THE UNIVERSITY *of* EDINBURGH

A novel amplicon-based sequencing method for HRSV genotyping

Goncalo Teixeira Fernandes

THE UNIVERSITY *of* EDINBURGH

MScR Infectious Diseases

The University of Edinburgh

2023

Abstract

Introduction

Human Respiratory Syncytial Virus (HRSV) is considered the etiologic agent of diseases of the respiratory system with the greatest paediatric importance in the world. It is identified in children aged between zero and five years, who have respiratory failure, pneumonia, and bronchiolitis. It affects about 95% of children in their first two years of life and some adults although mainly the elderly and immunocompromised. Two antigenic subgroups of HRSV have been described, named A and B, both of which circulate separately or simultaneously during epidemics.

The COVID-19 pandemic has led to significant changes in the transmission dynamics of other respiratory viruses including HRSV in children. The autumn and winter seasons of 2020/21 saw almost no cases of HRSV in the United Kingdom and Ireland. An increase in HRSV cases was observed in July2021 and peaked in August, which is out of season for the UK as cases usually occur in November/December. In response to this, a National Institute for Health and Care Research (NIHR) study was funded called The BronchStart and which collects data from patients with HRSV infection such as age, sex, clinical severity on presentation, treatment received and outcomes (discharge/hospital admission/PICU admission). Also in 2021, the WHO has established an online repository of HRSV whole genome sequences (GISAID). To contribute to these studies and to aid laboratories worldwide to be able to readily sequence HRSV, we designed a novel short amplicon approach for sequencing complete HRSV genomes, which can be adapted from existing SARS-CoV-2 sequencing methodologies.

Methods

We developed our own primer scheme for a shorter amplicon-based sequencing approach to be used in the existing SARS-CoV-2 sequencing infrastructure. A total of 6 HRSV A and 6 HRSV B genomes, with collection dates spanning 2019 to 2021, and locations spanning the UK, Australia and South Africa were used to develop the primers using Primal Scheme, resulting in two tiled ~400bp amplicon schemes for HRSV A and HRSV B. To validate the method, we selected NHS Lothian (Edinburgh, Scotland) HRSV samples from 2019 to 2022 and followed the ARTIC-loCost-v3 nanopore library preparation method, with modifications. Data analysis was performed using an in-house version of the “fieldbioinformatics” pipeline.

Results and discussion

The short amplicon-based approach provided PCR products which could be taken forward to nanopore library preparation and sequencing. We tested these in a subsection of our sample pool from 2019 to 2022 with a range of different Ct values. From 45 clinical samples for HRSV A, 39 (86.7%) showed genome completeness above 90%, with a median genome completeness across all HRSV A samples of 96.6%. For HRSV B, 39 (86.7%) out of 45 clinical samples showed genome completeness above 90%, with a median genome completeness of 97.2%.

Future work includes determining the Ct cut off to find the upper limit for successful sequencing, testing of primer schemes on Illumina sequencing platforms, and efforts will be made to identify primers that can be spiked in to combat amplicon dropouts.

Conclusion

There is a clear need to develop more protocols for Whole Genome Sequencing for HRSV and to submit new data to GISAID aiming to contribute to the study of epidemiology of HRSV. This will assist in the knowledge of viral circulation and in the prevention of future public health problems. We developed an effective and easy-to-use tiled amplicon approach for HRSV which will be able to be used in existing SARS-CoV-2 sequencing infrastructures, allowing high throughput HRSV WGS.

Keywords

Human Respiratory Syncytial Virus; Next Generation Sequencing; GISAID; Whole Genome Sequencing

Lay Summary

Human Respiratory syncytial virus (HRSV) is a virus that belongs to the same family of viruses as mumps and measles virus. HRSV is a common respiratory virus that causes coughs and colds in winter. However, it can also cause severe respiratory infection in children particularly young children and is one of the leading global causes of infant mortality. In the United Kingdom, HRSV accounts for approximately 450,000 GP appointments, 29,000 hospitalisations and 83 deaths per year in children and adolescents. Oxygen is used to manage symptoms and feeding support is given to young children as there is currently no specific treatment for HRSV. To help prevent the infection, HRSV vaccines have been developed which are currently awaiting approval from Regulatory Agencies in 2023. Sequencing of HRSV virus can aid vaccine development which in turn will help with the management and prevention of the disease. This project aimed to develop a new method for sequencing HRSV which will be simple and easy to adapt in laboratories worldwide that have been upscaled as part of the COVID pandemic.

Acknowledgments

I want to start by thanking Rebecca and Danny who both had so much patience with me and spent so much time teaching me and encouraging me whenever I went off track or things did not go as planned, which happened a few times during the course of this project. It is something I truly appreciate, and I will never forget.

I also want to thank Kate Templeton for giving me the opportunity to do this project, and in doing so, she allowed me to further myself as a person, and my career, through acquiring a lot of new knowledge and a real feeling of accomplishment.

Abbie, who helped me navigate the department when I first started, and became a dear friend, who has always helped me and provided advice no matter how painfully silly my questions were. Bilkis has also provided comfort with snacks and hugs which I have needed over the last year.

I want to thank the whole team in the Virology Sequencing Department for their help, who have directly or indirectly helped me through this master's, and I hope you all know that you have all contributed to this project and truly made a difference. I look forward to a lot more years of VSS at work but also some fun socials.

Lastly, I am thanking my partner for being the most patient and supportive person I have ever met and ever had in my life. No words will ever be able to emphasize enough how much I love and appreciate you.

Also, thanks to my cats of course.

Table of Contents

Table of Figures

List of Tables

Introduction

History of human respiratory syncytial virus

Human Respiratory Syncytial Virus (HRSV) was isolated for the first time in 1956 by Morris and collaborators, from a colony of chimpanzees at the *Walter Reed Army Institute of Research*, United States of America (USA), which had clinical symptoms common to respiratory diseases, such as runny nose, sneezing and coughing. In the face of the epizootic, a swab collection was made from the throat of the animals and isolation was carried out in cell culture. After confirmation of viral infection, the pathogen was named as the *Chimpanzee Coryza Agent* (CCA) (Morris et al. 1956).

Chanock et al. (1957) inoculated the virus in some laboratory animals such as rabbits, mice and chimpanzees, but it was observed that only the primates developed the symptoms of the disease. In the same year of isolation, a laboratory technician who worked with the CCA and the chimpanzees showed to be potentially infected by the virus and serological data suggested that tested humans had antibodies against this virus or some others closely related to it (Morris et al. 1956).

Subsequently, a similar virus was isolated from children hospitalized in the USA and due to its ability to form syncytia and to have affinity to the respiratory tract it was renamed Respiratory Syncytial Virus (Chanock et al. 1957).

HRSV is considered the etiological agent of diseases of the respiratory tract with the greatest paediatric importance in the world, routinely identified in children aged between zero and five years, who have respiratory failure, pneumonia and bronchiolitis (Shi et al. 2017b). It affects about 95% of children in their first two years of life, adults, mainly elderly and immunocompromised; the occurrence of lifelong reinfections are common (Anderson et al. 1990, Hall et al. 1991, Falsey et al. 2005).

It presents tropism almost exclusively restricted to the respiratory tract, infecting mainly the ciliated epitheliums of small bronchioles and pneumocytes

in the alveoli. The most common clinical symptoms of HRSV arise, in large part, as a result of airway obstruction, where remnants of inflammatory cells, fibrin, mucus and lymphoid aggregates lead to restricted airflow, hypoxia and wheezing (Johnson et al. 2007).

HRSV classification

HRSV is classified as a member of the *Mononegavirales* order, which has other important representatives such as the rabies virus (family *Rhabdoviridae*) and the Ebola virus (family *Filoviridae*). It is included in the family *Pneumoviridae*, which has two genera: *Metapneumovirus* infecting mammals and birds; and *Orthopneumovirus*, to which the Respiratory Syncytial Virus belongs, of the *human Orthopneumovirus* species. Within the species there is also *bovine Orthopneumovirus* and *Murine Orthopneumovirus* (Amarasinghe et al. 2019). It is suggested that there was a jump between species during the evolution of these viruses and so far, an animal reservoir for HRSV has not been described (Collins & Graham 2008).

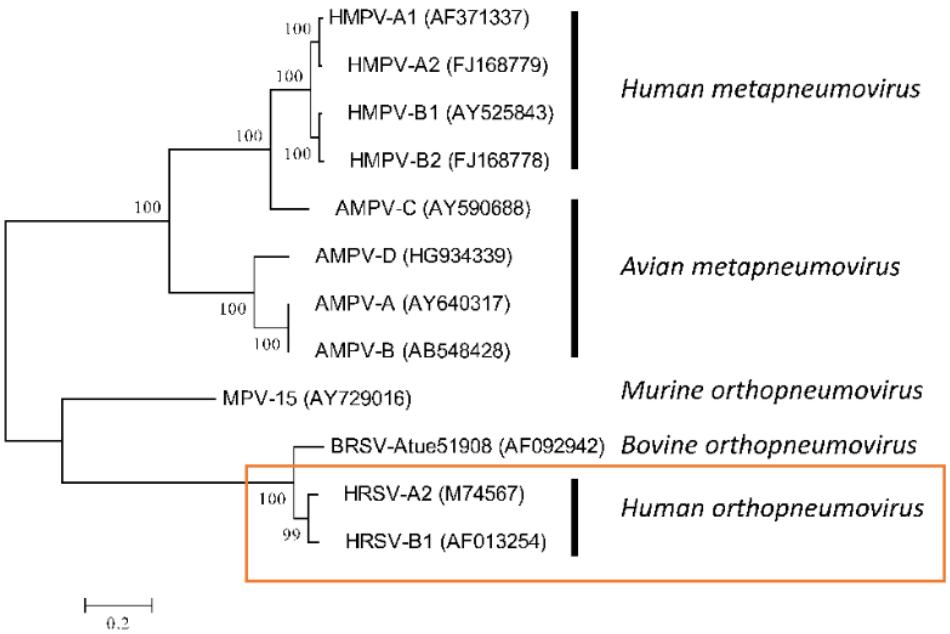

Figure 1 - Phylogenetic tree of Pneumoviridae
Genus Human orthopneumovirus in focus. Source: Rima et al. 2017

Members of the *Pneumoviridae* family have some common characteristics such as: the virions can be both spherical and filamentous; have non-segmented negative sense RNA genome; replication occurs in the cytoplasm; RNA polymerase is RNA dependent and comprises proteins P and L; the translation is done sequentially, and the machinery of translation translates the capped and polyadenylated messenger RNAs into the cytoplasm (Rima et al. 2017).

Two antigenic subgroups of HRSV have been described, called A and B, based on in vitro analyses of neutralisation assays and reactivity patterns with monoclonal antibodies, with higher variance observed in the G protein (Coates et al. 1966, Anderson et al. 1985, Maurice Mufson et al. 1985). It is known that the two subgroups can circulate separately during epidemics (Johnson et al. 1987).

Some studies suggest that HRSV-A infections are frequent when compared to HRSV-B and its transmissibility seems to be larger (White et al. 2005). In addition, isolates from subgroup A may present different clinical manifestations and be associated with worse disease outcomes when compared to B (Martinello et al. 2002, Liu et al. 2016).

In 2013, data from the largest epidemiological study of HRSV showed that both subtypes (A and B) were detected, however, more patients were infected with HRSV-A than with HRSV-B (20.4% versus 10.9%, respectively) and patients with positive HRSV-A were more likely to be admitted to a hospital or intensive care unit (47.7% A) compared to HRSV-B (35.8%), upholding the hypothesis that the subtype A may be more virulent and cause more severe disease. Nevertheless, more studies are needed to determine whether actually the different HRSV subtypes are associated with the severity of the disease (Jafri et al. 2013).

Fodha et al. (2007) when evaluating 81 babies in their study, concluded that severity of the disease was not significantly associated with the subgroup of the HRSV, and it was probably determined by an interaction between

pathogen and host factors, such as viral load. Such correlation between the severity of the disease and viral load remains controversial, as the methods of measuring viral load vary widely between studies (Vianna et al. 2021)

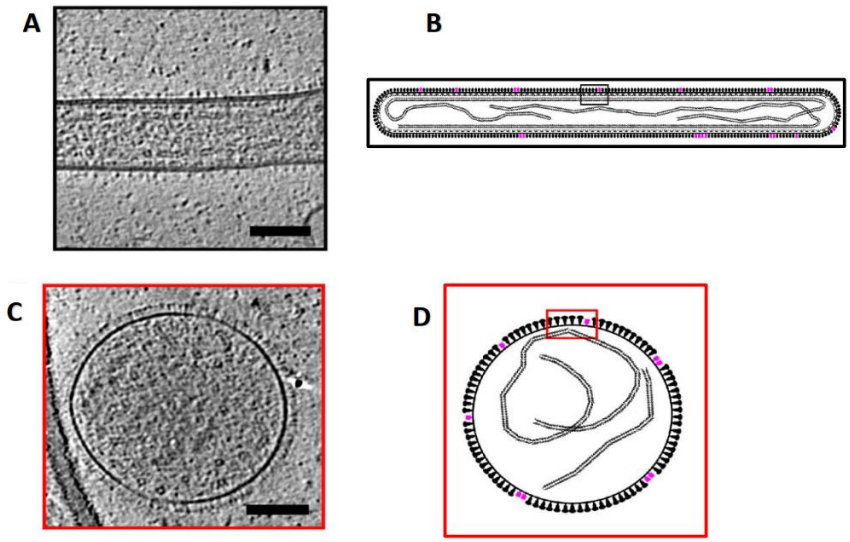

Figure 2- Structural comparison between the filamentous and spherical particles of HRSV and their schematic representations.

Representation of a filamentous particle (A) and its scheme (B); Structure and scheme of a spherical particle of HRSV (C and D). The boxes in B and D show the arrangement of surface glycoproteins F and G. Source: Adapted from Ke et al. 2018

Structure: Genome and Morphology

HRSV is a pleomorphic, enveloped virus that may present in irregular spherical structures from 100 nm to 350 nm in diameter or particles in filamentous forms with a diameter of 60 nm to 200 nm (Norrby et al. 1970), as shown in Figure 2. It is suggested that the matrix protein regulates the morphology of HRSV, however, there are challenges in elucidating some of the issues related to this topic due to the difficulty in isolating it in cell cultures because of the physical instability of the particle and the low viral yield (Liljeroos et al. 2013, Ke et al. 2018).

The virion has an external envelope, derived from the plasma membrane of the host cell by budding, in which there are three inserted surface

glycoproteins: fusion protein (F), binding glycoprotein (G) and the small hydrophobic protein (SH), which form similar projections to spikes of 10-14 nm long, spaced in HRSV with 8-11 nm (Rima et al. 2017).

In the matrix, two proteins are present, called M and M2-1, which cover the interior of the viral membrane (Kiss et al. 2014). The nucleocapsid, a protein complex associated with viral RNA (vRNA), is composed of Phosphoprotein (P), Nucleoprotein (N), polymerase subunit 4 (L) and Transcription factor M2-2. There are also two non-structural proteins 1 and 2 (NS1 and NS2) (Collins & Graham 2008). The functions of the different proteins are presented summarized in Figure 5 and in the morphological representation in Figure 3.

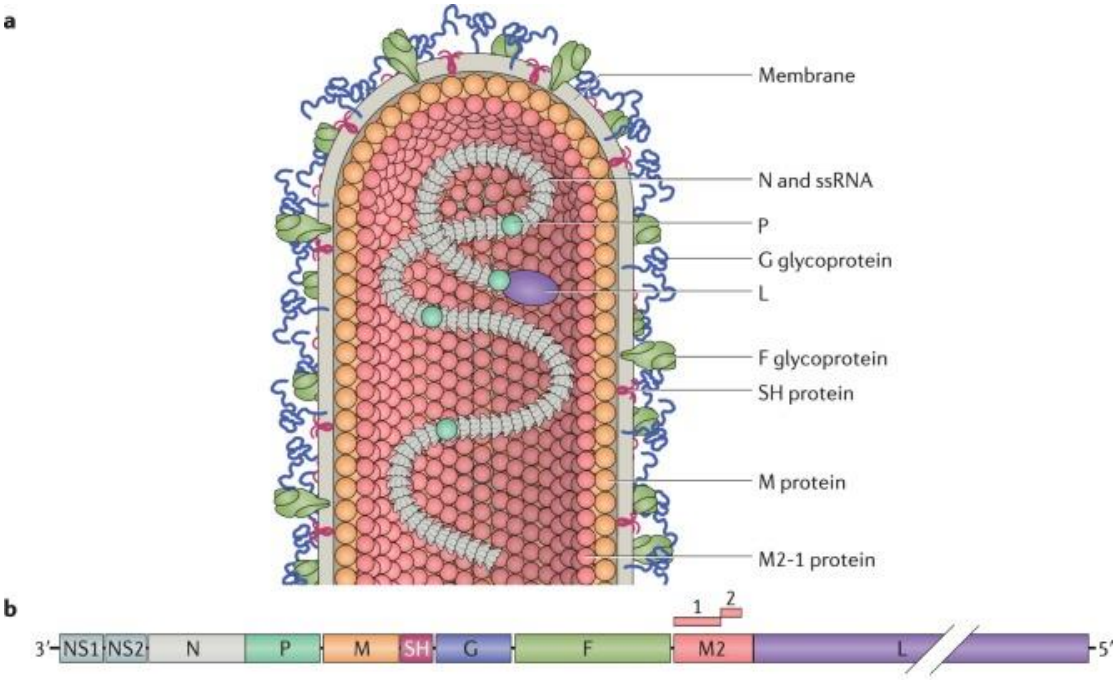

Figure 3 - Representative model of the morphology of filamentous HRSV.

Shown are the binding (G) and fusion (F) glycoproteins, as well as the small hydrophobic protein (SH) in the viral membrane. A layer of matrix protein (M) lies under the viral membrane and gives the virion its filamentous form. M2-1 protein - a transcription processability factor - interacts both with the M protein and the nucleoprotein (N) that surrounds the viral RNA genome. The polymerase (L) and the phosphoprotein polymerase (P) cofactor are also Associated with N. Source: Battles & McLellan 2019.

The viral genome consists of a simple, unsegmented RNA tape of negative polarity, composed of approximately 15,000 nucleotides, with 10 genes encoding 11 proteins, including structural proteins and non-structural (Huang & Wertz 1982, Collins et al. 1984). Each infectious particle of HRSV has only one functional copy of the genome (Collins et al. 2008).

HRSV encodes 10 sub genomic messenger RNAs, each containing an ORF (Open Reading Frame), with the exception of M2 that has two, encoding the M2-1 and M2-2 proteins (Gould & Easton 2007). In the part 3' of the viral genome, there is a region of 44 nucleotides called the leader region, followed by the ten viral genes and at the end of the genome there is a region of 155 nucleotides, known as the trailer region (Collins et al. 2001).

Genes are separated by intergenic regions that can vary up to 56 nucleotides and are poorly conserved regions between samples (Figure 4). The first two transcribed genes are non-structural proteins NS1 and NS2, this being one of the main differences between the genomes of the genera *Orthopneumovirus* and *Metapneumoviruses* that do not possess them. The M2 and L overlap 68 nucleotides at the end of the genome, causing L to have the beginning of its transcription within the M2 gene and consequently, the M2 termination and polyadenylation signal is located within the L gene (Collins et al. 2013).

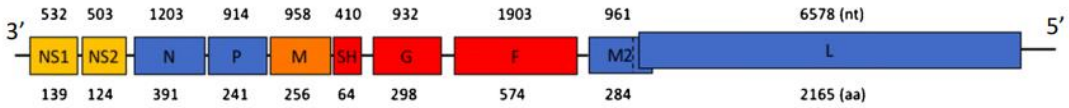

Figure 4 - Organization of genes in the HRSV genome.

In yellow: non-structural protein genes (NS1 and NS2); genes of nucleocapsid-associated proteins (N, P, M2 and L); in orange the matrix protein gene (M); in red the genes of glycoproteins (SH, F and G). The length of the genome is indicated by nt (nucleotides) on the top and the length of proteins per aa (amino acids), on the bottom. Source: Taleb et al. 2018.

Viral proteins

HRSV encodes 11 proteins, differing from most members of its family, which typically have 6 to 7 mRNAs encoding 7 to 9 separate proteins (Collins et al. 2013).

The first encoded proteins are non-structural proteins NS1 and NS2, which have 139 and 124 amino acids, respectively. In the literature, it is described that they can intervene in the innate immune responses of the host, inhibiting interferon induction and signalling (Spann et al. 2005, Sedeyn et al. 2019). In addition, they inhibit apoptosis, prolonging cellular life and increasing viral production (Bitko et al. 2007). However, more studies should be focused on the elucidation of their functions.

Like other members of the *Pneumoviridae* family, HRSV has its genome complexed with proteins N, P, L and M2-1 for the formation of nucleocapsid (Bakker et al. 2013).

The N protein, is the main structural protein of the nucleocapsid, composed of 391 amino acids, which binds to the genome to form helical nucleocapsids, serving as a template for RNA synthesis (Bakker et al. 2013, Collins et al. 2013).

The Phosphoprotein (P) has 241 amino acids, is an essential cofactor of polymerase, interacts with nucleoprotein (N) and polymerase (L) and has an important role for transcription and viral replication (Collins et al. 2013). It is the main phosphorylated protein of HRSV containing phosphate at different sites and some experiments in which P phosphorylation was reduced have shown that it is an important factor for virus replication (Lu et al. 2002).

Protein M, composed of 256 amino acids, is a non-glycosylated structural protein that plays an important role in assembling the virus (Marty et al. 2004). It is also associated with the nucleocapsid, and studies suggest that it may inhibit viral transcription activity (Ghildyal et al. 2002). M2-1 intercedes the

association between M and ribonucleoprotein complexes (RNPs) comprising viral genomic RNA strongly associated with N (Battles & Mclellan 2019).

The internal protein M2-1, of 194 amino acids, is an essential factor in transcription of HRSV, interacting with the other proteins that make up the nucleocapsid (Collins et al. 2001). In the literature, there is evidence showing M2-1 is located between RNP and M in isolated viral particles and that M and M2-1 regulate the organization of the virus (Kiss et al. 2014).

In turn, M2-2 (approximately 90 amino acids), is encoded by the penultimate HRSV gene, which has two overlapping ORFs, generating M2-1 and M2-2 (Bermingham & Collins 1999). It is involved in the regulation of transcription and replication of viral RNA and studies suggest that the level of production of M2-2 protein in infected cells may be critical for the replication of the RSV (Cheng et al. 2005).

The L protein is the largest protein in HRSV, consisting of 2,165 Amino acids. It has enzymatic activities of RNA-dependent polymerase. Like in other unsegmented and negative sense RNA genome viruses, six conserved regions in L are described (Collins et al. 2013).

The small surface protein, SH, is expressed by all members of the family *Pneumoviridae* and has about 64 amino acids (can differ from HRSV-A to HRSV-B). Its function is not yet well clarified, but SH appears to be a viroporin, a class of small viral proteins that can modify membrane permeability and may affect budding and apoptosis (Fuentes et al. 2007).

The two main antigens, F and G, are surface glycoproteins abundant in the virion and play important roles in the entry of the virus into the cell. Glycoprotein G acts primarily as a fixation protein that connects the virions to the target cells and glycoprotein F has the main function of mediating the fusion of viral membranes with the host cell (Walsh & Hruska 1983, Levine et al. 1987). Both are the only target proteins for neutralising antibodies, with protein

F being the most potent candidate for studies that are seeking development of drugs and vaccines against HRSV (Graham et al. 2015).

Protein F is a class I fusion protein of 574 amino acids that forms a trimeric structure with at least two well-known states, a metastable pre-fusion state and a stable post-fusion state (Melero et al. 2017). Structural comparisons between filamentous and spherical virions revealed significant differences associated with surface glycoproteins, where the filamentous particles of HRSV were covered with F in the pre-fusion state and spherical particles predominantly had the post-fusion state. This provides a molecular mechanism that relates morphology and infectivity of the virus, suggesting that filamentous HRSV is the infectious form (Liljeroos et al. 2013; Ke et al. 2018).

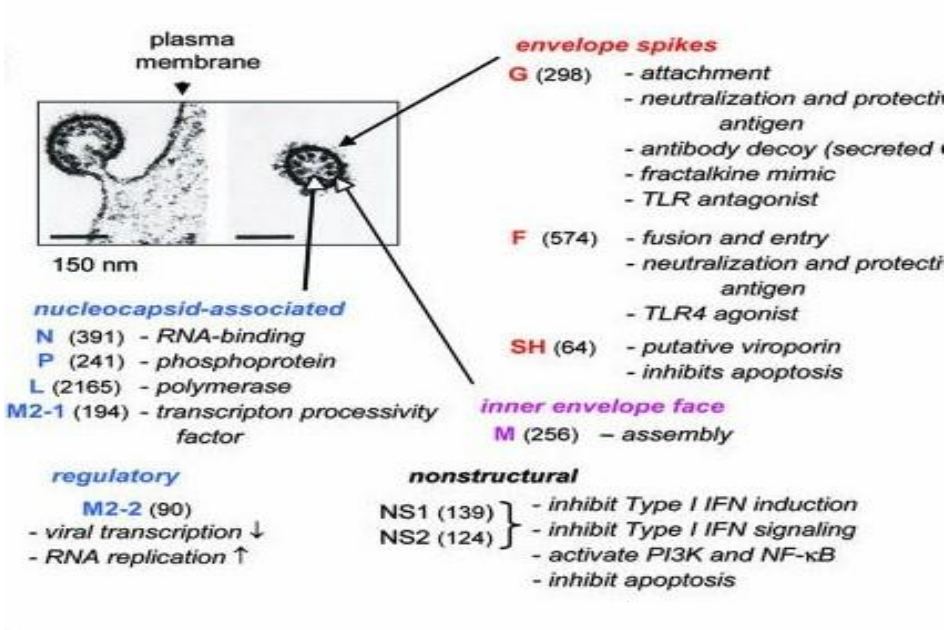

Figure 5 - Summary of HRSV proteins and their functions.

Source: Collins et al. 2008

Among viral proteins, the G protein, responsible for binding of the virus to the cell has been the subject of many studies for interacting with the host cell receptors, being the target for neutralizing antibodies, and being used for genotyping of the virus, presenting highly variable regions (Johnson et al.

1987, Levine et al. 1987). It is known to be a protein analogous to hemagglutinin from other viruses in the *Paramyxoviridae* family but does not present hemagglutinating action and is one of the most abundant in the virion along with the protein F (Hall et al. 1991, Collins & Graham 2008).

G protein can be found in the virus in two ways: anchored to the membrane (Gm), responsible for viral binding to the host cell; or in a soluble form (Gs) that is related to the mediation of immune evasion (Levine et al. 1987, Hendricks et al. 1988). The gene encoding it has approximately 950 nucleotides (Wertz et al. 1985).

In its membrane-bound form, G is a type II glycoprotein of approximately 300 amino acids, depending on the viral strain, consisting of three regions: a cytoplasmic tail (AA 1-38) in the amino-terminal portion (N- terminal), a single hydrophobic domain between *waste* 38 and 66 which acts as a membrane anchor domain and an extracellular domain (AA 66-298), where two variable regions responsible for divergence between samples of groups A and B are present, in addition to some conserved regions (Figure 6) (Wertz et al. 1985; Johnson et al. 1987; Hendricks et al. 1988). Its molecular mass is only 32,588 kDa, but post-translational changes in the Golgi complex can lead to the formation of a peptide of 80 to 90 kDa (Satake & Venkatesan 1984, Palomo et al. 1991).

The soluble form of G arises due to the initiation of translation in an alternative codon (Met48 in Gm), located in the middle of the transmembrane domain that excludes the cytoplasmic domain and the remaining portion of the transmembrane domain after a proteolytic event, allowing the ectodomain to be secreted by the cell (Roberts et al. 1994). Therefore, this form does not remain anchored to the membrane, as it does not have the initial 65 amino acid residues of the N-terminal portion where the transmembrane region of the protein is located. It is not clear why such a proteolytic event occurs only in the shortened G protein, since the complete G protein contains the same sequence. In addition, while Gs is monomeric, Gm is oligomeric (Scribe-Romero et al. 2004).

Gs's role has not yet been well clarified, but Bukreyev et al. (2008) suggest that the protein in its soluble form acts as a bait of antibodies, helping HRSV escape the replication restriction dependent on antibodies and as a leukocyte modulator containing Gamma Fc receptors.

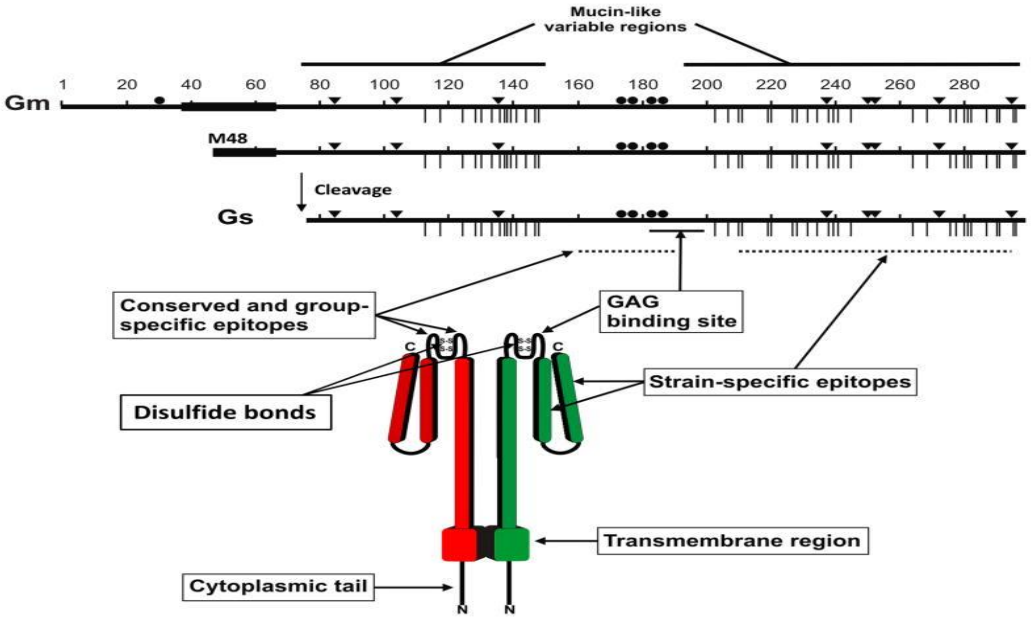

Figure 6 – Scheme of the membrane-anchored (Gm) form and soluble (Gs) form of Glicoprotein G of HRSV.

Hydrophobic regions are represented by thick lines. Triangles represent N-bound glycosylation sites and vertical lines indicate locations of glycosylation linked to O. Conserved cysteine residues are represented by circles. Variable regions are shown with horizontal bars upon Gm. The soluble form is formed by initiation of alternative translation in M48, followed by cleavage after residue 65. Structure model of membrane-anchored G. Source: Melero et al. 2017

Molecular analyses of the second hypervariable region of G protein have been used to characterize HRSV genotypes, divided into two groups, HRSV-A and B (Peret et al. 2000). It is believed that this variability may contribute to the ability of the virus to cause annual outbreaks and is related to virulence factors between the strains, in addition to this the changes in the sequences of amino acids can provide an advantage to the virus possibly by altering epitopes and

assisting in the escape of the host's pre-existing immune response (De-Paris et al. 2014).

In the ectodomain, studies suggest that the conserved regions found are responsible for the connection of the virus to the cell receptor and candidates for important domains in the transport, processing, biological activity and neutralization, such as the central region of 13 amino acids covering residues 164 to 176 and a region with four spaced cysteines (residues 173, 176, 182 and 186) connected to each other by disulphide bridges, which are preserved in all viral strains (Johnson et al. 1987, Palomo et al. 1991).

In turn, the variable regions are composed of a sequence of amino acids with a high content of serine, threonine and proline, resembling a protein class called mucins (Wertz et al. 1985). These regions are glycosylated by the post-translational addition of various sugars bound to -N and oligosaccharides linked to -O, contributing to the antigenic structure of the protein and viral infectivity (Eshaghi et al. 2012).

The first variable region precedes the cluster of conserved cysteines and has a stem-shaped structure due to multiple O-glycosylation, which possibly facilitates virus-cell interaction due to making the interaction region far enough between the receptor and the viral envelope (Martinelli et al. 2014, Melero et al. 2017). The second variable region, which corresponds to the C-terminal region of the G protein, is the most variable region of the protein and reflects the overall variability of the gene, being the most widely used in molecular epidemiology studies (Melero & Moore 2013).

Viral replication and transcription

The process of the virus entering into the cell (Figure 7) consists of two main steps: attachment of the virion to the host cell and fusion of the viral membranes with the host cell (Battles & Mclellan 2019). Viral infection begins when HRSV G protein interacts with the ciliated cell receptors in the airways,

having tropism by the cells of the apical surface (Collins & Graham 2008, Villenave et al. 2012). Some described receptors that interact with HRSV by G protein binding are Toll-like 4 receptor (TLR4), receptor 1 of chemokine (CX3CR1) and Heparan Sulphate (HSPG) (Krusat & Streckert 1997, Kurt-Jones et al. 2000).

Subsequently, the viral envelope is incorporated into the cell membrane of the host occurring fusion by action of protein F, which in addition to this action facilitates micro pinocytosis, further promoting the fusion of infected cells with adjacent non-infected cells, resulting in syncytia, a characteristic that named the virus (Levine et al. 1987, Kahn et al. 1999, Krzyzaniak et al. 2013).

After fusion, the helical ribonucleoprotein complex (RNP) is released into the host cell cytoplasm, and transcription and replication occur in the cytoplasm without the participation of the cell nucleus (Collins et al. 2001).

Transcription of the viral genome is initiated at 3' 5' direction from a promoter located at the 3' non-coding termination of the genome called leader (Fearns et al. 2002). The RNAs produced are polyadenylated in the 3' end and receive a "cap" at the 5' end. Each gene is transcribed into a separate and limited polyadenylate mRNA, which encodes a single viral protein, with the exception of the M2 gene, which encodes two distinct proteins, M2-1 and M2-2 (Collins et al. 2001). At the beginning of transcription, the RpRd enzyme binds to the genome-promoting region and at the junction between that region and the first gene (NS1), the transcript release occurs and the polymerase resumes transcription sequentially from start to finish, guided by the start signals (leader) and end of the gene (trailer), represented in Figure 8, resulting in the production of copies of messenger RNAs equal to the original gene (Collins et al. 1984, Fearns et al. 2002, Collins et al. 2007).

Replication occurs with the generation of an antigenic RNA of positive sense which will serve as a template for the synthesis of copies of the viral genome (Noton et al. 2019).

Finally, viral maturation occurs with the assembly of the nucleocapsid, by combining N, P and L proteins with genomic RNA, close to the cell membrane. For the assembly of the envelope, when passing through the endoplasmic reticulum and the Golgi complex, the surface proteins go through glycosylation and the matrix proteins are incorporated into the internal part of the envelope being formed. The nucleocapsid reaches the cell surface and is released by budding, carrying a part of the plasma membrane in a process opposite to fusion penetration (Collins & Graham 2008).

The virions are released as filamentous particles and some studies indicate that over time, M dissociates itself from the viral membrane, creating non-filamentous regions in the virion that ultimately lead to the formation of spherical particles or pleomorphic that are considered less infectious, probably due to a premature conversion of pre-fusion protein F in post fusion conformation (Ke et al. 2018).

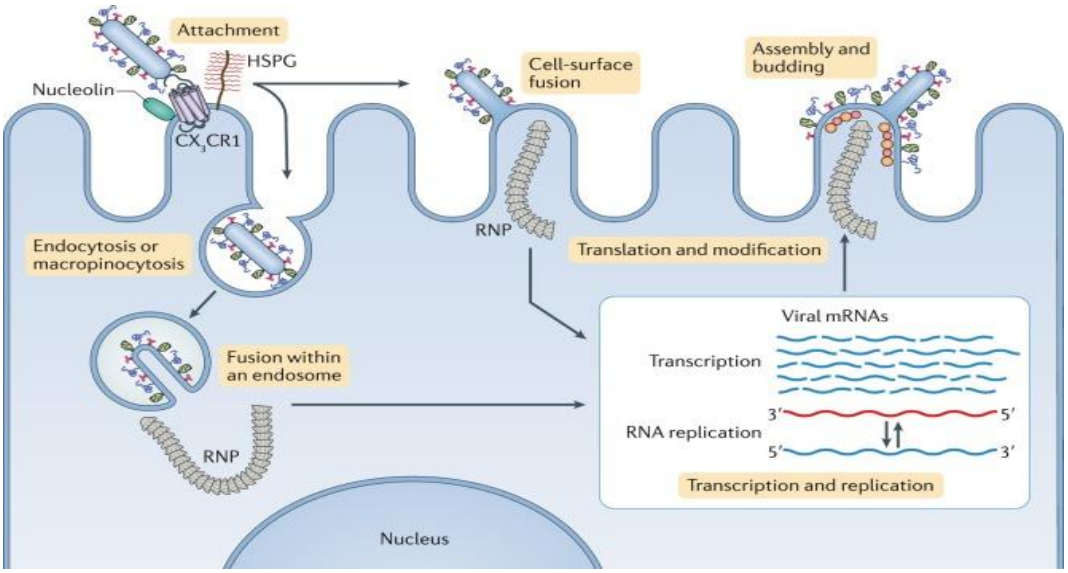

Figure 7 – HRSV replication cycle.

Entry of the virus into the cell by endocytosis, fusion of membranes and release of helical ribonucleoprotein complex (RNP) in the cytoplasm lead to the transcription and replication processes of the genome. Subsequently, the new viral particles are released by budding. Source: (Battles & Mclellan 2019).

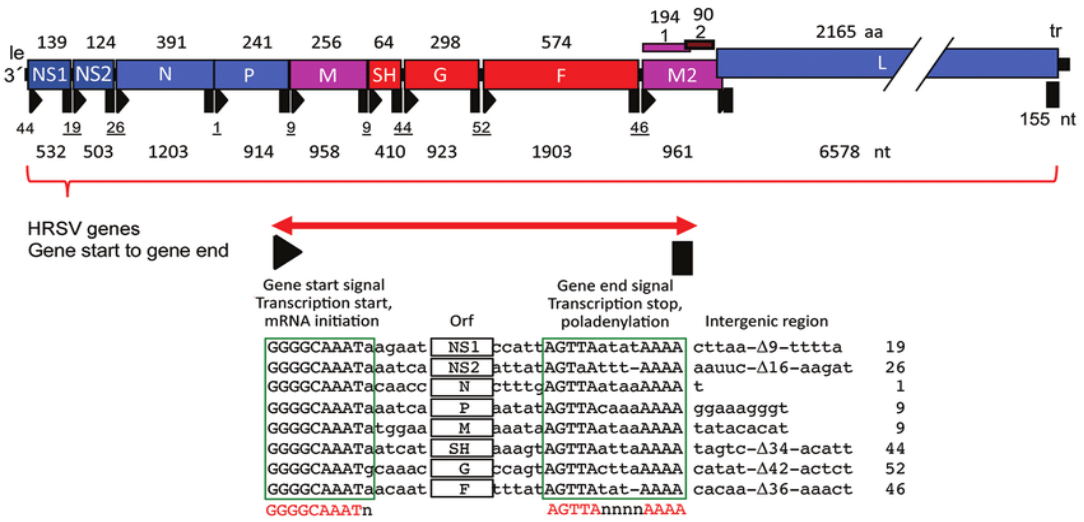

Figure 8 - Schematic representation of the location of the signs of onset and end of genes of HRSV.

The leader signs are represented by black triangles and the gene end signs (trailer) are shown as black rectangles, separated by intergenic regions (underlined). Source: Salimi et al. 2021.

Gene and antigenic characterization

RNA genome viruses have high mutation rates along their sequences, determined by the lack of revision by the RNA polymerase and high replication rates, and may result in benefits for the virus such as greater virulence and ability to evolve and adapt to the host (Holland et al. 1982, Belshaw et al. 2007).

In 1966, Coates and collaborators showed antigenic variability in HRSV isolates by serum neutralization tests of infected ferrets, and later, it was possible to definitively confirm the existence of these antigenic groups through use of assays with monoclonal antibodies (mAbs) obtained from mice immunized with different strains of HRSV, such as A2, Long and CH18537 (Anderson et al. 1985), naming the different groups as A and B and demonstrating that both evolved separately over time (Mufson et al. 1985). The designation of the HRSV Group A and B is also referred to as antigenic "subgroups" in the literature (Hall et al. 1990, Matheson et al. 2006).

With the advent of molecular biology and sequencing techniques based on analyses of the sequences of genes N, P, SH, and G it was possible to confirm the division of HRSV in groups A and B and identify the circulating genotypes (Sullender et al. 1993, Peret et al. 1998) and it was seen that the different genotypes co-circulated throughout the seasons and that there may be different patterns of movement between groups (Cane et al. 1992, Venter et al. 2001).

The first HRSV genotype classification system, proposed in 1998, had sequencing information from the second hypervariable region of the G gene and genotypes were named on the basis of the gene used for classification, followed by subtype HRSV (A or B) and an increasing number (Peret et al. 1998). Later, the previous classification method was refined to include genetic distance as a metric for defining genetic groups and there were changes in the nomenclature system, to include the country where the genotype was described (Venter et al. 2001).

Initially, six genotypes were identified, four for HRSV A, called SHL1 to SHL6 and two identified for HRSV B, called NP1 and NP3 (Cane et al. 1992). Over time, other genotypes have been described such as GA1 to GA5 and GB1 to GB4, GA6 and GA7 (Peret et al. 2000); SAA1 and SAB1 to SAB3 were described for the first time in South Africa (Venter et al. 2001); BA1 to BA10 were described in Argentina and these had as its main characteristic a duplication of 60 nucleotides in G (Trento et al. 2006); URU1 and URU2 were described in Uruguay (Arbiza et al. 2005); NA1 and NA2 identified in Japan (Shobugawa et al. 2009); and the most recent the ON1 genotype identified in Ontario, Canada and which presents a doubling of 72 nt (Eshaghi et al. 2012).

Today, at least 13 HRSV A genotypes and 37 genotypes for HRSV B are described (Muñoz-Escalante et al. 2019b, 2021). Due to its high variability, the genetic diversity of HRSV was most often based on in the G gene (García et al. 1994). However, there are no established criteria for the definition of the genotype, which affects the understanding of viral evolution, immunity and vaccine development.

Several groups advocate the need for the unification of the classification of HRSV genotypes, but so far there is no consensus. (Muñoz-Escalante et al. 2019b, 2021, Goya et al. 2020, Ramaekers et al. 2020).

Author	Target Region	Classification Proposal	Access
Peret et al. 1998	HVR2 of G protein	5 genotypes and 22 subtypes of HRSVA; 4 genotypes and 6 subtypes of HRSVB	10.1099/0022-1317-79-9-2221
Venter et al. 2001	HVR2 of G protein	Definition of new genotypes: SAA1 and SAAB1-4	10.1099/0022-1317-82-9-2117
Trento et al. 2006	HVR2 of G protein	Definition of genotypes BA1-6	10.1128/JVI.80.2.975-984.2006
Shobugawa et al. 2009	HVR2 of G protein	Definition of NA1-2	10.1128/JCM.00115-09
Depat et al. 2010	HVR2 of G protein	Definition of BA7-10	10.1128/JCM.00646-10
Eshaghi et al. 2011	HVR2 of G protein	Definition of genotype ON1	10.1371/journal.pone.0032807
Cui et al. 2013	HVR2 of G protein	Definition of NA3, NA4, BA-C and CB1	10.1371/journal.pone.0075020
Trento et al. 2015	G ectodomain	Reclassification of HRSVA using G	10.1128/JVI.00467-15
MuñozEscalante et al. 2019	G ectodomain	Reclassification of HRSVA to 9 genotypes: GA1-GA7, SAA1 and NA1	10.1038/s41598-019-56552-2
Goya et al. 2020	G ectodomain	Reclassification of HRSVA to 3 genotypes (GA1-GA3) and 7 for HRSVB (GB1-GB7)	10.1111/irv.12715
Ramaekers et al. 2020	Whole genome	Reclassification of HRSVA and HRSVB: 23 genotypes for A and 6 for B	10.1093/ve/veaa052
MuñozEscalante et al. 2021	G ectodomain	Reclassification of HRSVB from 37 genotypes to 15	10.1038 / s41598-021-83079-2.

Table 1 - HRSV's main classification proposals over the years.

Several HRSV research groups have come together to propose a definition of common nomenclature for samples and HRSV isolates and also the use of standard names and appropriate annotations for its genes. Prioritizing a short, concise and well-defined line that is easy to use in the laboratory, easy-to-read and a uniform system for HRSV in public health databases (Salimi et al. 2021).

The sequence definition elements proposed by Salimi et al. 2021 are:

I. Name of the organism; Virus name abbreviation: HRSV;

II. HRSV subgroup: A or B; X, if unknown;

III. Geographical identifier for the sampling site (e.g., city or state);

IV. Unique isolated identifier;

V. Year of sampling; YYYY or XXXX, if unknown.

Examples of sequence definition lines using naming proposal: HRSV / A / USA / 001/2011; HRSV / B / Denver.USA / 14617/1985 or HRSV / X / IRN / 001/2017 (subgroup unknown).

HRSV epidemiology

Respiratory diseases are responsible for a large volume of hospital visits and hospitalizations per year worldwide, causing morbidity and mortality in the population (WHO 2014, Kabir 2018).

Viruses are one of the main etiological agents involved in these infections and have as characteristics easy transmission and the ability to affect individuals in all age groups, being children, elderly and immunocompromised the most vulnerable (Falsey et al. 2005).

Viral diagnosis is often neglected and difficult when evaluated only at community level due to the presence of a large number of pathogens in the respiratory airways, however, the correct diagnosis is crucial in the identification of circulating respiratory viruses, and transmission, control and prevention of outbreaks (Kabir 2018). The most common viruses associated with these infections are Influenza (A and B) (Cox & Bender 1995), HRSV, Rhinovirus (RV), Coronavirus (CoV), Parainfluenza (PIV) 1 - 4, Adenovirus

(HAdV), Metapneumovirus (HMPV) (Broor et al. 2008) and Bocavirus (HBoV) (Essa et al. 2015).

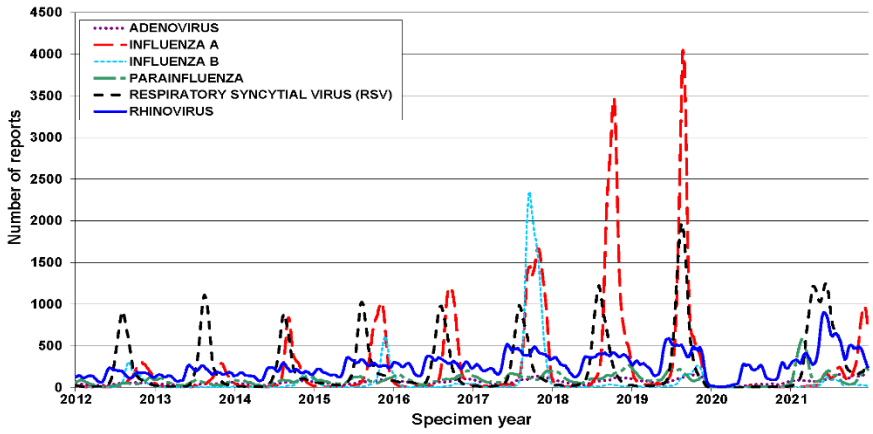

Figure 9 – Number of reports of select respiratory viruses per year

Six major respiratory viruses (positive numbers) reported from UKHSA and NHS laboratories (SGSS) in England and Wales between week 18, 2012 and week 16, 2022 (3-week moving average). Source: UKHSA 2022

HRSV is the most important paediatric virus, being the etiological agent that causes most acute respiratory infections and death in children under 5 years of age, and can also affect adults, immunocompromised, and mainly elderly (Falsey et al. 2005, Meissner 2016, Tseng et al. 2020).

A systematic review estimated that lower respiratory tract infection (LRTI) associated with HRSV is responsible for between 94,600 and 149,400 deaths annually worldwide; in addition, about 33 million LRTI associated with HRSV resulted in 3.2 million hospital admissions in 2015, with low- and middle-income countries being the most affected (Shi et al. 2017).

LRTI correspond to the leading cause of infant mortality in low-income countries and the fourth in middle-income countries (Ferone et al. 2014). The main factors associated with the severity of infection in children are

prematurity, congenital heart disease, chronic lung disease, low birth weight, having siblings, maternal smoking and non-breastfeeding (Shi et al. 2017).

In the United Kingdom, HRSV accounts for approximately 450,000 GP appointments, 29,000 hospitalisations and 83 deaths per year in children and adolescents.

Several countries and regions have described the occurrence of HRSV such as USA (Rose et al. 2018), Africa (Umuhoza et al. 2021), Canada (Mitchell et al. 2017), England and other European countries (Reeves et al. 2020), Russia (Krivitskaya et al. 2021), Japan (Hibino et al. 2018), China (Zhang et al. 2015; Luo et al. 2020), Brazil (Moura et al. 2006), Portugal (Sáez-López et al. 2020) and others. It is observed that seasonality patterns and the duration of outbreaks vary considerably between regions.

Much of the studies done in the USA describe the seasonality of HRSV from autumn to spring (Haynes et al. 2014). However, when using a new statistical method that analyses laboratory PCR detections reported to the National Respiratory and Enteric Virus Surveillance System (NREVSS) to determine the seasonality of HRSV, this showed that the beginning of the 2014 to 2017 seasons occurred approximately 2 weeks earlier than the 2012-2014 seasons, which were determined using antigen data, showing this to be a model of greater confidence to be used (Rose et al. 2018).

In general, countries in the southern hemisphere show an increase in the number of cases between March and June and countries in the northern hemisphere report it between September and December (Obando-pacheco et al. 2018). In temperate climate countries epidemics have been described in the winter months and in the regions of tropical climate, these seem to occur in the rainy seasons (Law et al. 2002).

In addition to seasonality, the circulation pattern of each of the HRSV groups (A and B) also differs in each location and seasons. The co-circulation of both groups is well reported, as well as its genotypes, but in recent years it has

been noted that there is a predominance of the HRSV-A group and its genotypes in relation to B (Al-Sharif et al. 2020).

Clinical aspects, Treatment and Vaccines

The transmission of HRSV can happen from person to person through aerosols, or by contact with contaminated secretions, objects and surfaces, being very infectious when in direct contact with mucous membranes, such as ocular or nasal. The onset of symptoms varies according to age and health conditions and usually occurs within six days after infection (Hall et al. 1991). Symptoms of the upper respiratory tract usually begin to appear after the fourth day, with fever, rhinorrhea and congestion being common, followed by symptoms of the lower respiratory tract. Increased cough is usually the first sign of involvement of the lower respiratory tract; the symptoms include tachypnea, dyspnoea, increased respiratory effort, and difficulty eating (Meissner 2016).

Johnson et al. (2007) demonstrated that the virus infects both the bronchial epithelial and the alveolar, and that the obstruction of the small airways from epithelial projections, inflammatory debris, oedema or compression by lymphoid aggregates may be an important component of the pathogenesis of the disease.

The clinical aspect of the infection may vary according to age and individual's health conditions and the main symptoms usually include runny nose, loss of appetite, cough, sneezing and fever, which may appear isolated or together. In very young babies, symptoms may be irritability, reduced activity and breathing difficulties (Ghazaly & Nadel 2018). In more severe cases, HRSV may cause bronchiolitis and pneumonia, these being the most common consequences in children under 1 year of age. In addition, the formation of syncytia and the presence of excessive secretions represent the main marks of HRSV infection (Glezen et al. 1986).

Bronchiolitis is a common infection of the lower respiratory tract in babies and HRSV is the most common cause (Smith et al. 2017). Most of the children recover without interference, but chronic lung disease, congenital anomalies, a state of immunocompromise, and haemodynamically significant heart disease in patients with bronchiolitis can be associated with progression to severe disease or mortality (Ralston et al. 2014).

In children, the only risk factors reported for hospitalization for HRSV infection are prematurity and young age and one of the main reasons why this occurs is the high proportion between the surface area and the volume of the airways that are still developing, and in addition to that when compared to the respiratory tract of adults, the bronchioles of infants and children are relatively smaller and therefore more subject to obstructions (Griffiths et al. 2017).

Another group very affected by HRSV are the elderly. According to Falsey et al. (2014), this virus was the third leading cause of respiratory infections in the elderly, after Enterovirus and Influenza, and the second largest cause of hospitalization, where elderly patients diagnosed with HRSV were twice as likely to be hospitalized than positive patients for other respiratory viruses.

In the current scenario, with the COVID-19 pandemic, it is difficult not to compare respiratory diseases such as HRSV and SARS-CoV-2, because both viruses disproportionately affect older adults, making diagnosis only with the clinical evaluation difficult (Falsey & Walsh 2020). However, there are differences in the pathogenesis of diseases caused by these viruses. In HRSV infection, nasal congestion and upper respiratory symptoms are common, and may progress to cough, dyspnoea and wheezing but fever is not a common feature, differing on that from influenza and SARS-CoV-2 for example (Tseng et al. 2020).

Despite affecting millions of people each year, there is still no specific medicine to treat HRSV infection. The default for the treatment in adults is mainly limited to supportive care with the use of bronchodilators, supplemental oxygen, intravenous fluid and antipyretics (Borchers et al. 2013). In patients at risk of

acute infection of HRSV it is recommended to use Palivizumab, a monoclonal antibody for immunoprophylaxis, approved by the Food and Drug Administration (FDA) in 1998 (IMpact-RSV Study Group 1998). For treatment, Ribavirin, a drug used to treat infections caused by various respiratory viruses may be advised, however, its use in new-borns and hospitalized children should be carried out with caution due to its side effects (Chidgey & Broadley 2005).

Palivizumab is a monoclonal antibody composed of 95% human amino acids and 5% murine, which binds to glycoprotein F blocking binding and absorption by host cell receptors, inhibiting replication and propagation of HRSV (Johnson et al. 1997, Subramanian et al. 1998). In the United Kingdom, the British National Formulary (BNF) by the National Institute for Health and Care Excellence (NICE), states that Palivizumab is “licensed for the prevention of serious lower respiratory-tract disease caused by respiratory syncytial virus (RSV) in children under 6 months of age (at the start of the RSV season) and born at less than 35 weeks corrected gestational age, or in children under 2 years of age who have received treatment for bronchopulmonary dysplasia in the last 6 months, or in children under 2 years of age with haemodynamically significant congenital heart disease”.

The high cost of palivizumab usually limits its use to high-risk children only, including premature new-borns, ones with congenital heart disease or chronic lung disease (Subramanian et al. 1998).

The only medicine licensed for the treatment of HRSV infection is ribavirin in aerosol, a broad-spectrum antiviral agent approved by the FDA, used in patients at higher risk of progression to severe disease, acting on inhibiting viral replication, inhibiting viral structural protein synthesis (Conrad et al. 1987, Borchers et al. 2013). The guidelines of the American Academy of Paediatrics (AAP) do not recommend its use routinely due to insecurity as to its effectiveness, its possible side effects and its high cost (Jartti et al. 2009).

HRSV has some challenges for vaccine development, such as the possibility of life-long reinfection because immunisation is not complete or lasting (Hall et al. 1991). However, repeated infections are associated with decreased risk of lower respiratory tract infections, even if secondary infection occurs in the first year of life (Ohuma et al. 2012).

The first attempt to obtain an HRSV vaccine was made using the virus inactivated with formalin in the 1960s, administered to children in a three-dose regimen, based on vaccine tests for other viruses. Unfortunately, despite being immunogenic, the experiment failed and resulted in an increase in respiratory diseases following natural infection with the virus, in which several vaccinated children were hospitalized and two died (Kim et al. 1969).

Since the discovery of HRSV, several studies have focused on development of a safe and effective vaccine. In recent years, with the improvement of understanding of HRSV biology and technological advances in vaccine research, there have been significant progresses and in 2020 there were approximately 19 vaccines in clinical trials and two candidates for new generation monoclonal antibodies (Hirve et al. 2020).

Although glycoproteins F and G are the main targets of antibodies, protein F is the target in the development of vaccines, as it is more conserved and due to having multiple neutralising epitopes when compared to G (Taleb et al. 2018)

Different approaches to the prevention of HRSV disease are being studied and the two most promising published trials are: one on maternal vaccination (Madhi et al. 2020) and one on passive immunisation of infants with a long-term monoclonal antibody (Griffin et al. 2020). The justification for maternal vaccination is to stimulate the production of IgG for the target of HRSV (protein F), which can then be transferred through the placenta to babies and protect them from diseases in the first months of life, when they are vulnerable to serious diseases and there is the highest risk of deaths (Madhi et al. 2020).

Respiratory Virus Surveillance

Since 1952, the WHO has carried out global influenza surveillance with the WHO's Global Influenza Surveillance and Response System - GISRS, whose objective is to identify circulating strains and provide information to subsidize recommendations for the composition of vaccines, monitor the epidemiology of influenza virus globally and alert to new mutations and other pathogens (WHO 2017, Broor et al. 2020).

Given the epidemiological importance and impact that HRSV causes and importance of obtaining a vaccine against this virus, the WHO designed a pilot project to strengthen HRSV genomic surveillance. With the aim of taking advantage of GISRS's already well-established structure, in 2016 a HRSV surveillance pilot project was created with the aim of better describing the epidemiological situation, seasonal variations, incidence in different regions and to support information that assists in the development of a vaccine. Fourteen countries, which were already members of GISRS, were selected for the pilot in 2016. In 2021, the project was in phase II, after a three-year extension (until October 2021) and aimed to consolidate its achievements and proposed to expand HRSV monitoring (WHO 2017).

GISAID has been an important partner of GISRS, and its EpiFlu database has been a key component of influenza surveillance since 2008. At the start of the COVID-19 pandemic GISAID rapidly launched the EpiCoV platform, where the first complete genome of SARS-CoV-2 was shared globally, and various analytical tools were developed to support the rapid sharing and interpreting of SARS-CoV-2 data.

GISRS-GISAID collaboration on influenza has been expanded to other respiratory viruses including SARS-CoV-2 and Respiratory Syncytial Virus (RSV). A joint bioinformatics training programme was developed to support

WHO Member States to expedite the effective genomic surveillance of SARS-CoV-2 using influenza surveillance systems (WHO).

In December 2019, cases were reported of severe acute respiratory syndrome (SARS) in pneumonia patients in Wuhan, China (Zhou et al. 2020). Following this, it was confirmed to be an infection with a novel coronavirus (SARS-CoV-2) (Wu et al. 2020). COVID-19 was declared a pandemic on March 11, 2020 (WHO 2020a). In the United Kingdom, the first recorded case was reported on January 30th, 2020, and as of April 2022 the country had reported approximately 22 million confirmed cases and 190,000 deaths (https://coronavirus.data.gov.uk/).

The COVID-19 pandemic has led to significant changes in the transmission dynamics of other respiratory viruses, most notably in HRSV in children, with an autumn and winter of 2020/21 that saw almost no HRSV cases in the UK and Ireland. The increase in HRSV in 2021 was seen in July 2021 and peaking in August 2021, while in Scotland, the hospitalisation rate for RSV peaked in week 38 and 39 in 2021 (4.5 per 100,000) (PHS 2022), with incidence rates shown in Figure 10. This is out of season for the United Kingdom as this usually occurs in November/December.

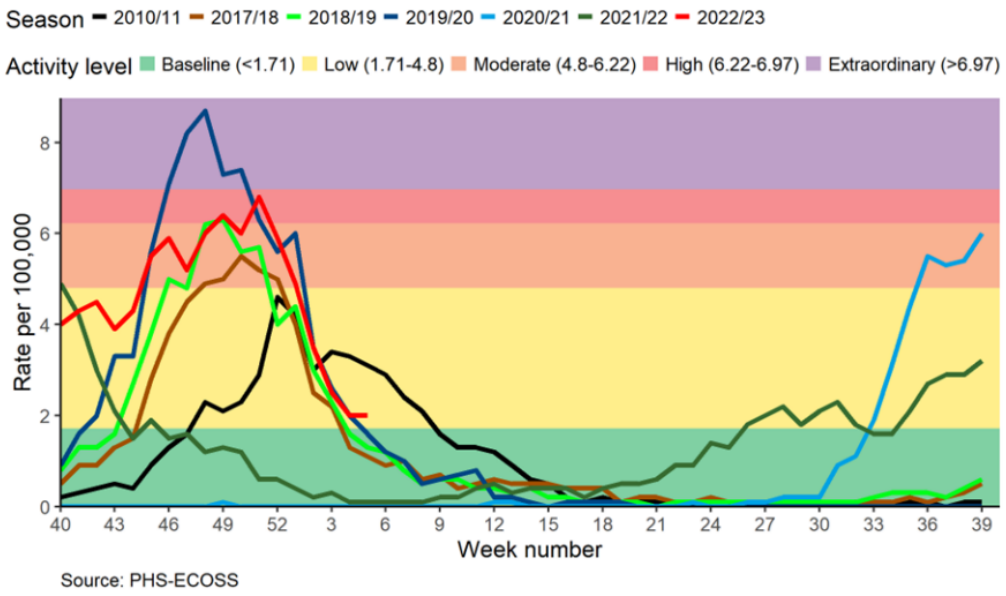

Figure 10 - HRSV incidence rate (per 100,000 population)

Comparison of incidence rates for the years 2022/2023 to the years 2010/2011, 2017/2018, 2018/2019, 2019/2020, 2020/2021 and 2021/2022. Source: PHS-ECOSS

In response to this a National Institute for Health and Care Research (NIHR) study was funded called The BronchStart. The BronchStart project has over 50 centres and includes the paediatric departments in Edinburgh. They have been collecting age, sex, clinical severity on presentation, treatment received and outcomes (discharge/hospital admission/PICU admission). Summary data on presentations is already being presented in real-time on a dashboard hosted by Microreact (Williams et al. 2021). It is a surveillance platform study run by the PERUKI network (Paediatric Emergency Research UK and Ireland) across paediatric emergency departments throughout the UK. This capability has enabled BronchStart through PERUKI to commence surveillance activities at speed and low cost. Also in 2021, the WHO has established an online repository of HRSV whole genome sequences (WGS, available at www.gisaid.org).

The impact of non-pharmaceutical interventions for COVID-19 has disrupted HRSV transmission, and with it, our ability to predict timing of RSV epidemics.

BronchStart is now able to deliver information on RSV case frequency and distribution in populations across the UK. Importantly BronchStart can also provide information on age distribution of cases and disease severity. These functions are reported live (through the microreact website) and can alert clinicians to anticipate changes in workload and prepare for out of season RSV surges. To date over 11000 children have been recruited. BronchStart is now interrogating genomic aspects of RSV spread and severity, adding another layer of information to our understanding of RSV disease.

There is a clear need to develop more protocols for WGS for HRSV and to submit new data to GISAID aiming to contribute to the epidemiology of HRSV during the COVID-19 pandemic, hoping to assist in the knowledge of viral circulation and in the prevention of future public health problems.

Next Generation Sequencing (NGS) and its applications

The first DNA sequencing method was described by Sanger and collaborators in 1977. The general principle of the enzymatic sequencing technique developed by Sanger is to carry out the synthesis of a marked strand of DNA, complementary to the chain whose sequence you want to determine. The first step is to obtain chains of the DNA you want to sequence. These chains can be separated by heating, which breaks the hydrogen bonds between the DNA bases, which keeps the joined strands. With the strands separated, a synthetic oligonucleotide, called a primer, binds to a predetermined region of one of the strands. The elongation of primer is an enzymatic process, i.e., the enzyme DNA polymerase adds deoxyribonucleotides (dNTPs) sequentially to the 3' end of the primer (primer), complementary to the template DNA (Shendure & JI, 2008).

In practice, to the reaction tube, in addition to the DNA to be sequenced, the DNA polymerase enzyme, the appropriate primer and the four deoxyribonucleotides (dNTPs): ATP, TTP, CTP and GTP are also added. The

reaction mixture also includes a small amount of dideoxynucleotides (ddNTPs) labelled with a fluorescent molecule. The dideoxynucleotides lack the free OH group at the 3' end of the deoxyribose. Thus, when this nucleotide is incorporated, the polymerase cannot proceed, because the addition of the next nucleotide depends on this OH on the 3' carbon. The fluorescent marking imparts a different colour to each of the four dideoxynucleotides (Carraro & Kitajima, 2002). The marking is then made according to the last nucleotide added to the fragment. The various fragments marked with one of the four fluorochromes can be separated according to their sizes by gel electrophoresis. In gels of polyacrylamide, it is possible to separate DNA molecules that differ in only one nucleotide, thus presenting the necessary resolution for the sequencing of DNA molecules (Sanger, 1988; Huang et al., 1992) (Figure 4).

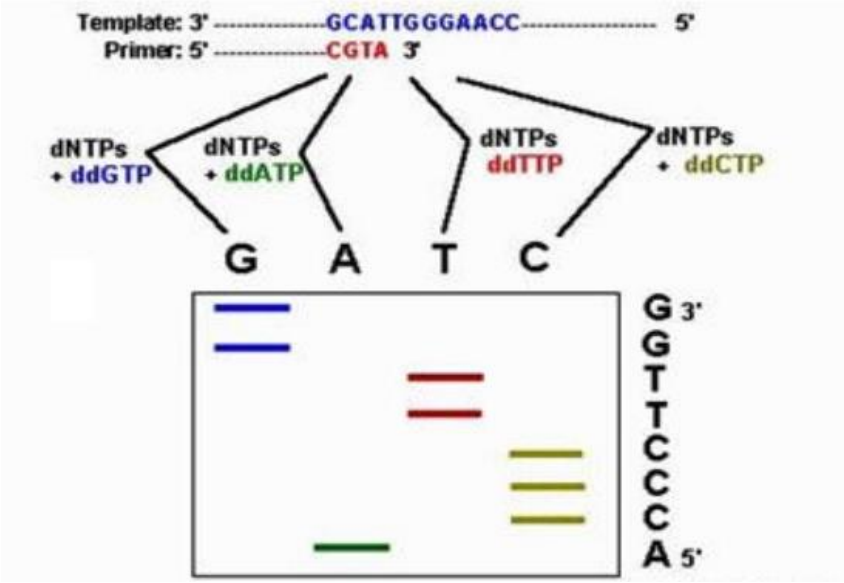

Figure 11 – Sequencing method described by Sanger

Nucleotides marked with fluorochromes and identification of the fragments by size on a polyacrylamide gel. Source: http://web.indstate.edu/thcme/mwking/sangersequencing.gif

A few years after the development of the Sanger technique, in 1986, the first automatic DNA sequencer was launched, the ABI 370, and in 1998, the first

capillary electrophoresis sequencer, the ABI 3100. With the automation of the process, it was possible to carry out large sequencing projects, such as the human genome (Lander et al. 2001; Venter et al. 2001), that of Arabidopsis (Garcia-Hernandez et al. 2002), among others. After the advent of automatic sequencers, these machines kept being improved and from 2005 new sequencing methods appeared: the next generation technologies. Next generation sequencers use methodologies different from Sanger's, with the aim of accelerating and lowering the cost of the sequencing. Despite being different from each other, all next generation sequencers (NGS) are based on the massive processing of DNA fragments.

Basically, all these new methodologies use different strategies to eliminate the most laborious steps of the Sanger method, which involve cloning in bacterial vectors and electrophoresis. While a capillary electrophoresis sequencer processes at most 96 fragments at a time, next-generation sequencers can sequence over 500 million of fragments at the same time (Chevreaux, 2004).

The first new generation platform to be commercialized was the 454, from Roche company. This platform performed synthesis-based sequencing through a technique known as pyrosequencing (Margulies et al. 2005). From this, several others emerged using different sequencing techniques, including Illumina's MiSeq and HiSeq platforms; Life Technologies' SoliD and Ion Torrent; PacBio from Pacific Bioscience, among others. Among the mentioned platforms, Illumina's were the standout for the large volume of sequences obtained, wide spectrum of usefulness, in addition to having the lowest cost per sequenced base (Glenn, 2011) in the market, which lead to the company dominating the sequencing industry in a large scale (Quail et al, 2012).

Illumina sequencing method

The sequencing method acquired by Illumina was developed by the company Solexa and is constantly being improved. Like Sanger sequencing, the Illumina

sequencing is performed by synthesis using DNA polymerase and marked terminator nucleotides. The novelty of this platform consists of the in vitro cloning of fragments on a solid surface (glass slide), a process also known as Solid phase PCR and bridge PCR (Fedurco et al. 2006; Turcatti et al. 2008). Another great innovation of the method is the modified nucleotide with a reversible terminator and cleavable fluorochrome.

Sample DNA fragments are ligated to adapters on both ends, which allows their attachment to the sequencing slide by hybridization to one of the oligonucleotides fixed on the support. After the amplification step by bridge PCR, a sequencing surface (flow cell) has more than 300 million of clusters, where each cluster is composed of about 1000 clonal copies of a single molecule (Bentley, 2006).

The cloned DNA molecules in the clusters are sequenced in parallel using a DNA synthesis approach that employs special nucleotides containing reversible terminator and fluorescent radical with removable section. An incorporation of these nucleotides prevents new bases from being incorporated until the terminator is removed. The DNA polymerase used was also modified to allow the incorporation of these modified nucleotides. The nucleotides are tagged with fluorescence, with four different colours to distinguish between the different bases. The position and template sequence of each cluster are deduced by reading the colour identified in high-resolution images obtained at each nucleotide addition step (Bentley, 2006).

Although the Illumina approach is more effective in sequencing homopolymeric stretches than the pyrosequencing technique, it produces and reads shorter sequences. Initially, the methodology generated sequences of only 36 bp (Bentley, 2006) but this has greatly improved with time. And the cost per bp generated is also more than four orders of magnitude lower with the Illumina methodology compared to other platforms, which has caused a true revolution in several areas of Biological Sciences.

The Illumina platforms have been used in several studies with different purposes and in different organisms. As an example, there have been the sequencing of microbial DNA (Caporaso et al. 2012), de-novo assembly and characterization of the sweet potato (Wang et al. 2010) and sesame (Wei et al. 2011) transcriptome, assembly de novo of genes expressed in Eucalyptus (MIZRACHI et al., 2010), sequencing of the mitochondrial genome of whales (Thompson et al. 2014), molluscs (Williams et al. 2014) and humans (King et al. 2014), detection of SNPs in sequencing data (Yi et al. 2014), among thousands of other studies.

Bioinformatics

Bioinformatics is a multidisciplinary science, which involves the sciences of computing, statistics, and mathematics, along with information technology, to decipher biological data. Bioinformatics began with the sequencing of proteins. In the 1950s, Sanger deciphered the sequence of the 51 amino acids that form insulin in different species (pig, sheep, horse, and whale). At the time, one of his students, Hans Tuppy, sequenced parts of cytochrome c in different species (pig, horse, chicken, and beef) (Strasser, 2009).

After Sanger's sequencing, new techniques emerged, such as Edman degradation reaction, in which it is possible to identify individual amino acids from the amino-terminal region in short peptides. In addition, improvements in chromatographic and electrophoretic techniques facilitated the sequencing of amino acids, making it more efficient (Fruton, 1999). In a short time, the identification of amino acids was becoming automated, which made Stanford Moore and William Stein succeed in sequencing Ribonuclease, with 124 amino acids, in half the time that Sanger had sequenced insulin (Moore & Stein, 1973).

In the late 1950s and early 1960s, computers began to be used by academic researchers. In 1957, the first high-level programming language, Fortran, was

developed which was produced by the IBM Corporation. This language was easy to learn compared to older ones, and it was suitable for scientific applications, which helped in the growth of biological computation (Ledley, 1959). In the early 1960s, John Kendrew made use of computers in an attempt to elucidate the three-dimensional structure of proteins (Perutz, 1985) However, computers began to make real contributions with Margaret Oakley Dayhoff, also in the early 1960s (Ddayhoff, 1965).

Dayhoff's early work was to formulate a series of Fortran programmes to determine the amino acid sequences of proteins. These programmes used fragments of digested peptides, which overlapped like a jigsaw puzzle, to determine protein sequences. The Dayhoff programmes reached the correct sequence of a small ribonuclease in a few minutes (Dayhoff, 1965). In addition, Dayhoff created the Structure and Sequence Atlas of Proteins with the aim of cataloguing all known peptides, this being considered the first bank of data for research in molecular biology (Fitch, 1972). Both Dayhoff and other researchers of his time realized that the logic used in the programmes they built to correctly sequence the ribonuclease, could also be used in nucleic acids (Hagen, 2000).

With new nucleic acid sequencing methods emerging and with the advent of the internet in the 1980s and 1990s, computational biology gained strength. These two factors, together with new computational methods, facilitated the creation of several nucleotide and protein databases that allow detecting structural and functional similarity between sequences. Genome projects, in particular the human genome project, also contributed to the creation, use and popularization of new tools in bioinformatics (Kanehisa & Bork, 2003).

Currently, with the advent of high-performance sequencing technologies, a large volume of data is generated and deposited daily in databases for genomic data. As a result, the biological sciences have increasingly depended on computational tools. With the development of genomic analysis platforms and bioinformatics tools, today, it is possible to carry out numerous analyses and studies, such as alignment and de novo assembly of sequences, detection

of polymorphisms and other structural variations in genomes, transcriptome analyses and metagenome analyses, development of molecular markers, among other applications (Giachetto & Higa, 2015). Thus, we can say that bioinformatics has become a fundamental tool in molecular biology research (Petrey & Honig, 2014).

Aims

To develop an effective, easy-to-use tiled amplicon approach to HRSV sequencing, suitable for a range of samples (wide range of collection dates, Ct values, storage methods, etc) which can be used with existing SARS-CoV-2 sequencing methodologies. It is likely to be of use to laboratories sequencing RSV globally, particularly those in low- and middle-income countries. Thus, this project aims to contribute to the epidemiology of HRSV particularly during the pandemic period of COVID-19, with the aim of assisting in the global knowledge of viral circulation and in the prevention of future public health problems.

Materials and methods

Collection and diagnosis of samples by Virology Diagnostics, NHS Lothian

Nasopharyngeal and throat swab samples were referred to virology diagnostic laboratory, NHS Lothian, in tubes containing 3 mL of Viral Transport Medium (VTM) or aspirate and sent to the labs. HRSV A or B positive samples from the NHS Lothian area were identified via Luminex NxTag Respiratory Pathogen Panel testing. After diagnosis, an aliquot of around 250 µl of the samples positive for HRSV were kept and stored at -80°C until use.

Viral RNA extraction

Nucleic acid was extracted from clinical specimens using the NucliSENS EMAG automated extraction platform (Biomerieux), following the manufacturer's instructions.

In a Class 1 Safety Cabinet, 2 mL of Lysis buffer was added to each well of new and sterile sample input segments. The samples were vortexed prior to use, and 200 µL from each was added to the segment and left for 10 minutes to lyse. Following instructions for machine setup with loading of input segments and aspirators, at the end of the process I obtained 110 µL of extracted RNA material. For each set of 24 samples, I included one negative control of VTM. The extracted material was kept in a cold room at 5ºC until used in molecular assays.

Real time RT-PCR for CT values

Cycle threshold (CT) values were then established via RT-PCR using the methodology outlined in Templeton et al., 2004. Reactions were performed on plates of 96 wells with a final volume of 10 µl, being 5 µl of mix (Invitrogen

QPCR Supermix), 0.5 µl of Primers/Probes, 1 µl of Enzyme (Invitrogen Superscript mix) and 3.5 µl of RNA. The Real Time RT-PCR reaction was conducted in the PCR platform Abi7500 FAST PLATFORM, using a cycle of: 50ºC for 15 min, 95ºC for 20 sec, (95º for 3 sec, 55ºC for 30 sec, 72ºC for 10 sec) x 45.

Inclusion criteria

HRSV positive samples were selected for sequencing using the cycle threshold value (Ct) as a criterion, and those that had the Ct lower than 30 were chosen, due to the difficulty in sequencing samples with higher Ct as discussed in different SARS-CoV-2 studies (Chrzastek et al, 2022) and through our Virology Sequencing Service in-house validation of SARS-CoV-2 sequencing method. A Ct cut-off study for our samples was not performed at this time. In total 45 HRSV A and 45 HRSV B samples were chosen for Nanopore sequencing, ~50% of which were samples from the 2019/2020 HRSV season with the remaining half from 2021/2022.

Ethical aspects

Southeast Scotland; with appropriate tissue bank approval; Lothian NRS BioResource RTB approval (REC ref – 20/ES/0061).

One-step RT-PCR for WGS with Illumina system (long amplicon method)

For this method the provided protocol from the Victorian Infectious Diseases Reference Laboratory at The Peter Doherty Institute for Infection and Immunity in Melbourne, Australia, was used. Each HRSV sample required four independent sets of primers/PCR reactions. One primer set for each fragment

of the genome (Table 2). In the clean room, a mix of primers and water was created for every fragment. Working concentration of the primers was 20µM.

Fragment	Primer set (20µM)	Sequence (5'-3')
1	RSV_v2_49F	GGC AAA TAA GAA TTT GAT AAG TAC
	RSV_v2_4049R	TTT GAT TGM AAA WCG TGT AGC TG
2	RSV_v2_3944F	GCC ACA RAG TCA ATT YAT AGT AG
	RSV_v2_7528R	TGT RAC TGG TGT GYT TYT GG
3	RSV_Seg3_7215F	TGA TGC ATC AAT ATC TCA AGT CA
	RSV_Seg3_11165R	GRC CTA TDC CTG CAT ACT C
4	RSV_Seg4_10959F	TGG ACC ATW GAA GCY ATA TCA
	RSV_Seg4_15333R	AGT GTC AAA AAC TAA TRT CTC GT

Table 2 – List of Primer sets from VIDRL

Primer sequences from the protocol provided by the Victorian Infectious Diseases Reference Laboratory at The Peter Doherty Institute for Infection and Immunity in Melbourne, Australia

5 µL of extract was added to the primer mix and the reaction was run on Bio-rad thermocycler (Bio-rad, C1000) for a short denaturation cycle. Subsequently, 40 µL of a second mix containing 0.5 µL of Enzime (SuperScript IV RT Mix), 25 µL of Master Mix (2X Platinum SuperFi RT-PCR Master Mix) and 14.5 µL of nuclease free water was added to each well containing the denatured genetic material, for a total reaction volume of 50 µL.

The second reaction was run on the Bio-rad thermocycler under the following conditions: 55ºC for 10 min, 98ºC for 2 min, (98º for 10 sec, 55ºC for 10 sec, 72ºC for 2min) x 40, 72ºC for 5 min and 4ºC ∞.

Electrophoresis and normalisation

The PCR products were then checked for fragment amplification by running the products on E-Gels (E-Gel EX 2% Agarose, Invitrogen, Thermo Fisher Scientific). Expected results for the amplicons from each of the four independent PCR for Illumina sequencing should be a single band and

approximately 4kb in size. The PCR products were also run on Agilent 4200 Tape station using D5000 screen tape.

Following confirmation of amplification by electrophoresis, prior to library preparation, the protocol required quantification of amplicons through Qubit fluorometer or Tape station, followed by a normalisation step where all 4 fragments should be of the same concentration for every sample.

Primer scheme design for use with adapted Artic Lo-Cost V3 protocol

For our second approach to sequencing HRSV, initial work focused on selecting geographically diverse high quality RSV genomes from the GISAID database in order to design a tiled amplicon primer scheme. In total 6 RSV A and 6 RSV B genomes were selected, with collection dates spanning 2019 to 2021, and locations spanning the UK, Australia and South Africa (Table 3).

Location	Date	HRSV A Reference	HRSV B Reference
England	2020	hRSV/A/England/RS20000581/2020|EPI_ISL_1647421|2020-02-04 *	hRSV/B/England/RE20000104/2020|EPI_ISL_1647600|2020-01-30 *
England	2019	hRSV/A/England/RE19003299/2019|EPI_ISL_1520435|2019-12-18	hRSV/B/England/194520782/2019|EPI_ISL_1834174|2019-11-04
Australia	2021	hRSV/A/Australia/VIC-RCH035/2021|EPI_ISL_2543800|2021-02-28	hRSV/B/Australia/VIC-RCH060/2021|EPI_ISL_2543853|2021-02-22
Australia	2019	hRSV/A/Australia/VIC-RCH072/2019|EPI_ISL_1834101|2019-05-13	hRSV/B/Australia/VIC-RCH077/2019|EPI_ISL_1760384|2019-06-02
South Africa	2021	hRSV/A/South_Africa/NICD-R04470/2021|EPI_ISL_12529646|2021-03-30	hRSV/B/South_Africa/NICD-R09064/2021|EPI_ISL_11055801|2021-06-21
South Africa	2019	hRSV/A/South_Africa/NICD-R06229/2019|EPI_ISL_9003918|2019-05-31	hRSV/B/South_Africa/NICD-R06224/2019|EPI_ISL_9003920|2019-05-31

Table 3 – Selected genomes from GISAID for primer design through Primal Scheme.

References marked with a * represent the genomes selected as primary references.

Draft primer schemes were generated using Primal Scheme (https://github.com/aresti/primalscheme 18; Quick et al., 2017). The most recent UK based sample was used as the primary reference for primal scheme design for each scheme (Table 3). An amplicon size of 400bp was targeted, to facilitate a drop-in replacement to existing SARS-CoV-2 lab protocols, resulting in two distinct primer schemes containing 50 amplicons each, one for HRSV A and another for HRSV B (Table 4 and Table 5).

Name	Pool	Primer Sequence	Name	Pool	Primer Sequence	Name	Pool	Primer Sequence
RSVA_1_LEFT	1	TTGGTTAGAGATGGGCAGCAAC	RSVA_18_LEFT	2	ATCCAACCTGCTGGGCTATCT	RSVA_35_LEFT	1	CTGATAGTGATAAATCAAGAAGAGTATTAGAGT
RSVA_1_RIGHT	1	AGGTCAAATCCAAGTAATTCAGATAATTGA	RSVA_18_RIGHT	2	AGATTGTGATGGGTACTCGGATG	RSVA_35_RIGHT	1	CGAAATGCTTGATTGAATTTGCTGA
RSVA_2_LEFT	2	TGGCCTAATAGATGACAATTGTGAAA	RSVA_19_LEFT	1	GCTATCCAAGCCCATCACAAGT	RSVA_36_LEFT	2	AGCAGGAATAAGTAACAAATCAAATCGT
RSVA_2_RIGHT	2	TGACCAGAAATGTAAATGTGGCCT	RSVA_19_RIGHT	1	AGTTATAACACTAGTATACCAACCAGTTCT	RSVA_36_RIGHT	2	TGTCACCATTAATTAAAGCAGTAATTGAGA
RSVA_3_LEFT	1	TCTAACCAGAGACATCATAACACATAAATT	RSVA_20_LEFT	2	TCCAGTCAAAACATCACTGAAGAATT	RSVA_37_LEFT	1	AGGGTGGTGTCAAAAACTATGGA
RSVA_3_RIGHT	1	CCCATCTTTCATCTTATGTCTCTCCT	RSVA_20_RIGHT	2	CAGGACCTTGGATACGGCAAT	RSVA_37_RIGHT	1	TCTTGTGTCAAACTACCTATAGATTCTAGA
RSVA_4_LEFT	2	ACAACTTTATGCATAATCACACTCCA	RSVA_21_LEFT	1	ACACTCAACAATACCAAAAACACCA	RSVA_38_LEFT	2	ACCCAGCTAGTATAAAGAAAGTCCTAA
RSVA_4_RIGHT	2	TGTTACATCCACTCCATTTGCCT	RSVA_21_RIGHT	1	ACATATAAGTGCTTACAGGTGTAGTTAC	RSVA_38_RIGHT	2	TGAACACAGAGTGAACTATAGCCT
RSVA_5_LEFT	1	GCTATGTCTAGATTAGGAAGAGAAGACA	RSVA_22_LEFT	2	AGTTCCAACAAAAGAACAACAGACT	RSVA_39_LEFT	1	ATTTGCCCATGTTATTTGGTGGT
RSVA_5_RIGHT	1	TCCTTGGGTAATAAACCTTTATAACGTT	RSVA_22_RIGHT	2	TGTTTCAGCTTGTGGGAAAAAGG	RSVA_39_RIGHT	1	CTGCATAATATCATTAAGATCTATCTCTGTAGT
RSVA_6_LEFT	2	AGCAGCAGGAGATAGATCAGGT	RSVA_23_LEFT	1	CAACCAACACAAAGGAAGGATCC	RSVA_40_LEFT	2	TAGACTGGCAGTTACCGAGGTT
RSVA_6_RIGHT	2	TGCTTTTGGGTTGTTCAATATATGGT	RSVA_23_RIGHT	1	TACAGACACAGTATCCACCCCC	RSVA_40_RIGHT	2	GGATAAAGACCAAGATCTTTCTCTAACG
RSVA_7_LEFT	1	ATGGAACAAGTTGTGGAGGTGT	RSVA_24_LEFT	2	ACAGCATCCAATAAAAATCGTGGG	RSVA_41_LEFT	1	ACCATTAGATTGTAACAGAGATAAAAGAGA
RSVA_7_RIGHT	1	CTTCTCCATGGAATTCAGGAGCA	RSVA_24_RIGHT	2	GTGCTTCTGGCCTTGCAGTATA	RSVA_41_RIGHT	1	AACCCAAGAGTTCCTATGCTAAGT
RSVA_8_LEFT	2	TGACAGCAGAAGAACTAGAGGC	RSVA_25_LEFT	1	TCCACCACAAATATCATGATAACTACCA	RSVA_42_LEFT	2	TCTATTAGCAAAATTGGATTGGGTGT
RSVA_8_RIGHT	2	TGAAAAAGGATTATCACTTGGCGT	RSVA_25_RIGHT	1	TTGCAAGGATTCCTTCGTGACA	RSVA_42_RIGHT	2	TGGGAGGTTTCATCAAATGTATCTCA
RSVA_9_LEFT	1	AGATGATACTGTAGGGAACAAGCC	RSVA_26_LEFT	2	AACCATCTCACTTACACTTTTTAAGTAGA	RSVA_43_LEFT	1	AGCTTTGGCCTTAGCTTAATGTCT
RSVA_9_RIGHT	1	TCATTCCTGAGTCTTGCCATAGC	RSVA_26_RIGHT	2	CAGTAAGGAGTTTGCTCATGGC	RSVA_43_RIGHT	1	TGATCAGTTATATATCCCTCTCCCCA
RSVA_10_LEFT	2	GTGAAATACTAGGAATGCTTCACACA	RSVA_27_LEFT	1	AGAGTATGCCCTCGGTGTAGTT	RSVA_44_LEFT	2	ACTAATTTAGCTGGACATTGGATTCTT
RSVA_10_RIGHT	2	TTTTGGCTGGTTGGCTAATCAG	RSVA_27_RIGHT	1	GGAATTTATACTACAAGGATATTTGTCAGGT	RSVA_44_RIGHT	2	TAACAACCCAAGGGCAAACTGT
RSVA_11_LEFT	1	ACACAACACCAATAGAAAACCAACA	RSVA_28_LEFT	2	AGCATAACCATCAATAACCCAAAAGA	RSVA_45_LEFT	1	ACATAGAGTGAAGGGATGTCATAGC
RSVA_11_RIGHT	1	TGCTAGCACTGCACTTCTTGAG	RSVA_28_RIGHT	2	AGCAGAATTTCCGCTAATAATGGGA	RSVA_45_RIGHT	1	ACCTATACAATAGTCACTCAGTGTCTT
RSVA_12_LEFT	2	AGCCAATGTCAATATACTAGTGAAACAA	RSVA_29_LEFT	1	AGTGTCATAATACTCAATCCTAATACTTACCA	RSVA_46_LEFT	2	AGGATTGCTAATTCCGAATTAGAAAGT
RSVA_12_RIGHT	2	TGGCATTTTTGAATTCAGTGGTTGT	RSVA_29_RIGHT	1	CTTTTTGAGTAAATTAGTAGTAGTAGTCTGTTC	RSVA_46_RIGHT	2	CAACCTGTAGAACTAAATACAAAATTGAATCTG
RSVA_13_LEFT	1	ACCAACATACCTAAGATCCATCAGC	RSVA_30_LEFT	2	GTCCTTAATATCCAAGTATCATAAAGGTGA	RSVA_47_LEFT	1	CTTCTCATCAAATATCCTTAGTGCACA
RSVA_13_RIGHT	1	TGAGTTTTGTTTGATTGATTGAACCAC	RSVA_30_RIGHT	2	AACCATGATGGAGGATGTTGCA	RSVA_47_RIGHT	1	ACAAAAAGATTGATAGGTTCAGCAAACT
RSVA_14_LEFT	2	ACATCAATGAGTAGATTCATACAAACTTTC	RSVA_31_LEFT	1	GCAGACAAAAATCAATCCACAAAACA	RSVA_48_LEFT	2	GGTGAAAATTTGACCATTCCTGCTA
RSVA_14_RIGHT	2	TGGAATACATTATATTCGCAGAGTTTGT	RSVA_31_RIGHT	1	ACATTATTGAATCCACATCTTAAGCCT	RSVA_48_RIGHT	2	GGCATGATGAAATTTTTGGTTCTTGA
RSVA_15_LEFT	1	TCTGGCCTTACTTTACACTAATACACA	RSVA_32_LEFT	2	TCTTGACATGGAAAGATATTAGCCTTAGT	RSVA_49_LEFT	1	GGTCCTGCAAATGTGTTCCCA

RSVA_15_RIGHT	1	AGATGATTGAGAGTGTCCCAGGT	RSVA_32_RIGHT	2	ACTTACTCAATAGAATTATCCATCTGCCA	RSVA_49_RIGHT	1	GGTTGTCAAGCTGTTTAACAATTCAC
RSVA_16_LEFT	2	GGGCAAATGCAAACATGTCCAA	RSVA_33_LEFT	1	AAAAATCTGCTATCAAGAGTATGCCA	RSVA_50_LEFT	2	AGAGTGTTGTTAGTGGAGATATACTATCA
RSVA_16_RIGHT	2	TTGTGGATTGTGGGGTTGACTC	RSVA_33_RIGHT	1	TCCAACAAGGAAGGATAAGTGTTTAGT	RSVA_50_RIGHT	2	CCTAGATCAAAATGATAATTTTAGGATTGGTTC
RSVA_17_LEFT	1	CTTCTCCAATCTGTCCGGAACT	RSVA_34_LEFT	2	ACAGATGGCCTACTTTAAGAAATGC			
RSVA_17_RIGHT	1	TGATGGTTGGCTTTCCTGTAGG	RSVA_34_RIGHT	2	ACATGATTCGGGTTGTTAAGATAACTT			

Table 4 – Primer list for HRSV A

Name	Pool	Primer Sequence	Name	Pool	Primer Sequence	Name	Pool	Primer Sequence
RSVB_1_LEFT	1	TGATGAGTGCTATTTAAGTCTAACCTTT	RSVB_18_LEFT	2	ACCAACTACAAAACCCACAAACAA	RSVB_35_LEFT	1	GCAACCAGGTATGTTTAGGCAAAT
RSVB_1_RIGHT	1	AGTCATTACTGAGTCACTTAGTCTTTTAGA	RSVB_18_RIGHT	2	ATCACGGTTCTCTGCTAAGATGT	RSVB_35_RIGHT	1	ACTTCATTAAGATTGACAACATGATCCT
RSVB_2_LEFT	2	ACAATGCCAATATTACAAAATGGAGGA	RSVB_19_LEFT	1	CGAGCCCTCCACATCAAATTCT	RSVB_36_LEFT	2	TGGTTGCATTTAACAATACCTCTTGT
RSVB_2_RIGHT	2	TGTGTGTGATGATTTCTTTGGTGAG	RSVB_19_RIGHT	1	CTGTAATTCAGTTACTGCATTCTTATACTTATC	RSVB_36_RIGHT	2	TCGAGATATATAGGTCTCAGTTCCCT
RSVB_3_LEFT	1	AACGACAACACCACCATGCAAA	RSVB_20_LEFT	2	TTGGTATACCAGTGTCATAACAATAGAATT	RSVB_37_LEFT	1	CAGTCAATTGATATAAGTAAACCAGTTAGAC
RSVB_3_RIGHT	1	TTGGAGAATGGCTTGGTTTGGA	RSVB_20_RIGHT	2	ACACTGACCCCATTTGATAGACTG	RSVB_37_RIGHT	1	ACATAAAGCATGATTCCGGAGTTG
RSVB_4_LEFT	2	GCATTAAGCCTACAAAACATACTCCT	RSVB_21_LEFT	1	ACACCTTGAAGGAGAAGTGAACA	RSVB_38_LEFT	2	ACACAGGAGTTAGAATACAGAGGAGA
RSVB_4_RIGHT	2	TCTGGACATAGCATATAGCATACCTATT	RSVB_21_RIGHT	1	AGACATGATAGAATAACTTTGTTGCCTT	RSVB_38_RIGHT	2	ACACAAGTCAAGAATTTGTTCAGTCT
RSVB_5_LEFT	1	ACACCTAAACAAACTATGTGGTATGC	RSVB_22_LEFT	2	ACATGTTAACAAACAGTGAGTTACTATCA	RSVB_39_LEFT	1	TGTTATATCGAAGCTTTTATAGGAGAACTC
RSVB_5_RIGHT	1	ACCTGATCTATCTCCTGCTGCT	RSVB_22_RIGHT	2	AGTGTTACAAAGGCTGACTTCACT	RSVB_39_RIGHT	1	TCATAAACAACTCTTAATCCATGAGGG
RSVB_6_LEFT	2	GCATGATTCTCCAGACTGTGGG	RSVB_23_LEFT	1	TCCTTCTTTCCACAAGCTGACA	RSVB_40_LEFT	2	GTGCGCAACACTATACTACCACT
RSVB_6_RIGHT	2	TTGCTCCATTTCTGCTTGGACA	RSVB_23_RIGHT	1	CATCAGAAGGAAACACTAGAGGGT	RSVB_40_RIGHT	2	TGTCCATTGTGAACATAATACTTGGTG
RSVB_7_LEFT	1	AGGACAAGTAATGCTAAGATGGGG	RSVB_24_LEFT	2	ATGTCAACAAGCTGGAAGGCAA	RSVB_41_LEFT	1	AGCATAACTGAATTAAGCAAGTATGTAAGA
RSVB_7_RIGHT	1	TCTTTGGGGTTGAGTTGATGCT	RSVB_24_RIGHT	2	TGGTCAGCAGATTGTTGTTGGA	RSVB_41_RIGHT	1	TGTGGAAACAACTTTTTGGCCT
RSVB_8_LEFT	2	CAGAGCAACTCAAAGAAAATGGAGT	RSVB_25_LEFT	1	AGCAAAGACCAACTAAGTGGAATCA	RSVB_42_LEFT	2	TGCATCCATAGACAACAAAGATGAAT
RSVB_8_RIGHT	2	GCTTACTAGGGGTTTTCTTGGGT	RSVB_25_RIGHT	1	CATGAAGTTTTGCCTCACTAGCA	RSVB_42_RIGHT	2	TCAACTTGATGATATCAACATCTCCTG
RSVB_9_LEFT	1	CCGATAACATCTGGCACCAACA	RSVB_26_LEFT	2	TGCTTGAATGGTAGAAGATGCCA	RSVB_43_LEFT	1	TCTTAGCTTGATGTCAGTTGTGGA
RSVB_9_RIGHT	1	CATTGGTCATTAATGCTTCTGCTCT	RSVB_26_RIGHT	2	AGATGGATGGTTTGCTTGCTGT	RSVB_43_RIGHT	1	TGAACATATGATCAGTTATATACCCCTCT
RSVB_10_LEFT	2	TACTTCGGCTCGTGACGGAATA	RSVB_27_LEFT	1	ACCCAATTCACCTAAGATAAGAGTGT	RSVB_44_LEFT	2	GCTGGACATTGGATTCTGATTATTCA
RSVB_10_RIGHT	2	TTGTCAGGGTGCTGATTGATCA	RSVB_27_RIGHT	1	TGGCATCTAATAAGTTTTTAGGTGTCC	RSVB_44_RIGHT	2	GTTAACAACCCAAGGGCATACG
RSVB_11_LEFT	1	CCATTGAATCAATTGCCAGACTGA	RSVB_28_LEFT	2	AACCACAACCATTTAGATAACCACC	RSVB_45_LEFT	1	AGTTTGCATAGAATAAAAGGTTGTCACA
RSVB_11_RIGHT	1	TGGTGAAATTACTAGGCATTTGAGC	RSVB_28_RIGHT	2	TGCTCTATTAGTGGGCTTTGTCT	RSVB_45_RIGHT	1	CCACTTATACAAAATTTAGGTTTGTTTCCA
RSVB_12_LEFT	2	CAGATCTCTACGCCCAAAGGAC	RSVB_29_LEFT	1	TTTAGGGAGTTACCTTTTTAACGGC	RSVB_46_LEFT	2	ATATCACCCAACCCCAGAAGCT
RSVB_12_RIGHT	2	AGCATTGGTGATAGCATTTTTGAATTC	RSVB_29_RIGHT	1	TTCCACAGCTGAGAGTATATCATCTT	RSVB_46_RIGHT	2	ATATTCTATACTGATCTTGCATCCTGTG
RSVB_13_LEFT	1	ACCAACCTATCTAAGATCAATCAGTGT	RSVB_30_LEFT	2	GGAAAAGGACAGAGTTAAGCCCA	RSVB_47_LEFT	1	TGCATCACTTTATTGCATGCTTCC
RSVB_13_RIGHT	1	TGAGTGGGATTGGTTTGTGTGA	RSVB_30_RIGHT	2	GTCAAAAATTGATTGTATGTGGTAGTTGT	RSVB_47_RIGHT	1	ACAGGTAATTCAGCATCACAGACA
RSVB_14_LEFT	2	AGGTTCACATATATCCTCAACTGCA	RSVB_31_LEFT	1	AGATAATCAAACTTTGAGTGGTTTTCAGT	RSVB_48_LEFT	2	TCCTGCTACAGATGCAACTAACA
RSVB_14_RIGHT	2	GAGTTTTGTTGCAGAATATTTTATGTTCAC	RSVB_31_RIGHT	1	TGAGCCTTAATAGCTGCATCTGT	RSVB_48_RIGHT	2	AGATTCCTTGTCAATCTTTTTAGGCATA
RSVB_15_LEFT	1	TTGGCCCTATTTTACACTAATACATATGA	RSVB_32_LEFT	2	GAAGGCTTTTACATAATAAAAGAAGTAGAAGG	RSVB_49_LEFT	1	ACTTCCTGTTTTTGATGTTGTGCA

RSVB_15_RIGHT	1	GTATCCCAGGTCTTTTCTAGAGTCC	RSVB_32_RIGHT	2	GCACCTCTCAACGTACTTAGGC	RSVB_49_RIGHT	1	GCTTCTTGAGCTCATTGGTTGT
RSVB_16_LEFT	2	GGGATCAAAAACAACATTGGGGC	RSVB_33_LEFT	1	TGGACATCCAATGGTTGATGAAAGA	RSVB_50_LEFT	2	GCAACAAGCTTATAAACCACAAGC
RSVB_16_RIGHT	2	TTGGGTGATATTGTGGCTGAGT	RSVB_33_RIGHT	1	TGTGATGGCATGTAATTTCTAGGAAA	RSVB_50_RIGHT	2	TGTCTCGTTGTGTTGTAAATGCAC
RSVB_17_LEFT	1	GTCTCACCAGAAAGGGTTAGCC	RSVB_34_LEFT	2	TCTATCGTGAGTTTCATCTGCCT			
RSVB_17_RIGHT	1	TTTTTCGGTGTTTTGGCTGGTG	RSVB_34_RIGHT	2	GTCAAACTCTCAGGGAAGAATTGTAA			

Table 5 – Primer list for HRSV B

Synthesis of complementary DNA (cDNA) for Nanopore sequencing

For the synthesis of complementary DNA (cDNA) from the extracted RNA, the NEBNext® ARTIC SARS-CoV-2 Companion Kit (Oxford Nanopore Technologies) was used with the adapted Artic Lo-Cost V3 protocol. Starting with 2 µL of the Enzyme LunaScript RT SuperMix, added across the necessary wells of a Hard-Shell 96 well Low Skirted plates (Bio-rad) for the number of samples being tested (2 separate sets of 48 samples, of which 45 were clinical samples and 2 were extraction negative controls of VTM and 1 RT-PCR negative control of nuclease free water). This was then followed by transfer of the plate to a PCR room where 8 µL of sample was added to the reaction for a total volume of 10 µL. The plate was then sealed, inverted, pulse centrifuged and loaded into a Bio-rad thermocycler (Bio-rad, C1000) with the following cycling conditions used: 25ºC for 2 min, 55ºC for 20 min, 95º for 1 min and 4ºC ∞.

Genome amplification for Nanopore sequencing

The Artic Lo-Cost V3 protocol was adapted and used for this method. Two different PCR reactions were set-up for HRSVA (pool 1 and pool 2) and a separate set used for HRSVB (pool 1 and pool 2) (Table 4 and Table 5, respectively). Primers were pooled together before this step, with 50 primers per tube/pool. The PCR reaction contained the following components; 6.25 µl of Q5 Hot Start High-Fidelity 2X Master Mix, 2 µl of primer and 0.25 µl of nuclease free water.

In the PCR room, 4 µL of RT extract was added to the primer mix in each well, then the plate was sealed, inverted to mix and centrifuged to remove droplets from lids, and the plate was then loaded into the Bio-rad thermocycler (Bio-rad, C1000). The cycler conditions were as follows: 98ºC for 30 sec, (95º for 15 sec, 63ºC for 5 min) x 35 and 4ºC ∞.

Nanopore Sequencing

The PCR products were diluted in 45 µl of nuclease free water. The two PCR reactions were then pooled. Using a multichannel pipette, 2.5 µl of Pool 1 PCR product and 2.5 µl of Pool 2 PCR product were aliquoted into the same well. The plates were re-sealed with seals, inverted to mix, and centrifuged to remove droplets from lids. The positive control, the negative control, the negative extraction controls and one sample at random were quantified using the Qubit dsDNA HS kit.

For the end repair and dA-tailing of DNA step, NEBNext Ultra II End Prep was used, with the following components making up the mastermix: nuclease free water, Ultra II End Prep Reaction Buffer and Ultra II End Prep Enzyme mix. After placing the PCR plate on a cold block, 10 µl of the End Prep mastermix were aliquoted to each well. Using a multichannel pipette, 5 µl of PCR dilution from the previous step were aliquoted into the End Prep mastermix. The plate was again re-sealed, inverted to mix, and centrifuged. The PCR plate was loaded on a thermocycler using the following cycling programme: 20ºC for 10 min, 65ºC for 10 min and 4ºC ∞.

For the attachment of native barcodes NEBNext Ultra II Ligation was used. The EXP-NBD196 (1-96) barcoding kit was thawed at room temperature then centrifuged at 2000rpm to remove drops from lids. A Ligation mastermix was prepared using Nuclease free water, Ultra II Ligation mix and Ligation Enhancer. The PCR plate was placed on a cold block, and 7.3 µl of the Ligation mastermix were aliquoted into each well. Using a multichannel pipette, 1.2 µl of NBXX barcode was added per well. Using a multichannel pipette, 1.5 µl of End Prep reaction mixture from previous step was aliquoted into the plate containing Ligation mastermix and barcodes. Plate was re-sealed, inverted to mix, and centrifuged. The PCR plate was then placed on a thermocycler using the following conditions: 20ºC for 40 min, 65ºC for 10 min and 4ºC ∞.

For DNA purification, NEBNext Sample Purification beads (SFB and ethanol) was used. After mixing the NEBNext beads thoroughly, each barcoding

reaction was pooled into the one tube. 192 µl of 0.4x SPRI beads was added to tube with barcoded samples and mixed gently by flicking. An SFB wash was performed where the tube was pulse spun to 3000 rpm and incubated for 5 min at room temperature. It was then placed on a magnetic rack and incubated for 2-5 min until beads had pelleted and supernatant was completely clear. Using a P200, the supernatant was slowly removed and discarded, without touching the pellet. 250 µl of SFB was added to the tube without touching the pellet and the beads were resuspended completely by gently flicking tube and pipette mixing. The tube was again pulse spun to 3000 rp placed on a magnetic rack and incubated for 2-5 min until beads had pelleted and the supernatant was completely clear. Using a P200, the supernatant was slowly removed and discarded without touching the pellet. The previous steps for SFB wash were repeated for a second wash.

After the second SFB wash, the tube was pulse spun to 3000 rpm, placed on a magnet and, using a P10, residual SFB was removed. Keeping the tube on the magnet, 200 µl of fresh 70% ethanol was slowly added over the pellet to bathe (without touching pellet with tip). The ethanol was then slowly removed with a P200 and the tube was pulse spun to 3000 rpm and placed on the magnetic rack again. Using a P10, any residual ethanol at the bottom of the tube was removed. The pellet was air dried at room temperature for 2-5 min. 30µl of Elution Buffer were added to the tube without touching the pellet and the beads were resuspended completely by gently flicking the tube and pipette mixing. One final pulse spin to 3000 rpm was performed and the tube was incubated for 2 min at room temperature. It was then placed back on the magnetic rack and incubated for 5-10 min until beads had pelleted and the supernatant was completely clear. 0.3 ml of supernatant was transferred to a PCR tube without touching or disrupting the beads and for quantification, 1 µl of barcoded amplicon pool was used with the Qubit fluorometer mentioned in a previous step.

For the attachment of sequencing adapters, in the 0.3 ml PCR tube with the barcoded amplicon pool, NEBNext Quick Ligation Reaction Buffer, Adapter

Mix and Quick T4 DNA Ligase were added. The PCR tube was then placed on a thermocycler at 25^{o}C for 20 min.

After retrieving and placing the MinION flow cell into one of the five GridION ports, a Flow Cell Check was performed to check the number of pores available. A minimum of 800 pores was required before proceeding.

For DNA purification, NEBNext beads were vortexed thoroughly before use and 50 µl were added on a 1:1 ratio to the tube with barcoded samples, mixing gently by flicking. The tube was pulse spun to 3000rpm and incubated for 5 min at room temperature. It was then placed on magnetic rack and incubated for further 2-5 min until beads had pelleted and supernatant was completely clear. Using a P200, supernatant was slowly removed and discarded without touching the pellet, and 250 µl of SFB were added to the tube, again without touching pellet, and beads were resuspended completely by gently flicking tube and pipette mixing. Tube was pulse spun to 3000rpm, placed on magnetic rack and incubated for 2-5 min until beads had pelleted and supernatant was completely clear. Using P200, the supernatant was slowly removed and discarded without touching the pellet.

Previous steps were repeated for a second SFB wash. The tube was then pulse spun to 3000 rpm and using P10, residual SFB at bottom of the tube was removed. 15 µl of Elution Buffer was added to the tube without touching the pellet and beads were resuspended completely by gently flicking and pipette mixing. Once more, the tube was pulse spun to 3000 rpm and incubated for 2 min at room temperature. It was then placed on magnetic rack and incubated for 5-10 min until beads had pelleted and the supernatant was completely clear. The supernatant was then transferred to a new tube without touching or disrupting the beads. 1 µl of barcoded amplicon pool was used for quantification using the Qubit fluorometer.

Volume for loading was calculated noting that 20 ng are required for sequencing in volume of 12 µl. For priming the flow cell and preparing/loading the sequencing library, Sequencing buffer, Loading beads, Flush buffer and

Flush tether were used. Final libraries were loaded on R9.4.1 flow cells and run for 16 hours.

Consensus sequence generation

We assessed negative controls and calculated overall coverage following the sequencing runs using RAMPART (GitHub - artic-network/rampart: Read Assignment, Mapping, and Phylogenetic Analysis in Real Time 3). Negative controls passed if fewer than 20 reads mapped to the HRSV genome, which is the cut-off used for the NHS Lothian SARS-CoV-2 Nanopore sequencing protocol. To generate consensus sequences we used an in-house version of the “field bioinformatics” pipeline (v1.2.1) for tiled amplicon consensus sequence generation (GitHub - artic-network/fieldbioinformatics: The ARTIC field bioinformatics pipeline 2) with a 20x depth threshold and using Nanopolish for variant calling.

Results

Identification of HRSV subgroups (A and B) and CT values

HRSV samples were subjected to a real-time RT-PCR assay to determine CT values, where a wide range of values was obtained that did not seem to directly correlate to sample age (Figure 12). When looking at the subtypes for each sample for the period between 2019 and 2021, a difference in circulation was observed over the three years (Figure 13). For 2019 (from April to December), 790 samples (90%) corresponded to HRSV-A and 82 to HRSV-B (10%), showing HRSV A was the predominant strain by a large margin in circulation in Lothian at the time. But unlike in 2019, the 2021 wave showed a prevalence of group B over A.

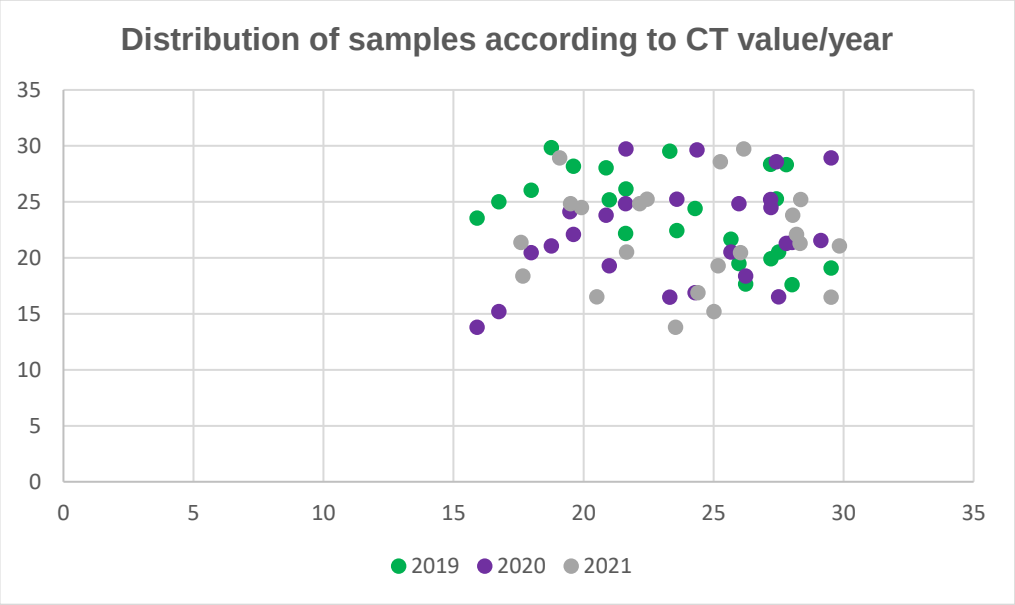

Figure 12 – Distribution of samples according to CT value/year

Representation of the different samples used in this study according to Ct and to what year they were collected. Showing that for the 3 different years (2019, 2020, 2021) Ct values ranged from 15 to 30, without correlation to sample age.

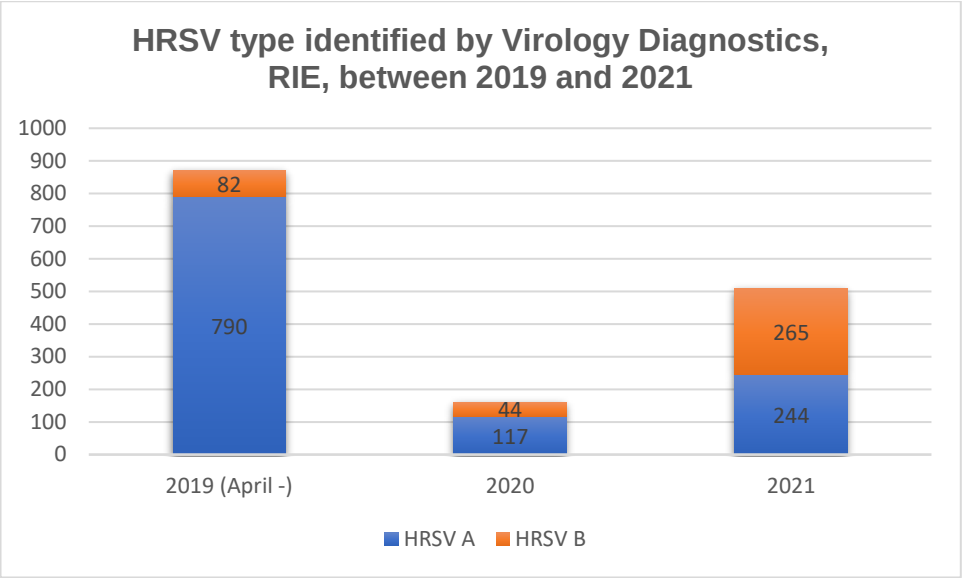

Figure 13 – HRSV type identified by Virology Diagnostics, RIE, between 2019 and 2021
Representation of the total amount of samples identified as HRSV A and B in the years 2019 (April to December), 2020 and 2021, in Lothian. In blue HRSV-A and HRSV-B in orange.

Sequencing with long amplicon approach

Initially, our goal was to sequence our samples using the MiSeq (Illumina) system to sequence the complete genome to assure the robustness of our data and contribute globally with more British HRSV genomes, since there are still not big numbers of published complete genomes in global databases.

Therefore, for the amplification of the complete HRSV genome, we used an adapted protocol from previous versions already described in the literature (Bose et al., 2015, Schobel et al., 2016). We used the primers present in the protocol which can be seen in Table 2, and samples previously identified as belonging to either groups HRSV-A or HRSV-B, from the 2019-2021 period, with a range of different CT values under 25.

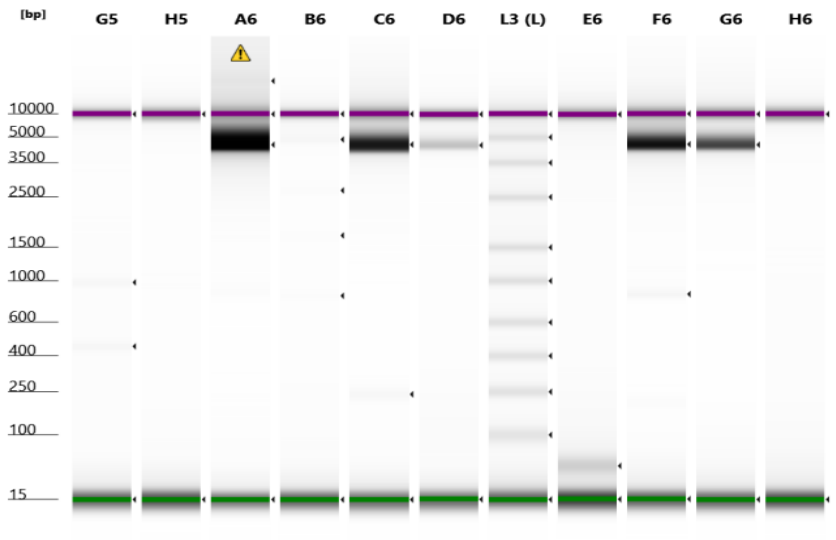

Figure 14 – Tape station results
Results for fragment 4 from a subset of samples after PCR was done using the VIDRL protocol. Samples are labelled at the top (G5, H5, A6, B6, C6, D6, E6, F6, G6, H6) and a ladder is shown in L3. Gel bands for fragment 4 (4000bp) can be seen for samples A6, C6, D6, F6 and G6, between the 3500bp and the 5000bp mark. On every other sample no 4000 bp bands are visible.

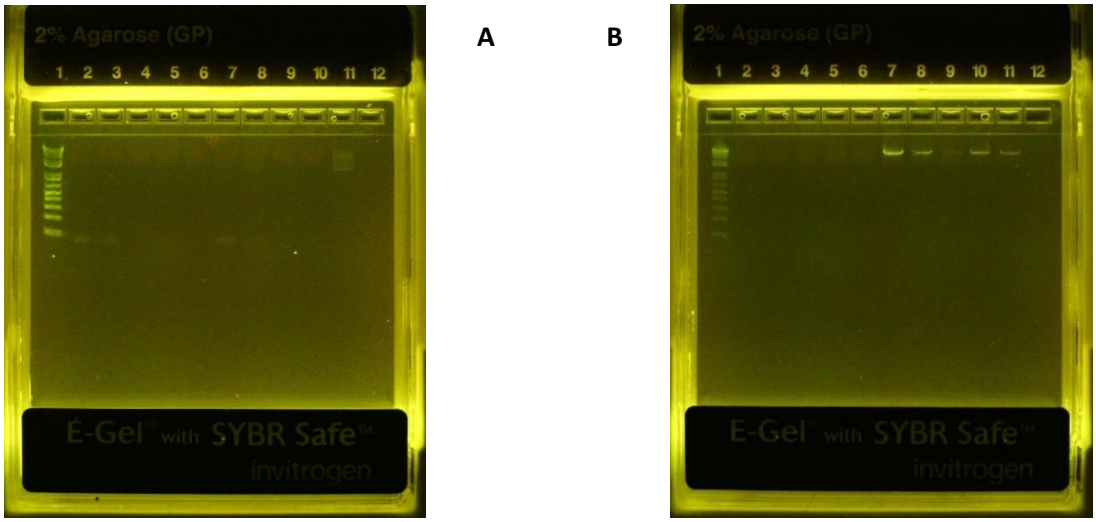

Figure 15 – E-gel results for long amplicon PCR method
For the same subset of samples present in Figure 14, A) shows results for fragments 1 and 2 post-PCR, where no 4000bp bands are visible. B) shows results for fragments 3 and 4, where amplification bands are only present for fragment 4 at the 4000bp mark. For guidance, a ladder was used in the first well for A) and B).

Even after attempts at re-extraction and reamplification, we did not obtain the expected PCR result of the samples for any of the years selected (Figure 14 and Figure 15) and consequently did not proceed to sequencing. The conditions of handling and storage of the samples, even prior to their arrival to the Virology Diagnostics department, may have influenced this result.

Therefore, when analysing the results obtained, in total only 5 samples out of 20, with different collection dates ranging from 2019-2021, had any positive result after PCR, where they amplified one of the four amplicons (amplicon 4). All other samples used in this assay amplified no fragments, making it impossible for us to use them to sequence the complete genome. Because of this, we decided to attempt to design and adapt our own sequencing method with a shorter amplicon approach, from the methods we had in place in our laboratory for SARS-COV-2 in hopes of getting full genome coverage.

HRSV complete genome sequencing by NGS using a shorter amplicon approach

Among the sequenced samples, 45 belonged to HRSV-B and 45 to HRSV-A. For 2019/2020 the majority were collected in the seasonal period of HRSV and other respiratory viruses in the United Kingdom, between the months of the year corresponding to the fall and winter seasons of this region. For 2020/2021 the majority of samples were collected in the out of season wave of summer 2021.

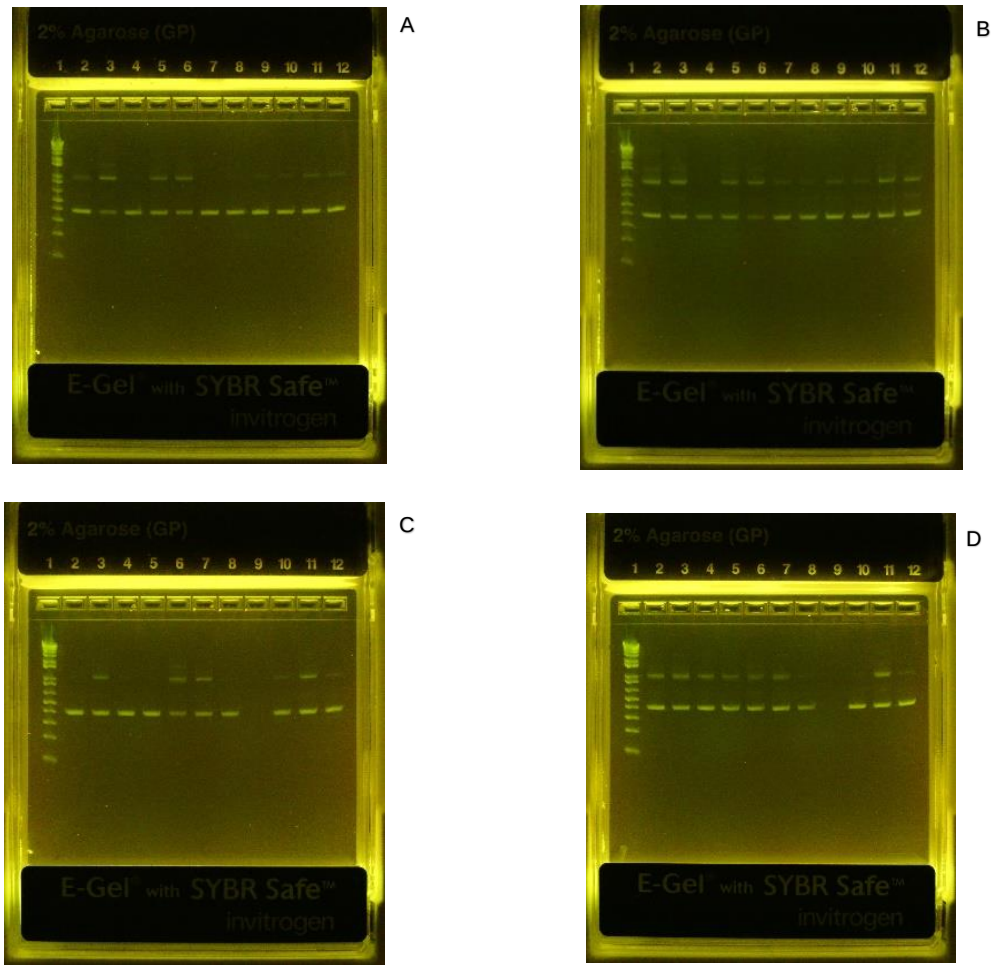

Figure 16 – E-gel PCR results for short amplicon approach
Pictures shown (A-D) demonstrate amplification of the correct bands at 400bp on E-Gels for selection of samples (21 samples of HRSV A with Ct under 30 and collection dates ranging from 2019 to 2021), with the only two gaps at well 9 of C) and well 9 of D), being negative controls. A) and C) show fragments for pool 1 and B) and D) show pool 2. Each gel contained a ladder in the first well for guidance on band size.

We successfully amplified the desired fragments using the primer schemes obtained from the Primal Scheme tool (Tables 3 and 4), with an adapted version of the Artic Lo-Cost V3 protocol we have in place in the laboratory for SARS-COV-2 sequencing. Results were confirmed on E-gels as seen on Figure 16, where all samples presented the 400bp band on the gels, and negatives did not. This allowed us to then proceed with library preparation and sequencing via Nanopore on the GridION, in conjunction with RAMPART for a

real-time overview of genome coverage and reference matching for each barcode (Figure 22).

Genome coverage analysis for short amplicon method

39 out of 45 (86.7%) clinical samples for HRSV A showed genome completeness above 90% (Figure 17). For HRSV B samples, 39 (86.7%) out of 45 clinical samples, showed genome completeness above 90% (Figure 18). After analysing results for genome completeness and Ct no correlation was found between higher Ct values and loss in genome completeness (Figure 19). Phylogenetic trees were then produced by uploading our sequencing data on Nextclade (Figure 20 and Figure 21).

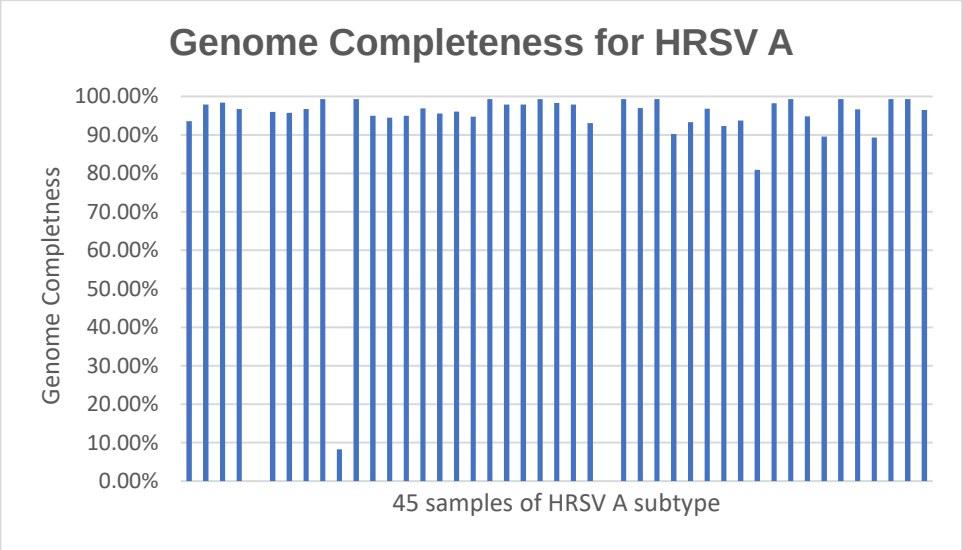

Figure 17 - Genome completeness for HRSV A
Results for genome completeness for 45 samples selected for sequencing of HRSV A with novel short amplicon method with our own developed primers from Primal scheme, showing a majority of samples with genome completeness above 90%.

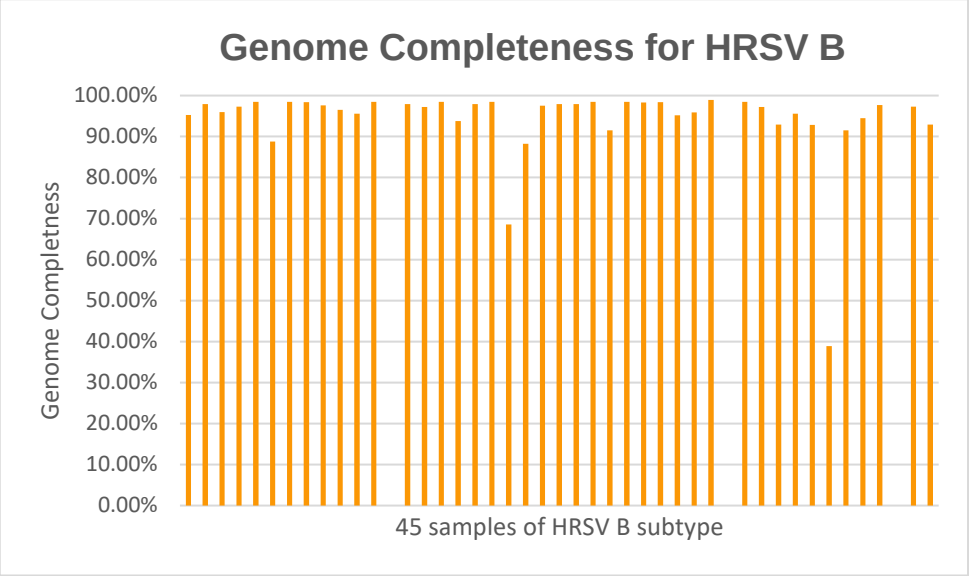

Figure 18 – Genome completeness for HRSV B
Results for genome completeness for 45 samples selected for sequencing of HRSV B with novel short amplicon method with our own developed primers from Primal scheme, showing a majority of samples with genome completeness above 90%.

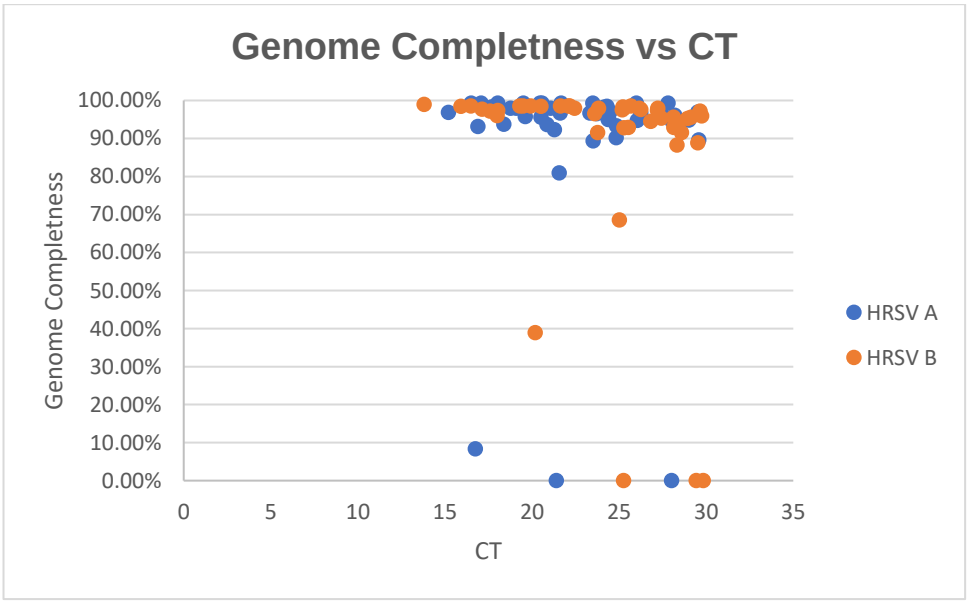

Figure 19 - Genome Completeness vs Ct value.
As shown, Ct range for samples in the study ranged from 15 to 30 and genome completeness was over 90% for most samples, with 45 being subgenotype A and 45 being subgenotype B. Subgenotype also did not seem to influence results.

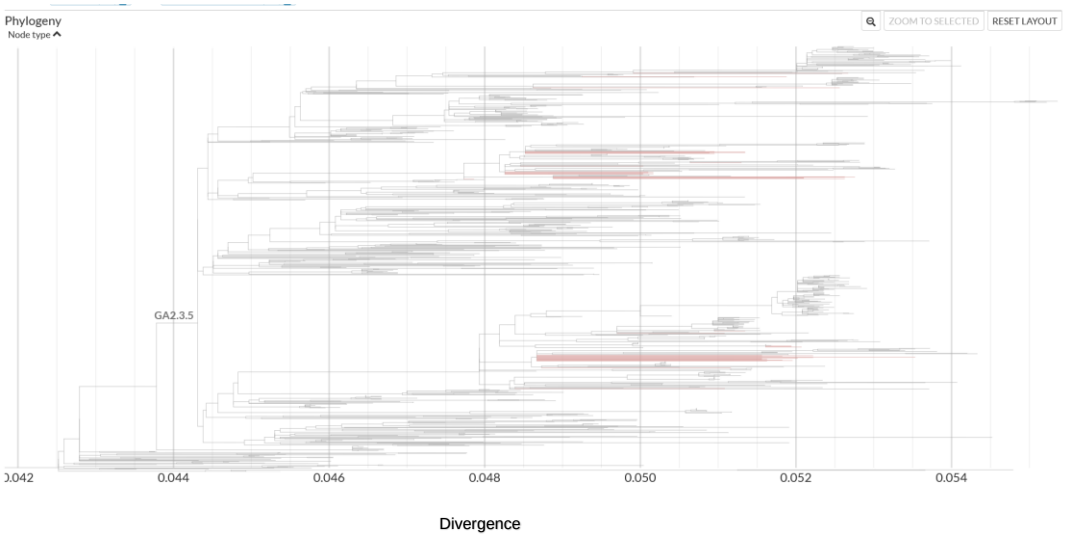

Figure 20 - Phylogenetic tree of 37 HRSV A genomes
Using Nextclade we obtained a phylogenetic tree for HRSV A. Samples uploaded from our study are represented by red lines. Clade GA2.3.5 location shown as all our HRSV A samples belong to this subgenotype

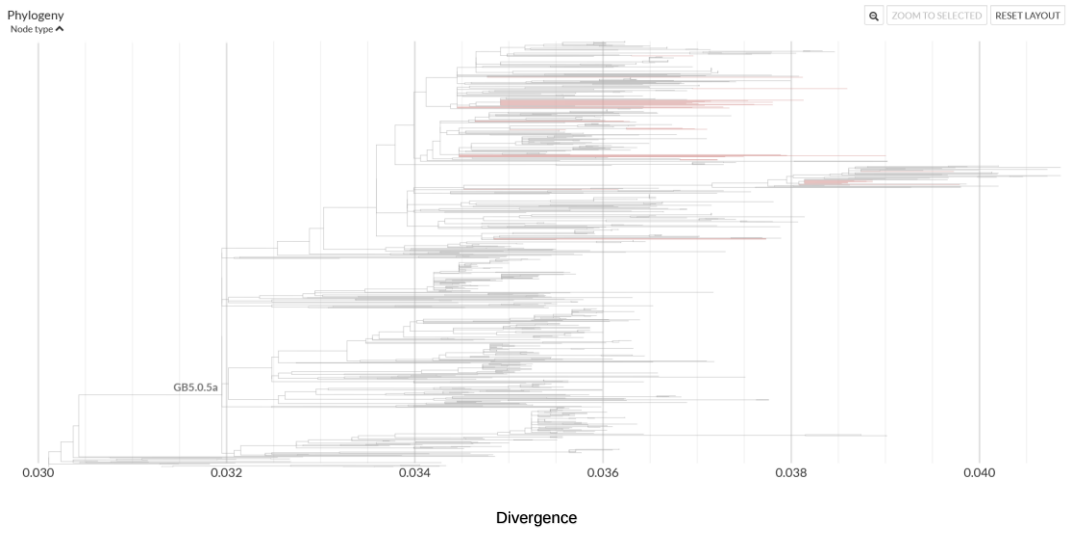

Figure 21 - Phylogenetic tree of 36 HRSV B genomes
Using Nextclade we obtained the represented phylogenetic tree for HRSV B. Samples uploaded from our study are represented by red lines. Clade GB5.0.5a location shown as all our HRSV B samples belong to this subgenotype

Discussion

In the United Kingdom, HRSV is among the most common respiratory viruses detected in individuals who have acute infections of the respiratory tract and is one of the main causes of childhood infections, where pneumonia, bronchiolitis and wheezing are characteristic (Suryadevara & Domachowske, 2021). Also in the United Kingdom, HRSV infections cause approximately 83 deaths every year among children and adolescents. HRSV can be classified into two subgroups: HRSV-A and HRSV-B and the two subgroups can circulate separately or simultaneously during epidemics, varying according to region and year (Anderson et al., 1985, Johnson et al., 1987). Some authors have pointed out that HRSV-A is more virulent than HRSV-B, while other studies found no differences between them (Papadopoulos et al., 2004, Laham et al., 2017). Vianna et al. (2021) investigated, during the period of 2016 to 2018, the prevalence of HRSV in children up to 3 years of age with acute respiratory infection. In it, children infected with HRSV-A presented greater clinical severity, remained longer in the hospital and required intensive care when compared to infected children with HRSV-B.

Current sequencing methods available for HRSV genotyping are few, costly and time consuming, so to aid in readily gaining further insights into how HRSV spreads and its biology, we developed a new sequencing approach, adapted from existing methods already in place for SARS-CoV-2 sequencing, which should particularly help low and middle income countries to easily adapt their current viral surveillance efforts to include HRSV and by doing so, helping to reduce infant and elderly morbidity and mortality caused by HRSV.

In this study, the co-circulation of HRSV-A and HRSV-B was observed between 2019 and 2021, with group A being the most found in 2019-2020 (Figure 13). On the other hand, in 2021, despite the lower number of samples, HRSV-B predominated over group A, where 52% of detected samples were from group B (Figure 13). Consideration should be given to the fact that sampling in the years 2020 and 2021 may have been influenced by the efforts

that were made at the time to primarily perform diagnosis of SARS-CoV-2, compromising the identification of other respiratory viruses and consequently the screening of HRSV samples for this and other studies. Initially, the project aimed to genetically characterise HRSV samples from the years 2019 to 2021, with particular interest towards the out of season wave in the summer of 2021 in the United Kingdom. As a result, a large number of samples from Lothian, Scotland, were selected but, due to the ongoing efforts at the time to respond to the COVID 19 pandemic in the country, and as our laboratory is one of the main sites for SARS-COV-2 surveillance in Scotland, only a small subsection was included at the end of this work.

Although HRSV A is more frequently detected than HRSV B, some studies show the predominance of HRSV-B in certain outbreaks (Martínez et al. 1999). Different patterns of HRSV circulation are described in the literature, where after every one or two years of group A's dominance there is one year of B dominance (Hall et al. 1990). A 10-year study in Finland showed the cyclic occurrence of the virus and that every two years the prevalence alternates between A and B (Waris 1991).

The frequency of HRSV, as well as other respiratory viruses, is clearly influenced by a seasonal pattern that may vary according to the geographical location and climatic conditions of each country (Gurgel et al. 2016). In temperate countries, the circulation of respiratory viruses intensifies in the winter period, while in countries with equatorial climates, circulation is higher during the rainy season (Obando-pacheco et al. 2018). In the study conducted by Hervás et al. (2012), in Spain, it was observed that the annual rates of hospitalization by HRSV were strongly associated with weather conditions, where atmospheric pressure and average temperature were the main factors related to increase and reduction of cases. Research focused on the epidemiology and seasonality of respiratory viruses in the United Kingdom have suggested a seasonal pattern for HRSV infections covering the fall and winter, and fewer cases in the summer months, mainly from November to February (Reeves et al. 2020).

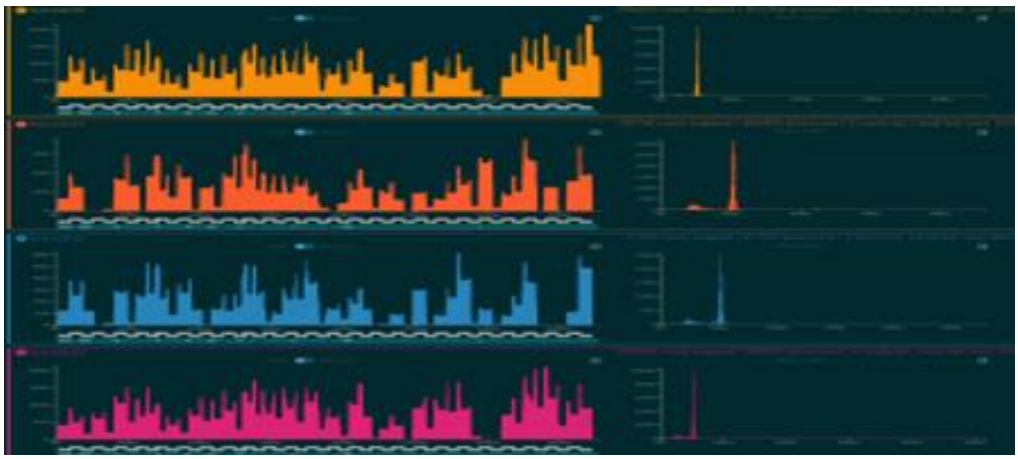

Figure 22 - Genome coverage for HRSV B using RAMPART
Genome coverage for 4 samples belonging to HRSV B, on RAMPART, showing the tiled amplicons along the bottom of each genome (in grey on the left side). Each colour represents a different sample and dropouts are shown as the empty areas where coverage failed to reach optimal levels. On the right side, fragment size of sequenced data is shown, with each spike showing results that we have sequenced the expected size of 400bp for our amplicons.

When analysing data from RAMPART, amplicon 3 from the HRSV A primer scheme was the amplicon with the biggest dropout as coverage did not reach the 20x depth threshold for 24 (53.3%) out of 45 clinical samples. For HRSV B primer scheme, amplicon 45 failed to reach the 20x depth threshold for any sample, being the only amplicon from this new scheme that failed to amplify successfully. These results are encouraging and future work will involve developing new primers that can be added to our current scheme in order to improve genome coverage.

On defining the different genotypes, Goya et al. (2020) compared the patterns of grouping of complete genome sequences with the three surface glycoprotein genes (SH, G and F), the ectodomain, and the 2nd hypervariable region of the G gene of HRSV. These authors argue that, although complete genome analysis has achieved the best outcomes, the ectodomain phylogeny of F and G, showed statistical support comparable to the complete genome for defining genotype. Thus, they report that using the sequences of the ectodomain of G would provide the minimum region suitable for HRSV

genotyping. Based on their analyses, they proposed that the number of genotypes can be reduced from 13 to 3 for HRSV-A (GA1-GA3) and from 20 to 7 for HRSV-B (GB1-GB7). Figures 23 and 24 show phylogenic data available as a tree that can be used to infer divergence trends over time worldwide and used for comparison with our own results shown in Figures 20 and 21. Looking at preliminary results from Nexclade these indicated that for the G_Clade classification proposed by Goya et al. (2020), our HRSV A samples belong to subgenotype GA2.3.5 and our HRSV B samples belong to GB5.0.5a (Figure 20 and Figure 21, respectively).

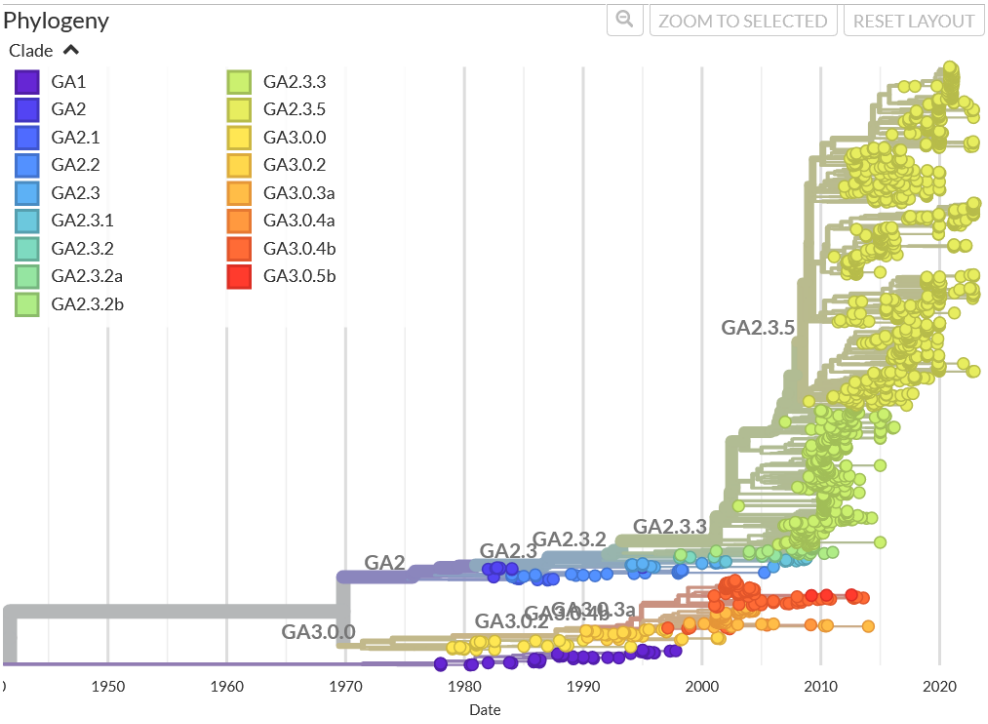

Figure 23 – Phylogeny tree for HRSV A available on Nextclade
Showing 1939 genomes enabled on Nextclade from Genbank, sampled between December 1977 and December 2022. Divergence shown over time with clade GA2.3.5 containing the most recent samples uploaded to the platform

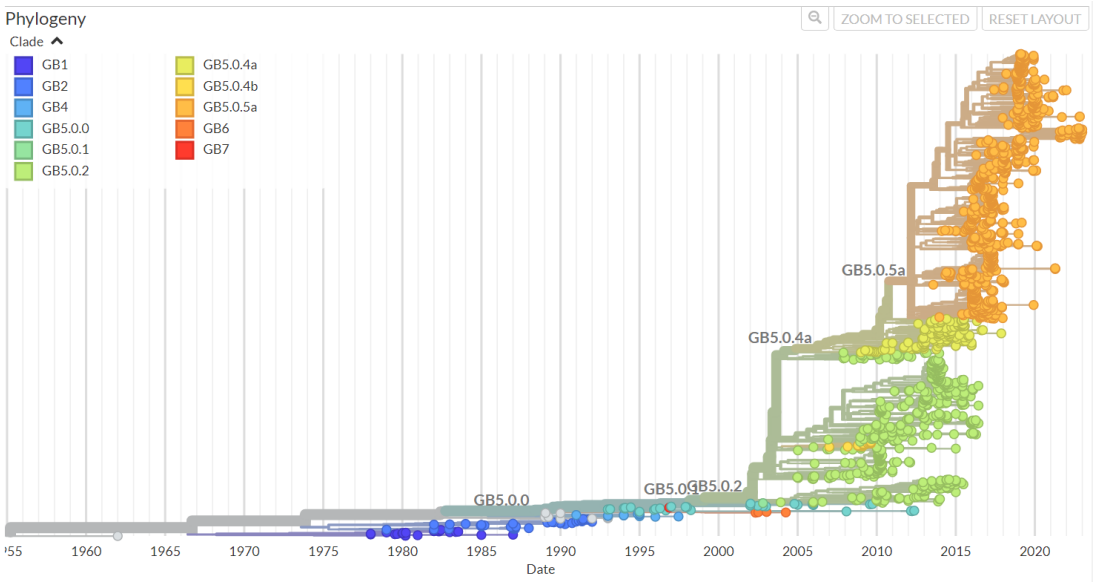

Figure 24 – Phylogeny tree for HRSV B available on Nextclade
Showing 1904 genomes enabled on Nextclade from Genbank, sampled between January 1962 and December 2022. Divergence shown over time with clade GB5.0.5a containing the most recent samples uploaded to the platform

The definition of HRSV genotypes is often performed on the basis of the alignment of the 270 nucleotides of the second variable region of the gene that encodes the G protein or the complete G (Mufson et al., 1985). The ON1 genotype was first detected in Ontario, Canada, in 2010 by Eshaghi et al. (2012) and has a unique genetic characteristic, a duplication of 72 tandem nucleotides in the 2nd hypervariable region of the gene encoding for the G protein. Since its discovery, it has been reported in several countries around the globe, in addition to this, several studies also indicated that the emergence of ON1 quickly replaced the NA1 genotype previously dominant (Auksornkitti et al. 2014, Hirano et al. 2014, Yoshihara et al. 2016). The BA genotype was first isolated in samples collected in Buenos Aires in 1999 by Trento et al. (2003) and had a rapid worldwide dispersion, becoming the predominant genotype of group B, identified in several countries such as Kenya (Agoti et al. 2013), Japan (Dapat et al. 2010), Spain (Gimferrer et al. 2016), China (Song et al. 2017, Zhao et al. 2019), Africa (Kenmoe et al. 2018) and Brazil (Vianna

et al. 2021). Therefore, it is expected that most of the HRSV-B samples in our study belong to this genotype.

Although several genotypes have been described for HRSV, there is no consensus related to genotype attribution criteria, impacting directly in understanding viral evolution, developing detection methods and design of vaccines (Muñoz-Escalante et.al 2021). The definition criteria most commonly used for genotyping of HRSV was proposed by Peret et al. (1998) and Venter et al. (2001), based on HVR2 of the G gene, in which a genotype is distinguished on the basis of a phylogenetic cluster with bootstrap support values of at least 70% and a distance of <0.07 subst./site for all other members of the cluster. It is important to point out that this definition was formulated at a time when genetic sequencing was a time-consuming and costly process (Ramaekers et al. 2020).

However, Ramaekers et al. (2020) showed in their analyses that neither the HVR2 fragment nor the G gene contain sufficient phylogenetic signal to perform a reliable phylogenetic reconstruction and therefore recommend the use of full-genome HRSV alignments to determine the Genotypes. They propose the following criteria: a ≥ 70% support bootstrap for the respective clade and a maximum patristic distance between all members of the clade of ≤0.018 site replacements for HRSV-A or ≤0.026 replacements per site for HRSV-B, this way distinguishing twenty-three genotypes of A and six B Genotypes.

Another group proposed new classifications for HRSV-A (Muñoz-Escalante et al. 2019) and HRSV-B (Muñoz-Escalante et al. 2021), evaluating complete genome sequences, of ectodomain of G and the other individual genes of HRSV, using Mean Variance (MV) analyses and Bayesian phylogenetics, as well as matrices of intergenotypic and intragenotypic distance. They indicate that the 13 genotypes described for HRSV-A can be classified as 9 and the 37 described previously for HRSV-B can be classified into 15.

Efforts to establish criteria for the definition of HRSV have mobilized several research groups that are working together to resolve this issue. As the virus continues to evolve, HRSV genotyping designations will need to be defined and be adaptable to changes in the future and should also be periodically reassessed by a global consortium (Salimi et al. 2021).

When looking at genetic divergence for the different subtypes, a study in Australia, when analysing a set of local data, identified that both subgroups of HRSV evolved rapidly in a similar way to a clock, with average rates of approximately 6–8 × 10^{-4} nucleotide substitutions per site per year. Contextualizing on a global scale, they analysed a set of data by the least square dating (LSD) method and the result was 5.72 × 10^{-4} for HRSV-A and 6.41 × 10^{-4} for HRSV-B (Di Giallonardo et al. 2018). These rates are within the range previously reported for other paramyxovirus, ranging from 6.585 to 11.350 x 10^{-4} substitutions by site and year (Pomeroy et al. 2008). For HRSV-A, Otieno et al. (2018) estimated the rate of genome-wide replacement for ON1 viruses as 5.97 × 10^{-4} substitutions of nucleotides per site/per year (95% HPD: 5.42-6.58 × 10^{-4}). Previously, higher evolution rates for HRSV-B had already been described suggesting that there are different selective pressures between groups (Schobel et al. 2016).

Di Giallonardo et al. (2018) points out in their study that highly skewed global HRSV sampling prevents extensive analysis of the distribution and the dynamics of transmission of the virus and they stress that increased molecular surveillance is essential to improve understanding of epidemiology and evolution of HRSV and to assist in vaccine development projects. The estimated origin of HRSV introductions in Lothian might be possible to infer once we look at our data in more detail, but due to the scarcity of complete genomes deposited in the online databanks such as GenBank for Scotland, there is a risk that our results will be heavily influenced by a specific region due to site and method of sampling. Robertson et al. (2021) in their study, reports that their analyses were also biased to regions that have been highly sampled, such as the US and Kenya, and that the lack of data from many territories

around the world undermine inferences about the global and local viral traffic. To elucidate such issues in the Australian context, they used phylogenetic analysis of MV in a better data set derived from sequences of the G gene region only instead of the complete genome and although this analysis has improved the location sampling, there was worse phylogenetic outcome. Vianna et al. (2021) say it is necessary to expand sequencing of HRSV samples in an overall manner to check if there is a formation of local genetic groups or whether this observation is caused by a bias of sampling.

GISAID (Global Initiative on Sharing Avian Influenza Data), is a worldwide effort created for global virus data sharing for influenza virus in 2006 (gisaid.org) with the epiflu genomic database. In 2020 it expanded its activities and offers of analysis tools on due to the pandemic of SARS-CoV-2, and today the EpiCov database already has more than 15 million of genomes shared by groups around the world. Among the available tools there is CoVsurver (gisaid.org/covsurver) to assist the scientific community in the identification, analysis and interpretation of mutations in coronavirus genomes. Due to the epidemiological importance and impact of HRSV in the world, GISAID together with WHO has also created the database EpiRSV which was released in June 2021. EpiRSV also has genomic surveillance tools and analyses for HRSV, helping to elucidate knowledge gaps in relation to transmission, risk factors for infection and, above all, from genetic data it will assist in the identification of the viral variants circulating in real time and in the interpretation of the impact of mutations on the genome.

For further analyses of the sequences obtained in this study with the HRSV survey tool (gisaid.org), it will be possible to quickly identify the mutations found in our sequences compared to the sequences of references used by GISAID. However, we did not have enough time to use this tool, so it was not possible to analyse and interpret the results in time for the writing of this thesis. But from a similar tool from Nextclade we have observed a greater number of mutations in the G protein.

It is important to stress the need to increase globally the number of HRSV genomes to understand viral dynamics better and to better understand the meaning and consequence of mutations found in terms of virulence, escape from the immune response, escape to immunoprophylactics or antivirals, and in that way aiding vaccine development. Additionally, correlating this data with the phenotypic and antigenic characterization of these viruses is also important to support genomic findings.

Conclusions

On the basis of the samples analysed in this study, in the period 2019 to 2021, both subgroups of HRSV (A and B) circulated in Lothian. However, the circulation dominance has been different over the years.

In 2019 and 2020 the frequency of HRSV-A was higher in the samples studied, and in 2021, HRSV-B circulated more frequently, corroborating studies indicating the cyclic circulation of HRSV, changing at least every two years.

In total, we sequenced 90 HRSV genomes by NGS from Lothian, Scotland, ranging from 2019 to 2021 in this study, with 45 HRSV-B genomes and 45 HRSV-A genomes.

This preliminary testing of amplicon-based primer schemes for RSV A and RSV B has proved successful at harnessing existing SARS-CoV-2 sequencing infrastructure and has shown significant promise for allowing large-scale high-quality RSV genome production.

The sequenced samples of HRSV-A belonged to the GA2.3.5 genotype and all sequenced HRSV-B samples belonged to the genotype GB5.0.5a.

In this study, we did not list the present changes in amino acids throughout the genome for our sequenced samples, so further analyses will be necessary to understand the role of these mutations and their importance. However, a high number of mutations was observed in the G gene.

Work will continue to fully characterise our primer schemes, and to make them as efficient as possible. In particular, future work will increase the Ct cut off to find an upper limit for successful sequencing, the primer schemes will be tested on Illumina sequencing platforms, and efforts will be made to identify primers that can be spiked in to combat amplicon dropouts. In order to achieve this fully, a broader sample of RSV genomes needs to be sequenced, in particular as our current sample selection is geographically limited.

References

Agoti CN, Gitahi CW, Medley GF, Cane PA, Nokes DJ (2013). Identification of group B respiratory syncytial viruses that lack the 60-nucleotide duplication after six consecutive epidemics of total BA dominance at coastal Kenya. Influenza Other Respi. Viruses 7: 1008–1012.

Amarasinghe GK, Ayllón MA, Bào Y, Basler CF, Bavari S, Blasdell KR, Briese T, Brown PA, Bukreyev A, Balkema-Buschmann A, Buchholz UJ, Chabi-Jesus C, Chandran K, Chiapponi C, Crozier I, Swart RL de, Dietzgen RG, Dolnik O, Drexler JF, Dürrwald R, Dundon WG, Duprex WP, Dye JM, Easton AJ, Fooks AR, Formenty PBH, Fouchier RAM, Freitas-Astúa J, Griffiths A, Hewson R, Horie M, Hyndman TH, Jiāng D, Kitajima EW, Kobinger GP, Kondō H, Kurath G, Kuzmin I V., Lamb RA, Lavazza A, Lee B, Lelli D, Leroy EM, Lǐ J, Maes P, Marzano SYL, Moreno A, Mühlberger E, Netesov S V., Nowotny N, Nylund A, Økland AL, Palacios G, Pályi B, Pawęska JT, Payne SL, Prosperi A, Ramos-González PL, Rima BK, Rota P, Rubbenstroth D, Shī M, Simmonds P, Smither SJ, Sozzi E, Spann K, Stenglein MD, Stone DM, Takada A, Tesh RB, Tomonaga K, Tordo N, Towner JS, Hoogen B van den, Vasilakis N, Wahl V, Walker PJ, Wang LF, Whitfield AE, Williams J V., Zerbini FM, Zhāng T, Zhang YZ, Kuhn JH (2019). Taxonomy of the order Mononegavirales: update 2019. Arch. Virol. 164: 1967–1980.

Anderson LJ, Hierholzer JC, Tsou C, Michael Hendry R, Fernie BF, Stone Y, McIntosh K (1985). Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. J. Infect. Dis. 151: 626–633.

Anderson LJ, Parker RA, Strilms RL (1990). Association between respiratory syncytial virus outbreaks and lower respiratory tract deaths of infants and young children. J. Infect. Dis. 161: 640–646. Arbiza J, Delfraro A, Frabasile S (2005). Molecular epidemiology of human respiratory syncytial virus in Uruguay: 1985-2001 - A review. Mem. Inst. Oswaldo Cruz 100: 221–230.

Auksornkitti V, Kamprasert N, Thongkomplew S, Suwannakarn K, Theamboonlers A, Samransamruajkij R, Poovorawan Y (2014). Molecular characterization of human respiratory syncytial virus, 2010-2011: Identification of genotype ON1 and a new subgroup B genotype in Thailand. Arch. Virol. 159: 499–507.

Avila V, Calvo E, Barbosa J, Velandia-Romero ML, Castellanos JE (2018). Description of respiratory syncytial virus genotypes circulating in Colombia. J. Infect. Dev. Ctries. 12: 1112–1118.

Bakker SE, Duquerroy S, Galloux M, Loney C, Rey A, Bhella D, Conner E, Eléouet J-F (2013). The respiratory syncytial virus nucleoprotein – RNA complex forms a lefthanded helical nucleocapsid. J. Virol.: 1734–1738. Battles MB, Mclellan JS (2019). Respiratory syncytial virus entry and how to block it. Nat. Rev. Microbiol. 17.

Belshaw R, Pybus OG, Rambaut A (2007). The evolution of genome compression and genomic novelty in RNA viruses. Genome Res. 17: 1496–1504.

Bermingham A, Collins PL (1999). The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and 91 transcription. Proc. Natl. Acad. Sci. U. S. A. 96: 11259–11264.

Bitko V, Shulyayeva O, Mazumder B, Musiyenko A, Ramaswamy M, Look DC, Barik S (2007). Nonstructural Proteins of Respiratory Syncytial Virus Suppress Premature Apoptosis by an NF- B-Dependent, Interferon-Independent Mechanism and Facilitate Virus Growth J. Virol. 81: 1786–1795.

Borchers, AT, Chang, C., Gershwin, ME et al. (2013) Respiratory Syncytial Virus - A Comprehensive Review. Clinic Rev Allerg Immunol 45, 331-379. https://doi.org/10.1007/s12016-013-8368-9.

Broor S, Bharaj P, Chahar HS (2008). Human metapneumovirus: A new respiratory pathogen. J. Biosci. 33: 483–493.

Broor S, Campbell H, Hirve S, Hague S, Jackson S, Moen A, Nair H, Palekar R, Rajatonirina S, Smith PG, Venter M, Wairagkar N, Zambon M, Ziegler T, Zhang W (2020). Leveraging the Global Influenza Surveillance and Response System for global respiratory syncytial virus surveillance—opportunities and challenges. Influenza Other Respi. Viruses 14: 622–629.

Bukreyev A, Yang L, Fricke J, Cheng L, Ward JM, Murphy BR, Collins PL (2008). The Secreted Form of Respiratory Syncytial Virus G Glycoprotein Helps the Virus Evade Antibody-Mediated Restriction of Replication by Acting as an Antigen Decoy and through Effects on Fc Receptor-Bearing Leukocytes. J. Virol. 82: 12191–12204.

Cane PA, Matthews DA, Pringle CR (1992). Analysis of relatedness of subgroup A respiratory syncytial viruses isolated worldwide.

Chanock R, Roizman B, Myers R (1957). Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA): Isolation, properties and characterization. Am. J. Epidemiol. 66: 281–290.

Cheng X, Park H, Zhou H, Jin H (2005). Overexpression of the M2-2 Protein of Respiratory Syncytial Virus Inhibits Viral Replication. 79: 13943–13952.

Chidgey SM, Broadley KJ (2005). Respiratory syncytial virus infections: characteristics and treatment. J. Pharm. Pharmacol. 57: 1371–1381.

Chrzastek K, Tennakoon C, Bialy D, Freimanis G, Flannery J, Shelton H. A random priming amplification method for whole genome sequencing of SARS-CoV-2 virus. BMC Genomics. 2022 May 30;23(1):406. doi: 10.1186/s12864-022-08563-z.

Coates H V., Alling DW, Chanock RM (1966). An antigenic analysis of respiratory syncytial virus isolates by a plaque reduction neutralization test. Am. J. Epidemiol. 83: 299–313.

Collins PL, Graham BS (2008). Viral and Host Factors in Human Respiratory Syncytial Virus Pathogenesis. J. Virol. 82: 2040–2055.

Collins PL, Huang YT, Wertz GW (1984). Identification of a tenth mRNA of respiratory syncytial virus and assignment of polypeptides to the 10 viral genes. J. Virol. 49: 572–578.

Collins PL, Fearns R GB (2013). Respiratory syncytial virus: virology, reverse genetics, and pathogenesis of disease. Curr Top Microbiol Immunol 372: 3–38.

Cox NJ, Bender CA (1995). The molecular epidemiology of influenza viruses. Semin. Virol. 6: 359–370.

Cui, G., Zhu, R., Qian, Y., Deng, J., Zhao, L., Sun, Y., & Wang, F. (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(9), 92 e75020. https://doi.org/10.1371/journal.pone.0075020

Dapat IC, Shobugawa Y, Sano Y, Saito R, Sasaki A, Suzuki Y, Kumaki A, Zaraket H, Dapat C, Oguma T, Yamaguchi M, Suzuki H (2010). New genotypes within respiratory syncytial virus group B genotype BA in Niigata, Japan. J. Clin. Microbiol. 48: 3423–3427.

Escribano-Romero E, Rawling J, García-Barreno B, Melero JA (2004). The Soluble Form of Human Respiratory Syncytial Virus Attachment Protein Differs from the Membrane-Bound Form in Its Oligomeric State but Is Still Capable of Binding to Cell Surface Proteoglycans. J. Virol. 78: 3524–3532.

Eshaghi AR, Duvvuri VR, Lai R, Nadarajah JT, Li A, Patel SN, Low DE, Gubbay JB (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.

Essa S, Owayed A, Altawalah H, Khadadah M, Behbehani N, Al-Nakib W (2015). The prevalence of human bocavirus, human coronavirus-NL63, human metapneumovirus, human polyomavirus KI and WU in respiratory tract infections in Kuwait. Med. Princ. Pract. 24: 382–387.

Falsey AR, Hennessey PA, Formica MA, Cox C, Walsh EE (2005). Respiratory Syncytial Virus Infection in Elderly and High-Risk Adults.

Falsey AR, Mcelhaney JE, Beran J, Essen GA Van, Duval X, Esen M (2014). Respiratory Syncytial Virus and Other Respiratory Viral Infections in Older Adults With Moderate to Severe Influenza-like Illness. J. Infect. Dis. 209: 1873–1881.

Falsey AR, Walsh EE (2020). Respiratory syncytial virus: An old foe in a new era. J. Infect. Dis. 222: 1245–1246.

Fearns R, Peeples ME, Collins PL (2002). Mapping the Transcription and Replication Promoters of Respiratory Syncytial Virus. J. Virol. 76: 1663–1672.

Ferone EA, Berezin EN, Durigon GS, Finelli C, Felício MCC, Storni JG, Durigon EL, Oliveira DBLD (2014). Clinical and epidemiological aspects related to the detection of adenovirus or respiratory syncytial virus in infants hospitalized for acute lower respiratory tract infection. J. Pediatr. (Rio. J). 90: 42–49.

Frabasile S, Delfraro A, Facal L, Videla C, Galiano M, Sierra MJ De, Ruchansky D, Vitureira N, Berois M, Carballal G, Russi J, Arbiza J (2003). Antigenic and genetic variability of human respiratory syncytial viruses (group A) isolated in Uruguay and Argentina: 1993-2001. J. Med. Virol. 71: 305–312.

Fuentes S, Tran KC, Luthra P, Teng MN, He B (2007). Function of the Respiratory Syncytial Virus Small Hydrophobic Protein. J. Virol. 81: 8361–8366.

García O, Martín M, Dopazo J, Arbiza J, Frabasile S, Russi J, Hortal M, Perez-Breña P, Martínez I, García-Barreno B (1994). Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. J. Virol. 68: 5448–5459.

Ghazaly M, Nadel S (2018). Characteristics of children admitted to intensive care with acute bronchiolitis. Eur. J. Pediatr. 177: 913–920.

Ghildyal R, Mills J, Murray M, Vardaxis N, Meanger J (2002). Respiratory syncytial virus matrix protein associates with nucleocapsids in infected cells. : 753–757.

Giallonardo F Di, Kok J, Fernandez M, Carter I, Geoghegan JL, Dwyer DE, Holmes EC, Eden JS (2018). Evolution of human respiratory syncytial virus (RSV) over multiple seasons in New South Wales, Australia. Viruses 10: 1–13.

Gimferrer L, Andrés C, Campins M, Codina MG, Rodrigo JA, Melendo S, Martin MC, Fuentes F, Saiz MR, Esperalba J, Bruguera A, Vilca LM, Armadans L, Pumarola T, Antón A (2016). Circulation of a novel human respiratory syncytial virus Group B genotype during the 2014-2015 season in Catalonia (Spain). Clin. Microbiol. Infect. 22: 97.e5-97.e8.

Glezen PW, Taber LH, Frank AL, Kasel JA (1986). Risk of Primary Infection and Reinfection With Respiratory Syncytial Virus. Am J Dis Child 140: 543–546.

Gould PS, Easton AJ (2007). Coupled Translation of the Second Open Reading Frame of M2 mRNA Is Sequence Dependent and Differs Significantly within the Subfamily Pneumovirinae. 81: 8488–8496.

Goya S, Galiano M, Nauwelaers I, Trento A, Openshaw PJ, Mistchenko AS, Zambon M, Viegas M (2020). Toward unified molecular surveillance of RSV: A proposal for genotype definition. Influenza Other Respi. Viruses 14: 274–285.

Graham BS, Modjarrad K, Mclellan JS, Diseases I, Spring S (2015). Novel Antigens for RSV Vaccines. Curr. Opin. Immunol. 35: 30–38.

Griffin MP, Yuan Y, Therese Takas BS, Domachowske JB, Madhi SA, Manzoni P (2020). Single-Dose Nirsevimab for Prevention of RSV in Preterm Infants. new Engl. J. of Med. 383: 415–425.

Griffiths C, Drews SJ, Marchant DJ (2017). Respiratory syncytial virus: Infection, detection, and new options for prevention and treatment. Clin. Microbiol. Rev. 30: 277–319.

Haider MSH, Khan WH, Deeba F, Ali S, Ahmed A, Naqvi IH, Dohare R, Alsenaidy HA, Alsenaidy AM, Broor S, Parveen S (2018). BA9 lineage of respiratory syncytial virus from across the globe and its evolutionary dynamics. PLoS One 13: 1–22.

Hall CB, Walsh EE, Long CE, Schnabel KC (1991). Immunity to and frequency of reinfection with respiratory syncytial virus. J. Infect. Dis. 163: 693–698.

Hall CB, Walsh EE, Schnabel KC, Long CE, McConnochie KM, Hildreth SW, Anderson LJ (1990). Occurrence of groups A and B of respiratory syncytial virus over

15 years: Associated epidemiologic and clinical characteristics in hospitalized and ambulatory children. J. Infect. Dis. 162: 1283–1290.

Haynes AK, Prill MM, Iwane MK, Gerber SI (2014). Respiratory Syncytial Virus — United States, July 2012 – June 2014. 63: 2012–2015.

Hendricks DA, Mclntosh K, Patterson JL (1988). Further characterization of the soluble form of the G glycoprotein of respiratory syncytial virus. Virus Res. 11: 15. 94

Hervás D, Reina J, Hervás JA (2012). Meteorologic conditions and respiratory syncytial virus activity. Pediatr. Infect. Dis. J. 31: 176–181.

Hibino A, Saito R, Taniguchi K, Zaraket H, Shobugawa Y, Matsui T, Suzuki H, Hrsv J (2018). Molecular epidemiology of human respiratory syncytial virus among children in Japan during three seasons and hospitalization risk of genotype ON1.: 1–19.

Hirano E, Kobayashi M, Tsukagoshi H, Yoshida LM, Kuroda M, Noda M, Ishioka T, Kozawa K, Ishii H, Yoshida A, Oishi K, Ryo A, Kimura H (2014). Molecular evolution of human respiratory syncytial virus attachment glycoprotein (G) gene of new genotype ON1 and ancestor NA1. Infect. Genet. Evol. 28: 183–191.

Hirve S, Crawford N, Palekar R, Zhang W (2020). Clinical characteristics, predictors, and performance of case definition — Interim results from the WHO global respiratory syncytial virus surveillance pilot. Influ. Other Respi Viruses.: 647–657.

Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, Vandepol S (1982). Mutation Rates of RNA Viruses Rapid Evolution of RNA Genomes.

Huang YT, Wertz GW (1982). The genome of respiratory syncytial virus is a negativestranded RNA that codes for at least seven mRNA species. J. Virol. 43: 150–157.

IMpact-RSV Study Group (1998) Palivizumab, a Humanized Respiratory Syncytial Virus Monoclonal Antibody, Reduces Hospitalization From Respiratory Syncytial Virus Infection in High-risk Infants. Pediatrics 102: 531-537. Doi: 10.1542/peds.102.3.531

Jafri, Hasan S; MD Wu, Xionghua; Makari, Doris MD; Henrickson, Kelly J. MD (2013 Distribution of Respiratory syncytial virus subtypes A and B among infants presenting to the emergency department with lower respiratory tract infection or apnea The Pediatric Infectious Disease Journal: abril de 2013 – Vol 32 - Edição 4 - p 335-340 doi: 10.1097 / INF.0b013e318282603a

Jartti, T., Lehtinen, P., Vuorinen, T., & Ruuskanen, O. (2009). Bronchiolitis: age and previous wheezing episodes are linked to viral etiology and atopic characteristics. The Pediatric infectious disease journal, 28(4), 311–317. https://doi.org/10.1097/INF.0b013e31818ee0c1

Johnson JE, Gonzales RA, Olson SJ, Wright PF, Graham BS (2007). The histopathology of fatal untreated human respiratory syncytial virus infection. Mod. Pathol. 20: 108– 119.

Johnson PR, Olmsted RA, Prince GA, Murphy BR, Alling DW, Walsh EE, Collins PL (1987). Antigenic relatedness between glycoproteins of human respiratory syncytial

virus subgroups A and B: evaluation of the contributions of F and G glycoproteins to immunity. J. Virol. 61: 3163–3166.

Johnson S, Oliver C, Gregory A. Prince, Val G. Hemming, David S. Pfarr, Sheau-Chiann Wang, Melissa Dormitzer, John O'Grady, Scott Koenig, James K. Tamura, Robert Woods, Geetha Bansal, Debra Couchenour , Eric Tsao, William C. Hall, James F. Young, (1997). Development of a Humanized Monoclonal Antibody (MEDI-493) with Potent In Vitro and In Vivo Activity against Respiratory Syncytial Virus. The Journal of Infectious Diseases, Volume 176, Issue 5, Novembro de 1997, Páginas 1215–1224, https://doi.org/10.1086/514115

Kabir MS (2018). Molecular methods for detection of pathogenic viruses of respiratory tract - A review. Asian Pac. J. Trop. Biomed. 8: 237–244.

Kahn JS, Schnell MJ, Buonocore L, Rose JK (1999). Recombinant Vesicular Stomatitis Virus Expressing Respiratory Syncytial Virus (RSV) Glycoproteins: RSV Fusion Protein Can Mediate Infection and Cell Fusion. Virology 254: 81–91.

Katoh K, Standley DM (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30: 772–780.

Ke Z, Dillard RS, Chirkova T, Leon F, Stobart CC, Hampton CM, Strauss JD, Rajan D, Rostad CA, Taylor J V., Yi H, Shah R, Jin M, Hartert T V., Peebles RS, Graham BS, 95 Moore ML, Anderson LJ, Wright ER (2018). The morphology and assembly of respiratory syncytial virus revealed by cryo-electron tomography. Viruses 10.

Kenmoe S, Vernet MA, Miszczak F, Dina J, Schoenhals M, Beng VP, Vabret A, Njouom R (2018). Genetic diversity of human respiratory syncytial virus isolated among children with acute respiratory infections in Southern Cameroon during three consecutive epidemic seasons, 2011-2013. Trop. Med. Health 46.

Kiss G, Holl JM, Williams GM, Alonas E, Vanover D, Lifland AW, Gudheti M, GuerreroFerreira RC, Nair V, Yi H, Graham BS, Santangelo PJ, Wright ER (2014). Structural Analysis of Respiratory Syncytial Virus Reveals the Position of M2-1 between the Matrix Protein and the Ribonucleoprotein Complex. J. Virol. 88: 7602–7617.

Krivitskaya V, Komissarova K, Pisareva M, Sverlova M, Fadeev A, Petrova E, Timonina V, Sominina A, Danilenko D (2021). Respiratory syncytial virus g protein sequence variability among isolates from st. Petersburg, russia, during the 2013–2014 epidemic season. Viruses 13.

Krusat T, Streckert HJ (1997). Heparin-dependent attachment of respiratory syncytial virus (RSV) to host cells. Arch. Virol. 142: 1247–1254.

Krzyzaniak MA, Zumstein MT, Gerez JA, Picotti P, Helenius A (2013). Host Cell Entry of Respiratory Syncytial Virus Involves Macropinocytosis Followed by Proteolytic Activation of the F Protein. 9.

Kurt-Jones EA, Popova L, Kwinn L, Haynes LM, Jones LP, Tripp RA, Walsh EE, Freeman MW, Golenbock DT, Anderson LJ, Finberg RW (2000). Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. Nat. Immunol. 1: 398–401.

Laham FR, Mansbach JM, Piedra PA, Hasegawa K, Sullivan AF, Espinola JA, Camargo CA (2017). Clinical Profiles of Respiratory Syncytial Virus Subtypes A and B among Children Hospitalized with Bronchiolitis. Pediatr. Infect. Dis. J. 36: 808–810.

Law BJL, Carbonell-Estrany X, Simões EAFS (2002). An update on respiratory syncytial virus epidemiology: a developed country perspective. Respir. Med. 96: 1–7.

Levine S, Kaliaber-Franco R, Paradiso PR (1987). Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. J. Gen. Virol. 68: 2521–2524.

Liljeroos L, Krzyzaniak MA, Helenius A, Butcher SJ (2013). Architecture of respiratory syncytial virus revealed by electron cryotomography. Proc. Natl. Acad. Sci. U. S. A. 110: 11133–11138.

Liu W, Chen D, Tan W, Xu D, Qiu S, Zeng Z, Li X, Zhou R (2016). Epidemiology and clinical presentations of respiratory syncytial virus subgroups A and B detected with multiplex real-time PCR. PLoS One 11: 1–13.

Lu B, Ma C, Brazas R, Jin H (2002). The Major Phosphorylation Sites of the Respiratory Syncytial Virus Phosphoprotein Are Dispensable for Virus Replication In Vitro. J. Virol. 76: 10776–10784.

Madhi SA, Polack FP, Piedra PA, Munoz FM, Trenholme AA, Simões EAF, Swamy GK, Agrawal S, Ahmed K, August A (2020). Respiratory Syncytial Virus Vaccination during Pregnancy and Effects in Infants. new Engl. J. of Med. 383: 426–439. 96

Martinelli M, Frati ER, Zappa A, Ebranati E, Bianchi S, Pariani E, Amendola A, Zehender G, Tanzi E (2014). Phylogeny and population dynamics of respiratory syncytial virus (Rsv) A and B. Virus Res. 189: 293–302.

Martinello RA, Chen MD, Weibel C, Kahn JS (2002). Correlation between Respiratory Syncytial Virus Genotype and Severity of Illness. J. Infect. Dis. 186: 839–842.

Martínez I, Valdés O, Delfraro A, Arbiza J, Russi J, Melero JA (1999). Evolutionary pattern of the G glycoprotein of human respiratory syncytial viruses from antigenic group B: The use of alternative termination codons and lineage diversification. J. Gen. Virol. 80: 125–130.

Marty A, Meanger J, Mills J, Shields B, Ghildyal R (2004). Association of matrix protein of respiratory syncytial virus with the host cell membrane of infected cells Brief Report. : 199–210.

Matheson JW, Rich FJ, Cohet C, Grimwood K, Huang QS, Penny D, Hendy MD KJ (2006). Distinct patterns of evolution between respiratory syncytial virus subgroups A and B from New Zealand isolates collected over thirty-seven years. J. Med. Virol. 78: 1354–1364.

Maurice Mufson BA, Orvell C, Rafnar B, Norrby E (1985). Two Distinct Subtypes of Human Respiratory Syncytial Virus.

Meissner HC (2016). Viral Bronchiolitis in Children (JR Ingelfinger, Ed.). N. Engl. J. Med. 374: 62–72.

Melero JA, Mas V, McLellan JS (2017). Structural, antigenic and immunogenic features of respiratory syncytial virus glycoproteins relevant for vaccine development. Vaccine 35: 461–468.

Melero JA, Moore ML (2013). Influence of respiratory syncytial virus strain differences on pathogenesis and immunity. Curr. Top. Microbiol. Immunol. 372: 59–82.

Mitchell I, Defoy I, Grubb E (2017). Burden of Respiratory Syncytial Virus Hospitalizations in Canada. 2017.

Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA (2008). Database indexing for production MegaBLAST searches. Bioinformatics 24: 1757–1764.

Morris JA, Blount RE, Savage RE (1956). Recovery of Cytopathogenic Agent from Chimpanzees with Goryza. Proc. Soc. Exp. Biol. Med. 92: 544–549.

Muñoz-Escalante JC, Comas-García A, Bernal-Silva S, Noyola DE (2021). Respiratory syncytial virus B sequence analysis reveals a novel early genotype. Sci. Rep. 11.

Muñoz-Escalante JC, Comas-García A, Bernal-Silva S, Robles-Espinoza CD, GómezLeal G, Noyola DE (2019). Respiratory syncytial virus A genotype classification based on systematic intergenotypic and intragenotypic sequence analysis. Sci. Rep. 9: 1–12.

Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274 97

Norrby E, Marusyk H, Orvell C (1970). Morphogenesis of respiratory syncytial virus in a green monkey kidney cell line (Vero). J. Virol. 6: 237–242.

Noton SL, Tremaglio CZ, Fearns R (2019). Killing two birds with one stone: How the respiratory syncytial virus polymerase initiates transcription and replication. PLoS Pathog. 15: 1–8.

Obando-pacheco P, Justicia-grande AJ, Rivero-calle I, Rodríguez-tenreiro C, Sly P, Ramilo O (2018). Respiratory Syncytial Virus Seasonality: A Global Overview. J. Infect. Dis. 217: 1356–1364.

Ohuma EO, Okiro EA, Ochola R, Sande CJ, Cane PA, Medley GF, Bottomley C (2012). The Natural History of Respiratory Syncytial Virus in a Birth Cohort: The Influence of Age and Previous Infection on Reinfection and Disease. Am. J. Epidemiol. 176: 794–802.

Otieno JR, Kamau EM, Oketch JW, Ngoi JM, Gichuki AM, Binter S, Otieno GP, Ngama M, Agoti CN, Cane PA, Kellam P, Cotten M, Lemey P, Nokes DJ (2018). Whole genome analysis of local Kenyan and global sequences unravels the epidemiological and molecular evolutionary dynamics of RSV genotype ON1 strains. Virus Evol. 4: 1–13.

Palomo C, Garcia-Barreno B, Penas C, Melero JA (1991). The G protein of human respiratory syncytial virus: Significance of carbohydrate side-chains and the Cterminal end to its antigenicity. J. Gen. Virol. 72: 669–675.

Papadopoulos NG, Gourgiotis D, Javadyan A, Bossios A, Kallergi K, Psarras S, Tsolia MN, Kafetzis D (2004). Does respiratory syncytial virus subtype influences the severity of acute bronchiolitis in hospitalized infants? Respir. Med. 98: 879–882.

Peret TCT, Hall CB, Hammond GW, Piedra PA, Storch GA, Sullender WM, Tsou C, Anderson LJ (2000). Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. J. Infect. Dis.: 1891– 1896.

Peret TCT, Hall CB, Schnabel KC, Golub JA, Anderson LJ (1998). Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. J. Gen. Virol. 79: 2221–2229.

Pomeroy LW, Bjørnstad ON, Holmes EC (2008). The evolutionary and epidemiological dynamics of the paramyxoviridae. J. Mol. Evol. 66: 98–106.

Quick, J., Grubaugh, N. D., Pullan, S. T., Claro, I. M., Smith, A. D., Gangavarapu, K., Oliveira, G., Robles-Sikisaka, R., Rogers, T. F., Beutler, N. A., Burton, D. R., Lewis-Ximenez, L. L., de Jesus, J. G., Giovanetti, M., Hill, S. C., Black, A., Bedford, T., Carroll, M. W., Nunes, M., Alcantara, L. C., Jr, Loman, N. J. (2017). Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nature protocols, 12(6), 1261–1276.

Ralston SL, Lieberthal AS, Meissner HC, Alverson BK, Baley JE, Gadomski AM, Johnson DW, Light MJ, Maraqa NF, Mendonca EA, Phelan KJ, Zorc JJ, StankoLopp D, Brown MA, Nathanson I, Rosenblum E, Sayles S, Hernandez-Cancio S (2014). Clinical Practice Guideline: The Diagnosis, Management, and Prevention of Bronchiolitis abstract. Pediatrics 134: 1474–1502.

Ramaekers K, Rector A, Cuypers L, Lemey P, Keyaerts E, Ranst M van (2020). Towards a unified classification for human respiratory syncytial virus genotypes. Virus Evol. 6.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. (2018). Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. https://doi.org/10.1093/sysbio/syy032 98

Reeves RM, Wijhe M Van, Tong S, Lehtonen T, Stona L, Teirlinck AC, Fernandez V, Li Y (2020). Respiratory Syncytial Virus-Associated Hospital Admissions in Children Younger Than 5 Years in 7 European Countries Using Routinely Collected Datasets. PLoS One 13.

Rima B, Collins P, Easton A, Fouchier R, Kurath G, Lamb RA, Lee B, Maisner A, Rota P, Siddell SG, Simmonds P (2017). ICTV Virus Taxonomy Profile: Pneumoviridae. 98: 10–14.

Roberts SR, Lichtenstein D, Ball LA, Wertz GW (1994). The membrane-associated and secreted forms of the respiratory syncytial virus attachment glycoprotein G are synthesized from alternative initiation codons. J. Virol. 68: 4538–4546.

Robertson M, Eden J, Levy A, Carter I, Tulloch RL (2021). The spatial-temporal dynamics of respiratory syncytial virus 2 infections across the east-west coasts of Australia during 2016-17. : 1–22.

Rojo GL, Goya S, Orellana M, Sancilio A, Rodriguez Perez A, Montali C, García C, Sosa L, Musto A, Alvarez D, Castello A, Viegas M (2017). Unravelling respiratory syncytial virus outbreaks in Buenos Aires, Argentina: Molecular basis of the spatiotemporal transmission. Virology 508: 118–126.

Rose EB, Wheatley A, Langley G, Gerber S, Haynes A (2018). Respiratory Syncytial Virus Seasonality — United States, 2014 – 2017. 67: 2014–2017.

Sáez-López E, Cristóvão P, Costa I, Pechirra P, Conde P, Guiomar R (2020). Epidemiology and genetic variability of respiratory syncytial virus in Portugal, 2014-2018. J. Clin. Virol. 121.

Salimi V, Viegas M, Trento A, Agoti CN, Anderson LJ, Avadhanula V, Bahl J, Bont L, Brister JR, Cane PA, Galiano M, Graham BS, Hatcher EL, Hellferscee O, Henke DM, Hirve S, Jackson S, Keyaerts E, Kragten-tabatabaie L, Lindstrom S, Nauwelaers I, Nokes DJ, Openshaw PJ, Peret TC, Piedra PA, Ramaekers K, Rector A, Trovão NS, Gottberg A Von, Zambon M, Zhang W, Williams TC, Barr IG, Buchholz UJ (2021). Proposal for Human Respiratory Syncytial Virus Nomenclature below the Species Level. 27: 1–9.

Satake M, Venkatesan S (1984). Nucleotide sequence of the gene encoding respiratory syncytial virus matrix protein. J. Virol. 50: 92–99.

Schobel SA, Stucker KM, Moore ML, Anderson LJ, Larkin EK, Shankar J, Bera J, Puri V, Shilts MH, Rosas-Salazar C, Halpin RA, Fedorova N, Shrivastava S, Stockwell TB, Peebles RS, Hartert T V., Das SR (2016). Respiratory Syncytial Virus wholegenome sequencing identifies convergent evolution of sequence duplication in the C-terminus of the G gene. Sci. Rep. 6: 1–11.

Sedeyn K, Schepens B SX (2019). Respiratory syncytial virus nonstructural proteins 1 and 2 : Exceptional disrupters of innate immune responses. PLoS Pathog. 15: 1–18.

Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, Polack FP, Balsells E, Acacio S, Aguayo C, Alassani I, Ali A, Antonio M, Awasthi S, Awori JO, Azziz-Baumgartner E, Baggett HC, Baillie VL, Balmaseda A, Barahona A, Basnet 99 S, Bassat Q, Basualdo W, Bigogo G, Bont L, Breiman RF, Brooks WA, Broor S, Bruce N, Bruden D, Buchy P, Campbell S, Carosone-Link P, Chadha M, Chipeta J, Chou M, Clara W, Cohen C, Cuellar E de, Dang DA, Dash-yandag B, Deloria-Knoll M, Dherani M, Eap T, Ebruke BE, Echavarria M, Freitas Lázaro Emediato CC de, Fasce RA, Feikin DR, Feng L, Gentile A, Gordon A, Goswami D, Goyet S, Groome M, Halasa N, Hirve S, Homaira N, Howie SRC, Jara J, Jroundi I, Kartasasmita CB, Khuri-Bulos N, Kotloff KL, Krishnan A, Libster R, Lopez O, Lucero MG, Lucion F, Lupisan SP, Marcone DN, McCracken JP, Mejia M, Moisi JC, Montgomery JM, Moore DP, Moraleda C, Moyes J, Munywoki P, Mutyara K, Nicol MP, Nokes DJ, Nymadawa P, Costa Oliveira MT da, Oshitani H, Pandey N, Paranhos-Baccalà G, Phillips LN, Picot VS, Rahman M, Rakoto-Andrianarivelo M, Rasmussen ZA, Rath BA, Robinson A, Romero C, Russomando G, Salimi V, Sawatwong P, Scheltema N, Schweiger B,

Scott JAG, Seidenberg P, Shen K, Singleton R, Sotomayor V, Strand TA, Sutanto A, Sylla M, Tapia MD, Thamthitiwat S, Thomas ED, Tokarz R, Turner C, Venter M, Waicharoen S, Wang J, Watthanaworawit W, Yoshida LM, Yu H, Zar HJ, Campbell H, Nair H (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.

Shobugawa Y, Saito R, Sano Y, Zaraket H, Suzuki Y, Kumaki A, Dapat I, Oguma T, Yamaguchi M, Suzuki H (2009). Emerging genotypes of human respiratory syncytial virus subgroup A among patients in Japan. J. Clin. Microbiol. 47: 2475–2482.

Smith DK, Seales S, Budzik C, Jacksonville NH (2017). Respiratory Syncytial Virus Bronchiolitis in Children. Am J Dis Child 95: 94–99.

Spann KM, Tran KC, Collins PL, Irol J V (2005). Effects of Nonstructural Proteins NS1 and NS2 of Human Respiratory Syncytial Virus on Interferon Regulatory Factor 3, NF-B, and Proinflammatory Cytokines. J. Virol. 79: 5353–5362.

Strimmer K, Haeseler A von (1997). Likelihood-mapping: A simple method to visualize phylogenetic. Proc Natl Acad Sci U S A 94: 6815–6819.

Suchard, M. A. et al. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 4, vey016

Sullender WM, Sun L, Anderson LJ (1993). Analysis of respiratory syncytial virus genetic variability with amplified cDNAs. J. Clin. Microbiol. 31: 1224–1231.

Suryadevara M, Domachowske JB (2021). Epidemiology and seasonality of childhood respiratory syncytial virus infections in the tropics. Viruses 13.

Taleb SA, Thani AA Al, Ansari K Al, Yassine HM (2018). Human respiratory syncytial virus: pathogenesis, immune responses, and current vaccine approaches. Eur. J. Clin. Microbiol. Infect. Dis. 37: 1817–1827. 100

Templeton K, Scheltinga, SA, Beersma, MF, Kroes AC, Claas EC. (2004). Rapid and sensitive method using multiplex real-time PCR for diagnosis of infections by influenza a and influenza B viruses, respiratory syncytial virus, and parainfluenza viruses 1, 2, 3, and 4. Journal of clinical microbiology, 42(4), 1564–1569. https://doi.org/10.1128/JCM.42.4.1564-1569.2004

Trento A, Viegas M, Galiano M, Videla C, Carballal G, Mistchenko AS, Melero J. (2006). The rich-poor divide on America's college campuses is. J. od Virol. 80: 975–984.

Trento A, Galiano M, Videla C, Carballal G, García-Barreno B, Melero JA, Palomo C (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.

Trento A, Ábrego L, Rodriguez-Fernandez R, González-Sánchez MI, GonzálezMartínez F, Delfraro A, Pascale JM, Arbiza J, Melero JA. (2015). Conservation of G-Protein Epitopes in Respiratory Syncytial Virus (Group A) Despite Broad Genetic Diversity: Is Antibody Selection Involved in Virus Evolution?. Journal of virology, 89(15), 7776–7785. https://doi.org/10.1128/JVI.00467-15

Tseng HF, Sy LS, Ackerson B, Solano Z, Slezak J, Luo Y, Fischetti CA, Shinde V (2020). Severe morbidity and short- And mid- And long-term mortality in older adults hospitalized with respiratory syncytial virus infection. J. Infect. Dis. 222: 1298–1310.

Umuhoza T, Id WDB, Oyugi J, Musabyimana JP, Kinengyere AA, Mancuso JD (2021). Prevalence of human respiratory syncytial virus, parainfluenza and adenoviruses in East Africa Community partner states of Kenya, Tanzania, and Uganda: A systematic review and meta-analysis (2007 – 2020). PLoS One 16: 6–21.

Venter M, Madhi SA, Tiemessen CT, Schoub BD (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: 2117– 2124.

Vianna LA, Siqueira MM, Volpini L, Louro I, Resende PC (2021). Seasonality, molecular epidemiology, and virulence of Respiratory Syncytial Virus (RSV): A perspective into the Brazilian Influenza Surveillance Program. PLoS One 16: 1–20.

Villenave R, Thavagnanam S, Sarlang S, Parker J, Douglas I, Skibinski G, Heaney LG, McKaigue JP, Coyle P V., Shields MD, Power UF (2012). In vitro modeling of respiratory syncytial virus infection of pediatric bronchial epithelium, the primary target of infection in vivo. Proc. Natl. Acad. Sci. U. S. A. 109: 5040–5045.

Walsh EE, Hruska J (1983). Monoclonal Antibodies to Respiratory Syncytial Virus Proteins: Identification of the Fusion Protein. 47: 171–177.

Waris M (1991). Pattern of respiratory syncytial virus epidemics in Finland: Two-year cycles with alternating prevalence of groups A and B. J. Infect. Dis. 163: 464–469.

Weekly national respiratory report (publichealthscotland.scot)

Wertz GW, Collins PL, Huang Y, Gruber C, Levine S, Andrew Ball1' L (1985). Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein (mRNA sequence/glycoprotein structure/paramyxovirus/O-linked oligosaccharides).

White LJ, Waris M, Cane PA, Nokes DJ, Medley GF (2005). The transmission dynamics of groups A and B human respiratory syncytial virus (HRSV) in England & Wales and Finland: Seasonality and cross-protection. Epidemiol. Infect. 133: 279–289.

WHO (2017). Information for the molecular detection of Influenza viruses. Who: 1–60. 101 World Health Organization - WHO (2014). Infection prevention and control of epidemicand pandemic-prone acute respiratory infections in health care. WHO Guidel.: 1– 156.

World Health Organization (2020a). Novel coronavirus (2019-nCoV): situation reports -20.https://www.who.int/docs/defaultsource/coronaviruse/situationreports/20200209-sitrep-20-ncov.pdf?sfvrsn=6f80d1b9_4.

World Health Organization (2020b). WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. https://www.who.int/director-general/speeches/detail/who-director-general-sopening-remarks-at-the-media-briefing-on-covid-19---11-march-2020.

Wu F, Zhao S, Bin Y, Chen Y-M, Wang W, Zhi-Gang S (2020). A new coronavirus associated with human respiratory disease in China. Nature 579: 265–269.

Yoshihara K, Le MN, Okamoto M, Wadagni ACA, Nguyen HA, Toizumi M, Pham E, Suzuki M, Nguyen ATT, Oshitani H, Ariyoshi K, Moriuchi H, Hashizume M, Dang DA, Yoshida LM (2016). Association of RSV-A ON1 genotype with Increased Pediatric Acute Lower Respiratory Tract Infection in Vietnam. Sci. Rep. 6: 1–10.

Zhang Z, Schwartz S, Wagner L, Miller W (2000). A greedy algorithm for aligning DNA sequences. J. Comput. Biol. 7: 203–214.

Zhou P, Yang X, Wang X, Hu B, Zhang L, Zhang W, Guo H, Jiang R, Liu M, Chen Y, Shen X, Wang X, Zhan F, Wang Y, Xiao G, Shi Z (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579.