



HAL
open science

Molecular Epidemiology of Human Respiratory Syncytial Virus Subgroups A and B Identified in Adults with Haematological Malignancy Attending an Irish Hospital Between 2004 - 2009.

Brendan Crowley

► **To cite this version:**

Brendan Crowley. Molecular Epidemiology of Human Respiratory Syncytial Virus Subgroups A and B Identified in Adults with Haematological Malignancy Attending an Irish Hospital Between 2004 - 2009.. *Journal of Medical Virology*, 2010, 83 (2), pp.337. 10.1002/jmv.21957 . hal-00602308

HAL Id: hal-00602308

<https://hal.science/hal-00602308>

Submitted on 22 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Molecular Epidemiology of Human Respiratory Syncytial Virus Subgroups A and B Identified in Adults with Haematological Malignancy Attending an Irish Hospital Between 2004 – 2009.

Journal:	<i>Journal of Medical Virology</i>
Manuscript ID:	JMV-10-1937.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	24-Aug-2010
Complete List of Authors:	Crowley, Brendan; National Virus Reference Laboratory, University College Dublin; St James's Hospital, Department of Microbiology
Keywords:	Aisling Salter 1,2, Bairbre Ni Laoi 2, Brendan Crowley 1,3*

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

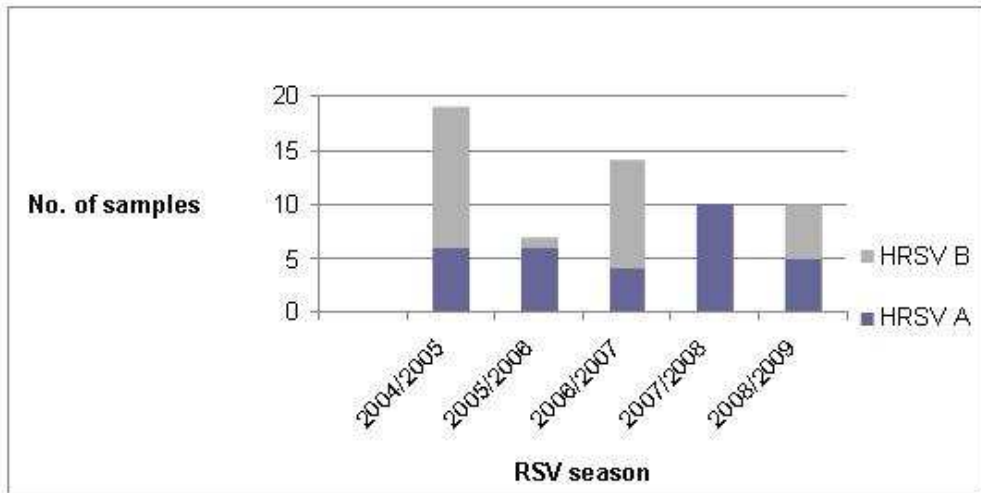


Figure 1. Distribution of HRSV subgroups in each winter season.
143x101mm (96 x 96 DPI)

Review

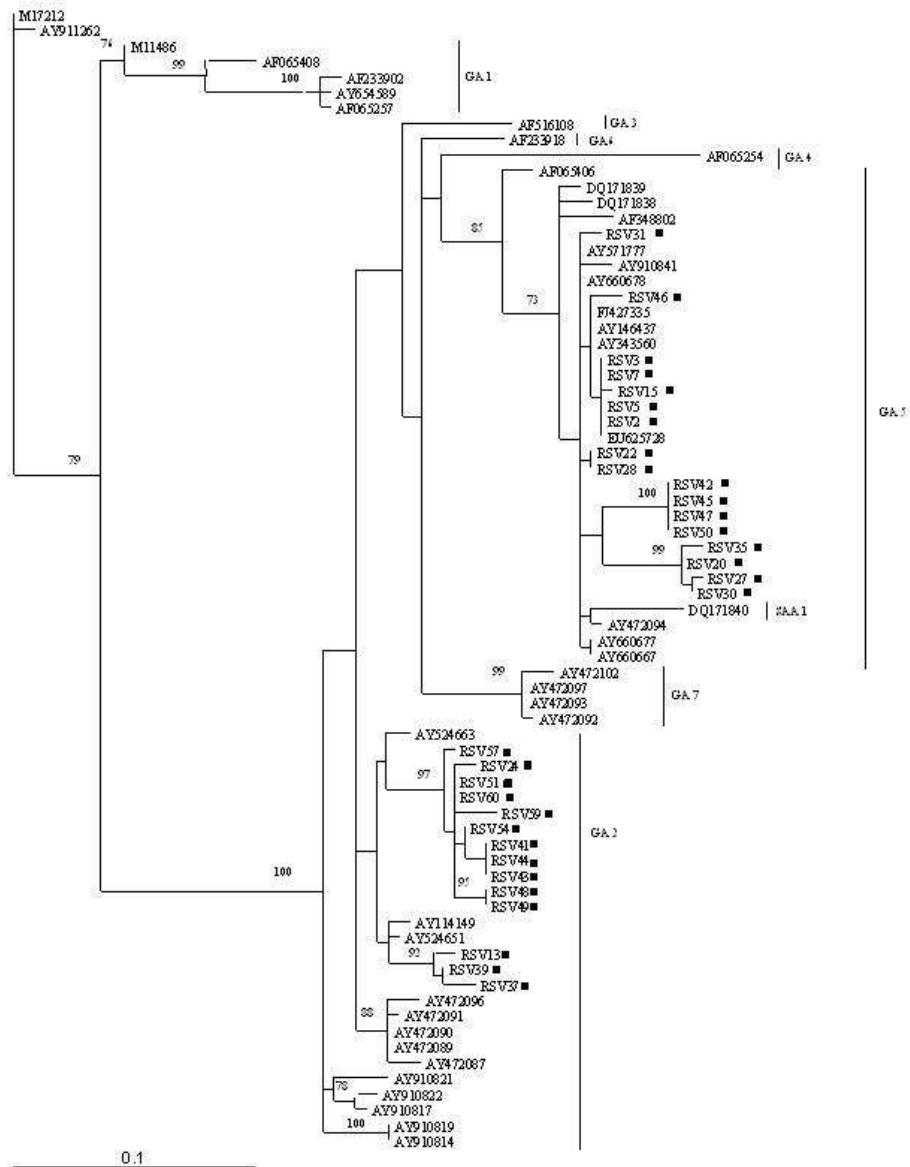


Figure 2 A maximum likelihood tree based on nucleotide sequences from the second variable region of the G protein gene for Irish HRSV subgroup A isolates from St. James's Hospital between 2004/2005 and 2008/2009. The sequences of the Irish strain are indicated by solid squares. HRSV A control sequence, A Strain Long (GenBank M17212) as used as an outgroup. Bootstrap resampling was carried out using 1,000 replicates of the dataset and the values are written as a percentile beside the appropriate branch. Bootstrap values <70% are not shown. Scale bars indicate 0.1 nucleotide substitutions per site.
164x212mm (96 x 96 DPI)

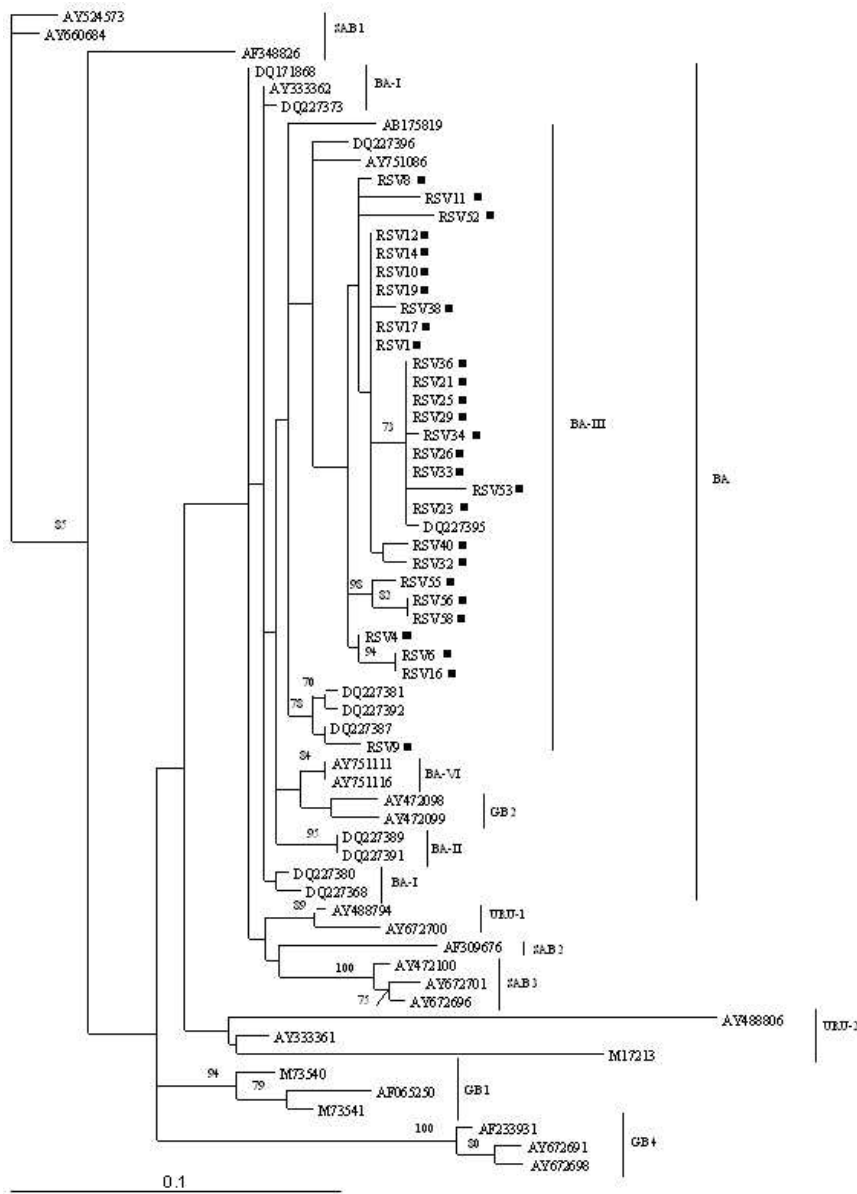


Figure 3. A maximum likelihood tree based on nucleotide sequences from the second variable region of the G protein gene for Irish HRSV subgroup B isolates from St. James's Hospital between 2004/2005 and 2008/2009. The sequences of the Irish strain are indicated by solid squares. HRSV B sequence, Ken/109/02 (GenBank AY524573) as used as an outgroup. Bootstrap resampling was carried out using 1,000 replicates of the dataset and the values are written as a percentile beside the appropriate branch. Bootstrap values <70% are not shown. Scale bars indicate 0.1 nucleotide substitutions per site.
161x207mm (96 x 96 DPI)

BA4128(BA)	Year	219	230	240	250	260	270	280	290	300	310
		L AKTLKKETTIN	PTKKPTPKTT	ERDTSTSQST	VLDTTTSKHT	ERDTSTSQST	VLDTTTSKHT	IQRRSLHSTT	PENTPNSTQT	PTASEPSTSN	STQKL*
RSV 1	2004/2001	P		P		I	A	Y			
RSV 4	2004/2001	P		P		I		Y			QSYA*
RSV 6	2004/2001	P		P		R	I	Y			IQSYA*
RSV 8	2004/2001	P		P		I	A	Y			
RSV 9	2004/2001		A	P		I			N		
RSV10	2004/2001	P		P		I	A	Y			
RSV11	2004/2001	P		P		I	A	Y			
RSV12	2004/2001	P		P		I	A	Y			
RSV14	2004/2001	P		P		I	A	Y			
RSV16	2004/2001	P		P		R	I	Y			IQSYA*
RSV17	2004/2001	P		P		I	A	Y			
RSV19	2004/2001	P		P		I	A	Y			
RSV23	2004/2004	P		P		L	I	Y	F		
RSV21	2004/2007	P		P		L	I	Y	F		
RSV25	2004/2007	P		P		L	I	Y	F		
RSV 26	2004/2007	P		P		L	I	Y	F		
RSV28	2004/2007	P		P		L	I	Y	F		
RSV32	2004/2007	P		A	P	I	A	Y	F		
RSV 33	2004/2007	P		P		L	I	Y	F		
RSV34	2004/2007	P		P		L	I	Y	F		
RSV36	2004/2007	P		P		L	I	Y	F		
RSV38	2004/2007	P		P		G	I	Y	F		
RSV40	2004/2007	P		A	P	I	A	Y			
RSV52	2003/2009	P		P		I	A	Y	L		QSYA*
RSV53	2003/2009	P		P		I		Y	L		
RSV55	2003/2009	P		P		P	N	Y	G		
RSV56	2003/2009	P		P		P	I	Y	G		
RSV58	2003/2009	P		P		P	I	Y	G		

Table 1. Deduced amino acid alignment of the second variable region of the G protein gene from HRSV subgroup B Irish strains. Alignments are shown relative to the sequence of genotype BA strain BA4128/99B (GenBank AY333364). The amino acids shown correspond to strain BA4128/99B G protein positions 219 to 315 for the subgroup B viruses. Identical residues are indicated by dots. Stop codons are indicated by asterisks. Potential N-glycosylation sites (NXT, where X is not proline) are underlined. The two copies of the duplicated 20 amino acid region in these RSV B viruses are indicated by rectangles.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

A2	Year	T	220	230	240	250	260	270	280	290	Q*																		
RSV2	2004/2005	P	A	K	<u>I</u>	P	R	<u>NS</u>	L	H	E	L	T	I	P	S	<u>ITD</u>	*					
RSV3	2004/2005	P	A	K	<u>I</u>	P	R	<u>NS</u>	L	H	E	L	T	I	Y	L	P	S	<u>ITD</u>	*		
RSV5	2004/2005	P	A	K	<u>I</u>	P	R	<u>NS</u>	L	H	E	L	T	I	P	S	<u>ITD</u>	*		
RSV7	2004/2005	P	A	K	<u>I</u>	P	R	<u>NS</u>	L	H	E	L	T	I	Y	L	P	S	<u>ITD</u>	*		
RSV13	2004/2005	P	L	K	P	R	H	K	L	T	SL	S	<u>TK</u>	*				
RSV15	2004/2005	P	A	K	<u>I</u>	P	R	<u>NS</u>	L	H	E	L	T	I	Y	L	P	S	<u>ITD</u>	*		
RSV20	2005/2006	P	A	S	D	K	P	R	<u>NS</u>	L	H	E	K	L	T	I	P	S	<u>ITD</u>	*		
RSV22	2005/2006	P	A	K	<u>I</u>	P	R	<u>NS</u>	L	H	E	L	T	I	Y	L	P	S	<u>ITD</u>	*		
RSV24	2005/2006	P	I	K	D	P	R	H	E	P	T	SP	SS	<u>TK</u>	*					
RSV27	2005/2006	P	A	S	D	K	P	R	<u>NS</u>	L	H	N	E	K	L	T	I	P	S	<u>ITD</u>	*	
RSV28	2005/2006	P	A	K	<u>I</u>	P	R	<u>NS</u>	L	H	E	L	T	I	Y	L	P	S	<u>ITD</u>	*		
RSV30	2005/2006	P	A	S	D	K	P	R	<u>NS</u>	L	H	N	E	K	L	T	I	P	S	<u>ITD</u>	*	
RSV31	2006/2007	P	A	K	<u>I</u>	P	R	<u>NS</u>	L	R	E	L	T	I	Y	L	P	S	<u>ITG</u>	*		
RSV35	2006/2007	P	A	S	D	K	P	R	<u>N</u>	L	K	H	E	K	L	T	I	Y	P	S	<u>ITD</u>	*
RSV37	2006/2007	P	L	K	P	R	<u>S</u>	H	K	L	T	SL	S	<u>TK</u>	*				
RSV39	2006/2007	P	L	K	P	R	<u>N</u>	H	K	L	T	Y	L	SL	S	<u>TK</u>	*			
RSV41	2007/2008	P	L	K	D	P	R	<u>A</u>	H	E	L	T	Y	L	SP	SS	<u>ATK</u>	*				
RSV42	2007/2008	P	A	D	K	<u>I</u>	P	R	<u>NS</u>	H	E	L	F	T	I	Y	L	P	S	<u>TDL</u>	*
RSV43	2007/2008	P	L	K	D	P	R	<u>A</u>	H	E	L	T	Y	L	SP	SS	<u>ATK</u>	*				
RSV44	2007/2008	P	L	K	D	P	R	<u>A</u>	H	E	L	T	Y	L	SP	SS	<u>ATK</u>	*				
RSV45	2007/2008	P	A	D	K	<u>I</u>	P	R	<u>NS</u>	H	E	L	F	T	I	Y	L	P	S	<u>TDL</u>	*
RSV46	2007/2008	P	A	K	<u>I</u>	P	R	<u>NS</u>	L	H	E	L	F	T	I	Y	L	P	LS	<u>ITD</u>	*	
RSV47	2007/2008	P	A	D	K	<u>I</u>	P	R	<u>NS</u>	H	E	L	F	T	I	Y	L	P	S	<u>TDL</u>	*
RSV48	2007/2008	P	L	K	S	VD	R	H	E	L	T	LN	Y	L	SP	SS	<u>TK</u>	*			
RSV49	2007/2008	P	L	K	S	VD	R	H	E	L	T	LN	Y	L	SP	SS	<u>TK</u>	*			
RSV50	2007/2008	P	A	D	K	P	R	<u>NS</u>	H	E	L	F	T	I	Y	L	P	S	<u>TDL</u>	*
RSV51	2008/2009	P	L	K	D	P	R	H	E	L	T	Y	L	SP	SS	<u>TK</u>	*				
RSV54	2008/2009	P	L	K	D	P	R	H	E	L	T	Y	L	SP	SS	<u>TK</u>	*				
RSV57	2008/2009	P	L	K	D	P	R	H	E	L	T	Y	L	SP	SS	<u>TK</u>	*				
RSV59	2008/2009	P	L	K	D	P	R	<u>Y</u>	H	E	L	T	Y	L	SP	SS	<u>TK</u>	*				
RSV60	2008/2009	P	L	K	D	P	R	H	E	L	T	Y	L	SP	SS	<u>TK</u>	*				

Table 2. Deduced amino acid alignment of the second variable region of the G protein gene from HRSV group A Irish strains. Alignments are shown relative to the sequence of prototype strain A2 (GenBank M11486). The amino acids shown correspond to strain A2 G protein positions 220 to 298 for the group A viruses. Identical residues are indicated by dots. Stop codons are indicated by asterisks. Potential N-glycosylation sites (NXT, where X is not proline) are underlined. Potential sites for extensive O-glycosylation KPX---TTKX motifs are indicated by highlighting.

1
2
3 **Molecular Epidemiology of Human Respiratory Syncytial Virus Subgroups A and B**
4 **Identified in Adults with Haematological Malignancy Attending an Irish Hospital Between**
5 **2004 – 2009.**
6
7
8
9

10
11
12 Aisling Salter^{1,2}, Bairbre Ni Laoi², Brendan Crowley^{1,3*}
13
14

15 ¹ Department of Clinical Microbiology, Trinity College Dublin, Dublin, Ireland
16

17 ² Department of Microbiology, Central Pathology Laboratory, St James's Hospital, Dublin 8,
18
19 Ireland
20

21 ³ National Virus Reference Laboratory, University College Dublin, Dublin 4, Ireland
22
23
24
25
26

27 *Corresponding author: bcrowley@stjames.ie
28
29
30
31
32
33

34 **Running title:** Genetic variation of HRSV in patients with haematological malignancy
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Human respiratory syncytial virus (HRSV) is an important cause of respiratory infection in patients with haematological malignancy, particularly haematopoietic stem cell transplant - recipients. This study investigated the genetic variability of the attachment (G) protein gene among HRSV isolates collected from adult patients with haematological malignancy. Between December 2004 and March 2009, 60 samples collected from 58 adults attending an Irish hospital were positive for HRSV by direct immunofluorescence. Nucleotide sequence analysis of the G gene showed a slightly higher frequency of HRSV subgroup A (52%) than HRSV subgroup B (48%). Genetic variability was higher among subgroup A viruses (up to 13% at nucleotide level) than among subgroup B viruses (up to 4%). Phylogenetic analysis revealed two genotypes of HRSV subgroup A, GA2 and GA5, which cocirculated between 2004/2005 and 2007/2008, although GA2 alone was identified in season 2008/2009. Genotype BA was the only genotype of HRSV subgroup B identified. Genotype-specific amino acid substitutions were identified, with two and seven changes for GA2 and GA5, respectively. Furthermore, one to four potential N-glycosylation sites were found among HRSV subgroup A isolates while two to three were identified in HRSV B isolates. Predicted O-glycosylation sites included 25 to 34 and 40 to 43 in HRSV subgroups A and B, respectively. The average synonymous mutation-to-nonsynonymous mutation ratios (dS/dN) implied neutral selection pressure on both HRSV subgroup isolates. This study provides data for the first time on the molecular epidemiology of HRSV isolates over five successive epidemic seasons among patients attending an Irish hospital.

Key words: haematological malignancy; HRSV subgroups; G-protein gene; phylogenetics; genotypes

Introduction

Human respiratory syncytial virus (HRSV) is recognised as a significant pathogen among the elderly, in particular those individuals with cardiopulmonary disease, and also immunocompromised hosts [Falsey and Walsh, 2000]. The main routes of transmission are by contact and droplet spread [Hall, 2001]. It is associated with significant morbidity and mortality in haematopoietic stem cell transplant recipients [Avetisyan *et al.*, 2009]. In a study involving post-bone marrow transplant patients, it was found that nosocomial infection with HRSV occurred in almost half (48%) of them [Whimbey *et al.*, 1996]. Furthermore, progression of HRSV infection from upper respiratory tract infection to lower respiratory tract infection is associated with increased risk of death [Avetisyan *et al.*, 2009].

HRSV strains have been classified on the basis of reactivity of the virus with monoclonal antibodies directed against the attachment glycoprotein, a type II integral membrane protein, which is highly glycosylated with both N-linked and O-linked sugars [Wertz *et al.*, 1989] and is the main target for neutralising and protective humoral immune responses [Galiano *et al.*, 2005]. Therefore, antigenic variability as a consequence of mutations in the two hypervariable regions of the G protein contributes to the ability of the virus to cause HRSV reinfections throughout life. These two hypervariable regions are separated by a central conserved region of 13 amino acids [Johnson *et al.*, 1987]. The second hypervariable region of the C-terminal of the G protein has only 44% amino acid homology between the two groups of HRSV while the transmembrane and cytoplasmic domains of the G protein have 83% identity on comparison of amino acid

1
2
3 sequences [Johnson *et al.*, 1987]. Furthermore, it has been shown that there is approximately
4
5 20% sequence variation in this region of the G protein among RSV isolates belonging to the
6
7 same antigenic group [Melero *et al.*, 1997]. Not surprisingly, sequencing of the variable regions
8
9 of the G protein gene has been used widely in subdividing further the two groups into genotypes
10
11 and therefore facilitating differentiation between RSV isolates.
12
13
14
15
16

17 Eight HRSV subgroup A genotypes have been described so far, namely GA1 to GA7 and South
18
19 Africa A1 (SAA1) [Peret *et al.*, 2000, Peret *et al.*, 1998, Venter *et al.*, 2001]. HRSV B viruses
20
21 have been divided into genotypes GB1-4 [Peret *et al.*, 2000, Peret *et al.*, 1998], SAB1-3 [Venter
22
23 *et al.*, 2002, Venter *et al.*, 2001], and BA [Trento *et al.*, 2003]. Of these HRSV B genotypes, the
24
25 BA genotype is relatively new and is characterised by a 60 nucleotide duplication in the G
26
27 glycoprotein gene of HRSV group B [Trento *et al.*, 2003]. An increasing incidence of strains
28
29 containing this duplication has been reported from Buenos Aires [Trento *et al.*, 2006]. This
30
31 variant of HRSV probably arose when the viral polymerase duplicated sequences within the G
32
33 protein gene. It originated from an ancestor between 1998 and 1999, just before the first isolation
34
35 of these variants in 1999, and can be used as a natural tag to study the evolution and
36
37 dissemination of HRSV worldwide [Trento *et al.*, 2006]. Phylogenetic analysis of this strain of
38
39 HRSV B determined that there was sequence divergence within the BA genotype and further
40
41 subdivided it into subgroups BA-I to BA-VI [Trento *et al.*, 2006]. There have been reports of the
42
43 BA genotype circulating in Argentina, Australia, Belgium, Canada, Kenya, New Zealand, Japan
44
45 and Uruguay [Matheson *et al.*, 2006, Tozer *et al.*, 2008, Trento *et al.*, 2003, Trento *et al.*, 2006].
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Data on HRSV infection based on GP and hospital sentinel surveillance programmes in the Republic of Ireland is published weekly by the Health Protection Surveillance (HPSC) (<http://www.ndsc.ie/hpsc/A-Z/Respiratory/Influenza/SeasonalInfluenza/InfluenzaSurveillanceReports>). This information shows that HRSV infection occurs predominantly in children less than five years of age and in a seasonal pattern, mainly between October and March each year. An active surveillance programme in the adult haematopoietic stem cell transplant facility in St James's Hospital targets immunocompromised haematology patients with the aim of early identification of influenza A and B viruses, HRSV, parainfluenza viruses types 1 to 4 and adenovirus. There is no information published regarding the molecular characterisation of HRSV in the Republic of Ireland. Therefore, this study investigated the distribution of HRSV subgroups A and B virus infections among adult patients with haematological malignancy attending St. James's Hospital over between November 2004 and March 2009. Furthermore, phylogenetic analysis was performed to determine the different genotypes of HRSV in this cohort during those five winter seasons.

Materials and Methods

Sample selection and Storage

A total of 60 samples collected from patients attending St. James' Hospital (SJH) over the period from November 2004 to March 2009 were selected for this study. Samples included nasopharyngeal aspirates and bronchoalveolar lavages (BALs). These samples were confirmed positive for human respiratory syncytial virus (HRSV) by direct immunofluorescence (ImagenTM RSV A & B, Oxoid, UK). An aliquot of NPA or BAL (approx 1.5ml) from each sample was frozen at -80°C in an RNase and DNase free tubes until required.

Human respiratory syncytial virus (HRSV) RNA extraction

Template RNA was extracted from 200µl of each sample using the QIAamp MinElute Virus Spin Kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer's instructions.

Reverse Transcription Polymerase Chain Reaction (RT-PCR)

The template RNA was converted to complementary DNA (cDNA) using the High Capacity Archive kit (Applied Biosystem Warrington, UK) following the manufacturer's instructions.

Amplification the C-terminal of G protein gene

To amplify the C-terminal region of the G protein gene, hemi-nested primers (Gr5, FV and F1) were used [da Silva *et al.*, 2008] (Supplementary Table 1). Primer pair Gr5 and FV amplified the first round product that results in a 912bp amplicon. Any isolates that failed to amplify in

1
2
3 round 1 were tested by hemi-nested PCR using the primer pair Gr5 and F1 with an amplicon size
4 of 828bp. PCR amplification was carried out on cDNA extracts using the HotStar polymerase
5 kit (Qiagen GmbH, Hilden, Germany). First and second round PCR reaction components,
6
7
8 volumes used and final concentrations achieved are outlined in Supplementary Table 2. A non-
9
10
11 template control and pre-extracted negative control were included with each run. Both first and
12
13
14 second round cycling parameters were identical and consisted of an initial temperature of 94°C
15
16
17 for 3 minutes, followed by 45 cycles of 94°C for 30 seconds (denaturation), 57°C for 30 seconds
18
19
20 (primer annealing) and 72°C 1 minute (extension) and a final elongation step of 72°C for 10
21
22
23 minutes. The PCR product was purified using the Qiaquick® PCR Purification kit (Qiagen
24
25
26 GmbH, Hilden, Germany) following the manufacturer's instructions.

27 28 29 ***DNA Sequencing and Phylogenetic analysis***

30
31 Automated DNA sequencing was performed on all samples using the ABI Prism 3130xl
32
33
34 Analyser (Applied Biosystems, Warrington, UK). Sequence traces were analysed and a
35
36
37 consensus sequence for all samples were obtained using the Lasergene® SeqMan sequence
38
39
40 analysis software (DNASTAR, Madison, Wisconsin). All nucleotide and amino acid sequences
41
42
43 were aligned using Clustal W [Thompson *et al.*, 1997] and sequence polymorphisms within the
44
45
46 amplified region were identified by comparison with a variety of sequences collected from
47
48
49 GenBank which represented circulating subtypes of HRSV A and B during the defined season.
50
51
52 Aligned sequences were then translated into protein using the protein translation function within
53
54
55 the Lasergene® MegAlign software (DNASTAR, Madison, Wisconsin). Amino acid sequences
56
57
58 were compared to the GenBank sequences at known mutation sites and any changes in amino
59
60
61 acids were noted. The sequences analysed correspond to amino acid residues 220 to 298 of the

1
2
3 prototype A2 strain (GenBank M11486) [Wertz *et al.*, 1985] and amino acid residues 219 to 315
4
5 of the BA4128/99B strain (GenBank AY333364) [Trento *et al.*, 2003]. Known substitution sites
6
7 and any changes in amino acids were noted.
8
9

10
11
12 Phylogenetic reconstructions, all employing maximum likelihood, were carried out using
13
14 PAUP*4.0b10 (Swofford, 2002). The 'best fit' models of evolution for the aligned datasets were
15
16 determined using the Modeltest version 3.06 [Posada and Crandall, 1998]) by comparing the
17
18 likelihood of the different parameters for all models of substitution. Model parameters were then
19
20 estimated using a heuristic search strategy and selecting those with the highest likelihood.
21
22 Bootstrap resampling was carried out using PAUP* with a 1000 replicates performed to examine
23
24 the robustness of the trees produced. In total, two phylogenetic trees were constructed.
25
26
27
28
29
30
31
32
33

34 ***Identification of potential glycosylation sites in the G protein of HRSV A and HRSV B isolates***

35
36 Potential N-glycosylation sites were identified in both HRSV A and HRSV B isolates from the
37
38 amino acid sequences, where the amino acid motif, asparagines X threonine (NXT), was seen,
39
40 where X is not proline. To estimate the numbers of potentially O-glycosylated serine and
41
42 threonine residues, the NetOglyc software (version 3.1) [Hansen *et al.*, 1995] was used.
43
44
45
46
47

48 ***Analysis of Synonymous vs Non-Synonymous mutations***

49
50 Synonymous and non-synonymous mutations were analysed by the method of Nei and Gojobori
51
52 [Nei and Gojobori, 1986]. The program SNAP (Synonymous/Nonsynonymous Analysis
53
54 Program) provided by the HIV database website
55
56
57
58
59
60

1
2
3 (<http://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html>) was used for analysis of
4
5
6 synonymous mutations versus nonsynonymous mutations.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

Results

Study population and clinical details

A total of 980 samples, including nasopharyngeal aspirates (98%) and bronchial washings (2%) were collected for laboratory diagnosis of respiratory viruses at St. James's Hospital Dublin between December 2004 and March 2009. All samples were collected from adult patients, aged between 19 and 77 years, who had haematological malignancies and were on immunosuppressive therapy. Most patients (85%) had undergone haematopoietic stem cell transplantation. Eighty-three samples were confirmed positive for HRSV infection by direct immunofluorescence during this time period but only 60 samples from 58 patients were available for analysis in this study. Interestingly, all HRSV-positive samples were consistent with community-acquired HRSV infection and collected from patients attending the Outpatient Clinic. However, almost 50% (n=31) of these patients were hospitalised for treatment with aerosolised ribavirin 2g every 8 hours because of risk factors for severe complications of HRSV infection. The latter included patients who were within 2 months of haematopoietic stem cell transplant, on immunosuppression for acute or chronic GVHD, lymphopenic ($< 0.3/\text{mm}^3$) or received alemtuzumab or anti-thymocyte globulin (ATG) as part of conditioning regimen. The underlying haematological diseases in these patients, 70% of whom were male, included chronic myeloid leukaemia (n=10), chronic lymphocytic leukaemia (n=10), acute lymphoblastic leukaemia (n=10), acute myeloid leukaemia (n=9), non-Hodgkins lymphoma (n=8), multiple myeloma (n=7), mantle cell lymphoma (n=2) and aplastic anaemia (n=2). The main clinical presentations involved the upper respiratory tract, and included cough (100%), rhinorrhoea (100%) and fever (50%). However, three patients (2%) who were diagnosed with pneumonia and respiratory

1
2
3 failure were transferred to the intensive care unit (ICU). None of them died of HRSV lower
4
5 respiratory tract infection.
6
7

10 ***Human respiratory syncytial virus (HRSV) subgroups***

11
12 Of these 60 isolates, 31 (52%) belonged to HRSV subgroup A and 29 (48%) belonged to HRSV
13
14 subgroup B. An overview of the distribution of HRSV subgroups over the five seasons is shown
15
16 in Figure 1. HRSV subgroup B predominated in 2004/2005 and 2006/2007 while in the
17
18 2005/2006 and 2007/2008 seasons HRSV A were more prevalent. The 2008/2009 season
19
20 showed an even distribution of isolates of both subgroups.
21
22
23
24
25
26
27
28

29 ***Phylogenetic analysis and genotype distribution of HRSV subgroup A and B sequences.***

30
31 The G protein gene in 59 of the 60 subgroup A and B HRSV isolates were successfully amplified
32
33 and sequenced. One isolate, RSV 18, failed to amplify successfully but was shown to belong to
34
35 subgroup B based on molecular analysis of the SH gene (data not shown). A maximum
36
37 likelihood phylogenetic tree was constructed of the G protein gene of the 31 subgroup A HRSV
38
39 isolates collected from St. James's Hospital (SJH) over that past five winter seasons, including a
40
41 selection of strains that represented the different genotypes of HRSV A isolates collected from
42
43 the GenBank database (Figure 2). The reference strain Long (GenBank M17212) was used as an
44
45 outgroup in the phylogenetic tree. In this study the HRSV A isolates belonged to two genotypes,
46
47 14 (45%) were genotype GA2, and 17 (55%) belonged to genotype GA5. The rates of
48
49 divergence between prototype A2 and the Irish strains were 12% to 18% at nucleotide level, and
50
51 13% to 18% at the amino acid level. Differences of 13% were observed at both the nucleotide
52
53
54
55
56
57
58
59
60

1
2
3 and amino acid level among the Irish HRSV A strains. Differences of 8% were observed among
4
5 the genotype GA2 strains at both nucleotide and amino acid level. Genotype GA5 strains
6
7 showed as much as 6% nucleotide and amino acid differences.
8
9

10
11
12 The phylogenetic tree also demonstrated that both of the HRSV A genotypes co-circulated in
13
14 each of the epidemic seasons with the exception of the 2008/2009 season, where all of the HRSV
15
16 A isolates belonged to genotype GA2 (RSV 51, RSV 54, RSV 57, RSV 59 and RSV 60) (Figure
17
18 2). In 2004/2005, of the six HRSV A isolates, one (RSV 13) belonged to genotype GA2 and five
19
20 (RSV 2, RSV 3, RSV 5, RSV 7 and RSV 15) to GA5. The 2005/2006 season had five (RSV 20,
21
22 RSV 22, RSV 27, RSV 28 and RSV 30) isolates that were genotype GA5 and one (RSV 24) was
23
24 GA2. The 2006/2007 season, showed an even distribution of isolates between the two genotypes
25
26 GA2 and GA5, with two (RSV 31 and RSV 35) belonging to GA5 and two (RSV 37 and RSV
27
28 39) belonging to GA2. The 2007/2008 season also showed an even distribution between the two
29
30 genotypes, with five (RSV 41, RSV 43, RSV 44, RSV 48 and RSV 49) identified as genotype
31
32 GA2 and five (RSV 42, RSV45, RSV 46, RSV 47 and RSV 50) in genotype GA5 cluster. The
33
34 isolates from this study that belonged to genotype GA5 showed close similarity to isolates from
35
36 Uruguay (AY571777), Argentina (AY910841), Kenya (AY660678), Brazil (FJ427335 and
37
38 EU625728), South Africa (AY146437) and Belgium (AY343560) (Figure 2). Among the GA2
39
40 isolates a high similarity was seen to isolates from Kenya (AY524663 and AY524651) and
41
42 Singapore (AY114149) (Figure 2).
43
44
45
46
47
48
49

50
51
52 Maximum likelihood phylogenetic trees were constructed for the G protein gene of all the HRSV
53
54 B isolates in this study along with a variety of strains collected from the GenBank database that
55
56
57
58
59
60

1
2
3 represented each of the known HRSV B genotypes (Figure 3). The reference strain Ken/109/02
4 (GenBank AY524573) was used as an outgroup in the phylogenetic tree. In this study, all of the
5 HRSV subgroup B isolates examined belonged to the newly identified genotype, genotype BA.
6 This genotype is defined by a 60 nucleotide (nt) duplication at position 781-840 in the second
7 variable region of the G protein gene. All of the HRSV B isolates from this study belonged to
8 subgroup BA-III within the BA genotype, the first one being detected in December 2004. There
9 was 5% to 9% divergence at both nucleotide and amino acid level, respectively, between the
10 Irish HRSV B strains and the prototype BA strain. Among the Irish HRSV B strains, up to 4%
11 divergence was observed at both nucleotide level and amino acid level. The majority of the
12 isolates had a high similarity with two isolates from a study in Spain, DQ227395 and DQ227396
13 [Trento *et al.*, 2006], and an isolate (AY751086) from a Belgian study in 2001 [Zlateva *et al.*,
14 2005] (Figure 3). However, one isolate, RSV 9, clustered with strains of HRSV subgroup B
15 including DQ227381, DQ227392 and DQ227387, all isolated in Spain [Trento *et al.*, 2006] and
16 another isolate (AB175819) from a study in Japan [Sato *et al.*, 2005] (Figure 3). It was observed
17 that all of the HRSV B isolates in this study were quite conserved from season to season with
18 very little variability among the amino acid sequences observed (Table 1). However, all of the
19 2006/2007 isolates (RSV 21, RSV 23, RSV 25, RSV 26 and RSV 29, RSV 33, RSV 34 and RSV
20 36) formed a distinct cluster with the strain DQ227395 (Figure 3) each had amino acid
21 substitution S297F, which occurred with the potential N-glycosylation site at position 296 (Table
22 1). Also in this cluster of isolates was isolate HRSV 53 (Figure 3), which had a different amino
23 acid substitution, S297L (Table 1).
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Sequence analysis of the G glycoprotein

Predicted amino acid sequences of the C-terminal end of the G protein gene in all the HRSV A isolates in this study were compared with prototype A2 strain. Seventeen of the HRSV A isolates were predicted to encode G proteins of 298 amino acids in length and 14 of them encoded a G protein of 297 amino acids (Table 2). Interestingly, the length of the G protein was genotype-specific in this study as G proteins of amino acid length of 297 belonged to the GA2 genotype while G proteins of 298 amino acids in length belonged to genotype GA5. Furthermore, genotype-specific amino acid substitutions were observed at several positions among the HRSV subgroup A isolates. For instance, genotype GA2 was characterised by S269T (serine to threonine) and P289S (proline to serine). Similarly, GA5 was distinguished by amino acid substitutions V225A (valine to alanine), S250N (serine to asparagine), N251S (asparagine to serine), P274T (proline to threonine) and V279I (valine to isoleucine).

In this study, the partial G protein gene sequences of the BA genotypes were predicted to encode proteins of three different lengths, 312, 315 and 319 amino acids (Table 1). The majority of the isolates (23/28) had predicted amino acid lengths of 312 residues, one isolate encoded a protein of 315 amino acids and four isolates had 319 amino acids (Table 1). A serine (S) to proline (P) amino acid change occurred within the 20-amino acid duplicated region at position 247 in all of the Irish BA strains compared to the sequence of the prototype BA strain. A second conserved amino acid substitution was also observed at position 270 where a threonine (T) was substituted for an isoleucine (I) when compared with the prototype BA strain. The majority of isolates also had amino acid substitutions at positions 271 (V271A) and 287 (H287Y) (Table 1).

Glycosylation sites

Three putative N-glycosylation sites (NXT, where X is not proline) were identified on the second variable region of the G protein among the majority (19/31) of the HRSV A strains (Table 2). Two potential sites were identified in 10 of the 31 isolates, one potential site in one isolate (RSV 59) and four potential sites in one isolate (RSV 13) (Table 2). The potential N-glycosylation site, at position 294 was conserved among all of the isolates while that at residue 237 was also observed among the majority of the isolates (20/31) (Table 2). Among the GA2 genotypes, 12 of the isolates had a conserved N-glycosylation pattern at residue 251 consistent with the site being genotype-specific. The remaining two GA2 isolates (RSV 37 and RSV 59) lost this site due to an amino acid substitution at position 251 (N251S and N251Y, respectively) (Table 2). Furthermore, three isolates (RSV 13, RSV 37 and RSV 39) were observed to have gained an N-glycosylation site at position 244, where isoleucine was replaced by asparagine (Table 2).

Two N-glycosylation sites were identified among the group B strains at the C-terminal end of the G protein gene. One of these, at amino acid 296, was conserved among all isolates examined while the second site, amino acid 310, was found in the majority (26/28) of HRSV subgroup B isolates (Table 1). Two of the isolates analysed, RSV 9 and RSV 55, acquired N-glycosylation sites as a result of amino acid substitutions at K278N and D253N, respectively.

One of the features of the G protein is the large number of serine and threonine residues that are potential O-linked sugar acceptors [Wertz *et al.*, 1985]. The analysis of the deduced 78 to 79 amino acid sequences of the group A strains by the program NetOglyc predicted 25 to 34 potential serine and threonine residues to be O-glycosylated with G scores of between 0.5 and

1
2
3 0.8. The same program also predicted 40 to 43 serine and threonine residues to be potentially O-
4 glycosylated with the score predictors (G scores) of between 0.5 and 0.8 in the deduced 94 to
5
6
7
8 101 amino acids of group B strains.
9

10
11
12 In addition to serine and threonine residues, two repeats of the KPX---TTKX were present
13 among subgroup A strains and may be associated with extensive O-glycosylation of the G
14 protein (Table 2) [Cane *et al.*, 1991, Roca *et al.*, 2001]. This motif was not identified among the
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Synonymous vs Non-Synonymous mutations

To study the evolutionary divergence of the HRSV A and B sequences the number of
synonymous (silent) and non-synonymous (amino acid altering) nucleotide substitutions was
estimated using SNAP (www.hiv.lanl.gov) [Korber, 2000]. On average, the synonymous
mutation/non-synonymous mutation (dS/dN) ratio for all the HRSV groups A samples (i.e. GA5
and GA2) in this study was 1.59. The average synonymous mutation/non-synonymous mutation
(dS/dN) ratios for genotypes GA2 and GA5 were 1.64 and 3.13, respectively. Ratios greater than
one demonstrated a high abundance of synonymous mutations. This suggested neutral selection
pressure on the variable region for all of the HRSV A isolates. Similarly, the mean dS/dN ratio
for all of the HRSV B samples in the study was 2.04 which again suggested neutral selection.

Discussion

In this study, the C-terminal region of the G protein was examined in all isolates collected over five seasons between 2004/2005 and 2008/2009 to obtain important phylogenetic and antigenic information of the circulating HRSV viruses in Ireland. Phylogenetic analysis of HRSV positive isolates in this study revealed that the HRSV A subgroup occurred at a slightly higher frequency over the five winter seasons from 2004/2005 to 2008/2009, with HRSV A and B at 52% and 48%, respectively. It was observed that HRSV B predominated in two of the five seasons (2004/2005 and 2006/2007) investigated in this study, although HRSV A was the only subgroup found in the 2005/2006 and 2007/2008 seasons when it was the only HRSV subgroup detected. During the 2008/2009 season an even distribution of HRSV A and HRSV B isolates was observed. Although shifts in HRSV subgroup dominance have been reported worldwide, HRSV subgroup A viruses were detected more frequently as found in this study. Studies in Germany, Sweden, Belgium, Argentina, Japan, Kenya and India [Ostlund *et al.*, 2008, Parveen *et al.*, 2006, Reiche and Schweiger, 2009, Sato *et al.*, 2005, Scott *et al.*, 2004, Viegas and Mistchenko, 2005, Zlateva *et al.*, 2007] have reported a predominance of circulating HRSV A between 2004 and 2009.

Two of the eight known genotypes of HRSV A, namely GA2 and GA5, were shown to be circulating in Ireland between 2004/2005 and 2008/2009. The GA5 genotype predominated at 55% compared to the GA2 genotype (45%) in the HRSV subgroup A isolates examined. The identification of GA2 and GA5 genotypes in this study is consistent with the results of long-term studies from Europe and other geographic areas. In countries like Sweden, Belgium, Uruguay, Argentina and the United States genotypes GA2 and GA5 and, at a lower frequency GA7, have

1
2
3 been cocirculating since 1998 [Frabasile *et al.*, 2003, Ostlund *et al.*, 2008, Parveen *et al.*, 2006,
4
5 Sato *et al.*, 2005, Zlateva *et al.*, 2004]. This continuous and predominant circulation of the
6
7 genotypes GA2 and GA5 demonstrate that these strains are stable and have become epidemic in
8
9 many countries.
10

11
12
13
14
15 In this study, HRSV subgroup A amino acid sequences were predicted to be 297 to 298 residues
16
17 in length. This was in agreement with previous studies in Uruguay, Argentina and Germany
18
19 where these changes in protein length were also observed and they correlated with the different
20
21 genotypes [Frabasile *et al.*, 2003, Reiche and Schweiger, 2009]. It was found that G proteins of
22
23 297 amino acids grouped into genotypes GA1 and GA2, whereas G proteins of 298 amino acids
24
25 in length represented the GA3 and GA5 genotypes [Frabasile *et al.*, 2003]. The results in the
26
27 current study supported this observation, however, since several genotypes can have the same
28
29 number of amino acid residues, sequence analysis is an absolute requirement for confirmation of
30
31 HRSV A genotype.
32
33
34
35
36
37
38

39 Genotype-specific amino acid substitutions facilitated differentiation between genotypes GA2
40
41 and GA5 in this study. Two were identified in the GA2 genotypes (S269T and P289S) while
42
43 seven were observed in the GA5 isolates (V225A, S250N, N251S, P274T, V279I, T295I and
44
45 R297D). All of these substitutions were identified previously in the corresponding genotypes in
46
47 a study of German isolates examined from the 1998 season to the 2007 season (Reiche &
48
49 Schweiger, 2009). Interestingly, some of these substitutions were also observed in strains from
50
51 Argentina, South Africa and New York [Galiano *et al.*, 2005, Peret *et al.*, 1998, Venter *et al.*,
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

2001]. This implies that there is a global circulation of HRSV strains over different winter seasons.

The relatively new genotype of HRSV B genotype, BA, was the only one identified in the current study. The BA strains had three G protein lengths of 312, 315 and 319 amino acids reflecting the 60-nucleotide duplication in the G protein gene. Similarly, G protein lengths of 312 [Zlateva *et al.*, 2005], 315 [Trento *et al.*, 2003], 317 [Zlateva *et al.*, 2005] and 319 [Kuroiwa *et al.*, 2005, Zlateva *et al.*, 2005] have been previously reported. Analysis of the duplicated regions of the Irish HRSV B isolates revealed substitutions in the amino acid sequences when compared to the originally described BA viruses, supporting the view that mutations have accumulated over time and that positive selection influenced the evolution of HRSV [Trento *et al.*, 2006]. The BA genotype has been reported in several countries worldwide since its emergence, Japan [Kuroiwa *et al.*, 2005, Nagai *et al.*, 2004, Sato *et al.*, 2005], China [Zhang *et al.*, 2010], Kenya [Scott *et al.*, 2004], Belgium [Zlateva *et al.*, 2005], Canada, Brazil the United Kingdom and the United States [Trento *et al.*, 2006]. This rapid and global spread implies that the BA genotype of the HRSV B virus may have a selective advantage over other circulating strains. It can be hypothesised that the 60-nucleotide duplication in this genotype may result in an antigenic change in the G protein and avoidance of host immune responses which would result in a selective advantage. However, it is possible that there may be mutations elsewhere in the genome that have conferred on these viruses more replicative capacity than other HRSV B genotypes.

1
2
3 The frequency and pattern of glycosylation sites identified in this study was different between
4 the HRSV A and HRSV B subgroups. One to four potential (three in the majority of isolates) N-
5 glycosylation sites were identified in Irish HRSV A isolates while two to three potential sites
6 were identified in the HRSV B isolates. More extensive N-glycosylation among the HRSV A
7 subgroup of isolates may influence antibody binding to the protein [Cane *et al.*, 1991]. This in
8 turn, may explain the greater antigenic variability observed for the G glycoprotein [Garcia-
9 Barreno *et al.* 1989]. The predicted O-glycosylation sites in the C-terminal region of the G
10 protein were identified in both HRSV A and HRSV B isolates where there were 25 to 34 and 40
11 to 43 identified, respectively. The number of sites identified in the HRSV A isolates in this
12 study are very similar to the number identified in HRSV A isolates in another study in Germany
13 which varied from 22 to 34 [Reiche and Schweiger, 2009]. Furthermore, the greater number of
14 potential sites observed in the HRSV B BA genotype isolates may be due to the 60-nucleotide
15 duplication in the G protein where 16 to 19 potential sites were identified among the duplicated
16 region alone. Similarly, these additional O-glycosylation residues have also been identified in
17 the duplicated region in a previous study [Zlateva *et al.*, 2005]. Therefore, these additional O-
18 linked glycosylation residues in the BA genotype may influence the expression of some of the
19 antigenic epitopes by either masking the antigenic sites or contributing to antibody recognition,
20 thus giving these strains an evolutionary advantage over other HRSV subgroup B viruses. In
21 conclusion, the number and location of the glycosylation sites in the G protein of the HRSV
22 virus may cause a long lasting circulation of certain genotypes by avoiding the neutralization by
23 pre-existing antibodies in the host. Further work is however necessary to understand the
24 functional importance of glycosylation sites in the G protein of HRSV which may be important
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 in predicting possible changes in virulence and immune responses to the HRSV virus in future
4
5 vaccination studies.
6
7
8
9

10 The average dS/dN ratios for the HRSV A and HRSV B isolates in this study were 1.59 and 2.04,
11
12 respectively, which implied neutral selection pressure on the variable region of the G protein for
13
14 genotypes GA2, GA5 and BA in this study. Ratios of greater than or equal to one (≥ 1)
15
16 demonstrate a high abundance of synonymous mutations [Reiche and Schweiger, 2009].
17
18 However, when the individual genotypes of the RSV A isolates were compared with each other,
19
20 the average dS/dN ratios were calculated to be 1.64 and 3.13 for GA2 and GA5 isolates,
21
22 respectively. A study in Germany found that there was neutral selection pressure on the GA5
23
24 isolates which were found to have an average dS/dN value of 4.34 similar to a dS/dN of 3.13 in
25
26 the current study, but a positive selection pressure on the GA2 isolates with a dS/dN of 1.18
27
28 [Reiche and Schweiger, 2009]. However, in another study in India, the average dN/dS ratio was
29
30 1.16 and the dS/dN ratio was 1.37 for their HRSV subgroup A isolates and they report that these
31
32 values suggest positive and negative selective pressures, respectively [Parveen *et al.*, 2006].
33
34 These results are similar to the results of isolates in the present study as the dS/dN ratios appear
35
36 to be lower for the GA2 isolates (1.64) than those of GA5 isolates. The observation that a dS/dN
37
38 of 1.16 was evidence of positive selection [Reiche and Schweiger, 2009] but a dS/dN of 1.37
39
40 was evidence of neutral selection [Parveen *et al.*, 2006] implies that the dS/dN of 1.64 for the
41
42 GA2 isolates in this study indicated neutral selection. In this study, the dS/dN for the HRSV B
43
44 isolates was 2.04. No reports were found to compare this value with for other HRSV B,
45
46 genotype BA samples in other countries but this dS/dN value does suggest that there was a
47
48 neutral selection pressure in these isolate as the value is greater than one. However, due to the
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 small sample number used in this study, further work on a larger sample number from other
4
5 population groups is needed to confirm this value. Further worldwide data on the dS/dN values
6
7 of this genotype would also be beneficial in determining the true selection pressure on this
8
9 genotypes.
10
11

12
13
14
15 The use of pre-emptive aerosolised ribavirin treatment in this study appeared to be effective in
16
17 preventing progression from upper to lower respiratory tract infection with HRSV. Most samples
18
19 were collected from adults with HRSV upper respiratory tract infection. Those patients at high
20
21 risk of progression to HRSV pneumonia were treated pre-emptively with aerosolised ribavirin 2g
22
23 every 8 hours for 5 days. None of this cohort progressed to lower respiratory tract infection,
24
25 supporting findings of a previous study (Boeckh M et al Clinical Infectious Diseases 2007;
26
27 44(2); 245-9 Randomized controlled multicenter trial of aerosolized ribavirin for respiratory
28
29 syncytial virus upper respiratory tract infection in hematopoietic cell transplant recipients).
30
31 However, three patients who presented to hospital with HRSV pneumonia were admitted to the
32
33 intensive care unit (ICU) and survived following treatment with intravenous ribavirin and
34
35 supportive care. Although all the samples examined in this study were collected from adult
36
37 patients, these infections were community-acquired reflecting strains circulating in the
38
39 community, including among children who are recognised as potential sources of HRSV. The
40
41 molecular epidemiology of the HRSV isolates circulating in Ireland, based on this study, only
42
43 reflected isolates from a Dublin hospital from the 2004/2005 to 2008/2009 seasons. The adult
44
45 HSCT facility in St James's Hospital is a national unit with at least 50% of patients referred from
46
47 outside of Dublin.
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 In summary, HRSV A viruses were observed to be marginally predominant over HRSV B in
4 Ireland over five winter season from 2004/2005 to 2008/2009 and both the HRSV subgroups
5 were observed to be cocirculating in each of the seasons, except in 2008/2009. Two genotypes,
6 GA2 and GA5 were identified among the HRSV A isolates and have been reported to be
7 circulating in Europe for more than 20 years [Rafiefard *et al.*, 2004]. All of the HRSV B isolates
8 belonged to genotype BA. Phylogenetic analysis of the C-terminal regions of the G protein of
9 the two HRSV subgroups has shown that closely related strains of HRSV can be found in
10 different parts of the world at different times and the same genotypes have been circulating for a
11 long time.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- 1
2
3
4
5
6
7 Avetisyan, G., Mattsson, J., Sparrelid, E., Ljungman, P. 2009. Respiratory syncytial virus
8 infection in recipients of allogeneic stem-cell transplantation: a retrospective study of the
9 incidence, clinical features, and outcome. *Transplantation* 88:1222-6.
- 10
11
12
13
14 Bukreyev, A., Whitehead, S.S., Murphy, B.R., Collins, P.L. 1997. Recombinant respiratory
15 syncytial virus from which the entire SH gene has been deleted grows efficiently in cell
16 culture and exhibits site-specific attenuation in the respiratory tract of the mouse. *J Virol*
17 71:8973-82.
- 18
19
20
21
22
23 Cane, P.A. 2001. Molecular epidemiology of respiratory syncytial virus. *Rev Med Virol* 11:103-
24 16.
- 25
26
27
28 Cane, P.A., Matthews, D.A., Pringle, C.R. 1991. Identification of variable domains of the
29 attachment (G) protein of subgroup A respiratory syncytial viruses. *J Gen Virol* 72 (Pt
30 9):2091-6.
- 31
32
33
34
35 Cane, P.A., Pringle, C.R. 1991. Respiratory syncytial virus heterogeneity during an epidemic:
36 analysis by limited nucleotide sequencing (SH gene) and restriction mapping (N gene). *J*
37 *Gen Virol* 72 (Pt 2):349-57.
- 38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Chen, M.D., Vazquez, M., Buonocore, L., Kahn, J.S. 2000. Conservation of the respiratory
syncytial virus SH gene. *J Infect Dis* 182:1228-33.
- Collins, P.L., Mottet, G. 1993. Membrane orientation and oligomerization of the small
hydrophobic protein of human respiratory syncytial virus. *J Gen Virol* 74 (Pt 7):1445-50.
- Collins, P.L., Olmsted, R.A., Johnson, P.R. 1990. The small hydrophobic protein of human
respiratory syncytial virus: comparison between antigenic subgroups A and B. *J Gen*
Virol 71 (Pt 7):1571-6.

- 1
2
3
4
5 da Silva, L.H., Spilki, F.R., Riccetto, A.G., de Almeida, R.S., Baracat, E.C., Arns, C.W. 2008.
6
7 Genetic variability in the G protein gene of human respiratory syncytial virus isolated
8
9 from the Campinas metropolitan region, Brazil. *J Med Virol* 80:1653-60.
10
11
12 Falsey, A.R., Walsh, E.E. 2000. Respiratory syncytial virus infection in adults. *Clin Microbiol*
13
14 *Rev* 13:371-84.
15
16
17 Frabasile, S., Delfraro, A., Facal, L., Videla, C., Galiano, M., de Sierra, M.J., Ruchansky, D.,
18
19 Vitureira, N., Berois, M., Carballal, G., Russi, J., Arbiza, J. 2003. Antigenic and genetic
20
21 variability of human respiratory syncytial viruses (group A) isolated in Uruguay and
22
23 Argentina: 1993-2001. *J Med Virol* 71:305-12.
24
25
26 Galiano, M.C., Palomo, C., Videla, C.M., Arbiza, J., Melero, J.A., Carballal, G. 2005. Genetic
27
28 and antigenic variability of human respiratory syncytial virus (groups a and b) isolated
29
30 over seven consecutive seasons in Argentina (1995 to 2001). *J Clin Microbiol* 43:2266-
31
32 73.
33
34
35 Garcia-Barreno, B., Paloma, C., Penas, C., Delgado, C., Perez-Brena, P., Melero, J. A. 1989.
36
37 Marked differences in the antigenic structure of human respiratory syncytial virus F and
38
39 G glycoproteins. *J virol* 63:925-932.
40
41
42 Hall, C.B. 2001. Respiratory syncytial virus and parainfluenza virus. *N Engl J Med* 344:1917-28.
43
44
45
46 Hansen, J.E., Lund, O., Engelbrecht, J., Bohr, H., Nielsen, J.O. 1995. Prediction of O-
47
48 glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc:polypeptide
49
50 N-acetylgalactosaminyltransferase. *Biochem J* 308 (Pt 3):801-13.
51
52
53 HPSC, Health Protection Surveillance Centre.
54
55
56
57
58
59
60

- 1
2
3 Johnson, P.R., Spriggs, M.K., Olmsted, R.A., Collins, P.L. 1987. The G glycoprotein of human
4 respiratory syncytial viruses of subgroups A and B: extensive sequence divergence
5 between antigenically related proteins. *Proc Natl Acad Sci U S A* 84:5625-9.
6
7
8
9
10
11
12 Karron, R.A., Buonagurio, D.A., Georgiu, A.F., Whitehead, S.S., Adamus, J.E., Clements-Mann,
13 M.L., Harris, D.O., Randolph, V.B., Udem, S.A., Murphy, B.R., Sidhu, M.S. 1997.
14 Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral
15 replication in vitro: clinical evaluation and molecular characterization of a cold-passaged,
16 attenuated RSV subgroup B mutant. *Proc Natl Acad Sci U S A* 94:13961-6.
17
18
19
20
21
22
23 Korber, B. 2000. HIV Signature and Sequence Variation Analysis. *Computational Analysis of*
24 *HIV Molecular Sequences*. Allen G. Rodrigo and Gerald H. Learn, eds. Dordrecht,
25 Netherlands: Kluwer Academic Publishers Chapter 4:55-72.
26
27
28
29
30 Kuo, L., Grosfeld, H., Cristina, J., Hill, M.G., Collins, P.L. 1996. Effects of mutations in the
31 gene-start and gene-end sequence motifs on transcription of monocistronic and
32 dicistronic minigenomes of respiratory syncytial virus. *J Virol* 70:6892-901.
33
34
35
36
37 Kuroiwa, Y., Nagai, K., Okita, L., Yui, I., Kase, T., Nakayama, T., Tsutsumi, H. 2005. A
38 phylogenetic study of human respiratory syncytial viruses group A and B strains isolated
39 in two cities in Japan from 1980-2002. *J Med Virol* 76:241-7.
40
41
42
43
44 Lim, C.S., Kumarasinghe, G., Chow, V.T. 2003. Sequence and phylogenetic analysis of SH, G,
45 and F genes and proteins of Human respiratory syncytial virus isolates from Singapore.
46 *Acta Virol* 47:97-104.
47
48
49
50
51 Matheson, J.W., Rich, F.J., Cohet, C., Grimwood, K., Huang, Q.S., Penny, D., Hendy, M.D.,
52 Kirman, J.R. 2006. Distinct patterns of evolution between respiratory syncytial virus
53 subgroups A and B from New Zealand isolates collected over thirty-seven years. *J Med*
54 *Virol* 78:1354-64.
55
56
57
58
59
60

- 1
2
3
4
5 Melero, J.A., Garcia-Barreno, B., Martinez, I., Pringle, C.R., Cane, P.A. 1997. Antigenic
6 structure, evolution and immunobiology of human respiratory syncytial virus attachment
7 (G) protein. *J Gen Virol* 78 (Pt 10):2411-8.
8
9
10
11
12 Nagai, K., Kamasaki, H., Kuroiwa, Y., Okita, L., Tsutsumi, H. 2004. Nosocomial outbreak of
13 respiratory syncytial virus subgroup B variants with the 60 nucleotides-duplicated G
14 protein gene. *J Med Virol* 74:161-5.
15
16
17
18
19 Nei, M., Gojobori, T. 1986. Simple methods for estimating the numbers of synonymous and
20 nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3:418-26.
21
22
23
24 Ostlund, M.R., Lindell, A.T., Stenler, S., Riedel, H.M., Wirtgart, B.Z., Grillner, L. 2008.
25 Molecular epidemiology and genetic variability of respiratory syncytial virus (RSV) in
26 Stockholm, 2002-2003. *J Med Virol* 80:159-67.
27
28
29
30
31
32 Parveen, S., Sullender, W.M., Fowler, K., Lefkowitz, E.J., Kapoor, S.K., Broor, S. 2006. Genetic
33 variability in the G protein gene of group A and B respiratory syncytial viruses from
34 India. *J Clin Microbiol* 44:3055-64.
35
36
37
38
39 Peret, T.C., Hall, C.B., Hammond, G.W., Piedra, P.A., Storch, G.A., Sullender, W.M., Tsou, C.,
40 Anderson, L.J. 2000. Circulation patterns of group A and B human respiratory syncytial
41 virus genotypes in 5 communities in North America. *J Infect Dis* 181:1891-6.
42
43
44
45
46 Peret, T.C., Hall, C.B., Schnabel, K.C., Golub, J.A., Anderson, L.J. 1998. Circulation patterns of
47 genetically distinct group A and B strains of human respiratory syncytial virus in a
48 community. *J Gen Virol* 79 (Pt 9):2221-9.
49
50
51
52
53 Rafiefard, F., Johansson, B., Teclé, T., Orvell, C. 2004. Molecular epidemiology of respiratory
54 syncytial virus (RSV) of group A in Stockholm, Sweden, between 1965 and 2003. *Virus*
55 *Res* 105:137-45.
56
57
58
59
60

- 1
2
3
4
5 Reiche, J., Schweiger, B. 2009. Genetic variability of group A human respiratory syncytial virus
6 strains circulating in Germany from 1998 to 2007. *J Clin Microbiol* 47:1800-10.
7
8
9
10 Roca, A., Loscertales, M.P., Quinto, L., Perez-Brena, P., Vaz, N., Alonso, P.L., Saiz, J.C. 2001.
11 Genetic variability among group A and B respiratory syncytial viruses in Mozambique:
12 identification of a new cluster of group B isolates. *J Gen Virol* 82(Pt 1):103-11.
13
14
15
16
17 Sato, M., Saito, R., Sakai, T., Sano, Y., Nishikawa, M., Sasaki, A., Shobugawa, Y., Gejyo, F.,
18 Suzuki, H. 2005. Molecular epidemiology of respiratory syncytial virus infections among
19 children with acute respiratory symptoms in a community over three seasons. *J Clin*
20 *Microbiol* 43:36-40.
21
22
23
24
25
26 Scott, P.D., Ochola, R., Ngama, M., Okiro, E.A., Nokes, D.J., Medley, G.F., Cane, P.A. 2004.
27 Molecular epidemiology of respiratory syncytial virus in Kilifi district, Kenya. *J Med*
28 *Virol* 74:344-54.
29
30
31
32
33 Techaarpornkul, S., Barretto, N., Peeples, M.E. 2001. Functional analysis of recombinant
34 respiratory syncytial virus deletion mutants lacking the small hydrophobic and/or
35 attachment glycoprotein gene. *J Virol* 75:6825-34.
36
37
38
39
40 Tozer, S., Whiley, D.M., Lum, G., Rawlinson, W.D., Sammels, L.M., Sloots, T.P., Nissen, M.D.
41 2008. Identification of Australian human respiratory syncytial virus strains containing a
42 60-nucleotide duplication within the G glycoprotein gene. *Pathology* 40:632-5.
43
44
45
46
47 Trento, A., Galiano, M., Videla, C., Carballal, G., Garcia-Barreno, B., Melero, J.A., Palomo, C.
48 2003. Major changes in the G protein of human respiratory syncytial virus isolates
49 introduced by a duplication of 60 nucleotides. *J Gen Virol* 84(Pt 11):3115-20.
50
51
52
53
54 Trento, A., Viegas, M., Galiano, M., Videla, C., Carballal, G., Mistchenko, A.S., Melero, J.A.
55 2006. Natural history of human respiratory syncytial virus inferred from phylogenetic
56
57
58
59
60

- 1
2
3 analysis of the attachment (G) glycoprotein with a 60-nucleotide duplication. *J Virol*
4 80:975-84.
5
6
7
8
9 Tripp, R.A., Jones, L., Anderson, L.J. 2000. Respiratory syncytial virus G and/or SH
10 glycoproteins modify CC and CXC chemokine mRNA expression in the BALB/c mouse.
11 *J Virol* 74:6227-9.
12
13
14
15
16 Venter, M., Collinson, M., Schoub, B.D. 2002. Molecular epidemiological analysis of
17 community circulating respiratory syncytial virus in rural South Africa: Comparison of
18 viruses and genotypes responsible for different disease manifestations. *J Med Virol*
19 68:452-61.
20
21
22
23
24
25 Venter, M., Madhi, S.A., Tiemessen, C.T., Schoub, B.D. 2001. Genetic diversity and molecular
26 epidemiology of respiratory syncytial virus over four consecutive seasons in South
27 Africa: identification of new subgroup A and B genotypes. *J Gen Virol* 82(Pt 9):2117-24.
28
29
30
31
32 Viegas, M., Mistchenko, A.S. 2005. Molecular epidemiology of human respiratory syncytial
33 virus subgroup A over a six-year period (1999-2004) in Argentina. *J Med Virol* 77:302-
34 10.
35
36
37
38
39 Wertz, G.W., Collins, P.L., Huang, Y., Gruber, C., Levine, S., Ball, L.A. 1985. Nucleotide
40 sequence of the G protein gene of human respiratory syncytial virus reveals an unusual
41 type of viral membrane protein. *Proc Natl Acad Sci U S A* 82:4075-9.
42
43
44
45
46 Wertz, G.W., Krieger, M., Ball, L.A. 1989. Structure and cell surface maturation of the
47 attachment glycoprotein of human respiratory syncytial virus in a cell line deficient in O
48 glycosylation. *J Virol* 63:4767-76.
49
50
51
52
53 Whimbey, E., Champlin, R.E., Couch, R.B., Englund, J.A., Goodrich, J.M., Raad, I., Przepioraka,
54 D., Lewis, V.A., Mirza, N., Yousuf, H., Tarrand, J.J., Bodey, G.P. 1996. Community
55
56
57
58
59
60

1
2
3 respiratory virus infections among hospitalized adult bone marrow transplant recipients.
4 Clin Infect Dis 22:778-82.
5
6
7

8
9 Whitehead, S.S., Bukreyev, A., Teng, M.N., Firestone, C.Y., St Claire, M., Elkins, W.R.,
10 Collins, P.L., Murphy, B.R. 1999. Recombinant respiratory syncytial virus bearing a
11 deletion of either the NS2 or SH gene is attenuated in chimpanzees. J Virol 73:3438-42.
12
13

14
15
16 Zhang, Z.Y., Du, L.N., Chen, X., Zhao, Y., Liu, E.M., Yang, X.Q., Zhao, X.D. 2010. Genetic
17 Variability of Respiratory Syncytial Viruses Prevalent in Southwestern China from 2006
18 - 2009: Emergence of RSV BA as Dominant Strains. J Clin Microbiol. 48:1201-7.
19
20

21
22
23 Zlateva, K.T., Lemey, P., Moes, E., Vandamme, A.M., Van Ranst, M. 2005. Genetic variability
24 and molecular evolution of the human respiratory syncytial virus subgroup B attachment
25 G protein. J Virol 79:9157-67.
26
27

28
29
30 Zlateva, K.T., Lemey, P., Vandamme, A.M., Van Ranst, M. 2004. Molecular evolution and
31 circulation patterns of human respiratory syncytial virus subgroup a: positively selected
32 sites in the attachment G glycoprotein. J Virol 78:4675-83.
33
34

35
36
37 Zlateva, K.T., Vijgen, L., Dekeersmaeker, N., Naranjo, C., Van Ranst, M. 2007. Subgroup
38 prevalence and genotype circulation patterns of human respiratory syncytial virus in
39 Belgium during ten successive epidemic seasons. J Clin Microbiol 45:3022-30.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60