Epidemiol. Infect. (2009), **137**, 1302–1308. © 2009 Cambridge University Press
doi:10.1017/S0950268808001878 Printed in the United Kingdom

Genetic relationship in southern African Crimean-Congo haemorrhagic fever virus isolates: evidence for occurrence of reassortment

F. J. BURT[1]*, J. T. PAWESKA[2], B. ASHKETTLE[3] AND R. SWANEPOEL[2]

[1] *Department of Medical Virology, National Health Laboratory Services Universitas and Faculty of Health Science, University of the Free State, Bloemfontein, South Africa*
[2] *Special Pathogens Unit, National Institute for Communicable Diseases, Sandringham, Johannesburg, South Africa*
[3] *Department of Medical Virology, Faculty of Health Science, University of the Free State, Bloemfontein, South Africa*

(*Accepted 3 December 2008; first published online 23 January 2009*)

SUMMARY

Crimean-Congo haemorrhagic fever (CCHF) is a tick-borne viral zoonosis widely distributed in Africa, Asia and eastern Europe. Reassortment of CCHF genome segments has been shown to occur in nature. We therefore investigated the genetic relationship of southern African isolates using partial sequence data for each RNA segment, S, M and L, and comparing the tree topologies constructed using a neighbour joining method. A total of 21 southern African isolates were studied. The incongruencies which were identified in S, M and L sequence datasets involved group switching implying reassortment for 15 isolates. A higher fatality rate occurred in patients infected with isolates which had apparently acquired M segments from a group in which predominantly Asian strains are usually found. This suggests that reassortment may affect the pathogenicity of the virus.

Key words: CCHF, genetic reassortment.

INTRODUCTION

Crimean-Congo haemorrhagic fever (CCHF) is a tick-borne viral zoonosis widely distributed in Africa, Asia and eastern Europe within the distribution range of ticks belonging to the genus *Hyalomma* [1]. The virus is a member of the *Nairovirus* genus of the family Bunyaviridae and has a single-stranded, negative-sense RNA genome consisting of three segments L (large) $4{\cdot}1–4{\cdot}9 \times 10^6$ Da, M (medium) $1{\cdot}5–1{\cdot}9 \times 10^6$ Da and S (small) $0{\cdot}6–0{\cdot}7 \times 10^6$ Da, each contained in a separate nucleocapsid within the virion [2]. The virions contain three major structural proteins: two envelope glycoproteins, G_N and G_C encoded by the M segment, and a nucleocapsid protein, encoded by the S segment, plus minor quantities of viral transcriptase or L (large) protein. The M segment of CCHF virus has one open reading frame which encodes a precursor polypeptide, with a highly variable amino-terminal domain and fairly conserved carboxyl-terminal region. The precursor polyprotein is cleaved to form the two mature glycoproteins, G_N (37 kDa) and G_C (75 kDa) [3].

Recent reports on molecular characterization of CCHF virus include the first complete sequence data for all three RNA segments [4–8], characterization of the M segment [6], genetic analyses based on

* Author for correspondence: Professor F. J. Burt, Department of Medical Virology, National Health Laboratory Services Universitas and Faculty of Health Science, University of the Free State, Bloemfontein, South Africa.
(Email: burtfj.md@ufs.ac.za)

Table 1. *Origins of CCHF virus isolates included in the present study*

Isolate	Locality (province or city, country)	Source of infection	Outcome	GenBank accession no.		
				S segment	M segment	L segment
SPU247/85	Kwazulu Natal, SA	Unknown	Survived	FJ435410	FJ435390	FJ435385
SPU415/85	Northern Cape, SA	Nosocomial	Died	AY905627*	FJ435406	FJ435378
SPU196/86	Northern Cape, SA	Sheep	Died	AY905629*	FJ435404	FJ435379
SPU536/86	Windhoek, Namibia	Unknown	Died	AY905632*	FJ435407	FJ435377
SPU566/86	Free State, SA	Tick bite	Survived	AY905633*	FJ435405	FJ435375
SPU103/87	Northern Cape, SA	Tick bite	Survived	AY905634*	FJ435393	FJ435369
SPU383/87	Northern Cape	Tick bite	Survived	FJ435416	FJ435397	FJ435373
SPU18/88	Northern Cape, SA	Tick bite	Died	AY905635*	FJ435401	FJ435376
SPU45/88	Free State, SA	Tick bite	Died	FJ435422	FJ435403	FJ435367
SPU71/88	Northern Cape, SA	Livestock	Died	FJ435423	FJ435402	FJ435365
SPU203/88	Gauteng, SA	Tick bite	Died	FJ435413	FJ435387	FJ435381
SPU273/88	Northern Cape, SA	Tick bite	Survived	FJ435420	FJ435394	FJ435372
SPU497/88	Namibia	Livestock/ticks	Died	FJ435419	FJ435399	FJ435371
SPU498/88	Northern Cape, SA	Tick bite	Survived	FJ435412	FJ435409	FJ435382
SPU130/89	Northern Cape, SA	Tick bite	Survived	FJ435417	FJ435396	FJ435374
SPU281/89	Kwazulu Natal, SA	Unknown	Survived	FJ435411	FJ435391	FJ435386
SPU337/89	Gauteng, SA	Abattoir worker	Survived	FJ435421	FJ435389	FJ435366
SPU372/89	Northern Cape, SA	Tick bite	Survived	FJ435414	FJ435388	FJ435368
SPU48/90	North West Province, SA	Unknown, rural resident	Survived	FJ435415	FJ435392	FJ435364
SPU380/90	Free State, SA	Tick bite	Survived	FJ435418	FJ435398	FJ435370
SPU51/01	Western Cape, SA	Abattoir worker	Survived	AY905649*	FJ435400	FJ435380
SPU246/02/18	Quetta, Pakistan	Unknown	Unknown	AY905661*	FJ435395	FJ435383
SPU246/02/20	Quetta, Pakistan	Unknown	Unknown	AY905662*	FJ435408	FJ435384

* Sequence data for S segment determined in this study retrieved from GenBank [12].

nucleotide sequence data from the S and M segments from various geographic regions [5–14], and evidence for the natural occurrence of reassortment and recombination [10, 14, 15]. RNA viruses with segmented genomes have the ability to reassort when dual infection occurs, and this can play a role in pathogenicity and epidemiology of the viruses. This in turn implies that phylogeny based on nucleotide sequence data for the S segment alone [12] may be inadequate. Hence we extended our earlier observations to the M and L segments of the genome to determine whether reassortment plays a role in viral pathogenicity, or correlates with geographic distribution within southern Africa.

METHODS AND MATERIALS

Virus isolates

Details of the origin of 23 CCHF isolates included in the study are summarized in Table 1; 21 isolates were from southern Africa and two were from Pakistan. Sequence data were determined retrospectively for isolates collected in southern Africa from 1985 to 2001 and stored at −70 °C as freeze-dried 10% suckling mouse brain suspensions at the level of mouse brain passage 2–3. The suspensions were inoculated into Vero cell cultures and total RNA extracted from the infected cells using the acid guanidium thiocyanate–phenol–chloroform method. For two isolates viral RNA was extracted directly from serum samples submitted from patients in Pakistan using the QIAamp viral RNA isolation kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer's protocol.

Reverse transcriptase polymerase chain reaction (RT–PCR) and nucleotide sequencing of amplicons

Partial nucleotide sequences were determined for each segment of the genome. Pairs of primers were identified from alignment of sequence data retrieved from GenBank for five complete M and L segments of southern African isolates (SPU41/84, SPU128/84, SPU97/85, SPU415/85, SPU103/87) (accession numbers AY900142, AY900141, DQ211633, DQ211635,

https://doi.org/10.1017/S0950268808001878 Published online by Cambridge University Press

Table 2. *Data retrieved from GenBank, accession numbers provided*

Isolate	Country and year of isolation	Source of isolate	GenBank accession no.		
			S segment	M segment	L segment
HD 38562	Burkina Faso, 1983	Human	U15093	n.a.	n.a.
ArB 604	Central African Republic, 1973	*Hyalomma nitidum*	U15092	n.a.	n.a.
7001	China, 1970	Human	AF415236	AB069670	n.a.
7803	China, 1978	Human	AF354296	AB069672	n.a.
8402	China, 1984	Tick	AJ010649	AB69674	n.a.
66019	China, 1965	Human	AJ010648	AB069669	n.a.
75024	China, 1975	Human	AF362080	AB069671	n.a.
79121	China, 1979	Long-eared jerboa	AF358784	AB69673	n.a.
88166	China, 1988	Human	AY029157	AB069675	n.a.
C-68031	China, 1968	Sheep	M86625	DQ211629	DQ211616
HY13	China, 1968	*Hyalomma asiaticum*	U88413	n.a.	n.a.
AP92	Greece, 1975	*Rhipecephalus bursa*	U04958	DQ211625	DQ211612
Baghdad 12	Iraq, 1976	Human	AJ538196	AJ538197	AY647890
HD49199	Mauritania, 1988	Human	U15023	n.a.	n.a.
ArD39554	Mauritania, 1984	*Hyalomma marginatum rufipes*	U15089	DQ211628	DQ211615
IbAr10200	Nigeria, 1966	*Hyalomma excavatum*	NC005302	AF467768	AY947891
Oman	Oman, 1997	Human	DQ211645	DQ211632	DQ211619
Matin	Pakistan, 1976	Human	AF527810	AF67769	AY422208
JD206	Pakistan, 1965	*Hyalomma anatolicum*	U88414	n.a.	n.a.
UG3010	Republic of Congo, 1956	Human	U88416	DQ211637	DQ211624
Drosdov	Russia, 1967	Human	U88412	DQ211630	DQ211617
VLG/TI29414	Russia, 2000	*Hyalomma marginatum*	n.a.	Ay179961	NA
VLV100	Russia, 2003	Human	n.a.	n.a.	AY955166
ArD97264	Senegal, 1993	*Hyalomma marginatum rufipes*	U15090	n.a.	n.a.
ArD97268	Senegal, 1993	*Hyalomma truncatum*	U15091	n.a.	n.a.
ArD8194	Senegal, 1969	*Hyalomma truncatum*	U15021	DQ211626	DQ211613
ArD15786	Senegal, 1972	Goat	U15020	DQ211627	DQ211614
TADJ/HU8966	Tajikistan, 1990	Human	n.a.	AY179962	AY720893

n.a., Not available.

DQ211634, DQ076417, DQ076414, DQ211620, DQ211622 and DQ211621). To confirm that incongruencies resulted from segment reassortment, two regions of the M segment were sequenced and the topology of the trees constructed from the data were compared.

The M primer sequences (positions relative to SPU415/85) were:

MF3 (4804–4821): 5′-G(C/T)T GCT GG(C/T) TAG AGT CAG-3′
MR3 (5365–5346): 5′-CTC AAA GAT ATA GTG GCT GC-3′
MF2 (2756–2774): 5′-CCC TGG A(C/T)T GTC CAT TTG C-3′
MR2 (3392–3372): 5′-CTT (A/G)TT GCC TCT GTG TTC (C/G)AC-3′.

L segment primer sequences were:

LF1 (2120-2139): 5′-CAG TTG CAT CCA GAG TTC AG-3′
LR1 (2969-2951): 5′-CTT TGC GCA TTG CCT GTT C-3′.

Partial sequence data for the S segment of southern African isolates were obtained from a previous study using published primer pairs F2 and R3 [12, 16, 17]. In addition, sequence data were retrieved from GenBank for isolates from other geographic regions, to determine their relationships to southern African isolates (Table 2).

For all amplifications RT–PCR was performed using the Titan One Tube RT–PCR system (Roche Diagnostics GmbH, Mannheim, Germany) according

https://doi.org/10.1017/S0950268808001878 Published online by Cambridge University Press

to the manufacturer's instructions and cycled using standard cycling conditions and annealing temperatures determined from the T_m of the oligonucleotide primers. The nucleotide sequences of the amplicons were determined using Big Dye™ Terminator Sequencing Ready Reaction kits (Applied Biosystems, CA, USA) according to manufacturer's instructions.

Data analysis

After editing of the data and allowing for inclusion of data retrieved from GenBank, the phylogenetic analyses were performed using a 413 base pair (bp) region from the conserved 3′-carboxyl terminal of the M segment, a 590 bp region targeting the central region of the M gene, a 436 bp region of the S segment and a 754 bp region of the L segment. The data were analysed using Molecular Evolutionary Genetics Analysis (MEGA) version 3.1 and bootstrap neighbour-joining method with 1000 replicates [18]. Sequence divergence was determined using MEGA to calculate the average *P* distances within groups and between groups.

RESULTS

Genetic analysis

The genetic relationship of isolates determined using neighbour-joining analysis is shown in Figure 1 based on partial nucleotide sequence data for (*a*) S segment (436 bp), (*b*) M segment, 413 bp 3′-region (the tree topology was identical for a 590 bp region from the middle of the M segment, data not shown), and (*c*) L segment (754 bp) of CCHF virus. The phylogenetic relationships determined for CCHF isolates using partial segment data were in agreement with findings previously obtained with complete segment sequence data [14, 19], and the groups were designated I–VIII following the nomenclature of Deyde *et al.* [14]. Tree topologies based on the S and L segment data showed good correlation, with no incongruencies in the grouping of the southern African isolates. Group switching of non-southern African isolates ArD8194 and ArD15786 from S group I to L group III was previously described by Deyde *et al.* [14] who concluded that it was probably the result of L segment reassortment. Comparison of the present tree topologies for each segment showed incongruencies in the M segment grouping for 15 southern African and one Asian isolate. Two isolates, SPU247/85 and SPU281/89 (shown in bold in Fig. 1) from group II in both S and L trees acquired M segments from group III; one isolate (SPU498/88, shown in bold in Fig. 1) from group III in both S and L trees acquired M segment from group VII; one Asian isolate (SPU246/02/18) from group IV in both S and L trees acquired M segment from group III; and 12 isolates (shown in upper case in Fig. 1) from group III in both S and L trees (SPU380/90, SPU497/88, SPU51/01, SPU130/89, SPU383/87, SPU18/88, SPU71/88 (group I in S), SPU45/88, SPU566/86, SPU196/86, SPU536/86, SPU415/85) acquired M segments from group IV. In the absence of complete nucleotide sequence data for the CCHF strains included in the present study, and to support the hypothesis that the discrepancies in grouping were the result of segment reassortment and not recombination events, an additional section of the M segment was sequenced and the tree topology based on this region was identical to that shown in Figure 1 (data not shown). Group IV in both S and L trees comprised isolates solely from Asia. The apparent acquisition of Asian group IV M segments by southern African isolates suggests movement of CCHF virus between continents.

A high degree of nucleotide diversity was evident when sequence divergence was determined using MEGA to calculate the average *P* distances within groups and between groups for the southern African isolates. The S segment was the most diverse, with nucleotide distances between groups ranging from 29·6% to 43·1%, whereas the amino-acid distances ranged from 2·8% to 4·9%, suggesting that the majority of changes were synonymous. Within each S segment group nucleotide distances were <10·1% and predicted amino-acid distances 1·7%. The L segment was more conserved, with nucleotide distances between groups ranging from 4·1% to 6·6% and amino-acid distances ranging from 7·9% to 11·6%, suggesting that more non-synonymous changes had occurred. Within L segment groups the nucleotide distances were <1·1% and the predicted amino-acid distances 2·1%. The M segment showed the largest number of non-synonymous changes between groups with nucleotide distances ranging from 10·0% to 33·9%, translating to predicted amino-acid differences of 15·3% to 37·3%, while within groups nucleotide distances were as low as 0·6% to 4·3% and predicted amino-acid differences ranged from 1·3% to 5·4%.

An interesting finding is the possible association between M segment reassortment and increased pathogenicity. Of 21 randomly selected CCHF isolates from

https://doi.org/10.1017/S0950268808001878 Published online by Cambridge University Press

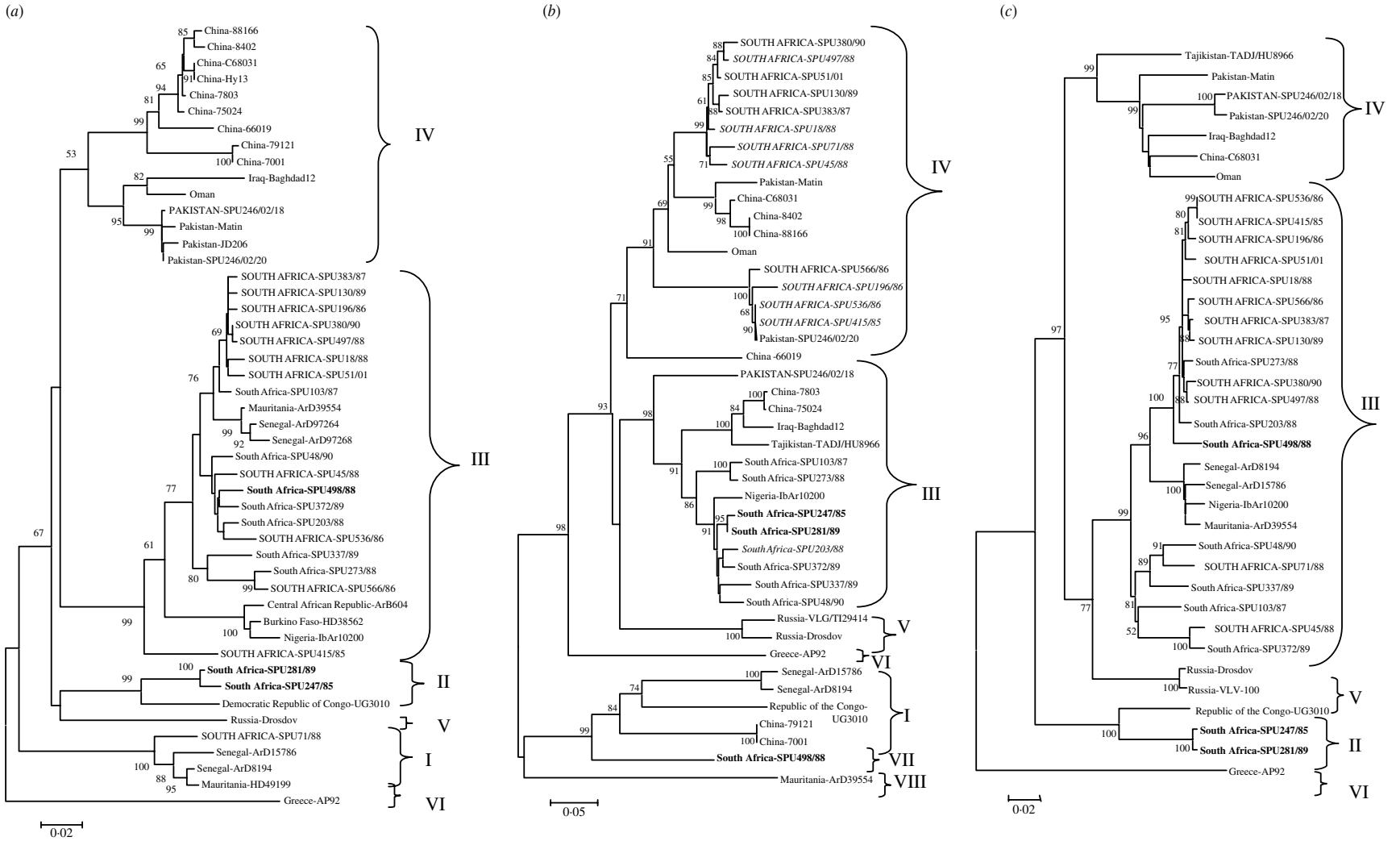

Fig. 1. Phylogenetic analysis of worldwide strains of CCHF virus based on partial nucleotide sequence data from (*a*) S segment (436 bp), (*b*) M segment, 413 bp 3′-region (590 bp from the middle region of the M segment, tree topology identical, not shown) and (*c*) L segment (754 bp) of CCHF virus. Data analysed using a bootstrap neighbour-joining method with 1000 replicates using Molecular Evolutionary Genetics Analysis (MEGA) version 3.1. Patients with a fatal outcome illustrated in italics.

southern Africa, 7/8 isolates from patients with fatal outcome had evidence suggesting the acquisition of M segment RNA from group IV viruses, compared to 5/13 isolates from non-fatal infections, but the difference is not highly significant ($P = 0{\cdot}027$).

DISCUSSION

Reassortment and recombination events have been described for CCHF virus and other bunyaviruses. Although confirmation of genomic RNA segment exchange and recombination would theoretically require full-length sequencing, previous publications have consistently shown close relationships between segment topologies based on partial and complete sequence data [14]. Hence we focused on sequencing selected regions. The L segment is largely conserved with three variable regions, the N terminal, C terminal and a variable region from bases 2346–2488 which we selected for sequence determination [8]. S segment sequence data were available from a previous study [12]. The M segment has a highly variable region at the N terminal and a more conserved region coding for the glycoprotein precursor. Based on the assumption that if phylogenies generated from different regions of the same segment have different topologies then the incongruencies in grouping are more likely to be attributable to recombination than reassortment, we selected to confirm the consistency of the topology generated for the M segment by determining partial sequence data for two distinct regions.

In a previous study based solely on partial S segment data of southern African isolates it was concluded that despite the potential which exists for dispersal of the virus between Africa and Eurasia, circulation is largely compartmentalized within the two land masses, with the inference that the geographic distribution of phylogenetic groups is largely related to the distribution of tick vectors of the virus [12]. Evidence that multiple strains were clustered geographically, and that in some instances similar subtypes existed in distant geographic locations, was difficult to explain. It was postulated that genetic diversity within regions resulted from the introduction of multiple lineages of virus over millennia through carriage of infected immature ticks on migrating birds, or through the movement of livestock, as well as from evolutionary pressure within ecological niches. Evidence for the occurrence of reassortment and recombination indicates the existence of additional mechanisms for the generation of genetic diversity, and implies that phylogeny based on nucleotide sequence data from the S segment alone is not adequately informative. The present findings indicate that M segment reassortment is not uncommon and tends to confirm the occurrence of spillover of virus circulation between southern Africa and Asia. It has been suggested, that co-infection and reassortment is more likely to occur over time within vectors than within mammalian hosts, partly because ticks remain infected for longer and are potentially exposed to successive infected hosts [10, 14, 19].

Reassortment has previously been associated with changes in pathogenicity of bunyaviruses. Ngari virus, for example, was found to be a reassortant orthobunyavirus in which the L and S segments are derived from Bunyamwera virus and the M segment from Batai virus. While Bunyamwera and Batai viruses cause benign disease, Ngari was found to be associated with cases of haemorrhagic fever in East Africa [20–22]. The determinants of fatal outcome in CCHF infection are largely unknown, but it is thought that factors such as age and underlying health play a role. Clearly, the possibility that segment reassortment may be associated with increased severity of CCHF infection merits further investigation. While no differences in antigenicity between CCHF isolates has been reported, genetic variation must be taken into consideration when developing new-generation diagnostic assays or vaccines.

ACKNOWLEDGEMENTS

The project was funded by the Polio Research Foundation, South Africa.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Hoogstraal H.** The epidemiology of tick-borne Crimean-Congo haemorrhagic fever in Asia, Europe, and Africa. *Journal of Medical Entomology* 1979; **15**: 307–417.
2. **Clerx JP, Casals J, Bishop DH.** Structural characteristics of nairoviruses (genus *Nairovirus, Bunyaviridae*). *Journal of General Virology* 1981; **55**: 165–178.
3. **Marriott AC, Nuttall PA.** Molecular biology of nairoviruses. In: Elliott RM, ed. *The Bunyaviridae*. New York: Plenum Press, 1996, pp. 91–104.
4. **Marriott AC, Nuttall PA.** Comparison of S RNA segments and nucleoprotein sequences of Crimean-Congo

hemorrhagic fever, hazara and Dugbe viruses. *Virology* 1992; **189**: 795–799.
5. **Sanchez AJ, Vincent MJ, Nichol ST.** Characterization of the glycoprotein of Crimean-Congo hemorrhagic fever virus. *Journal of Virology* 2002; **76**: 7263–7265.
6. **Papa A, *et al.*** Genetic characterization of the M RNA segment of Crimean Congo hemorrhagic fever virus strains, China. *Emerging Infectious Diseases* 2002; **8**: 50–53.
7. **Honig JE, Osborne JC, Nichol ST.** Crimean-Congo hemorrhagic fever virus genome L RNA segment and encoded protein. *Virology* 2004; **321**: 29–35.
8. **Meissner JD, *et al.*** Complete L segment coding-region sequences of Crimean-Congo haemorrhagic fever virus strains from the Russian Federation and Tajikistan. *Archives of Virology* 2006; **151**: 465–475.
9. **Yashina L, *et al.*** Genetic variability of Crimean-Congo haemorrhagic fever virus in Russia and Central Asia. *Journal of General Virology* 2003; **84**: 1199–1206.
10. **Hewson R, *et al.*** Evidence of segment reassortment in Crimean-Congo hemorrhagic fever virus. *Journal of General Virology* 2004; **85**: 3059–3070.
11. **Seregin SV, *et al.*** Genetic characterization of the M RNA segment of Crimean-Congo hemorrhagic fever virus strains isolated in Russia and Tajikistan. *Virus Genes* 2004; **28**: 187–193.
12. **Burt FJ, Swanepoel R.** Molecular epidemiology of African and Asian Crimean-Congo haemorrhagic fever isolates. *Epidemiology and Infection* 2005; **133**: 659–666.
13. **Tonbak S, *et al.*** Crimean-Congo hemorrhagic fever virus: genetic analysis and tick survey in Turkey. *Journal of Clinical Microbiology* 2006; **44**: 4120–4124.
14. **Deyde VM, *et al.*** Crimean-Congo hemorrhagic fever virus genomics and global diversity. *Journal of Virology* 2006; **80**: 8834–8842.
15. **Lukashev AN.** Evidence for recombination in Crimean-Congo hemorrhagic fever virus. *Journal of General Virology* 2005; **86**: 2333–2338.
16. **Schwarz TF, *et al.*** Polymerase chain reaction for diagnosis and identification of distinct variants of Crimean-Congo hemorrhagic fever virus in the United Arab Emirates. *American Journal of Tropical Medicine and Hygiene* 1996; **55**: 190–196.
17. **Burt FJ, *et al.*** The use of a reverse transcription-polymerase chain reaction for the detection of viral nucleic acid in the diagnosis of Crimean-Congo haemorrhagic fever. *Journal of Virological Methods* 1998; **70**: 129–137.
18. **Kumar S, *et al.*** MEGA3.1: Molecular evolutionary genetics analyses software. Version 2.0. University Park: Pennsylvania State University, 2001.
19. **Morikawa S, Saijo M, Kurane I.** Recent progress in molecular biology of Crimean-Congo hemorrhagic fever. *Comparative Immunology, Microbiology and Infectious Diseases* 2007; **30**: 375–389.
20. **Sall AA, *et al.*** Genetic reassortment of Rift valley fever virus in nature. *Journal of Virology* 1999; **73**: 8196–8200.
21. **Bowen MD, *et al.*** A reassortant bunyavirus isolated from acute hemorrhagic fever cases in Kenya and Somalia. *Virology* 2001; **291**: 185–190.
22. **Briese T, *et al.*** Batai and Ngari viruses: M segment reassortment and association with severe febrile disease outbreak in East Africa. *Journal of Virology* 2006; **80**: 5627–5630.