Identification of Novel genotypes of Human Respiratory Syncytial Virus (RSV) A strain circulating in Sudan

Sahar O Khalil1*, Hisham N Altayb2, Khalid A Enan3, Ali. Y. H4, Isam M Elkhidir5 and Mohamed A. Hassan6

1 Department of Microbiology, University of science and technology, P.O. Box: 30, Omdurman, Sudan.
2 College of medical laboratory Science, Sudan. University of Science and Technology, Sudan.
3 Department of Virology, Central laboratory, Ministry of science and Technology, P.O. Box 7099, Khartoum, Sudan.
4 Department of Virology, Veterinary research Institute, P.O. Box 80 67, Khartoum, Sudan.
5 Department of Microbiology and Parasitology, Faculty of Medicine, University of Khartoum, Khartoum, Sudan.
6 Department of Bioinformatics, Africa City of Technology. Division of Molecular Genetics, Institute of Human Genetics, University of Tubingen, Germany.

* Corresponding author: Sahar O Khalil; e-mail: saharosmankhalil@yahoo.com

ABSTRACT
Respiratory syncytial virus (RSV) is the major cause of acute lower respiratory tract infection in children and vulnerable adults, but little is known regarding RSV infection in Sudan. 224 throat swab specimens were collected from children less than 5 years old, with respiratory tract infections admitted at Khartoum Hospitals in winter season (2011-2012), were screened for RSV using direct immunofluorescence assay (DFA), Reverse transcription-polymerase chain reaction (RT-PCR) and Nested RT-PCR. Nucleotide sequencing and bioinformatics analysis based on the G gene were done for nine cell culture isolated of RSV strains and were positive by DFA, RT-PCR and Nested RT-PCR. Out of 224 patients, RSV infections were detected in 136 (60.7%) patients, using DFA technique, 44 (19.6%) patients using RT-PCR and 14 (41.2%) patients, using Nested RT-PCR. Multiple sequence alignment of Sudanese RSV sequences showed, substitution of amino acid serine (S) to Cysteine (C) at position 11 in isolate 176, In isolate 316 also there was substitution of amino acid Asparagine (N) to lysine (K) at position 73. Insertion of G at position 84 and A at position 103 in isolate 377 cause frame shift mutations, when compared to other RSV G genes from database. The G protein novel mutations in isolates 176, 316 and 377 were damaging the protein in addition to change in functions and 3D structure of mutant protein of the Sudanese RSV strains, this may have implications for RSV vaccine development in Sudan.

Keywords: Respiratory tract infection, respiratory syncytial viruses, reverse transcription polymerase chain reaction, direct immunofluorescence assay.

1. INTRODUCTION
Respiratory syncytial virus (RSV), a Pneumovirus, family Paramyxoviridae, remains the most significant cause agent of severe lower respiratory tract disease in infants, young children, and the elderly worldwide. However, there is no safe and effective RSV vaccine licensed for human use [1, 2, 3]. RSV has two antigenic subgroups (A and B) exist with partial cross-protection [4, 5]. In general, subtype A strains are thought to be more virulent and usually the predominant circulating strains compared with subtype B strains [6, 7]. The most two immunogenic RSV proteins are the Fusion (F) and Glycoprotein (G) that are expressed on the virion surface, responsible for inducing production of neutralizing antibodies [8, 9], with resultant
accumulation of mutations in response to host immunological pressure [10]. Glycoprotein (G) has the major antigenic differences, a type II glycoprotein with a single N-terminal hydrophobic region (amino acids 36–66) that serves as a signal peptide and membrane-anchor [9, 11, 12]. Proximal to the membrane anchor region is an extracellular ectodomain containing four cysteine residues that are highly conserved in all RSV isolates [9,13,14] and a putative receptor binding site, flanked by two hypervariable regions (HVRs) [9,15]. The second HVR (HVR2) carries the C terminus of the protein and is commonly sequenced to examine the genetic variability of RSV strains within a given population [15, 16, 17, 18, 19]. The cysteine region of RSV contains a CX3C chemokine motif (amino acids 182–186) that may facilitate virus attachment to cells expressing the CX3C chemokine receptor, and modify CX3CL1 (fractalkine)-mediated responses as an immune evasion strategy (20).Several studies of RSV strains have described major differences in G protein length due to duplications, deletions, and insertions [21,22,23]. Much less information is available about the Incidence of RSV infection in children in Sudan. In fact there is no previous study have provided significant insight in to the detection and typing of RSV circulating in Sudan. We believe that it is important to compare our strain with other strains in data base to have an effective diagnosis, prevention, and treatment solution for RSV in near future. The current study was designed to detect and type RSV by molecular techniques and to describe identification of G protein novel mutations that were significantly change the structure and functions of the RSV strains in children hospitalized with respiratory tract infection in Sudan.

2. MATERIALS AND METHODS

2.1. Study Area
This study was conducted in Khartoum State of Central Sudan. Patients involving children under 5 years of age with respiratory tract infection (RTI) seen at the emergency department in Khartoum Hospitals (Gaffer Ebinuff and omdorman children hospitals), were recruited in winter seasons 2011-2012.

2.2. Data Collection:
Demographic data of the patients were collected using a structured questionnaire, which included the following criteria: Age, gender, and clinical symptoms (pneumonia, bronchiolitis, history of asthma and allergy). The specimens were collected within 1-7 days of disease onset.

2.3. Collection of specimens:
A total of 224 throat swab samples were collected from patients by using sterile nylon swabs (Regular Flocked swab, Cat. No. 520CS01, Copan Diagnostics Inc., Murrieta, Calif, USA), in 3 ml of transport media (UTM-RT, Cat.No. 92562, Copan Diagnostics Inc., Murrieta, Calif, USA). The collected samples were transported on ice bags at the same day of collection to the Department of Virology, Central Laboratory (Ministry of Science and Technology) and stored at −80°C until processed.

2.4. Direct immunofluorescence assay (DFA)
DFA was carried out using specific fluorescein conjugated monoclonal antibody (ARGENE Respiratory Syncytial Virus DFA kit, Varialhes, France) in accordance with the manufacturer's instructions, for the detection and identification of RSV antigen in direct respiratory specimen.

2.5. RNA extraction and cDNA synthesis
RNA was extracted using the Riboprep nucleic acid extraction kit (Ecoli.s.r.o., Bratislava, Slovak Republic), in accordance with the manufacturers protocol. The RNA was extracted from 100 μl of specimen and 10 μL of internal control was added to each sample. cDNA synthesis was performed using Reverta-L reagents kit, according to the protocol of the manufacturer (Ecoli.s.r.o., Bratislava, Slovak Republic) under the following conditions: The tested tubes were placed in the thermocycler and incubated at 37°C for 30 min. The cDNA samples were diluted by adding 20 μl from DNA-buffer.

2.6. RT-PCR for RSV
The reverse transcription- polymerase chain reaction (RT-PCR) assay targeting G gene was performed using a commercial kit following the manufactures instructions (hRSV-Eph-PCR kit, Ecoli, Slovak Republic). The reaction was conducted in a thermocycler using the following protocol; 95°C for 5 min for initial denaturation, 42 cycles of 95°C for 45 s, 56°C for 45 s, and 72°C for 45 s, final extension at 72°C for 1 min. All the amplified products were subjected to 1.5% agarose gel electrophoresis. Positive specimens for RSV resulted in PCR fragments of 298 bp.

2.7. Virus Isolation
Viral culture was done to the positive RSV samples by RT- PCR, in which 1ml of each sample was inoculated in HEp-2 monolayer(60%–70% confluent) cultures,obtained from (Vircell, Granada, Spain) and was maintained in modified Eagle's medium supplemented with 2% fetal bovine serum and 1% penicillin-streptomycin. Following adsorption, fresh medium was added, and inoculated cells were observed daily for cytopathic effect(CPE) (syncytia) for up to 7 days. Virus was harvested when the monolayer demonstrated approximately 75% CPE then stored at −80°C. RT- PCR for (RSV) was then carried out to confirm virus identity.

2.8. Nested RT-PCR
The Nested RT-PCR was carried out for the RT-PCR positive (RSV) samples including cell culture isolate using primer pairs located on the RSV G gene (Oligomicrogen, Korea). The following primers were used (24):

http://crmb.aizeonpublishers.net/content/2016/3/crmb874-883.pdf
G1- CCA TTC TGG CAA TGA TAA TCT C
G2- GTT TTT TGT TGG GTA TTC TTT TGC GA
G3- CGG CAA ACC ACA AAG TCA CAC
G4- GGG TAC AAA GTT AAA CAC TTC

The first round of association was performed by phylogeny.fr, and the sequences were achieved from NCBI and subjected to multiple sequence alignment using BioEdit software. Protein PDB built online by phylogeny.fr. The nine sequences have been deposited with the Gene Bank under accession numbers (AB979187, AB979188, AB979189, AB979190, AB979191, AB979192, AB979193, KP853084 and KP853085).

2.10. Statistical analysis
The collected data was analyzed using statistical package for social science (SPSS) version 20, Chi square test was used, a p-value of <0.05 was considered significant.

3. RESULTS AND DISCUSSION
3.1. Detection of RSV by DFA
During the study period, 224 children patients were enrolled. Out of these, 151 children were under 2 years of age and 73 were between 2-5 years of age. RSV infections were detected in 136 (60.7%) of these patients, using DFA technique (Table 1, Fig 1).

The figure show percentage of positive and negative RSV cases detected by direct immunofluorescence assays (DFA) and reverse transcription-polymerase chain reaction (RT-PCR).

3.3. Detection of RSV by Nested- PCR
34 throat swab samples were showed to be positive by the three techniques (tissue culture, DFA and RT-PCR) were confirmed using Nested-PCR, out of this 14 (41.2%) were positive and 20 (58.8%) were negative (Table 2, fig 2).

Table 1: Age-related Respiratory Syncytial Virus Infection in Children in Khartoum state using DFA and RT-PCR techniques.

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>DFA Positive</th>
<th>DFA Negative</th>
<th>DFA Total</th>
<th>RT-PCR Positive</th>
<th>RT-PCR Negative</th>
<th>RT-PCR Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;2 years</td>
<td>92</td>
<td>59</td>
<td>151</td>
<td>30</td>
<td>121</td>
<td>151</td>
</tr>
<tr>
<td>% Of Total</td>
<td>41.1%</td>
<td>26.3%</td>
<td>67.4%</td>
<td>13.4%</td>
<td>54.0%</td>
<td>67.4%</td>
</tr>
<tr>
<td>2-5 years</td>
<td>44</td>
<td>29</td>
<td>73</td>
<td>14</td>
<td>59</td>
<td>73</td>
</tr>
<tr>
<td>% Of total</td>
<td>19.6%</td>
<td>12.9%</td>
<td>32.6%</td>
<td>6.2%</td>
<td>26.3%</td>
<td>32.6%</td>
</tr>
<tr>
<td>Total</td>
<td>136</td>
<td>88</td>
<td>224</td>
<td>44</td>
<td>180</td>
<td>224</td>
</tr>
<tr>
<td>% Of total</td>
<td>60.7%</td>
<td>39.3%</td>
<td>100%</td>
<td>19.6%</td>
<td>80.4%</td>
<td>100%</td>
</tr>
<tr>
<td>P-value</td>
<td>0.925</td>
<td>-</td>
<td>0.903</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 1: Detection of Respiratory Syncytial Virus infection in children by RT-PCR and DFA.

Table 2: Age-related Respiratory Syncytial Virus Infection of Children in Khartoum state using Nested-PCR.

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Positive</th>
<th>Negative</th>
<th>Total</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;2 years</td>
<td>7</td>
<td>16</td>
<td>23</td>
<td>0.066</td>
</tr>
<tr>
<td>% of total</td>
<td>20.6%</td>
<td>47.1%</td>
<td>67.6%</td>
<td></td>
</tr>
<tr>
<td>2-5 years</td>
<td>7</td>
<td>4</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>% of total</td>
<td>20.6%</td>
<td>11.8%</td>
<td>32.4%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>20</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>% of total</td>
<td>41.2%</td>
<td>58.8%</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Nested-PCR amplification of G gene of RSV.

Nested-PCR amplification of G gene of RSV on 1.5% agarose gel electrophoresis. Lane M DNA ladder: MW 100 bp. Lane PC; positive control. Lane NC; negative control, Lane 1, 2, 6 and 7 showing typical band size of (326pb) corresponding to the molecular size of G genes, 3, 4, 5 and 8 negative samples.

3.4. Nucleotide sequence and phylogenetic analysis of RSV

Successful sequencing of HRSV based on G gene was done for nine samples. The nucleotide sequences and their accession numbers in gene bank are shown in Table 3.

Multiple sequence alignment of RSV sequences obtained from Sudanese isolates with sequences that obtained from data base showed, transversions mutation in isolate176, in which A was replaced by T at position 31, substitution of the codon AGC to TGC that resulted in substitution of amino acid serine (S) to Cysteine (C) at position 11 (Fig. 3). In isolate 316 also there was a transversions mutation in which C was
replaced by A at position 220, a substitution of the codon AAC to AAA resulted in substitution of amino acid Asparagine (N) to lysine (K) in isolate 316 at position 73 (Fig. 3). Insertion of G at position 84 and A at position 103 in isolate 377 cause frame shift mutations (Fig. 3).

### Table 3: Sudanese Respiratory Syncytial virus (RSV) Sequences and Related information.

<table>
<thead>
<tr>
<th>Sample name</th>
<th>Date of collection</th>
<th>Age</th>
<th>Gender</th>
<th>Genebank accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isolate 17</td>
<td>29-12-2011</td>
<td>4 y</td>
<td>M</td>
<td>AB979187</td>
</tr>
<tr>
<td>Isolate 127</td>
<td>29-01-2011</td>
<td>4 y</td>
<td>M</td>
<td>AB979188</td>
</tr>
<tr>
<td>Isolate 129</td>
<td>29-01-2011</td>
<td>2 y</td>
<td>F</td>
<td>AB979189</td>
</tr>
<tr>
<td>Isolate 316</td>
<td>16-01-2012</td>
<td>5 y</td>
<td>F</td>
<td>AB979190</td>
</tr>
<tr>
<td>Isolate 354</td>
<td>07-02-2012</td>
<td>2 m</td>
<td>F</td>
<td>AB979191</td>
</tr>
<tr>
<td>Isolate 372</td>
<td>11-02-2012</td>
<td>3 m</td>
<td>M</td>
<td>AB979192</td>
</tr>
<tr>
<td>Isolate 395</td>
<td>19-02-2012</td>
<td>1 y</td>
<td>M</td>
<td>AB979193</td>
</tr>
<tr>
<td>Isolate 176</td>
<td>20-02-2012</td>
<td>2 y</td>
<td>M</td>
<td>KP853084</td>
</tr>
<tr>
<td>Isolate 377</td>
<td>23-02-2012</td>
<td>3 y</td>
<td>M</td>
<td>KP853085</td>
</tr>
</tbody>
</table>

![Figure 3: Multiple sequence alignment.](http://crmb.aizeonpublishers.net/content/2016/3/crmb874-883.pdf)

(A) BioEdit multiple sequence alignment, the transversion mutations in isolate 176, 316 and the inserted G and A in isolate 377 illustrated by the arrows. (B) Amino acid multiple sequence alignment of Sudanese mutant G genes compared to other G genes from database. Substitution of the amino acid serine (S) to cystine (C) in isolate 176. And substitution of amino acid Asparagine (N) to lysine (K) in isolate 316. The insertion of G and A in isolate 377 illustrated by the arrows. The alignment was performed using the ClustalW2 sequence alignment.

### 3.4.1. Protein tertiary structure

The prediction of 3D structure of the wild and mutant G genes of isolate 176 revealed that the mutant residue is more hydrophobic than the wild-type residue. This mutation is probably damaging to the protein. The mutant residue hasn’t similar properties observed at this position in other homologous sequences (Fig. 4).

The prediction of 3D structure of the wild and mutant G genes of isolate 316 revealed that the mutant residue is bigger than the wild-type residue. The wild-type residue was neutral, while the mutant residue is positively charged. This mutation can occur at this position and is probably not damaging to the protein. The mutant residue is located near a highly conserved position (Fig. 5).
Figure 4: Tertiary protein structure of wild and mutant G genes of Isolate 176
The wild type of predicted amino acid serine (S) at position 11 of G gene (Isolate 176) was replaced by mutant type of amino acid cysteine (C) illustrated by the arrows. Tertiary protein structure of wild and mutant genes was drawn by Chimera software version 1.9.

Figure 5: Tertiary protein structure of wild and mutant G genes of Isolate 316.
Left: wild type (A) of predicted amino acid Asparagine (N) at position 73 of G gene (Isolate 316). Right: mutant type (B) of predicted amino acid Lysine (K) at position 73. (A, B) Tertiary protein structure of wild and mutant genes that drawn by Chimera software version 1.9.
3.4.2. Phylogenetic tree

Phylogenetic analysis of nucleotide sequences revealed that all analyzed RSV belonged to the RSV-A strain (Fig. 6). The nucleotide sequences were compared with other sequences published in GenBank. Phylogenetic tree that were generated for RSV subtype-A are shown in Figure 6.

Figure 6: Phylogenetic analysis of the RSV G gene.

Alignment of G gene sequences was undertaken with BioEdit software. Sequences were aligned with 8 sequences of RSV subgroup A downloaded from Genbank. Red dots indicate the Sudanese RSV-A mutant strains. The phylogenetic tree was designed using phylogeny.fr software.

This study provides the first report on molecular detection and characterization of RSV in Sudan. We analyzed G genes of HRSV isolates from throat swab samples tested during winter 2011–2012 in Sudan. All sequences were analyzed using various bioinformatics methods in order to understand genotype variability, molecular epidemiology and evolutionary adaptability of circulating strains.

Two hundred and twenty four samples were screened for RSV infections using DFA, in which 136 (60.7%) positive cases were detected. This result is slightly different from a results reported by LRA Vaz-de-Lima [31], in which 113 (36.9%) specimens were found to be positive HRSV by DFA. The differences between their result and ours may be due to the lees sensitivity of DFA technique [32, 33]. In addition, there is differential detection related to viral load. Infants usually have severe RSV disease associated with higher viral load hence a better chance of detection [34].

Out of 224 samples were collected from children with RTI and significantly high rate of RSV infection was detected by RT-PCR, whereby 44 [19.6%] were reported as positive cases. Comparable results were recorded in Egypt in which, RSV was detected in 16.4% of the cases [35]. RT-PCR has shown improved sensitivity in the detection of RSV infection [32, 33]. The high rate of RSV in this work explains the increase in RTI cases during winter seasons in Khartoum state.

Our study shows that, out of 34 cases 14 (41.2%) were positive by Nested-PCR. This is parallel to previous study [36] were 15 out of 45 samples was positive by nested PCR. Nested-PCR provides a sensitive tool for both detection and typing of RSV into groups. In
addition, Nested-PCR products can be further characterized to study RSV molecular epidemiology.

The results reported here represent the first genetic characterization of RSV mutations in Sudan. Transversion and inserted mutations in G gene was identified in three of the Sudanese strains (Isolate 176, 316 and 377). Several studies of RSV strains have described many mutations in G gene in different positions [37, 38 and 39], suggesting that the G gene has a high nucleotide sequence variability correlated with an elevated amino acid substitution rate. In addition, the G genes showed a high frequency of nucleotide insertions and deletions within the G ectodomain. The highly variable and immunogenic G protein has been previously suggested to be susceptible to neutralizing antibodies; this might drive the selection of strains carrying immune escape mutations within the ectodomain [40].

The current study showed that, the mutation in isolate 176 introduces a more hydrophobic residue. This is probably damaging the protein and can result in loss of hydrogen bonds and/or disturb correct folding [30]. Our results showed that, in isolate 316 the wild-type and mutant amino acids differ in size. The mutant residue is bigger, this might lead to bumps. Moreover, there is a difference in charge between the wild-type and mutant amino acid. The mutation introduces a charge; this can cause repulsion of ligands or other residues with the same charge [30].

The sequence data indicate that one of the Sudanese strains (Isolate 377) originated by insertions of A and G in the G protein gene, generating frame shift mutations responsible for resistant phenotype. Various studies reported frame shift mutation in G gene [41, 42]. Moreover, we recognized two reading frame shifts in isolate 377; this can create viruses with fitness analogous to that of the parental strain [41].

We documented circulation of one genotype (RSV-A) in Sudan during winter 2011–2012. Previous studies have also documented that Subtype A usually is the predominant circulating strains compared with subtype B [43, 44].

Five Sudanese RSV strains are grouped together, this clarify that the Sudanese strains are similar to each other and differ from other strains found in gene bank. Three strains were closer to the previously described strain in Belgium, a fact that is hard to explain, in view of the fact that Belgium is far than many other European countries in which other different RSV strains are circulating [45, 46 and 47]. Furthermore, one strain (Isolate 377) was far from the other Sudanese strains and belongs to separated branch, this was due to the generating of frame shift mutations which resulted in a unique strain. Moreover, no identical sequences to the Sudanese isolates could be found in Gene Bank. This is plausible because limited sequences data is available at present on the prevalence of RSV in Africa, which is more likely to share the same genotypes.

In this application for detection and typing of RSV in throat swab samples, we sought increased sensitivity. A combination of the two techniques (RT-PCR and Nested-PCR) was of value in increasing the sensitivity of detection. These findings highlight the need for the molecular detection and typing of RSV in Sudan for better management of respiratory infections in children.

Further work must be done on molecular genotyping of RSV in Sudan to have implications for the utilization of certain attenuated strains as vaccine candidates in children.

4. CONCLUSION

The results of this study emphasize the importance of early detection and characterization of newly emerging genotypes. Understanding the effect of the novel RSV-A genotype on fitness, virulence and transmissibility could help predict changes in viral phenotype and immunogenicity. It will also provide insight into vaccine potential of the G gene protein. Continued genotyping and molecular epidemiological surveillance of RSV are essential to further understanding RSV evolution and transmission in communities and healthcare settings.

5. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Staff of Khartoum Hospital for allowing us to collect the samples, and Central lab for their help in practical work. We are grateful to the children and parents that approved to provide samples for this study from Khartoum hospital.

6. REFERENCES

8. McLeLan JS, Chen M, Chang JS, Yang YP, Kim A, Graham BS, Kwong PD. (2010). Structure of a major antigenic site on the...


© 2016; AIZEON Publishers; All Rights Reserved

This is an Open Access article distributed under the terms of the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.