DOCTORAL THESIS

Molecular evolution of human respiratory syncytial virus attachment

glycoprotein (*G*) gene of new genotype ON1 and ancestor NA1.

（ RS ウイルス・遺伝子型 NA1 ならびに遺伝子型 ON1 における

付着糖タンパク（*G*）遺伝子の分子進化 ）

March, 2023

（ 2023 年 3 月 ）

Eiko Igarashi

五十嵐　映子

Department of Microbiology and Molecular Biodefense Research
Yokohama City University Graduate School of Medicine

横浜市立大学　大学院医学研究科 分子生体防御学

（　Doctoral Supervisor：Akihide Ryo, Professor　）

（　指導教員：梁　明秀 教授　）

Molecular evolution of human respiratory syncytial virus attachment glycoprotein (*G*) gene of new genotype ON1 and ancestor NA1 ☆

Eiko Hirano [a], Miho Kobayashi [b], Hiroyuki Tsukagoshi [b], Lay Myint Yoshida [c], Makoto Kuroda [d], Masahiro Noda [e], Taisei Ishioka [e], Kunihisa Kozawa [b], Haruyuki Ishii [f], Ayako Yoshida [g], Kazunori Oishi [e], Akihide Ryo [h,*], Hirokazu Kimura [e,h,*]

[a] *Fukui Prefectural Institute of Public Health and Environmental Science, 39-4 Harame-cho, Fukui-shi, Fukui 910-8851, Japan*
[b] *Gunma Prefectural Institute of Public Health and Environmental Sciences, 378 Kamioki-machi, Maebashi-shi, Gunma 371-0052, Japan*
[c] *Department of Pediatric Infectious Diseases, Institute of Tropical Medicine, Nagasaki University, 1-12-4 Sakamoto, Nagasaki-shi, Nagasaki 852-8523, Japan*
[d] *Pathogen Genomics Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan*
[e] *Infectious Disease Surveillance Center, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama-shi, Tokyo 208-0011, Japan*
[f] *Department of Respiratory Medicine, Kyorin University, School of Medicine, 6-20-2 Shinkawa, Mitaka-shi, Tokyo 181-8611, Japan*
[g] *Aomori Prefectural Institute of Public Health and Environment, 1-1-1 Higashitsukurimichi, Aomori-shi, Aomori 030-8566, Japan*
[h] *Department of Microbiology and Molecular Biodefense Research, Yokohama City University, Graduate School of Medicine, 3-9 Fukuura, Kanagawa-ku, Yokohama-shi, Kanagawa 236-0004, Japan*

ARTICLE INFO

ABSTRACT

We conducted a comprehensive genetic analysis of the C-terminal 3rd hypervariable region of the attachment glycoprotein (*G*) gene in human respiratory syncytial virus subgroup A (HRSV-A) genotype ON1 (93 strains) and ancestor NA1 (125 strains). Genotype ON1 contains a unique mutation of a 72 nucleotide tandem repeat insertion (corresponding to 24 amino acids) in the hypervariable region. The Bayesian Markov chain Monte Carlo (MCMC) method was used to conduct phylogenetic analysis and a time scale for evolution. We also calculated pairwise distances (*p*-distances) and estimated the selective pressure. Phylogenetic analysis showed that the analyzed ON1 and NA1 strains formed 4 lineages. A strain belonging to lineage 4 of ON1 showed wide genetic divergence (*p*-distance, 0.072), which suggests that it might be a candidate new genotype, namely ON2. The emergence of genotype NA1 was estimated to have occurred in 2000 (95% of highest probability density, HPD; 1997–2002) and that of genotype ON1 in 2005 (95% HPD; 2000–2010) based on the time-scaled phylogenetic tree. The evolutionary rate of genotype ON1 was higher than that of ancestral genotype NA1 (6.03×10^{-3} vs. 4.61×10^{-3} substitutions/site/year, $p < 0.05$). Some positive and many negative selection sites were found in both ON1 and NA1 strains. The results suggested that the new genotype ON1 is rapidly evolving with antigenic changes, leading to epidemics of HRSV infection in various countries.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Human respiratory syncytial virus (HRSV) of genus *Pneumovirus* in family *Paramyxoviridae* is a major causative agent of acute lower respiratory infections. Specifically, primary HRSV infections may be responsible for about 50–90% of bronchitis, bronchiolitis, and pneumonia cases in children under 2 years of age (Leung et al., 2005; Shay et al., 1999; Yorita et al., 2007). Moreover, HRSV reinfections can occur throughout life and may result in bronchitis and pneumonia in elderly people (Lee et al., 2013). Thus, the impact of the disease burden of HRSV infection may be comparable to that of influenza (Lee et al., 2013; Weiss and McMichael, 2004).

HRSV contains two major antigens, fusion protein (F protein) and attachment glycoprotein (G protein) (Collins and Crowe, 2006). F protein is conserved, while rapid evolution (mutation) may be seen in the C-terminal 3rd hypervariable region of the G protein (Melero et al., 1997). This variable region contains some epitopes that induce neutralizing antibodies (Palomo et al., 1991). Thus, the ability of HRSV to establish reinfections throughout life may be due to the evolution of G protein, similar to the reinfections that occur with influenza virus subtype A(H1N1) due to the evolution of hemagglutinin (HA) gene (Hall et al., 1991; Taubenberger and Kash, 2010). HRSV have been classified into two subgroups, HRSV-A and HRSV-B, by genetic analysis of G gene/G protein (Mufson et al., 1985). Furthermore, HRSV-A and -B have been subdivided into 11 genotypes (GA1-GA7, SAA1, NA1, NA2, and ON1) and 20 genotypes (GB1-4, BA1-10, SAB1-4, and URU1-2), respectively (Cui et al., 2013; Trento et al., 2006). Of them, HRSV-A genotype ON1 was initially detected in 2010 in Ontario, Canada (Eshaghi et al., 2012). The results of genetic analyses suggest that new genotype ON1 evolved from genotype NA1 (Eshaghi et al., 2012). Notably, a tandem replication (corresponding to 24 amino acid residues) of 72 nucleotides was found in the C-terminal 3rd hypervariable region of the G gene of ON1 strains (Eshaghi et al., 2012). Furthermore, this genotype is rapidly replacing other genotypes, such as NA1, in some countries (Agoti et al., 2014; Kim et al., 2014; Pierangeli et al., 2014; Tsukagoshi et al., 2013). Another HRSV-B genotype BA first emerged in 1999 in Buenos Aires, Argentina (Galiano et al., 2005) and rapidly spread to various countries resulting in the current prevalence of HRSV-B infections (Trento et al., 2010). A tandem duplication of 60 nucleotides (corresponding to 20 amino acid residues) in the G gene C-terminal 3rd hypervariable region was also found in this genotype. Previous reports have suggested that the ancestor of genotype BA was another genotype, namely GB3 of HRSV-B (Galiano et al., 2005). BA has been subdivided into 10 genotypes (BA1-10) during the last 15 years (Dapat et al., 2010; Trento et al., 2006). However, the molecular evolution of the G gene in ON1 is not precisely known.

In general, the evolution of the virus may be associated with nucleic acid type (DNA or RNA), genome structure, and genome size (Gago et al., 2009; Holmes, 2011; Sanjuán et al., 2010). To gain a better understanding of HRSV epidemics, it is essential to analyze the G gene, the major antigen coding gene. Therefore, we conducted detailed genetic analyses of the global molecular evolution of the G gene of new prevalent HRSV genotype ON1 and its ancestor, NA1.

2. Materials and methods

2.1. Clinical samples

We collected nasopharyngeal swabs from patients with acute respiratory infection after verbal informed consent was obtained from the patients or their guardians. Samples were obtained between January 2009 and December 2013 in Fukui Prefecture. Patients were mainly diagnosed as having bronchitis. The study protocol was approved by the Ethics Committee of the National Institute of Infectious Diseases (approval No. 417).

2.2. RNA extraction, RT-PCR, sequencing, and BLAST search

RNA extraction was performed using a QIAamp Viral RNA Mini kit (QIAGEN, Valencia, CA), according to the manufacturer's instructions. Primescript RT reagent kit (Takara Bio, Otsu, Japan) was used to synthesize cDNA. We performed PCR using Takara ExTaq (Takara Bio) to amplify a part of the G gene, as previously described (Peret et al., 1998). The primer sequences used for PCR were as follows: first PCR primer set, forward primer F1 (5′-CAA CTCCATTGTTATTGGC-3′), reverse primers GPA (5′-GAAGTGTTCA ACTTTGTACC-3′) and GPB (5′-AAGATGATTACCATTTTGAAG-3′); second PCR primer set, forward primer F1, reverse primers GSA (5′-AACCACCACCAAGCCCACAA-3′) and GSB (5′-AAAACCAACCAT CAAACCCAC-3′). PCR was carried out under the following conditions of 95 °C for 5 min, 30 cycles at 94 °C for 1 min, 50 °C for 1 min, and 72 °C for 2 min, followed by a final extension at 72 °C for 7 min.

Amplicons were purified with MinElute PCR purification kit (QIAGEN), and cycle sequenced by BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) using the second PCR primer sets. The products were purified with BigDye XTerminator® Purification Kit (Applied Biosystems). Sequence analysis was performed by an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems).

We conducted a BLAST search for the segments and genotyped all strains of HRSV-A. BLAST analyses indicated that there were 22 and 58 HRSV strains of ON1 and NA1 in the present samples, respectively. Of them, we omitted the strains with 100% nucleotide identity. Thus, we used 3 strains of ON1 and 20 strains of NA1 in this study.

2.3. Other strains used in this study

To estimate the global evolution of HRSV-A G gene in genotypes ON1 and its ancestor NA1, we obtained a comprehensive collection of the target region (C-terminal 3rd hypervariable region of G gene) from GenBank and the sequences were added to the dataset of the present strains. After alignment of the G gene sequences, we omitted the strains with 100% nucleotide identity. As a result, we analyzed 93 strains of genotype ON1 and 125 strains of NA1. Detailed data of the strains are shown in Table S1.

2.4. Calculation of pairwise distances

The nucleotide sequences of segments of the G gene (positions 658–894, 237 bp for strain AUS/A2/61, Genbank: Nucleotide M11486; positions 673–984, 309 bp for strain ON67-1210A, Genbank: Nucleotide JN257693) were aligned. The frequency distribution of pairwise distances (p-distances) among the strains of genotypes ON1 and NA1 were calculated using MEGA 6.0 (Tamura et al., 2013).

2.5. Phylogenetic analyses by the Bayesian Markov chain Monte Carlo and neighbor joining methods

Phylogenetic analyses by the Markov chain Monte Carlo (MCMC) method were performed as previously described (Kushibuchi et al., 2013). Briefly, we used Kakusan4 (http://www.fifthdimension.jp/products/kakusan/) to select the nucleotide substitution model (Tanabe, 2011). The datasets were analyzed by BEAST package program v1.7.5 (Drummond and Rambaut, 2007) under an uncorrelated lognormal relaxed clock model or a strict clock model (Drummond et al., 2006). The MCMC chains were run to achieve convergence with sampling every 1000 steps. Convergence was confirmed using Tracer v1.6.0 (http://tree.bio.ed.ac.uk/software/tracer/). We accepted parameters with effective samples size above 200 after 10% burn-in. The maximum clade credibility tree was generated by Tree Annotator v 1.7.4 after removing the first 10% of trees as burn-in and the phylogenetic tree was viewed in FigTree v.1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/). The evolutionary rates of genotypes NA1 and ON1 were also calculated as described above. Detailed conditions are shown in Table 1. We also constructed a phylogenetic tree by the neighbor

Table 1
Conditions for the estimation of the evolutionary time scale.

Genotype	No. of strain	Substitution model[a]	Clock model	Length of chain
ON1	93	HKY85-Γ	Lognormal relaxed clock	30,000,000
NA1	125	HKY85-Γ	Lognormal relaxed clock	15,000,000
All strains	236	HKY85-Γ	Strict clock	40,000,000

[a] HKY85: Hasegawa, Kishino and Yano 1985 model.

joining (NJ) method based on the analyzed region (Saitou and Nei, 1987). Evolutionary distances were estimated using Kimura's two-parameter method (Kimura, 1980). The reliability of the tree was estimated using 1,000 bootstrap replications.

2.6. Selective pressure analysis

To evaluate the selective pressures on the partial *G* gene among the strains, positive selection sites in each genotype were estimated by Datamonkey (http://www.datamonkey.org/), as previously described (Pond and Frost, 2005a). The synonymous (*dS*) and nonsynonymous (*dN*) substitution rates at every codon were calculated using the following three methods: single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), and internal fixed effects likelihood (IFEL). The cut-off *p*-value was set at 0.1 (Kushibuchi et al., 2013).

2.7. Statistical analysis

Statistical analyses were performed using Welch's test by EZR v1.24 (Kanda, 2013). A *p*-value of <0.05 was considered to be statistically significant.

3. Results

3.1. Phylogenetic analyses of the global ON1 and NA1 strains by MCMC and NJ methods

We constructed global time-scaled phylogenetic trees by the MCMC method using strains detected in various countries, including Japan, which were comprehensively collected from GenBank (Fig. 1a and b). We also constructed a phylogenetic tree by the NJ method to obtain a clear presentation of the genetic distances (Fig. 2). First, the phylogenetic trees obtained by MCMC method estimated that genotype ON1 diverged from genotype NA1 in 2005, while NA1 diverged from GA2 in 2000 (Table 2). An ancestor of all present strains could be dated back to 1953 (Fig. 1a and Table 2). These NA1 strains could be classified into 4 lineages (lineages 1–4, Fig. 1a). The new genotype ON1 strains were also classified into 4 lineages (lineages 1–4) on the phylogenetic trees constructed by both the MCMC and NJ methods (Figs. 1a and b and 2). The ON1 strains may have derived from lineage 1 of genotype NA1 (Fig. 1a). The ON1 strains belonging to lineage 1 (45 strains) were the dominant strains detected in many countries, namely Germany, Italy, Canada, Croatia, Thailand, Japan, South

Korea, Philippines, China, and South Africa. Strains of lineage 2 (21 strains) were detected in Germany, Italy, Japan, and South Korea. Lineage 3 strains (25 strains) were from Kenya, Germany, Japan, Italy, and Croatia. Notably, the 2 strains detected in Italy (1301-118RM: Genbank: Nucleotide KC858255, 1301-125RM: Genbank: Nucleotide KC858256) independently formed lineage 4.

Next, the estimated divergence times for each lineage of new genotype ON1 were April 2009 (lineage 1), February 2010 (lineage 2), April 2010 (lineage 3), and August 2010 (lineage 4) (Fig. 1b). In addition, the evolutionary rate of ON1 strains in the analyzed region was higher than that of NA1 (mean rate, 6.03×10^{-3} vs. 4.61×10^{-3} substitutions/site/year, $p < 0.05$) (Table 2). The results suggested that new genotype ON1 evolved rapidly and spread quickly throughout many countries.

3.2. The p-distance values and phylogenetic locations of lineages 1 to 3 of the ON1 strains
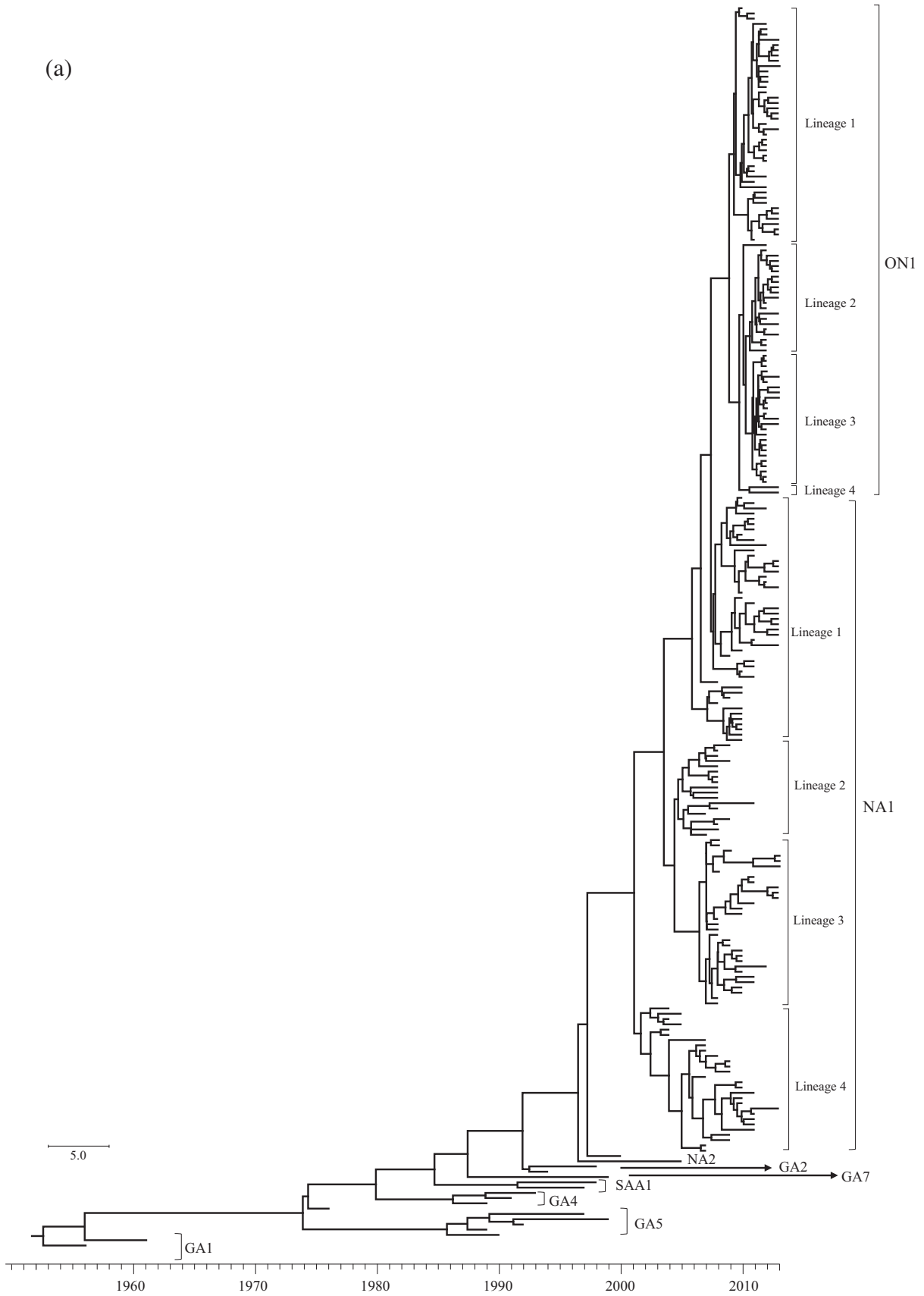
The distributions of the *p*-distances are shown in Fig. 3a and b. First, the mean values of the *p*-distances of genotypes ON1 and NA1 were relatively short (less than 0.025) and no significant differences were found; however, the distribution patterns differed. Next, the *p*-distance values of the present ON1 strains belonging to lineages 1–3 were <0.062.

3.3. New genotype candidate strain ON2 belonging to lineage 4 based on p-distance values and phylogenetic analyses by MCMC and NJ methods

It has been proposed that the assignment of each HRSV genotype corresponds to a *p*-distance value of less than 0.07 (Venter et al., 2001). To clearly demonstrate the *p*-distances of the present ON1 strains, we constructed a phylogenetic tree by the NJ method (Fig. 2). In the present study, two strains of ON1 formed an independent cluster as lineage 4 on the phylogenetic trees (Figs. 1a and b and 2). Of them, one strain (1301-118RM) had a large *p*-distance value of 0.072 (this value was calculated by the prototype ON1 strain, ON67-1210A). The other strain, 1301-125RM, had a relatively large *p*-distance value of 0.065. These results implicated that strain 1301-118RM is a candidate new genotype, namely ON2.

3.4. Positive and negative selection sites in the present strains

We estimated the positive and negative selection sites in the C-terminal 3rd hypervariable region of *G* gene in the present strains (Tables 3 and 4). In the ON1 strains, some sites under positive selection were found (Table 3). Of them, 2 substitutions were genotype specific (Asn251Asp, Asn251Tyr, or Asn251Ser; and Tyr297His). In particular, an amino acid substitution (Tyr297His) was found in ON1 strains in the newly inserted sites of the region (72 nt duplication). In the NA1 strains, 3 genotype specific sites under positive selection (Thr253Ile or Thr253Lys; Pro276Leu or Pro276Ser; and Thr296Ser, Thr296Ile, or Thr296Ala) were found (Table 3). Many sites under negative selection were found in both NA1 and ON1 strains (Table 4). These results suggested that

Table 2
Evolutionary rates and branched years of the analyzed HRSV-A genotypes.

Genotype	Mean rate[a] (95% HPD)	Branched year (95% HPD)
ON1	6.03×10^{-3} (3.43–9.10×10^{-3})[*]	2005 (2000–2010)
NA1	4.61×10^{-3} (3.33–5.98×10^{-3})[*]	2000 (1997–2002)
All strains[b]	5.36×10^{-3} (4.42–6.39×10^{-3})	1953 (1949–1956)

[*] $p < 0.05$.
[a] Substitutions/site/year.
[b] All strains are in the phylogenetic tree in this study.

Fig. 1. Phylogenetic trees for *G* gene of HRSV-A (a) and the expanded genotype ON1 (b) constructed by the Bayesian Markov chain Monte Carlo (MCMC) method. Scale bars represent unit of time (year). Gray bars indicate 95% highest probability density (HPD) for the branched year.

(b)

Expanded

Fig. 1 (*continued*)

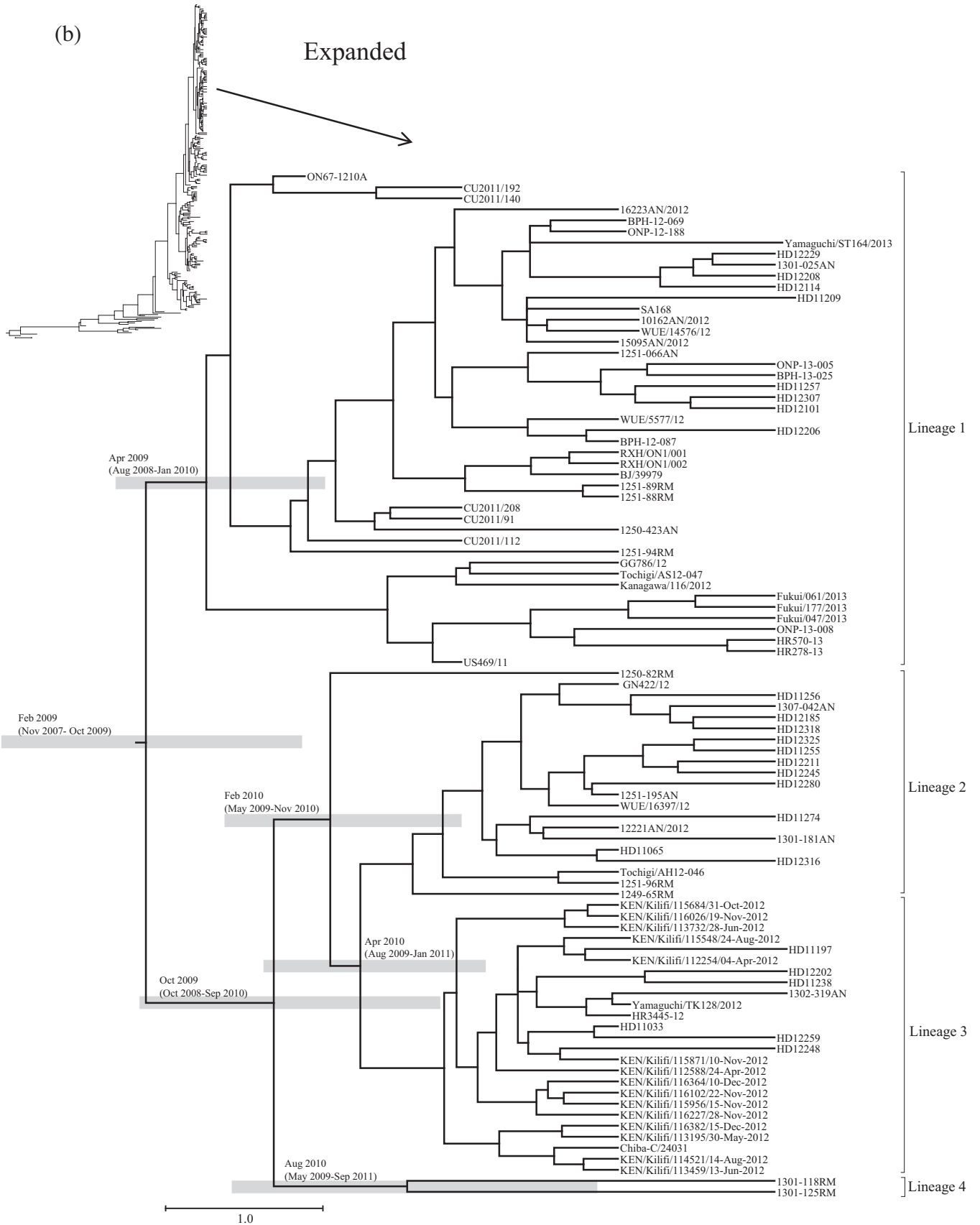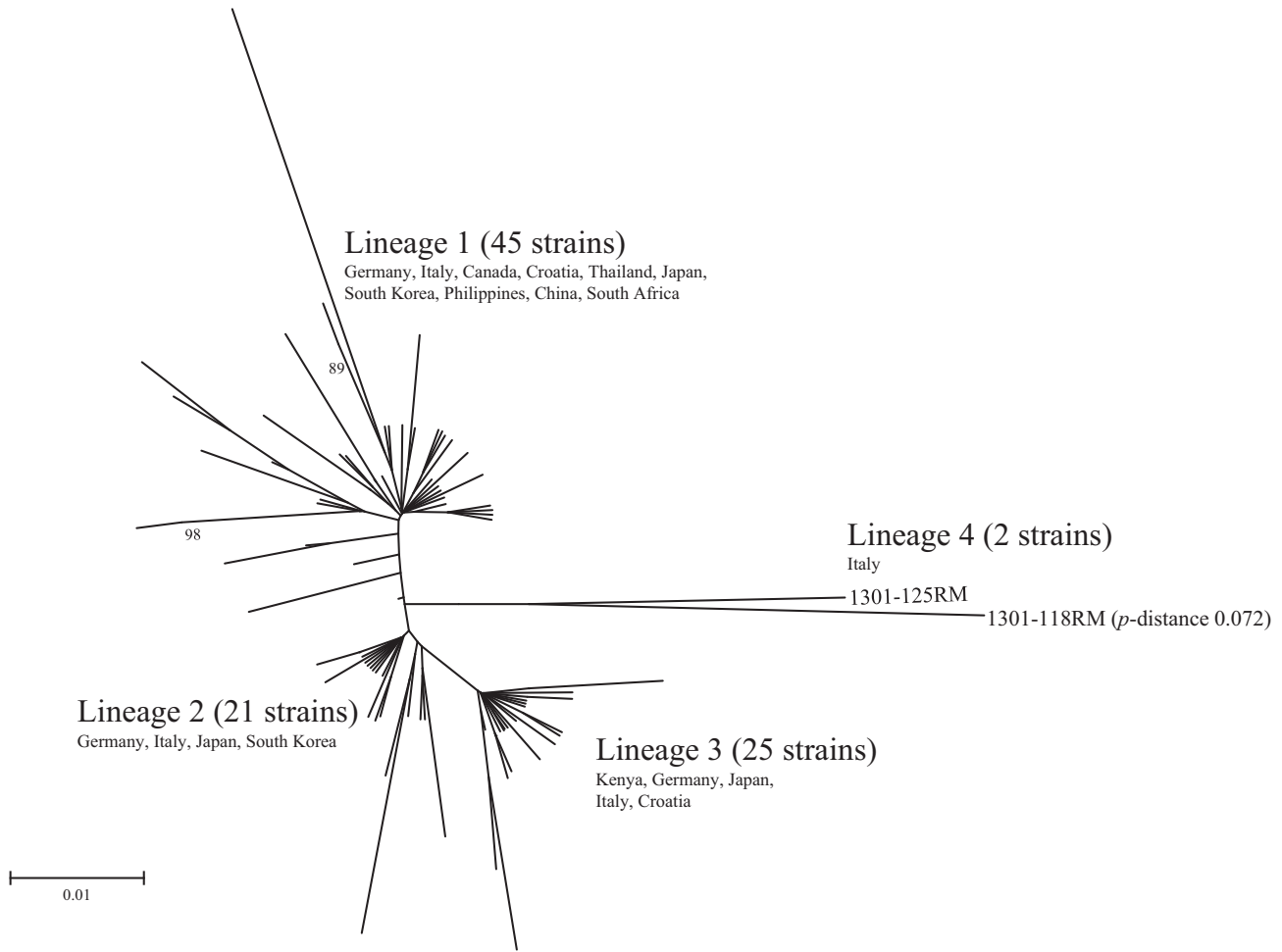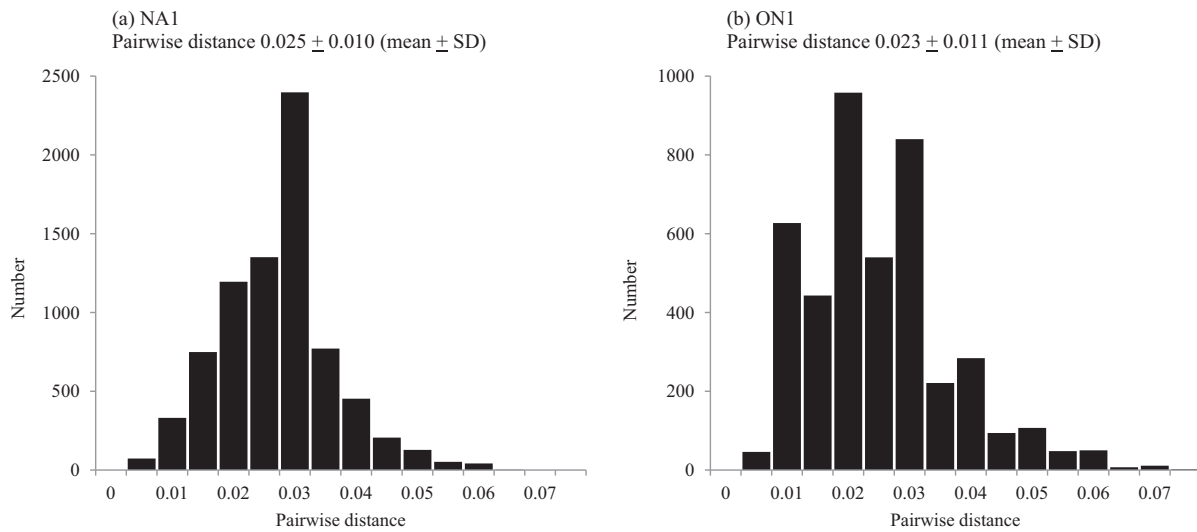**Fig. 2.** Phylogenetic tree for *G* gene of HRSV-A genotype ON1constructed by the neighbor joining (NJ) method. Labels at the branch nodes show at least 70% bootstrap support. Scale bar indicates nucleotide substitutions per site.

Fig. 3. Distribution of pairwise distances (*p*-distances) for HRSV-A genotype NA1 (a) and genotype ON1 (b) based on the nucleotide sequences of the *G* gene.

Table 3
Positive selection sites in C-terminal hypervariable region of G gene.

Genotype	Model	Positive selection site[a]	Mean dN/dS
ON1	SLAC	L274P, L274R	0.822
	FEL	N251D, N251Y, N251S, L274P, L274R, Y297H*	
	IFEL	S260N, Y273H, Y297H*	
NA1	SLAC	L274P	0.818
	FEL	T253I, T253K, S260N, S260I, N273Y, N273H, Y273N	
	IFEL	N273Y, N273H, Y273N, P276L, P276S, T296S, T296I, T296A	

p-Value < 0.1.
[a] Sites inside 72 duplication position are indicated asterisks.

Table 4
Negative selection sites in C-terminal hypervariable region of G gene.

Genotype	Model	Negative selection site[a]
ON1	SLAC	T227, P230, T231, T239, T245, S291*
	FEL	P222, E224, P230, T231, I243, T245, T259, S277, Q285, S291*, S307*
	IFEL	T231, S277, Q285, S291*
NA1	SLAC	T245, T268
	FEL	E232, T239, L248, G254, E257, L265, L266, T268, S270, T282, S287
	IFEL	L248, G254, T268, S270, S283

p-Value < 0.1.
[a] Sites inside 72 duplication position are indicated asterisks.

frequent amino acid substitutions have occurred in the analyzed region of genotypes NA1 and ON1.

4. Discussion

We studied the molecular evolution of the C-terminal 3rd hypervariable region of HRSV-A G gene in new genotype ON1 and its ancestor NA1. The phylogenetic trees with evolutionary time scales constructed by the MCMC method indicated that genotype ON1 diverged in 2005 from a lineage (lineage 1) of genotype NA1 (Fig. 1a). Four lineages of ON1 have emerged in various countries over some years and show a significantly rapid evolutionary rate in the analyzed region compared with genotype NA1 (Fig. 1a and Table 2). Furthermore, p-distance values and phylogenetic analyses suggested that a candidate new genotype (namely ON2), which likely derived from ON1, diverged in 2010. Some sites under positive selection and many under negative selection were found in both genotypes. The results suggested that new genotype ON1 is rapidly evolving with essential amino acid substitutions in the hypervariable region of the G gene.

Previous reports have deduced that new genotype ON1 of HRSV-A emerged in 2010 in Canada (Eshaghi et al., 2012) and may have been rapidly spreading and replacing another prevalent HRSV-A genotype, NA1, in many Asian, European, American, and African areas over a period of 3 years (Agoti et al., 2014; Auksornkitti et al., 2013; Cui et al., 2013; Eshaghi et al., 2012; Forcic et al., 2012; Kim et al., 2014; Pierangeli et al., 2014; Prifert et al., 2013; Tsukagoshi et al., 2013; Valley-Omar et al., 2013). This genotype has unique nucleotide insertions (72 nt duplication) in the C-terminal 3rd hypervariable region of HRSV-A G gene (Eshaghi et al., 2012). Genotype NA1, the likely ancestral strain of ON1, emerged in 2004 and became a prevalent type of HRSV-A infection in many areas in Asia, Europe, America, and Africa over a period of 10 years (Cui et al., 2013; de-Paris et al., 2014; Etemadi et al., 2013; Forcic et al., 2012; Khor et al., 2013; Kushibuchi et al., 2013; Pretorius et al., 2013; Rebuffo-Scheer et al., 2011;Shobugawa et al., 2009; Tran et al., 2013; Yamaguchi et al., 2011). A similar insertion has been confirmed in HRSV-B

genotype BA (60 nt, corresponding to 20 amino acid insertions) (Trento et al., 2003). This genotype emerged in 1999 in Argentina (Galiano et al., 2005). The ancestral strain is thought to be genotype GB3 of HRSV-B, a prevalent type during the 1990s (Galiano et al., 2005). Furthermore, many divergent genotypes (BA1-10) have evolved from genotype BA over a period of 10 years, and are prevalent types of HRSV-B (Cui et al., 2013; Dapat et al., 2010; Trento et al., 2006). Previous reports have suggested that such nucleotide insertions in the G gene might be linked to changes in the antigenicity of the G protein (Trento et al., 2003). Thus, the mutations seen in HRSV-A genotype ON1 and HRSV-B genotype BA may lead to future epidemics of HRSV infections.

The analyzed region of ON1 strains showed a high evolutionary rate in comparison with ancestral genotype NA1 (Table 2). In addition, lineage 4 of ON1 showed wide genetic divergence in the phylogenetic tree (Fig. 2). Notably, the p-distance of one strain was calculated to be 0.072, based on the sequence of the prototype ON1 (ON67-1210A). This strain was detected in Rome, Italy in 2013. Previous reports proposed that the genetic distance (p-distance) range in the same genotype was <0.07 (Cui et al., 2013; Venter et al., 2001). When we apply the p-distance value, this strain may be a candidate new genotype (namely ON2) that evolved from ON1 (Figs. 1b and 2). Moreover, this strain formed a unique lineage (lineage 4) that may have emerged in August 2010 as estimated by the present phylogenetic tree (Fig. 1b). However, since we analyzed and evaluated only a part of the G gene, additional analysis of the strain, including whole genome analysis, may be needed.

The evolution of antigens of various respiratory viruses may be involved in the infectivity toward the host, including the ability to establish reinfections (Domingo, 2006). For example, the rapid evolution of HA gene in seasonal influenza viruses, such as subtypes A(H3N2) or A(H1N1), is closely related to the ability of influenza to reinfect the host (Taubenberger and Kash, 2010). Similarly, the evolution of HRSV G gene might be associated with the ability of HRSV to reinfect humans (Botosso et al., 2009; Collins and Melero, 2011). In the present study, we analyzed the evolution of the G gene in the prevalent genotypes of HRSV-A. Rapid rates of evolution were found in the analyzed region in both ON1 and NA1 strains, although the rate of ON1 was faster than that of NA1 (Table 2). The C-terminal 3rd hypervariable regions are known to be involved in the function of epitopes against neutralizing antibodies (Palomo et al., 1991). Thus, the high evolutionary rate of the analyzed region of HRSV may be associated with the ability of the virus to reinfect the host (Botosso et al., 2009; Collins and Melero, 2011).

It has been suggested that the evolution of major antigens of various respiratory viruses including HRSV is associated with selective pressure in the host (Botosso et al., 2009). Furthermore, negative selection may be associated with preventing deterioration of viral functions (Domingo, 2006). Thus, we analyzed sites under positive and negative selection in the analyzed regions. In both ON1 and NA1 strains, some sites under positive selection were

found, with two unique sites found in the ON1 strains (Asn251Asp, Asn251Tyr, or Asn251Ser; and Tyr297His). Of them, Tyr297His was located in the new tandem duplication regions of the genotype. In this study, we analyzed sites under positive selection using SLAC, IFL, and IFEL methods (Botosso et al., 2009; Kushibuchi et al., 2013). SLAC is the more conservative of the three methods and appropriate for large alignments (Pond and Frost, 2005b). However, the number of positively selected sites may be underestimated (Pond and Frost, 2005b). In contrast, FEL and IFEL methods take synonymous and nonsynonymous rate variations into account and may be efficiently parallelized (Pond and Frost, 2005b). Thus, we used the three different methods to obtain an accurate estimate of sites under positive selection in the present study (Botosso et al., 2009; Kushibuchi et al., 2013).

Furthermore, many sites under negative selection were found in both ON1 and NA1 strains (Table 4). In general, negative selection may act to prevent deterioration of various viruses (Domingo, 2006). For example, the sites under negative selection in neutralization epitopes of polioviruses may be involved in receptor recognition and in the formation of altered particles (Domingo et al., 1993). Although the roles of many sites under negative selection in HRSV G protein are not exactly known, it is possible that these amino acid substitutions are involved in preventing the deterioration of antigenic function (Domingo, 2006; Kushibuchi et al., 2013).

5. Conclusion

A prevalent new genotype, ON1 of HRSV-A, with some positively selected amino acid substitutions emerged during a few years of rapid evolution. Although we analyzed only a part of the G gene, this genotype may have diverged to 4 lineages, including a lineage with new genotype ON2. Genotypes ON1 and ON2 may be potential agents of continuous epidemics of HRSV-A strains in the future.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.meegid.2014.09.030.

References

Agoti, C.N., Otieno, J.R., Gitahi, C.W., Cane, P.A., Nokes, D.J., 2014. Rapid spread and diversification of respiratory syncytial virus genotype ON1, Kenya. Emerg. Infect. 20, 950–959.

Auksornkitti, V., Kamprasert, N., Thongkomplew, S., Suwannakarn, K., Theamboonlers, A., Samransamruajkij, R., Poovorawan, Y., 2013. Molecular characterization of human respiratory syncytial virus, 2010–2011: identification of genotype ON1 and a new subgroup B genotype in Thailand. Arch. Virol. 159, 499–507.

Botosso, V.F., Zanotto, P.M., Ueda, M., Arruda, E., Gilio, A.E., Vieira, S.E., Stewien, K.E., Peret, T.C., Jamal, L.F., Pardini, M.I., Pinho, J.R., Massad, E., Sant'anna, O.A., Holmes, E.C., Durigon, E.L., VGDN Consortium, 2009. Positive selection results in frequent reversible amino acid replacements in the G protein gene of human respiratory syncytial virus. PLoS Pathog. 5, e1000254.

Collins, P.L., Crowe Jr., J.E., 2006. Respiratory syncytial virus and metapneumovirus. In: Knipe, D.M., Howley, P.M., Griffin, D.E., Martin, M.A., Lamb, R.A., Roizman, B., Straus, S.E. (Eds.), Fields Virology, fifth ed. Lippincott Williams & Wilkins, Philadelphia, pp. 1601–1646.

Collins, P.L., Melero, J.A., 2011. Progress in understanding and controlling respiratory syncytial virus: still crazy after all these years. Virus Res. 162, 80–99.

Cui, G., Zhu, R., Qian, Y., Deng, J., Zhao, L., Sun, Y., Wang, F., 2013. Genetic variation in attachment glycoprotein genes of human respiratory syncytial virus subgroups A and B in children in recent five consecutive years. PLoS One 8, e75020.

Dapat, I.C., Shobugawa, Y., Sano, Y., Saito, R., Sasaki, A., Suzuki, Y., Kumaki, A., Zaraket, H., Dapat, C., Oguma, T., Yamaguchi, M., Suzuki, H., 2010. New genotypes within respiratory syncytial virus group B genotype BA in Niigata, Japan. J. Clin. Microbiol. 48, 3423–3427.

de-Paris, F., Beck, C., de Souza, N.L., Machado, A.B., Paiva, R.M., da Silva, M.D., Pires, M.R., dos Santos, R.P., de Souza, K.R., Barth, A.L., 2014. Evaluation of respiratory syncytial virus group A and B genotypes among nosocomial and community-acquired pediatric infections in Southern Brazil. Virol. J. 11, 36.

Domingo, E., Diez, J., Martinez, M.A., Hernandez, J., Holguin, A., Borrego, B., Mateu, M.G., 1993. New observations on antigenic diversification of RNA viruses: antigenic variation is not dependent on immune selection. J. Gen. Virol. 74, 2039–2045.

Domingo, E., 2006. Virus evolution. In: Knipe, D.M., Howley, P.M., Griffin, D.E., Martin, M.A., Lamb, R.A., Roizman, B., Straus, S.E. (Eds.), Fields Virology, fifth ed. Lippincott Willams & Wilkins, Philadelphia, pp. 389–421.

Drummond, A.J., Ho, S.Y., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Drummond, A.J., Rambaut, A., 2007. BEAST: bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Eshaghi, A., Duvvuri, V.R., Lai, R., Nadarajah, J.T., Li, A., Patel, S.N., Low, D.E., Gubbay, J.B., 2012. Genetic variability of human respiratory syncytial virus A strains circulating in ontario: a novel genotype with a 72 nucleotide G gene duplication. PLoS One 7, e32807.

Etemadi, M.R., Sekawi, Z., Othman, N., Lye, M.S., Moghaddam, F.Y., 2013. Circulation of human respiratory syncytial virus strains among hospitalized children with acute lower respiratory infection in Malaysia. Evol. Bioinform. 9, 151–161.

Forcic, D., Ivancic-Jelecki, J., Mlinaric-Galinovic, G., Vojnovic, G., Babic-Erceg, A., Tabain, I., 2012. A study of the genetic variability of human respiratory syncytial virus in Croatia, 2006–2008. J. Med. Virol. 84, 1985–1992.

Gago, S., Elena, S.F., Flores, R., Sanjuán, R., 2009. Extremely high mutation rate of a hammerhead viroid. Science 323, 1308.

Galiano, M.C., Palomo, C., Videla, C.M., Arbiza, J., Melero, J.A., Carballal, G., 2005. Genetic and antigenic variability of human respiratory syncytial virus (groups A and B) isolated over seven consecutive seasons in Argentina (1995 to 2001). J. Clin. Microbiol. 43, 2266–2273.

Hall, C.B., Walsh, E.E., Long, C.E., Schnabel, K.C., 1991. Immunity to and frequency of reinfection with respiratory syncytial virus. J. Infect. Dis. 163, 693–698.

Holmes, E.C., 2011. What does virus evolution tell us about virus origins? J. Virol. 85, 5247–5251.

Kanda, Y., 2013. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. Bone Marrow Transplant. 48, 452–458.

Khor, C.S., Sam, I.C., Hooi, P.S., Chan, Y.F., 2013. Displacement of predominant respiratory syncytial virus genotypes in Malaysia between 1989 and 2011. Infect. Genet. Evol. 14, 357–360.

Kim, Y.J., Kim, D.W., Lee, W.J., Yun, M.R., Lee, H.Y., Lee, H.S., Jung, H.D., Kim, K., 2014. Rapid replacement of human respiratory syncytial virus A with the ON1 genotype having 72 nucleotide duplication in G gene. Infect. Genet. Evol. 26C, 103–112.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Kushibuchi, I., Kobayashi, M., Kusaka, T., Tsukagoshi, H., Ryo, A., Yoshida, A., Ishii, H., Saraya, T., Kurai, D., Yamamoto, N., Kanou, K., Saitoh, M., Noda, M., Kuroda, M., Morita, Y., Kozawa, K., Oishi, K., Tashiro, M., Kimura, H., 2013. Molecular evolution of attachment glycoprotein (G) gene in human respiratory syncytial virus detected in Japan 2008–2011. Infect. Genet. Evol. 18, 168–173.

Lee, N., Lui, G.C., Wong, K.T., Li, T.C., Tse, E.C., Chan, J.Y., Yu, J., Wong, S.S., Choi, K.W., Wong, R.Y., Ngai, K.L., Hui, D.S., Chan, P.K., 2013. High morbidity and mortality in adults hospitalized for respiratory syncytial virus infections. Clin. Infect. Dis. 57, 1069–1077.

Leung, A.K., Kellner, J.D., Davies, H.D., 2005. Respiratory syncytial virus bronchiolitis. J. Natl Med. Assoc. 97, 1708–1713.

Melero, J.A., García-Barreno, B., Martínez, I., Pringle, C.R., Cane, P.A., 1997. Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. J. Gen. Virol. 78, 2411–2418.

Mufson, M.A., Orvell, C., Rafnar, B., Norrby, E., 1985. Two distinct subtypes of human respiratory syncytial virus. J. Gen. Virol. 66, 2111–2124.

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

Peret, T.C., Hall, C.B., Schnabel, K.C., Golub, J.A., Anderson, L.J., 1998. Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. J. Gen. Virol. 79, 2221–2229.

Pierangeli, A., Trotta, D., Scagnolari, C., Ferreri, M., Nicolai, A., Midulla, F., Marinelli, K., Antonelli, G., Bagnarelli, P., 2014. Rapid spread of the novel respiratory syncytial virus A ON1 genotype, central Italy, 2011 to 2013. Euro. Surveill. 19, 20843.

Pond, S.L., Frost, S.D., 2005a. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21, 2531–2533.

Pond, S.L., Frost, S.D., 2005b. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol. Biol. Evol. 22, 1208–1222.

Pretorius, M.A., van Niekerk, S., Tempia, S., Moyes, J., Cohen, C., Madhi, S.A., Venter, M., SARI Surveillance Group, 2013. Replacement and positive evolution of

subtype A and B respiratory syncytial virus G-protein genotypes from 1997–2012 in South Africa. J. Infect. 208, S227–237.

Prifert, C., Streng, A., Krempl, C.D., Liese, J., Weissbrich, B., 2013. Novel respiratory syncytial virus A genotype, Germany, 2011–2012. Emerg. Infect. Dis. 19, 1029–1030.

Rebuffo-Scheer, C., Bose, M., He, J., Khaja, S., Ulatowski, M., Beck, E.T., Fan, J., Kumar, S., Nelson, M.I., Henrickson, K.J., 2011. Whole genome sequencing and evolutionary analysis of human respiratory syncytial virus A and B from Milwaukee, WI 1998–2010. PLoS One 6, e25468.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sanjuán, R., Nebot, M.R., Chirico, N., Mansky, L.M., Belshaw, R., 2010. Viral mutation rates. J. Virol. 84, 9733–9748.

Shay, D.K., Holman, R.C., Newman, R.D., Liu, L.L., Stout, J.W., Anderson, L.J., 1999. Bronchiolitis-associated hospitalizations among US children, 1980–1996. JAMA 282, 1440–1446.

Shobugawa, Y., Saito, R., Sano, Y., Zaraket, H., Suzuki, Y., Kumaki, A., Dapat, I., Oguma, T., Yamaguchi, M., Suzuki, H., 2009. Emerging genotypes of human respiratory syncytial virus subgroup A among patients in Japan. J. Clin. Microbiol. 47, 2475–2482.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Tanabe, A.S., 2011. Kakusan4 and Aminosan: two programs for comparing nonpartitioned, proportional and separate models for combined molecular phylogenetic analyses of multilocus sequence data. Mol. Ecol. Resour. 11, 914–921.

Taubenberger, J.K., Kash, J.C., 2010. Influenza virus evolution, host adaptation, and pandemic formation. Cell Host Microbe 7, 440–451.

Tran, D.N., Pham, T.M., Ha, M.T., Tran, T.T., Dang, T.K., Yoshida, L.M., Okitsu, S., Hayakawa, S., Mizuguchi, M., Ushijima, H., 2013. Molecular epidemiology and disease severity of human respiratory syncytial virus in Vietnam. PLoS One 8, e45436.

Trento, A., Casas, I., Calderón, A., Garcia-Garcia, M.L., Calvo, C., Perez-Breña, P., Melero, J.A., 2010. 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. 84, 7500–7512.

Trento, A., Galiano, M., Videla, C., Carballal, G., García-Barreno, B., Melero, J.A., Palomo, C., 2003. Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. J. Gen. Virol. 84, 3115–3120.

Trento, A., Viegas, M., Galiano, M., Videla, C., Carballal, G., Mistchenko, A.S., Melero, J.A., 2006. Natural history of human respiratory syncytial virus inferred from phylogenetic analysis of the attachment (G) glycoprotein with a 60-nucleotide duplication. J. Virol. 80, 975–984.

Tsukagoshi, H., Yokoi, H., Kobayashi, M., Kushibuchi, I., Okamoto-Nakagawa, R., Yoshida, A., Morita, Y., Noda, M., Yamamoto, N., Sugai, K., Oishi, K., Kozawa, K., Kuroda, M., Shirabe, K., Kimura, H., 2013. Genetic analysis of attachment glycoprotein (G) gene in new genotype ON1 of human respiratory syncytial virus detected in Japan. Microbiol. Immunol. 57, 655–659.

Valley-Omar, Z., Muloiwa, R., Hu, N.C., Eley, B., Hsiao, N.Y., 2013. Novel respiratory syncytial virus subtype ON1 among children, Cape Town, South Africa, 2012. Emerg. Infect. Dis. 19, 668–670.

Venter, M., Madhi, S.A., Tiemessen, C.T., Schoub, B.D., 2001. Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. J. Gen. Virol. 82, 2117–2124.

Weiss, R.A., McMichael, A.J., 2004. Social and environmental risk factors in the emergence of infectious diseases. Nat. Med. 10, S70–76.

Yamaguchi, M., Sano, Y., Dapat, I.C., Saito, R., Suzuki, Y., Kumaki, A., Shobugawa, Y., Dapat, C., Uchiyama, M., Suzuki, H., 2011. High frequency of repeated infections due to emerging genotypes of human respiratory syncytial viruses among children during eight successive epidemic seasons in Japan. J. Clin. Microbiol. 49, 1034–1040.

Yorita, K.L., Holman, R.C., Steiner, C.A., Effler, P.V., Miyamura, J., Forbes, S., Anderson, L.J., Balaraman, V., 2007. Severe bronchiolitis and respiratory syncytial virus among young children in Hawaii. Pediatr. Infect. Dis. J. 26, 1081–1088.

論文目録

I. 主論文
Molecular evolution of human respiratory syncytial virus attachment glycoprotein (G) gene of new

genotype ON1 and ancestor NA1.

Hirano, E., Kobayashi, M., Tsukagoshi, H., Yoshida, LM., Kuroda, M., Noda, M., Ishioka, T.,

Kozawa, K., Ishii, H., Yoshida, A., Oishi, K., Ryo, A., Kimura, H.:

Infect Genet Evol., 28:183-91, 2014

II. 副論文
なし

III. 参考論文

1 Surveillance of adenovirus D in patients with epidemic keratoconjunctivitis from Fukui Prefecture, Japan, 1995-2010.

Nakamura, M., **Hirano, E**., Kowada, K., Ishiguro, F., Yamagishi, Z., Adhikary, A.K., Hanaoka, N., Okabe, N., Taniguchi, K., Fujimoto, T.:

J Med Virol., 84(1):81-6, 2012

2   Molecular epidemiology of human metapneumovirus from 2005 to 2011 in Fukui, Japan.

Nakamura, M., **Hirano, E**., Ishiguro, F., Mizuta, K., Noda, M., Tanaka, R., Tsukagoshi, H., Kimura, H.:

Jpn J Infect Dis., 66(1):56-9, 2013

3   Molecular evolution of the hypervariable region of the attachment glycoprotein gene in human respiratory syncytial virus subgroup B genotypes BA9 and BA10.

Nagasawa, K., **Hirano, E**., Kobayashi, M., Ryo, A., Oishi, K., Obuchi, M., Ishiwada, N., Noda, M., Kuroda, M., Shimojo, N., Kimura, H.:

Infect Genet Evol., 36:217-23, 2015

4   Molecular evolution of the fusion protein (F) gene in human respiratory syncytial
     virus subgroup B.
     Kimura, H., Nagasawa, K., Kimura, R., Tsukagoshi, H., Matsushima, Y., Fujita, K.,
     Hirano, E., Ishiwada, N., Misaki, T., Oishi, K., Kuroda, M., Ryo, A.:
     Infect Genet Evol., 52:1-9, 2017

5   Molecular epidemiology of the attachment glycoprotein (G) gene in respiratory
     syncytial virus in children with acute respiratory infection in Japan in 2009/2010.
     Yoshida, A., Kiyota, N., Kobayashi, M., Nishimura, K., Tsutsui, R., Tsukagoshi, H.,
     Hirano, E., Yamamoto, N., Ryo, A., Saitoh, M., Harada, S., Inoue, O., Kozawa, K.,
     Tanaka, R., Noda, M.,
     Okabe, N., Tashiro, M., Mizuta, K., Kimura, H.:
     J Med Microbiol, 61(Pt 6), 820-9, 2012