A Molecular Epidemiological Study of Human Parainfluenza 4 in the Western Cape, South Africa.

Jane Parsons

PRSHIL002

This thesis is submitted in fulfilment of the requirements for the degree of Master of Virology in the Division of Medical Virology, Department of Pathology in the Faculty of Health Sciences at the University of Cape Town.

Supervisor: Professor Diana Hardie

Co-Supervisor: Dr Heidi Smuts

January 2021

Dedication

For my late husband Christopher Parsons, who is my guardian angel. He was an example of courage, commitment, endurance and hard work. I am a better person for knowing him.

ABSTRACT

Background

Human parainfluenza 4 (HPIV 4) is a recognised cause of acute respiratory infection (ARI). However, there is no published data on the epidemiology of this virus in South Africa. This thesis describes the molecular epidemiology of HPIV 4 over a 4-year period (2014-2017). Respiratory samples from infants, children and adults presenting with respiratory illness in the Western Cape, South Africa were studied.

Method

A retrospective 4-year study using routine diagnostic samples from patients with ARI was conducted in Western Cape, South Africa. A database search of positive HPIV 4 samples detected by the Seegene Anyplex RV 16 diagnostic assay was extracted. Epidemiological information was recorded to determine age, gender, hospital ward (used as a proxy for disease severity), specimen type (upper or lower respiratory tract) and collection date (to indicate seasonality).

To determine genetic evolution, novel primers targeting the haemagglutinin-neuraminidase (HN) in both HPIV 4 subtypes were designed to amplify a 733 bp and 738 bp sequence for HPIV 4A and HPIV 4B respectively. This product was then sequenced and aligned with known reference sequences from GenBank, using BioEdit. These aligned sequences were analysed using the phylogenetic analysis tool, MEGA 6, and Highlighter plots to determine sequence divergence events and evolution.

A real-time PCR assay, targeting the phosphoprotein, was developed to rapidly distinguish subtype A and B viruses.

Results

HPIVs were the 6th most common respiratory viruses detected in diagnostic samples. In all, there were 312/7456 (4.2 %) HPIV 4 positive samples in patients with a median age of 12 months. Males had a higher infection rate. HPIV 4 was the most prevalent of the HPIVs accounting for 47% of all HPIVs. Respiratory infections due to HPIV 4 were seasonal, peaking in autumn and mid-winter (March to August). The overall prevalence of HPIV 4 increased over the study period.

Of the HPIV 4-positive samples that were subtyped, 59 were subtype A and 26 subtype B. Both subtypes co-circulated during each season.

71 % of patients who were positive for HPIV 4 were co-infected with one or more additional respiratory virus with Adenovirus (27 %), Human Rhinovirus (23 %) and Bocavirus (19 %) as the most common. HPIV 1 and HPIV 3 were both able to co-infect patients with HPIV 4, but no co-infections with HPIV 2 were detected.

Phylogenetic trees constructed using neighbour joining (NJ) method showed that most of the South African HPIV 4 subtypes did not group with the closest significant reference sequences from GenBank. The phylogenetic tree for HPIV 4A revealed 4 genetic groupings. There were many nucleotide changes increasing with time as well as a non-synonymous change in HPIV 4A, at location N161D. HPIV 4B had an amino acid change in location G198R in the HN protein sequenced.

Conclusion

HPIV 4 with an overall prevalence of 4 % over the study period was identified as a significant cause of ARI in the Western Cape, South Africa. Mono-infection with HPIV4 was associated with severe disease. In hospitalized infants who were HPIV 4 positive, between ¼ to 1/3 were from patients in ICU. Of these almost half (46 %) had HPIV 4 as a single infection. Further studies are needed to fully understand the molecular epidemiology of this infection.

Abbreviations

Adeno	adenovirus
ARI	Acute respiratory infection
ARIs	Acute respiratory infections
amps	ampere
AMV	Avian myeloblastosis virus
Boc	bocavirus
bp	base pair
BHQ1	Black Hole Quencher 1
BPIV 3	Bovine parainfluenza type 3
C	Capsid protein
CAP	Community acquired pneumonia
Cor	coronavirus
CPE	Cytopathic effect
Ct	Cycle threshold
DNA	Deoxyribose nucleic acid
dNTPs	deoxynucleotides
ddNTPs	dideoxynucleotides
EPI	Epidemiological
EQA	External quality assessment
ESCRT	Endosomal sorting complex required for transport
EV	enterovirus
F	Fusion protein
FAM	Fluorescein amidite
Flu	influenza A and B
FRET	Fluorescence resonance energy transfer
GBD	Global burden of disease
GC content	Guanine/ cytosine content
GE	Gel electrophoresis
HA	Haemagglutinin
HEX	Hexachloro-fluorescein
HR-C	C-terminal heptad repeat region
HR-N	N-terminal heptad repeat region
HN	Haemagglutinin-neuraminidase

HN gene	Haemagglutinin-neuraminidase gene
HPIV 4	human parainfluenza 4
HPIV 4A	human parainfluenza 4A
HPIV 4B	human parainfluenza 4B
HPIVs	human parainfluenzas
HPL	Highlighter plot
ICU	Intensive care unit
ICTV	International Committee on Taxonomy of Viruses
IDT	Integrated DNA Technologies
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IFN	Interferon
kb	kilobase
L	Large protein
LOD	Limit of detection
LRT	Lower respiratory tract
LRTI	Lower respiratory tract infection
M	Matrix protein
MEGA	Molecular Evolutionary Genetic Analysis
Meta	metapneumovirus
MMVL	Moloney murine leukaemia virus
mRNA	messenger RNA
MSc	Masters
N	Nucleoprotein
NA	Neuraminidase
NDV	Newcastle disease virus
NEG	Negative
NJ	Neighbor Joining
NHLS	National Health Laboratory Services
N-P-L	Nucleoprotein- Phosphoprotein-Large protein
ORF	Open reading frame
PMV	*Paramyxoviridae*
P-L	Phosphoprotein-Large protein
P/V	Phosphoprotein
PCR	Polymerase chain reaction

PI	Principle investigator
POS	Positive
QCMD	Quality Control for Molecular Diagnostics
qPCR	quantitative Polymerase chain reaction
Rhino	human rhinovirus
RNA	Ribonucleic acid
RSV	respiratory syncytial virus
RT-PCR	Reverse transcription polymerase chain reaction
R-T PCR	Real-time PCR
RTIs	Respiratory tract infections
SeV	Sendai virus
SV 5	simian virus 5
SV 41	simian virus 41
TAE	Tris-Acetate- EDTA
*T*m	melting temperature
V	Virial envelope protein
ViPR	Virus Pathogen Data base Analysis Resource
WHO	World Health Organisation
WMA	World Medical Association

Declaration

I, Hilda Jane Parsons, do declare that the work on which this thesis is based is my own original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part thereof has been, is being, or will be submitted for another other degree in this or any other University.

I empower the University of Cape Town to reproduce for purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signed Signed by candidate

Dated 28/01/2021

Acknowledgements

I would like to express my sincere appreciation for my supervisors: Prof Diana Hardie and Dr Heidi Smuts, for their encouragement, patience and guidance. I am grateful for the opportunity to do this work and for the valuable assistance given by both supervisors in writing-up and proofreading this thesis.

The NHLS was also very generous in funding the study, for which I am very grateful and for the University of Cape Town for giving me the opportunity to pursue a Master's degree.

A special thanks for the technologists, especially Ms Ntombi Booi, Ms Nomathemba Mgazulwa, Ms Tathym Gelderblom and Ms Pam Topham, who helped me with the bench work.

A big thank you to my friends, especially Dr Annabel Enoch and Wendy Thom who supported, encouraged and gave me the strength to endure.
I am truly grateful for my family and my faith that gave me the desire to continue in moments of doubt and disappointment when times were stressful and our lives were centred on SARS-CoV-2 testing.

List of Tables

List of Figures

Chapter 1 INTRODUCTION AND LITERATURE REVIEW

1. Overview of various aetiologies causing respiratory illness/disease

1.1. General

Acute respiratory infections (ARIs) are one of the most common causes of illness and disease worldwide affecting both the upper and lower respiratory tract (Assane et al., 2018; GBD 2016 Lower Respiratory Infections Collaborators, 2018; Ortín & Martín-Benito, 2015; Rhedin, 2017; Thomazelli, et al., 2017; Dasaraju & Liu, 1996). Bacteria, viruses and fungi are the main causes of ARIs (Table 1) (Cilloniz et al., 2016; Dasaraju & Liu, 1996). A Global Burden of Disease (GBD) study conducted from 1990 to 2016 showed lower respiratory tract infections (LRTI) were the leading cause of mortality in people of all ages. However, the number of deaths has decreased in the last 20 years due to the availability of vaccines and antibiotics (GBD 2016 Lower Respiratory Infections Collaborators, 2018). Similarly, the Atlas of African Health Statistics 2018 also showed that LTRIs followed by HIV/AIDS and gastroenteritis were significant causes of mortality (Brini Khalifa et al., 2018). The World Health Organisation (WHO) reported that ARIs were responsible for 1.9 to 2 million childhood deaths annually with almost ¾ of these deaths occurring in Africa and Southeast Asia (Assane et al., 2018; World Health Organisation – Africa 2018, GBD 2016 Lower Respiratory Infections Collaborators 2018; Schmidt et al., 2011).

1.1.1. Bacteria and viruses associated with respiratory tract infections in humans

ARIs are predominantly caused by viruses, bacteria and fungi (Table 1) (Alimia et al., 2017; Cilloniz et al., 2016; Seema et al., 2015[a]; Seema et al., 2015[b]). The most common bacterium associated with LRTI, specifically pneumonia, is *Streptococcus pneumoniae* though its incidence has declined since the introduction of conjugated pneumococcal vaccines (Cilloniz et al., 2016). Other significant bacterial pathogens include *Bordetella pertussis*, and *Haemophilus influenzae* (Cilloniz et al., 2016). A systemic review and meta-analysis of 28 European studies of community acquired pneumonia (CAP) in adults, detected respiratory viruses in 22% of cases which is similar to recent major studies from Asia and North America where viruses were detected in 27% of the adult patients with CAP. The most prevalent viruses were influenza virus (A & B), rhinovirus, coronaviruses, human parainfluenza viruses (HPIV), human metapneumovirus, respiratory syncytial virus (RSV) and adenovirus (GBD 2016 Lower Respiratory Infections Collaborators 2018; Alimia et al., 2017; Cilloniz et al., 2016; Zar et al., 2016; Seema et al., 2015[a]; Seema et al., 2015[b]; Dasaraju & Liu, 1996).

RSV is frequently associated with pneumonia and ARIs in the elderly, but is most common among children < 5 years old (28- 54%) (Cilloniz et al., 2016; Seema et al., 2015[a]; Porotto et al., 2012). In a USA paediatric population with CAP human metapneumovirus, adenovirus, HPIVs, and coronaviruses, were most commonly detected, (Seema et al., 2015[a]). The Global Burden of Disease (GBD) study, reported that the most common organism causing LRI deaths was pneumococcal pneumonia followed by RSV with 54% of these infections occurring in children under 5 years (Porotto et al., 2012). In the African context a retrospective study testing children under 5 years', showed that viruses were the most prevalent organisms causing respiratory illness, accounting for 79% of the infections (Bhuyan et al., 2017). In a local study, Paarl, South Africa, Zar et al. (2016) showed that pneumonia was very common in <6-month-old infants, with RSV the most significant pathogen although *Bordetella pertussis*, *Haemophilus influenzae*, influenza virus, HPIV, adenovirus, bocavirus and cytomegalovirus, were also frequently detected.

Table 1 Common Miro-organisms Causing Upper and Lower Respiratory Infection (Dasaraju & Liu, 1996)

Clinical Illness	Bacteria	Viruses	Fungi	Other
Common cold (rhinitis, coryza)	Rare	Rhinovirus Coronavirus Human parainfluenza viruses Adenovirus Respiratory syncytial virus Influenza virus	Rare	Rare
Pharyngitis and tonsillitis	Group A β-haemolytic streptococcus *C. diphtheriae* *N. gonorrhoea* *M. pneumoniae* *M. hominis (type 1)* Mixed anaerobes	Adenovirus Coxsackievirus A Rhinovirus Coronavirus Human parainfluenza viruses Epstein Barr virus Herpes simplex virus	*Candida albicans*	Rare
Epiglottitis & laryngotracheitis (croup)	*H. influenzae type b* *C. diphtheriae*	Respiratory syncytial virus Human parainfluenza viruses	Rare	Rare
Bronchitis and bronchiolitis	*H. influenzae* *S. pneumoniae* *M. pneumoniae*	Human parainfluenza viruses Respiratory syncytial virus Adenovirus Herpes simplex virus	Rare	Rare
Pneumonia	*S. pneumoniae* *S. aureus* *S. pyogenes* *H. influenzae* *K. pneumoniae* *E. coli* *P. aeruginosa* *M. pneumoniae* *Legionella spp.* *Anaerobic bacteria* *Mycobacterium tuberculosis* *C. burnetii* *C. psittaci* *C. trachomatis* *C. pneumoniae*	Adenovirus Human parainfluenza viruses Respiratory syncytial virus Influenza Varicella zoster virus Measles Cytomegalovirus Herpes simplex virus Hantavirus	*H. capsulatum* *B. dermatitis* *P. brasiliensis* *C. mimuli* *Candida albicans* *Filobasidiella C. neoformans* *A. fumigatus* *Aspergillus spp*	*P. carinii*

1.1.2. Paramyxoviruses

Viruses from the family *Paramyxoviridae* (PMV) are known to infect both humans and animals (Table 2). In 2018 the taxonomic name “Parainfluenza” changed. Viruses in the family PMV that cause infections in humans sorts into 2 subfamilies namely Orthoparamyxovirinae (Genus Respirovirus) and Rubulavirinae (Genus Orthorubulavirus). The species names of *Human parainfluenza 1 and 3* changed to *Human respirovirus 1 and 3,* while *Human parainfluenza 2 and 4* are now referred to as *Human orthorubulavirus 2 and 4* (Table 2) (International Committee on Taxonomy of Viruses (ICTV), 2019). For this thesis I will continue using human parainfluenza virus (HPIV) as the species name.

HPIV types 1, 2 and 3 were first detected in 1956 when cell culture and hemadsorption techniques were used to study paediatric respiratory samples (Ren et al., 2011; Karron & Collins, 2007; Henrickson, 2003). HPIV 4 was isolated in 1959 from upper respiratory tract samples from children and young adults (Karron & Collins, 2007). This new virus met the same criteria as the other 3 in that it grew poorly in embryonic eggs and did not share many of the antigenic sites on the influenza viruses and led to a new taxonomic group called “Parainfluenza viruses” (Henrickson, 2003). HPIV 4 was later divided into subtypes HPIV 4A and HPIV 4B (Karron & Collins, 2007).

There are several PMVs that infect animals, with the first identified in 1926 in Java, Indonesia and concurrently in Newcastle-upon-Tyne, England and named Newcastle disease virus (NDV). The origin of this virus is unknown but thought to originate from wild birds. With the increase of poultry farming and evolution of this bird virus a new disease was detected when many poultry died. Since then many more avian PMVs have been detected (Table 2) (Karron & Collins, 2007).

Sendai virus (SeV) was discovered in 1952 in Japan when an autopsy specimen from an infant was inoculated into a mouse. SeV appeared to be related to HPIV 1 but did not cause disease in humans. In 1959 Bovine parainfluenza type 3 (BPIV 3) was isolated from cattle known to have shipping fever (a respiratory disease). Many animal viruses were isolated from primary monkey kidney cell cultures, including Simian virus 5 (SV 5) (isolated in 1954) and Simian virus 41 (SV 41) (isolated in 1961). HPIV 2 was shown to be closely related to SV 5 and SV 41, with SV 5 causing croup in dogs (Karron & Collins, 2007).

PMVs infecting animals can also infect humans (Rissanen et al., 2016; Albariño et al., 2014; Marsh et al., 2012; Woo et al., 2012 Bowden et al., 2001). *Menangle pararubulavirus* (1994) and *Sosuga pararubulavirus* (2012) viruses from the *Pararubulavirus* genus have been reported to cause sterility in pigs and fatal respiratory disease in horses and humans (Menangle virus) and febrile illness in humans (Sosuga virus) (Albariño et al., 2014; Bowden, et al., 2001). Henipaviruses (1990), Hendravirus and Nipahvirus, have been the cause of disease in various animals and humans with a mortality rate of 40% – 100% in both (Marsh et al., 2012). Another Henipavirus found in rats, the Mojiang virus (2012), was also found to be lethal when Chinese miners died from severe pneumonia (Rissanen et al., 2016). However, research using a non-pathogenic Henipavirus namely bat-borne Cedar virus has the potential to provide future studies into the pathogenesis of Henipaviruses (Marsh et al., 2012). Over the years many more non-human viruses have been added to the *Paramyxoviridae* family and may continue to grow (Table 2) (International Committee on Taxonomy of Viruses (ICTV), 2019).

1.2. Classification of *Paramyxoviridae*

The name parainfluenza was thus termed, due to similar disease symptoms that are influenza-like and because the virus particle is medium-sized, has a lipid envelope, and hemagglutinin and neuraminidase activity (Karron & Collins, 2007). With the innovation of molecular diagnostic tools, the molecular study of viruses has significantly improved. Due to these innovations both the taxonomy and nomenclature for *Paramyxoviridae* and all other viruses have dramatically changed over the last few years (Table 2 and Table 3) (International Committee on Taxonomy of Viruses (ICTV), 2019).

Table 2 Taxonomic relationships of HPIVs as members of Family *Paramyxoviridae* (adapted from ICTV website - https://talk.ictvonline.org).

Family	Subfamily	Genus	Species
Paramyxoviridae	*Avulavirinae*	*Metaavulavirus*	*Avian metaavulavirus 2,5,6,7,8,10,11,14,15,20*
		Orthoavulavirus	*Avian Orthoavulavirus 1,9,12,13,16,17,18,19*
		Paraavulavirus	*Paraavulavirus 3,4*
	Metaparamyxovirinae	*Synodonvirus*	Synodus paramyxovirus
	Orthoparamyxovirinae	*Aquaparamyxovirus*	*Salmon aquaparamyxovirus*
		Ferlavirus	*Reptilian ferlavirus*
		Henipavirus	***Hendravirus henipavirus**** ***Nipahvirus henipavirus**** *Cedarvirus henipavirus* *Ghanaian bat henipavirus* ***Mojiang henipavirus****
		Jeilongvirus	*Beilong jeilongvirus* *Jun jeilongvirus* *Lophuromys jeilongvirus 1* *Lophuromys jeilongvirus 2* *Myodes jeilongvirus* *Tailam jeilongvirus*
		Morbillivirus	***Measles morbillivirus**** *Canine morbillivirus* *Rinderpest morbillivirus* *Cetacean Morbillivirus* *Feline Morbillivirus* *Porcine Morbillivirus* *Small ruminant Morbillivirus*
		Narmovirus	*Mossman narmovirus* *Myodes narmovirus* *Nariva narmovirus* *Tupaia narmovirus*
		Respirovirus	***Human respirovirus 1 (HPIV 1)**** ***Human respirovirus 3 (HPIV 3)**** *Bovine respirovirus 3* *Sendai respirovirus* *Murine respirovirus* *Porcine respirovirus*
		Salamvirus	*Salamvirus salamvirus*
	Rubulavirinae	***Orthorubulavirus***	***Human orthorubulavirus 2 (HPIV 2)**** ***Human orthorubulavirus 4 (HPIV 4)**** ***Mumps orthorubulavirus**** *Bat mumps orthorubulavirus* *Simian orthorubulavirus* *Mammalian orthorubulavirus* *Mapuera orthorubulavirus* *Porcine orthorubulavirus*
		Pararubulavirus	*Achimota pararubulavirus 1* *Achimota pararubulavirus 2* ***Menangle pararubulavirus**** ***Sosuga pararubulavirus**** *Teviot pararubulavirus* *Tuhoko pararubulavirus 1,2,3*

*Animal viruses that infect humans; ** Viruses that infect humans (Virus taxonomy adapted from ICTV web site)

Table 3 Virus taxonomic releases between 1970 to 2018 showing changes in classification and the increase of all known viruses (adapted from ICTV website - https://talk.ictvonline.org).

VIRUS TAXONOMY RELEASES									
YEAR	REALM	PHYLUM	SUB-PHYLUM	CLASS	ORDER	FAMILY	SUB-FAMILY	GENUS	SPECIES
2018	1	1	2	4 & 2	14	150	79	1019	5560
2015					7	111	27	609	3704
1998					3	63	9	233	2370
1990					3	40	9	137	1290
1975					3	17	1	63	309
1971					2	2	0	43	290

1.2.1. Old vs new classification

Over the years' classification of viruses has changed. The old classification system was based on phenotypic characteristics that include:

- Type of the nucleic acid: Deoxyribose nucleic acid (DNA) or ribonucleic acid (RNA), single- or double-stranded genome, linear or circular genome and segmented or not
- Morphology: envelope present/absent, icosahedral/helical/complex symmetry, size, number of capsomers, surface structure
- Epidemiology: geographic distribution, host, seasonal spread, age groups, types of transmission
- Diseases and pathology caused: e.g. hepatitis and syncytia
- Antigenic properties: neutralization via known antibodies
- Sensitivity to various agents: Ultraviolet light, chlorine and ether

(International Committee on Taxonomy of Viruses (ICTV), 2019; Mandal, 2019; World Health Organisation (WHO) – Africa 2018; OpenStax College, Biology 2013; Korsman, et al., 2012).

With the discovery of nucleic acid sequencing in the 1970s genomics has played an important role in taxonomy. The International Committee on the Taxonomy of Viruses (ICTV) has developed and is maintaining a complete database for virus genomes. The genome carries the blueprint necessary to produce new viruses and has been considered the most important characteristic for classification. As seen in Table 3 the number of viral species that have been identified and classified over a >40 years period (1971-2018) has increased dramatically from 290 species to 5590 species, with the inclusion of a Realm and various other categories. As

previously mentioned, naming of HPIVs has changed to *Human Respirovirus* 1 and 3 and *Human orthorubulavirus* 2 and 4 (International Committee on Taxonomy of Viruses (ICTV), 2019; Mandal, 2019; World Health Organisation – Africa 2018; OpenStax College, Biology 2013).

1.3. Virion structure

1.3.1. Size

HPIVs are medium-sized virions. They are pleomorphic particles when viewed by negative staining electron microscopy (Figure 1A) (Henrickson, 2003). However, cryoelectron microscopy images show perfect spheres (Figure 1B) with diameters varying from 129 to 360 nm (average of 217 nm) and elongated particles about 445 nm (Karron & Collins, 2007).

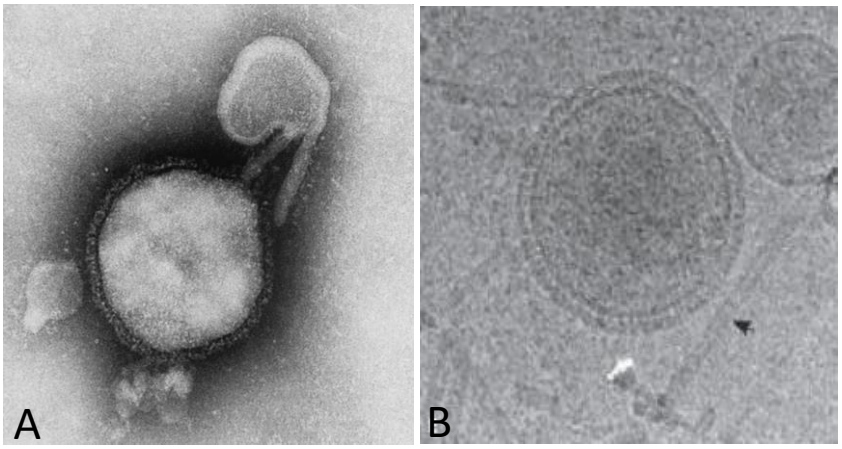


Figure 1 A: Electron photomicrograph of a parainfluenza virion (Henrickson, 2003) **B**: Cryomicrograph of ice-embedded PIV 5 with intact virions and free nucleocapsids (arrow) (Karron & Collins, 2007).

1.3.2. Genome

The HPIV genome is a single-stranded negative-sense RNA genome ranging in length from 14.9 to 17.3 kb (Karron & Collins, 2007). The differences in length occur mainly in the noncoding regions. The genome is not capped or adenylated but has a short 3' extragenic leader region of about 55 nucleotides (nt) followed by six genes encoding the nucleoprotein (N), phosphoprotein (P), matrix (M), fusion (F), hemagglutinin-neuraminidase (HN) and large (L) proteins followed by an extragenic trailer region of 21 to 291 nt (Thomazelli et al., 2017; Karron, & Collins, 2007) (Figure 2).

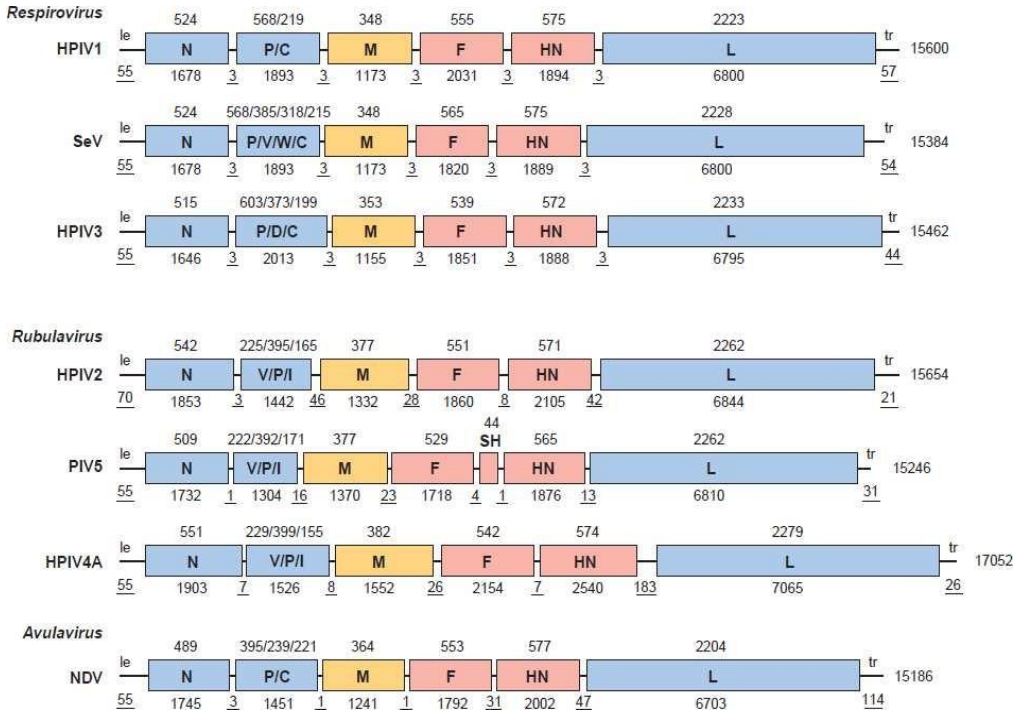


Figure 2 Selected gene maps of HPIVs (not to scale). The genome is drawn 3' to 5' (direction of transcription). Nucleotide lengths are shown below each rectangle with amino acid lengths shown above each rectangle (Karron & Collins, 2007). (N-Nucleoprotein; P-Phosphoprotein; M-Matrix protein; F-Fusion protein; HN-Haemagglutinin/neuaminidase; L-Large protein; C/V/D/W/I-Accessory proteins)

1.4. Proteins

1.4.1. Nucleoprotein (N)

The N protein is a conserved protein. It associates with genomic and antigenomic RNA forming highly stable RNase-resistant helical nucleocapsids. The N-terminal (75%) is highly conserved and plays a part in forming a soluble complex with P and associating with other N monomers and RNA to form the nucleocapsid. The variable C-terminal (25%) is not required in nucleocapsid formation (Lamb & Parks, 2007; Karron & Collins, 2007).

1.4.2. Phosphoprotein (P)

The P protein is highly conserved within a genus and there is little sequence variability between genera. It consists of N- and C terminal functional sections divided by a variable spacer region. The P protein assists in the formation of the nucleocapsid during RNA transcription. The N-terminal region of P binds free N protein monomers and maintains their solubility to enable addition to the RNA template. The C-terminal region of P is necessary for transcription. It binds to the nucleocapsid and the polymerase (L protein) (Lamb & Parks,

2007; Karron & Collins, 2007). The P gene also encodes accessory proteins C and V. These are products of the P gene and differ as to which is expressed in the various HPIVs. C is not well conserved and V is highly conserved between viruses. The C is expressed in carboxy-co-terminal forms with translation occurring at one or more translational start sites in the open reading frame (ORF) and is usually non-structural. The V protein has an N-terminal that is fused to a C-terminal V specific domain containing serval cysteine residues. The main functions of the two proteins are to interfere with the hosts innate immunity – namely the type 1 interferon (IFN) response and to downregulate viral RNA synthesis (Lamb & Parks, 2007; Karron & Collins, 2007)

1.4.3. Matrix (M)

The viral envelope is double-layered and contains lipoproteins, glycoproteins, and lipids derived mainly from the host cell (Lamb & Parks, 2007; Karron & Collins, 2007). Attached to the inner side of the envelope is the non-glycosylated M protein (Lamb & Parks, 2007). This is the most abundant protein in the virion and plays a role in with budding and virus particle assembly (Harrisons, 2010; Karron & Collins, 2007; Vainionpaa & Hyypia,1994).

1.4.4. Fusion and Haemagglutinin/neuraminidase (F and HN)

All *Paramyxoviridae* possess two membrane glycoproteins, the haemagglutinin-neuraminidase (HN) forming a globular head and the fusion protein (F) forming spikes on the viral membrane (Figure 3.) (Gaymard et al., 2016; Goya et al., 2016; Karron & Collins, 2007; Lamb & Parks, 2007; Moscona, 2005).

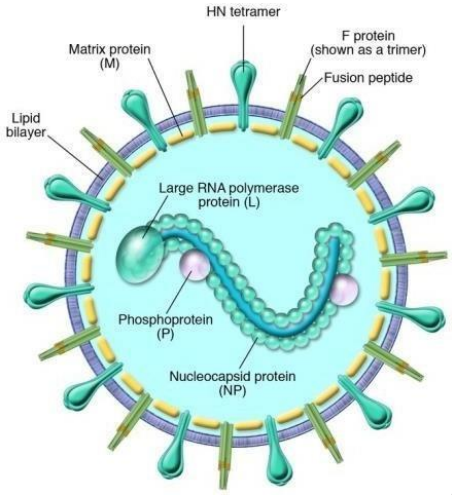
The HN proteins are positioned so that their amino termini extend into the cytoplasm and the C termini are extracellular (Figure 3). It is found on the surface of infected cells and on the virion as a tetramer with disulphide-linked dimers. The HN protein comprises a cytoplasmic domain, a membrane-spanning region, a stalk region and a globular head. The globular head consists of the primary sialic-acid binding site and the neuraminidase (NA) active site (Moscona, 2005). It has the largest antigenic and genetic variation (Goya et al., 2016). Crystallographic studies have shown that the HN of NDV and HPIV 3 have a single active/catalytic site on the globular head called “site I”. X-ray crystal structures on the NDV HN globular head also observed a second binding site called “site II”. Site II has a higher avidity than site I when activated (Porotto et al., 2012). Both hemagglutinin (HA) and NA recognise sialic acid (Vainionpaa & Hyypia, 1994). The HA facilitates cell attachment by binding via the receptor-binding site to sialic acids present in the hosts cell membranes. The

NA serves as a receptor-destroying enzyme that cleaves linked sialic acids, to facilitate the release of newly synthesised virions (Gaymard et al., 2016; Goya et al., 2016; Karron & Collins, 2007; Lamb & Parks, 2007; Moscona, 2005; Vainionpaa & Hyypia, 1994).

The fusion protein (F) directs the fusion of the virus to the uninfected epithelial cells membrane. The F protein is a type I glycoprotein consisting of a cleaved N-terminal hydrophobic signal peptide and a C-proximal membrane anchor. The anchor consists of 2 hydrophobic domains the one adjacent to the N-terminal heptad repeat region (HR-N) and the other adjacent the C-terminal heptad repeat region (HR-C) (section 1.5.1.). The F is synthesized as an inactive precursor and when triggered by the HN protein in *Paramyxoviruses*, is converted into the fusion-ready conformation, thus able to merge with the cell membrane. The efficiency of this triggering is important in influencing the extent of fusion mediated by the F protein, thus extent of viral entry. The accumulation of the F protein, in the late stages of infection, on infected cells also leads to fusion of surrounding cells forming syncytia (Moscona, 2005; Porotto et al., 2012).

1.4.5. Large protein (L)

The L protein is the largest viral protein (approximately 2.200 amino acids in length) with the N- and C-terminal regions being diverse compared to the 6 highly conserved regions in the middle of the protein (Karron & Collins, 2007). The L protein is the viral polymerase and is also responsible for mRNA capping and methylation (Karron & Collins, 2007; Lamb & Parks, 2007). It forms a functional complex with the P (Figure 3) (Karron & Collins, 2007; Lamb & Parks, 2007); Vainionpaa & Hyypia, 1994).



Parainfluenza virion· Photo Credit: JCI 2017

Figure 3 Schematic diagram depicting the genome organisation of a Paramyxovirus (Schmidt et al., 2011)

1.5. HPIV life cycle

1.5.1. Virus attachment and fusion

There are two variants of sialic acid-linked glycans: namely α 2,6 and α 2,3 sialic acid-linked glycans. The first is widespread within the human upper respiratory tract and the latter is widespread on alveolar cells, and both are the target receptors for many respiratory viruses (Rhedin, 2017). The first step of virus entry into the cytoplasm is attachment to these host cell receptors. Entry requires the concerted action of 2 envelope glycoproteins (HN and F). This is initiated by the HN protein which binds to the sialic acid-containing molecules on the host cell membrane. However, even though receptor binding is important, attachment is insufficient. The F protein requires the presence of the HN protein to mediate fusion. Once attached to sialic acid receptor, the HN activates F protein to a fusion-ready state (Moscona, 2005; Palermo et al., 2016). The fusion peptides buried in the F protein trimer are exposed. These peptides are then inserted into the host cell generating a momentary “prehairpin” intermediate anchoring to both viral and host cell membranes state. The F protein folds into a fuso-genic 6-helix structure and the N-terminal heptad repeat region (HR-N) and HR-C (refer to 1.4.4. for F protein activation) regions snap together bringing the virus and host cell membranes together leading to fusion (Moscona, 2005). The fusion process occurs at the cell surface at a neutral pH (Porotto et al., 2012). After fusion the F protein trimer forms a stable post fusion 6-helix structure. The virion can now enter the host cell and begin replicating (Karron & Collins, 2007; Lamb & Parks, 2007; Vainionpaa & Hyypia, 1994).

1.5.2. Replication

The single-stranded negative-sense RNA genome cannot serve as messenger RNA (mRNA) and must be transcribed into mRNA by an RNA-dependent RNA polymerase (Karron & Collins, 2007; Lamb & Parks, 2007; Henrickson, 2003; Vainionpaa & Hyypia, 1994). The host cell does not contain this enzyme thus the virion brings this activity into the cell via the P-L (Karron & Collins, 2007; Lamb & Parks, 2007; Vainionpaa & Hyypia, 1994). Genome replication is carried out by the N-P-L complex. The genomic promoter, at the 3’ end of the genome, serves as the binding site for the RNA polymerase, where it begins copying the genome into a complementary strand termed the antigenome (positive-sense RNA). The antigenome strand acts as a mRNA for initial gene transcription and also as a template for new genome transcripts. The N-terminal of P plays a role in binding free N subunits with the genome for nucleocapsid formation during RNA replication (Karron & Collins, 2007; Lamb & Parks, 2007).

1.5.3. Gene expression

Primary transcription of negative-sense RNA follows the “stop start” model and is synthesized from original genome (El Najjar et al., 2014; Ortin & Martín-Benito, 2015; Vainionpaa & Hyypia, 1994). This results in a gradient of mRNA abundance where the genes at the 3’ end are transcribed more abundantly than genes on the 5’ end (El Najjar et al., 2014; Vainionpaa & Hyypia, 1994). These transcripts contain a methylated cap structure on the 5’ end and a polyadenylated 3’ end (typical mRNA structure) (Vainionpaa & Hyypia, 1994). For transcription of the viral genes the viral polymerase binds to the 3’ end of the antigenome (antigenomic promoter) systematically making copies of the genes with each gene being separately transcribed. In order to synthesize full length genomic RNA, the polymerase must ignore the transcription stop signals at the end of each gene. The mRNAs are then translated via the host cell ribosomes into complete functional proteins. The F protein is synthesized as a precursor, then cleaved to its active form by the host cell proteolytic enzymes. The glycoproteins mature as they are processed through the Golgi apparatus to the host cell membrane (Vainionpaa & Hyypia, 1994). Secondary transcription occurs from newly synthesised viral genomes (Figure 4) which are incorporated in the newly formed virions (Lamb & Parks, 2007).

1.5.4. Virus assembly

Virus assembly occurs at the cell membrane. Viral glycoproteins (F & HN), are transported from the endoplasmic reticulum to the Golgi apparatus, accumulate in the cell membrane forming the envelope (Karron & Collins, 2007). The M protein is responsible for paramyxovirus assembly. The M protein associates with the cytoplasmic domains of the HN and F proteins and the ribonucleoprotein (RNP) capsids to enable complete particle assembly (Moscona, 2005). Fusion and infection of adjacent cells may occur due to the expression of viral F proteins on the surface of infected cells. This can lead to the formation of syncytia in an infected sheet of cells.

1.5.5. Virus release

Once assembly is completed budding occurs through the cell membrane studded with glycoproteins (Figure 4) (Karron & Collins, 2007; Lamb & Parks, 2007). The mechanism of budding is still mostly unknown in HPIVs and there are many theories of release. The budding of HPIVs is driven mainly by the M protein. The M protein of numerous PIVs can induce viral-like particles (VLP) when expressed by itself which may be enough to drive membrane

deformation and outward budding. Alternatively, the L protein (polymerase) interacts with and may recruit cellular proteins of the endosomal sorting complex required for transport (ESCRT) leading to virus particle release. Another theory is that clustering of glycoproteins in the lipid raft micro-domains may create a pulling force on the plasma membrane inducing membrane deformation that is further lengthened by oligomerization of the M protein (El Najjar et al., 2014). All these scenarios lead to the budding process and virus release from the hostcell.

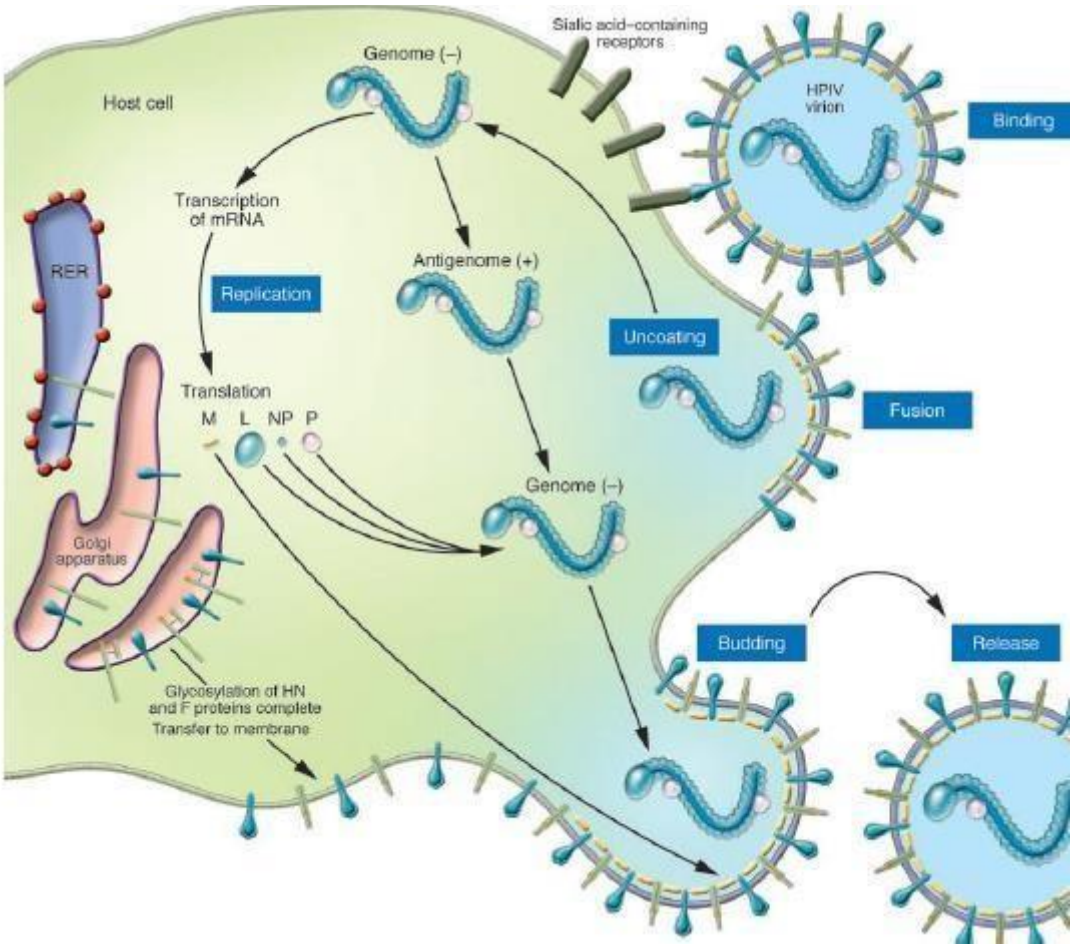


Figure 4 A schematic diagram of the paramyxovirus life cycle (Lamb & Parks, 2007).

1.6. Pathogenesis

HPIVs normally initiate localized infections in the upper and lower respiratory tracts without causing systemic infection, although viremia may occur. Local antibodies IgM, IgA and IgG1 to HPIV 4 develop after primary infection (Henrickson, 2003). However, previous infection does not adequately prevent re-infection but does provide some protection against disease (Lamb & Parks, 2007).

The ciliated epithelial cells of the nose and throat are the primary site of infection, which may extend to the paranasal sinuses, the middle ear, and sometimes to the lower respiratory tract.

Progeny viruses spread among cells both extracellularly and intracellularly due to direct fusion with adjacent cells. Shedding of virus in the respiratory secretions may be for 3 to 16 days following primary infection and 1 to 4 days following re-infection. It starts shortly before the symptoms appear and ends with development of specific local antibody. An inflammatory response causes the main pathogenic changes in the superficial layers of the mucous membranes (Lamb & Parks, 2007).

Clinical syndromes mostly associated with HPIV infection are croup, bronchiolitis, and pneumonia. Virus specific IgE antibodies and release of histamine are associated with the development of croup. Severe manifestation of infection with types HPIV 1 and HPIV 2 is usually croup, whereas HPIV 3 causes all three syndromes. HPIV 4 infections are usually mild, but there have been cases of bronchiolitis and pneumonia reported. In immunosuppressed individuals infection with any of the HPIVs can lead to serious illness (Karron & Collins, 2007; Vainionpaa & Hyypia, 1994). Croup caused by HPIV cannot be distinguished clinically from other viruses causing the same infection such as RSV or measles virus (Lamb & Parks, 2007).

1.7. Anatomy of respiratory tract and disease manifestations

The respiratory tract is divided into an upper and lower region. Most of the respiratory tract is lined with a ciliated pseudostratified columnar epithelial cell layer, but the oral cavity, tonsils and epiglottis have a stratified squamous non-keratinized epithelium layer (Figure 5). Tissue tropism and species-specificity for various viruses is dependent on the glycans (sialic-acid) covering the epithelial cell layers (Rhedin, 2017; Dasaraju & Liu, 1996).

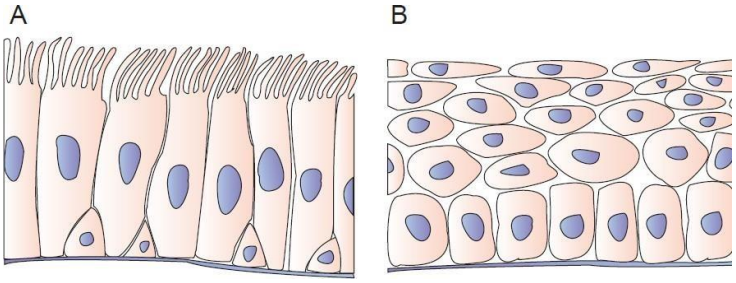


Figure 5 Epithelial cells in respiratory tract**. A:** ciliated pseudostratified columnar epithelial cell layer. **B:** stratified squamous non-keratinized epithelium (Rhedin, 2017; Dasaraju & Liu, 1996).

The upper respiratory tract is composed of the oral and nasal cavities, the sinuses, the pharynx, epiglottis, the Eustachian tube and the tonsils. Upper respiratory tract infections are common in children (Figure 6).

The lower respiratory tract is composed of the trachea, the bronchi, the bronchioli and the lungs (Figure 6). It has been considered a sterile site, but many different organisms reside here challenging this theory.

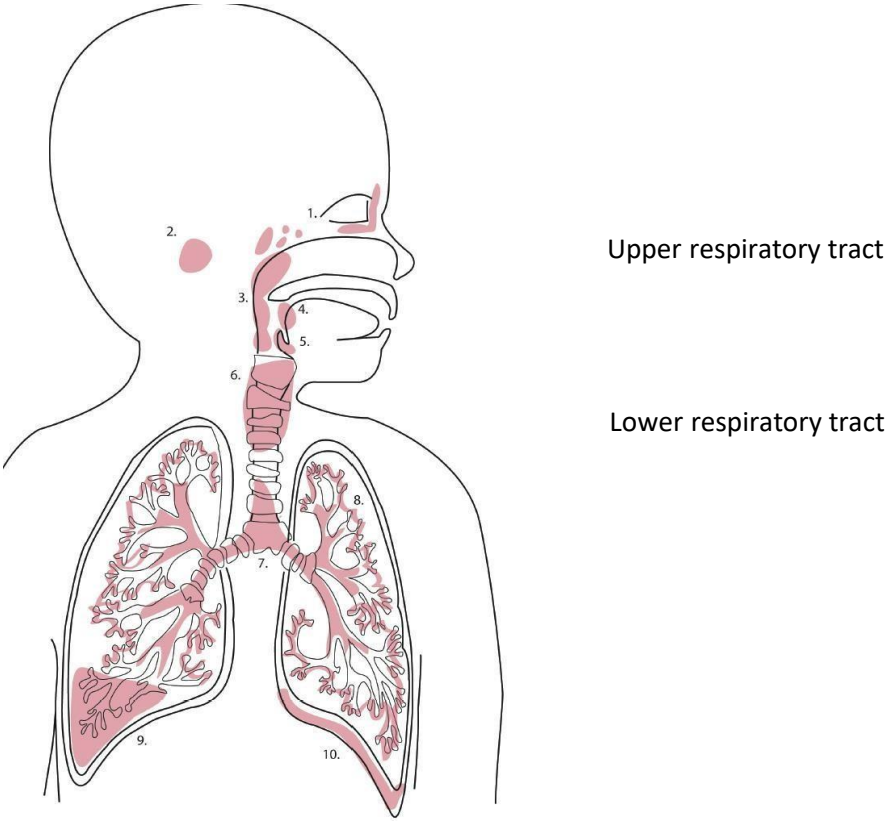


Figure 6 Upper and lower respiratory tract (Rhedin, 2017).

1.7.1. Upper respiratory tract infections

1.7.1.1. Common cold

The common cold is mainly of viral aetiology, that is predominantly a mild upper respiratory tract infection (URTI). The infection is mostly self-limiting and clears within 2 weeks of onset. There is no treatment, other than treating the associated symptoms of nasal congestion, coryza, sore throat or cough. Babies, young children, the elderly and those who are immunocompromised are more severely affected. Due to aerosolization and antigenic variation which is caused by polymerase error the viruses are able to re-infect the same individual many times (Palermo et al., 2016; El Najjar et al., 2014; Vainionpaa & Hyypia, 1994).

1.7.1.2. Croup (laryngotracheitis)

Viral croup is also a self-limiting infection, infecting the larynx and the upper respiratory tract (URT) (Rhedin, 2017). It mainly affects children between the ages of 6 months to 3 years. Symptoms include difficulty breathing, barky cough, and a vibrating rasp caused by inflammation of the mucus membranes in the larynx. During the night hours the symptoms worsen which may be due to low levels of cortisol which plays a role in anti-inflammatory processes (Frost et al., 2013]. The duration is usually less than 2 days (Rhedin, 2017; Palermo et al., 2016). HPIVs are the most common viral cause of croup, with HPIV 1 being the most common cause (26 to 74%) and followed by HPIV 2 & 3 (10%) (Henrickson, 2003). HPIV 4 is the only HPIV that does not cause croup (Pawełczyk & Kowalski, 2017; Frost et al., 2013).

1.7.1.3. Otitis media

Acute otitis media is common in young children it presents with severe earache which is accompanied by a fever and vomiting. Historically considered a bacterial infection but viruses have been recognised to also play a role in 30 to 60% of cases (Goya et al., 2016; Henrickson, 2003). Frost et al. (2013) reported that the most common HPIV associated with this condition was HPIV 4, compared to none associated with HPIV 1 and 2 and less than half with HPIV 3. It presents as a bulging of the erythematous tympanic membrane and may lead to a more serious outcome if an obstruction of the Eustachian tube occurs due to an accumulation of fluid in the middle ear (Rhedin, 2017; Dasaraju & Liu, 1996).

1.7.2. Lower respiratory tract infections

1.7.2.1. Asthma/wheezing

A variety of aetiologies trigger asthma. It is characterized as a hyper-reactivity of the airways and may be allergic, mediated by IgE or non-allergic. Wheezing refers to a high-pitched whistling sound heard in the lungs of children with asthma. Asthma is not an infectious disease but attacks may be triggered by viral infection. All HPIVs are associated with asthma attacks, with HPIV 1 & 4 being the most common (Rhedin, 2017). Treatment with antibiotics has been shown to shorten episodes of asthma-like symptoms, suggesting a bacterial infection may play a contributing role but this may also be attributed to the anti-inflammatory effect of the treatment.

Exposure to a large variety of bacteria during childhood has a protective effect, whereas early exposure to viruses may lead to asthma in later years of life. Viral induced wheezing may not be asthma but the first symptoms of an underlying susceptibility (Rhedin, 2017).

1.7.2.2. Bronchiolitis

Bronchiolitis is caused by a viral lung infection and is characterized by inflammation and congestion in the small airways (bronchioles) (Rhedin, 2017; Dasaraju & Liu, 1996). It is unique to infants due to the size of the terminal airways. The typical symptoms of bronchiolitis are coryza, cough, low-grade fever and severe breathing problems leading to hypoxia (Rhedin, 2017; Henrickson, 2003). This is due to acute inflammation with increased mucus production (Rhedin, 2017). It affects mainly children < 1 years old and disappears almost completely by the time children reach school age (Henrickson, 2003). The peak period for bronchiolitis is during winter. The most common organisms causing this illness are RSV, Human Metapneumovirus (HMPV) and rhinovirus (Rhedin, 2017; Dasaraju & Liu, 1996). All 4 types of HPIV can cause bronchiolitis, with HPIV 1 and 3 being the most common (Parija, 2019; Henrickson, 2003).

1.7.2.3. Pneumonia

Pneumonia is the infection of the lung parenchyma and is accompanied by a high fever, tachypnea, rales, productive cough and lethargy (Rhedin, 2017; Henrickson, 2003). Viruses are associated with 90% of LRTIs in the first year of life, decreasing to 50% by school age. It still plays a significant role in older children but declines to 12% by adulthood. All 4 HPIVs cause pneumonia, with HPIV 1 and 3 causing 10% of outpatient pneumonia, and HPIV 3 being the highest cause in hospitalised patients (Henrickson, 2003).

1.8. Host defences/Immune system

Epithelial cells of the respiratory tract are tightly arranged and secrete antimicrobial peptides preventing viral entry into the cells (Rhedin, 2017). Resistance to HPIVs may be due to innate defences, including type 1 interferon. Natural infection with HPIV in infants and children is not well understood. Protection does not last even though type-specific secretory and humoral immune responses occur. Re-infection with the same serotype may occur within 3 months to several years after primary infection (Lamb & Parks. 2007). Inflammation of the respiratory epithelium is cytokine mediated (Pawełczyk & Kowalski, 2017). A humoral immune response is generated to the viral structural proteins HN and F proteins on the infected cell or virus (Pawełczyk & Kowalski, 2017; Goya et al., 2016). Clinical disease seems to depend mainly

on the concentration of secretory IgA antibodies that possess neutralizing activity, and this neutralizing antibody is found in infants or young children for a short time after primary infection. The presence of serum antibodies in high titres may restrict local virus multiplication. Maternal antibodies do not completely protect against infection, but may influence disease severity (Lamb & Parks, 2007).

Most respiratory viruses, including HPIVs are enveloped and are susceptible to inactivation by complement (Pawełczyk & Kowalski, 2017). Specific defense against HPIV is primarily mediated by both the innate and adaptive immune response. Cellular response aids in restricting viral replication and clearing of HPIV infection. Antibodies directed against the surface glycoproteins HN and F also play a role in conferring long-term protection against HPIVs (Rhedin, 2017). The prevalence of antibodies against HPIV 1, 2 and 3 increases with age, and are twice as prevalent in adults. Immunity to HPIV 1 and 2 develops in the second and third years of life, whereas HPIV 3 occurs in the first year of life. HPIV 4 immunity develops later, occurring in school-aged children virus (Pawełczyk & Kowalski, 2017).

Mucosal and serum antibodies provide lasting immunity, where nasal antibodies were found to be a good correlate of protection, the serum antibodies had to be in high titres in order to provide similar protection (Rhedin, 2017). The T-lymphocyte responses are directed against epitopes in the HN, P and NP proteins of HPIVs. They play a role in clearing HPIV from LRT.

1.9. Vaccines and antivirals

The first attempts to develop a vaccine against HPIV was in 1960 (Parija, 2019). The question that needs to be asked, “Is there a need for HPIV vaccine or antivirals?” “Absolutely” (Parija, 2019).

Even though HPIVs appear an easy target for developing a vaccine due to their self-limiting disease and not establishing persistent infection, none are available. Vaccine development has been rather difficult, and many obstacles need to be overcome. Infection with wild type HPIV does not prevent reinfection. Adults and children experience multiple reinfections but are generally milder. Severe infections with HPIV 3 occur in infants less than 6 months old and a robust immune response at this age is less likely to be induced. Infants also have a less diverse B-cell population and thus less efficient antibody affinity maturation. Maternal antibodies can also interfere with immune system responses to parenterally administered attenuated and mucosally delivered live vaccines (Munoz, 2018; Schmidt et. al., 2011). The

best HPIV vaccine should be immunogenic in infants in the presence of maternal antibodies, protect against RTIs with first contact of wild type HPIVs, plus it needs to be safe and well-tolerated.

HPIVs cause just as many RTIs as influenza. However, influenza vaccine is available whereas there is no licensed vaccine for HPIVs (Durbin & Karron, 2003). The development of HPIV vaccines started in 1960s with a mixture of inactivated HPIV 1, 2 and 3. An antibody response was produced but was not able to protect against an *in vitro* challenge. Many other studies have been performed with reverse genetic technology which has led to second stage vaccine trials e.g. the rHPIV-3-cp-45 has shown promise. A two dose regime was tested on infants 6 to 36 months of age, with good immunogenic outcomes, but a third dose may be needed for those less than 6 months old. (Munoz, 2018)

However, due to lack of HPIV vaccines medications such as corticosteroids and nebulized epinephrine are administered to treat the respiratory symptoms (Munoz, 2018; Schmidt et. al., 2011).

1.10. Laboratory diagnosis

1.10.1. Sample types

HPIVs are shed in the respiratory tract during acute infection and respiratory samples (naso-pharyngeal swabs or aspirates, sputum) are therefore the most appropriate samples to use for HPIV screening.

1.10.2. Cell culture

Virus isolation in cell culture has been the gold standard for detecting respiratory viruses (Bhuyan et al., 2017). Most HPIVs were previously identified by cell culture from respiratory specimens. These viruses are difficult to culture and can take up to 24 days to show cytopathic effect (CPE) thus of no clinical relevance. Importantly, CPE is detectable in only a minority of infected cultures and infection should be confirmed by hemadsorption and immunofluorescent assays using monoclonal antibodies to HPIV antigens (Vachon et al., 2006; Billaud et al., 2005). There have been improvements in turn-around times to 24-48 hours, but still much longer than molecular techniques (3 hours) (Bhuyan et al., 2017; Loeffelholz & Chronmaitree, 2010). HPIV 4 has not been routinely tested for in cell culture as the antigenic typing reagents for HPIV 4 are expensive and not widely available (International Committee on Taxonomy of Viruses (ICTV), 2019; Billaud et al., 2005).

1.10.3. Serology

Diagnostic assays are limited to acute or post infection sera to detect seroconversion or require a 4-fold increase in antibody titre. The detection of IgG also has little significance in patient management as it signifies immunity or passed infection. Testing of IgM antibodies is less sensitive due repeated circulation virus thus decreasing antibody titres (Loeffelholz & Chronmaitree, 2010). Neutralization assays have limited use in a diagnostic laboratory and are mostly used for vaccine trials, retrospective and epidemiological studies (Bhuyan et al., 2017). Further serological testing is not useful as most people are serologically positive.

1.10.4. Molecular Diagnosis

1.10.4.1. "In-house PCR" and Real-time PCR

Molecular testing has revolutionized the detection of HPIVs from clinical samples. Polymerase chain reaction (PCR) enhances sensitivity and increases the speed of obtaining a diagnosis from clinical samples (Wang et al., 2012). Due to HPIV being an RNA virus, a reverse transcription step needs to be included to convert the RNA to copy DNA (cDNA) for the amplification to occur. Real-time PCR (R-T PCR) is widely used as it is less time-consuming and needs less hands-on work and interpretation than "In-house PCR" which requires a nested PCR, which may lead to contamination, then running gel-electrophoresis to obtain an end point result that needs to be interpreted. Many published studies use the P gene to detect HPIV as a diagnostic tool, either using "In-house" or R-T PCR methods (Wang et al., 2012; Lau et al., 2009; Yea et al., 2009; Billaud et al., 2005; Lau et al., 2005). It is a very conserved region and many primer sequences are publicly available. Molecular amplification assays including real-time RT-PCR, provide a very sensitive detection method for diagnosing respiratory virus disease.

1.10.4.2. Commercial assays

Multiplex real-time PCR is now widely used in clinical laboratories and enables the simultaneous testing for multiple respiratory viruses from a clinical sample (World Health Organisation – Africa 2018; Thomazelli et al., 2017; Zhang et al., 2014; Ren et al., 2011; Ren et al., 2009). These commercial assays are also able to separately detect all four HPIVs in one assay. Testing for multiple viruses can create a diagnostic conundrum in knowing what is clinically significant, especially when two or more viruses have been detected.

1.11. Molecular Epidemiology

Molecular epidemiology of parainfluenza is the study of viral epidemiological patterns combined with molecular biology. As technology advances have improved so has the genetic characterisation and molecular knowledge of all living organisms including viruses. This in turn has led to improved typing methods of all micro-organisms. The sequencing of the complete HPIV genome, including HPIV 4B in 2009, has provided a framework to compare molecular sequences for all gene-encoding proteins of the *Paramyxoviridae* family (Yea et al., 2009). Phylogenetic tree analysis are able to show distinct lineages of all the HPIVs. This also facilitates molecular investigations into understand the diversity HPIV 4. Of the 6 genes (M, P, M, F, HN and L) encoding proteins the HN and F genes are the most informative. (Gaymard et al., 2016; Goya et al., 2016; Ortín & Martín-Benito, 2015; Porotto et al., 2012; Schmidt et al., 2011; Vainionpaa & Hyypia, 1994).

HPIVs are endemic globally and may sometimes reach epidemic proportions (Lamb & Parks, 2007). In a study done in the United States approximately 6.8% of all respiratory infections in hospitalized patients were due to HPIV infection (Karron and Collins, 2007).

1.11.1. Age

HPIVs have a worldwide distribution and cause respiratory disease in populations of all ages, which is usually mild in healthy individuals, but may be more severe in children, older adults and immunocompromised individuals (Pawełczyk & Kowalski, 2017; Karron and Collins, 2007; Fry et al., 2006). Symptomatic infection is particularly common in children under 2 years of age but can also cause complications of pneumonia in elderly patients (Yano et al., 2014; Fry et al., 2006).

1.11.2. Transmission, disease caused and duration

In primary infection, virus is shed in the respiratory secretions for 3 to 16 days and 1 to 4 days following re-infection, and is transmitted via contact, fomites and respiratory droplets infecting the epithelial cells of the airways (Pawełczyk & Kowalski, 2017; Lamb & Parks). Only a small inoculum is required to infect. As a result, HPIVs are a major contributor to upper and lower respiratory disease worldwide. (Yano et al., 2014; Zhang et al., 2014; Hsieh et al., 2010; Ren et al., 2009; Karron and Collins, 2007; Billaud et al., 2005). Symptoms appear from day 3 to 4 after exposure and can last up to 17 days with an average of 4 days for HPIV 1 and 6 to 13

days for HPIV 2 and HPIV 4 (Pawełczyk & Kowalski, 2017).

1.11.2.1. Worldwide

1.11.2.2. HPIV's Seasonal trends and prevalence

Many countries including France, China, Canada, Japan, Brazil, England, Wales and the United States (US) reported that HPIV 1 peaks during the autumn/ winter season. However, the US and England reported a biennial increase of HPIV 1 during odd years. This could be due to the cross reactivity of HPIV 1 and HPIV 3. A prevalence rate of 20% to 26% for HPIV 1 (calculated as a % of total HPIV positives) was reported in studies conducted in Brazil, England and the US. During the 2009 to 2011 period, Southern China and Japan had higher prevalence rates of 33% and 52% respectively. (Horton et al., 2017; Thomazelli, et al., 2017; Zhao et al., 2017; Abedi et al., 2016; Yano et al., 2014; Abiko et al., 2013 ; de Mello Freitas, 2013; Liu et al., 2013; Lau et al., 2009; Vachon et al., 2006; Billaud et al., 2005) (Table 4).

HPIV 2 has very similar seasonal patterns to HPIV 1. However, in Japan in a study conducted from 2009 to 2011 reported a peak during late summer early autumn was reported (Yano et al., 2014). The prevalence rate of HPIV 2 was between 7% and 14% in most countries except Brazil which had a prevalence of 40% (2000-2010). However, Wang et al. (2012) showed a 20-fold decrease in the prevalence during the last 2 years of this study. This could be due to the population having a high serological immunity.

HPIV 3 has differing seasonal patterns worldwide. In the US and England, a seasonal spring peak occurs. However, when there is a seasonal decrease in HPIV 1, HPIV 3 increases extending the season into autumn and may also have a small peak in the winter. Brazil, Canada, China and France are similar, peaking from autumn into winter. Japan is the exception with a seasonal peak in summer. HPIV 3 was the most prevalent HPIV worldwide ranging from 25% to 68% (Horton et al., 2017; Zhao et al., 2017; Thomazelli, et al., 2017; Abedi et al., 2016; Yano et al., 2014; de Mello Freitas, 2013; Liu et al., 2013; Lau et al., 2009; Vachon et al., 2006; Billaud et al., 2005) (Table 4).

HPIV 4 the seasonal trend peaked from autumn to winter in most countries worldwide except for Brazil which occurred during winter to spring. A prevalence of 2.1 to 7.7 % was reported in England, US and Japan having the lowest rate of the HPIVs.

Table 4 Various studies showing prevalence of HPIVs worldwide.

Study	Number and type of samples tested	Number of HPIV infections	Age	Season	Ref
France study from 1998 to 2002 of HPIV 4 infections in 20 hospitalized patients	20 samples Sputum, NPA swabs and BALs	20	16 infants 11 < 12 months 5 < 4yrs 3 adults > 65 yrs 1 adult 43	Autumn to winter	Billaud et al., 2005
China (Hong Kong) study from 2004 to 2005 of HPIV 4 infections in young children and immunocompromised patients	2912 NPAs negative for other respiratory virus, were tested for HPIV 4	35 (1.2%) HPIV 4 pos 19 HPIV 4a 16 HPIV 4b	Median age 2yrs Range 1 month to 94 yrs 4 adults (3 >70yrs)	All year round But had peaks in summer and late autumn	Lau et al., 2009
Southern China (Guangdong) study 2009 to 2011 analysing the epidemiology of HPIVs	4755 throat swabs	178 HPIV infections 99 (2.1%) HPIV 3 58 (1.2%) HPIV 1 19 (0.4%) HPIV 2 8 (0.2%) HPIV 4	160 paediatric patients were < 5yrs HPIV 1: evenly distributed over age but highest > 12 months HPIV2: 0-3 months & 3-6 months HPIV 3: < 5yrs but > 12 months HPIV 4: 5 months to 8 yrs	HPIV 1 & 3 Autumn to winter Or summer to autumn HPIV 2 & 4 Winter to spring	Liu et al., 2013
Japan (Mei Prefecture)study 2009 to 2011 measuring antibody titres in samples from subjects who underwent medical treatment at their workplace	725 serum samples	142 HPIV positive HPIV 1: 52% HPIV 2: 14.8% HPIV 3: 25.4% HPIV 4 7.7%	<1 yr: 30 1 yr: 69 2 yrs: 42 3-4 yrs: 57 5-9yrs: 71 10-14 yrs: 72 15-19 yrs 60 20-29 yrs: 138 30-39 yrs: 83 40-49 yrs: 54 >50 yrs: 49	HPIV 1: throughout the year peak in autumn HPIV 2: post summer HPIV 3: early summer HPIV 4 autumn/ winter	Yano et al., 2014
Japan study (Yamagata) 2011 to 2012 epidemiology of ARIs in children	1499 NPAs on paediatric ARI patients	43 PIV 4 positive HPIV 4a : 7 HPIV 4b: 11 unidentified: 25	9.3% < 1yr 23.2 % 1yr 4.7% 2yrs 7% 3 yrs 11.6% 4 yrs 14% 6 yrs 14% 6-9 yrs 16.2% 10-15 yrs	86% during winter	Abiko et al., 2013
Canada study of HPIV from 2004 to 2005	1424 NPAs	371 (26%) were positive for a virus 9 (43%)HPIV 4 (most frequent HPIV) HPIV 4a: 8 HPIV 4b: 1	6 children (5 < 6 months) 3 adults	Autumn to winter	Vachon et al., 2006
Brazil surveillance study of respiratory viruses from 2000 to 2010	37120 samples 6421 were virus +	HPIV 1437 (22.4%) HPIV 1: 277 HPIV 2: 571 HPIV 3: 589 HPIV 4: not tested	Age: HPIV 1,2,3 0-4 yrs: 54, 114, 178 5-14 yrs: 116,143,243 15-24 yrs: 20,76,35 25-59 yrs: 80,211,109 >60 yrs: 7,27,24	HPIV 1: May to Sept HPIV 2: March to Sept HPIV 3: July to Nov HPIV 4: not tested	de Mello Freitas, 2013
Brazil surveillance study of respiratory viruses from Feb 2008 to Aug 2010	1002	HPIV: 104(10.4%) HPIV 1: 12 (11.5%) HPIV 2: 2 (1.9%) HPIV 3: 60 (57.7%) HPIV 4: 30 (28.8%)	Median age HPIV 4: 7m HPIV 1: 11m HPIV 2: 5m HPIV 3: 8m HPIV 4: 3-6m Other RV's: 473 (47.2%) RSV: 244 (24.3%) Adeno: 121 (12.1%) Influenza: 58 (5.8%) HMPV: 50 (5%)	HPIV 1: Autumn HPIV 2: March to Sept HPIV 3: July to Nov HPIV 4: not tested	Thomazelli, et al., 2017

In Brazil, HPIV 4 was second to HPIV 3 with a 30% prevalence of total HPIVs. (Horton et al., 2017; Zhao et al., 2017; Thomazelli, et al., 2017; Abedi et al., 2016; Yano et al., 2014; de Mello Freitas, 2013; Liu et al., 2013; Lau et al., 2009; Vachon et al., 2006; Billaud et al., 2005) (Table 4).

1.11.2.3. Africa

The burden of respiratory disease attributed to viral aetiologies and specifically HPIV 4 have been hindered in less-developed countries due to access to health-care and doctors just treating the patients and not finding out the pathogen-specific cause of the disease. There have been very few published studies showing the presence of HPIV 4 in respiratory samples on the African continent. A study by Brini Khalifa et al. (2018) in Tunisia, on neonates, reported that 10.9% of the cohort had HPIV infection, with rhinovirus the predominant virus causing respiratory infection. A study in Kenya, on children > 5 years with ARI, reported a 7 % HPIV prevalence with RSV predominating followed by Adenovirus (Breiman et al., 2015). In a third study conducted in Gambia RSV was the dominant virus causing respiratory infection with HPIVs reported at 11 % after Influenza (Mackenzie et al., 2019). A South African surveillance study only testing HPIV 1, 2 and 3 reported a 6.5 % HPIV prevalence with HPIV 3 being most prevalent (Cohen et. al., 2015). Based on this assessment there have only been 3 studies in Africa excluding South Africa.

1.11.2.4. South Africa

Despite the widespread use of real-time assays that detect HPIV 4 in the clinical setting, little is known of the molecular evolution of HPIV 4, or if there are differences in the clinical manifestation with specific variants. The prevalence of HPIV 4 associated with respiratory infections in South Africa is largely unknown.

At the Groote Schuur National Health Laboratory Services (NHLS) laboratory, a real-time PCR assay (Anyplex II RV16; Seegene, South Korea) has been used for detection of respiratory viruses including all 4 HPIVs in patients hospitalised with LRTI's.

Although a number of papers have been published over the last few years showing the increasing global interest in this HPIV 4 as an emerging virus, reporting the prevalence, symptoms, hospitalization and the complications of HPIV 4 (Pawełczyk & Kowalski, 2017; Thomazelli et al., 2017; Yano et al., 2014; Zhang et al., 2014; Liu et al., 2013; Ren et al.,

2009; Lau et al., 2009; Vachon et al., 2006). Peer-reviewed literature on the epidemiology of HPIV 4 in South Africa are not available.

1.12. Aim

The aim of this study is to characterize the molecular epidemiology of HPIV 4 causing respiratory infections in the Western Cape, South Africa, 2014 to 2017. This will involve the development of in- house PCR assays to characterize HPIV 4 viruses in respiratory samples and to establish the seasonal occurrence and demographics of patients infected with HPIV 4.

1.12.1. Objectives

The specific objectives are as follows:

a. Determine the seasonal occurrence of infection and demographics of patients infected with HPIV 4.

   Rationale: HPIVs cause seasonal illnesses, however, not much is known of this emerging virus in South Africa. This will be achieved by implementing this test as a routine diagnostic test with all results captured in a data base. This will be evaluated to establish the seasonal occurrence and demographics of the patients infected with HPIV 4.

b. Develop in-house PCR and real-time PCR-based assays to detect and subtype HPIV 4 viruses in respiratory clinical samples over 4 successive winter seasons in Cape Town. *Rationale: HPIV 4 is not routinely diagnosed in virology laboratories. There is also very little knowledge of the prevalence of this emerging virus due to it not being routinely diagnosed as there are very few real-time or traditional PCR assays available. Developing an in-house real-time PCR assay will aid in detecting this virus at a much cheaper cost and will generate a wealth of information as to best respiratory clinical sample to be used and determining the prevalence of this emerging virus.*

c. Determine the genetic variation in the hemagglutinin-neuraminidase (HN) gene in viruses over successive winter seasons.

   Rationale: HPIVs are RNA viruses with an error prone polymerase. Point mutations are likely to accumulate in key genes and may be selected for at a population level if advantageous to the virus (e.g. antigenic drift driven by herd immunity).

CHAPTER 2 MATERIALS and METHODS

2.1. Study Design

This is a retrospective study undertaken over a 4-year period (2014-2017) that investigated the prevalence, seasonality and genetic variation of HPIV 4 in respiratory samples from infants, children and adults presenting at a hospital or clinic with respiratory illness in the Cape Town area. The study also developed a real-time PCR assay to subtype HPIV 4.

2.1.1. Ethics

This study was approved by the Human Research Ethics Committee (605/2017) of the University of Cape Town (Faulty of Health Sciences), South Africa (see appendix).

Consent was not obtained as samples were not taken from patients specifically for the study. This retrospective study was done on genetic material obtained from a bank of stored samples already tested diagnostically at the National Health Laboratory Services (NHLS) laboratory. Confidentiality was signed and maintained by all working on the study and no patient details were published. All patient details and results were password protected and only the PI and supervisors had access to this information. The study complied with the WMA Declaration of Helsinki 2013 (HREC605/18).

2.2. Inclusion and Exclusion criteria.

Over the 4-year study period a total of 7456 respiratory samples were received in the Virology laboratory for routine respiratory virus diagnosis including HPIV 4. Clinical and demographic information available for each HPIV 4 positive case was extracted from laboratory information system.

2.3. Study population and samples

There were 633 HPIV 4 positive results over the 4- year period studied. Respiratory samples were submitted to the NHLS Virology diagnostic laboratory for routine investigation. These samples were sourced from infants and children admitted to Red Cross Children's War Memorial Hospital in Cape Town and adults from other hospitals or clinics. Only samples for acute respiratory infections were tested. External quality control samples (QCMD) from the corresponding study period were also retrieved to verify the real-time PCR developed.

A search on the TrakCare data system for HPIV 4 positive cDNA and extracted eluates was

conducted. These previously extracted samples were retrieved from diagnostic storage, placed in cryo-boxes clearly labelled with the year tested and then stored at -20 °C for further analysis.

2.4. Epidemiological analysis for HPIV 4 prevalence and seasonality

A database search of HPIV 4 positive samples by month and year was performed to determine seasonality over 4 years. Epidemiological curves were constructed from this data using Excel. Analysis of the following data was performed: age, gender, hospital ward (used as a proxy for disease severity) and collection date (to indicate seasonality). This data and all other generated for this retrospective study had no effect on the clinical management of patients.

2.5. Genetic analysis of HPIV 4 positive samples

2.5.1. Nucleic acid extraction

Nucleic acid extraction of routine diagnostic respiratory samples (nasopharyngeal swabs, nasopharyngeal aspirates, tracheal aspirates, bronchiolar lavage, sputum) had previously been performed as follows: samples were added to a mucolytic solution (0.5 ml of a 0.1 % dithiothreitol solution (Appendix A) and thermomixed at 37°C for 15 minutes to aid liquefying of these samples. Samples were then centrifuged for 3 minutes at 4000rpm and 0.5 ml of supernatant was added to lysis buffer and extracted on the EasyMag extraction platform as per manufacturers' instructions (Biomerieux, South Africa).

2.5.2. cDNA synthesis

In order to amplify RNA, it must first be converted to complementary DNA (cDNA). This is performed via reverse transcription using a reverse transcriptase enzyme (RNA-dependent DNA polymerase) which acts on single-stranded RNA. This enzyme is derived from the polymerase gene from either the Avian myeloblastosis virus (AMV) or Moloney murine leukaemia virus (MMLV). Initial reverse transcription requires heating to denature secondary structures or double-stranded RNA. As the reaction cools, primers bind to complementary sequences on the genome (mRNA, prokaryotic or viral RNA) forming RNA-DNA hybrids which are the starting point of new strand synthesis. The second step involves the elongation of these hybrids when the reverse transcriptase adds dNTPs in a 3' to 5' direction forming various sized fragments of cDNA. There are different types of primers: oligo-dT primers, random primers and sequence-specific primers. The oligo-dT primers anneal to the poly- A

tail at the end of any messenger RNA (mRNA). This is advantageous as full-length copies from many different transcripts can be synthesized. However, a disadvantage is when RNA is degraded or does not have a poly- A tail. This can be overcome by using random primers. These are random sequences 6 nucleotides long and are able to bind at many different positions on the RNA template and also bind to any type of RNA, generating short cDNA fragments. Thus, they can bind RNA without poly-A tails or if RNA is degraded or has a secondary structure. Sequence-specific primers bind to a specific known target region usually at the 3' terminal end of the region of interest and result in a pure cDNA population (cDNA of only one specific RNA sequence is synthesized). The single-stranded cDNA can also be converted to double-strand cDNA for use in real-time PCR, sequencing and constructing cDNA libraries (Bustin & Mueller, 2005; Resuehr & Spiess, 2003).

2.5.2.1. Method

cDNA was generated with random hexamer primers using the RevertAid First Strand cDNA synthesis kit (Thermofisher Scientific, Lithuania), as per manufacturer's instructions with minor modifications**.** Briefly, 10 µl of nucleic acid eluate was added to 5 µl mastermix 1 (1,25 µl of random hexamer primer mix,1,25 µl RV16 internal control and 5 µl H_2O) and incubated for 3 minutes at 80°C. The mixture was cooled to 37°C before the addition of 10 µl mastermix 2 (2,5 µl dNTP, 1,25 µl RNase, 1,25 µl reverse transcriptase and 5 µl buffer) and incubated for 90 minutes at 37°C followed by 2 min incubation at 94°C and cooled to 4°C. Eluates were stored at -80°C. Residual eluate and cDNA of HPIV 4 positive samples was used in the current study.

2.6. Primer design for HN gene PCR

In-house PCR primers were developed (a) to amplify a region of the HN gene for phylogenetic analysis and (b) to develop an in-house subtyping protocol for HPIV 4 in clinical samples.

2.6.1. Primer design for HN region targeted

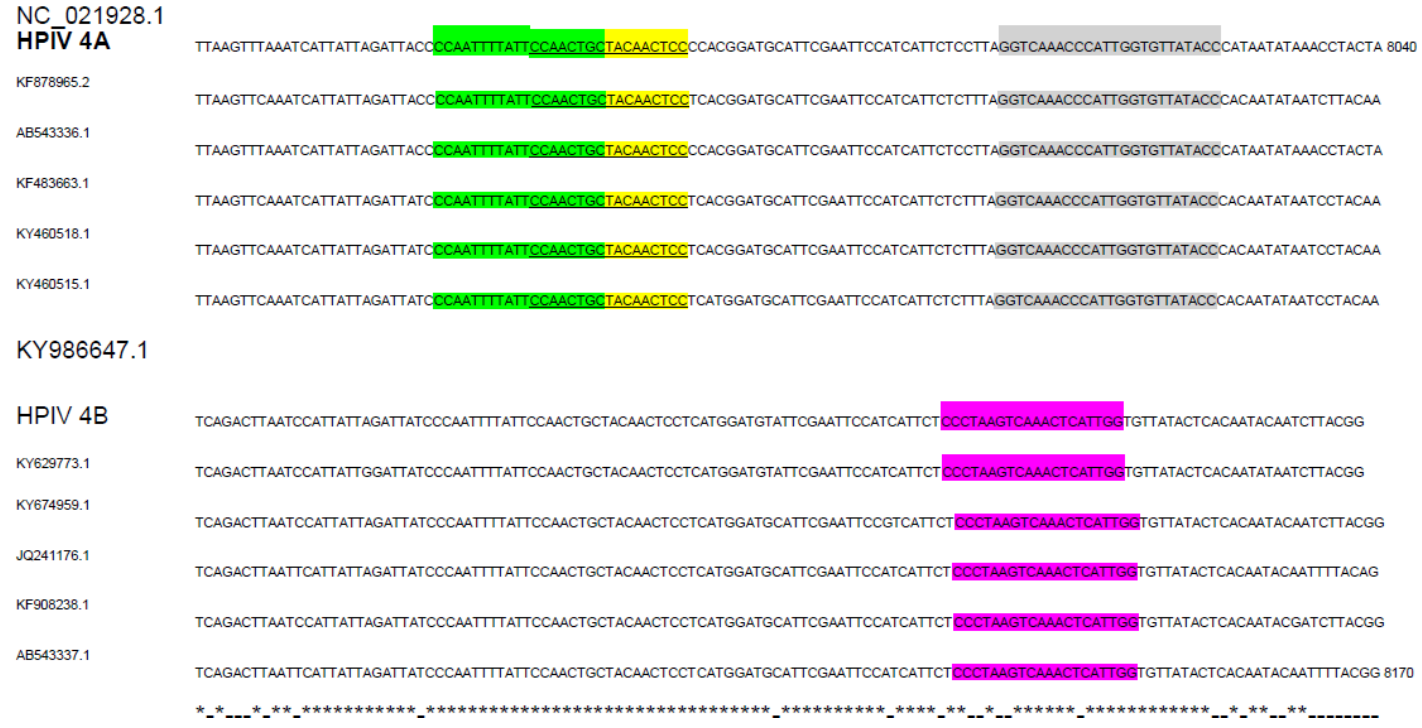
Primers are short strands of single-stranded DNA (or RNA) that serve as the initial foundation for the DNA replication process and are used to demarcate the segment of the DNA template to be amplified. In the PCR process, two primers (forward and reverse primer) are matched to the segment of DNA in a complementary manner with one primer located upstream of the region to be amplified and the second downstream. These primers are usually 18 to 24 bases

long. The melting temperature (*T*m) of the primers (as determined primarily by GC content) should be similar and cannot be much higher or lower than the PCR annealing temperature as this results in non-specific binding of primers to non-specific targets.

2.6.1.1. Method

HPIV 4 complete genome sequences were obtained from GenBank and copied to a text file. Using the BioEdit software version 7.2.5 (Hall, 1999) the sequences were imported and aligned (Figure 7) using the ClustalW Multiple alignment program. Once aligned, conserved regions within the globular head domain (Figure 8) (Ping et al. 2005), which spans 450 aa (1350 bp) in length within the 3' prime NA region of the HN gene, were identified and forward and reverse primer chosen. This region of the HN gene was targeted because of the high genetic variation within this region which allowed discrimination of the subtypes (Goya et al., 2016). These primers (Table 5) were purchased through Whitehead Scientific Ltd, South Africa but synthesized by Integrated DNA Technologies (IDT) using standard desalting with a final concentration of 100 picomoles.

A: Location of forward primers



B: Location of reverse primer

Figure 7 Sequence alignment of HPIV 4 subtype A and B HN genes showing the region used to design forward and reverse primers (Table 5). The star* represents aligned nucleotides. The dash – represents non-aligned nucleotides. **A:** The highlighted nucleotides indicate the forward primers designed to detect HPIV 4 and subtypes. **B:** The highlighted nucleotides indicates the universal reverse primer.

HPIV 4 (F1) – highlighted green

HPIV 4 (F2) – highlighted green/yellow

HPIV 4 (R) – highlighted blue

HPIV 4A (F) – highlighted grey

HPIV 4B (F) – highlighted pink

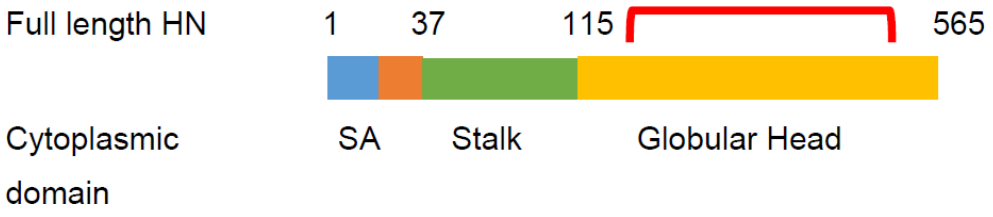


Figure 8 Diagrams representing the stalk and NA regions within the HN region. Area highlighted in red indicates region amplified by in-house PCR (adapted from Ping et al., 2005).

2.7. Amplification of HN gene regions

2.7.1. Polymerase chain reaction (PCR)

PCR is an *in vitro* replication process using a thermostable DNA-dependent polymerase to make large amounts of DNA from a DNA template. The template may be double-stranded, single-stranded or an RNA/DNA hybrid. There are three phases to each cycle of amplification: denaturation of the template at 95 ºC, annealing of primers at 50-60 ºC to the single-stranded template in a site-specific manner and then elongation, in 5' to 3' direction, at 72 ºC. This process is repeated over 35 to 45 cycles, resulting in an exponential increase in amplified product of specific targeted nucleic acid (Bustin & Mueller, 2005). To increase sensitivity a second round (nested) of amplification with different primers internal to the first set can be performed. A modification of the second-round amplification is a hemi/semi-nested PCR where one primer is common to the outer/initial reaction and the second primer is internally located (Korsman et al., 2012).

2.7.1.1. Method

Primer sequences used for first round amplification were: HPIV 4 (F1) and HPIV 4 (R). These generated a product of 796 bp. Primer sequences used for hemi-nested amplification and sequencing: HPIV 4 (F2) and HPIV 4 (R) generated a product of 784 bp (Table 5). These target the globular head of the HN gene of HPIV 4 (Figure 8 and Figure 9). The PCR was performed in a 50 µl reaction containing 37.1 µl water (Sigma, South Africa), 5 µl 10 x buffer, 3 µl 25 mM $MgCl_2$, 0.3 µl (250 U/µl) Super-Therm polymerase (JMR Holdings (JMR-801), Separation Scientific, South Africa), 0.1 µl of a 25 µmol per dNTP (Biolab, Inqaba, South Africa), 1 µl of 50 picomoles for each primer.

Amplification was as follows: For the initial amplification, 4 µl of cDNA was added to 46 µl of mastermix and incubated through 45 cycles at 94º C for 10 s, 55º C for 30 s, and 72º C for 45 s using the Bio-Rad 2700 thermal cycler (Bio-Rad) resulting in a 796 bp PCR fragment. For second round amplification 2 µl of the initial product were then added to 48 µl of the same mastermix with F2 and R primers for HPIV 4 and an additional 8 µl of water (Sigma, South Africa). The nested amplification conditions was the same as for the initial amplification, but with an annealing temperature of 60° C. Visible 784 bp PCR fragments were used for sequencing (Figure 9).

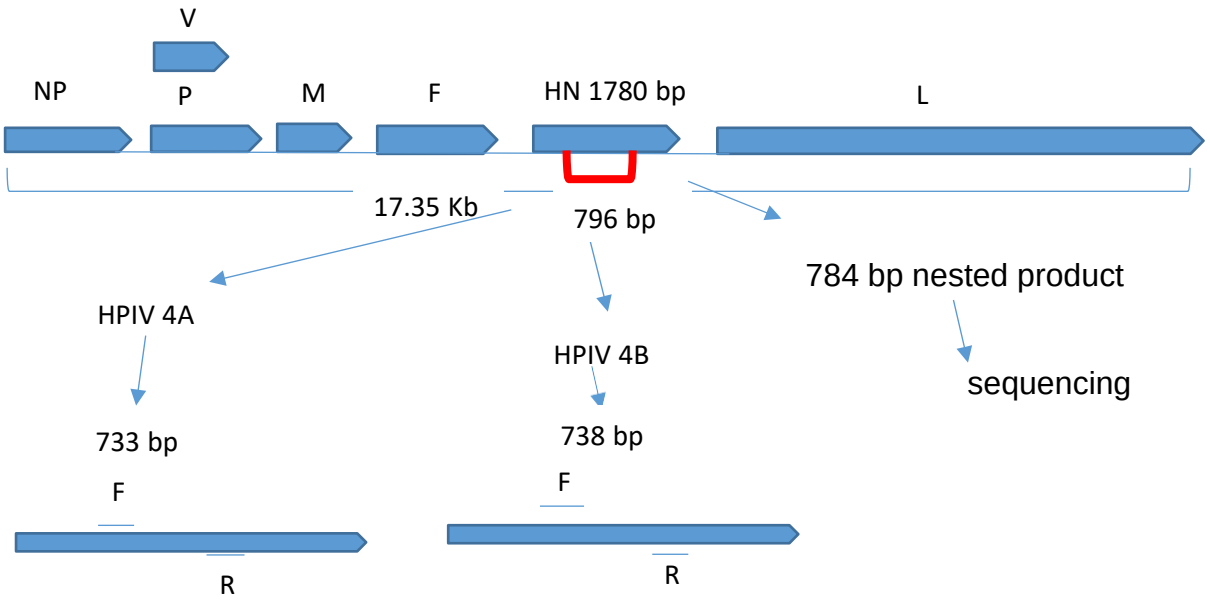


Figure 9 Schematic diagram used to show the in-house generated PCR products based on HN gene (based on NCBI reference sequence NC_021928.1). A 796 bp HN gene fragment (indicated in red) is the region amplified and the nested product of 784 bp sequenced for phylogenetic analysis. Nested PCR of the outer 796 bp product produced fragments of 733 bp for subtype A and 738 bp for subtype B (Kb: kilobases bp: base pair, HN: Haemagglutinin neuraminidase gene).

2.7.2. In-house subtyping PCR

The same 796 bp outer PCR product generated by primers F1 and R was used in an in-house subtyping PCR protocol. For each sample, 2 nested PCRs (in separate tubes) were performed using either primers F4A and R, or F4B and R. Two microliters of the outer amplified product (initial amplification product) were added to 48 µl mastermixes containing forward primers for each serotype. HPIV 4A (F) and HPIV 4 (R) resulted in a 733 bp PCR fragment. The HPIV 4B primers used were: HPIV 4B (F) and HPIV 4R resulting in a 738 bp PCR fragment (Table 5) (Figure 9). This was incubated through 45 cycles at 94º C for 10 s, 60º C for 30 s, and 72º C for 45 s using the Bio-Rad 2700 thermal cycler (Bio-Rad). PCR subtyping results were compared with those obtained by real-time PCR and phylogenetic analysis.

Table 5 Forward and reverse primers for HPIV 4 and subtypes HPIV 4A and HPIV 4B targeting the HN gene.

REGION	PCR	NAME	SEQUENCE	SIZE bp	POSITION	ACCESSION # & REFERENCE
PCR PRIMERS for phylogenetic analysis						
HN gene	Hemi-nested	HPIV 4 (F1)	CCCAATTTTATTCCAACTGC	796	8066-8085	NC_021928.1 This study
HN gene	Hemi-nested	HPIV 4 (F2)	CCAACTGCTACAACTCC	784	8078-8094	NC_021928.1 This study
HN gene	Hemi-nested	HPIV 4 (R)	GGATCCAGCTGGTGGCCC		8846-8862	NC_021928.1 This study
PCR PRIMERS for in-house typing method						
HN gene	Hemi-nested	HPIV 4A (F)	GGTCAAACCCATTGGTGTTATACC	733	8129-8152	NC_021928.1 This study
HN gene	Hemi-nested	HPIV 4B (F)	CCCTAAGTCAAACTCATTGG	738	8124-8143	NC_021928.1 This study
HN gene	Hemi-nested	HPIV 4 (R)	GGATCCAGCTGGTGGCCC		8846-8862	NC_021928.1 This study

2.7.3. Agarose gel electrophoresis

Gel electrophoresis separates nucleic acids by charge and size. An electric current is applied to a solid support (agarose gel) and the negatively charged nucleic acids migrate from the negative anode to the positive cathode. Smaller fragments run faster than larger fragments thus differentiating molecules by size. The NA is visualized with an intercalating dye e.g. ethidium bromide, which illuminates when exposed to ultraviolet light. A molecular marker of known band sizes together with known positive controls are run to determine the result (Lodish, et al., 2000).

2.7.3.1. Method

A 2 % agarose/ethidium bromide gel (Appendix A) was prepared. Melted agarose was poured into a gel tray with combs placed in the gel to create loading wells. The amplified product, positive controls and marker were mixed with loading buffer (Appendix A), and applied to each well, aiding in visualization and separation. Current of 115 amps was applied. The gel was stopped after 1 hour and exposed to ultraviolet light using the Gel doc platform (Whitehead Scientific Ltd, South Africa). A photo was captured to determine the results.

2.8 Real-time typing PCR for rapid subtyping of clinical samples

Phylogenetic analysis effectively distinguished the subtype A and B viruses but would not be practical for monitoring large numbers of clinical isolates for epidemiological purposes. For this reason, a rapid, in-house multiplex real-time assay was set up and validated using the study samples.

2.8.1. Real-time PCR (qPCR) of the P gene

The principle of qPCR is different from that of conventional PCR in that the amplification process is monitored continuously in 'real time' by measuring an increase in a fluorescent dye. The dye may be included in the qPCR mastermix (intercalating dye-based method) or be attached to an oligonucleotide (hydrolysis probe-based detection method). Dye-based qPCR measures fluorescence during each cycle. The amount of fluorescent signal measured is proportional to the amount of amplified DNA. However, only one target can be measured and the dye binds to all dsDNA in the sample thus is not specific (Neidler, 2019).

The most common method used for qPCR is probe-based PCR e.g. Fluorescence resonance energy transfer (FRET). It is more sensitive and specific, and many targets can be detected simultaneously (multiplex PCR). However, it is more expensive than dye-based PCR and traditional PCR. This method employs a probe labelled fluorophore and a quencher. If the probe anneals to the specific targeted sequence the Taq polymerase cleaves the probe separating it from the quencher resulting in fluorescence when exposed to a specific light wavelength. A qPCR platform monitors the amplification in real-time using software to interpret the results. The greater the initial viral load the sooner it is detected as amplified product accumulates increasing the fluorescence (Bustin, 2005; Bustin & Mueller, 2005).

2.8.1.1. Method

The primers of Wang et al. (2012) that targeted the P gene were used to set up an in-house multiplex qPCR assay to detect both subtypes of HPIV 4 circulating in the Western Cape (Table 6). The end-point subtyping assay based on HN gene was used to validate the multiplex real-time assay for future use in the diagnostic virology laboratory.

Retrospective testing was performed on 66 samples, comprising 16 negatives, 46 positive HPIV 4 clinical samples and 4 positive EQA samples (positive controls). To calculate the sensitivity and specificity the following formula was used: Sensitivity = [TP/TP+FN] and Specificity = [TN/TN+FP] (Mallett et.al. 2012). 10x dilution was performed on 2 EQA samples.

Each dilution was run in duplicate to measure limit of detection (LOD), reproducibility, specificity and sensitivity.

The HPIV 4A/B qPCR assay master mix consisted of 4 µl of forward and reverse primers for both HPIV 4A and HPIV 4B. One µl of each probe, labelled with FAM on 5' end and quencher BHQ1 on 3' end (green spectrum at 530 nm) for HPIV 4A and HEX on 5' end and quencher BHQ1 on 3' end for HPIV 4B (yellow spectrum at 580nm) (Table 5). To this 15 µl of PrecisionPLUS Onestep RT-qPCR Master Mix was added (Primer Design, Celtic Molecular Diagnostics Ltd, South Africa). To optimise this assay 4 different volumes of cDNA (2 µl, 4 µl, 6 µl and 8 µl) from HPIV 4 positive EQA samples (QCMD PINFRNA17S-08, PINFRNA16S-02 and PINFRNA14S-05) were added to the mastermix to determine the most efficient amount needed. The optimum volume was added and incubated through 55 cycles at 94º C for 10 s, 50º C for 10 s, and 60º C for 60 s using the RotorGene (Qiagen). The results for HPIV 4A/B were analysed at different wave lengths/ colour spectrums with HPIV 4A at 530 nm (green light) and HPIV 4B at 580 nm (yellow light). If the sigmodal curve crossed the threshold a positive result was obtained. If below threshold the result was negative. All samples were run with controls including a known positive and a blank to rule out contamination.

Table 6 Primer probes targeting P gene for real-time PCR.

REGION	PCR	NAME	SEQUENCE	SIZE bp	POSITION	ACCESSION # & REFERENCE
PCR PRIMERS & PROBES for real-time typing method						
P Gene	RT PCR	HPIV 4A 01 (F)	GCAATTAAGGCAYTAGAAGTRA	129	2790-2811	JN651405-JN651406 Wang et al., 2012
P Gene		HPIV 4A 02 (R)	AATTGTGGCAAGTGAACC		2901-2918	Wang et al., 2012
P Gene		HPIV 4A TM Probe	FAM-TTTGCTAACTTTCCCYTCAATCCTG-BHQ1		2939-2963	Wang et al., 2012
P Gene	RT PCR	HPIV 4B 01 (F)	TCCHATAATCGTCACTGGYA	154	3008-3027	JQ406185–JQ406193 Wang et al., 2012
P Gene		HPIV 4B 02 (R)	TATTTTAAGTGCATCTAT ACG AAC		3138-3161	Wang et al., 2012
P Gene		HPIV 4B TM Probe	HEX-ACAAAATGGGTCTTGCTARCGG-BHQ1		3079-3099	Wang et al., 2012

The acceptance criterion for the validation of the HPIV 4 qPCR assay was to detect a positive sensitivity rate of ≥ 95 % in respiratory samples and to show better evidence of detection sensitivity of the qPCR versus the in-house PCR assay.

The LOD was determined as the lowest dilution with a positive result. To determine reproducibility all duplicates were analysed. The qPCR assay results were compared to the in-house HN gene-based PCR to determine clinical sensitivity and specificity of the HPIV 4 using a 2 by 2 table (Parikh et al., 2008).

HPIV 4 was analysed with RotorGene software (Qiagen South Africa), using the green and yellow channel for HPIV 4A and HPIV 4B respectively. Firstly, the input sample volume needed to be optimised which was performed by adding 2 µl, 4 µl and 6 µl of eluate to the PCR mastermix. The optimal volume was determined to be 4 µl as it gave better results than the 2 µl volume and very similar results to 6 µl volumes i.e. in sample PINFRNA17S-08 the cycle threshold (Ct) values at 4 µl and 6 ul of input produced very similar results of Ct 27,12 and 27,5 respectively, while 2 µl input showed a much higher Ct value of 29.33 (Table 7).

Table 7 Determination of optimal input sample volume

OPTIMISED INPUT VOLUME		
VOLUME µl	#	Ct
2	PINFRNA17S-08	29,33
4	PINFRNA17S-08	27,12
6	PINFRNA17S-08	27,5
2	PINFRNA16S-02	25,98
4	PINFRNA16S-02	26,32
6	PINFRNA16S-02	26,6
2	PINFRNA14S-05	20,28
4	PINFRNA14S-05	19,37
6	PINFRNA14S-05	20,57

2.9. Genetic analysis of HN gene

2.9.1. Sequencing

Sequence analysis of the 784 bp PCR product of the HN gene was performed. The most common method used to determine the nucleotide sequence of a PCR product is the “Sanger method” commonly referred to as chain termination or dideoxynucleotide sequencing. Together with normal dinucleotides (dNTP’s), dideoxynucleotides (ddNTP’s) labelled with a

fluorescent dye are used.

First step in the sequencing reaction is the denaturation the double-stranded DNA, followed by a specific primer annealing to the targeted sequence on the template strand. Complimentary strand synthesis occurs with the known fragment of DNA polymerase incorporating either a dNTP or one of the four fluorescently labeled ddNTPs (present at 100-fold lower concentration) in separate reactions. As synthesis occurs the ddNTP's are incorporated in place of the dNTP's resulting in a chain terminating event. The products are run on a machine which analyses and determines the result according to the wavelength at which each ddNTP fluoresces. The result is shown as a chromatograph, with each coloured peak depicting the specific nucleotide in the sequence (Obenrader, 2003).

Sequencing is the best method used for virus typing. The sequence is downloaded into a program e.g. BLASTn which will show how closely related the sequences are to known sequences stored in a database e.g. GenBank. By comparing the sequences to those that have statistical and significantly similar sequences to the viral sequence studied, the virus can be typed and placed in a gene family. It also aids in showing functional and evolutionary changes that may occur over time.

2.9.1.1. Method

PCR products were cleaned up using the Zymoclean Gel DNA Recovery kit (Zymo Research Corp. Irvine, California, USA) and sequenced directly with the BigDye terminator cycle sequencing kit (Applied Biosystems, Foster City CA, USA) using forward (F2) and reverse (R) PCR-specific primers (Table 5). The sequencing was performed by Inqaba Biotechnology, South Africa and assembled contigs were provided in a FASTA format.

The sequence files of 80 positive clinical samples were assembled and edited to correct any sequence anomalies using BLAST® (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine the sequence

identity and the closest match to HPIV 4 sequences in the GenBank database. These sequences were then aligned together with reference sequences from GenBank (Figure 7) using ClustalW Multiple alignment within BioEdit program, version 7.2.5 (Hall, 1999). To facilitate analysis, each clinical sample was named according to the year and month of collection and this information was used to highlight seasonal genetic characteristics.

2.9.2. Phylogenetic analysis

Phylogenetics is the study of relationships between organisms and how they evolve. Sequencing of genes shows similarities and differences, and with the ability to illustrate sequence relatedness, an understanding of their relationships to ancestral species and how they have evolved can be demonstrated. It also aids taxonomy in the classification of organisms into a realm, phylum, subphylum, class, order, family, genus and species (ICTV website - https://talk.ictvonline.org, 2019; Barry, 2013; Tamura et al., 2013).

A phylogenetic tree is constructed using a software program e.g. MEGA 6, to determine sequence divergence events in species. The tree is used to illustrate research results. "The main question a phylogenetic tree answers are: how are the organisms in the tree related?" The tree consists of branches, nodes and tips where branches connect to form a node with the tips of each node representing the sequence or taxa. The length of the branches represents the approximate time since divergence or changes that occurred between sequences e.g. the longer the branch the longer the time since divergence or a change occurred. A clade is a group of taxa around one divergent node. A sister group are taxon's that have split from the same node and have the same ancestor. An outgroup has the same distant ancestor but is not closely related to the group of interest. However, the outgroup can add a bigger picture as to where the main group of interest falls and is useful to show evolution (Understanding Evolution 2019; Tamura et al., 2013) (Figure 10).

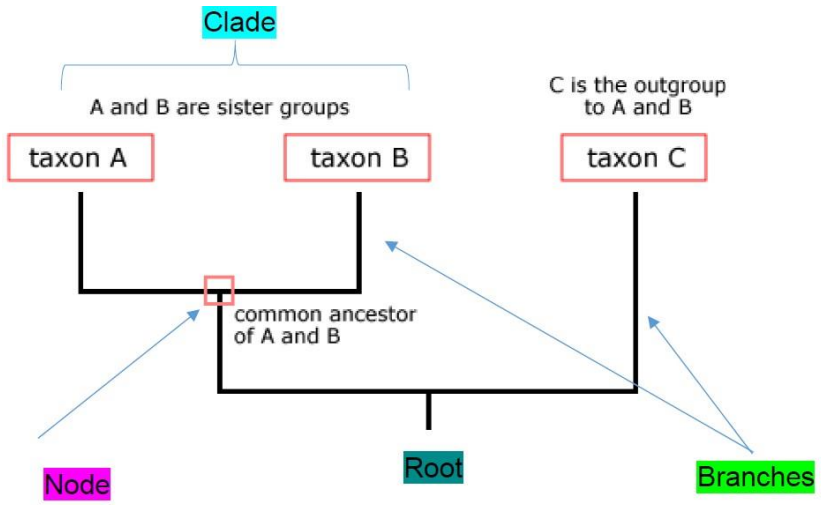


Figure 10 An illustrated diagram of how a phylogenetic tree is interpreted (adapted from website: Understanding Evolution, 2019).

2.9.2.1. Method

Once sequences were blasted (refer 2.9.2 Method) the phylogenetic analysis of a 784 bp

region of the HN gene (Figure 9) was performed. Sequences from clinical samples were aligned with reference HPIV sequences using the BioEdit software version 7.2.5 (Hall, 1999), ClustalW Multiple alignment program (refer 2.9.2 Method). This was then converted to a MEGA format. This file was loaded into MEGA and analyzed using the Neighbor Joining (NJ) algorithm from Molecular Evolutionary Genetic Analysis (MEGA) version 6.06 (Tamura et al., 2013). The robustness of the phylogenetic tree was tested by bootstrapping with 1000 replicates and bootstrap values of 70% were considered statistically robust and included on the branch and nodes of the phylogenetic tree.

2.9.3. Highlighter plot for amino acid (protein) and nucleotide analysis

Highlighter plots are used to visualize mutations in aligned amino acid (aa) or nucleic acid (NA) sequences. Using a software program e.g. ViPR or Highlighter in the HIV Sequence Database these mutations can be observed. Matches, mismatches, transition and transversion mutations, silent and non-silent mutations are some of the features visualized with this software.

To interpret these changes a colour coded “l” is used to depict the change. The colour code should be the same as that used in the alignment software for consistency and easier interpretation. Each NT or aa is allocated a colour. This aids in visualizing the changes. Mismatches are aligned sequences that do not match the master sequence. Transitions are changes in either purine for purine or pyrimidine for pyrimidine. Transversion is where there is a change in purine for pyrimidine. Silent and non-silent mutations occur when a mutation in a nucleotide does not or does change the amino acid in the protein sequence (Keele, et al., 2008).

2.9.4. Subtype variation in HPIV 4 positive clinical samples over the four-year study period

To determine the epidemiology of the HPIV 4A and B subtypes over the study period, epidemiology curves were constructed in EXCEL using the collection dates of the subtyped clinical samples.

Chapter 3 Results

Chapter 3 Results

3.1. Epidemiological analysis of HPIV 4 in the Western Cape

3.1.1. HPIV prevalence

HPIV 4 was the most common of all the HPIVs detected during the study period (Figure 11). During this time, a total of 7456 clinical samples were tested for respiratory infections. These 633 (8.5 %) samples tested positive for HPIV, with 312 (4.2 %), 194 (2.6 %) positive for HPIV 4, 80 (1.1 %) for HPIV 3 and 73 (1.0 %) for HPIV 1 (Figure 10). Over the 4- years HPIV 3 waned with HPIV 4 becoming the most common HPIV and HPIV 1 and 2 had biennial increases with HPIV 1 in odd years and HPIV 2 in even years (Figure 12). These results were not tested statistically.

HPIVs were 6th most common (8.5 %; n = 633/7456) of the viruses detected, with rhinovirus (31.8 %; n = 2367/7456), being the most common respiratory infection followed by adenovirus (24.5%; n = 1823/7456), respiratory syncytial virus (18.1 %; n = 1351/7456), and bocavirus (10.7 %; n = 794/7456) (Figure 13). HPIV 4 (4,2%; n = 312/7456) had a similar prevalence as human metapneumovirus (Figure 13).

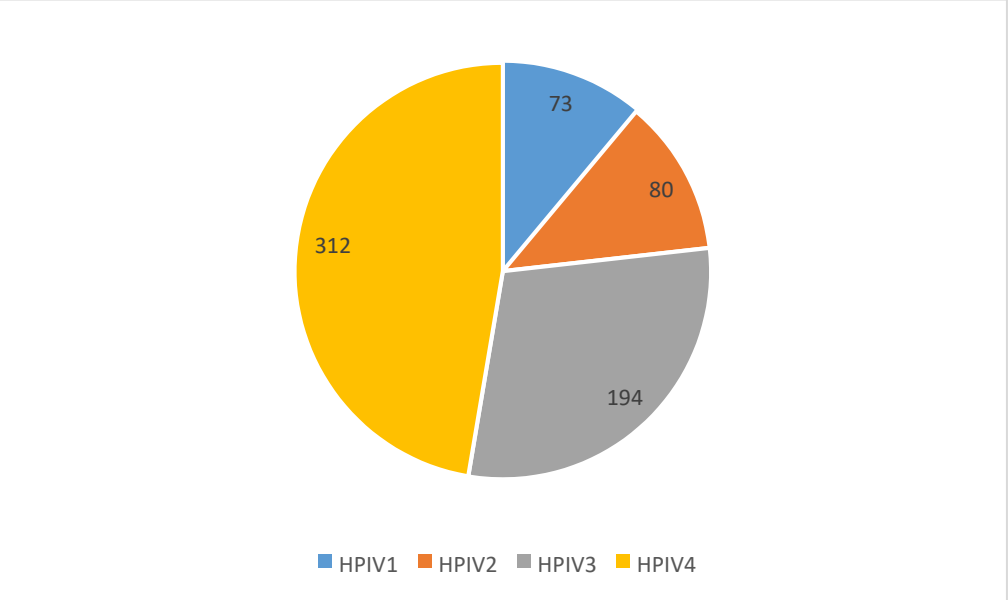


Figure 11 Pie chart showing the number of HPIVs detected in clinical samples during the study period (2014 – 2017) with HPIV 4 being the most common

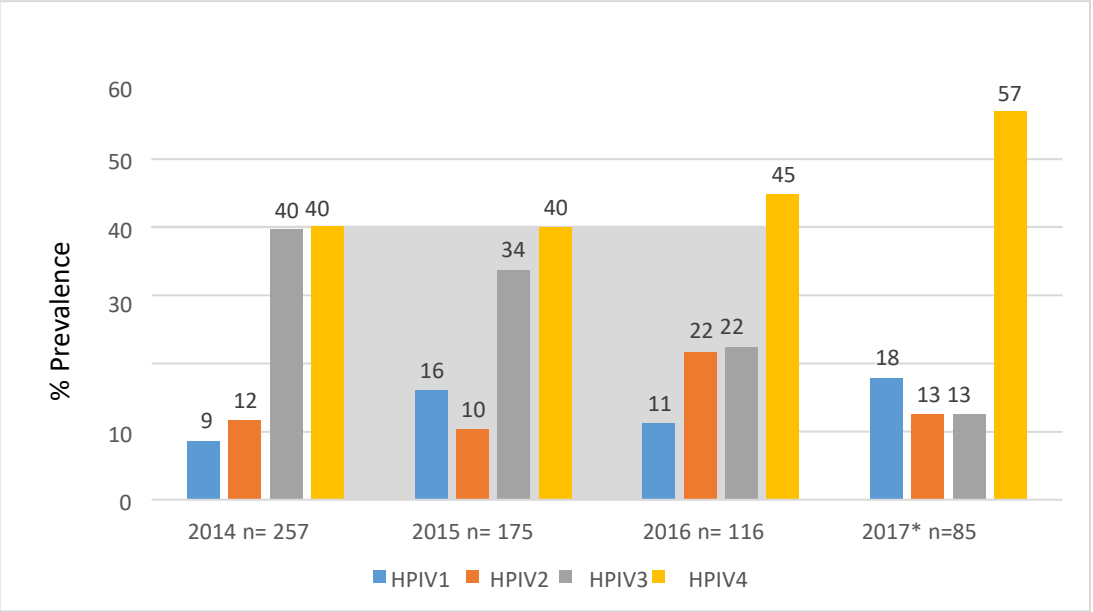


Figure 12 A comparison of HPIVs prevalence during 2014 to 2017 (n=633) (*2017 ended August)

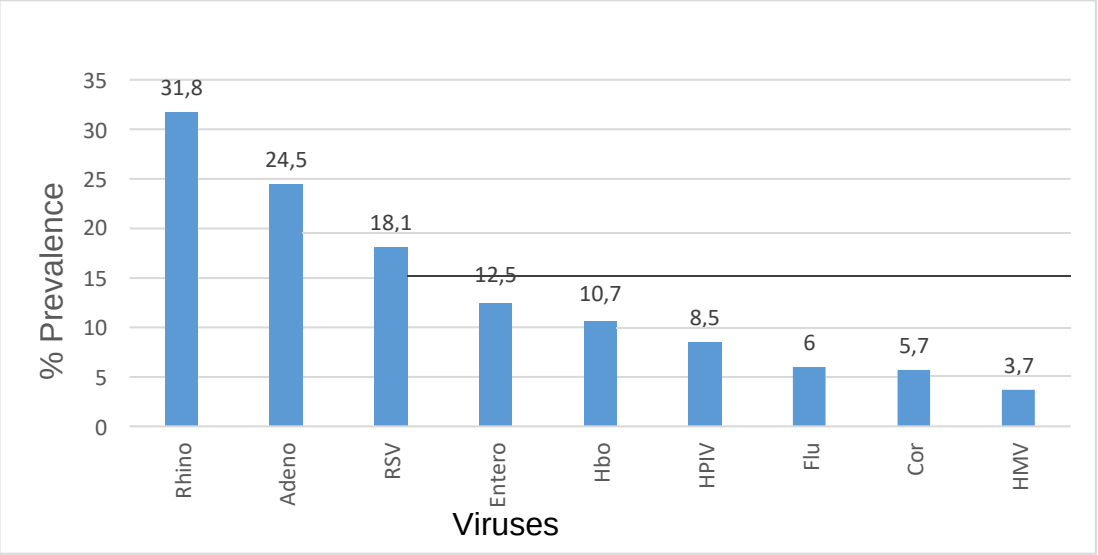


Figure 13 Prevalence of respiratory viruses detected from 2014 to 2017 from clinical samples of infants with SARI. (Rhino = human rhinovirus, Adeno = adenovirus, RSV = respiratory syncytial virus A and B, EV = enterovirus, Boc = bocavirus, HPIV = human parainfluenza 1, 2, 3 and 4, Flu = influenza A and B, Cor = coronavirus OC43, NL63 and 229E, Met = metapneumovirus).

3.1.2. Demographic characteristics

Patient data for 312 respiratory samples positive for HPIV 4 were extracted from the NHLS Virology data capturing system and analyzed. 95 % (295/312) had residual sample available for HPIV 4 subtyping and genetic analysis of the HN gene. More than 98 % (295/312) of the

samples were received from infants and children from Red Cross Children's War Memorial Hospital and 2 % from adults attending other clinics and hospitals. The median patient age was 12 months (range 1 month to 63 years). There were slightly more samples collected for males than for females (ratio 1: 1.3) with males having an overall higher infection (Table 8). Only in 2016 the proportion of HPIV 4 positive cases were higher among females (Table 8).

Table 8 Epidemiological data for HPIV 4 from 2014 to 2017

	2014	2015	2016	2017	Total	Overall %
	n=114	n=72	n=55	n=71	312	312
Female	51 (44,7 %)	24 (33,3 %)	29 (52,7 %)	28 (39,4 %)	132	42.3
Male	63 (55,3%)	48 (66,7%)	26 (47,3 %)	43 (60,3%)	180	57.6
Median age (months)	12	12	17	15	12	

3.1.3. Typing of HPIV 4A and HPIV 4B

Positive patient samples were amplified (nested PCR), using specific novel HN primers for HPIV 4A (733 bp) and HPIV 4B (738 bp) (Figure 14 and Figure 15). Both subtype primers did not cross react with HPIV 1, 2 and 3 as well as with each other (Figure 15). Of the 77 samples tested 51 (64 %) were HPIV 4A and 26 (33 %) HPIV 4B. The positive inner amplicons from the HPIV 4 amplification (section 3.2) (784 bp fragment), were sent for sequencing.

Non-specific banding was observed in Figure 15 and Figure 23. These were clearly distinguishable from the positive controls in that they did not align.

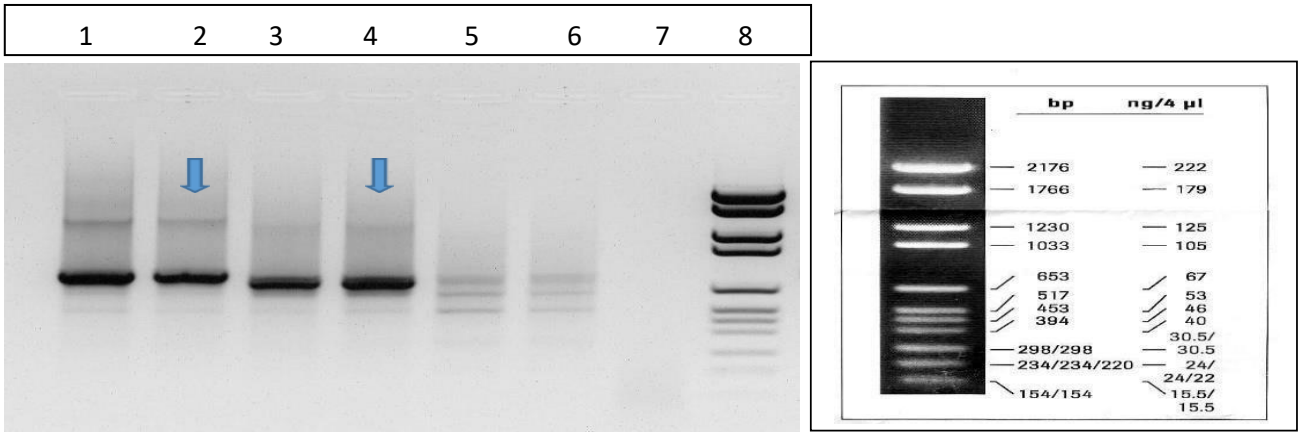


Figure 14 Agarose electrophoresis gel showing HPIV 4A primers (Table 5) had no cross reactivity with HIPV 4B positive samples (positive result in lanes 1 and 3 and negative results in lanes 5 and 6). Positive HPIV 4 and 4A controls - 784 bp and 733 bp respectively (blue arrow). Lane 7 negative control. Lane 8 molecular weight marker.

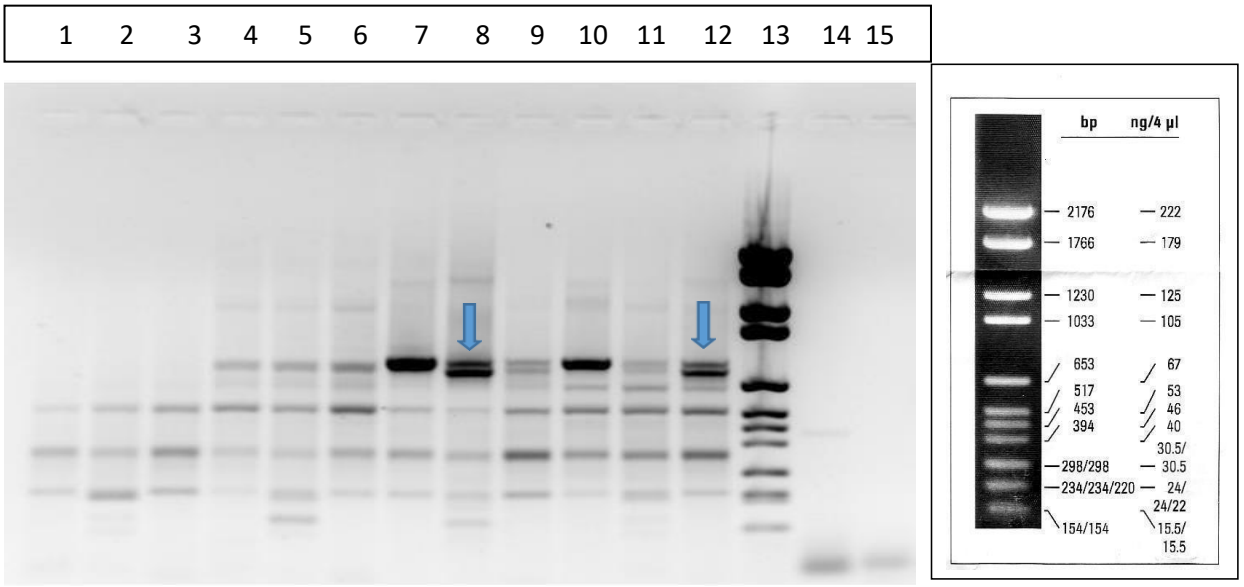


Figure 15 Agarose electrophoresis gel showing both subtypes had no cross reactivity with HPIV 1, 2 and 3 (lane 1 HPIV 1, lane 2 HPIV 2 and lane 3 HPIV 3 are all negative for cross reactivity with HPIV 4A and lane 4 HPIV 1, lane 5 HPIV 2 and lane 6 HPIV 3 are all negative for cross reactivity with HPIV 4B (non-specific binding in lanes 4-6 are higher than HPIV 4 B positive control)). There was also no cross reactivity between both subtypes primers as observed in lane 9 and 11 (non-specific bands are observed which did not align with positive controls). Lane 8 and 12 positive controls for HPIV 4A and HPIV 4B 733 bp and 738 bp respectively (blue arrow). Lane 7 and 10 are HPIV 4 positive controls. Lane 13 is the molecular weight marker and 14 and 15 the negative controls.

3.1.4. Seasonality

Over the 4-year study period (2014-2017) HPIV epidemiological (EPI) curves from children and adults with severe acute respiratory infection (SARI) showed that infections due to HPIV 4 was seasonal, peaking earlier in autumn and winter (March to August) (Figure 16). HPIV 1 and 2 infections did not have a distinct seasonal pattern and occurred sporadically all year round (Figure 17). Whereas HPIV 3 showed a seasonal trend, peaking during autumn to winter (March to July) (Figure 18).

Of the 295 HPIV 4-positive samples that were subtyped by gel based assay, 77 were sequenced, were 51 were typed as A and 26 as B and 218 could not be typed. From the data subtypes co-circulated during each season. Overall, HPIV 4A was the predominant subtype circulating, especially in 2014 and 2017 (Figure 19). In 2015 and 2016 the prevalence of HPIV 4 was lower with HPIV 4A continuing to predominate. (Figure 19).

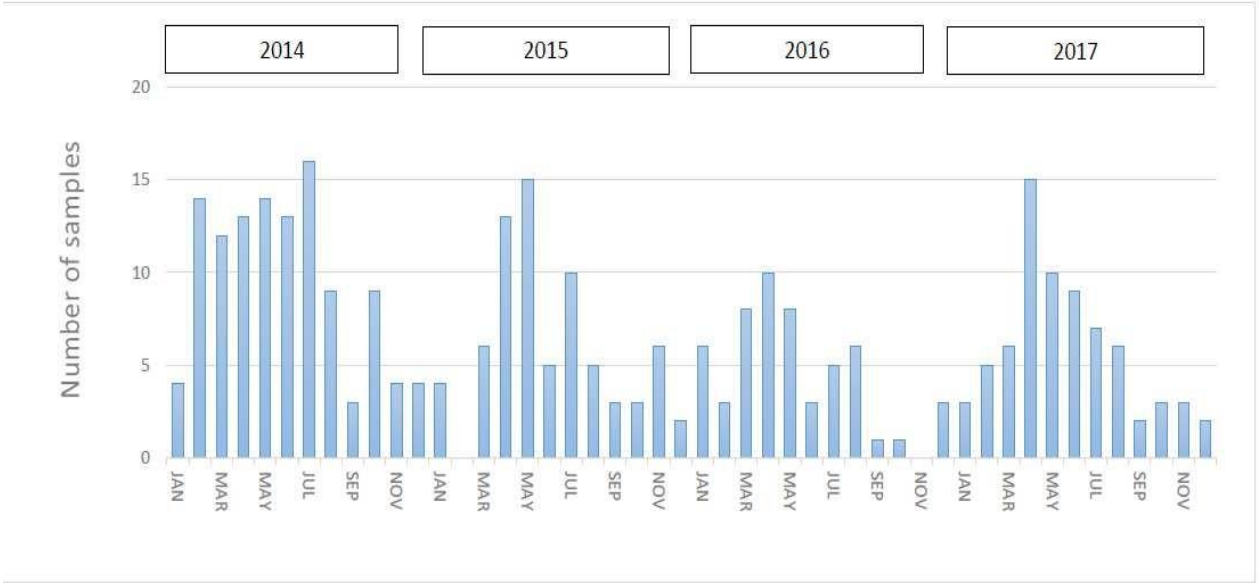


Figure 16 Epidemiological curve showing seasonal HPIV 4 pattern of prevalence with the peak in autumn and early winter (n=312).

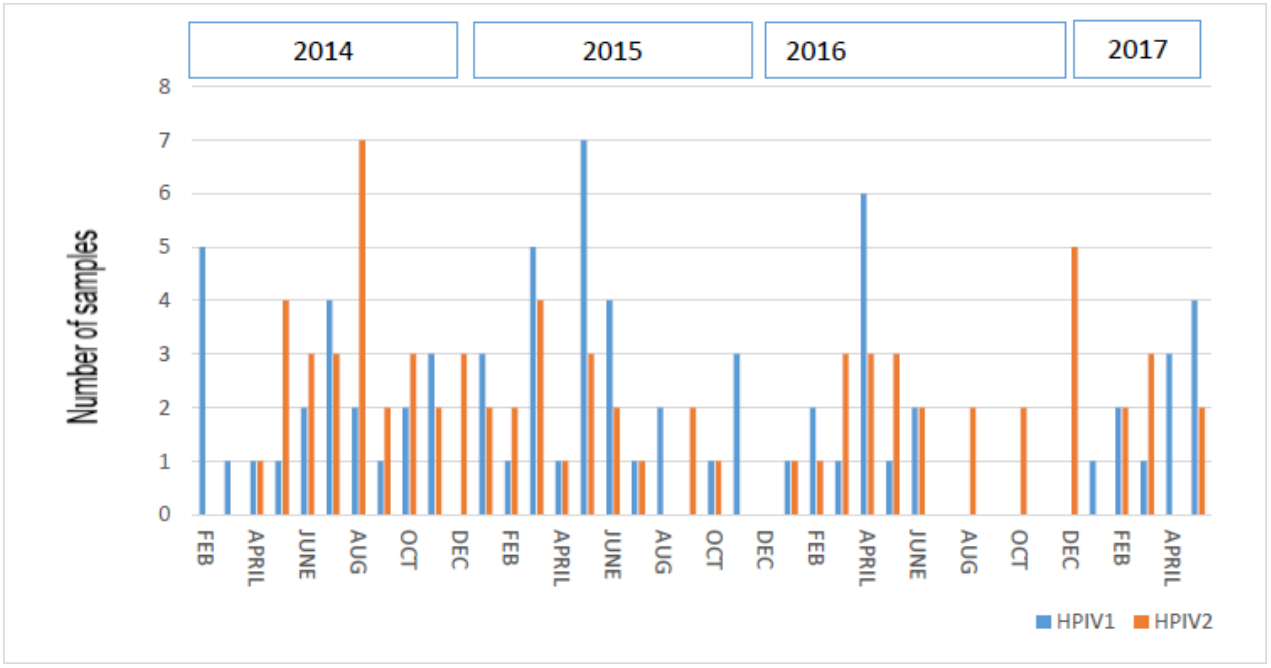


Figure 17 An epidemiological curve of HPIV 1 and HPIV 2 detection over a 4-year period showing a year-round prevalence with no clear seasonal peak (n=266).

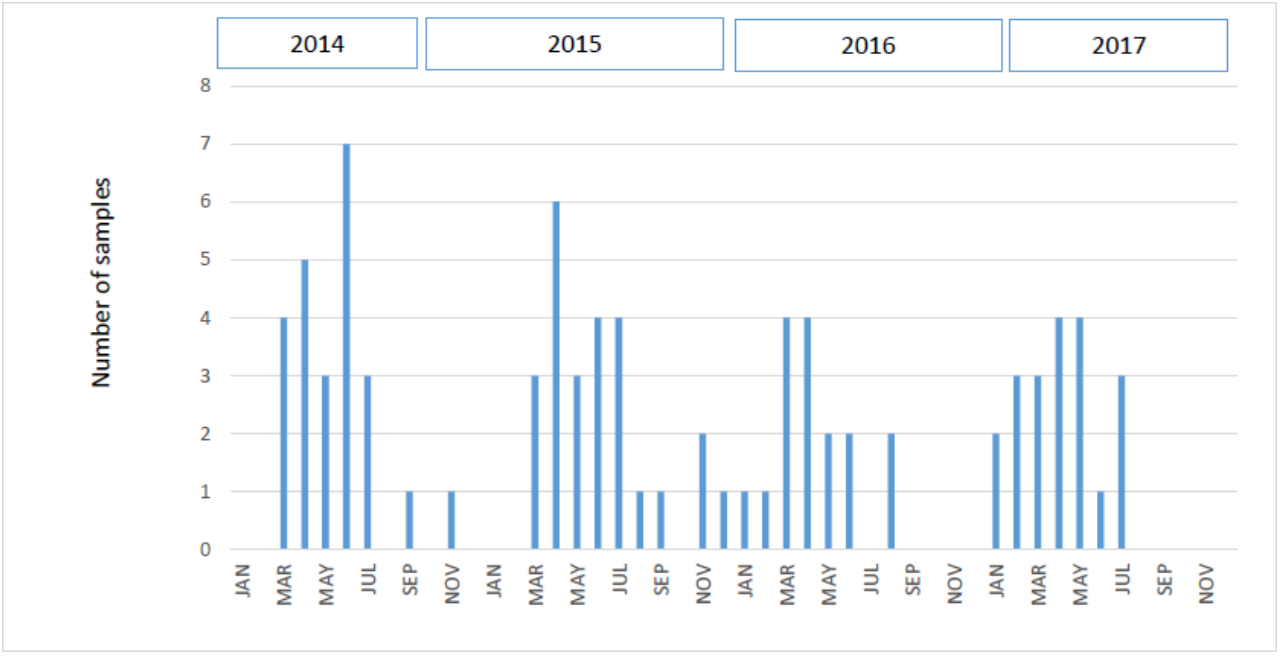


Figure 18 An epidemiological curve of HPIV 3 detection over a 4-year period showing seasonal trend during autumn to winter (n=188).

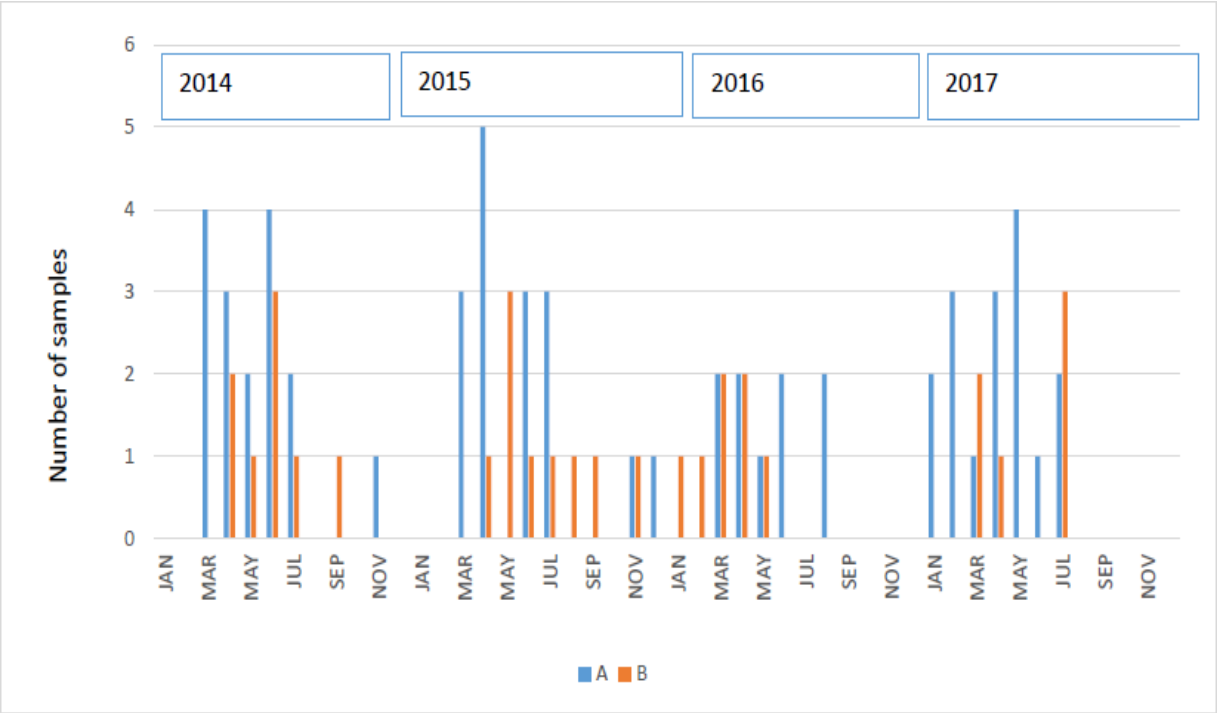


Figure 19 Seasonality of HPIV 4A and HPIV 4B subtypes from 2014 to 2017 in the Western Cape showing co-circulation of the subtypes in all 4 years (n=77).

3.1.5. Co-infections

80 % of positive HPIV 4 samples were co-infected with one or more other respiratory viruses. The most common viruses associated with HPIV 4 co-infection were Adenovirus (27 %), Human Rhinovirus (23 %) and Bocavirus (19 %) (Figure 18). Interestingly, HPIV 1 and HPIV 3 were both able to co-infect HPIV 4 patients, but no co-infections with HPIV 2 were detected (Figure 18).

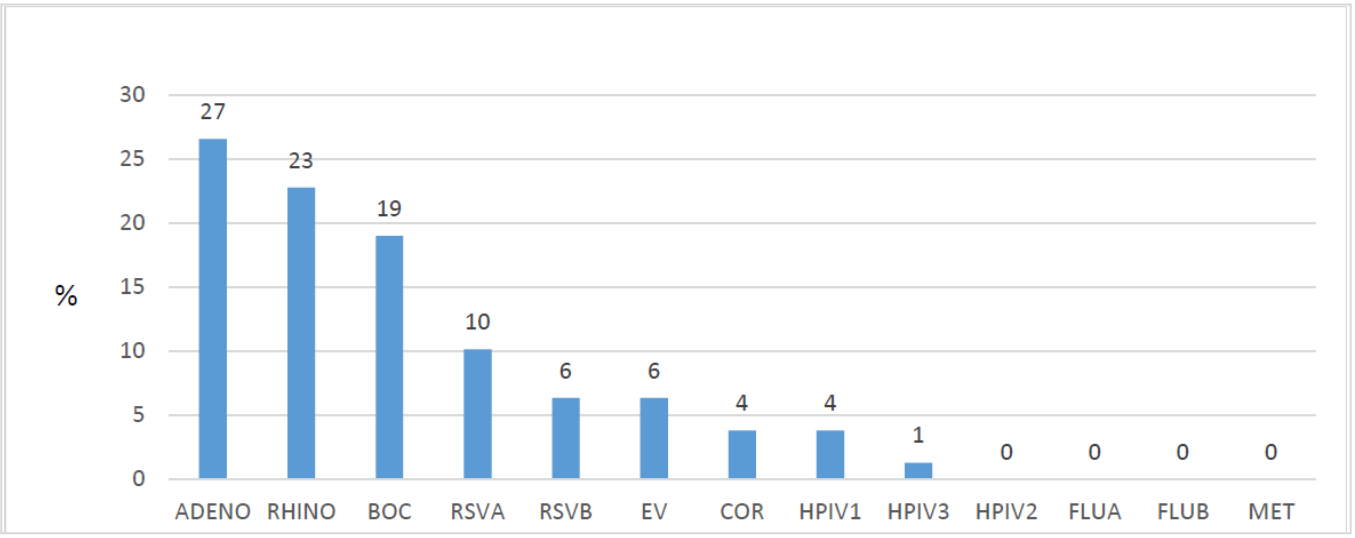


Figure 20 Most prevalent viral co-infections in HPIV 4 infected patients over the study period (n=7456).

3.1.6. Severity of disease

Of those samples sent from hospitalised children, 28 % (87/312) were from patients in ICU. Of these 46 % (40/87) had HPIV 4 as a single infection, indicating that HPIV 4 was associated with severe lower respiratory tract infection in hospitalised infants (Figure 19). There was no apparent difference in clinical severity associated in samples that were subtyped, with HPIV 4A at 55 % (12/22) and HPIV 4B at 45 % (10/22).

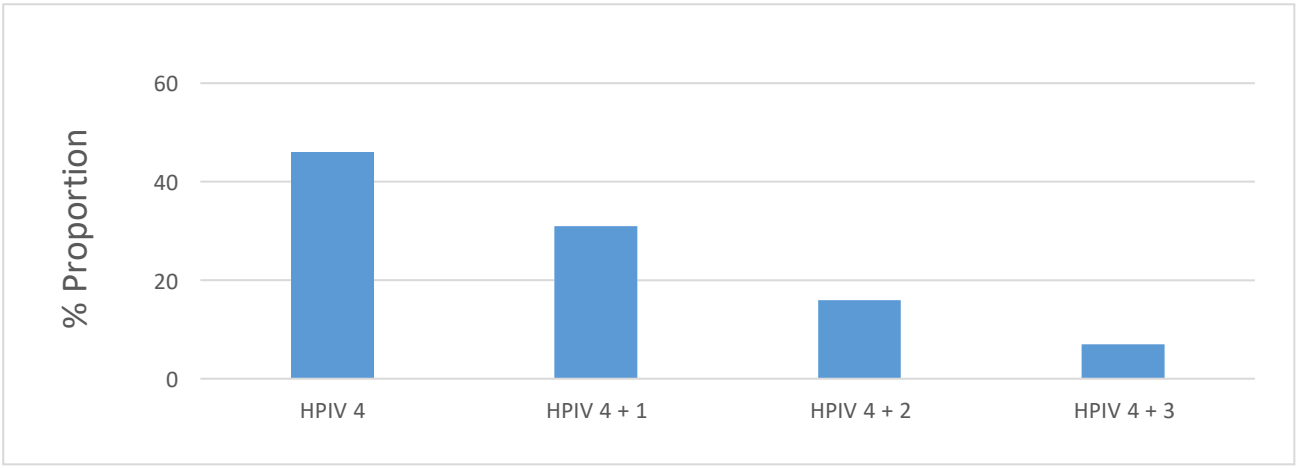


Figure 21 The proportion of samples from ICU patients with HPIV 4 as a mono-infection, or with one or more co-pathogens.

3.3. Genetic analysis of HPIV 4

Positive HPIV 4 patient samples (n=312), as detected by RV screening, were amplified (nested PCR) using novel primers targeting the HN region of HPIV 4. Of these, 77 (25 %) were positive for HN gene of HPIV 4, producing an inner band of 784 bp on agarose gel electrophoresis (Figure 20). Only one patient sample had high enough viral loads to produce an outer PCR product of 796 bp (Figure 20 Lane 2). Further, these HPIV 4 primers did not cross react with any of the other three HPIV positive samples (Figure 23). The weak bands seen were non-specific and did not align with the positive control.

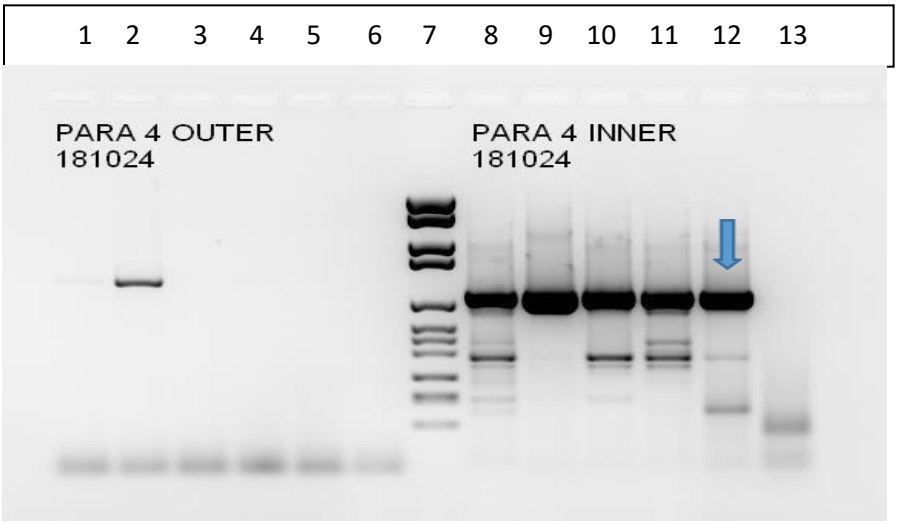


Figure 22 Gel electrophoresis gel showing positive HPIV 4 patient samples with an inner fragment of 784 bp. Lanes 1-6 outer reaction, lanes 9-13 nested reaction with positive control (lane 5 & 12), negative control (lane 6 & 13). Lane 7 molecular weight marker. Positive control (blue arrow).

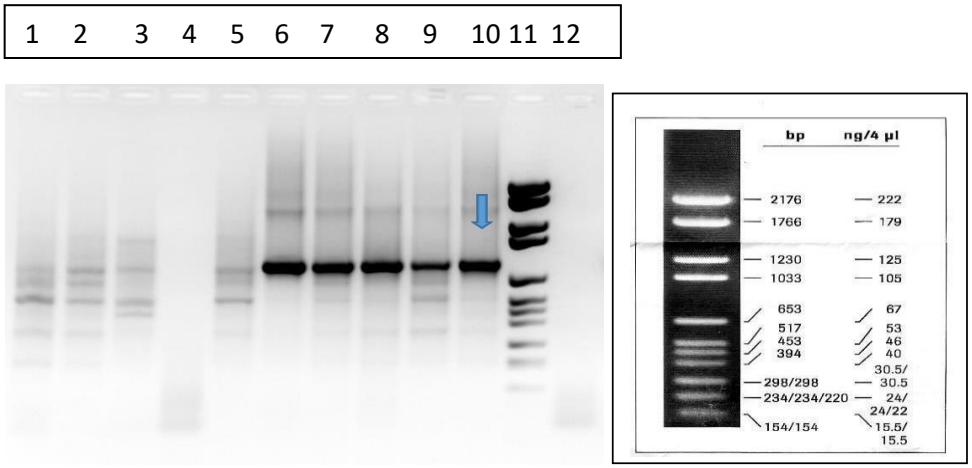
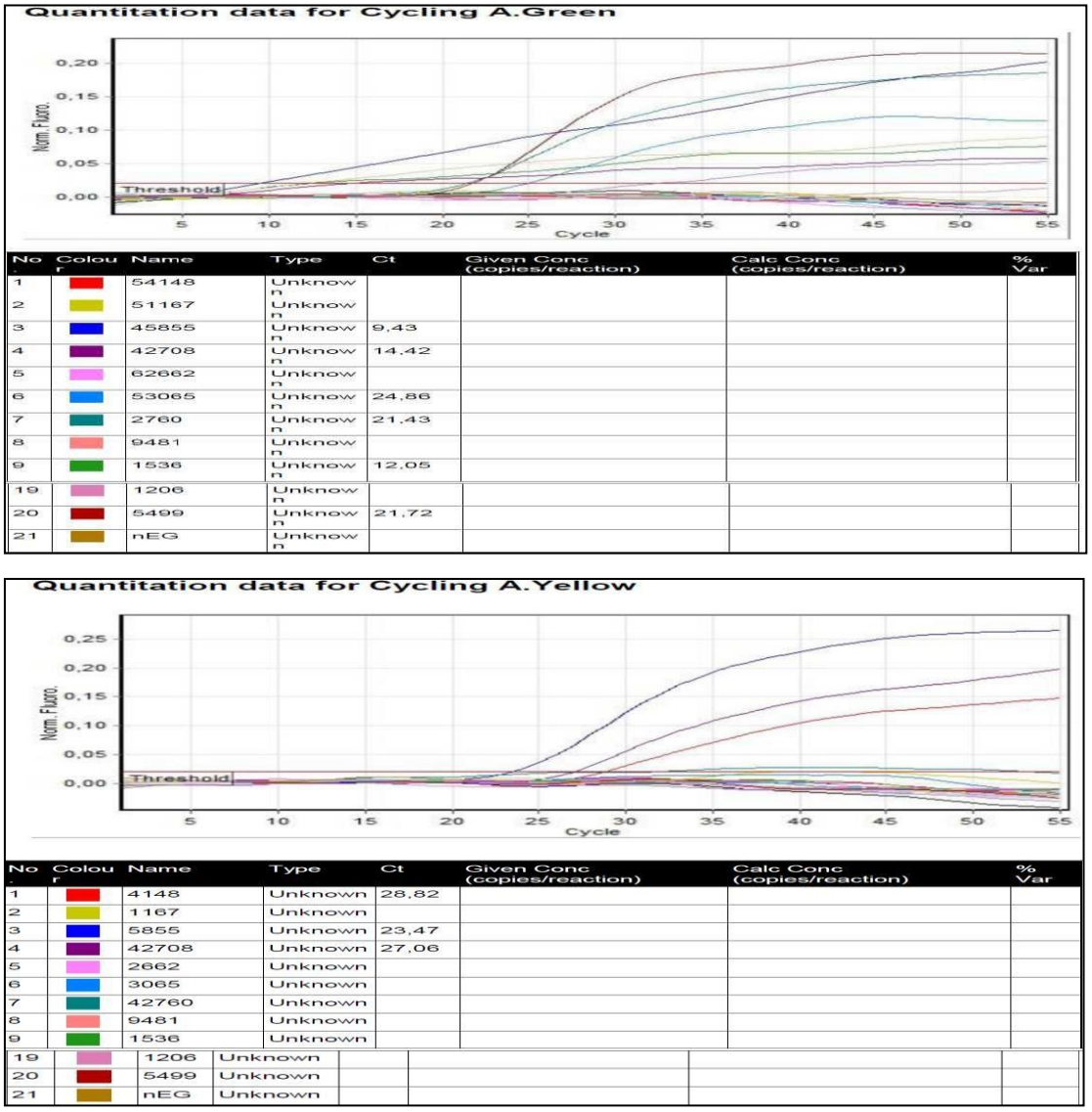


Figure 23 Agarose electrophoresis gel showing HPIV 4 sequences (Table 5) having no cross reactivity with HIPV 1 (lane 1), HPIV 2 (lane 2) and HPIV 3 (lane 3) (non-specific bands are observed which do not align with positive controls). Lanes 6-9 positive control 784 bp (blue arrow) (lane 10), negative controls (lane 4, 5 and 12). Lane 11 molecular weight marker.

3.4. Real-time PCR for rapid subtyping of clinical samples

A differentiation of the HPIV 4A and HPIV 4B was observed, when 46 positive clinical HPIV 4 samples were subtyped, using rapid real-time PCR (Figure 24). The green channel (FAM)

detected HPIV 4A and the yellow channel (HEX) detected HPIV 4B. Where the green channel was positive for HPIV 4A the yellow channel was negative for HPIV 4B with the same outcome for the HPIV 4B results (Figure 24). In Figure 24 the green channel no3, the result was negative as the curve is straight (non-specific fluorescence) and not sigmoidal whereas the result in yellow channel is positive as a sigmoidal cure is observed in the same position.



No.	Colour	Name	Type	Ct	Given Conc (copies/reaction)	Calc Conc (copies/reaction)	% Var
1		54148	Unknown				
2		51167	Unknown				
3		45855	Unknown	9,43			
4		42708	Unknown	14,42			
5		62662	Unknown				
6		53065	Unknown	24,86			
7		2760	Unknown	21,43			
8		9481	Unknown				
9		1536	Unknown	12,05			
19		1206	Unknown				
20		5499	Unknown	21,72			
21		nEG	Unknown				

No.	Colour	Name	Type	Ct	Given Conc (copies/reaction)	Calc Conc (copies/reaction)	% Var
1		4148	Unknown	28,82			
2		1167	Unknown				
3		5855	Unknown	23,47			
4		42708	Unknown	27,06			
5		2662	Unknown				
6		3065	Unknown				
7		42760	Unknown				
8		9481	Unknown				
9		1536	Unknown				
19		1206	Unknown				
20		5499	Unknown				
21		nEG	Unknown				

Figure 24 Real-time PCR results for green and yellow channels showing differentiation between HPIV 4A and HPIV 4B.

Table 9 Clinical sensitivity and specificity of the HPIV 4

		IN-HOUSE PCR		
		POS	NEG	TOTAL
REAL-TIME RESULTS	POS	TP=49	FP=0	49
	NEG	FN=1	TN=16	17
	TOTAL	50	16	66

(TP= true positive, TN=true negative, FP=false positive, FN=false negative)

Reproducibility showed a 100 % correlation between the qPCR assay and gel electrophoresis run in parallel on consecutive days (Table 10 and 11). The highest positive dilution was 10^{-4} giving an LOD of log 4 (Table 11).

Table 10 Sensitivity and specificity of HPIV 4 EQA samples

EQA			
SAMPLE #	Expected result	Real-time RESULT	In-house PCR RESULT
PINFRNA17S-08	POS	POS	POS
PINFRNA16S-02	POS	POS	POS
PINFRNA14S-05	POS	POS	POS
PINFRNA14S-06	NEG	NEG	NEG

Table 11 Six 10-fold dilutions used to show reproducibility and LOD from 2 EQA samples.

6 x 10 fold dilutions	CT day 1 P gene	CT day 2 P gene	GE day 1 HN gene	GE day 2 HN gene
19 neat	21,89	21,43	POS	POS
19 10^{-1}	24,23	24,76	POS	POS
19 10^{-2}	27,49	27,43	POS	POS
19 10^{-3}	30,21	30,16	POS	POS
19 10^{-4}	32,79	32,31	POS	POS
19 10^{-5}	NEG	NEG	NEG	NEG
98 neat	23,81	23,67	POS	POS
98 10^{-1}	25,65	25,63	POS	POS
98 10^{-2}	29	28,29	POS	POS
98 10^{-3}	31,86	31,58	POS	POS
98 10^{-4}	33,25	34,96	POS	POS
98 10^{-5}	NEG	NEG	NEG	NEG

CT = crossing threshold, GE = gel electrophoresis

To validate the real-time PCR, 13 positive gel-based subtyping samples and 7 negatives were used. These results correlated (Table 10).

Table 12 Real-time PCR results compared to gel-based PCR /sequencing results (n=20).

Number	Gel electrophoresis/ sequencing HPIV 4A or 4B +	Real-time PCR HPIV 4A or 4B +
SA4148	4B	4B
SA2760	4A	4A
SA3065	4A	4A
SA5855	4B	4B
SA9481	4A	4A
SA4820	4A	4A
SA1536	4A	4A
SA6198	4A	4A
SA5499	4A	4A
SA4420	4A	4A
SCH6577	4A	4A
SG5855	4B	4B
SG42708	4B	4B
SA5230	N	N
SCH9380	N	N
SCH9104	N	N
SG4470	N	N
SCH2171	N	N
TD5866	N	N
SCH5310	N	N

The assay was optimized using HPIV 4 positive clinical samples and EQA samples. The best performance was obtained using an eluate volume of 4 µl in the master mix (Table 7) . The assay had a sensitivity and specificity of 98 % and 100 % respectively, thus meeting validation criteria, even though it was slightly less sensitive than the in-house nested PCR assay (Table 9 and Table 10) . This was probably due to low copy numbers, where qPCR only amplifies for 45 cycles while nested reactions usually have between 70 - 80 cycles of amplification. The reproducibility and LOD also met validation criteria with 100 % reproducibility and LOD of log 4 (Table 11).

3.5. Phylogenetic analysis

The sequencing data received was analysed using BLAST® (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine the viral sequence identity.

3.5.1. Sequencing of HN gene fragment to identify HPIV 4 subtypes

A total of 80/295 samples amplified gave a visible PCR product that were sequenced. Of these, 77 were successfully assembled HN gene sequence contigs. Each assembled sequence was blasted using BLAST® (https://blast.ncbi.nlm.nih.gov/Blast.cgi). A total of 77

HPIV 4 HN specific sequences were identified (51 were HPIV 4A and 26 HPIV 4B) and 3 were non-specific sequences of human origin. HPIV 4A and HPIV 4B sequences separated into 2 distinct lineages with 87 % bootstrap support (Figure 25).

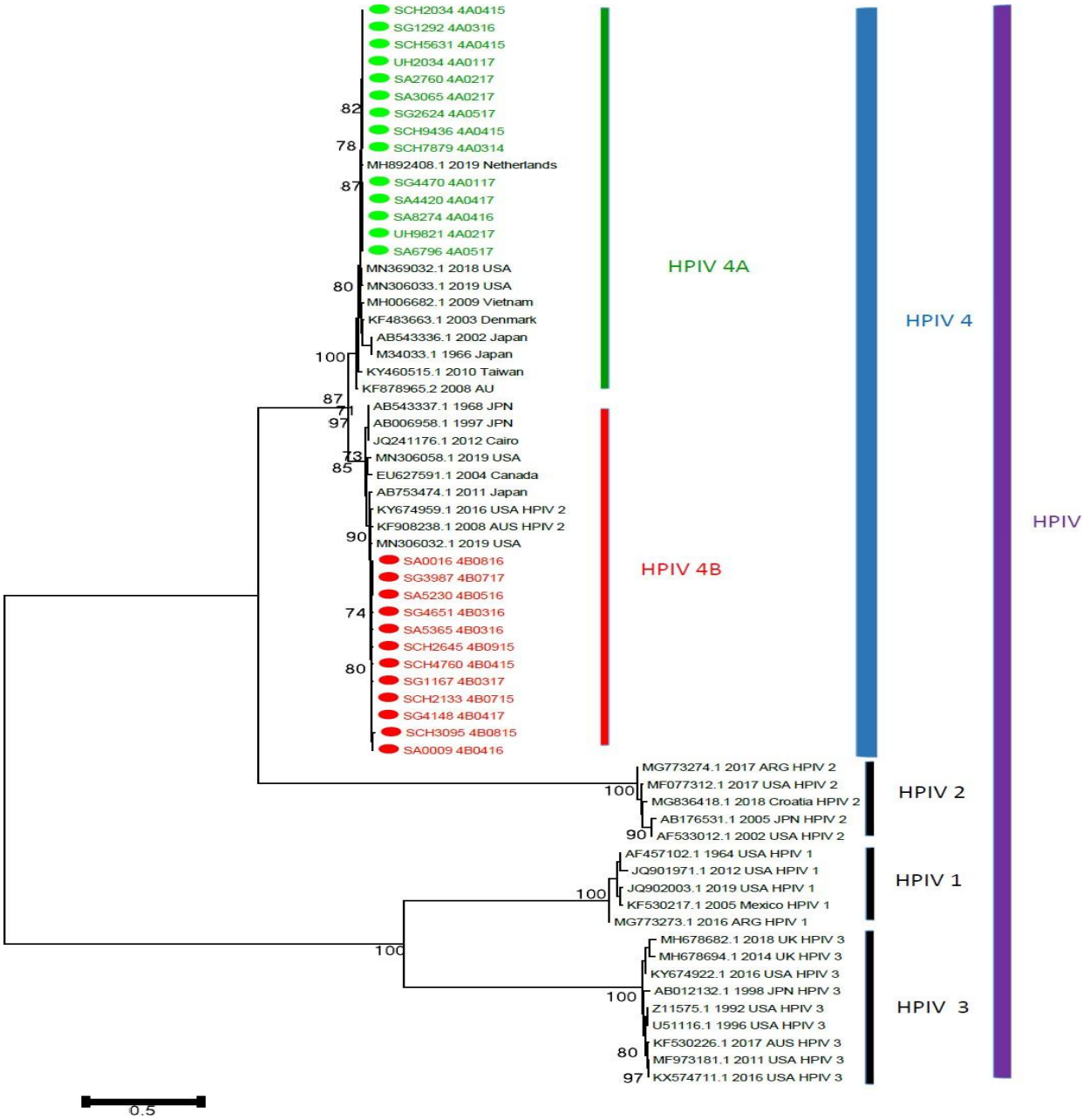


Figure 25 Neighbour-joining phylogenetic tree of HPIV 4 partial HN gene sequences from South Africa patients (n=77) and reference sequences from Genbank (black). In green are South African sequences that group with HPIV 4A and in red those that group with HPIV 4B. A neighbour-joining phylogenetic tree was constructed and showed distinct groupings with high bootstrap values for all four species of HPIV, including both HPIV 4 subtypes (Figure 25). Not many reference sequences grouped with the South African HPIV 4s indicating that

these subtypes are unique.

More detailed phylogenetic trees for HPIV 4A and HPIV 4B were constructed (Figure 26 and Figure 27), where a few observations were revealed. When analyzing the bases on the electropherogram there where definite changes in the bases which could lead to the long branches observed in the phylogenetic trees. In the HPIV 4A tree (Figure 26) 4 general clusters were observed, however the bootstrap values were below 70%. Cluster 1 comprised only South African sequences while Clusters 2 - 4 included reference sequences as well. Of the 14, 2014 samples sequenced the majority (78.5 %) grouped with Cluster 1 and the remaining with Cluster 2 (14.4 %) and Cluster 4 (7.1 %). Clusters 2, 3 and 4 included single reference sequences from Netherlands, Taiwan and Japan respectively.

Of note are 2 sample groupings (highlighted in purple and red in Figure 26) with bootstrap values of 89 % and 82% respectively. When examining the sample demographics, it was found that the samples from the "purple" grouping were from the same patient taken a day apart, and no differences in the viral genome were noted (Figure 26 & Figure 28). The second "red" group were from unrelated patients admitted a month apart to the same hospital, but to different wards (Figure 26). Another grouping with a high bootstrap value of 88 % in Cluster 2 comprised 3 samples. Two of the samples were from siblings aged 2 years and 4 months, respectively. They had been admitted to hospital at the same time, indicating possible transmission within the home setting. The third case within this grouping was apparently unrelated and admitted to a different hospital (Figure 28).

The HPIV 4B tree (Figure 27) revealed 2 separate clusters. The first cluster (Cluster 1) comprised samples from all 4 years of the study group while no 2014 sequences were located within the second cluster (Cluster 2). Two reference sequences from Malaysia and Japan were interspersed within the Cluster 1, whereas no reference sequences grouped with Cluster 2, indicating a possible unique South African cluster. Also observed in phylogenetic tree for HPIV 4B, all the samples with the 12 nucleotide changes (Figure 25) had the amino acid change at G198R which all grouped in cluster 2 (Figure 27) In Cluster 1, two closely related sequences formed a grouping with a bootstrap value of 94 %, however these were from unrelated patients from two different geographic locations (George and Cape Town).

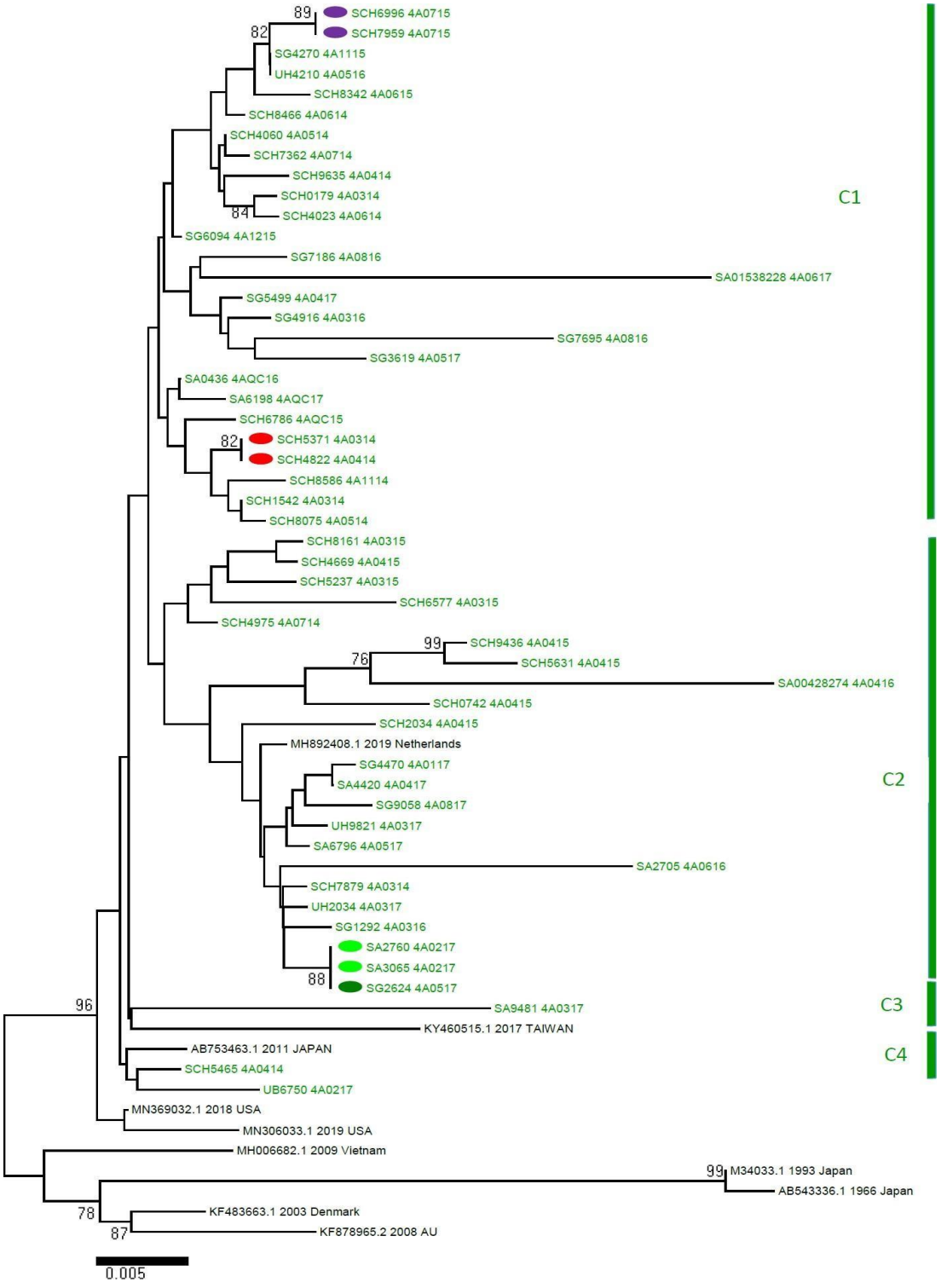


Figure 26 Neighbour-joining phylogenetic tree of HPIV 4A sequences and reference sequences from GenBank. Sequences in green HPIV 4A study patient samples (n=56). Only a bootstrap values above 70% are indicated. All sequences in black were obtained from GenBank. Clusters 1, 2, 3 and 4 (C1-4) are indicated.

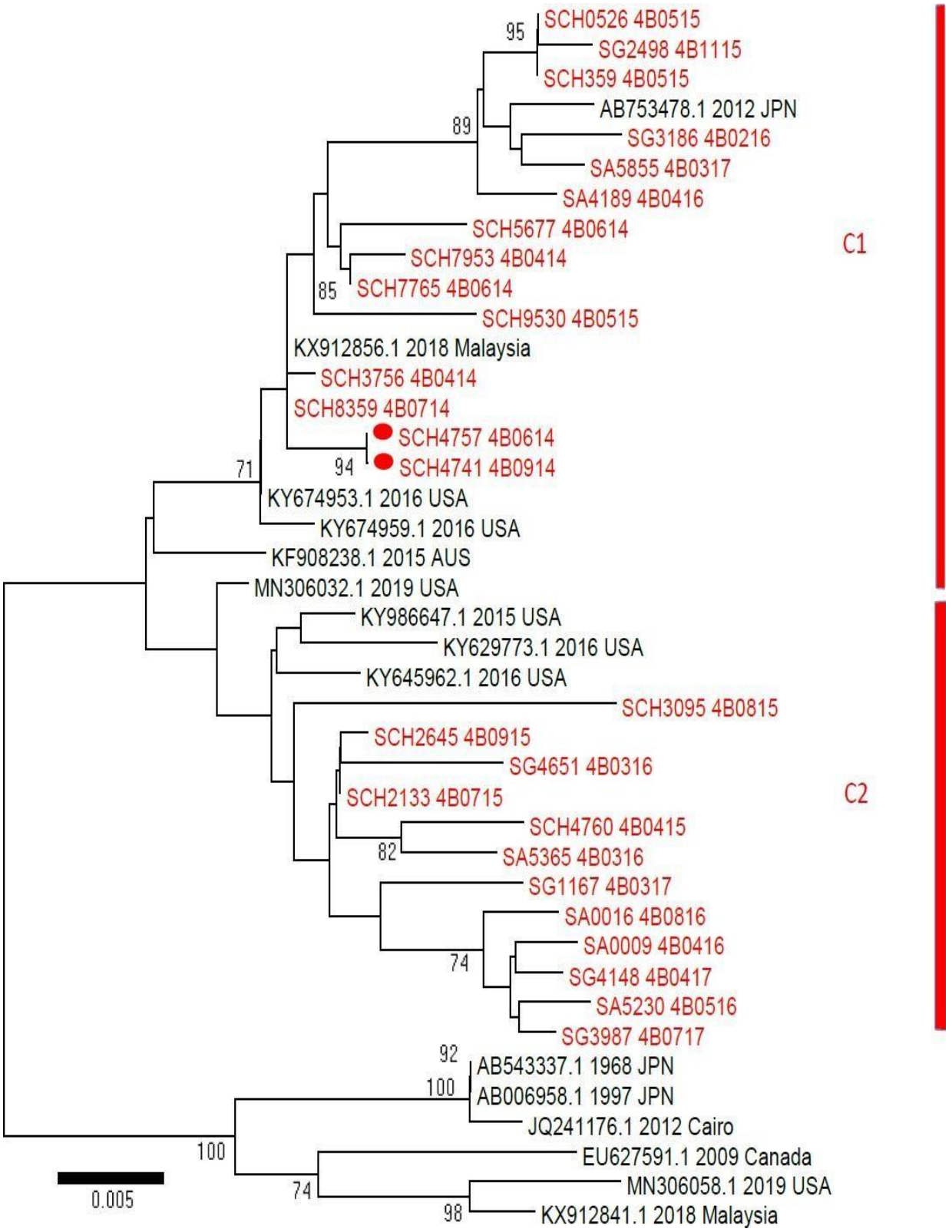


Figure 27 Neighbour-joining tree of HPIV 4B sequences and reference sequences from GenBank. Sequences in red HPIV 4B study patient samples (n=21). Only a bootstrap of 70% are indicated. All sequences in black were obtained from GenBank.

Highlighter plot (HLP) of HPIV 4A nucleotide sequences was drawn (Figure 28) using the earliest sequence as the comparator, SCH0179 March 2014. In Cluster 2 (17 samples) there was 1 sample from 2014 (SCH7879) and the rest from 2015 to 2017. This suggests that Specimen SCH7879 was probably an ancestor to cluster 1 with a year between this sample and the next, (SCH5237, 2015) which was similar to it (Figure 28 blue dots). Over time there was an accumulation of unique mutations compared to the early reference sequence, increasing from 14 nucleotide changes in 2014 to 32 in 2017 (Figure 29). A nucleotide change at location T54C (marked with blue arrow on HLP) was always accompanied with a change at position A482G (marked with orange arrow on HLP) (Figure 28), indicating possible co-evolution. Further, certain nucleotide changes became more common over time e.g. T42C, T240C, A427G, A446G, A461G, G502A, C520T, T553C. A very conserved region in the HN gene between location 205 and 390 was also observed (Figure 28 red bracket).

Using the same comparator as above, the HPIV 4A amino acid HLP showed three main changes (Figure 30). These occurred at amino acid positions M85I, N161D and G164E (corresponding to M292I, N368D and G371E in complete HN protein respectively). The N161D change resulted in the loss of a glycosylation site (Figure 29, seen in samples with*). The N161D and G164E were more frequently detected in samples from 2016 and 2017 than 2014/2015 (32% versus 12% of samples).

The HPIV 4B nucleotide HLP revealed a similar pattern to HPIV 4A. (Figure 31) Over time more mismatches to the reference sequence were found, increasing from 4% in 2014 to 40% in 2017(Figure 32). At location T525 the earliest sample following the comparator had a nucleotide change, T525C, which remained present in all subsequent samples over the 4-year study. A pattern of 10 nucleotide changes occurred in 12 samples starting at SCH4760 (*) with the exception for SCH3095 with 2 unchanged nucleotides (**) (Figure 30).

The amino acid HLP of HPIV 4B showed one major mismatch occurring at location G198R (corresponding to G405R in complete HN protein), changing it from a neutral to positively charged residue. The mutation was first detected in 2015 and increased in dominance over the following years of the study (Figure 33, blue arrow).

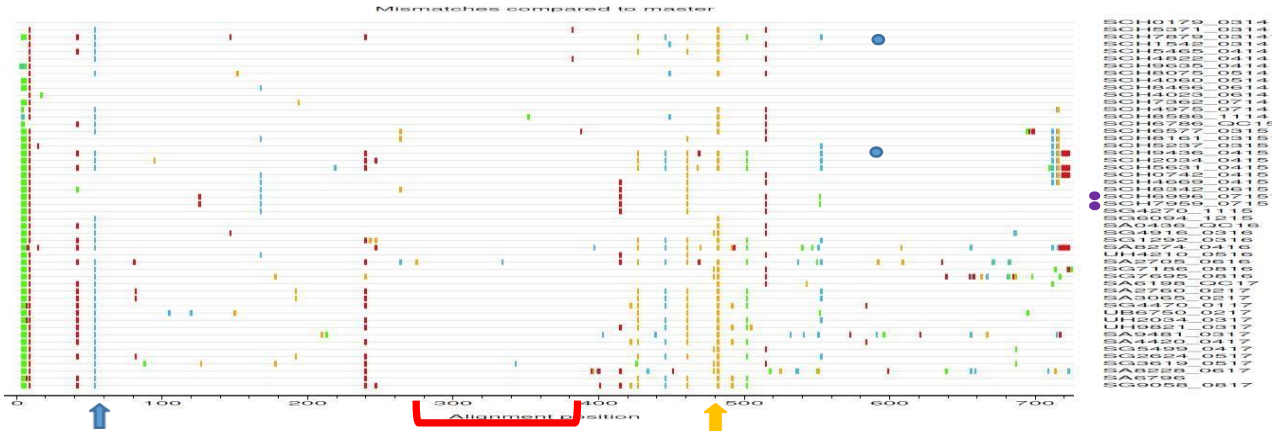


Figure 28 Highlighter plot of only South African HPIV 4A sequences showing nucleotide changes over time with SCH0179 used as the comparator. SCH7879 ancestor to SCH5237 (blue dots). SCH6996 and SCH7959 are the same patient a day apart (purple dots). T54C (marked with blue arrow) was always accompanied with a change at position A482G (marked with orange arrow).

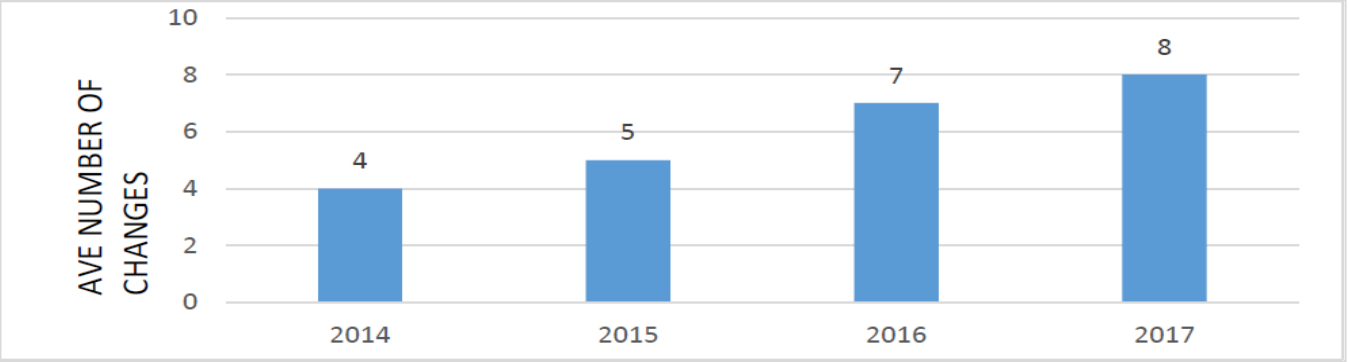


Figure 29 Average number of nucleotide changes in HPIV 4A over time, showing 2017 with the most changes with sequence SCH0179 (March 2014) used as comparator.

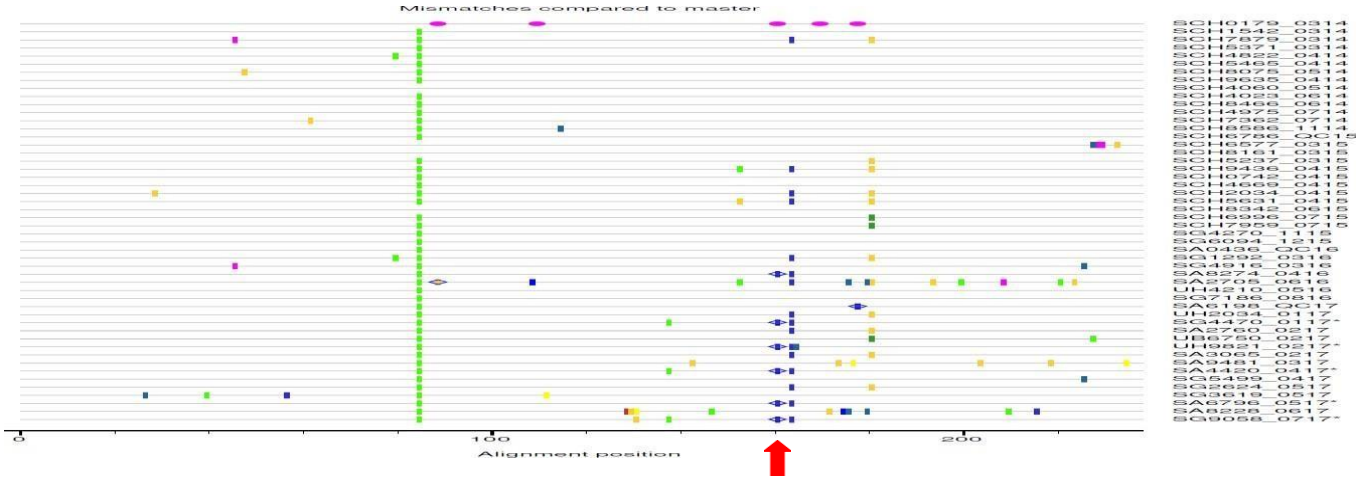


Figure 30 Highlighter plot of only South African HPIV 4A sequences showing synonymous and nonsynonymous amino acid changes over time with SCH0179 used as comparator. * samples with loss of glycosylation site at N161D (red arrow).

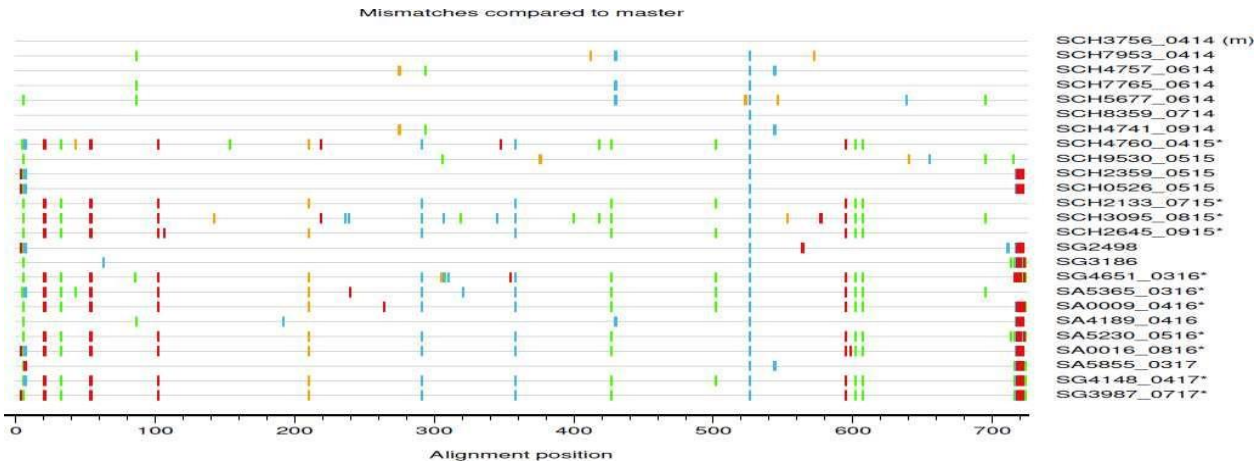


Figure 31 Highlighter plot of only South African HPIV 4B sequences showing nucleotide changes over time with SCH3756 (April 2014) used as the comparator. 10 nucleotide changes occurred in 12 samples starting at SCH4760 (*) with the exception for SCH3095 with 2 unchanged nucleotides (**)

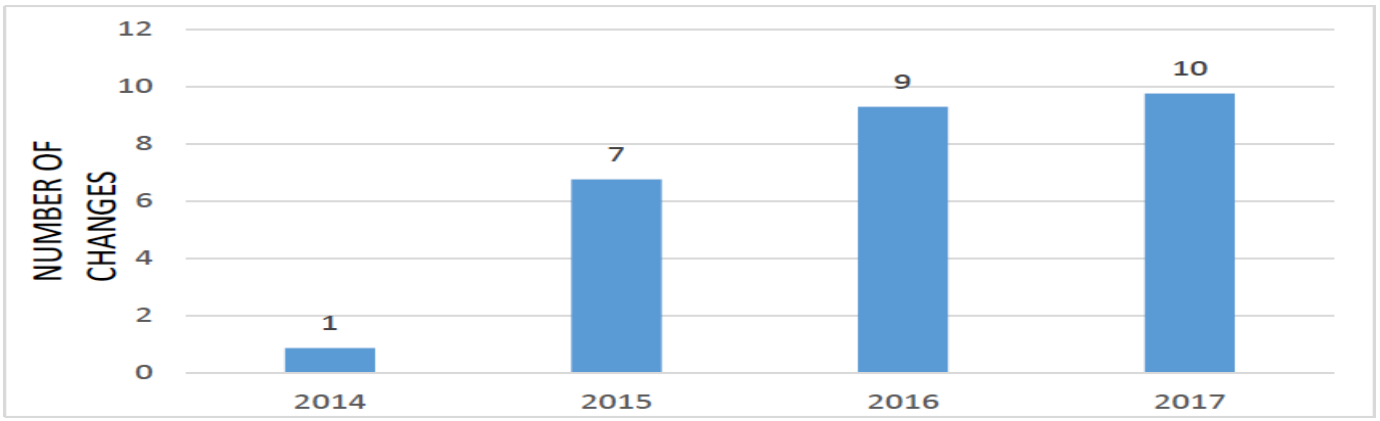


Figure 32 An average number of nucleotide changes in HPIV 4B over time, showing 2017 with the most changes with SCH3756 used as the comparator.

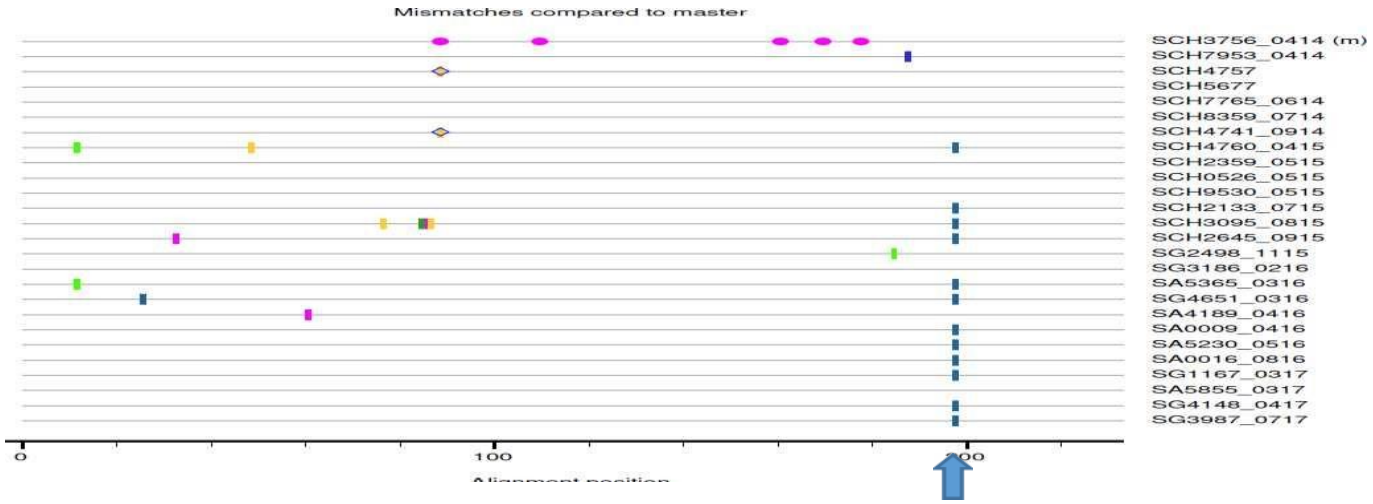


Figure 33 Highlighter plot of only South African HPIV 4B showing synonymous and non-synonymous amino acid changes over time with SCH3756 used as the comparator with one major mismatch occurring at location G198R (blue arrow).

Chapter 4 Discussion

Chapter 4 Discussion

With the wider use of molecular diagnostic assays that are able to detect this virus in clinical samples, the importance of HPIV4 as a cause of respiratory tract infection in infants has received better recognition. However, very little epidemiological and molecular information is available on these infections in South Africa where no studies have been published. This study describes the molecular epidemiology of HPIV 4 in hospitalised infants and children with lower respiratory tract infections from Red Cross Children's War Memorial Hospital and from adults attending other clinics and hospitals in the Western Cape province, South Africa, 2014 -2017. It highlights the burden of disease and the role this virus plays as a cause of severe acute respiratory infection.

4.1. HPIV 4 epidemiology and seasonality

Of all respiratory viruses detected HPIVs were the 6th most viruses commonly detected after rhinovirus, adenovirus, RSV, enterovirus and bocavirus (Figure 13) which is similar to results reported by Lui et al. (2013). However, HPIV 4 was detected at frequencies similar to metapneumovirus, common human coronaviruses and influenza viruses.

The median age of HPIV 4 - positive patients was 12 months which coincided with findings of many other studies. (Zhao et al., 2017, Frost et al., 2013, Billaud et al. 2005). This reflects the fact that these viruses are very prevalent, and infants are likely to be exposed in the first year of life when infection can be more severe, requiring hospitalization. The median age is older than for some other respiratory viruses such as RSV. This may be due to the presence of maternal antibodies that provide protection from infection in the first 6 months of life in infants. Most of the samples collected were from a children's hospital. Males had an overall higher HPIV 4 infection. This has been observed for other respiratory viruses as well and in other studies (Shi et al., 2020; Liu et al., 2013; Weinberg et al. 2009).

Overall, HPIV 4 was the most prevalent HPIV detected in diagnostic samples over the 4-year study period, accounting for almost half (47.5 %, 312/633) of positive HPIV detections. This contrasts with findings in other parts of the world, where HPIV 3 followed by HPIV 1 are more predominant (Thomazelli et al., 2017; Zhao et al., 2017; Lui et al., 2013; Ren et al.,

2011). The prevalence of HPIV 4 increased over the study period from 40 % (103/257, 2014) to 57 % (32/56, 2017) (Figure 12). There could be many reasons for this increase. One that molecular detection techniques are more sensitive and together with better quality sample collection, prevalence improved. The results may also be skewed as most samples were from children that were hospitalised, where nosocomial spread could occur. Over the same period, the prevalence of HPIV 3 decreased (Figure 12). The prevalence of HPIV 1 and HPIV 2 infections varied, showing a biennial pattern, with higher HPIV 1 detection in odd-numbered years and higher HPIV 2 prevalence during even-numbered years (Figure 12). These findings were similar to the biennial prevalence in other studies (Pawelczyl & Kowalski, 2017; Fry et al., 2006).

Respiratory infections due to HPIV 4 were also seasonal, peaking in autumn and mid-winter (March to August) (Figure 16). This was in agreement with most countries worldwide (Horton et al., 2017; Zhao et al., 2017; Abedi et al., 2016; Yano et al., 2014; de Mello Freitas, 2013; Frost et al., 2013; Lau et al., 2009; Vachon et al., 2006; Billaud et al., 2005) except in Brazil and Guangdong province in China where HPIV 4 occurred during late winter to spring, similar to HPIV 3 (Thomazelli et al. 2017 and Liu et al., 2013). Over the 4-year study period (2014-2017) HPIV 1 and 2 infections occurred sporadically all year round (Figure 17, whereas, both HPIV 3 and 4 showed a distinct seasonal trend. HPIV 3 peaked during autumn to winter (March to July) (Figure 18) which was similar to the US and England seasonal trend (Zhao et al., 2017; Fry et al., 2006).

Both HPIV 4A and HPIV 4B co-circulated at varying frequencies during each season (2014 - 2017) (Figure 19). Overall, HPIV 4A was the predominant subtype even for years when the annual prevalence of HPIV 4 was lower. These results were consistent with a 1 year study in Hong Kong, 2004 (Lau et al., 2009).

4.2. Co-infections and disease severity

71 % of patients who were positive for HPIV 4 were co-infected with one or more additional respiratory virus. This is higher than co-infection rates observed in Ren et al. (2011), but similar to results in Hasman et al. (2009). The most common viruses associated with HPIV 4 co-infection were adenovirus (27 %), human rhinovirus (23 %), bocavirus (19 %) and RSVA and RSVB (10 % and 6 % respectively) (Figure 20). This is likely due to adenovirus, rhinovirus and bocavirus having no distinct seasonality and are the most prevalent respiratory viruses

in young children. These results were similar to other studies (Linster et al., 2018; Frost et al., 2013; Lui et al., 2013; Ren et al., 2011). Less commonly, co-infections with other HPIVs was seen. In particular, HPIV 1 and HPIV 3 were both able to co-infect patients with HPIV 4, but no co-infections with HPIV 2 were detected (Figure 20). It is possible that viral interference or immune cross reactivity could account for this being a less common occurrence, or it could simply be that HPIV 2 was an uncommon infection.

The clinical spectrum and disease severity of HPIV 4 infection has been reported to be very similar to that caused by HPIV 3 (Zhang et al., 2014; Frost et al., 2013). In this study, of those samples sent from hospitalised children, between ¼ to 1/3 were from patients in ICU. Of these almost half (46 %) had HPIV 4 as a single infection, indicating that HPIV 4 alone could be responsible for severe lower respiratory tract infection in hospitalised infants (Figure 21). There was no apparent difference in clinical severity associated with single subtype A or B infections which was also observed by Lau, et al., 2009, Hong Kong.

4.3. Real-time versus traditional PCR

Rapid, sensitive and specific molecular methods are continuously evolving, aiding in better turnaround times thus improving patient care and treatment. To provide a rapid subtyping assay for HPIV 4, a real-time multiplex assay targeting the HPIV 4 P gene (Wang et al., 2012) was designed and used to monitor clinical isolates for epidemiological purposes. The qPCR assay is much easier to set-up together with a quicker TAT, makes it feasible alternative assay to test clinical samples.

In conclusion the qPCR is equivalent to the nested in-house PCR for rapid monitoring of clinical isolates for epidemiological purposes in HPIV 4A and HPIV 4B positive samples.

4.4. Genetic analysis

HPIV 4 has two major surface glycoproteins hemagglutinin and neuraminidase (HN). This HN region is comprised of the globular head and stalk which are the active sites involved in viral infection (Moscona, 2005). In this study a 240 amino acid fragment of the globular head was targeted for amplification and genetic analysis as it is the region with the largest the antigenic and genetic diversity (Goya et al., 2016).

The phylogenetic tree for HPIV 4A revealed 4 genetic groupings with Cluster 1 being the largest (31 samples), comprising sequences from all 4 years and none had any close homology to the GenBank reference sequences. In Cluster 2 (17 samples) there was 1 sample from 2014 (SCH7879) and the rest from 2015 to 2017, suggesting SCH7879 was probably an ancestor to cluster 1 with a year between this sample and the next, (SCH5237, 2015) which was similar to it (Figure 26, blue dots). The 2 samples in Cluster 4, from 2014 and the other 2017 clustered with a reference sequence from Japan 2011 (AB753463.1) (Figure 26). It is not clear how these sequences are linked to the 2011 sequence from Japan. However, bootstrap support for all of the clusters were below 70 % and full gene or genome analysis may be required to correctly identify distinct linages within subtype 4A. The many nucleotide changes in HPIV 4B may be related to the amino acid change (G198R) as it occurred in all the same samples with this change.

Due to many of the HPIV 4 positive samples having a semi-quantitative value of 1+ (low copy numbers), only 26% (77/295) of the positive HPIV 4 samples could be amplified sufficiently to produce a PCR product which could be used for genetic analysis.

4.4.1. Phylogenetics

Phylogenetic trees constructed using NJ method showed that most of the South African subtypes did not group with the closest significant reference sequences from GenBank, suggesting unique clusters for both subtypes suggesting local inter-epidemic spread. This is unlike influenza, where global circulation plays a role in reintroducing the virus from different countries. Whereas this data suggests that HPIV 4 is sustained by local inter-epidemic circulation. Only HPIV 4A reference sequence, MH892408.1 (2019) from Netherlands, and 2 HPIV 4B reference sequences, (AB753478.1, 2012; KX912856.1, 2018) from Japan and Malaysia respectively clustered with HPIV strains from South Africa. Two of these sequences were added to the database after the study period suggesting that they may be linked to South Africa. The long branches observed in both subtypes in the South Africa samples also show true diversity within these sequences.

A nonsynonymous amino acid change in HPIV 4A, at location N161D (amino acid position 330 in complete HN protein), resulted in the loss of a potential N-glycosylation site in 6 samples from 2016 (Figure 30 red arrow) and becoming more prominent in 2017. This loss may result in a conformational change in HN protein structure, which could help the virus to

evade the immune system and could possibly have played a role in the higher prevalence of HPIV 4A in 2017 (Shental-Benchor & Levy, 2008; Lis & Sharon, 1993; Gorman et al., 1991). This may also account for increased clinical severity, as all of these patients were hospitalised, with almost a 1/3 in ICU. The other 3 amino acid changes M85I, G164E, V181A, corresponding to positions 254, 333 and 350 in the complete HN protein did not alter potential N-linked glycosylation sites but may affect either escape or reversion mutations in potential T-cell epitopes (Figure 30).

In contrast HPIV 4B had relatively few scattered nonsynonymous amino acid changes compared to HPIV 4A, with a single amino acid change at G198R (corresponding to amino acid position 367 in complete HN protein) which corresponded with the 12 nucleic acid changes for the same patients, becoming predominant in 2016/2017 samples, indicating viral evolution possibly due to immune evasion (Figure 33). With the change of a neutral residue (asparagine) to a charged residue (aspartic acid) the virus could also evade the humoral response by antagonizing the IFN-I production (Wei, et al., 2020).

Chapter 5 Conclusion

In summary, this is the first study examining the molecular epidemiology of HPIV 4 in South Africa. It showed that HPIV 4 is a common respiratory infection in young infants and its prevalence is increasing. This may be due to more social contacts especially in a hospital setting. The assay used was a very sensitive test thus increasing the number of positive results. In this study it was found to be the most common HPIV detected in hospitalized children. There was a definite seasonal trend with infection occurring from autumn to winter. Use of an in-house developed real-time PCR assay for subtyping, revealed that both subtypes A and B co-circulate during periods of seasonal prevalence. We showed that HPIV 4 mono-infections could be responsible for severe disease requiring ICU admission. Most of the South African HPIV 4 sequences clustered together suggesting unique subtypes. There were also many amino acid changes together with a glycosylation change in HPIV 4A suggesting viral evolution. More molecular epidemiological studies need to be done in order to provide more information about viruses causing SARI and the impact they have on the burden of disease.

Appendices

Appendix A1 Ethics approval

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E53-46 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6626
Email: shuretta.thomas@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

21 August 2017

HREC REF: 605/2017

A/Prof Diana Hardie
Pathology
Virology
C-18 New Groote Schuur Hospital

Dear A/Prof Hardie

PROJECT TITLE: MOLECULAR EPIDEMIOLOGICAL STUDY OF PARAINFLUENZA 4 IN THE WESTERN CAPE (MSc-candidate-J Parsons)

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 30 August 2018.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

Please quote the HREC REF in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate institutional approval before the research may occur.

The HREC acknowledge that the student, Jane Parsons will also be involved in this study.

Yours sincerely

Signature Removed

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938
This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical

HREC 605/2017

Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP), South African Good Clinical Practice Guidelines (DoH 2006), based on the Association of the British Pharmaceutical Industry Guidelines (ABPI),and Declaration of Helsinki (2013) guidelines.
The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

Appendix A2 Buffers and solutions

A2.1. 0.1 % Dithiothreitol

Dissolve 10,5 g DL-dithiothreitol, anhydrous (Sigma) in 800 ml of autoclaved, reverse osmosis water then top up to 1 L. Store at 4 ºC.

A2.2. Tris-Acetate- EDTA (1 x TAE)

Dissolve 25 g Tris base (Sigma) in 500 ml reverse osmosis water, add 6 ml glacial acetic (Sigma) and 100 ml 0,5 M EDTA (pH 7.2) (Sigma) top up to 1 L then autoclave.

A2.3. 2 % Agarose/ ethidium bromide gel

6 g agarose tablets (Celtic) dissolved in 300 ml TAE, add 6 µl ethidium bromide (Sigma), bring to boil in microwave, cool in 60 º water bath, once cooled pour into prepared gel tray.

A2.4. Loading buffer

Add the following: Bromophenol Blue 200 mg (Sigma), Xylene Cyanol FF 50 mg (Sigma), 0.5M EDTA (~pH 8) 20 ml (Sigma) and glycerol 90 ml, top up with water, filter sterilize.

References

Abedi, G.R., et al., (2016). Estimates of Parainfluenza Virus-Associated Hospitalizations and Cost Among Children Aged Less Than 5 Years in the United States, 1998–2010 *Journal Paediatric Infection Disease Society*. March 5(1) pp. 7–13. doi:10.1093/jpids/piu047.

Abiko, C. et al., (2013). An Outbreak of Parainfluenza Virus Type 4 Infections among Children with Acute Respiratory Infections during the 2011-2012 Winter Season in Yamagata, Japan. *Japanese Journal of Infectious Diseases*, 66(1) pp.76-78.

Albariño, C.G., et al., (2014). Novel Paramyxovirus Associated with Severe Acute Febrile Disease, South Sudan and Uganda, 2012. *Emerging Infectious Diseases*. 20(2) pp. 211- 216. DOI: http://dx.doi.org/10.3201/eid2002.131620.

Alimia, Y., et al., (2017). Systematic review of respiratory viral pathogens identified in adults with community-acquired pneumonia in Europe. *Journal of Clinical Virology* 95 pp. 26–35. http://dx.doi.org/10.1016/j.jcv.2017.07.019.

Assane, D., et al., (2018). Viral and Bacterial Aetiologies of Acute Respiratory Infections Among Children Under 5 Years in Senegal. *Microbiology Insights*, *11*, 1178636118758651. doi:10.1177/1178636118758651

Bhuyan, G.S., et al., (2017). Bacterial and viral pathogen spectra of acute respiratory infections in under-5 children in hospital settings in Dhaka city*. PLoS ONE,* 12(3): e0174488. https://doi.org/10.1371/journal. pone.0174488

Billaud, G et al., (2005). Human parainfluenza virus type 4 infections: A report of 20 cases from 1998 to 2002. *Journal of Clinical Virology*, 34(1), pp.48-51.

Bowden, T.R., et al., (2001). Molecular Characterization of *Menangle Virus,* a Novel Paramyxovirus Which Infects Pigs, Fruit Bats, and Humans. *Virology* 283 pp. 358-373. doi:10.1006/viro.2001.0893

Brini Khalifa, I., et al., (2018). Demographic and seasonal characteristics of respiratory pathogens in neonates and infants aged 0 to 12 months in the Central-East region of Tunisia. *Journal of Medical Virology,* 91(4), pp. 570 – 58. https://doi.org/10.1002/jmv.25347

Breiman, R. et al., (2015). Severe acute respiratory infection in children in a densely populated urban slum in Kenya 2001-2011. *BiomedCentral Infectios Diseases.* 15 pp. 95- 106. DOI10.1186/s12879-015-0827-x

Bustin, S.A. (2005). Real-Time Reverse Transcription PCR. *Encyclopedia of Diagnostic Genomics and Proteomics*. 113, DOI: 10.1081/E-EDGP 120020680

Bustin S.A. & Mueller R., (2005). Real-time reverse transcription PCR (qRT-PCR) and its potential use in clinical diagnosis. *Clinical Science* 109 pp. 365–379 doi:10.1042/CS20050086 365

Cilloniz, C., et al., (2016). Microbial Aetiology of Pneumonia: Epidemiology, Diagnosisand Resistance Patterns. *International Journal of Molecular Sciences* 17 (12) pp. 2120; doi:10.3390/ijms17122120.

Cohen, A., et al (2015). Parainfluenza virus infection among human immunodeficiency virus (HIV)-infected and HIV-uninfected children and adults hospitalized for severe acute respiratory illness in South Africa, 2009-2014. *Open Forum Infectious Diseases*. Dec. 2(4) ofv139. doi: 10.1093/ofid/ofv139. PMID: 26566534; PMCID: PMC4630450.

Dasaraju P.V. & Liu C., (1996). Infections of the Respiratory System. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; Chapter 93. https://www.ncbi.nlm.nih.gov/books/NBK8142/

de Mello Freitas, F.T., (2013). Sentinel surveillance of influenza and other respiratory viruses, Brazil, 2000–2010. *The Brazilian Journal of Infectious Diseases*, 17(1) pp.62-68. DOI: 10.1056/NEJMoa1405870.

Durbin A. & Karron R., (2003). Progress in the Development of Respiratory Syncytial Virus and Parainfluenza Virus Vaccines. *Clinical Infectious Disease*. (37) pp. 1668 - 1677.

El Najjar F., et al., (2014). Paramyxovirus Glycoprotein Incorporation, Assembly and Budding: A Three Way Dance for Infectious Particle Production. *Viruses 6*, 3019-3054; doi:10.3390/v6083019

Frost, H.M., et al., (2013). Epidemiology and Clinical Presentation of Parainfluenza Type 4 in Children: A 3-Year Comparative Study to Parainfluenza Types 1–3. *The Journal of Infectious Diseases*. 209 pp. 695–702*;* DOI: 10.1093/infdis/jit552

Fry, A., et al., (2006). Seasonal Trends of Human Parainfluenza Viral Infections: United States, 1990–2004. *Clinical Infectious Diseases*, 43(8), pp.1016-1022.

Gaymard, A., et al., (2016). Functional balance between neuraminidase and haemagglutinin in influenza viruses. *Clinical Microbiology and Infection*, 22, pp. 975-983.

Gorman, W.L., et al. (1991). Glycosylation of the hemagglutinin-neuraminidase glycoprotein of human parainfluenza virus type 1 affects its functional but not its antigenic properties. *Virology,* 183 pp. 83-90. https://doi.org/10.1016/0042-6822(91)90120-Z.

GBD 2016 Lower Respiratory Infections Collaborators (2018). Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infectious Diseases.* 18 pp. 1191–210. http://dx.doi.org/10.1016/S1473-3099(18)30310-4

Goya, S., et al., (2016). Phylogenetic and molecular analyses of human parainfluenza type 3 virus in Buenos Aires, Argentina, between 2009 and 2013: The emergence of new genetic lineages. Infection, *Genetics and Evolution* 939, pp. 85-91.

Harrisons, M.S., et al., (2010). Paramyxovirus Assembly and Budding: Building

Particles that Transmit Infections. *The International Journal of Biochemistry & Cell Biology*, September; 42(9): 1416–1429; doi: 10.1016/j.biocel.2010.04.005.

Henrickson, K., (2003). Parainfluenza Viruses. *Clinical Microbiology Reviews*, 16(2), pp.242-264.

Horton, K.C., et al., (2017). Viral etiology, seasonality and severity of hospitalized patients with severe acute respiratory infections in the Eastern Mediterranean Region, 2007-2014. *PLoS One*. 12 (7): e0180954. https://doi.org/10.1371/journal.pone.0180954.

Hsieh, Y., et al., (2010). Hospitalized Paediatric Parainfluenza Virus Infections in a Medical Center. *Journal of Microbiology, Immunology and Infection*, 43(5), pp.360-365.

International Committee on Taxonomy of Viruses (ICTV), 2019, *Taxonomic information - Historical Taxonomy releases,* viewed, 04 August 2019, <https://talk.ictvonline.org/taxonomy/https://en.wikipedia.org/wiki/Virus_classification>

Karron, R. A. & Collins, P.L. In: Fields, B., Knipe, D. and Howley, P. (2007). *Fields' virology*. 34 Parainfluenza viruses, pp.996-1023, 6th ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Korsman S, et al. (2012). *Virology an illustrated colour text.* Churchill Livingston ELSEVIER. London pp. 30.

Lamb R.A. & Parks G.D., Paramyxoviridae. In: Fields, B., Knipe, D. and Howley, P. (2007). *Fields' virology*. 33 Parainfluenza viruses, pp. 957-994 6th ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Lau, S., et al., (2009). Clinical and Molecular Epidemiology of Human Parainfluenza Virus 4 Infections in Hong Kong: Subtype 4B as Common as Subtype 4A. *Journal of Clinical Microbiology*, 47(5), pp.1549-1552.

Lau, S., et al., (2005). Human Parainfluenza Virus 4 Outbreak and the Role of Diagnostic Tests. *Journal of Clinical Microbiology*, 43(9), pp.4515-4521.

Li, Y., et al., (2019). Global patterns in monthly activity of influenza virus, respiratory syncytial virus, parainfluenza virus, and metapneumovirus: a systematic analysis. *Lancet Global Health*. 7, pp. 1035 – 1045. www.thelancet.com/lancetgh Vol 7 August 2019 e1031.

Linster, M., et al., (2018). Clinical and Molecular Epidemiology of Human Parainfluenza Viruses 1–4 in Children from Viet Nam. *Scientific Reports*. 8, pp.6833-6841 https://doi.org/10.1038/s41598-018-24767-4.

Liu, W., et al., (2013). Epidemiology and clinical presentation of the four human parainfluenza virus types. *BMC Infectious Diseases*, 13 (1) pp. 13-28. doi:10.1186/1471- 2334-13-28.

Lis, H. & Sharon, N., (1993). Protein glycosylation Structural & functional aspects. *European Journal of Biochemistry.* 218 pp. 1-27.

Loeffelholz, M. & Chonmaitree, T., (2010). Advances in Diagnosis of Respiratory Virus Infections. *International Journal of Medical Microbiology.* 2010 pp. 1-5; doi: 10.1155/2010/126049.

Lodish, H., et al., (2000). Molecular Cell Biology. 4th edition. New York: W. H. Freeman; 2000. Section 7.3, Identifying, Analyzing, and Sequencing Cloned DNA. Available from: https://www.ncbi.nlm.nih.gov/books/NBK21505/

Mackenzie, G., et al., (2019). Respiratory syncytial, parainfluenza and influenza

virus infection in young children with acute lower respiratory infection in rural Gambia. *Scientific Reports nature research.* 9 , pp. 17965-17975. http://doi.org/10.1038/s41598-019-54059- 4.

Mandal, C., (2019). *Virus Classification*. News Medical Life Sciences, viewed, 04 August 2019, <https://www.news-medical.net/health/Virus-Classification.aspx#>

Marsh, G.A., et al., (2012). Cedar Virus: A Novel Henipavirus Isolated from Australian Bats. *PLoS Pathogens* 8(8): e1002836. doi:10.1371/journal.ppat.1002836.

Mellett, S., et.al. (2012). Interpreting diagnostic accuracy studies for patient care. *The British Medical Association.* 344 pp. 1-7. doi: https://doi.org/10.1136/bmj.e3999.

Moscona, A., (2005). Entry of parainfluenza virus into cells as a target for interrupting childhood respiratory disease. *Journal of clinical investigation.* 115 (7), pp. 1688–1698; doi:10.1172/JCI25669.

Munoz F.M., (2018). *Parainfluenza viruses in children,* UpToDate, viewed, 10 December 2019, https://www.uptodate.com/contents/parainfluenza-viruses-in-children/contributors.

Obenrader, S., (2003). *The Sanger Method,* Dept. of Biology, Davidson College, viewed, December 2019, http://www.bio.davidson.edu/Courses/Molbio/MolStudents/spring2003/Obenrader/sanger_method_page.htm

OpenStax College, Biology *(2013). *Viral Evolution, Morphology, and Classification,* Lumen Boundless Biology, viewed, 04 August 2019, https://www.news-medical.net/health/Virus- Classification.aspx#.

Ortín, J. & Martín-Benito, J., (2015). The RNA synthesis machinery of negative-stranded RNA viruses. *Virology*, 479, pp. 532–544. http://dx.doi.org/10.1016/j.virol.2015.03.018

Palermo, L.M., et al., (2016). Features of Circulating Parainfluenza Virus Required for Growth in Human Airway. *mBio* 7(2): e00235-16; doi:10.1128/mBio.00235-16.

Parija, S.C., (2019). Human Parainfluenza Viruses (HPIV) and Other Parainfluenza Viruses Medication. Viewed 04 August 2019, *https://emedicine.medscape.com/article/224708-treatment.*

Parikh, R., et al., (2008). Understanding and using sensitivity, specificity and predictive values. *Indian Journal of Ophthalmology* vol. 56 (1) pp. 45-50. doi:10.4103/0301- 4738.37595.

Pawełczyk, M. & Kowalski, M., (2017). The Role of Human Parainfluenza Virus Infections in the Immunopathology of the Respiratory Tract. *Current Allergy and Asthma Reports*, 17(3) pp. 16: doi: 10.1007/s11882-017-0685-2.

Porotto, M. et al., (2012). The Second Receptor Binding Site of the Globular Head of the Newcastle Disease Virus Hemagglutinin-Neuraminidase Activates the Stalk of Multiple Paramyxovirus Receptor Binding Proteins to Trigger Fusion. *Journal of Virology*. pp. 5730– 5741; doi:10.1128/JVI.06793-11.

Ren, L., et al., (2009). Prevalence of human respiratory viruses in adults with acute respiratory tract infections in Beijing, 2005–2007. *Clinical Microbiology and Infection*, 15(12), pp.1146-1153.

Ren, L., et al., (2011). Human parainfluenza virus type 4 infection in Chinese children with lower respiratory tract infections: A comparison study. *Journal of Clinical Virology*, 51(3), pp.209-212.

Resuehr, D. & Spiess, A., (2003). A real-time polymerase chain reaction-based evaluation of cDNA synthesis priming methods. *Analytical Biochemistry.* Nov. 322 (2) pp. 287-291. DOI: 10.1016/j.ab.2003.07.017

Rhedin, S.A., (2017). Severe viral respiratory tract infections in children. Published: Karolinska Institutet. ISBN: 978-91-7676-497-8;

http://hdl.handle.net/10616/45606

Rissanen, I., et al., (2016). Idiosyncratic Mojiang virus attachment glycoprotein directs a host-cell entry pathway distinct from genetically related Henipaviruses. *Nature Communications*. 8:16060. DOI: 10.1038/ncomms16060

Schmidt, A., et al., (2011). Progress in the development of human parainfluenza virus vaccines. *Expert review respiratory medicine*., 5, pp. 515-526

Shental-Benchor, D. & Levy, L., (2008). Effect of glycosylation on protein folding: A close look at thermodynamic stabilization. *Proceedings of the National Academy of Sciences of the United States of America.* 105(24) pp. 8256-8261.

Shi, P., et al., (2020). Age- and gender-specific trends in respiratory outpatient visits and diagnoses at a tertiary pediatric hospital in China: a 10-year retrospective study. *BioMed Central Pediatrics*. 20, pp. 115-125. https://doi.org/10.1186/s12887-020-2001-x.

Tamura, K., et al., (2013). Mega6 : Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology Evolution.* 30(12) pp. 2725-2729. doi:10.1093/molbev/mst197.

Thomazelli, L.M., et al., (2017). Human parainfluenza virus surveillance in paediatric patients with lower respiratory tract infections: a special view of parainfluenza type 4. *Journal Paediatrics (Rio J)*. 94(5), pp. 554 – 558

Vachon, M., et al., (2006). Human Parainfluenza Type 4 Infections, Canada. *Emerging Infectious Diseases*, 12(11), pp.1755-1758.

Vainionpaa, R. & Hyypia T., (1994). Biology of Parainfluenza Viruses. *Clinical Microbiology Reviews*, Apr. pp. 265-275.

Wang, C., et al., (2012). A novel duplex real-time PCR for HPIV-4 detects co-

circulation of both viral subtypes among ill children during 2008. *Journal of Clinical Virology*, 54(1), pp.83- 85.

Wei Z.Y., et al., (2020). Aspartic acid at residue 185 modulates the capacity of HP-PRRSV nsp4 to antagonize IFN-I expression. *Virology*. 2020;546:79-87. doi:10.1016/j.virol.2020.04.007

Weinberg, G.A., et al. (2009). New Vaccine Surveillance Network. Parainfluenza virus infection of young children: estimates of the population-based burden of hospitalization. *The Journal Pediatrics*. 154(5) pp. 694-9. doi: 10.1016/j.jpeds.2008.11.034. Epub 2009 Jan 21. PMID: 19159905.

World Health Organisation – Africa 2018, *Atlas of African Health Statistics 2018: universal health coverage and the Sustainable Development Goals in the WHO African Region*, viewed 04 August 2019, <http://www.aho.afro.who.int/en/atlas/atlas-african-health- statistics-2018>

Yano, T., et al., (2014). Epidemiological Investigation and Seroprevalence of Human Parainfluenza Virus in Mie Prefecture in Japan during 2009-2013. *Japanese Journal of Infectious Diseases*, 67(6), pp.506-508.

Yea, C., et al., (2009). The Complete Sequence of a Human Parainfluenza virus 4 Genome. *Viruses*, 1(1), pp.26-41.

Zhang, D., et al., (2014). Epidemiology characteristics of respiratory viruses found in children and adults with respiratory tract infections in southern China. *International Journal of Infectious Diseases*, 25, pp.159-164.

Zhao, H., et al., (2017). Epidemiology of parainfluenza infection in England and Wales, 1998- 2013: any evidence of change? *Epidemiology Infection*. Apr. 145 (6) pp. 1210-1220.