Co-circulation of multiple rubella virus strains in Belarus forming novel genetic groups within clade 1

Judith M. Hübschen,1 Marina Yermalovich,2 Galina Semeiko,2 Elena Samoilovich,2 Elena Blatun,3 Sébastien De Landtsheer1 and Claude P. Muller1

1Institute of Immunology and WHO Collaborative Centre for Measles and WHO European Regional Reference Laboratory for Measles and Rubella, Laboratoire National de Santé, 20A rue Auguste Lumière, L-1011 Luxembourg
2Research Institute for Epidemiology and Microbiology, Minsk, Belarus
3Infectious Disease Hospital, Minsk, Belarus

Although the WHO recommends a comprehensive genetic characterization, little is known about circulating strains and genotypes of rubella virus (RUBV) for many European countries. Studies investigating the genetic diversity of a sizeable number of strains from a certain location are rare. This study presents the first molecular characterization of isolates from Belarus. Throat swab and urine samples were collected between 2004 and 2005 from patients presenting in two infectious disease hospitals and three outpatient clinics in and around Minsk. In total, 14 isolates were obtained from this clinical material. Phylogenetic analysis of the E1 gene sequence of these isolates showed that three distinct groups of RUBV strains co-circulated. One group of isolates was assigned to genotype 1E, whereas the other two did not group with any of the recognized genotypes but grouped with a strain belonging to the provisional genotype 1g. Detailed analysis showed that the group comprising 1g strains also contained sequences formerly attributed to genotype 1B and could be divided into four subgroups, one of which might represent a putative novel provisional genotype of clade 1. These findings show that three distinct strains with limited variability are present in Belarus, suggesting independent introductory events. As there currently seem to be misattributions of strains to genotypes and unclear phylogenetic relationships, criteria for genotyping of RUBV should be clarified further.

INTRODUCTION

Acute infection with rubella virus (RUBV) normally causes only mild symptoms and may even remain asymptomatic. However, serious birth defects known as congenital rubella syndrome (CRS) are frequently reported as a consequence of infection during early pregnancy. Currently, it is estimated that more than 100 000 CRS cases occur worldwide every year (Robertson et al., 2003), despite the availability of an effective live-attenuated vaccine since 1969. Many countries have included rubella vaccines in their routine vaccination programme and the majority of European countries use combined measles/rubella vaccines (WHO, 2005a). The World Health Organization (WHO) European Region aims to eliminate measles and rubella and to prevent congenital rubella infection by 2010 (WHO, 2005a). To support these control activities, WHO considers the molecular epidemiology of RUBV to be of increasing importance (WHO, 2005c). Recently, a standardized nomenclature for the classification and designation of wild-type RUBV was proposed. Based on a part of the E1 gene sequence (nt 8731–9469), two distinct clades comprising seven recognized and three provisional genotypes were defined (WHO, 2005c). However, the geographical distribution of genotypes in the WHO European Region is poorly understood. For a number of countries, no information on circulating RUBV genotypes is yet available.

In Belarus, vaccination against rubella was introduced in 1996. One dose of measles/mumps/rubella (MMR) combined vaccine was administered at the age of 12 months (Samoilovich et al., 2000). Since the year 2000, revaccination of 6-year-old children with the same vaccine has been implemented (Samoilovich, 2005). For both doses, a vaccination coverage of more than 98 % has been reported during the last few years and rubella has virtually disappeared among children born after 1995. However, epidemics continue to occur every 5 years, although with considerably lower case numbers than observed before the start of vaccination (65 562 in 1994, 44 443 in 1999, and 4492 and 3812 in 2004/2005). During the pre-vaccination
period, rubella occurred mainly among 1–14-year-olds, whereas currently mostly 10–19-year-old adolescents and young adults are affected (Samoilovich et al., 2005). Such a development is often observed during the implementation period of vaccination programmes and can even lead to an increase in the rate of CRS cases (Reef et al., 2002). Therefore, an additional vaccination with monovalent RUBV vaccine was offered to this age group in October 2005. Despite the high but decreasing number of rubella cases, the virus strains circulating in Belarus have never been genotyped or characterized on a molecular level.

METHODS

Specimens. Thirty-three urine and 27 nasopharyngeal swab samples were collected from 45 patients presenting with a rash in two infectious disease hospitals and three outpatient clinics in Minsk city and the Minsk region between June 2004 and July 2005. Infection by RUBV was confirmed serologically in all cases by specific IgM (Dade Behring Enzynost immunoassay). Specimens were inoculated onto Vero cell cultures and passaged up to six times. To avoid cross-contamination, only one sample was handled at a time. Cell culture supernatant was harvested and was either used immediately for RNA extraction or stored at −70 °C. Fourteen RUBV isolates were obtained from these clinical specimens and characterized further.

RNA extraction and PCR. RNA was extracted according to the manufacturer’s protocol from 140 μl virus culture supernatant using a QIAamp Viral RNA Mini kit (Qiagen). Reverse transcription was carried out in a 20 μl reaction containing 1 μl SuperScript III reverse transcriptase (200 U), 0.5 μl RNaseOUT Recombinant RNase Inhibitor (20 U), 2 μl 0.1 M DTT, 4 μl 5× First-strand Buffer, 1 μl 10 mM dNTP mix (Invitrogen), 1 μl 40 μM gene-specific reverse primer 3′E1 (Zheng et al., 2003b), 5.5 μl water and 5 μl of the extracted RNA. The initial denaturation at 65 °C for 5 min was carried out without the enzymes and was followed by 80 min at 55 °C and 10 min at 72 °C. For the diagnostic PCR, previously published primers (Eggerding et al., 1991) were used as well as to generate fragments fit for sequencing (Bosma et al., 1996; Cooray et al., 2006; Eggerding et al., 1991; Katow et al., 1997; Vyse & Jin, 2002; Zheng et al., 2003b). In addition, reverse primers RVir (5′-TTTCTTATRC-AGCAGCGGTCG-3′) and R9124a (5′-GATTCCGGACTTGCGGCGG- CTT-3′, a modification of primer R9124 published previously; Zheng et al., 2003b) were employed. All primers had a concentration of 40 μM and were from Eurogentec. PCRs were performed in a 25 μl volume with 0.5 U Platinum Taq DNA polymerase (Invitrogen) per reaction. The equivalent of 1 μl of the first-round reaction mix was transferred to a new tube for the nested reaction. Diagnostic PCR conditions were 94 °C for 3 min, followed by 40 cycles of 94 °C for 30 s, 61 °C for 30 s and 72 °C for 45 s (30 s for the nested reaction), with a final incubation at 72 °C for 5 min. To generate fragments for sequencing, samples were incubated at 94 °C for 3 min, followed by 40 cycles of 94 °C for 1 min, 61 °C for 1 min and 72 °C for 1 min, with a final incubation at 72 °C for an additional 5 min. The RA27/3 vaccine strain (RUDIVAX; Aventis Pasteur MSD) served as a positive control. PCRs were performed in either a Mastercycler Gradient (Eppendorf) or a DNA Engine Opticon 2 System (Bio-Rad Laboratories). Amplification products were analysed in a 1.5% agarose gel stained with ethidium bromide, using 1× TAE as the electrophoresis running buffer.

Sequencing. PCR products were either purified directly using the Jet Quick PCR Purification Spin kit (Genomed) or, when multiple bands were visible, a gel-purification step was included (QIAquick Gel Extraction kit; Qiagen). Purified products were sequenced in both directions using a Big Dye Terminator v.3.1 Cycle Sequencing kit (Applied Biosystems) on a capillary sequencer (Model 3100 Avant; Applied Biosystems) using the PCR primers as sequencing primers. In case of nucleotide ambiguity, sequencing was repeated.

Data analysis. Sequences were analysed with the help of the SEQSCAPE v2.5 program (Applied Biosystems). Phylogenetic analysis using MEGA v3.1 (Kumar et al., 2004) was based on the entire E1 gene sequences comprising 1443 nt, as well as on the 739 nt corresponding to the minimum acceptable window defined by WHO (2005c). Reference sequences (WHO, 2005c) were included in each analysis. Methods employed were neighbour-joining, minimum evolution, maximum parsimony and UPGMA. Bootstrap values (Felsenstein, 1985) above 50 (500 replications) are shown on each phylogenetic tree. A total of 139 variable positions were identified in the nucleotide sequence over the complete E1 gene of the novel isolates from Belarus but only seven in the amino acid sequence. Of the variable nucleotides, 129/139 were found at position 3 of a codon and only five each at codon positions 1 and 2. Of the seven non-synonymous mutations, five occurred at positions 1 and 2 at codon position 2. The maximal observed distance between the Belarusian isolates in the complete E1 gene sequence was 6.17%, a percentage comparable to that between the most diverse reference sequences of clade 1 (6.03% between RVi/SLV/02[1C] and RVi/Shandong.CHN/02[1E]).

The new sequences were compared with 90 complete E1 gene sequences comprising the 22 reference strains, the provisional 1g strain RVi/UGA/20.01 and all sequences currently (April 2006) available on GenBank. Twenty-seven positions were identified where one or more of the Belarusian sequences showed a specific nucleotide not exhibited by any of the other 90 sequences, including six strains from Russia. Twenty-four of the 27 specific mutations (88.89%) occurred at codon position 3 and remained silent. The other three mutations, two of which appeared at codon position 1 and the third at codon position 2, led to changes in the amino acid sequence and were observed in individual isolates only (113, 121 and 006). Whilst some mutations were specific for certain groups of our isolates, there was none affecting all new sequences simultaneously.
Phylogenetic analysis

The Belarussian RUBV isolates clustered into three distinct phylogenetic lineages, irrespective of the method used for analysis and whether the whole E1 gene sequence was utilized or the sequence of the minimum acceptable window (data not shown). According to WHO, the phylogenetic analysis of RUBV sequences is considered valid if the accepted set of reference viruses falls into the accepted groups (WHO, 2005c). Therefore, our analysis performed with the UPGMA method was considered invalid. The phylogenetic tree obtained by the neighbour-joining algorithm is shown in Fig. 1.

Isolates 006, 072, 113 and 120 (group 1) could clearly be identified as genotype 1E strains, with the intra-group variation within the E1 gene ranging from 0.14 to 2.98 % (2 to 43 nt difference). Isolates 050, 087, 095 and 121 (group 2) were very similar to each other, with a maximal distance in the E1 gene sequence of 0.28 %. These sequences did not cluster with any of the official reference sequences, but instead clustered with a genotype 1g sequence from Uganda that was also included in the analysis (Fig. 1). The isolates in the third group (023, 040, 041, 042, 046 and 070) showed a maximal distance of 1.25 %, with sequences 040, 042 and 046 being identical. They branched off together with the 1g strain and the sequences in group 2 (Fig. 1).

When all 104 complete E1 gene sequences were included in the analysis, the Belarussian isolates of groups 2 and 3 also clustered with the strain from Uganda (data not shown). Isolates 006, 072, 113 and 120 (group 1) could clearly be identified as genotype 1E strains, with the intra-group variation within the E1 gene ranging from 0.14 to 2.98 % (2 to 43 nt difference). Isolates 050, 087, 095 and 121 (group 2) were very similar to each other, with a maximal distance in the E1 gene sequence of 0.28 %. These sequences did not cluster with any of the official reference sequences, but instead clustered with a genotype 1g sequence from Uganda that was also included in the analysis (Fig. 1). The isolates in the third group (023, 040, 041, 042, 046 and 070) showed a maximal distance of 1.25 %, with sequences 040, 042 and 046 being identical. They branched off together with the 1g strain and the sequences in group 2 (Fig. 1).

When all 104 complete E1 gene sequences were included in the analysis, the Belarussian isolates of groups 2 and 3 also clustered with the strain from Uganda (data not shown). Isolates from group 3 were more closely related to this 1g strain (maximal observed distance 2.84 %) than isolates belonging to group 2 (maximal observed distance 3.74 %). None of the new sequences branched off with any of the six full-length E1 gene sequences from Russia available on GenBank (data not shown).

Together with our isolates and the 1g strain from Uganda, there are currently (April 2006) 210 sequences for the complete diagnostic window available, among them two new sequences from Russia isolated in 2004 and 2005. Phylogenetic analysis revealed that our group 2 sequences clustered most closely with the Russian strain from 2004 (GenBank accession no. DQ454162), whereas two strains isolated in Germany in 1995 and 1998 (GenBank accession nos AF039133 and AY326342) were most closely related to the isolates belonging to our group 3 (Fig. 2).

DISCUSSION

Our study showed that RUBV strains of different genotypes co-circulated in the Minsk region in 2004/2005 and that the 14 isolates were more diverse than the 18 reference sequences (including the 1g strain) of clade 1. These reference strains come from nine different countries in Europe, America, Asia and Africa, whereas the new isolates were collected essentially within a single city in Belarus. There are only few reports of a sizeable number of RUBV strains
analysed from a single location. In a study from Italy, RUBV strains from Milan (GenBank accession nos AY161370, AY161371–73 and AY161379) exhibited an even higher maximal distance (10 %) within the minimum acceptable window sequence than that observed in Minsk. This is explained by strains belonging to different genotypes and clades (Zheng et al., 2003b). The variation observed in our study corresponds to multiple chains of transmission, which could be interpreted either as endemic co-circulation or as a reintroduction of different strains. RUBV is considered to be quite stable in its nucleotide sequence during cell culture (Frey & Abernathy, 1993; Frey et al., 1998; Hofmann et al., 2003). Nevertheless, the small variation comprising up to 4 nt observed among similar sequences within each group of the Belarussian samples may be explained by the up to six sequential passages, thus possibly overestimating the duration of the strain circulation. Therefore, the strains within each group do not seem diverse enough to be compatible with long-term endemic circulation, but rather to correspond to three independent introductions. However, the high number of cases during the last few years, mostly among people born before the introduction of the rubella vaccine in 1996, shows that there is a reservoir of susceptible individuals, which could sustain endemic transmission of the virus. One scenario could be that vaccination has considerably reduced the number of susceptible individuals, decreasing or eliminating the endemic virus and allowing new strains to be imported and established. As there is no genotype or sequence information available from the pre-vaccination era and/or the years before 2004, an unequivocal conclusion is not possible.

The phylogenetic clustering of the new isolates was independent of the different algorithms (except UPGMA) used on either the minimum acceptable window or the whole E1 gene sequence. The viruses investigated in this study segregated into three distinct phylogenetic lineages, all of which were supported by high bootstrap values. One group belonged to genotype 1E and the other two branched off with a sequence of genotype 1g, which is considered provisional because of its unclear relationship to 1B sequences. When all available sequences of the minimum acceptable window (n=210, including the 1g strain from Uganda and our isolates) were used for analysis, the closest relative of our group 2 sequences was the 2004 strain from Russia (GenBank accession no. DQ454162), which has been assigned to genotype 1g. However, the German strain isolated in 1995 (RVi/Stuttgart.DEU/95/1B/CRS, GenBank accession no. AF039133), which is one of the two strains most closely related to our group 3 sequences, has previously been attributed to genotype 1B. In the phylogenetic tree, there is a second group containing 1B sequences, indicating either that genotype 1B is very diverse and closely related to genotype 1g or that the sequences in the group close to the 1g strain might have been misattributed to genotype 1B. In order to clarify the phylogenetic relationship of the sequences in the ‘1g cluster’ (Fig. 2), we made pairwise comparisons of the closest strains that clearly belonged to different genotypes using MEGA v3.1 (Kumar et al., 2004). For this purpose, we selected one strain of each of the genotypes as well as four strains representing the different subgroups of the ‘1g cluster’ in Fig. 2 (Table 2). Considering that genotypes 1D/1E and 1B/1D are separated by a minimum of 14 and 16 nt, respectively, it should be considered whether at least subgroup 3 represents a putative new provisional genotype of clade 1. This subgroup differs by 17 nt from its closest relative in subgroup 1 (Table 2) and is supported by a high bootstrap value of 95. Subgroup 1, which includes the 1g strain from Uganda and is separated by at least 20 nt from the genotypes 1a–1F, might represent the true 1g genotype. The status of subgroup 2 should be evaluated again when more strains become available. Subgroup 4 is most closely related to 1B, despite a 13 nt distance. In this context, the assignment of the sequences with GenBank accession nos. AF039128 and AF039133 to genotype 1B should be reconsidered. The phylogenetic relationships within clade 1 are more complex than those of clade 2, where genotypes are very clearly separated from each other with at least 50 nt difference. If higher threshold values were applied to
Fig. 2. Phylogenetic tree showing the 14 new RUBV isolates (●) and two recent isolates from Russia (◆) together with all strains attributed to genotypes 1B and 1g and reference sequences for all genotypes. The analysis was done using the neighbour-joining method and sequences covering the minimum acceptable window.
Table 2. Pairwise observed distance (nt) between representatives of the different RUBV genotypes (1a–1g, 2A–2c) and four subgroups of the ‘1g cluster’ (sg1–sg4) calculated using MEGA v3.1 (Kumar et al., 2004)

<table>
<thead>
<tr>
<th>Strain name (GenBank accession no.)</th>
<th>Genotype</th>
<th>[1a]</th>
<th>[1B]</th>
<th>[1C]</th>
<th>[1D]</th>
<th>[1E]</th>
<th>[1F]</th>
<th>[lg]</th>
<th>[sg1]</th>
<th>[sg2]</th>
<th>[sg3]</th>
<th>[sg4]</th>
<th>[2A]</th>
<th>[2B]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matsuura Progenitor (AB003337)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RVi/ISR/75[1B] (AY968207)</td>
<td></td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-1 (L16232)</td>
<td></td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RVi/Vancouver.CAN/8 [1D] (AF039114)</td>
<td></td>
<td>15</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6496-ITA97 (AY161378)</td>
<td></td>
<td>22</td>
<td>20</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AH5/AH-CHN99 (AY326351)</td>
<td></td>
<td>15</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RVi/UGA/20.01[lg]</td>
<td></td>
<td>28</td>
<td>28</td>
<td>33</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4844-ITA93 (AY161361)</td>
<td></td>
<td>21</td>
<td>20</td>
<td>28</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I76/EV-ISR92 (AY326340)</td>
<td></td>
<td>29</td>
<td>26</td>
<td>34</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4008-ITA91 (AY161357)</td>
<td></td>
<td>30</td>
<td>25</td>
<td>31</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tomi London UK 86 (AF039122)</td>
<td></td>
<td>18</td>
<td>13</td>
<td>20</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRD2 vac Japan (AB003341)</td>
<td></td>
<td>58</td>
<td>57</td>
<td>56</td>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RVi/TelAviv.ISR/68[2B] (AY968219)</td>
<td></td>
<td>56</td>
<td>55</td>
<td>61</td>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4-RUS67 (AY247015)</td>
<td></td>
<td>45</td>
<td>45</td>
<td>52</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

NOTE ADDED IN PROOF

The subdivision of the genotype 1g cluster and the distance calculations presented in this paper are currently considered by WHO and other authorities as a basis for establishing two new provisional genotypes (1i and 1j), as well as upgrading two new isolates to provisional status (1i and 1j). The criteria for assigning strains to genotypes and subtypes should be clarified further. In particular, genotypes 1D and 1E are difficult to assign to any of the genotypes. As more strains become available, criteria for genotyping of RUBV should be refined.
ACKNOWLEDGEMENTS

We thank Emily S. Abernathy from CDC, Atlanta, USA, for providing the E1 gene sequence of the provisional Ig strain RVi/UGA/20.01. This work was supported by the Ministry of Foreign Affairs, Luxembourg.

REFERENCES


