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 increas-
ing importance (WHO, 2005c). Recently, a standardized

The GenBank/EMBL/DDBJ accession numbers for the sequences
reported in this paper are AM258944–AM258957.
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 Enzymognost immunoassay). Specimens were inoculated onto Vero cell cultures and passed 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., 1996) were used as well 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 RVi (5’-TCTTCTATRC-AGCAACRGGTG-3’) and R9124a (5’-GATCCGACCTTGCGG-CCT-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 v3.1 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. In addition, a maximum-likelihood analysis was carried out with all complete E1 gene sequences using PHYLIP v3.6a2 (Felsenstein, 1993).

RESULTS

E1 gene sequence analysis

The 14 RT-PCR-positive samples were obtained from patients of 11–29 years of age from clinical material collected within 6 days after the onset of a rash. Nine viruses were isolated from nasopharyngeal swabs and five from urine (Table 1). Strains were named following the WHO nomenclature for RUBV (WHO, 2005c), but for ease of reference, the original ID numbers (Table 1) are used throughout this manuscript. 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).

Table 1. Characteristics of the novel rubella virus isolates from Belarus

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Age of patient (years)</th>
<th>Place of sample collection</th>
<th>Origin of patient</th>
<th>Genotype</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>006</td>
<td>18</td>
<td>NPS, Infectious Disease Hospital, Minsk</td>
<td>Minsk city</td>
<td>IE</td>
<td>AM288954</td>
</tr>
<tr>
<td>023</td>
<td>29</td>
<td>NPS, Infectious Disease Hospital, Minsk</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288955</td>
</tr>
<tr>
<td>041</td>
<td>12</td>
<td>NPS, Outpatient Clinic No 26</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288956</td>
</tr>
<tr>
<td>042</td>
<td>14</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288957</td>
</tr>
<tr>
<td>046</td>
<td>11</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288958</td>
</tr>
<tr>
<td>070</td>
<td>19</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288959</td>
</tr>
<tr>
<td>087</td>
<td>20</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288960</td>
</tr>
<tr>
<td>095</td>
<td>13</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288961</td>
</tr>
<tr>
<td>113</td>
<td>12</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288962</td>
</tr>
<tr>
<td>120</td>
<td>13</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288963</td>
</tr>
<tr>
<td>121</td>
<td>18</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288964</td>
</tr>
<tr>
<td>006</td>
<td>18</td>
<td>NPS, Infectious Disease Hospital, Minsk</td>
<td>Minsk city</td>
<td>IE</td>
<td>AM288954</td>
</tr>
<tr>
<td>023</td>
<td>29</td>
<td>NPS, Infectious Disease Hospital, Minsk</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288955</td>
</tr>
<tr>
<td>041</td>
<td>12</td>
<td>NPS, Outpatient Clinic No 26</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288956</td>
</tr>
<tr>
<td>042</td>
<td>14</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288957</td>
</tr>
<tr>
<td>046</td>
<td>11</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288958</td>
</tr>
<tr>
<td>070</td>
<td>19</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288959</td>
</tr>
<tr>
<td>087</td>
<td>20</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288960</td>
</tr>
<tr>
<td>095</td>
<td>13</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288961</td>
</tr>
<tr>
<td>113</td>
<td>12</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288962</td>
</tr>
<tr>
<td>120</td>
<td>13</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288963</td>
</tr>
<tr>
<td>121</td>
<td>18</td>
<td>NPS, Urine</td>
<td>Minsk city</td>
<td>IE, subgroup 2</td>
<td>AM288964</td>
</tr>
</tbody>
</table>

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 (\( \times 10 \% \)) 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. 1.** Phylogenetic analysis of the complete E1 gene of the 14 new RUBV isolates (●) using the neighbour-joining algorithm and including all accepted reference strains as well as sequences of the two provisional genotypes 1g and 2c.
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.
Genotypes 1E and 1g have been reported from different European countries (WHO, 2005b, c), but no complete E1 gene sequences are available from 2004 or 2005 in GenBank. None of the viruses from Minsk clustered with any of the full-length E1 gene sequences of Russian origin from the 1960s, 1970s and 1997 (data not shown), although sequences from the two countries shared some unique nucleotides exhibited by no other strain. Two new isolates from Western Siberia from 2004 and 2005 for which the minimum acceptable window sequence is available clustered very closely with our group 2 sequences (Fig. 2). The two Russian isolates differed by 11 nt within the minimum acceptable window sequence, whereas our strains exhibited a maximal distance of 4 nt. Interestingly, the variation between the Russian isolate from 2004 and our isolates increased with time between sample collections (i.e. isolate 050 collected 03.05: 4 nt; 087 and 095 collected 06.05: 6 nt; isolate 121 collected 07.05: 7 nt). Therefore, these Minsk isolates may have evolved from a strain from Russia.

Nearly 95% (132/139) of the variable positions among the new isolates were silent at the amino acid level, which is clearly a higher percentage than in earlier studies from 1993 (71%) (Frey & Abernathy, 1993), 1997 (78.8%) (Katow et al., 1997) and 1998 (83%) (Frey et al., 1998). The very low variation observed is in line with previous reports (Donadio et al., 2003; Frey et al., 1998; Zheng et al., 2003a) and seems to indicate that RUBV is not under high selective pressure. The different percentages may be explained partially by the numbers and domains of sequences analysed, the time and place of sample collection and whether sequences of only one or both clades were included in the analyses. In the older studies, mutations may be attributed partially to a poorer sequencing performance.

In conclusion, our findings show that three distinct strains with limited variability were 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.

**NOTE ADDED IN PROOF**

The subdivision of the genotype 1g cluster and the distance calculations presented in this paper are currently considered by WHO as a basis for establishing two new provisional genotypes (1i and 1j), corresponding to our subgroups 3 and 4. Subgroup 1 representing genotype 1g is upgraded from a provisional to a recognized genotype (1G). Subgroup 2 is currently considered an outlier, as only one sequence is available. AF039128 and AF039133 are reassigned to genotype 1G as suggested in this manuscript.
ACKNOWLEDGEMENTS

We thank Emily S. Abernathy from CDC, Atlanta, USA, for providing the E1 gene sequence of the provisional Ig strain RV1/UGA/20.01. This work was supported by the Ministry of Foreign Affairs, Luxembourg.

REFERENCES


