Archives of Virology (2024) 169:106
https://doi.org/10.1007/s00705-024-06036-0

BRIEF REPORT

Genomic characterization of human respiratory syncytial virus circulating in Islamabad, Pakistan, during an outbreak in 2022-2023

Syed Adnan Haider[1] · Zunera Jamal[1] · Faheem Tahir[1] · Muhammad Salman[1] · Massab Umair[1]

Received: 27 November 2023 / Accepted: 26 February 2024

Abstract

In this study, conducted at the National Institute of Health, Islamabad, during an outbreak of human respiratory syncytial virus (hRSV) from December 2022 to January 2023, the first whole-genome sequences of hRSV isolates from Islamabad, Pakistan, were determined. Out of 10 positive samples, five were sequenced, revealing the presence of two genotypes: RSV-A (GA2.3.5, ON1 strain) and RSV-B (GB5.0.5.a, BA-10 strain). A rare non-synonymous substitution (E232G) in G the protein and N276S in the F protein were found in RSV-A. In RSV-B, the unique mutations K191R, Q209R, and I206M were found in the F protein. These mutations could potentially influence vaccine efficacy and viral pathogenicity. This research underscores the importance of genomic surveillance for understanding RSV diversity and guiding public health responses in Pakistan.

Introduction

Human respiratory syncytial virus (hRSV) is a respiratory pathogen that affects the lungs and breathing passages, often leading to hospital admissions in severe cases. It is the second leading cause of deadly acute lower respiratory tract infections (ALRTIs) in children under five years old [1]. In 2019, hRSV was associated with approximately 25.4 to 44.6 million infections, 2.9 to 4.6 million hospital admissions, and between 15,100 and 49,100 fatalities in hospitalized children less than 2 years old [2]. RSV infections place a considerable strain on public health systems and the global economy, with their impact on young children surpassing that of influenza [3].

RSV is an enveloped virus with a single-stranded negative-sense RNA genome of approximately 15.2 kb that encodes 11 proteins. The glycoprotein G and fusion protein F are essential for entry of the virus into host cells. RSV has two main subtypes, RSV-A and RSV-B, which are distinguished primarily by variations in their G and F proteins and their genetic makeup [4]. These subtypes are further divided into 45 genotypes based on genetic differences in hypervariable region 2 (HVR2) of the G protein. RSV-A has 15 genotypes, including GA1-7, NA1-4, ON1-2, SAA1, and CBA, while RSV-B has 30 genotypes, including GB1-4, BA1-14, BAc, SAB1-4, URU1-2, CB1 (GB5), CBB, BA-CCA, BA-CCB, and THB. Since 2015, the dominant global genotypes have been ON1 (RSV-A) and BA9 (RSV-B) [5]. In an outbreak in the USA in 2022-2023, the prevalent strains belonged to RSV-A genotypes GA2.3.5 and GA2.3.6b (both ON1) and RSV-B genotype GB5.0.5.a (BA) [6].

Pakistan, a developing country with one of the highest infant mortality rates, lacks comprehensive RSV genomic surveillance and disease burden data [7]. While some molecular epidemiological data are available [8–10], there is still a notable knowledge gap, particularly regarding the complete genome sequences of circulating strains. Previous research has mainly focused on partial sequences of RSV-A isolates, particularly the G protein, revealing the prevalence of the GA2/NA1 genotype during the 2010-2013 [9, 10]. In this study, for the first time, we determined the complete genome sequences of five RSV isolates from Pakistan.

Materials and methods

From Dec 2022 to Jan 2023, throat/nasopharyngeal swabs from 10 RSV-positive patients at a tertiary care hospital in Islamabad were sent to the Department of Virology at the National Institutes of Health (NIH), Islamabad, for analysis. The samples were collected in viral transport medium

Handling Editor: Bert K. Rima

✉ Massab Umair
massab.umair@yahoo.com

[1] Department of Virology, National Institute of Health, Park Rd, Chak Shahzad, Islamabad 45500, Pakistan

Published online: 22 April 2024

(VTM), stored at 4°C, and processed for whole-genome sequencing within 24 hours. The study was conducted with the approval of the Institutional Review Board of the National Institutes of Health.

Total viral RNA was extracted using a MagMAX Viral/Pathogen Nucleic Acid Isolation Kit and a KingFisher Flex Purification System (Thermo Fisher Scientific), following the manufacturer's guidelines. For whole-genome sequencing, the extracted RNA was prepared for unbiased paired-end sequencing (2 × 74 bp) using the Illumina RNA prep method with a Respiratory Virus Oligo Panel v2 (RVOP2) Enrichment Kit (Illumina Inc), following the guidelines provided by the manufacturer. The libraries were sequenced using an Illumina MiSeq system, employing a MiSeq Reagent Kit v3 (150 cycles).

The quality of the raw NGS reads was assessed using FastQC v0.11.9. Low-quality sequences and adapters were removed using Trimmomatic v0.39, followed by removal of duplicates, using the PICARD tool's 'MarkDuplicates' function. Contigs were assembled using SPAdes v3.15.5 [11] and aligned using the Burrows-Wheeler Aligner (BWA) against the NCBI Non-redundant (NR) database. Consensus genome sequences were generated using Geneious Prime v2022.2 with a minimum coverage >10 [12]. The genome sequences were uploaded to GISAID under the accession numbers EPI_ISL_18089334, EPI_ISL_18094390, EPI_ISL_18094391, EPI_ISL_18094392, and EPI_ISL_18094393.

RSV genotypes and clades were identified using Nextclade v2.14.1 (Goya et al. classification) [13]. For phylogenetic analysis, four complete RSV-A genome sequences were used in a BLAST search to identify closely related sequences in the NCBI GenBank database as of August 25, 2023. Sequences with more than 2% N's or nearly identical matches were excluded. Multiple sequence alignment was made using MAFFT, and the general time-reversible model (GTR+G) was selected in MEGA11 for phylogenetic analysis. Phylogenetic tree was constructed using IQ-TREE (v2.2.0) with 1000 bootstrap replicates [14] and visualized using FigTree (v1.4.4).

To identify amino acid sequence variations, the G and F regions were aligned separately using ClustalW, translated into amino acids using MEGA11, and compared with RSV-A (GenBank accession no. NC_038235.1) and RSV-B (GenBank accession no. NC_001781.1) reference sequences.

Results

Between December 2022 and January 2023, NIH received throat/nasopharyngeal swabs from 10 RSV-positive individuals from a tertiary care hospital in Islamabad, five of which were successfully sequenced. Common symptoms included sore throat, chest pain, shortness of breath, and fever, accompanied by sputum production and nasal congestion. Additional metadata are presented in Table 1.

Four of the isolates were of the RSV-A subtype, identified as genotype GA2.3.5 (ON1 strains), with 96.0% to 98.3% coverage, and the other was of the RSV-B subtype, identified as genotype GB5.0.5.a (BA-10 strain), with 98.3% coverage, according to Nextclade (v2.14.1) (Fig. 1). Notably, the RSV-A isolates exhibited a high degree of sequence similarity (99% to 99.85% identity) to strains detected in the USA in 2022 (NCBI accession numbers OP890340.1, OR143212.1, and OQ024128.1). One of the patients with RSV-A (GISAID EPI_ISL_18094391) was found to be coinfected with SARS-CoV-2, and another (EPI_ISL_18094392) was found to be coinfected with Human adenovirus C2 (HAdV-C2).

A comparison of the predicted amino acid sequences of the G and F proteins of the RSV-A (GA2.3.5 genotype; ON1 strains) and RSV-B (GB5.0.5.a genotype, BA-10 strain) isolates to those of reference sequences (NC_038235.1 for RSV-A and NC_001781.1 for RSV-B) revealed 32 mutations in the G protein and 21 in the F protein of RSV-A and 24 mutations in the G protein and 10 in the F protein of RSV-B. The rare amino acid

Table 1 Metadata for RSV samples analyzed via whole-genome sequencing

Lab ID	GISAID ID	Age	Gender	Date of Collection	District	Symptoms	Genotype	Subclade (Goya et al.)
RSV-1	EPI_ISL_18089334	18 days	Female	22-Dec-22	Islamabad	Sore throat, sputum production, chest pain, shortness of breath	A	GA2.3.5 (ON1)
RSV-2	EPI_ISL_18094390	7 years	Female	06-Jan-23		Not available		
RSV-3	EPI_ISL_18094391	4 months	Male	26-Dec-22		Nasal congestion, sore throat, tachypnea, abnormal breath sounds, shortness of breath		
RSV-4	EPI_ISL_18094392	3 months	Female	28-Dec-22		Chest pain		
RSV-5	EPI_ISL_18094393	10 days	Male	26-Dec-22		Shortness of breath	B	GB5.0.5.a (BA-10)

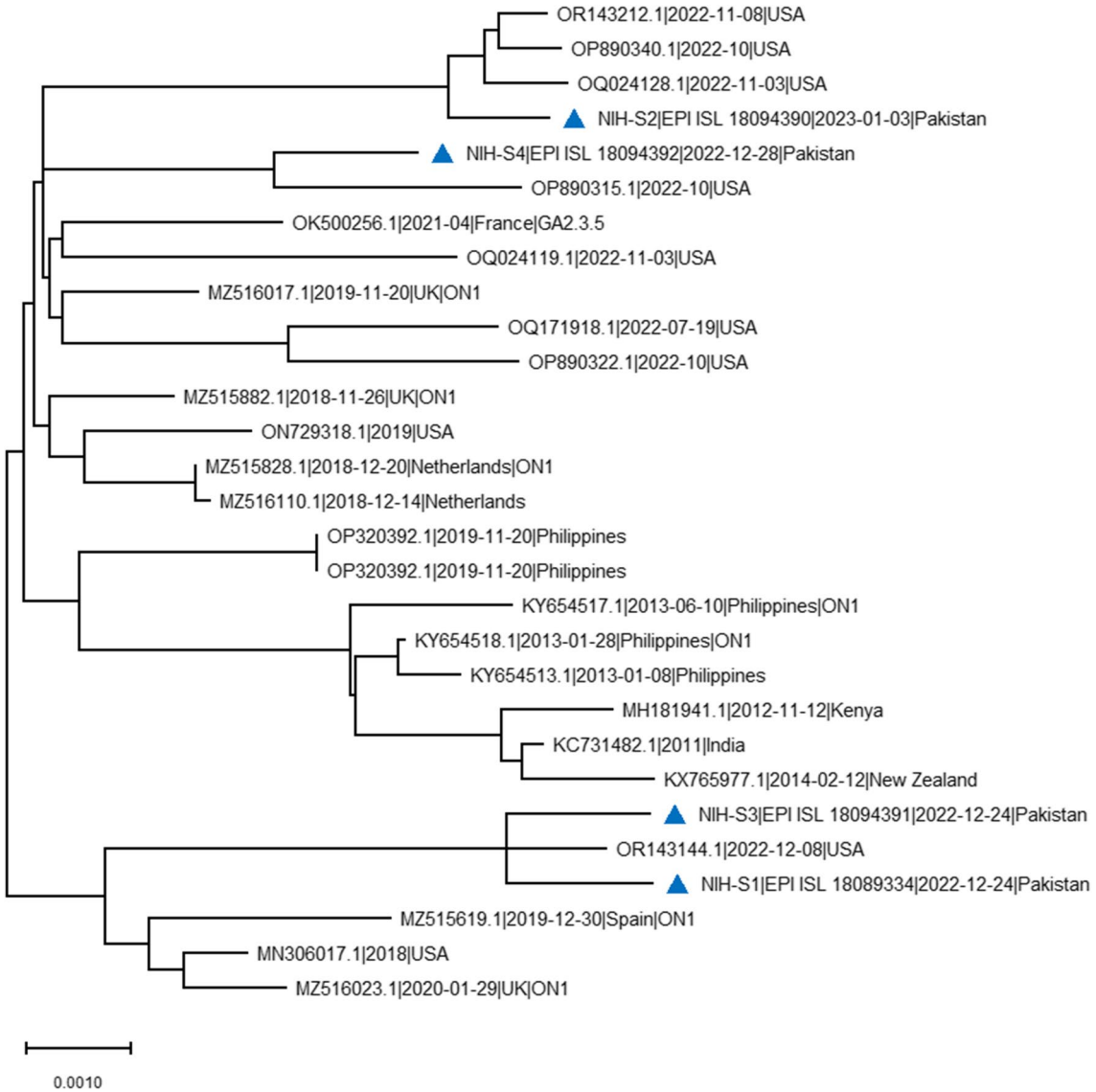

Fig. 1 Unrooted phylogenetic tree based on complete genome sequences of RSV-A isolates. The tree was constructed in MEGA 11 by the maximum-likelihood method using IQ-TREE software, with the general time-reversible model (GTR+G) with 1000 bootstrap replicates. Sequences obtained from samples collected in Islamabad, Pakistan, during 2022 and 2023 are indicated by blue triangles.

substitution E232G in the G protein, which has been observed only twice worldwide (as of August 29, 2023, according to GISAID data), was observed in one RSV-A sequence (EPI_ISL_18089334), and the substitution N276S in the F protein was found in two RSV-A strains (EPI_ISL_18094390, EPI_ISL_18094392). Likewise, our RSV-B isolate contained the rare mutations K191R and I206M in the F protein, which have been reported twice worldwide (as of August 29, 2023), with Q209R being unique to our Pakistani RSV-B sequence EPI_ISL_18094393 (Fig. 2).

Discussion

RSV presents a significant global challenge, especially in developing countries, causing severe respiratory illness. The number of RSV-related fatalities in children under 5 is estimated to range from 66,000 to 199,000 annually, contributing to 15-40% of pediatric hospitalizations for acute respiratory infection worldwide [15, 16]. Pakistan is a developing country with one of the highest infant mortality rates [7]. Data from 2005 indicated a 20% (n = 80) RSV infection rate in Islamabad, and in a three-year study

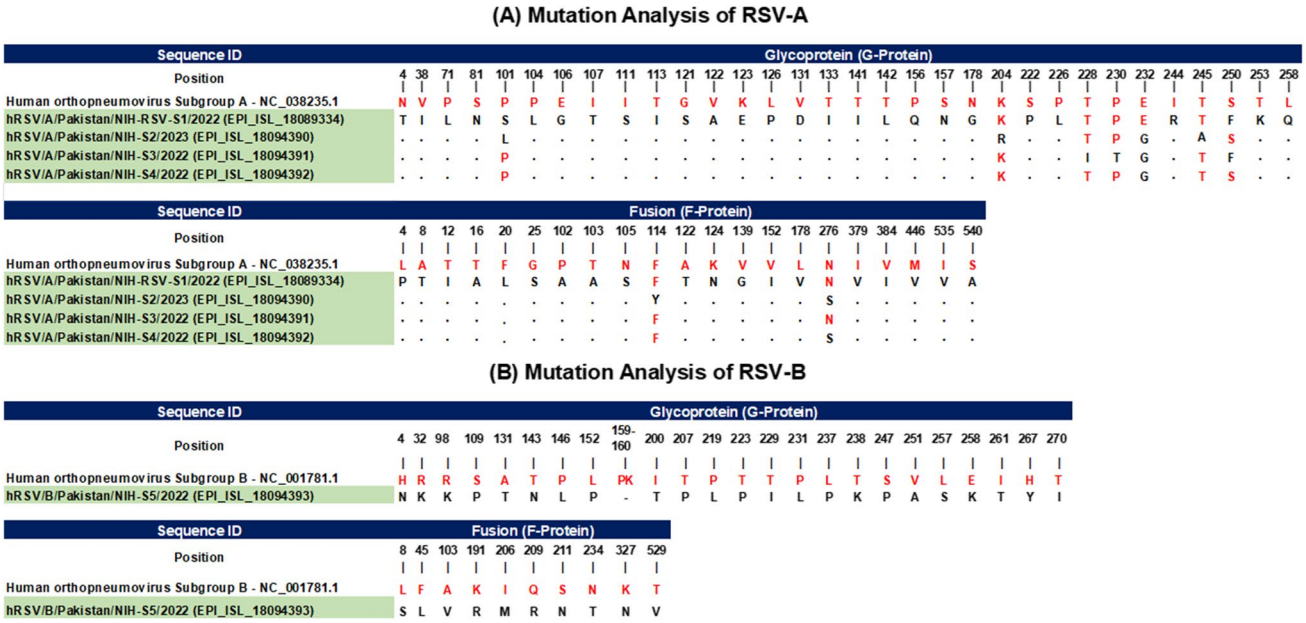

Fig. 2 Nonsynonymous mutations in the G and F proteins of Pakistani RSV-A (EPI_ISL_18089334, EPI_ISL_18094390, EPI_ISL_18094391, EPI_ISL_18094392) and RSV-B (EPI_ISL_18094393) isolates. Pakistani samples are highlighted in green, with dots indicating shared mutations differing from the reference sequence. The figure displays only positions where mutations are present.

(2009-2012) in Karachi, 223 cases were reported [17, 18]. Another study in Karachi from November 2010 to September 2011 that reported an infection rate of 17.8% (n = 30) emphasized the importance of annual surveillance for RSV [19]. There were 320 cases of RSV reported in 2019 (pre-COVID-19) and only 21 cases in 2020 in Pakistan [20]. Previous reports of RSV strains in Pakistan have provided only partial genome sequence data [8–10]. As of September 15, 2023, over 100 RSV G protein sequences had been recorded in the NCBI GenBank database, but complete genome sequences of isolates from Pakistan are lacking. Our study, therefore, helps to fill this vital gap by providing the first complete RSV genome sequences from Pakistan, including those of four RSV-A isolates and one RSV-B isolate obtained from an outbreak in Islamabad from December 2022 to January 2023.

In Pakistan, there is a notable gap in the literature on RSV-A genotypes from 2013 to present. Earlier studies (2010-2013) reported the prevalence of the GA2.3.3 (GA2/NA1) genotype of RSV-A [9, 10]. In our study, we identified ON1 strain, but the lack of data from 2014 to 2021 makes it difficult to determine when it emerged in Pakistan. In the case of RSV-B, we identified the GB5.0.5.a (BA-10) genotype. Historically, RSV-B strains in Pakistan were classified as genotype BA, which has been prevalent globally since its 1999 discovery in Argentina [21].

Phylogenetic analysis showed that the GA2.3.5 genotype (ON1 strains) of RSV-A and the GB5.0.5a genotype (BA-10) of RSV-B were present in Pakistan from 2022 to 2023. These isolates showed 99% to 99.85% nucleotide sequence identity to isolates from the 2022 Washington State, USA, outbreak, in which over 126,000 cases involving similar strains were reported [22, 23]. This surge in cases was suggested to be due to diminished population immunity resulting from low RSV exposure during the COVID-19 pandemic [6]. Similarly, Pakistan observed a 93.4% decrease in RSV cases during the pandemic from 2019 to 2020 [20]. Furthermore, genome-level analysis, including pre and post-COVID-19 comparisons, showed the consistent presence of the amino acid substitutions A103T and T122A in the RSV-A F protein, suggesting that the COVID-19 pandemic mitigation measures may have only influenced seasonality [6, 24]. Despite the lack of pre-pandemic F protein sequence data from Pakistan (as of Dec 8, 2023), these mutations were also found in our samples, underscoring the need for continuous surveillance to monitor seasonal variations in RSV.

The pre-fusion conformation of the F protein, which is a key target of the Arexvy and Abrysvo vaccines, contains key antigenic sites that are important for immune defense [24–26]. Host antibodies target six antigenic sites in the F protein, denoted as sites Ø-V [27]. In our study, significant amino acid substitutions were found in the RSV-A F protein, including N276S (in samples EPI_ISL_18089334 and EPI_ISL_18094391), which affects antigenic site II, and I379V (in all four RSV-A samples) which affects site I. The N276S mutation, had been reported twice globally as of Sep

15, 2023, and it might reduce the neutralizing activity of palivizumab [28, 29]. The isolate RSV-B EPI_ISL_18094393 contains a K191R mutation within antigenic site V, which is also rare globally and is linked to increased pathogenicity [30]. The mutations I206M and Q209R in RSV-B enhance neutralization by nirsevimab and MEDI8897, emphasizing their potential role in vaccine effectiveness [31, 32]. Understanding such mutations is crucial for assessing their impact on the host defense against RSV infection and developing vaccines and therapeutic agents.

Sequencing the entire RSV genome provides more comprehensive information than sequencing only the F and/or G gene. The F and G genes are the main targets of the host immune response, which causes them to mutate frequently [33, 34]. Therefore, these genes alone may not be sufficient for diagnostics. In contrast, the N (nucleoprotein) and M (matrix) genes, which are more conserved, may be more suitable as targets for developing molecular assays [35, 36]. The complete genome sequences of RSV strains circulating in Pakistan reported here may therefore serve as valuable references for future studies and provide information about mutations in other essential genes, such as M2-1, M2-2, NS1, and NS2, which are involved in viral replication, transcription, and immune evasion [37] and may also be of interest in the development of antiviral drugs and vaccines against RSV.

In this study, the RVOP v2 panel detected coinfections (RSV with HAdV-C2 in a three-month-old child and with SARS-CoV-2 in a four-month-old child). While RSV and HAdV coinfections in young children often cause severe disease, clinically, they resemble single infections [38, 39]. It has been reported that coinfection with RSV and SARS-CoV-2 results in symptoms such as nasal congestion and sore throat but generally does not lead to severe outcomes, and this was also the case with our coinfected patient [40]. These cases highlight the importance of enhanced NGS capability to detect multiple viruses simultaneously.

One limitation of this study was the small number of RSV-positive samples (n = 10) analyzed. Only half of these samples could be sequenced due to poor sample quality and a high Ct value (>28). Another limitation was the lack of a national surveillance or diagnostic system for RSV in Pakistan like the one that has been established for influenza. This resulted in a lack of RSV data, especially for the study period in Islamabad. We therefore recommend integrating RSV testing into the routine influenza surveillance program, using available RT-PCR kits that allow testing of multiple pathogens.

The current study provides the first whole-genome sequence data of RSV from the 2022-2023 outbreak in Pakistan, demonstrating the presence of the genotypes GA2.3.5 (ON1 strains) of RSV-A and GB5.0.5.a (BA-10 strain) of RSV-B. Notably, these same genotypes were also reported in an outbreak in Washington, USA, in 2022. The detection of both RSV types and their presence in coinfections highlights the critical need for precise laboratory diagnostics and robust genomic surveillance. The discovery of significant mutations in antigenic sites in the F protein suggests the importance of RSV vaccination in our immunization program [18]. More local efforts need to be made to generate RSV genomic surveillance data and establish robust surveillance systems for monitoring outbreaks and informing public health policies in Pakistan.

Author's contribution Conceptualization: M.U, M.S, and F.T. Methodology: S.A.H and Z.J. Formal analysis: S.A.H and Z.J. Resources: M.U, M.S, and F.T. Writing—original draft preparation: S.A.H and Z.J. Writing—review and editing: M.U, S.A.H, and Z.J. All authors have read and agreed to the published version of the manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this article.

Data availability The RSV datasets generated in the current study are available in the GISAID repository under the accession numbers EPI_ISL_18089334, EPI_ISL_18094390, EPI_ISL_18094391, EPI_ISL_18094392, and EPI_ISL_18094393.

Declarations

Ethical approval This study was approved by the Institutional Review Board of the National Institute of Health (NIH), Islamabad.

Patient consent Informed consent was obtained from all individual participants included in the study.

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

1. Li Z-J et al (2021) Etiological and epidemiological features of acute respiratory infections in China. Nat Commun 12(1):5026
2. Li Y et al (2021) National burden estimates of hospitalisations for acute lower respiratory infections due to respiratory syncytial virus in young children in 2019 among 58 countries: a modelling study. Lancet Respir Med 9(2):175–185
3. Di Giallonardo F et al (2018) Evolution of human respiratory syncytial virus (RSV) over multiple seasons in New South Wales, Australia. Viruses 10(9):476
4. Mufson MA et al (1985) Two distinct subtypes of human respiratory syncytial virus. J Gen Virol 66(10):2111–2124
5. Rojo S et al (2023) Respiratory syncytial virus incidence and typing in the last six seasons in the north of Spain (Asturias). Genetic characterization during the SARS-CoV-2 pandemic. J Med Virol 96(3):e29499
6. Goya S et al (2023) Genomic Characterization of Respiratory Syncytial Virus during 2022–23 Outbreak, Washington, USA. Emerg Infect Dis 29(4):865
7. Pasha O et al (2015) Maternal and newborn outcomes in Pakistan compared to other low and middle income countries in the

Global Network's Maternal Newborn Health Registry: an active, community-based, pregnancy surveillance mechanism. Reprod Health 12:1–10
8. Bashir U et al (2017) Molecular detection and characterization of respiratory syncytial virus B genotypes circulating in Pakistani children. Infect Genet Evol 47:125–131
9. Aamir UB et al (2020) Molecular characterization of circulating respiratory syncytial virus genotypes in Pakistani children, 2010–2013. J Infect Public Health 13(3):438–445
10. Aamir UB 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(9):e74018
11. Bankevich A et al (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19(5):455–477
12. Kearse M et al (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28(12):1647–1649
13. Goya S et al (2020) Toward unified molecular surveillance of RSV: a proposal for genotype definition. Influenza Other Respir Viruses 14(3):274–285
14. Nguyen L-T et al (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32(1):268–274
15. McLellan JS et al (2010) Structural basis of respiratory syncytial virus neutralization by motavizumab. Nat Struct Mol Biol 17(2):248–250
16. Venter M et al (2001) Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. J Gen Virol 82(9):2117–2124
17. Tariq W-U-Z et al (2005) Winter peak of respiratory syncytial virus in Islamabad. Trop Doct 35(1):28–29
18. Ali A et al (2017) RSV associated hospitalizations in children in Karachi, Pakistan: implications for vaccine prevention strategies. J Med Virol 89(7):1151–1157
19. Ali A et al (2013) Role of human metapneumovirus, influenza A virus and respiratory syncytial virus in causing WHO-defined severe pneumonia in children in a developing country. PLoS ONE 8(9):e74756
20. Rana MS et al (2021) Impact of COVID-19 preventive measures on other infectious and non-infectious respiratory diseases in Pakistan. J Infect 82(5):e31–e32
21. Trento A et al (2010) Ten years of global evolution of the human respiratory syncytial virus BA genotype with a 60-nucleotide duplication in the G protein gene. J Virol 84(15):7500–7512
22. Rachael Zimlich R. (2023) BSN. RSV surges early in 2022, but appears to make early retreat. Available from: https://www.contemporarypediatrics.com/view/rsv-surges-early-in-2022-but-appears-to-make-early-retreat. Accessed 15 Sept 2023
23. Adams G et al (2023) Viral lineages in the 2022 RSV surge in the United States. N Engl J Med 388(14):1335–1337
24. Holland LA et al (2023) Genomic sequencing surveillance and antigenic site mutations of respiratory syncytial virus in arizona, USA. medRxiv pp 2023–06
25. Papi A et al (2023) Respiratory syncytial virus prefusion F protein vaccine in older adults. N Engl J Med 388(7):595–608
26. Kampmann B et al (2023) Bivalent prefusion F vaccine in pregnancy to prevent RSV illness in infants. N Engl J Med 388(16):1451–1464
27. Gilman MS et al (2016) Rapid profiling of RSV antibody repertoires from the memory B cells of naturally infected adult donors. Sci Immunol 1(6):eaaj1879
28. Choi S-H, Park KS, Kim Y-J (2019) Analysis of respiratory syncytial virus fusion protein from clinical isolates of Korean children in palivizumab era, 2009–2015. J Infect Chemother 25(7):514–519
29. Respiratory Syncytial Virus (RSV) There are FDA-approved vaccines and monoclonal antibodies for RSV now available. [09/11/2023 19 Oct 2023]; Available from: https://www.fda.gov/consumers/covid-19-flu-and-rsv/respiratory-syncytial-virus-rsv
30. Hotard AL et al (2015) Identification of residues in the human respiratory syncytial virus fusion protein that modulate fusion activity and pathogenesis. J Virol 89(1):512–522
31. Ahani B et al (2023) Molecular and phenotypic characteristics of RSV infections in infants during two nirsevimab randomized clinical trials. Nat Commun 14(1):4347
32. Tabor DE et al (2020) Global molecular epidemiology of respiratory syncytial virus from the 2017–2018 INFORM-RSV Study. J Clin Microbiol 59(1):10–1128. https://doi.org/10.1128/jcm.01828-20
33. McLellan JS, Ray WC, Peeples ME (2013) Structure and function of respiratory syncytial virus surface glycoproteins. Challenges and opportunities for respiratory syncytial virus vaccines 83–104
34. Tripp RA et al (2018) Respiratory syncytial virus: targeting the G protein provides a new approach for an old problem. J Virol 92(3):10–1128. https://doi.org/10.1128/jvi.01302-17
35. Jung HE, Kim TH, Lee HK (2020) Contribution of dendritic cells in protective immunity against respiratory syncytial virus infection. Viruses 12(1):102
36. Pickens JA, Tripp RA (2018) Verdinexor targeting of CRM1 is a promising therapeutic approach against RSV and influenza viruses. Viruses 10(1):48
37. Rios Guzman E, Hultquist JF (2022) Clinical and biological consequences of respiratory syncytial virus genetic diversity. Ther Adv Infect Dis 9:20499361221128092
38. Liu C et al (2015) Adenovirus infection in children with acute lower respiratory tract infections in Beijing, China, 2007 to 2012. BMC Infect Dis 15(1):1–9
39. Palomino MA et al (2004) Adenovirus and respiratory syncytial virus-adenovirus mixed acute lower respiratory infections in Chilean infants. Pediatr Infect Dis J 23(4):337–341
40. Swets MC et al (2022) SARS-CoV-2 co-infection with influenza viruses, respiratory syncytial virus, or adenoviruses. The Lancet 399(10334):1463–1464