

Circulation of Human Respiratory Syncytial Virus Strains Among Hospitalized Children with Acute Lower Respiratory Infection in Malaysia

Mohammad Reza Etemadi¹, Zamberi Sekawi¹, Norlijah Othman², Munn-Sann Lye³
and Faezeh Yazdani Moghaddam⁴

¹Department of Medical Microbiology and Parasitology, Faculty of Medicine and Health Sciences, University Putra Malaysia. ²Department of Pediatrics, Faculty of Medicine and Health Sciences, University Putra Malaysia, Serdang, Selangor, Malaysia. ³Department of Community Health, Faculty of Medicine and Health Sciences, University Putra Malaysia. ⁴Department of Biology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran.
Corresponding author email: zamberi@medic.upm.edu.my

Abstract: *Human respiratory syncytial virus* (RSV) is a major viral pathogen associated with acute lower respiratory tract infections (ALRTIs) among hospitalized children. In this study, the genetic diversity of the RSV strains was investigated among nasopharyngeal aspirates (NPA) taken from children less than 5 years of age hospitalized with ALRTIs in Hospital Serdang, Malaysia. A total of 165 NPA samples were tested for the presence of RSV and other respiratory viruses from June until December 2009. RSV was found positive in 83 (50%) of the samples using reverse transcription polymerase chain reaction (RT-PCR). Further classification of 67 RSV strains showed that subgroups A and B comprised 11/67 (16.4%) and 56/67 (83.6%) of the strains, respectively. The second hypervariable region at the carboxyl-terminal of the G gene was amplified and sequenced in order to do phylogenetic study. The phylogenetic relationships of the samples were determined separately for subgroups A and B using neighbor joining (NJ), maximum parsimony (MP), and Bayesian inference (BI). Phylogenetic analysis of the 32 sequenced samples showed that all 9 RSV-A strains were clustered within NA1 genotype while the remaining 23 strains of the RSV-B subgroup could be grouped into a clade consisted of strains with 60-nucleotide duplication region. They were further classified into newly discovered BA10 and BA9 genotypes. The present finding suggests the emergence of RSV genotypes of NA1 and BA. This is the first documentation of the phylogenetic relationship and genetic diversity of RSV strains among hospitalized children diagnosed with ALRTI in Serdang, Malaysia.

Keywords: human respiratory synctial virus (RSV), acute lower respiratory tract infections (ALRTIs), Malaysia

Evolutionary Bioinformatics 2013:9 151–161

doi: [10.4137/EBO.S10999](https://doi.org/10.4137/EBO.S10999)

This article is available from <http://www.la-press.com>.

© the author(s), publisher and licensee Libertas Academica Ltd.

This is an open access article published under the Creative Commons CC-BY-NC 3.0 license.



Introduction

The role of the *human respiratory syncytial virus* (RSV) has been demonstrated as a major viral pathogen associated with acute lower respiratory tract infections (ALRTI) and frequent cause of hospitalization among children less than 5 years.¹⁻⁴ RSV belongs to the genus *Pneumovirus* in the family *Paramyxoviridae*. The genome consists of nonsegmented single-stranded negative sense RNA that encodes 11 separate proteins, 8 of which are structural (L, G, F, N, P, M, M2-1, and SH) and 2 of which are nonstructural (NS1 and NS2, respectively).^{5,6} RSV was initially divided into 2 distinct major subgroups, A and B, using monoclonal antibodies against virus proteins.^{7,8} Nucleotide and amino acid (AA) sequence analysis showed that G protein is significantly more divergent than for other RSV proteins for both within and between major groups.⁹ Extensive nucleotide and AA variability of the G protein suggests that the virus may reinfect people and cause yearly outbreaks through escaping previously induced immunity.^{10,11} Extracellular domain of the membrane-bound G protein contains two hypervariable regions separated by a conserved 13 AA motif among all strains. Nucleotide mutations are preferentially accumulated in these two hypervariable regions of the ectodomain. The second hypervariable region of the G protein ectodomain located at the C terminal region provides a reliable surrogate for the variability of the entire G gene. It has been used for phylogenetical and molecular epidemiological studies of the RSV.^{10,12} Being as a target for neutralizing antibody as well as a major constituent for protective antibody is considered another reason for sequence study of the G protein.¹³ Phylogenetic analysis using G protein gene revealed that distinct genotypes consist of phylogenetically related strains.^{10,12} RSV group A has been clustered into 10 genotypes: GA1-GA7, SAA1, NA1, and NA2.^{10,12,14,15} For the classification of group B, the following genotypes have been proposed: BA1-BA10, GB1-GB4, and SAB1 to SAB3.^{10,12,13,16,17} Determination of circulation pattern of strains is not easy because of continuous shifting in circulating strains, their multiplicity, and cocirculation during a given season.^{14,18-20} In recent years, there has been an increasing interest in the genotype study and phylogenetic analysis of RSV strains. However, far too little attention has been paid to phylogenetic

study of respiratory viruses in Malaysia. In the current study, we evaluated the genetic variability of the second hypervariable region of G protein gene and deduced AA sequence of both RSV A and B subgroups, detected from hospitalized children with acute lower respiratory tract infections (ALRTI) in Serdang, Malaysia.

Materials and Methods

Clinical samples

Nasopharyngeal aspirates (NPA) were taken from children more than 1 month of age and less than 5 years of age who were admitted from the emergency department to the pediatric wards at Hospital Serdang with the diagnosis of ALRTI within 24 hours of admission from June 16 through December 1, 2009. All potential subjects (including caregivers of the patients) were briefed on the study before written informed consent was obtained. The final selection of the subjects into the study was based on predetermined inclusion and exclusion criteria. Approval from the following authorities was obtained prior to the embarkation of the study: the Medical Research and Ethics committee (MREC) of the Ministry of Health Malaysia and the Medical Research Ethics Committee, University Putra Malaysia. NPAs were transported to the laboratory and refrigerated at 4 °C to 8 °C. In order to avoid repeated freezing and thawing, all NPAs were processed upon receipt. Demographic, clinical features, hospital course, and chest X-ray findings were collected from patients; however, they were not analyzed here.

RSV detection

The D3 Ultra 8 Direct Immunofluorescence Assay (DFA) Respiratory Virus Screening and Identification kit (Diagnostic Hybrids Inc. [DHI], USA) was used as a first step of detection, which made it possible to detect 8 common viruses including RSV, human metapneumovirus (HMPV), influenza virus (IFV) types A and B, parainfluenzavirus 1-3 (PIV1-3), and human adenovirus (HAdV) using direct immunofluorescence assay (DFA). Positive samples for RSV by DFA were further classified into A and B groups using multiplex reverse transcription polymerase chain reaction (RT-PCR).

RNA extraction and cDNA synthesis

Viral genome was extracted from 0.4 mL of each NPA supernatant by MagMAX Viral RNA Isolation Kit



(Applied Biosystems, California, USA) according to the manufacturer's instructions. Reverse transcription was performed on RNA extracts in a final volume of 20 μ L by random hexamer primer using RevertAid H Minus First Strand cDNA synthesis kit (Fermentase, Hanover, MD, USA). The samples were incubated first for 5 minutes at 25 °C and followed by 60 minutes at 42 °C. The reaction was terminated by heating at 70 °C for 5 minutes. Seminested multiplex RT-PCR was applied for the detection of a panel of respiratory viruses including RSV, HMPV, IFV types A and B, PIV1-4, and human coronaviruses (HCoV) OC43 and 229E as previously reported. An internal control consisting of glyceraldehydes-3-phosphate dehydrogenase (GAPDH) was included.²¹ The presence of human bocavirus (HBoV) and HAdV were individually investigated in the samples by singleplex PCR and nested PCR respectively.^{22,23} Due to financial constraints, the samples with RSV as a sole agent detected by both methods of DFA and RT-PCR were randomly chosen and proceeded for genotyping study. The second hypervariable region at the carboxy-terminal of G gene of subgroups A and B isolates was amplified and sequenced using primer sets GPA/F1 and GPB/F1, complementary to the G and F gene nucleotide sequences, respectively.¹² The GPA and GPB primers correspond to the bases 511–530 and 494–515 of the G gene of RSV-A prototype strains A2 and RSV-B prototype strain of 18537, respectively. The F primer corresponds to bases 3–22 of the prototype strains. The PCR was performed in following cycling parameters: 94 °C for 10 minutes denaturation followed by 35 cycles of 94 °C for 50 seconds, annealing at 48 °C for 30 seconds for GPA/F1 and at 52 °C for 30 seconds for GPB/F1 primer sets, 72 °C for 45 seconds, with a final extension of 72 °C for 10 minutes. The PCR products were fractionated by electrophoresis of 2.5% agarose gel and visualized using ethidium bromide under UV light.

Sequencing and alignments

PCR purification was carried when the fragment of expected size was obtained using QIAquick Gel Extraction Kit (QIAGEN, USA), as per manufacturer's instruction. Both strands of PCR products were sequenced using the same forward and reverse RSV primer set with ABI 3730 xl DNA Analyzer (Applied Biosystems) commercially (Repfon Glamor Sdn Bhd,

Serdang, Malaysia). Nucleotide sequences of RSV subgroups A and B were assembled separately and manually edited with Bioedit version 7.0.1. Clustal X version 3.06 program of MEGA software, version 4.0 was used for alignment of the sequences.²⁴ Unique sequences for both A and B subgroup viruses representative of known genotypes obtained from GenBank were included in the phylogenetic analysis. A total of 34 partial sequences (270 nucleotide long) of RSV A and 55 sequences (270 and 330 nucleotide long) of RSV B, previously assigned to specific genotypes, were downloaded from GeneBank for comparative studies. Prototype strains Long (Accession number M17212) for subgroup A and CH18537 (Accession number M17213) for subgroup B were included as outgroups in the phylogenetic analysis.

Phylogenetic analysis

MEGA5 software was used for phylogenetic analyses using maximum parsimony and neighbor joining methods.²⁵ Genetic distances within/among genotypes were also calculated separately for subgroups A and B based on the sequences used to construct the phylogenetic tree at nucleotide acid level by Kimura's 2-parameter (K2P) models.²⁶ For Bayesian analysis, Modeltest 3.7²⁷ was used to determine the optimal model of nucleotide evolution. The (GTR+G) substitution model was selected using an Akaike Information Criterion as implemented in Modeltest. Bayesian Inference (BI) was carried out with MrBayes 3.0²⁸ to calculate posterior probabilities of recovered clades, with the optimal model of sequence evolution determined from the likelihood ratio test (LRT)s. MrBayes 3.0 was run with 8×10^6 generation Markov chain with trees sampled every 100th generation (resulting in 10,000 trees) using default priors. The analyses began on a random starting tree. We discarded the first 5000 trees as a conservative "burn-in," and the posterior probability values were calculated from the remaining trees. Stationarity was assumed when the cumulative posterior probabilities of all clades were stabilized. A cluster was defined as a bootstrap value of more than 50%. The deduced AA sequences and polymorphisms of the second hypervariable region of subgroup A and B were compared with those of the prototype long strain spanning 270 bp and AY333364 strain (BA strain spanning 330 nucleotide), respectively with MEGA4 software according



to the standard genetic codes. The sequences generated in this study have been submitted to GenBank under accession numbers JQ933942–JQ933973.

Results

Occurrence of RSV

A total of 165 children < 5 years of age who fulfilled the inclusion criteria were enrolled and screened for the presence of RSV and other respiratory viruses in the study. RSV was detected in 67 (41%) and 83 (50.3%) of the samples using DFA and RT-PCR, respectively. Using RT-PCR as detection method, RSV was the sole viral agent in 49 (29.7%) of the samples, while 34 (20.6%) samples were coinfecting with other viruses including double infections in 29 (17.6%) cases and triple infections in 5 (3.0%) cases. The most prevalent combinations were related to RSV with HAdV (14/34 or 41%) and with HRV (13/34 or 38%). Other viruses including HMPV, HCoV-OC43, IFV-A, and HBoV were codetected in the rest of the samples. RSV positive samples using DFA were further divided into separate antigenic subgroups A and B. Of 67 RSV positive samples with DFA, 56 (83.6%) of the samples categorized as subgroup B and 11 (16.4%), as subgroup A. Thirty-two of 49 samples (9 RSV-A and 23 RSV-B) were RSV was detected as a sole agent were randomly chosen and partial sequences were obtained using GPA/F1 and GPB/F1 primers.

Phylogenetic reconstruction

RSV group A strains

The results of model-based methods, Bayesian analyses are largely congruent with NJ and MP analyses for RSV-A and B. We obtained sequence data totaling 270 base pairs for RSV-A and 43 individuals. Of these, 135 characters were constant, and 82 were parsimony informative. MP searches recovered 115 equally most likelihood trees with a consistency index of 0.63 and retention index of 0.87. Phylogenetic analysis showed that all the RSV-A strains were clustered into the recently described NA1 genotype, supported with high posterior probability and high bootstrapping values in maximum parsimony and neighbor-joining (100%, 96% and 98%, respectively) (Fig. 1).

RSV subgroup B strains

Relationships of all taxa derived from MP, NJ and partitioned Bayesian analyses of sequences were

nearly identical. We obtained sequence data totaling 330 base pairs for RSV-B and 78 individuals. Of these, 151 characters were constant, and 112 were parsimony informative. MP searches recovered 69 equally most likelihood trees with a consistency index of 0.56 and retention index of 0.83.

All the University Putra Malaysia (UPM) strains as well as representative strains with duplicate region were clustered in a clad with bootstrapping value of 86% in a Bayesian analyses. The UPM strains were classified into 3 different genotypes including 19 strains that belonged to the BA10, supported by 100% bootstrapping value, 3 strains belonged to BA9 with 100% bootstrapping value in a Bayesian analysis and one strain that more likely belonged to BAIV genotype (Fig. 2).

Nucleotide and amino acid analysis of subgroup A strains

Nine of 11 group A samples were retrieved in sequencing. The average p-distance within the UPM subgroup A strains was 0.01, and it ranged 0.0 to 0.058 within the representatives of other group A genotypes (Table 1). The average intergenotype p-distances between UPM strains and the other genotypes ranged from 0.016 to 0.214 (Table 2). Inferred AA sequences of the second variable region of carboxyl terminal of the G protein for nine RSV-A samples were compared with prototype Long strain (Fig. 3). The results showed that all RSV-A strains included in this analysis exhibit changes in the stop codon position (Gln298 stop) and resulted in predicted G protein of 297 amino acids. Several other AA substitutions occurred in other parts of the studied region is compared with Long strain. AA substitutions specific for NA1 genotype including Asp (D) 237 (except for JQ933945 which was N), Leu (L) 274 (except for JQ933943 which was P), and Ser (S) 292 were demonstrated for these strains. Similar to the GA2 genotype, AA substitutions at Thr (T) 269 and Ser (S) 289, previously assigned as GA2 specific amino acids, were observed in our samples.

Nucleotide and amino acid analysis of group B strains

The sequences of the 23 subgroup B strains detected from the patients in this study were included in the phylogenetic analysis. The results showed that all the strains of BA genotypes consisted of a 60-nucleotide

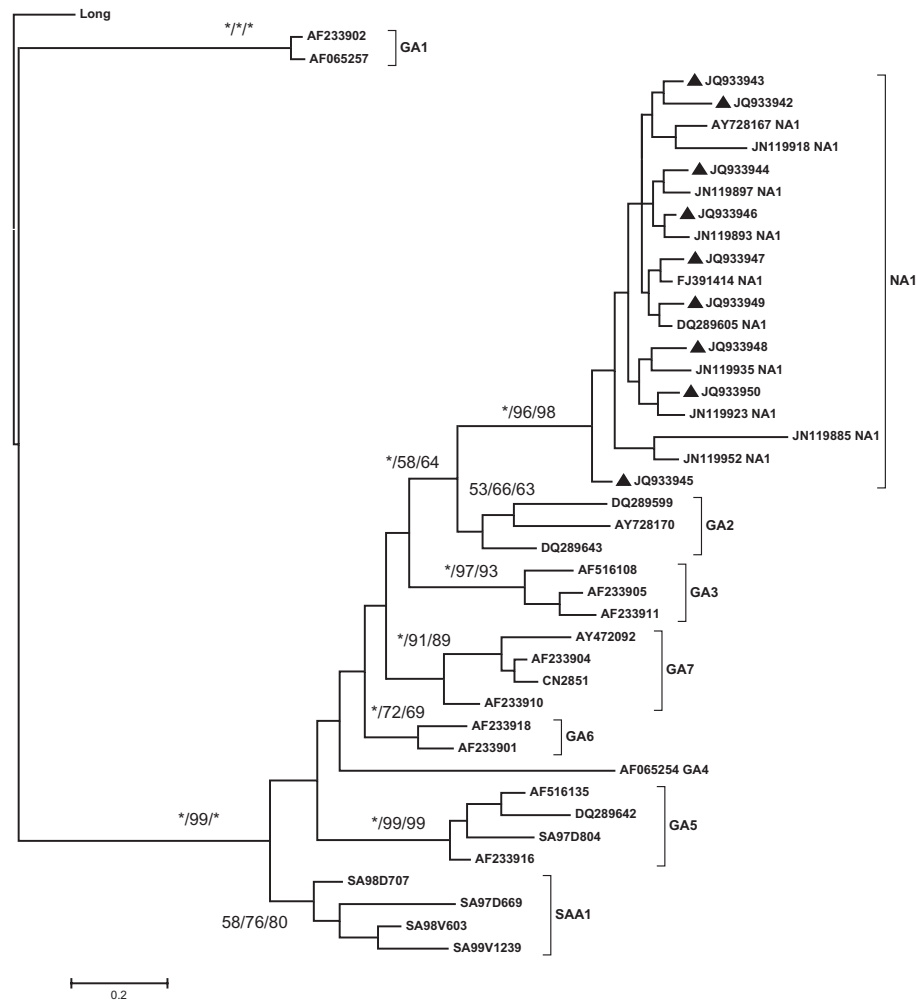


Figure 1. Phylogenetic tree for RSV group A nucleotide sequences based on the second variable region of the G protein (270 bp) constructed by the Bayesian analysis method using MrBayes 3.0 software.

Notes: Nodes represent posterior probability/maximum parsimony/neighbor-joining bootstrap values. Numbers at the branches are bootstrap values determined for 1000 replicates. A bootstrap value of more than 50% was depicted in the tree. The representative sequences for each genotype (GA1-7, SAA1, and NA1) were retrieved from GenBank and included in the tree. UPM strains are indicated in dark triangles. Strain Long was used as the outgroup sequences. The asterisks represent bootstrap value equal 100%. The genotypes are shown on the right.

duplication region at the second hypervariable region of the G protein gene. The average p-distance within the UPM strains was 0.03 and ranged from 0.002 to 0.042 within the representatives of other group B genotypes (Table 3). The average intergenotype p-distance between UPM strains and other genotypes ranged 0.023 to 0.174 (Table 4). Deduced G protein length and AA alignment of second hypervariable region and stop codons of the strains are shown in Figure 4 for RSV B strains in comparison with the representative sequences. All strains had predicted length of 312 a.a except for 3 sequences (JQ933955, JQ933961, and JQ933962) with predicted length of 319 a.a. In the alignment of deduced a.a sequences with the duplicated segment, AA substitutions were

observed in duplicated region as well as in the 60 nucleotide segment exactly upstream. Residue Thr 229 located outside the duplicated region was changed to the Ile. This change was not observed in other retrieved GenBank strains without duplicated region. All the BA10 strains contained E292G substitution. Strain JQ933961 showed E226D substitution which was not observed in the other BA10 strains.

Discussion

The present study was designed to provide a preliminary report on the genetic diversity and molecular epidemiology of RSV detected from children hospitalized with ALRTIs in Malaysia, and to date no similar study has been conducted in this country. Previously

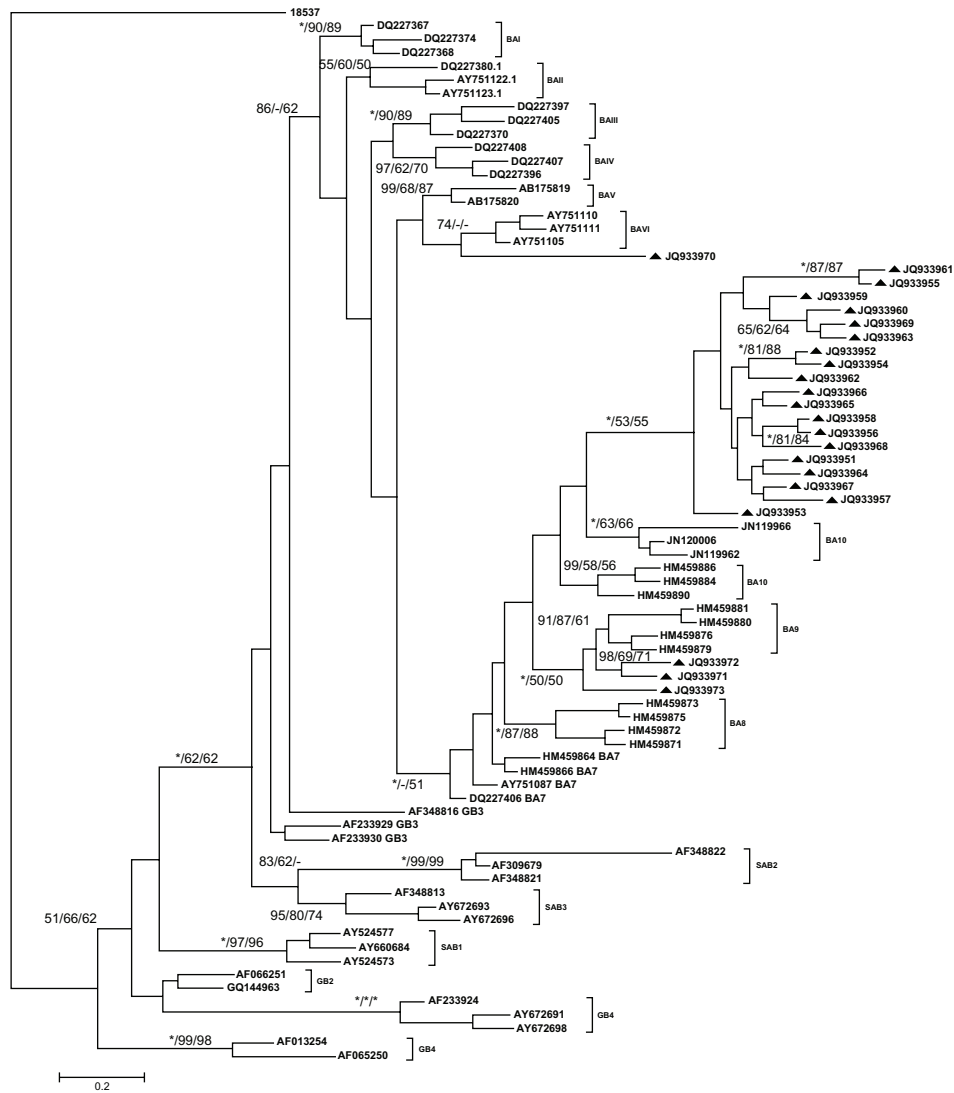


Figure 2. Phylogenetic tree for RSV group B nucleotide sequences based on the second variable region of the G protein (270 and 330 bp) constructed by the Bayesian analysis method using MrBayes 3.0 software.

Notes: Nodes represent posterior probability/maximum parsimony/neighbor-joining bootstrap values. Numbers at the branches are bootstrap values determined for 1000 replicates. A bootstrap value of more than 50% was depicted in the tree. The representative sequences for each genotype (GB1-4, SAB1-3, and BA1-10) were retrieved from GenBank and included in the tree. UPM strains are indicated in dark triangles. Strain CH18537 was used as the outgroup sequences. The asterisks represent bootstrap value equal 100%. Strain CH18537 was used as the outgroup sequences. The genotypes are shown on the right.

Table 1. RSV-A intragenotype pairwise distances from variation in G gene derived from the Kimura's 2-parameter methods.

UPM	0.01
NA1	0.021
GA2	0.058
GA3	0.022
GA7	0.026
GA6	0.023
SAA1	0.034
GA5	0.029
GA1	0

published findings have frequently confirmed RSV to be the major viral pathogen associated with ALRTI in children.¹⁻³ RSV was the major reason for hospitalization of children with ALRTI in this study as it represented 50% of the NPA samples tested positive using RT-PCR of which 30% of the cases were single infections. This study is in agreement with previous local studies by Chan²⁹ and Zamberi,³⁰ which showed RSV as a major causative agent in Malaysian children. The results further support that RSV infection is a frequent cause of hospitalization among

**Table 2.** RSV-A intergenotype pairwise distances from variation in G gene derived from the Kimura's 2-parameter methods.

	1	2	3	4	5	6	7	8	9	10	11
1. UPM											
2. NA1	0.016										
3. GA2	0.08	0.085									
4. GA3	0.087	0.091	0.089								
5. GA7	0.107	0.111	0.087	0.085							
6. GA6	0.088	0.092	0.074	0.065							
7. SAA1	0.108	0.115	0.097	0.098	0.087	0.063					
8. GA5	0.133	0.139	0.125	0.122	0.109	0.088	0.101				
9. GA4	0.16	0.162	0.151	0.135	0.125	0.12	0.123	0.15			
10. GA1	0.214	0.213	0.139	0.191	0.189	0.177	0.19	0.239	0.24		
11. Outgroup	0.165	0.17	0.158	0.156	0.146	0.133	0.127	0.162	0.171	0	

children in tropical and developing countries.⁴ In the majority of the RSV studies in developing countries, the prevalence of RSV associated with ALRTI was underestimated due to the application of less sensitive detection methods such as ELISA or immunofluorescence assays.³¹ During the 6-month study from May to November, a continuous activity of RSV was seen with a small peak in July, and then RSV activity started to increase from August to October with a relatively big peak in October. The finding seems to be consistent with RSV outbreaks during the wet season in tropical and subtropical countries with seasonal rainfall.³² However, this study only presents a small picture of the circulation pattern of the RSV genotypes in Malaysia. For a more representative result, a study spanning several years is required since variation in prevalence from year to year is common in respiratory infections.

RSV strains were further classified into A and B using RT-PCR and further sequencing of the partial G protein. During the study period, both antigenic subgroups of A and B cocirculated with predominance of subgroup B.

These results differ from some published research that shows the predominance of subgroup A virus among the strains.^{19,20} However the results of these studies are consistent with those of other published studies that show RSV-B as the dominant group. On the other hand, Gilca (2006) also found rapid changes from A to B during 2 consecutive seasons. This suggests that there is a variation in the pattern of RSV circulation in different regions and seasons.

RSV subgroup B viruses are generally divided into the genotypes comprising nonduplicated regions (GB1-4 and SAB1-3) and genotypes with a duplicated segment namely BA genotype. Since the report of BA genotypes with 60-nucleotide insertions in the second hypervariable region of G gene by Trento,¹⁷ the partial sequences of this genotype has been reported worldwide.^{19,33-36} Results of different tree methods for RSV-B confirmed the K2P distance method and defined that the least mean intergenotype distance was found between genotypes with duplicated regions and UPM strains and the most mean intergenotype distance belonged to genotypes that were not duplicated. The

```

.....221.....231.....241.....251.....261.....271.....281.....291.....298
Long      KKDHPQTTK PKEVPTTKPT EEPPTINTTKT NIITPLLTNN TTGNPKLTSQ METFPHSTSE GNLSPSQVST TSEHPSQPSS FPNTTRQ*-- [298aa]Long
JN119918  ...P..... ..L..... .K...D.... .R.....S. ....EH... E..L...T... ..P....Y. ...YL..YP. SS...K*-- [297aa]NA1
FJ391414  ...P..... ..L..... .K...D.... .R.....S. ....EH... E..L...T... ..P....Y. ...YL..SP. SS...K*.SL [297aa]NA1
JQ933943  ...P..... ..L..... .K...D.... .R.....S. ....EH... E..L...T... ..P....Y. ...YL..SP. SS...K*-- [297aa]UPM
JQ933942  ...P..... ..L..... .K...D.... .K....SH. .I...EH... E..L...T... ..P....Y. ...YL..SP. SS...K*-- [297aa]UPM
JQ933943  ...P..... ..L..... .K...D.... .R.....S. ....EH... E..L...T... ..Y....Y. ...YL..SP. SS...K*-- [297aa]UPM
JQ933945  ...P..... ..L..... .K...D.... .R.....S. ....EH... E..L...T... ..Y....Y. ...YL..SP. SS...K*-- [297aa]UPM
JQ933946  ...P..... ..L..... .K...D.... .R.....S. ....EH... E..L...T... ..Y....Y. ...YL..SP. SS...K*-- [297aa]UPM
JQ933947  ...P..... ..L..... .K...D.... .R.....S. ....EH... E..L...T... ..Y....Y. ...YL..SP. SS...K*-- [297aa]UPM
JQ933948  ...P..... ..L..... .K...D.... .R.....S. ....EH... E..L...T... ..Y....Y. ...YL..HSP. SS...E*-- [297aa]UPM
JQ933949  ...P..... ..L..... .KS...D.... .R.....S. ....EH... E..L...T... ..Y....Y. ...YL..SP. SS...K*-- [297aa]UPM
JQ933950  ...P..... ..L..... .K...D.... .R.....S. ....EH... E..L...T... ..Y....Y. ...YL..SP. SS...K*-- [297aa]UPM
DQ289599  ...L..... ..L..... .K...D.... .R.....S. ....EH... E..L...T... ..Q....Y. ...YL..SL. .S...Y*-- [297aa]GA2
AY728170  ...L..... ..L.....K. KQ..... .R.....S. ....EH... E..L...T... ..P....Y. ...Y....SL. .S...YY-- [298aa]GA2

```

Figure 3. Deduced AA alignment of the second hypervariable region of the G protein gene of RSV A for 9 UPM detected strains and 5 retrieved GeneBank sequences.

Notes: The amino acids shown correspond to the second variable region of the G protein strain Long for the subgroup A. Identical residues are shown by dots. Stop codons are indicated by asterisks. Length numbers indicate the position of residues relative to positions 222 to 298 of the Long strain.



Table 3. RSV-B intragenotype pairwise distances from variation in G gene derived from the Kimura's 2-parameter methods.

BA1	0.011
BA2	0.019
BA3	0.019
BA4	0.017
BA5	0.008
BA6	0.005
BA10	0.026
BA9	0.012
BA7	0.002
BA8	0.02
UPM	0.03
GB1	0.025
GB2	0.02
GB3	0.034
GB4	0.022
SAB1	0.016
SAB2	0.042
SAB3	0.022

detection of the BA strains in this study further supports the widespread dissemination of the BA strains in the community. It was speculated that all BA genotypes with duplicated regions shared a common ancestor that was separate from other group B genotypes.¹³ It was classified into BA-I to BA-VI using G protein gene analysis. Rapid generation of the genetic variation among BA strains is partly related to the immunological selection at duplicated regions in the ectodomain of the G protein. The fast dissemination of genotypes with duplicated regions may be related to the fact that the human immune system may not be exposed to these new genotypes.¹³

In a detailed phylogenetic analysis of BA strains by Dapat in Japan,³⁷ additional new BA genotypes including BA7, BA8, BA9, and BA10 were introduced that were previously described as BA4 genotypes. These newly introduced genotypes comprised the majority of the RSV strains from 2005 to 2010. The different methods of phylogenetic trees largely supported that UPM B strains were separated into 2 different clades, namely BA9 and BA10 genotypes. These findings of the current study are consistent with the Dapat study, which found cocirculation of several BA genotypes in all studied periods with domination of 1 genotype in each season. The present finding seems to be also consistent with other research that found BA genotypes are replacing the previously

Table 4. RSV-B intergenotype pairwise distances from variation in G gene derived from the Kimura's 2-parameter methods.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1. BA1																				
2. BA2	0.035																			
3. BA3	0.034	0.029																		
4. BA4	0.043	0.038	0.037																	
5. BA5	0.035	0.034	0.032	0.041																
6. BA6	0.028	0.028	0.026	0.035	0.011															
7. BA10	0.055	0.054	0.05	0.05	0.046	0.04														
8. BA9	0.046	0.047	0.042	0.039	0.038	0.032	0.032													
9. BA7	0.036	0.034	0.029	0.036	0.026	0.02	0.024	0.017												
10. BA8	0.065	0.063	0.057	0.063	0.049	0.043	0.051	0.044	0.027											
11. UPM	0.058	0.054	0.05	0.055	0.047	0.04	0.039	0.037	0.03	0.057										
12. GB1	0.094	0.092	0.094	0.111	0.093	0.085	0.105	0.094	0.082	0.117	0.108									
13. GB2	0.056	0.05	0.054	0.067	0.054	0.048	0.081	0.072	0.059	0.088	0.081	0.07								
14. GB3	0.035	0.042	0.041	0.049	0.043	0.036	0.062	0.052	0.043	0.07	0.066	0.089	0.054							
15. GB4	0.149	0.145	0.142	0.167	0.152	0.148	0.183	0.171	0.156	0.2	0.174	0.152	0.101	0.142						
16. SAB1	0.081	0.072	0.08	0.084	0.083	0.075	0.105	0.088	0.083	0.115	0.104	0.093	0.051	0.079	0.131					
17. SAB2	0.101	0.109	0.107	0.118	0.109	0.101	0.13	0.122	0.109	0.137	0.117	0.137	0.097	0.097	0.161	0.131				
18. SAB3	0.054	0.051	0.054	0.063	0.052	0.046	0.074	0.065	0.053	0.083	0.076	0.106	0.071	0.052	0.155	0.097	0.098			
19. Outgroup	0.158	0.144	0.142	0.164	0.166	0.157	0.17	0.148	0.154	0.195	0.167	0.158	0.125	0.155	0.241	0.159	0.218	0.173		

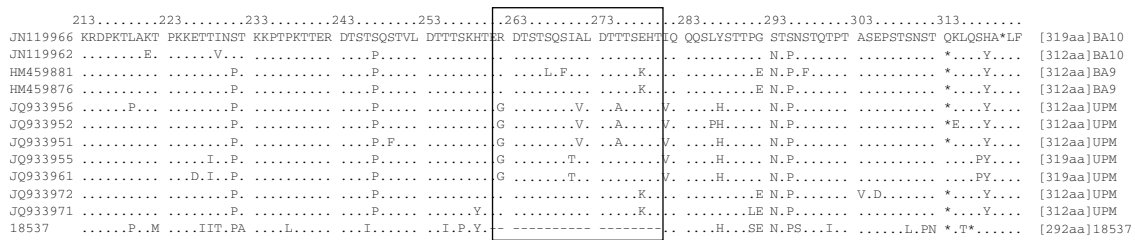


Figure 4. Deduced AA alignment of the second hypervariable region of the G protein gene of RSV B for 7 representative UPM strains and 5 retrieved GeneBank sequences.

Notes: The amino acids shown correspond to the second variable region of the G protein strain BA prototype strain JN119966 for the group B. Identical residues are shown by dots. Stop codons are indicated by asterisks. Length numbers indicate the position of residues relative to positions 213 to 315 of the BA prototype strain JN119966. Box denotes the 60 nt duplication in BA strains.

dominant RSV B genotypes.^{13,35} Among all RSV-B sequenced strains, residue Thr229, located outside the duplicated portion was changed to the Ile. This change was not observed in other retrieved GenBank strains without duplicated regions. This finding is in agreement with Trento et al¹³ who showed that T229I change occurred in all BA strains with duplicated regions. Consistence with that study, the S247P AA substitution was found in all B strains. The strains JQ933971, JQ933972, and JQ933973, which were classified in clad-included BA9 strains, had a substitution of V271A. Except for strain JQ933953, the rest of the strains clustered in the BA10 had a E292G change. Also E226D change was seen in JQ933961. The results are consistent with the study by Dapat.³⁷ However, other changes including T289I and S269P were not observed in this study. G protein length diversity due to the position of the final stop codon was observed among RSV B strains, which leads to a deduced G protein length of 312 or 319 amino acids.

The novel emerging of NA1 and NA2 genotypes was reported in 2009.¹⁴ These 2 recently reported genotypes were recognized as variants of GA2. In spite of genetic relatedness between these novel genotypes, they are different antigenically.¹⁴ Different methods of phylogenetic tree (BI, MP, and NJ) suggested that the RSV-A strains detected in this study were clustered within the NA1 genotype. There is remarkably good agreement between the result of pairwise distance methods and molecular trees of different methods (BI, MP, and NJ). There is a remarkable match between the molecular phylogeny and genetic distance among UPM subgroup A and NA1 genotype in this study. However, all RSV-A strains in this study showed a 297 AA in length.

Circulation of different genotypes of RSV groups A and B in the study emphasize further epidemiological and clinical surveillance of RSV in Malaysia. Host immune system evading and reinfections by RSV could be due to the within and between subgroup antigenic variations.¹⁴ The presence of emerging viruses and the lack of relative protective immunity to the new strains may facilitate effective spread of these new strains.¹² The variation in the G protein length may contribute to the antigenic changes and escape from the immune system.³⁸ However, further studies are required to determine the role of the G protein length on RSV immunity. The findings obtained in this study have important implications for the understanding of the evolutionary pattern of RSV as well as vaccine development. Previous studies have shown that genotype shifting during consecutive years is a common feature of the molecular epidemiology of RSV infections.¹² However, caution must be taken as study at multiple locations and for several consecutive years with a large sample size is required to better appreciate the epidemiology of the RSV. Long-term epidemiological surveillance of RSV strains is important to illuminate antigenic and genetic bases of the infections by RSV.

In conclusion, our study is the first report of the genetic diversity of the RSV strains according to the G gene second hypervariable region during a period of 7 months in Malaysia. Both group A and B subgroups cocirculated in the studied patients, although group B dominated. The BA strains, with 60-nucleotide duplicated region, were identified as the dominant genotype in the samples. The majority of BA strains comprised recently emerging BA10 genotype. However, further long-term surveillance during several epidemic seasons are required to understand the



genetic variability and circulation pattern of the RSV strains in the Malaysia.

Acknowledgment and Funding

This work was partially supported by grants from the Ministry of Science, Technology and Innovation Malaysia (grant number 5450401). We are grateful to all the doctors and nurses of Pediatrics Department, Hospital Serdang for providing nasopharyngeal aspirates for the study.

Author Contributions

MRE performed laboratory methods, analyzed the data and wrote the first draft of the manuscript. ZS as a corresponding author designed the study and verified all the data. NO was involved in obtaining ethics approval from Hospital Serdang and the Ministry of Health Malaysia and conducted the study design. MSL conducted study design. FYM involved in phylogenetic analysis design and data interpretation. All authors have read and approved the final manuscript.

Competing Interests

Author(s) disclose no potential conflicts of interest.

Disclosures and Ethics

As a requirement of publication the authors have provided signed confirmation of their compliance with ethical and legal obligations including but not limited to compliance with ICMJE authorship and competing interests guidelines, that the article is neither under consideration for publication nor published elsewhere, of their compliance with legal and ethical guidelines concerning human and animal research participants (if applicable), and that permission has been obtained for reproduction of any copyrighted material. This article was subject to blind, independent, expert peer review. The reviewers reported no competing interests.

References

- Grimwood K, Cohet C, Rich FJ, et al. Risk factors for respiratory syncytial virus bronchiolitis hospital admission in New Zealand. *Epidemiol Infect.* 2008;136:1333–41.
- Hall CB, Weinberg GA, Iwane MK, Blumkin AK, Edwards KM, Staat MA, et al. The burden of respiratory syncytial virus infection in young children. *N Engl J Med.* 2009;360:588–98.
- Richard N, Komurian-Pradel F, Javouhey E, et al. The impact of dual viral infection in infants admitted to a pediatric intensive care unit associated with severe bronchiolitis. *Pediatr Infect Dis J.* 2008;27:213–7.
- Weber M, Hartley J, Ashworth M, Malone M, Sebire N. Virological investigations in sudden unexpected deaths in infancy (SUDI). *Forensic Sci Med Pathol.* 1998;6:261–7.
- Peter LC. Respiratory Syncytial Virus and Metapneumovirus. In: Knipe DM, editor. *Fields Virology*, Vol. 1. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2007:1603–45.
- Prober CG. Respiratory Syncytial Virus. In: Long SS, editor. *Principles and Practice of Pediatric Infectious Disease*, 2nd ed. Philadelphia, PA: Churchill Livingstone; 2003:1645.
- Mufson MA, Orvell C, Rafnar B, Norrby E. Two distinct subtypes of human respiratory syncytial virus. *J Gen Virol.* 1985;66:2111–24.
- Anderson LJ, Hierholzer JC, Tsou C, et al. Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J Infect Dis.* 1985;151:626–33.
- Johnson PR, Spriggs MK, Olmsted RA, Collins PL. The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proc Natl Acad Sci U S A.* 1987;84:5625–9.
- Peret TCT, Hall CB, Hammond GW, et al. Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. *J Infect Dis.* 2000;181:1891–6.
- Botosso VF, Paolo MA, Ueda M, et al. Positive selection results in frequent reversible amino acid replacements in the G protein gene of human respiratory syncytial virus. *PLoS Pathog.* 2009;5:e1000254.
- Peret TC, Hall CB, Schnabel KC, Golub JA, Anderson LJ. Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. *J Gen Virol.* 1998;79:2221–9.
- Trento A, Viegas M, Galiano M, et al. Natural History of Human Respiratory Syncytial Virus Inferred from Phylogenetic Analysis of the Attachment (G) Glycoprotein with a 60-Nucleotide Duplication. *J Virol.* 2006;80:975–84.
- Shobugawa Y, Saito R, Sano Y, et al. Emerging Genotypes of Human Respiratory Syncytial Virus Subgroup A among Patients in Japan. *J Clin Microbiol.* 2009;47:2475–82.
- Venter M, Madhi SA, Tiemessen CT, Schoub BD. 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.* 2001;82:2117–24.
- Dapat IC, Shobugawa Y, Sano Y, et al. New genotypes within respiratory syncytial virus (RSV) Group B genotype BA in Niigata, Japan. *J Clin Microbiol.* 2010;48(9):3423–7.
- Trento A, Galiano MN, Videla C, et al. Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. *J Gen Virol.* 2003;84:3115–20.
- Arnott A, Vong S, Mardy S, Chu S, Naughtin M, Sovann L, et al. A study of the genetic variability of human respiratory syncytial virus (HRSV) in Cambodia reveals the existence of a new HRSV group B genotype. *J Clin Microbiol.* 2011;49(10):3504–13.
- Zhang Y, Xu W, Shen K, et al. Genetic variability of group A and B human respiratory syncytial viruses isolated from 3 provinces in China. *Arch Virol.* 2007;152:1425–34.
- Zhang Z, Du L, Chen X, et al. Genetic variability of respiratory syncytial viruses (RSV) prevalent in Southwestern China from 2006 to 2009: emergence of subgroup B and A RSV as dominant strains. *J Clin Microbiol.* 2010;48:1201–7.
- Bellau-Pujol S, Vabret A, Legrand L, et al. Development of three multiplex RT-PCR assays for the detection of 12 respiratory RNA viruses. *J Virol Methods.* 2005;126:53–63.
- Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A, Andersson B. Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proc Natl Acad Sci U S A.* 2005;102:12891–6.



23. Lu X, Erdman DD. Molecular typing of human adenoviruses by PCR and sequencing of a partial region of the hexon gene. *Arch Virol.* 2006;151:1587–602.
24. Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 1987;4:406–25.
25. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 2011;28:2731–9.
26. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32:1792–7.
27. Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. *Bioinformatics.* 1998;14:817–8.
28. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics.* 2003;19:1572–4.
29. Chan P, Goh A, Chua KB, Kharullah NS, Hooi PS. Viral aetiology of lower respiratory tract infection in young Malaysian children. *J Paediatr Child Health.* 1999;35:287–90.
30. Zamberi S, Zulkifli I, Ilina I. Respiratory viruses detected in hospitalised paediatric patients with respiratory infections. *Med J Malaysia.* 2003;58:681.
31. Nair H, Nokes DJ, Gessner BD, et al. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *Lancet.* 2010;375:1545–55.
32. Martin WW, Mulholland EK, Brian MG. Respiratory syncytial virus infection in tropical and developing countries. *Trop Med Int Health.* 1998;3:268–80.
33. Kuroiwa Y, Nagai K, Okita L, et al. A phylogenetic study of human respiratory syncytial viruses group A and B strains isolated in two cities in Japan from 1980–2002. *J Med Virol.* 2005;76:241–7.
34. Nagai K, Kamasaki H, Kuroiwa Y, Okita L, Tsutsumi H. Nosocomial outbreak of respiratory syncytial virus subgroup B variants with the 60 nucleotides—duplicated G protein gene. *J Med Virol.* 2004;74:161–5.
35. Parveen S, Sullender WM, Fowler K, Lefkowitz EJ, Kapoor SK, Broor S. Genetic variability in the G protein gene of group A and B respiratory syncytial viruses from India. *J Clin Microbiol.* 2006;44:3055–64.
36. Scott PD, Ochola R, Ngama M, et al. Molecular epidemiology of respiratory syncytial virus in Kilifi district, Kenya. *J Med Virol.* 2004;74:344–54.
37. Dapat IC, Shobugawa Y, Sano Y, et al. New genotypes within respiratory syncytial virus (RSV) Group B genotype BA in Niigata, Japan. *J Clin Microbiol.* 2010;48(9):3423–7.
38. Melero JA, Garcia-Barreno B, Martinez I, Pringle CR, Cane PA. Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. *J Gen Virol.* 1997;78:2411–8.