SEQUENCE AND PHYLOGENETIC ANALYSIS OF SH, G, AND F GENES AND PROTEINS OF HUMAN RESPIRATORY SYNCYTIAL VIRUS ISOLATES FROM SINGAPORE

C.S. LIM[1], G. KUMARASINGHE[2], V.T.K. CHOW[1*]

[1]Programme in Infectious Diseases, Department of Microbiology, Faculty of Medicine, National University of Singapore, Kent Ridge 117597, Singapore; [2]Department of Laboratory Medicine, National University Hospital, Kent Ridge, Singapore

Received March 26, 2003; accepted May 14, 2003

Summary. – To study the genetic variability and molecular epidemiology of Human respiratory syncytial virus (HRSV) occurring in Singapore, nucleotide sequencing of three membrane-associated genes (SH, G and F) of four local isolates was performed. Comparison of their nucleotide and amino acid sequences with those of the prototype strains A2 (subgroup A) and CH-18537 (subgroup B) indicated that the Singapore isolates belong to the subgroup A. Comparison of the Singapore isolates with the reference strain A2 showed that whereas the G protein was the most divergent with up to 15% difference, the F and SH proteins showed less diversity of only up to 4%. Each gene exhibited its distinct variable and conserved regions. The *N*- and *O*-glycosylation sites within the G protein of the isolates were analyzed to ascertain their potential implications on the antigenicity of the viral glycoprotein. Based on the second variable region of the G protein, phylogenetic analysis of the Singapore isolates with 91 previously identified genotypes of subgroup A revealed that more than one genotype (GA2 and GA5) may circulate in the local population at a given time. This epidemiological study reflects the pattern of genetic relationships between the HRSV isolates from Singapore to those from other parts of the world.

Key words: F, G, SH genes/proteins; genotypes; phylogeny; Human respiratory syncytial virus; sequence analysis

Introduction

HRSV is one of the major causative agents of lower respiratory tract infections in young children worldwide. Mortality rates of up to 5% have been reported in infants and children with underlying heart or lung disease, and they may even exceed 90% in severely immunocompromised children and adults. In Singapore, HRSV infection accounts for about 72% of all infants hospitalized for lower respiratory tract infections (Chew *et al*., 1998). HRSV (the species *Human respiratory syncytial virus*, the genus *Pneumovirus*, the family *Paramyxoviridae*) is an enveloped virus with a linear non-segmented single-stranded negative-sense RNA genome. The viral genome comprises 10 genes in the following order: 3'-NS1-NS2-N-P-M-SH-G-F-M2-L-5'. All RSV genes initiate with a nine-nucleotide conserved gene-start sequence and terminate with a 12–13-nucleotide semi-conserved gene-end/polyadenylation sequence, which serve as transcriptional signals for direct synthesis of individual mRNAs. The complete SH gene of 405 nucleotides includes an ORF spanning nt 85–276 that encodes a protein of 64 amino acids. The G gene of 917 nucleotides and the F gene

*Corresponding author. E-mail: micctk@nus.edu.sg; fax: +65-6776 6872.

Abbreviations: CPE = cytopathic effect; DEPC = diethylpyrocarbonate; dNTP = deoxyribonucleoside triphosphate; EDTA = ethylenediamine tetraacetate; F = fusion; G = attachment glycoprotein; MAb = monoclonal antibody; PCR = polymerase chain reaction; HRSV = Human respiratory syncytial virus; RT = reverse transcription; SH = small hydrophobic

of 1899 nucleotides contain ORFs spanning nucleotides 470–1363 and 1443–3164 that encode the corresponding proteins of 298 and 574 amino acids, respectively.

Based on their reactivity with monoclonal antibodies (MAbs), HRSV isolates are classified into two antigenic subgroups, A and B (Anderson *et al*., 1985; Mufson *et al*., 1985). Three viral proteins are associated with the viral envelope, namely the fusion glycoprotein (F), attachment glycoprotein (G) and small hydrophobic protein (SH). The G protein is involved in binding the virus to the cellular receptor, while the F protein is responsible for promoting membrane fusion and virus penetration. However, the exact function of the SH protein is so far unknown. In response to HRSV infection, the two most important neutralizing antibodies are generated against F and G proteins (Taylor *et al*., 1984; Hall *et al*., 1991). Antibody studies have shown that the F protein is the most efficient HRSV antigen, as the protective immunity induced by F protein is equally effective against viruses of both the subgroups A and B, consistent with their structural similarities (Olmsted *et al*., 1986). In contrast to the F protein, the G protein is the most divergent gene product both between and within HRSV subgroups. It contains two hypervariable regions separated by a conserved motif comprising aa 164–176, which is responsible for receptor binding. At amino acid level, the G protein shows only 53% identity to the prototype strains of the subgroups A and B, and up to 20% sequence diversity within the same antigenic subgroup (Johnson *et al*., 1987; Cane *et al*., 1991, 1994). The second variable region, consisting of the C-terminal region of the G protein, has been reported to serve as a reliable proxy for variability of the entire G gene, and has been used in phylogenetic analysis for molecular epidemiological studies in various countries (Peret *et al*., 1998, 2000; Choi and Lee, 2000; Venter *et al*., 2001). Thus, this formed the basis for the molecular epidemiological analysis of the Singapore HRSV isolates. In addition to the G protein, the variability of the HRSV F and SH proteins was also evaluated at nucleotide and amino acid levels. Analyses of these three viral proteins resulted in a better understanding of the antigenic relatedness of HRSV strains circulating in Singapore compared with previously identified genotypes.

Materials and Methods

Patients. Four local HRSV strains have been isolated from nasopharyngeal aspirates collected from three children and an adult in 2000–2001. All four patients had symptoms of an upper respiratory tract infection and recovered without any complications. They were symptomatically treated without antiviral therapy. The isolate LLC1144-115 was from a 15-week-old Indonesian boy with bilateral hydronephrosis who presented with cough, dyspnoea and tachycardia. The isolate LLC235-267 was from a 5-month-old Chinese girl who was born prematurely due to placental insufficiency, with frequent respiratory tract infections since birth. She presented with fever and cough of four days' duration. The isolate LLC242-282 was from a 20-month-old Chinese girl with a history of fever, cough and irritability of one week's duration. The isolate LLC62-111 was from a 53-year-old Indian man, a heavy smoker who presented with cough and chest pain, and was diagnosed to have atypical angina with a mild respiratory tract infection.

Table 1. Nucleotide sequences of the primers used for amplification and sequencing of HRSV genes

Primer	Nucleotide sequence (5' – 3')
For amplification of SH-G fragment	
RSVP3F	GGAAGCACACAGCTACACGA
RSVP6R	AACAATGGAGTTGCCAATCC
For sequencing of SH-G fragment	
RSVP3F	GGAAGCACACAGCTACACGA
RSVP7F	CAACATCTCACCATGCAAGCC
RSVP8F	CAGCTTGGAATCAGCTTCTCCA
For amplification of F fragment	
RSVP5F	TCCATCTCCATCCAACACAA
RSVP4R	CCAACTCTGCAGCTCCACTT
For sequencing of F fragment	
RSVP9F	CCACAACAAGGCTGTAGTCA
RSVP10R	ACGAATAAATGCTAGGCTCTGG
RSVP4R	CCAACTCTGCAGCTCCACTT

Virus isolates. The specimens were initially tested for a routine panel of respiratory viruses, i.e. HRSV, influenza viruses A and B, parainfluenza viruses 1, 2, and 3, and adenoviruses by propagating them in LLC-MK2 and MDCK cells in Minimal Essential Medium at 35°C. The cultures were observed daily for cytopathic effect (CPE). FITC-conjugated MAbs confirmed the presence of HRSV in cell cultures. Uninfected cells and known positives were also used regularly as controls.

Total RNA was extracted from virus-infected cell culture supernatants according to Chungue *et al*. (1993) with minor modifications. Briefly, 500 µl of each supernatant was treated with 500 µl of lysis buffer (8 mol/l guanidine isothiocyanate, 0.1 mol/l Tris-HCl pH 6.4, 35 mmol/l EDTA, and 0.02% Triton X-100) and acid-washed silica at room temperature for 10 mins. After centrifugation at 12,000 rpm for 15 secs, the resulting RNA-bound silica pellet was washed twice with a wash buffer (50% ethanol, 10 mmol/l Tris-HCl, 1 mmol/l EDTA, and 50 mmol/l NaCl) and once with the DEPC-treated water. The silica was pelleted and RNA was eluted with 50 µl of nuclease-free water at 56°C for 10 mins.

cDNA synthesis was performed in a 20 µl reaction mixture containing 0.1 µg of total RNA, 1 x first-strand buffer, 0.5 mmol/l dNTPs 2.5 µmol/l random hexamers, 10 U of RNase inhibitor and 100 U of SuperScript II reverse transcriptase. The mixture was then incubated at 25°C for 20 mins, at 42°C for 1 hr, at 99°C for 1 min, and the cDNA was stored at -20°C prior to use.

RT-PCR and nucleotide sequencing. The primers RSVP3F and RSVP6R were employed to amplify the contiguous SH and G ge-

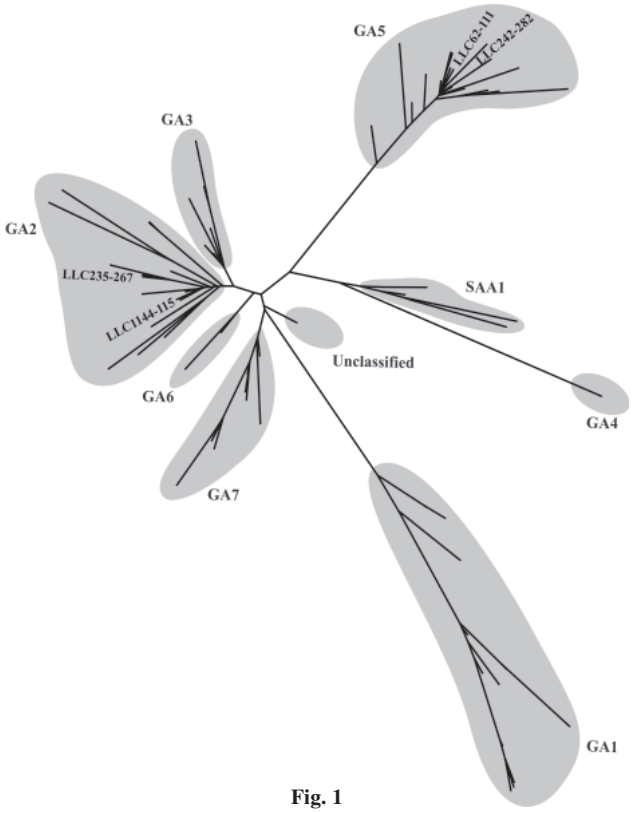

Fig. 1
Phylogenetic analysis of 95 HRSV isolates belonging to the subgroup A

The tree was constructed on the basis of nt 649–918 of the G gene. Distribution of the isolates according to genotype and country of origin is listed in Table 3.

nes, while the F gene was amplified by the primers RSVP5F and RSVP4R (Table 1). The amplification mixture contained a cDNA template, 1 x PCR buffer, dNTPs, the primer pair and *Taq* DNA polymerase. The reaction was run for 35 cycles each at 95°C for 1 min, at 56–58°C for 0.5 min and at 72°C for 1 min, followed by a final extension at 72°C for 1 min. The RT-PCR products were electrophoresed in ethidium bromide-stained agarose gels and visualized under an ultra-violet transilluminator. Amplified target product bands were excised from agarose gels, and purified using the QIAquick Gel Extraction Kit (Qiagen, Germany). The purified fragments were directly sequenced in both directions with the corresponding primers (Table 1) using the ABI PRISM BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, USA).

Sequence data and phylogenetic analysis. The derived nucleotide sequences were translated using an on line translation tool (http://www.expasy.org/tools/dna.html). The pI values of viral proteins were predicted using the bioinformatics software from the same website (http://www.expasy.org). The deduced amino acid sequences of the SH, G and F proteins of the Singapore isolates were aligned and compared with the prototype strains A2 and CH-18537 using the CLUSTAL W software version 1.8. The percentage similarity of the four Singapore isolates was compared and calculated by the GAP analysis in the GCG package (Chow and Leong, 1999). Using the PHYLIP Version 3.5 phylogenetic trees were constructed on the basis of the sequences of a 270-nucleotide fragment spanning the positions 649–918 of the G gene of the Singapore isolates and previously identified subgroup A isolates available at the GenBank. These included two prototype strains A2 and Long; 35 isolates from USA and Canada (AL – Alabama, CH – Rochester, MO – Missouri, NY – New York, TX – Texas, WV – West Virginia, CN – Canada); 20 isolates from Korea; 13 isolates from Africa (MOZ – Mozambique, SA – South Africa); 11 isolates from Uruguay (MON – Montevideo); and 10 isolates from Spain (MAD – Madrid).

Results

Nucleotide and amino acid sequence analyses of SH, G and F genes and proteins

The full-length sequences of SH, G and F genes of the four Singapore isolates were amplified by RT-PCR. The primer pair RSVP3F and RSVP6R amplified a product of 1.6 kbp comprising the entire SH and G genes. The F gene was amplified using the primers RSVP5F and RSVP4R that generated a product of 2.2 kbp. The full-length nucleotide sequences of the SH, G and F genes of the four Singapore isolates were deposited at the GenBank under Acc. Nos. AF512538 (LLC1144-115), AY114149 (LLC235-267),

Table 2. Percentage similarity of nucleotide and amino acid sequences of the coding region of G gene

HRSV isolates	LLC1144-115	LLC242-282	LLC62-111	LLC235-267	A2	18537
LLC1144-115		87.9	87.2	97.0	88.9	60.9
LLC242-282	93.4		99.3	86.9	85.2	56.7
LLC62-111	93.3	99.6		86.2	84.5	56.7
LLC235-267	97.2	92.8	92.6		88.9	60.5
A2	92.6	91.3	91.1	92.3		60.0
18537	69.1	67.9	68.0	69.6	67.7	

Nucleotide level

Amino acid level

Table 3. Distribution of 95 HRSV subgroup A isolates according to genotype and country of origin

Genotype	No. of isolates	Country of origin
GA1	2	(Prototype strains)
	5	Uruguay
	8	USA
GA2	8	Korea
	2	Mozambique
	2	Singapore
	2	South Africa
	2	Spain
	3	Uruguay
	5	USA
GA3	1	Canada
	2	Korea
	1	Spain
	1	Uruguay
	4	USA
GA4	1	USA
GA5	2	Canada
	2	Korea
	3	Mozambique
	2	Singapore
	1	South Africa
	5	Spain
	2	Uruguay
	8	USA
GA6	2	USA
GA7	2	Canada
	6	Korea
	1	South Africa
	2	Spain
	1	USA
SAA1	2	Korea
	4	South Africa
Unclassified	1	USA

AY114150 (LLC242-282) and AY114151 (LLC62-111), respectively.

Among the four isolates compared by GAP analysis, SH and F genes were the least diverse. The SH gene showed only up to 4% and 2% divergence among the four isolates at nucleotide (coding region) and amino acid levels, respectively, similarly to the F gene which exhibited about 3% and 2% divergence, respectively. In contrast, the G gene displayed only 92.6–99.6% and 86.2–99.3% similarity at nucleotide and amino acid levels, respectively. Comparison of the G gene sequences revealed that the isolates LLC242-282 and LLC62-111 were the most closely related, in contrast to the isolates LLC62-111 and LLC235-267 which were the least similar. Comparison of the Singapore isolates and reference strains from the two subgroups also revealed that our isolates belong to subgroup A with up to about 15% difference from the reference subgroup A strain A2, and up to about 43% difference from the reference subgroup B strain CH-18537 (Table 2). It can be further observed that for the G gene, the similarity between strains at amino acid level was lower than that at nucleotide level. However, the reverse was true for the SH and F genes (data not shown). The G protein is known to be the most variable gene product both between and within subgroups (Johnson *et al*., 1987; Garcia *et al*., 1994) and is used in the classification of both the genotypes and subtypes of HRSV (Peret *et al*., 1998, 2000). Consequently, the second variable region of the G gene (nt 649–918) was selected for the phylogeny analysis to study the variability of the Singapore HRSV isolates with respect to other 91 HRSV strains from subgroup A (Fig. 1 and Table 3). Notable from the phylogenetic tree was the clustering of the isolates LLC1144-115 and LLC235-267 to genotype GA2, while the isolates LLC62-111 and LLC242-282 were closely related to genotype GA5 (Peret *et al*., 1998, 2000). The isolates LLC62-111 and LLC242-282 were more closely related to each other than were the isolates LLC1144-115 and LLC235-267.

The nucleotide and amino acid sequences of the SH, G and F genes of the Singapore isolates compared to the reference strain A2 were analyzed. Zheng *et al*. (1998) have reported that the SH, G and F genes have defined regions of variability and conservation. Both the isolates LLC242-282 and LLC62-111 were almost identical except for three disparities at nt 39, 289 and 894, with the latter two causing amino acid substitutions at positions 97 and 292 of the G protein. Our study of the Singapore isolates localized variable regions at amino acid positions 101–133 and from 181 to the end of the C-terminus of the G protein. A conserved 13-amino acid motif spanning positions 164–176 (Fig. 2) was found between these two hypervariable regions, a finding that corroborated earlier ones (Cane *et al*., 1991; Sullender *et al*., 1991).

Similarly to the G gene, several changes were observed in the F gene of the isolates LLC242-282 and LLC62-111. Of the four changes at nt 56, 711, 993 and 1467 only one resulted in an amino acid alteration at position 19. However, comparison of the Singapore isolates and the reference strain A2 illustrated that the F gene was variable throughout, although the mutations observed were mainly silent (Fig. 3). Majority of the silent mutations were located particularly at the region believed to be crucial for the virus-cell fusion process (aa 137–154), as well as in two reported neutralization epitopes around aa 221–232 and aa 262–268 of the F1 subunit, known to be conserved among HRSV strains (Scopes *et al*., 1990; Cane and Pringle, 1992). Nonetheless, only the putative signal peptide region (aa 1–23) showed the amino acid diversity.

Similarly to earlier studies on the SH gene, a relatively conserved region (nt 1–123) encoding the cytoplasmic and

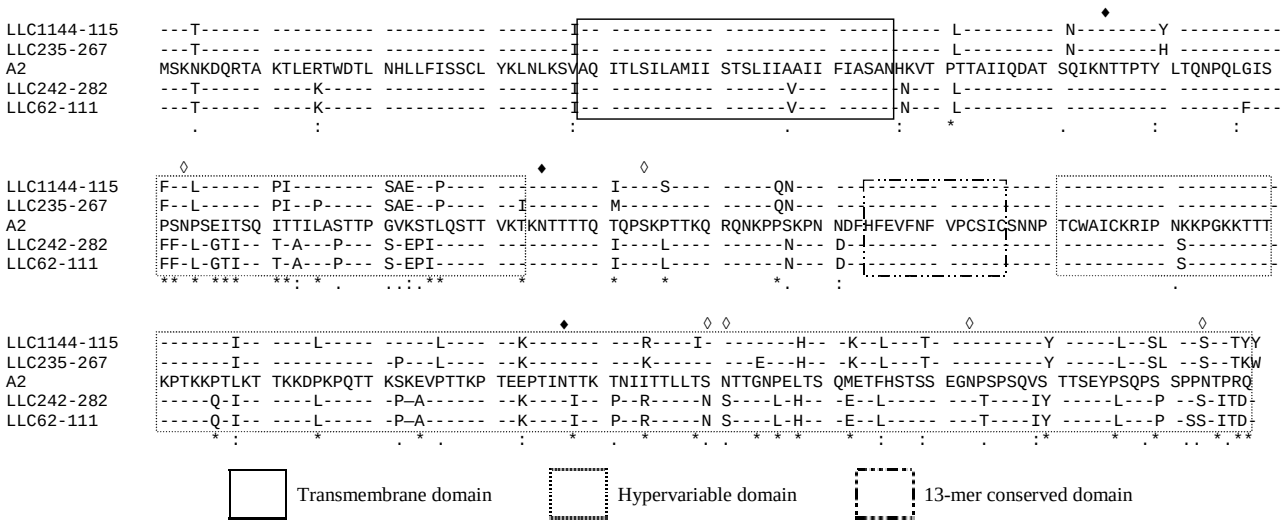

Fig. 2
Multiple alignment of deduced amino acid sequences of the G proteins of the HRSV reference strain A2 (Acc. No. AAB59857) and the four Singapore isolates

Weakly conserved group (.), strongly conserved group (:), non-conserved group (*), conserved *N*-glycosylation sites (◆), non-conserved *N*-glycosylation sites (◇).

transmembrane domains (Cane and Pringle, 1991) was identified in the Singapore isolates. Although the Singapore isolates exhibited several nucleotide differences with respect to the extracellular domain (nt 124–405) of the SH protein compared to strain A2, only three amino acid substitutions were identified, including one weakly and two strongly conserved ones (Fig. 4).

Glycosylation sites of the G protein

An obvious feature found in the G protein of the Singapore isolates and known HRSV strains was the presence of several neighbor threonines. It has been reported earlier that the repetitive sequence K-P-X*n*-T-T-K can occur up to five times within the region spanning aa 193–240, which may be implicated in extensive *O*-glycosylation of the protein (Cane *et al*., 1991). These motifs were found located at positions 193–201, 201–212, 216–221, 221–229 (except LLC1144-115) and 233–240 (except LLC242-282 and LLC62-111). Similarly to earlier studies, the four Singapore isolates showed a fair degree of amino acid conservation at the N-terminus (aa 1–66), while the region spanning aa 67–100 exhibited relative variability. This was followed by a highly variable domain spanning aa 101–133, flanked by two conserved *N*-glycosylation sites (N-X-S/T where X excludes proline) at aa 85 and 135. A third conserved *N*-glycosylation site was located at aa 237. Based on the alignment of various HRSV strains, further six potential non-conserved *N*-glycosylation sites have been proposed earlier (Cane *et al*., 1991). None of the isolates had the potential site at aa 144–146. In addition, the isolates LLC1144-115 and LLC235-267 did not harbor the sites located at aa 250–252 and 273–275, while the site at aa 251–253 was absent from the other isolates. Additional conserved features in this variable domain were two amino acid triplets TTP and TTV at positions 118–120 and 129–131, respectively (Cane *et al*., 1991).

Discussion

Ever since the accidental infection by an agent causing respiratory infection of the co-discoverer, Dr. Blount, it has been then determined that this agent, HRSV (designated in this way because the infected cells fuse together to form syncytia), is also a major human pathogen. Being highly contagious among infants and young children, it had prompted the establishment of the HRSV Protection Program to educate parents and to provide support services. HRSV has a substantial medical and social impact that justifies the importance of epidemiological and prevention studies, which motivated our study. The four Singapore isolates of HRSV belong to the subgroup A due to their high homology with the prototype strain A2. This homology was significant as there was about 43% difference in amino acid sequence between strains of the two subgroups. Subgroup A is predominant, circulates in many communities (Peret *et al*., 1998, 2000; Zheng *et al*.,

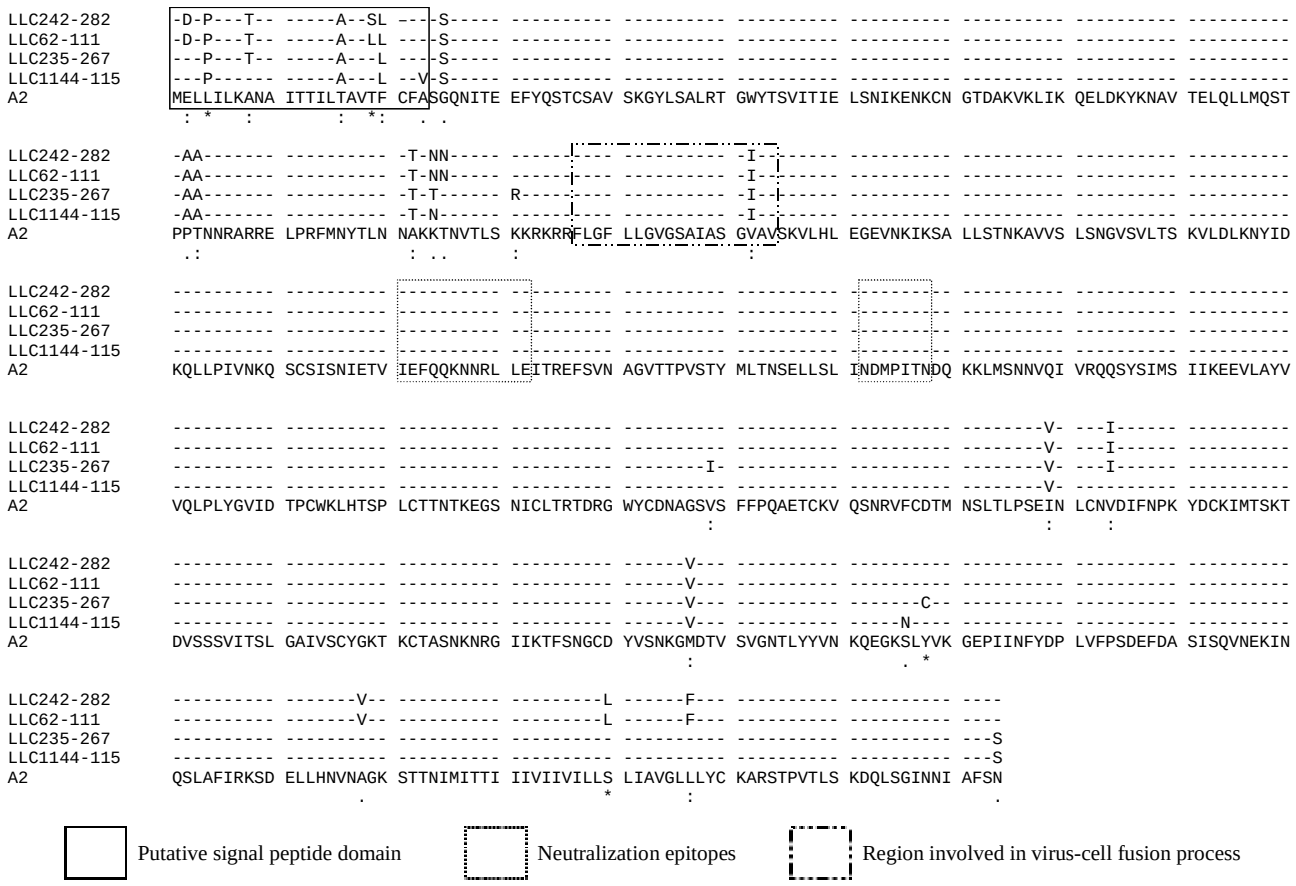

Fig. 3
Multiple alignment of deduced amino acid sequences of the F proteins of HRSV reference strain A2 (Acc. No. AAB59858) and the four Singapore isolates

Weakly conserved group (.), strongly conserved group (:), non-conserved group (*).

1999; Choi and Lee, 2000; Roca *et al*., 2001; Venter *et al*., 2001, 2002) and accounts for more severe cases (Walsh *et al*., 1997), thereby reflecting the virulence of the four Singapore isolates.

Similarly to earlier analyses of known HRSV strains, the G gene of the Singapore isolates was the least conserved compared to the SH and F genes (Johnson *et al*., 1987; Garcia *et al*., 1994). Analysis of both the G and F genes revealed a close relationship between the isolates LLC242-282 and LLC62-111, while the isolates LLC1144-115 and LLC235-267 were similar to each other. Thus, we assume that there have been at least two co-circulating genotypes in Singapore. Compared to the SH and F genes, the unique feature of the G gene is that the similarity between strains at amino acid level is lower than that at nucleotide level, and serves as a good candidate for understanding the phylogeny of HRSV strains. The phylogenetic analysis at nucleotide level included the four Singapore isolates together with 91 subgroup A isolates from Mozambique, South Africa, South Korea, Spain, Uruguay, USA and Canada that clustered within genotypes GA1 to GA 7 and SAA1 (Peret *et al*., 1998, 2000; Venter *et al*., 2001). Two Singapore isolates each clustered with genotypes GA2 and GA5. Interestingly, partial G gene sequences of the Singapore isolate LLC242-282 were identical to that of the Mozambique isolate MOZ/169/99 (Roca *et al*., 2001). Isolation of closely related or even identical genotypes from various countries across the continents suggests that the increasing air travel may enhance the spread and evolution of HRSV strains. Even though HRSV lineages appear to be distributed worldwide, their significance in terms of the degree of virulence and individual and herd immunity have not been completely elucidated (Cane and Pringle, 1992). Recently, Martinello *et al*. (2002) have reported that the GA3 clade is associated with significantly greater severity of illness compared with the clades GA2 and GA4.

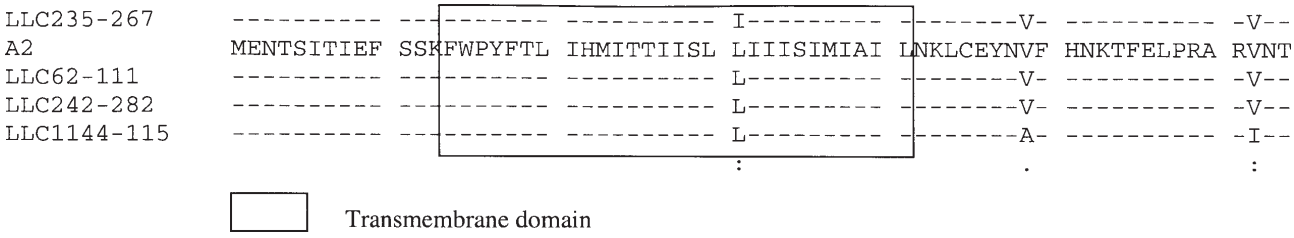

Fig. 4
Multiple alignment of deduced amino acid sequences of the SH proteins of HRSV reference strain A2 (Acc. No. AAB59857) and the four Singapore isolates

Weakly conserved group (.), strongly conserved group (:).

Little is known about the role played by the SH protein in the infection process of HRSV, given especially that it is dispensable for viral infection *in vitro* and *in vivo* (Bukreyev *et al*., 1997). Although several nucleotide substitutions at the extracellular domain were found between the Singapore isolates and the reference strain A2, the amino acid sequences were relatively conserved, showing only three changes. The infectivity of HRSV is largely determined by the glycoproteins G and F, targets for antibodies to HRSV (Taylor *et al*., 1984; Hall *et al*., 1991). Both the G and F proteins facilitate the cell-to-cell transmission of the virus through virus attachment to cells and by promoting fusion of infected cells, respectively. We found out that the F gene of the Singapore isolates and strain A2 displayed several nucleotide disparities. However, the F protein sequences remained highly conserved across the two subgroups, making it an appropriate target for the anti-RSV MAb. A good example of a successful passive RSV immunoprophylactic agent is the Palivizumab (Synagis), a humanized MAb ($IgG1_k$) produced by recombinant DNA technology and directed against an epitope in the antigenic A site of the F protein (Pollack *et al*., 2002). Although there are specific antibodies directed against the G protein, its high variability renders it unsuitable as a vaccination target. The F gene of the Singapore isolates contained several silent mutations that were limited to regions likely to be involved in virus-cell fusion as well as in neutralization. Its conservation in the Singapore isolates and other HRSV strains makes this target a more viable option for vaccine development.

Among the HRSV subgroup A, two stretches of the G protein sequences (297 and 298 aa) have been described (Garcia *et al*., 1994); the Singapore isolates shared the stretch of 298 residues with the A2 reference strain. The G protein is a type II integral membrane protein with a relatively conserved N-terminus (aa 1–38) located inside the viral envelope. In contrast, the ectodomain of 232 aa has two regions of marked sequence variation that contain most of the potential sites for glycosylation. We demonstrated that up to 14% of the G protein sequence diversity could be observed in the Singapore isolates. Both the isolates LLC242-282 and LLC62-111 shared the same pI value of 9.76, while the other two isolates had pI values of 9.90 and 9.88, further emphasizing their relatedness. The native G protein of 80–90 K (Palomo *et al*., 1991) has a high content of both *N*- and *O*-linked oligosaccharides, which may play a significant role in determining the antigenic properties of HRSV. Virus infectivity is sensitive to limited removal of *N*- and *O*-linked oligosaccharides by endoglycosidases, signifying that carbohydrate moieties are required for the G protein function (Lambert, 1988). Whereas only the isolate LLC235-267 possessed the five potential repetitive elements (K-P-X_n-T-T-K) potentially involved in *O*-glycosylation, other isolates showed only four potential sites. For the *N*-glycosylation sites, the Singapore isolates contained the three conserved sites (N-X-S/T) and up to six potential non-conserved sites. The occupancy status of predicted glycosylation sites in the conserved sequence or around the cysteines of the ectodomain is undefined. Nonetheless, crucial to the study of HRSV infectivity is knowledge of the occupancy status particularly of the conserved portion of the ectodomain, which may be involved in ligand interactions with a cellular receptor for the G protein.

Our analysis of the three surface proteins of the Singapore isolates reiterated the existence of distinct variable and conserved regions. Previous studies have also alluded that glycosylation is a probable a determinant of antigenicity (Roca *et al*., 2001). Hence, an in-depth understanding of the glycosylation status and disulfide linkage patterns can help to better formulate a more effective vaccination target (Mader *et al*., 2000). The phylogeny studies as well as the genetic data demonstrated the HRSV sequence variability and genetic relationships of the HRSV isolates from Singapore with those around the world. Together with the prevalence of more than one circulating genotype at any given time, such findings facilitate monitoring of the spread of existing HRSV strains and the emergence of new ones.

Acknowledgement. This study was supported by a grant from the Biomedical Research Council, Singapore.

References

Anderson LJ, Hierholzer JC, Tsou C, Hendry RM, Fernie BF, Stone Y, McIntosh K (1985): Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J. Infect. Dis.* **151**, 626–633.

Bukreyev A, Whitehead SS, Murphy BR, Collins PL (1997): Recombinant respiratory syncytial virus from which the entire SH gene has been deleted grows efficiently in cell culture and exhibit site-specific attenuation in the respiratory tract of the mouse. *J. Virol.* **71**, 8973–8982.

Cane PA, Pringle CR (1991): Respiratory syncytial virus heterogeneity during an epidemic: analysis by limited nucleotide sequencing (SH gene) and restriction mapping (N gene). *J. Gen. Virol.* **72**, 349–357.

Cane PA, Pringle CR (1992): Molecular epidemiology of human respiratory syncytial virus: rapid identification of subgroup A lineages. *J. Virol. Methods* **40**, 297–306.

Cane PA, Matthews DA, Pringle CR (1991): Identification of variable domains of attachment (G) protein of subgroup A respiratory syncytial viruses. *J. Gen. Virol* **72**, 2091–2096.

Chew FT, Doraisingham S, Ling AE, Kumarasinghe G, Lee BW (1998): Seasonal trends of viral respiratory tract infections in the tropics. *Epidemiol. Infect.* **121**, 121–128.

Choi EH, Lee HJ (2000): Genetic diversity and molecular epidemiology of the G protein of subgroups A and B of respiratory syncytial viruses isolated over 9 consecutive epidemics in Korea. *J. Infect. Dis.* **181**, 1547–1556.

Chow VTK, Leong PWF (1999): Complete nucleotide sequence, genomic organization and phylogenetic analysis of a novel genital human papillomavirus type, HLT7474-S. *J. Gen. Virol.* **80**, 2923–2929.

Chungue E, Roche C, Lefevre M, Barbazan P, Chanteau S (1993): Ultra-rapid, simple, sensitive, and economical silica method for extraction of dengue viral RNA from clinical specimens and mosquitoes by reverse transcriptase-polymerase chain reaction. *J. Med. Virol.* **40**, 142–145.

Garcia O, Martin M, Dopazo J, Arbiza J, Frabasile S, Russi J, Hortal M, Perez-Brena P, Martinez I, Garcia-Barreno B, Melero JA (1994): Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. *J. Virol.* **68**, 5448–5459.

Hall CB, Walsh EE, Long CE, Schnabel KC (1991): Immunity to and frequency of reinfection with respiratory syncytial virus. *J. Infect. Dis.* **163**, 693–698.

Johnson PR, Spriggs MK, Olmsted RA, Collins PL (1987): The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proc. Natl. Acad. Sci. USA* **84**, 5625–5629.

Lambert DM (1988): Role of oligosaccharides in the structure and function of respiratory syncytial virus glycoproteins. *Virology* **164**, 458–466.

Mader D, Huang Y, Wang C, Fraser R, Issekutz AC, Stadnyk AW, Anderson R (2000): Liposome encapsulation of a soluble recombinant fragment of the respiratory syncytial virus (RSV) G protein enhances immune protection and reduces lung eosinophilia associated with virus challenge. *Vaccine* **18**, 1110–1117.

Martinello RA, Chen MD, Weibel C, Kahn JS (2002): Correlation between respiratory syncytial virus genotype and severity of illness. *J. Infect. Dis.* **186**, 839–842.

Mufson MA, Orvell C, Rafnar B, Norrby E (1985): Two distinct subtypes of human respiratory syncytial virus. *J. Gen. Virol.* **166**, 2111–2124.

Olmsted RA, Elango N, Prince GA, Murphy BR, Johnson PR, Moss B, Chanock RM, Collins PL (1986): Expression of the F glycoprotein of respiratory syncytial virus by a recombinant vaccinia virus: comparison of the individual contributions of the F and G glycoproteins to host immunity. *Proc. Natl. Acad. Sci. USA* **83**, 7462–7466.

Palomo C, Garcia-Barreno B, Penas C, Melero JA (1991): The G protein of human respiratory syncytial virus: significance of carbohydrate side-chains and the C-terminal end to its antigenicity. *J. Gen. Virol.* **72**, 669–675.

Peret TCT, Hall CB, Schnabel KC, Golub JA, Anderson LJ (1998): Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. *J. Gen. Virol.* **79**, 2221–2229.

Peret TCT, Hall CB, Hammond GW, Piedra PA, Storch GA, Sullender WM, Tsou C, Anderson LJ (2000): Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. *J. Infect. Dis.* **181**, 1891–1896.

Pollack P, Groothuis JR, Barbarotto GM (2002): Development and use of palivizumab (Synagis): a passive immunoprophylactic agent for RSV. *J. Infect. Chemother.* **8**, 201–206.

Roca A, Loscertales MP, Quinto L, Perez-Brena P, Vaz N, Alonso PL, Saiz JC (2001): Genetic variability among group A and B respiratory syncytial viruses in Mozambique: identification of a new cluster of group B isolates. *J. Gen. Virol.* **82**, 103–111.

Scopes GE, Watt PJ, Lambden PR (1990): Identification of a linear epitope on the fusion glycoprotein of respiratory syncytial virus. *J. Gen. Virol.* **71**, 53–59.

Sullender WM, Mufson MA, Anderson LJ, Wertz GW (1991): Genetic diversity of the attachment protein of subgroup B respiratory syncytial viruses. *J. Virol.* **65**, 5425–5434.

Taylor G, Stott EJ, Bew M, Fernie BF, Cote PJ, Collins AP, Hughes M, Jebbett J (1984): Monoconal antibodies protect against respiratory syncytial virus infection in mice. *Immunology* **52**, 137–142.

Venter M, Madhi SA, Tiemessen CT, Schoub BD (2001): Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. *J. Gen. Virol.* **82**, 2117–2124.

Venter M, Collinson M, Schoub BD (2002): Molecular epidemiological analysis of community circulating respiratory syncytial virus in rural South Africa: comparison of viruses and genotypes responsible for different disease manifestations. *J. Med. Virol.* **68**, 452–461.

Walsh EE, McConnochie KM, Long CE, Hall CB (1997): Severity of respiratory syncytial virus infection is related to virus strain. *J. Infect. Dis.* **175**, 814–820.

Zheng H, Storch GA, Zang C, Peret TCT, Park CS, Anderson LJ (1999): Genetic variability in envelope-associated protein genes of closely related groups A strains of respiratory syncytial virus. *Virus Res.* **59**, 89–99.