



Genotyping of Type A Human Respiratory Syncytial Virus Based on Direct F Gene Sequencing

Daifullah Al Aboud ¹, Nora M. Al Aboud ², Mater I. R. Al-Malky ³ and Ahmed S. Abdel-Moneim ^{1,4,*}

- ¹ College of Medicine, Taif University, Al-Taif 21944, Saudi Arabia; d.alaboud@tu.edu.sa
- ² Department of Biology, College of Applied Sciences, Umm Al Qura University, Makkah 21955, Saudi Arabia; nmaboud@uqu.edu.sa
- ³ Hospital of Paediatrics, Al-Taif 21944, Saudi Arabia; mater1386@hotmail.com
- ⁴ Department of Virology, Faculty of Veterinary Medicine, Beni-Suef University, Beni-Suef 62511, Egypt
- * Correspondence: asa@bsu.edu.eg or asa@tu.edu.sa

Received: 23 March 2019; Accepted: 14 May 2019; Published: 20 May 2019



Abstract: *Background and objectives:* The human respiratory syncytial virus (hRSV) is among the important respiratory pathogens affecting children. Genotype-specific attachment (G) gene sequencing is usually used to determine the virus genotype. The reliability of the fusion (F) gene vs. G gene genotype-specific sequencing was screened. *Materials and Methods:* Archival RNA from Saudi children who tested positive for hRSV-A were used. Samples were subjected to a conventional one-step RT-PCR for both F and G genes and direct gene sequencing of the amplicons using the same primer sets. Phylogeny and mutational analysis of the obtained sequences were conducted. *Results:* The generic primer set succeeded to amplify target gene sequences. The phylogenetic tree based on partial F gene sequencing resulted in an efficient genotyping of hRSV-A strains equivalent to the partial G gene genotyping method. NA1, ON1, and GA5 genotypes were detected in the clinical samples. The latter was detected for the first time in Saudi Arabia. Different mutations in both conserved and escape-mutant domains were detected in both F and G. Conclusion: It was concluded that a partial F gene sequence can be used efficiently for hRSV-A genotyping.

Keywords: hRSV; F gene; G gene; children; respiratory diseases; genotyping; Saudi Arabia

1. Introduction

Human respiratory syncytial virus (hRSV) causes considerable respiratory distress with variable severity in infancy and early childhood [1,2]. It is also responsible for the induction of respiratory illness in elderly and immunocompromised patients [3], in addition to being a common nosocomial pathogen [4]. The virus belongs to the family *Pneumoviridae*, genus *Orthopneumovirus*, and possesses a single-stranded negative-sense RNA genome that encodes eleven proteins [5], including both attachment (G) and fusion (F) envelope proteins that work together to attach to the target cell membrane by the G protein [6], while fusion of the viral and cell membranes occurs through the action of the F glycoprotein [7].

The G glycoprotein is a type II surface protein which is highly glycosylated and possesses a considerable degree of nucleotide variability [8]. It possesses two hypervariable regions (HVR1 and HVR2); the HVR2 in the C-terminal region is used to screen the G gene variability of different genotypes [9,10]. Five N-glycosylation sites were detected in HVR2 of the hRSV-A attachment protein [11]. The F protein is more conserved than the G glycoprotein and it is translated as an *F0* precursor that is cleaved twice, resulting in disulfide-linked *F1* (aa 137–574) and *F2* (aa 1–109) subunits along with a short peptide, pep27 (aa 110–136). *F2* possesses heptad repeat C (HRC), while *F1* possesses,



at its N terminal, a hydrophobic fusion peptide (FP) followed by two heptad repeats, A and B (HRA and HRB). HRA, HRB, and HRC are essential for envelope fusion to the host cell membrane [12,13]. *F2* possesses five N-glycosylation sites, while HRB, part of *F1*, possesses a single N-glycosylation site; however, *F2* contains more conserved sequence than *F1*, which is characterized by a high homology among different hRSV genotypes [14]. The highly conserved nature of *F2* qualifies it for being a potential target for diagnostic assays.

The hRSV is classified into two major groups: hRSV A and hRSV B [15]. To date, 20 genotypes of hRSV-A and 36 genotypes of hRSV-B are known to exist based on the sequence variation of the G gene [16,17]. Although both hRSV types circulate worldwide, type A was found to be dominant in certain countries [18,19]. Novel genotypes may appear and could replace prevalent genotypes, and many genotypes may circulate together; however, some new genotypes may dominate [20,21]. In 1999, the B/BA genotype, with a duplication of 60 nucleotides of the G gene, appeared [22]; then, the A/ON1 genotype, with a duplication of 72 nucleotides (nt) in the G gene, appeared [23]. Both genotypes are currently prevalent worldwide. The severity of the disease is dependent on the virus genotype, as genotypes A/GA2, A/GA4, A/ON1, and B/BA were found to be of relatively low virulence in comparison to A/NA1, A/GA3, and A/GA5 [24–26].

Typing of hRSV is mainly based on direct G sequencing of sequences flanking HVR2; however, genotype-specific amplification is needed. The F gene could constitute an alternative method for hRSV genotyping; it was found to successfully genotype RSV strains in comparison to the G gene and resulted in the same phylogenetic clusters based on a full genome sequence of the F gene [27]. Accordingly, the aim of the current study was to screen the efficiency of genotyping hRSV-A based on gene sequencing of *F*2 and the first part of *F*1 (FP and HRA) and compare it with genotype-specific partial G gene sequencing. Genotyping and sequence analysis of the Saudi hRSV-A strains is another objective of the current study.

2. Materials and Methods

2.1. Ethical Statement

Ethical release was obtained from both by the College of Medicine, Taif University (TU/1652/1433/1, approved on 29th April 2011), and also from the hospital of Paediatrics ethical committee based on the human subjects protection guidelines. The first part of the study was previously published [28]. The guardians of the children involved in the study signed informed written consent about the study.

2.2. Samples

A total of 59 archived RNA/DNA samples that showed positive real-time RT-PCR results for hRSV-A, determined in our previous study [28], were used as targets for amplification of F and G genes. The RNA samples were stored at -80 °C. All samples were collected previously from children with respiratory distress from hospital of Paediatrics, Taif, Saudi Arabia, during the period from January to May 2012. The viral load in the samples (n = 56) samples were found to be weak, with a range of 1×10^{2} to $1 \times 10^{4.5}$ copies/mL [29].

2.3. One-Step Conventional RT-PCR

An F gene-specific oligonucleotide primer set: Fus-For (1–29) 5'-ATGGAGTTGCCAATCCTCA AAGCAAATGC-3' and Fus-Rev (622–643) 5'-ATATGCTGCAGCTTTGCTTGTT-3', that flanks F2 and the first part of F1 (FP and HRA), was designed based on hRSV-A-GZ08-0 (KP218910). G gene genotype-specific primers for ON1/NA1 and GA5 that flank HVR2 in the C-terminal region of the G gene were also designed and used for amplification of the partial G gene based on the result of F gene sequencing. G gene-specific primer sets were as follows: ON1/NA1 (365-384)-For 5'-CTGAGTC AACCCCACAATCC-3', ON1/NA1-Rev (24–43 G–F intergene UTR) 5'-ATTTGGTCATG GCTTTTTGC-3' based on hRSV-A-GZ08-0 (KP218910) and GA5-For (374–395) 5'-TCCT

GCAATCTACAACAGTCAA -3' and GA5-Rev-(867–889) 5'-CTGTTATGTTGGATGGAGATG-GA-3' based on RSVA/Homo sapiens/ITA/120/2009 (KF826832) were designed. Reverse transcription was conducted using GoTaq[®] 1-Step RT-PCR (Promega, Southampton, UK) that was based on GoScript[™] Reverse Transcriptase. Briefly, the program was adjusted for reverse transcription at 45 °C for 45 min, followed by an initial denaturation step at 95 °C for 5 min and 35 cycles of 95 °C for 30 s, 50 °C for 30 s, and 72 °C for 1 min. This was followed by a final 10 min elongation step at 72 °C. The RT-PCR amplicons were subjected to 1.5% agarose gel electrophoresis.

2.4. Direct Sequencing and Gene Sequence Analysis

Amplicons with the expected size were excised from the gel and then purified using a gel-extraction kit (Koma Biotek, Seoul, Korea). The purified amplicons were used as templates for direct sequencing using the same primers as those used for the F and G genes' amplification. Sequencing was performed commercially using the BigDye v.3.1 Applied Biosystems (Foster City, CA, USA) kit according to the manufacturer's protocol. Raw sequences were visualized and analyzed using MEGA 5.2. Homology BLASTn searches from different strains were performed using highly similar sequences (megablast) against published hRSV sequences in the GenBank databases using default algorithm parameters. The gene sequences of both F and G genes were deposited in GenBank under the accession numbers: F (KU924011-KU924023) and G (MK182708-MK182720). CLUSTAL W multisequence analysis was performed and the phylogenetic tree was constructed using the maximum likelihood statistical method with 1000 bootstrap replicates and Tamura–Nei model substitution model. Nucleotide variability among different strains was calculated using the CLUSTALW available in GenomeNet Database [30].

2.5. Deduced Amino Acid Sequence and Sequence Analysis

Deduced amino acid sequences of both F and G amino acid sequences were compared using MEGA 5.2. Amino acid substitutions at different functional and structural protein fragments were analysed. The polyphen (polymorphism phenotyping) prediction webtool [31] was used to screen the potential effect of amino acid substitutions among strains on the function.

3. Results

3.1. Direct Sequencing and Phylogenetic Analysis

Both generic primers for F and G genes were successful in amplification of the strains, showing relatively high viral load for 13 of the 59 samples ($\geq 1 \times 10^{3.5}$ copies/mL). The F gene domains of the hRSV-A isolates possess about 10% nucleotide variability. The signal peptide, however, showed higher nucleotide variability. Partial F gene sequencing revealed that the Saudi strains in the current study belong to the NA1 genotype (6/13, 46.15%), ON1 (5/13; 38.46%), and GA5 genotype (2/13; 15.38%) (Figure 1a). Partial G gene sequencing using strain-specific primers for these strains confirmed the correct strain classification obtained by using F gene sequencing, and similar clustering of the hRSV strains was obtained using both genes (Figure 1a,b). The F gene domains of the hRSV-A isolates possess about 10% nucleotide variability. The signal peptide and transmembrane domains, however, showed nucleotide variability to be higher than detected in the HVRs of the G gene (data not shown).



Figure 1. Phylogenetic trees of the nucleotide sequences of partial fusion (F) and attachment (G) genes of human respiratory syncytial virus A (hRSV-A) strains from Saudi Arabia in comparison to relevant published strains. (a) F gene; (b) G gene. The maximum likelihood tree was constructed using MEGA 5.2 freeware. Strains detected in the current study are shown in blue color. Only bootstrap values above 70% are shown.

3.2. Deduced Amino Acid Sequence and Sequence Analysis

Five N-glycosylation sites, N27, N70, N116, N120, and N126, were found to be conserved among all the thirteen Saudi strains, except for strain Taif-103, that has lost one of the N-glycosylations at N120 (Figure 2). The fusion protein of Taif-20, Taif-23, Taif-110, and Taif-112 showed S-to-N substitution at position 105 just before the first cleavage (Figure 2), with no potential effect on the function as demonstrated by the polymorphism phenotyping prediction tool (data not shown). Seven amino acid substitutions—L3F, T11I, I19T, T13A, L15F, A16T, and I19T—were detected in the signal peptide, and three—T120A, T125N, and V127I—amino acid substitutions were detected in the p27 domain (Figure 2).

A duplication of amino acids (QEETLHSTTSEGYLSPSQVYTTS) was found in two regions within the G protein which reflects an 72 nt insert characteristic to the ON1 genotype (Figure 3). L298P, V303A, and Y304H amino acid substitutions were detected in all ON1 Saudi strains, Y273N was detected in Taif 23, and L274P was detected in Taif-24 and Taif-103 strains (Figure 3). The P310L substitution, equivalent to P286L in viruses that have no duplication, was recorded in 10 out of the 13 Saudi strains. There is an existence of accumulated signature coding changes in the epitope regions of the G protein (Figure 3). Variations among the N-glycosylation site of HVR2 of the Saudi strains' attachment proteins were detected. Interestingly, all Saudi NA1 and ON1 strains except Taif-24 and Taif-103 showed N237D amino acid substitution that resulted in loss of the N237 N-glycosylation site. N250 was found only in Saudi GA5 strains, while N251 was detected in 2/5 ON1 and 6/6 of the NA1 Saudi strains. N273 was detected only in Saudi GA5 strains, as well as a single ON1 Saudi strain. N294 (N318 in ON1 genotypes) was detected in the majority of Saudi ON1 and NA1 strains, but none of the Saudi GA5 strains or the A2 prototype strain (Figure 3).

	Signal Pe	eptide							
	13	10 12	20 1	30	40	50	60	70	80
JX627336 GN435/11	MELPILKTNA	ITTILAAVTL	CFASSONITE	EFYOSTCSA	V SKGYLSALR	GWYTSVITIE	LSNIKENKC	N GTDAKVKLJ	IK
M114861A2	T A	F		[~]				<u> </u>	
KU924012 Taif-7	A								
KI19240171Taif=20	2		· · · · · · · · · · · · · · · · · · ·						
KU924019/Taif 23	E A		· · · · · · · · · · · · · · · · · · ·					<u></u>	••
KU924010 Tall-25			· · · · · · · · · · · · · · · · · · ·				•••••	<u></u>	••
K0924019 1a11-24			· · · · · · · · · · · · · · · · · · ·				•••••	<u></u>	•••
KU924021 Tair-103		· · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·				•••••	<u></u>	•••
KU924011 Tair-1	A	· · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·			• • • • • • • • • • •	•••••	<u></u>	• •
KU924013 Tait-13	A		· · · · · · · <u>· · ·</u> ·	• • • • • • • • • •	• • • • • • • • • • • • •	• • • • • • • • • • •	••••••	<u></u>	• •
KU924014 Taif-16	A	I.	· · · · · · <u>· · ·</u> ·				•••••	<u></u>	• •
KU924015 Taif-17	A	.II.	· · · · · · <u>· · ·</u> ·				• • • • • • • • • •	<u></u>	• •
KU924016 Taif-18	A	I.	<u></u> .					<u></u>	
KU924020 Taif-47	A								
KU924022 Taif-110	A	A.FTS.						. <u></u>	
KU924023 Taif-112	A	A.FTS.							
								<u> </u>	
					p27		FP		
		90 13	LOO <u> </u>]	110	120	130	140	150	160
JX627336 GN435/11	QELDKYKNAV	TELQLLMQST	PAANSRARRE	LPRFM <u>NYT</u> L	N NTKNT <u>NVT</u> LS	5 KKRKR <mark>r</mark> flGf	LLGVGSAIA	S GIAVSKVLF	AL
M11486 A2			.PT.N	<u></u> .	A.K			v	
KU924012 Taif-7									
KU924017 Taif-20			N						
KU924018 Taif-23			N						
KU924019 Taif-24									
KU924021 Taif-103					A				
KT192/011/Taif-1				····· <u>···</u> ·	· ····································				
KU9240131T-if-13				····· <u>···</u> ·	· · · · · · · · · · · · · · · · · · ·				••
KU924013 Tail 15				····· <u>···</u> ·	• • • • • • • <u>• • •</u> • •	• • • • • • • • • • • • •			••
K0924014 1a11-10				···· <u>···</u> ·	• • • • • • • <u>• • •</u> • •	• • • • • • • • • • • • •			••
K0924015 1a11-17				· · · · · <u>· · · ·</u> ·	• • • • • • • <u>• • • •</u> • •	• • • • • • • • • • • • •			••
KU924016 1a11-18				· · · · · · <u>· · · ·</u> ·	• • • • • • • <u>• • • •</u> • •	• • • • • • • • • • • • •			••
K0924020 Talf-47				· · · · · <u>· · ·</u> ·	· · · · · · · · · · · · · · · · · · ·	• • • • • • • • • • • •			••
KU924022 Taif-110			· · · · N · · · ·	•••• <u>•••</u> •	• • • • • • • • • • • • • • • • • • •	• • • • • • • • • • •			••
KU924023 Taif-112			N		• • • • • • • • • • • • • • • • • • •				••
	11	70 11	80 11	190	1200 1	210			
TV6273361CM435/11	FCEIMUTUCA		T CNCVCVT TC			CCCT			
M11406120	EGEVINATASA	TTO TRIVE V O	19100030113	KV LDLKMI II	, võnnei viivõ	1 2021			
M11486 A2		• • • • • • • • • • •		•••••					
R0924012 Tall-7									
R0924017 Taif-20									
KU924018 Taif-23									
KU924019 Taif-24									
KU924021 Taif-103									
KU924011 Taif-1									
KU924013 Taif-13									
KU924014 Taif-16									
KU924015 Taif-17									
KU924016 Taif-18									
KU924020 Taif-47					Y				
KII9240221Taif=110									
KU924022 Iaii = 110									
RUJZAUZJIIAII-IIZ									

Figure 2. Deduced amino acid sequences of partial fusion proteins from different Saudi hRSV-A strains. Highlighted are the signal peptide (1–22) (blue box), heptad repeat domain C (75–97), cleavage site-1 (CS-1) and cleavage site-2 (CS-2) at Arg109 (NSRARR↓E) and Arg136 (KKRKRR↓F) (red color), p27 (110–136) (green box), fusion peptide (FP) (137–155), and heptad repeat domain A (153–end of the current sequence). The N-glycosylation sites, NXT/S, where X is not a proline, are underlined.

	130	140	150	160	170	180	190	200	210	220
JX627336 GN435/11	TTVKIKNTTT	TQILPSKPTT	KQRQNKPQNK	PNNDFHFEVF	NFVPCSICSN	NPTCWAICKR	IPNKKPGKKT	TTKPTKKPTL	KTTKKDPKPQ	TTKPKEVLTT
M11486 A2	т. <u></u> .	TQ	PS.							SP
MK182709 Taif-7	<u></u> .									
MK182714 Taif-20	<u></u> .							Q		
MK182715 Taif-23	<u></u> .							I		
MK182716 Taif-24	<u></u> .	ML	Q					Q		
MK182718 Taif-103	<u></u> .	MQL	Q					Q		
MK182708 Taif-1		P						I		
MK182710 Taif-13		Q						I		
MK182711 Taif-16		Q						I		
MK182712 Taif-17		Q						RI		
MK182713 Taif-18		Q					•••••	I		
MK182717 Taif-47	s. <u></u> .	Q						RI		
MK182719 Taif-110	T. <u>.S.</u> .	MQL	Q.KP	.KD			s	Q.I	L	AP
MK182720 Taif-112	T. <u>.S.</u> .	MQL	Q.KP	.KD			s	Q.I	L	AP
	1000	1040	1050	10.00	1070	1000	1000	1000	1010	
	230	240	250	260	270	280	290	300	310	
JX627336 GN435/11	230 KPTGKPTI <u>NT</u>	240 <u>T</u> KTNIRTTLL	250 TSNTKGNPEH	260 TSQEETLHST	270 TSEGYLSPSQ	280 VYTTSGQEET	290 LHSTTSEGYL	300 SPSQVYTTSE	310 YLSQSLSSS <u>N</u>	TTK
JX627336 GN435/11 M11486 A2	230 KPTGKPTI <u>NT</u>	240 <u>TKTNIRTTLL</u> I	250 TSNTKGNPEH <u>T</u> L	260 TSQEETLHST MF	270 TSEGYLSPSQ SNP	280 VYTTSGQEET .S	290 LHSTTSEGYL	300 SPSQVYTTSE	310 YLSQSLSSS <u>N</u> .PPS.PP.	TTK .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7	230 KPTGKPTINT EED.	240 <u>TKTNIRTTLL</u> I Q.	250 TSNTKGNPEH <u>T</u> L <u>T</u> L	260 TSQEETLHST MF	270 TSEGYLSPSQ SNP	280 VYTTSGQEET .S	290 LHSTTSEGYL	300 SPSQVYTTSE 	310 YLSQSLSSS <u>N</u> .PPS.PP.	TTK .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20	230 KPTGKPTINT ED. ED.	240 <u>T</u> KTNIRTTLL Q.	250 TSNTKGNPEH <u>T</u> L <u>T</u>	260 TSQEETLHST MF	270 TSEGYLSPSQ SNP	280 VYTTSGQEET .S	290 LHSTTSEGYL P P	300 SPSQVYTTSE AH	310 YLSQSLSSSN .PPS.PP. P	TTK .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23	230 KPTGKPTI <u>NT</u> ED. ED. ED.	240 <u>T</u> KTNIRTTLL Q.	250 TSNTKGNPEH <u>.T</u> L <u>.T</u> <u>Y</u> .T	260 TSQEETLHST MF	270 TSEGYLSPSQ SNP	280 VYTTSGQEET .S	290 LHSTTSEGYL P P P	300 SPSQVYTTSE 	310 YLSQSLSSSN .PPS.PP. P P P Y	<u>TT</u> K .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23 MK182716 Taif-24	230 KPTGKPTI <u>NT</u> ED. ED. ED.	240 <u>T</u> KTNIRTTLL Q. Q.	250 TSNTKGNPEH 	260 TSQEETLHST MF	270 TSEGYLSPSQ SNP 	280 VYTTSGQEET .S	2 90 LHSTTSEGYL P P P P	300 SPSQVYTTSE 	310 YLSQSLSSSN .PPS.PP. P P PY	<u>TT</u> K .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23 MK182716 Taif-24 MK182718 Taif-103	230 KPTGKPTI <u>NT</u> ED. ED. ED.	240 <u>TKTNIRTTLL</u> Q. Q. Q.	250 TSNTKGNPEH 	260 TSQEETLHST MF	270 TSEGYLSPSQ SNP P	280 VYTTSGQEET .S	2 90 LHSTTSEGYL P P P P P P	300 SPSQVYTTSE 	310 YLSQSLSSS <u>N</u> .PPS.PP. P PY PY	<u>TTK</u> .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23 MK182716 Taif-24 MK182718 Taif-103 MK182708 Taif-1	230 KPTGKPTINT ED. ED. ED. ED.	240 <u>TKTNIRTTLL</u> Q. Q. Q. Q.	250 TSNTKGNPEH 	260 TSQEETLHST MF	270 TSEGYLSPSQ SNP P NP	280 VYTTSGQEET .S	2 90 LHSTTSEGYL P P P P P P	300 SPSQVYTTSE 	310 YLSQSLSSSN .P.PS.PP. P PY PY	<u>TTK</u> .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23 MK182716 Taif-24 MK182718 Taif-103 MK182708 Taif-1 MK182710 Taif-13	230 KPTGKPTINT ED. ED. ED. ED. ED.	240 <u>TKTNIRTTLL</u> Q. Q. Q.	250 TSNTKGNPEH <u>.T</u> L <u>.T</u> Y.T <u>.</u> T	260 TSQEETLHST MF	270 TSEGYLSPSQ SNP P P NP	280 VYTTSGQEET	2 90 LHSTTSEGYL P P P P P P	300 SPSQVYTTSE 	310 YLSQSLSSSN .P.PS.PP. P PY PY PF PF	<u>TT</u> K .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23 MK182716 Taif-24 MK182718 Taif-103 MK182708 Taif-1 MK182710 Taif-13 MK182711 Taif-16	230 KPTGKPTINT ED. ED. ED. ED. ED. ED. ED. ED.	240 <u>TKTNIRTTLL</u> Q. Q. Q. I. 	250 TSNTKGNPEH <u>.T</u> L <u>.T</u> Y.T <u>.</u> T <u>.</u> T	260 TSQEETLHST MF	270 TSEGYLSPSQ SNP NP NP NP NP	280 VYTTS GQEET .S	2 90 LHSTTSEGYL P P P P P 	300 SPSQVYTTSE AH AH AH AH AH AH AH	310 YLSQSLSSSN .P.PS.PP. P P .P .P .P .P .P .P .P .P .P .P .P .P .P .P .P	<u>TTTK</u> .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23 MK182716 Taif-24 MK182718 Taif-103 MK182708 Taif-1 MK182710 Taif-13 MK182711 Taif-16 MK182712 Taif-17	230 KPTGKPTINT ED. ED. ED. ED. ED. ED. E.S.D. E.S.D.	240 <u>TKTNIRTTLL</u> Q. Q. I. Q. I 	250 TSNTKGNPEH <u>.T</u> L <u>.T</u> Y.T <u>.</u> T <u>.</u> T <u>.</u> T <u>.</u> T <u>.</u> T <u>.</u> T	I260 TSQEETLHST	270 TSEGYLSPSQ SNP NP NP NP NP NP NP	280 VYTTS GQEET .S	2 90 LHSTTSEGYL P P P P P 	300 SPSQVYTTSE AH AH AH AH AH 	310 YLSQSLSSSN .P.PS.PP. P P .P.Y .P.Y .P.Y	<u>TT</u> K .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23 MK182716 Taif-24 MK182718 Taif-103 MK182708 Taif-1 MK182710 Taif-13 MK182711 Taif-16 MK182712 Taif-17 MK182713 Taif-18	230 KPTGKPTINT ED. ED. ED. ED. ED. E.S.D. E.S.D. E.S.D.	240 <u>TKTNIRTTLL</u> Q. Q. Q. I Q. Q. Q. Q. Q. Q. Q. Q.	250 TSNTKGNPEH <u>.T</u> L <u>.T</u> Y.T <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u>	I260 TSQEETLHST	270 TSEGYLSPSQ SNP NP NP NP NP NP NP NP NP	280 VYTTS GQEET .S	2 90 LHSTTSEGYL P P P P P 	300 SPSQVYTTSE AH AH AH AH AH AH 	310 YLSQSLSSSN .P.PS.PP. P P .P.Y .P.Y .P.Y	<u>TT</u> K .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23 MK182716 Taif-24 MK182718 Taif-103 MK182708 Taif-1 MK182710 Taif-13 MK182711 Taif-16 MK182712 Taif-17 MK182713 Taif-18 MK182717 Taif-47	230 KPTGKPTINT ED. ED. ED. ED. ED. E.S.D. E.S.D. E.S.D. E.S.D.	240 <u>TKTNIRTTLL</u> Q. Q. Q. I I I Q. Q. Q. Q. 	250 TSNTKGNPEH <u>.T</u> L <u>.T</u> Y.T <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> 	1260 TSQEETLHST	270 TSEGYLSPSQ SNP P P NP NP NPP. NP.P. NP.P.	280 VYTTS GQEET .S	2 90 LHSTTSEGYL P P P P P P	300 SPSQVYTTSE AH AH AH AH AH AH 	310 YLSQSLSSSN .PPS.PP. P P .PY .PY .P .PF HPF HPP HPP HSP	TTK .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23 MK182716 Taif-24 MK182718 Taif-103 MK182708 Taif-1 MK182710 Taif-13 MK182711 Taif-16 MK182712 Taif-17 MK182713 Taif-18 MK182717 Taif-47 MK182719 Taif-110	230 KPTGKPTINT ED. ED. ED. ED. E.S.D. E.S.D. E.S.D. E.S.D. E.S.D. E.S.D. E.S.D.	240 <u>TKTNIRTTLL</u> Q. Q. Q. I I I T T T T T T	250 TSNTKGNPEH <u>.T</u> L <u>.T</u> Y.T <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> <u>.T</u> 	1260 TSQEETLHST	270 TSEGYLSPSQ SNP P P NP NP NPP. NP.P. NP.P. NP.P. NP.P. NP.P. NP.P. NP.P.	280 VYTTS GQEET . S	2 90 LHSTTSEGYL P P P P P P	300 SPSQVYTTSE AH AH AH AH AH AH 	310 YLSQSLSSSN .PPS.PP. P P .PY PF PF HPF HPP HPP HSP P.P	TTK .PR
JX627336 GN435/11 M11486 A2 MK182709 Taif-7 MK182714 Taif-20 MK182715 Taif-23 MK182716 Taif-24 MK182718 Taif-103 MK182708 Taif-1 MK182710 Taif-13 MK182711 Taif-16 MK182712 Taif-17 MK182713 Taif-18 MK182717 Taif-47 MK182719 Taif-110 MK182720 Taif-112	230 KPTGKPTINT ED. ED. ED. ED. E.S.D. E.S.D. E.S.D. E.S.D. E.S.D. E.S.D. E.S.D. E.S.D.	240 <u>TKTNIRTTLL</u> Q. Q. I. I T. T. T. T. T. T. T. 	250 TSNTKGNPEH <u>.T</u> L <u>.T</u> Y.T <u>.T</u> <u>.</u> T <u>.T</u> <u>.</u> T <u>.T</u> <u>T</u> <u>T</u> <u>T</u> 	260 TSQEETLHST	270 TSEGYLSPSQ SNP P P NP NP NPP. NP.P. NP.P. NP.P. NP.P. NP.P. NP.P. NP.SFNT	280 VYTTS GQEET . S	2 90 LHSTTSEGYL P P P P P P	300 SPSQVYTTSE AH AH AH AH AH AH AH AH 	310 YLSQSLSSSN .PPS.PP. P .P .PY PF PF HPF HPP HPP HPP HPP P.P	TTK .PR

Figure 3. Deduced amino acid sequences of the partial attachment protein (G) from different Saudi hRSV-A strains. The N-glycosylation sites, NXT/S, where X is not a proline, are underlined. Duplicated regions in the ON1 strains are shown in the boxes.

4. Discussion

Both of the designed generic primers were successful in amplification of the three different genotypes of hRSV-A, ON1, NA1, and GA5, that were detected among Saudi archival RNA samples. It is also assumed that the generic F primer set could detect all hRSV-A strains, since it showed complete identities to most of the published hRSV-A strains except for a single mismatch detected occasionally. Genotyping based on partial gene sequencing of both F and G genes showed similar results, which denotes that the designed F gene could be used efficiently in genotyping of hRSV-A strains.

The Saudi strains in the current study are classified as NA1, ON1, and GA5 genotypes. Both NA1 and GA5 showed higher virulence to the children in comparison to ON1 and other hRSV-A strains [24–26]. Two out of the thirteen strains were related to GA5, and this represents the first record of this subtype in Saudi Arabia. In Saudi Arabia, only a few studies have been conducted on hRSV genotypes; however, the hRSV-A type, mostly the NA-1 subtype, is the dominant genotype [32,33]. The hRSV-A ON1 genotype was detected for the first time in Canada in 2010 [23]. A Saudi strain in the GenBank database was found to be closely related to the ON1 strain isolated in 2009 in Riyadh and also closely related to ON1 strains detected in the current study from Al-Taif. This finding reveals that the ON1 strains have circulated in Saudi Arabia since 2009. This assumption is confirmed by the global data that suggested that the ancestor of ON1 emerged during 2008–2009 [34].

It is known that the F protein is focused on the signal peptide, p27, transmembrane domain, and ø antigenic site [35]. Similarly, the highest variability among amino acids among Saudi strains was detected in the signal peptide, followed by the p27 domain of the fusion protein. The fusion protein of hRSV strain harbors six N-glycosylation sites [9,36], which were confirmed to be conserved among the current Saudi strains, except for strain Taif-103 (A/ON1), which has lost one of the N-glycosylations at N120. It is known that N-glycosylation is important for the folding and transport of viral proteins and hence virus infectivity [37]. It is worth mentioning here that the three N-glycosylation sites—N116, N120, and N126—are found in close proximity to the proteolytic cleavage site [38,39]. It seems that the T122A amino acid substitution that resulted in a loss of the N-glycosylation site in Taif-103 does not affect the fusion or the proteolytic affinity of the protein, as confirmed previously [38].

The deduced amino acid sequence of the G gene revealed the presence of L298P, V303A, and Y304H amino acid substitution in all ON1 Saudi strains. Y273N was detected in Taif 23 and L274P was detected in the Taif-24 and Taif-103 strains. Such substitutions are considered noteworthy as they are present next to aa 265–273 (antigenic site) [40]. The P310L was recorded in the majority of the Saudi strains. Positively selected residues reported by Botosso et al., were found to be conserved in the Saudi ON1 and NA1 strains, except at the residues T237D and L274P [41] that constituted escape-mutant screened monoclonal antibodies [42,43]. The N273 N-glycosylation site was lost in the Saudi NA1 and most of the ON1 strains, which was also observed in NA1 strains from Japan and China [11,20]. Absence of the N250 site in all Saudi strains except for Saudi GA5 strains was expected, since it is a specific N-glycosylation site for both GA5 and SAA1 of hRSV-A [11]. The variation of the number of N-glycosylation sites was found to be associated with altered antigenicity [44].

5. Conclusions

Direct partial F gene sequencing represents an accurate method for hRSV-A genotyping that matched partial G gene sequencing results. The current study provides evidence of the circulation of GA5, NA1, and ON1 genotypes in Saudi Arabia. Although mutations in conserved or escape-mutant domains were detected in both F and G proteins, most of them do not affect the virus' virulence.

Author Contributions: D.A.A., N.M.A.A., and A.S.A.-M. designed the primers, conducted the RT-PCR, and analyzed the sequences. M.I.R.A.-M. and A.S.A.-M. conceived the study design. D.A.A. and N.M.A.A. drafted the manuscript. A.S.A.-M. and M.I.R.A.-M. critically revised the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

- 1. Nair, H.; Nokes, D.J.; Gessner, B.D.; Dherani, M.; Madhi, S.A.; Singleton, R.J.; O'Brien, K.L.; Roca, A.; Wright, P.F.; Bruce, N.; et al. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: A systematic review and meta-analysis. *Lancet* **2010**, *375*, 1545–1555. [CrossRef]
- 2. Weber, M.W.; Mulholland, E.K.; Greenwood, B.M. Respiratory syncytial virus infection in tropical and developing countries. *Trop. Med. Int. Health* **1998**, *3*, 268–280. [PubMed]
- 3. Falsey, A.R.; Hennessey, P.A.; Formica, M.A.; Cox, C.; Walsh, E.E. Respiratory syncytial virus infection in elderly and high-risk adults. *N. Engl. J. Med.* **2005**, *352*, 1749–1759. [CrossRef] [PubMed]
- French, C.E.; McKenzie, B.C.; Coope, C.; Rajanaidu, S.; Paranthaman, K.; Pebody, R.; Nguyen-Van-Tam, J.S.; Noso-RSV Study Group; Higgins, J.P.; Beck, C.R. Risk of nosocomial respiratory syncytial virus infection and effectiveness of control measures to prevent transmission events: A systematic review. *Influ. Resp. Virus.* 2016, 10, 268–290. [CrossRef] [PubMed]
- 5. Collins, P.L.; Melero, J.A. Progress in understanding and controlling respiratory syncytial virus: Still crazy after all these years. *Virus Res.* **2011**, *162*, 80–99. [CrossRef] [PubMed]
- 6. Levine, S.; Klaiber-Franco, R.; Paradiso, P.R. Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. *J. Gen. Virol.* **1987**, *68*, 2521–2524. [CrossRef] [PubMed]
- 7. Walsh, E.E.; Hruska, J. Monoclonal antibodies to respiratory syncytial virus proteins: Identification of the fusion protein. *J. Virol.* **1983**, 47, 171–177. [PubMed]
- Wertz, G.W.; Krieger, M.; Ball, L.A. Structure and cell surface maturation of the attachment glycoprotein of human respiratory syncytial virus in a cell line deficient in O glycosylation. *J. Virol.* 1989, 63, 4767–4776. [PubMed]
- 9. Melero, J.A. Molecular biology of human respiratory syncytial virus. *Persp. Med. Virol.* 2007, 14, 1–42.
- 10. Melero, J.A.; Mas, V.; McLellan, J.S. Structural, antigenic and immunogenic features of respiratory syncytial virus glycoproteins relevant for vaccine development. *Vaccine* **2017**, *35*, 461–468. [CrossRef] [PubMed]
- 11. Cui, G.; Zhu, R.; Qian, Y.; Deng, J.; Zhao, L.; Sun, Y.; Wang, F. Genetic variation in attachment glycoprotein genes of human respiratory syncytial virus subgroups a and B in children in recent five consecutive years. *PLoS ONE* **2013**, *8*, e75020-e. [CrossRef]
- 12. Bermingham, I.M.; Chappell, K.J.; Watterson, D.; Young, P.R. The heptad repeat C domain of the respiratory syncytial virus fusion protein plays a key role in membrane fusion. *J. Virol.* **2018**, *92*, e01323-17. [CrossRef] [PubMed]
- 13. Zhao, X.; Singh, M.; Malashkevich, V.N.; Kim, P.S. Structural characterization of the human respiratory syncytial virus fusion protein core. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 14172–14177. [CrossRef] [PubMed]
- 14. Lopez, J.A.; Bustos, R.; Orvell, C.; Berois, M.; Arbiza, J.; Garcia-Barreno, B.; Melero, J.A. Antigenic structure of human respiratory syncytial virus fusion glycoprotein. *J. Virol.* **1998**, *72*, 6922–6928. [PubMed]
- 15. Anderson, L.J.; Hierholzer, J.C.; Tsou, C.; Hendry, R.M.; Fernie, B.F.; Stone, Y.; McIntosh, K. Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J. Infect. Dis.* **1985**, *151*, 626–633. [CrossRef]
- Malekshahi, S.S.; Razaghipour, S.; Samieipoor, Y.; Hashemi, F.B.; Manesh, A.A.R.; Izadi, A.; Faghihloo, E.; Ghavami, N.; Mokhtari-Azad, T.; Salimi, V. Molecular characterization of the glycoprotein and fusion protein in human respiratory syncytial virus subgroup A: Emergence of ON-1 genotype in Iran. *Infect. Genet. Evol.* 2019, 71, 166–178. [CrossRef] [PubMed]
- 17. Gaymard, A.; Bouscambert-Duchamp, M.; Pichon, M.; Frobert, E.; Vallee, J.; Lina, B.; Casalegno, J.S.; Morfin, F. Genetic characterization of respiratory syncytial virus highlights a new BA genotype and emergence of the ON1 genotype in Lyon, France, between 2010 and 2014. *J. Clin. Virol.* **2018**, *102*, 12–18. [CrossRef] [PubMed]
- Meng, J.; Stobart, C.C.; Hotard, A.L.; Moore, M.L. An overview of respiratory syncytial virus. *PLoS Pathog.* 2014, 10, e1004016. [CrossRef] [PubMed]
- Pankovics, P.; Szabo, H.; Szekely, G.; Gyurkovits, K.; Reuter, G. Detection and molecular epidemiology of respiratory syncytial virus type A and B strains in childhood respiratory infections in Hungary. *Orvosi Hetilap* 2009, 150, 121–127. [CrossRef]
- 20. Shobugawa, Y.; Saito, R.; Sano, Y.; Zaraket, H.; Suzuki, Y.; Kumaki, A.; Dapat, I.; Oguma, T.; Yamaguchi, M.; Suzuki, H. Emerging genotypes of human respiratory syncytial virus subgroup A among patients in Japan. *J. Clin. Microbiol.* **2009**, *47*, 2475–2482. [CrossRef] [PubMed]

- 21. Katzov-Eckert, H.; Botosso, V.F.; Neto, E.A.; Zanotto, P.M. Phylodynamics and dispersal of HRSV entails its permanence in the general population in between yearly outbreaks in children. *PLoS ONE* **2012**, *7*, e41953. [CrossRef] [PubMed]
- 22. Trento, A.; Galiano, M.; Videla, C.; Carballal, G.; Garcia-Barreno, B.; Melero, J.A.; Palomo, C. Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. *J. Gen. Virol.* **2003**, *84*, 3115–3120. [CrossRef] [PubMed]
- 23. Eshaghi, A.; Duvvuri, V.R.; Lai, R.; Nadarajah, J.T.; Li, A.; Patel, S.N.; Low, D.E.; Gubbay, J.B. Genetic variability of human respiratory syncytial virus A strains circulating in Ontario: A novel genotype with a 72 nucleotide G gene duplication. *PLoS ONE* **2012**, *7*, e32807. [CrossRef] [PubMed]
- 24. Midulla, F.; Nenna, R.; Scagnolari, C.; Petrarca, L.; Frassanito, A.; Viscido, A.; Arima, S.; Antonelli, G.; Pierangeli, A. How respiratory syncytial virus genotypes influence the clinical course in infants hospitalized for bronchiolitis. *J. Infect. Dis.* **2019**, *219*, 526–534. [CrossRef] [PubMed]
- 25. Rodriguez-Fernandez, R.; Tapia, L.I.; Yang, C.F.; Torres, J.P.; Chavez-Bueno, S.; Garcia, C.; Jaramillo, L.M.; Moore-Clingenpeel, M.; Jafri, H.S.; Peeples, M.E.; et al. Respiratory syncytial virus genotypes, host immune profiles, and disease severity in young children hospitalized with bronchiolitis. *J. Infect. Dis.* **2017**, 217, 24–34. [CrossRef]
- 26. Martinello, R.A.; Chen, M.D.; Weibel, C.; Kahn, J.S. Correlation between respiratory syncytial virus genotype and severity of illness. *J. Infect. Dis.* **2002**, *186*, 839–842. [CrossRef] [PubMed]
- 27. Song, J.; Wang, H.; Ng, T.I.; Cui, A.; Zhu, S.; Huang, Y.; Sun, L.; Yang, Z.; Yu, D.; Yu, P.; et al. Sequence analysis of the fusion protein gene of human respiratory syncytial virus circulating in China from 2003 to 2014. *Sci. Rep.* **2018**, *8*, 17618. [CrossRef]
- 28. Abdel-Moneim, A.S.; Kamel, M.M.; Al-Ghamdi, A.S.; Al-Malky, M.I. Detection of bocavirus in children suffering from acute respiratory tract infections in Saudi Arabia. *PLoS ONE* **2013**, *8*, e55500. [CrossRef]
- 29. Abdel-Moneim, A.S.; Shehab, G.M.; Alsulaimani, A.A.; Al-Malky, M.I.R. Development of TaqMan RT-qPCR for the detection of type A human respiratory syncytial virus. *Mol. Cell. Probes* **2017**, *33*, 16–19. [CrossRef] [PubMed]
- 30. GenomeNet. Available online: https://www.genome.jp (accessed on 24 May 2019).
- 31. PolyPhen-2 (Polymorphism Phenotyping v2). Available online: http://genetics.bwh.harvard.edu/pph2/ (accessed on 10 October 2017).
- Almajhdi, F.N.; Farrag, M.A.; Amer, H.M. Genetic diversity in the G protein gene of group A human respiratory syncytial viruses circulating in Riyadh, Saudi Arabia. *Arch. Virol.* 2014, 159, 73–81. [CrossRef] [PubMed]
- AlMajhdi, F.N.; Al-Jarallah, A.; Elaeed, M.; Latif, A.; Gissmann, L.; Amer, H.M. Prevalence of respiratory syncytial virus infection in Riyadh during the winter season 2007–2008 and different risk factors impact. *Int. J. Virol.* 2009, *5*, 154–163.
- 34. Duvvuri, V.R.; Granados, A.; Rosenfeld, P.; Bahl, J.; Eshaghi, A.; Gubbay, J.B. Genetic diversity and evolutionary insights of respiratory syncytial virus A ON1 genotype: Global and local transmission dynamics. *Sci. Rep.* **2015**, *5*, 14268. [CrossRef] [PubMed]
- 35. Hause, A.M.; Henke, D.M.; Avadhanula, V.; Shaw, C.A.; Tapia, L.I.; Piedra, P.A. Sequence variability of the respiratory syncytial virus (RSV) fusion gene among contemporary and historical genotypes of RSV/A and RSV/B. *PLoS ONE* **2017**, *12*, e0175792-e.
- 36. Kornfeld, R.; Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Ann. Rev. Biochem.* **1985**, *54*, 631–664. [CrossRef]
- 37. Bieberich, E. Synthesis, processing, and function of N-glycans in N-glycoproteins. *Adv. Neurobiol.* **2014**, *9*, 47–70. [PubMed]
- 38. Zimmer, G.; Trotz, I.; Herrler, G. N-glycans of F protein differentially affect fusion activity of human respiratory syncytial virus. *J. Virol.* **2001**, *75*, 4744–4751. [CrossRef] [PubMed]
- 39. Lambert, D.M. Role of oligosaccharides in the structure and function of respiratory syncytial virus glycoproteins. *Virology* **1988**, *164*, 458–466. [CrossRef]
- 40. Cane, P. Molecular Epidemiology and Evolution of RSV. In *Respiratory Syncytial Virus*; Elsevier: Amsterdam, The Netherlands, 2007; pp. 89–113.

- 41. Botosso, V.F.; Zanotto, P.M.; Ueda, M.; Arruda, E.; Gilio, A.E.; Vieira, S.E.; Stewien, K.E.; Peret, T.C.; Jamal, L.F.; Pardini, M.I.; et al. Positive selection results in frequent reversible amino acid replacements in the G protein gene of human respiratory syncytial virus. *PLoS Pathog.* **2009**, *5*, e1000254. [CrossRef] [PubMed]
- 42. Garcia, O.; Martin, M.; Dopazo, J.; Arbiza, J.; Frabasile, S.; Russi, J.; Hortal, M.; Perez-Breña, P.; Martínez, I.; García-Barreno, B.; et al. Evolutionary pattern of human respiratory syncytial virus (subgroup A): Cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. *J. Virol.* **1994**, *68*, 5448–5459.
- Martinez, I.; Dopazo, J.; Melero, J.A. Antigenic structure of the human respiratory syncytial virus G glycoprotein and relevance of hypermutation events for the generation of antigenic variants. *J. Gen. Virol.* 1997, 78, 2419–2429. [CrossRef]
- 44. Palomo, C.; Cane, P.A.; Melero, J.A. Evaluation of the antibody specificities of human convalescent-phase sera against the attachment (G) protein of human respiratory syncytial virus: Influence of strain variation and carbohydrate side chains. *J. Med. Virol.* **2000**, *60*, 468–474. [CrossRef]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).