Molecular Characterization of Human Respiratory Syncytial Virus in the Philippines from 2012-2015

博士論文

Molecular Characterization of Human Respiratory Syncytial Virus in the Philippines, 2012-2015

(フィリピンにおける Respiratory Syncytial ウイルスの遺伝子解析、2012～2015 年)

東北大学大学院医学系研究科医科学専攻

病理病態学講座 微生物学分野

Rungnapa Malasao

Contents page

SUMMARY

Human respiratory syncytial virus (HRSV) is one of leading causes of acute lower respiratory tract infection especially in infants and young children. Molecular characterization of HRSV may provide useful information in vaccine development and disease management. However, genetic data of HRSV infection among children in developing countries is limited. This study aimed to characterize HRSV in children in the Philippines, from 2012-2015. Nasopharyngeal swabs (NPSs) were collected from hospitalized children and screened for HRSV using real-time polymerase chain reaction (PCR). Positive samples were tested by conventional PCR and sequenced for the second hypervariable region (2nd HVR) of the G gene and complete genome for representative strains. Among a total of 2,308 samples collected from June 2012 – December 2015, 647 samples were positive for HRSV (28.0%), of which 424 (65.5%) and 223 (34.5%) were identified as HRSV-A and HRSV-B, respectively. Of the 424 HRSV-A, 102 (24.1%) were genotype NA1 and 322 (76%) were ON1, while almost all of the 223 HRSV-B were genotype BA9 except for three samples identified as GB2. In 2012/2013 season, two genotypes (NA1 and ON1) of HRSV-A and two genotypes (GB2 and BA9) of HRSV-B co-circulated. The detection of the novel genotype ON1 virus with a 72-nucleotide duplication in the 2nd HVR of the G gene increased rapidly and became the predominant genotype in early 2013. Comparison of whole genomic sequences between six strains of NA1 and seven strains of ON1 showed differences in the G (amino acid positions 232, 237, and 290/314) and the L (amino acid position 598) proteins. Among HRSV-B, BA9 was the predominant genotype circulating in the Philippines. However, two sporadic cases of GB2 genotype were found, which share a common ancestor with other Asian strains. In 2014, NA1 disappeared and ON1 continued to be the predominant strain. A novel ON1 variant, which was phylogenetically distinct from other ON1 strains, was detected in 2014. These novel ON1 variants, ON1-Bi, possess three unique amino acid substitutions (K212R, K213G, and K223G)

in the 2nd HVR of the G gene, which had not been found in other countries. In 2015, the proportion of ON1 viruses decreased while BA9 became the predominant strain. Most of amino acid sites of the 2nd HVR of the G gene were under purifying/negative selection with only <10% of sites under positive selection in both subgroups. These findings suggest that HRSV is an important cause of severe acute respiratory infection among children in the Philippines and revealed the dynamic evolutionary changes of the virus.

BACKGROUND

One of the most common causes of acute lower respiratory tract infection (ALRI) is human respiratory syncytial virus (HRSV), which is particularly important in infants and young children as a main cause of morbidity and mortality around the world[1,2]. Globally, 2.8 to 4.3 million children are admitted to hospitals and approximately 66,000 to 199,000 children infected with HRSV aged less than 5 years old die annually, particularly in developing countries[3]. HRSV is a member of family *Paramyxoviridae*, genus *Pneumovirus*. The HRSV virion is 100–350 nm in diameter. It has an envelope and a linear, negative-sense, and single-stranded RNA genome of approximately 15.2 kb with helical nucleocapsid. The genome has 10 genes, which encodes 11 proteins. The G protein, which is heavily glycosylated with N-linked and, especially, O-linked sugars[4], contains two hypervariable regions (HVR), the 1st HVR and 2nd HVR, flanking a central conserved region, which is the receptor binding site. The HVRs are ectodomains of the G protein and are "mucin-like" structures rich in serine, threonine, and proline. Moreover, the 2nd HVR of the G gene has the highest genetic divergence and has been commonly used for genotyping of HRSV[5-7]. Neutralizing antibodies against purified G proteins were not cross-reactive for HRSV-A and HRSV-B while neutralizing antibodies against purified F proteins can demonstrate cross-reactivity for both subgroups[4].

The etiological significance of HRSV, which is further classified into 2 subgroups including HRSV-A and HRSV-B based initially on their antigenic differences[4,8], has been shown in many countries worldwide, including in developing countries. In the Philippines, a previous study conducted in 2008-2009 revealed that HRSV was the second most common virus detected (24.1%) among hospitalized children with acute respiratory infection, next to rhinoviruses[9]. In our previous study from 2008 to 2012, HRSV was detected in 415 children admitted to hospitals with acute respiratory infection[10]. Almost half of the HRSVs were

identified as HRSV-A, and all of them were genotype NA1. The remaining HRSVs were HRSV-B genotype BA9[10].

Structure of HRSV virion

HRSV is a member of family *Paramyxoviridae*, genus *Pneumovirus*. It has an envelope and a linear, negative-sense, and single-stranded RNA genome of approximately 15.2 kb with helical nucleocapsid. When observed by electron microscopy, two forms of the HRSV virions are irregular spherical form of 100-350 nm in diameter and long filamentous form of 60-200 nm in diameter and up to 10,000 nm in length. Both forms mostly associate with cells during virus budding. The structure of nucleocapsid looks like herringbone which is characterization of viruses in family *Paramyxoviridae*[11-13].

The nine structural proteins containing Nucleoprotein (N), Phosphoprotein (P), Matrix (M), Small hydrophobic (SH), Glycoprotein (G), which is a type II glycoprotein on the virus envelope involved in attachment during virus entry, Fusion (F), M2 ORF1 (M2-1), M2 ORF2 (M2-2), and Large (L) proteins comprise a virion and the remaining two non-structural proteins containing Nonstructural-1 (NS1) and Nonstructural-2 (NS2) are not included in the virion but these proteins are important for replication cycle of HRSV[14]. The schematic structure of the HRSV virus is shown in Fig.1 and the functions of all 11 proteins are shown in Table 1.

Replication cycle of HRSV

The surface G and F glycoproteins mediate virus attachment and penetration[4]. Viral envelope fused with host cell plasma membrane during penetration. There is also evidence of penetration by non-required endosomal acidification clathrin-mediated endocytosis[15,16]. After viral genome releases into cytoplasm called uncoating step, viral RNA dependent RNA polymerase (RDRP) is important in primary transcription to synthesize mRNA which encodes

11 proteins from negative single strand RNA genome and then viral proteins translation occur in cytoplasm. Also, antigenome was synthesized to replicate viral genome for assembling into progeny. The newly synthesized soluble viral N and P proteins promotes elongation of RNA replication products, leading to the production of full-length encapsidated antigenomes and genomes. The genome transcription and replication occur in the cytoplasm with lack of essential nuclear involvement. mRNAs and proteins of HRSV can be detected intracellularly at 4–6 h after infection and reach a peak accumulation by 15–20 h. After that, the M2-2 protein mediates to downregulate transcription in favour of RNA replication and the production of genomes needed for packaging[17]. In the assembly step, nine structural proteins and viral genome package in cytoplasm to be progeny virus. HRSV appears to hijack cellular apical recycling endosomes for budding, a pathway that is distinct from that described for a number of other enveloped RNA viruses[18,19]. The release of progeny virus begins by 10–12 h post infection, reaches a peak after 24 h, and continues until the cells deteriorate by 30–48 h. The schematic diagram of replication cycle of HRSV was shown in Fig. 2. The changing of infected cells develop filamentous surface projections and syncytia which can result from multiple cell fusions of infected uninuclear cells and lead to the destruction of the monolayer[20,21].

Massive glycosylation in the G protein

The mature G protein migrates in gel electrophoresis with band of relative molecular mass (M_r) 80,000-90,000, while the G polypeptide backbone has approximately M_r 32,000. The difference of M_r between the mature and backbone forms of the G protein is from N- and O-linked sugars. N-linked glycosylation can increase the M_r to 45,000. The regions of the G protein having a lot of glycosylations are called mucin-like region including the 1st and 2nd HVRs. The pattern of N-glycosylation is that the 12-15 long chain glycan attaching to nitrogen atom of amino acid Asparagine with configuration as Asparagine (Asn)-Any amino acids

except proline (Pro)-Threonine (Thr). N-glycosylation usually contains 12-15 sugar residues and it is essential for its normal folding and extracellular trafficking to cell surface[22]. The pattern of O-glycosylation is that the 3-4 short chain glycan attaching to oxygen atom of amino acid Serine (Ser) and threonine (Thr)[4]. O-glycosylation contains 3-6 sugar residues and it is important for rigid conformations and protein stability[23]. The G protein with glycosylation sites was shown in Fig. 3.

Genotypes of HRSV-A

HRSV-A can be divided into 11 genotypes: GA1-GA7, SAA1, NA1-NA2 and the 72-nucleotide duplication genotype, ON1 based on the sequence variability of the 2nd HVR of the G gene. The HRSV-A strain A2 is the prototype strain isolated from infant with bronchiolitis and bronchopneumonia in Australia in 1961[24]. Another strain is also used as a prototype named strain Long that was isolated from young male with pneumonia in USA in 1956[25]. In early 1990s, the G gene restriction pattern for subgroup A were mapped for identifying lineages[26], however this method was not popular to define lineages of subgroup A. After that, bioinformatics played an important role for classifying genotypes of HRSV-A. In 1998, the molecular analysis of HRSV-A based on the 2nd HVR region of the G gene in the United States during five successive seasons (season 1990/1991 to season 1994/1995) showed that subgroup A can divided into five genotypes (GA1-GA5) with >96% sequence similarity within genotype[27]. A few years later, two new genotypes (GA6 and GA7) were classified based on the same region[28]. Then, the genotype named SAA1 emerged in South Africa over four consecutive years (1997–2000)[29]. In season 2004/05, the NA1 and NA2 genotypes were phylogenetically closed to GA2, meaning that the variants of GA2 emerged in Niigata, Japan[30]. Also, NA1 was found in different parts of the world including India[31], Thailand[32], China[33], and

Germany[34], and became the predominant genotype in 2008-2009 in the Philippines[10], 2009-2010 and 2010-2011 in Latvia[35], and 2011-2012 in Pakistan[36].

Recently, the novel genotype named ON1 containing 72 nucleotide duplication in the 2nd HVR of the G gene emerged firstly in Ontario, Canada in 2010[5]. ON1 is phylogenetically different from NA1 and supposed to be another variant of GA2. In 2011, it was detected in other countries such as South Korea, Thailand, Malaysia, and India[6,31,32,37]. In 2012, the ON1 genotype HRSVs had spread further to Japan, South Africa, China, Kenya, Germany, Latvia, Italy, Croatia, and Cyprus[7,35,38-44]. The ON1 genotype became the predominant genotype in several countries including South Korea, Kenya, Cyprus, and Italy[37,40,43,44]. In the Philippines, NA1 had circulated since surveillance was conducted in 2008[10]. The phylogenetic tree of HRSV-A showing genotypes was shown in Fig. 4.

Genotypes of HRSV-B

HRSV-B can be divided into 9 genotypes based on the sequence variability of the 2nd HVR of the G gene: GB1-GB4, SAB1-SAB4 and the 60-nucleotide duplication genotype, BA, which has 12 minor groups (BA1-12)[5,6,10,45,46]. After the first prototype HRSV-B strain, SW8/60, was detected in Sweden in 1960[47] was found in 1960, the subgroup B can divided into four genotypes (GB1-GB4) in the United States during 1990-1995[27]. After that, three genotypes named SAB1-SAB3 emerged in South Africa over four consecutive years (1997–2000)[29]. The SAB4 genotype was designated as the new genotype which was found in Cambodia in 2008/2009[48].

Insertions and deletions in the G protein of HRSV have been reported previously[47,49,50]. A noteworthy G gene modification in HRSV-B, which is a 60-nucleotide duplication in the second hypervariable region, was detected in 1999 in Buenos Aires, Argentina[51]. This new genotype, named BA, became predominant worldwide and has even replaced previously

circulating HRSV-B strains in most countries by mid 2000s [51,52]. However, in recent years, a re-emergence of non-BA genotypes has been observed[53]. The first BA genotype of HRSV-B with a 60-nucleotide duplication in the G gene spread throughout the world[51]. Until now, BA1-BA12 subtypes were designated and BA became the predominant genotype in many countries including in 2001-2002 and 2004-2005 in Belgium[52], and in 2002-2003 in Buenos Aires, Argentina[54] and Niigata, Japan[55]. BA9 was first described in Niigata, Japan[45]. After that, some countries including Vietnam[56], India[31,57], China[7,58], Croatia[59], and Thailand[32] reported BA9 as the predominant genotype among HRSV-B viruses. Some countries including Malaysia[6], Latvia[35], and Saudi Arabia[60] reported other BA genotypes as the predominant genotype, possibly indicating that this genotype was circulating globally. In Philippines, BA9 genotype, a subtype of HRSV-B had circulated since surveillance was conducted in 2008[10]. The phylogenetic tree of HRSV-B showing genotypes was shown in Fig. 5.

OBJECTIVE

This study aimed to characterize circulating HRSV strains in the Philippines between June 2012 and December 2015. Whole-genome analysis of representative HRSV-A (NA1 and ON1) strains was also conducted in this study.

MATERIALS AND METHODS

Clinical specimens

Nasopharyngeal swabs (NPSs) were collected from children aged less than 14 years who sought hospital care or referred to the hospital with clinical diagnosis of acute respiratory infection by staff of the Eastern Visayas Regional Medical Center (EVRMC) in Tacloban City of Leyte Island, the Ospital Ng Palawan (ONP) in Puerto Princesa City of Palawan Island, the Biliran Provincial Hospital (BPH) in Naval City of Biliran Island, as well as the Research Institute for Tropical Medicine (RITM) in Metro Manila based on the Integrated Management of Childhood Illness (IMCI) guidelines[61] in the Philippines between June 2012 and December 2015 (Fig. 6).

Study sites

The study was conducted in four study sites including the EVRMC in Leyte Island, the ONP in Palawan Island, the BPH in Biliran Island, and the RITM in Metro Manila. Geographic and demographic data were obtained from the websites https://psa.gov.ph/ and http://nap.psa.gov.ph/activestats/psgc/ (accessed on December 13, 2016). Leyte Island is the regional center for Region VIII (Eastern Visayas). This island has a total area of 6,515 square kilometers and compose of 43 municipalities. The capital of the region is Tacloban City with an area of 202 square kilometers. The climate of Leyte Island is tropical monsoon, and it rains throughout the year with higher rainfall between November and January. The average annual temperature is 27.1°C in Tacloban. Total population in 2015 census was 1,967,000, and of those approximately 242,000 people were living in Tacloban city. Palawan Island is in Region IV-B (Mimaropa) and this island is the fifth largest Island of the Philippines with an area of 17,030 square kilometers. This island consists of 24 municipalities and the capital city is Puerto

Princesa, which is the largest and highly urbanized city. The climate of Palawan Island is very different from other islands in the Philippines. The northern and southern have six months dry and six months wet. The eastern coast has a short dry season (one to three months) and no rainy period during the rest of the year. The average maximum temperature is 31^{O}C. Total population in 2015 census was 1,104,000, of those approximately 255,000 people are living in Puerto Princesa. Biliran Island is in Region VIII (Eastern Visayas) as same region as Leyte Island. This island has a total area of 536 square kilometers and has of 8 municipalities. The capital urbanized city is Naval City with area of 108 square kilometers. The climate of Biliran Island combines between warm and cool with no distinct dry season but heavy wet season occurs in December. The range annual temperature is 24-32^{O}C. Total population in 2015 census was 161,000, of those approximately 49,000 people were living in Naval City. Metro Manila is located in the southwestern portion of Luzon and belongs to the National Capital Region (NCR) of the Philippines, which is a totally urbanized area. This region has a total area of 613.9 square kilometers and composes of 17 municipalities. The capital is Manila with area of 42.8 square kilometers. The climate of this region is wet and dry seasons and classified as tropical monsoon. The short dry season is from January to May. The wet season is from June to December. The range annual temperature is 21-34^{O}C. Total population in 2015 census was 12 million, of those approximately 1,780,000 people were living in Manila.

In 2012/2013 season, samples were collected from four study sites including EVRMC in Tacloban City of Leyte Island, ONP in Puerto Princesa City of Palawan Island, BPH in Naval City of Biliran Island, and RITM in Metro Manila. While in 2014 and 2015 seasons, samples were collected from Biliran Provincial Hospital (BPH) in Naval City of Biliran Island. NPS samples were placed in viral transport medium (VTM) and transported on ice to the Research Institute for Tropical Medicine (RITM) in Metro Manila for further analysis.

Virus isolation

VTM containing NPS samples was inoculated into HEp-2 cells maintained in Eagle's Minimum Essential Medium (MEM) containing 2% inactive calf serum and 1.7% glucose with the addition of penicillin/streptomycin at 35^{O}C under atmosphere of 5% CO_2. Cytopathic effect (CPE) was observed every day after inoculation. If cells showed CPE >80%, supernatant was collected and kept in -80^{O}C. Before viral RNA extraction was performed, isolate samples, which were positive for HRSV-A, were re-inoculated in HEp-2 cells. Three days after inoculation, if more than 80% of HEp-2 cells showed cytopathic effect (CPE), supernatant was collected and kept in -80^{O}C freezer for further study. If cells showing CPE were less than 80%, then cells were kept longer for one or two days until CPE shown >80%.

Viral RNA extraction, cDNA synthesis, Polymerase Chain Reaction (PCR) and Sequencing

NPS samples were centrifuged and the supernatant was collected. Viral RNA was extracted using QIAamp® MinElute® Virus Spin kit. (Qiagen, Hilden, Germany). The viral RNA was reverse-transcribed to complementary DNA (cDNA) using Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase and random primers (Invitrogen Carlsbad, CA, USA). Real-time PCR was performed for HRSV screening[62] and the primers used for this study are shown in Supplementary document 1. The viral RNA extraction, cDNA synthesis, and real-time PCR were done by staff of the RITM in Metro Manila and then cDNAs were transferred to Department of Virology, Graduate School of Medicine in Tohoku University. HRSV positive samples were further genotyped into HRSV-A or HRSV-B by hemi-nested PCR shown in Supplementary document 2. Whole genome was amplified using primers shown in Supplementary document 3. PCR products were purified using QIAquick® PCR Purification

Kit (Qiagen, Hilden, Germany) and then nucleotide sequencing by Sanger dideoxy technique was performed using Big Dye Terminator version 1.1 cycle sequencing kit and Genetic Analyzer 3730 (Applied Biosystems, Foster City, USA).

For complete genome and complete G gene were processed as described above, except differences in some steps used in purification of PCR product by illustra^{TM} ExoProStar^{TM} (GE Healthcare, Buckinghamshire, UK) and nucleotide sequencing using Big Dye Terminator version 3.1 cycle sequencing kit.

Deduced amino acid alignment and Entropy analysis

To evaluate the amino acid variability across the 2nd HVR of the G gene, Shannon entropy analysis as implemented in the BioEdit software was performed[63,64]. The entropy plot is a measure of variability at a particular amino acid in the alignment. The entropy values were exported and plotted in Microsoft Excel. In this analysis, the range of the Shannon entropy values was from 0 to 0.8 with the Shannon entropy threshold value set to 0.2, which was determined from the highest entropy value representing amino acid residues in the conserved region (amino acid position 164-186 of the G protein). Amino acids with entropy value <0.2 are considered conserved. On the other hand, amino acids with entropy value >0.2 are considered variable.

For whole genome analysis, consensus sequences were created with threshold frequency for inclusion in consensus of 90% and then comparison of differences in each protein was performed after sequence alignment.

Glycosylation analysis

To predict potential N-linked and O-linked glycosylation sites of the 2nd HVR of the G protein, NetNGlyc 1.0 and NetOGlyc 4.0 (http://www.cbs.dtu.dk/services/NetNGlyc/ and http://www.cbs.dtu.dk/services/NetOGlyc/) were used. N-linked glycosylation sites were estimated based on the amino acid configuration Asn-Xaa-Ser/Thr except proline (Pro) and O-linked glycosylation sites were estimated based on amino acids serine (Ser) and threonine (Thr).

Phylogenetic analysis

HRSV G nucleotide sequences were aligned using ClustalW implemented in MEGA6.06 software. To infer the genetic relationship among viruses, maximum-likelihood (ML) analysis and the best-fit nucleotide substitution models were performed to estimate unknown parameters of a probability model using MEGA6.06 software[65]. Phylogenetic trees of Philippine sequences of HRSV-A (NA1 and ON1) and HRSV-B (GB2 and BA9) were constructed. BLASTn search results of the nearest relatives of Philippine HRSV strains were included in the analysis. Statistical support was evaluated using bootstrap method with 100 replicates and bootstrap values $\geq$70% are shown on the branches of the consensus tree[66].

To determine the evolutionary dynamics of HRSV-A and HRSV-B in the Philippines, Bayesian Markov Chain Monte Carlo (MCMC) method was used as implemented in the BEAST V.1.8.0 software package[67]. The datasets include only sequence data of genotypes NA1, ON1, and BA9 Philippine HRSVs. Sequences of Philippine HRSVs collected between June 2012 and July 2013, as well as sequences from our previous study collected between May 2008 to April 2012 were included in the analysis[10]. The best-fitting nucleotide substitution model on each dataset, which was selected using MEGA 5.2 software, was used in the analysis.

To estimate the rate of mutation and time to most recent common ancestor (TMRCA), an uncorrelated relaxed clock model was selected. To estimate the genetic diversity of population over time, a skyline plot analysis was performed. For each dataset, 600 million MCMC chains and sampling every 60,000 generations were performed. Convergence was assessed based on the effective sample size (ESS) value >200 of each parameter on every run using TRACER V.1.5 software[67]. The best-fitting model of each dataset was selected based on the lowest Akaike's information criterion through MCMC (AICM) value. The maximum clade credibility (MCC) tree was inferred using TreeAnnotator v.1.8.0 and visualized using FigTree v.1.4.0 software. To estimate the uncertainty in the year for each node, the 95% highest posterior density (HPD) intervals are indicated on the node bars.

To infer the evolutionary relationship of Philippine GB2 strains with strains from other countries, sequences of two GB2 strains detected in the Philippines were used as queries for BLASTn search (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome). The top 100 results and previously reported GB2 strains (BE/90/92, accession number AY751247)[49] were included in the analysis. About 100 million MCMC chains and sampling every 10,000 generations were performed. Assessment of convergence and visualization of MCC tree were performed as described above.

As an independent estimate of the temporal signal of sequence data, root-to-tip linear regression analysis was performed as implemented in the Path-O-Gen program (tree.bio.ed.ac.uk/software/pathogen). The software assesses how genetic distance correlates with sampling date.

Selective pressure analysis

The selective pressure of deduced amino acid sequences of each genotype was evaluated using the estimate selection for each codon method in Hyphy program[68], which is

available in MEGA 5.2 software[69]. If the ratio of the proportion of nonsynonymous substitutions to synonymous substitutions, dN/dS, is > 0, = 0, or < 0, this would indicate that the site is under positive, neutral, or negative selection, respectively.

Two data sets (HRSV-A and HRSV-B) and ML trees for each data set were analyzed using PAML V4.9a (http://abacus.gene.ucl.ac.uk/software/paml.html) under different models and parameters in the site models[70,71]. Three pairs of models were compared including M0 (one-ratio) vs M3 (discrete), M1a (Nearly neutral) vs M2a (Positively selection), and M7 (β) vs M8 (β & ω). The M0 model calculates single nonsynonymous/synonymous substitution ratio (ω) for all sites and assumes that there is no variation among sites. The M3 model allows the three heterogeneous ratios among sites using an unconditional discrete distribution including ω_0 (purifying/negative selection), ω_1 (neutral selection), and ω_2 (positive selection). The M1a model estimates two categories of sites: negatively selected sites with $\omega_0 = 0$ and neutral sites with $\omega_1 = 1$. The M2a model allows a third category of site under positive selection with $\omega_2 > 1$. The M7 model incorporates a continuous beta distribution with estimated parameters p and q among negative or neutral selected sites, and the M8 model specifies a similar model with an additional category of site allowing positive selection with $\omega > 1$. The log-likelihood scores (l) were obtained from each model and compared between paired models (M0 vs M3, M1a vs M2a, and M7 vs M8) by calculating the Likelihood Ratio Tests (LRT, $2\Delta l$) against chi-square distributions with four or two degrees of freedom (d.f.) to test whether the null model was to be rejected[72]. Positive selection is inferred in the event that LRT is statistically significant ($p<0.05$). If LRT is significant, the Bayes Empirical Bayes (BEB) analysis for M2a and M8 models and the Naïve Empirical Bayes (NEB) analysis for M3 model are used for calculating the posterior probabilities. Sites with more than 50.0% posterior probability were considered positively selected sites.

RESULTS

Incidence of HRSV in the Philippines

Two thousands three hundred-eight samples were collected from children admitted with acute respiratory infection in the Philippines between June 2012 and December 2015. Three consecutive epidemic seasons were observed in this study. Both subgroups co-circulated over three epidemic seasons, more detail was shown in Table 2. In 2012/2013 season, ON1 genotype belonged to HRSV-A was predominant. In 2014 season, ON1 genotype was also predominant and, interestingly, the novel ON1 variants, tentatively named ON1-Bi, emerged and circulated in this season with three unique amino acid substitutions, which are not found in the sequences from other countries. In 2015 season, BA9 genotype belonged to HRSV-B was predominant while ON1 strains were detected sporadically.

Temporal distribution of genotypes among three seasons is shown in Fig. 7. ON1 genotype was detected for the first time in October 2012 in Palawan, an island located separately from the main group of islands. Consequently, this genotype also detected in other parts of the Philippines including Biliran, Leyte, and Metro Manila. NA1 genotype was detected only in 2012/2013 season but not in 2014 and 2015 seasons. Moreover, temporal distributions of each study site were shown in Fig. 7A-7D. In Metro Manila from October 2012 to July 2013, small number of samples was collected but most of them were BA9. In Palawan Island from August 2012 to July 2013, most cases were ON1 followed by BA9 and NA1. In Leyte Island, NA1 and ON1 co-circulated predominantly followed by BA9 and sporadic GB2. In Biliran Island from September 2012 to December 2015, four genotypes were detected in 2012/2013 season while ON1 and BA9 co-circulated in 2014 and 2015 seasons.

Phylogenetic analysis of HRSV

HRSV-A were distributed into two clusters, NA1 and ON1 (Fig. 8). The phylogenetic tree was constructed by including strains from other countries that were identified by BLASTn search. The topology of the tree showed a monophyletic group of Philippine strains. The Philippine NA1 strains were closely related to the previously circulating strains in the country in 2008 and 2011, and strains from other countries. The Philippine ON1 strains also clustered with strains detected in other countries including Canada, USA, Panama, Paraguay, South Korea, Japan, Thailand, Malaysia, India, Germany, Spain, Italy, Croatia, and South Africa.

Interestingly, the 13 ON1 strains detected in 2014, tentatively named "ON1-Bi", formed a separate clade (Fig. 8). This novel ON1 variant has three unique amino acid substitutions (K212R, K213G, and K223G) in the 2nd HVR of the G gene (Fig. 9A), which have not been found in the sequences from other countries. BLASTn search showed that the 13 ON1-Bi strains were phylogenetically related to ON1 strains from the Philippines in 2013-2014 (Fig. 8). Based on whole G gene, the phylogenetic tree was constructed by using four ON1-Bi isolates and 100 reference strains identified by BLASTn search. Our four ON1-Bi isolates were phylogenetically near ON1 strains from the Philippines in 2013 (Fig. 9B).

All HRSV-B strains formed a monophyletic group and were identified as BA9 genotype except for three strains, which were identified as GB2 genotype (Fig. 10). The tree was generated by adding strains from other countries that were identified by BLASTn search. All BA9 strains clustered with strains from Argentina, Japan, China, and India, while the three Philippine GB2 strains clustered with Chinese strains.

Amino acid variability of HRSV

Shannon entropy analysis of the 2nd HVR of the G protein of Philippine HRSV-A showed that variable sites are located within and outside of the duplicated regions (Fig 11A).

In 2012/2013 season, five amino acid positions (positions 274, 280, 298, 303, and 304) in both analogous and duplication sites of ON1 strains were variable. Variable sites of HRSV-B are also distributed within and outside the duplicated regions (Fig 11B). Four amino acid positions (positions 242, 257, 267, and 270) in both analogous and duplication sites of BA9 strains were variable. Furthermore, complete genome sequences of HRSV-A, which was predominant in this season, was compared between six isolates of NA1 and seven isolates of ON1. Differences in consensus sequences between NA1 and ON1 were found at position 232, 237, and 290/310 in G protein and position 598 in L protein (Fig. 12).

The deduced amino acid sequences of the ON1 were aligned in each season shown in Fig. 13. Shanon entropy analysis on ON1 amino acid sequences showed variability in nine positions (V225A, T245I, H258Q, L274P, Y297H, L298P, Y304H, L310P, and T319A/K) in 2012/2013 season, nine positions (K212R, K213G, K223G, V225A, L226F, K233E, T252N, H258Q, and T281I) in 2014 season, and five positions (V225A, T239A, R244I, S316P, and T319K) in 2015 season. Most of amino acid variabilities were found in the positions upstream of the duplication site, especially in 2014 season, which was season of emerging ON1-Bi variants. Only in 2012/2013 season, the amino acid variabilities were detected in the duplication region (L298P and Y304H). Moreover, the three amino acid positions (K212R, K213G, and K223G) in 2014 season showed variabilities, which were found in the ON1-Bi variant.

In the BA9 amino acid sequence alignment shown in Fig. 14, there were five positions (S267P, I270T, Y287H/N, T290I, and T302A) in 2013, twelve positions (T227N/P, K233Q, T254I, S267P, I270T, A271V, Y287H/N, T290I, P291S, T302A, A303T/V, and T312I) in 2014, eight positions (P216L, T254I, S265N, S267P, I270T, Y287H/N, T290I, and T312I) in 2015, which showed amino acid variability. Some amino acid positions were variable during

the three seasons including the positions inside the duplication site (S267P and I270T) and the positions downstream of the duplication site (Y287H/N and T290I).

N- and O-glycosylation of HRSV

Six patterns of N-glycosylation were found in during the three seasons among HRSV-A strains. Four patterns were found in the NA1 strains while three patterns were found in the ON1 strains (Table 3). N-glycosylation at positions 237 and 318 was common among ON1 strains including the novel ON1-Bi variant. This glycosylation pattern was not detected in the NA1 strains (Table 3). Three patterns for 2013, three patterns for 2014, and seven patterns for 2015 were found among HRSV-B strains, with the most N-glycosylation pattern at amino acid positions 296 and 310 which was detected mostly in the 2013 season (Table 4).

A total of 21 O-glycosylation patterns were observed among HRSV-A and 37 O-glycosylation patterns among HRSV-B. Three patterns were found in NA1 strains while 18 patterns were found in ON1 strains. The shared pattern between NA1 and ON1 was not found in this study (Table 5). Only one pattern shared among BA9 strains in 2013 and 2014 seasons (Table 6).

Phylodynamics of HRSV in the Philippines

Markov Chain Monte Carlo (MCMC) analysis was performed for nucleotide sequences of HRSVs detected in the Philippines from May 2008 to April 2012 from our previous report[10] and HRSV sequences from viruses collected from June 2012 to July 2013 in the present study. Dated phylogeny analysis of ON1 genotype (204 sequences) and NA1 genotype (198 sequences) viruses collected during May 2008 – July 2013 showed that the ON1 genotype HRSVs might have emerged in January 2012 (95% HPD, July 2011 to June 2012), while the NA1 genotype HRSVs might have emerged in March 2005 (95% HPD, June 2001 to August

2007) (Fig. 15). The NA1 genotype viruses had been circulating since 2008, and its population slightly decreased in late 2012, suddenly decreased in early 2013, and then was steady until the middle of 2013. In contrast, the ON1 genotype viruses emerged in late 2012, suddenly increased in early 2013, and became the predominant genotype of HRSV-A in a short period (October 2012 – July 2013) with a higher evolutionary rate [6.30 x 10^{-3} substitutions/site/year (95% HPD, 3.63 x 10^{-3} to 9.37 x 10^{-3})] than the NA1 genotype [5.40 x 10^{-3} substitutions/site/year (95% HPD, 3.30 x 10^{-3} to 7.74 x 10^{-3})].

Dated phylogeny analysis of BA9 genotype viruses (186 sequences) detected during May 2008 – July 2013 showed that BA9 genotype HRSVs might have emerged in January 2005 (95% HPD, July 2000 to January 2008) (Fig. 16). The BA9 genotype viruses were circulating from 2008 to 2013 then suddenly increased in number in early 2013, with an evolutionary rate of 9.62 x10^{-3} substitutions/site/year (95% HPD, 6.04 x 10^{-3} to 1.36 x 10^{-2}). Coalescent analysis was not possible for the GB2 genotype because there were only two strains detected.

Root-to-tip linear regression analysis of HRSV–A and –B sequence data from 2012/2013 season by Path-O-Gen software revealed no strong temporal signal in both subgroups with correlation coefficient (*r*) of -0.071 for HRSV-A and 0.089 for HRSV-B (Supplementary document 4-5). When the dataset were grouped according to genotypes, higher *r*-values were obtained with $r = 0.429$ for NA1 and 0.315 for ON1 but no temporal signal for BA9 with $r = -0.3722$ (Supplementary document 6-8).

Phylogenetic analysis of GB2 strains detected in the Philippines

Two of the three non-identical GB2 strains detected in 2012/2013 season were analyzed including 97 GenBank strains with known collection dates obtained from a BLASTn result and a reference strain known as GB2 strain (BE/90/92: AY751247). GB2 genotype HRSVs in the

Philippines might have emerged in 2011 and are closely related with strains detected in several locations in Asia, including China, Hong Kong, Thailand, India, and Japan (Fig. 17).

Selective pressure along the 2nd HVR of the G gene of HRSV

Selection analysis of the sequence data obtained from June 2012 – July 2013 identified a positive selected site at L274P found among ON1 strains but not found in other genotypes[73].

Furthermore, I used the sequence data obtained from Biliran Island to analyze selective pressure of HRSV over time. Among HRSV-A detected in three seasons, only a paired model, M0 vs M3, was best-fit and found 16 positively selected sites (Table 7). From 16 positive selected sites, three sites including T245I/S/H, Y273H, and L274P were likely to be under positive selection because the high posterior probabilities for site class were 95.1-97.5%. The 13 remaining sites were under positive selection with range of posterior probability from 50.2 to 89.9%. The proportion of the presence of positive selected sites ranged from 1.59% (model M2a) to 1.73% (model M8). The mean *d*N/*d*S rate ratio ranged from 0.4808 to 0.6641 ($\omega < 1$), thus the 2nd HVR of the G protein was under purifying/negative selection.

Among HRSV-B detected in three seasons, M0 vs M3 paired model was best-fit and found consistently two positive selected sites (V251G/A and Y287H/N). Both sites were shown very high posterior probabilities (>95%). The proportion of the presence of positive selected sites ranged from 2.13% (model M2a) to 2.24% (model M3). The mean *d*N/*d*S rate ratio ranged from 0.5524 to 0.7604 ($\omega < 1$), indicating that the 2nd HVR of the G protein of HRSV-B was under purifying/negative selection (Table 7).

DISCUSSION

HRSV is a major cause of hospitalizations among children with childhood pneumonia. Globally, 22% of children with ALRI were infected with HRSV[3]. From 2008 to 2012, an etiological study on acute respiratory infection in the Philippines showed that HRSV (19.3%) was one of the most common virus detected from hospitalized children with severe pneumonia[10]. During three consecutive seasons, 28.0% children admitted to hospitals were positive for HRSV, indicating etiological significance of HRSV in pediatric population in the Philippines. The positive rates of 2012/2013, 2014, and 2015 seasons were 28.8% (465/1614), 31.3% (121/387), and 19.9% (61/307), respectively. In this study, HRSV circulated over three seasons and was a major cause of children with acute respiratory infection admitted to hospital. Both HRSV subgroups are usually co-circulating during each epidemic season[74,75]. Circulation of HRSV genotypes among four study sites were different. Geographic, climate, and demographic factors might be involved in such differences. But it is unknown how these factors affected epidemiological patterns in each site. The novel ON1 variant named ON1-Bi emerged in a study site (Biliran Island) where data was analyzed until 2014 and 2015 seasons but the study was not conducted in three remaining study sites during that period. Therefore, it is not known if ON1-Bi was also circulating in other sites.

This study was conducted for hospitalized children with a clinical diagnosis in four study sites in different islands of the Philippines. However, from August 2013, the data from only one site (Biliran Island) were included in the analysis. It is possible that other genotypes were circulating in other parts of the country. This study focused on viruses from hospitalized children with a clinical diagnosis of severe pneumonia that might differ with viruses circulating in the community, which may cause less severe infection.

Variants of genotype GA2 circulated in over three epidemic seasons in the Philippines including NA1 and ON1. In this study, NA1 and ON1 co-circulated in 2012/2013 season with

first emergence of ON1 detected in October 2012 in Palawan Island and ON1 continued to circulate in the Philippines until 2015. Among variants of GA2, N- and O-glycosylations conserved at amino acid positions but diversity of amino acids along the 2nd HVR of the G protein were observed over three seasons, especially in 2014 season. The emergence of novel ON1 variant named ON1-Bi has been detected in Biliran Island in 2014 season containing three new amino acid substitutions (K212R, K213G, and K223G) in the G protein have been detected first in this study and were not found in sequences from other countries. If we looked back to genetic codes of three amino acids, we found only one type of mutation called transitions: a → g (adenine → guanine). All three sites were considered as variable and also in the antigenic site[76,77], suggesting that these mutations may contribute to viral fitness. However, fitness-adjusted virus may grow well or less than parent virus due to the nucleotide substitutions in the identified amino acid region or to mutations elsewhere in other unknown regions[78].

Based on whole G gene, ON1-Bi was related to previous ON1 strains found in the Philippines (Biliran and Palawan Islands) with 98% nucleotide sequence identity based on whole G gene, implying that the variant might be local evolution among HRSV-A in the Philippines. However, novel ON1-Bi variants were found only in 2014 in hospitalized children with acute respiratory infection, this variant was still circulating in Biliran Island in non-hospitalized children with mild symptom through the community-based cohort study in 2015. The nucleotide sequence data was similar to ON1-Bi circulating in 2014 season (data not shown).

The possible reasons why genotype of HRSV-A switched from NA1 to ON1 have been unknown. However my molecular analysis may contribute to understand genotype switching. Diversification within the GA2 genotype may play the important role for appearing the distinct variants, for example NA1 variant emerged in Niigata, Japan and contribute worldwide[30] and ON1 variant emerged in Ontario, Canada and spread throughout many

countries[5]. Even though the NA1 and ON1 genotypes formed a monophyletic group, ON1 is phylogenetically different from NA1 by including 72 nucleotide duplication in the G gene which also was found variability in this region among ON1. ON1 emerged, increased, and became predominant among HRSV-A with faster evolutionary rate than NA1. Another protein except the G protein also found the differences between NA1 and ON1 was the L protein which acts as the RNA polymerase. The mutation Y598H showed that 85.7% (12/14) of complete ON1 genomic sequences possess the Y598H mutation in the L protein while absent in the other genotypes (NA1, GA1, GA2, GA5, and GA7) (data not shown).

BA9 genotype circulated in over three epidemic seasons in the Philippines. However BA9 was still having diversity in amino acid over three seasons, especially in 2014 season which only included strains from Biliran Island. The multiple introduction from other places in the Philippines were phylogenetically found in 2014 season, suggesting that movement of disease vector, especially human, from other places to Biliran Island. For example, poultry movements and cross-border viral transmission of pathogenic avian influenza H5N1 virus can drive an intensive evolutionary process[79].

The variants of GA2 (NA1 and ON1) and BA9 were less likely to escape immunity, even though positive selection at codon sites was indicated in the 2nd HVR of the G protein. It might because the persistence of these genotypes circulated over eight years in the Philippines for NA1 (2008 to 2013), ON1 (2013 to 2015), and BA9 (2008 to 2015). The variants persisted in this study might be a stochastic phenomenon to maintain the chain of transmission or due to capability of variants to avoid the immune response[80]. Moreover, the positive selected sites along the 2nd HVR of the G protein at positions 245, 273, and 274 found in HRSV-A and positions 251 and 287 found in HRSV-B might not be in the antigenic sites on the G protein.

Some studies revealed that infants infected with HRSV-A were more severe than HRSV-B and showed that HRSV-A was predominant[81-83]. However, this study showed that HRSV-B was predominant in 2015 season. It might because of ON1 strains belonged to HRSV-A suddenly decreased in 2015, resulting in the high proportion of the BA9 strains belonged to HRSV-A was signified. Some viruses appeared for increasing in numbers over many years and then decline, it is possible that accumulation of resistance in the community to that viruses[84]. One study has found that HRSV subgroups between A and B were not significantly correlated with disease severity[85]. Moreover, HRSV-B was shown as the predominant genogroup, for example HRSV-B was predominant in 2002/2003 and 2007/2008 seasons in Niigata, Japan[45]. However, predominant pattern in each countries was different because not only circulating viruses but also other factors including climate and social contact among school children involved in epidemics of HRSV[86]. At least, the evolutionary rate of BA9 (7.76×10^{-4} to 5.89×10^{-3} substitutions/site/year)[39,87,88] was the fastest followed by ON1 and NA1. The varieties of N- and O-glycosylation patterns were observed more than HRSV-A. The average entropy value counting only in sites with variable of BA9 (0.43) was higher than ON1 (0.31) and the average dN/dS ratio from M0 model of BA9 (0.6281) were also higher than HRSV-A (0.4808). Two positive selected sites were found with strong statistic significant. The first site was amino acid position 251 (V251A/G) defined as an invariable site based on entropy. This site was Glycine in the cluster A and B which were two distinct clusters belonged to HRSV-B exhibiting unique amino acid substitutions and low homology comparing with BA strains detected in 2008 and 2011[10], suggesting that this site exhibited reversible amino acid substitution because some strains might gain some benefits from changing back. The second site was amino acid position 287 (Y287H/N) also defined as a high variable site among all seasons and exhibited reversible amino acid substitution as well[10].

GB2 genotype is one of the genotypes of HRSV-B that has no 60-nucleotide duplication in the 2nd HVR of the G gene. The first GB2 virus was detected in Belgium in 1987[49] indicating that genotype GB2 viruses have been circulating for more than 20 years. In Asia, GB2 genotype viruses have been circulating for more than 10 years, especially in China, albeit at low incidence rates with only a few samples detected in India and Thailand[31,32]. GB2 viruses may have some mechanism to survive, may be endemic somewhere in the world where no sequence data is available, or maintenance of the chain of transmission[80]. The nucleotide sequences of GB2 viruses available in GenBank are also limited. The BLASTn algorithm was used for finding nucleotide sequences that are closely related to the GB2 viruses from the Philippines in order to define its nearest relatives. Phylogenetic analysis showed that Philippine GB2 strains are most closely related to strains from China.

Nevertheless, the positive selected sites with proportion <10% were found along the 2nd HVR of the G protein, both subgroups potentially preferred to stabilize this region ($\omega < 1$). Generally, viruses prefer to fix the deleterious variations since the mutations can lead to incomplete structure or discontinued infection[89-91].

This molecular finding of HRSV revealed that HRSV subgroups co-circulated in every epidemic season of HRSV and gave us the evolutionary trajectory of HRSV genotypes, especially ON1 belonged to HRSV-A and BA9 belonged to HRSV-B with the duplication in G protein, in the Philippines during 2012-2015. Even though, main target of RSV vaccines is the F protein, it is still useful to monitor virus evolution after future introduction of RSV vaccines.

Molecular characterization of HRSV is very important to define in detail the emergence, diversity, evolution dynamics, and possible mechanisms of virus to persist or escape host immunity. The efficient surveillance system allowed us to obtain adequate information to use for analysis.

CONCLUSION

HRSV is the major cause of hospitalization in children with acute respiratory infection in the Philippines. Based on molecular analysis, the genotypic changing pattern of HRSV over three epidemic seasons was observed including long term circulation of GA2 variants and BA9, the emergence of novel ON1 variant with three new amino acid substitutions, and sporadic cases of GB2 sharing common ancestor from other Asian strains.

ACKNOWLEDGEMENTS

We sincerely thank the Tohoku-RITM Collaborative Research Team and the staff of the Biliran Provincial Hospital (BPH), the Eastern Visayas Regional Medical Center (EVRMC), the Ospital Ng Palawan (ONP), and the Research Institute for Tropical Medicine (RITM) on Emerging and Reemerging Diseases.

REFERENCES

1 Black, RE, Cousens, S, Johnson, HL *et al.* Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet* 2010; **375**, 1969-1987.

2 Collins, PL & Melero, JA. Progress in understanding and controlling respiratory syncytial virus: still crazy after all these years. *Virus research* 2011; **162**, 80-99.

3 WHO. 1-42 (WHO Press, 2013).

4 Collins, PL & Karron, RA. in *Fields virology* Vol. 1 (eds D. M. Knipe & P. Howley) 1087-1123 (Lippincott Williams & Wilkins, 2013).

5 Eshaghi, A, Duvvuri, VR, Lai, R *et al.* Genetic variability of human respiratory syncytial virus A strains circulating in Ontario: a novel genotype with a 72 nucleotide G gene duplication. *PLoS One* 2012; **7**, e32807.

6 Khor, CS, Sam, IC, Hooi, PS & Chan, YF. Displacement of predominant respiratory syncytial virus genotypes in Malaysia between 1989 and 2011. *Infect Genet Evol* 2013; **14**, 357-360.

7 Cui, G, Zhu, R, Qian, Y *et al.* Genetic variation in attachment glycoprotein genes of human respiratory syncytial virus subgroups a and B in children in recent five consecutive years. *PLoS One* 2013; **8**, e75020.

8 Mufson, MA, Belshe, RB, Orvell, C & Norrby, E. Respiratory syncytial virus epidemics: variable dominance of subgroups A and B strains among children, 1981-1986. *J Infect Dis* 1988; **157**, 143-148.

9 Suzuki, A, Lupisan, S, Furuse, Y *et al.* Respiratory viruses from hospitalized children with severe pneumonia in the Philippines. *BMC Infect Dis* 2012; **12**, 267.

10 Ohno, A, Suzuki, A, Lupisan, S *et al.* Genetic characterization of human respiratory syncytial virus detected in hospitalized children in the Philippines from 2008 to 2012. *J Clin Virol* 2013; **57**, 59-65.

11 Bachi, T & Howe, C. Morphogenesis and ultrastructure of respiratory syncytial virus. *J Virol* 1973; **12**, 1173-1180.

12 Jeffree, CE, Rixon, HW, Brown, G, Aitken, J & Sugrue, RJ. Distribution of the attachment (G) glycoprotein and GM1 within the envelope of mature respiratory syncytial virus filaments revealed using field emission scanning electron microscopy. *Virology* 2003; **306**, 254-267.

13 Norrby, E, Marusyk, H & Orvell, C. Ultrastructural studies of the multiplication of RS (respiratory syncytial) virus. *Acta pathologica et microbiologica Scandinavica. Section B: Microbiology and immunology* 1970; **78**, 268.

14 Bawage, SS, Tiwari, PM, Pillai, S, Dennis, V & Singh, SR. Recent advances in diagnosis, prevention, and treatment of human respiratory syncytial virus. *Advances in virology* 2013; **2013**, 595768.

15 Kolokoltsov, AA, Deniger, D, Fleming, EH *et al.* Small interfering RNA profiling reveals key role of clathrin-mediated endocytosis and early endosome formation for infection by respiratory syncytial virus. *J Virol* 2007; **81**, 7786-7800.

16 Srinivasakumar, N, Ogra, PL & Flanagan, TD. Characteristics of fusion of respiratory syncytial virus with HEp-2 cells as measured by R18 fluorescence dequenching assay. *J Virol* 1991; **65**, 4063-4069.

17 Bermingham, A & Collins, PL. The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription. *Proceedings of the National Academy of Sciences of the United States of America* 1999; **96**, 11259-11264.

18 Brock, SC, Goldenring, JR & Crowe, JE, Jr. Apical recycling systems regulate directional budding of respiratory syncytial virus from polarized epithelial cells.

Proceedings of the National Academy of Sciences of the United States of America 2003; **100**, 15143-15148.

19 Utley, TJ, Ducharme, NA, Varthakavi, V *et al.* Respiratory syncytial virus uses a Vps4-independent budding mechanism controlled by Rab11-FIP2. *Proceedings of the National Academy of Sciences of the United States of America* 2008; **105**, 10209-10214.

20 Johnson, CR, Yu, W & Murtaugh, MP. Cross-reactive antibody responses to nsp1 and nsp2 of Porcine reproductive and respiratory syndrome virus. *J Gen Virol* 2007; **88**, 1184-1195.

21 Zhang, L, Peeples, ME, Boucher, RC, Collins, PL & Pickles, RJ. Respiratory syncytial virus infection of human airway epithelial cells is polarized, specific to ciliated cells, and without obvious cytopathology. *J Virol* 2002; **76**, 5654-5666.

22 Chen, Q, Miller, LJ & Dong, M. Role of N-linked glycosylation in biosynthesis, trafficking, and function of the human glucagon-like peptide 1 receptor. *American journal of physiology. Endocrinology and metabolism* 2010; **299**, E62-68.

23 Van den Steen, P, Rudd, PM, Dwek, RA & Opdenakker, G. Concepts and principles of O-linked glycosylation. *Critical reviews in biochemistry and molecular biology* 1998; **33**, 151-208.

24 Cameron, R, Buck, C, Merrill, D & Luttick, A. Identification of contaminating adenovirus type 1 in the ATCC reference strain of respiratory syncytial virus A2 (VR-1302). *Virus research* 2003; **92**, 151-156.

25 Chanock, R & Finberg, L. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). II. Epidemiologic aspects of infection in infants and young children. *American journal of hygiene* 1957; **66**, 291-300.

26 Cane, PA & Pringle, CR. Molecular epidemiology of respiratory syncytial virus: rapid identification of subgroup A lineages. *Journal of virological methods* 1992; **40**, 297-306.

27 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 (Pt 9)**, 2221-2229.

28 Peret, TC, 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**, 1891-1896.

29 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**, 2117-2124.

30 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**, 2475-2482.

31 Choudhary, ML, Anand, SP, Wadhwa, BS & Chadha, MS. Genetic variability of human respiratory syncytial virus in Pune, Western India. *Infect Genet Evol* 2013; **20**, 369-377.

32 Auksornkitti, V, Kamprasert, N, Thongkomplew, S *et al.* Molecular characterization of human respiratory syncytial virus, 2010-2011: identification of genotype ON1 and a new subgroup B genotype in Thailand. *Arch Virol* 2014; **159**, 499-507.

33 Ren, L, Xia, Q, Xiao, Q *et al.* The genetic variability of glycoproteins among respiratory syncytial virus subtype A in China between 2009 and 2013. *Infect Genet Evol* 2014; **27**, 339-347.

34 Tabatabai, J, Prifert, C, Pfeil, J, Grulich-Henn, J & Schnitzler, P. Novel Respiratory Syncytial Virus (RSV) Genotype ON1 Predominates in Germany during Winter Season 2012-13. *PLoS One* 2014; **9**, e109191.

35 Balmaks, R, Ribakova, I, Gardovska, D & Kazaks, A. Molecular epidemiology of human respiratory syncytial virus over three consecutive seasons in Latvia. *J Med Virol* 2013.

36 Aamir, UB, Alam, MM, Sadia, H, Zaidi, SS & Kazi, BM. Molecular characterization of circulating respiratory syncytial virus (RSV) genotypes in Gilgit Baltistan Province of Pakistan during 2011-2012 winter season. *PLoS One* 2013; **8**, e74018.

37 Kim, YJ, Kim, DW, Lee, WJ *et al.* Rapid replacement of human respiratory syncytial virus A with the ON1 genotype having 72 nucleotide duplication in G gene. *Infect Genet Evol* 2014; **26C**, 103-112.

38 Tsukagoshi, H, Yokoi, H, Kobayashi, M *et al.* Genetic analysis of attachment glycoprotein (G) gene in new genotype ON1 of human respiratory syncytial virus detected in Japan. *Microbiol Immunol* 2013; **57**, 655-659.

39 Pretorius, MA, van Niekerk, S, Tempia, S *et al.* Replacement and positive evolution of subtype A and B respiratory syncytial virus G-protein genotypes from 1997-2012 in South Africa. *J Infect Dis* 2013; **208 Suppl 3**, S227-237.

40 Agoti, CN, Otieno, JR, Gitahi, CW, Cane, PA & Nokes, DJ. Rapid Spread and Diversification of Respiratory Syncytial Virus Genotype ON1, Kenya. *Emerg Infect Dis* 2014; **20**, 950-959.

41 Valley-Omar, Z, Muloiwa, R, Hu, NC, Eley, B & Hsiao, NY. Novel respiratory syncytial virus subtype ON1 among children, Cape Town, South Africa, 2012. *Emerg Infect Dis* 2013; **19**, 668-670.

42 Prifert, C, Streng, A, Krempl, CD, Liese, J & Weissbrich, B. Novel respiratory syncytial virus a genotype, Germany, 2011-2012. *Emerg Infect Dis* 2013; **19**, 1029-1030.

43 Pierangeli, A, Trotta, D, Scagnolari, C *et al.* Rapid spread of the novel respiratory syncytial virus A ON1 genotype, central Italy, 2011 to 2013. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* 2014; **19**.

44 Panayiotou, C, Richter, J, Koliou, M *et al.* Epidemiology of respiratory syncytial virus in children in Cyprus during three consecutive winter seasons (2010-2013): age distribution, seasonality and association between prevalent genotypes and disease severity. *Epidemiol Infect* 2014, 1-6.

45 Dapat, IC, Shobugawa, Y, Sano, Y *et al.* New genotypes within respiratory syncytial virus group B genotype BA in Niigata, Japan. *J Clin Microbiol* 2010; **48**, 3423-3427.

46 Baek, YH, Choi, EH, Song, MS *et al.* Prevalence and genetic characterization of respiratory syncytial virus (RSV) in hospitalized children in Korea. *Arch Virol* 2012; **157**, 1039-1050.

47 Akerlind, B, Norrby, E, Orvell, C & Mufson, MA. Respiratory syncytial virus: heterogeneity of subgroup B strains. *J Gen Virol* 1988; **69 (Pt 9)**, 2145-2154.

48 van Niekerk, S & Venter, M. Replacement of previously circulating respiratory syncytial virus subtype B strains with the BA genotype in South Africa. *J Virol* 2011; **85**, 8789-8797.

49 Zlateva, KT, Lemey, P, Moës, E, Vandamme, AM & Van Ranst, M. Genetic variability and molecular evolution of the human respiratory syncytial virus subgroup B attachment G protein. *J Virol* 2005; **79**, 9157-9167.

50 Cane, PA, Matthews, DA & Pringle, CR. Frequent polymerase errors observed in a restricted area of clones derived from the attachment (G) protein gene of respiratory syncytial virus. *J Virol* 1993; **67**, 1090-1093.

51 Trento, A, Casas, I, Calderón, A *et al.* Ten years of global evolution of the human respiratory syncytial virus BA genotype with a 60-nucleotide duplication in the G protein gene. *J Virol* 2010; **84**, 7500-7512.

52 Zlateva, KT, Vijgen, L, Dekeersmaeker, N, Naranjo, C & Van Ranst, M. Subgroup prevalence and genotype circulation patterns of human respiratory syncytial virus in Belgium during ten successive epidemic seasons. *J Clin Microbiol* 2007; **45**, 3022-3030.

53 Agoti, CN, Gitahi, CW, Medley, GF, Cane, PA & Nokes, DJ. Identification of group B respiratory syncytial viruses that lack the 60-nucleotide duplication after six consecutive epidemics of total BA dominance at coastal Kenya. *Influenza Other Respir Viruses* 2013; **7**, 1008-1012.

54 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**, 975-984.

55 Sato, M, Saito, R, Sakai, T *et al.* Molecular epidemiology of respiratory syncytial virus infections among children with acute respiratory symptoms in a community over three seasons. *J Clin Microbiol* 2005; **43**, 36-40.

56 Tran, DN, Pham, TM, Ha, MT *et al.* Molecular epidemiology and disease severity of human respiratory syncytial virus in Vietnam. *PLoS One* 2013; **8**, e45436.

57 Choudhary, ML, Wadhwa, BS, Jadhav, SM & Chadha, MS. Complete Genome Sequences of Two Human Respiratory Syncytial Virus Genotype A Strains from India, RSV-A/NIV1114046/11 and RSV-A/NIV1114073/11. *Genome Announc* 2013; **1**.

58 Cui, G, Qian, Y, Zhu, R *et al.* Emerging human respiratory syncytial virus genotype ON1 found in infants with pneumonia in Beijing, China. *Emerg Microbes Infect* 2013, e22.

59 Forcic, D, Ivancic-Jelecki, J, Mlinaric-Galinovic, G *et al.* A study of the genetic variability of human respiratory syncytial virus in Croatia, 2006-2008. *J Med Virol* 2012; **84**, 1985-1992.

60 Almajhdi, FN, Farrag, MA & Amer, HM. Group B strains of human respiratory syncytial virus in Saudi Arabia: molecular and phylogenetic analysis. *Virus Genes* 2013.

61 Gove, S. Integrated management of childhood illness by outpatient health workers: technical basis and overview. The WHO Working Group on Guidelines for Integrated Management of the Sick Child. *Bulletin of the World Health Organization* 1997; **75 Suppl 1**, 7-24.

62 Bonroy, C, Vankeerberghen, A, Boel, A & De Beenhouwer, H. Use of a multiplex real-time PCR to study the incidence of human metapneumovirus and human respiratory syncytial virus infections during two winter seasons in a Belgian paediatric hospital. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2007; **13**, 504-509.

63 Shannon, CE. A mathematical theory of communication. *Bell Labs Journals* 1948; **27**, 379-423, 623-356.

64 Hall, TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 1999; **41**, 95-98.

65 Tamura, K, Stecher, G, Peterson, D, Filipski, A & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution* 2013; **30**, 2725-2729.

66 Hillis, DMaJ, J.J. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* 1993; **42**, 182-192.

67 Drummond, AJ & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology* 2007; **7**, 214.

68 Pond, SL, Frost, SD & Muse, SV. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 2005; **21**, 676-679.

69 Tamura, K, Peterson, D, Peterson, N *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution* 2011; **28**, 2731-2739.

70 Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer applications in the biosciences : CABIOS* 1997; **13**, 555-556.

71 Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology and evolution* 2007; **24**, 1586-1591.

72 Nielsen, R & Yang, Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 1998; **148**, 929-936.

73 Malasao, R, Okamoto, M, Chaimongkol, N *et al.* Molecular Characterization of Human Respiratory Syncytial Virus in the Philippines, 2012-2013. *PLoS One* 2015; **10**, e0142192.

74 Imaz, MS, Sequeira, MD, Videla, C *et al.* Clinical and epidemiologic characteristics of respiratory syncytial virus subgroups A and B infections in Santa Fe, Argentina. *J Med Virol* 2000; **61**, 76-80.

75 Gilca, R, De Serres, G, Tremblay, M *et al.* Distribution and clinical impact of human respiratory syncytial virus genotypes in hospitalized children over 2 winter seasons. *J Infect Dis* 2006; **193**, 54-58.

76 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. *Proceedings of the National Academy of Sciences of the United States of America* 1987; **84**, 5625-5629.

77 Plotnicky-Gilquin, H, Goetsch, L, Huss, T *et al.* Identification of multiple protective epitopes (protectopes) in the central conserved domain of a prototype human respiratory syncytial virus G protein. *J Virol* 1999; **73**, 5637-5645.

78 Zhao, X, Liu, E, Chen, FP & Sullender, WM. In vitro and in vivo fitness of respiratory syncytial virus monoclonal antibody escape mutants. *J Virol* 2006; **80**, 11651-11657.

79 Marinova-Petkova, A, Feeroz, MM, Rabiul Alam, SM *et al.* Multiple introductions of highly pathogenic avian influenza H5N1 viruses into Bangladesh. *Emerg Microbes Infect* 2014; **3**, e11.

80 Agoti, CN, Otieno, JR, Ngama, M *et al.* Successive Respiratory Syncytial Virus Epidemics in Local Populations Arise from Multiple Variant Introductions, Providing Insights into Virus Persistence. *J Virol* 2015; **89**, 11630-11642.

81 Falsey, AR & Walsh, EE. Respiratory syncytial virus infection in adults. *Clinical microbiology reviews* 2000; **13**, 371-384.

82 McConnochie, KM, Hall, CB, Walsh, EE & Roghmann, KJ. Variation in severity of respiratory syncytial virus infections with subtype. *The Journal of pediatrics* 1990; **117**, 52-62.

83 Walsh, EE, McConnochie, KM, Long, CE & Hall, CB. Severity of respiratory syncytial virus infection is related to virus strain. *J Infect Dis* 1997; **175**, 814-820.

84 Cane, PA, Matthews, DA & Pringle, CR. Analysis of respiratory syncytial virus strain variation in successive epidemics in one city. *J Clin Microbiol* 1994; **32**, 1-4.

85 Bergstrasser, E, Zbinden, R, Minder, C & Gnehm, HE. [Severity of respiratory syncytial virus infection influenced by clinical risk factors and subtype A and B in hospitalized children]. *Klinische Padiatrie* 1998; **210**, 418-421.

86 Weber, A, Weber, M & Milligan, P. Modeling epidemics caused by respiratory syncytial virus (RSV). *Mathematical biosciences* 2001; **172**, 95-113.

87 Tan, L, Coenjaerts, FE, Houspie, L *et al.* The comparative genomics of human respiratory syncytial virus subgroups A and B: genetic variability and molecular evolutionary dynamics. *J Virol* 2013; **87**, 8213-8226.

88 Kushibuchi, I, Kobayashi, M, Kusaka, T *et al.* Molecular evolution of attachment glycoprotein (G) gene in human respiratory syncytial virus detected in Japan 2008-2011. *Infect Genet Evol* 2013; **18**, 168-173.

89 Balasuriya, UB, Nadler, SA, Wilson, WC *et al.* The NS3 proteins of global strains of bluetongue virus evolve into regional topotypes through negative (purifying) selection. *Veterinary microbiology* 2008; **126**, 91-100.

90 Ngandu, NK, Scheffler, K, Moore, P *et al.* Extensive purifying selection acting on synonymous sites in HIV-1 Group M sequences. *Virology journal* 2008; **5**, 160.

91 Lima, CJ & Bandiera-Paiva, P. Group A Rotavirus NSP4 is Under Negative Selective Pressure. 1-16 2014.

92 Atreya, PL, Peeples, ME & Collins, PL. The NS1 protein of human respiratory syncytial virus is a potent inhibitor of minigenome transcription and RNA replication. *J Virol* 1998; **72**, 1452-1461.

93 Fuentes, S, Coyle, EM, Beeler, J, Golding, H & Khurana, S. Antigenic Fingerprinting following Primary RSV Infection in Young Children Identifies Novel Antigenic Sites and Reveals Unlinked Evolution of Human Antibody Repertoires to Fusion and Attachment Glycoproteins. *PLoS pathogens* 2016; **12**, e1005554.

94 Kumaria, R, Iyer, LR, Hibberd, ML, Simoes, EA & Sugrue, RJ. Whole genome characterization of non-tissue culture adapted HRSV strains in severely infected children. *Virology journal* 2011; **8**, 372.

95 Rebuffo-Scheer, C, Bose, M, He, J *et al.* Whole genome sequencing and evolutionary analysis of human respiratory syncytial virus A and B from Milwaukee, WI 1998-2010. *PLoS One* 2011; **6**, e25468.

96 Tan, L, Lemey, P, Houspie, L *et al.* Genetic variability among complete human respiratory syncytial virus subgroup A genomes: bridging molecular evolutionary dynamics and epidemiology. *PLoS One* 2012; **7**, e51439.

97 Eckardt-Michel, J, Lorek, M, Baxmann, D *et al.* The fusion protein of respiratory syncytial virus triggers p53-dependent apoptosis. *J Virol* 2008; **82**, 3236-3249.

98 Krzyzaniak, MA, Zumstein, MT, Gerez, JA, Picotti, P & Helenius, A. Host cell entry of respiratory syncytial virus involves macropinocytosis followed by proteolytic activation of the F protein. *PLoS pathogens* 2013; **9**, e1003309.

99 Collins, PL, Hill, MG, Cristina, J & Grosfeld, H. Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. *Proceedings of the National Academy of Sciences of the United States of America* 1996; **93**, 81-85.

100 Hardy, RW & Wertz, GW. The product of the respiratory syncytial virus M2 gene ORF1 enhances readthrough of intergenic junctions during viral transcription. *J Virol* 1998; **72**, 520-526.

101 Fearns, R & Collins, PL. Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J Virol* 1999; **73**, 5852-5864.

102 Barik, S. The structure of the 5' terminal cap of the respiratory syncytial virus mRNA. *J Gen Virol* 1993; **74 (Pt 3)**, 485-490.

FIGURES

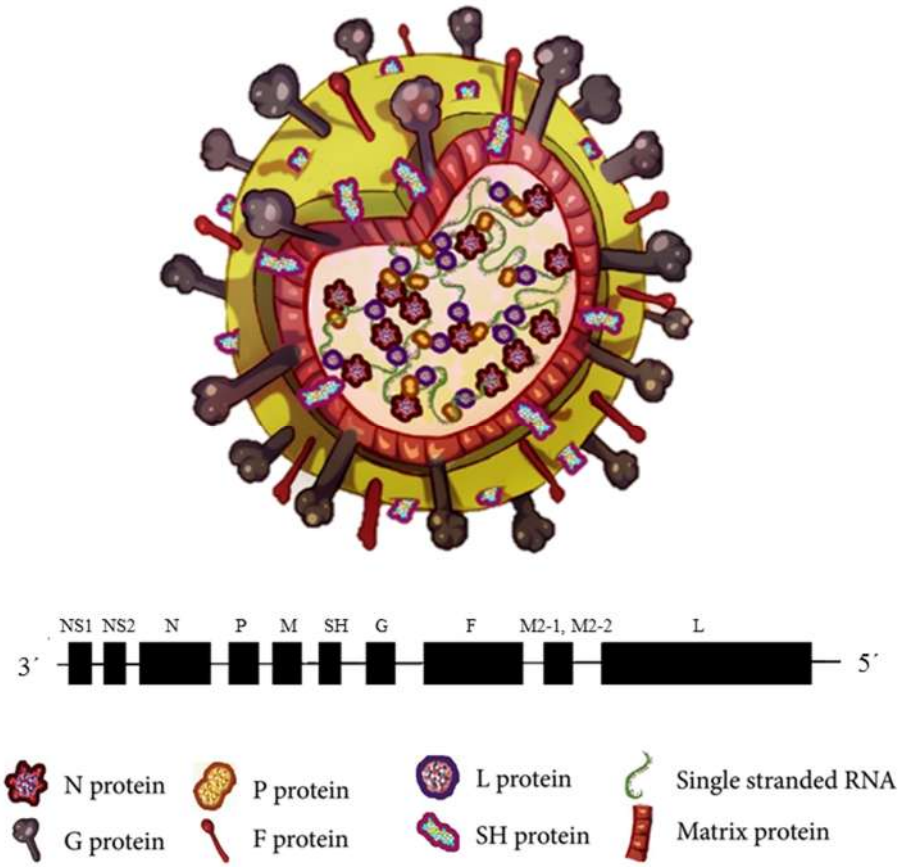


Fig. 1 Cartoon diagram of HRSV structure and negative single-strand RNA which encodes two non-structural proteins: NS1 and NS2 and nine structural proteins: N, P, M, SH, G, F, M2-1, M2-2, and L. The surface glycoproteins contain G, F, and SH. The nucleocapsid consists of genomic RNA, N, P, M2-1, and L. This figure was adapted from Bawage S.S. *et al*, Adv in Virol. 2013.

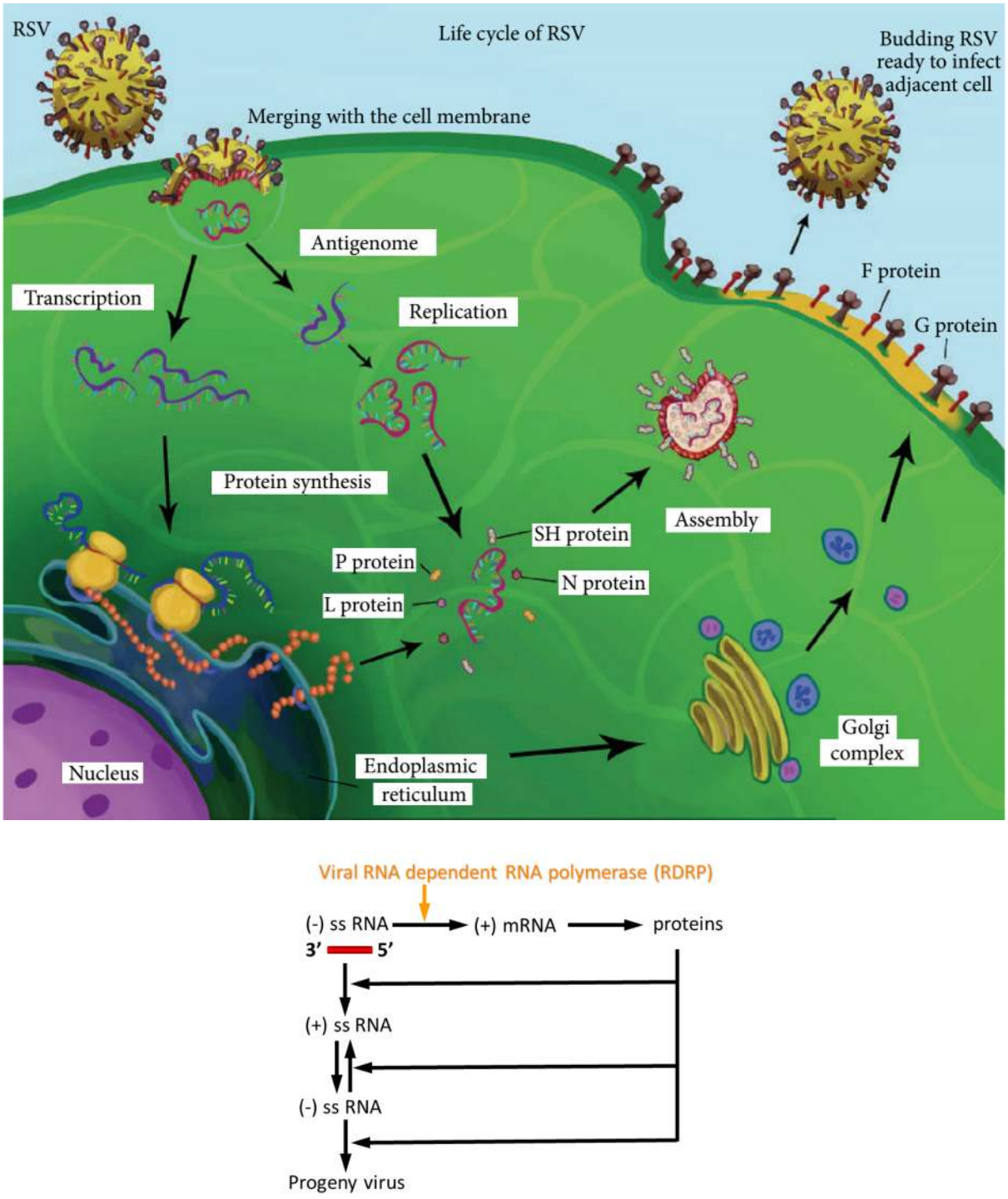


Fig. 2 Replication cycle of HRSV (upper) Bawage S.S. *et al*, Adv in Virol. 2013 and simple life cycle diagram (lower) of HRSV. The viral replication performs in cytoplasm and viral RNA dependent RNA polymerase is enzyme that virus need to carry it.

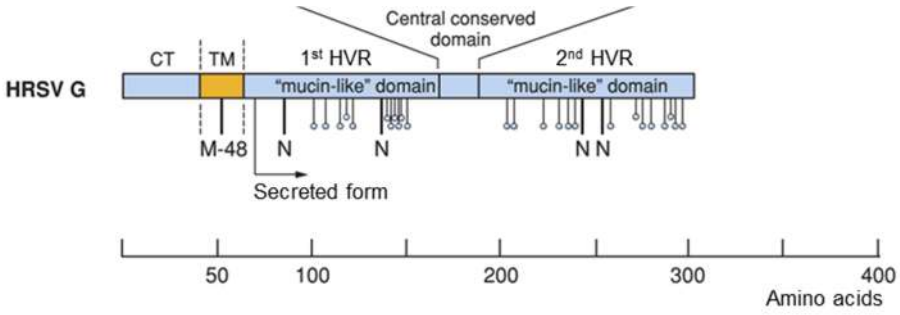


Fig. 3 Primary structure of the G surface glycoprotein of HRSV-A strain A2 (approximately to scale). The blue bar is the G protein; CT, cytoplasmic tail; TM (brown box), transmembrane anchor; 1st HVR, the first hypervariable region; 2nd HVR, the second hypervariable region; stalks with N, N-linked sugars; stalks with O, O-linked sugars. Adapted from Collins et al., 2013[4].

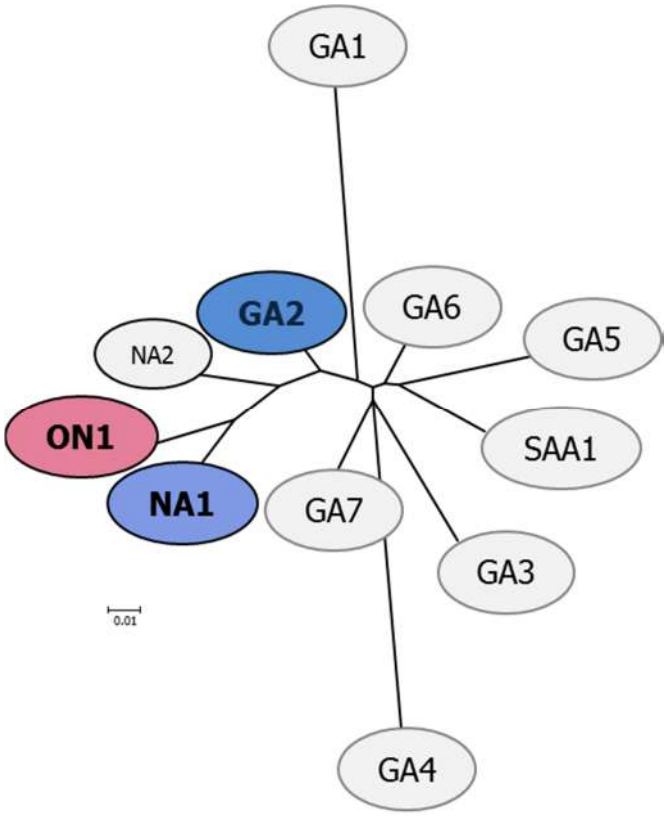


Fig. 4 The phylogenetic tree of HRSV-A genotypes. The colored circles showed genotypes which found in this study (NA1 and ON1). The variants of GA2 including NA1, NA2, and ON1.

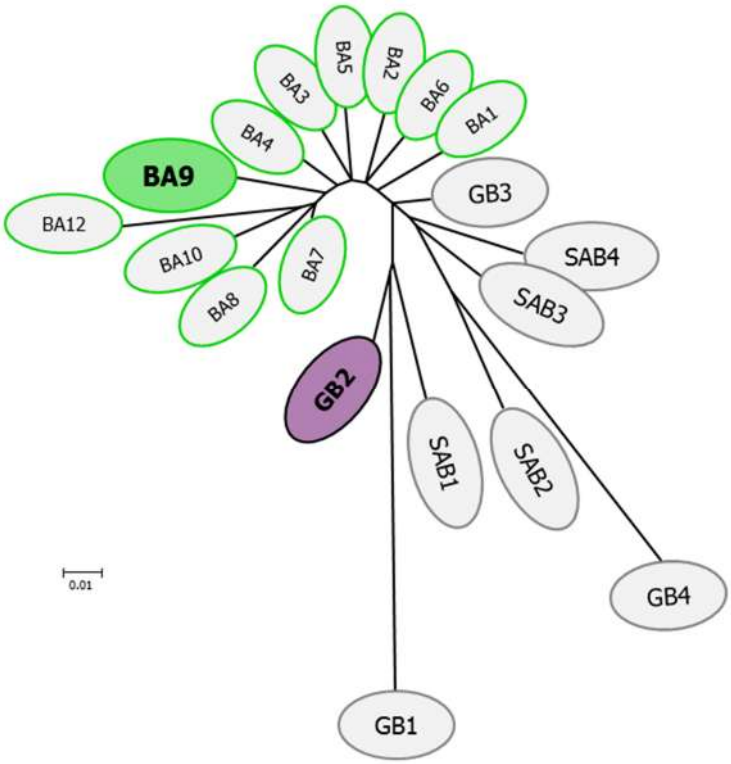


Fig. 5 The phylogenetic tree of HRSV-B genotypes. The colored circles showed genotypes which found in this study including BA9 and GB2.

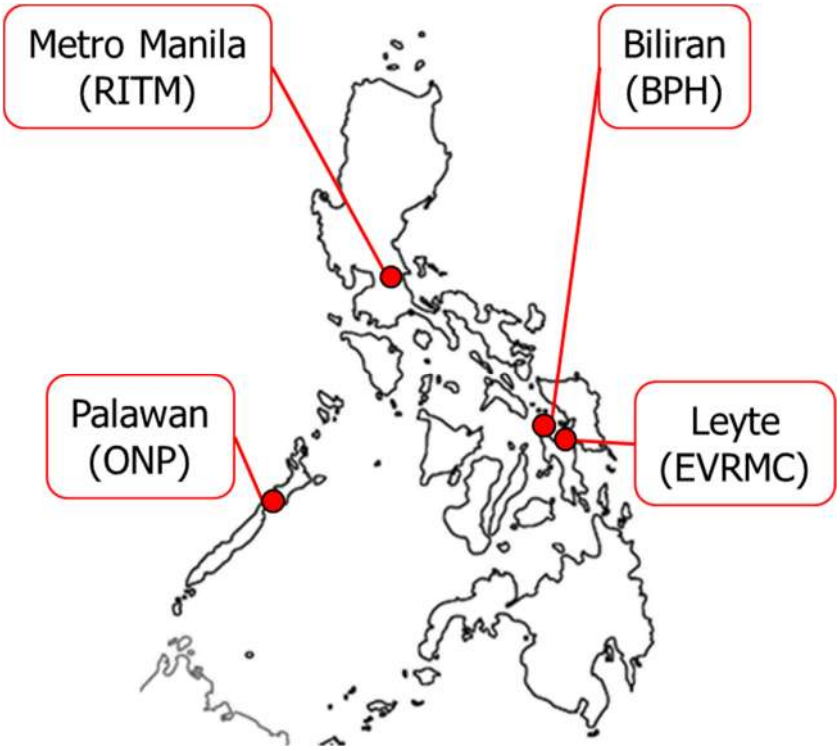


Fig. 6 The study sites including the Research Institute for Tropical Medicine (RITM) in Metro Manila, the Ospital Ng Palawan (ONP) in Puerto Princesa City of Palawan Island, the Eastern Visayas Regional Medical Center (EVRMC) in Tacloban City of Leyte Island, and the Biliran Provincial Hospital (BPH) in Naval City of Biliran Island.

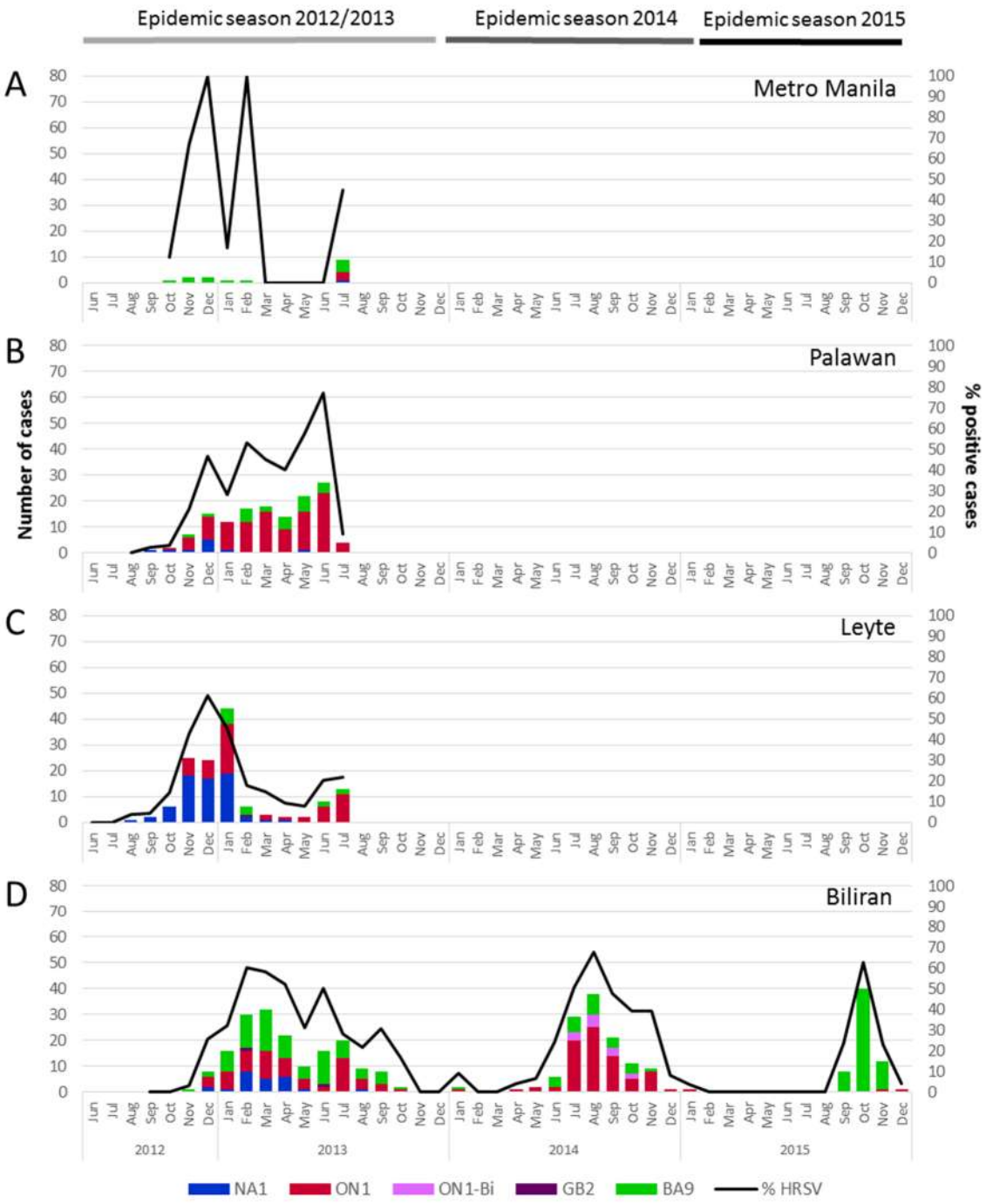


Fig. 7 Temporal distribution of HRSV genotypes showing the three consecutive epidemic seasons 2012/2013, 2014, and 2015 in the Philippines. Periods of each epidemic season were indicated using horizontal lines colored by light gray for 2012/2013, gray for 2014, and black for 2015. Note that the 2012/2013 season was conducted in four study sites while the 2014 and 2015 seasons were conducted in one study site. (A) was conducted in the Research Institute for Tropical Medicine (RITM) in Metro Manila, (B) was conducted in the Ospital Ng Palawan (ONP) in Puerto Princesa City of Palawan Island, (C) was conducted in the Eastern Visayas Regional Medical Center (EVRMC) in Tacloban City of Leyte Island, and (D) was conducted in the Biliran Provincial Hospital (BPH) in Naval City of Biliran Island.

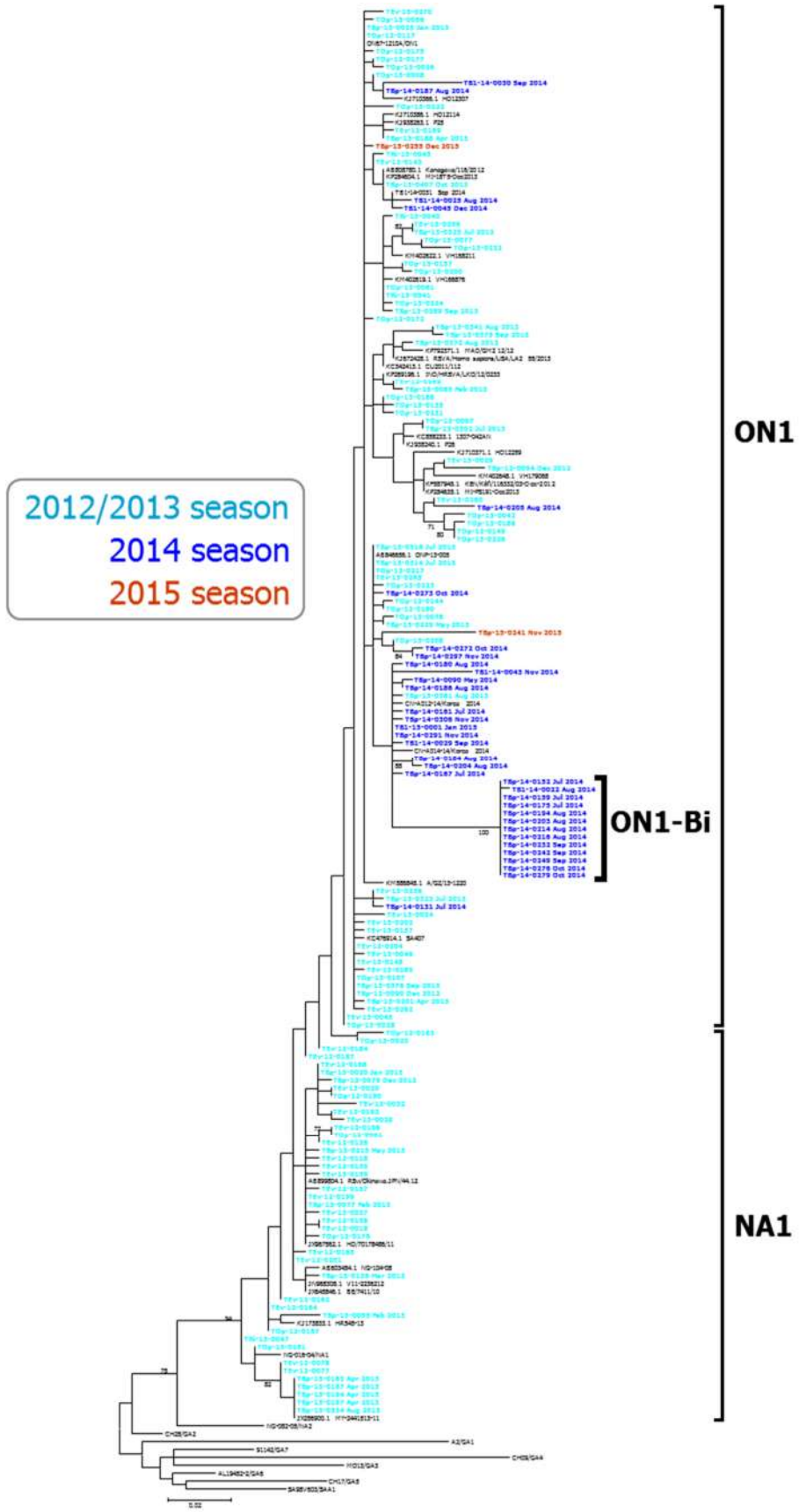


Fig. 8 Phylogenetic tree of HRSV-A circulating in the Philippines from June 2012-December 2015. NA1 circulated in 2012/2013 season while ON1 circulated over three epidemic seasons. All NA1 and ON1 strains were form monophyletic group which were GA2 variants. Interestingly, the novel ON1 variants, named ON1-Bi, by supporting with high value of bootstrap (99%) were found in 2014 season. (n=189)

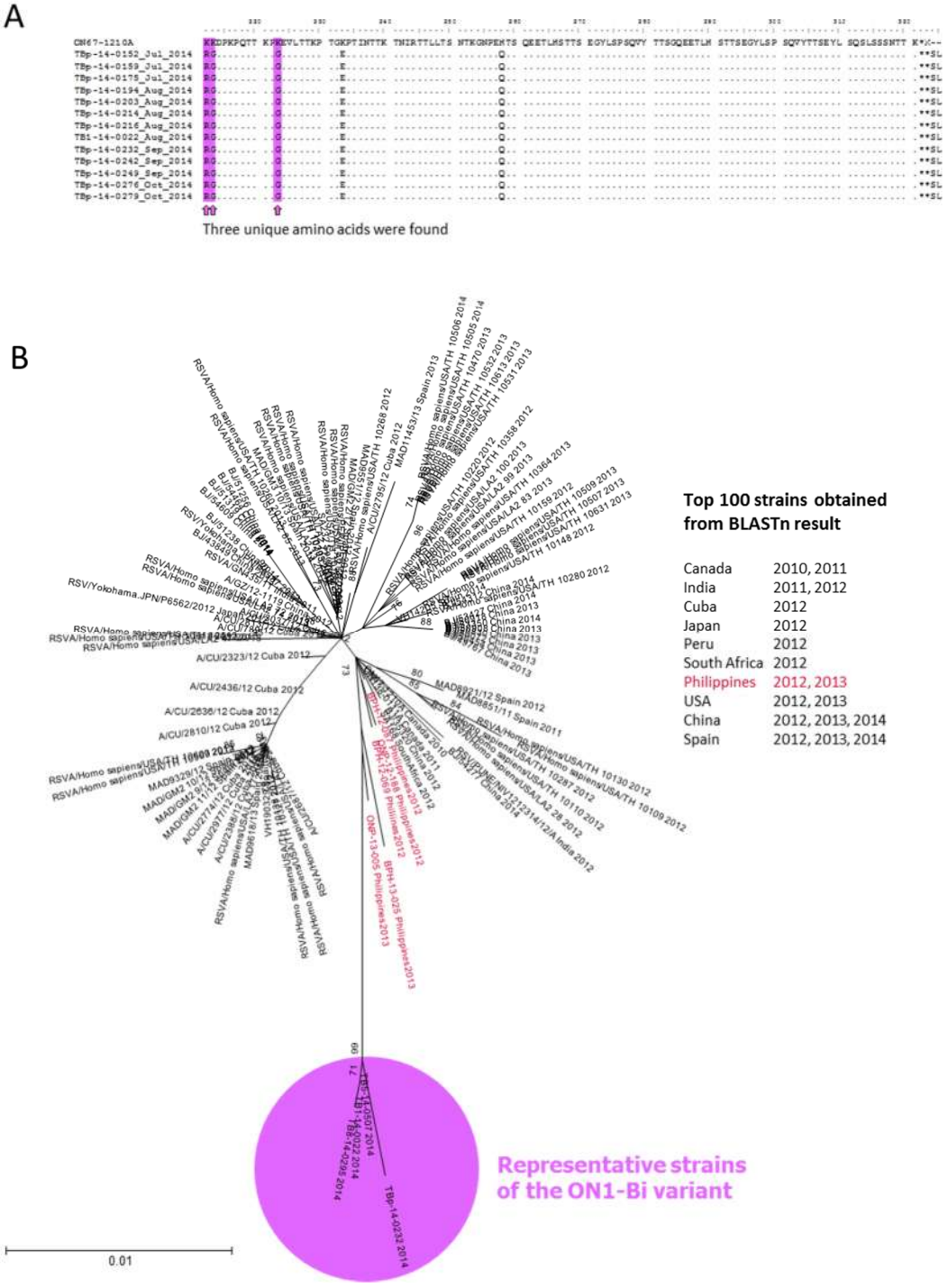


Fig. 9 Amino acid sequence alignment and Phylogenetic tree of novel ON1-Bi variants. (A) was deduced amino acid sequence alignment along the 2nd HVR of the G protein among 13 ON1-Bi variants. Three new amino acid substitutions consisted K212R, K213G, and K223G comparing with prototype strain (ON67-1210A). (B) was a phylogenetic tree constructed using nucleotide sequences of whole G gene among four isolates of ON1-Bi and top 100 strains obtained from BLASTn result. The nearest strain was from the Philippines in 2013.

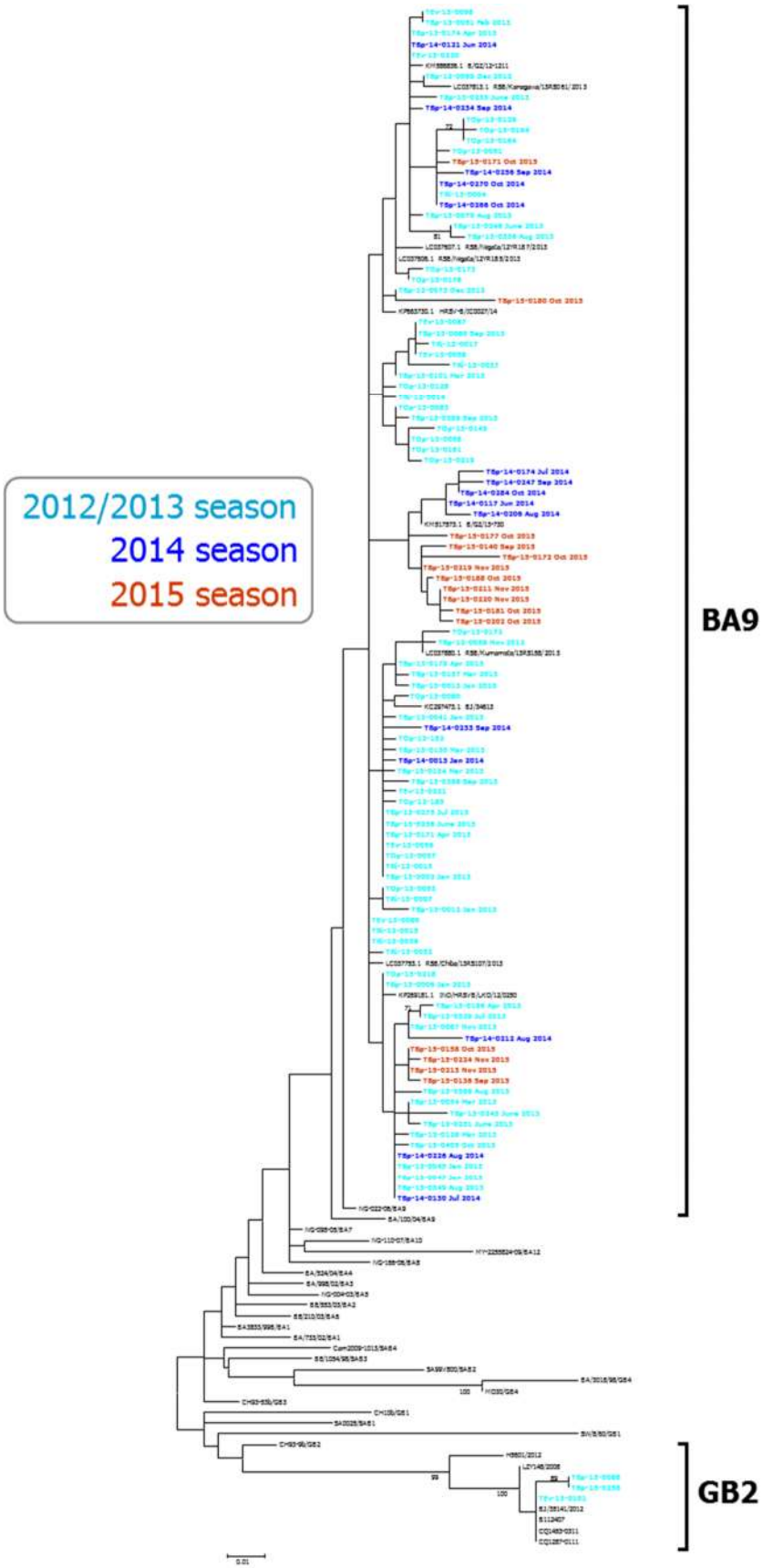


Fig. 10 Phylogenetic tree of HRSV-B circulating in the Philippines from June 2012-December 2015. GB2 was sporadic detection in 2013 season while most of them were BA9 which circulated among three epidemic seasons in the Philippines. (n=144)

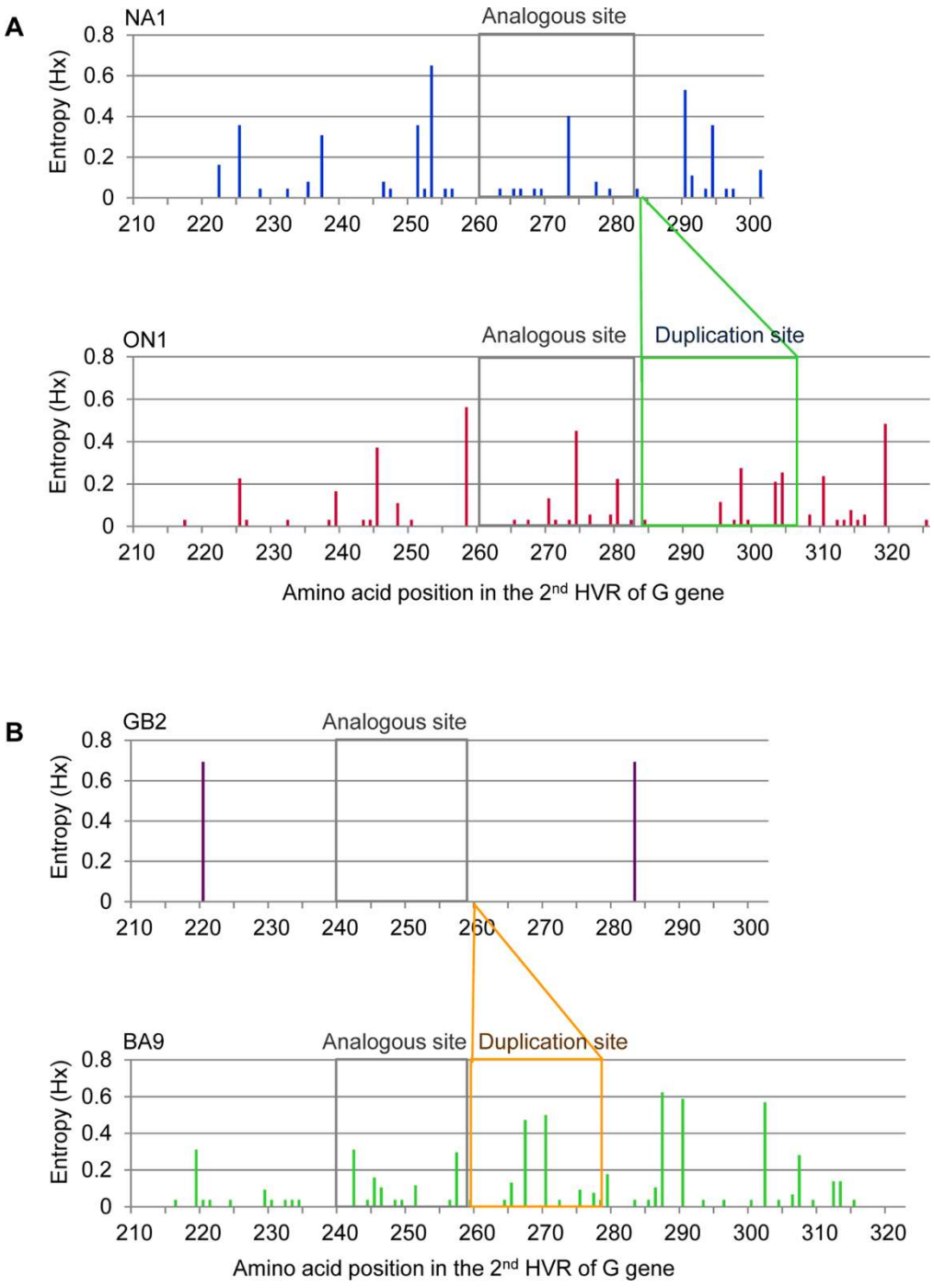
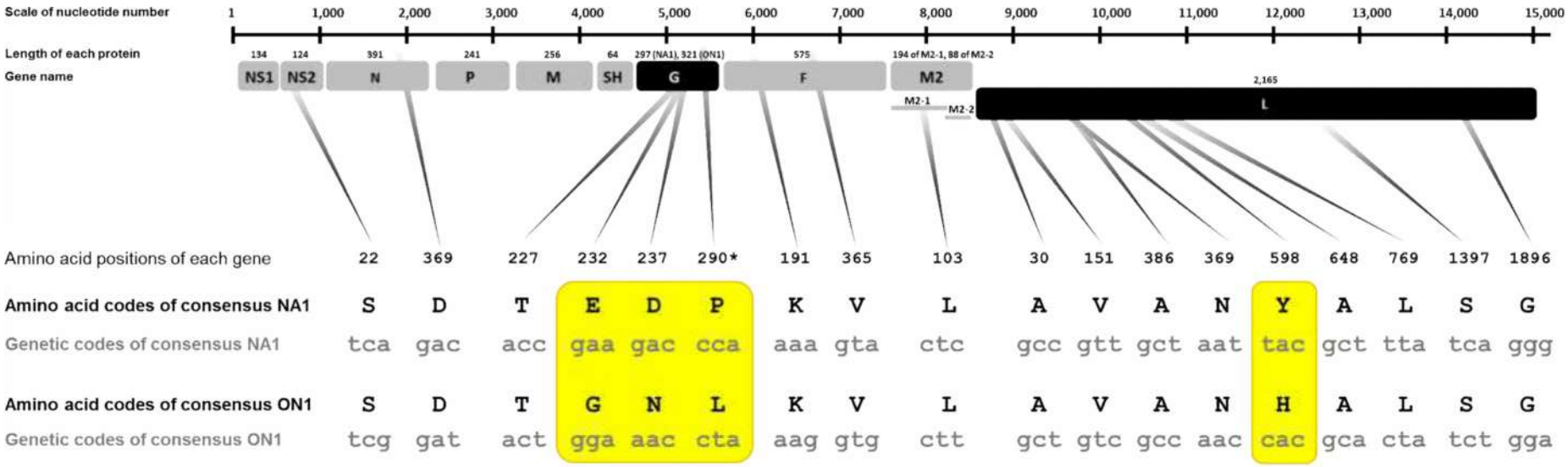


Fig. 11 Comparison of entropy value among HRSV. (A) was compared between NA1 and ON1 while (B) was compared between GB2 and BA9. Note: this analysis was performed by using HRSV sequences from June 2012 – July 2013. (n of NA1 = 101; n of ON1 = 204; n of GB2 = 2; n of BA9 = 116)



* The position 290 in the G protein of NA1 is equal to the position 314 in the G protein of ON1.

Fig. 12 Comparison of complete genome between the consensus NA1 and ON1 sequences represented in the Philippines. The black boxes for genes containing non-synonymous substitutions and the yellow boxes for codes having non-synonymous substitutions were indicated. (n of NA1 = 6; n of ON1 = 7)

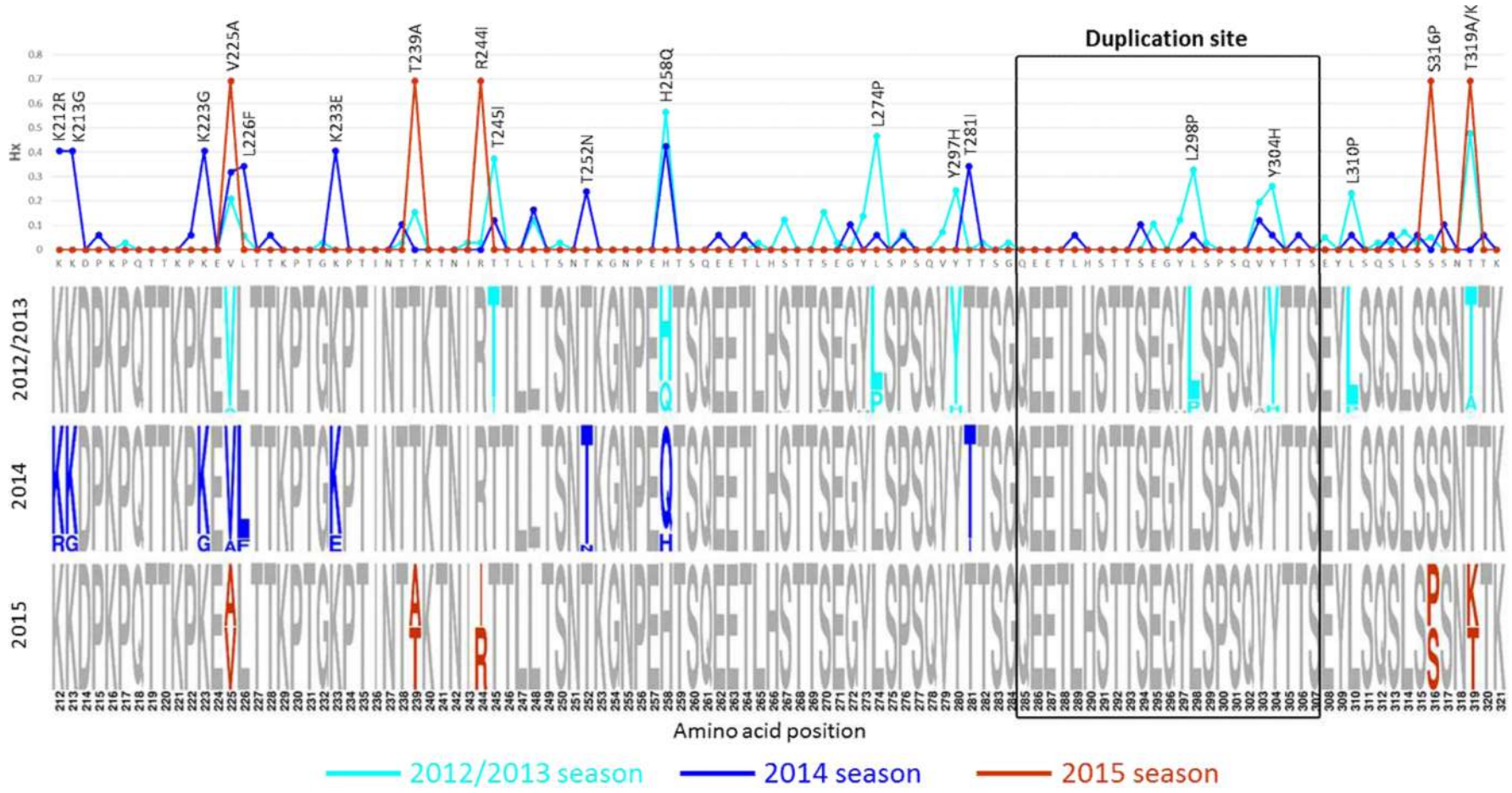


Fig. 13 Entropy plots (upper) and sequence logos (lower) along the 2nd HVR of the G protein (position 212-321) of all ON1 strains of HRSV-A among three epidemic seasons. The amino acid positions with variable (entropy value >0.2) were marked. Sequence logos of each epidemic season showed the height of symbols within the stack indicates the frequency of sequence conservation at that position. Colored symbols were sites with variability for each season. (n of 2012/2013 season = 329; no of 2014 season = 93; n of 2015 season = 2)

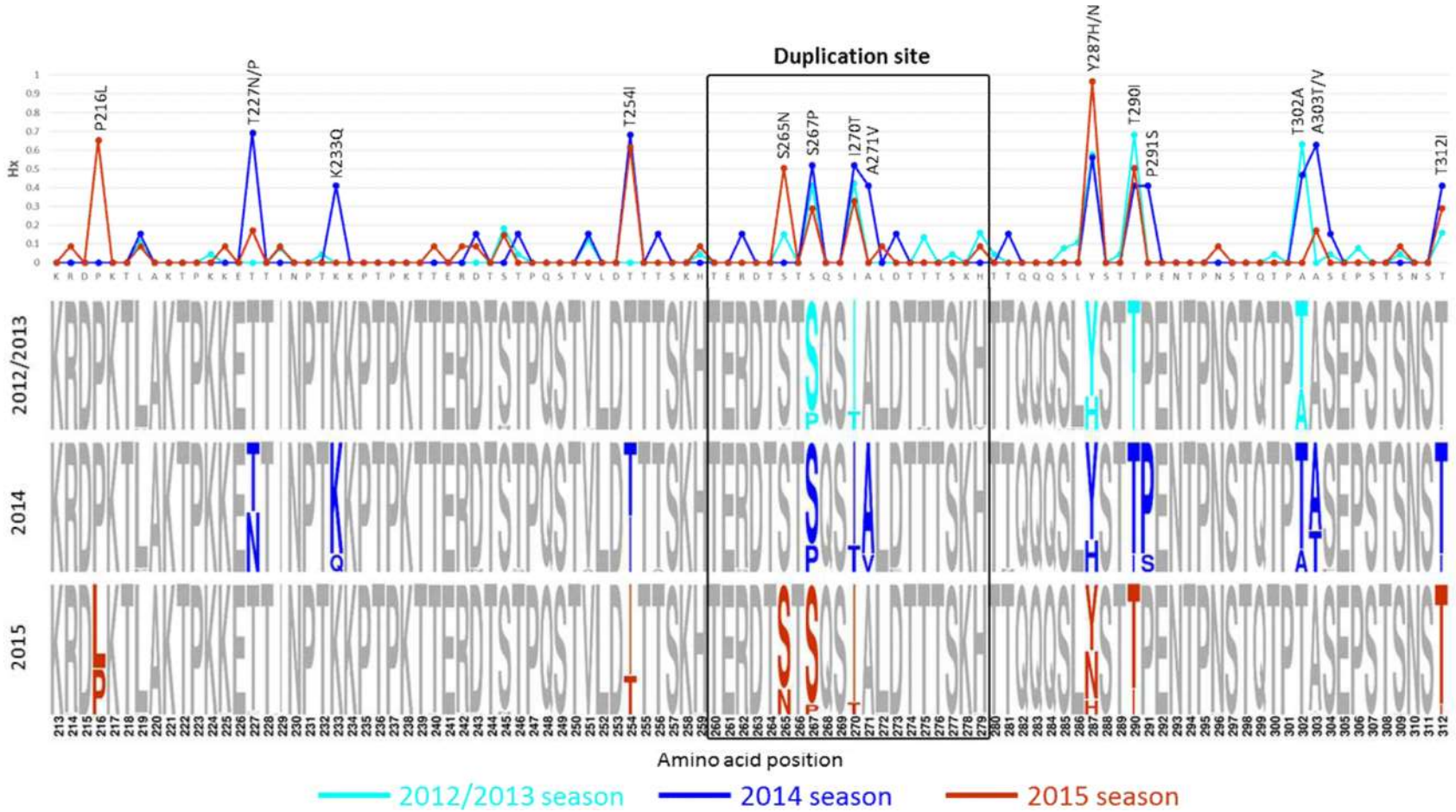


Fig. 14 Entropy plots (upper) and sequence logos (lower) along the 2nd HVR of the G protein (position 213-312) of all BA9 strains of HRSV-B among three epidemic seasons. The amino acid positions with variable (entropy value >0.2) were marked. Sequence logos of each epidemic season showed the height of symbols within the stack indicates the frequency of sequence conservation at that position. Colored symbols were sites with variability for each season. (n of 2012/2013 season = 136; no of 2014 season = 28; n of 2015 season = 59)

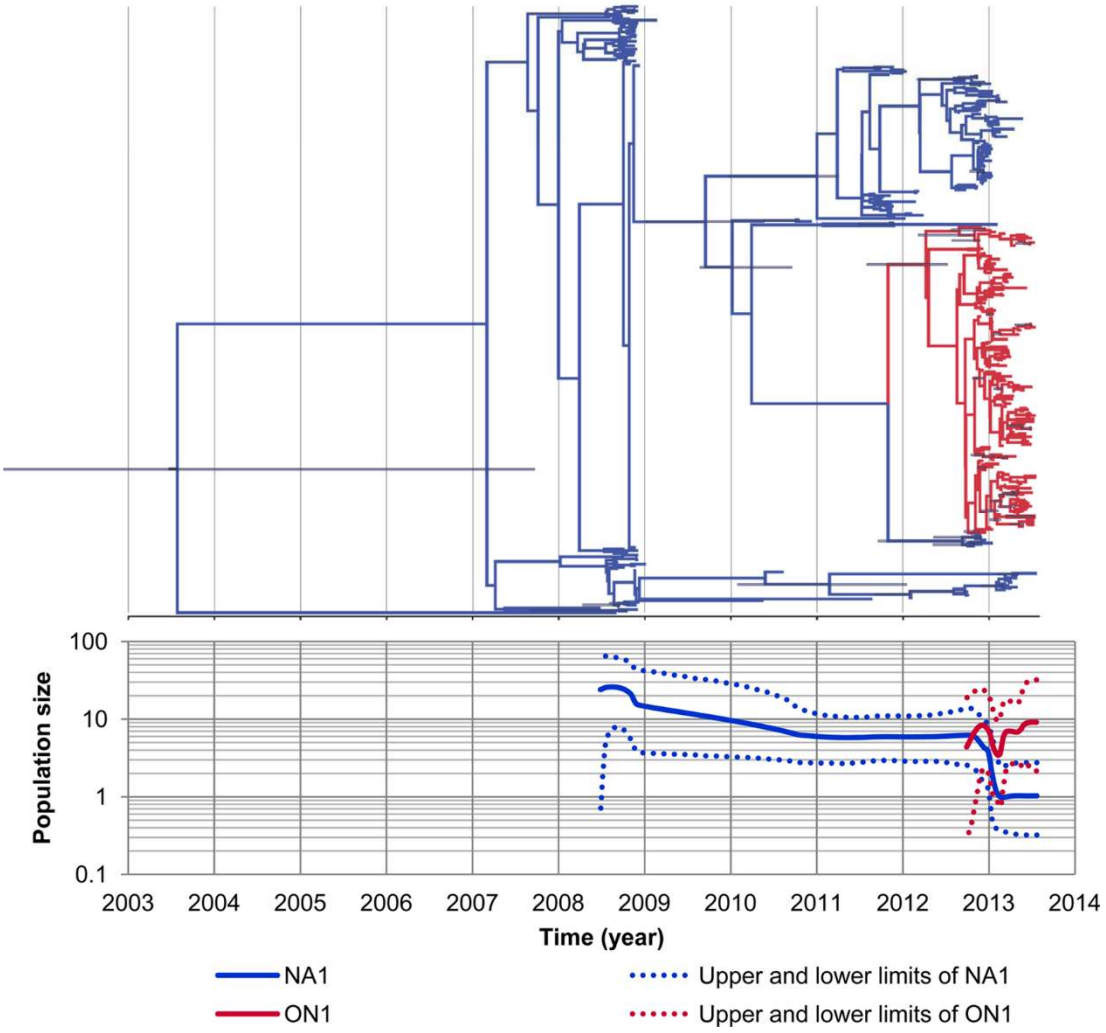


Fig. 15 Phylodynamic analysis of HRSV-A from 2008 to 2013. Dated phylogeny (upper) and Bayesian skyline plot (lower) of the NA1 and ON1 genotypes. The NA1 genotype was predominant and its population size showed a decreasing trend from 2009 to late 2012 and then exhibited a sharp decline in 2013. The ON1 genotype emerged around late 2012 and became the predominant genotype with an increasing trend in population size in 2013. (n of NA1 = 198; n of ON1 = 204)

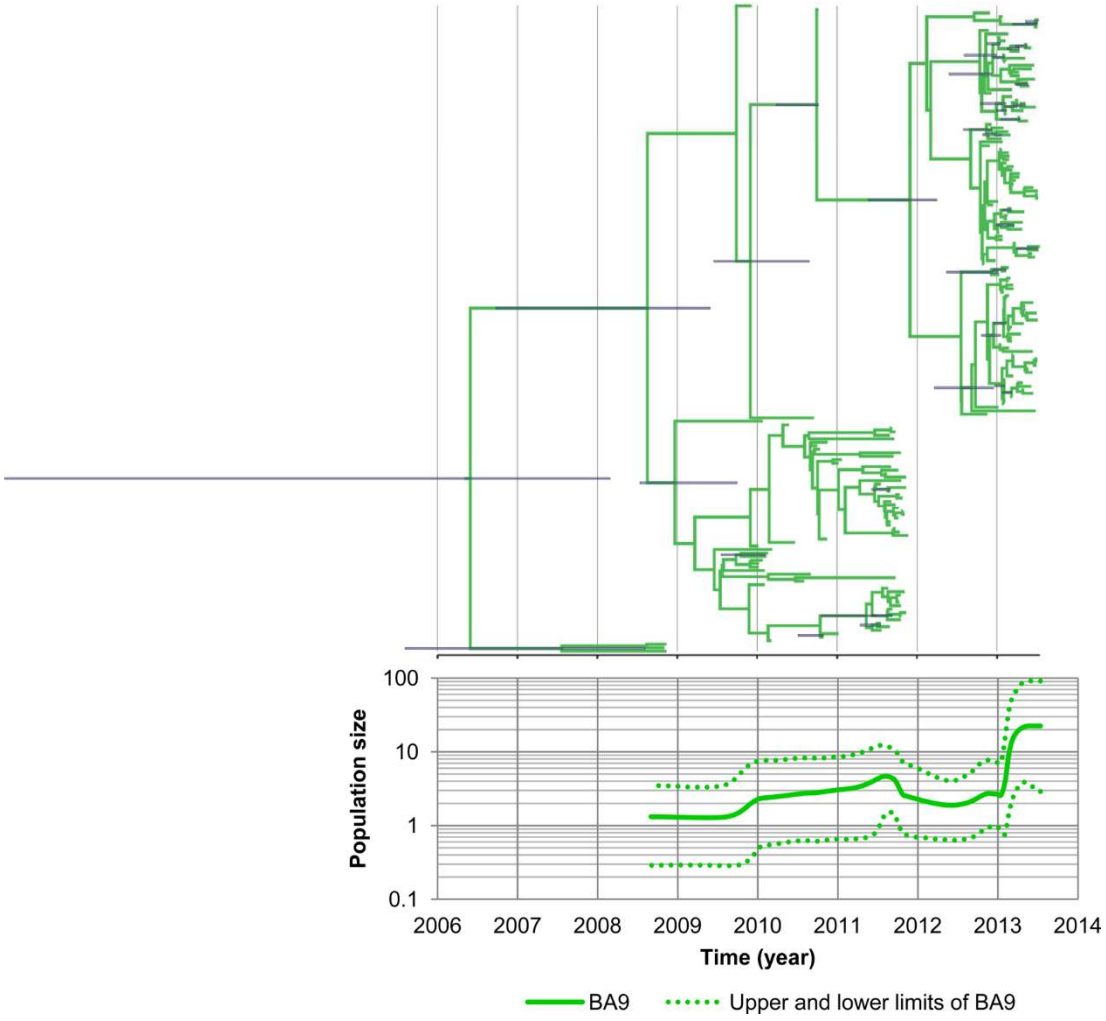


Fig. 16 Phylodynamic analysis of HRSV-B from 2008 to 2013. Dated phylogeny (upper) and Bayesian skyline plot (lower) of the BA9 genotype. The population size of the BA9 genotype showed an increasing trend from 2009 and a sharp increase in 2013. (n of BA9 = 186)

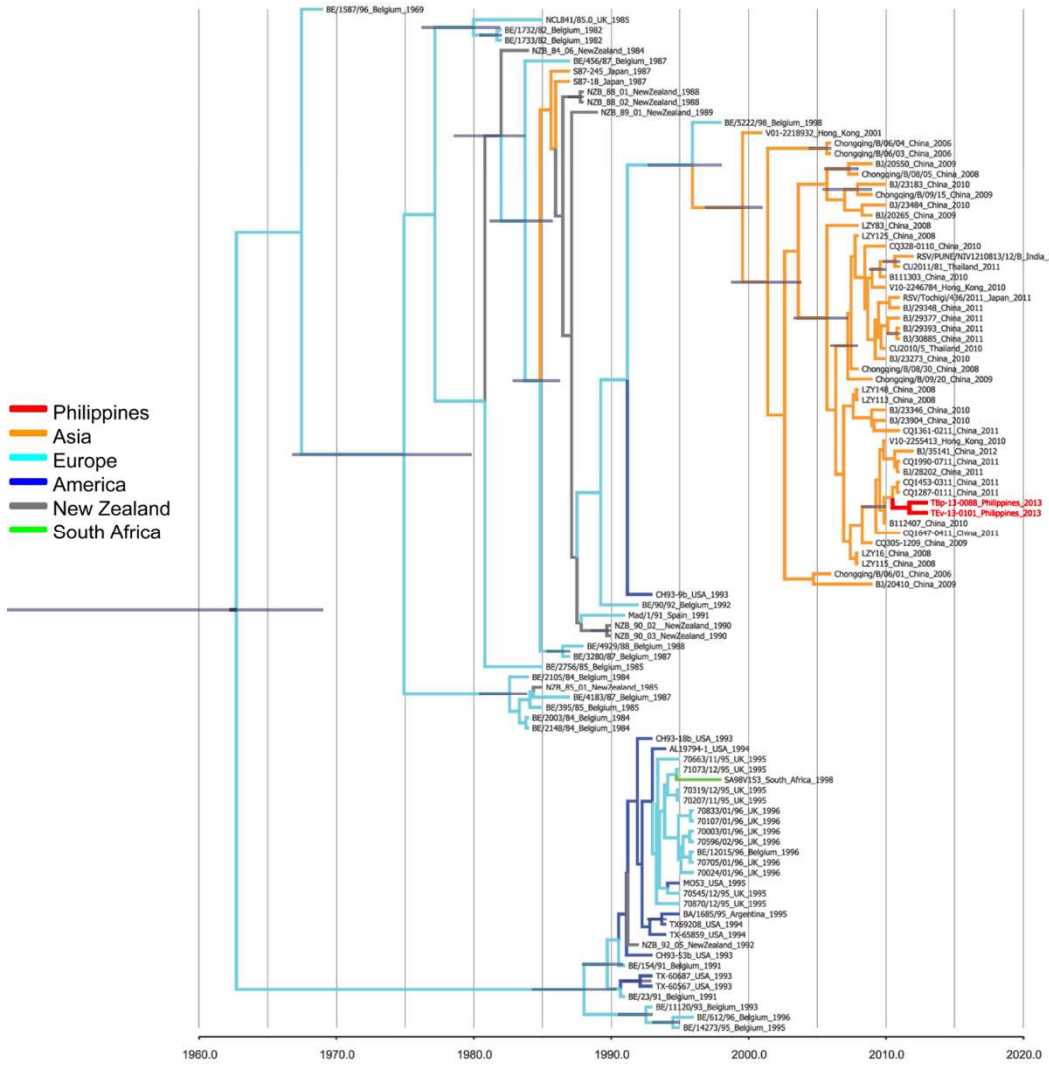


Fig. 17 Phylogenetic analysis of GB2 genotype from the Philippines. The phylogenetic tree was constructed by adding 98 strains from other countries that were identified by BLASTn search. The Philippine GB2 strains may share a common ancestor with strains from Asia. (n = 100)

TABLES

Table 1 The functions of all 11 HRSV proteins

Proteins	Functions	References
Nonstructural-1 (NS1)	Acts cooperatively with NS2 to repress activation and nuclear translocation of host IFN-regulatory factor IRF-3. Plays a major role in antagonizing the type I IFN-mediated antiviral response. Interacts with host MAVS and prevents the interaction with its upstream partner DDX58/RIG-I in the signaling pathway leading to interferon production. Mediates the proteasomal degradation of host STAT2 with elongin-cullin E3 ligase. Suppresses premature apoptosis by an NF-kappa-B-dependent, interferon-independent mechanism and thus facilitates virus growth. Additionally, NS1 may serve some inhibitory role in viral transcription and RNA replication.	81-83, 85, 92
Nonstructural-2 (NS2)	Plays a major role in antagonizing the type I IFN-mediated antiviral response. Acts cooperatively with NS1 to repress activation and nuclear translocation of host IFN-regulatory factor IRF-3. Interacts with the host cytoplasmic sensor of viral nucleic acids DDX58/RIG-I and prevents the interaction with its downstream partner MAVS. Mediates the proteasomal degradation of host STAT2 with Elongin-Cullin E3 ligase. Induces activation of NF-kappa-B. Suppresses premature apoptosis by an NF-kappa-B-	78, 81, 82, 85, 93

Proteins	Functions	References
	dependent, interferon-independent mechan-ism and thus facilitates virus growth. May also inhibit viral transcription and RNA replication.	
Nucleoprotein (N)	Encapsidates the genome, protecting it from nucleases. The nucleocapsid (NC) has a helical structure. The encapsidated genomic RNA is termed the NC and serves as template for transcription and replication. During replication, encapsidation by protein N is coupled to RNA synthesis and all replicative products are resistant to nucleases.	79
Phosphoprotein (P)	Acts as a cofactor serving both to stabilize the protein L and to place the polymerase complex on the N protein: RNA template.	4
Matrix (M)	The matrix interacts with the RNP complex and this association serves two functions: facilitate virion assembly and inhibit the viral transcriptase activity. Has a crucial role in virus assembly and budding. Early in infection, M is localized to the nucleus and may inhibit host cell transcription. Later on, M can associate with lipid rafts supposely by interacting with the cytoskeleton and with the cytoplasmic tail of glycoprotein G. The binding of M to host membrane is stabilized by the surface expression of the viral glycoproteins. These interactions may allow	84

Proteins	Functions	References
	virus formation by mediating association of the nucleocapsid with the nascent envelop.	
Small hydrophobic (SH)	May form a proton-selective ion channel, playing a role in budding and /or during virus entry. May also play a role in counteracting host innate immunity.	80
Glycoprotein (G)	Attaches the virion to the host cell membrane by interacting with heparan sulfate, initiating the infection. Interacts with host CX3CR1, the receptor for the CX3C chemokine fractalkine, to modulate the immune response and facilitate infection. Unlike the other paramyxovirus attachment proteins, lacks both neuraminidase and hemagglutinating activities. Secreted glycoprotein G helps RSV escape antibody-dependent restriction of replication by acting as an antigen decoy and by modulating the activity of leukocytes bearing Fc gamma receptors	4
Fusion (F)	During virus entry, induces fusion of viral and cellular membranes leading to delivery of the nucleocapsid into the cytoplasm. The fusogenic activity is inactive until entry into host cell endosome, where a furin-like protease cleaves off a small peptide between F1 and F2. Interacts directly with heparan sulfate and may participates in virus	94-98

Proteins	Functions	References
	attachment. Furthermore, the F2 subunit was identifed as the major determinant of RSV host cell specificity. Later in infection, proteins F expressed at the plasma membrane of infected cells can mediate fusion with adjacent cells to form syncytia, a cytopathic effect that could lead to tissue necrosis. The fusion protein is also able to trigger p53-dependent apoptosis.	
M2 ORF 1 protein (M2-1)	Acts as a transcriptional elongation factor to prevent premature termination during transcription thus allowing complete synthesis of RSV mRNAs. Functions also as a processivity and antitermination factor to permit transit of the polymerase through intergenic regions to access promoter distal genes. Plays a role in the association of the matrix protein with the nucleocapsid, which initiates assembly and budding. Also, can activate NF-kappa-B through association with host RELA.	99-101
M2 ORF 2 protein (M2-2)	Mediates the regulatory switch from transcription to RNA replication. Acts late in infection by inhibiting viral transcription and up-regulating RNA replication.	17
Large (L)	Displays RNA-directed RNA polymerase, mRNA guanylyl transferase, mRNA (guanine-N(7)-)-methyltransferase and poly(A) synthetase activities. The viral mRNA guanylyl transferase displays a different	102

Proteins	Functions	References
	biochemical reaction than the cellular enzyme. The template is composed of the viral RNA tightly encapsidated by the nucleoprotein (N). Functions either as transcriptase or as replicase. The transcriptase synthesizes subsequently the subgenomic RNAs, assuring their capping and polyadenylation by a stuttering mechanism. The replicase mode is dependent on intracellular protein N concentration. In this mode, the polymerase replicates the whole viral genome without recognizing the transcriptional signals (By similarity).	

Table 2 Incidence of HRSV in the Philippines from June 2012-December 2015; the 2012/2013 season from June 2012-Demcember 2013, the 2014 season from January 2014-January 2014, and the 2015 season from February 2015-December 2015. Surveillance was conducted in four study sites in 2012/2013 season while in one study site in 2014 and 2015 seasons.

Season	Total sample	HRSV (%)	Subgroup (%)		HRSV-A			HRSV-B	
			A	B	NA1 (%)	ON1 (%)	ON1-Bi (%)	GB2 (%)	BA9 (%)
2012/2013	1614	465 (28.8)	329 (70.8)	136 (29.2)	102 (31.0)	227 (69.0)	-	3 (2.2)	133 (97.8)
2014	387	121 (31.3)	93 (76.9)	28 (23.1)	-	80 (86.0)	13 (14.0)	-	28 (100.0)
2015	307	61 (19.9)	2 (3.3)	59 (96.7)	-	2 (100.0)	-	-	59 (100.0)
Overall	2308	647 (28.0)	424 (65.5)	223 (34.5)	102 (24.1)	309 (72.9)	13 (3.1)	3 (1.3)	220 (98.7)

Table 3 N-glycosylation patterns of HRSV-A among three epidemic seasons. Six patterns were shown with "N" for possibly attaching glycan to nitrogen atom of amino acid asparagine and "-" for non-glycosylation. A number of strains having patterns were counted for each epidemic season and genotype. Note: this analysis was performed in Biliran Island.

Amino acid position*			Epidemic season				
237	251	318	2013		2014		2015
			NA1	ON1	ON1	ON1-Bi	ON1
-	-	-	14				
-	-	N	6				1
-	N	N	3				
N	N	N	1				
N	-	-			1		
N	-	N		64	79	13	1

* Amino acid position 318 in ON1 shown in the table was equal to position 294 in NA1.

Table 4 N-glycosylation patterns of HRSV-B among three epidemic seasons. Nine patterns were shown with "N" for possibly attaching glycan to nitrogen atom of amino acid asparagine and "-" for non-glycosylation. A number of strains having patterns were counted for each epidemic season and genotype.

Amino acid position*								Epidemic season			
243	258	265	275	279	287	296	310	2013		2014	2015
								GB2	BA9	BA9	BA9
-	N	-	-	-	-	N	N	2			
-	-	-	N	-	-	N	N		2		
N	-	-	-	-	-	N	N			1	1
-	-	-	-	-	-	N	-			4	5
-	-	N	-	-	-	N	N				12
-	-	-	-	N	-	N	N				1
-	-	-	-	-	N	N	N				21
-	-	-	-	-	N	-	N				1
-	-	-	-	-	-	N	N		82	23	18

* Amino acid positions 287, 296, and 310 in BA9 shown in the table was equal to position 267, 276, and 290 in GB2, respectively.

Table 5 O-glycosylation patterns of HRSV-A among three epidemic seasons. Twenty-one patterns were shown with “O” for possibly attaching glycan to oxygen atom of amino acids Serine or Threonine and “-” for non-glycosylation. A number of strains having patterns were counted for each epidemic season and genotype.

Amino acid position*																																															Epidemic season				
219	220	222	227	228	231	235	238	239	241	243	245	246	249	250	252	253	259	260	264	267	268	269	270	272	275	277	281	282	283	**288**	**291**	**292**	**293**	**294**	**299**	**301**	**305**	**306**	**307**	311	313	315	316	317	319	320	2013		2014		2015
																																															NA1	ON1	ON1	ON1-Bi	ON1
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O											O	O	O	O	O	O	O	10				
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	Gap										O	O	O	O	O	O	O	13				
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O											O	O	O	O	O	O	-	1				
O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O		1			
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O		4			
O	O	-	O	O	O	O	O	O	O	-	-	O	O	O	O	-	O	O	O	O	O	O	-	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O		2			
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	-	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O		1			
O	O	-	O	O	O	O	O	O	O	-	-	O	O	O	O	-	O	O	O	O	O	O	-	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O		1			
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O		1			
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O			10		
O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O			1		
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O			1		
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O			1		1
O	O	-	O	O	O	O	-	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O			1		
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O	-			1		
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O		-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O			5		
O	O	-	O	O	O	O	-	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	-	-	O	O	O	O			1		
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	-	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O			1		
O	O	-	O	-	O	O	O	O	O	-	-	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O			1		
O	O	-	O	O	O	O	O	O	-	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O					1
O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O		54	57	13	

* Amino acid positions 311, 313, 315, 316, 317, 319, and 320 in ON1 shown in the table was equal to position 287, 289, 291, 292, 293, 295, and 296 in NA1, respectively. The amino acid positions 288, 291, 292, 293, 294, 301, 305, 306, and 307 with bold shown in table were in duplication site. White gap box for aligned gap is indicated.

Table 6 O-glycosylation patterns of HRSV-B among three epidemic seasons. Thirty-seven patterns of were shown with "O" for possibly attaching glycan to oxygen atom of amino acids Serine or Threonine and "-" for non-glycosylation. A number of strains having patterns were counted for each epidemic season and genotype.

Amino acid position*																																																				Epidemic season			
218	220	222	227	228	229	232	236	239	240	242	244	245	246	247	249	250	254	255	256	257	**260**	**262**	**264**	**265**	**266**	**267**	**269**	**270**	**274**	**275**	**276**	**277**	280	281	285	288	289	290	291	294	297	298	300	302	303	304	307	308	309	311	312	2013		2014	2015
																																																				GB2	BA9	BA9	BA9
O	O	O	-	O	O	O	O	-	O	-	O	O	O	O	O	O	O	O	O	-	Gap												O	-	O	O	O	O	-	O	O	O	O	O	O	-	O	O	O	O	O	2			
O	-	O	O	O	O	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O		4	1	
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	-	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	-	-	O	O	O	O	O	-	O	O	O	O	O	O		3		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	-	O	-	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O		3		
O	-	O	O	O	-	O	O	O	O	-	O	-	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	-	-	O	O	O	O	O	-	O	O	O	O	O	O		3		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	-	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	-	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	-	-	O	O	O	O	O	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	-		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	-	-	O	O	O	O	O	-	O	O	O	O	O	-		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	-	-	O	O	O	O	-	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	-	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	-	-	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	-	-	O	O	O	O	O	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	-	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	O	-	O	O	-	-	O	O	O	O	O	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	-	O	O	O	-	O	-	O	O	O	O	-	-	O	O	O	O	O	O		1		
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	-	-	O	O	O	O	O	O			1	
O	-	O	-	O	-	O	O	O	O	-	O	O	O	-	O	O	-	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O			4	
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	-	-	O	O	O	O	O	O			4	
O	-	O	-	O	-	O	O	O	O	-	O	O	O	-	O	O	-	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O			8	
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O			1	
O	-	O	-	O	-	O	O	O	O	-	O	O	O	-	O	O	-	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	-			1	
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	-	O	O	-	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	-			1	
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	-			3	
O	-	O	-	O	-	O	O	O	O	O	O	O	O	-	O	O	-	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O				1
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	-				5
O	-	O	O	O	-	O	O	O	-	-	O	O	O	-	O	O	-	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O				1
O	-	O	O	O	-	O	O	O	O	-	O	-	O	-	O	O	-	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O				2
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	-	O	O				1
O	-	O	-	O	-	O	O	O	O	-	O	O	O	-	O	O	-	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	O	O				1
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	-	O	O	O	-	O	O	O	O	O	O	O	O	O	-	-	O	O	O	O	O	-	O	O	O	O	O	O				11
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	-	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O				37
O	-	O	O	O	-	O	O	O	O	-	O	O	O	-	O	O	O	O	O	O	O	-	O	O	O	O	O	-	O	O	O	O	O	O	O	O	O	-	-	O	O	O	O	O	-	O	O	O	O	O	O		40	4	

* Amino acid positions 280, 281, 283, 288, 289, 290, 291, 294, 297, 298, 300, 302, 303, 304, 307, 308, 309, 311, and 312 in BA9 shown in the table was equal to position 260, 261, 265, 268, 269, 270, 271, 274, 277, 278, 280, 282, 283, 284, 287, 288, 289, 291 and 292 in GB2, respectively. The amino acid positions 260, 262, 264, 265, 266, 267, 269, 270, 274, 275, 276, and 277 in bold shown in table were in duplication site. White gap box for aligned gap is indicated.

Table 7 Parameters in the site model including Model parameter estimates, dN/dS ratios, positive selected sites with % posterior probability, log-likelihood scores (*l*), the Likelihood Ratio tests (LRT), and *p* for site model (n of HRSV-A = 183 ; n of HRSV-B = 171)

Subgroup	Model	dN/dS	Parameter estimates	Positive selected sites (% posterior probability)	l	LRT: 2Δl, d.f.	p
A	M0	0.4808	$\omega = 0.48079$	not allowed	-1208.82	M0 vs. M3: 18.06, df=4	0.0012
	M3	0.5835	$p_0 = 0.09248$ $p_1 = 0.68569$ $(p_2 = 0.22183)$ $\omega_0 = 0.29785$ $\omega_1 = 0.29787$ $\omega_2 = 1.58530$	K213G (50.2%), K223G (82.3%), V225A (81.4%), L226F/P (59.3%), D237N (54.5%), T245I/N/S (95.1%), L248F/I (88.2%), G272S (68.0%), Y273H/N (95.9%), L274P (97.5%), Y280H (81.2%), L298P (89.9%), V303A/I (65.5%), L310P (87.0%), S313I (78.9%), L314P (58.2%)	-1199.79		
	M1a	0.5291	$p_0 = 0.59200$ $(p_1 = 0.40800)$ $\omega_0 = 0.20451$ $\omega_1 = 1.00000$	not allowed	-1201.24	M1a vs. M2a: 0.82, df=2	0.66
	M2a	0.5837	$p_0 = 0.77839$ $p_1 = 0.00000$ $(p_2 = 0.22161)$ $\omega_0 = 0.29818$ $\omega_1 = 1.00000$ $\omega_2 = 1.58660$	T245I/N/S (57.8%), Y273H/N (63.5%)	-1200.83		
	M7	0.5993	$p = 0.38194$ $q = 0.25534$	not allowed	-1001.53	M7 vs. M8: 1.96, df=2	0.38
	M8	0.6641	$P_0 = 0.76163$ $(P_1 = 0.23837)$ $p = 47.86248$ $q = 96.75661$ $\omega = 1.72864$	T245I/N/S (69.2%), Y273H/N (74.6%), L274P (61.6%)	-1000.55		
B	M0	0.6281	$\omega = 0.62813$	not allowed	-1011.70	M0 vs. M3: 17.06, df=4	0.002
	M3	0.7579	$p_0 = 0.18097$ $p_1 = 0.79663$ $(p_2 = 0.02240)$ $\omega_0 = 0.00000$ $\omega_1 = 0.67213$ $\omega_2 = 9.93401$	V251A/G (99.7%), Y287H/N (96.1%)	-1003.17		
	M1a	0.5524	$p_0 = 0.59175$ $(p_1 = 0.40825)$ $\omega_0 = 0.24367$ $\omega_1 = 1.00000$	not allowed	-1008.39	M1a vs. M2a: 10.36, df=2	0.006
	M2a	0.7598	$p_0 = 0.68871$ $p_1 = 0.28996$ $(p_2 - 0.02133)$ $\omega_0 = 0.36166$ $\omega_1 = 1.00000$ $\omega_2 = 10.34951$	V251A/G (98.5%), Y287H/N (95.1%)	-1003.21		
	M7	0.5823	$p = 0.37846$ $q = 0.27143$	not allowed	-1008.72	M7 vs. M8: 11.06, df=2	0.004
	M8	0.7604	$P_0 = 0.97824$ $(P_1 = 0.02176)$ $p = 1.10045$ $q = 0.90049$ $\omega = 10.20322$	V251A/G (98.9%), Y287H/N (97.5%)	-1003.19		

SUPPLEMENTARY DOCUMENTS

Supplementary document 1:

Nucleotide sequences of G gene-specific primers used in real-time polymerase chain reaction for HRSV screening.

Primer name	Sequence (5'→3')	Position
Forward	GCTCTTAGCAAAGTCAAGTTRAATGATACA	1143-1172
Reverse[a]	GTTTYTGCACATCATAATTRGGAGT	1266-1242
Probe[b]	VIC-CTRTCATCCAGCAAATAYACYATCCAACGKAGYACAGG-MGB	1191-1228

Nucleotide position of primer annealing site based on HRSVA strain A2 (M74568)

[a, b] modified from Bonroy C, Vankeerberghen A, Boel A, De Beenhouwer H. Use of a multiplex real-time PCR to study the incidence of human metapneumovirus and human respiratory syncytial virus infections during two winter seasons in a Belgian pediatric hospital. Clin Microbiol Infect. 2007; 13: 504-9.

Supplementary document 2:

Primers used in hemi-nested PCR for detecting the 2nd HVR region of the G gene

Primer name*	Polar	Sequence (5'→3')
GPA	+	GAAGTGTTCAACTTTGTACC
GPB	+	AAGATGATTACCATTTTGAAGT
nRSAG	+	TATGCAGCAACAATCCAA
nRSBG	+	GTGGCAACAATCAACTCTGC
F1	-	CAACTCCATTGTTATTTGCC

*The first PCR using GPA/F1 for HRSV-A to get PCR product sizes of 487 bp for genotype without 72 nucleotide duplication or 559 bp for genotype with 72 nucleotide duplication

The first PCR using GPB/F1 for HRSV-B to get PCR product sizes of 507 bp for genotype without 60 nucleotide duplication or 567 bp for genotype with 60 nucleotide duplication

The second PCR using nRSAG/F1 for HRSV-A to get PCR product sizes of 459 bp for genotype without 72 nucleotide duplication or 531 bp for genotype with 72 nucleotide duplication

The second PCR using nRSBG/F1 for HRSV-B to get PCR product sizes of 459 bp for genotype without 60 nucleotide duplication or 519 bp for genotype with 60 nucleotide duplication

References: Trento A, *et al*., J Gen Virol 2003; Dapat IC, *et al*., J Clin Microbiol 2010; Peret TC, *et al*., J Gen Virol 1998.

Supplementary document 3:

List of primers for amplifying complete genome of HRSV-A

code[#]	Primer name	Polar	Sequence (5'→3')	Position
1F	110984SP-1F	+	ACG CGA AAA AAT GCG TAC AAC	1-21
2R	S1R	-	GCC ATA TTT TGT RTT RTA TTC AG	872-894
3F	RSVA_NS_2F	+	CAC CAT TAG TTA ATA TAA AA	557-576
4R	110983SP-1RR	-	TGG ATG GTA TAT TTG CTG GA	1,197-1,216
5F	RSV-A_rt_outside_F	+	TGG AGC CTG AAA ATT ATA GTA	1,070-1,091
6R	RSV-A_rt_outside_R	-	ACA TGA TAT CCC GCA TCT CTG AG	1,386-1,408
7F	RSVA_NS_2R	+	CAC AGA AGA TGC TAA TCA TA	1,298-1,317
8R	RSV_NIIDAB_1stR_RSVAB2	-	GGG CTT TCT TTG GTT ACT TC	2,490-2,509
9F	RSVA_N_2F	+	AAC AGC TTC TAT GAA GTG TT	1,758-1,777
10F	111057SP-F3A for	+	GCA GAA GAA CTA GAG GCT ATC	2,253-2,273
11R	110979SP-3R	-	CCC TTG GGT GTG GAT ATT TG	3,441-3,460
12F	111175F4b-forF	+	GAA GCT ATG GCA AGA CTC AG	2,922-2,941
13R	110977SP-4R	-	AGG CCA GAA TTT GCT TGA GA	4,331-4,350
14F	M_in_F	+	CAC ACC CTG TGA AAT TAA GG	3,581-3,600

code[#]	Primer name	Polar	Sequence (5'→3')	Position
15F	P-5F	+	ACM AAC MCT CTG TGG TTC AA	4,105-4,124
16R	SH_end_R	-	ATG ATT GAG AGT GTC CCA G	4,735-4,753
17F	SH10-29	+	CAT CCA TCA CAA TAG AAT TC	4,313-4,332
18R	F164_modify	-	GTT ATA ACA CTR GTA TAC CAA CC	5,811-5,833
19R	F1	-	CAA CTC CAT TGT TAT TTG CC	5,650-5,669
20F	RSVA_FAF_5654_Venter_md (F5654)	+	TGG GGC AAA TAA CAA TGG AKT TGC	5,647-5,670
21R	RSVA_FAR_7422_Venter_md (F7422)	-	AAC YAT TGT AAG AAY ATG ATT ARGT	7,397-7,421
22R	F1 in R	-	GAT ACT GAT CCT GCA TT	6,693-6,709
23F	F1 in F	+	GAT ATG CCT ATA ACA AAT GA	6,447-6,466
24F	RSV_FF	+	CCA TGA CCA ACT CAA ACA GAA TC	5,613-5,635
25R	RSV_NIID_nestR_22K4	-	GCA ACA CAT GCT GAT TGT	7,881-7,998
26F	RSV_(7323-11375)_F	+	CAT CAA TAT CTC AAG TCA AYG AGA	7,129-7,152
27R	S7(9108)R	-	AAC CAT GAT GGA GGA TGT TGC AT	9,086-9,108
28R	S6-7R	-	GGG TTA TTG ATG GTT ATG CTC TTG	8,109-8,132
29F	S8F	+	CCY AAG ATA AGA GTG TAC AAT ACT G	7,972-7,996
30F	L_beg_F	+	AAT GGA TCC CAT TAT TAR TGG A	8,497-8,518

code[#]	Primer name	Polar	Sequence (5'→3')	Position
31R	P-9R	-	TTT ATT ATG TAG AAC CCC TCA TTG TG	9,470-9,495
32F	110965SP-10FF	+	CRA TGC AAC ATC CTC CAT CA	9,084-9,103
33R	S8RR	-	CTT AAA GTA GGC CAT CTG TTG	9,928-9,948
34F	S9(9715)F	+	ATT RAT TAA GCT TGC AGG TGA	9,715-9,735
35R	110964SP-10R	-	GGT TGC ATT GCA AAC ATT CTA	10,375-10,395
36F	110963SP-11FF	+	CGA GAG TTT CGG TTG CCT A	10,064-10,082
37R	L_mid_R	-	CAA ATA ATC TGC TTG AGC ATG RG	10,980-11,002
38F	S10F	+	CTA ATA TCT CTC AAA GGG AAA TTC	10,880-10,903
39R	110962SP-11R	-	GGG ATC ACC ACC ACC AAA TA	11,448-11,467
40F	110961SP-12F	+	AGT GGG ACC GTG GAT AAA CA	11,164-11,183
41R	110960SP-12RR	-	TGA CTG TAA GAC GAT GCA A	12,506-12,524
42F	L_in1_F	+	ACA TTG ATG AGA GAT CCT CA	11,663-11,682
43F	110959SP-13FF	+	TGG ACA TAA AAT ATA CWA CAA GCA	12,180-12,203
44R	110958SP-13R	-	TTA ACA ACC CAA GGG CAA AC	13,361-13,380
45F	L_in2_F	+	AAG TAT GGT GAT GAA GAT AT	12,620-12,639
46F	110957SP-14F	+	AAA AAG ATT GGG GAG AGG GAT A	13,041-13,062

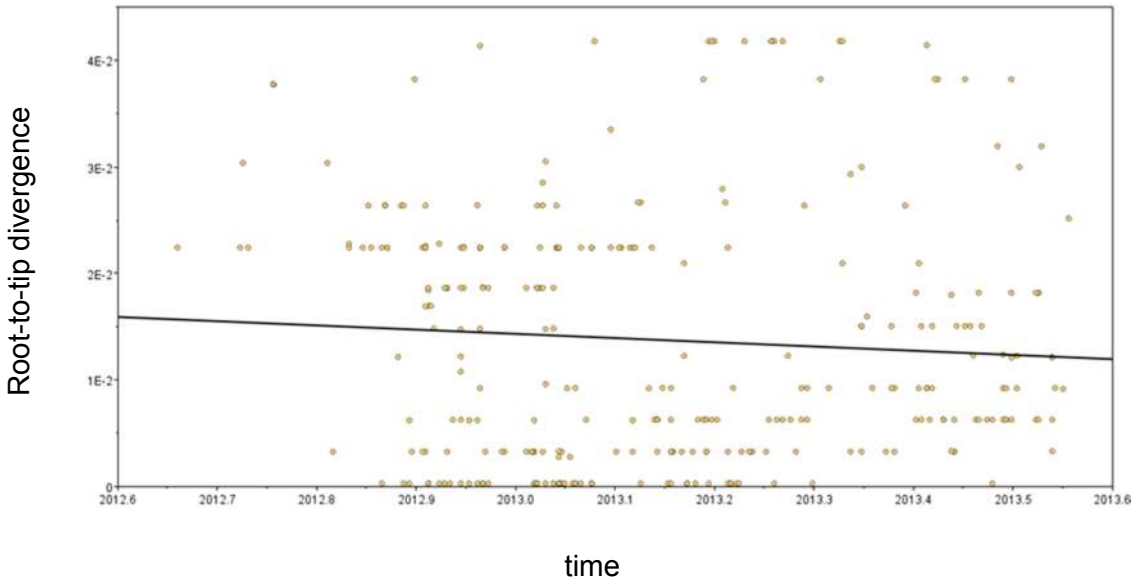
code[#]	Primer name	Polar	Sequence (5'→3')	Position
47R	110956SP-14R	-	TGC AYT TTC TTA CAT GCT TGC	14,355-14,375
48F	L_in3_F	+	AAC AAA TTA TAT CAT CCY ACA	13,604-13,624
49F	110955SP-15F	+	GGT GAA GGA GCA GGG AAT TT	14,054-14,073
50R	110954SP-15R	-	ACG AGA AAA AAA GTG TCA AAA ACT	15,199-15,222
51F	L_in4_F	+	ATT GAT GCA AAT ATT AAR AGT TT	14,636-14,658

Simple codes of primers

IUPAC codes: B = C, G or T; H = A, C or T; N = any base; R = A or G; S = G or C; V = A, C or G; W = A or T; Y = C or T. Each primer position was aligned with HRSV-A strain A2 (M74568)

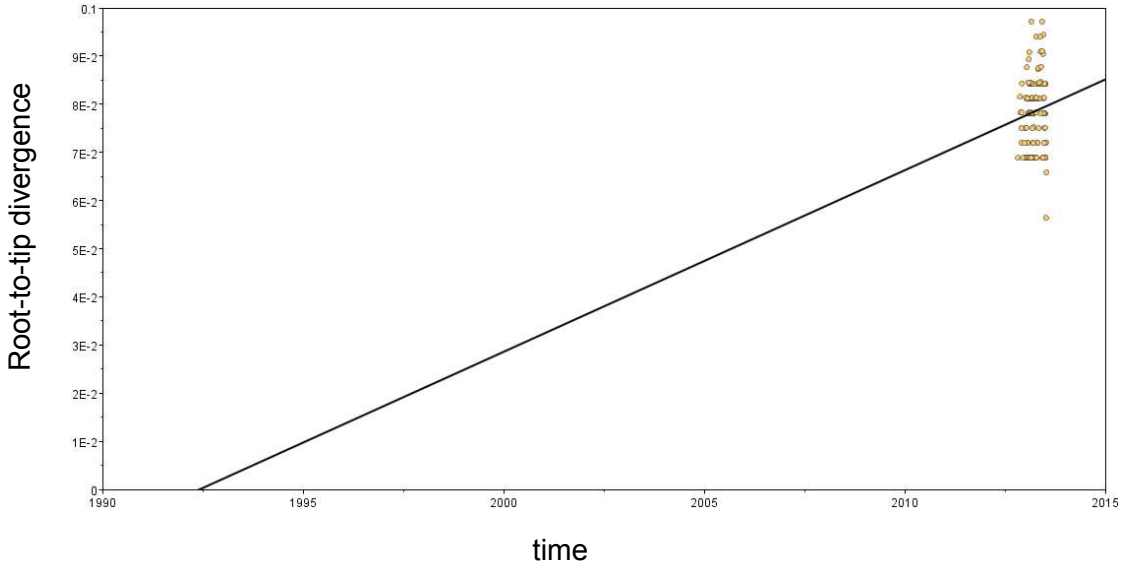
Supplementary document 4:

Root-to-tip linear regression of HRSV-A strains circulating from Jun 2012-July 2013. The correlation coefficient (r) value was -0.071 and the R-squared (r^2) value was 0.005.



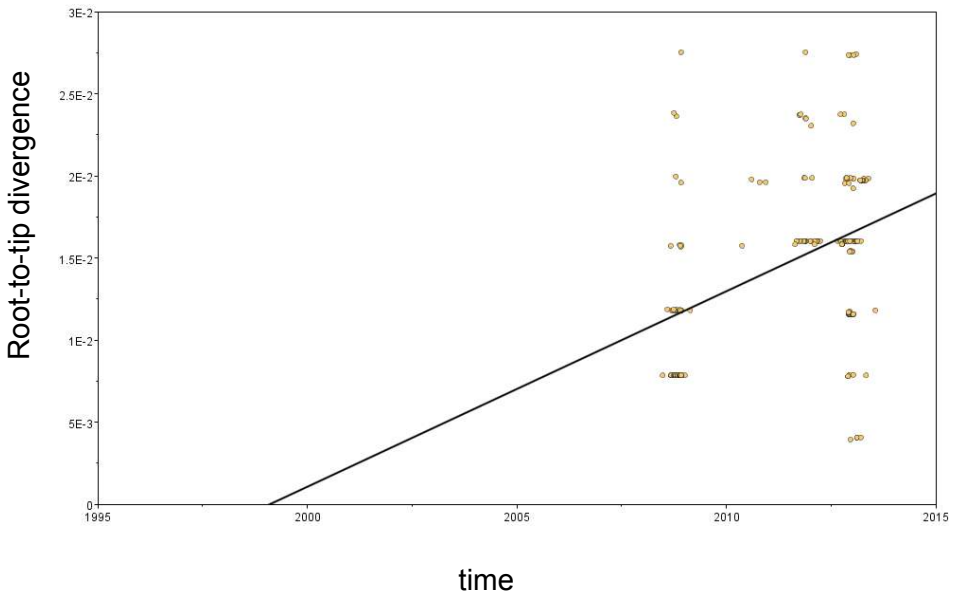
Supplementary document 5:

Root-to-tip linear regression of HRSV-B strains circulating from Jun 2012-July 2013. The *r* value was 0.089 and the r^2 value was 0.008.



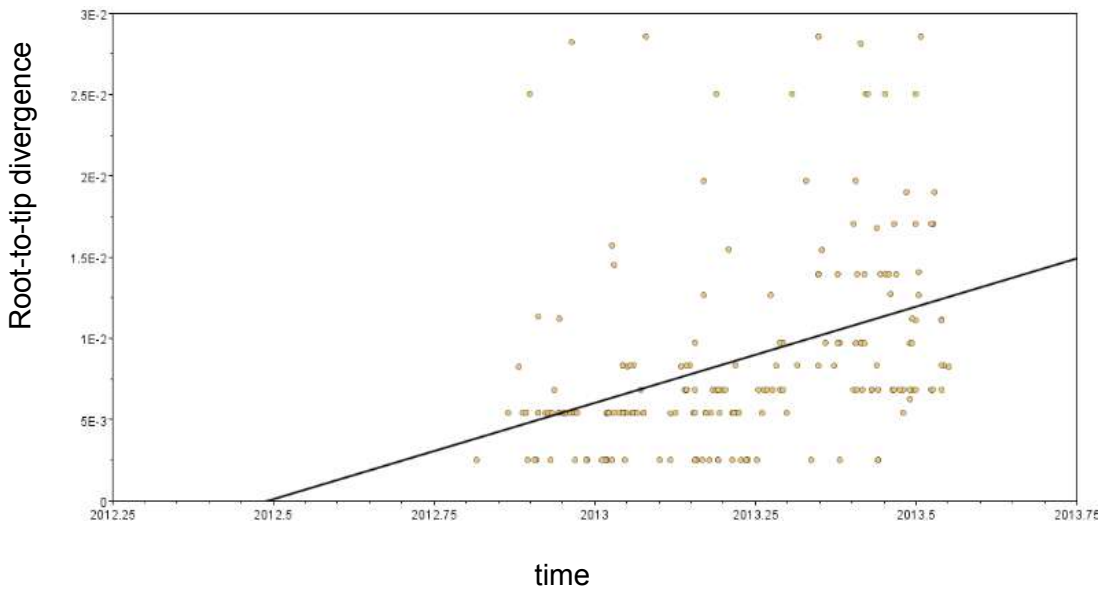
Supplementary document 6:

Root-to-tip linear regression of NA1 strains circulating from May 2008 – April 2012. The correlation coefficient (r) value was 0.4296 and the R-squared (r^2) value was 0.1846.



Supplementary document 7:

Root-to-tip linear regression of ON1 strains circulating from October 2012 – April 2012. The correlation coefficient (r) value was 0.3951 and the R-squared (r^2) value was 0.1561.



Supplementary document 8:

Root-to-tip linear regression of BA9 strains circulating from May 2008 - July 2013. The *r* value was -0.3722 and the r^2 value was 0.1386.