รายงานวิจัยฉบับสมบูรณ์

โครงการการศึกษาพันธุกรรมและคุณลักษณะของไวรัสชิคุนกุนยาที่มีการปรับตัวเข้ากับเซลล์เพาะเลี้ยง

โดย ดร.รจนภร พูลมานะอุสาหะกุล

สิงหาคม ๒๕๕๘

สัญญาเลขที่ MRG5480128

รายงานวิจัยฉบับสมบูรณ์

โครงการการศึกษาพันธุกรรมและคุณลักษณะของไวรัสชิคุนกุนยาที่มีการปรับตัวเข้ากับเซลล์เพาะเลี้ยง

ดร.รจนภร พูลมานะอุสาหะกุล
สถาบันชีววิทยาศาสตร์โมเลกุล
มหาวิทยาลัยมหิดล

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย

Abstract

Project Code : MRG5480128

Project Title : Genetic and phenotypic characterization of tissue cultured-adapted chikungunya viruses

Investigator : Rojjanaporn Pulmanausahakul
Institute of molecular biosciences, Mahidol University

E-mail Address : rojjanaporn.pul@mahidol.ac.th

Project Period : 15 มิถุนายน 2554 –
(ระยะเวลาโครงการ)

Chikungunya virus (CHIKV), a mosquitoes-borne *Alphavirus*, is an etiological agent of chikungunya fever. The virus is an enveloped positive single stranded RNA virus. In 2005-2006, CHIKV re-emerged and caused the major outbreak started from the Indian Ocean islands and spread to many countries in Asia including Thailand. In this project, the whole genome of CHIKV isolated from a Thai patient was sequenced. The results showed the virus belongs to the East Central South African (ECSA) genotype with an alanine to valine substitution at position 226 of the E1 protein, which is consistent with other isolates in the same outbreak. The CHIKV isolates was then undergone through two independent plaque purification series in Vero and C6/36 cells, resulting in the viruses that produced large and small plaque size, respectively. The viruses exhibited different levels of pathogenicity after intracranially infected to suckling mice in which the large plaque virus is highly pathogenic whereas the small plaque virus is almost non-pathogenic. Comparison of the genetic sequences of the Vero adapted-virus and parental isolates revealed 3 amino acid differences, nsP2: P618R, nsP3: G117R and E2: N187K, while the C6/36 adapted-virus differ from the parental virus only one position at nsP2: D546G.

Keywords: Chikungunya virus, Thai isolate, Tissue culture-adaptation, Plaque purification, Viral pathogenicity

บทคัดย่อ

ชื่อโครงการ: การศึกษาพันธุกรรมและคุณลักษณะของไวรัสชิคุนกุนยาที่มีการปรับตัวเข้ากับเซลล์เพาะเลี้ยง

ไวรัสชิคุนกุนยาเป็นสาเหตุของโรคไข้ปวดข้อชิคุนกุนยาซึ่งมียุงเป็นพาหะ ไวรัสนี้จัดอยู่ในวงศ์ *alphavirus* ซึ่งเป็นไวรัสที่มีเปลือกหุ้มและมีจีโนมเป็นอาร์เอ็นเอสายบวก ในช่วงปีค.ศ. 2005 ถึง 2006 ได้เกิดการระบาดของไวรัสชิคุนกุนยาครั้งใหญ่ โดยเริ่มเกิดขึ้นที่หมู่เกาะในมหาสมุทรอินเดีย และแพร่ระบาดมาจนถึงทวีปเอเชีย รวมทั้งประเทศไทย ซึ่งโครงการนี้ได้ทำการศึกษาไวรัสชิคุนกุนยาที่แยกได้จากผู้ป่วยในประเทศไทย โดยทำการหาลำดับเบสทั้งจีโนมของไวรัส หลังจากนำผลที่ได้มาทำการเปรียบเทียบกับฐานข้อมูล พบว่าไวรัสนี้จัดอยู่ในกลุ่ม ECSA และมีการเปลี่ยนแปลงกรดอะมิโนที่ตำแหน่ง 226 ของโปรตีน E1 จาก alanine เป็น valine ซึ่ง สอดคล้องกับผลของผู้วิจัยอื่นที่ทำการศึกษาไวรัสชิคุนกุนยาจากการระบาดครั้งเดียวกันนี้ จากนั้นผู้วิจัยได้นำไวรัสชิคุนกุนยานี้ไปเลี้ยงในเซลล์ไตลิง (Vero) และมาทำให้บริสุทธิ์ด้วยวิธีการเลือก plaque ขนาดใหญ่ หรือเลี้ยงในเซลล์ยุง (C6/36) แล้วเลือก plaque ขนาดเล็ก โดยไวรัสที่มีลักษณะ plaque ต่างกันนี้แสดงความรุนแรงต่างกันหลังจากฉีดเข้าสมองหนูแรกเกิด ไวรัสที่สร้าง plaque ขนาดใหญ่มีความรุนแรงมาก ในขณะที่ไวรัสที่สร้าง plaque ขนาดเล็กเกือบจะไม่มีอันตรายต่อหนู เมื่อเปรียบเทียบลำดับเบสของไวรัสทั้งสองกับไวรัสต้นแบบที่แยกจากผู้ป่วย พบว่าไวรัสที่สร้าง plaque ใหญ่มีการสร้างกรดอะมิโนแตกต่างจากไวรัสต้นแบบ 3 ตำแหน่ง คือ ที่โปรตีน nsP2 ตำแหน่ง 618 จากโพรลีนเป็นอาร์จีนีน ที่โปรตีน nsP3 ตำแหน่ง 117 จากไกลซีนเป็นอาร์จีนีน และที่โปรตีน E2 ตำแหน่ง 187 จากแอสพาราจีนเป็นไลซีน ส่วนไวรัสที่สร้าง plaque ขนาดเล็กมีความแตกต่างจากไวรัสต้นแบบแค่ หนึ่งตำแหน่งที่ 546 ของโปรตีน nsP2 จากกรดแอสพาติกเป็นไกลซีน

Executive summary MRG5480128

ชื่อโครงการ

(ภาษาไทย) การศึกษาพันธุกรรมและคุณลักษณะของไวรัสชิคุนกุนยาที่มีการปรับตัวเข้ากับเซลล์เพาะเลี้ยง

(ภาษาอังกฤษ) Genetic and phenotypic characterization of tissue cultured-adapted chikungunya viruses

ชื่อหัวหน้าโครงการ

(ภาษาไทย) นางสาว รจนภร พูลมานะอุสาหะกุล

(ภาษาอังกฤษ) Miss Rojjanaporn Pulmanausahakul

(ตำแหน่งวิชาการ) อาจารย์

สถานที่ติดต่อ

ที่ทำงาน สถาบันชีววิทยาศาสตร์โมเลกุล มหาวิทยาลัยมหิดล
25/25 หมู่ 3 ถ.พุทธมณฑล สาย 4 ต.ศาลายา อ.พุทธมณฑล จ.นครปฐม 73170

โทรศัพท์ 02-441-9003-7 ต่อ 1479

email rojjanaporn.pul@mahidol.ac.th

ชื่อนักวิจัยที่ปรึกษา

(ภาษาไทย) นายประเสริฐ เอื้อวรากุล

(ภาษาอังกฤษ) Mister Prasert Auewarakul

(ตำแหน่งวิชาการ) ศาตราจารย์

สถานที่ติดต่อ

ที่ทำงาน สถาบันชีววิทยาศาสตร์โมเลกุล มหาวิทยาลัยมหิดล
25/25 หมู่ 3 ถ.พุทธมณฑล สาย 4 ต.ศาลายา อ.พุทธมณฑล จ.นครปฐม 73170

โทรศัพท์ 02-441-9003

โทรสาร 02-441-1013

email prasert.aue@mahidol.ac.th

ชื่ออธิการบดี หรือผู้รับมอบอำนาจให้ลงนามแทนอธิการบดี

ศาสตราจารย์คลินิก นายแพทย์อุดม คชินทร

1. ความสำคัญและที่มาของปัญหาที่ทำการวิจัย

Chikungunya is the dengue-like illness that produces a high fever, myalgia, skin rashes, severe and sometimes prolonged arthritis. Although it originated in Africa, chikungunya became epidemic throughout Asia including Thailand. Recently, the 2005-2006 massive outbreak of chikungunya in La Reunion Islands lead to 270,000 infected cases and more than a million cases were reported all over Asia in 2007. The situation of chikungunya outbreak in Thailand became more significant in the October of 2008. By the end of 2009, the total of 49,000 confirmed cases were reported throughout the country, particularly in the South. Unlike the prior outbreaks that caused by the Asian genotype of chikungunya virus, the severity of the disease in this most recent outbreak, which caused by the ECSA genotype with A226V mutation, was shown to be elevated. Hence, a better understanding about the pathogenicity of the virus is a great concern, as ultimately, it might lead to a chikungunya vaccine development that is still unavailable at present. Serially passages of the virus in non-natural host cells can lead to the virus-host adaptation.

2. วัตถุประสงค์

1. To obtain the complete genomic sequence of wild-type clinical isolate of chikungunya virus and perform phylogenetic analysis.
2. To obtain tissue culture-adapted chikungunya viruses that produce homogenous large or small plaques.
3. To genotypically and phenotypically characterize tissue culture-adapted chikungunya viruses in comparison to wild-type chikungunya virus and identify tentative genotypic markers of the plaque morphologies.

3. ระเบียบวิธีวิจัย

Complete genomic sequencing of clinical isolated virus

The wild-type virus, isolated from the patient from Phang-nga, was kindly given from Professor Sukathida Ubol, faculty of science, Mahidol University. Viral RNA was extracted using Trizol and used for first strand cDNA synthesis. Nine overlapping fragments covering the entire genome were amplified and cloned into sequencing plasmids. 5'end of the genome was amplified using RACE PCR. The sequences from 3 clones of each fragment were aligned to determine the consensus sequence. Then, the sequences of 9 fragments were joined to construct the complete genomic sequence.

Tissue culture-adapted viruses

Tissue cultured-adapted CHIKVs derived from several passages of viral propagation and plaque purified in suitable cell-lines. Vero cells were used for large plaque purification while C6/36 cells were used to select small plaque virus. Wild-type CHIKV was used in plaque assay and large or small plaques from terminal dilution were picked by using sterile pipette tips, suspended in serum-free medium and inoculated directly onto the Vero or C6/36 cells, respectively. The recovered viruses were harvested when CPE was observed. Plaque purifications were done several times until homogenous small or large size plaques were obtained.

In vitro characterization

Plaque morphologies of tissue cultured-adapted CHIKVs in comparison to wild-type virus were determined by standard plaque assay using Vero cells. The growth kinetics of the viruses was studied in C6/36 cells.

In vivo characterization

Mouse neurovirulent assay were performed to study the pathogenecities of the viruses. Groups of suckling mice were intracranially infected with wild-type or tissue cultured-adapted viruses and monitored for the survival for 21 days.

Determination of genotypic markers of plaque morphology

The whole genome sequencing of tissue cultured-adapted viruses were performed and the results were compared to that of wild-type virus to determine the tentative genotypic markers. The mutations were confirmed in 2 additional clones.

4. ผลการทดลอง

In this project, the whole genome of chikungunya virus (CHIKV) isolated from a Thai patient, WT-CHK025, was sequenced. The results showed the viral genome is 11,811 nucleotides in length. The virus belongs to the East Central South African (ECSA) genotype with an alanine to valine substitution at position 226 of the E1 protein, which is consistent with other isolates in the same outbreak. The CHIKV isolate was then undergone through two independent plaque purification series in Vero and C6/36 cells, resulting in the viruses that produced large and small plaque size, respectively. The viruses exhibited different levels of pathogenicity after intracranially infected to suckling mice in which the large plaque virus is highly pathogenic whereas the small plaque virus is almost non-pathogenic. Comparison of the genetic sequences of the Vero adapted-virus and parental isolates revealed 3 amino acid differences, nsP2: P618R, nsP3: G117R

and E2: N187K, while the C6/36 adapted-virus differ from the parental virus only one position at nsP2: D546G.

5. Output

อยู่ระหว่างการเตรียม manuscript สำหรับตีพิมพ์ในวารสารวิชาการระดับนานาชาติ โดยชื่อเรื่องที่คาดว่าจะตีพิมพ์คือ Genetic and phenotypic characterization of tissue cultured-adapted chikungunya viruses.

วัตถุประสงค์

1. To obtain the complete genomic sequence of wild-type clinical isolate of CHIKV and perform phylogenetic analysis.
2. To obtain tissue culture-adapted CHIKVs that produce homogenous large or small plaques.
3. To genotypically and phenotypically characterize tissue culture-adapted CHIKVs in comparison to wild-type CHIKV and identify tentative genotypic markers of the plaque morphologies.

วิธีทดลอง

Cells and viruses:

C6/36 cells, an *Aedes albopictus* larvae cells, were maintained at 28ºC in minimum essential medium (MEM; Gibco, CA, USA) supplemented with 10% heat inactivated fetal bovine serum (FBS; Gibco, CA, USA or HyClone, Cramlington, UK) and 100 unit of penicillin/streptomycin per ml. Vero cells, African green monkey kidney epithelial cells, and MRC-5 cells, human lung fibroblast, were maintained at 37 ºC, 5% CO_2 in Dulbecco's modified Eagles medium (DMEM; Gibco, CA, USA) supplemented with 5% FBS and 100 unit of penicillin/streptomycin per ml.

CHIKV was kindly given from Prof. Sukathida Ubol, Department of Microbiology, Faculty of Science, Mahidol University. The virus was isolated from a patient in Phang-nga, Thailand in 2009 and passaged two times in C6/36 cells. The virus was designated as WT-CHK025. CHIKV stock was propagated in C6/36 cells. The culture medium and the infected cells was harvested at 40 hr after infection and stored at -80 ºC until use.

Standard plaque assay:

Vero cells were used for the plaque assay since they display distinct cytopathic effect (CPE). The cells were seeded in 6 wells plate with DMEM, supplemented with 5% heat inactivated FBS. Viruses were 10-fold serially diluted in BA-1 medium (1xM-199E, 1 M Tris-HCl pH 7.6, 2% (w/v) BSA, 100 units of penicillin/streptomycin per ml, 0.075% (w/v) NaHCO3) and 200 µl of each dilution was used to infect 90-95% confluent of monolayer Vero cells. The plates were incubated and rocked every 10 min for 2 hr at 37 ºC. After that, each well was overlaid with 4 ml of nutrient agarose [0.8%agarose (SeaKem LE, USA) containing Earle's balanced salts supplemented with 0.5% (w/v) yeast extract, 2.5% lactalbumin hydrolysate, 3% FBS] and incubated at 37ºC, 5% CO_2 for 3 or 6 days until plaque formation. The cell layer was stained with the second

overlay containing 1% neutral red and incubated as above condition for 16 hr for plaque observation. To determine virus titers, plaque assay was done in duplicated. The plaques were counted and calculated as plaque forming unit per ml (pfu/ml).

Tissue cultured-adapted viruses:

Tissue cultured-adapted CHIKV derived from several passages of viral propagation and plaque purified in suitable cell-lines. Vero cells were used for large plaque purification while C6/36 cells were used to select small plaque virus. WT-CHK025 was used in plaque assay and large or small plaques from terminal dilution were picked by using sterile pipette tips, suspended in serum-free medium and inoculated directly onto the Vero or C6/36 cells, respectively. The recovered viruses were harvested when CPE was observed. Plaque purifications were done several times until homogenous small or large size plaques were obtained.

For Vero-adapted CHIKV, isolated CHIKV was propagated three times in C6/36 cells and two times in Vero cells (WT-CHK025 passage 5), followed by 6 plaque purifications alternated with propagation in Vero cells for the total of 17 passages. The virus showed homogenous large plaque morphology and was designated as CHK-L.

To obtain CHIKV with small plaque morphology, the process was started with WT-CHK025 passage 4 (isolated CHIKV that propagated two times in C6/36 cells and two times in Vero cells). Plaque assay was performed to separate small plaques. One clone of transferred small plaque virus was propagated in three different monolayer cell lines; Vero, MRC-5 and C6/36 to select optimum cell line for small plaque purification and C6/36 was selected. Plaque purifications were continued two times in Vero cells alternated with three times virus propagation in C6/36. Finally, homogenous small plaque morphology was observed after three times plaque purification and total 14 times viral propagations. This virus was designated as CHK-S.

Virus growth kinetics:

Monolayer of C6/36 cells were infected with CHK-L or CHK-S at MOI of 5. The viruses were adsorbed at room temperature for 2 hours. Then the cells were washed with PBS for 3 times and maintained in fresh media at 37ºC with 5% CO_2. Supernatants of infected cells were collected every 6 hours post-infection (hpi) until 48 hpi. The titers of the infectious viruses were determined by standard plaque assay using Vero cells.

Mouse Neurovirulence:

This study was kindly performed by the Center for Vaccine Development, Institute of Molecular Biosciences, Mahidol University. Three groups of suckling mice (>24 to 72 hr-old), 50, 55 and 45 mice/group, were intracranially injected (i.c.) with 10^3 pfu of WT-CHK025, CHK-S or CHK-L, respectively. The mice were observed daily for

21 days. The results were used to plot the survival curves and calculate for average survival times.

RNA extraction, cDNA synthesis and PCR

RNA was extracted from the supernatant or cell lysate of CHIKV passage 2 infected cells using Trizol reagent (Invitrogen, USA.) following the manufacturer's protocol. The RNA was resuspended in 50 µl of DEPC-treated water. One µg of RNA was reverse transcribed using Improm II™ reverse transcriptase and random hexamer or oligo-dT as a primer (Promega, USA.).

To design specific primers for PCR, several sequences of CHIKV isolates were aligned to get the consensus sequence using clustalW web online. Nine primer pairs were designed to amplify 9 overlapping fragments covering the entire 12 kb of CHIKV genome using ApE (Davis M. W.: http://www.biology.utah.edu/jorgensen/wayned/ape/) and Primers3 (*Lincoln S., et al.*: http://frodo.wi.mit.edu/primer3/) v.0.4.0. software (figure 1). Some primers were synthesized following the paper from E.Sreekumar, et al.,2010. Table 1 shows sequences and priming positions of the primer pair for each fragment.

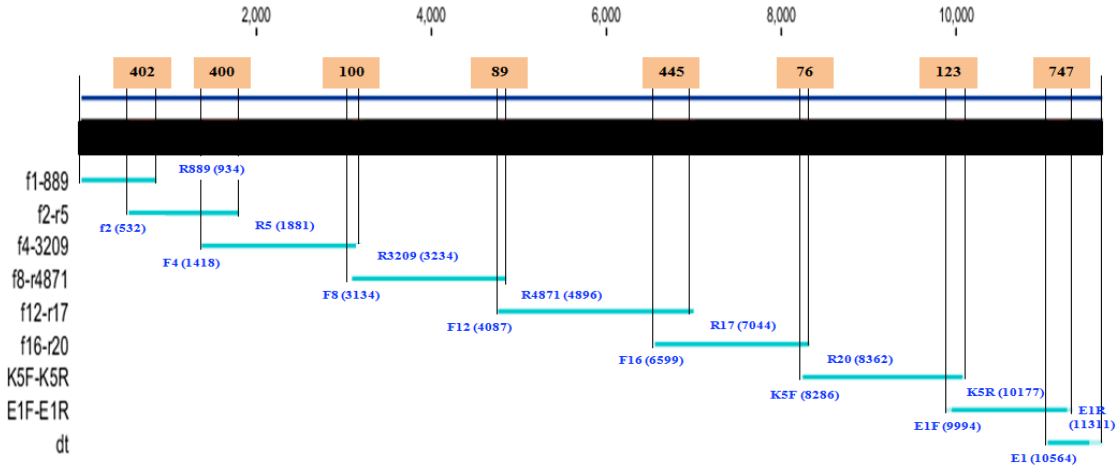

Figure 1: PCR primer pairs were designed to amplify nine overlapping fragments covering the entire 12 kb of CHIKV genome.

The cDNA was amplified by PCR with the following reaction mixture; cDNA, 1X buffer (20mM Tris-HCL (pH 8.4), 50 mM KCL), 1.5 mM MgCl2, 0.2 mM dNTP, 0.2 µM of each forward and reverse primers, 1U Platinum@ Taq DNA polymerase and the volume was adjusted to 50 µl with ddH_2O. The PCR condition is as follow;

94 ºC	2 min	
94 ºC	30 sec	35 cycles
52-60 ºC	30 sec	
68 ºC	2-3 min	
68 ºC	7 min	

The amplified PCR products were analyzed by gel electrophoresis and purified with Gel/PCR DNA fragments extraction kit (Geneaid, Taipei, Taiwan) according to manufacturer's protocol.

Table 1: Primer pairs for amplifying 9 fragments of CHIKV genome

Fragment	Primers	Primer sequence (5'-3')	nt position
Fragment1	F:anchor primer	GAC CAC GCG TAT CGA TGT CGA CTT TTT TTT TTT TTT TT	-
	R:R889	ACG TAG CCC TCA CAC GAA ACC	913-993
Fragment2	F:chf2	CGT CTA TGC TGT ACA CGC A	532-549
	R:chr5	CGT CGT ACG CTT CGA CCG	1863-1880
Fragment3	F:chf4	GAC AGC TTT GTG GTA CCG A	1418-1436
	R:R3209	GGG CTA CTT CAG GTG ACT ATG	3213-3233
Fragment4	F:chf8	CCT ATC CTC GAA ACA GCG	3134-3151
	R:R4871	TGA CGC GTT CTG GAG TCA TA	4876-4895
Fragment5	F:chf12	GAA ATG CCC GGT GGA TGA	4807-4824
	R:chr17	ACG CGG ATT TTG TCA GAC G	7025-7043
Fragment6	F:chf16	ATA CAG GCG GCT GAA CCC T	6599-6617
	R:chr20	GAT GGC AAG ACT CCA CTC T	8343-8361
Fragment7	F:K5F	CTC GGT GGT GAC CTG GAA TA	8286-8305
	R:K5R	TGT TGG CTC CAA AGT GAC TG	10097-10116
Fragment8	F:E1F	AAA CAT ATG TAC GAA CAC GTA ACA G	9994-10009
	R:E1R	AAA GGA TCC GTG CCT GCT GAA CGA CA	11294-11313
Fragment9	F:E1(10564)	GGC GCC TAC TGC TTC TG	10564-10582
	R:oligo dT	Oligo dT (16 mer)	-

The 5' end of the viral RNA was amplified using 5'/3' RACE, 2nd Generation Kit (Roche, Switzerland) following the manufacturer's instruction. Briefly, the RNA was reverse transcribed using specific primer chr5. RT reaction mixture was composed of 1xcDNA synthesis buffer, 0.5 mM dNTP, 12.5 µM chr5 primer, RNA 0.2-2 µg, 1 unit Transcriptor reverse transcriptase enzyme and DEPC-water up to 20 µl. RT reaction was incubated at 55 ºC for 60 minutes followed by additionally incubation at 85 ºC for 5 minutes. The cDNA product was purified by High pure PCR product Purification kit (Roche, Switzerland) according to the manufacturer's protocol. After that, a homopolymeric A-tail was added to the 3'end of purified cDNA product using recombinant terminal transferase enzyme. The reaction mixture (1xbuffer, 0.2 mM dATP and purified cDNA up to 24 µl) was incubated at 94 ºC for 3 minutes and immediately chilled on ice. Then 1µl of 80 unit/µl terminal transferase enzyme was added into the reaction and incubated at 37 ºC for 30 minutes. Finally, the enzyme was heat inactivated at 70 ºC for 10 minutes. The dA-tail cDNA was amplified using oligo d(T) anchor primer and specific primers R889 using Platinum@ Taq DNA polymerase enzyme with the condition described above. The PCR product was monitored by gel electrophoresis and purified by Gel/PCR DNA fragments extraction kit.

Cloning and sequencing

Each purified PCR fragment was cloned into the sequencing vector. Fragments 1-5 were cloned into p-SC-A vector (Takara, Japan) while fragments 6-8 were cloned into pCR@2.1vector (Invitrogen, USA.) and fragment 9 was cloned into pCR4-TOPO vector (Invitrogen, USA.).

Since the PCR fragment was amplified with Taq DNA polymerase, it can be directly cloned into the pCR@2.1 vector using TA cloning technique. The purified PCR product was ligated to the vector using T4 DNA ligase enzyme. The ligation reaction was composed of 1x ligation buffer (6 mM Tris-HCl, pH 7.5, 6 mM MgCl2, 5 mM NaCl, 0.1 mg/ml bovine serum albumin, 7 mM β-mercaptoethanol, 0.1 mM ATP, 2 mM dithiothreitol and 1 mM spermidine), 2 µl of pCR@2.1 vector (25 ng/µl plasmid DNA in 10 mM Tris-HCl, 1 mM EDTA, pH 8), 50 ng of purified PCR product, sterile-water up to 9 µl and 1 µl of 1unit/µl T4 DNA ligase. Then, this reaction was incubated at 16 ºC for overnight.

The PCR products were cloned into pCR4-TOPO and p-SC-A vectors using Topoisomerse I enzyme, which covalently bound to the vector. Therefore, the ligation reaction was incubated at room temperature for only 5 min. In detail, the ligation reaction of pCR4-TOPO vector was composed of 0.5-4.0 µl of purified PCR product (about 10-50 ng), 1 µl of salt solution (1.2 M NaCl, 0.06 M MgCl2) and 1µl of pCR4-TOPO vector (10 ng/µl plasmid DNA in 50% glycerol, 50 mM Tris-HCl pH 7.4, 1 mM

EDTA, 2 mM DTT, 0.1%Triton X-100, 100 µg/ml BSA and 30 µM phenol red). The p-SC-A ligation reaction was composed of 3µl of StrataClone Cloning buffer, 2 µl of purified PCR product (5-50 ng) and 1 µl of StrataClone vector. The ligation reaction was mix gently by pipetting and incubated as described. When the incubation was completed, the reaction was placed on ice.

Every ligation product was transformed into Escherichia coli DH5α competent cells using heat-shock method at 42 ºC for exactly 90 seconds. The transformant was grown on LB agar plate containing 100 ng/µl of amplicilin with 4 µl of 400 mM IPTG and 20 µl of 20 mg/ml X-gal. The bacterial colonies were initially screened using blue/white colony method. The white colonies were picked and cultured in LB broth containing 100 ng/µl of ampicillin at 37ºC for 12-16 hr. Then, the plasmid DNA was extracted using High-speed plasmid mini kit (Geneaid, Taipei, Taiwan) and the insert fragment was confirmed by restriction enzymes digestion analysis.

Three clones of recombinant plasmids of each fragment were sequenced at Tech Dragon Limited (Hong Kong) and Macrogen (Korea). The plasmids were sequenced with two universal primers, M13F and M13R, and one specific internal primer with the exception of fragment 2 that was sequenced with only two specific PCR primers, chf2 and chr5. Finally, fragment 9 was sequenced with one universal primer at 3´end of this fragment. The sequences of the specific internal primers are shown in table 2. Three clones of recombinant plasmids of each fragment were aligned to confirm and generate the consensus sequence. Then the sequences of 9 fragments were combined to reconstruct the whole CHIKV genome using BioEdit Sequence Alignment Editor Program (version 7.0.5.3). The CHIKV nucleotide and amino acid sequence were subjected to the phylogenetic analysis using CLC program based on complete genome sequences.

Table 2: Sequencing primers

Fragment	Sequencing primers	Primer sequence (5'-3')	nt position
Fragment1	R: Chr2	TGC GTG TAC AGC ATA GAC G	532-549
Fragment3	R: Chr5	CGT CGT ACG CTT CGA CCG	1863-1880
Fragment4	F: FS3584	AAC CTT GCA CTG CCT ACT AAG AGA G	3608-3632
Fragment5	F: FS5229	TAA TGA GCA CCG TAG CTG TC	5253-5272
Fragment6	R: Chr19	ACC TTC GTG CTT GAC TTC G	7911-7929
Fragment7	F: FS8793	GAT TAC TGG AAC AAT GGG ACA C	8817-8838
Fragment8	F: E1F10564	CCC TTT GGC GCA GGA AGA C	10564-10582

The sequences of CHK-L and CHK-S were obtained from RT-PCR directly. The viral RNAs were extracted, cDNAs were generated and then 9 fragments were amplified by PCR using the primer pairs shown in table 1. The products were sequenced using the same PCR primers excepted fragment 9 that was cloned and sequenced with universal primer. The sequences of plaque-purified viruses were reconstructed the same as those of WT-CHK025.

To identify the tentative genotypic markers for plaque morphologies, the sequences of CHK-L and CHK-S were compared to WT-CHK025 sequence using BioEdit Sequence Alignment Editor software. The differences were confirmed with two others plaque purified CHIKV clones.

ผลการทดลอง

Complete genomic sequencing of wild type CHIKV isolates 025

To sequence the genome of WT-CHK025, nine overlapping fragments covering the entire 12 kb were amplified by RT-PCR as described earlier. Figure 2 shows the result of agarose gel electrophoresis of all 9 fragments.

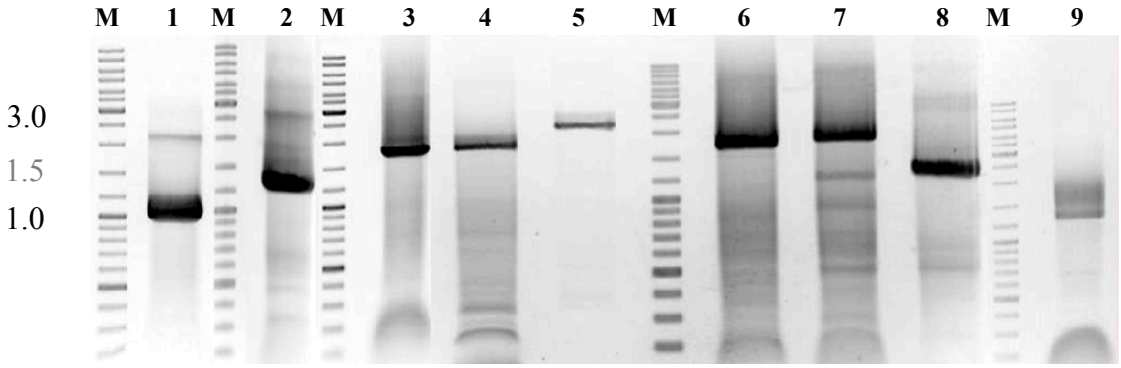

Figure 2: Agarose gel electrophoresis of the RT-PCR products of fragment 1-9 of CHIKV genome. The sizes are about 1, 1.3, 1.8, 1.8, 2.2, 1.8, 1.8, 1.3 and 1.3 kb, respectively. M: DNA marker; lane1-9: RT-PCR products of fragment 1-9.

The products of each fragment were cloned into sequencing vectors and checked by restriction enzyme digestion analysis. Three clones of recombinant plasmids of each fragment were sequenced and aligned to determine the consensus sequence and reconstructed to obtain complete genomic sequence of WT-CHK025. The result was submitted to GenBank, NCBI. The accession number is KP164869.

The size of the entire genome of WT-CHK025 is 11,811 bp with the 5' and 3' NTR of 76 and 498 bp, respectively. The length of the first open reading frame (ORF) is

7,422 nucleotides while the second ORF is 3,747 nucleotides. In addition, there is a 68 nucleotide-untranslated junction between the two ORFs.

The first ORF comprises four non-structural genes; nsP1 to nsP4, which are 1,605, 2,394, 1,590, and 1,833 nucleotides in length, respectively. After being translated and proteolytically cleaved, the sizes of nsP1-4 are 535, 798, 529 and 611 amino acids, respectively. At C-terminal of nsP3, there is an opal stop codon (UGA) at nucleotide position 5,645-5,647. The second ORF consists of capsid, E3, E2, 6k and E1 genes, which are 783, 192, 1,269, 138 and 1,320 nucleotides long or 261, 64, 423, 46 and 440 amino acids after being expressed, respectively. The amino acid at position 226 of E1 protein is valine, which is hypothesized to affect the vector specificity that responsible for this massive outbreak. At the 3'end, there are poly (A) tracts and 19-nucleotides of conserved sequence element (CSE), 5'-AUUUUGUUUUUAAUAUUC-3', which is conserved among alphavirus. Moreover, 3'NTR contains three repeats sequence elements (RSEs) at nucleotides position 11,384-11,418, 11,513-11,547 and 11,598-11,632. The whole genome sequence of WT-CHK025 is shown in figure 3.

1	ATGGCTGCGT	GAGACACACG	TAGCCTACCA	GTTTCTTACT	GCTCTACTCT	GCAAAGCAAG
61	AGATTAATAA	CCCATCATGG	ATCCTGTGTA	CGTGGACATA	GACGCTGACA	GCGCCTTTTT
121	GAAGGCCCTG	CAACGTGCGT	ACCCCATGTT	TGAGGTGGAA	TCAAGGCAGG	TCACACCGAA
181	TGACCATGCT	AATGCTAGAG	CGTTCTCGCA	TCTAGCTATA	AAACTAATAG	AGCAGGAAAT
241	TGACCCCGAC	TCAACCATCC	TGGATATCGG	CAGTGCGCCA	GCAAGGAGGA	TGATGTCGGA
301	CAGGAAGTAC	CACTGCGTCT	GCCCGATGCG	CAGTGCGGAA	GATCCCGAGA	GACTCGCTAA
361	TTATGCGAGA	AAGCTAGCAT	CTGCCGCAGG	AAAAGTCCTG	GACAGAAACA	TCTCTGGAAA
421	GATCGGGGAC	TTACAAGCAG	TAATGGCCGT	GCCAGACAAG	GAGACGCCAA	CATTCTGCTT
481	ACACACAGAC	GTCTCATGTA	GACAGAGAGC	AGACGTCGCT	ATATACCAAG	ACGTCTATGC
541	TGTATACGCA	CCCACGTCGC	TATACCACCA	GGCGATTAAA	GGGGTCCGAG	TGGCGTACTG
601	GGTTGGGTTC	GACACAACCC	CGTTCATGTA	CGATGCCATG	GCGGGTGCCT	ACCCCTCATA
661	CTCGACAAAC	TGGGCAGATG	AGCAGGTACT	GAAGGCTAAG	AACATAGGAT	TATGTTCAAC
721	AGACCTGACG	GAAGGTAGAC	GAGGCAAGTT	GTCTATTATG	AGAGGGAAAA	AGCTAAAACC
781	GTGCGACCGT	GTGCTGTTCT	CAGTAGGGTC	AACGCTCTAC	CCGGAAAGCC	GCAAGCTACT
841	TAAGAGCTGG	CACCTGCCAT	CGGTGTTCCA	TTTAAAGGGC	AAACTCAGCT	TCACATGCCG
901	CTGTGATACA	GTGGTTTCGT	GTGAGGGCTA	CGTCGTTAAG	AGAATAACGA	TGAGCCCAGG
961	CCTTTATGGA	AAAACCACAG	GGTATGCGGT	AACCCACCAC	GCAGACGGAT	TCCTGATGTG
1021	CAAGACTACC	GACACGGTTG	ACGGCGAAAG	AGTGTCATTC	TCGGTGTGCA	CATACGTGCC
1081	GGCGACCATT	TGTGATCAAA	TGACCGGCAT	CCTTGCTACA	GAAGTCACGC	CGGAGGATGC
1141	ACAGAAGCTG	TTGGTGGGGC	TGAACCAGAG	AATAGTGGTT	AACGGCAGAA	CGCAACGGAA
1201	TATGAACACC	ATGAAAAATT	ATCTGCTTCC	CGTGGTCGCC	CAAGCCTTCA	GTAAGTGGGC
1261	AAAGGAGTGC	CGGAAAGACA	TGGAAGATGA	AAAACTCCTG	GGGGTCAGAG	AAAGAACACT
1321	GACCTGCTGC	TGTCTATGGG	CATTCAAGAA	GCAGAAAACA	CACACGGTCT	ACAAGAGGCC
1381	GGATACCCAG	TCAATTCAGA	AGGTTCAGGC	CGAGTTTGAC	AGCTTTGTGG	TACCGAGTCT
1441	GTGGTCGTCC	GGGTTGTCAA	TCCCTTTGAG	GACTAGAATC	AAATGGTTGT	TAAGCAAGGT
1501	GCCAAAAACC	GACCTGATCC	CATACAGCGG	AGACGCCCGA	GAAGCCCGGG	ACGCAGAAAA
1561	AGAAGCAGAG	GAAGAACGAG	AAGCAGAACT	GACTCGCGAA	GCCCTACCAC	CTCTACAGGC
1621	AGCACAGGAA	GATGTTCAGG	TCGAAATCGA	CGTGGAACAG	CTTGAGGACA	GAGCGGGCGC

1681 AGGAATAATA GAGACTCCGA GAGGAGCTAT CAAAGTTACT GCCCAACCAA CAGACCACGT
1741 CGTGGGAGAG TACCTGGTAC TCTCCCCGCA GACCGTACTA CGTAGCCAGA AGCTCAGTCT
1801 GATTCACGCT TTGGCGGAGC AAGTGAAGAC GTGCACGCAC AACGGACGAG CAGGGAGGTA
1861 TGCGGTCGAA GCGTACGACG GCCGAGTCCT AGTGCCCTCA GGCTATGCAA TCTCGCCTGA
1921 AGACTTCCAG AGTCTAAGCG AAAGCGCAAC GATGGTGTAT AACGAAAGAG AGTTCGTAAA
1981 CAGAAAGCTA CACCATATTG CGATGCACGG ACCAGCCCTG AACACCGACG AAGAGTCGTA
2041 TGAGCTGGTG AGGGCAGAGA GGACAGAACA CGAGTACGTC TACGACGTGG ATCAGAGAAG
2101 ATGCTGTAAG AAGGAAGAAG CCGCAGGACT GGTACTGGTG GGCGACTTGA CTAATCCGCC
2161 CTACCATGAA TTCGCATATG AAGGGCTAAA AATCCGCCCT GCCTGCCCAT ACAAAATTGC
2221 AGTCATAGGA GTCTTCGGAG TACCGGGATC TGGCAAGTCA GCTATTATCA AGAACCTAGT
2281 TACCAGGCAG GACCTGGTGA CTAGCGGAAA GAAAGAAAAC TGCCAAGAAA TCACCACCGA
2341 CGTGATGAGA CAGAGAGGTC TAGAGATATC TGCACGTACG GTTGACTCGC TGCTCTTGAA
2401 TGGATGCAAC AGACCAGTCG ACGTGTTGTA CGTAGACGAG GCGTTTGCGT GCCACTCTGG
2461 AACGCTACTT GCTTTGATCG CCTTGGTGAG ACCAAGGCAG AAAGTTGTAC TTTGTGGTGA
2521 CCCGAAGCAG TGCGGCTTCT TCAATATGAT GCAGATGAAA GTCAACTATA ATCACAACAT
2581 CTGCACCCAA GTGTACCACA AAAGTATCTC CAGGCGGTGT ACACTGCCTG TGACCGCCAT
2641 TGTGTCATCG TTGCATTACG AAGGCAAAAT GCGCACCACG AATGAGTACA ACAAGCCGAT
2701 TGTAGTGGAC ACTACAGGCT CAACAAAACC TGACCCTGGA GACCTCGTGT TAACGTGCTT
2761 CAGAGGGTGG GTTAAACAAC TGCAAATTGA CTATCGTGGA TACGAGGTCA TGACAGCAGC
2821 CGCATCCCAA GGGTTAACCA GAAAAGGAGT TTACGCAGTT AGACAAAAAG TTAATGAAAA
2881 CCCGCTCTAT GCATCAACGT CAGAGCACGT CAACGTACTC CTAACGCGTA CGGAAGGTAA
2941 ACTGGTATGG AAGACACTTT CCGGCGACCC GTGGATAAAG ACGCTGCAGA ACCCACCGAA
3001 AGGAAACTTC AAAGCAACTA TTAAGGAGTG GGAGGTGGAG CATGCATCAA TAATGGCGGG
3061 CATCTGCAGT CACCAAATGA CCTTCGATAC ATTCCAAAAT AAAGCCAACG TTTGTTGGGC
3121 TAAGAGCTTG GTCCCTATCC TCGAAACAGC GGGGATAAAA CTAAATGATA GGCAGTGGTC
3181 TCAGATAATT CAAGCCTTCA AAGAAGACAA AGCATACTCA CCTGAAGTAG CCCTGAATGA
3241 AATATGTACG CGCATGTATG GGGTGGATCT AGACAGCGGG CTATTTTCTA AACCGTCGGT
3301 GTCTGTGTAT TACGCGGATA ACCACTGGGA TAATAGGCCT GGAGGGAAAA TGTTCGGATT
3361 TAACCCCGAG GCAGCATCCA TTCTAGAAAG AAAGTACCCA TTCACAAAAG GGAAGTGGAA
3421 CATCAACAAG CAGATCTGCG TGACTACCAG GAGGATAGAA GACTTTAACC CTACCACCAA
3481 CATCATACCG GCCAACAGGA GACTACCACA CTCATTAGTG GCCGAACACC GCCCAGTAAA
3541 AGGGGAAAGA ATGGAATGGC TGGTTAACAA GATAAACGGC CACCACGTGC TCCTGGTCAG
3601 TGGCTATAAC CTTGCACTGC CTACTAAGAG AGTCACTTGG GTAGCGCCGT TAGGTGTCCG
3661 CGGAGCGGAC TACACATACA ACCTAGAGTT GGGTCTGCCA GCAACGCTTG GTAGGTATGA
3721 CCTTGTGGTC ATAAACATCC ACACACCTTT TCGCATACAC CATTACCAAC AGTGCGTCGA
3781 CCACGCAATG AAACTGCAAA TGCTCGGGGG TGACTCATTG AGACTGCTCA AACCGGGCGG
3841 CTCTCTATTG A TCAGAGCAT ATGGTTACGC AGATAGAACC AGTGAACGAG TCATCTGCGT
3901 ATTGGGACGC AAGTTTAGAT CGTCTAGAGC GTTGAAACCA CCATGTGTCA CCAGCAACAC
3961 TGAAATGTTT TTCCTATTCA GCAACTTTGA CAATGGCAGA AGGAATTTCA CAACTCATGT
4021 CATGAACAAT CAACTGAATG CAGCCTTCGT AGGACAGGTC ACCCGAGCAG GATGTGCACC
4081 GTCGTACCGG GTAAAACGCA TGGACATCGC GAAAAACGAT GAAAAGTGCG TAGTCAACGC
4141 CGCTAACCCT CGCGGGTTAC CGGGTGACGG TGTTTGCAAG GCAGTATACA AAAAATGGCC
4201 GGAGTCCTTT AAGAACAGTG CAACACCAGT GGGAACCGCA AAAACAGTTA TGTGTGGTAC
4261 GTATCCAGTA ATCCACGCTG TTGGACCAAA CTTCTCTAAT TATTCGGAGT CTGAAGGGGA
4321 CCGGGAATTG GCAGCTGCCT ATCGAGAAGT CGCAAAGGAA GTAACTAGGC TGGGAGTAAA
4381 TAGTGTAGCT ATACCTCTCC TCTCCACAGG TGTATACTCA GGAGGGAAAGACAGGCTGAC
4441 CCAGTCACTG AACCACCTCT TTACAGCCAT GGACTCGACG GATGCAGACG TGGTCATCTA
4501 CTGCCGCGAC AAAGAATGGG AGAAGAAAAT ATCTGAGGCC ATACAGATGC GGACCCAAGT
4561 AGAGCTGCTG GATGAGCACA TCTCCATAGA CTGCGATATT GTTCGCGTGC ACCCTGACAG
4621 CAGCTTGGCA GGCAGAAAAG GATACAGCAC CACGGAAGGC GCACTGTACT CATATCTAGA
4681 AGGGACCCGT TTTCATCAGA CGGCTGTGGA TATGGCGGAG ATACATACTA TGTGGCCAAA
4741 GCAAACAGAG GCCAATGAGC AAGTCTGCCT ATATGCCCTG GGGGAAAGTA TTGAATCTAT
4801 CAGGCAGAAA TGCCCGGTGG ATGATGCAGA CGCATCATCT CCCCCCAAAA CTGTCCCGTG
4861 CCTTTGCCGT TACGCTATGA CTCCAGAACG CGTCACCCGG CTTCGCATGA ACCACGTCAC
4921 AAGCATAATT GTGTGTTCTT CGTTTCCCCT CCCAAAGTAC AAAATAGAAG GAGTGCAAAA
4981 AGTCAAATGC TCTAAGGTAA TGCTATTTGA CCATAACGTG CCATCGCGCG TAAGTCCAAG
5041 GGAATATAGA TCTTCCCAGG AGTCTGCACA GGAGGCGAGT ACAATCACGT CACTGACGCA
5101 TAGTCAATTC GACCTAAGCG TTGATGGCGA GATACTGCCC GTCCCGTCAG ACCTGGATGC

5161 TGACGCCCCA GCCCTAGAAC CAGCACTAGA CGACGGGGCG ACACACACGC TGCCATCCAC
5221 AACCGGAAAC CTTGCGGCCG TGTCTGACTG GGTAATGAGC ACCGTACCTG TCGCGCCGCC
5281 CAGAAGAAGG CGAGGGAGAA ACCTGACTGT GACATGTGAC GAGAGAGAAG GGAATATAAC
5341 ACCCACGGCT AGCGTCCGAT TCTTTAGGGC AGAGCTGTGT CCGGTCGTAC AAGAAACAGC
5401 GGAGACGCGT GACACAGCAA TGTCTCTTCA GGCACCACCG AGTACCGCCA CGGAACCGAA
5461 TCATCCGCCG ATCTCCTTCG GAGCATCAAG CGAGACGTTC CCCATTACAT TTGGGGACTT
5521 CAACGAAGGA GAAATCGAAA GCTTGTCTTC T GAGCTACTA ACTTTCGGAG ACTTCTTACC
5581 AGGAGAAGTG GATGACTTGA CAGACAGCGA CTGGTCCACG TGCTCAGACA CGGACGACGA
5641 GTTATGACTA GACAGGGCAG GTGGGTATAT ATTCTCGTCG GACACCGGTC CAGGTCATTT
5701 ACAACAGAAG TCAGTACGCC AGTCAGTGCT GCCGGTGAAC ACCCTGGAGG AAGTCCACGA
5761 GGAGAAGTGT TACCCACCTA AGCTGGATGA AGCAAAGGAG CAACTATTAC TTAAGAAACT
5821 CCAGGAGAGT GCATCCATGG CCAACAGAAG CAGGTATCAG TCGCGTAAAG TAGAAAACAT
5881 GAAAGCAGCA ATCATCCAGA GACTAAAGAG AGGCTGTAGA CTATACTTAA TGTCAGAGAC
5941 CCCAAAAGTC CCTACTTACC GGACTACATA TCCGGCGCCT GTGTACTCGC CTCCGATCAA
6001 CGTCCGATTG TCCAATCCCG AGTCCGCAGT GGCAGCATGC AATGAGTTCT TAGCTAGAAA
6061 CTATCCAACT GTCTCGTCAT ACCAAATTAC CGACGAGTAT GATGCATATC TAGACATGGT
6121 GGACGGGTCG GAGAGTTGCC TGGACCGAGC GACATTCAAT CCGTCAAAAC TCAGGAGCTA
6181 CCCGAAACAG CACGCTTACC ACGCGCCCTC CATCAGAAGC GCTGTACCGT CCCCATTCCA
6241 GAACACACTA CAGAATGTAC TGGCAGCAGC CACGAAAAGA AACTGCAACG TCACACAGAT
6301 GAGGGAATTA CCCACTTTGG ACTCAGCAGT ATTCAACGTG GAGTGTTTCA AAAAATTCGC
6361 ATGCAACCAA GAATACTGGG AAGAATTTGC TGCCAGCCCT ATTAGGATAA CAACTGAGAA
6421 TTTAGCAACC TATGTTACTA AACTAAAAGG GCCAAAAGCA GCAGCGCTAT TCGCAAAAAC
6481 CCATAATCTA CTGCCACTAC AGGAAGTACC AATGGATAGG TTCACAGTAG ATATGAAAAG
6541 GGACGTGAAG GTGACTCCTG GTACAAAGCA TACAGAGGAA AGACCTAAGG TGCAGGTTAT
6601 ACAGGCGGCT GAACCCTTGG CGACAGCATA CCTATGTGGG ATTCACAGAG AGCTGGTTAG
6661 GAGGCTGAAC GCCGTCCTCC TACCCAATGT ACATACACTA TTTGACATGT CTGCCGAGGA
6721 TTTCGATGCC ATCATAGCCG CACACTTTAA GCCAGGAGAC ACTGTTTTGG AAACGGACAT
6781 AGCCTCCTTT GATAAGAGCC AAGACGATTC GCTTGCGCTT ACTGCTTTGA TGCTGTTAGA
6841 GGATTTAGGG GTGGATCACT CCCTGCTGGA CTTGATAGAG GCTGCTTTCG GAGAGATTTC
6901 CAGCTGTCAC CTACCGACAG GTACGCGCTT CAAGTTCGGC GCCATGATGA AATCAGGTAT
6961 GTTCCTAACT CTGTTCGTCA ACACATTGTT AAACATCACC ATCGCCAGCC GAGTGCTGGA
7021 AGATCGTCTG ACAAAATCCG CGTGCGCGGC CTTCATCGGC GACGACAACA TAATACATGG
7081 AGTCGTCTCC GATGAATTGA TGGCAGCCAG ATGTGCCACT TGGATGAACA TGGAAGTGAA
7141 GATCATAGAT GCAGTTGTAT CCTTGAAAGC CCCTTACTTT TGTGGAGGGT TTATACTGCA
7201 CGATACTGTG ACAGGAACAG CTTGCAGAGT GGCAGACCCG CTAAAAAGGC TTTTTAAACT
7261 GGGCAAACCG CTAGCGGCAG GTGACGAACA AGATGAAGAT AGAAGACGAG CGCTGGCTGA
7321 CGAAGTGATC AGATGGCAAC GAACAGGGCT AATTGATGAG CTGGAGAAAG CGGTATACTC
7381 TAGGTACGAA GTGCAGGGTA TATCAGTTGT GGTAATGTCC ATGGCCACCT TTGCAAGCTC
7441 CAGATCCAAT TTCGAGAAGC TCAGAGGACC CGTCATAACT TTGTACGGCG GTCCTAAATA
7501 GGTACGCACT ACAGCTACCT ATTTTGCAGA AGCCGACAGC AAGTATCTAA ACACTAATCA
7561 GCTACAATGG AGTTCATCCC AACCCAAACT TTTTACAATA GGAGGTACCA GCCTCGACCC
7621 TGGACTCCGC GCTCTACTAT CCAAATCATC AGGCCCAGAC CGCGCCCTCA GAGGCAAGCT
7681 GGGCAACTTG CCCAGCTGAT CTCAGCAGTT AATAAACTGA CAATGCGCGC GGTACCCCAA
7741 CAGAAGCCAC GCAGGAATCG GAAGAATAAG AAGCAAAAGC AAAAACAACA GGCGCCACAA
7801 AACAACACAA ATCAAAAGAA GCAGCCACCT AAAAAGAAAC CGGCTCAAAA GAAAAAGAAG
7861 CCGGGCCGCA GAGAGAGGAT GTGCATGAAA ATCGAAAATG ATTGTATTTT CGAAGTCAAG
7921 CACGAAGGTA AGGTAACAGG TTACGCGTGC CTGGTGGGGG ACAAAGTAAT GAAACCAGCA
7981 CATGTAAAGG GGACCATCGA TAACGCGGAC CTGGCCAAAC TGGCCTTTAA GCGGTCATCT
8041 AAGTATGACC TTGAATGCGC GCAGATACCC GTGCACATGA AGTCCGACGC TTCGAAGTTC
8101 ACCCATGAGA AACCGGAGGG GTACTACAAC TGGCACCACG GAGCAGTACA GTACTCAGGA
8161 GGCCGGTTCA CCATCCCTAC AGGTGCTGGC AAACCAGGGG ACAGCGGCAG ACCGATCTTC
8221 GACAACAAGG GACGCGTGGT GGCCATAGTC TTAGGAGGAG CTAATGAAGG AGCCCGTACA
8281 GCCCTCTCGG TGGTGACCTG GAATAAAGAC ATTGTCACTA AAATCACCCC CGAGGGGGCC
8341 GAAGAGTGGA GTCTTGCCAT CCCAGTTATG TGCCTGTTGG CAAACACCAC GTTCCCCTGC
8401 TCCCAGCCCC CTTGCACGCC CTGCTGCTAC GAGAAGGAAC CGGAGGAAAC CCTACGCATG
8461 CTTGAGGACA ACGTCATGAG ACCTGGGTAC TATCAGCTGC TACAAGCATC TTTAACATGT
8521 TCTCCCCACC GCCAGCGACG CAGCACCAAG GACAACTTCA ATGTCTATAA AGCCACAAGA
8581 CCATACTTAG CTCACTGTCC CGACTGTGGA GAAGGGCACT CGTGCCATAG TCCCGTAGCA

Figure 3: The genomic sequence of WT-CHK025 with the color-coding of each genetic element.

The result of phylogenetic analysis based on whole genome sequence using the neighbor-joining method calculated with CLC Sequence Viewer 6.6.1 program showed WT- CHK025 is classified in ECSA genotype, similar to other recent isolates in Thailand (Fig. 4).

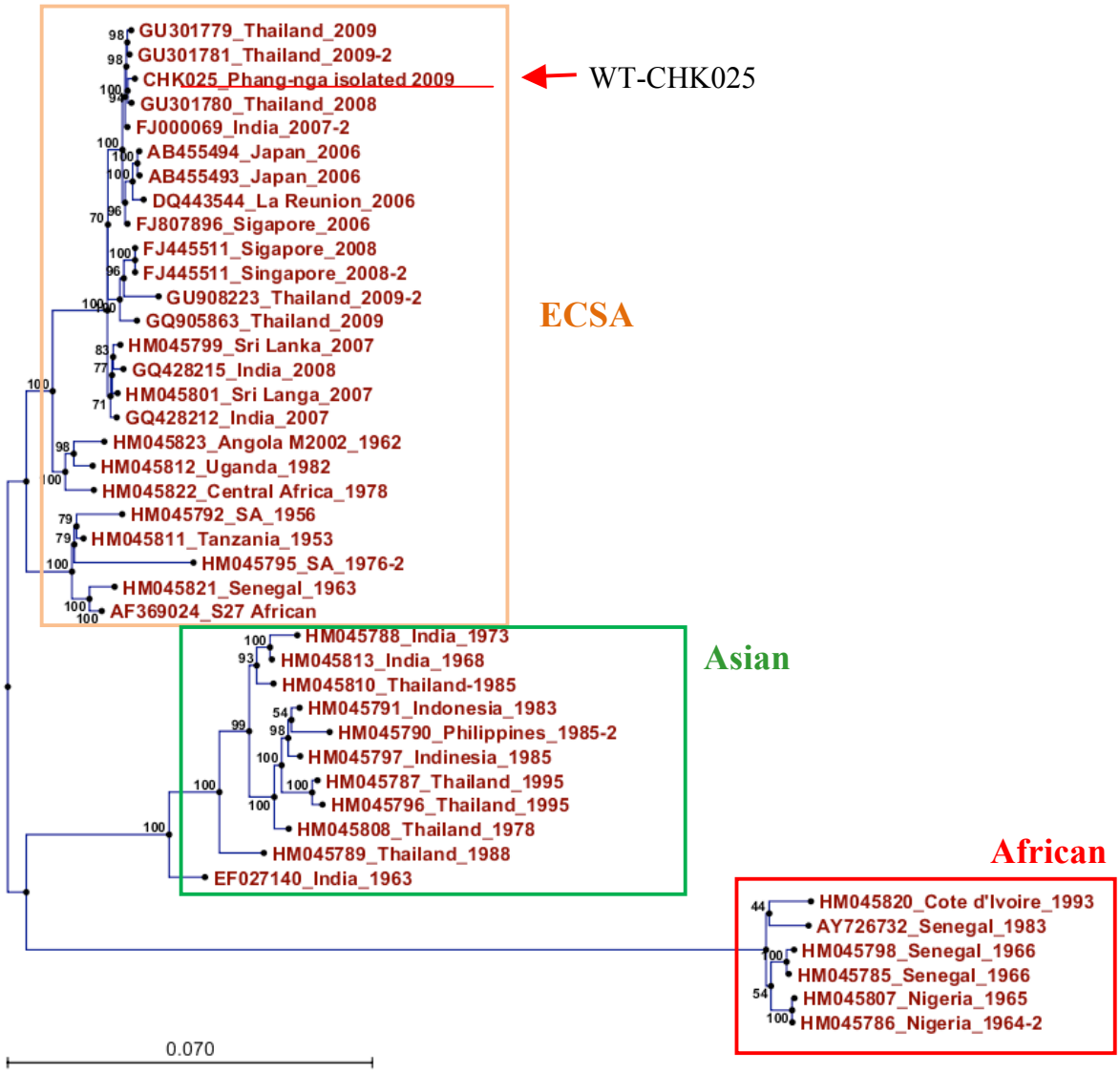

Figure 4: Phylogenetic analysis of WT-CHK025 based on complete genome sequence.

Tissue culture-adapted CHIKVs

To obtain tissue cultured-adapted CHIKV, WT-CHK025 was serially passaged in particular cell line. In addition, the virus was purified using plaque purification technique to get the virus that produces certain plaque morphology.

It was noticed that CHIKV produced from Vero cells showed larger plaque morphology when compared to its previous passage. So WT-CHK025 was subjected to the total of 17 passages alternating with 6 large plaque purifications in Vero cells resulting in CHIKV that produced uniformly large plaques, designated as CHK-L (Fig. 5).

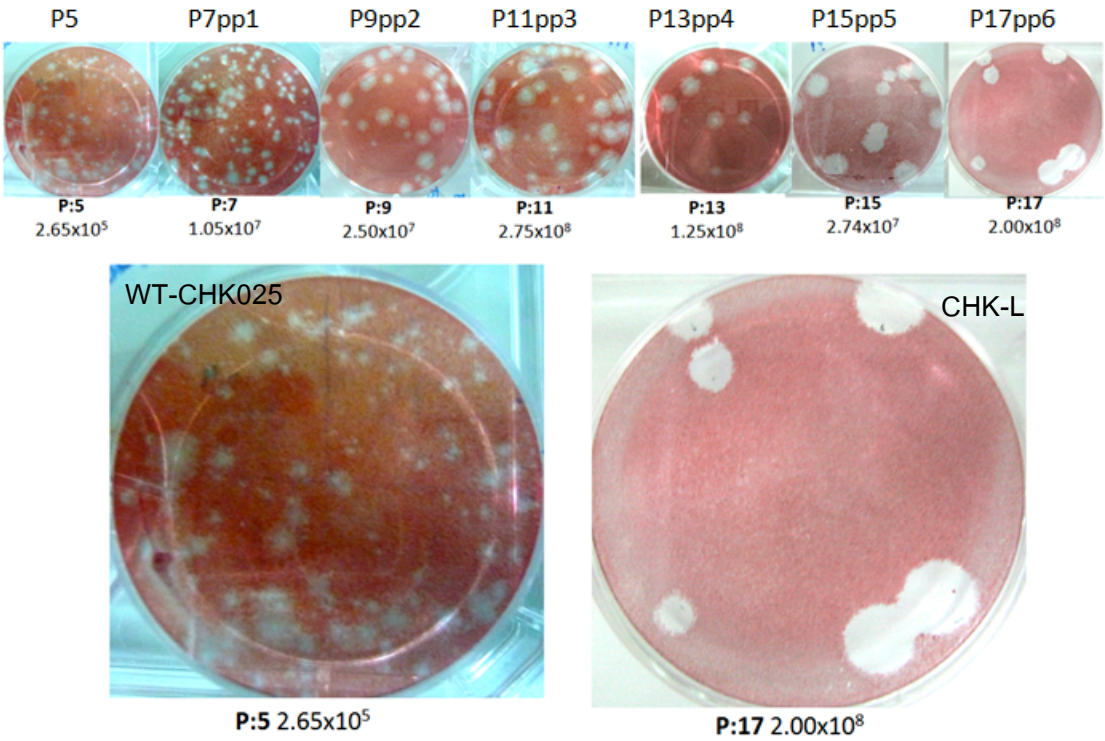

Figure 5: Plaque morphology of Vero-adapted CHIKV. WT-CHK025 was undergone through 17 passages with 6 plaque purifications, resulting in CHK-L. WT-CHK025 and CHK-L were subjected to plaque assay and monitored at 4 dpi. P: passage; pp: plaque purification.

Since it was not possible to maintain the small plaque phenotype of CHIKV in Vero cells, the procedure to obtain the virus began with selecting the cell line that provides optimal condition for small plaque virus propagation. The 4th passage of WT-CHK025 was subjected to standard plaque assay for small plaque purification. After one small plaque was picked and propagated, this virus was used to infect three cell lines, Vero, MRC-5 and C6/36. The results showed that C6/36 cell produced more homogenous small plaques when compared to the others (fig.6).

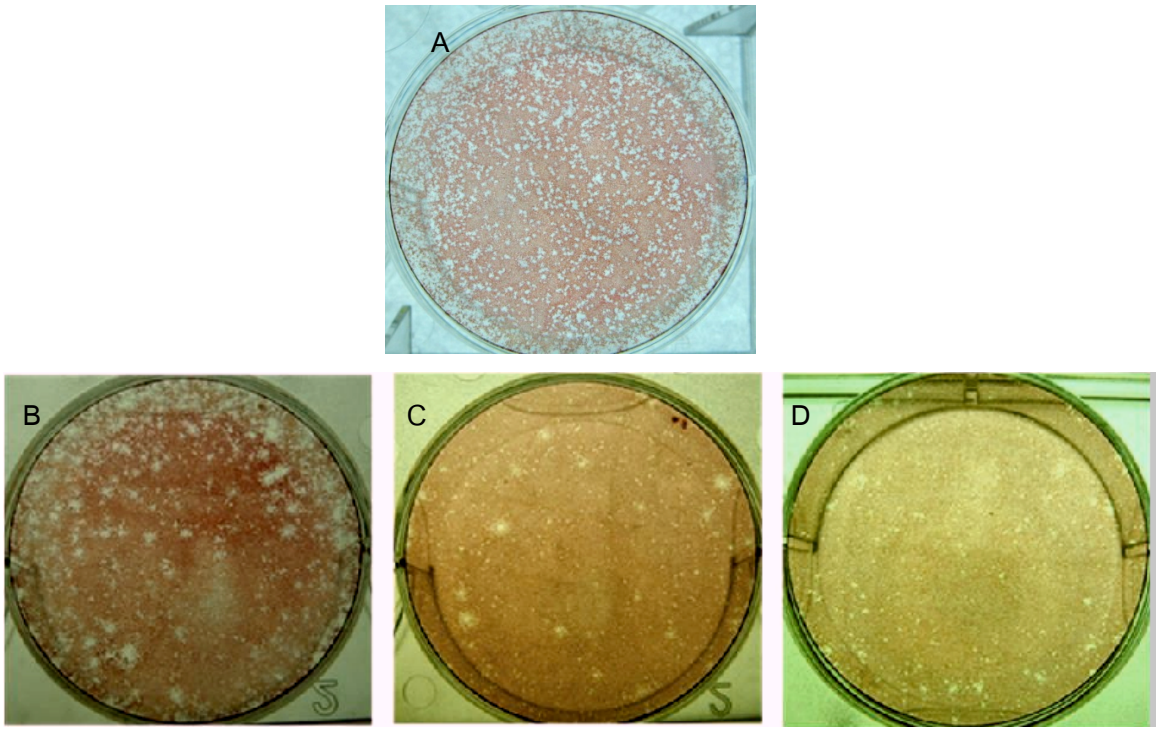

Figure 6: Plaque morphology of CHIKV that was plaque purified for small plaque phenotype (A) then propagated in Vero (B), MRC-5 (C) and C6/36 (D) cells. The viruses were subjected to standard plaque assay and monitored at 7 dpi.

After propagated in C6/36 cells alternating with plaque purified for 3 times to the total of 14 passages, the virus showed the homogeneous small plaque phenotype and was designated as CHK-S (fig.7).

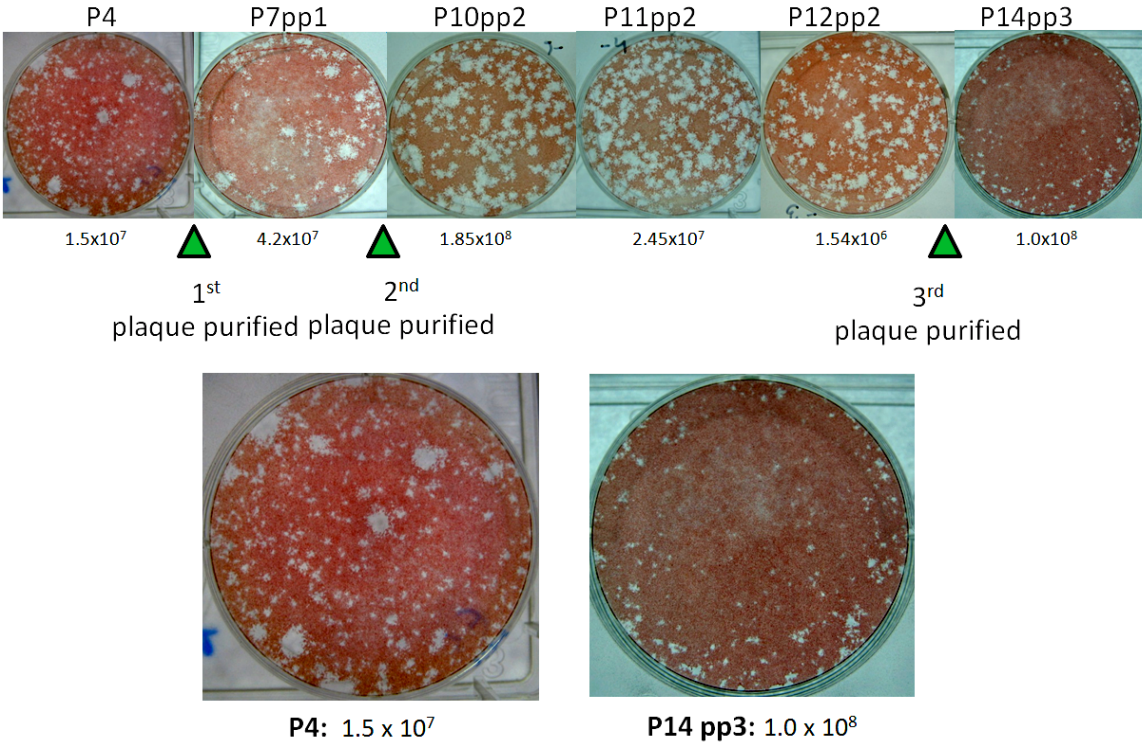

Figure 7: Plaque morphology of C6/36-adapted CHIKV. WT-CHK025 was undergone through 14 passages with 3 plaque purifications, resulting in CHK-S. WT-CHK025 and CHK-S were subjected to plaque assay and monitored at 7 dpi. P: passage; pp: plaque purification.

To compare parental and both tissue cultured-adapted CHIKVs in terms of the plaque morphology, WT-CHK025, CHK-L and CHK-S were subjected to the plaque assay in the same experiment. While WT-CHK025 produced plaques with different sizes, Vero-adapted CHK-L and C6/36 adapted CHK-S showed homogeneously large and small plaques, respectively (fig. 8).

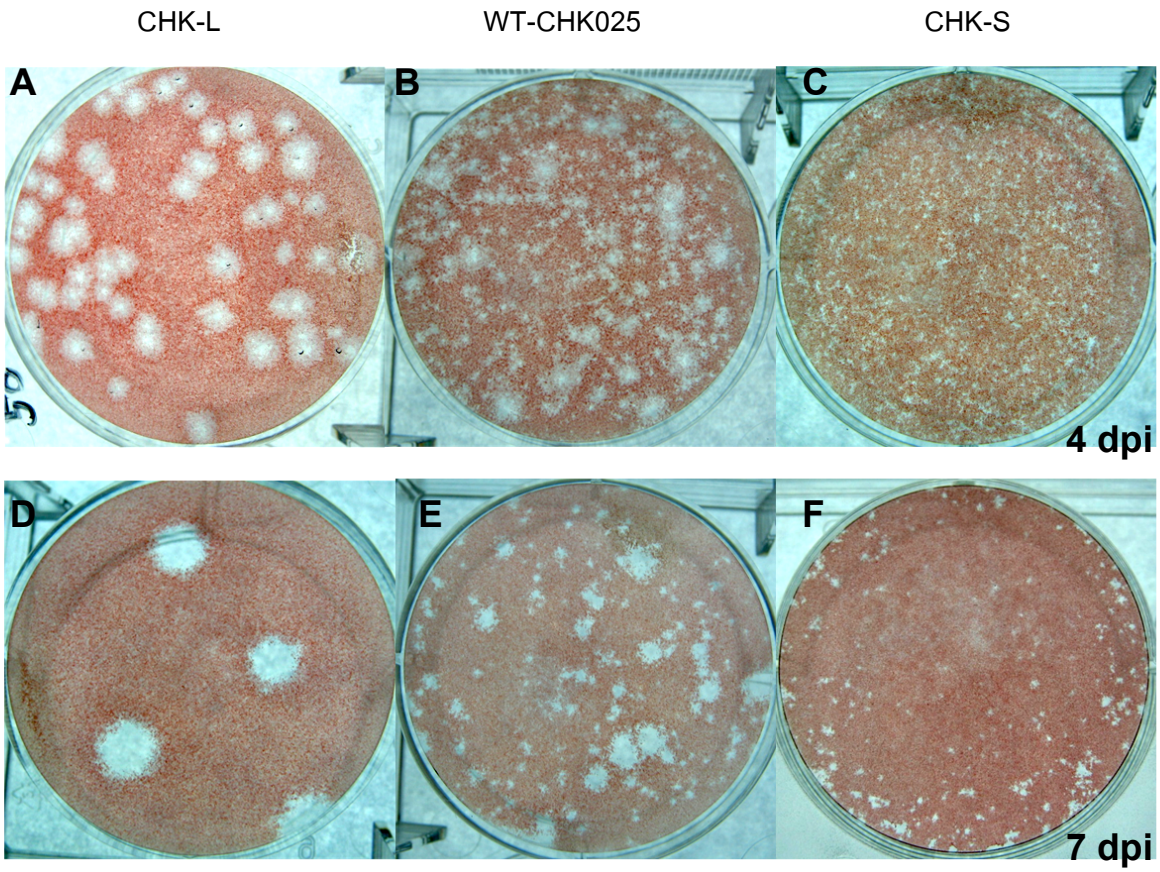

Figure 8: Comparison of plaque morphology of WT-CHK025 (B and E), CHK-L (A and D) and CHK-S (C and F). The viruses were subjected to plaque assay and monitored on day 4 (A, B, C) or day 7 (D, E, F) post-infection.

Phenotypic characterization of tissue culture-adapted CHIKVs

The growth kinetics is one of the distinct phenotypic characters of each virus. However, due to the limited amount of WT-CHK025 virus, only tissue cultured-adapted viruses were used in this experiment. Single-step growth kinetics was performed in C6/36 cells and the viral titers were determined in Vero cells. Although both viruses reached the plateau phase in 18 hpi, CHK-L showed the maximum titer of log10 pfu/ml whereas the maximum titer of CHK-S is approximately 1,000 times lower than that indicating the higher growing efficiency of CHK-L over CHK-S (fig.9).

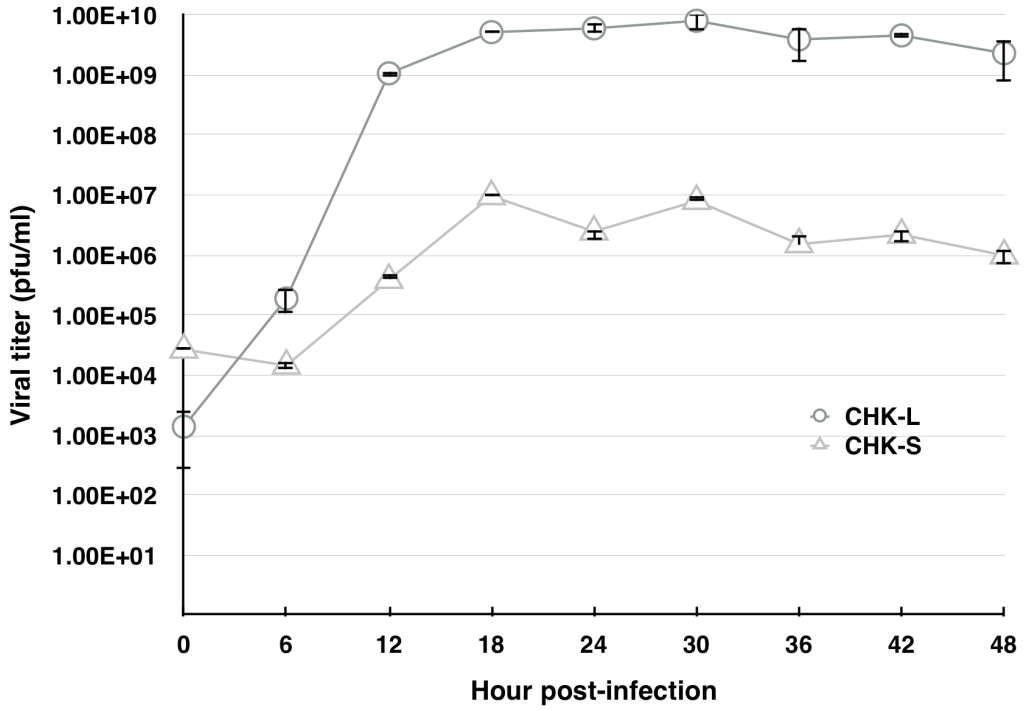

Figure 9: Growth kinetics of CHK-L (opened circle) and CHK-S (opened triangle). C6/36 cells were infected with the viruses at MOI of 5. The supernatants were collected every 6 hours until 48 hpi and subjected to plaque assay in Vero cells in duplicate. The values are means from 2 independent experiments. The sample was done in duplicate for each experiment. The error bars indicate SD.

Next, the pathogenicities of the viruses were compared by mouse-neurovirulence study in suckling mice. Three groups of newborn mouse were intracranially injected with 10^3 pfu of WT-CHK025, CHK-L or CHK-S. The mice were kept for observation for 3 weeks and the numbers of survived mouse were used to plot survival curves (fig.10). Mice that received CHK-L succumbed to the infection from 4 dpi and only 17.8% of them survived by the end of the experiment. On the other hand, only one mouse of the group that infected with CHK-S was dead, making the survival rate of this group to be 98.2%. The group that got WT-CHK025 showed survival rate in between of those 2 groups with 74% of mouse survived at 21 dpi. In addition, the average survival times of the mice were calculated. The mice that infected WT-CHK025, CHK-L and CHK-S showed average survival times of 17.68, 8.18 and 20.82 days, respectively. These results indicate the correlation between CHIKV plaque size and pathogenecity in which the virus that produce large plaque is more pathogenic than the virus with small plaque.

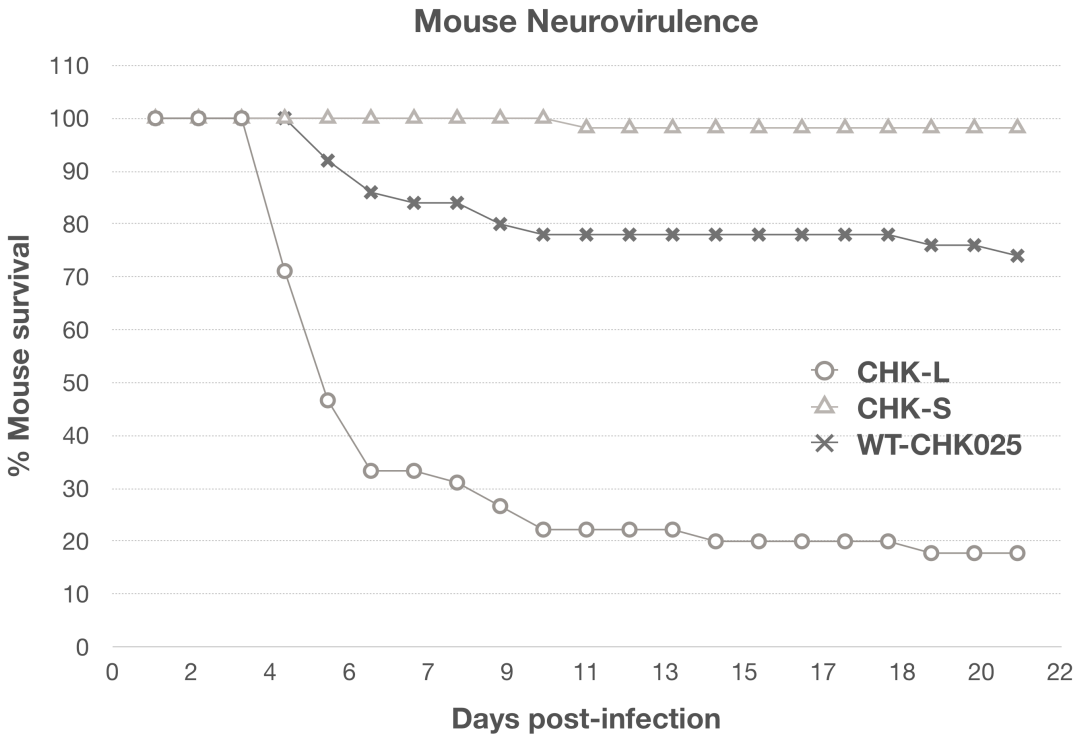

Figure 10: Mouse neurovirulence of CHIKVs. Three groups of suckling mouse were intracranially infected with WT-CHK025 (cross), CHK-L (opened circle) or CHK-S (opened triangle). Mice were observed for survival for 21 days. The graph showed percentage of survived mouse.

Genotypic characterization of tissue culture-adapted CHIKVs

The whole genomes of CHK-L and CHK-S were sequenced as described in the previous section. The resulting sequences were compared to WT-CHK025 sequence.

Seven nucleic acid differences between the genome of WT-CHK025 and CHK-L were found at 5'NTR, nsP1, nsP2, nsP3, E2 and 3'NTR regions, resulting in 5 amino acid differences. To confirm these differences, two additional clones of large plaque purified-Vero adapted CHIKV were sequenced at those particular regions. The result showed 4 consensus mutations, which lead to 4 amino acid changes at nsP1: G230R, nsP2: P618L, nsP3: G117R and E2: N187K (table 3).

Table 3 Differences between WT-CHK025 and 3 clones of CHK-L when complete genomic sequences were compared.

Gene/ position	virus	Nucleotides	Amino acid	Nucleotides and amino acid comparison	WT	L-17	L-17/2	L17/3
nsP1 134	WT	TGC	C		C TGC	S AGC	C TGC	C TGC
	Large	AGC	S					
nsP1 230	WT	GGG	G		G GGG	R AGG	R AGG	R AGG
	Large	AGG	R					
nsP2 618	WT	CCA	P		P CCA	L CTA	L CTA	L CTA
	Large	CTA	L					
nsP3 117	WT	GGG	G		G GGG	R AGG	R AGG	R AGG
	Large	AGG	R					
E2 187	WT	AAC	N		N AAC	K AAA	K AAA	K AAA
	Large	AAA	K					
5' NTR	WT	A	-		A	G	A	A
	Large	G	-					
3' NTR	WT	A	-		A	G	A	A
	Large	G	-					

The nucleotide sequence alignment between WT-CHK025 and CHK-S showed 4 nucleotide differences in nsP1, nsP2 and nsP3, which are consensus in all 3 clones of CHK-S. However, the changes in nsP2:761 (A→G) and nsP3:39 (A→G) resulted in synonymous mutations while one mutation in nsP1 turned out to be the same mutation as CHK-L. Therefore, the mutation that might responsible for the small plaque phenotype remain only at position 546 of nsP2 in which Aspartic acid is changed to Glycine (table 4).

Table 4 Differences between WT-CHK025 and 3 clones of CHK-S when complete genomic sequences were compared.

Gene/ position	virus	Nucleotides	Amino acid	Nucleotides and amino acid comparison	WT	S14	S14/2	S14/3
nsP1 230	WT	GGG	G		G GGG	R AGG	R AGG	R AGG
	Small	AGG	R					
nsP2 546	WT	GAT	D		D GAT	G GGT	G GGT	G GGT
	Small	GGT	G					
nsP2 761	WT	GAA	E		E GAA	E GAG	E GAG	E GAG
	Small	GAG	E					
nsP3 39	WT	AAA	K		K AAA	K AAG	K AAG	K AAG
	Small	AAG	K					

In summary, 3 mutations in nsP2, nsP3 and E2 might be responsible for large plaque phenotype while only single mutation in nsP2 seemed to be determinant of small plaque. The nsP1: G230R seemed to be the tissue culture adapted mutation since this mutation is present in both CHK-L and CHK-S (table 5).

Table 5 Summary of amino acid differences among WT-CHK025, CHK-L and CHK-S.

aa position	Gene	WT-CHK025	CHK-L	CHK-S
230	nsP1	G GGG	R AGG	R AGG
546	nsP2	D GAT	D GAT	G GGT
618	nsP2	P CCA	L CTA	P CCA
117	nsP3	G GGG	R AGG	G GGG
187	E2	N AAC	K AAA	N AAC

สรุปและวิจารณ์ผลการทดลอง

The complete genomic sequence of CHIKV clinical isolate, WT-CHK025, was determined. The length of the genome is 11,811 bp. Phylogenetic analysis revealed this virus is classified in ECSA genotype based on whole genome sequence analysis and amino acid position 226 of E1 protein is Valine, which is similar to other CHIKV isolates from the recent outbreaks in Thailand and neighboring countries. In the previous study, E1:A226V mutation was suggested to enhance viral vector adaptation to *A.albopictus*. This could be a major factor leading to an epidemic in the area that is predominated by *A.albopictus* including the southern part of Thailand.

Through serially passages and plaque purifications, CHIKV was tissue cultured-adapted in mammalian cells, Vero, and mosquito cells, C6/36, resulting in large and small plaque viruses, respectively. Vero cells are capable of adapting the virus to produce large plaque. In addition to our result, other study also showed that Vero cell adapted-enterovirus 71 produced larger plaque size than the original enterovirus 71. In contrast, C6/36 cells showed the ability to stabilize the small plaque morphology of the virus.

The viruses with different plaque morphology also display different degrees of virulent in mouse. Percentages of the survived suckling mouse after infected with WT-CHK025, CHK-S and CHK-L are 74.0%, 98.2% and 15.6%, respectively. CHK-L shows the highest pathogenicity while CHK-S is almost nonpathogenic. Our result also concurs with the general acceptance that the virus with larger plaque size is more pathogenic than the virus with smaller one.

Comparison of entire genomic sequences of WT-CHK025 and CHK-L reveals 3 mutations at nsP2: P618L, nsP3: G117R and E2: N187K. Previous report showed that a point mutation in the protease domain of nsP2 could affect the infectivity and pathogenesis of sindbis virus (SINV), the other virus in the same *Alphavirus* family. The nsP2: P618L mutation is also located in nsP2 protease domain. Hence, this change from proline to leucine might be able to alter the infectivity, resulting in large plaque phenotype.

The nsP3: G117R mutation is located in the macro domain, which is believed to display an ADP-ribose binding capacity. The RNA binding property can be enhanced by the presence of positively charged patches of amino acid at protein surface or outside of the active site pocket of this domain. The mutation increases positive charge of this macro domain, which might lead to an increase in ADP-ribose binding activity and consequently increase viral transcription. Moreover, nsP3 macro domain was shown to be significant in an age-dependent susceptibility of mice to SINV encephalomyelitis. Thereby, this mutation might also involve in the increase of viral pathogenicity.

The viral E2 protein functions as a receptor binding protein. Previous study in chimeric alphaviruses suggested that the mutations in E2 that increase positive charge of the protein can enhance the binding of the protein to heparan sulfate, one of the putative alphavirus receptor, which lead to more efficient viral entry. E2: N187K mutation, which increases positive charge of CHK-L, might also work the same way in enhancing the viral entry and lead to the enlarge plaque.

Only one nonsynonymous mutation of CHK-S was identified, nsP2: D546G. This mutation is located in the nsP2 protease domain, very close to catalytic site of the protein. The nsP2 amino acid residues 546, 547 and 548 of CHIKV (or 544, 545 and 546 of VEEV crystal structure) form a thumb that regulate grooved binding and it is also responsible for substrate specificity of nsP2 protease enzyme. The efficiency of the protease activity of nsP2 has a high impact on the viral replication since the processes in viral RNA replication depend largely on the cleavage of the viral replication complex. Therefore, any change that might interrupt the protease activity will alter the viral replication as well. However, the contribution of these mutations on the viral phenotype has to be confirmed by reverse genetics.