Molecular Epidemiology of Respiratory Syncytial Virus Infections among Children with Acute Respiratory Symptoms in a Community over Three Seasons

Mizuho Sato,1 Reiko Saito,1 Takatsugu Sakai,1 Yasuko Sano,2 Makoto Nishikawa,3
Asami Sasaki,1 Yugo Shobugawa,1 Fumitake Gejyo,4 and Hiroshi Suzuki1*

Division of Public Health, Department of Infectious Control and International Medicine, and Division of Clinical Nephrology, Rheumatology, Respiratory Medicine and Infection Control and Prevention, Niigata University Graduate School of Medical and Dental Sciences, Sano Clinic, and Niigata Prefectural Institute of Public Health and Environmental Sciences, Niigata, Japan

Received 13 May 2004/Returned for modification 6 July 2004/Accepted 2 September 2004

To study the molecular epidemiology of respiratory syncytial virus (RSV) in a community, children with acute respiratory symptoms at a pediatric outpatient clinic in Niigata, Japan, were analyzed over three seasons from November 2001 to July 2004. Of 499 nasopharyngeal aspirate specimens, 185 (37.1%) were RSV positive, and only 8 (4.5%) of 177 patients were shown by the reverse transcription (RT)-PCR method to be reinfected. RSV infection occurred beginning in the early winter, and the rates declined in the spring. The predominant subgroup changed from A to B and returned to A over the three seasons. Phylogenetic analysis also revealed that multiple genotypes cocirculated each year, with genotype GA5 of subgroup A predominating in the 2001-2002 and the 2003-2004 seasons. A new genotype of subgroup B (named BA, according to the nomenclature for viruses) with a 60-nucleotide insertion in the second variable region of the attachment glycoprotein protein was predominant as an emerging strain in the 2002-2003 season, but this was not associated with new epidemiological or clinical features, unlike the cases of disease caused by other genotypes in the other seasons. In conclusion, our molecular analysis of RSV confirms that multiple genotypes cocirculate each year and that the genotype predominating may shift with the season. Support for determination of the genotype by RT-PCR as an effective tool for characterization of RSV circulation patterns in the community is provided.

Respiratory syncytial virus (RSV) is a major cause of serious acute lower respiratory disease in infants and young children and is found mainly in the late fall, winter, and spring (2, 8, 13, 15). Bronchiolitis and pneumonia are observed most frequently during the first few months of life (2, 8, 17). Almost all children are infected with RSV by age 2 years, and half have experienced two infections (8, 11, 17).

RSV strains have been classified into antigenic groups A and B (RSV-A and RSV-B, respectively) on the basis of the activities of the viruses with monoclonal antibodies directed against the attachment glycoprotein (G protein) (1, 7, 12, 19) and also by genetic analyses (9, 25, 26). G protein is the most variable RSV protein (10, 14, 19), and its C-terminal region (the second hypervariable region) accounts for strain-specific epitopes (3, 4, 6, 9, 14, 22, 23). The molecular epidemiology and evolutionary patterns of G protein have provided important information about the clinical and epidemiological features of RSV: several different genotypes cocirculate, and the one that predominates in a community every year may change (5, 21, 22, 28). However, the importance of strain diversity to the clinical and epidemiological features of RSV has yet to be elucidated in detail.

Our objectives in the present study were to clarify the molecular epidemiology of RSV in a community over three seasons, determine the relationship between genotypes and circulation patterns, and assess clinical features and repeated infections.

MATERIALS AND METHODS

Study population and clinical samples. The study was conducted over three seasons from November 2001 to July 2004 at one pediatric outpatient clinic in Niigata City, Japan. Niigata City is the prefectural capital of Niigata Prefecture, and its total population is approximately 0.5 million. The clinic had an average of 2,300 patient visits per month during the study periods. Children with acute lower respiratory symptoms, such as wheezing, cough, rhinorrhea, and fever, were evaluated for RSV infections; and their clinical data were recorded at the clinic. Informed consent was obtained from all patients or their families. This study was approved by the medical faculty ethics committee of Niigata University.

Nasopharyngeal aspirates were obtained from the patients and were stored at 4°C in the clinic. The specimens were transported to the Department of Public Health, Niigata University School of Medical and Dental Sciences, within 5 days of sampling and were kept frozen at −80°C until further examination.

RT-PCR and nucleotide sequencing. Viral RNA was extracted from 100-μl samples of the nasopharyngeal aspirates with 500 μl of Trizol (Invitrogen Corp., Carlsbad, Calif.) in 100 μl of chloroform and centrifuged. RNA from the upper aqueous phase was precipitated with 100% isopropanol and purified with ethanol. Reverse transcription (RT) was then performed for cDNA synthesis by using 3 μg of random primers (Invitrogen Corp.) and 200 U of Moloney murine leukemia virus (Invitrogen Corp.) by incubation at 37°C for 1 h. First and heminested PCRs targeting the second hypervariable region of G protein were performed. The subgroup A-specific forward primer used for the first PCR was primer GPA (nucleotide positions 511 to 530), based on the primer described by Peret et al. (22), and that used for the heminested PCR was primer nRSAG (5′-TATGCAGCAACAATCCAAACC-3′; nucleotide positions 539 to 558). Subgroup B-specific forward primer GPB was used for the first PCR (nucleotide positions 495 to 515) (22), and primer nRSBG (5′-GGTGCCAACTACCAACTTGC-3′; nucleotide positions 512 to 531) was used for the heminested PCR. The reverse primer specific for both subgroups A and B was primer...
TABLE 1. Distribution of RSV subgroups and genotypes and demographic details for patients infected with RSV over three epidemic seasons from November 2001 to July 2004

<table>
<thead>
<tr>
<th>Subgroup and genotype</th>
<th>2001-2002 season&lt;sup&gt;a&lt;/sup&gt;</th>
<th>2002-2003 season&lt;sup&gt;b&lt;/sup&gt;</th>
<th>2003-2004 season&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/total (%)</td>
<td>Age (yr [mean ± SD])</td>
<td>n/total (%)</td>
</tr>
<tr>
<td>A</td>
<td>30/74 (40.5)</td>
<td>0.8 ± 0.6</td>
<td>9/187 (4.8)</td>
</tr>
<tr>
<td>GA2</td>
<td>0/74 (0.0)</td>
<td>1.0 ± 0.0</td>
<td>0/187 (0.0)</td>
</tr>
<tr>
<td>GA5</td>
<td>29/74 (39.2)</td>
<td>0.8 ± 0.6</td>
<td>6/187 (3.2)</td>
</tr>
<tr>
<td>GA7</td>
<td>1/74 (1.4)</td>
<td>0.3 ± 0.0</td>
<td>3/187 (1.6)</td>
</tr>
<tr>
<td>B</td>
<td>2/74 (2.7)</td>
<td>0.7 ± 0.5</td>
<td>45/187 (24.1)</td>
</tr>
<tr>
<td>GB3</td>
<td>0/74 (0.0)</td>
<td>1.0 ± 0.0</td>
<td>4/187 (2.1)</td>
</tr>
<tr>
<td>SAB3</td>
<td>2/74 (2.7)</td>
<td>0.7 ± 0.5</td>
<td>0/187 (0.0)</td>
</tr>
<tr>
<td>BA virus&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0/74 (0.0)</td>
<td>1.0 ± 0.0</td>
<td>41/187 (21.9)</td>
</tr>
<tr>
<td>Total</td>
<td>32/74 (43.2)</td>
<td>0.8 ± 0.6</td>
<td>54/187 (28.9)</td>
</tr>
</tbody>
</table>

<sup>a</sup> The 2001-2002 season was from November 2001 to July 2002.
<sup>b</sup> The 2002-2003 season was from August 2002 to July 2003.
<sup>c</sup> The 2003-2004 season was from August 2003 to July 2004.
<sup>d</sup> n, number of confirmed cases.
<sup>e</sup> Total, number of suspected cases.
<sup>f</sup> BA viruses are strains with a 60-nucleotide insertion in the second variable region of G protein.

RESULTS

Nasopharyngeal aspirate samples were obtained from 74 children from November 2001 to July 2002 (2001-2002 season), 187 children from August 2002 to July 2003 (2002-2003 season), and 238 children from August 2003 to July 2004 (2003-2004 season). The average age of the children with RSV infection was 0.93 ± 0.84 years. Thirty (40.5%) of the 74 children, 9 (4.8%) of the 187 children, and 83 (34.9%) of the 238 children were identified as having subgroup A infections in the 2001-2002, 2002-2003, and 2003-2004 seasons, respectively, while 2 (2.7%), 45 (24.1%), and 16 (6.7%) of the children identified were as having subgroup B infections in the three seasons, respectively (Table 1).

Phylogenetic analysis revealed that 122 subgroup A strains clustered as three genotypes (117 strains in genotype GA5, 4 strains in genotype GA7, and 1 strain in genotype GA2) during the 2001-2002 to 2003-2004 seasons; the bootstrap values were 70 to 100% (Table 1 and Fig. 1). GA5 was the predominant genotype among subgroup A isolates during the three seasons, and the genetic distances (p distances) among the genotype GA5 strains ranged from 0.004 to 0.059. The number of genotype GA5 strains decreased in the 2002-2003 season but increased again in the 2003-2004 season (Table 1; Fig. 2). The 117 genotype GA5 isolates that caused infections could be classified into 28 distinct strains (Fig. 1). In the 2001-2002 season, 15 strains were genotype GA5 strains identical to strain NG-001-02. The numbers of strains identical to strain NG-065-02 were 1, 3, and 63 in the 2001-2002, 2002-2003, and 2003-2004 seasons, respectively. Genotype GA7 strains comprised one, three, and none of the strains in the 2001-2002, 2002-2003, and 2003-2004 seasons, respectively; and only one genotype GA2 strain was detected, which was in the 2003-2004 season (Table 1; Fig. 1 and 2).

Phylogenetic analysis also revealed that 63 subtype B strains were clustered in three genotypes (4 strains in genotype GB3, 2 strains in genotype SAB3, and 57 strains in a new genotype with a 60-nucleotide insertion in the second variable region of G protein) during the three seasons, with bootstrap values of...
70 to 100% (Table 1; Fig. 1). Genotypes similar those of the strains of subgroup B of the new genotype with the 60-nucleotide insertion (named BA viruses) have been reported in Buenos Aires (Buenos Aires [BA] virus) in 1999 (27). All of our BA-like viruses demonstrated a Ser247Pro amino acid change compared with the sequence of BA virus in the region with the insert. Two genotype SAB3 strains of subgroup B were found only in the 2001-2002 season, 4 genotype GB3 strains were found only in the 2002-2003 season (Table 1, Fig. 1 and 2), 41 strains of the new genotype were found in the 2002-2003 season, and 16 strains of the new genotype were found in the 2003-2004 season. Genetic distances (p distances) among the strains of the new genotype ranged from 0.003 to 0.064, and 23 strains from the 2002-2003 season and 1 strain from the 2003-2004 season were identical to strain NG-006-03. Furthermore, strains in both subgroups A and B tended to cluster adjacently by year of collection by phylogenetic tree analysis.

RSV infections were detected from November 2001 to February 2002 and were detected again starting in September 2002 (Fig. 2A). The peak month for RSV infections was December in the 2001-2002 and 2002-2003 seasons and November in the 2003-2004 season. The predominant genotype shifted from genotype GA5 to BA viruses of subgroup B in the 2002-2003 season and returned to genotype GA5 in the 2003-2004 season (Table 1; Fig. 2B and C). The average age of the patients demonstrated no significant linkage with the subgroup or genotype infecting the patients in any of the 3 years (Table 1).

The numbers of hospitalized patients infected with genotype GA5 were 3 (10.3%) of 29, 2 (20.0%) of 10, and 3 (3.7%) of 82 in the 2001-2002, 2002-2003, and 2003-2004 seasons, respectively. The values for BA viruses were 5 (12.5%) of 40 and 2 (12.5%) of 16 in the 2002-2003 and 2003-2004 seasons, respectively, and that for genotype GB3 was 1 (25.0%) of 4 in the 2002-2003 season.

Eight (4.3%) of 177 patients with RSV infections were reinfected over the study period. Two patients infected with genotype GA5 were reinfected with the same genotype over 2- to 24-month periods. One patient infected with genotype GA5 virus was reinfected with BA viruses after 1 year. Four patients infected with BA viruses in the 2002-2003 season were reinfected with genotype GA5 over 4- to 14-month periods.
DISCUSSION

The present investigation of the patterns of circulation of RSV infections in a community over three seasons by genotype typing of the second hypervariable region of G protein (21, 22, 28) demonstrated that multiple genotypes cocirculate each year. In the present study RSV infections started in early winter, and the rates declined in the spring. The predominant subgroup changed from subgroup A to subgroup B over the three epidemic seasons, in line with the findings of earlier reports (4, 8, 15). In our study, we monitored patients over three seasons; however, the 2001-2002 season began in November 2001. The peak month for RSV infection was December 2001 in Niigata City, as was the case in a national survey of RSV infection (20), and we considered that our analysis may have been developed or implemented partially in the 2001-2002 season.

Our phylogenetic analysis revealed that genotype GA5 of subgroup A was predominant in the 2001-2002 and 2003-2004 seasons, while a new genotype of subgroup B, which featured a 60-nucleotide insertion in the second variable region of G protein (BA viruses), was predominant in the 2002-2003 season. Our observations indicate that multiple genotypes cocirculate in a single epidemic and that the genotypes in each epidemic may differ, as described previously (22, 28). Genotypes of both subgroups A and B showed temporal clustering by year of detection, which supported previous findings (6). Strains detected at the end of the previous season tended to be predominant in the next season, which might be associated with antigenic evasion from host immunity.

Viruses of genotypes similar to our BA viruses were also reported in Sapporo, Japan, in 2002 (GenBank accession number AB117522). The BA virus reported by Trento et al. (27) had an exact duplication of 60 nucleotides in the C-terminal one-third of the G-protein gene and illustrated a new type of drastic change introduced in G protein during the natural propagation of RSV. Our strains of the new genotype and strains from Sapporo were demonstrated to have 1 amino acid substitution in the insert region compared to the sequence of the BA strain. It is noted that the specific strains circulated in two countries, one in South America and another in Northeast Asia, after approximately 3 years with only a minor modification of the amino acid, which could support the robustness of the virus. The C terminus of the G-protein molecule has been shown to be immunologically relevant. Therefore, it is suggested that the 60-nucleotide insertion in the C-terminal one-third of the G-protein gene and the amino acid replacement compared with the amino acid in prototype BA strains change its antigenic structure, which confers an evolutionary advantage that allows reinfection of individuals previously exposed to the ancestor virus. However, as an emerging strain, our strain of the new genotype of subgroup B was not associated with new epidemiological or clinical features compared with those of the other clades during the three seasons that we studied. Further studies are required to determine the effect of the insertion on the immune response to RSV and susceptibility to infection and disease.

It has been reported that the severity of RSV infection may vary with the specific virus genotype (16); however, in the present study, no differences in the epidemiological or clinical manifestations, such as age or an illness that required hospitalization, were detected among these genotypes. Thus, we need continued observations to determine whether the greater severity of illness is associated with specific genotypes.

The variability of RSV strains may contribute to the cause of repeated infections, and children infected with subgroup A strains appear to be more likely to be reinfected than those
infected with subgroup B strains (18). Only 8 (4.3%) of our 177 RSV patients became reinfection over the study period. With such a small number of patients, it is impossible to discuss the relationship between reinfection and genetic diversity, even with the new genotype of subgroup B strains. Furthermore, small numbers of reinfections may have been detected in our study because the patients visited other medical care facilities or the patients may have had mild symptoms during the second infections.

In conclusion, our molecular analysis of RSV in Niigata, Japan, confirmed that plural genotypes cocirculate each year and that the predominant genotype may shift with the season. A new genotype of subgroup B with a 60-nucleotide insertion, named BA-like virus, was found to be a predominant genotype, but it was not associated with new epidemiological or clinical features compared with those of the other genotypes that were present during the three seasons that we studied. Finally, our results provide support for genotype designation by RT-PCR methods as an effective tool for characterization of RSV circulation patterns in communities.

ACKNOWLEDGMENTS

This investigation was supported by a research grant for science and welfare, Subcommittee of Emerging and Re-Emerging Diseases, Ministry of Health, Labor and Welfare, Tokyo, Japan.

We thank A. Watanabe, A. Obata, and S. Aida for technical expertise.

REFERENCES