Infection, Genetics and Evolution 71 (2019) 166–178

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Molecular characterization of the glycoprotein and fusion protein in human respiratory syncytial virus subgroup A: Emergence of ON-1 genotype in Iran

Somayeh Shatizadeh Malekshahi[a], Shaghayegh Razaghipour[b], Yazdan Samieipoor[c], Farhad B. Hashemi[b], Ali Akbar Rahbari Manesh[d], Anahita Izadi[d], Ebrahim Faghihloo[e], Nastaran Ghavami[c], Talat Mokhtari-Azad[c], Vahid Salimi[c,*]

[a] Department of Virology, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran
[b] Department of Microbiology, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran
[c] Virology Department, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
[d] Bahrami Children Hospital, Tehran University of Medical Sciences, Tehran, Iran
[e] Department of Microbiology, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Keywords:
RSV
Iran
Fusion gene
Molecular epidemiology
Genotype

ABSTRACT

HRSV is a principle cause of infant hospitalization, childhood wheezing and a common pathogen in the elderly. Limited information exists regarding HRSV genotypes in Iran. In order to better understand HRSV strain diversity, we performed an in-depth evaluation of the genetic variability of the HRSV F protein detected in children under two years of age that, presented with acute respiratory symptoms during 2015–2016 in Tehran. A total of 180 nasopharyngeal swabs were evaluated. The HRSV positive samples were genotyped for G and F gene sequences using RT-PCR and sequencing methods. Phylogenetic analysis was performed using the neighbor-joining and maximum likelihood methods. Genetic and antigenic characteristics of the F gene, nucleotide and amino acids in significant positions and immune system binding regions, as well as the p-distance, positive/negative selection site, linear epitopes and glycosylation sites were investigated in all selected sequences. Among the 83 HRSV positive samples, the Fifty-five cases were successfully sequenced. All of them were classified as subgroup A and belonged to the ON-1 genotype, which possessed 72-nt duplication in the G gene. This study is the first report on the emergence of ON-1 in Iran. ON-1 Iranian sequences clustered in three lineages according to virus fusion (F) gene variations. F gene sequence analysis showed that all genetic changes in the isolates from Iran were base substitutions and no deletion/insertions were identified. The low dN/dS ratio and lack of positively selected sites showed that the fusion genes found in the strains from Iran are not under host selective pressure. Continuing and long-term molecular epidemiological surveys for early detection of circulating and newly emerging genotypes are necessary to gain a better understanding of their epidemic potential.

1. Introduction

Human respiratory syncytial virus (HRSV) infects about 60–70% of children by the end of the first year of their life, and nearly all children by three years of age (Nair et al., 2010). HRSV is a principle cause of infant hospitalization and childhood wheezing, and a common pathogen in elderly and immunocompromised patients, all with severe consequences (Shi et al., 2017; Scheltema et al., 2017). HRSV is a member of the genus *Ortopneumovirus* in the family *Pneumoviridae* with non-segmented, a negative-sense RNA genome of approximately 15, 200 nucleotides encoding 11 proteins. The G, SH and F proteins are three structural trans-membrane glycoproteins that are embedded in the viral envelope (Gimferrer et al., 2015). The attachment (G) glycoprotein is the most variable protein both between and within the major antigenic groups of HRSV (Cane, 2001).HRSV has two antigenic subgroups, A and B, the differentiation of which is based on antibody cross-reactivity patterns against the G glycoprotein and genetic analysis (Tan et al., 2012). These subgroups A and B have been further classified into 20 genotypes (GA1-7, SAA1-2, NA1-4, ON1-4, CB-A and TN1-2) and at least 36 genotypes (GB1–13, THB, BA1–14, BA-CCA, BA-CCB, SAB1–4, and URU1–2), respectively, based on variability in the distal third of G gene (Kimura et al., 2016; Gaymard et al., 2018). BA and ON-1 are two novel HRSV genotypes with unique genetic characteristics. ON-1 possesses72-nucleotide tandem duplication within the G protein 2nd

* Corresponding author.
E-mail address: vsalimi@tums.ac.ir (V. Salimi).

https://doi.org/10.1016/j.meegid.2019.03.026
Received 27 August 2018; Received in revised form 5 March 2019; Accepted 29 March 2019
Available online 01 April 2019

hypervariable region of HRSV-A and BA has 60-nucleotide duplication within the c-terminal region of the G gene of HRSV-B, which has now spread globally (Trento et al., 2003; Eshaghi et al., 2012). Epidemiological studies have demonstrated that more than one genotype from the same HRSV subgroup co-circulate in a community (Kuroiwa et al., 2004; Salimi et al., 2016). There is still little information regarding HRSV circulating genotypes in Iran. One preliminary study about HRSV genotypes reported a prevalence of GA1, GA2, and GA5 during 2007–2008 in Iran (Faghihloo et al., 2011a). The same group identified GA1, GA2, and BA between 2009 and 2013 (Faghihloo et al., 2011b; Faghihloo et al., 2014).

The HRSV F protein mediates not only virus-cell membrane fusion but also acts on the syncytium formations between infected cells (Kimura et al., 2016). This type I membrane glycoprotein is a key antigenic determinant that elicits neutralizing antibodies and cytotoxic T-lymphocyte immunity (Papenburg et al., 2012). Moreover, the F protein has gained attention as a major antigen for HRSV vaccine design and antibody development (Gálvez et al., 2017). Currently no licensed vaccine exists for HRSV; however considerable amount of vaccine candidates are in development (Hause et al., 2017). Likewise; Palivizumab, a humanized neutralizing monoclonal antibody which reduces the incidence of both lower and upper severe respiratory disease caused by HRSV (Smith et al., 2012). Motavizumab is an enhanced second generation humanized mAb which is under investigation for preventing the HRSV infection. Notably, both aforementioned candidates target conserved epitopes in the HRSV F protein (Simões et al., 2018). The HRSV F protein has been synthesized as an inactive precursor (F0) of the574 amino acid homotrimers, which is cleaved by furin in the trans-Golgi at two sites, generating two disulfide-linked polypeptides (F2-F1) (González-Reyes et al., 2001; Bolt et al., 2000). Despite the important roles of HRSV F in pathogenesis, immunity and prevention strategies, its molecular mechanism of action has yet to be characterized (Papenburg et al., 2012). Interestingly, we found many variations in the F gene among clinical strains; especially among HRSV subtypes, despite previous reports of F gene relatively conserved sequence. Due to the importance of elucidating variations in HRSV F protein epitopes, the F gene was the focus of current study. Furthermore, little is known about the prevalence of mutations in key amino acid residues within the palivizumab binding site and their potential association with resistance to palivizumab in clinical samples. Since genetic and antigenic variations have been shown to occur more frequently in the G protein rather than F protein, most previous phylogenetic studies have been conducted based on the analysis of HRSV-G glycoprotein (Gilca et al., 2006; Peret et al., 2000). In order to better understand the HRSV strain diversity; apart from phylogenetic analysis based on the G gene, in this study we evaluated in-depth the genetic variability and evolutionary patterns inherent to the HRSV F protein detected in children less than two years old presenting with acute respiratory symptoms. The nucleotide sequences of Long strain, the lab adapted historical reference strain of HRSV subgroup A, were also used for comparison in order to increase the power of study. This study provides the first genetic analysis of the F gene of HRSV isolates from Iran.

2. Materials & methods

2.1. Patients and samples

Between December 2015 and April 2016, nasopharyngeal swabs were collected from 180 children less than two years of age with symptoms of acute lower respiratory tract infection (i.e., sneezing, runny nose, cough, dyspnea, and fever) in the pediatric ward of Bahrami Children's Hospital in Tehran, Iran, and stored at −80 °C. The current human study was approved by the ethics committee of Tehran University of Medical Sciences (No. 26939). The details of demographic and clinical data according to patient history are shown in Table 1 of our previously published paper (Tahamtan et al., 2018). None of these patients had already received the monoclonal antibody palivizumab.

Table 1
RSV F gene amplification, primers and products.

Primer	Sequence (5′-3′)	Location	Tm	Product size
F-RSVF1	TCAATCAACATGCAGTGCAGT	274–294	57.88	1196
R-RSVF1	ATTGTCACAGTACCATCCTCT	1470–1451	57.88	
F-RSVF2	TGGTACTGTGACAATGCAGGA	1242–1262	57.88	436
R-RSVF2	GAGGGGAACACTAATGGATCA	1677–1655	57.88	
F-RSVF3	GTGGATACTGTGTCTGTAGGT	1560–1580	57.88	973
R-RSVF3	GGTGAGTTTGGCTCTTCATTG	2532–2511	57.88	

Viral RNA extraction, HRSV detection and genotyping of the strains based on G gene:

Viral RNA was extracted from 200 µl of each sample using the High Pure Viral Nucleic Acid Kit (Roche Diagnostics, Mannheim, Germany). Nucleoprotein (N) gene were amplified by one-step RT-PCR (Qiagen) for HRSV detection, as previously described (Malekshahi et al., 2010). G gene strain-specific RT-PCR for both A and B subgroups were performed using the method of Sato M et al. (Sato et al., 2005). DNA was sequenced using the ABI Big Dye Terminator cycle sequencing kit v3.1 on the ABI 3130 genetic analyzer (Applied Biosystems Foster City, California, USA).

2.2. F gene primers

Primers were designed on the basis of highly conserved genomic regions of the F gene (Table 1) in order to enhance the efficiency of the primers to recognize all HRSV-A strains.

2.3. RT-PCR and genotyping of the strains based on F gene sequences

Selected positive samples were tested for the HRSV-F gene using One-Step RT-PCR Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. PCR analysis was performed as follows: the anterior half of the F gene (nucleotides 274–1451) was amplified using the F-RSVF1 and R-RSV-F1 primer, the central part (nucleotides 1242–1677) was amplified using F-RSVF2 and R-RSVF2 primers, which overlapped at nucleotides 1560–1677.The C-terminal half (1560-2532 nt) of the F gene was amplified using the F-RSVF3 and R-RSVF3 primers. The PCR products were detected via gel electrophoresis on 1.0% agarose gel with SYBR®Safe DNA Gel Stain. Discrete bands corresponding to the sizes of the fragments were obtained for all of the PCR products. The PCR products were purified using QIA quick Spin Columns (Qiagen, Valencia, CA). Positive samples were subjected to sequencing of the F gene using the same primers which were applied for amplification except for the F3 fragment (reverse primer: GGTCAAAT AGCGAACCATTG) by the ABI Prism Big Dye Terminator Cycle Sequencing Kit v3.1 on the ABI 3130 genetic analyzer (Applied Biosystems Foster City, California, USA).

Nucleotide sequences were assembled and manually rechecked to find possible base errors using BioEditv.7.2.5software. To assign the genotypes of the Iranian strains based on the F gene sequences, phylogenetic tree based on clustering was employed using references sequences. The reference strains were A2 (accession number KJ155694.1), Long (accession number JX198112.1) and ON1 (USA. Texas 2012).

The detected viruses were named using the nomenclature virus name/city. Country / subtype/ week. Year/ isolate number/genotype, for example: HRSVs/Tehran.IRN/A/53.15/2 (ON-1).

2.4. Phylogenetic analysis of G and F gene sequences

As the first step, phylogenetic trees were separately constructed using positive sequences by comparing them to the G and F gene

reference strains. The amplified region of HRSV genome was located in the hypervariable region 2 of the G gene.

The phylogenetic trees were constructed using MEGA X software using the neighbor joining (NJ) and maximum likelihood (ML) methods related to the sequence dataset. The topological accuracy and statistical significance of each tree were evaluated using the bootstrap method (1000 replicates) (Efron et al., 1996) with a cut-off of 70%. The best substitution model was determined by selecting best-fit substitution models as implemented in MEGA x software package. Each dataset analyzed for its own best fit model. For this purpose, pairwise distances between the genotypes at the nucleotide level were calculated using Tamura-Nei (TN93 + G) model for G gene and Kimura's2 parameter (K2 + G) model for F gene. Mean nucleotide distances within and among different lineages and sub-lineages were also estimated using the Kimura model for F gene.

Next, Phylogenetic trees were constructed for the designated Iranian F gene sequences, as well as the major international sequences of the F gene reported from various parts of the world (overall 206 sequences). Phylogenetic analysis of F Iranian sequences was performed by comparing the ON-1 sequences obtained from Iranian samples with those of other countries. Mean nucleotide distances within sub-lineages were ⩽0.004, while mean nucleotide distances between lineages were ⩾0.011. The nucleotide sequences of G and F genes described in this report were deposited at GenBank.

2.5. F gene sequence analysis

Nucleotide and amino acid sequences of the entire F gene were analyzed and compared to the related reference sequences via multiple sequence alignments using BioEdit v.7.2.5 software. The F gene ON1 references sequences obtained from confirmed G gene ON1 sequences. Specific mutations and proportion of differences (p-distance) were identified at the nucleotide and amino acid levels.

2.6. Positive selection site analysis

To test whether any sites in the F gene were under positive selection, the synonymous (dS) and nonsynonymous (dN) mutation rates were determined at every codon in the alignment. Ratios (dN/dS mutation between species) of > 1 were considered as positive selection. Analyses were performed on the Datamonkey website interface (http://www.datamonkey.org) (Pond and Frost, 2005) using the following four methods: SLAC (single likelihood ancestor counting), FEL (fixed effects likelihood), IFEL (internal fixed effects likelihood), and MEME (mixed effects model for evolution).

2.7. Amino acid analysis of published CTL (cytotoxic T cell) and neutralizing antibody epitopes

We predicted the B-cell epitope for HRSV F gene sequences using Bepipred Linear Epitope Prediction method (http://tools.iedb.org/main/bcell/). Two websites www.allelefrequencies.net and www.bimas.cit.nih.gov/molbio/hla-bind were used to predict putative CTL epitopes for the HRSV F protein. For the predictions, only HLA types were used which were shown to be dominant in at least 10% of the different ethnic groups of Iran.

2.8. N-glycosylation and O-glycosylation sites analysis

The acquisition or loss of potential N- and O-glycosylation sites in the complete amino acid sequence of the F protein was tracked using NetNGlyc1.0Server (www.cbs.dtu.dk/services/NetNGlyc) and NetOGlyc1.0 Server (www.cbs.dtu.dk/services/NetOGlyc), respectively.

2.9. Secondary structure &3D structure

Secondary (linear) and 3D structure predictions (α-helices, β-strands and turns) for the Fprotein amino acid sequences were performed with I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER) database.

3. Results

3.1. HRSV detection and genotyping

Of the 180 nasopharyngeal swab samples (90 inpatients and 90 outpatients), HRSV was detected in 83 (46.1%). The mean age of the patients was 4.65 months and among HRSV positive individuals males (57.8%) were more affected than females (42.2%). The results showed that the prevalence of HRSV was highest in January, followed by February and March. The most common clinical symptoms were defined as cough, dyspnea, sneezing, the runny nose and fever. In children with HRSV infection, runny nose was the leading symptom (97.6%) followed by a cough (96.4%), sneezing (94%) and nasal congestion (84.3%). The G gene was sequenced for 55 outof 83 positive samples. All of these were classified as subgroup A and belonged to the ON-1 genotype, which possessed 72-nt duplication in the G gene (Fig. 1). In addition, all HRSV-positive samples were selected for F gene amplification, phylogenetic analysis and genetic characterization, and 27 sequences underwent successful F gene sequencing. (Fig. 2). Accession numbers for Iranian HRSV G sequences are MH174870 - MH174924 and accession numbers for Iranian HRSV F sequences are MH187282-MH187308.

3.2. HRSV phylogeny according to F gene

All 27 F gene sequences from Iranian patient samples were compared with GA1-GA7, ON-1 and SAA1 reference sequences and a complete sequence of the F gene (n = 179) that had been deposited in GenBank in 2017 as representative sequences of samples from the United States, Canada, Mexico, Argentina, Peru, South Africa, Kenya, Netherlands, Germany, Italy, Australia, China, Philippines, Singapore, India, New Zealand and Saudi Arabia (Fig. 3). Fig. 2 compares the 72 ON-1 sequences from Iran and those from the rest of the world with using the ML method. Three lineages and three sublineages were identified among these ON-1 sequences (Fig. 4). All sequences from Iran were in lineage 1 and most were in sublineage 1. Only one sequence from Iran along with ON-1 sequences reported from the USA and the Philippines, were in sublineage 2 (I-II). The position of some strains relatively differed when comparing the NJ and ML trees (Data not shown). However, there was no significant difference between these tree topologies.

3.3. F gene nucleotide analysis

The complete sequence (1725 nt) of the F gene of all 27 isolates from Iran were aligned and compared with the three reference strains and analyzed for the presence of possible mutations. Only nucleotide mutations, which occurred in > 50% of the isolates, were reported. All genetic changes in the isolates from Iran were base substitutions; no deletions, insertions, or frame-shift mutations were identified. Most of the F gene sequences exhibited mutations in the heptad repeat-2 (HR2) position when compared with the ON-1 reference sequence. These domains are known to be involved in trimerization of the F protein and potentially could influence the viral pathogenesis. Compared to the ON-1 reference strain, the only major mutation was A1041G at a frequency of 63%.

The average nucleotide similarity among the sequences from Iran was 99.3% with a range of 98.4% to 99.9% and, an average p-distance of 0.007. Given that the A2 and long strains are from the GA1 genotype,

Fig. 1. RSV G gene of strains from Iran compared with the RSV group A reference sequences displayed in a phylogenetic tree determined using the Neighbor-joining method with MEGA X software. Only bootstrap values > 70% are displayed at the branch nodes. The ON-1 genotypes of samples from Iran are indicated as solid black triangles.

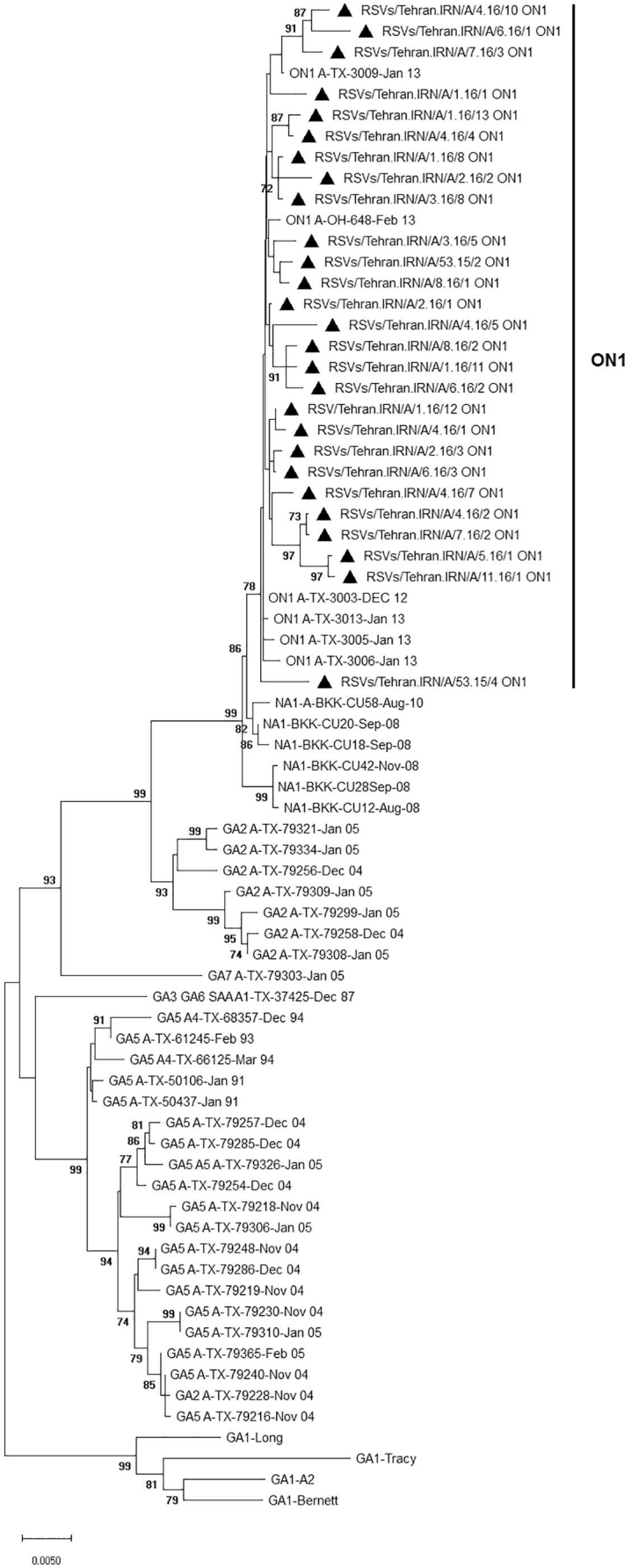

Fig. 2. RSV F gene of strains from Iran compared with the RSV group A reference sequences displayed in a phylogenetic tree determined using the Neighbor-joining method with MEGA X software. Only bootstrap values > 70% are displayed at the branch nodes. The ON-1 genotypes of samples from Iran are indicated as solid black triangles.

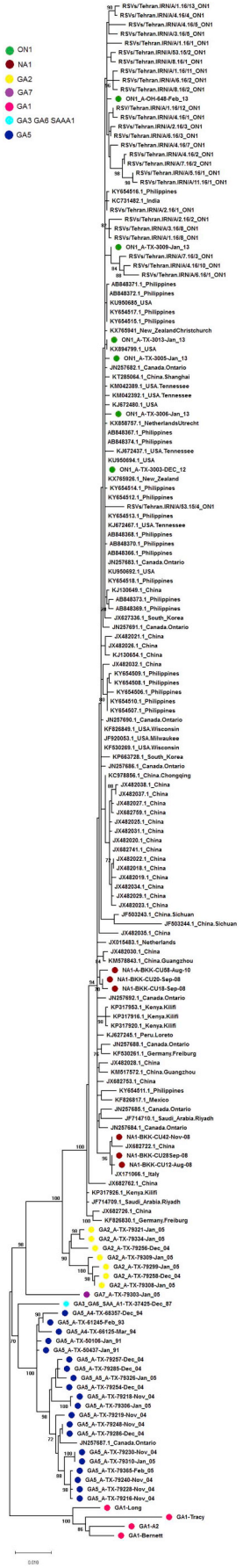

Fig. 3. Phylogenetic analysis of 179 major international sequences of the F gene reported from various parts of the world and group A reference sequences gene along with 27 sequences from Iran. A phylogenetic tree was drawn with F genes by using the maximum-likelihood method using MEGA X software. Only bootstrap values > 70% are displayed at the branch nodes.

one would expect changes over time; thus, Iranian sequences were compared with the A2 and Long strain. The results of pair-wise comparisons of the nearly whole F gene sequences show the following, nucleotide similarities: 94.8% between Iranian isolates and A2 strain, 99.5% between Iranian isolates and ON-1 strain and 95% between the Iranian isolates and the Long strain.

3.4. Amino acid analysis of F gene

The predicted amino acid sequences of the HRSV F strains from Iran were compared with those of the A2, ON-1 and Long reference strains. All subgroup A Iranian viruses showed high similarity to the reference strains with minimal variation. The high level of amino acid sequence similarity between strains suggests that the present HRSV-A F gene is highly conserved. No mutation was found in the palivizumab and motavizumab binding site (antigenic site II) in any of the specimens (Fig. 5). Furthermore, no mutation was found in the cysteine residues (C21, C37, C69, C212, C313, C322, C333, C343, C358, C367, C382, C393, C416, C422, C439) of any of the specimens. The average similaritybetween sequences from Iran at the amino acid level was 99.4% with a range of 98.4%–100%. The average p-distance between Iranian isolates was 0.006.

When compared with the Long and A2 strains, only the N276S mutation in antigenic site II was observed in all sequences. When antigenic site Ø was compared with the A2 strain, the only mutation observed among the strains from Iran was K66E. When antigenic site I was compared with the Long strain, only the V384I mutation was found in all sequences from Iran (Table 2).

3.5. Selection pressure on F gene

Analysis of the codons under positive selection in the F gene was undertaken in order to further assess the evolutionary pressures acting on HRSV. The 335, 356 and 546 codons were reported as episodic diversifying selection sites by the MEME method and the IFEL, FEL and SLAC methods showed negative pressure at the 53, 17 and 8 sites, respectively. Overall these three methods showed seven common codons (73, 347, 378, 454, 550, 560, 569) with potential negative selection. None of these codons are related to the palivizumab binding site. No positively selected codons were identified by the IFEL, FEL and SLAC methods. It could be suggested that the F genes among the samples from Iran are not under host selective pressure; however this study is from just one epidemic with a small sample size and this subject deserves further evaluation.

3.6. Neutralizing epitope variation on F gene

With regard to the monoclonal antibody binding sites reported in the literature, the sequences in this study showed several important mutations (V284I, S285I, N276S, R213S, and K66E) that were present in all sequences. Tables 3 summarizes the comparison of the Long, A2 and ON-1 reference strain sequences within the neutralizing monoclonal antibodies versus the sequences obtained in this study. Multiple mutations were also observed in the predicted linear epitopes (Fig. 6).

3.7. F gene CTL epitope variation

In Table 4, the mutations of Iranian sequences were compared with the Long, ON-1 and A2 reference strains reported in the global MHC-I

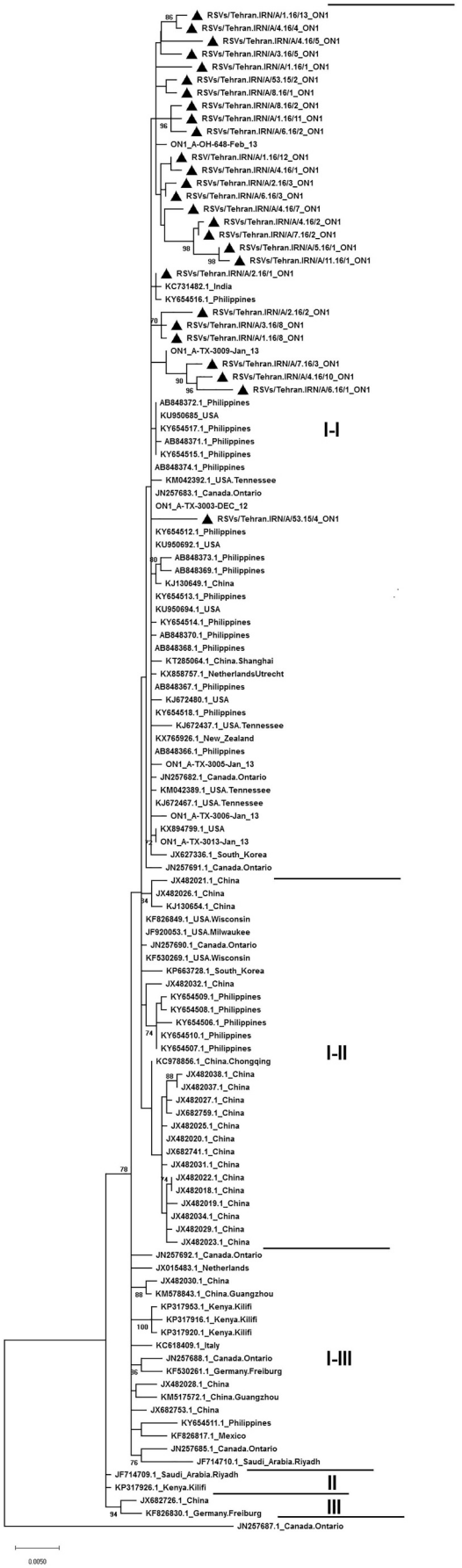

Fig. 4. Phylogenetic analysis of 27 full sequence of RSV-A F gene along with 72 ON-1 reference strains and other strains. A phylogenetic tree was drawn with F genes by using the Maximun-likelihood method using MEGA X software. Only bootstrap values > 70% are displayed at the branch nodes. The black Roman numerals indicate the lineages (I,II and III) and sublineages (I—I, I-II and I-III) among the RSV-A sequences. JN257687.1 is a GA5 reference strain as an out of ON-1cluster.

restricted CTLs. The predicted putative CTL epitopes in the Iranian sequences are summarized in Table 5. Compared to the ON-1 reference strain, Iranian sequences showed E356V and T122A mutations in the predicted putative CTL epitopes in most sequences.

3.8. F gene N-glycosylationand O-glycosylation sites analysis

Analysis of the amino acid sequences of the Iranian and reference strains (Long, ON-1 and A2) showed different patterns from potential sites for N-glycosylation. Six N-glycosylation sites (N27, N70, N116, N120, N126, N500) were similar between the Long and ON-1 reference strains. The HRSV A2 strain had five potential sites for N-glycosylation (N27, N70, N116, N126, N500). The predicted sites with N-glycosylation potential in most of our samples were similar to those of the Long and ON-1 reference strains.

Different patterns were observed from potential sites for O-glycosylation. The ON-1 strain showed two more possible sites with O-glycosylation potential when compared with the Long and A2 reference strains. As shown in Table 6, the predicted sites with O-glycosylation potential in Iranian samples were similar to the ON-1 reference strain at six amino acids 101, 103, 109, 121, 131, and 254.

4. Discussion

In order to explore the molecular epidemiology of HRSV in Tehran (Iran), we determined the genotypes of the strains circulating from December 2015 to April 2016. Phylogenetic analysis revealed that all of the isolated strains were clustered in the ON-1 genotype first described by Eshaghi et al. in Ontario, Canada in 2010 (Eshaghi et al., 2012). A total of 83 out of 180 patients with HRSV infection (46.1%) were detected with HRSV group A by a strain-specific RT-PCR test. Previous phylogenetic analyses of the HRSV distribution in Iran have reported that the majority of viruses belonged to the group A of HRSV viruses. A total of 55 out of 83 positive samples were successfully sequenced and all belonged to the ON-1 genotype (Fig. 1). While the previous phylogenetic analyses of HRSV genotype distribution in Iran from 2007 to 2013 has been reported to be 85 HRSV-A (52 strains were GA-2, 32 GA-1 and 1 GA-5) and all 9 cases of HRSV-B were BA. However, we found a striking shift in HRSVA circulation patterns from GA2 to ON-1.

Our findings demonstrate that four distinct lineages of the ON-1 genotype (Fig. 4) co-circulate in Tehran, supporting previous reports of simultaneous co-existence of multiple ON-1 lineages in a given season (Agoti et al., 2014; Duvvuri et al., 2015). Detailed analysis of the genetic characteristics of the novel HRSV-A genotype is provided herein. Given that the genotypes prevailing in samples from Iran were reported to be GA1, GA2 and BA (Faghihloo et al., 2014), this study is the first to report the emergence of ON-1 in Iran. Since its identification in Ontario, Canada in 2010, ON-1 was detected in Malaysia, India, South Korea, Germany, Italy, South Africa, Japan, China, and Kenya from 2010 to 2012 (Agoti et al., 2014). To date, the highest prevalence of the ON-1 genotype has been reported in Kenya (Agoti et al., 2014), Spain (Gimferrer et al., 2015) and the USA (Avadhanula et al., 2015). These findings confirmed that the ON-1 viruses have rapidly disseminated globally and demonstrates their strong ability to survive and replace other HRSV genotypes (Zheng et al., 2017).

The rapid distribution of BA and ON-1 genotypes harboring unique insertions suggest that these duplications may provide advantages to

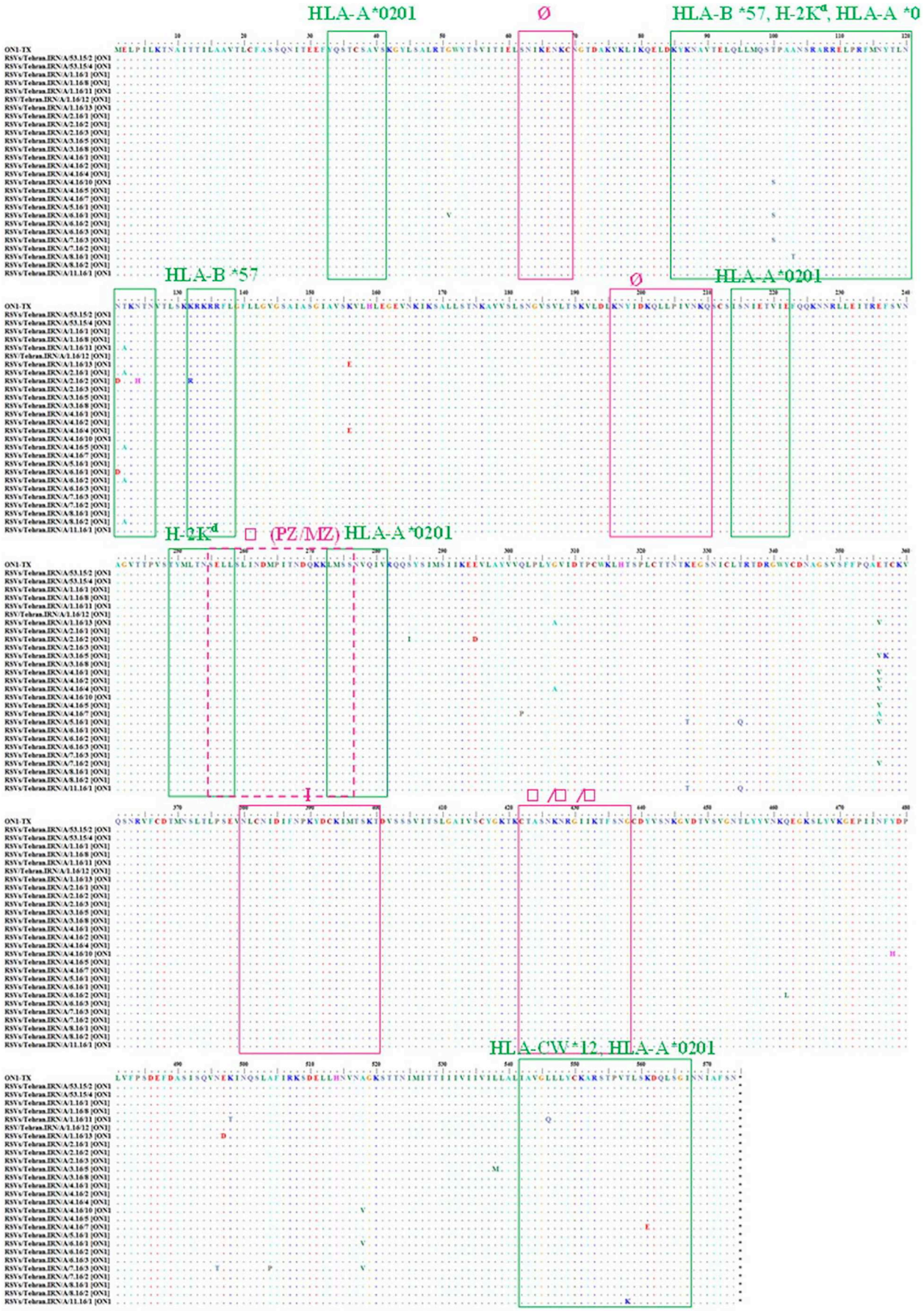

Fig. 5. Comparison of F protein sequences (574 amino acids) with ON1 reference strain. Similar residuals with dot and codon end are marked with star. Amino acid sequence of monoclonal antibody site "Palivizimab "(IIF255–276) is indicated in a pink box with a dotted point. Green boxes represent CTL epitopes and pink boxes representing antigenic sites. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Variations in HRSV F gene sequences compared to those of reference strains at antigenic sites.

Domains (amino acid no.)	Mutation of Iranian sequences compared to reference sequences		
	Long strain	A2 strain	ON1 strain
Antigenic site II PZ and MZ binding sites (255–276)	N276S(All)	N276S(All)	No mutation
Antigenic site I (380–400)	V384I(All)	V384I(All)	No mutation
Antigenic site IV,V,VI (422–438)	No mutation	No mutation	No mutation
Antigenic site Ø (62–69), (196–210)	No mutation	K66E(All)	No mutation

the virus (Korsun et al., 2017). The longer attachment protein of the 72-nt duplication in the ON-1 viruses appears to offer increased fitness over previous group A genotypes (Agoti et al., 2014). In this regard, the possibility of an ON-1 outbreak in Tehran should be expected. Whether the non-ON-1 genotypes are replaced by a variation of the ON-1 genotypes found in the Iranian population during consecutive seasons or if its emergence is temporary is an issue of great interest. As present, the data indicate that the ON-1 genotype of HRSV has entered the largest city of Iran and become the most common circulating virus and that this virus is undergoing active genetic evolution. HRSV-A and -B genotypes show that complex fluctuating dynamics and consistent shifts in the dominance of an HRSV group can occur at different times in a community (Cui et al., 2013). We speculate that high levels of herd immunity against the GA1, GA2 and GA5 subgroups that were prevalent in Tehran (Salimi et al., 2016) might be a reason for the disappearance of these genotypes by restricting their effective population size.

It should be mentioned that the failure of sequencing of the positive samples could result from the decrease in the RNA concentration in the samples due to RNA degradation because of improper sample transport conditions and reduced viral load due to the late referral of patients (Chu et al., 2013; Chu et al., 2014).

Further surveillance of circulating genotypes in unanalyzed regions is needed to determine the future trend and distribution of HRSV genotypes. Besides the epidemiological impact, novel variants may influence clinical outcomes, which underlines the importance of monitoring their spread. Diverse sets of circulating viruses can adapt to herd immunity, but the ability of HRSV to evade previously induced herd immunity is a challenge to the development of an effective vaccine (Schobel et al., 2016).

Molecular epidemiological data based on the F protein characteristics of HRSV strains from various regions of the world is scarce. To date, the molecular evolution and phylogenetic analyses of the HRSV F protein have not been reported for viruses circulating in Iran or any other Middle Eastern country. Our molecular analysis found that the F protein was highly conserved among the strains from Iran, which confirms previous findings in other countries (Tapia et al., 2014; Xia et al., 2014). It was found that the similarity at the nucleotide level between Iranian sequences was 99.3% and the similarity between these sequences with reference strains in the subtype A was 98%. Because the mean p-distance between the Iranian sequences at the nucleotide level was 0.007 and at the amino acid level was 0.006, it shows less amino acid divergence when compared with the nucleotide divergence. Moreover, most of the changes resulted in silent mutations (synonymous substitutions) and, to a lesser extent, led to amino acid alterations which implies an absence of positive immune selection within the protein. Additionally, despite the reports that show F gene mostly contains conserved regions, our results indcate that, this gene can also be used for HRSV genotyping.

Sequencing of the second hypervariable region of the G gene has been widely used to further subdivide the HRSV-A and HRSV-B subgroups into genotypes and facilitated differentiation of the RSV isolates. Because duplication in the second hypervariable region of the G gene is characteristic of the ON-1 genotype and a major difference from the previously circulating NA-1 genotype and because the F gene is known to contain more conserved regions, the exclusive analysis of the F gene alone is not an appropriate method for HRSV genotyping.

All sequences in the present study showed the N276S mutation in antigenic site II, which is the target site for palivizumab monoclonal antibody (Hause et al., 2017). Residue 276 is important because it is located next to the palivizumab binding site (Song et al., 2018). Fortunately, N276S substitution does not affect the efficacy of the palivizumab antibody (Ren et al., 2014) and it binds to the defined epitopes (Asn258 - Val278) (Smith et al., 2012). This is an important finding, as palivizumab prophylaxis at specific doses is recommended during the HRSV season for high-risk children with preterm births ($\leq$32 weeks of gestation and a subset of those between 32 and 35 weeks) and infants under two years of age with chronic lung problems or congenital heart disease at specific times (Boivin et al., 2008).

In the F1 polypeptide, five antigenic sites (I, II, IV, V and VI) have been mapped and are composed of conformational epitopes that are involved in the neutralization and fusogenic activity of the protein

Table 3
Mutation of Iranian F sequences compared to ON1, Long and A2 strains in reported and predicted B cell epitopes.

Domains (nucleotide)	Mutation of Iranian sequences compared to reference sequences		
	Long strain	A2strain	ON1strain
Reported neutralizing mAb epitopes	R213S(All)	K66E(All)	E295D
(62–69,196–210)	N276S(All)	N276S(All)	
(205–225)(221–236)	E295D	E295D	
(262–276)(262–268)	V384I(All)	V384I(All)	
(289–298)(380–400)			
(422–438)			
Predicted B-Cell epitopes with IEDB	F20L(All), T100S	G25S(All), K66E(All)	T100S, A103T
(25–35)(64–74)	S101P(All), A103T	T100S, T103A(All)	N121D, T122A
(83–89)(99–110)	N105S(All),N121D	N105S(All), N121D	K132R, K327T
(119–132)(238–248)	T122A	A122T, K124N(All)	T335Q, E356V
(265–273)(321 − 331)	K124N(All), K124H	K124H, K132R	E356A, T357K
(338–361)(397–405)	K132R, K327T	K327T, E356V	Q462L, N496T
(419–430)(436–452)	E356V, E356A	E356A, T357K	E497D, K498T
(460–470)(482–500)	T357K, Q462L	Q462L, N496T	A518V, T558K
(517–523)(553–566)	N496T, E497D	E497D, K498T	K561E
	K498T, A518V	A518V, T558K	
	T558K, K561E	K561E	

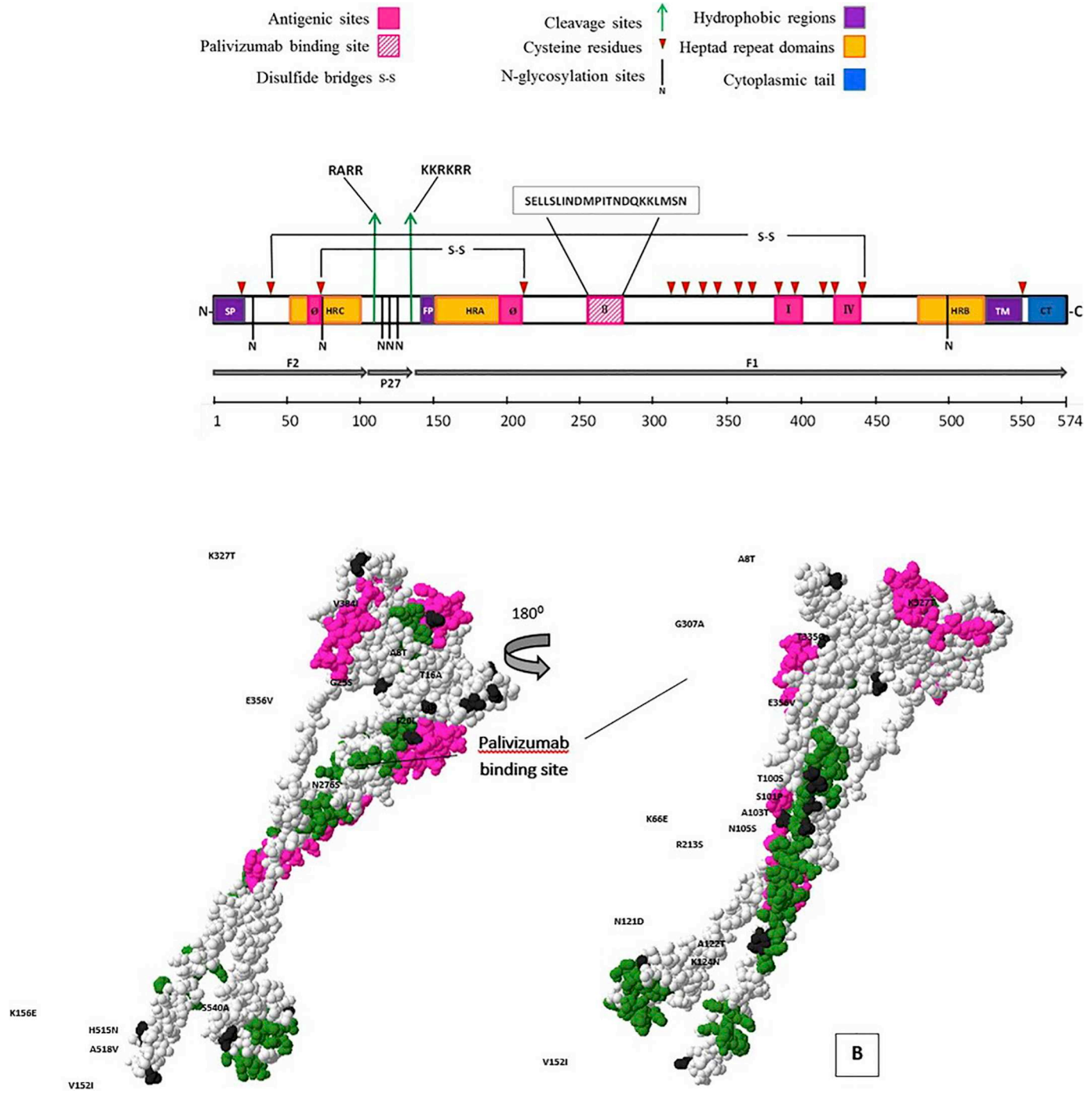

Fig. 6. A: Linear schematic of fusion protein. Palivizumab binding site and its sequences are shown. B: The three-dimensional structure of the F protein of ON1 strain from two angles. Monoclonal antibodies epitope is shown in pink and CTL epitopes are shown in green. The nucleotides that were mutated in ≥50% of the sequences were marked with black solid circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Tomé et al., 2012). Previous reports have stated that the F1 subunit was more conserved than the F2 subunit (Chi et al., 2013; Plows and Pringle, 1995). The F2 subunit showed more variation than the F1 subunit, because of the large number of sites with potential for glycosylation in this subunit (usually 4 or 5 sites, depending on the strains) (Plows and Pringle, 1995). In a study in South Korea the highest genetic divergence also was observed in F2 (Kim et al., 2007). The cleavage sites were unchanged in the Iranian sequences. Also, the number and position of cysteine residues were conserved among all sequences, which is similar to studies in the UK in 1995 (Plows and Pringle, 1995) and South Africa (Agenbach et al., 2005). Genetic variation in the HRSV for each of its genes occurs with different patterns and can occur during viral RNA synthesis or by the nucleotide hypermutation. A high degree of conservation in the F1 domain is required to retain a functional fusion mechanism (Schobel et al., 2016).

Overall, the IFEL, FEL and SLAC methods of positive selection site analysis showed seven common codons (73, 347, 378, 454, 550, 560 and 569) in the Iranian sequences with potential negative selection. A low dN/dS ratio and lack of positively selected sites indicate that the fusion genes in strains found in Iran are not under host selective pressure and this negative selection pressure only leads to variability. This might indicate that changes in the F protein are deleterious, probably

Table 4
Predicted epitopes of MHC-I restricted CTLs for common MHC-I allels in Iran.

HLA type	Population	Score result	Start	Sequence (9mer)
B*51:02	Iranian	55	7	KANAITTIL
A*01	Iranian	90	58	TIELSNIKK
A*03	Iranian	180	77	KLIKQELDK
A*24	Iranian	440	85	KYKNAVTEL
CW*04:01	Iranian	440	85	KYKNAVTEL
B*51:02	Iranian	181.5	88	NAVTELQLL
B*51:01	Iranian	60.5	88	NAVTELQLL
B*51:01	Iranian	157.3	111	LPRFMNYTL
CW*04:01	Iranian	88	111	LPRFMNYTL
B*51:02	Iranian	60.5	111	LPRFMNYTL
B*35:01	Iranian	60	111	LPRFMNYTL
B*51:01	Iranian	73.205	121	NAKKTNVTL
CW*04:01	Iranian	50	136	RFLGFLLGV
A*02:01	Iranian	177.566	140	FLLGVGSAI
B*51:01	Iranian	57.2	144	VGSAISAGV
B*51:01	Iranian	110	146	SAIASGVAV
B*51:02	Iranian	300	146	SAIASGVAV
B*51:03	Iranian	110	146	SAIASGVAV
B*51:02	Iranian	72.6	150	SGVAVSKVL
B*51:01	Iranian	65	152	VAVSKVLHL
B*51:02	Iranian	150	152	VAVSKVLHL
A*02:01	Iranian	257.342	170	ALLSTNKAV
A*02:01	Iranian	83.527	187	VLTSKVLDL
A*02:01	Iranian	266.579	191	KVLDLKNYI
CW*04:01	Iranian	200	222	FFQQKNNRL
A*24	Iranian	330	249	TYMLTNSEL
CW*04:01	Iranian	240	249	TYMLTNSEL
A*02:01	Iranian	128.67	250	YMLTNSELL
A*02:01	Iranian	135.105	272	KLMSNNVQI
A*02:01	Iranian	80.527	273	LMSNNVQIV
B*52:01	Iranian	100	284	QSYSIMSII
B*51:01	Iranian	173.03	297	LAYVVQLPL
B*51:02	Iranian	302.5	297	LAYVVQLPL
B*51:03	Iranian	54.45	297	LAYVVQLPL
CW*04:01	Iranian	50	298	AYVVQLPLY
B*52:01	Iranian	720	301	VQLPLYGVI
B*51:02	Iranian	880	352	FPQAETCKV
B*51:01	Iranian	260	352	FPQAETCKV
CW*04:01	Iranian	380	365	VFCDTMNSL
A*01	Iranian	50	383	NVDIFNPKY
B*35:01	Iranian	180	388	NPKYDCKIM
A*02:01	Iranian	123.846	394	KIMTSKTDV
B*51:01	Iranian	242	423	TASNKNRGI
B*51:02	Iranian	266.2	423	TASNKNRGI
B*51:03	Iranian	121	423	TASNKNRGI
A*24	Iranian	75	467	LYVKGEPII
A*01	Iranian	112.5	470	KGEPIINFY
CW*04:01	Iranian	75	487	EFDASISQV
CW*04:01	Iranian	50	504	RFLGFLLGV
B*51:01	Iranian	242	517	NAGKSTTNI
B*51:02	Iranian	242	517	NAGKSTTNI
B*51:03	Iranian	110	517	NAGKSTTNI
A*02:01	Iranian	80.527	525	IMITTIIIV
A*02:01	Iranian	138.354	536	VILLSLIAV

due to stringent structural or functional constraints (Gaunt et al., 2011).

It is important to determine whether genotype-specific mutations that occur in the F protein altered the antibody and CTL epitopes of HRSV. Changes such as G25S, T103A, and A122T, which are specific to subtype A, have been observed in the Iranian sequences and may affect the secondary protein structure as they can cause biochemical changes. Table 3 shows that, at specific sites coding for the neutralization of monoclonal antibody-binding sites, several important mutations, including K66E at the DM25 binding region, R213S at the RS-348 binding region, and N276S at the palivizumab binding region, occurred in all sequences. Lawlor et al. in 2013 reported that a change at aa 66 in HRSV F in the F2 fragment alters fusion activity. As mentioned, the importance of residue 276 is due to the fact that it located next to the palivizumab binding site (Song et al., 2018). Recent studies have concluded that the N276S substitution does not affect the efficacy of the palivizumab antibody (Zhu et al., 2012; Xia et al., 2014), but single amino acid substitutions at residues 262, 272, 275 could confer resistance to palivizumab (Zhu et al., 2012). There were multiple mutations in the predicted linear epitopes for B lymphocytes; these regions are considered good candidates for vaccine development; further research on the effects of these mutations on the pathogenicity of the virus, as well as the efficacy of the antibodies are needed.

In the current study, in addition to the MHCI-restricted CTL epitopes previously reported in various parts of the world, the predicted CTL epitopes related to predominant HLAs in Iran were investigated. The Iranian sequences showed E356V and T122A mutation in the predicted putative CTL epitopes in most sequences. The description of the CTL epitopes in context of the Iranian strains is new and is an important addition to the literature. This can aid in the design of a new vaccine and determine antibody candidates against HRSV.

Table 5
Mutation of Iranian sequences compared to reference strains in reported restricted cytotoxic T-lymphocyte MHC-I in the world and predicted putative Cytotoxic T-lymphocyte epitopes in Iran.

Domains (nucleotide)	Mutation of Iranian sequences compared to reference sequences		
	Longstrain	A2strain	ON1strain
ReportedMHC-IRestricted CTL Epitopes	T100S	T100S, T103A(All)	T100S, A103T
H-2K[d](85–93/92–106/249–258)	S101P(All)[1], A103T	N105S(All),N121D	N121D, T122A
HLA-B *57(106–114,118-126,132–138)	N105S(All),N121D	A122T, K124 N(All)	K132R, L546Q
HLA-A *01(109–120)	T122A, K124 N(All)	K124H, K132R	T558K, K561E
HLA-A*0201(33–41,214-222,273-281,559–567)	K124H, K132R	N276S(All),L546Q	
HLA-CW*12 (542–550,551–559)	N276S(All), L546Q	T558K, K561E	
	T558K, K561E		
Predicted putative CTL epitopes in Iran	A8T(All), N121D	A8T(All),K66E(All)	N121D, T122A
	T122A, K124 N(All) K124H, K156E	N121D, A122T	K156E, S285I
	N276S(All),S285I	K124 N(All), K124H	Q302P, G307A
	Q302P, G307A	V152I(All), K156E	E356V, E356A
	E356V, E356A	N276S(All), S285I	T357K, Y478H
	T357K,V384I(All)	Q302P, G307A	A504P, A518V
	Y478H,A504P	E356V, E356A	L538 M
	A518V, L538 M	T357K, V384I(All)	
	S540A(All)	Y478H, A504P	
		A518V, L538 M	
		S540A(All)	

Table 6
Number of positions with potential O-glycosylation for Long, ON-1 and A2 reference strains.

Reference strain	Source	Number of sites	Number of amino acid with potential O-glycosylation
Long	NetOGlyc-4.0.0.13	4	99, 101, 118, 128
A2	NetOGlyc-4.0.0.13	4	99, 100, 103, 244
ON1	NetOGlyc-4.0.0.13	6	101, 103, 109, 121, 131,254

The variability of positions with the potential of N-glycosylation and O-glycosylation in the fusion protein, which seems to play a role in evading the humoral immune response, may explain the prolonged duration of some genotypes in HRSV outbreak seasons (Reiche and Schweiger, 2009). We found that in most Iranian sequences (74%), there is a consistent profile of 6 positions with the potential for N-linked glycosylation in the F protein (N27, N70, N116, N120, N126 and N500). There were no mutations in these positions in Iranian sequence. Modulation of the number and distribution patterns of the N- and O-linked glycosylation sites can influence epitope expression either by masking or by contributing to recognition by carbohydrate specific antibodies, thereby aiding in the escape of HRSV from the host immune responses (Reiche and Schweiger, 2009; Bashir et al., 2013; Baek et al., 2012). However there were no mutations in these glycosylation positions in any of the Iranian sequences.

In conclusion, this study confirms the predominance of the novel ON-1 genotype in Tehran. It is evident that the ON-1 genotype is disseminating quickly throughout the world, highlighting the significance of studying HRSV genotypes at the local level. Data regarding selective pressures driving amino acid changes, and glycosylation sites can potentially help determine antigenicity of HRSV protein and its susceptibility to palivizumab treatment, as well as the development of vaccines. It is reasonable to postulate that the prevalence of dominant new HRSV genotypes can occur by evading host immunity and lessening the potential of vaccine-induced herd immunity in a population. Continuous long-term molecular epidemiological surveillance that aim at early detection of circulating HRSV strains and newly emerging genotypes will ensure a better understanding of HRSV potential epidemic impact as well as improve the prospects of developing targeted therapies.

Declarations of interest

None.

Acknowledgments

We thank the following people for their contributions to this work: Mr. Ali Ramezani, Ms. Maryam Naseri, for their technical support, and Pedro A. Piedra for the HRSV reference sequences. This work was supported by School of Public Health, Tehran University of Medical Sciences (No. 27253).

References

Agenbach, E., Tiemessen, C.T., Venter, M., 2005. Amino acid variation within the fusion protein of respiratory syncytial virus subtype A and B strains during annual epidemics in South Africa. Virus Genes 30, 267–278.
Agoti, C.N., Otieno, J.R., Gitahi, C.W., et al., 2014. Rapid spread and diversification of respiratory syncytial virus genotype ON1, Kenya. Emerg. Infect. Dis. 20, 950–959.
Avadhanula, V., Chemaly, R.F., Shah, D.P., et al., 2015. Infection with novel respiratory syncytial virus genotype Ontario (ON1) in adult hematopoietic cell transplant recipients, Texas, 2011–2013. J. Infect. Dis. 211, 582–589.
Baek, Y.H., Choi, E.H., Song, M.S., et al., 2012. Prevalence and genetic characterization of respiratory syncytial virus (RSV) in hospitalized children in Korea. Arch. Virol. 157, 1039–1050.
Bashir, U., Alam, M.M., Sadia, H., et al., 2013. Molecular characterization of circulating respiratory syncytial virus (RSV) genotypes in Gilgit Baltistan Province of Pakistan during 2011–2012 winter season. PLoS One 8, e74018.
Boivin, G., Caouette, G., Frenette, L., et al., 2008. Human respiratory syncytial virus and other viral infections in infants receiving palivizumab. J. Clin. Virol. 42, 52–57.
Bolt, G., Pedersen, L.O., Birkeslund, H.H., 2000. Cleavage of the respiratory syncytial virus fusion protein is required for its surface expression: role of furin. Virus Res. 68, 25–33.
Cane, P.A., 2001. Molecular epidemiology of respiratory syncytial virus. Rev. Med. Virol. 11, 103–111.
Chi, H., Liu, H.F., Weng, L.C., et al., 2013. Molecular epidemiology and phylodynamics of the human respiratory syncytial virus fusion protein in northern Taiwan. PLoS One 8, e64012.
Chu, H.Y., Kuypers, J., Renaud, C., et al., 2013. Molecular epidemiology of respiratory syncytial virus transmission in childcare. J. Clin. Virol. 57, 343–350.
Chu, H.Y., Englund, J.A., Podczervinski, S., et al., 2014. Nosocomial transmission of respiratory syncytial virus in an outpatient cancer center. Biol. Blood Marrow Transplant. 20, 844–851.
Cui, G., Zhu, R., Qian, Y., et al., 2013. Genetic variation in attachment glycoprotein genes of human respiratory syncytial virus subgroups A and B in children in recent five consecutive years. PLoS One 8, e75020.
Duvvuri, V.R., Granados, A., Rosenfeld, P., et al., 2015. Genetic diversity and evolutionary insights of respiratory syncytial virus A ON1 genotype: global and local transmission

dynamics. Sci. Rep. 5, 14268.
Efron, B., Halloran, E., Holmes, S., 1996. Bootstrap confidence levels for phylogenetic trees. Proc. Natl. Acad. Sci. U. S. A. 93, 13429–13434.
Eshaghi, A., Duvvuri, V.R., Lai, R., et al., 2012. Genetic variability of human respiratory syncytial virus A strains circulating in Ontario: a novel genotype with a 72 nucleotide G gene duplication. PLoS One 7, e32807.
Faghihloo, E., Rezaie, F., Salimi, V., et al., 2011a. Molecular epidemiology of human respiratory syncytial virus in Iran. ActaVirol 55, 81–83.
Faghihloo, E., Salimi, V., Rezaei, F., et al., 2011b. Genetic diversity in the G protein gene of human respiratory syncytial virus among Iranian children with acute respiratory symptoms. Iran. J. Pediatr. 21, 58–64.
Faghihloo, E., Yavarian, J., Jandaghi, N.Z., et al., 2014. Genotype circulation pattern of human respiratory syncytial virus in Iran. Infect. Genet. Evol. 22, 130–133.
Gálvez, N.M.S., Soto, J.A., Kalergis, A.M., 2017. New insights contributing to the development of effective vaccines and therapies to reduce the pathology caused byhRSV. Int. J. Mol. Sci. 18 (pii: E1753).
Gaunt, E.R., Jansen, R.R., Poovorawan, Y., Simmonds, P., et al., 2011. Molecular epidemiology and evolution of human respiratory syncytial virus and human metapneumovirus. PLoS One 6, e17427.
Gaymard, A., Bouscambert-Duchamp, M., Pichon, M., et al., 2018. 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. 102, 12–18.
Gilca, R., De Serres, G., Tremblay, M., et al., 2006. Distribution and clinical impact of human respiratory syncytial virus genotypes in hospitalized children over 2 winter seasons. J. Infect. Dis. 193, 54–58.
Gimferrer, L., Campins, M., Codina, M.G., et al., 2015. Molecular epidemiology and molecular characterization of respiratory syncytial viruses at a tertiary care university hospital in Catalonia (Spain) during the 2013–2014 season. J. Clin. Virol. 66, 27–32.
González-Reyes, L., Ruiz-Argüello, M.B., García-Barreno, B., et al., 2001. Cleavage of the human respiratory syncytial virus fusion protein at two distinct sites is required for activation of membrane fusion. Proc. Natl. Acad. Sci. U. S. A. 98, 9859–9864.
Hause, A.M., Henke, D.M., Avadhanula, V., et al., 2017. Sequence variability of the respiratory syncytial virus (RSV) fusion gene among contemporary and historical genotypes of RSV/A and RSV/B. PLoS One 12, e0175792.
Kim, Y.K., Choi, E.H., Lee, H.J., 2007. Genetic variability of the fusion protein and circulation patterns of genotypes of the respiratory syncytial virus. J. Med. Virol. 79, 820–828.
Kimura, H., Nagasawa, K., Tsukagoshi, H., et al., 2016. Molecular evolution of the fusion protein gene in human respiratory syncytial virus subgroup A. Infect. Genet. Evol. 43, 398–406.
Korsun, N., Angelova, S., Tzotcheva, I., et al., 2017. Prevalence and genetic characterization of respiratory syncytial viruses circulating in Bulgaria during the 2014/15 and 2015/16 winter seasons. Pathog. Glob. Health 111, 351–361.
Kuroiwa, Y., Nagai, K., Okita, L., et al., 2004. Genetic variability and molecular epidemiology of respiratory syncytial virus subgroup a strains in Japan determined by heteroduplex mobility assay. J. Clin. Microbiol. 42, 2048–2053.
Lawlor, H.A., Schickli, J.H., Tang, R.S., 2013. A single amino acid in the F2 subunit of respiratory syncytial virus fusion protein alters growth and fusogenicity. J. Gen. Virol. 94, 2627–2635.
Malekshahi, S.S., Azad, T.M., Yavarian, J., et al., 2010. Molecular detection of respiratory viruses in clinical specimens from children with acute respiratory disease in Iran. Pediatr. Infect. Dis. J. 29, 931–933.
Nair, H., Nokes, D.J., Gessner, B.D., et al., 2010. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. Lancet 375, 1545–1555.
Papenburg, J., Carbonneau, J., Hamelin, M.È., et al., 2012. Molecular evolution of respiratory syncytial virus fusion gene, Canada, 2006–2010. Emerg. Infect. Dis. 18, 120–124.
Peret, T.C., Hall, C.B., Hammond, G.W., et al., 2000. Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. J. Infect. Dis. 181, 1891–1896.
Plows, D.J., Pringle, C.R., 1995. Variation in the fusion glycoprotein gene of human respiratory syncytial virus subgroup A. Virus Genes 11, 37–45.
Pond, S.L., Frost, S.D., 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21, 2531–2533.
Reiche, J., Schweiger, B., 2009. Genetic variability of group A human respiratory syncytial virus strains circulating in Germany from 1998 to 2007. J. Clin. Microbiol. 47, 1800–1810.
Ren, L., et al., 2014. The genetic variability of glycoproteins among respiratory syncytial virus subtype A in China between 2009 and 2013. Infect. Genet. Evol. 27, 339–347.
Salimi, V., Tavakoli-Yaraki, M., Yavarian, J., et al., 2016. Prevalence of human respiratory syncytial virus circulating in Iran. J. Infect. Public Health 9, 125–135.
Sato, M., Saito, R., Sakai, T., et al., 2005. Molecular epidemiology of respiratory syncytial virus infections among children with acute respiratory symptoms in a community over three seasons. J. Clin. Microbiol. 43, 36–40.
Scheltema, N.M., Gentile, A., Lucion, F., et al., 2017. Global respiratory syncytial virus-associated mortality in young children (RSV GOLD): a retrospective case series. Lancet Glob. Health 5, e984–e991.
Schobel, S.A., Stucker, K.M., Moore, M.L., et al., 2016. Respiratory Syncytial Virus whole-genome sequencing identifies convergent evolution of sequence duplication in the C-terminus of the G gene. Sci. Rep. 6, 26311.
Shi, T., McAllister, D.A., O'Brien, K.L., et al., 2017. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. Lancet 390, 946–958.
Simões, E.A.F., Bont, L., Manzoni, P., et al., 2018. Past, present and future approaches to the prevention and treatment of respiratory syncytial virus infection in children. Infect. Dis. Ther. 7, 87–120.
Smith, G., Raghunandan, R., Wu, Y., et al., 2012. Respiratory syncytial virus fusion glycoprotein expressed in insect cells form protein nanoparticles that induce protective immunity in cotton rats. PLoS One 7, e50852.
Song, J., Wang, H., Ng, T.I., et al., 2018. Sequence analysis of the fusion protein gene of human respiratory syncytial virus circulating in China from 2003 to 2014. Sci. Rep. 8, 17618.
Tahamtan, A., Samieipoor, Y., Nayeri, F.S., et al., 2018. Effects of cannabinoid receptor type 2 in respiratory syncytial virus infection in human subjects and mice. Virulence 9, 217–230.
Tan, L., Lemey, P., Houspie, L., et al., 2012. Genetic variability among complete human respiratory syncytial virus subgroup a genomes: bridging molecular evolutionary dynamics and epidemiology. PLoS One 7, e51439.
Tapia, L.I., et al., 2014. Gene sequence variability of the three surface proteins of human respiratory syncytial virus (HRSV) in Texas. PLoS One 9 (3), e90786.
Tomé, L., Frabasile, S., Candia, C., et al., 2012. Selection and characterization of human respiratory syncytial virus escape mutants resistant to a polyclonal antiserum raised against the F protein. Arch. Virol. 157, 1071–1080.
Trento, A., Galiano, M., Videla, C., et al., 2003. Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. J. Gen. Virol. 84, 3115–3120.
Xia, Q., et al., 2014. Detection of respiratory syncytial virus fusion protein variants between 2009 and 2012 in China. Arch. Virol. 159 (5), 1089–1098.
Zheng, Y., Liu, L., Wang, S., et al., 2017. Prevailing genotype distribution and characteristics of human respiratory syncytial virus in northeastern China. J. Med. Virol. 89, 222–233.
Zhu, Q., Patel, N.K., McAuliffe, J.M., et al., 2012. Natural polymorphisms and resistance-associated mutations in the fusion protein of respiratory syncytial virus (RSV): effects on RSV susceptibility to palivizumab. J. Infect. Dis. 15 (205), 635–638.