Reinis Balmaks

HUMAN RESPIRATORY SYNCYTIAL VIRUS CAUSED LOWER RESPIRATORY TRACT INFECTION: CLINICAL AND MOLECULAR CHARACTERIZATION IN HOSPITALIZED CHILDREN IN LATVIA

Summary of the Doctoral Thesis

Speciality – Pediatrics

Riga, 2014

The study was conducted at:

Department of Pediatrics, Rīga Stradiņš University
Latvian Biomedical Research and Study Centre

Supervisors:

Dr. habil. med., Professor **Dace Gardovska**,
Department of Pediatrics, Rīga Stradiņš University

Dr. biol. **Andris Kazāks**,
Latvian Biomedical Research and Study Centre

Official reviewers:

Dr. biol., Associate Professor **Edvīns Miklašēvics**,
RSU Institute of Oncology

Dr. med., Associate Professor **Enoks Biķis**,
University of Latvia, Faculty of Medicine, Department of Pediatrics

Ph. D., Assistant Professor **Matti Warris**,
University of Turku, Department of Virology

The Doctoral Thesis will be defended on 11th of March, 2014 at 15.00 during Rīga Stradiņš University Medical Degree Committee open meeting in Lecture theatre Hippocrates, Rīga Stradiņš University, 16 Dzirciema Street, Riga

The Doctoral Thesis is available at RSU library and RSU website: www.rsu.lv

The thesis was co-funded by the ESF project "Support for Doctoral Students in Mastering the Study Programme and Acquisition of a Scientific Degree in Rīga Stradiņš University", agreement No. 2009/0147/1DP/1.1.2.1.2/09/IPIA/VIAA/009

Secretary of the Degree Committee:

Dr. med., Associate Professor **Angelika Krūmiņa**

CONTENTS

ABBREVIATIONS

Abbreviation	Explanation
95% HPD	95% highest probability density interval
aa	Amino acid(s)
bp	Base pairs
BSP	Bayesian skyline plot
cDNA	Complementary DNA
CDPC	The Centre for Disease Prevention and Control of Latvia
CSS	Clinical Severity Score
dN	Non-synonymous substitution
dN/dS	Non-synonymous to synonymous substitution rate ratio
DNA	Deoxyribonucleic acid
dS	Synonymous substitution
FEL	Fixed effects likelihood
Fr	French catheter gauge system unit
HRSV	Human respiratory syncytial virus
HRSV-A	Group A human respiratory syncytial virus
HRSV-B	Group B human respiratory syncytial virus
HVR	Hypervariable region
IFA	Immunofluorescence assay
IFEL	Internal fixed effects likelihood
IQR	Interquartile range
kB	Kilobase
LRTI	Lower respiratory tract infection
MCC	Maximum clade credibility tree
MCMC	Markov chain Monte Carlo method
MEME	Mixed effects model of evolution
ML	Maximum likelihood
Ne	Effective population size
NJ	*Neighbor-joining* ātrās klasterizācijas algoritms
NPA	Nasopharyngeal aspirate
nt	Nucleotide(s)
PCR	Polymerase chain reaction
pMol	Picomole
RDAI	Respiratory Distress Assessment Instrument
REL	Random effects likelihood
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
SD	Standard deviation
SIRS	Systemic inflammatory response syndrome
SLAC	Single likelihood ancestor counting

ssRNA	Single stranded RNA
tMRCA	Time of most recent common ancestor
π	Nucleotide sequence polymorphisms
τ	Generation length in years

1. INTRODUCTION

Human respiratory syncytial virus (HRSV) has been recognized as the most important cause of lower respiratory tract infections (LRTI) in children. It is estimated that HRSV causes 33.8 million new infections each year in children less than five years old [Nair et al., 2010]. About 10% of the infected children experience severe LRTI requiring admission to hospital and 234,000 annual deaths have been estimated globally in this age group [Lozano et al., 2012]. In developed countries HRSV mortality is rare, but the associated costs pose a great burden to the health care budgets. During the first year of life up to 70% of all children get infected, but by the age of two even more than 90% [Glezen et al., 1986; Simões and Carbonell-Estrany, 2003]. HRSV is detected in 45% hospitalized children with LRTI that are less than two years old [Simões and Carbonell-Estrany, 2003]. Children with preterm birth, cyanotic or complicated congenital heart disease and chronic lung disease of prematurity are at increased risk of severe disease and mortality [Shay et al., 2001]. A substantive disease burden is also associated with HRSV in vulnerable adult and elderly populations, where mortality is even higher than in children [Falsey et al., 2005]. Although HRSV has a single serotype, infection does not induce protective immunity, therefore can occur multiple times throughout one's life and even within a single season. [Henderson et al., 1979; Simões and Carbonell-Estrany, 2003]. Epidemiologic data available from Latvia are insufficient. HRSV has been detected in 15–47% of samples positive for respiratory viruses in Infectology Center of Latvia [Nikiforova et al., 2011], however incidence, mortality and seasonality data are not known.

Prevention and treatment of HRSV infections pose several challenges. Respiratory support and hydration remain the cornerstone of the therapy [American Academy of Pediatrics, Subcommittee on Diagnosis and

Management of Bronchiolitis, 2006]; other interventions (including antivirals) have shown limited or no effect in randomized controlled trials. The decrease in HRSV mortality in developed countries is largely related to overall improvement of pediatric care. Although passive immunoprophylaxis with monoclonal antibody palivizumab is safe and effective, unfavorable cost-effectiveness ratio prevents its use in the general population [American Academy of Pediatrics Committee on Infectious Diseases, 2009]. There is an obvious medical need and economic rationale [Meijboom et al., 2012] for vaccine development, however no licensed product is available yet [Anderson et al., 2013]. One of the main hardships in vaccine development is the high variability of the virus.

HRSV is a member of the *Pneumovirus* genus that is classified within the subfamily *Pneumovirinae* of the family *Paramyxoviridae*. Accordingly, it is a cytoplasmic, enveloped virus with linear, negative sense, ssRNA genome [Wang et al., 2012]. The viral RNA of HRSV is approximately 15.2 kB in size and encodes 11 viral proteins [Collins and Crowe, 2007]. Two surface glycoproteins, G and F, are antigenically significant because they induce neutralizing antibody responses. Based on the reaction with monoclonal antibodies, HRSV strains are separated into two major groups, HRSV-A and HRSV-B [Mufson et al., 1985], which are genetically divergent viruses that have evolved separately [Zlateva et al., 2005]. Viruses from both groups co-circulate simultaneously during epidemic seasons with alternating predominance. Typically, there is a cyclic pattern whereby several predominant HRSV-A seasons are followed by a single HRSV-B dominant season [Venter et al., 2001; Zlateva et al., 2007]. HRSV viruses also vary considerably within the groups, with several distinct genotypes in each group accounting for clusters of circulating strains. Several genotypes co-circulate in the same community and are replaced by new ones in successive seasons [Peret et al., 2000; Venter et al.,

2001; Zlateva et al., 2007]. The most extensive differences are found in the gene encoding G protein, and the genotype classification based on partial sequencing of this gene is now widely used in molecular epidemiologic studies of HRSV. It has been confirmed by genome wide analysis that genotyping based G gene variability represents overall virus variability [Rebuffo-Scheer et al., 2011; Tan et al., 2012].

The G protein is a transmembrane glycoprotein and its large ectodomain consists of two mucin-like hypervariable regions (HVR1 & 2) [Johnson et al., 1987]. Sequence diversity in the HVRs is among the most extensive found in human viruses. This variability represents both immune-driven selection and structural plasticity of the protein [Collins and Melero, 2011]. The mucin-like regions are heavily glycosylated with *O*- and *N*-linked oligosaccharide chains. The number and positions of glycosylation sites are poorly conserved among the strains, also contributing to their antigenetic differences [Johnson et al., 1987; Martinez et al., 1997; Palomo et al., 1991].

Local HRSV molecular surveillance is important for virologic characterization – prediction of novel virulence factors and future outbreak strains. By combining regional and global data, it is possible to reconstruct the population size of the virus and the geographic spread. From public health point of view, the identification of the main transmission routes could lead to implementation of efficient prevention strategies. The molecular epidemiology and circulation patterns of HRSV in Latvia have not been studied before.

1.1. Aim of the Study

To characterize the clinical manifestations of HRSV caused LRTI and molecular epidemiology of strains circulating in tertiary level pediatric hospital in Latvia over three consecutive seasons.

1.2. Objectives

1. To develop polymerase chain reaction (PCR) HRSV diagnostic and group differentiation test.
2. To determine the proportion of HRSV infections in young children hospitalized with LRTI.
3. To elucidate HRSV seasonality over three years;
4. To detect a possible link between HRSV group or genotype with LRTI severity.
5. To analyze phylogenetics and variability of HRSV strains.
6. By using the genealogy of HRSV G gene, to reconstruct the population size and geographic spread of the virus over time.

1.3. Hypotheses

1. Molecular epidemiology of HRSV in Latvia is not significantly different from other countries, but unique local strains are possible.
2. There is a correlation between HRSV molecular characteristics and disease severity.

3. By using molecular clock approach on time-stamped HRSV sequences with location trait, it is possible to reconstruct its global spread.

1.4. Scientific Novelty

This is the first HRSV molecular epidemiologic study in Latvia and its findings are of local and global scientific importance. This work emphasizes the need for precise HRSV seasonality data in Latvia. Several unique HRSV strains were discovered and their sequences were deposited in GenBank database. This is the first study that estimates global dissemination hypothesis and population size dynamics of genotype ON1. The data presented here can be used to optimize the timing of immunoprophylaxis in high risk infants in Latvia and development of public health interventions.

All the data presented here are the results of author's own research under the supervision of the mentors. The doctoral thesis "Human respiratory syncytial virus caused lower respiratory tract infection: clinical and molecular characterization in hospitalized children in Latvia" was presented at the extended faculty meeting in Department of Pediatrics, Riga Stradiņš University on September 2, 2013.

2. METHODS

2.1. Study Population

This prospective cohort study was conducted at Children's Clinical University Hospital in Riga from July 1, 2009 to June 30, 2012. The study protocol was approved by the Ethics Committee of Riga Stradiņš University, and written informed consent was obtained from the parents of all the participating children. Inclusion criteria were: (1) age: 2 to 24 months and (2) LRTI according to the World Health Organization case definition [Wright and Cutts, 2000]. Exclusion criteria were: (1) chronic central nervous system or cardiopulmonary disease and (2) symptoms more than ten days.

Clinical data were standardized by using different scales: Respiratory Distress Assessment Instrument (RDAI; grades respiratory distress from 0 to 17) [Lowell et al., 1987], Clinical Severity Score (CSS; disease severity from 0 to 6) [Martinello et al., 2002], and systemic inflammatory response syndrome (SIRS) criteria [Goldstein et al., 2005].

2.2. Clinical Samples

Nasopharyngeal aspirate (NPA) was collected from each of the enrolled patient. NPAs were frozen and stored at -70°C. Total RNA was extracted from 140 µl of NPA specimen with QIAamp Viral RNA Mini Kit (QIAGEN), according to the manufacturer's instructions.

2.3. HRSV Detection, Group Differentiation, and Sequencing

HRSV specific cDNA was synthesized by RevertAid™ Premium reverse transcriptase (Fermentas) with 12.5 µl of extracted RNA and 20 pmol of primer F164_Rv (Figure 3.1) [Sullender et al., 1993] added to the reaction mixture, according to the manufacturer's instructions.

First, HRSV was detected by PCR-amplification of a conserved fragment in the non-coding sequence between the P and M genes (Figure 3.1). Second, HRSV-positive samples were differentiated into groups A and B by group-specific PCR targeting the HVR2 segment of the G gene. In this reaction reverse primer F_Rv was cross-reactive, while forward primers were group-specific: Ga_Fw for group A and Gb_Fw for group B (Figure 3.1). The amplified products were analyzed by electrophoresis in ethidium bromide stained 1% agarose gels.

The amplified fragments were purified from agarose gel and sequenced by the same set of primers as in the group differentiation reaction. 336-nucleotide (nt)-long HRSV-A sequences (corresponding to codon positions 187 to 299 of reference strain A2) and 516-nt-long HRSV-B sequences (corresponding to codon positions 140 to 293 of reference strain B1) were retrieved. The unique sequences were deposited in GenBank database under accession numbers JF979145–57 and KF030137–85.

2.4. Phylogenetic and Adaptive Evolutionary Analysis

The alignments of nt and deduced amino acid (aa) sequences were prepared using ClustalW2 algorithm [Larkin et al., 2007]. The phylogenetic trees were constructed using the neighbor-joining (NJ) algorithm [Saitou and

Nei, 1987] and genetic distances (the number of nt and aa substitutions per site from averaging over all sequence pairs) were calculated under the best-fit substitution models in MEGA v5.1 software [Tamura et al., 2001]. Bootstrapping with 1,000 replicates was performed for each analysis to evaluate confidence estimates. The *O*-glycosylation sites were determined using NetOGlyc 3.1 Server neural network predictions [Julenius et al., 2005], while acceptance of the *N*-linked oligosaccharides using NetNGlyc 1.0 Server [Gupta et al., 2004].

Positively selected sites were identified by estimating site-specific non-synonymous (*dN*) to synonymous (*dS*) substitution rate ratios (*dN/dS*) with five different algorithms available on the Datamokey web server [Delport et al., 2010]: SLAC, FEL, IFEL, REL, and MEME. The site was considered under positive selection (*dN/dS* >1) when two or more methods reached agreement with statistical significance ($p < 0.1$ or Bayes factor >20) as recommended by [Kosakovsky Pond et al., 2005]. The mean *dN/dS* ratio was estimated using the SLAC algorithm.

2.5. Evolutionary Rate, Population Dynamics and Phylogeographic Analysis

Nucleotide substitution rate per site, the time of most recent common ancestor (tMRCA), changes in the population size and discrete phylogeographic analysis were estimated from time-stamped sequences with location trait using Bayesian Markov chain Monte Carlo (MCMC) method in BEAST v2.0.2 software package [Drummond et al., 2012]. The results obtained from MCMC analysis were assessed using Tracer v1.5 and the maximum clade credibility (MCC) tree was inferred using TreeAnnotator v2.0.2. The MCC tree

was visualized using FigTree v1.4.0 software. Marginal posterior distributions for data were summarized using medians and 95% highest posterior density intervals (HPDs).

2.6. Statistical Analysis

The statistical analysis was performed using SPSS 17.0 and *Microsoft Excel* programs. Interval and ordinal scale data were compared using Mann-Whitney test, while nominal scale data were compared using Pearson χ^2 or Fisher exact test. Statistical significance was defined as $p < 0.05$.

3. RESULTS

3.1. Study Cohort

207 patients were enrolled in the study (67–71 per season) with the median age 8 (interquartile range; IQR: 4–14) months. 36% of the patients were less than 6 months old and 68% were less than 12 months old. A nomenclature for patient samples was adapted where LV indicates Latvia followed by the laboratory log number and the isolation year.

3.2. HRSV Detection and Group Differentiation

In the primary screening, the amplified 509 bp PCR fragment was well-detectable on agarose gels (Figure 3.1). HRSV specific RNA was detected in 88 (42.5%) of 207 tested samples. Group-specific primers were designed for discrimination of HRSV-A and B amplification products by their length (Figure 3.1). By this means, of 88 HRSV-positive samples 53 (60.2%) and 35 (39.8%) had amplified fragment lengths specific for groups A and B, respectively. There were no cases of co-infection with the two groups.

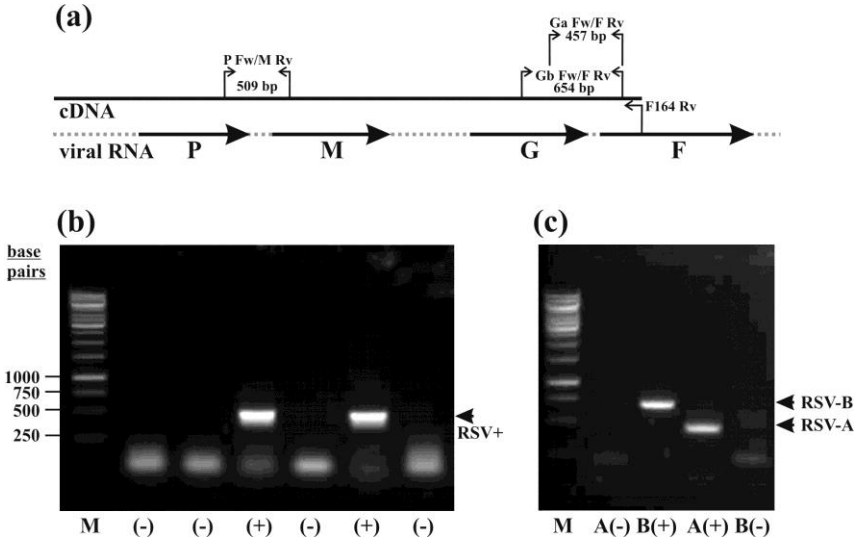

Figure 3.1. Molecular Diagnosis of HRSV

Schematic presentation of RSV genome fragment and general strategy of RT-PCR (a). Agarose gel electrophoresis illustrating primary HRSV screening with primers P_Fw/M_Rv (b) and genotyping of HRSV-positive samples (c). M, DNA molecular mass marker (Fermentas); (+) positive reaction; (–) negative reaction.

In addition to NPA samples for RT-PCR analysis, 110 patients had also off-study nasal swab samples ordered by attending physician for direct immuno-fluorescent assay (IFA). Comparison of RT-PCR results to IFA as a "gold standard" generated a sensitivity, a specificity, a positive predictive, and a negative predictive value of 95.2, 91.2, 87, and 97%, respectively, for RT-PCR.

3.3. HRSV Seasonality

The proportion of HRSV-positive samples in each season ranged from 33% to 57%, but during the peak activity could reach even 90% (Figure 3.2).

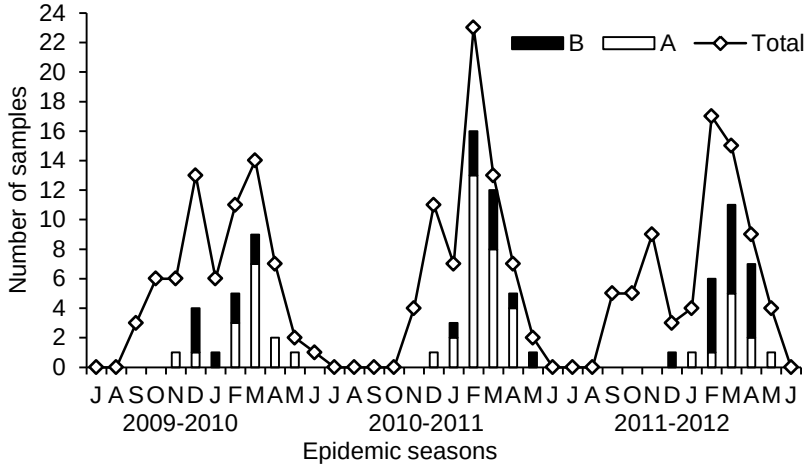

Figure 3.2. Distribution of HRSV throughout Seasons
Number of total samples tested (curve) and number of HRSV-A (white bars) and B (black bars) positive samples per month. Months are indicated by the initial letter starting from July, 2009.

The onset of HRSV season was defined as the first of two consecutive weeks with ≥10% samples positive for HRSV and the offset was defined as the last of two consecutive weeks with ≥10% samples positive for HRSV. Analyzing the summary data over three years, HRSV season was estimated from week 51 to 19. During this period HRSV caused 57% (86/151) of LRTIs, 92.9% (26/28) of bronchiolitis, and 51.4% (18/35) of pneumonia. Outside this period HRSV caused only 3.6% (2/56) of LRTIs.

3.4. Clinical Comparisons

Clinical comparison of HRSV and non-HRSV LRTIs showed several distinctions between the two groups (Table 3.1). Patients infected with HRSV were: (i) significantly younger, (ii) more likely to have atopy in the family history, but (iii) less likely to have SIRS. Discharge diagnoses were different,

as well – HRSV more frequently caused bronchiolitis, but was not detected in any of the patients with croup. However, HRSV-caused LRTI was not different in clinical severity measured as requirement for oxygen, RDAI, CSS, and length of stay in the hospital. One (0.5%) patient not infected with HRSV died during the study period.

Table 3.1

Clinical Comparison of HRSV and non-HRSV Infections

Characteristic	HRSV	Non-HRSV	P value
Number of patients	88	119	
Girls, no. (%)	38 (43.2)	42 (35.3)	0.249
Age in months, median (IQR)	6 (3–9.25)	10 (6–16.5)	**<0.001**
Weight-for-age, *z*-value median (IQR)	0.69 (-0.17–1.24)	0.84 (0.04–1.36)	0.252
Passive smoking, no. (%)	59 (67)	68 (57.1)	0.148
Crowdedness index, median[a] (IQR)	1.7 (1.3–2)	1.7 (1.3–2)	0.424
Atopy in family history, no. (%)	50 (56,8)	51 (42,9)	**0,047**
Systemic inflammatory response syndrome (SIRS), no. (%)	17 (19.3)	79 (66.4)	**0.023**
Oxygen requirement, no. (%)	27 (30.7)	46 (38.7)	0.235
Respiratory Distress Assessment Instrument, median (IQR)	6 (5–8)	6 (4–8)	0.839
Clinical Severity Score, mode (range)	1 (1–6)	1 (0–5)	0.21
Treatment with antibiotics, no. (%)	31 (35.2)	45 (37.8)	0.703
Length of stay in the hospital, median (IQR)	4 (3–6)	4 (2–6)	0.651
Diagnosis			**<0.001**
Bronchitis	40 (45.5)	55 (46.2)	0.913
Bronchiolitis	27 (30.7)	13 (10.9)	**0.001**
Pneumonia	19 (21.6)	33 (27.7)	0.398
Asthma	2 (2.3)	7 (5.9)	0.306
Croup	0	11 (9,2)	**0.002**

[a] number of family members/number of rooms in the housing ratio. P values <0.05 are shown in bold.

58% (n=51) of patients with HRSV infection were less than 6 months old compared to 29.4% (n=35) of patients not infected with HRSV (p<0.001). 79.5% (n=80) of patients with HRSV infections were less than 12 months old compared to 64.7% (n=64) of patients not infected with HRSV (p=0.03).

No statistically significant differences were found between the HRSV groups in baseline characteristics of the patients or severity of the disease (Table 3.2).

Table 3.2

Clinical Comparison of HRSV-A and B Infections

Characteristic	HRSV-A	HRSV-B	P value
Number of patients	53	35	
Girls, no. (%)	19 (35.8)	19 (54.3)	0.087
Age in months, median (IQR)	6 (3–9)	6 (3.75–12)	0.518
Weight-for-age, *z*-value median (IQR)	0.53 (-0.2–1.1)	0.91 (-0.01–1.29)	0.279
Passive smoking, no. (%)	37 (69.8)	22 (62.9)	0.497
Crowdedness index, median (IQR)	1.7 (1.3–2.5)	1.6 (1.4–2)	0.552
Atopy in family history, no. (%)	23 (43.4)	13 (37.1)	0.633
Systemic inflammatory response syndrome (SIRS), no. (%)	12 (22.6)	5 (14.3)	0.331
Oxygen requirement, no. (%)	17 (32.1)	10 (28.6)	0.727
Respiratory Distress Assessment Instrument, median (IQR)	6 (4–8)	7 (5.5–8)	0.928
Clinical Severity Score, mode (range)	1 (1–5)	1 (1–6)	0.303
Treatment with antibiotics, no. (%)	20 (37.7)	11 (31.4)	0.544
Length of stay in the hospital, median (IQR)	4 (3–6)	3 (2–4.75)	0.163
Diagnosis			0.316
Bronchitis	26 (49.1)	14 (40)	0.538
Bronchiolitis	18 (34)	9 (25.7)	0.559
Pneumonia	8 (15.1)	11 (31.4)	0.119
Asthma	1 (1.9)	1 (2.9)	1

3.5. Phylogenetic Analysis

By comparing nt sequences, 53 HRSV-A and 35 HRSV-B isolates grouped into 29 and 23 different strains, respectively. Representative gene sequences of strains retrieved in this study were aligned and included in

phylogenetic analysis together with sequences of reference strains obtained from GenBank representing 11 HRSV-A genotypes (GA1–7 [Peret et al., 2000], SAA1 [Venter et al., 2001], NA1 and 2 [Shobugawa et al., 2009], and ON1 [Eshaghi et al., 2012]) and 13 HRSV-B genotypes (GB1–4 [Peret et al., 2000], SAB1–3 [Venter et al., 2001], BA-I–VI [Trento et al., 2006]).

Two HRSV-A genotypes were identified, with 28 strains (51 isolates) clustering in genotype NA1 (Figure 3.3). One strain representing two identical isolates from March, 2012 clustered in the recently described NA1-related genotype ON1. All HRSV-B strains had a 60-nt duplication in the G gene, characteristic of the BA genotypes, and clustered within the clade BA-IV (Figure 3.4).

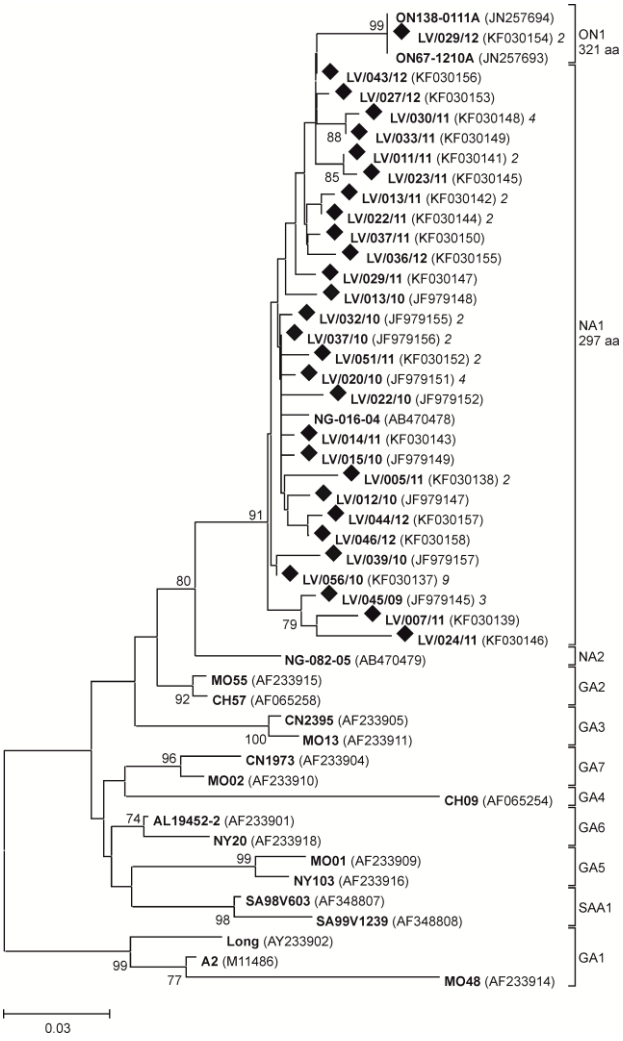

Figure 3.3. Phylogenetic Tree for HRSV-A

Latvian strains are marked by (◆). Taxa description include: strain name in bold, GenBank accession number in brackets and number of identical isolates in italics. Bootstrap values of ≥ 70 are shown at nodes; the bar denotes nucleotide substitutions per site.

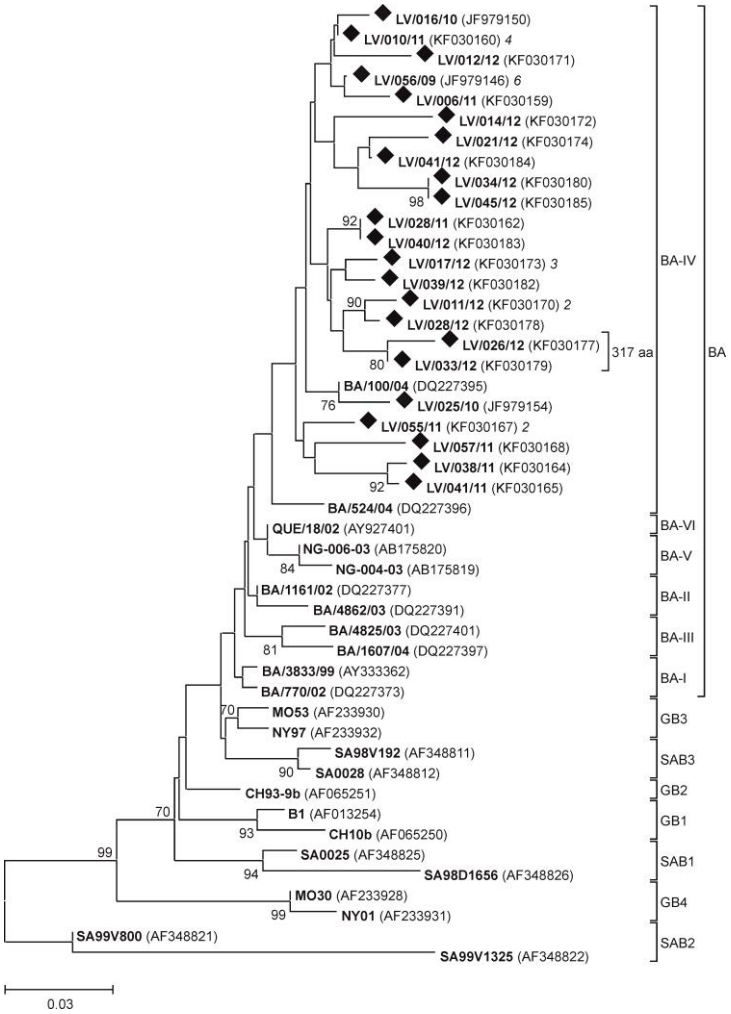

Figure 3.4. Phylogenetic Tree for HRSV-B

Description as in Figure 3.3. Strains with 317-aa-long G protein are indicated separately; other Latvian HRSV-B strains had 310-aa-long G protein.

3.6. Molecular Analysis of Genotype NA1

The HVR2 sequences of the G gene of Latvian strains of genotype NA1 were closely related, with 2.1% (± 0.4) and 4.1% (± 0.9) divergence at nt and aa levels, respectively. All HRSV-A strains had a premature termination codon (TGA) at position 298 and preserved the original termination codon (TAG) at position 299 (Figure 3.5). These strains also had the following G protein mutations: *P*226→*L*, *E*233→*K*, *L*258→*H/Y*, *M*262→*E/R*, *F*265→*L*, *S*269→*T*, *S*280→*Y*, *P*289→*S*, *S*290→*P/L*, *P*292→*S*, *P*293→*S*, *P*296→*T*, and *R*297→*K*. Other common mutations (detected in >50% of strains) included *L*208→*I*, *S*222→*P*, *N*237→*D*, *I*244→*R*, *N*273→*Y/K*, *P*274→*L*, and *P*286→*L*.

Using NetOGlyc software, 40 serine and threonine residues of the consensus sequence of Latvian strains were predicted to be potentially *O*-glycosylated (range 36 to 43 among individual strains). In the same sequence strain A2 is predicted to have 39 *O*-glycosylation sites, with 36 of them being in common with the Latvian consensus sequence. There are two potential *N*-glycosylation motifs in the HVR2 segment of strain A2 (N237 and N251), and two additional motifs were identified among Latvian strains (N273 and N294). NetNGlyc software predicted only N251 and N294 glycosylation of the Latvian consensus sequence. Among individual strains, *N*-glycosylation sites ranged from 0 to 3. None of the sites was conserved among all isolates.

In total, 28 sites showed elevated rates of non-synonymous substitutions (*dN*) and the mean non-synonymous/synonymous substitution rate ratio (*dN/dS*) was 1.12, suggesting presence of selection pressure. However, only two positively selected sites, M262 and N273, were identified based on the consensus of two or more of the methods.

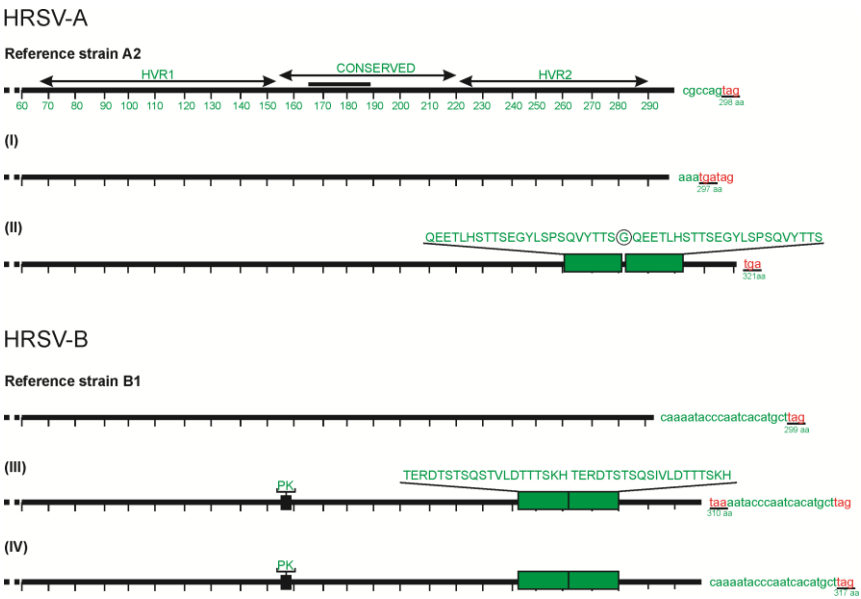

Figure 3.5. HRSV G Protein Ectodomain Variability among Latvian Isolates
The linear presentations are given according to the reference strains A2 and B1. I, premature stop codon (underlined nucleotides) terminates 297-aa-long protein in all Latvian HRSV-A isolates. II, two HRSV-A isolates contained 72-nt duplication (green box) in the second hypervariable region. III, 6-nt deletion (■), 60-nt duplication, and premature termination codon resulted in 310-aa-long protein that was characteristic for the most Latvian HRSV-B isolates. IV, two HRSV-B isolates had retained the original termination codon of strain B1, these proteins were 317-aa-long. Schematic diagrams and symbols are used as in [Zlateva et al., 2005].

3.7. Molecular Analysis of Genotype BA-IV

Latvian strains of genotype BA-IV were more diverse than NA1 strains, with a calculated mean distances of 3.3% (± 0.4) and 4.2 % (± 0.7) at nt and aa levels, respectively. In addition to the characteristic 20-aa in-frame duplication in the HVR2 segment, all Latvian BA-IV strains had deletion P159_K160 (Figure 3.5). Two alternative termination codon positions were observed either at position 293 (TAA) or at the original position 300 (TAG) of the reference

strain B1 (299 aa long). Thus, the Latvian HRSV-B viruses produced either a 310- or 317-aa-long G protein (Figure 3.5). All Latvian BA-IV strains had *K*218→*T*, *L*223→*P*, and *S*247→*P* mutations. In addition, the following mutations were common (>50%): *L*219→*P*, *T*270→*I/F*, *V*271→*A*, *H*287→*Y*, and *Q*313→*Stop*.

NetOGlyc software predicted 44 *O*-glycosylation sites in the consensus sequence of Latvian strains (range 36–50) and 41 in reference strain BA/3833/99 (originally described strain with 60-nt duplication). Only 16 of the predicted *O*-glycosylation sites were in common between consensus sequences of HRSV-A and B. There are two potential *N*-glycosylation motifs in the HVR2 segment of strain BA/3833/99 (N296 and N310). N296 was conserved and predicted to be glycosylated in all strains, while N310 was predicted to glycosylated only in 317-aa-long sequences. Strain LV/057/11 had three additional motifs (N212, N230 and N253), two of which (N212 and N230) were predicted to be *N*-glycosylated.

Thirty-two sites showed elevated rates of *dN* and the overall *dN/dS* ratio was 0.45. Two positively selected sites, L219 and T270, were found by two-method agreement.

3.8. Phylodynamics and Phylogeography of Genotype ON1

Two identical HRSV-A isolates (LV/029/12) belonged to the recently described genotype ON1 and were identical to the originally isolated sequence from Canada. Genotype ON1 is characterized by 72-nt duplication in the HVR2 that leads to indel mutation E284G and duplication of Q261_S283, elongating G protein by 24 aa [Eshaghi et al., 2012]. Because of the premature termination codon like in NA1 strains the resulting protein length is 321 aa.

Since the discovery in Canada in 2011 this novel genotype has been isolated in at least nine different countries in North America, Europe, Africa, Asia and Oceania with the total of 42 sequences available in the GenBank, 16 of these sequences were identical to the strain ON67-1210A [Cui et al., 2013; Eshaghi et al., 2012; Khor et al., 2013; Lee et al., 2012; Prifert et al., 2013; Valley-Omar et al., 2013].

The evolutionary rate of the HVR2 was estimated to be 7.92×10^{-3} nt substitutions/site/year (highest probability density; HPD 95%: 2.97×10^{-3} – 1.28×10^{-2}) and the tMRCA was estimated August, 2010 (HPD 95%: September, 2009 – January, 2011).

Discrete phylogeographic analysis suggests spread from Canada to Italy and Germany, initially, and subsequent global spread to the rest of the world (posterior value >0.7). Also, dissemination from Germany to Japan and re-entry in Germany from South Korea are well supported. However, the migration of this virus to Latvia through Malaysia has poor support (posterior value 0.03; Figure 3.6). The Bayesian skyline model indicates that the effective population size of genotype ON1 had slow expansion, reached plateau early in 2012 and decreased slightly before the beginning of the season 2012-2013 (Figure 3.6).

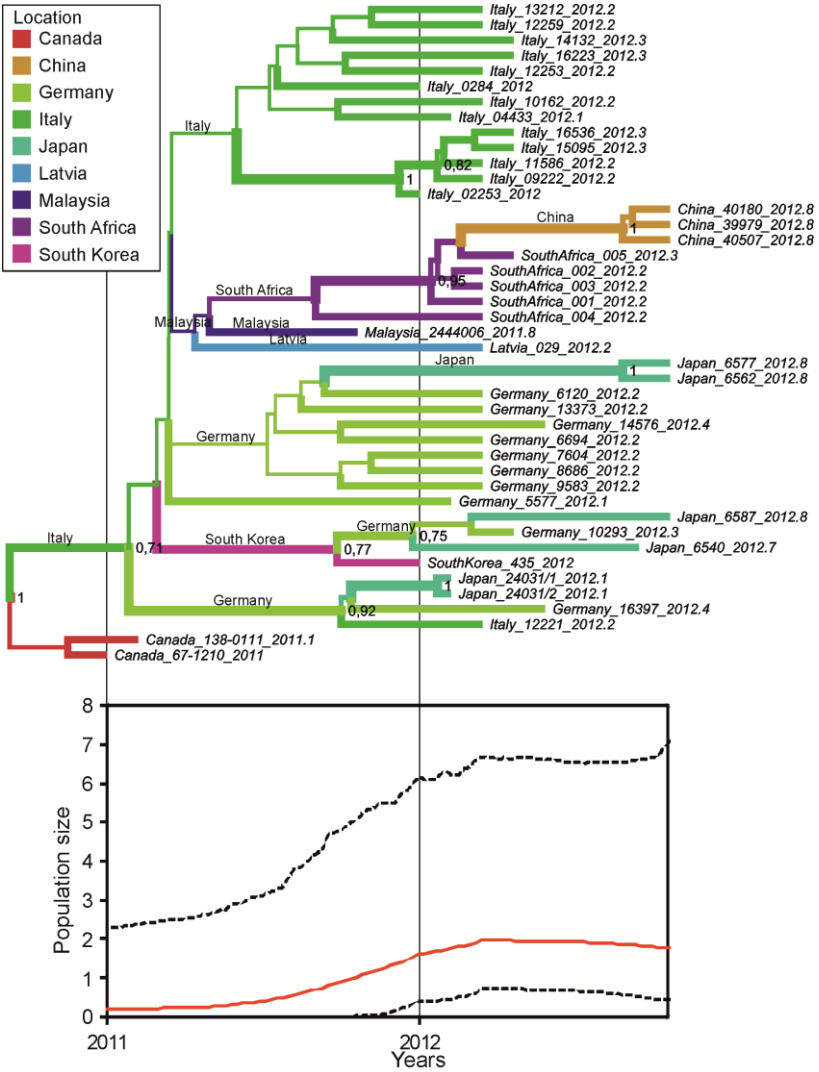

Figure 3.6. Discrete phylogeographic analysis and demographic history of genotype ON1

Branch color and width in the MCC tree represent location (see legend) and posterior support for the branch (values ≥0.7 are shown at nodes), respectively. The *y* axis of the BSP represents the population size, which is equal to the product of the effective population size (Ne) and the generation length in years (τ). The red solid line represents the median estimate and the area between the dotted lines represents the 95% HPD limits. Both graphs are drawn in the same time scale. Taxa names code for country of isolation, isolation number and time.

4. DISCUSSION

4.1. Study Cohort

In this prospective study at Children's Clinical University Hospital, the only tertiary level pediatric hospital in Latvia, patient population with LRTI according to the World Health Organization case definition were enrolled. Only previously healthy 2–24-month-old children were included for several reasons: (1) virtually all children older than 2 years have had HRSV infection at least once; (2) work of breathing in infants younger than 2 months and children with chronic conditions is influenced by multiple factors. Prematurity, congenital heart disease and chronic lung disease are significant risk factor for HRSV associated morbidity and mortality [Simões and Carbonell-Estrany, 2003], however most children with HRSV infection have been previously healthy and control strategies targeting only high-risk children will have a limited effect on the total disease burden [Anderson et al., 2013; Hall et al., 2009].

4.2. HRSV Detection

For the study purposes a new clinical sampling method was introduced in Children's Clinical University Hospital – NPA. By using NPA it is possible to increase test sensitivity by up to 30% when compared to nasal swabs [Macfarlane et al., 2005].

The RT-PCR protocol developed in this study not only precisely detects HRSV infection, but is also suitable for group differentiation. The main advantage of the RT-PCR when compared to IFA test (the standard at

Children's Clinical University Hospital at the time of study) is detection of the genome present at very low titer and/or replication non-competent virus.

HRSV epidemiology is fairly constant in Europe. In children ≤2 years of age hospitalized with LRTI HRSV is detected in 42–45% [Simões and Carbonell-Estrany, 2003]. Similarly, in this study HRSV caused 42.5% of LRTIs.

4.3. HRSV Seasonality

Seasonal activity in Latvia consistently started later in winter and lasted later in spring (February–April) compared to the average (December–February) of the Northern hemisphere [Bloom-Feshbach et al., 2013], but was similar to that observed in Russia [Tatochenko et al., 2010] and in biennial "late" seasons described in several European countries – Austria, Croatia, Finland, Germany, Sweden, and Switzerland [Mlinaric-Galinovic et al., 2012; Reyes et al., 1997]. The reasons for geographic HRSV seasonality differences are not well defined, but are thought to be related to local climate and host behavior factors [Stensballe et al., 2003].

2009–2012 cumulative data from this study (88 HRSV-positive samples) and The Centre for Disease Prevention and Control (CDPC) of Latvia (1297 HRSV-positive nationwide samples from patients of all ages; http://www.spkc.gov.lv/) are compared in Figure 4.1.

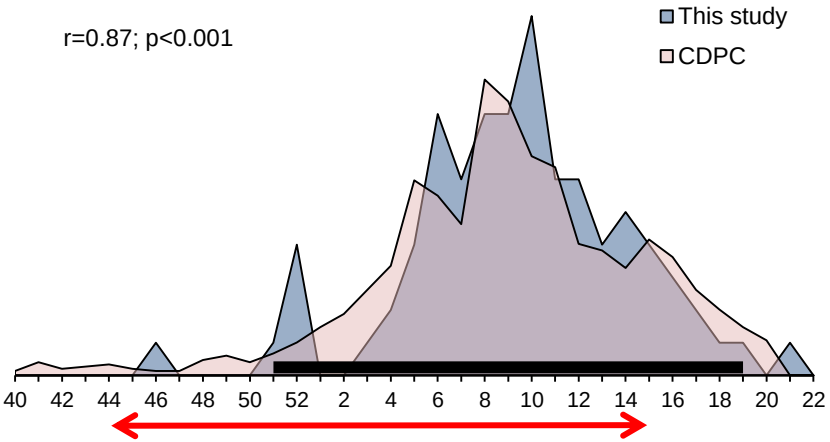

Figure 4.1. Seasonality of This Study and The Centre for Disease Prevention and Control

Normalized histograms represent relative HRSV-positive sample density distributions in the same scale. *X* axis represents weeks of the year. Thick black line indicates HRSV seasonal activity period over three years in this study (weeks 51–19). Red arrow represents palivizumab protection if applied according to American Academy of Pediatrics guidelines, assuming the effect lasts 45 days after the last injection [American Academy of Pediatrics Committee on Infectious Diseases, 2009; Panozzo et al., 2010].

Density distributions correlated well between these two data sets (r=0.87; p<0.001). CDPC detected HRSV-positive samples starting from week 40, but overall the peak activity (–2SD to +2SD) was similar – weeks 50 to 20. HRSV season was 21 weeks long, which correlates with the global median [Bloom-Feshbach et al., 2013].

4.4. HRSV Caused LRTI

During the annual seasons HRSV caused more than 90% of bronchiolitis and 50% of pneumonia that underlines the role of this pathogen in childhood morbidity. Patients with HRSV infection were younger, less likely to have SIRS, and were more likely to have atopy in the family history. There are

several hypotheses why infants during the first few months of life are at higher risk for severe HRSV infection: (i) immunologic processes (immune complex reaction in the lung; lack of secretory IgA antibodies; cell mediated Type IV reaction in the lung; T-cell independent mechanisms); (ii) physiological conditions (narrow airways, where small obstruction cause proportionally greater changes in cross-sectional area) [Sommer et al., 2011]. Relationship between HRSV and asthma is unclear. On one hand, asthma and recurrent wheezing is more commonly observed after severe HRSV disease. On the other, HRSV infection is more severe in atopic children. SIRS was more common in patients not infected with HRSV, because this group probably included patients with bacterial LRTIs.

4.5. Clinical Comparison of HRSV Groups and Genotypes

HRSV-positive samples were analyzed with group A and B specific primers. Specificity of the method was tested by nucleotide sequencing and in all cases it confirmed the group detected. Viruses of both groups co-circulated in each of the three seasons. HRSV-A strains were isolated more frequently and dominated in the first two of the analyzed seasons. Different patterns of group circulation have been described, but generally 1–3 successive predominant HRSV-A seasons are followed by a single season of HRSV-B dominance [Venter et al., 2001; Zlateva et al., 2007]. The reason for this alternating pattern is not entirely clear, but lower host susceptibility to homologous than heterologous strain re-infection has been observed and suggested to play a role [Sande et al., 2013; White et al., 2005].

Characteristics of patients and disease severity of HRSV-A and B infections are compared in Table 3.2. Since 97.7% of isolates represented the two dominant genotypes, NA1 and BA-IV, this table essentially gives the comparison of these genotypes (exclusion of genotype ON1 from the analysis did not significantly change the results). In this analysis length of hospital stay was measured as the primary outcome, while CSS was used as an indicator of disease severity. There were no statistically significant differences in these two parameters between the groups, thus the 2nd hypothesis of the study was rejected. It is possible that clinical differences of infections by different groups (and genotypes) were not observed because only hospitalized children were enrolled in the study excluding less acute patients.

4.6. Molecular Analysis of HRSV Strains

HVR2 of HRSV G gene is very convenient for molecular epidemiologic studies, because a relatively small fragment of the genome (less than 2%) represents the overall variability of the virus [Rebuffo-Scheer et al., 2011; Tan et al., 2012]. All the unique sequences of the analyzed samples were deposited in GenBank database. A single genotype in each group, NA1 in HRSV-A and BA-IV in HRSV-B, remained dominant in all three of the seasons. Very similar genotype assignments were reported in similar studies elsewhere (Figure 4.2). Thus the 1st hypothesis of the study was not rejected.

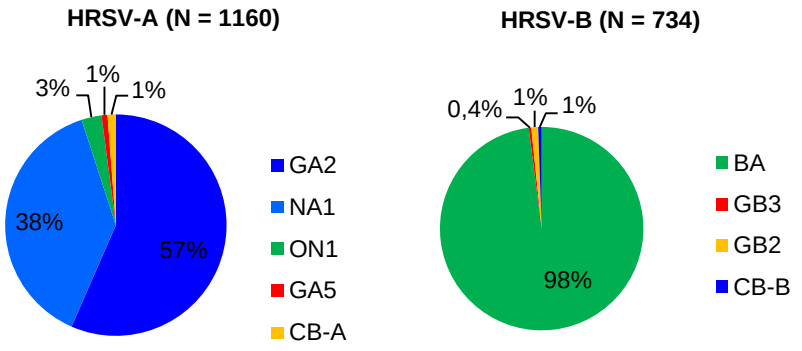

Figure 4.2. Global Distribution of HRSV Genotypes over Seasons 2009-2012
17 reports from PubMed search and the current study are summarized. Proportions of genotypes within groups are expressed as percentage. Genotypes GA2 and NA1 were not distinguished universally, therefore are presented in the same color.

Genotype NA1 was described first in Japan during the 2004–2005 season [Shobugawa et al., 2009], and was reported subsequently in numerous molecular epidemiologic studies as the dominant HRSV-A genotype (Figure 4.2). The NA1 genotype is related closely to GA2 and is not distinguished universally [van Niekerk and Venter, 2011; Yoshida et al., 2012]. Genotype GA5 was not detected in Latvia, although it was observed in other countries during the study period [Baek et al., 2012; Eshaghi et al., 2012; Houspie et al., 2013; Khor et al., 2013].

All HRSV-B sequences had a 60-nt duplication in the G gene, characteristic of BA genotypes first described in Buenos Aires (BA) in 1999 [Trento et al., 2003]. Since 2005 BA-IV has been the most common HRSV-B genotype detected worldwide and has replaced all other lineages [Agrawal et al., 2009; Houspie et al., 2013; van Niekerk and Venter, 2011; Trento et al., 2010].

In 2012, a year after its first description, another HRSV-A genotype, ON1, was detected. ON1 is characterized by a 72-nt in-frame duplication in the G gene. This strain emerged in Canada in the 2010–2011 season, where it accounted for 10% of HRSV-A infections [Eshaghi et al., 2012]. A similar distribution was reported from Germany in the following season (10%) [Prifert et al., 2013], and in 2012-2013 season in China [Cui et al., 2013] and Japan (13.4%) [Tsukagoshi et al., 2013]. In the 2011–2012 season in Latvia only ten HRSV-A viruses were detected, of which two (20%) were ON1.

There was a significant G protein length polymorphism among the isolates, with predicted lengths varying from 297 to 321 aa, accounted for by the following: (i) all NA1 sequences had a premature termination codon and were 297 aa long; (ii) the ON1 strain had a 321-aa-long sequence, resulting from a 24-aa insertion; (iii) BA-IV lengths were 310 and 317 aa, resulting from the use of alternative termination codons. The use of different reversible termination codons, including *Stop*313→*Q*, in HRSV-B has been described before [Botosso et al., 2009]. *In vitro*, the reduction of G protein size by in-frame premature termination codons allows neutralization escape by monoclonal antibodies without affecting HRSV infectivity [Rueda et al., 1991].

Strain specific monoclonal antibodies react only with glycosylated G protein [Martinez et al., 1997], therefore the glycosylation properties were evaluated by software predictions. Because of substitutions directly in glycosylation sites or surrounding sequences, *O*- and *N*-glycosylation sites were conserved poorly among the strains of both groups, further diversifying strain phenotypes. Despite sequence diversity and length polymorphism among the strains, the overall proline/serine/threonine proportions remained stable (roughly 10, 10 and 25%, respectively), maintaining mucin-type characteristics [Julenius et al., 2005].

Mucin-like regions, a pathogenicity factor in a variety of negative-stranded RNA viruses, are able to retain functional stability in the face of extreme genetic divergence [Wertheim and Worobey, 2009]. Previous studies have suggested that positive selection pressure is one of the main forces that drive the genetic variability of HRSV [Cane and Pringle, 1995], therefore in-depth natural selection analysis was also performed. Although the overall *dN/dS* ratio of NA1 strains was >1, indicative of evolutionary selection pressure, only two positively selected sites were identified. Both sites are involved in B and T cell epitopes of strain A2 [Cane, 1997; Hancock et al., 2003; Norrby et al., 1987]. These findings are not consistent among different studies, i.e. M262 and N273 were also predicted by [Eshaghi et al., 2012; Houspie et al., 2013; Kushibuchi et al., 2013; Yoshida et al., 2012; Zlateva et al., 2004], but not confirmed by [Botosso et al., 2009; Tan et al., 2012] using similar methodology. BA-IV strains had higher diversity that appeared to be driven less by positive selection pressure, as the overall *dN/dS* was 0.45. The two positively selected sites, L219 and T270, have been reported before by [Botosso et al., 2009; Kushibuchi et al., 2013; Zlateva et al., 2005] while other codons (142, 227, 235, 258, 259, 311) found to be positively selected in the same studies using stringent criteria were conserved completely in Latvian HRSV-B strains. There are several possible reasons for inconsistency among the studies. In this study, only one genotype in each group was studied. Other studies have extended over several decades and included other circulating genotypes and extinct lineages. This might also suggest strain adaptation to varying levels of herd immunity among the different populations. Finally, this might indicate a lack of positive selection and evolution of the virus under relaxed purifying selection, i.e. HVR2 as long as it retains mucin-like characteristics is not too restricted in its variability [Wertheim and Worobey, 2009].

4.7. Global Dissemination of HRSV Strains

Since the original description in Canada, ON1 strains have been isolated in many parts of the world. Prifert and coworkers [2013] found a strong association of this genotype with intensive care unit admission. However, this association has not been validated in other studies. The ON1-infected patients in this study had lower respiratory tract infections and, even though one of them required oxygen therapy for a brief period of time, they did not require intensive care. Eshaghi and coworkers [2012] suggested that the 72-nt duplication occurred because of RNA polymerase backtracking at an upstream 7-nt repeat motif and a stem loop in the secondary RNA structure. This is the largest natural insertion in the G gene described to date. Mutations like this are rare and unlikely to appear simultaneously in various locations; therefore, they can be used as a "natural tag" to reconstruct HRSV dissemination [Trento et al., 2010].

This is the first study that estimates population dynamics and discrete phylogeographic analysis of genotype ON1 based on globally isolated sequences from December, 2010 to June, 2013. The evolutionary rate determined based on HVR2 sequences (7.92×10^{-3} nt/site/year) was considerably higher than previously reported rates of the HVR2 segment of other HRSV-A and B genotypes ($3.6–4.7\times10^{-3}$) [Kushibuchi et al., 2013; Tan et al., 2012]. This indicates relaxed selective pressure operating in the duplicated region [Wertheim and Worobey, 2009].

The demographic history of ON1 so far has been considerably different than that of the BA genotype, which showed global expansion for the first 10 years after undergoing a 60-nt duplication in the G gene [Trento et al., 2010]. Although the effective population size decreased before the 2012–2013 season,

these data might change when new sequences become available. It is possible that as in the case of the BA genotypes, other mutations besides the duplication are required for optimal adaptation for replication. The mechanism by which duplication provides selective advantage for BA strains is unknown. It has been speculated that it modifies antigenic characteristics of the G protein, allowing escape from neutralizing antibodies. In the case of ON1, this duplication interferes with a region that contains several B and T cell epitopes [Cane, 1997] and changes the *O*- and *N*-glycosylation properties of the G protein.

Discrete phylogeographic analysis showed two main transmission patterns: (i) wave-like transmission from Canada through Italy to the rest of the world in West-to-East direction and (ii) gravity-like dynamics where perhaps following patterns of human travel routes the virus re-entered Germany [Holmes et al., 2008]. Although several transmissions lacked statistical support, this hypothesis is supported by several additional pieces of evidence. First, all of ON1 genomes available in GanBank also have *E*232→*G* and *T*253→*K* mutations. When NCBI-BLAST was queried for ON1 protein sequences without the duplicated region, none of the retrieved strains had these two mutations; however, *E*232→*G* was described in an NA1 strain isolated in the same study as genotype ON1 in Canada, suggesting possible ancestry. Second, there was another NA1 strain (LV/043/12) that was isolated in Canada and in the following season in Italy, Germany and Latvia. Third, the calculated tMRCA (August, 2010) matches well the estimation by Eshaghi et al. [2012] that this genotype emerged several months prior to winter 2010–2011. In summary, the 3rd hypothesis of the study was not rejected – Bayesian evolutionary analysis yielded credible reconstruction of genotype ON1 global spread.

5. CONCLUSIONS

1. For the study purposes a new PCR-based HRSV diagnostic tool was developed that also allowed to discriminate the main groups, HRSV-A and B. When compared to IFA as the „gold standard" this method showed sensitivity and specificity of 95.2 and 91.2%, respectively.
2. Among 2–24-month-old infants hospitalized with LRTI, HRSV caused 42.5% of the infections.
3. HRSV seasonal activity in Latvia was estimated from week 51 to 19.
4. No statistically significant differences in severity of disease were found between infections caused by HRSV-A and B viruses or among their genotypes (2nd hypothesis of the study was rejected).
5. Several strains detected previously elsewhere and unique strains from Latvian were isolated; their molecular epidemiology matched the global circulation patterns (1st hypothesis of the study was not rejected). The variability of the glycoprotein G was explained only partially by host immune-driven positive selection pressure.
6. By using molecular clock approach in Bayesian evolutionary analysis, the global spread and population size dynamics of genotype ON1 were reconstructed credibly (3rd hypothesis of the study was not rejected).

6. RECOMMENDATIONS

1. Considering the high proportion of HRSV-caused LRTI among hospitalized infants during the seasonal activity, it is recommended to implement a rapid diagnostic test in tertiary level pediatric hospital in Latvia for optimal cohorting of the patients. Rapid diagnosis could potentially save the HRSV-associated use of resources.
2. It is likely that HRSV seasonal activity starts and ends later than the average for the Northern hemisphere and guidelines from other countries may be of limited value. This finding should be taken in account when national immunoprohylaxis strategy is developed.

7. AUTHOR'S PUBLICATIONS

1. Balmaks R, Ribakova I, Gardovska D, Kazaks A. 2013. Molecular epidemiology of human respiratory syncytial virus over three consecutive seasons in Latvia. Journal of Medical Virology doi:10.1002/jmv.23855.
2. Balmaks R, Ribakova I, Gardovska D, Kazaks A. 2013. Molecular epidemiology of respiratory syncytial virus during the 2009–2010 season in Latvia. Archives of Virology 158(5):1089-1092.
3. Eihvalde L, Ansone S, Balmaks R, Ribakova I, Kazāks A, Grope I, Kozlovska L, Livdāne J, Gardovska D. 2012. Viral lower respiratory tract infections in children under 2 years of age in the winter season of 2010/2011 in Children's University Hospital, Riga (Latvia). Lietuvos Bendrosios Praktikos Gydytojas 16(9):612-617.
4. Balmaks R, Kazaks A, Likopa Z, Grope I, Rasnacs O, Gardovska D. 2011. A new tool for molecular monitoring of respiratory syncytial virus in children with lower respiratory tract infections. Riga Stradins University Collection of Scientific Papers 2010:100-105.

8. REFERENCES

1. Agrawal AS, Sarkar M, Ghosh S et a. 2009. Prevalence of respiratory syncytial virus group B genotype BA-IV strains among children with acute respiratory tract infection in Kolkata, Eastern India. Journal of Clinical Virology 45(4):358-361.

2. American Academy of Pediatrics Committee on Infectious Diseases. 2009. Modified recommendations for use of palivizumab for prevention of respiratory syncytial virus infections. Pediatrics 124(6):1694-1701.

3. American Academy of Pediatrics Subcommittee on Diagnosis and Management of Bronchiolitis. 2006. Diagnosis and management of bronchiolitis. Pediatrics 118(4):1774-1793.

4. Anderson LJ, Dormitzer PR, Nokes DJ, Rappuoli R, Roca A, Graham BS. 2013. Strategic priorities for respiratory syncytial virus (RSV) vaccine development. Vaccine 31, Supplement 2(0):B209-B215.

5. Baek Y, Choi E, Song M-S et al. Prevalence and genetic characterization of respiratory syncytial virus (RSV) in hospitalized children in Korea. Archives of Virology 157(6):1039-1050.

6. Bielejec F, Rambaut A, Suchard MA, Lemey P. 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics 27(20):2910-2912.

7. Bloom-Feshbach K, Alonso WJ et al. 2013. Latitudinal variations in seasonal activity of influenza and respiratory syncytial virus (RSV): a global comparative review. PLoS ONE 8(2):e54445.

8. Botosso VF, Zanotto PMdA, Ueda M et al. 2009. Positive selection results in frequent reversible amino acid replacements in the G protein sene of human respiratory syncytial virus. PLoS Pathog 5(1):e1000254.

9. Cane PA. 1997. Analysis of linear epitopes recognised by the primary human antibody response to a variable region of the attachment (G) protein of respiratory syncytial virus. Journal of Medical Virology 51(4):297-304.

10. Cane PA, Pringle CR. 1995. Evolution of subgroup A respiratory syncytial virus: evidence for progressive accumulation of amino acid changes in the attachment protein. Journal of Virology 69(5):2918-2925.

11. Collins P, Crowe J. 2007. Respiratory syncytial virus and metapneumovirus. In: Knipe D, Howley P, editors. Fields virology. 5th ed. Philadelphia: Lippincott Williams & Wilkins. p 1602-1646.

12. Collins PL, Melero JA. 2011. Progress in understanding and controlling respiratory syncytial virus: Still crazy after all these years. Virus Research 162(1–2):80-99.

13. Cui G, Qian Y, Zhu R et al. 2013. Emerging human respiratory syncytial virus genotypeON1 found in infants with pneumonia in Beijing, China. Emerging Microbes and Infections 2(4):e22.

14. Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26(19):2455-2457.

15. Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29(8):1969-1973.

16. Eshaghi A, Duvvuri VR, Lai R et al. 2012. Genetic variability of human respiratory syncytial virus A strains circulating in Ontario: A novel genotype with a 72 nucleotide G gene duplication. PLoS ONE 7(3):e32807.

17. Falsey AR, Hennessey PA, Formica MA, Cox C, Walsh EE. 2005. Respiratory syncytial virus infection in elderly and high-risk adults. New England Journal of Medicine 352(17):1749-1759.

18. Frost SDW, Volz EM. 2010. Viral phylodynamics and the search for an 'effective number of infections'. Philosophical Transactions of the Royal Society B: Biological Sciences 365(1548):1879-1890.

19. Glezen W, Taber L, Frank A, Kasel J. 1986. Risk of primary infection and reinfection with respiratory syncytial virus. American Journal of Diseases of Children 140(6):543-546.

20. Gupta R, Jung E, Brunak S. 2004. Prediction of N-glycosylation sites in human proteins. http://www.cbs.dtu.dk/services/NetNGlyc/.

21. Goldstein B, Giroir B, Randolph A, Sepsis MotICCoP. 2005. International pediatric sepsis consensus conference: Definitions for sepsis and organ dysfunction in pediatrics. Pediatric Critical Care Medicine 6(1):2-8.

22. Hall CB, Weinberg GA, Iwane MK et al. 2009. The burden of respiratory syncytial virus Infection in young children. New England Journal of Medicine 360(6):588-598.

23. Hancock GE, Tebbey PW, Scheuer CA, Pryharski KS, Heers KM, LaPierre NA. 2003. Immune responses to the nonglycosylated ectodomain of respiratory syncytial virus attachment glycoprotein mediate pulmonary eosinophilia in inbred strains of mice with different MHC haplotypes. Journal of Medical Virology 70(2):301-308.

24. Hart GW. 1992. Glycosylation. Current Opinion in Cell Biology 4(6):1017-1023.

25. Henderson FW, Collier AM, Clyde WA, Denny FW. 1979. Respiratory-syncytial-virus infections, reinfections and immunity. New England Journal of Medicine 300(10):530-534.

26. Holmes E. 2008. Evolutionary history and phylogeography of human viruses. Annual Review of Microbiology 62:307-328.

27. Hornsleth A, Klug B, Nir M, Johansen J, Hansen K, Christensen L, Larsen LB. 1998. Severity of respiratory syncytial virus disease related to type and genotype of virus and to cytokine values in nasopharyngeal secretions. The Pediatric Infectious Disease Journal 17(12):1114-1121.

28. Houspie L, Lemey P, Keyaerts E et al. 2013. Circulation of HRSV in Belgium: from multiple genotype circulation to prolonged circulation of predominant genotypes. PLoS ONE 8(4):e60416.

29. Imaz MS, Sequeira MD, Videla C, Veronessi I, Cociglio R, Zerbini E, Carballal G. 2000. Clinical and epidemiologic characteristics of respiratory syncytial virus subgroups A and B infections in Santa Fe, Argentina. Journal of Medical Virology 61(1):76-80.

30. Jafri HS, Wu X, Makari D, Henrickson KJ. 2013. Distribution of respiratory syncytial virus subtypes A and B among infants presenting to the emergency department with lower respiratory tract Infection or apnea. The Pediatric Infectious Disease Journal 32(4):335-340.

31. Johnson LW, Robles J, Hudgins A, Osburn S, Martin D, Thompson A. 2013. Management of bronchiolitis in the emergency department: impact of evidence-based guidelines? Pediatrics 131(Supplement 1):S103-S109.

32. Johnson PR, Spriggs MK, Olmsted RA, Collins PL. 1987. The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. Proceedings of the National Academy of Sciences 84(16):5625-5629.

33. Julenius K, Mølgaard A, Gupta R, Brunak S. 2005. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. Glycobiology 15(2):153-164.

34. Katzov-Eckert H, Botosso V, Neto E, Zanotto P, VGND consortium. 2012. Phylodynamics and dispersal of HRSV entails its permanence in the general population in between yearly outbreaks in children. PLoS One 7(10):e41953.

35. Khor C-S, Sam IC, Hooi P-S, Chan Y-F. 2013. Displacement of predominant respiratory syncytial virus genotypes in Malaysia between 1989 and 2011. Infection, Genetics and Evolution 14(0):357-360.

36. Kneyber M, Brandenburg A, Rothbarth P, de Groot R, Ott A, van Steensel-Moll H. 1996. Relationship between clinical severity of respiratory syncytial virus infection and subtype. Archives of Disease in Childhood 75(2):137-140.

37. Kosakovsky Pond SL, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21(5):676-679.

38. Kushibuchi I, Kobayashi M, Kusaka T et al. Molecular evolution of attachment glycoprotein (G) gene in human respiratory syncytial virus detected in Japan 2008–2011. Infection, Genetics and Evolution 18(0):168-173.

39. Larkin MA, Blackshields G, Brown NP et al. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23(21):2947-2948.

40. Lee W-J, Kim Y-j, Kim D-W, Lee HS, Lee HY, Kim K. 2012. Complete genome sequence of human respiratory syncytial virus genotype A with a 72-nucleotide duplication in the attachment protein G gene. Journal of Virology 86(24):13810-13811.

41. Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11):1451-1452.

42. Lowell DI, Lister G, Von Koss H, McCarthy P. 1987. Wheezing in infants: the response to epinephrine. Pediatrics 79(6):939-945.

43. Lozano R, Naghavi M, Foreman K et al. 2012. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet 380(9859):2095-2128.

44. Macfarlane P, Denham J, Assous J, Hughes C. 2005. RSV testing in bronchiolitis: which nasal sampling method is best? Archives of Disease in Childhood 90(6):634-635.

45. Martinello RA, Chen MD, Weibel C, Kahn JS. 2002. Correlation between respiratory syncytial virus genotype and severity of illness. Journal of Infectious Diseases 186(6):839-842.

46. Martinez I, Dopazo J, Melero JA. 1997. Antigenic structure of the human respiratory syncytial virus G glycoprotein and relevance of hypermutation events for the generation of antigenic variants. Journal of General Virology 78(10):2419-2429.

47. McIntosh E, De Silva L, Oates R. 1993. Clinical severity of respiratory syncytial virus group A and B infection in Sydney, Australia. Pediatric Infectious Disease Journal 12(10):815-819.

48. Meijboom MJ, Rozenbaum MH, Benedictus A et al. 2012. Cost-effectiveness of potential infant vaccination against respiratory syncytial virus infection in The Netherlands. Vaccine 30(31):4691-4700.

49. Mlinaric-Galinovic G, Tabain I, Kukovec T et al. 2012. Analysis of biennial outbreak pattern of respiratory syncytial virus according to subtype (A and B) in the Zagreb region. Pediatrics International 54(3):331-335.

50. Mufson MA, Örvell C, Rafnar B, Norrby E. 1985. Two distinct subtypes of human respiratory syncytial virus. Journal of General Virology 66(10):2111-2124.

51. Nair H, Nokes DJ, Gessner BD et al. 2010. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. The Lancet 375(9725):1545-1555.

52. Nikiforova R, Lucenko I, Zamjatina N. 2011. Pārskats par saslimstību ar akūtām augšējo elpceļu infekcijām (AAEI) un gripu 2010. - 2011. gada epidēmiskajā sezonā. Epidemioloģijas biļetens 34(1205):1-21.

53. Norrby E, Mufson MA, Alexander H, Houghten RA, Lerner RA. 1987. Site-directed serology with synthetic peptides representing the large glycoprotein G of respiratory syncytial virus. Proceedings of the National Academy of Sciences 84(18):6572-6576.

54. Palomo C, García-Barreno B, Peñas C, Melero JA. 1991. The G protein of human respiratory syncytial virus: significance of carbohydrate side-chains and the C-terminal end to its antigenicity. Journal of General Virology 72(3):669-675.

55. Panozzo CA, Stockman LJ, Curns AT, Anderson LJ. 2010. Use of respiratory syncytial virus surveillance data to optimize the timing of immunoprophylaxis. Pediatrics 126(1):e116-e123.

56. Peret TCT, Hall CB, Hammond GW et al. 2000. Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. Journal of Infectious Diseases 181(6):1891-1896.

57. Prifert C, Streng A, Krempl C, Liese J, Weissbrich B. 2013. Novel respiratory syncytial virus a genotype, Germany, 2011-2012. Emerging Infectious Diseases 19(6):1029-1030.

58. Rebuffo-Scheer C, Bose M, He J et al. 2011. Whole genome sequencing and evolutionary analysis of human respiratory syncytial virus A and B from Milwaukee, WI 1998-2010. PLoS ONE 6(10):e25468.

59. Reyes M, Eriksson M, Bennet R, Hedlund K, Ehrnst A. 1997. Regular pattern of respiratory syncytial virus and rotavirus infections and relation to weather in Stockholm, 1984--1993. Clinical Microbiology and Infection 3(6):640-646.

60. Reynolds EOR, Cook CD. 1963. The treatment of bronchiolitis. The Journal of pediatrics 63(6):1205-1207.

61. Rueda P, Delgado T, Portela A, Melero JA, García-Barreno B. 1991. Premature stop codons in the G glycoprotein of human respiratory syncytial viruses resistant to neutralization by monoclonal antibodies. Journal of Virology 65(6):3374-3378.

62. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4(4):406-425.

63. Sande CJ, Mutunga MN, Medley GF, Cane PA, Nokes DJ. 2013. Group- and genotype-specific neutralizing antibody responses against respiratory syncytial virus in infants and young children with severe pneumonia. Journal of Infectious Diseases 207(3):489-492.

64. Shay DK, Holman RC, Roosevelt GE, Clarke MJ, Anderson LJ. 2001. Bronchiolitis-associated mortality and estimates of respiratory syncytial virus-associated deaths among US Children, 1979–1997. Journal of Infectious Diseases 183(1):16-22.

65. Shobugawa Y, Saito R, Sano Y et al. 2009. Emerging genotypes of human respiratory syncytial virus subgroup A among patients in Japan. Journal of Clinical Microbiology 47(8):2475-2482.

66. Simões EAF, Carbonell-Estrany X. 2003. Impact of severe disease caused by respiratory syncytial virus in children living in developed countries. The Pediatric Infectious Disease Journal 22(2):S13-S20.

67. Sommer C, Resch B, Simões EAF. 2011. Risk factors for severe respiratory syncytial virus lower respiratory tract infection. The Open Microbiology Journal 5:144-154.

68. Stensballe L, Devasundaram J, Simões E. 2003. Respiratory syncytial virus epidemics: the ups and downs of a seasonal virus. The Pediatric Infectious Disease Journal 22(2):S21-S32.

69. Sullender WM, Sun L, Anderson LJ. 1993. Analysis of respiratory syncytial virus genetic variability with amplified cDNAs. Journal of Clinical Microbiology 31(5):1224-1231.

70. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28(10):2731-2739.

71. Tan L, Lemey P, Houspie L et al. 2012. Genetic variability among complete human respiratory syncytial virus subgroup A genomes: bridging molecular evolutionary dynamics and epidemiology. PLoS ONE 7(12):e51439.

72. Tatochenko V, Uchaikin V, Gorelov A et al. 2010. Epidemiology of respiratory syncytial virus in children ≤2 years of age hospitalized with lower respiratory tract infections in the Russian Federation: a prospective, multicenter study. Clinical Epidemiology(2):221-227.

73. Trento A, Casas I, Calderón A et al. 2010. Ten years of global evolution of the human respiratory syncytial virus BA genotype with a 60-nucleotide duplication in the G protein gene. Journal of Virology 84(15):7500-7512.

74. Trento A, Galiano M, Videla C et al. 2003. Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. Journal of General Virology 84(11):3115-3120.

75. Trento A, Viegas M, Galiano M et al. 2006. Natural history of human respiratory syncytial virus inferred from phylogenetic analysis of the attachment (G) glycoprotein with a 60-nucleotide duplication. Journal of Virology 80(2):975-984.

76. Tsukagoshi H, Yokoi H, Kobayashi M et al. 2013. Genetic analysis of attachment glycoprotein (G) gene in new genotype ON1 of human respiratory syncytial virus detected in Japan. Microbiology and Immunology. Epub: June 11, 2013. doi: 10.1111/1348-0421.12075.

77. Valley-Omar Z, Muloiwa R, Hu N, Eley B, Hsiao N. 2013. Novel respiratory syncytial virus subtype ON1 among children, Cape Town, South Africa, 2012. Emerging Infectious Diseases 19(4):668-670.

78. van Niekerk S, Venter M. 2011. Replacement of previously circulating respiratory syncytial virus subtype B strains with the BA genotype in South Africa. Journal of Virology 85(17):8789-8797.

79. Venter M, Madhi SA, Tiemessen CT, Schoub BD. 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. Journal of General Virology 82(9):2117-2124.

80. Wang E, Law B, Stephens D. 1995. Pediatric Investigators Collaborative Network on Infections in Canada (PICNIC) prospective study of risk factors and outcomes in patients hospitalized with respiratory syncytial viral lower respiratory tract infection. The Journal of Pediatrics 126(2):212-219.

81. Wang L-F, Collins PL, Fouchier RAM et al. 2012. Family Paramyxoviridae. In: King AMQ, Lefkowitz E, Adams MJ, Carstens EB, editors. Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. San Diego: Elsevier Academic Press. p 672-685.

82. Wertheim JO, Worobey M. 2009. Relaxed selection and the evolution of RNA virus mucin-like pathogenicity factors. Journal of Virology 83(9):4690-4694.

83. White L, Waris M, Cane P, Nokes D, Medley G. 2005. The transmission dynamics of groups A and B human respiratory syncytial virus (hRSV) in England & Wales and Finland: seasonality and cross-protection. Epidemiology and Infection 133(2):279-289.

84. Wright P, Cutts F. 2000. Generic protocol to examine the incidence of lower respiratory infection due to respiratory syncytial virus in children less than five years of age: field test version. Geneva: Department of Vaccines and Other Biologicals, World Health Organization. 1-34 p.

85. Yoshida A, Kiyota N, Kobayashi M et al. 2012. Molecular epidemiology of the attachment glycoprotein (G) gene in respiratory syncytial virus in children with acute respiratory infection in Japan in 2009/2010. Journal of Medical Microbiology 61(6):820-829.

86. Zhang R-F, Jin Y, Xie Z-P et al. 2010. Human respiratory syncytial virus in children with acute respiratory tract infections in China. Journal of Clinical Microbiology 48(11):4193-4199.

87. Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. Journal of Computational Biology 7(1-2):203-2014.

88. Zlateva KT, Lemey P, Moës E, Vandamme A-M, van Ranst M. 2005. Genetic variability and molecular evolution of the human respiratory syncytial virus subgroup B attachment G protein. Journal of Virology 79(14):9157-9167.

89. Zlateva KT, Lemey P, Vandamme A-M, van Ranst M. 2004. Molecular evolution and circulation patterns of human respiratory syncytial virus subgroup A: positively selected sites in the attachment G glycoprotein. Journal of Virology 78(9):4675-4683.

90. Zlateva KT, van Ranst M. 2004. Detection of subgroup B respiratory syncytial virus in the cerebrospinal fluid of a patient with respiratory syncytial virus Pneumonia. The Pediatric Infectious Disease Journal 23(11):1065-1066.

91. Zlateva KT, Vijgen L, Dekeersmaeker N, Naranjo C, van Ranst M. 2007. Subgroup prevalence and genotype circulation patterns of human respiratory syncytial virus in Belgium during ten successive epidemic seasons. Journal of Clinical Microbiology 45(9):3022-3030.