GENETIC ANALYSIS OF SAFFOLD VIRUS-PENANG IN RELATION TO OTHER NEWLY DISCOVERED SAFFOLD VIRUSES

Kenny Voon[1], Qi Mei Ng[2], Meng Yu[3], Lin-Fa Wang[3] and Kaw Bing Chua[2]

[1]International Medical University, Kuala Lumpur, Malaysia;
[2]Temasek Life Science Laboratory, Temasek, Singapore; [3]CSIRO Livestock Industries, Australian Animal Health Laboratory, Geelong, Victoria, Australia

Abstract. Viruses in the family Picornaviridae are classified into nine genera. Within the family Picornaviridae, two species: *Encephalomyocarditis virus* and *Theilovirus*, are listed under the genus *Cardiovirus*. A novel *Theilovirus*, Saffold virus (SAFV), was first reported in 2007. Since then, numerous SAFV isolates have been detected around the world and genetic recombinations have been reported among them. In 2009, SAFV-Penang was isolated from a febrile child with influenza-like illness in Malaysia. SAFV-Penang is a genotype 3 SAFV. In this study we investigated the genome features of SAFV-Penang to exclude the possibility it is a recombinant variant. SAFV-Penang was found not to be a recombinant variant but to have three unique non-synonymous substitutions, alanine [A^{689}], lysine [K^{708}] and isoleucine [I^{724}] in the VP1 protein.

Keywords: Saffold virus, SAFV-Penang, genetic analysis

INTRODUCTION

The Picornaviridae family of viruses encompasses nine genera: *Enterovirus*, *Hepatovirus*, *Rhinovirus*, *Kobuviru*s and *Parechovirus*, which are human pathogens, and *Aphtovirus*, *Erbovirus*, *Teschovirus* and *Cardiovirus*, which are other animal pathogens (King *et al*, 2000). *Cardiovirus* consists of two species: *Encephalomyocarditis virus* (EMCV) and *Theilovirus* (ThV). Both species can infect animals ranging from rodents to macaques, and may infect humans (Lipton *et al*, 1983; Craighead *et al*, 1990; Martin *et al*, 2000; Ohsawa *et al*, 2003). A novel cardiovirus infecting human, named Saffold virus (SAFV) was reported in 2007. This virus was isolated from a child with fever of unknown origin in 1981 (Jones *et al*, 2007). Since then, numerous SAFVs have been detected from human clinical specimens around the world (Jones *et al*, 2007; Abed and Boivin, 2008; Blinkova *et al*, 2009; Zoll *et al*, 2009; Chiu *et al*, 2010; Ren *et al*, 2010). With the increasing number of molecular detection methods and virus isolation, Drexler *et al* (2008) proposed classifying these viruses into different genetic lineages based on VP1 protein. There are eight genetic lineages found circulating in the world (Blinkova *et al*, 2009). Recently, SAFV Nig239 was proposed as a new genotype 9 SAFV (Blinkova *et al*, unpublished data). In Malaysia, SAFV-Penang (the name re-

Correspondence: Kenny Voon, International Medical University, Jalan Jalil Perkasa 19, Bukit Jalil, 57000, Kuala Lumpur, Malaysia.
Tel: +603 27317245; Fax: +603 86567229
E-mail: kenny_voon@imu.edu.my

flects the locality of isolation) was isolated from a child in 2009 (Chua *et al*, 2011).

For picornaviruses, inter- and intra-species recombinations are often reported as the mechanism of virus evolution and antigenic variability due to the nature of their single-stranded positive RNA genomes (Liang *et al*, 2008; Blinkova *et al*, 2009). Recombination events have also been reported among cardiovirus, both SAFV and ThV (Liang *et al*, 2008; Blinkova *et al*, 2009). Drexler *et al* (2010) reported the SAFV D/VI2229/2004 isolate (Germany) is a recombinant closely related to BR/118/2006 (Brazil) and UC1 (USA). The ThV BeAn 8386 isolate is a recombinant virus related to ThV TO Yale and DA strain (Liang *et al*, 2008). In this study, the genomic features of SAFV-Penang were analyzed with respect to other SAFVs in order to investigate whether SAFV-Penang is a recombinant virus or a new virus isolate.

MATERIALS AND METHODS

The complete genomic sequence of SAFV-Penang has been documented previously and is deposited in GenBank under accession no. HQ162476 (Chua *et al*, 2011). Complete genome sequences of other SAFV, ThV and EMCV viruses were obtained from GenBank and are tabulated in Table 1. The sequences were screened to exclude patented and artificial mutants prior to analysis and alignment using MUSCLE and Clustal W. Phylogenetic analysis was conducted using MEGA4 (Tamura *et al*, 2007). The sequences were edited and managed using BioEdit 7.0.05 and Jalview 2.4 (Hall, 1999; Waterhouse *et al*, 2009). Detection of potential recombination events, identification of potential parental sequences and localization of potential recombination break points were determined using the Recombination Detection Program (RDP), GENECONV, BOOTSCAN, MaxChi, CHIMERA, and SISCAN methods embedded in RDP4 (Martin *et al*, 2005).

RESULTS

SAFV-Penang is a genotype 3 SAFV (Chua *et al*, 2011). Since the genetic mapping of the putative proteins has not been documented previously, the respective length of putative proteins and their deduced cleavage sites for SAFV-Penang were carried out and are shown in Table 2.

The L peptide, L* peptide, CD loop of VP1 and EF loop of VP2 are reported to play important roles in viral persistence and pathogenesis in cardioviruses. The L protein has been shown to block beta interferon gene transcription (Stavrou *et al*, 2010). Previous findings showed the Ser/Thr phosphorylation domain of L protein of SAFV is shorter than the respective proteins of animal cardioviruses, such as Theiler's murine encephalomyelitis virus (TMEV) and the Theiler-like virus.

The L* peptide, on the other hand, is a 156 amino acid (aa) peptide which is encoded by an alternative start codon 13 nt downstream from the cognate polyprotein AUG of the TMEV. The putative L* protein of SAFV-1 is translated with the start codon ACG instead of the classical AUG observed with TMEV (Jones *et al*, 2007). The putative L* peptide of SAFV-1 is shorter than the TMEV by 99 aa. The L* peptides for SafV-2 -3, -5, and -6 were subsequently identified and are shorter than the TMEV by 123 aa (Drexler *et al*, 2008; Blinkova *et al*, 2009; Zoll *et al*, 2009). Similar findings have also been noted for the L and L* proteins for SAFV- Penang (Fig 1A and 1B).

Mutations in the CD and EF loops

Table 1
Accession number for sequences of cardioviruses used in this study.

Strain	Name	Country	Source	Accession No.
Penang	SAFV	Malaysia	Human	HQ162476
-	SAFV	USA	Human	EF165067
BCH1031	SAFV	China	Human	GU943513
BCH115	SAFV	China	Human	GU943514
BCHU79	SAFV	China	Human	GU943518
Nijmegen2007	SAFV	Netherlands	Human	FM207487
Nijmegen2008	SAFV	Netherlands	Human	FN999911
Can112051-06	SAFV	Canada	Human	AM922293
BR/118/2006	SAFV	Germany	Human	EU681177
D/VI2229/2004	SAFV	Germany	Human	EU681176
D/VI2223/2004	SAFV	Germany	Human	EU681179
D/VI2273/2004	SAFV	Germany	Human	EU681178
NL2005-1035	SAFV	Netherlands	Human	HM181999
NL1999-590	SAFV	Netherlands	Human	HM181996
NL2007-2686	SAFV	Netherlands	Human	HM181997
NL2007-2690	SAFV	Netherlands	Human	HM181998
Fin2008	SAFV	Finland	Human	FR682076
Pak5003	SAFV	Pakistan	Human	FJ463615
Pak5152	SAFV	Pakistan	Human	FJ463616
Pak6572	SAFV	Pakistan	Human	FJ463617
Pak5842*	SAFV	Pakistan	Human	FJ463606
Pak12*	SAFV	Pakistan	Human	FJ463600
Pak1411*	SAFV	Pakistan	Human	FJ463604
Pak962*	SAFV	Pakistan	Human	FJ463603
Afg1449*	SAFV	Afghanistan	Human	FJ463602
Nig329*	SAFV	Nigeria	Human	FJ997532
CMH045/07*	SAFV	Thailand	Human	HQ668172
CMH038/07*	SAFV	Thailand	Human	HQ668171
CMH143/07*	SAFV	Thailand	Human	HQ668173
CMH023/07*	SAFV	Thailand	Human	HQ668170
JPN084/2004	SAFV	Japan	Human	HQ902242
UC6	SAFV	USA	Human	GU595289
UC1	SAFV	USA	Human	EU376394
-	VHEV	Siberia	Human	M94868
GDVII	TMEV	UK	Mouse	X56019
DA	TMEV	USA	Mouse	M20301
NGS910	TRV	Japan	Rat	AB090161
-	TRV	USA	Rat	EU542581
PV2	EMCV	-	Mouse	X87335

*Partial sequences for VP1
SAFV, Saffold virus; VHEV, Viliuisk's human encephalomyelitis virus; TMEV, Theiler's murine encephalomyelitis virus; TRV; Theiler' rat virus; EMCV, Encephalomyocarditis virus

Table 2
SAFV-Penang genome and the corresponding translated proteins.

Genome section	Nucleotide position	Deduced translated protein	Length (aa)	Cleavage site *
5′ Untranslated region (UTR)	1-1064			
	1065-1277	L	71	*L*/*VP4* **Q/G**
P1	1278-1493	VP4	72	*VP4*/*VP2* **M/D**
	1494-2306	VP2	271	*VP2*/*VP3* **Q/S**
	2307-3002	VP3	232	*VP3*/*VP1* **Q/G**
	3003-3818	VP1	272	*VP1*/*2A* **N/P**
P2	3819-4241	2A	141	*2A*/*2B* **Q/G**
	4242-4622	2B	127	*2B*/*2C* **Q/G**
	4623-5597	2C	325	*2C*/*3A* **Q/S**
P3	5598-5855	3A	86	*3A*/*3B* **Q/A**
	5856-5915	3B	20	*3B*/*3C* **Q/G**
	5916-6566	3C	217	*3C*/*3D* **Q/G**
	6567-7949	3D	461	
3′ UTR	7950-8076			

[a] Amino acid cleavage site (bold) between each protein (italics). aa, amino acids

have been shown to alter the tropism and sialic acid binding capacity of TMEV, which further alters the potential for viral persistence and neurovirulence (Jnaoui *et al*, 2002). As for SAFV, these loops are unique and conserved for each genotype even though there are regions of great divergence (Fig 1C and 1D). SAFV- Penang has been found to have a diverse sequence from other isolates of human cardioviruses (Fig 1C and 1D) with aa similarity ranging from 45% to 55% (Blinkova *et al*, 2009). The conserved pattern sequences of CD and EF loops observed in SAFV-Penang along with other SAFV-3 strains further strengthen SAFV-Penang in genotype 3.

Cardioviruses are known to undergo recombination involving part of their genome to produce new virus variants (Drexler *et al*, 2010). Recombination has been shown to happen within and between genotypes. An example of the former is SAFV-2 D/V12229/2004 (Germany), which is most likely a recombinant of UC1 (United States) and Br/118/2006 (Brazil) (Drexler *et al*, 2010). Likewise, SAFV-5 Pak 5152 (Pakistan) is most likely a recombinant of SAFV-5 Pak 5003 and SAFV-6 Pak 6572 (Blinkova *et al*, 2009). In this analysis, no recombination event was found with SAFV-3 Penang.

Besides recombination, the cardiovirus genome has a high substitution rate due to its RNA nature. Both non-synonymous and synonymous substitutions were more common in the P1 genome region than the P2 and P3 regions of SAFV, irrespective of their genotypes (Drexler *et al*, 2010). Paired comparisons of the VP1 proteins within SAFV-3 showed unique non-synonymous substitution sites in the SAFV-Penang virus (Fig 2). Alanine [A^{689}], lysine [K^{708}] and isoleucine [I^{724}] were found only in SAFV-Penang while

Table 3
Relative synonymous codon usage (RSCU) among SAFV-1, -2 and -3 viruses.

Codon	RSCU			Codon	RSCU		
	SAFV-1	SAFV-2	SAFV-3		SAFV-1	SAFV-2	SAFV-3
UUU(F)	1.22	1.26	1.22	CAA(Q)	1.40	1.38	1.44
UUC(F)	0.78	0.74	0.78	CAG(Q)	0.60	0.62	0.56
UUA(L)	0.86	0.91	0.86	AAU(N)	1.08	1.10	1.16
UUG(L)	1.71	1.70	1.71	AAC(N)	0.92	0.9	0.84
CUU(L)	1.32	1.27	1.32	AAA(K)	1.26	1.26	1.26
CUC(L)	0.89	0.83	0.89	AAG(K)	0.74	0.74	0.74
CUA(L)	0.64	0.58	0.64	GAU(D)	0.99	1.00	0.96
CUG(L)	0.58	0.7	0.58	GAC(D)	1.01	1.00	1.04
AUU(I)	1.72	1.54	1.72	GAA(E)	1.59	1.54	1.52
AUC(I)	0.55	0.71	0.55	GAG(E)	0.41	0.46	0.48
AUA(I)	0.73	0.75	0.73	UGU(C)	1.14	1.46	1.42
AUG(M)	1.00	1.00	1.00	UGC(C)	0.86	0.54	0.58
GUU(V)	1.67	1.58	1.67	UGG(W)	1.00	1.00	1.00
GUC(V)	0.69	0.7	0.69	CGU(R)	0.38	0.67	0.64
GUA(V)	0.78	0.76	0.78	CGC(R)	0.51	0.48	0.43
GUG(V)	0.87	0.96	0.87	CGA(R)	0.51	0.39	0.48
UCU(S)	2.20	2.06	1.96	CGG(R)	0.19	0.21	0.26
UCC(S)	1.22	1.07	1.14	AGA(R)	3.54	3.44	3.56
UCA(S)	1.40	1.63	1.77	AGG(R)	0.88	0.81	0.63
UCG(S)	0.45	0.43	0.31	GGU(G)	1.38	1.33	1.12
AGU(S)	0.42	0.47	0.57	GGC(G)	0.78	0.99	0.98
AGC(S)	0.31	0.34	0.25	GGA(G)	1.59	1.34	1.55
CCU(P)	1.60	1.44	1.47	GGG(G)	0.25	0.34	0.35
CCC(P)	0.58	0.92	0.91	GCU(A)	1.55	1.61	1.62
CCA(P)	1.45	1.28	1.18	GCC(A)	0.80	0.76	0.82
CCG(P)	0.37	0.36	0.44	GCA(A)	1.39	1.31	1.25
ACU(T)	1.58	1.60	1.68	GCG(A)	0.26	0.32	0.31
ACC(T)	0.71	0.90	0.84	UAU(Y)	1.01	0.89	0.97
ACA(T)	1.47	1.22	1.28	UAC(Y)	0.99	1.11	1.03
ACG(T)	0.24	0.28	0.20	UAA(*)	0.00	0.00	0.00
CAU(H)	0.87	1.00	0.89	UAG(*)	0.00	0.38	0.00
CAC(H)	1.13	1.00	1.11	UGA(*)	3.00	2.63	3.00

serine [S], glutamic acid [E] and valine [V] were in the other SAFV-3 strains, respectively. Whether these substitutions suggest tropism or play a role in virulence or persistence of the SAFV-Penang virus remains to be investigated. Substitution of serine with alanine was also observed in VP1 proteins of SAFV-2 strains, suggesting these amino acids may play a role in virulence or persistence.

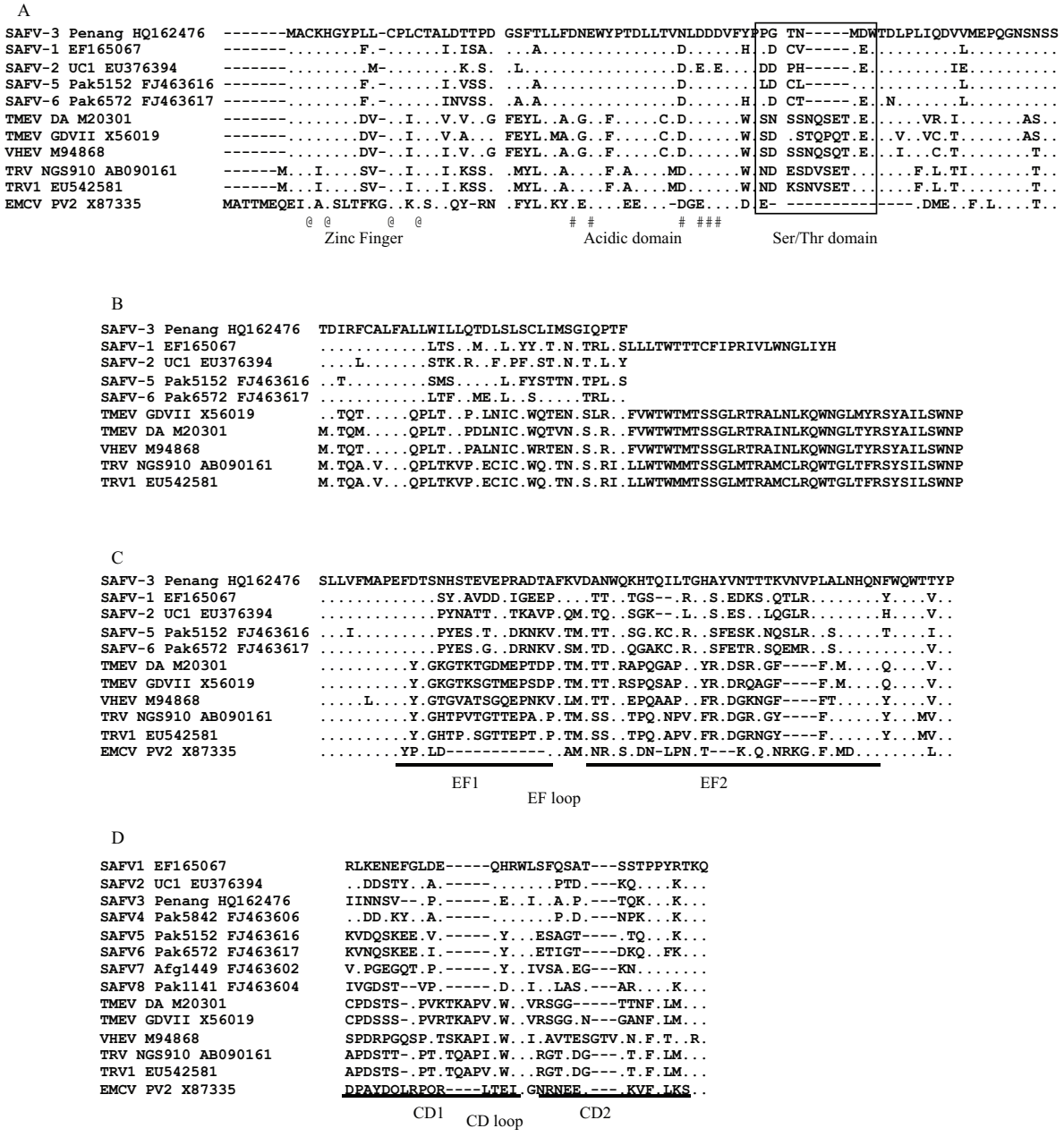

Fig 1–Alignment of SAFV-Penang proteins with homologues from other cardioviruses. (A) L proteins showing different domains with conserved amino acid motifs (label under alignment). (B) The putative L* protein in SAFV-3 Penang is only 34 amino acids in length. The SAFV -2, SAFV-5 and SAFV-6 viruses have the same length but SAFV-1 does not. (C) EF loops in VP2 protein of the SAFV-Penang virus. (D) CD loops in VP1 protein of the SAFV-Penang virus. Amino acids identical to the SAFV-Penang virus are dotted. Selected strains of different genotypes are shown in comparison.

```
                          689                 708                 724
                           *                   *                   *
Penang HQ162476           TLASVLHSTSDVSSKFTPSTAKNLQNSILLTPLPSDIINN
BCH103 GU943513           ..S.................E.................V..
BCH115 GU943514           ..S..........R......E.................V..
BCH350_GU126465           ..S.................E.................V.S
UC5 EU604746              ..S.................E.................V..
UC2 EU604745              ..S.................E.................V..
Nijmegen2007 FM207487     ..S...Q.............................V..
D/VI2223/2004 EU681179    ..S.................E.................V..
D/VI2273/2004 EU681178    ..S.................E.................V..
NL2005-1035 HM181999      ..S...R.............E.................V..
NL2007-2686 HM181997      ..S.................E.................V..
NL1999-590  HM181996      ..S.................E.................V..
NL2007-2690 HM181998      ..S.................E.................V..
JPN084/2004 HQ902242      ..S.................E.................V..
Pak2678 FJ463605          ..S.................SE................V..
Nig266 FJ997531           ..S..........F......VD................V..
07-Aichi10345 AB542807    ..S.................E.................V..
07-Aichi10247 AB542806    ..S.................V.................V..
```

Fig 2–Partial alignment of the VP1 protein gene in SAFV-Penang with other SAFV-3 strains. Unique non-synonymous substitutions in the SAFV-Penang virus are highlighted with a*. The positions of the amino acids are numbered referring to the SAFV-Penang polyprotein AEM00022 deposited in GenBank. Alanine [A^{689}], lysine [K^{708}] and isoleucine [I^{724}] were observed only in the SAFV-Penang virus while serine [S], glutamic acid [E] and valine [V] were in the other SAFV-3 strains, respectively. Identical amino acids to the SAFV-Penang virus are dotted.

DISCUSSION

Isolation and adaptation of novel viruses in cultured cells is an important prerequisite for study of these viruses with respect to their genetic and biological characteristics, sero-epidemiology and pathogenesis. SAFV-Penang was first isolated in Hep2 cells, but not Vero or MDCK cells. However, subsequent passages showed it was able to adapt better in Vero cells with higher titers and showing cytophatic effects (CPE) in shorter periods with the culture (Chua *et al*, 2011). Similarly, SAFV-3 NL2007 was reported to show full CPE in cell culture within two to three days post-inoculation (Chiu *et al*, 2010). Unlike SAFV type 3, other genotypes replicate either poorly in culture cells or not at all. A well adapted SAFV-2 UC6 grew poorly in rhesus monkey kidney LLC-MK2 cells and only produced moderate CPE with lower viral titers after a prolonged period in culture (Zoll *et al*, 2009). SAFV-1 has been reported to grow well in human fetal diploid kidney cell culture but poorly in human fetal diploid lung cell culture (Jones *et al*, 2007). SafV-4, -5, -6, -7 and -8 have yet to be isolated and grown in cell culture. SAFV-3 is more easily cultured, indicating distinct phenotypic characteristic features of SAFV-3 genotype in comparison to other genotypes of SAFV.

The ability of SAFV-3 to grow in cell culture enables increased accessibility to study the sero-epidemiology of SAFV. Zoll *et al* (2009) postulated each genotype of SAFV had distinct serological types. This notion is based on the neutralization patterns of human enteroviruses delineated into various serotypes with respect to the identities of their VP1 protein genes in which serotype boundaries are present at about 25% at the nucleotide level and 12% at the amino acid level (Oberste *et al*, 1999; Liang *et al*, 2008). Based on the published significance of CD and EF loops in the VP1 protein among enteroviruses in determining the receptor tropism of the virus, the unique sequences observed in the CD and EF loops of the SAFV-3 Penang virus compared with the other strains, shows these may have contributed to its ability to replicate well in cell culture.

Since only a handful of SAFVs are able to be cultured in cells, data regarding sero-epidemiology are limited. To date,

serological data have been generated based on SAFV-3. The prevalence rate of infection with SAFV-3 NL2007 and SAFV-3 Penang strains in a population of healthy children has been reported to be >67% among 12 year old children (Zoll *et al*, 2009; Chua *et al*, 2011). Positive antibodies against SAFV-3 NL2007 were found among children as young as nine months old and the serological data corroborated with PCR detection of SAFV, which supports that children are exposed to SAFV early in life (Zoll *et al*, 2009).

While only one genotype of SAFV was isolated in Malaysia, molecular detection of more than one genotype of SAFV has been reported in a neighboring country, Thailand (Khamrin *et al*, 2011). Two strains, CMH023/07 and CMH143/07, are closely related to SAFV-1 and another two strains, CMH038/07 and CMH045/07, are more closely related to SAFV-2 (Khamrin *et al*, 2011). Blinkova *et al* (2009) reported recombination may occur between two genotypes of SAFV in the same country. It is possible for recombinant variants of SAFV to circulate in Malaysia and Thailand. Further surveillance studies may provide more insights into the genotypes and variants of SAFVs in Malaysia.

Other than recombination, SAFVs were found to have a low level of bias in codon usage and GC content of the genome, which is correlated to mutation and selection pressure (Liu *et al*, 2011). Liu *et al* (2011) suggested the serotype may be another factor affecting codon usage bias. We tabulated the relative synonymous codon usage (RSCU) among SAFV-1, -2 and -3. SAFV-5 (Pak5003 and Pak5152) and -6 (Pak6572) were not included, since the C- terminals of the polypeptides were truncated. The RSCU simply measures the non-uniform usage of synonymous codons in a coding sequence (Sharp *et al*, 1986). There is uneven distribution of bias in codon usage among SAFV-1, -2 and -3 (Table 3). The differences in codon use suggest the existence of a different SAFV genotype is a potential factor influencing codon usage. However, without concrete evidence to support the correlation between genotype and serotype, it is too early to establish bias in synonymous codon usage occurs among the genotypes or serotypes in SAFVs.

SAFV, a member of *Theilovirus*, shares a similar genome structure and gene organization with TMEV. Downstream to the 5′ UTR are the L, P1, P2 and P3 coding regions with 3′ UTR at the end. The L and L* proteins in TMEV inhibit interferon production in infected cells and prevents apoptosis, respectively (Chen *et al*, 1995; Michiels *et al*, 1995; van Eyll and Michiels, 2002). SAFVs share 77% and 24% similarity in L and L* proteins, respectively, with TMEV. The L and L* proteins in SAFVs are shorter (99 or 123 aa, depending on genotype) than in TMEV and utilize the ACG threonine start codon instead of the AUG methionine start cordon (Fig 1B). Whether the L* proteins are functionally similar in the TMEV remains to be determined.

The genome of the SAFV-3 Penang was fully mapped in this study. SAFV-3 Penang is not a recombinant variant since no recombination event was detected. The SAFV3- Penang has three unique non-synonymous substitution sites, which are alanine (A^{689}), lysine (K^{708}) and isoleucine (I^{724}) in the VP1 proteins.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Liew Siaw Cheok and Lim Kim Chooi for proof reading this manuscript and to the International Medical University for supporting this project.

REFERENCES

Abed Y, Boivin G. New Saffold cardioviruses in 3 children, Canada. *Emerg Infect Dis* 2008; 14: 834-6.

Blinkova O, Kapoor A, Victoria J, *et al.* Cardioviruses are genetically diverse and cause common enteric infections in South Asian children. *J Virol* 2009; 83: 4631-41.

Chen HH, Kong WP, Zhang, L, Ward PL, Roos RP. A picornaviral protein synthesized out of frame with the polyprotein plays a key role in a virus-induced immune-mediated demyelinating disease. *Nature Med* 1995; 1: 927-31.

Chiu CY, Greninger AL, Chen EC, *et al.* Cultivation and serological characterization of a human Theiler's-like cardiovirus associated with diarrheal disease. *J Virol* 2010; 84: 4407-14.

Chiu CY, Greninger AL, Kanada K, *et al.* Identification of cardioviruses related to Theiler's murine encephalomyelitis virus in human infections. *Proc Nat Acad Sci USA* 2008; 14124-9.

Chua KB, Voon K, Yu M, Ali WN, Kasri AR, Wang LF. Saffold virus infection in children, Malaysia, 2009. *Emerg Infect Dis* 2011; 17: 1562-4.

Craighead JE, Huber SA, Haynes MK. Diverse patterns of immune and non-immune-mediated disease in EMC M-variant-infected mice. *J Autoimmun* 1990; 3 (suppl 1): 27-9.

Drexler JF, Luna LK, Stocker A, *et al.* Circulation of 3 lineages of a novel Saffold cardiovirus in humans. *Emerg Infect Dis* 2008; 14: 1398-405.

Drexler JF, Baumgarte S, Luna LK, *et al*. Genomic features and evolutionary constraints in Saffold-like cardioviruses. *J Gen Virol* 2010; 91: 1418-27.

Hall TA. BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp Ser* 1999; 4: 95-8.

Jnaoui K, Minet M, Michiels T. Mutations that affect the tropism of DA and GDVII strains of Theiler's virus in vitro influence sialic acid binding and pathogenicity. *J Virol* 2002; 76: 8138-47.

Jones MS, Lukashov VV, Ganac RD, Schnurr DP. Discovery of a novel human Picornavirus in a stool sample from a pediatric patient presenting with fever of unknown origin. *J Clin Microbiol* 2007; 45: 2144-50.

Khamrin P, Chaimongkol N, Nantachit N, Okitsu S, Ushijima H, Maneekarn, N. Saffold cardioviruses in children with diarrhea, Thailand. *Emerg Infect Dis* 2011; 17: 1150-2.

King AMQ, Brown F, Christian P, *et al.* Family picornaviridae. In: van Regenmortal MHV, Fauquet CM, Bishop DHL, *et al*, eds. Virus taxonomy classification and nomenclature of viruses: seventh report of the International Committee on Taxonomy of Viruses. San Diego: Academic Press, 2000: 657-78.

Liang Z, Kumar AS, Jones MS, Knowles NJ, Lipton HL. Phylogenetic analysis of the species *Theilovirus*: Emerging murine and human pathogens. *J Virol* 2008; 82: 11545-54.

Lipton HL, Friedmann A, Sethi P, Crowther JR. Characterization of *Vilyuisk virus* as a Picornavirus. *J Med Virol* 1983; 12: 195-203.

Liu WQ, Zhang J, Zhang YQ, *et al*. Compare the differences of synonymous codon usage between the two species within *Cardiovirus*. *Virol J* 2011; 8: 325.

Martin DP, Williamson C, Posada D RDP2: Recombination detection and analysis from sequence alignments. *Bioinformatics* 2005; 21: 260-2.

Martin LR, Neal ZC, McBride MS, Palmenberg AC. Mengovirus and Encephalomyocarditis virus poly(C) tract lengths can affect virus growth in murine cell culture. *J Virol* 2000; 74: 3074-81.

Michiels T, Jarousse N, and Brahic M. Analysis of the leader and capsid coding regions of persistent and neurovirulent strains of Theiler's virus. *Virology* 1995; 214: 550-8.

Oberste MS, Maher K, Kilpatrick DR, Pallansch MA. Molecular evolution of the human Enteroviruses: correlation of serotype with VP1 sequence and application to picornavirus classification. *J Virol* 1999; 73: 1941-8.

Ohsawa K, Watanabe Y, Miyata H, Sato H. Genetic analysis of a Theiler-like virus isolated from rats. *Compar Med* 2003; 53: 191-6.

Ren L, Gonzalez R, Xie Z, *et al*. Saffold cardioviruses of 3 lineages in children with respiratory tract infections, Beijing, China. *Emerg Infect Dis* 2010; 16: 1158-61.

Sharp PM, Tuohy TM, Mosurski KR. Codon usage in yeast: Cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res* 1986; 14: 5125-43.

Stavrou S, Baida G, Viktorova E, Ghadge G, Agol VI, Roos RP. Theiler's murine encephalomyelitis virus L* amino acid position 93 is important for virus persistence and virus-induced demyelination. *J Virol* 2010; 84: 1348-54.

Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evolut* 2007; 24: 1596-9.

van Eyll O, Michiels T. Non-AUG-initiated internal translation of the L* protein of Theiler's virus and importance of this protein for viral persistence. *J Virol* 2002; 76: 10665-73.

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 2009; 25: 1189-91.

Zoll J, Erkens Hulshof S, Lanke K, *et al*. Saffold virus, a human Theiler's-like cardiovirus, is ubiquitous and causes infection early in life. *PLoS Pathogens* 2009; 5: e1000416.