Molecular Analysis of Respiratory Syncytial Virus Reinfections in Infants from Coastal Kenya

Paul D. Scott,1,a Rachel Ochola,1 Mwanajuma Ngama,1 Emelda A. Okiro,4 D. James Nokes,2,4 Graham F. Medley,2 and Patricia A. Cane1,3

1Division of Immunity and Infection, University of Birmingham, Birmingham, 2Ecology and Epidemiology Group, Department of Biological Sciences, University of Warwick, Warwick, and 3Centre for Geographic Medicine Research–Coast, Kenya Medical Research Institute, Kilifi, Kenya

Background. Individuals are reinfected with respiratory syncytial virus (RSV) repeatedly. The nature of re-infection, in relation to RSV genetic and antigenic diversity, is ill defined and has implications for persistence and vaccine control.

Methods. We examined the molecular relatedness of RSV causing primary and repeat infections, by phylogenetic analysis of the attachment (G) gene in 12 infants from a birth cohort in rural Kenya, using nasal wash samples collected during a 16-month period in 2002–2003, which spanned 2 successive epidemics.

Results. Six infants were infected during both epidemics, 4 with RSV-A in the first epidemic followed by RSV-B during the second epidemic and 2 with RSV-A during both epidemics, with no significant G gene sequence variability between samples. Two infants were infected and reinfected with different RSV-A strains during the same epidemic. Possible viral persistence was suspected in the remaining 4 infants, although reinfection with the same variant cannot be excluded.

Conclusions. These are the first data that specifically address strain-specific reinfections in infancy in relation to the primary infecting variant. The data strongly suggest that, following primary infection, some infants lose strain-specific immunity within 7–9 months (between epidemics) and group-specific immunity within 2–4 months (during an epidemic period).

Respiratory syncytial virus (RSV) causes a significant burden of acute respiratory disease during infancy and childhood in both the developed and the developing worlds, characteristically occurring in recurrent, discrete epidemics [1–3]. First (primary) infection with RSV occurs early in life, usually before the age of 3 years, and is associated with a high risk of lower respiratory tract infection (LRTI) [4]. Immunity following primary exposure, however, does not prevent secondary or subsequent infections [5], and reinfections with RSV have been recorded throughout life [6], although with a lower risk of severe disease [4]. The absence of solid immunity to RSV infection is likely to be of importance to the persistence of infection within a population and of relevance to the potential impact of future vaccines. However, this phenomenon is inadequately understood. A possible role for RSV antigenic and genetic diversity in the reinfection process exists, although, presently, this is poorly defined. RSV isolates can be divided into 2 groups, A and B, that are distinct at both the antigenic and the nucleotide sequence levels. Both groups often cocirculate during epidemics, but group A isolates more commonly predominate. In addition, the groups can be subdivided into several strains or genotypes that also cocirculate during the epidemics; the predominant strain is typically replaced each year, possibly as a result of localized herd immunity to particular variants [7–10]. The molecular epidemiological evidence suggests that group or
genotype infection prevalence influences future transmission of the homologous and heterologous variants within a population—a notion that is supported by recent mathematical modelling studies [11].

At the individual level, there have been few studies of RSV reinfection, particularly in the context of the group or genotype of infecting strains. Natural reinfections with the homologous and heterologous group of RSV have been shown. Mufson et al. [12] reported 13 children (age range, 6–49 months) with RSV infection who were subsequently reinfected at least 9 months after the first infection. Six reinfections with group B occurred in 10 children initially infected with group A, and 2 reinfections with group B occurred in 3 children initially infected with group B; these data provide little support for greater homologous, compared with heterologous, protection. Sullender et al. [6] studied 2 children (aged 7 and 23 months at the time of the first infection), each with sequential (separated by >1 year) group A virus infections, for whom the attachment (G) protein amino acid sequences of the reinfecting viruses were up to 15% different. In a study of adult volunteers who were repeatedly challenged with the same strain of virus [13], Hall et al. found that reinfection could occur within a few months of initial exposure. Although experimental, these data raise questions about the strength of homologous protection. Furthermore, a study of hospitalized patients from 1981 to 1990 in Finland [14] found no direct evidence for homologous protection in the 18 cases of reinfection that were typed—the reinfecting strain reflected the predominant group in circulation. Readmissions with infections with the same RSV group were always during an epidemic. Differentiating repeat infection from delayed clearance of the same infection, in the absence of genotyping, was not possible.

We have previously described the incidence of RSV infection and the risk of associated disease in a birth cohort from Kilifi District, Kenya, during the first year of life [15]. We now report on a molecular analysis of the viruses from infants in this cohort who experienced primary followed by repeat infections during the course of 2 sequential epidemics and compare these data with those on the viruses characterized during these epidemics from the whole cohort [16].

PATIENTS, MATERIALS, AND METHODS

Study population and samples. The study was performed in Kilifi District, a rural coastal area of Kenya, and forms part of an epidemiological investigation of RSV through the intensive surveillance of a birth cohort, details of which have been described elsewhere [15, 16]. Infants were recruited at or near birth, between January and June 2002, and were monitored through active household visits weekly during the epidemic periods and otherwise every 4 weeks and through passive referral to a research outpatient clinic at the Kilifi District Hos-

...
AY524625, AY660677, AY660679, AY660681, and AY773286–AY773301 [inclusive]).

RESULTS

Details of age at infection, the length of time between each RSV infection, and the clinical status of each RSV infection for the 12 infants studied (Ken01–Ken12) are given in figure 1 and table 1. Six infants were found to have separate RSV infections during each of the 2 epidemics, and 6 infants were found to have 2 or, in 1 instance, 3 separate RSV infections during the first epidemic alone. No child experienced reinfections in both epidemics. All infants in the present study were in their first year of life; 2 infants were <40 days old. Of the 12 primary infections, 7 were LRTIs, 5 of which were severe. Of the 13 repeat infections (1 child had 2), 2 involved mild pneumonia, and 2 caused severe pneumonia, and only 1 of the 6 reinfections (infants aged 10 or 11 months) during the second epidemic involved the lower respiratory tract (severe case in Ken08).

Strain typing by RFLP. Results of N gene RFLP typing are presented in table 1. All infants were infected with RSV-A strains during the first epidemic (figure 1). Of the 12 primary infections, 4 were of genotype NP2, and 8 were of genotype NP4. Of the 6 infants reinfected during the second epidemic, 4 had RSV-B genotype NP3 infections, and 2 had RSV-A genotype NP4 infections (same as initial infection). Of these 6 infants, 0 of 2 infected with the same strain developed LRTI, and 1 of 4 infected with group B developed severe LRTI. Two of the 6 infants who received repeat diagnoses of RSV infection during the same epidemic were reinfected with a different RSV-A strain (Ken04, NP4 to NP2; Ken09, NP2 to NP4). The remaining 4 infants (Ken01, Ken06, Ken07, and Ken10) were infected with a virus that was of the same group and genotype, and, in Ken04, a repeat infection with NP4 preceded reinfection with NP2. Of the 4 reinfections with the same genotype, 2 resulted in mild pneumonia, and 1 of the 3 reinfections with a different genotype resulted in severe pneumonia. The overall distribution of genotypes of viruses was similar in the infections and reinfections reported here compared with those described previously for these epidemics [16].

Sequencing and phylogenetic analysis of the G gene: comparisons with circulating strains in each epidemic. The G genes (nt 284–912 for RSV-A and nt 194–915 for RSV-B) from all RSV-positive samples (n = 25) from the 12 infants were sequenced and analyzed phylogenetically. Comparisons were made with other samples from Kilifi District, details of which have been published elsewhere [16], as well as with reference isolates from around the world. All of these samples were found to cluster with others from Kilifi District collected during the same period (figures 2 and 3). Six (29%) of 21 samples clustered with others in cluster A1, and 15 (71%) of 21 clustered with
Table 1. Clinical information for each child.

<table>
<thead>
<tr>
<th>Child (sex), sample date</th>
<th>Age at diagnosis, days (months)</th>
<th>Time since last diagnosis, days (months)</th>
<th>Clinical status(^b)</th>
<th>NP type</th>
<th>Genotype(^c) (GenBank accession no.)</th>
<th>Type of infecting virus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ken01 (F)</td>
<td>May 2002 90 (3)</td>
<td>...</td>
<td>Severe pneumonia</td>
<td>NP2</td>
<td>A1 (AY524610)</td>
<td>Virtually identical virus over the course of 28 days</td>
</tr>
<tr>
<td></td>
<td>Jun 2002 118 (4)</td>
<td>28 (1)</td>
<td>URTI</td>
<td>NP2</td>
<td>A1 (AY524614)</td>
<td>Very similar virus during 2 epidemics</td>
</tr>
<tr>
<td>Ken02 (M)</td>
<td>Jun 2002 119 (4)</td>
<td>...</td>
<td>Mild pneumonia</td>
<td>NP4</td>
<td>A2 (AY773287)</td>
<td>Very similar virus during 2 epidemics</td>
</tr>
<tr>
<td></td>
<td>Jan 2003 344 (11)</td>
<td>225 (7)</td>
<td>URTI</td>
<td>NP4</td>
<td>A2 (AY660679)</td>
<td>Identical virus for 3 weeks and then a different group A genotype 6 weeks later</td>
</tr>
<tr>
<td>Ken03 (F)</td>
<td>Apr 2002 52 (2)</td>
<td>...</td>
<td>Severe pneumonia</td>
<td>NP4</td>
<td>A2 (AY773286)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
<tr>
<td></td>
<td>Jan 2003 328 (11)</td>
<td>276 (9)</td>
<td>URTI</td>
<td>NP4</td>
<td>A2 (AY660677)</td>
<td>Identical virus for 7 weeks</td>
</tr>
<tr>
<td>Ken04 (M)</td>
<td>Apr 2002 59 (2)</td>
<td>...</td>
<td>Severe pneumonia</td>
<td>NP4</td>
<td>A2 (AY773299)</td>
<td>Identical virus for 24 days</td>
</tr>
<tr>
<td></td>
<td>May 2002 86 (3)</td>
<td>35 (1)</td>
<td>Mild pneumonia</td>
<td>NP4</td>
<td>A2 (AY773297)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
<tr>
<td>Ken05 (M)</td>
<td>Apr 2002 68 (2)</td>
<td>...</td>
<td>Severe pneumonia</td>
<td>NP4</td>
<td>A2 (AY773296)</td>
<td>Identical virus for 24 days</td>
</tr>
<tr>
<td></td>
<td>May 2002 92 (3)</td>
<td>24 (1)</td>
<td>URTI</td>
<td>NP4</td>
<td>A2 (AY773295)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
<tr>
<td>Ken06 (M)</td>
<td>May 2002 61 (2)</td>
<td>...</td>
<td>Severe pneumonia</td>
<td>NP4</td>
<td>A2 (AY773290)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
<tr>
<td></td>
<td>Dec 2002 292 (10)</td>
<td>231 (8)</td>
<td>Severe pneumonia</td>
<td>NP3</td>
<td>B1 (AY773291)</td>
<td>Different group A viruses (infected twice; the second infection was more severe) during the first epidemic</td>
</tr>
<tr>
<td>Ken07 (F)</td>
<td>Mar 2002 25 (1)</td>
<td>...</td>
<td>URTI</td>
<td>NP2</td>
<td>A1 (AY524606)</td>
<td>Identical virus for 19 days</td>
</tr>
<tr>
<td></td>
<td>Jul 2002 138 (5)</td>
<td>113 (4)</td>
<td>Severe pneumonia</td>
<td>NP4</td>
<td>A2 (AY524621)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
<tr>
<td>Ken08 (F)</td>
<td>May 2002 75 (2)</td>
<td>...</td>
<td>URTI</td>
<td>NP4</td>
<td>A2 (AY773294)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
<tr>
<td></td>
<td>Jun 2002 94 (3)</td>
<td>19 (1)</td>
<td>Mild pneumonia</td>
<td>NP4</td>
<td>A2 (AY524622)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
<tr>
<td>Ken09 (F)</td>
<td>Jun 2002 84 (3)</td>
<td>...</td>
<td>URTI</td>
<td>NP2</td>
<td>A1 (AY524625)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
<tr>
<td></td>
<td>Feb 2003 321 (11)</td>
<td>231 (8)</td>
<td>URTI</td>
<td>NP3</td>
<td>B1 (AY773289)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
<tr>
<td>Ken10 (F)</td>
<td>Apr 2002 39 (1)</td>
<td>...</td>
<td>URTI</td>
<td>NP4</td>
<td>A2 (AY773293)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
<tr>
<td></td>
<td>Jan 2003 326 (11)</td>
<td>287 (9)</td>
<td>URTI</td>
<td>NP3</td>
<td>B2 (60 nt(^d)) (AY660681)</td>
<td>Group A virus during the first epidemic, group B virus during the second epidemic</td>
</tr>
</tbody>
</table>

**NOTE.** NP, nucleoprotein; URTI, upper respiratory tract infection.

\(^a\) Months were rounded to the nearest integer.

\(^b\) Clinical status is indicated according to the definitions of Nokes et al. [15]. No child was admitted to the hospital.

\(^c\) G gene genotypes are indicated according to the systems described elsewhere [16].

\(^d\) The sample has a 60-nt duplication, as described by Trento et al. [24].
Figure 2. Phylogenetic tree comparing the respiratory syncytial virus–A reinfection samples from Kilifi District, Kenya, with other samples from Kilifi District and sequences from GenBank (nt 284–912 of the G gene). Internal node labels represent the bootstrap values for 1000 iterations. Reinfection samples are represented as KenXX, where XX is a 2-digit number. The scale bar represents 0.1 nt substitutions/site. Other samples from Kenya during the same epidemics are represented as Ken/xxx/yy, where xxx represents a unique sample identification and yy represents the epidemic year.

A2 for RSV-A (figure 2). Clusters A1 and A2 correspond to previously described RSV genotypes SHL1/3/4 and SHL2 [20] or GA2 and GA5, respectively [10, 16, 21–23].

All but 1 (3/4) of the RSV-B infection samples (all reinfections) clustered with other samples from Kilifi District (figure 3) within the B1 cluster (genotype SAB1 [22]). The reinfection sample that was distinct from the others was found to have a 60-nt duplication within the G gene, as described elsewhere by Trento et al. [24], and belonged to the B2 cluster (genotype BA).

**Infants reinjected during separate epidemics.** Ken02 had minor sequence variability between primary and secondary RSV-A infections, which were 7 months apart, with 3 of 629 noncoding nucleotide changes in the samples. The G gene sequences of the samples from Ken03 from each epidemic also had very similar sequences, with 3 of 629 noncoding nucleotide changes. Both of these infants thus appear to have been infected with the same RSV-A variant (A2) (figure 2), since the G gene sequences were highly similar.

Both Ken05 and Ken11 were infected with an RSV-A A1 variant during the first epidemic and then with an RSV-B B1
virus during the second epidemic (figures 2 and 3). Ken08 and Ken12 were infected with RSV-A A2 viruses during the first epidemic and then with RSV-B viruses during the second epidemic. Ken08 was reinfected with an RSV-B B1 virus, whereas Ken12 was reinfected with the only RSV-B B2 variant to be found in the present study.

**Infants reinfected during the same epidemic.** Ken04 was infected with an RSV-A A2 variant, then by A2 again after 3 weeks, then by an A1 variant within a period of almost 2 months from the primary infection, or 42 days from the date that the second A2 strain was identified. Conversely, Ken09 had a primary infection with an RSV-A A1 variant, followed by reinfection with an A2 variant during a period of nearly 4 months (table 1).

Four other infants from whom multiple samples were collected during the same epidemic (2002) (Ken01, Ken06, Ken07, and Ken10) had no significant phylogenetic differences in any of the samples (Ken06, Ken07, and Ken10 had no nucleotide changes, and Ken01 had only 1 coding nucleotide change between the samples). All of the samples from these infants were collected ~1 month apart, suggesting the possibility that their
infections were prolonged single infections rather than reinfections, as is likely for the 2 A2 variants identified in Ken04.

Amino acid comparisons between primary and secondary infections. Analysis of the amino acid sequences (codons 91–298 for RSV-A and 60–297 for RSV-B) of all primary and secondary infection samples was performed (data not shown). All RSV-A sequences were the same length, with the stop codon at the same position, and all had 6 potential N glycosylation sites. All but 1 of these glycosylation sites were shared by all samples examined. The exception was within the carboxy terminus of the gene product, and the potential N glycosylation site was moved by 1 aa downstream in samples corresponding to the A1 genotype, which was represented by NTT in the sequences. The A2 variants had NST in this area. The 2 infants who were reinfected with a different variant (Ken04 and Ken09) had clear differences in the G gene deduced amino acid sequences. Comparisons of the amino acid sequences within and between patients showed very little difference when the variant of RSV was the same genotype (maximum difference, 4% and <1% for the A1 and A2 variants, respectively). Intergenotype differences were found to be 19% in the region examined.

The RSV-B amino acid analysis demonstrated that the duplicated sequence was at codons 260–279 of the G gene product, and samples from variants without this duplication had a 4-aa extension to the end of the gene product, which was not present in RSV-B variants circulating during the previous epidemic but was representative of the current epidemic. The variants with the duplication had 4 potential N glycosylation sites, and those without the duplication had 6 sites. Comparisons of the amino acid sequences within and between patients showed very little difference when the variant of RSV was the same genotype (maximum difference, <1%). However, cross-genotype differences were much greater (maximum difference, 20% and 12% including and excluding the duplication, respectively).

DISCUSSION

The mechanism by which RSV reinfects the host is unclear. Given that RSV displays considerable genetic and antigenic variation [25], which is thought to be under immune selection [23, 26], there is good reason to suppose that the RSV group or genotype may play a role in the avoidance of immune protection, although evidence is very limited at present. The lack of solid immunity following primary infection is of particular importance for the development of effective infant vaccines. Data are therefore specifically required to characterize the relationship between the reinfected and the primary infecting variants. Despite the fact that the sample size was small, the results presented here, from a rural coastal region of Kenya, provide valuable information on the molecular characteristics of RSV reinfections, and the present study is, to our knowledge, the first such study to target the key infant age group. Furthermore, these data provide the most-detailed information available to date on repeat and persistent infections in the general community, compared with hospital inpatient- and outpatient-based studies [6, 12, 27] or an adult-based study [13]. We have shown that reinfection with RSV occurred not only during the first year of life but also during the same epidemic and, importantly, that reinfections can arise with the same variant within the space of 7–9 months and in the same group but with a different variant within the space of 2–4 months. The virus strains in these infants reflect those being detected generally at the time in Kilifi District, as has been shown by phylogenetic analysis of G gene data [16]. All RSV G gene sequences from the first epidemic were RSV-A variants, as was expected, since this group represented 98% of strains during the 2002 epidemic [16]. Similarly, RSV-B predominated during the second epidemic (61%) [16], and 4 of 6 samples from the reinfected infants contained RSV-B variants. One of these infants was reinfected with a variant of RSV-B that contained a 60-nt duplication within the G gene, which has been described elsewhere [16, 24, 28]. However, because of the small numbers of samples, it is not possible to state whether the presence of this emerging strain of RSV-B has any significance in this community.

Protective immunity can be transient to a virus of the same group in some individuals, as was suggested by the findings on Ken04 and Ken09, who were reinfected with a different strain of RSV-A ~2–3 months after the first infection (RSV-A), both during the first epidemic. However, this observation does not preclude the possible existence of some form of genotype-specific immunity in the infants reinfected with the heterologous group of virus.

A possible explanation for the apparent reinfection with the same virus 7 and 9 months later in 2 infants (Ken02 and Ken03, respectively) is that strain-specific immunity wanes in some individuals during this period, allowing reinfection. Alternatively, these 2 infants may not have been reinfected but rather had very prolonged viral persistence/shedding. However, this would seem unlikely: samples collected between infections from each infant were negative by RT-PCR (data not shown). Experimental reinfection with the same virus within 2 months has been demonstrated in adults [13], so, perhaps, it is possible that reinfection did occur in these infants. Comparisons of the amino acid sequences of the G genes from each of the infections between both infants showed that they are identical. Further analysis would be necessary to ascertain whether this particular RSV strain is prone to immune evasion, especially in light of the fact that intermediate samples were negative by both RT-PCR and IFAT.

The samples collected 1 month apart from the remaining 4 infants (and the first 2 samples from Ken04) with the same G gene amino acid sequences imply that viral persistence can occur for at least 1 month. For infants for whom prolonged
shedding was a possibility (Ken01–Ken04, Ken06, Ken07, and Ken10), we investigated potential predisposing factors from outpatient and Kilifi District Hospital admission records; no relevant underlying conditions—such as congenital disease, malnutrition, or immunocompromised state—were identified.

Repeat analysis of all nasal wash samples yielded the same sequences as those presented here, so it is unlikely that there were errors in RT-PCR and sequencing. Consequently, the 2 infants who were reinfected with a different RSV-A strain during the same epidemic (Ken04 and Ken09) were probably reinfected, rather than coinfected with both strains at the same time. This was also highlighted by the nucleoprotein typing of samples from these 2 infants: only 1 N gene pattern was noted at any time.

Most of the amino acid variation was associated with the variable regions flanking the conserved cysteine-rich domain, as was expected. The significance of occasional amino acid differences between samples of the same genotype is not known but may be a result of the host’s immune pressure on the virus during the course of infection.

Susceptibility to reinfection with RSV is a result of the individual’s immunity to previous infections and the currently circulating strains. Thus, individual risks of reinfection are combinations of factors both the individual child (immunity and previous exposure) and the population (previous circulating strains and current circulating strains). The factors at the population level are determined by the individuals within the population—for example, the most-parsimonious explanation for strain replacement between epidemics is that it is driven by individual (homologous) immunity. However, that the RSV variants in each of the 2 epidemics reported here are similar (and, in the case of 2 infants, reinfection occurred with identical variants) makes interpretation of the transmission and genetic dynamics less straightforward.

Further information on the pattern of infection and reinfection observed may be provided by the families of these infants, but these data are not available for this particular cohort. It should also be noted that, in the case of influenza, there is the hypothesis of “original antigenic sin,” which suggests that the response to secondary (and subsequent) infections can depend, at least partially, on the immune response toward the primary infection. This may be the case for RSV, but many more reinfections would need to be studied to elucidate this.

In summary, we have demonstrated multiple RSV infections during the first year of life in a community study of a birth cohort. During 2 consecutive epidemics, separated by a gap of 7–9 months, some infants were apparently reinfected with the same strain, whereas others were reinfected with the heterologous group (RSV-A followed by RSV-B). During the same epidemic, we observed infection and reinfection with different RSV-A strains within 2–4 months, suggesting transient immunity to the homologous group of RSV in some infants. These data are important in questioning the nature of immunity to reinfection in relation to the primary infecting and reinfecting variant in the infant.

**Acknowledgments**

We thank the study volunteers for generous participation. We thank also the respiratory syncytial virus team, for tireless work and enthusiasm, and the research outpatient–clinic field workers and clinical officers; the pediatric ward, maternity, and Maternity Child Health Clinic staff; and the senior hospital staff from Kilifi District Hospital, for support and cooperation in the running of this study.

**References**