Journal of General Virology (2010), **91**, 1067–1076 DOI 10.1099/vir.0.015743-0

Re-emergence of Chikungunya virus in South-east Asia: virological evidence from Sri Lanka and Singapore

H. C. Hapuarachchi,[1] K. B. A. T. Bandara,[2] S. D. M. Sumanadasa,[1] M. D. Hapugoda,[2] Yee-Ling Lai,[1] Kim-Sung Lee,[1] Li-Kiang Tan,[1] Raymond T. P. Lin,[3] Lisa F. P. Ng,[4] Göran Bucht,[1] Wimaladharma Abeyewickreme[2,5] and Lee-Ching Ng[1]

Correspondence
Lee-Ching Ng
ng_lee_ching@nea.gov.sg

[1]National Environment Agency, Environmental Health Institute, Singapore

[2]Molecular Medicine Unit, Faculty of Medicine, University of Kelaniya, Sri Lanka

[3]Ministry of Health, Singapore

[4]Singapore Immunology Network, Agency for Science, Technology and Research (A* STAR), Singapore

[5]Department of Parasitology, Faculty of Medicine, University of Kelaniya, Sri Lanka

Chikungunya fever swept across many South and South-east Asian countries, following extensive outbreaks in the Indian Ocean Islands in 2005. However, molecular epidemiological data to explain the recent spread and evolution of Chikungunya virus (CHIKV) in the Asian region are still limited. This study describes the genetic Characteristics and evolutionary relationships of CHIKV strains that emerged in Sri Lanka and Singapore during 2006–2008. The viruses isolated in Singapore also included those imported from the Maldives (n=1), India (n=2) and Malaysia (n=31). All analysed strains belonged to the East, Central and South African (ECSA) lineage and were evolutionarily more related to Indian than to Indian Ocean Islands strains. Unique genetic characteristics revealed five genetically distinct subpopulations of CHIKV in Sri Lanka and Singapore, which were likely to have emerged through multiple, independent introductions. The evolutionary network based on *E1* gene sequences indicated the acquisition of an alanine to valine 226 substitution (E1-A226V) by virus strains of the Indian sublineage as a key evolutionary event that contributed to the transmission and spatial distribution of CHIKV in the region. The E1-A226V substitution was found in 95.7 % (133/139) of analysed isolates in 2008, highlighting the widespread establishment of mutated CHIKV strains in Sri Lanka, Singapore and Malaysia. As the E1-A226V substitution is known to enhance the transmissibility of CHIKV by *Aedes albopictus* mosquitoes, this observation has important implications for the design of vector control strategies to fight the virus in regions at risk of chikungunya fever.

Received 8 August 2009
Accepted 28 November 2009

INTRODUCTION

Chikungunya fever is an acute illness caused by Chikungunya virus (CHIKV), an alphavirus of the family *Togaviridae*. CHIKV is transmitted to humans by mosquitoes of the genus *Aedes*, particularly *Aedes aegypti* and *Aedes albopictus*. Though generally a non-fatal condition characterized by high fever, severe arthralgia, polyarthritis and skin rash, CHIKV infections may rarely be associated with complications such as encephalopathy and hepatic failure (Schuffenecker *et al.*, 2006). The clinical illness is often associated with prolonged morbidity, which can impose enormous social and economic disadvantages on affected communities (Enserink, 2007).

The first notable emergence of CHIKV attributable to the recent wave of outbreaks occurred in Kenya in 2004 (Enserink, 2007). In 2005, CHIKV emerged in Comoros (Sergon *et al.*, 2007) followed by Reunion Island, the Seychelles, Mauritius and Mayotte. In Reunion Island

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this paper are GU013528–GU013530 and FJ513628–FJ513679.

Four supplementary tables and a supplementary figure are available with the online version of this paper.

alone, more than a third of the total population (267 000 individuals) was debilitated (Schuffenecker *et al.*, 2006). By late 2005, massive outbreaks of CHIKV erupted in India, affecting approximately 1.3 million individuals (Ravi, 2006). The re-emergence of CHIKV in Sri Lanka was confirmed in November 2006 (Hapuarachchi *et al.*, 2008), after four decades of absence (Munasinghe *et al.*, 1966). The estimated figures indicate that approximately 40 000 individuals in Sri Lanka were affected during 2006–2007 and a similar number in 2008. Singapore reported the first indigenous chikungunya fever outbreak in January 2008 (Leo *et al.*, 2009; Ng *et al.*, 2009), and subsequent outbreaks have so far resulted in approximately 1000 infections. Similarly, epidemics of chikungunya fever have also been reported from Malaysia since 2006 (AbuBakar *et al.*, 2007; Noridah *et al.*, 2007), causing approximately 7000 cases in 2008 (Sam *et al.*, 2009). In the Maldives, approximately 12 000 suspected chikungunya fever cases were reported during the 2006–2007 outbreak (Yoosuf *et al.*, 2009). Similarly, CHIKV outbreaks were also reported from Gabon (Peyrefitte *et al.*, 2008), Cameroon (Peyrefitte *et al.*, 2007) and Thailand (Theamboonlers *et al.*, 2009). Moreover, the first indigenous CHIKV outbreak in Europe was reported from Italy in 2007 (Rezza *et al.*, 2007). Several other European countries (Hochedez *et al.*, 2006; Parola *et al.*, 2006; Veber *et al.*, 2007) and the USA (Lanciotti *et al.*, 2007) have also reported imported infections, raising serious concerns about the potential burden of CHIKV infections.

Many studies conducted so far to understand the origin and evolution of CHIKV have focused on full genome and *E1* gene sequences of the virus (Powers *et al.*, 2000; Schuffenecker *et al.*, 2006; Arankalle *et al.*, 2007; Kumar *et al.*, 2007). Phylogenetic analyses of the CHIKV genome have shown three genetic lineages: Asian, West African, and East, Central and South African (ECSA) (Powers *et al.*, 2000). It is now clear that the emergence of CHIKV in the Indian Ocean Islands in early 2005 was due to a newer strain of the ECSA lineage (Schuffenecker *et al.*, 2006). Current epidemiological evidence suggests that this strain could have moved to the Indian Ocean Islands following CHIKV outbreaks in Kenya in 2004 (Kariuki Njenga *et al.*, 2008). Although the molecular epidemiology of CHIKV in India has been thoroughly described (Yergolkar *et al.*, 2006; Arankalle *et al.*, 2007), similar data to explain the evolutionary relationships of CHIKV that spread in the rest of the Asian region after 2005 are still limited.

In the present study, we sought to analyse the evolutionary relationships of CHIKV in South and South-east Asian regions after 2005, in order to understand the possible routes of spread of the virus to Sri Lanka, Singapore, Malaysia and the Maldives. Moreover, we present a hypothetical evolutionary pathway for CHIKV strains that emerged in the Indian Ocean Islands and Asian region after 2005. The implications of these findings with regard to the control of chikungunya fever in affected regions are also discussed.

RESULTS

Sri Lankan and Singapore CHIKV isolates evolved from Indian isolates

The *E1*-gene-based network (see Methods) revealed two distinct evolutionary paths within the ECSA lineage, indicating that strains from the Indian Ocean Islands were most likely to have evolved separately from the Indian isolates (Fig. 1). All Sri Lankan, Singapore, Malaysian and Maldivian CHIKV strains clustered within the Indian sublineage, reflecting their closer genetic relationship with the Indian than the Indian Ocean Islands isolates. This observation was corroborated by the full genome tree topology, which showed two distinct sublineages consisting of Indian Ocean Islands and Indian isolates within a large monophyletic clade of the ECSA lineage (Fig. 2). Comparative analysis of full genome sequences showed that Sri Lankan and Singapore CHIKV strains shared very high nucleotide (99.2–99.9 %) and amino acid (99.1–100 %) similarities with both the Indian Ocean Islands and Indian isolates. However, the Sri Lankan and Singapore isolates possessed nsP1-T128K, nsP1-T376M, Capsid-P23S and Capsid-V27I amino acid substitutions that were previously thought to be unique to Indian isolates (Arankalle *et al.*, 2007). On the other hand, synonymous mutations at nucleotide positions 2067 (T→C) and 2748 (T→C) in the structural polyprotein (Brisse *et al.*, 2007) and at positions 5172 (C→T) and 6471 (G→A) in the non-structural polyprotein were uniquely found in the Indian Ocean Islands isolates and were not observed in Sri Lankan and Singapore isolates. These observations explained the division observed between the Indian Ocean Islands and Indian isolates in the phylogenetic analyses and confirmed that the Sri Lankan and Singapore CHIKV strains, including those imported from Malaysia and the Maldives, were evolutionarily more related to the Indian than to the Indian Ocean Islands strains.

Moreover, the tree topology indicated further division of the study isolates into five genetically divergent groups with robust bootstrap support (Fig. 2). There was one subclade consisting of 2006–2007 Sri Lankan isolates (cluster A in Fig. 1). Two subclades were evident within the 2008 Sri Lankan CHIKV population and were designated 2008 group 1 and 2 (clusters B and C in Fig. 1). While group 1 isolates were detected in both study areas, 95.8 % (23/24) of group 2 isolates were confined to a single area (Eheliyagoda; Supplementary Table S3). In a similar way, the Singapore isolates belonged to three subclades: January 2008 (cluster A in Fig. 1), May–August 2008 (cluster D in Fig. 1) and a single isolate [SG(EHI)CHS394108] reported in early June 2008 of Sri Lankan group 2 subclade (linked to cluster C in Fig. 1).

Genetic differences that define diverging groups of Sri Lankan and Singapore CHIKV isolates

Among the Sri Lankan isolates, all four 2007 strains had the wild-type amino acid residue (alanine) at position 226 of

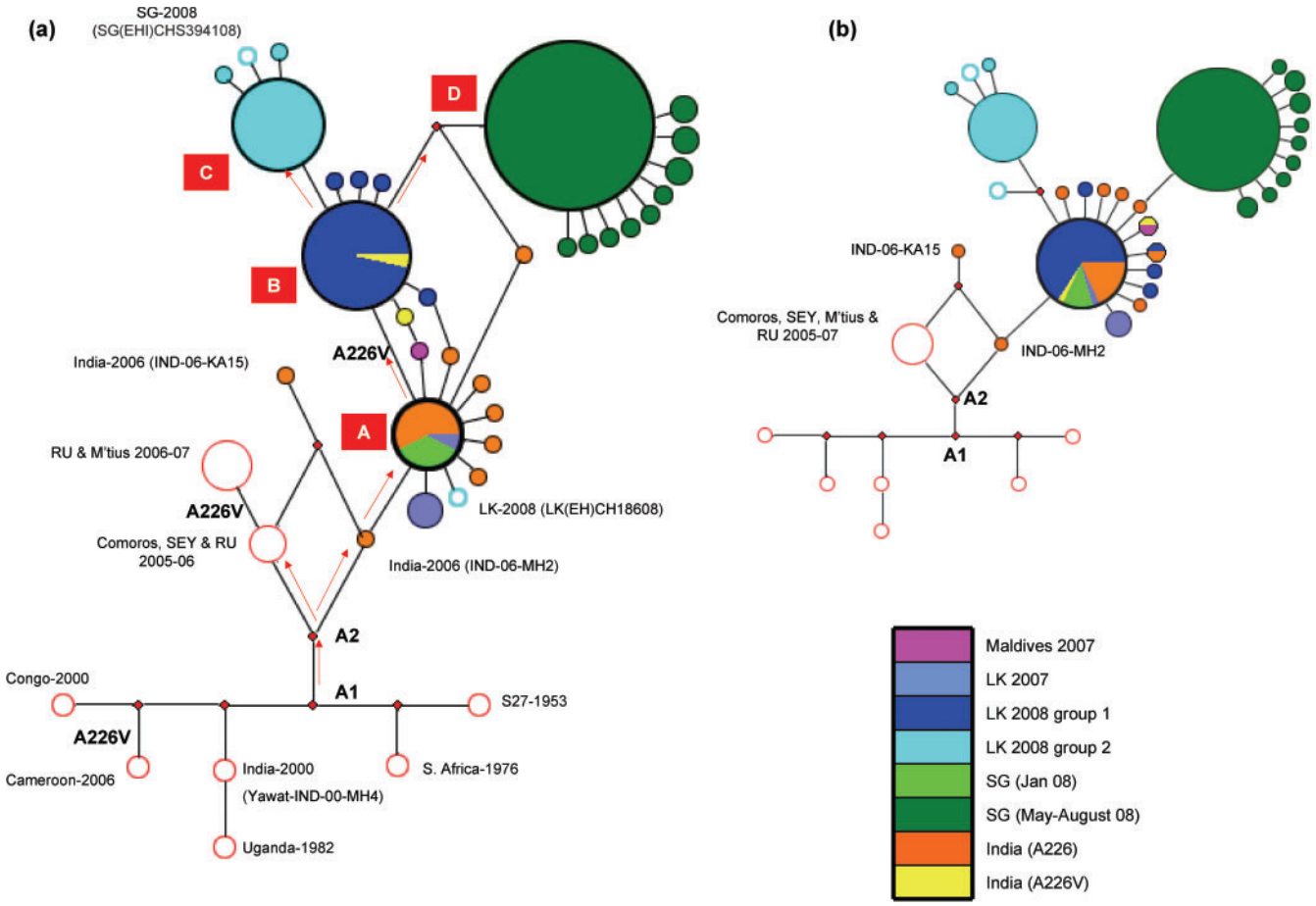

Fig. 1. Median-joining network of CHIKV *E1* gene sequences. (a) Using sequences that included amino acid residue 226. (b) Using sequences from which amino acid residue 226 has been removed. The median joining network was drawn using 1002 nt of the *E1* gene, spanning from amino acid residues 91 to 424. Circles represent either individual isolates or clusters. The diameter of each circle is proportional to the number of isolates within each circle. The length of lines linking circles is not proportional to the mutational distance between them. The red nodes represent hypothetical ancestral strains or strains present in the population but not sampled. The sequences retrieved from GenBank are listed in Supplementary Table S2. A–D represent clusters A–D described in the text. Cluster C includes a Singapore isolate [SG(EHI)CHS394108]. Cluster D also includes isolates imported from Malaysia. The red arrows show the suggested evolutionary path of the *E1* gene of CHIKV. A1, ancestral strain of the newer ECSA lineage; A2, common ancestor of Indian and Indian Ocean Islands isolates; RU, Reunion; SEY, Seychelles; M'tius, Mauritius; SG, Singapore; LK, Sri Lanka.

the E1 envelope protein (E1-A226). Of these four isolates, three formed a distinct cluster, which included two viruses isolated in Japan from travellers from Sri Lanka in 2006 (Fig. 2). The absence of nsP1-Q488R and nsP3-V331A amino acid substitutions in this distinct cluster was notable, as both of these substitutions were otherwise present in the rest of the analysed isolates from Sri Lanka, India and Indian Ocean Islands. All 2008 group 1 Sri Lankan isolates (n=33) showed a unique signature of nsP3-Y38H, nsP3-M394I and E3-S18F substitutions. On the other hand, group 2 isolates (n=24) uniformly showed the presence of nsP3-T444M, E2-V222I and E1-K211N substitutions and the absence of the E1-M269V substitution that was present in isolates of both Indian and Indian Ocean Islands sublineages. Of the 2008 Sri Lankan isolates, 98.3 % (57/58) possessed the E1-A226V substitution. The remaining 2008 Sri Lankan isolate [LK(EH)CH18608] did not possess the E1-A226V substitution, but it clustered with 2008 group 2 isolates that possessed this substitution (Fig. 2).

As in the Sri Lankan virus population, one of the notable genetic differences in Singapore isolates was the presence of E1-A226V substitution in all isolates of the May–August 2008 group, but not in the January 2008 group. The two main groups of viruses from Singapore had distinct signatures. All strains of the January 2008 group showed a combination of nsP1-T60A, nsP4-I563T and E2-R178H substitutions. On the other hand, isolates of the May–August 2008 group, which also included three cases

Fig. 2. Phylogenetic analysis of CHIKV full genomes. The maximum-likelihood method was used to construct the phylogenetic tree using 11 580 nt consisting of complete coding regions of full genome sequences. The numbers shown on branches are bootstrap values determined by using the NJ method under the ML criterion based on 1000 replicates. All sequences are labelled with the country of origin, reported year and GenBank accession number. Sequences of Asian and West African lineages were not included in the analysis.

imported from Malaysia, characteristically showed nsP2-L539S, nsP4-R82S and E2-K252Q substitutions (Table 1). The only Singapore isolate [SG(EHI)CHS394108] that clustered with Sri Lankan strains in the network (Fig. 1a) showed the E1-N211N and E1-M269 amino acid combination uniquely found in Sri Lankan group 2 isolates. All of these substitutions except E1-K211N, E1-M269 and nsP4-I563T were novel. As the above mutation profiles were unique and present uniformly in isolates of each group, they were defined as genetic signatures of the Sri Lankan and Singapore CHIKV isolates.

Diverging evolutionary pathways of Asian, Indian Ocean Islands and African CHIKV strains

The network analysis indicated two diverging pathways arising from the hypothetical ancestor A1, which represented the progenitor of the ECSA sublineage that emerged recently in Africa and Asia (Fig. 1). One path included the Yawat strain (IND-00-MH4, EF027139.1) and African isolates reported from Uganda (1982, AF192907.1), Congo (2000, AY549583.1) and Cameroon (2006, EF051584.1). The other path led to a common ancestor (A2) of Indian and Indian Ocean Islands isolates attributable to 2005–2006 outbreaks. According to the network, E1-D284E (D1093E in the structural polyprotein) was the only non-synonymous mutation in the *E1* gene observed between A1 and A2 ancestors.

Further evolution within Asian isolates resulted in distinct clusters that represented the five distinct subclades of Asian isolates described above (Fig. 1a). All analysed isolates of the Indian sublineage were genetically linked to either cluster A or cluster B. One of the characteristic differences between isolates of clusters A and B was the presence of the E1-A226V substitution in cluster B isolates. This was clearly evident when amino acid residue

Table 1. Novel amino acid substitutions observed in Singapore and Sri Lankan CHIKV isolates

A dot denotes similarity to amino acids of the reference strain at the corresponding positions. The amino acid positions correspond to the respective positions in the polyproteins and individual proteins. The coding sequence of the African prototype (GenBank accession no. NC_004162) was used as the reference sequence. SG, Singapore; LK, Sri Lanka.

Protein	Polypeptide position	Protein position	Ref. strain	SG Jan 2008	SG May–Aug 2008	LK 2007	LK 2008 group 1	LK 2008 group 2
Non-structural polyprotein								
nsP1	60	60	T	A	.	.	.	.
	85	85	R	.	.	S*	.	.
	230	230	G	.	.	.	.	R*
	251	251	S	.	.	.	.	N*
	320	320	T	.	.	M*	.	.
	456	456	W	.	.	C*	.	.
	512	512	P	.	.	R*	.	.
nsP2	714	179	I	.	.	T*	.	.
	888	353	G	.	R*	.	.	.
	1074	539	L	.	S	.	.	.
	1237	702	A	.	.	.	.	S*
nsP3	1371	38	Y	.	.	.	H	.
	1483	150	K	.	.	.	R*	.
	1727	394	M	.	.	I*	I	.
	1777	444	T	.	.	.	.	M
	1782	449	M	.	.	.	.	T*
nsP4	1945	82	R	.	S	.	.	.
	2282	419	G	.	.	.	.	S*
	2426	563	I	T	.	.	.	.
Structural polyprotein								
E3	279	18	S	.	.	.	F	.
E2	359	34	L	.	.	.	I*	.
	362	37	I	.	.	.	.	L*
	389	64	W	.	.	R*	.	.
	503	178	R	H	.	.	.	.
	523	198	R	.	.	.	Q*	.
	527	202	N	.	.	.	D*	.
	547	222	V	.	.	.	.	I
	558	233	K	.	.	.	.	E*
	577	252	K	.	Q	.	.	.
	587	262	A	.	.	.	.	V*
	677	352	P	L*	.	.	.	.
	691	366	M	.	.	.	T*	.
E1	919	110	K	.	.	.	.	R*
	1029	220	V	.	.	.	I*	.

*Indicates mutations distributed randomly among analysed isolates. The remaining mutations were found in all isolates of respective groups.

226 was removed from the analysis, which collapsed clusters A and B into a single cluster (Fig. 1b). All isolates beyond cluster B possessed the E1-A226V substitution and displayed a wide geographical distribution, indicating that acquisition of the E1-A226V substitution by the Indian sublineage could have played an important role in the recent spread of CHIKV in Asia. The network further indicated independent acquisition of the E1-A226V substitution by Indian, Indian Ocean Islands and African CHIKV populations, as they were not interlinked due to their genetic differences (Fig. 1a).

The hypothetical evolutionary pathway presented in Fig. 3 was constructed based on the compilation of evidence from network and full genome analyses. This was performed mainly to overcome the limited resolution possibly generated in the network due to the analysis of *E1* gene alone. As shown in the pathway, it was likely that three distinct groups of Sri Lankan isolates could have evolved independently. Similarly, two main clusters of Singapore isolates (Jan 2008 and May–Aug 2008) from two separate outbreaks in 2008 were also distinct from each other.

DISCUSSION

Phylogenetic analyses and genetic characterization of recent CHIKV isolates from Sri Lanka, Singapore, Malaysia and the Maldives revealed that they were genetically related to Indian isolates of the ECSA lineage. Besides the close proximity to India and frequent travel between India and these countries, the above observation was further strengthened by previous extensive chikungunya fever outbreaks in India since late 2005. Moreover, our findings revealed further evolution within the Indian sublineage that resulted in five distinct subclades of Sri Lankan and Singapore CHIKV isolates. A genome-wide analysis allowed us to redefine the previously reported *E1*-gene-based genetic signatures of these subclades and confirmed our earlier finding that the first three locally acquired chikungunya fever episodes in Singapore were most likely due to independent importations of CHIKV strains from India, Malaysia and Sri Lanka (Ng *et al.*, 2009). Based on this information, travel advisories were issued to the public during the peak of the chikungunya fever epidemic in Singapore in 2008. A similar approach may be useful as a strategic tool to reduce the imported case burden in other countries at risk of importing multiple CHIKV infections.

Apart from the novel mutations reported in the present study (Table 1), the distribution pattern of certain known amino acid substitutions was also noteworthy. The nsP1-Q488R and nsP3-V331A amino acid substitutions were absent in three out of four 2006–2007 Sri Lankan isolates (AB455493.1, AB455494.1 and FJ445427.2), although they were present in all 2008 Sri Lankan, Indian and Indian Ocean Islands isolates that emerged after 2005. The absence of these two mutations resulted in a separate subclade with robust bootstrap support (Fig. 2). However, the evolution

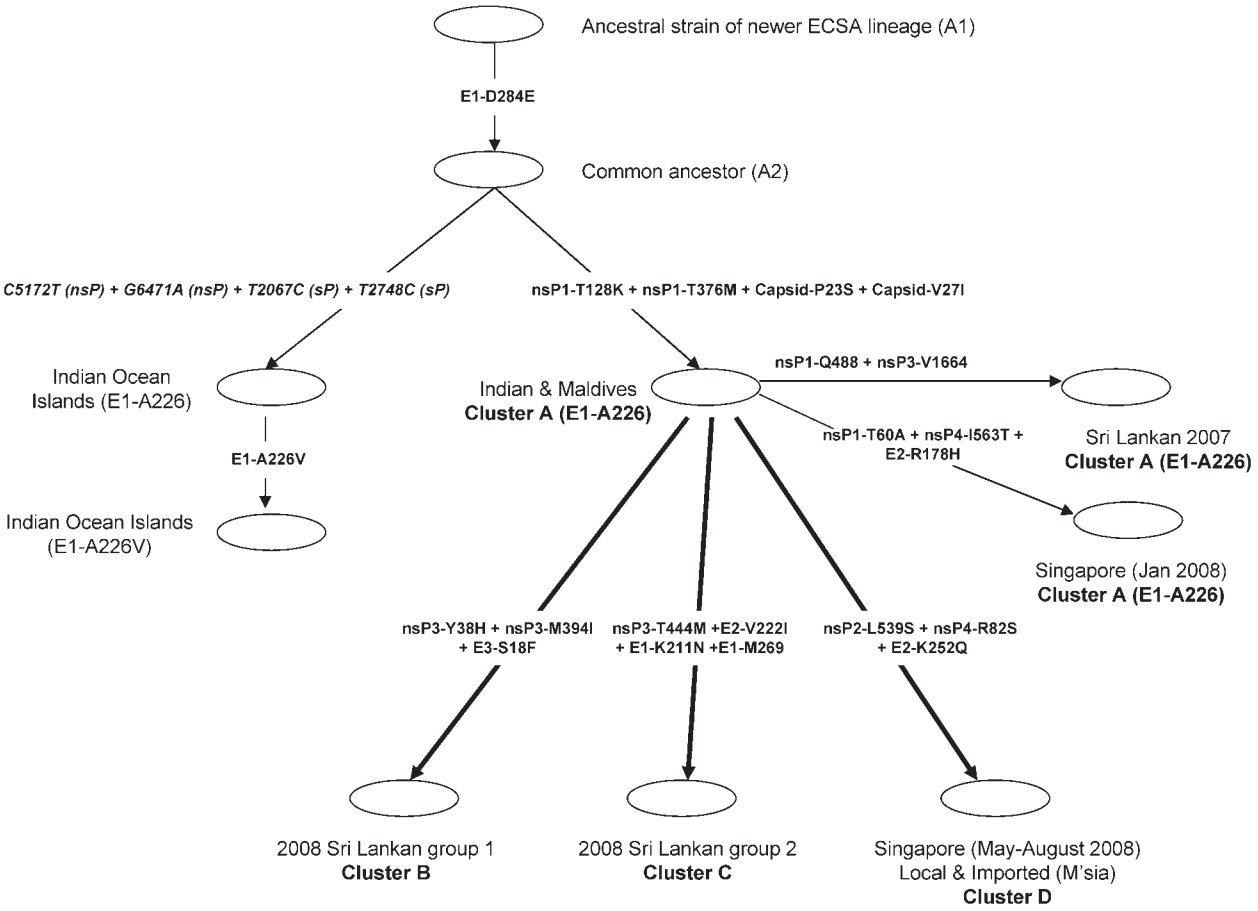

Fig. 3. Schematic illustration of hypothetical evolutionary pathway of Asian and Indian Ocean Islands CHIKV strains. Clusters A–D represent different clusters illustrated in Figs 1 and 2. The notable mutations involved in each step are shown on the arrows. Synonymous mutations involved in the Indian Ocean Islands sublineage are shown in bold italics with nucleotide position numbers. The numbers shown on all other steps are related to amino acid residues in the respective proteins. All isolates in cluster A showed the E1-A226 (wild-type) residue, while isolates in clusters B–D showed the E1-A226V substitution. Indian isolates with the E1-A226V substitution also grouped within the cluster B in the network. A1, ancestral strain of the newer ECSA lineage; A2, common ancestor of Indian and Indian Ocean Islands isolates; M'sia, Malaysia; nsP, non-structural polyprotein; sP, structural polyprotein.

within this subclade to acquire these changes was unclear. Similarly, the presence of E1-K211N and the absence of E1-M269V substitutions in all Sri Lankan group 2 isolates were notable. So far, the E1-K211N substitution has been reported only in two Indian isolates (IND-06-KA15, EF027135.1, Arankalle *et al.*, 2007; and IND_GJ51, Cherian *et al.*, 2009), yet the E1-M269V substitution has been detected in all Indian and Indian Ocean Islands isolates of the ECSA lineage. Hence, based on current sequence data, the evolutionary linkage of Sri Lankan group 2 isolates to existing CHIKV isolates from the region was unclear. However, the close clustering of one Sri Lankan isolate [LK(EH)CH18608, FJ513675] with group 2 isolates was noteworthy (Fig. 2). This particular isolate lacked the E1-M269V substitution that was characteristically absent in all group 2 isolates and was found in the same area from where group 2 isolates were almost exclusively found. Interestingly, as in the 2006–2007 Sri Lankan isolates, this isolate also did not possess the E1-A226V substitution, although this substitution was found in all group 2 isolates. These observations indicated a potential link between the previously existing strains and group 2 isolates in Sri Lanka and permitted two important evolutionary inferences: (1) group 2 isolates could have evolved from an existing CHIKV population circulated during early outbreaks in Sri Lanka; and (2) group 2 isolates were likely to have acquired the E1-A226V substitution in Sri Lanka. Furthermore, it was recently shown that amino acid substitutions in the E2 envelope protein of CHIKV could play a critical role in infection of *A. aegypti* and *A. albopictus* mosquitoes (Tsetsarkin *et al.*, 2009). Of two mutations (E2-G60D and E2-I211T) described by Tsetsarkin and colleagues, we observed only the E2-I211T substitution in all Sri Lankan and Singapore CHIKV isolates. In addition, we detected several novel amino acid substitutions in the E2 envelope protein of our study isolates (Table 1). Among these, E2-R178H, E2-K252Q and E2-V222I substitutions were prominent as they were uniformly distributed in respective groups of Sri Lankan and Singapore isolates (Table 1). Furthermore, the nsP1-T60A, nsP2-L539S, nsP3-Y38H, nsP3-T444M, nsP4-I563T and E3-S18F substitutions are likely to be structurally important due to their hydrophobic to hydrophilic amino acid replacements and vice versa. It will be of interest to study the effect of these novel mutations and combinations on the tertiary structure of the respective proteins and their significance for the phenotypic characteristics of the virus.

Another important genetic distinction noted was the distribution pattern of E1-A226V substitution in Sri Lankan and Singapore CHIKV isolates. The E1-A226V substitution is an adaptation of CHIKV that is known to enhance its transmissibility by *A. albopictus* (Tsetsarkin *et al.*, 2007; Vazeille *et al.*, 2007). Epidemiological data with regard to the spread of CHIKV in Singapore and Sri Lanka also strengthened the above notion. In Singapore, the mutant (E1-A226V) virus circulated in rural areas with abundance of *A. albopictus*. In contrast, the wild-type (E1-A226) virus was found in urban areas inhabited mainly by *A. aegypti* (Ng *et al.*, 2009). Likewise, spread of the mutant virus in Sri Lanka in 2008 was largely in banana and rubber plantation areas mainly inhabited by *A. albopictus*. In contrast, CHIKV spread mainly in coastal towns and urban areas in 2006 and 2007, during which period the presence of wild-type virus was evident. A similar scenario has been reported from India in 2007 (Kumar *et al.*, 2008; Santhosh *et al.*, 2008). Considering the phylogenetic and epidemiological evidence, we postulated that the acquisition of E1-A226V substitution by Indian virus strains was one of the important evolutionary events that contributed to the transmission and spatial distribution of CHIKV in the Asian region. Phylogenetic analysis also supported the recent evidence that the E1-A226V substitution could have appeared independently in the Indian Ocean Islands, India and Africa (Cameroon and Gabon) (de Lamballerie *et al.*, 2008). Based on the molecular epidemiological data, we suggest that Sri Lankan group 2 CHIKV isolates could also have acquired the E1-A226V substitution independently in Sri Lanka. The availability of more sequences from affected regions at different time points would provide more evidence for this rare phenomenon of genetic convergence in CHIKV (de Lamballerie *et al.*, 2008).

The evolutionary pathway presented here (Fig. 3) supports the assumption that Indian and Indian Ocean Islands CHIKV isolates could have emerged from a common ancestor descended from an ancestral strain of the ECSA lineage (Brisse *et al.*, 2007). Although a Uganda82-like CHIKV strain (AF192907.1) was recently implicated as the progenitor of the newer ECSA clade (Cherian *et al.*, 2009), phylogenetic analyses described previously (de Lamballerie *et al.*, 2008) and presented in the current study indicated that the Uganda-1982 strain was evolutionarily more related to African strains of the newer ECSA lineage than to the Indian and Indian Ocean Islands isolates (Fig. 1). This observation was supported by the presence of three non-synonymous mutations (E1-L219P, E1-A377T and E1-V410E) in the *E1* gene alone that were present only in Uganda-1982, Yawat-2000 (EF027139.1), Congo-2000 (AY549583.1) and Cameroon-2006 (EF051584.1) strains. Considering the temporal distribution, the network indicated an evolutionary direction from Uganda-1982 to Yawat-2000 strains (Fig. 1), which was epidemiologically supported by the coincidental circulation of Congo-2000 strains that were more closely related to Yawat-2000 than to Uganda-1982 strains. Consequently, it was unclear whether the forward evolutionary pathway bifurcated at the hypothetical node closest to the Yawat-2000 strain and led towards the Indian and Indian Ocean Islands isolates (Fig. 1a). Thus, the current analysis could not completely rule out the assumption made by Cherian *et al.* (2009). Although we could not ascertain the evolutionary position of 2004 Kenyan strains due to lack of their sequences in current databases, we assume that they could fall either on or after the hypothesized common ancestor of Indian and Indian Ocean Islands isolates (A2 in Figs 1 and 3). This was

because the Kenyan strains also possessed the E1-D284E substitution (Kariuki Njenga *et al.*, 2008). Even though the mean evolutionary rate of the *E1* gene has been shown to approximate that of the full genome of CHIKV (Cherian *et al.*, 2009), the availability of more full genome sequences could better reflect the genome-wide variability and would provide answers to current gaps of knowledge with regard to the evolution of CHIKV.

In conclusion, our findings showed that the E1-A226V variant predominantly circulated in Sri Lanka, Singapore and Malaysia in 2008. This indicates that the mutant CHIKV variant is well established in the Asian region. As *A. albopictus* is a common vector species in this region and is a rural and outdoor mosquito, these observations have important implications for the revision of current vector control strategies in areas affected by CHIKV in the Asian region. Moreover, as the global distribution of *A. albopictus* is greater than that of *A. aegypti*, CHIKV, especially the E1-A226V variant, will remain as a global threat in years to come.

METHODS

Study sites and source of viruses. All Sri Lankan CHIKV isolates analysed in the present study were obtained from sera of infected individuals presented to government hospitals in the Pallebedda (6° 55′ 0″ N, 81° 26′ 0″ E) and Eheliyagoda (6° 51′ 0″ N, 80° 16′ 0″ E) areas in March and April 2008 respectively. In 2008, Pallebedda and Eheliyagoda were among the worst chikungunya-affected areas in Sri Lanka. Both are rural areas with abundant banana and rubber plantations. Ten samples collected in 2007 from the Gampaha and Matale districts in Sri Lanka were also included in the analyses. All samples were taken after obtaining written informed consent from each individual. While all *E1* gene sequences of Singapore isolates have been published previously (Ng *et al.*, 2009), CHIKV strains obtained from chikungunya fever-suspected sera were used for full genome sequencing. The ethical approval for this study was obtained from the Ethics Committee, Faculty of Medicine, University of Kelaniya, Sri Lanka, and the Bioethics Committee of the National Environment Agency, Singapore.

Screening of serum samples for CHIKV. The methodology used to screen serum samples for CHIKV is fully described elsewhere (Ng *et al.*, 2009). Briefly, viral RNA was extracted by using the QIAamp viral RNA mini kit (Qiagen) according to the manufacturer's instructions. A 354 bp fragment of the non-structural protein 1 gene of CHIKV was amplified by a real-time reverse-transcriptase polymerase chain reaction (real-time RT-PCR) protocol as described previously (Hasebe *et al.*, 2002). The amplification was performed in a LightCycler 2.0 system, using the LightCycler RNA Master SYBR Green I kit (Roche Diagnostics) according to the manufacturer's instructions.

cDNA synthesis and sequencing. cDNA was synthesized from RNA extracted from sera, using the SuperScript III First-Strand synthesis system for RT-PCR (Invitrogen) according to the manufacturer's instructions. Thirteen overlapping fragments covering the full genome sequence of CHIKV were amplified by using oligonucleotides constructed based on full genome sequences of recent CHIKV strains of the ECSA lineage, using Gene Runner 3.05 (Hastings Software) and Primer Select 5.03 (DNASTAR) software (Supplementary Table S1). The 50 μl reaction consisted of 250 μM dNTPs, 0.5 μM of each primer, 1× PfuTurbo buffer and 2.5 units PfuTurbo DNA polymerase (Stratagene). The amplification was performed in a GeneAmp 9700 thermal cycler (Applied Biosystems) using the following protocol: initial denaturation at 94 °C for 1 min, followed by 40 cycles of 94 °C for 30 s, 55 °C for 30 s and 75 °C for 2 min, with a final extension step at 72 °C for 7 min. All templates were purified by using the QIAquick PCR purification kit (Qiagen). The sequencing was performed by a commercial laboratory according to the BigDye Terminator Cycle Sequencing kit (Applied Biosystems) protocol.

Sequence and phylogenetic analyses. The contiguous sequences of the full genome and *E1* gene for each isolate were assembled from overlapping sequence data by using SeqMan II version 5.03 (DNASTAR). The resulting sequences were aligned using CLUSTAL W software in the BioEdit Sequence Alignment Editor version 7.0.9.0 (Hall, 1999). The coding sequence of the African prototype (GenBank accession no. NC_004162) was used as the reference sequence in subsequent analyses.

Phylogenetic tree. Due to low bootstrap support obtained from the analysis of the *E1* gene alone (Supplementary Fig. S1), a phylogenetic tree (Fig. 2) was constructed based on 11 580 nt of Sri Lankan (n=17) and Singapore (10 locally acquired and 3 imported from Malaysia) CHIKV full genomes, together with 22 sequences reported previously from India (n=9), Indian Ocean Islands (n=10), Japan (imported from Sri Lanka, n=2) and Tanzania (n=1). The full genome tree was inferred by using the maximum-likelihood (ML) method implemented in PAUP* version 4.0b10 (Swofford, 1999). The Tamura–Nei model with gamma distribution of rate variation among sites (TrN+G) was chosen as the best-fit model of nucleotide substitutions based on hierarchical likelihood ratio test in Modeltest 3.7 (Posada & Crandall, 1998). A neighbour-joining (NJ) tree was used as the starting tree for tree bisection and reconnection (TBR) branch swapping. Three iterative rounds of ML analysis were carried out and the most likely tree obtained from each round was used as the starting tree in the next round of tree searching. Bootstrapping to access the robustness of the ML tree topology was performed using the NJ method under the ML criterion based on 1000 replicates.

Evolutionary network. The median joining network (Fig. 1) was constructed using the Network version 4.5.0.2 software (Bandelt *et al.*, 1999). The network presented was constructed using 1002 nt of the *E1* gene, spanning from amino acid residues 91 to 424 of Sri Lankan (n=60), Singapore (n=83), Indian, Indian Ocean Islands and African (n=35, Supplementary Table S2) CHIKV strains. Sri Lankan viruses consisted of 14 isolates collected from the Pallebedda area and 44 isolates from the Eheliyagoda area in 2008 and 2 isolates collected from the Matale and Ragama areas in 2007 (Supplementary Table S3). Singapore isolates represented 49 locally acquired and 34 imported strains detected from January 2007 to August 2008 (Supplementary Table S4). The imported strains were obtained from patients with a recent history of travel to Malaysia (n=31), the Maldives (n=1) and India (n=2). Only *E1* gene sequences were selected for the network analysis due to the limited number of full genome sequences available in sequence databases. The sequences belonging to Asian and West African lineages were not included in the network presented here due to their wide divergence from the study sequences.

The network allows visualization of mutational paths based on single nucleotide polymorphisms (SNPs) in analysed sequences. It clusters similar sequences together and different clusters are then linked by using the median joining algorithm to form an interconnecting network. In the network diagram, circles represent either individual isolates or clusters. The diameter of each circle is proportional to the number of identical sequences. The software (Network version 4.5.0.2) predicts hypothetical ancestral genotypes ('median vectors') if there are no candidates in the available set of sequences to represent certain mutational steps. This may happen either because the

particular genotype is extinct at present or simply because although extant, it had not been sampled,. The network diagram is especially helpful to determine the relationships between closely related sequences, as it shows mutational differences between two evolutionary steps. Sequences that are divergent even by a single mutation are shown isolated from their parent clusters and the distance between two linked isolates differs based on the number of different mutations. The direction of the evolutionary path can be inferred based on the temporal distribution of the respective sequences. When sequences represent different geographical regions collected at known time points, the network analysis can provide clues regarding possible origins of circulating viruses in different regions.

ACKNOWLEDGEMENTS

We thank Regional Health Authorities, Medical Officers and the staff at government hospitals in the study areas in Sri Lanka for their assistance in sample collection and processing. Our sincere thanks go to University of Kelaniya (Research Grant RP/03/04/16/01/2007) and A* STAR Biomedical Research Council, Singapore (Research Grant SIgN-07-010) for funding this study. The authors declare that they have no competing interests.

REFERENCES

AbuBakar, S., Sam, I. C., Wong, P. F., MatRahim, N. A., Hooi, P. S. & Roslan, N. (2007). Reemergence of endemic Chikungunya, Malaysia. *Emerg Infect Dis* **13**, 147–149.

Arankalle, V. A., Shrivastava, S., Cherian, S., Gunjikar, R. S., Walimbe, A. M., Jadhav, S. M., Sudeep, A. B. & Mishra, A. C. (2007). Genetic divergence of Chikungunya viruses in India (1963–2006) with special reference to the 2005–2006 explosive epidemic. *J Gen Virol* **88**, 1967–1976.

Bandelt, H. J., Forster, P. & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**, 37–48.

Brisse, S., Iteman, I. & Schuffenecker, I. (2007). Chikungunya outbreaks. *N Engl J Med* **356**, 2650–2652.

Cherian, S. S., Walimbe, A. M., Jadhav, S. M., Gandhe, S. S., Hundekar, S. L., Mishra, A. C. & Arankalle, V. A. (2009). Evolutionary rates, timescale comparison of Chikungunya viruses inferred from the whole genome/E1 gene with special reference to the 2005–07 outbreak in the Indian subcontinent. *Infect Genet Evol* **9**, 16–23.

de Lamballerie, X., Leroy, E., Charrel, R. N., Ttsetsarkin, K., Higgs, S. & Gould, E. A. (2008). Chikungunya virus adapts to tiger mosquito via evolutionary convergence: a sign of things to come? *Virol J* **5**, 33.

Enserink, M. (2007). Chikungunya: no longer a third world disease. *Science* **318**, 1860–1861.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**, 95–98.

Hapuarachchi, H. A. C., Bandara, K. B. A. T., Hapugoda, M. D., Williams, S. & Abeyewickreme, W. (2008). Laboratory confirmation of dengue and chikungunya co-infection. *Ceylon Med J* **53**, 104.

Hasebe, F., Parquet, M. C., Pandey, B. D., Mathenge, E. G. M., Morita, K., Balasubramaniam, V., Saat, Z., Yusop, A., Sinniah, M. & other authors (2002). Combined detection and genotyping of Chikungunya virus by a specific reverse transcription-polymerase chain reaction. *J Med Virol* **67**, 370–374.

Hochedez, P., Jaureguiberry, S., Debruyne, M., Bossi, P., Hausfater, P., Brucker, G., Bricaire, F. & Caumes, E. (2006). Chikungunya infection in travelers. *Emerg Infect Dis* **12**, 1565–1567.

Kariuki Njenga, M., Nderitu, L., Ledermann, J. P., Ndirangu, A., Logue, C. H., Kelly, C. H. L., Sang, R., Sergon, K., Breiman, R. & Powers, A. M. (2008). Tracking epidemic Chikungunya virus into the Indian Ocean from East Africa. *J Gen Virol* **89**, 2754–2760.

Kumar, C. V. M. N., Johnson, A. M. A. & Gopal, D. V. R. S. (2007). Molecular characterization of Chikungunya virus from Andhra Pradesh, India and phylogenetic relationship with Central African isolates. *Indian J Med Res* **126**, 534–540.

Kumar, N. P., Joseph, R., Kamaraj, T. & Jambulingam, P. (2008). A226V mutation in virus during the 2007 Chikungunya outbreak in Kerala, India. *J Gen Virol* **89**, 1945–1948.

Lanciotti, R. S., Kosoy, O. L., Laven, J. J., Panella, A. J., Velez, J. O., Lambert, A. J. & Campbell, G. L. (2007). Chikungunya virus in US travelers returning from India, 2006. *Emerg Infect Dis* **13**, 764–776.

Leo, Y. S., Chow, A. L. P., Tan, L. K., Lye, D. C., Lin, L. & Ng, L. C. (2009). Chikungunya outbreak, Singapore, 2008. *Emerg Infect Dis* **15**, 836–837.

Munasinghe, D. R., Amarasekera, P. J. & Fernando, C. F. (1966). An epidemic of dengue-like fever in Ceylon (Chikungunya) – a clinical and haematological study. *Ceylon Med J* **11**, 129–142.

Ng, L. C., Tan, L. K., Tan, C. H., Tan, S. S., Hapuarachchi, H. C., Pok, K. Y., Lai, Y. L., Lam-Phua, S. G., Bucht, G. & other authors (2009). Entomologic and virologic investigation of Chikungunya, Singapore. *Emerg Infect Dis* **15**, 1243–1249.

Noridah, O., Paranthaman, V., Nayar, S. K., Masliza, M., Ranjit, K., Norizah, I., Mustafa, B., Kumarasamy, V. & Chua, K. B. (2007). Outbreak of Chikungunya due to virus of Central/East African genotype in Malaysia. *Med J Malaysia* **62**, 323–328.

Parola, P., de Lamballerie, X., Jourdan, J., Rovery, C., Vaillant, V., Minodier, P., Brouqui, P., Flahault, A., Raoult, D. & Charrel, R. N. (2006). Novel Chikungunya virus variant in travelers returning from Indian Ocean Islands. *Emerg Infect Dis* **12**, 1493–1499.

Peyrefitte, C. N., Rousset, D., Pastorino, B. A. M., Pouillot, R., Bessaud, M., Tock, F., Mansaray, H., Merle, O. L., Pascual, A. M. & other authors (2007). Chikungunya virus, Cameroon, 2006. *Emerg Infect Dis* **13**, 768–771.

Peyrefitte, C. N., Bessaud, M., Pastorino, B. A. M., Gravier, P., Plumet, S., Merle, O. L., Moltini, I., Coppin, E., Tock, F. & other authors (2008). Circulation of Chikungunya virus in Gabon 2006–2007. *J Med Virol* **80**, 430–433.

Posada, D. & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Powers, A. M., Brault, A. C., Tesh, R. B. & Weaver, S. C. (2000). Reemergence of Chikungunya and o'nyong-nyong viruses: evidence for distinct geographical lineages and distant evolutionary relationships. *J Gen Virol* **81**, 471–479.

Ravi, V. (2006). Re-emergence of Chikungunya virus in India. *Indian J Med Microbiol* **24**, 83–84.

Rezza, G., Nicoletti, L., Angelini, R., Romi, R., Finarelli, A. C., Panning, M., Cordioli, P., Fortuna, C., Boros, S. & other authors (2007). Infection with Chikungunya virus in Italy: an outbreak in a temperate region. *Lancet* **370**, 1840–1846.

Sam, I. C., Chan, Y. F., Chan, S. Y., Loong, S. K., Chin, H. K., Hooi, P. S., Ganeswrie, R. & AbuBakar, S. (2009). Chikungunya virus of Asian and Central/East African genotypes in Malaysia. *J Clin Virol* **46**, 180–183.

Santhosh, S. R., Dash, P. K., Parida, M. M., Khan, M., Tiwari, M. & Lakshmana Rao, P. V. (2008). Comparative full genome analysis revealed E1 A226V shift in 2007 Indian Chikungunya virus isolates. *Virus Res* **135**, 36–41.

Schuffenecker, I., Iteman, I., Michault, A., Murri, S., Frangeul, L., Vaney, M. C., Lavenir, R., Pardigon, N., Reynes, J. M. & other authors (2006). Genome microevolution of Chikungunya viruses causing the Indian Ocean outbreak. *PLoS Med* **3**, e263.

Sergon, K., Yahaya, A. A., Brown, J., Bedja, S. A., Mlindasse, M., Agata, N., Allaranger, Y., Ball, M. D., Powers, A. M. & other authors (2007). Seroprevalence of Chikungunya virus infection on Grande Comoros Island, Union of the Comoros, 2005. *Am J Trop Med Hyg* **76**, 1189–1193.

Swofford, D. L. (1999). PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4.04beta. Sunderland, MA: Sinauer Associates.

Theamboonlers, A., Rianthavorn, P., Praianantathavom, K., Wuttirattanakowit, N. & Poovorawan, Y. (2009). Clinical and molecular characterization of Chikungunya virus in South Thailand. *Jpn J Infect Dis* **62**, 303–305.

Tsetsarkin, K. A., Vanlandingham, D. L., McGee, C. E. & Higgs, S. (2007). A single mutation in Chikungunya virus affects vector specificity and epidemic potential. *PLoS Pathog* **3**, e201.

Tsetsarkin, K. A., McGee, C. E., Volk, S. M., Vanlandingham, D. L., Weaver, S. C. & Higgs, S. (2009). Epistatic roles of E2 glycoprotein mutations in adaptation of Chikungunya virus to *Ae. albopictus* and *Ae. aegypti* mosquitoes. *PLoS One* **4**, e6835.

Vazeille, M., Moutailler, S., Coudrier, D., Rousseaux, C., Khun, H., Huerre, M., Thiria, J., Dehecq, J., Fontenille, D. & other authors (2007). Two Chikungunya isolates from the outbreak of La Reunion (Indian Ocean) exhibit different patterns of infection in the mosquito, *Aedes albopictus. PLoS One* **2**, e1168.

Veber, M., Arroe, G. R. & Barfod, T. S. (2007). Chikungunya fever "that which bends" epidemic in Africa and Asia – several cases imported to Denmark. *Ugeskr Laeger* **169**, 2311–2313.

Yergolkar, P. N., Tandale, B. V., Arankalle, V. A., Sathe, P. S., Sudeep, A., Gandhe, S. S., Gokhle, M. D., Jacob, G. P., Hundekar, S. L. & Mishra, A. C. (2006). Chikungunya outbreaks caused by African genotype, India. *Emerg Infect Dis* **12**, 1580–1583.

Yoosuf, A. A., Shiham, I., Mohamed, A. J., Ali, G., Luna, J. M., Pandav, R., Gongal, G. N., Nisaluk, A., Jarman, R. G. & other authors (2009). First report of Chikungunya from the Maldives. *Trans R Soc Trop Med Hyg* **103**, 192–196.