

Genetic Variability in the G Protein Gene of Group A and B Respiratory Syncytial Viruses from India

Shama Parveen,¹ Wayne M. Sullender,³ Karen Fowler,^{3,4,5} Elliot J. Lefkowitz,⁶
Suresh Kumar Kapoor,² and Shobha Broor^{1*}

Department of Microbiology¹ and Comprehensive Rural Health Services Project, Centre for Community Medicine,² All India Institute of Medical Sciences, New Delhi, India, and Departments of Pediatrics,³ Epidemiology,⁴ Maternal and Child Health,⁵ and Microbiology,⁶ University of Alabama at Birmingham, Birmingham, Alabama

Received 28 January 2006/Returned for modification 17 March 2006/Accepted 6 May 2006

Respiratory syncytial virus (RSV) is the most commonly identified viral agent of acute respiratory tract infection (ARI) of young children and causes repeat infections throughout life. Limited data are available on the molecular epidemiology of RSV from developing countries, including India. This study reports on the genetic variability in the glycoprotein G gene among RSV isolates from India. Reverse transcription-PCR for a region of the RSV G protein gene was done with nasopharyngeal aspirates (NPAs) collected in a prospective longitudinal study in two rural villages near Delhi and from children with ARI seen in an urban hospital. Nucleotide sequence comparisons among 48 samples demonstrated a higher prevalence of group A (77%) than group B (23%) RSV isolates. The level of genetic variability was higher among the group A viruses (up to 14%) than among the group B viruses (up to 2%). Phylogenetic analysis revealed that both the GA2 and GA5 group A RSV genotypes were prevalent during the 2002-2003 season and that genotype GA5 was predominant in the 2003-2004 season, whereas during the 2004-2005 season both genotype GA5 and genotype BA, a newly identified group B genotype, cocirculated in almost equal proportions. Comparison of the nonsynonymous mutation-to-synonymous mutation ratios (dN/dS) revealed greater evidence for selective pressure between the GA2 and GA5 genotypes (dN/dS , 1.78) than within the genotypes (dN/dS , 0.69). These are among the first molecular analyses of RSV strains from the second most populous country in the world and will be useful for comparisons to candidate vaccine strains.

Acute respiratory tract infection (ARI) is the leading killer of children in the world (1.9 million per year), with the greatest number of deaths occurring in developing countries (49). One-fourth (2.5 million) of the total deaths among children less than 5 years of age occur in India (1), and approximately 20% of these are due to ARI (0.5 million) (35, 49). Viruses are found in 20 to 40% of children hospitalized with ARIs in India, with respiratory syncytial virus (RSV) being one of the most frequently identified pathogens (15, 22, 23).

RSV strains vary genetically and antigenically and have been classified into two broad groups, groups A and B (2, 11, 16, 25), with additional variability detected within the groups (4, 40). Antigenic variability is thought to contribute to the capacity of the virus to establish reinfections throughout life and may pose a challenge to vaccine design. Future planning for vaccine development will require an understanding of the genetic composition of the RSV strains circulating among target populations.

The RSV G protein is a type II integral membrane protein (48) and shows the highest degree of divergence both between and within the two groups (16). The G protein is highly glycosylated, and it is the target for neutralizing and protective immune responses. Variability in the G protein gene is concentrated in the extracellular domain, which consists of two hypervariable regions separated by a central conserved region

of 13 amino acids (16). The second variable region, which corresponds to the C-terminal region of the G protein, reflects overall G protein gene variability and has been analyzed in molecular epidemiological studies (5, 33).

The objective of the present study was to evaluate the genetic diversity of RSV strains collected in a longitudinal study of ARI from young children in two rural villages in India and from children with ARI seen in an urban hospital. Information about distribution of RSV genotypes in India will be beneficial to the development and implementation of RSV vaccines.

(This study was presented in part at the Joint Meeting of the International Union of Microbiological Societies 2005 [San Francisco, Calif., 23 to 28 July 2005] and the RSV 2005 Symposium [Oxford, United Kingdom, 15 to 18 September 2005].)

MATERIALS AND METHODS

Clinical samples and diagnosis of RSV infection. The details of this epidemiological study will be described in another report that is under preparation; the results of the molecular epidemiologic study are provided here. Newborns from two rural villages, Nawadha and Mujheri, of Ballabgarh block near Delhi, were enrolled between October 2001 and December 2004 and monitored up to 3 years of age or to the end of the study, March 2005. Two hundred eighty-one children were enrolled, and the total follow-up was 441 child-years. The children enrolled in the study were seen weekly in their homes; approximately 85% of these visits were completed. ARIs were classified according to World Health Organization definitions (50), and nasopharyngeal aspirates (NPAs) were collected at each episode of ARI. ARI was defined as the presence of a cough or difficulty in breathing that had occurred within the previous 7 days or that was observed during the visit. Each episode of ARI was considered to last 2 weeks, a new episode was counted if new signs and symptoms developed after the child had been free of symptoms for at least 1 week. ARIs were classified as upper respiratory tract infections (URIs), acute lower respiratory tract infections

* Corresponding author. Mailing address: Department of Microbiology, All India Institute of Medical Sciences, New Delhi 110029, India. Phone: 91 11 2659 4926. Fax: 91 11 2658 8663. E-mail: shobha.broor@gmail.com.

(ALRIs), and severe ALRIs, according to World Health Organization definitions (50).

Specimens were collected through an infant feeding tube, which was placed in the posterior nares and attached to a mucous trap (Sterimed, New Delhi, India). The NPAs were obtained by using suction from a hand vacuum pump (Nalgene Mityvac), after which 2 to 3 ml cold viral transport medium (Hanks balanced salt solution containing penicillin at 1,000 U–streptomycin at 1,000 µg per ml and 2% bovine serum albumin or gelatin) was aspirated into the mucous trap. The samples were transported to the Virology Laboratory of the All India Institute of Medical Sciences on ice within 4 to 6 h and were processed on the same day. The samples were split into two parts of 1.5 ml for reverse transcription-PCR (RT-PCR) and 2 ml for virus isolation. The PCR aliquots were stored at -70°C until PCR was done.

RSV was detected on direct smears by direct fluorescence assay (DFA) with two different commercially available kits: the SimulFluor Respiratory Screen and the Respiratory Panel 1 DFA kit (Chemicon International, Inc., Temecula, CA). RSV was also isolated from all the NPAs by centrifugation-enhanced culture (CEC), followed by indirect immunofluorescence assay with monoclonal antibody for RSV, as described earlier (22). RT-PCR for the G protein gene was done with all NPAs positive by DFA and/or CEC for RSV. A subset of DFA- and/or CEC-negative samples from children with ARI were also tested by RT-PCR. Institutional ethics committees in India and the United States approved the studies. In addition, 30 clinical samples from children with ARIs that were sent to the Diagnostic Virology Laboratory at the All India Institute of Medical Sciences were also studied.

RT-PCR for G protein gene of RSV. RNA from the clinical samples was extracted with an RNeasy mini kit (QIAGEN GmbH, Hilden, Germany), according to the manufacturer's instructions, with the addition of RNasin (Promega, Madison, WI) and glycogen (Sigma Aldrich Corp., St. Louis, MO) during RNA extraction. cDNA was synthesized by using random primers and avian myeloblastosis virus reverse transcriptase (Promega). The second hypervariable region of the G protein gene of RSV was the target for the external and the seminested PCRs. External PCR was carried out with primers ABG490 and F164. The forward primer, ABG490 (ATGATTWYCA YTTTGAAGTGTTCC), corresponds to bases 497 to 519 of the G protein gene of strain A2 and bases 491 to 513 of the G protein gene of strain 18537. The reverse primer, F164 (GTTA TGACTGTTGATACCAACC), corresponds to bases 164 to 186 of the F protein gene of strain 18537 (with one mismatch with the G protein gene of strain A2) and has previously been used to amplify the G protein genes of both the groups (42). Three microliters of cDNA was added to 22 µl PCR master mixture containing 200 µM deoxynucleoside triphosphates, 1.5 mM MgCl₂, 1.5 U *Taq* DNA polymerase (Bangalore Genei Ltd., Bangalore, India), and 50 pM of each primer. Amplification was carried out at 94°C for 1 min, followed by 35 cycles of 94°C for 40 s, 50°C for 45 s, and 72°C for 45 s, with a final extension at 72°C for 10 min. The amplified products of 607/610 bp and 670 bp for group A/B and BA viruses, respectively, were analyzed by electrophoresis on a 2% agarose gel. One microliter of diluted external PCR products was used for the seminested PCR. AG655 for group A and BG517 for group B were used as forward primers and F164 was used as the reverse primer for the seminested PCR. Primer AG655 (GATCYCAAACCTCAAACCAC), modified from a previously described primer (5), corresponded to bases 655 to 674 of the G protein gene of strain A2. Group B-specific primer BG517 (TTYGTTCCTGTAGTATATGTG) corresponded to bases 517 to 538 of the G protein gene of strain 18537. Amplification was carried out at 94°C for 1 min, followed by 25 cycles of 94°C for 40 s, 58°C for 45 s, and 72°C for 45 s, with a final extension at 72°C for 10 min. The nested amplicons of 450/585 bp and 645 bp for group A/B and BA viruses, respectively, were visualized by agarose gel electrophoresis.

DNA sequencing. The nested primers, AG655 for group A and BG517 for group B, were used as forward primers and F164 was used as the reverse primer for sequence determination. The PCR products were purified with a QIAquick gel extraction kit (QIAGEN), according to the manufacturer's instructions. The purified PCR products were cycle sequenced in the forward and the reverse directions in an ABI PRISM 310 genetic analyzer (PE Applied Biosystems Inc., Foster City, CA) by using an ABI PRISM BigDye Terminator cycle sequencing ready reaction kit (PE Applied Biosystems Inc.).

Phylogenetic analysis. The nucleotide sequences obtained from within the second variable region of the G protein gene were manually edited in Genedoc (version 2.6.02) (28) and aligned with the available RSV sequences from the GenBank database by using the CLUSTAL X program (version 1.83) (43). Phylogenetic trees were constructed by the neighbor-joining method in MEGA software (version 3) (18). The statistical significance of the tree topology was tested by bootstrapping (1,000 replicas). Pairwise distances between and within the genotypes at the nucleotide level were calculated with Kimura 2 parameters

TABLE 1. RSV group A G protein gene GenBank sequences used in the study

Strain	GenBank accession no.	Country of isolation	Yr of isolation	Reference
A2	M11486	Australia	1961	48
NY/CH09/93	AF065254	United States	1993	33
AL19376-1	AF233900	United States	1994–1995	32
AL19452-2	AF233901	United States	1994–1995	32
AL19471-5	AF233902	United States	1994–1995	32
AL19556-3	AF233903	United States	1994–1995	32
MO01	AF233909	United States	1994–1995	32
MO02	AF233910	United States	1994–1995	32
MO16	AF233913	United States	1994–1995	32
MO48	AF233914	United States	1994–1995	32
MO55	AF233915	United States	1994–1995	32
NY103	AF233916	United States	1994–1995	32
NY108	AF233917	United States	1994–1995	32
NY20	AF233918	United States	1994–1995	32
TX67951	AF233919	United States	1994–1995	32
TX68481	AF233920	United States	1994–1995	32
TX69564	AF233923	United States	1994–1995	32
CN1973	AF233904	Canada	1994–1995	32
CN2395	AF233905	Canada	1994–1995	32
CN2708	AF233906	Canada	1994–1995	32
SA97D1289	AF348803	S. Africa	1997	47
SA99V360	AF348804	South Africa	1999	47
SA98V603	AF348807	South Africa	1998	47
SA99V1239	AF348808	South Africa	1999	47
Ab4026B01	AY146435	South Africa	2001	21
Ab5076Pt01	AY146437	South Africa	2001	21
LLC235-267	AY114149	Singapore	2000–2001	20
LLC242-282	AY114150	Singapore	2000–2001	20
LLC62-111	AY114151	Singapore	2000–2001	20
Sal/87/99	AY472086	Brazil	1999	24
Sal/173/99	AY472094	Brazil	1999	24
Ken/61/02	AY524651	Kenya	2002	38
Ken/7/00	AY524663	Kenya	2000	38
NG/009/02	AB175815	Japan	2002	37
SE/10/92	AF193310	Korea	1992	8
SE/02/98	AF193325	Korea	1998	8

and with Poisson correction at the amino acid level with MEGA software. NetOglyc software (version 3.1) (14) was used to predict the potentially O-glycosylated serine and threonine residues.

The nucleotide sequence spanned bases 673 to 912 (240 bp) of prototype strain A2 (GenBank accession number M11486) (48). For group B viruses the sequence corresponded to bases 670 to 963 (294 bp) of a genotype BA strain from Argentina (strain BA4128/99B; GenBank accession number AY333364) (44). Thirty-six published sequences of group A RSVs and 38 sequences of group B RSVs were used for comparison to the sequences derived in the present study. GenBank accession numbers and the year and country of isolation of the group A and B sequences are given in Table 1 and Table 2, respectively. The sequences were selected so that representatives of each genotype of group A RSV (eight genotypes) and group B RSV (eight genotypes) were included. Indian group A and B RSV strains were genotyped by phylogenetic clustering, and their sequences were compared to sequences previously assigned to a specific genotype (32, 33, 37, 47).

Synonymous and nonsynonymous mutations were analyzed by the method of Nei and Gojobori (27). The program SNAP (Synonymous/Nonsynonymous Analysis Program) provided by the HIV database website (<http://www.hiv.lanl.gov/content/hiv-db/SNAP/WEBSNAP/SNAP.html>) was used for analysis of synonymous mutations versus nonsynonymous mutations.

Nucleotide sequence accession numbers. The GenBank accession numbers of the nucleotide sequences obtained in the present study are D248894 to D248941.

RESULTS

RT-PCR for the G protein gene of RSV was done on 245 samples collected from the rural villages and 30 samples col-

TABLE 2. RSV group B G protein gene GenBank sequences used in the study

Strain	GenBank accession no.	Country of isolation	Yr of isolation	Reference
18537	M17213	United States	1962	16
QUE/85/02-03	AY927408	Canada	2002–2003	R. Gilcia, unpublished data
QUE/155/01-02	AY927412	Canada	2001–2002	R. Gilcia, unpublished data
S71/02	AB161414	Japan	2002	19
BE/12670/01	AY751086	Belgium	2001	51
BE/12817/03	AY751091	Belgium	2003	51
BE12370/01	AY751118	Belgium	2001	51
BE13417/99	AY751131	Belgium	1999	51
NG/004/03	AB175819	Japan	2003	37
NG/006/03	AB175820	Japan	2003	37
NG/153/03	AB175821	Japan	2003	37
S4/01	AB117522	Japan	2001	26
BA4128/99B	AY333364	Argentina	1999	44
Ken/109/02	AY524573	Kenya	2002	38
Ken/2/00	AY524575	Kenya	2000	38
Ken/29/03	AY660681	Kenya	2003	38
AL19734-4	AF233924	United States	1994–1995	32
AL19794-1	AF233925	United States	1994–1995	32
MO30	AF233928	United States	1994–1995	32
MO35	AF233929	United States	1994–1995	32
NY01	AF233931	United States	1994–1995	32
TX69208	AF233933	United States	1994–1995	32
Ch10b	AF065250	United States	1990	33
CH93/9b	AF065251	United States	1993	33
WV10010	M73541	United States	1983	41
WV15291	M73542	United States	1985	41
SA98V192	AF348811	South Africa	1998	47
SA99V429	AF348813	South Africa	1999	47
SA97D934	AF348817	South Africa	1997	47
SA99V800	AF348821	South Africa	1999	47
SA99V1325	AF348822	South Africa	1999	47
SA98V602	AF348824	South Africa	1998	47
SA0025	AF348825	South Africa	2000	47
SA98D1656	AF348826	South Africa	1998	47
Mon/2/99	AY488795	Uruguay	1999	3
Mon/7/99	AY488800	Uruguay	1999	3
Moz/205/99	AF309678	Mozambique	1999	36
Moz/204/99	AF309684	Mozambique	1999	36

lected from the hospital. Forty-eight NPAs, 33 from the rural villages and 15 from the hospital, were positive for the RSV G protein gene by RT-PCR. A subset (14) of these sequences have been described in an analysis of repeat infections due to RSV in a separate publication.

Phylogenetic analysis. Twenty-seven and 10 group A strains were identified from the rural community and the hospital, respectively. Phylogenetic analysis revealed that these 37 group A strains clustered in two genotypes: 29 (78%) strains in genotype GA5 and 8 (22%) strains in genotype GA2 (Table 3 and Fig. 1A). Of these 37 sequences, 19 unique group A sequences were identified, of which 13 sequences were found only once. Of the remaining 24 sequences, six groups of identical sequences were found, with 2 to 8 sequences per group. In all but one set of identical viruses the viruses were identified in more than 1 year, and two sets contained viruses from both the community and the hospital (Table 4).

The rates of divergence between prototype strain A2 and the Indian strains were 11% to 17% at the nucleotide level and 23% to 37% at the amino acid level. Differences of up to 14% at the nucleotide level and 30% at the amino acid level were observed among the group A Indian strains. Differences of up to 3% at the nucleotide level and 7% at the amino acid level

were observed among the genotype GA2 strains. Genotype GA5 strains showed up to 5% nucleotide differences and up to 12% amino acid differences.

The G protein gene sequence of one of the strains (strain DEL/3W/04/A) was found to be identical to three sequences, one reported from Singapore (LLC235-267) (20) and two from Belgium (strains BE/11091/00 and BE/11129/00) (52). The G protein gene sequence of strain DEL/606/03/A was found to be identical to the sequences of 55 strains reported from different parts of the world, including Singapore (strain LLC242-282) (20),

TABLE 3. Yearly distribution of RSV genotypes

Year	No. of positive samples			
	Total	Group A		Group B, BA
		GA5	GA2	
2001–2002	1	0	0	1
2002–2003	8	5	3	0
2003–2004	22	15	5	2
2004–2005	17	9	0	8
Total	48	29	8	11

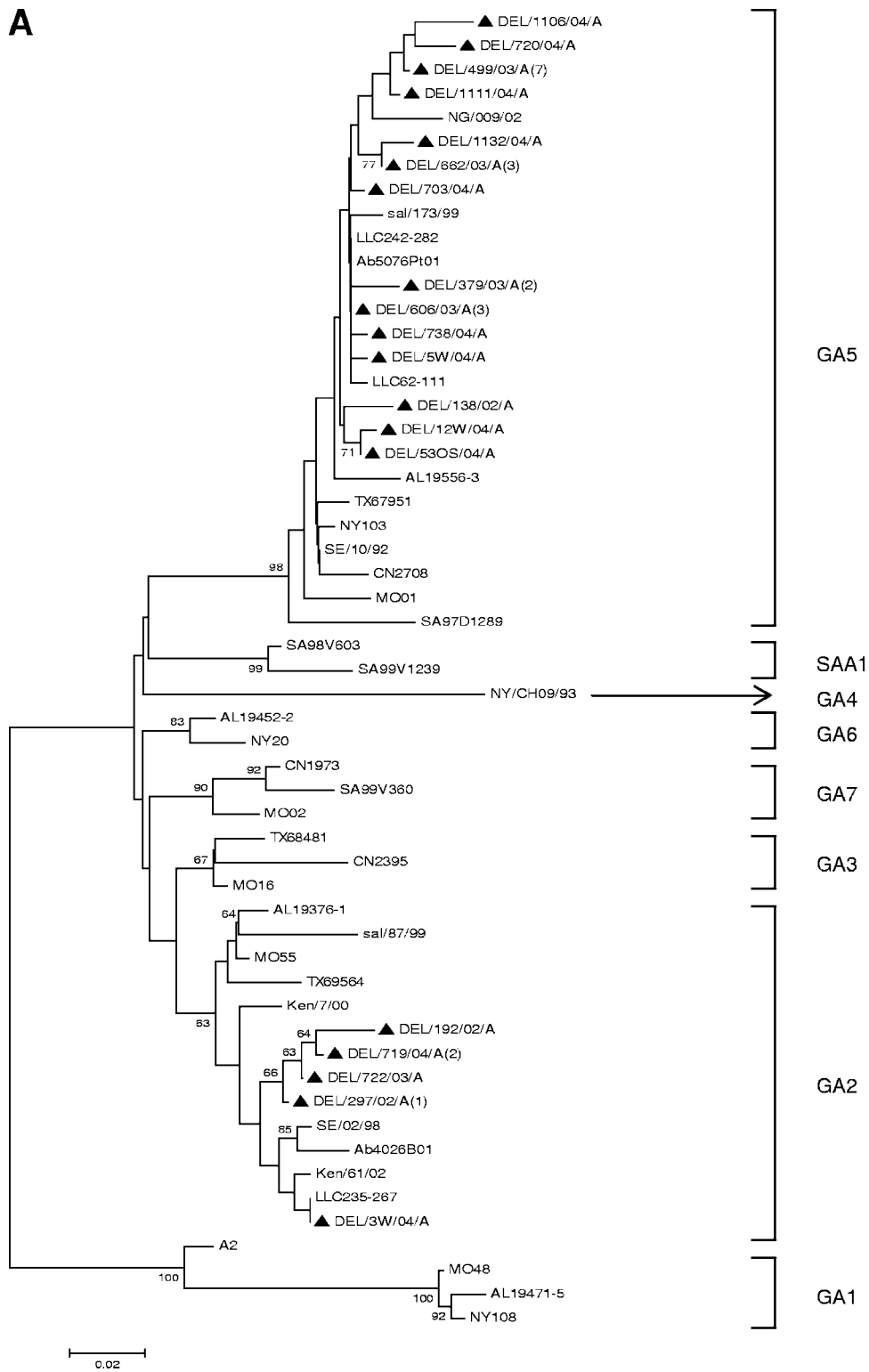
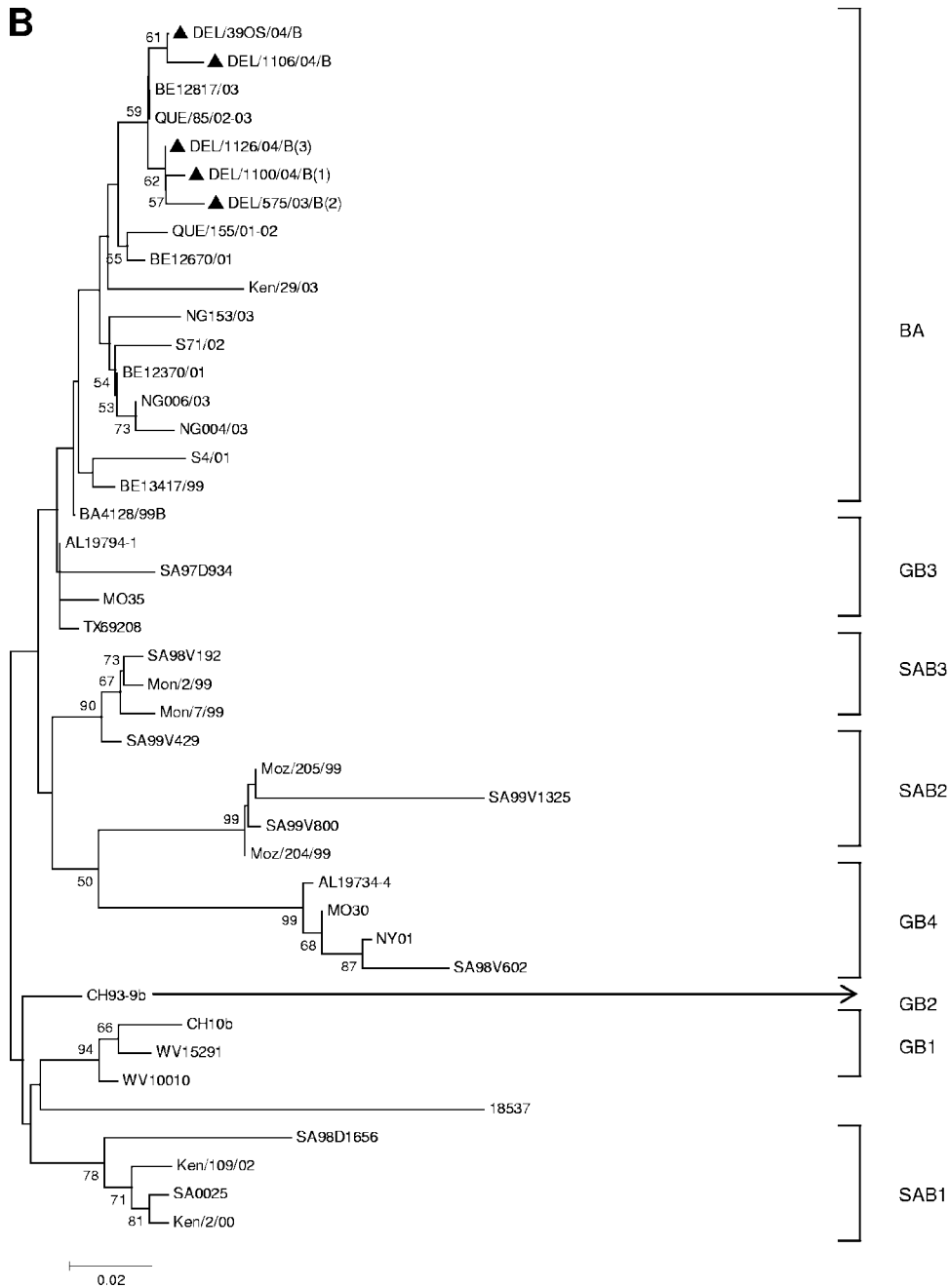


FIG. 1. Phylogenetic trees for Indian RSV group A (A) and group B (B) nucleotide sequences from the second variable region of the G protein gene. The numbers of identical sequences are indicated in parentheses. The sequences of the Indian strains are indicated by solid triangles. Details for the reference GenBank sequences used for phylogenetic analysis are given in Table 1 for the group A RSV strains and in Table 2 for the group B RSV strains. The nucleotide sequences were aligned with the CLUSTAL X program, and phylogenetic trees were constructed with MEGA 3 software with a neighbor-joining algorithm. Bootstrap values greater than 50% are shown at the branch nodes. The genotypes are indicated at the right by brackets. The following reference sequences were downloaded from GenBank and were used to construct the trees: MO48, AL19471-5, NY108, and A2 (GA1); LLC235-267, Ken/61/02, Ken/7/00, Ab4026B01, SE/02/98, TX69564, MO55, Sal/87/99, and AL19376-1 (GA2); MO16, CN2395, and TX68481 (GA3); NY/CH09/93 (GA4); SA97D1289, MO01, CN2708, SE/10/92, NY103, TX67951, AL19556-3, Sal/173/99,



Ab5076Pt01, LLC242-282, LLC62-111, and NG/009/02 (GA5); AL19452-2 and NY20 (GA6); MO02, SA99V360, and CN1973 (GA7); SA98V603 and SA99V1239 (SAA1); WV10010, WV15291, and CH10b (GB1); CH93-9b (GB2); AL19794-1, MO35, TX69208, and SA97D934 (GB3); AL19734-4, MO30, NY01, and SA98V602 (GB4); SA98D1656, Ken/109/02, SA0025, and Ken/2/00 (SAB1); SA99V800, SA99V1325, Moz/204/99, and Moz/205/99 (SAB2); SA99V429, SA98V192, Mon/7/99, and Mon/2/99 (SAB3); BA4128/99B, S4/01, S71/02, NG004/03, NG/006/03, NG153/03, Ken/29/03, Que/155/01-02, QUE/85/02-03, BE/13417/99, BE/12670/01, BE/12370/01, and BE/12817/03 (BA); and 18537.

Mozambique (strain Moz/169/99) (36), Belgium (strains BE/1556/01 and BE/2122/00) (52), Turkey (strains IST/A/35 and IST/A/36) (K. Midilli, unpublished data), Uruguay (strains Mon/1/96, Mon/5/96, and Mon/1/98) (13), Uruguay (strains Mon/6/97, Mon/7/97, and Mon/8/97) (12), South Africa (strains Ab5076Pt01 and Ab81J00) (21), and South Africa (strain 0240KS01) (E. Agenbach, unpublished data) and 40 strains from Kenya (strains Ken/16/03, Ken/262/03, Ken/9/02, Ken/9/

01, etc.) (38). The G protein gene sequence of another strain, strain DEL/499/03/A, was found to be identical to the sequences of five strains, one from Belgium (strain BE/13281/99) (52) and four from Turkey (strains IST/A/38, IST/A/9, IST/A/8, and IST/A/7) (K. Midilli, unpublished).

All 11 group B strains belonged to the newly identified BA genotype, with a 60-nucleotide duplication in the second variable region of the G protein gene (44) (Table 3 and Fig. 1B).

TABLE 4. Group A and group B RSV strains with identical G protein gene sequences identified during the study

Strains (identical sequences) ^a	Source	WHO classification ^b
DEL/606/03/A	Hospital	ALRI
DEL/649/03/A	Hospital	ALRI
DEL/725/04/A	Rural villages	ALRI
DEL/1102/04/A	Hospital	ALRI
DEL/499/03/A	Rural villages	URI
DEL/659/03/A	Rural villages	URI
DEL/717/04/A	Rural villages	ALRI
DEL/726/04/A	Rural villages	URI
DEL/727/04/A	Rural villages	URI
DEL/728/04/A	Rural villages	URI
DEL/754/04/A	Rural villages	ALRI
DEL/763/04/A	Rural villages	URI
DEL/662/03/A	Rural villages	Severe ALRI
DEL/974/04/A	Rural villages	URI
DEL/997/04/A	Hospital	ALRI
DEL/1213/05/A	Rural villages	ALRI
DEL/181/02/A	Rural villages	URI
DEL/191/02/A	Rural villages	ALRI
DEL/379/03/A	Rural villages	URI
DEL/218/02/A	Rural villages	URI
DEL/297/03/A	Rural villages	URI
DEL/708/04/A	Rural villages	URI
DEL/719/04/A	Rural villages	ALRI
DEL/723/04/A	Rural villages	URI
DEL/1126/04/B	Rural villages	URI
DEL/1137/04/B	Rural villages	URI
DEL/12W/04/B	Hospital	URI
DEL/14W/04/B	Hospital	URI
DEL/575/03/B	Rural villages	ALRI
DEL/609/03/B	Rural villages	ALRI
DEL/65/02/B	Rural villages	URI
DEL/1100/04/B	Rural villages	URI
DEL/16W/04/B	Hospital	URI

^a The RSV strains in boldface are shown in the phylogenetic trees, and the numbers of identical sequences are indicated in parentheses in the trees.

^b World Health Organization classification of disease severity (50).

Six group B strains were identified from the rural community and five were identified from the hospital. In the 11 group B sequences, there were 5 unique group B sequences, of which 2 were found once and 9 were grouped into three sets of identical sequences of 2 to 4 sequences each (Table 4). There was 2% to 3% divergence at the nucleotide level and 5% to 8% divergence at the amino acid level between the Indian strains and the prototype genotype BA strain. Among the Indian group B strains, up to 2% nucleotide differences and up to 7% amino acid differences were observed. Comparisons of the duplicated region within each virus revealed that in every instance there were 2 to 4 nucleotide sequence differences, each of which resulted in an amino acid coding change, between the two regions. Dual infection with both group A (genotype GA5) and group B (genotype BA) viruses was identified in two samples from hospitalized patients in the 2004-2005 season.

The study shows a higher prevalence of RSV group A (77%) than group B (23%): group A predominated during the 2002-2003 season and during the 2003-2004 season (Table 3). Only one virus which was in the BA genotype was evaluated from 2001; both the genotypes of group A, genotypes GA5 and GA2, were identified during the 2002-2003 season. All three genotypes were detected in 2003 and 2004, whereas during

2004 and 2005, genotypes GA5 and BA circulated in almost equal proportions. The period from October to January, which coincides with cooler weather in Delhi, accounted for 85% of RSV identifications. There were no obvious patterns of sequence differences between the community and the hospital viruses.

Group A RSV resulted in ALRI in 19 of 37 infections and group B RSV caused ALRI in 3 of 11 infections. Severe ALRI occurred in three group A infections but no group B infections. One set of genotype GA5 strains that had identical G protein gene sequences resulted in URI (strain DEL/974/04), ALRI (strains DEL/997/04 and DEL/1213/05), and severe ALRI (strain DEL/662/03) (Table 4). Three sets of identical G protein gene sequences were identified from different patients who had URI or ALRI. Thus, there were no obvious relationships among RSV group, genotype, or nucleotide sequence and the clinical classification.

Amino acid analysis. The predicted amino acid sequences of the group A and group B strains were compared to those of the prototype A2 and genotype BA strains, respectively (Fig. 2). The partial G protein gene sequences were predicted to encode G proteins of 298 amino acids for GA5 viruses and 297 amino acids for GA2 viruses; one GA2 strain (strain DEL/3W/04/A), however, had a G protein of 298 amino acids. The G protein genes of the BA genotype were predicted to encode proteins of two different lengths, 312 and 319 amino acids. A Ser 247 Pro amino acid change was observed in the 20-amino-acid duplicated region in all the Indian BA strains compared to the sequence of the prototype BA strain. This change was also reported in Japanese BA strains (37).

Glycosylation sites. Potential N-glycosylation sites (amino acids NXT, where X is not Pro) have been described for both groups (16, 36). Four putative N-glycosylation sites (Fig. 2A) were identified among group A strains, and the positions of the first and the fourth sites were conserved among all the strains. The second site was genotype specific and was present in genotype GA5 strains only, while the third site was identified in only one genotype GA2 strain, strain DEL/3W/04/A. Two N-glycosylation sites were identified among the group B strains at the C-terminal end of the G protein gene, and these were conserved among all the strains (Fig. 2B).

One of the striking features of the G protein is the large number of serine and threonine residues that are potential O-linked sugar acceptors (48). The program NetOglyc predicted 24 to 33 serine and threonine residues to be potentially O glycosylated with score predictors (G scores) of between 0.5 and 0.8 in the deduced 78- to 79-amino-acid sequence of group A strains (Fig. 2A). Among these potentially O-glycosylated residues, 8 to 10 residues were predicted to be most likely to contain O-linked sugars (10). The amino acid positions that were most likely to have O-linked side chains were 269, 270, 275, 283, and 287 for serine and 227, 231, 235, 253, and 282 for threonine (the amino acid positions refer to those in the prototype strain A2 sequence) (10). The same program also predicted 40 to 44 serine and threonine residues to be potentially O glycosylated with score predictors (G scores) of between 0.5 and 0.8 in the deduced 94 to 101 amino acids of group B strains (Fig. 2B). Among these potentially O-glycosylated residues, 14 to 15 residues were predicted to be most likely to contain O-linked sugars (17, 44, 51). The amino acid positions that

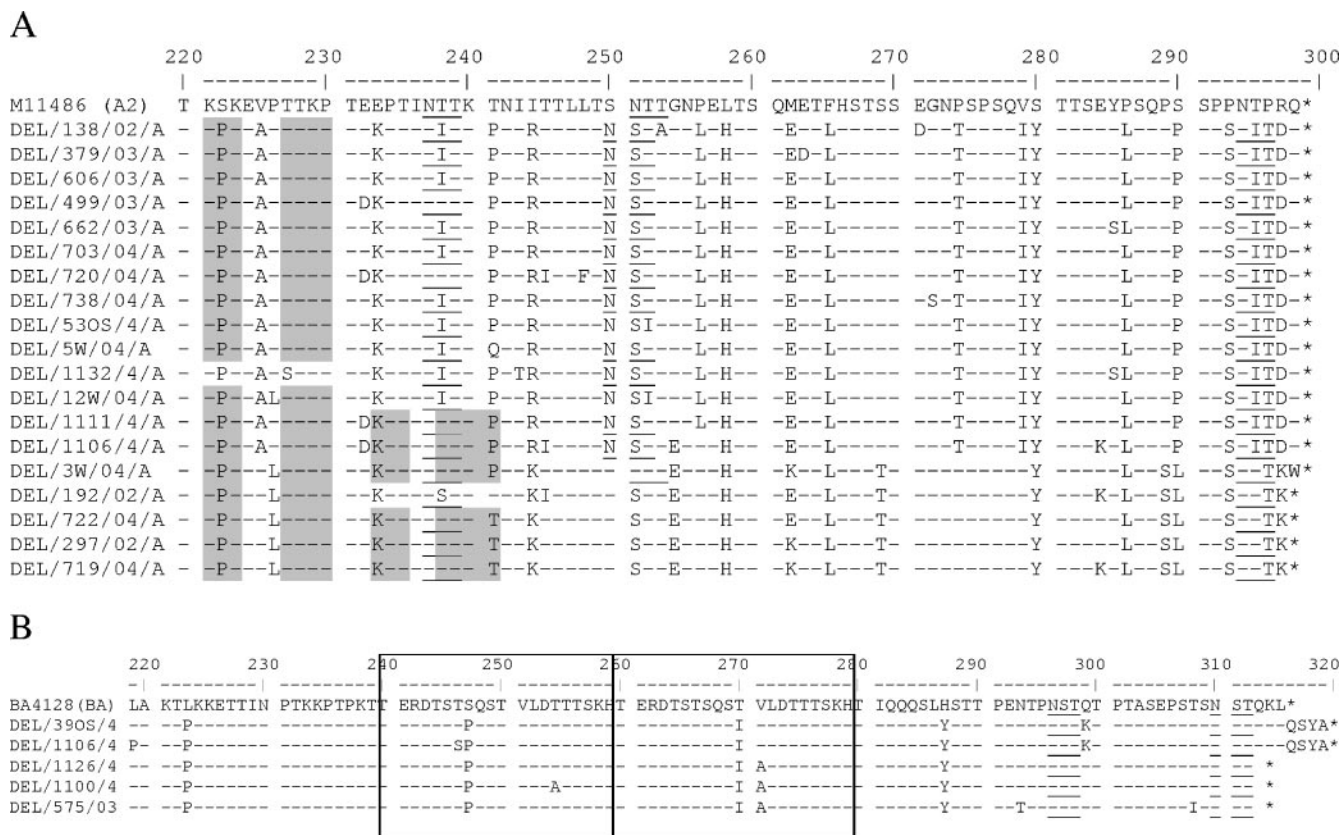


FIG. 2. Deduced amino acid alignments of the second variable region of the G protein gene from group A (A) and group B (B) Indian strains. Alignments are shown relative to the sequences of prototype strain A2 (GenBank accession number M11486) (48) and genotype BA strain BA4128/99B (GenBank accession number AY333364) (44). The amino acids shown correspond to strain A2 G protein positions 220 to 298 for the group A viruses or to strain BA4128/99B G protein positions 219 to 315 for the group B viruses. Identical residues are indicated by dashes. The two copies of the duplicated 20-amino-acid region in group B viruses are indicated by rectangles. Stop codons are indicated by asterisks. Potential N-glycosylation sites (NXT, where N is not proline) are underlined. Potential sites for extensive O-glycosylation KPX - - - TTKX motifs are indicated by gray shading.

were most likely to have O-linked side chains were 265, 267, 269, 304, and 307 for serine and were 228, 232, 236, 254, 264, 266, 274 to 276, and 280 for threonine (the amino acid positions refer to those of the prototype BA strain sequence from Argentina, strain BA4128/99B) (44). Furthermore, we identified nine serine and threonine residues (Fig. 2B) to be potentially O glycosylated in the repeat 20-amino-acid sequence in BA strains (amino acids 264 to 267, 269, 274 to 276, and 280) (44, 51). In addition to serine and threonine residues, two repeats of the motif KPX - - - TTKX (Fig. 2A) were present among group A strains and may be associated with extensive O glycosylation of the G protein (5, 36). This motif was not identified among group B strains.

Synonymous versus nonsynonymous mutations. On average, the nonsynonymous mutation/synonymous mutation (*dN/dS*) ratio for the New Delhi group A sequences was 1.16, while the average *dS/dN* ratio was 1.37. These values suggest the presence of both positive and negative selective pressures. Figure 3 displays a plot showing the distributions of all *dN/dS* pairwise comparisons versus the corresponding pairwise distance between the sequences. Two populations were revealed: one corresponded to higher *dN/dS* ratios (>1) and longer distances, demonstrating positive selection for amino

acid changes. These were intergenotype comparisons, which had a mean *dN/dS* ratio of 1.78. The other population showed generally lower *dN/dS* ratios (<1) and shorter distances, suggesting negative selection pressure for amino acid changes. These were intragenotype comparisons and had a mean *dN/dS* ratio of 0.69. There were too few group B sequences to make a meaningful comparison of the sequences.

DISCUSSION

RSV is a major respiratory tract viral pathogen among hospitalized children in India (15, 22, 23, 31, 34). Approximately 0.5 million children die due to ALRI in India each year, accounting for one-fourth of the 1.9 million deaths from ALRI that occur globally each year (1, 35, 49). Almost one-third of the global deaths from ALRI are estimated to be caused by viruses (RSV, influenza virus, parainfluenza virus), suggesting that there may be as many as 165,000 such deaths in India each year (39). Since 70% of India's population is rural, most of these deaths occur among children in rural areas, and most die without reaching a hospital. In the past three decades there have been no reports on RSV infections among children from rural India (30). Thus, we have included a very important and

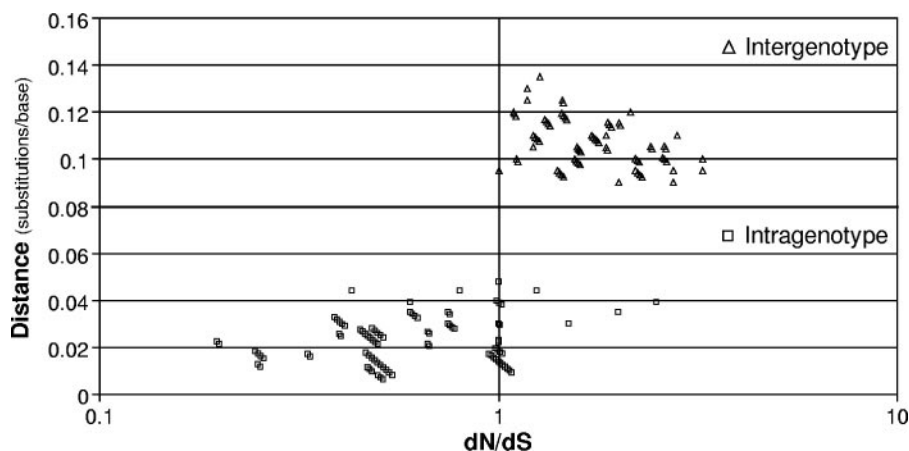


FIG. 3. Plot of the distribution of the dN/dS ratio versus the corresponding pairwise distances between the Delhi group A sequences. The program SNAP was used for analysis of nonsynonymous versus synonymous mutations (27). For each pairwise comparison, the dN/dS ratio is plotted along the x axis, while the distance between the two sequences is plotted along the y axis. Overlapping points along a diagonal represent multiple datum points that directly overlap with the top left-most point of the diagonal. They have been drawn as an extended series of points for clarity. All distances greater than 0.08 substitution/base represent intergenotype comparisons (genotype GA2 versus genotype GA5) and show higher dN/dS ratios, indicating the presence of positive selection pressure for amino acid change. On the other hand, all distances less than 0.08 substitutions/base represent intragenotype comparisons (genotype GA2 versus genotype GA2 and genotype GA5 versus genotype GA5) and show lower dN/dS ratios, suggesting neutral or negative selection pressure for amino acid changes.

yet poorly studied group, i.e., rural children from India, in our investigations. This report provides one of the first descriptions of the molecular analysis of RSV from children in India. In our earlier study the genetic variability among RSV strains from the hospital was studied by restriction enzyme analysis of RT-PCR products (34). We have now extended this earlier study (34) by including nucleotide sequence analysis to more precisely define the extent of genetic variability. In the present investigation, the samples were obtained from a longitudinal study of ARI in children in a rural community and from children in an urban teaching hospital.

Viruses collected over a 42-month period which included four RSV seasons were analyzed. The study demonstrated a higher prevalence of group A (77%) than group B (23%) infections. Among the Indian viruses, the genetic variability among the group A strains (up to 14%) was higher compared to that among group B strains (up to 2%). The higher genetic variability among group A viruses may be responsible for their predominance worldwide (9, 33). However, in our earlier hospital-based study, group B viruses predominated over a 2-year period (34). During the period of the current investigation three RSV genotypes were in circulation, and more than one genotype was found during each season in which more than one virus was characterized.

The G protein gene variable region sequences of some Indian strains were identical to the sequences of the strains reported from different parts of the world, including Singapore, Belgium, Uruguay, Turkey, South Africa, Mozambique, and Kenya. Thus, viruses isolated in distant places and several years apart may be more closely related than viruses isolated in the same place during the same epidemic (6, 32, 33). It has been suggested that the virus strains in a community may arise from the introduction of new viruses or the circulation of endemic strains. Local factors such as RSV strain-specific immunity and viral fitness then determine which virus strains predominate (29, 32).

A new group B genotype, BA, with a 60-nucleotide G protein gene duplication first appeared in Argentina in 1999 (44). The 60-nucleotide duplication starts from nucleotide 792 of the strain 18537 G protein gene and is predicted to lengthen the G protein by 20 amino acids. The BA prototype strain had a G protein of 315 amino acids (44); G protein lengths of 312 (51), 315 (19, 26, 51), 317 (51), and 319 (19, 51) amino acids have been reported. Indian genotype BA strains had predicted G proteins of two different lengths, 312 and 319 amino acids. Evaluation of the duplicated regions in the Delhi viruses revealed that mutations had occurred compared to the sequences of the originally described BA viruses. In addition, changes had occurred, so that for each of the Delhi viruses, the duplicated regions were no longer identical within individual viruses. Comparisons of BA viruses from around the world have shown that mutations are accumulating over time and suggest that positive selection is influencing the evolution of RSV (45).

The BA genotype appears to be spreading globally and has been reported from Japan (19, 26, 37), Kenya (38), Belgium (51), Canada, Brazil, the United Kingdom, and the United States (45). The rapid global spread of BA viruses suggests that these viruses may have a selective advantage over other circulating viruses. In keeping with this, all of the group B viruses in our study contained the 60-nucleotide duplication. Antigenic change in the G protein and avoidance of host immune responses may be the selective advantage (45). However, there could be aspects of G protein function, such as attachment, that are altered. Alternatively, the BA viruses might have as yet unrecognized mutations in other sites that have rendered these viruses more fit than other group B viruses.

It has been suggested that N and O glycosylation of the G protein of RSV helps the virus to evade the host immune response (5, 36). The frequency and pattern of glycosylation sites were different between the two antigenic groups. Four and two putative N-glycosylation sites were identified in group A and group B strains, respectively. More extensive N-glyco-

sylation of group A strains than of group B strains may contribute to additional antigenic variability in group A viruses (36). The predicted O-linked glycosylation sites in the second hypervariable region of the G protein gene analyzed in this study were 8 to 10 residues for group A viruses and 14 to 15 residues for group B viruses. Genotype BA strains had additional O-linked glycosylation residues (45 to 50%) in the duplicated 20-amino-acid region. These additional O-glycosylation residues have also been identified in the duplicated region in a previous study (51). The additional O-linked glycosylation residues in genotype BA strains due to 20-amino-acid duplication may influence the expression of some of the antigenic epitopes by either masking the antigenic sites or contributing to antibody recognition, thus giving these strains an evolutionary advantage over existing group B viruses. Very few studies have performed such detailed analysis of O-linked sugars for both groups of RSV (46, 51, 52).

The average dN/dS ratios for the group A viruses provided evidence both for and against selective pressure. The explanation for these seemingly disparate facts is that by calculating average values, we are obscuring the true explanation, which is that the sequence comparisons represent a nonnormal distribution of values that varies with the distance between any two sequences. Thus, when all dN/dS pairwise comparisons were plotted against the corresponding pairwise distance between the same sequences for the group A Delhi samples, two distinct populations were revealed. The intergroup comparisons showed evidence for selective pressure, with higher dN/dS ratios and more nucleotide changes in the pairwise comparisons. The intragroup comparisons showed clearly different results, with lower dN/dS ratios and fewer nucleotide changes. These results suggest that for closely related virus populations, only neutral or negative selection pressure on the variable region of the G glycoprotein is observed, whereas positive selection pressure for amino acid variation can be discerned only at greater sequence distances. This implies that environmental pressures such as immune evasion become apparent only when any one virus population has diverged sufficiently from another. This is compatible with and extends from the earlier observation that the ratio of synonymous to nonsynonymous substitutions within a sample of group A or B viruses did not suggest positive selection, whereas comparisons between genotypes did suggest positive selective pressure (46).

In conclusion, the molecular characterization of RSV strains from Delhi revealed variations in the proportion of infections caused by different RSV genotypes. Two group A genotypes and one group B genotype were identified, and all the group B strains clustered in the newly identified BA genotype. The present study supports previous suggestions that RSV has a high capability of spreading worldwide, as identical viruses with identical G protein gene variable regions were found in India and other countries (32, 33). This is one of the first reports of the molecular epidemiology of RSV strains from India and is the first description of the circulation pattern of RSV genotypes in both rural and urban Indian settings. As candidate RSV vaccines are considered for use, this information will allow comparison of vaccine viruses with the viruses present in India. In addition, further investigations of the BA viruses with the 60-nucleotide duplication should be very informative as to the epidemiology and evolution of RSV.

ACKNOWLEDGMENTS

The project described here was supported by the Department of Biotechnology (India), the National Institute for Allergy and Infectious Diseases (NIAID; grant AI50693) (United States), and the Indo-U.S. Vaccine Action Program. Financial support for Shama Parveen was received from the Council of Scientific and Industrial Research (India).

The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of NIAID.

REFERENCES

- Ahmad, O. B., A. D. Lopez, and M. Inoue. 2000. The decline in child mortality: a reappraisal. *Bull. W. H. O.* **78**:1175–1191.
- Anderson, L. J., J. C. Hierholzer, C. Tsou, R. M. Hendry, B. F. Fernie, Y. Stone, and K. McIntosh. 1985. Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J. Infect. Dis.* **151**:626–633.
- Blanc, A., A. Delfraro, S. Frabasile, and J. Arbiza. 2005. Genotypes of respiratory syncytial virus group B identified in Uruguay. *Arch. Virol.* **150**:603–609.
- Cane, P. A. 2001. Molecular epidemiology of respiratory syncytial virus. *Rev. Med. Virol.* **11**:103–116.
- Cane, P. A., D. A. Matthews, and C. R. Pringle. 1991. Identification of variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses. *J. Gen. Virol.* **72**(Pt 9):2091–2096.
- Cane, P. A., and C. R. Pringle. 1995. Evolution of subgroup A respiratory syncytial virus: evidence for progressive accumulation of amino acid changes in the attachment protein. *J. Virol.* **69**:2918–2925.
- Cane, P. A., and C. R. Pringle. 1995. Molecular epidemiology of respiratory syncytial virus: a review of the use of reverse transcription-polymerase chain reaction in the analysis of genetic variability. *Electrophoresis* **16**:329–333.
- Choi, E. H., and H. J. Lee. 2000. Genetic diversity and molecular epidemiology of the G protein of subgroups A and B of respiratory syncytial viruses isolated over 9 consecutive epidemics in Korea. *J. Infect. Dis.* **181**:1547–1556.
- Coggins, W. B., E. J. Lefkowitz, and W. M. Sullender. 1998. Genetic variability among group A and group B respiratory syncytial viruses in a children's hospital. *J. Clin. Microbiol.* **36**:3552–3557.
- Collins, P. L., R. M. Chanock, and B. R. Murphy. 2001. *Fields virology*, 4th ed. Lippincott Williams & Wilkins, Philadelphia, Pa.
- Cristina, J., J. A. Lopez, C. Albo, B. Garcia-Barreno, J. Garcia, J. A. Melero, and A. Portela. 1990. Analysis of genetic variability in human respiratory syncytial virus by the RNase A mismatch cleavage method: subtype divergence and heterogeneity. *Virology* **174**:126–134.
- de Sierra, M., and J. Arbiza. 2004. Genetic stability of the attachment glycoprotein of human respiratory syncytial viruses during serial passages in cell cultures. *Acta Virol.* **48**:115–121.
- Frabasile, S., A. Delfraro, L. Facal, C. Videla, M. Galiano, M. J. de Sierra, D. Ruchansky, N. Vitoreira, M. Berois, G. Carballal, J. Russi, and J. Arbiza. 2003. Antigenic and genetic variability of human respiratory syncytial viruses (group A) isolated in Uruguay and Argentina: 1993–2001. *J. Med. Virol.* **71**:305–312.
- Hansen, J. E., O. Lund, J. Engelbrecht, H. Bohr, and J. O. Nielsen. 1995. Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase. *Biochem. J.* **308**(Pt 3):801–813.
- John, T. J., T. Cherian, M. C. Steinhoff, E. A. Simoes, and M. John. 1991. Etiology of acute respiratory infections in children in tropical southern India. *Rev. Infect. Dis.* **13**(Suppl. 6):S463–S469.
- Johnson, P. R., M. K. Spriggs, R. A. Olmsted, and P. L. Collins. 1987. The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proc. Natl. Acad. Sci. USA* **84**:5625–5629.
- Karron, R. A., D. A. Buonagurio, A. F. Georgiu, S. S. Whitehead, J. E. Adamus, M. L. Clements-Mann, D. O. Harris, V. B. Randolph, S. A. Udem, B. R. Murphy, and M. S. Sidhu. 1997. Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral replication in vitro: clinical evaluation and molecular characterization of a cold-passaged, attenuated RSV subgroup B mutant. *Proc. Natl. Acad. Sci. USA* **94**:13961–13966.
- Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* **5**:150–163.
- Kuroiwa, Y., K. Nagai, L. Okita, I. Yui, T. Kase, T. Nakayama, and H. Tsutsumi. 2005. A phylogenetic study of human respiratory syncytial viruses group A and B strains isolated in two cities in Japan from 1980–2002. *J. Med. Virol.* **76**:241–247.
- Lim, C. S., G. Kumarasinghe, and V. T. Chow. 2003. Sequence and phylogenetic analysis of SH, G, and F genes and proteins of human respiratory syncytial virus isolates from Singapore. *Acta Virol.* **47**:97–104.
- Madhi, S. A., M. Venter, R. Alexandra, H. Lewis, Y. Kara, W. F. Karshagen, M. Greef, and C. Lassen. 2003. Respiratory syncytial virus associated illness

- in high-risk children and national characterisation of the circulating virus genotype in South Africa. *J. Clin. Virol.* **27**:180–189.
22. Maitreyi, R. S., S. Broor, S. K. Kabra, M. Ghosh, P. Seth, L. Dar, and A. K. Prasad. 2000. Rapid detection of respiratory viruses by centrifugation enhanced cultures from children with acute lower respiratory tract infections. *J. Clin. Virol.* **16**:41–47.
 23. Misra, P. K., R. S. Chaudhary, A. Jain, A. Pande, A. Mathur, and U. C. Chaturvedi. 1990. Viral aetiology of acute respiratory infections in children in north India. *J. Trop. Pediatr.* **36**:24–27.
 24. Moura, F. E., A. Blanc, S. Frabasile, A. Delfraro, M. J. de Sierra, L. Tome, E. A. Ramos, M. M. Siqueira, and J. Arbiza. 2004. Genetic diversity of respiratory syncytial virus isolated during an epidemic period from children of northeastern Brazil. *J. Med. Virol.* **74**:156–160.
 25. Mufson, M. A., C. Orvell, B. Rafnar, and E. Norrby. 1985. Two distinct subtypes of human respiratory syncytial virus. *J. Gen. Virol.* **66**(Pt 10):2111–2124.
 26. Nagai, K., H. Kamasaki, Y. Kuroiwa, L. Okita, and H. Tsutsumi. 2004. Nosocomial outbreak of respiratory syncytial virus subgroup B variants with the 60 nucleotides-duplicated G protein gene. *J. Med. Virol.* **74**:161–165.
 27. Nei, M., and T. Gojibori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**:418–426.
 28. Nicholas, K. B., H. B. J. Nicholas, and D. W. I. Deerfield. 1997. GeneDoc: analysis and visualization of genetic variation. *EMBLNEWS* **4**:14.
 29. Nokes, D. J., E. A. Okiro, M. Ngama, L. J. White, R. Ochola, P. D. Scott, P. A. Cane, and G. F. Medley. 2004. Respiratory syncytial virus epidemiology in a birth cohort from Kilifi district, Kenya: infection during the first year of life. *J. Infect. Dis.* **190**:1828–1832.
 30. Ota, W. K., and F. B. Bang. 1972. A continuous study of viruses in the respiratory tract in families of a Calcutta bustee. I. Description of the study area and patterns of virus recovery. *Am. J. Epidemiol.* **95**:371–383.
 31. Patwari, A. K., S. Bisht, A. Srinivasan, M. Deb, and D. Chattopadhyay. 1996. Aetiology of pneumonia in hospitalized children. *J. Trop. Pediatr.* **42**:15–20.
 32. Peret, T. C., C. B. Hall, G. W. Hammond, P. A. Piedra, G. A. Storch, W. M. Sullender, C. Tsou, and L. J. Anderson. 2000. Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. *J. Infect. Dis.* **181**:1891–1896.
 33. Peret, T. C., C. B. Hall, K. C. Schnabel, J. A. Golub, and L. J. Anderson. 1998. Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. *J. Gen. Virol.* **79**(Pt 9):2221–2229.
 34. Rajala, M. S., W. M. Sullender, A. K. Prasad, L. Dar, and S. Broor. 2003. Genetic variability among group A and B respiratory syncytial virus isolates from a large referral hospital in New Delhi, India. *J. Clin. Microbiol.* **41**:2311–2316.
 35. Reddaiah, V. P., and S. K. Kapoor. 1988. Acute respiratory infections in rural underfives. *Indian J. Pediatr.* **55**:424–426.
 36. Roca, A., M. P. Loscertales, L. Quinto, P. Perez-Brena, N. Vaz, P. L. Alonso, and J. C. Saiz. 2001. Genetic variability among group A and B respiratory syncytial viruses in Mozambique: identification of a new cluster of group B isolates. *J. Gen. Virol.* **82**:103–111.
 37. Sato, M., R. Saito, T. Sakai, Y. Sano, M. Nishikawa, A. Sasaki, Y. Shobugawa, F. Gejyo, and H. Suzuki. 2005. Molecular epidemiology of respiratory syncytial virus infections among children with acute respiratory symptoms in a community over three seasons. *J. Clin. Microbiol.* **43**:36–40.
 38. Scott, P. D., R. Ochola, M. Ngama, E. A. Okiro, D. J. Nokes, G. F. Medley, and P. A. Cane. 2004. Molecular epidemiology of respiratory syncytial virus in Kilifi district, Kenya. *J. Med. Virol.* **74**:344–354.
 39. Shann, F., and M. C. Steinhoff. 1999. Vaccines for children in rich and poor countries. *Lancet* **354**(Suppl. 2):SII7–SIII11.
 40. Sullender, W. M. 2000. Respiratory syncytial virus genetic and antigenic diversity. *Clin. Microbiol. Rev.* **13**:1–15.
 41. Sullender, W. M., M. A. Mufson, L. J. Anderson, and G. W. Wertz. 1991. Genetic diversity of the attachment protein of subgroup B respiratory syncytial viruses. *J. Virol.* **65**:5425–5434.
 42. Sullender, W. M., L. Sun, and L. J. Anderson. 1993. Analysis of respiratory syncytial virus genetic variability with amplified cDNAs. *J. Clin. Microbiol.* **31**:1224–1231.
 43. Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**:4876–4882.
 44. Trento, A., M. Galiano, C. Videla, G. Carballal, B. Garcia-Barreno, J. A. Melero, and C. Palomo. 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.
 45. Trento, A., M. Viegas, M. Galiano, C. Videla, G. Carballal, A. S. Mistchenko, and J. A. Melero. 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.
 46. Venter, M., M. Collinson, and B. D. Schoub. 2002. Molecular epidemiological analysis of community circulating respiratory syncytial virus in rural South Africa: comparison of viruses and genotypes responsible for different disease manifestations. *J. Med. Virol.* **68**:452–461.
 47. Venter, M., S. A. Madhi, C. T. Tiemessen, and B. D. Schoub. 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.
 48. Wertz, G. W., P. L. Collins, Y. Huang, C. Gruber, S. Levine, and L. A. Ball. 1985. Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein. *Proc. Natl. Acad. Sci. USA* **82**:4075–4079.
 49. Williams, B. G., E. Gouws, C. Boschi-Pinto, J. Bryce, and C. Dye. 2002. Estimates of world-wide distribution of child deaths from acute respiratory infections. *Lancet Infect. Dis.* **2**:25–32.
 50. World Health Organization. 2000. Integrated management of childhood illness handbook. Report WHO/FCH/CAH/00.12. World Health Organization, Geneva, Switzerland.
 51. Zlateva, K. T., P. Lemey, E. Moes, A. M. Vandamme, and M. Van Ranst. 2005. Genetic variability and molecular evolution of the human respiratory syncytial virus subgroup B attachment G protein. *J. Virol.* **79**:9157–9167.
 52. Zlateva, K. T., P. Lemey, A. M. Vandamme, and M. Van Ranst. 2004. Molecular evolution and circulation patterns of human respiratory syncytial virus subgroup a: positively selected site in the attachment G glycoprotein. *J. Virol.* **78**:4675–4683.