govt of India, New Delhi, 2007; http://www.nvbdcp.gov.in/Chikun-cases.html
3. Pavri, K., Disappearance of Chikungunya virus from India and South East Asia. *Trans. R. Soc. Trop. Med. Hyg.*, 1986, **80**, 491.
4. Yergolkar, P. N. *et al.*, Chikungunya outbreaks caused by African genotype, India. *Emerg. Infect. Dis.*, 2006, **12**, 1580–1583.
5. Hasebe, F. *et al.*, Combined detection and genotyping of Chikungunya virus by a specific reverse transcription-polymerase chain reaction. *J. Med. Virol.*, 2002, **67**, 370–374.
6. Kumar, S., Tamura, T. and Nei, M., MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefs Bioinf.*, 2004, **5**, 150–163.
7. Rozas, J., Sanchez-Delbarria, J. C., Messeguer, X. and Rozas, R., DnaSP, DNA polymorphism analysis by the coalescent and other methods. *Bioinformatics*, 2003, **19**, 2496–2497.

ACKNOWLEDGEMENTS. We are grateful to Dr J. Yuvaraj, Assistant Director, Vector Control Research Centre (VCRC), Puducherry for supply of some samples used in the study and to P. M. Ajithlal, VCRC for technical assistance.

Received 10 August 2007; accepted 5 October 2007

Comparative analysis of GPS baseline data using different stochastic modelling techniques

D. Chakravarty*, S. Venugopal and S. K. Ghosh

Department of Mining Engineering, Indian Institute of Technology, Kharagpur 721 302, India

Double-differenced Global Positioning System (GPS) carrier-phase observations produce accurate GPS-precise point positioning results if processed with least squares algorithms, which need a functional and a stochastic model to be properly defined, the latter being difficult to accomplish practically. The standard model, the MINQUE model and the simplified MINQUE model are widely used for estimating the variance–covariance components of GPS observations. In this communication, the outputs from these models are compared. It has been shown that both the MINQUE models are superior to the standard stochastic model and the simplified MINQUE model has the same accuracy as the MINQUE model, but with a drastically reduced computational time.

Keywords: Carrier phase, double difference, MINQUE model, stochastic modelling, variance–covariance.

IN principle, the Global Positioning System (GPS) is used for navigational purposes or for locating the position of any particular point on the earth's surface. However, with the application of mathematical and surveying concepts, high accuracy surveying and deformation studies up to centimetre or millimetre level accuracies can be done. The relative positioning of double-differenced (DD) GPS measurements is favoured because it cancels out many systematic errors and has a simplified functional model. Although major parts of the systematic errors can be modelled or cancelled out by double differencing the GPS measurements between the satellites and receivers,

*For correspondence. (e-mail: dc@mining.iitkgp.ernet.in)