ASSOCIATION OF RESPIRATORY SYNCYTIAL VIRUS GLYCO- AND NON-STRUCTURAL PROTEINS WITH DISEASE SEVERITY IN INFECTED CHILDREN

BY

STEPHANIE VAN NIEKERK

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

MAGISTER SCIENTIAE

MSC (MEDICAL VIROLOGY)

IN THE

DEPARTMENT OF MEDICAL VIROLOGY

FACULTY OF HEALTH SCIENCES

UNIVERSITY OF PRETORIA

2012

Declaration

I, Stephanie van Niekerk, declare that this thesis is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of Pretoria. It has not been submitted before for any degree or examination in any other Technikon or University. Confidentiality of patient's identity has been followed according to the ethical rules as prescribed by the Faculty of Health Science Research Ethics Committee, University of Pretoria.

_____

S van Niekerk

_____ day of _____ 2012

Acknowledgements

First and foremost I would like to thank God, the father of all, for giving me the strength and knowledge to finish my thesis.

I offer my sincerest gratitude to my supervisor, Prof M Venter, who has supported and guided me throughout my thesis. I want to thank the Poliomyelitis Research Foundation (PRF) for financial assistance and the University of Pretoria, Faculty of Health Sciences for providing financial means and laboratory facilities. In addition, many thanks to the NHLS for providing me with all my samples covered in this thesis. In my daily work I have been blessed with a friendly and cheerful group of fellow students: Dewald Zaayman, Charmaine van Eeden, Juliet Mentoor and Yvette Westerberg. You are not only colleagues but also my friends. Thank you to each and everyone one of you for the role you play in my studies and life.

I thank my parents for supporting and loving me throughout all my studies, for the financial and moral support. You are the reason that I am successful in what I do. Also, a special thanks to my in-laws who accepted me as their daughter and their continuous support.

Finally I thank my husband, Kobus, who inspired, encouraged and fully supported me throughout my thesis. I appreciate and love you immensely.

ASSOCIATION OF RESPIRATORY SYNCYTIAL VIRUS GLYCO- AND NON-STRUCTURAL PROTEINS WITH DISEASE SEVERITY IN INFECTED CHILDREN

by

STEPHANIE VAN NIEKERK

SUPERVISOR: PROF M VENTER
DEPARTMENT: MEDICAL VIROLOGY
DEGREE: MAGISTER SCIENTIAE (MEDICAL VIROLOGY)

SUMMARY

Respiratory Syncytial Virus (RSV) is a major cause of bronchiolitis and pneumonia in infants, immunocompromised and the elderly in both developed and developing countries. Re-infections are common and G protein variability is one mechanism to overcome herd immunity. This is illustrated by the appearance of the BA genotype with a 60 nucleotide duplication dominating the subtype B genotypes in epidemics worldwide. To investigate the evolution of subtype A and B in South Africa (SA) since 2002 the genetic variability in the G protein was analysed in all recent strains isolated over four years (2006-2009) in SA hospitals. Bayesian analysis revealed a replacement of all subtype B genotypes previously identified in SA with the BA genotype since 2006, while subtype A genotypes identified in previous years are still circulating. The evolutionary rate of the SA BA genotype was shown to be 2.305 x 10^{-3} nucleotide substitutions/site/year while subtype A was shown to have 3.382 x 10^{-3} substitutions/site/year and drift was evident. The most recent common ancestor (MRCA) of the SA BA viruses was determined to date back to 1997. All SA BA isolates clustered with the BA-IV sub-genotype and the appearance of new sub-genotypes within this branch may occur if drift continues. Sequencing of the complete G protein of selected BA SA strains revealed an additional 6 nucleotide deletion. When comparing G protein variability with disease severity, the dominating subtype/genotype in each season was also the subtype/genotype that was associated with increased hospitalizations. No direct

association between G protein variability and disease severity was seen. Subtype/genotype switching was evident over the four years most probably because of herd immunity. G protein variability may play a role in RSV's ability to re-establish annual epidemics by allowing immune evasion. Certain substitutions or alterations may enhance the fitness of viruses as is evident with the BA strains that replaced all other B genotypes previously identified in SA. The G protein's ability to accommodate such substantial changes and facilitate immune evasion may complicate vaccine development. It remains to be seen if this BA genotype will remain dominant or if the dominance will eventually fade because of herd immunity.

Subtyping of RSV strains over the four years identified G protein PCR amplicons significantly reduced in size in 2 out of 209 clinical specimens. Sequence analysis revealed subtype B strains lacking nearly the entire G protein ectodomain in one HIV positive and one HIV exposed child hospitalized with pneumonia. G protein deletion mutants replicate effectively in vitro but have not been detected in nature. This study suggests that RSV clinical strains that lack most of the G protein gene may occur in immunocompromised patients with lower respiratory tract infection (LRTI). The molecular mechanism whereby this occurs is not clear, however reduced immune pressure in these patients may allow these strains to utilise the F protein for binding and replication. Further characterization of such strains may elucidate the replication and pathogenic potential, however low viral load and long term storage of specimens complicated the isolation of such strains.

Acquisition of the 60 nucleotide duplication appeared to have improved the fitness of the BA viruses and more recent subtype B strains may need to be included in experimental vaccines to evaluate their efficacy in the current setting of evolved circulating strains. In addition, the association of clinical strains lacking most of the G protein with LRTI may have implications for the utilisation of certain attenuated strains in immunocompromised children.

RSV is unique among paramyxoviruses in having two non-structural (NS) proteins that play a major role in inhibiting the host's immune response. Sequence and quantification analysis of these proteins were performed to determine its role in disease severity. We

were unable to attribute specific protein polymorphisms with differences in disease severity but identified genome heterogeneity (quasispecies) within patients which may have an influence on the NS protein function, and also have an influence on the innate immune response and thus an effect on disease severity. When comparing patients with mild and severe disease, higher expression levels of NS1 were seen in the hospitalized group compared to the out-patient group, supporting our hypothesis that increased levels of NS may lead to enhanced suppression of the interferon pathway and in effect result in more severe disease. However, the opposite was found for NS2 with higher expression levels in the mild group. Genetic variability within the gene-end and gene-start sequences were not seen thus could not explain the differences in expression levels observed, although variability within the promoter area may need to be investigated.

PRESENTATIONS AND PUBLICATIONS

Presentations

Stephanie van Niekerk & Marietjie Venter; Poster presentation. Association of Respiratory Syncytial virus (RSV) G glycoprotein differences with disease severity in children. Federation of Infectious Diseases of Southern Africa, South Africa, Sun City, 23-25 August 2009.

Marietjie Venter; **Stephanie van Niekerk**; Andronica Rakgantso; Poster Presentation Identification of deletion mutant Respiratory Syncytial virus strains lacking most of the G-protein in children with pneumonia in South Africa. 7th International Respiratory Syncytial Virus Symposium, Rotterdam, the Netherlands, 2-5 December 2010.

Publications

Marietjie Venter, **Stephanie van Niekerk**, Andronica Rakgantso and Nicola Bent. (2011). Identification of deletion mutant Respiratory Syncytial virus strains lacking most of the G-protein in immunocompromised children with pneumonia in South Africa. *Journal of Virology* 85(**16**), p8453-8457.

Stephanie van Niekerk & Marietjie Venter. (2011). Replacement of previously circulating Respiratory Syncytial virus (RSV) subtype B strains with the BA genotype in South Africa. *Journal of Virology* 85(**17**), p8789-8797.

TABLE OF CONTENTS

CHAPTER 2: EVOLUTIONARY PATTERN AND DISEASE SEVERITY ASSOCIATION OF RESPIRATORY SYNCYTIAL VIRUS G GLYCOPROTEIN OVER FOUR CONSECUTIVE YEARS

ABBREVIATIONS

HRSV	= Human Respiratory Syncytial Virus
URTI	= Upper respiratory tract infection
LRTI	= Lower respiratory tract infection
WHO	= World Health Organization
aa	= Amino acid
Kb	= Kilo base
bp	= Base pair
MAbs	= Monoclonal antibodies
SA	= South Africa
ORF	= Open reading frame
Th2	= T helper 2
GAG	= Glycosaminoglycans
PIV	= Parainfluenza
TNF	= Tumour necrosis factor
GS	= Gene-start
GE	= Gene-end
NC	= Nucleocapsid
RNP	= Ribonucleoprotein
INF	= Interferon
CTL	= Cytotoxic T cell
PI3K	= Phosphoinositide 3-kinase
DC	= Dendritic cell
NK	= Natural killer
Ig	= Immunoglobulin
Th1	= T helper 1
IL	= Interleukin
HIV	= Human immunodeficiency virus
KMC	= Kangeroo Mother Care
SNP	= Single nucleotide polymorphism
US	= United States
EIA	= Enzyme immuno-assays

FA = Flourescent antibody

RT-PCR = Reverse transcription polymerase chain reaction

siRNA = Small interfering RNA

PFP = Purified F protein

FI-RSV = Formalin inactivated RSV

PVM = Pneumonia virus of mice

IF = Immunofluorescence

NHLS = National Health Laboratory Service

NPA = Nasopharyngeal aspirate

BEAST = Bayesian Evolutionary Analysis by Sampling Trees

MCMC = Markov Chain Monte Carlo

P-distance = Pairwise distance

MRCA = Most recent common ancestor

HPD = Highest probability density

ADV = Adeno virus

hMPV = Human Metapneumo virus

hBoV = Human Boca virus

hCoV = Human Corona virus

BT = Bovine turbinate

C_T = Cycle threshold

CP = Crossing point

FMDV = Foot-and-mouth disease virus

µl = Microliter

^oC = Degree Celsius

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Literature Review

1.1 INTRODUCTION

Human respiratory syncytial virus (HRSV) was first isolated from a laboratory chimpanzee with coryza (Morris et al., 1956) and one year later the first isolation from a human clinical sample was described (Chanock et al., 1957). The name describes the site of infection and the characteristic syncytium formation. RSV was rapidly determined to be of human origin and was shown to be a pathogenic virus that causes serious paediatric respiratory tract disease. It is one of the most contagious human pathogens and is readily introduced and spreads with ease in hospitals, nursing homes, families, and other close-contact settings. During the first year of life, approximately two thirds of infants are infected with RSV, and 90% have been infected one or more times by two years of age(Glezen et al., 1986). Most children with RSV infection suffer from mild upper respiratory tract infection (URTI); however up to 40% of children infected develop serious lower respiratory tract infection (LRTI) requiring hospitalization(Tripp, 2004). The virus accounts for approximately 50% of all pneumonia and up to 90% of the reported cases of bronchiolitis in infancy(Glezen, 1987).

Although RSV is primarily known as a paediatric pathogen, it is also a noteworthy cause of morbidity and mortality in the elderly(Falsey et al., 2005). Infection in adults can also cause LRTI, although the disease usually presents as a mild URTI. At these times of mild symptomatic disease, adults may serve as reservoirs for RSV and may cause infection in infants whose immune system is still immature(Hall, 2001). In moderate climates, RSV epidemics occur yearly in the winter months, whereas outbreaks are associated with the rainy season in humid climates(Cane, 2001). The disease burden of RSV is very high and globally, the Word Health Organization (WHO) estimates that RSV causes 64 million infections and 160 000 deaths annually (www.who.int/vaccine_research/diseases/ari/en/indexz.html.).

The respiratory epithelial cells are the primary target for RSV in the establishment of infection and re-infection are known to occur throughout life. This is mainly because children are infected in the presence of maternal antibodies and natural infection affords only partial protection(Wertz and Moudy, 2004). Evidence also shows that antigenic variability, especially in the G glycoprotein, plays a role in repeat infections by allowing immune evasion(Sullender, 2000). Animal models of RSV infection are very important for studying these mechanisms whereby the virus evades the immune system, although it is challenging because RSV induces different disease pathologies in different animal species(Moore and Peebles, 2006).

Despite the importance of this pathogen, neither a vaccine nor effective antiviral therapy is currently available. This is partly due to the fact that RSV pathogenesis is not fully understood. Education about the control of RSV infection, active surveillance for respiratory viruses and keeping records of the clinical condition of not only patients, but also mothers and medical personnel frequently in contact with the babies may serve as an early warning system and reduce the risk(Visser et al., 2008).

1.2 MOLECULAR BIOLOGY OF RSV

RSV is a member of the genus *Pneumovirus*, *Pneumovirinae* subfamily of the *Paramyxoviridae* family of viruses. RSV virions are membrane-enveloped capsids and possess a single-stranded negative strand RNA genome of 15.2 kilo bases (Kb) containing 10 genes which, in infected cells, are transcribed into positive mRNA molecules.

Five RSV proteins are involved in nucleocapsid structure and/or RNA synthesis: Nucleocapsid N protein for RNA binding, L protein which is a major polymerase subunit, the P phosphoprotein which is a cofactor in RNA synthesis, the M2-1 protein which is involved in transcription and the M2-2 protein which modulate the balance between transcription and RNA replication. The other four proteins form part of the viral envelope: The matrix M protein and the transmembrane surface glycoproteins G, F and SH. NS1 and NS2 are the two non-structural proteins that play a role in the pathogenesis of

RSV(Collins and Jr Crowe, 2007). Figure 1.1 shows the viral genome and the different proteins and their functions encoded by the 10 genes(Collins and Graham, 2008). The next section will deal with these proteins and their functions in more detail.

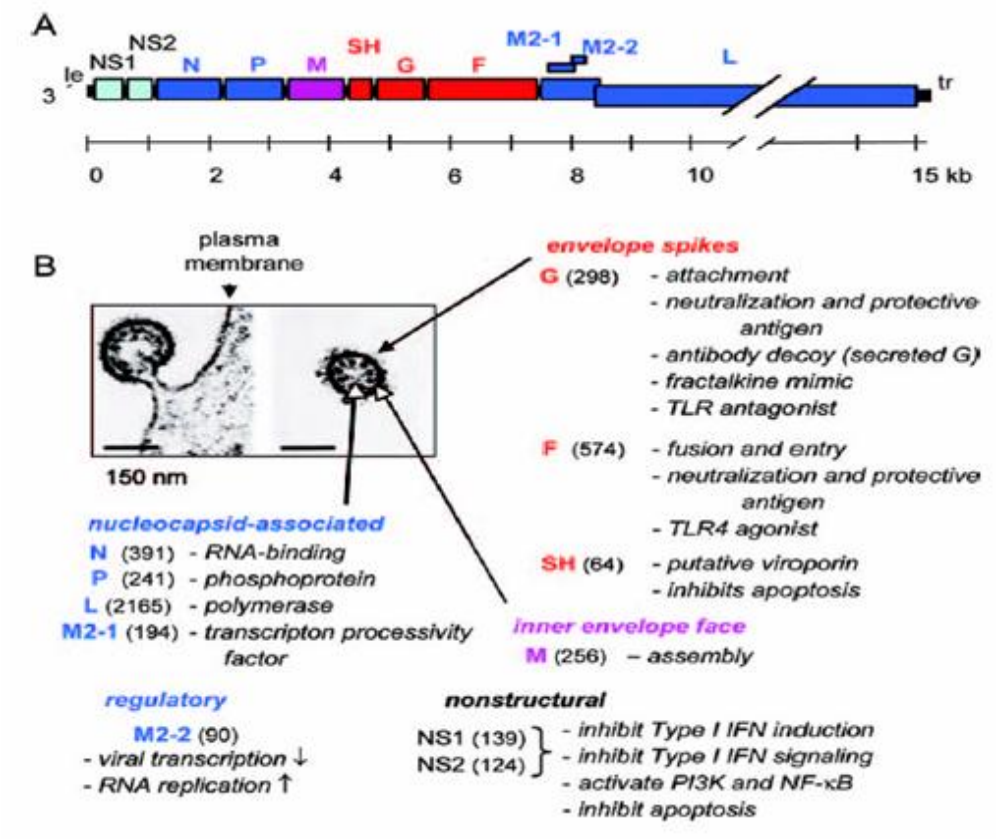


FIGURE 1.1 RSV virion, RNA genome and encoded proteins. A) The negative-sense RNA genome. B) Electron photomicrograph of the RSV virion. Protein functions and amino acid lengths are indicated(Collins and Graham, 2008).

1.2.1 Virus entry

1.2.1.1 The G glycoprotein

The G protein is a type II transmembrane glycoprotein that mediates viral attachment and is O- and N-glycosylated(Techaarpornkul et al., 2002, Teng and Collins, 2002). It is produced in two forms: a membrane-bound form and a soluble form which are generated by initiation of translation at alternative AUG codons(Roberts et al., 1994). It was shown that the soluble form (sG) helps RSV escape the antibody-dependent restriction of replication by acting as an antigen decoy and as a modulator of leukocytes bearing Fc gamma receptors(Bukreyev et al., 2008). The G protein has an ectodomain that contains a central region and four cystines which are a postulated receptor binding site. Flanking this region, are two hypervariable regions that have a high level of sequence variation(Roberts et al., 1994). Figure 1.2 illustrates the structural features of the G protein.

Two antigenic subtypes (A and B) have been identified, with little cross protection (Anderson et al., 1985). The two antigenic subtypes are estimated to have diverged 350 years ago and still retain a high level of antigenic relatedness(Zlateva et al., 2005b). The major antigenic differences between the two subtypes are a feature of the attachment protein G. It has been shown to be the most variable protein with 67% similarity at the nucleotide level and 53% similarity at the amino acid level between the two different subtypes(Johnson et al., 1987a). This high level of amino acid change is considered evidence for selective pressure for the change in G sequences(Wertz and Moudy, 2004). Because the G protein is so variable, G-specific monoclonal antibodies (MAbs) fall into three categories: (a) reaction with both antigenic subtypes recognize “conserved” epitopes, (b) those that react broadly within but not between subtypes are “subtype specific”, and (c) those that react with some strains but not others within a subtype are “strain specific”(Martinez et al., 1997).

Several genotypes have also been documented within each subtype: GA1-GA7 and SAA1 in subtype A and GB1-GB4 (Peret et al., 2000) and SAB1-SAB3 in subtype B (Venter et al., 2002a). This virus evolves at a rapid rate and can thus accommodate considerable changes in its surface proteins. Consequently a new genotype (BA) with a

60 nucleotide duplication in the C-terminal of the G protein have been discovered in Buenos Aires in 1999 that has subsequently been identified in many RSV outbreaks (Kuroiwa et al., 2005, Nagai et al., 2004, Sato et al., 2005, Scott et al., 2004b, Trento et al., 2003, Zlateva et al., 2005b, Zlateva et al., 2004). Trento et al. *(*Trento *et al*., 2006) investigated the evolution of this new BA genotype and their analysis showed that the BA genotype can be further divided into 6 sub-genotypes (I-VI). An in depth investigation into the evolution of this genotype revealed that since 2004 a relatively new sub-genotype IV has replaced all other group B viruses, suggesting further adaptation in its natural host (Trento *et al*., 2010). In the year 2006, Visser *et al*.(Visser *et al*., 2008) identified this new BA genotype for the first time in a RSV epidemic in SA. Different circulation patterns exist between all these different strains. These strain circulation patterns are complex and may be a mechanism by which RSV evades immunity(Moore and Peebles, 2006).

Three mechanisms of change in the variable regions of the G protein have been described:

1. Substitutions
2. Frameshifts into any of the three possible open reading frames (ORF)
3. Presence of alternate termination codons, resulting in proteins of differing length.

Frame shift mutations, premature stop codons and hypermutation events that caused severe truncations in the G protein have been shown to generate neutralization resistant mutants in serial passages but it remains to be seen whether these changes in the escape mutants are also observed among natural isolates(Garcia-Barreno et al., 1990, Martinez et al., 1997, Rueda et al., 1991). Studying the evolutionary patterns of the G protein of subtype B HRSV viruses, major differences in the G protein length were noted(Martinez et al., 1999). These differences were found to be caused by the usage of alternative termination codons which highlights yet another mechanism in the generation of HRSV diversity. These differences in length have been described, especially for RSV subtype B isolates. Truncated strains of up to 20 amino acid have been identified and Venter *et al*.(Venter *et al*., 2002a) identified a severely truncated wild-type G protein with a premature stop codon at position 262 that caused LRTI in a 32-month-old boy. There

was no evidence that this mutation resulted in an attenuated phenotype. This raised the question of the necessity of the G protein for efficient infection in vivo.

It has been shown that G protein subunit vaccines induce a T helper 2 (Th2) response. This is similar to the formalin inactivated vaccine that resulted in more severe disease upon re-exposure which indicate a role for the G protein in RSV pathogenesis(Srikiatkhachorn et al., 1999). According to Tripp *et al*.(Tripp et al., 2001) a potential contributor to immune modulation and disease pathogenesis is CX3C chemokine mimicry mediated by the CX3C motif in the central conserved region of the G protein of RSV. Upon live-attenuated RSV vaccine development the deletion of non-essential genes provides a method of attenuation. Karron *et al*.(Karron et al., 1997c) evaluated and characterized a cold-passaged subtype B virus which was found to contain a large deletion spanning most of the coding sequences for the small hydrophobic (SH) and attachment (G) proteins. This virus replicated efficiently in Vero cells, but was found to be overattenuated for RSV-seronegative infants and children. However, Teng and Collins (Teng and Collins, 2002) showed that the central conserved domain and cystine noose of the G protein can be deleted without affecting virus growth in vitro or in vivo (in mice).

Cell surface glycosaminoglycans (GAGs) are responsible for the majority of RSV attachment to cultured cells leading to infection(Feldman et al., 2000). Techaarpornkul *et al*.(Techaarpornkul et al., 2002) have shown that RSV virions containing the F protein as the sole surface protein, do bind to GAGs and that part of the F protein binding represents an interaction with a molecule other than a GAG. This demonstrates that the F protein may have an auxiliary role as an attachment protein.

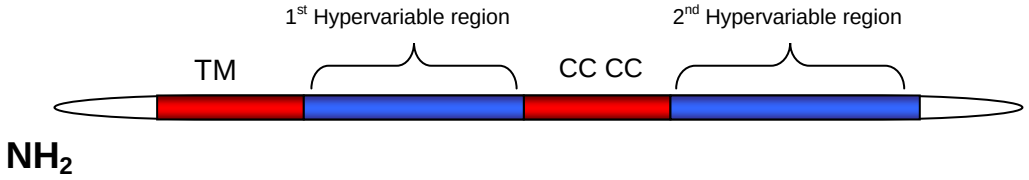


FIGURE 1.2 The structural features of the G glycoprotein. Areas of amino acid variability are indicated in blue and regions of conservation are indicated in red. NH_2, amino terminus; COOH, carboxy terminus; TM, trans membrane region; CC CC, Cystine noose. Adapted from Wertz and Moudy(Wertz and Moudy, 2004).

1.2.1.2 The F glycoprotein

Together with the G protein, the F protein is one of the major surface glycoproteins that are anchored in the viral membrane and is primarily N-glycosylated on asparagine residues(Techaarpornkul et al., 2002). The F protein is a type I integral membrane protein that is synthesized as a precursor of 574 amino acids, which is post-transcriptionally cleaved into two disulfide-linked subunits of 438 and 115 amino acids respectively(Collins et al., 1996b). This protein mediates fusion of the viral and cell membranes for virus entry into the cell as well as fusion of the infected cell membranes with that of adjacent cells which induce syncytial formation(Walsh and Hruska, 1983). Techaarpornkul *et al*.(Techaarpornkul et al., 2001) have also shown that the F protein can act as an auxiliary attachment protein where the G and SH proteins were absent.

Most phylogenetic studies have previously been conducted with the G protein because of its extensive variability. Kim *et al*.(Kim et al., 2007) has performed a study to characterize the molecular epidemiology and evolutionary patterns of the F protein throughout nine consecutive seasons and to clarify the relationship between the F and G protein with regard to genetic variation. Their phylogenetic analysis revealed nine genotypes within subtype A and four genotypes within subtype B and that the genetic variability of the F protein is associated with that of the G protein. In addition, the extent

of nucleotide and amino acid variability within the F protein neutralizing antibody and CTL (cytotoxic T cell) epitopes were investigated in SA (Agenbach et al., 2005). It was shown that the F protein was highly conserved between subtypes and suggested an absence of positive selection. Genotype-specific mutations were however found and secondary structure predictions identified minor amino acid changes that may influence the secondary structure of the protein.

1.2.1.3 The SH protein

The function of the small hydrophobic protein, which is a type II transmembrane protein, is largely unknown. Techaarpornkul *et al*.(Techaarpornkul et al., 2001) showed that RSV lacking the SH gene (RSV∆SH) is viable, causes syncytium formation and grows as well as the wild-type virus, thus indicating that the SH protein is not necessary for virus entry or syncytium formation. It was shown that the SH protein has a marked impact on the inflammatory and innate immune response to RSV infection(Tripp et al., 1999). Fuentes *et al*.(Fuentes et al., 2007) hypothesized that the SH protein of RSV may be functionally similar to other SH proteins from members of the *Paramyxoviridae* family. Their results indicated that the SH protein of RSV has a similar function to that of Parainfluenza virus 5 (PIV5) SH protein by inhibiting tumour necrosis factor alpha (TNF-α) signalling and consequently playing a role in blocking cell death.

1.2.2 Transcription and Replication

The RSV viral polymerase initiates at a single 3'-proximal promoter and transcribes 10 monocistronic mRNAs in an obligatorily sequential manner from 3' to 5'(Dickens et al., 1984). In this instance the polymerase must terminate transcription of each upstream gene to initiate transcription of the downstream gene. This is achieved by conserved gene-start and gene-end sequences flanking each gene. This transcriptional control mechanism is shown in figure 1.3. During RNA replication a genome-length positive strand replicative intermediate (antigenome) is produced that serves as the template for production of progeny genomes(Collins et al., 1996a). This viral RNA is enwrapped by the N protein in the helical nucleocapsid. The L, P and M2-1 proteins are bound to the nucleocapsid forming the ribonucleoprotein complex (RNP). This complex is the inner

part of the virion, surrounded by the M protein, that connects it to the membranous envelope in which the viral glycoproteins F, G and SH are inserted(Collins and Jr Crowe, 2007).

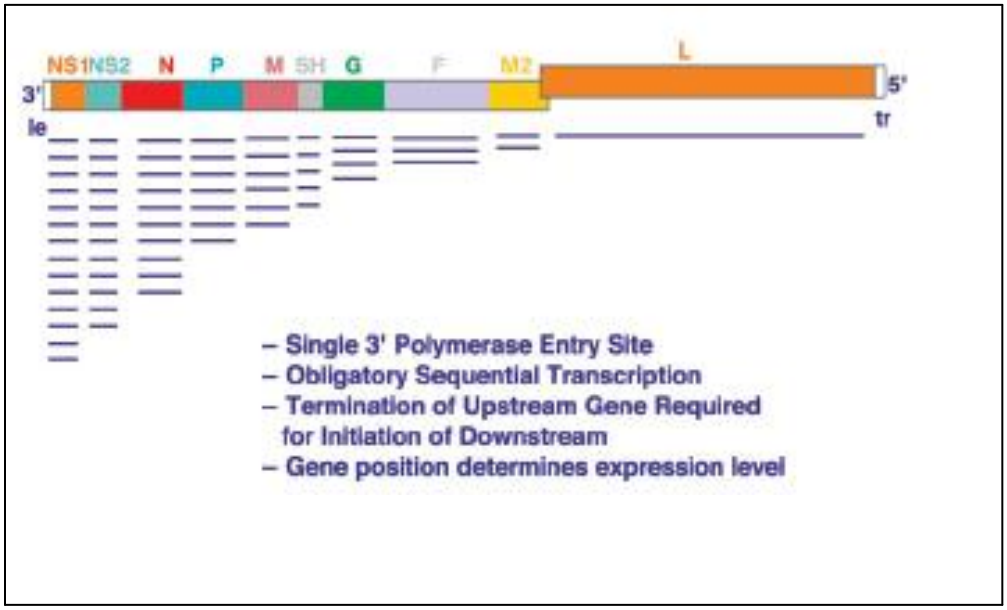


FIGURE 1.3 HRSV transcriptional control. Le, Leader; tr, trailer (Wertz and Moudy, 2004).

1.2.2.1 The RNA-dependent RNA polymerase (L)

The RSV L gene is 6578 nucleotides in length and contains a single large ORF that encodes a protein of 2165 amino acids(Stec et al., 1991). The L protein most likely catalyzes initiation, elongation, and termination of RNA synthesis although it may perform other functions during RNA replication because of its large size. Collins *et al*.(Collins et al., 1987) revealed that the L gene transcription initiates within its upstream neighbour, the 22K gene or otherwise known as the M2 mRNA and that the presence of the 22K gene-end signal within the RSV L gene appears to be a novel mechanism for obtaining additional attenuation of L gene transcription. Any increase or decrease in the availability of the polymerase L protein could have a dramatic effect on levels of transcription, replication and virion production, altering the balance between acute and persistent infection(Collins et al., 1987).

1.2.2.2 The nucleoprotein

The HRSV N protein (391amino acids) is smaller than its *Paramyxovirinae* counterparts. The RSV genome is fully encapsidated by the N proteins to form ribonucleocapsids that serve as templates for RNA synthesis(Grosfeld et al., 1995). This has several benefits for the virus: encapsidation prevents formation of secondary structures in the RNA template, it also protects the RNA from nuclease attack and reduces production of double-stranded RNA which reduces anti-viral responses(Le Mercier et al., 2002). Co-expression of HRSV N and P protein induces the formation of cytoplasmic aggregates although the relevance of these cytoplasmic inclusions for the infectious cycle remains to be elucidated(Garcia et al., 1993).

1.2.2.3 The phosphoprotein

HRSV P, an essential cofactor of the viral polymerase, is composed of 241 amino acids and is mainly phosphorylated at Ser-232(Barik et al., 1995). It also contains a central structured domain (amino acid 100-200), flanked by two disordered regions (amino acid 1-99 and 201-241)(Llorente et al., 2006). Its role during transcription and replication is to locate the L protein on the nucleocapsid template and to maintain the N protein in a functional RNA encapsidation state(Collins and Jr Crowe, 2007). Asenjo *et al*.(Asenjo et al., 2008) have also recently shown that phosphorylation at position S54 is required for viral uncoating that leads to the dissemination of the RNPs from the M protein, making them active for primary transcription.

1.2.2.4 The matrix protein

This protein has been found in cytoplasmic inclusions, which are RNP aggregates, in association with the N, P and M2-1 proteins(Li et al., 2008). Their findings suggest that the M protein associates with RNPs late in the virus life cycle in order to inhibit transcription activity, thus facilitating virus assembly. This study also showed that the M protein can only associate with these nucleocapsids in the presence of the M2-1 protein and that this interaction is strongly dependent on the N-terminal 110 amino acids of the M protein.

1.2.2.5 The M2 mRNA

This protein was shown by Collins *et al*.(Collins et al., 1990) to be one of the more conserved proteins of RSV with 92% amino acid identity between the two subtypes. The M2 mRNA contains two overlapping translational ORFs. The upstream translational ORF (ORF1) encodes a 194 amino acid protein called the M2-1 and the second translational ORF (ORF2) encodes an 83-90 amino acid protein called M2-2. This second ORF has three potential start sites at codons 1, 3 and 7(Bermingham and Collins, 1999).

1.2.2.5.1 The M2-1 protein

This protein consists of 194 amino acids and is a structural component of the RSV virion and has been shown in two studies to lead to an enhancement in transcriptional readthrough at gene-end signals, permitting access of the RSV polymerase to downstream transcriptional units(Collins et al., 1996a, Hardy and Wertz, 1998). Tran *et al*.(Tran et al., 2009) have shown that the M2-1 forms tetramers in solution and can be divided into four distinct parts: a putative ZnF (amino acid 1 to 32), the function of which remains unknown; a domain responsible for M2-1 oligomerization , situated between amino acid residues His33 and Gly62; a globular domain composed of α-helices, encompassing amino acid residues 63 to 170, capable of binding P and RNA with no sequence specificity; and a predicted unstructured C-terminal region (amino acid 171 to 194) with a low degree of conservation.

1.2.2.5.2 The M2-2 protein

The product of ORF2 (M2-2 protein) was shown to increase RNA replication and reduce transcription(Bermingham and Collins, 1999). This study suggested that the M2-2 protein mediates negative regulation of transcription late in infection. It has been proposed for Sendai virus and other non-segmented negative stranded RNA viruses(Vidal and Kalakofsky, 1989), that RNA synthesis switches from transcription to replication upon accumulation of N protein although with RSV this switch implicates the

M2-2 protein although the mechanism of action is still unclear (Bermingham and Collins, 1999).

1.2.3 Non-structural proteins

1.2.3.1 NS1 and NS2 proteins

NS1 and NS2 (139 and 124 amino acids, respectively) are non-structural proteins encoded from promoter-proximal genes to ensure high expression and has been shown to be highly conserved between the two subtypes at both nucleotide and amino acid level(Collins and Wertz, 1985, Johnson and Collins, 1989). Evans *et al*.(Evans et al., 1996) have found both proteins to be synthesised early in the infectious cycle with NS1 expression appearing to be stable while NS2 disappeared rapidly. Multimeric forms of both proteins have been demonstrated. Recent evidence indicates that they are interferon (IFN) antagonists thus suppressing the host's immune system. They inhibit the induction of IFN-α/β by blocking the activation of IFN regulatory factor 3 and by inhibiting type 1 IFN-induced signalling through the JAK/STAT pathway. Spann *et al.(*Spann *et al., 2004)* created recombinant RS viruses by deleting the NS1 and/or the NS2 proteins separately or in combination (∆NS1, ∆NS2 or ∆NS1/2). These recombinant viruses showed a decrease in replication efficiency in cells that produce IFN-α and IFN-β. Infection with the ∆NS1/2 virus resulted in an increase of both IFN-α and IFN-β as well as IFN-λ. They confirmed their findings with RT-PCR and ELISA. This supports the idea that RSV has two IFN antagonists, NS1 and NS2*(*Spann *et al., 2004)*.

Even though we know that NS1 and NS2 suppress IFN synthesis as well as function, their exact targets are still uncharacterized. Swedan *et al*.(Swedan et al., 2009) investigated whether either or both of the NS proteins affect the steady-state levels of key members of the IFN pathway. They found that both proteins decreased the levels of TRAF3, a strategic integrator of multiple IFN-inducing signals, although NS1 was more efficient. Only NS1 reduced IKKepsilon which is a key protein kinase that phosphorylates and activates IFN regulatory factor 3. NS2 was shown to decrease the levels of STAT2, the essential transcription factor for IFN-inducible antiviral genes. As a consequence of type I IFN suppression, it was shown that NS2 also suppresses the CTL component of the adaptive immune response (Kotelkin et al., 2006). This illustrates that

the two NS proteins work individually and together to regulate key signalling molecules of the IFN pathways.

In addition, it was recently found that the expression of NS1 and/or NS2 activates the phosphoinositide 3-kinase (PI3K) pathway. Suppression of the expression of NS1 and/or NS2 suppressed the activation of the PI3K pathway and resulted in accelerated apoptosis of the RSV-infected cells and a reduction in virus yield(Bitko et al., 2007). Munir *et al.(Munir et al., 2008)* also investigated whether the RSV NS1 and NS2 proteins affect the maturation of immature myeloid dendritic cells (DC) during RSV infection. It was found that these non-structural proteins mediate suppression of DC maturation which results in decreased antigen presentation and T lymphocyte activation leading to incomplete and/or weak immune responses that might contribute to RSV re-infection.

1.3 MOLECULAR EPIDEMIOLOGY OF RSV

It has been shown that the second hypervariable region (270 base pairs (bp)) which is at the C-terminal of the G protein, provide a reliable representation for the entire G gene variability(Peret et al., 1998). Sequence analysis of this hypervariable region is used in molecular epidemiological investigations of RSV epidemics. Previous studies conducted in SA from 1997-2000 indicated the circulation of three previously defined subtype A genotypes GA2, GA5 and GA7 and one distinct genotype named SAA1 (Venter et al., 2001). The subtype B strains clustered with two known genotypes GB3 and GB4, and three additional genotypes, SAB1-SAB3. It has been shown that these genotypes co-circulate but that dominant strains shift yearly probably due to herd immunity of the previous circulating strains which may cause the viruses to circulate more efficiently or be more pathogenic(Hall, 2001, Peret et al., 1998).Recent analysis of the circulation patterns and the evolution of subtype A and B strains have shown that both subtypes have a similar evolutionary rate however, subtype B strains can accommodate more drastic changes in their attachment proteins. In addition the most recent common ancestor of both subtypes was dated to the early 1940s (Zlateva et al., 2005b, Zlateva et al., 2004).

Many investigators have attempted to show an association of a specific strain with disease severity. Some have found no difference between subtype A and B, although others have reported that subtype A is more severe (Brandenburg et al., 2000, Cane, 2001, Savon et al., 2006). It was also shown that the two subtypes co-circulate but that subtype A was dominant in more epidemics(Cane, 2001). Venter et al. (Venter et al., 2002a) compared community circulating strains with strains from hospitalized patients and found that identical strains could cause LRTI as well as milder URTI. Viral load is another aspect that might influence disease severity. Gerna et al.(Gerna et al., 2008) found no difference in viral load between patients infected with RSV subtype A strains and those infected with subtype B strains while others have shown an association of subtype A strains with more severe disease (Walsh et al., 1997). Co-infection with patients with HIV AIDS also plays a role in disease severity. One study indicated the incidence for LRTI in children between 6 weeks and 2 years of age was much greater in HIV infected children (4040 per 100 000) then in HIV uninfected children (1753 per 100 000)(Madhi et al., 2001).

1.4 EPIDEMIOLOGY OF RSV

HRSV has a worldwide distribution. In the United States (US), approximately 75,000-125,000 hospitalizations associated with RSV occur each year of which the majority of children are under 6 months of age. Although the mortality rate for RSV infections has decreased significantly over the past 20 years, approximately 500 deaths still occur annually in the U.S., of which 80% of these deaths occur in children <1 year(Shay et al., 2001).

A few studies have characterized the disease burden of RSV in developing countries. One such study in Kilifi District, Kenya, has estimated that 85,000 infant cases of severe LRTI are due to RSV in Kenya(Nokes et al., 2008). Falsey *et al*.(Falsey et al., 2005) have shown that RSV infection in the elderly and high-risk adults, accounted for 10.6% of hospitalizations for pneumonia, 11.4% of hospitalizations for chronic obstructive pulmonary disease, 7.2% of hospitalizations for asthma, and 5.4% of hospitalizations for congestive heart failure. There has been a great deal of interest in a possible link

between RSV infection and asthma. Sigurs *et al*.(Sigurs et al., 2005) investigated this possible link, and their results suggested that severe RSV infection in the first year of life not only affected the predisposition to developing asthma, but to allergic disease as well.

Additionally RSV is an important nosocomial infectious agent and the rate of hospital acquired infection for infants and young children during a RSV season ranged from 26 to 47% in newborn units and from 20 to 40% for older children. Hospital staff members have been shown to play an important role in the transmission and it has been shown that they have an infection rate of 25% to 50% during RSV epidemics(Hall et al., 1975). HRSV has a clear seasonality and epidemics occur yearly in the winter months, whereas outbreaks are associated with the rainy season in humid climates(Cane, 2001). Mark Everard (Sheffield Children's Hospital, Sheffield, UK) observed RSV replication in dendritic cells during RSV seasons(Stevens et al., 2008). Following the administration of nitric oxide to the cultures he also detected viral replication during non-RSV seasons. He suggested that dendritic cells may function as reservoirs for RSV outside of the RSV season(Stevens et al., 2008).

1.5 CLINICAL INFECTION AND DISEASE PRESENTATION

Inoculation of the nose or eyes occurs by a colloidal dispersion of liquid particles, such as sneezing, or direct contact and results in viral replication in the nasopharynx, with an incubation period of 4 to 5 days, and can be followed over the next several days by spread to the lower respiratory tract(McNamara and Smyth, 2002). The first step in viral replication is attachment of the viral particle to a host cell, generally in the nasal epithelium. The viral RNA then enters the cell where it replicates and ultimately destroys the cell. The rapid destruction of ciliated epithelial cells lining the airways causes the symptoms characteristic of RSV infection(Domachowske and Rosenberg, 1999). RSV grows optimal at a PH of 7.5 and is inactivated at temperatures between 50-60ºC. Although temperature sensitive, it can be recovered from countertops and rubber gloves for several hours although it is readily inactivated with soap, water and disinfectants(McCarthy and Hall, 2003).

Airway dysfunction and inflammation are two important aspects of the pathophysiology of severe RSV disease. RSV-induced airway dysfunction includes epithelial cell sloughing in the small airways, mucus-overproduction and airway reactivity, which all ultimately leads to compromised ventilation. All of these could be directly a cause of viral replication but may also be facilitated by inflammation(McNamara et al., 2005). Primary manifestations of RSV infections are bronchiolitis and pneumonia. In addition rhinorrhea, cough, nasal congestion and low-grade fever are common. More severe LRTI are associated with an increase in asthma and allergic sensitization(Sigurs et al., 2005). Infants may also develop tachypnea, dyspnea and intercostals retractions as well as difficulty in feeding(McCarthy and Hall, 2003). In addition, otitis media is a common complication of HRSV. This is due to Eustachian tube dysfunction, resulting in bacterial stasis in the middle ear and subsequent otitis media(Tomochika et al., 2009).

Extrapulmonary manifestations have also been reported for severe RSV infections. The most commonly reported cases of cardiovascular manifestation of RSV were interstitial myocarditis(Eisenhut, 2006) and multifocal arterial tachycardia(Donnerstein et al., 1994). Acute neurological signs and symptoms such as central apneas, seizures, feeding/swallowing difficulties and muscle tone abnormalities were found in some patients with RSV infection(Kho et al., 2004).

1.6 IMMUNOPATHOGENESIS

1.6.1 Immunology (Role of innate/adaptive immunity)

Figure 1.4 shows the host defence mechanisms that are activated during RSV infection and the overlapping aspects of the innate and acquired immunity involved. Macrophages, eosinophils, basophils, neutrophils and natural killer (NK) cells are important innate immune cells that play a vital role in controlling infection, although during an imbalanced response may exacerbate inflammation and contribute to disease pathogenesis(Tripp, 2004). Acquired immunity is mediated by B cell (humoral) and T cell (cell-mediated) responses. Clinical studies of RSV-infected patients indicated increased levels of Th2 cytokines and immunoglobulin E (IgE) suggesting an allergy like reaction with infection(Becker, 2006). Immunocompetent infants infected with RSV stop shedding

virus within 21 days following infection while immunocompromised children can shed virus for several months indicating the importance of the immune system to clear the virus(Hall et al., 1986).

Normally the T-helper 1 (Th1) response is activated with RSV infection which is associated with INF-γ production and an earlier CD8+ CTL response. INF-γ stimulates the production of immunoglobulin (Ig) which leads to an increase in antigen presentation, cytolytic activity, antiviral cytokines and the end result is an increase in viral clearance. IgA is largely associated with defence of the mucosal surfaces and is rapidly secreted in the upper airways following primary RSV infection in mice(Singleton et al., 2003). Even though IgA plays a role in preventing an initial RSV infection of the upper airways, it wanes over time and are largely undetected by 8 weeks postinfection in mice(Valosky et al., 2005). The IgM antibody response is induced in 5-10 days, and can be detected for three months after infection(van Drunen Littel-van den Hurk et al., 2007). RSV-specific IgG antibodies are found in high titers in human serum and are believed to play an important role in the clearance of virus from the lower airways (Siber et al., 1994). The IgG consists mainly of IgG1 and IgG3 subclasses which reach peak titres 20-30 days after the onset of symptoms(van Drunen Littel-van den Hurk et al., 2007).

Factors such as atopy, RSV-specific antigen priming, other antigenic influences and prematurity promote the activation of the Th2 pathway. The Th2 response is autostimulatory and also inhibits the Th1 response, becoming the dominant response in subsequent infections. The Th2 response leads to a CD4+ CTL response and the production of interleukin 4 (IL-4), IL-5, IL-10 and IL-13. IL-4 promotes IgE production and indirectly leads to histamine production by activating mast cells. IL-5 on the other hand promotes the recruitment of eosinophils. Both the histamine and the eosinophil products cause the asthma like symptoms seen with a RSV infection(Graham, 1996). This suggest that an excess of type 2 and/or a deficient type 1 immune response play a role in RSV pathogenesis(Fig. 1.5).

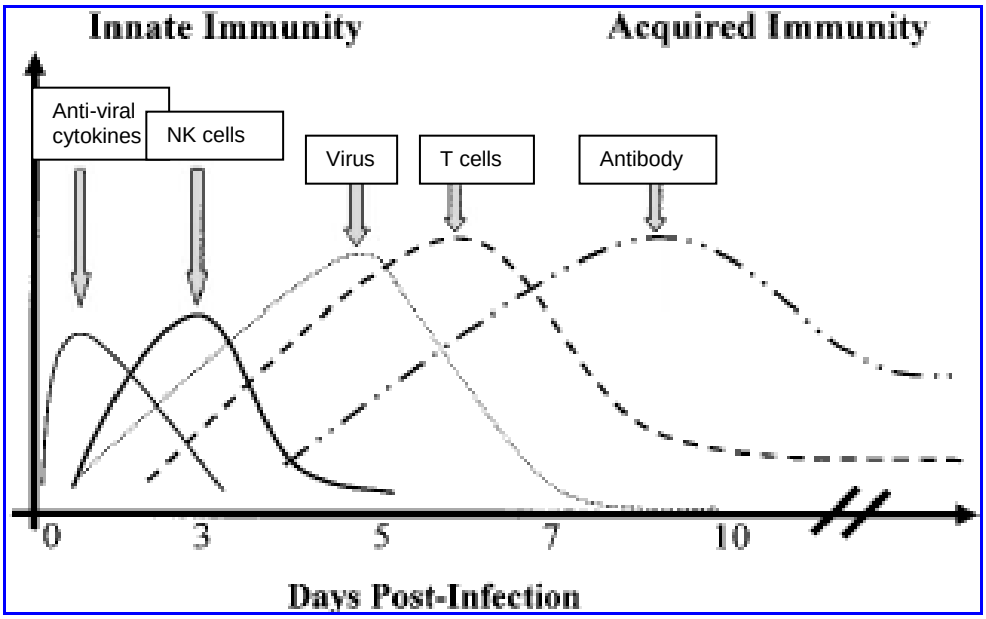


FIGURE 1.4 Host defence mechanisms against RSV infection involve overlapping aspects of innate and acquired immunity(Tripp, 2004).

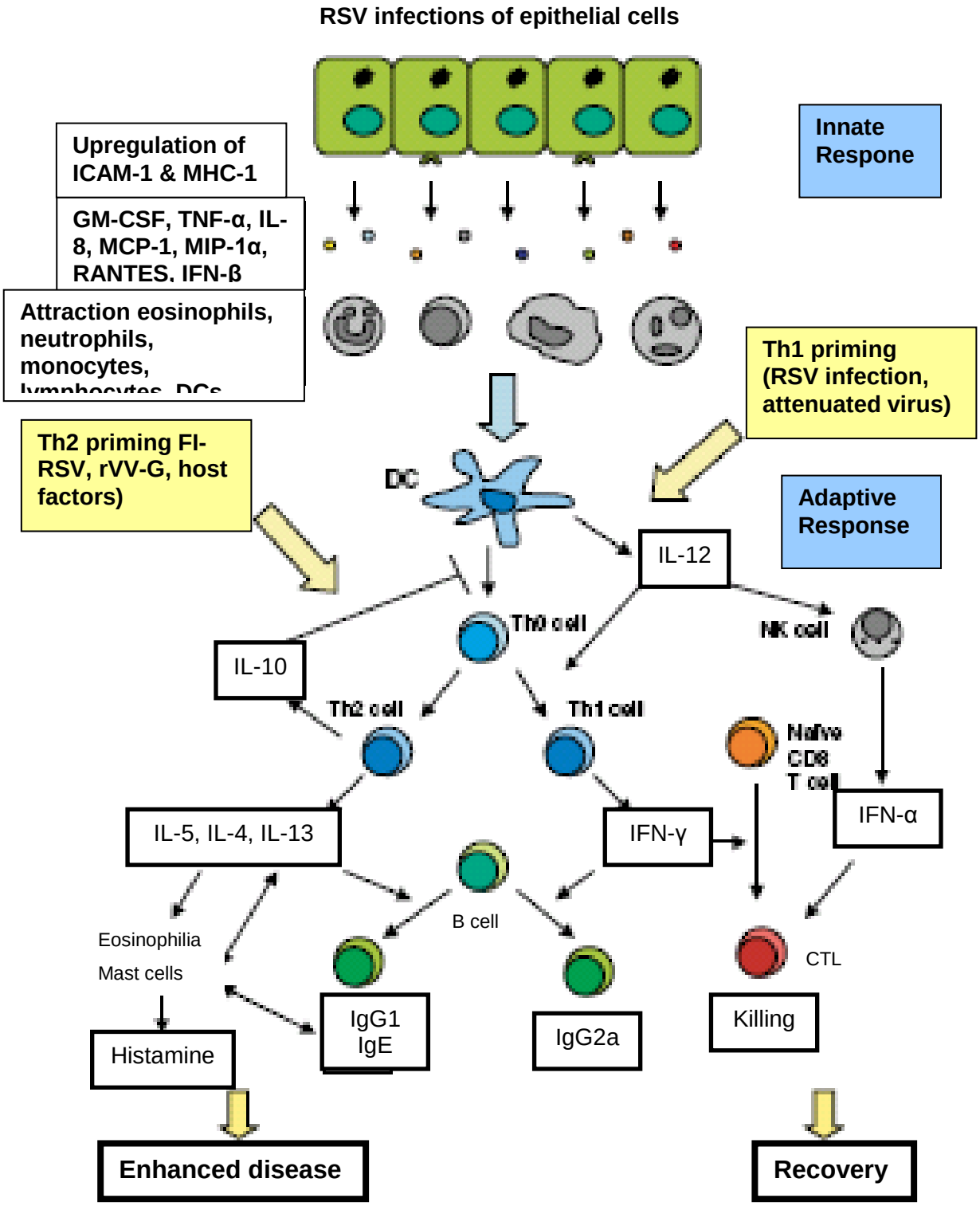


FIGURE 1.5 Response to RSV infection and vaccination. RSV infection leads to a balanced Th1 response, with production of IgG2a, Th1 cells and CTLs, as well as NK cell activation followed by IFN-γ production. Vaccination with FI-RSV, rVV or host factors that lead to reduced IL-12 production results in a Th2-biased response, characterized by Th2 cells, IgG1 and IgE production, and eosinophilia(van Drunen Littel-van den Hurk et al., 2007).

1.6.2 Risk factors for enhanced disease

The risk of severe RSV disease is increased by factors that compromise the ability to control and withstand a respiratory tract infection which include: young age (<6 months), premature birth (<35 weeks of gestation), bronchopulmonary dysplasia, congenital heart disease, immunodeficiency or immunosuppression, the first or second RSV infection in life, unusually narrow airways, low birth weight, male gender, a low titer of RSV-specific serum antibodies (maternal antibodies), and frail old age(Welliver, 2003). Environmental factors, including ones that affect lung function (e.g. exposure to tobacco) or increase exposure to infection (e.g. day care, hospitalization, multiple siblings), also play a role.

A major risk factor in SA is co-infection with Human Immunodeficiency virus (HIV). Clinical investigation by Madhi *et al.(Madhi et al., 2006)* showed that the incidence for RSV associated LRTI in children between 6 weeks and 2 years of age was much greater in HIV infected children (4040 per 100 000) than in HIV uninfected children (1753 per 100 000). Although a prominent risk factor, HIV seroconversion is not confirmed until 18 months because of the presence of maternal antibodies before this time(Rakusan et al., 1991).

A factor that also influences disease severity is pre-existing serum RSV-neutralizing antibody concentrations. Infants have varying concentrations of these antibodies, either obtained from previous infections or transplacentally from the mother. Various studies have demonstrated that higher titers of these antibodies are associated with less severe RSV disease(DeVincenzo, 2005). Premature infants acquire less transplacental antibodies than full-term infants, thus placing premature infants at increased risk for severe RSV disease(DeVincenzo, 2005).

Variation in susceptibility also has a genetic basis. Goetghebuer *et al.(Goetghebuer et al., 2004)* provided evidence that polymorphisms in certain genes, usually Th2 cytokine genes, contribute to the type and severity of clinical disease caused by RSV infection. Janssen *et al.(Janssen et al., 2007)* performed a large-scale genetic association study that used a candidate gene approach. In total, 22 single nucleotide polymorphisms (SNPs) in 21 genes, demonstrated a significant association with severe RSV

bronchiolitis either at the allele or genotype level, or at both. The 4 SNPs with the strongest association are located in genes involved in innate immunity (i.e., the VDR, JUN, NOS2A, and IFNA5 genes). This indicates that innate immunity may not only lead to viral clearance but may also be involved in the development of excessive pathology and disease.

1.7 DIAGNOSIS OF RSV INFECTION

Many different laboratory techniques exist for detection of RSV. All the techniques require sterile collection of nasal secretions which will be the starting material. Several factors exist that may affect the sensitivity and specificity of an assay for viral detection. These include quality of the specimen, transport conditions, quality of the reagents, laboratory technician experience, inter- and intra-laboratory standardization, suitability of the assay to specific populations and prevalence of the viral agent in the community(Henrickson and Hall, 2007). Clinicians need to recognize the benefits and limitations of existing tests which are described in the next section.

1.7.1 Viral culture

Viral culture requires viable virus that will grow in cell culture. Cell lines commonly used for RSV propagation are HEp-2, HeLa, A549 and Vero. Although considered the gold standard of diagnosing RSV, cell culture has various disadvantages. RSV is extremely unstable, heat-labile, and sensitive to freeze-thawing and is therefore difficult to culture. In addition syncytial cytopathic effect only appears after 3 to 7 days in cell culture(Noyola and Demmler, 2000). The financial and labour costs per test are also greater and cell culture has relative poor sensitivity compared to nucleic acid amplification(Henrickson and Hall, 2007).

1.7.2 Serology/Antigen detection

Serology is not very helpfull for diagnosis of acute respiratory virus infections beacause 10-30% of patients with documented respiratory viral infections will remain serologically negative(Henrickson, 2004). Antigen-based assays include indirect and direct

immunofluorescent antibody tests, enzyme immuno-assays (EIA) and optical immuno-assays. These assays are widely used because they are inexpensive, easy to perform, and easy to interpret(Henrickson and Hall, 2007). Kumar *et al*.(Kumar et al., 1987) evaluated the effieciency of two RSV rapid tests, an EIA (Abbott Laboratories, North Chicago, III.) and a fluorescent-antibody assay (FA). Their results indicated that both these tests proved to be highly specific and sensitive, although antigen detection methods are thousand times less sensitive than molecular assays.

1.7.3 RT-PCR and real-time PCR

Reverse transcription-PCR (RT-PCR) is a highly sensitive and quick method for diagnosis of viral infection and has been used successfully in children with RSV(Freymuth et al., 1995, Henkel et al., 1997). Falsey *et al*.(Falsey et al., 2002) have also shown that RT-PCR is a better and more sensitive method than cell culture when testing for RSV in adults due to low viral titers.

RT-PCR is generally available and used in research settings. Real time RT-PCR has become the method of choice for quantification of RSV viral load in nasopharyngeal aspirates due to the following theoretic advantages: (i) a lower threshold of detection, (ii) the potential stability of the assay after specimen freezing and thawing, (iii) a less subjective assay readout, and (iv) an assay that is unaffected by therapeutic passive neutralizing antibodies or experimental antiviral agents(Perkins et al., 2005).

For quantification purposes, the use of fluorescently labelled sequence-specific hybridization probes is very specific and sensitive, although the synthesis and optimization of each probe for a gene of interest can be time-consuming and expensive. An alternative method is double-stranded DNA binding dye SYBR Green I. The major advantage of this method is that it is sequence independent and can be used with any primer pair although non-specific amplification will prevent accurate quantification(Simpson et al., 2000).

1.8 TREATMENT AND PREVENTION OF RSV DISEASE

Currently, supportive care is the only rational treatment for RSV disease. Research has revolved around steroids, bronchodilators, antiviral medication and a variety of other therapies however no treatment currently exists. Transmission of RSV can however be prevented by adhering to basic hygiene practices including regular hand washing, prevention of fomite spread, and patient isolation(Simon et al., 2006).

1.8.1 Antiviral therapy

Corticosteroid therapy has not proven to be efficacious for RSV infection(Buckingham et al., 2002) and the use of ribavirin for RSV infection remains controversial. Ventre and Randoph (Ventre and Randolph, 2004) performed a meta-analysis of ribavirin treatment in children under 6 months of age with RSV associated LRTI and showed that there was statistically insignificant trends towards improvement in mortality, length of hospitalization and mechanical ventilation. A humanized IgG mAb called Palivizumab is being used and supported by medical aids as prophylaxis in high risk children, but has significant cost implications(Vogel *et al.*, 2002).

Promising research with regards to antiviral drug development is small interfering RNA (siRNA). Zhang and Tripp (Zhang and Tripp, 2008) showed that siRNAs can be developed as effective antiviral drugs that can reduce viral load and parameters of pathogenesis without limiting the induction of the memory immune response. Although all of this is promising, the best way to prevent RSV will be to design an effective vaccine.

1.8.2 Vaccines

Timing of vaccination, administration route, risks and efficacy remains important questions in this arena. A formalin-inactivated RSV (FI-RSV) vaccine was developed during the 1960s that resulted in an aggravated reaction in immunized children upon re-

exposure to RSV(Kim et al., 1969). Because of the disastrous results of the FI-RSV vaccine, it is increasingly difficult to test the vaccines because of safety issues as well as inadequate animal models where RSV induce different disease phenotypes in different animal species(van Drunen Littel-van den Hurk et al., 2007). The improvement in RSV reverse genetics development has however led to a new generation of live virus vaccine candidates with a balance of attenuation and immunogenicity(Collins and Murphy, 2005)

Clinical trials have been reported for two subunit RSV protein vaccines, all involving vaccinees of 24 months or older. The first consists of a purified full-length F protein from RSV-infected cells(Hancock et al., 1995). The second protein vaccine is a recombinant protein produced in bacteria that consists of amino acids 130-230 of the G protein, BBG2Na(Power et al., 2001). Both of these vaccines were shown to be safe and moderately immunogenic in adults however fears over Th2 type responses similar to the formalin inactivated vaccine may keep them out of clinical trials in younger infants. Recombinant RSV viruses that lack either NS1 or NS2 were attenuated both in the upper as well as the lower respiratory tract of chimpanzees but induced serum nAb titers similar to wild-type RSV(Teng et al., 2001). These recombinant viruses are under consideration for vaccine candidates. A vaccine which contains 3 mutations in the SH gene is the first RSV vaccine candidate that appears to be appropriately attenuated in young infants(Karron et al., 2005).

Live-attenuated vaccines have also been developed by subjecting wild-type RSV to extensive cold-passage in vitro as well as chemical mutagenesis. One of the sufficiently attenuated mutants was evaluated in 1- to 2-month-old RSV-naïve infants(Wright et al., 2000). It was shown that the vaccinnees shed a moderate amount of virus with a significant rise in RSV-specific serum IgA antibodies, and were highly resistant to re-infection with a second vaccine dose given 1 month later. However, a few of the vaccinnees experienced brief nasal congestion following the initial immunization indicating that the virus is still not sufficiently attenuated. In addition temperature sensitive, live-attenuated vaccine candidates have been generated for several viruses. Viruses with temperature sensitive phenotypes are able to replicate in the upper respiratory tract, stimulating anti-viral immunity, but are restricted in the lower respiratory tract, reducing the incidence of pneumonia and bronchiolitis(Tang et al., 2002). RSV is a

high priority for vaccine development but several challenges have limited the developmental process.

1.8.1.2 Obstacles for vaccine development

1.8.2.1.1 Maternal antibodies/Immunologic immaturity/Waning immunity in elderly

Maternal antibodies present in infants may confer partial protection against RSV infection, but it may also suppress antibodies and T cell responses to primary RSV infection(Crowe and Williams, 2003). A paediatric vaccine would be advantageous because the peak incidence of severe RSV is at 2-7 months of age, although immunological immaturity of infants limits Ab responses to RSV infection and limits the immunogenicity of vaccines. Murphy *et al*.(Murphy et al., 1986) have shown that infants less than 8 months old produced lower levels of serum and nasal wash anti-RSV IgG and IgA than infants 9-21 months old. Re-infection by RSV is a common phenomenon, and shows that natural immunity is incomplete which also poses a challenge for vaccine development(Collins and Murphy, 2005). In the elderly population it has been shown that RSV-specific antibody responses remain intact however there is a defect in the ability to maintain adequate numbers of RSV-specific memory CD8 T cells(Olson and Varga, 2008). It is likely that different immunization strategies will have to be utilized for infants and the elderly.

1.8.2.1.2 Economic factors

The cost of vaccine development has reached astronomic proportions, currently estimated at $1.7 billion and it is time consuming, taking up to 10 years to produce a new vaccine(Maggon and Barik, 2004).

1.8.2.1.3 Viral factors

RSV has a RNA genome that accumulates mutations at a high rate because of the lack of proofreading mechanisms which presents as a challenge for vaccine development. Especially antigenic variability in the G protein complicates vaccine development because both subtypes would be beneficial in a vaccine(Sullender, 2000). RSV does not

replicate efficiently in vitro and careful handling is necessary to maintain infectivity(Collins and Murphy, 2002)., RSV interacts with cells and specific cellular proteins, making it difficult to obtain cell-free viral material. In addition RSV-host interactions involve a large and complex network of signalling pathways that play a role in RSV disease manifestation. Finally, the immunopathology of RSV is equal complex and includes the relatively unique phenomenon of vaccine-enhanced disease or ‘immunopotentiation’(Maggon and Barik, 2004).

1.9 PROJECT AIMS AND OBJECTIVES

1.9.1 Hypothesis

We hypothesize that genetic differences in the G and NS proteins of RSV strains may be detected in children with mild relative to severe disease and that overcoming of herd immunity through evolution results in severe disease in immunologically naïve individuals in South Africa. In addition, expression level differences of the NS1 and NS2 proteins may influence disease severity in infected children through their capability to suppress the IFN response.

1.9.2 Long term aim

To characterize/determine the association of RSV protein variation with differences in disease severity in children

1.9.3 Short term aim

To investigate recent evolutionary patterns in the G protein as well as the role G and NS protein variability play in disease pathogenesis. To determine if NS protein expression level differences play a role in RSV disease severity in children.

1.9.4 Specific objectives

- To identify G protein genotypes associated with mild/severe RSV disease in children over four consecutive years
- To characterize deletion mutant strains of which most of the G protein ectodomain are deleted and to investigate if these strains circulate in other patients
- To investigate the NS proteins as a determinant of RSV pathogenesis with respect to:
    - Genetic differences
    - Expression level differences

CHAPTER 2

Replacement of previously circulating Respiratory Syncytial Virus (RSV) subtype B strains with the BA genotype in South Africa

---

2.1 INTRODUCTION

RSV is a major cause of severe paediatric respiratory tract disease in infants, the immunocompromised and the elderly (Boeck, 1996, Falsey et al., 2005, Welliver, 2003). In moderate climates, RSV epidemics occur yearly in the winter months, whereas outbreaks are associated with the rainy season in humid climates(Cane, 2001). Although the mortality rate for RSV infections has decreased significantly over the past 20 years, approximately 500 deaths still occur annually in the U.S., of which 80% occur in children <1 year(Shay et al., 2001). Globally, the WHO estimates that RSV causes 64 million infections and 160 000 deaths annually(WHO, 2009). A few studies have characterized the disease burden of RSV in developing countries. A study in Kilifi, Kenya, estimated that 85,000 infant cases of severe LRTI were due to RSV per year (Nokes et al., 2008). Re-infection is known to occur throughout life. Children are infected in the presence of maternal antibodies and natural infection affords only partial protection (Sullender, 2000, Wertz and Moudy, 2004).

RSV has a single-stranded negative sense RNA genome containing 10 genes encoding 11 proteins(Collins and Jr Crowe, 2007). Two antigenic subtypes (A and B) exist with little cross protection (Anderson et al., 1985). Major antigenic differences between subtypes are a feature of the attachment protein G, a type II transmembrane glycoprotein with a conserved central region with four cystines postulated to be a receptor binding site. Variability is concentrated in two hypervariable regions of the ectodomain(Johnson et al., 1987b). Several G protein genotypes within the two subtypes have been identified(Peret et al., 2000). These genotypes show complex circulation patterns likely facilitated by herd immunity to certain genotypes (Peret et al., 2000, Peret et al., 1998). SAA1, SAB1, SAB2 and SAB3 were identified in SA (Venter et

al., 2001) and subsequently in various other geographic locations(Riccetto et al., 2009, Shobugawa et al., 2009). Venter *et al.*(Venter et al., 2002a) also showed that identical RSV genotypes were identified in different regions in SA during one season. A new BA genotype has been identified in Buenos Aires in 1999 characterized by a 60 nucleotide duplication starting after residue 791 of the G protein(Trento et al., 2003). Subsequently strains with this duplication have been found in clinical specimens from distantly related places in the world(Kuroiwa et al., 2005, Nagai et al., 2004, Sato et al., 2005, Zlateva et al., 2005b, Zlateva et al., 2004) including Kenya in East Africa(Scott et al., 2004b). This BA genotype was first detected in SA during investigation of a nosocomial outbreak in Pretoria in 2006, and motivated us to re-evaluate the current RSV molecular epidemiology in SA (Visser et al., 2008).

It is unclear why some children experience severe disease and others develop milder disease. It may be due to host factors, maternal immunity or to differences in the virus itself. The risk of severe RSV disease is increased by factors that compromise the ability to control and withstand a respiratory tract infection. Many studies have investigated these risk factors and the association with severe disease including strain variability and infection with HIV. Some have found possible relationships while others have failed to do so (Brandenburg et al., 2000, Madhi et al., 2001, Savon et al., 2006, Venter et al., 2002a). However, most of these studies only concentrated on one epidemic season.

Despite the importance of RSV as a respiratory pathogen, there is currently no vaccine or effective therapy available. Experimental vaccines have been based on prototype strains identified shortly after its discovery in the 1960's(Collins and Murphy, 2002, Crowe, 2001). Protective vaccines may however have to take recent strains into account based on drift demonstrated relative to these strains. To determine the evolution of subtype A and B genotypes in SA hospitals, the molecular epidemiology of RSV epidemics in children was investigated. Correlation between disease severity and infecting strains was also investigated.

2.2 MATERIALS AND METHODS

2.2.1 Study population

The study population consisted of all children < 1 year of age who were diagnosed between February 2006 and May 2009 with RSV by immunofluorescence (IF) with the Respiratory Panel 1 IF Assay (Chemicon, Hampshire, UK) or with the Directigen^{TM} RSV rapid test (Becton Dickinson Microbiology Systems, Franklin Lakes, NJ), at the department Medical Virology University of Pretoria National Health Laboratory Service (NHLS), Tshwane academic division, and confirmed to have RSV by RT-PCR at our laboratory. The NHLS serves three secondary and tertiary hospitals in the Pretoria region and acts as referral hospitals for the Northern parts of SA: Kalafong Secondary-, Steve Biko Academic- and the 1-Military Hospital. Only children with single RSV infections were included in the disease severity assessments. Patients were divided into two classes: Mild disease was classified as outpatients not requiring hospitalization and severe disease were patients requiring hospitalization with severe LRTI or requiring intensive care treatment. This study was reviewed, approved and monitored by the ethics committee, University of Pretoria (25/2006).

2.2.2 Laboratory analysis

Specimens were confirmed by RT-PCR and screened for co-infections by multiplex PCR that detects 14 different viruses(Lassauniére et al., 2010). Specimens were labelled according to the following system: SA (isolate number) (Out-patient/Hospitalized) (year of isolation).

2.2.3 RNA isolation

RNA was extracted directly from nasopharyngeal aspirates (NPA) with QIAamp viral RNA mini kit according to the manufacturer's recommendations (Qiagen, Valencia, CA).

2.2.4 Subtype identification: Multiplex RT-PCR and Nested PCR

The specimens were subtyped with a multiplex RT-PCR method with the Titan One-tube RT-PCR system (Roche Applied Science, Mannheim, Germany) as described before (Venter et al., 2002a). Each PCR reaction contained 10µl of RNA template, 1µl dNTP mix (10mM), 2µl of 20pmol primer (F164, G32B, G267A), 2.5µl DTT (100mM), 0.25µl

RNase Inhibitor (40U/μl), 10μl of 5xRT-buffer and 1μl enzyme mix in a final reaction volume of 50μl. PCR cycling conditions: 50°C for 30min, 94°C for 2min, (94°C for 30sec, 54°C for 30sec, 68°C fir 1min) X35 cycles, 68°C for 7min.

The multiplex nested PCR were carried out using the Expand High Fidelity^{PLUS} PCR system (Roche Applied Science, Mannheim, Germany)(Venter et al., 2002a). Each nested PCR reaction contained 5μl of first round RT-PCR product, 2μl dNTP mix (10mM); 2μl of 20pmol primer (G52B, G283A, and F1), 20μl of 10x PCR reaction buffer and 0.75μl of enzyme mix in a final reaction volume of 100μl. PCR cycling conditions: 94°C for 2min, (94°C for 30sec, 55°C for 1min, 72°C for 1min) X35 cycles, 72°C for 7min. For specimens that could not be detected, the annealing temperature was lowered to 51°C.

2.2.5 Full length G protein amplification

Full length G protein were amplified with primers: G1-21(Trento et al., 2006) and F164(Sullender et al., 1993). In brief, 10μl of RNA was added to 10μl 10x reaction buffer, 1μl of dNTP mix (10mM), 2μl of 20pmol of each primer (G1-21, F164), 2.5μl DTT solution (100mM), 0.25μl RNase inhibitor (40U/μl), and 1μl of the Titan^{TM} enzyme mix; the volume made up the 50μl with distilled water. Cycling conditions were as follows: 50ºC for 30min, 94ºC for 2min, (94ºC for 10sec, 53ºC for 30sec, 68ºC for 1min) X35 cycles, 68ºC for 7min. Nested PCR for weak specimens with primers: G32B(Sullender et al., 1993) and F1(Peret et al., 1998). The nested PCR was conducted in a 100-μl reaction. 2μl of the RT-PCR product was mixed with 2μl of dNTP mix (10mM), 2μl of 20pmol of each primer (G32B, F1), 1μl of Expand high-fidelity PCR enzyme mix, 10x Expand HF buffer with 15mM $MgCl_2$ according to the following conditions: 94ºC for 2min, (94ºC for 30sec, 53ºC for 30sec, 72ºC for 1min) X35 cycles, 72ºC for 7min. All the PCR products were analyzed on a 1.5% agarose gel, using a 100-bp ladder as molecular weight marker (DNA molecular marker XIV, Roche Diagnostics, Mannheim, Germany).

2.2.6 Nucleotide sequencing

The Wizard SV Gel and PCR Clean-up system were used for PCR product purification (Promega, Madison, WI). Cycle sequencing was performed with the BigDye Terminator 3.1 Cycle Sequencing kit (Applied biosystems, Foster City, CA). Sequencing primers for the C-terminal of the G protein gene were as described in Venter *et al*.(Venter et al.,

2002a) and are shown in Table 2.1. For full G protein nucleotide sequencing, four primers were used to cover the entire G protein gene. Forward primers G32B as mentioned earlier and G604B (Venter et al., 2002a) and reverse primers F1 and G665R(Venter et al., 2002a) were used (Table 2.1). Nucleotide sequencing was carried out on both strands and the editing was performed with Sequencher^{TM} Version 4.6, (Gene Codes Corporation, Ann Arbor, MI).

TABLE 2.1 Primers selected for amplifying and sequencing the G protein gene.

Primer	Orientation	Sub-type	PCR/ Seq	Primer sequence (5' to 3')	Reference
G267A	Forward	A	RT	GATGCAACAAGCCAGATCAAG	(Sullender et al., 1993)
G32B	Forward	B	RT	GCAACCATGTCCAAACACAG	(Sullender et al., 1993)
F164	Reverse	A & B	RT	GTTATGACACTGGTATACCAACC	(Sullender et al., 1993)
G283A	Forward	A	N	CAAGAACACAACCCCAACAT	(Venter et al., 2002a)
G52B	Forward	B	N	AATCAACGCACTGCCAGKACTC (Where K=G/T)	(Christensen et al., 1999)
F1	Reverse	A & B	N&S	CAACTCCATTGTTATTTGCC	(Peret et al., 1998)
G1-21	Forward	A & B	RT	GGGGCAAATGCAACCATGTCC	(Trento et al., 2006)
G598A	Forward	A	S	GGAAAGAAAACCACCACCAA	(Venter et al., 2002a)
G604B	Forward	B	S	AAACCAACCATCAAACCCACA	(Venter et al., 2002a)
G665R	Reverse	A & B	S	TTTTGGGGCTCTTTTGTTTG	(Venter et al., 2002a)

G=Glycoprotein G ; RT=Reverse transcription ; N=Nest ; S=Sequencing

2.2.7 Multiplex real-time RT-PCR for identification of co-infections

In brief, cDNA was synthesized using Expand Reverse Transcriptase (Roche, Mannheim, Germany), according to the manufacturer's instructions. Multiplex Real-time PCR reactions were performed using the Light Cycler FastStart DNA Master^{PLUS} HybProbe Kit (Roche, Manheimm, Germany). These four multiplex real-time RT-PCR assays utilize fluorescence resonance energy transfer (FRET) hybridization probes to detect 13 different respiratory viruses(Lassauniére et al., 2010).

2.2.8 Phylogenetic Analysis

A region spanning 270 nucleotides (330 nucleotides for the BA strains), representing the second hypervariable region of the G protein gene was used for phylogenetic analysis as described before (Peret et al., 1998). This region corresponds to nucleotides 649-918 of prototype strain A2 for subtype A and 652-921 of prototype strain 18537 for subtype B(Johnson et al., 1987b). SA sequences were compared against reference sequences of each genotype, available in the GenBank database (http://www.ncbi.nlm.nih.gov/Genbank/index.html). Nucleotide sequences of subtype A and B viruses were aligned separately with Clustal X 1.81, using the multiple alignment option(Thompson et al., 1997). Midpoint-rooted maximum likelihood trees were constructed under the HKY codon position substitution model using PhyML(Guindon et al., 2005). Confidence estimates were based on bootstrap re-sampling carried out with 1000 replicates. Nucleotide accession numbers: HQ711628-HQ711839 (partial G protein sequences); JF704219-JF704234 (full G protein sequences).

Bayesian evolutionary analysis was carried out with BEAST (Bayesian Evolutionary Analysis by Sampling Trees) version 1.4.8(Drummond and Rambaut, 2007) using the Bayesian Markov chain Monte Carlo method (MCMC) to generate maximum clade credibility trees with the Tree annotator software and showing the highest posterior probability values for each node of the tree for the BA genotype identified in SA from 2006 to 2009 together with 63 unique BA sequences from the rest of the world since 1999. According to Trento *et al*.(Trento et al., 2006) the BA genotype could be further divided into six different branches (sub-genotypes) (I-VI). These guidelines were used to assign SA BA strains to specific sub-genotypes. All the trees were plotted using Figtree version 1.1.2 (http://O-tree.bio.ed.ac.uk.innopac.up.ac.za/software/figtree).

2.2.9 Nucleotide and amino acid sequence analysis

Pairwise distances (P-distance) were calculated between individual genotypes as well as within each genotype using Mega version 4(Tamura et al., 2007). Positive selection was calculated by Nei/Gojobori method indicated by Ka/Ks >1, where Ka is the non-synonymous substitution rate and Ks the synonymous substitution rate(Nei and Gojobori, 1986). BioEdit Sequence Alignment Editor version 7.0.4.1 was used for amino acid analysis(Hall, 1999).

2.2.10 Evolutionary rate and date of most recent common ancestor (MRCA)

The evolutionary rate was calculated using a Bayesian MCMC method available in the BEAST package, assuming an uncorrelated relaxed lognormal molecular clock that assumes independent rates on different branches(Drummond and Rambaut, 2007). MCMC chains were run for sufficient time to achieve convergence (as assessed using TRACER programme). Statistical uncertainty in parameter estimates is given by the 95% highest probability density (HPD) values. The root-to-tip regression plot displays the correlation between phylogenetic branch length and the time of sampling of the viral strains. This was performed with Path-O-Gen version 1.1.2 (http://tree.bio.ed.ac.uk/software/pathogen). The crossing point was taken as the date of the MRCA for all the sequences under analysis.

2.2.11 Statistical analysis

Association between genetic variability and disease severity was calculated with Fisher exact test using STATA 10 (Washington, USA). P values <0.05 were considered statistically significant.

2.3 RESULTS

2.3.1 Frequency of RSV subtypes and genotypes over four consecutive years in SA

Out of 621 specimens that tested positive for RSV at the NHLS Tshwane diagnostic laboratory over the four years (2006-2009), 245 specimens met the criteria for RT-PCR and 209 could be amplified. Of the 209 patients, 81 (39%) were HIV seropositive (54 were HIV DNA PCR negative and 27 DNA PCR positive for HIV); 26 (12%) were HIV negative and 102 (49%) had unknown HIV status. One hundred and eighty-five (89%) patients had single RSV infections, while 24 (11%) had co-infections with other viruses as tested with a respiratory multiplex PCR(Lassauniére et al., 2010).

Figure 2.1 indicate the RSV subtyping multiplex nested PCR on an agarose gel. Most clinical specimens are only visible in the nested PCR due to low viral loads. Subtype A and B can be distinguished on the basis of size. Positive specimens resulted in PCR fragments of 950bp for RSV subtype A and 1200bp for subtype B strains in the first round and 728 and 946bp respectively for the nested PCR corresponding with the RSV prototype strain A2 and B1. Lane 2 in figure 2.1 indicates a dual infection with both subtypes. As shown in Table 2.2, subtype A was the more prevalent subtype in 2006, 2007 and 2008 whereas subtype B was more prevalent in 2009. Several dual infections with both subtypes were also seen in two of the four years.

Two subtype A genotypes were detected throughout the four years: GA2 and GA5 (Figure 2.2A_/Table 2.3). Genotype GA5 predominated the subtype A genotypes in 2006 with 88%, whereas GA2 predominated in 2007 and 2008 with 82% and 83% respectively (Table 2.3). SA sequences formed three separate clusters within genotype GA2 and two separate clusters within GA5. All SA subtype B isolates clustered within the BA genotype and had the characteristic 60 nucleotide duplication (Figure 2.2B_/Table 2.3).

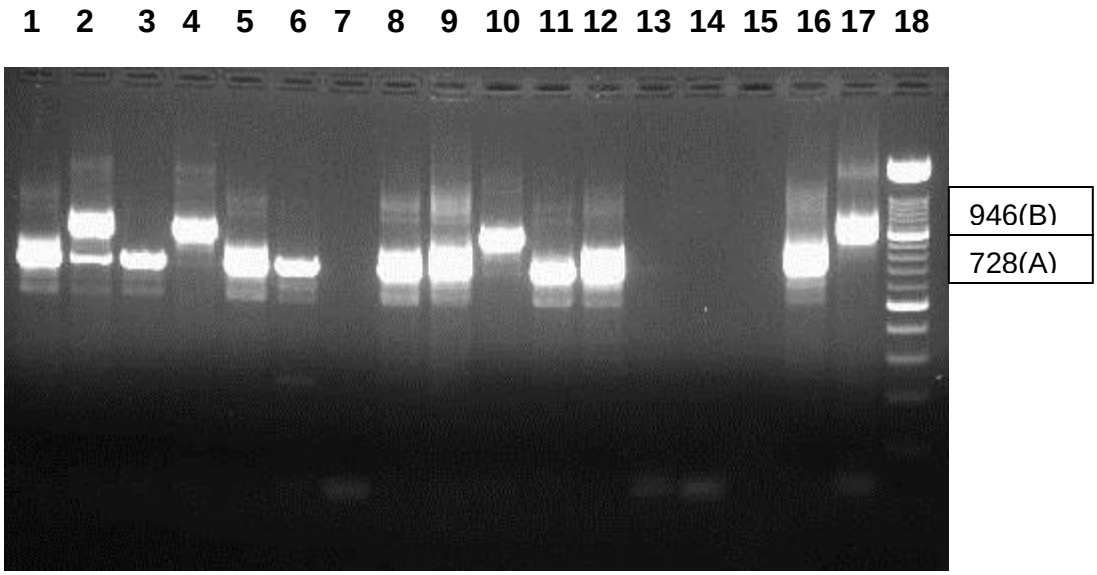


FIGURE 2.1 Agarose gel (1.5%) of the nested RT-PCR for RSV subtype A and B. Lanes 1-6 and 8-14: Clinical specimens; Lanes 7 and 15: Negative controls; Lane 16: Subtype A positive control (728bp); Lane 17: Subtype B positive control (946bp); Lane 18: 100bp Molecular weight marker.

TABLE 2.2 Frequency of RSV subtypes in SA during four consecutive years (2006-2009)

	2006 (76 SPECIMENS)	**2007 (74 SPECIMENS)**	**2008 (32 SPECIMENS)**	**2009 (25 SPECIMENS)**
Subtype A	**45 (59%)**	**67 (90%)**	**29 (91%)**	**4 (16%)**
Subtype B	**27 (36%)**	**5 (7%)**	**3 (9%)**	**21 (84%)**
Dual infections	**4 (5%)**	**2 (3%)**	**0**	**0**

TABLE 2.3 Frequency of RSV genotypes in SA during four consecutive years (2006-2009)

	2006	**2007**	**2008**	**2009**
Subtype A	**49**	**69**	**29**	**4**
GA2	**4 (8%)**	**56 (81%)**	**24 (83%)**	**4 (100%)**
GA5	**43 (88%)**	**13 (19%)**	**5 (17%)**	**0**
Untypable*	**2 (4%)**	**0**	**0**	**0**
Subtype B	**31**	**7**	**3**	**23**
BA	**30 (97%)**	**7 (100%)**	**3 (100%)**	**23 (100%)**
Untypable*	**1 (3%)**	**0**	**0**	**0**

*** Viral load too low for sequencing**

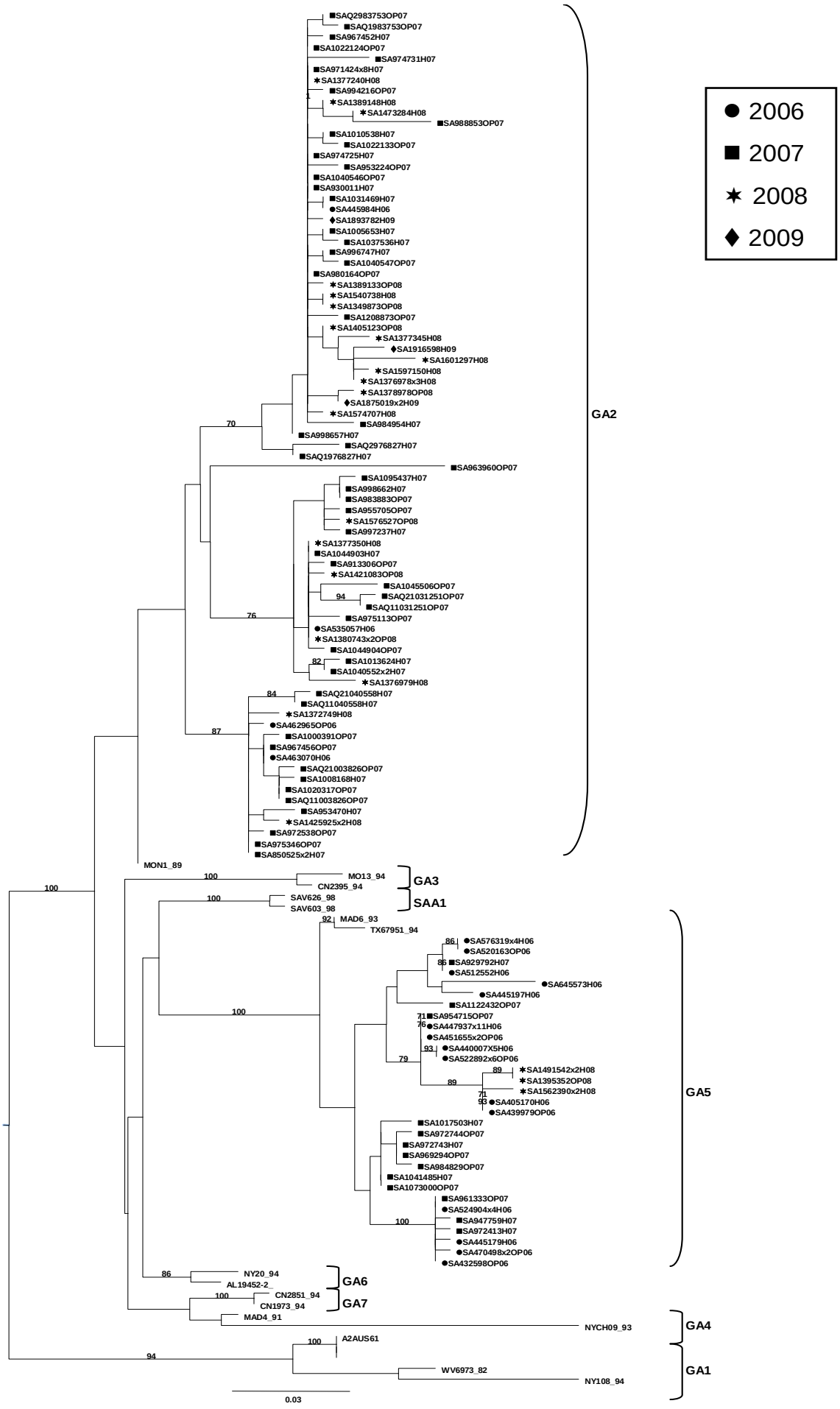

● 2006
■ 2007
✶ 2008
♦ 2009
GA2
GA3
SAA1
GA5
GA6
GA7
GA4
GA1
MON1_89
MO13_94
CN2395_94
SAV626_98
SAV603_98
MAD6_93
TX67951_94
NY20_94
AL19452-2_
CN2851_94
CN1973_94
MAD4_91
NYCH09_93
A2AUS61
WV6973_82
NY108_94
0.03

FIGURE 2.2 A

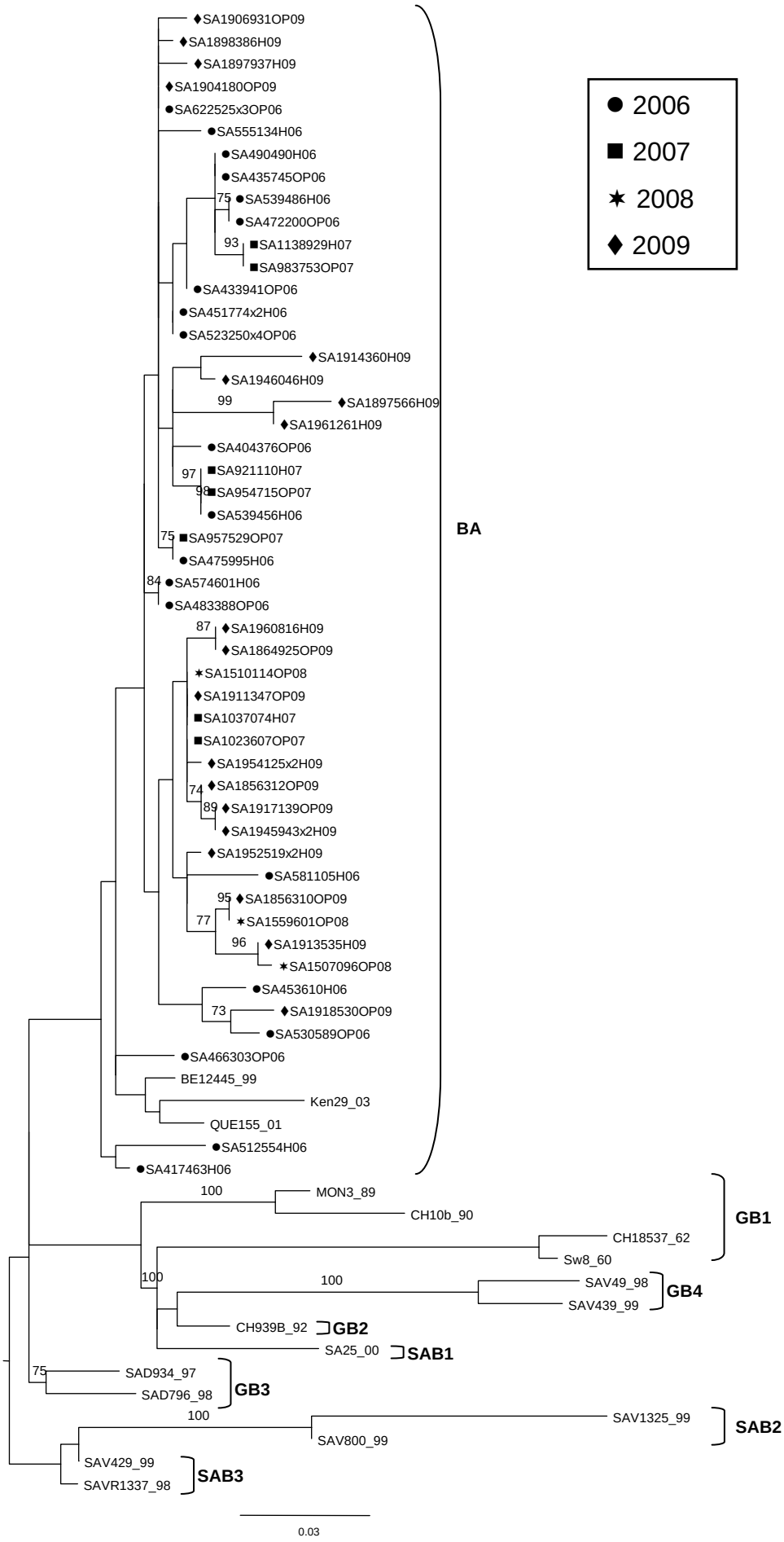

● 2006
■ 2007
✶ 2008
♦ 2009
♦SA1906931OP09
♦SA1898386H09
♦SA1897937H09
♦SA1904180OP09
●SA622525x3OP06
●SA555134H06
●SA490490H06
●SA435745OP06
75 ●SA539486H06
●SA472200OP06
93 ■SA1138929H07
■SA983753OP07
●SA433941OP06
●SA451774x2H06
●SA523250x4OP06
♦SA1914360H09
♦SA1946046H09
99 ♦SA1897566H09
♦SA1961261H09
●SA404376OP06
97 ■SA921110H07
98 ■SA954715OP07
●SA539456H06
75 ■SA957529OP07
●SA475995H06
84 ●SA574601H06
●SA483388OP06
87 ♦SA1960816H09
♦SA1864925OP09
✶SA1510114OP08
♦SA1911347OP09
■SA1037074H07
■SA1023607OP07
♦SA1954125x2H09
74 ♦SA1856312OP09
89 ♦SA1917139OP09
♦SA1945943x2H09
♦SA1952519x2H09
●SA581105H06
95 ♦SA1856310OP09
77 ✶SA1559601OP08
96 ♦SA1913535H09
✶SA1507096OP08
●SA453610H06
73 ♦SA1918530OP09
●SA530589OP06
●SA466303OP06
BE12445_99
Ken29_03
QUE155_01
●SA512554H06
●SA417463H06
BA
100 MON3_89
CH10b_90
CH18537_62
Sw8_60
GB1
100
100 SAV49_98
SAV439_99
GB4
CH939B_92
GB2
SA25_00
SAB1
75 SAD934_97
SAD796_98
GB3
100 SAV1325_99
SAV800_99
SAB2
SAV429_99
SAVR1337_98
SAB3
0.03

FIGURE 2.2 B

FIGURE 2.2 Midpoint-rooted maximum likelihood trees constructed under the HKY codon position substitution model using PhyML for **A)** subtype A and **B)** subtype B(Guindon et al., 2005). Trees were drawn to scale with the bars indicating 0.03 nucleotide substitutions. Estimates were based on bootstrap re-sampling carried out with 1000 replicates. Only bootstrap values >70 are shown. Names of the viruses refer to the place/number/disease severity/year of isolation. The number of identical sequences is indicated as x n. The genotypes assigned are indicated at the right by brackets. The reference sequences obtained from GenBank were as follows: from the USA (NY, New York; Al, Alabama; MO, Missouri; Tx, Texas; CN, Canada; CH, Rochester, New York)(Peret et al., 2000, Peret et al., 1998). WV, West Virginia(Sullender and Wertz, 1991); Montevideo, Uruguay (MON) and Madrid, Spain (MAD)(Garcia et al., 1994, Martinez et al., 1999). Prototype strains for subtype A: strain A2 (Australia)(Wertz et al., 1985); subtype B: Swed8-60 (Sweden)(Sullender and Wertz, 1991); USA (18537)(Johnson et al., 1987b).

2.3.2 Genetic characterization of the SA subtype A strains

Subtype A GA2 and GA5 strains identified in epidemics of 1997-2000 are still circulating in SA although the previously identified genotype SAA1 was not detected again. Indicated in table 2.4, intragenotypic P-distances for the SA strains were 0.040 and 0.030 for GA2 and GA5 respectively and the average intergenotypic P-distance between the different genotypes within subtype A fell in the range of 5 to 16% with 11% difference between the two dominant genotypes found during this study period, GA2 and GA5.

Figure 2.3 A and B indicate amino acid alignments of genotype GA2 and GA5 isolated during 2006-2009 in SA relative to the A2 reference strain. The names on the left are ordered according to genotype, year and disease severity. All the detected genetic changes were base substitutions, and no deletions, insertions, or frame shift mutations were observed. The percentage of non-synonymous substitutions relative to synonymous substitutions among SA subtype A isolates were 74% within the C-terminal of the G protein indicating a selective pressure for change. Amino acid changes have been associated with epitope loss in escape mutants and natural isolates of human RSV(Garcia et al., 1994). Three amino acid changes associated with epitope loss in natural isolates were seen in the SA strains: 233 E-K corresponding to position 21 in figure 2.3 mapped to MAb 68G(Rueda, Unpublished data); 280 S-Y and 293 P-S corresponding to position 68 and 81 respectively mapped to MAb 78G(Rueda, Unpublished data); and one amino acid change associated with epitope loss in escape mutants as well as natural isolates, 265 F-L corresponding to position 53 in figure 2.3 mapped to MAb 25G(Rueda et al., 1991). The average intragenotypic amino acid p-distances were 0.076 and 0.041 for GA2 and GA5 respectively while changes of 5 to 31% were seen between these and other subtype A genotypes. Another mechanism of RSV to generate diversity is the use of alternative termination codons. Subtype A isolates utilised four termination codons.

In total 18% used the UAG stop codon, 35% used UGA and 2% used UAA, all at position 909, while 45% used stop codon UAG at position 912. Using the Nei/Gojobori method, positive selection was shown to occur within genotype GA2 with Ka/Ks=1.08 but not for GA5 (Table 2.5). Using BEAST, the evolutionary rate of SA subtype A strains

were estimated as 3.382 x 10^{-3} (95% HPD: 1.911 to 4.954 x 10^{-3}) substitutions/site/year. A root-to-tip transgression plot estimation of SA subtype A strains indicated the MRCA to date back to 1980 (Figure 2.4).

TABLE 2.4 Average nucleotide p-distances between and within subtype A SA strains **(A)** separately and **(B)** together with the reference strains used in figure 2.2 A.
P-distances calculated between the individual genotypes are shown below the diagonal in each column of the table, and the p-distances within each genotype are in bold on the diagonal.

A) South African strains only

	GA2	**GA5**
GA2	**0.040**	
GA5	0.110	**0.030**

B) South African strains together with the reference strains

	GA1	**GA2**	**GA3**	**GA4**	**GA5**	**GA6**	**GA7**	**SAA1**
GA1	**0.041**							
GA2	0.130	**0.040**						
GA3	0.128	0.079	**0.015**					
GA4	0.164	0.128	0.122	**n/c**				
GA5	0.148	0.109	0.112	0.128	**0.031**			
GA6	0.119	0.074	0.065	0.106	0.082	**0.019**		
GA7	0.126	0.085	0.076	0.106	0.094	0.054	**0.004**	
SAA1	0.129	0.084	0.085	0.107	0.087	0.050	0.069	**0.007**

A)

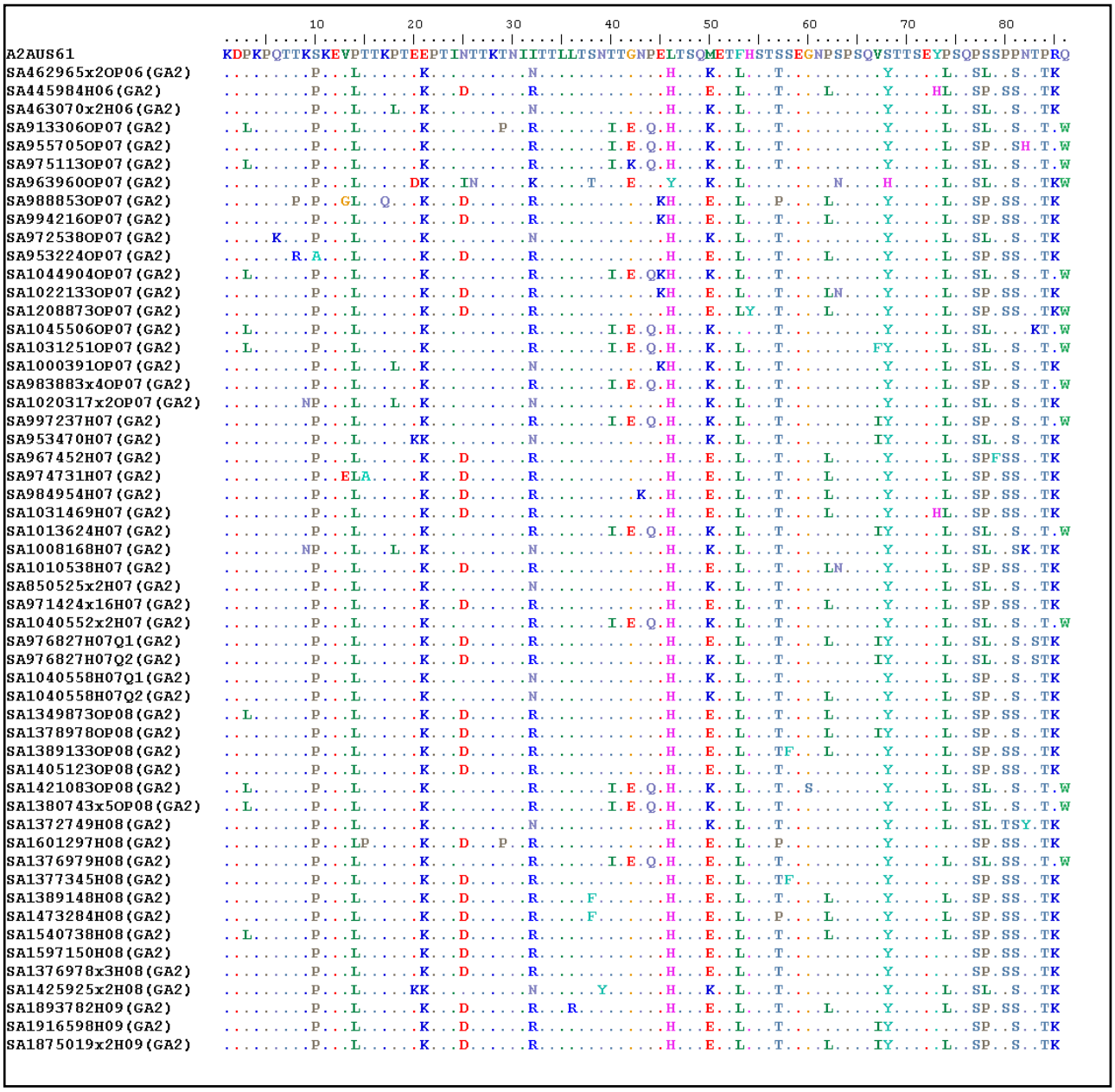


FIGURE 2.3 Amino acid alignment of the second variable region of the G protein for subtype A **A)** genotype GA2 and **B)** genotype GA5 SA strains relative to prototype sequence A2. H=Hospitalized patients with severe disease; OP=Out patients with mild disease.

B)

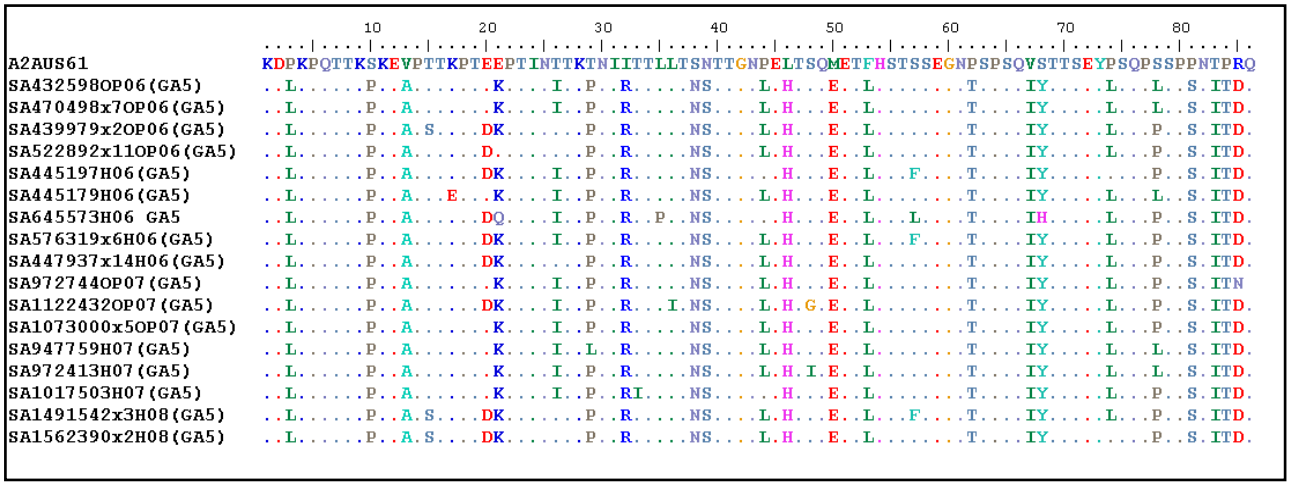


FIGURE 2.3 B Genotype GA5 strains detected in SA

TABLE 2.5 Estimates of selection using the modified Nei-Gojobori method(Nei and Gojobori, 1986) for Subtype A. Ka/Ks calculated between the individual branches are shown below the diagonal in each column of the table, and within each branch are in bold on the diagonal. Ka/Ks>1 are shaded.

Ka/Ks	GA1	GA2	GA3	GA4	GA5	GA6	GA7	SAA1
GA1	**0.356**							
GA2	1.032	**1.083**						
GA3	1.261	1.013	**1.417**					
GA4	1.190	1.439	1.206	**n/c**				
GA5	1.034	1.573	1.356	1.174	**0.404**			
GA6	0.784	0.946	0.725	1.202	1.153	**0.44**		
GA7	0.947	1.013	0.932	0.981	1.36	0.710	**0**	
SAA1	0.669	0.686	0.449	0.576	0.6	0.283	0.420	**0.5**

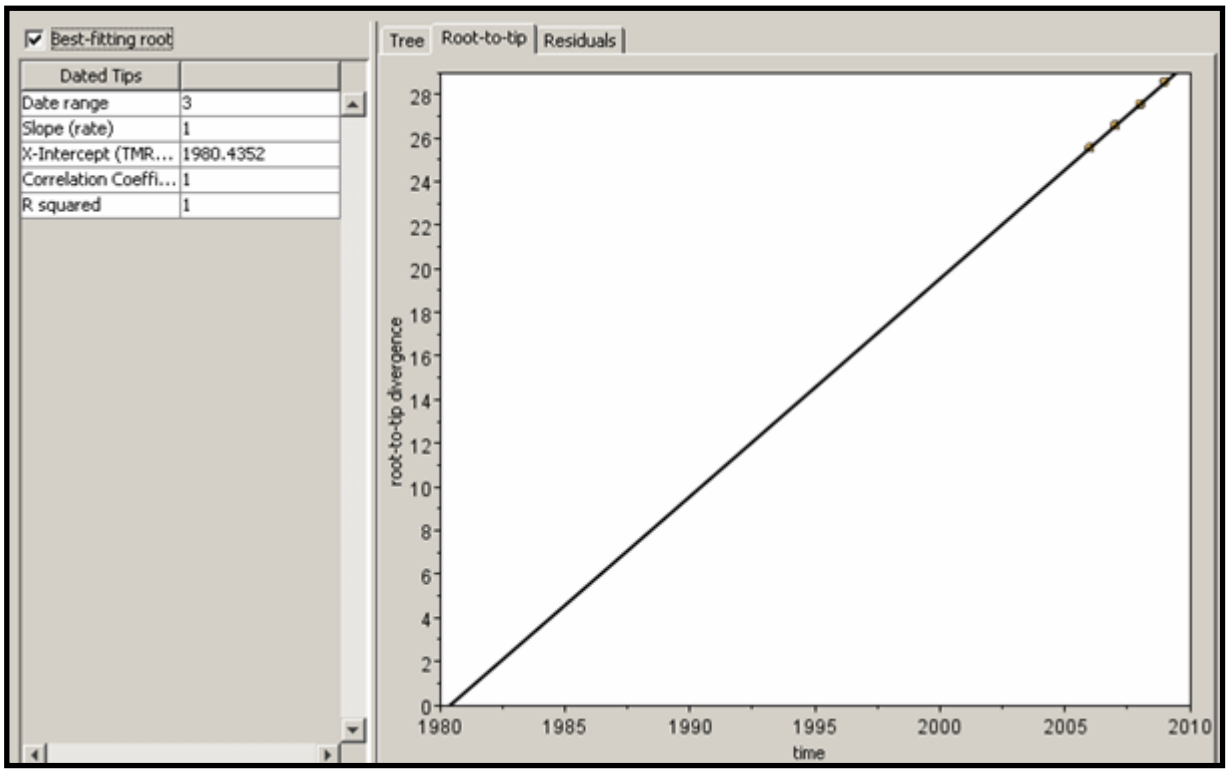


FIGURE 2.4 Root-to-tip transgression plot of subtype A SA strains. The phylogenetic branch length is plotted against the date of isolation. The crossing point is taken as the date of the most recent common ancestor.

2.3.3 Genetic characterization of the SA subtype B strains

All strains clustered with the BA genotype suggesting previously identified subtype B genotypes in SA have been replaced by the BA genotype since 2006. The overall nucleotide- and amino acid identity levels were shown to be 97.7% and 96.6% respectively for all the SA BA strains. It was shown that the same sequence in different patients was stable for three successive years. Using the reference strains in figure 2.1B, differences of 4 to 15% were seen between the subtype B genotypes. BA strains currently circulating in SA differs by 8%, 11% and 5% from SAB1, SAB2 and SAB3 that was previously identified in SA (Table 2.6).

Except for the 60 nucleotide duplication, only substitutions, and no deletions, insertions, or frame shift mutations were observed in the C-terminal of the G protein. The SA BA genotype had an average intragenotypic amino acid p-distance of 0.028 whereas the SA strains differed to other genotypes by between 6 and 27%. Figure 2.5 indicates the amino acid alignment of all the BA strains isolated during 2006-2009 in SA relative to the reference strains from each sub-genotype within the BA genotype. According to the modified Nei/Gojobori method, positive selection was not seen within the BA genotype (Table 2.7). The evolutionary rate of SA subtype B strains (BA genotype) were estimated as 2.305×10^{-3} (95% HPD: 1.112 to 3.665×10^{-3}) substitutions/site/year and a root-to-tip transgression plot estimation indicated the MRCA to date back to 1996 (Figure 2.6).

The full G protein was sequenced for selected BA strains of each of the four years including five sequences for 2006, three sequences for 2007, two for 2008 and six for 2009. Compared to other full G protein sequences on GenBank, all sixteen SA BA strains had a 6 nucleotide (2 amino acid) deletion at position 490 resulting in deletion of a Proline and Lysine. SA BA isolates used the UAA stop codon at nucleotide position 946 (79%) resulting in a protein length of 315 amino acid while 21% of the isolates used the UAG stop codon at position 967 resulting in a length of 322 amino acid.

TABLE 2.6 Average nucleotide p-distances between and within subtype B SA strains **(A)** separately and **(B)** together with the reference strains used in figure 2.2 B.
P-distances calculated between the individual genotypes are shown below the diagonal in each column of the table, and the p-distances within each genotype are in bold on the diagonal.

A) South African strains only

	BA
BA	**0.023**

Range: 0.003-0.064

B) South African strains together with the reference strains

	GB1	**GB2**	**GB3**	**GB4**	**SAB1**	**SAB2**	**SAB3**	**BA**
GB1	**0.087**							
GB2	0.079	**n/c**						
GB3	0.104	0.062	**0.036**					
GB4	0.153	0.097	0.130	**0.044**				
SAB1	0.099	0.048	0.073	0.111	**n/c**			
SAB2	0.148	0.097	0.104	0.139	0.117	**0.067**		
SAB3	0.103	0.060	0.042	0.121	0.075	0.087	**0.008**	
BA	0.101	0.060	0.051	0.136	0.083	0.114	0.050	**0.023**

FIGURE 2.5 Amino acid alignment of the second variable region of the G protein for subtype B SA strains relative to a reference strain from each sub-genotype within the BA genotype. H=Hospitalized patients with severe disease; OP=Out patients with mild disease

TABLE 2.7 Estimates of selection using the modified Nei-Gojobori method(Nei and Gojobori, 1986) for Subtype B. Ka/Ks calculated between the individual branches are shown below the diagonal in each column of the table, and within each branch are in bold on the diagonal. Ka/Ks>1 are shaded.

Ka/Ks	GB1	GB2	GB3	GB4	SAB1	SAB2	SAB3	BA
GB1	**0.763**							
GB2	0.937	**n/c**						
GB3	0.951	1.071	**0.973**					
GB4	1.256	1.259	1.085	**1.263**				
SAB1	1.163	1.486	1.074	1.269	**n/c**			
SAB2	0.793	0.924	0.585	0.958	0.900	**0.709**		
SAB3	1.143	1.271	0.796	1.054	1.491	0.658	**0.5**	
BA	1.021	1.163	0.852	1.278	1.45	0.674	1.163	**0.567**

Ka = non-synonymous substitution rate; Ks = synonymous substitution rate.

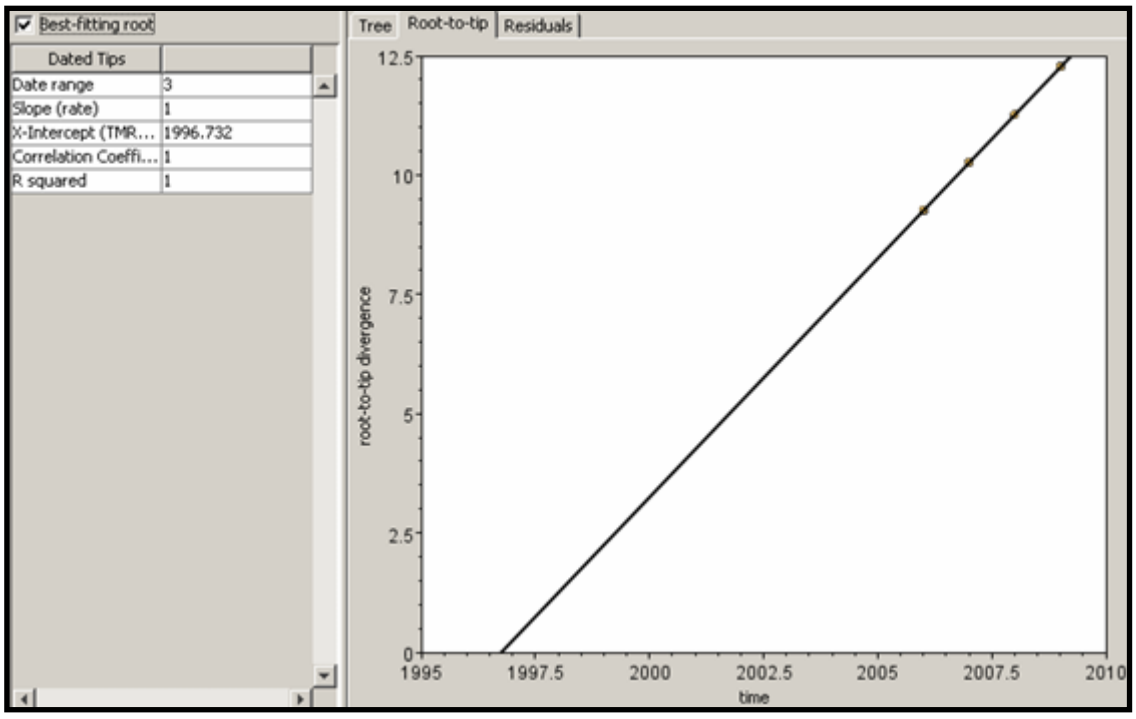


FIGURE 2.6 Root-to-tip transgression plot of all SA BA sequences used in this analysis. The phylogenetic branch length is plotted against the date of isolation. The crossing point is taken as the date of the most recent common ancestor.

2.3.4 Evolution of the newly identified BA genotype

Figure 2.7 illustrates Bayesian analysis of BA strains isolated over the world since 1999 together with all the BA strains isolated during this study period (2006-2009) in SA. All SA BA strains isolated during the four years clustered within sub-genotype four (BA-VI). Differences ranging from 0.3% to 6.4% were calculated between SA BA-IV strains. All the SA BA strains were unique with respect to the other strains within this group from other countries. SA strains were most closely related to strains from Brazil, Argentina and India. Strains from Belgium, Canada and Kenya within group IV clustered separately from the SA strains. Differences of between 1.7% and 4.7% were seen between the different BA branches. Positive selection was identified within BA-II, BA-III and BA-VI with Ka/Ks 1.25, 1 and 2 respectively. Positive selection was not identified within BA-IV (Table 2.8).

Genetic drift is visible within the BA genotype since the emergence eleven years ago as can be seen by the linear accumulation of genetic changes with time. The older strains are closer to the root then the more recently isolated strains. Drift is visible within the SA BA-IV strains. Numerous changes were observed between the two 60 nucleotide regions in the SA BA strains (Figure 2.8). Using BEAST, the average nucleotide substitution rate across all sites in the alignment for all BA sequences since 1999 was calculated as 3.143×10^{-3} (95% HPD: 2.41 to 3.911×10^{-3}) substitutions/site/year.

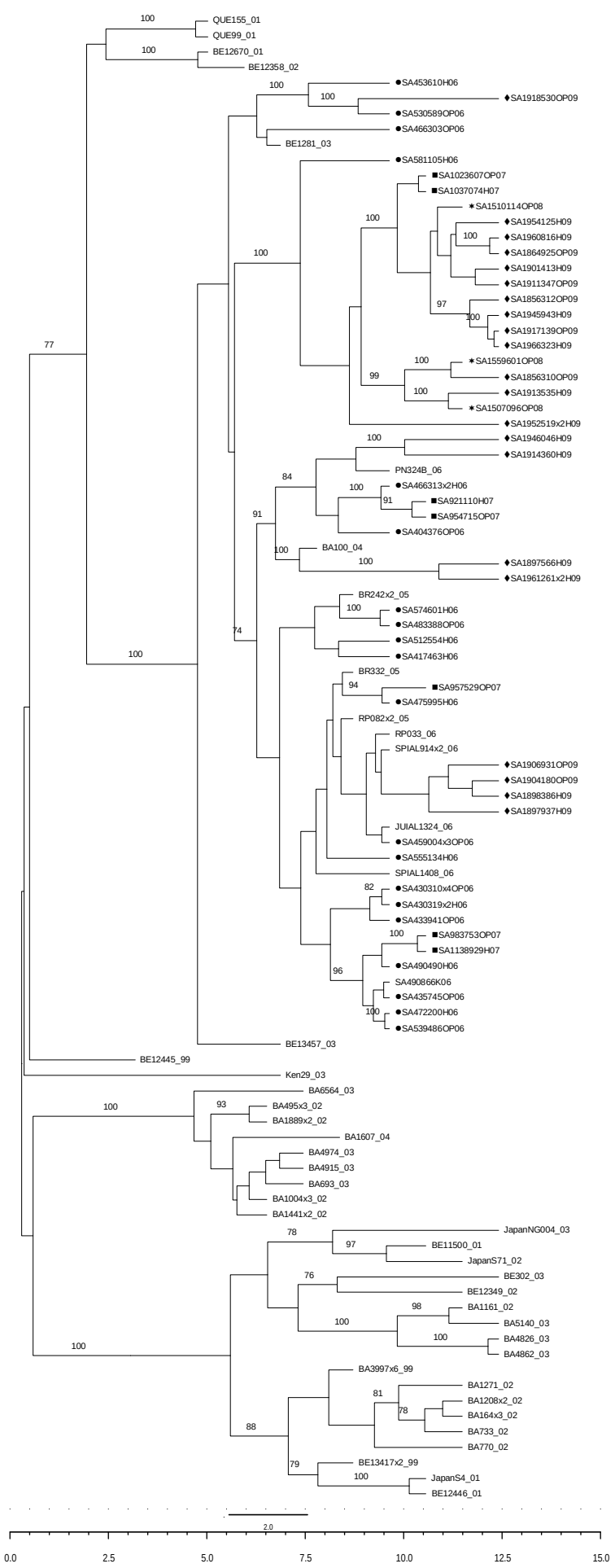

QUE155_01
QUE99_01
BE12670_01
BE12358_02
●SA453610H06
♦SA1918530OP09
●SA530589OP06
●SA466303OP06
BE1281_03
●SA581105H06
■SA1023607OP07
■SA1037074H07
✶SA1510114OP08
♦SA1954125H09
♦SA1960816H09
♦SA1864925OP09
♦SA1901413H09
♦SA1911347OP09
♦SA1856312OP09
♦SA1945943H09
♦SA1917139OP09
♦SA1966323H09
✶SA1559601OP08
♦SA1856310OP09
♦SA1913535H09
✶SA1507096OP08
♦SA1952519x2H09
♦SA1946046H09
♦SA1914360H09
PN324B_06
●SA466313x2H06
■SA921110H07
■SA954715OP07
●SA404376OP06
BA100_04
♦SA1897566H09
♦SA1961261x2H09
BR242x2_05
●SA574601H06
●SA483388OP06
●SA512554H06
●SA417463H06
BR332_05
■SA957529OP07
●SA475995H06
RP082x2_05
RP033_06
SPIAL914x2_06
♦SA1906931OP09
♦SA1904180OP09
♦SA1898386H09
♦SA1897937H09
JUIAL1324_06
●SA459004x3OP06
●SA555134H06
SPIAL1408_06
●SA430310x4OP06
●SA430319x2H06
●SA433941OP06
■SA983753OP07
■SA1138929H07
●SA490490H06
SA490866K06
●SA435745OP06
●SA472200H06
●SA539486OP06
BE13457_03
BE12445_99
Ken29_03
BA6564_03
BA495x3_02
BA1889x2_02
BA1607_04
BA4974_03
BA4915_03
BA693_03
BA1004x3_02
BA1441x2_02
JapanNG004_03
BE11500_01
JapanS71_02
BE302_03
BE12349_02
BA1161_02
BA5140_03
BA4826_03
BA4862_03
BA3997x6_99
BA1271_02
BA1208x2_02
BA164x3_02
BA733_02
BA770_02
BE13417x2_99
JapanS4_01
BE12446_01
2.0
0.0
2.5
5.0
7.5
10.0
12.5
15.0
●2006
■2007
✶2008
♦ 2009
BA-IV
BA-III
BA-VI
BA-II
BA-I

FIGURE 2.7

FIGURE 2.7 A midpoint-rooted BEAST phylogenetic tree of BA genotypes isolated over the world since 1999 together with all the BA genotypes isolated during this study (2006-2009) in SA(Drummond and Rambaut, 2007). The length of the horizontal lines is proportional to the genetic distance between viruses. The bar represents 2 substitutions per site. The numbers on the branches represent the highest posterior probability value for each node of the tree. Names of viruses refer to the place/number/disease severity/ year of isolation. SA strains were compared to strains from Buenos Aires (BA), Japan, Kenya (Ken), Quebec (QUE), Jundiai Brazil (JU), Sau Paulo Brazil (SP), Ribeirao Preto Brazil (RP), India (PN) and Belgium (BE)(Galiano et al., 2005, Scott et al., 2004b, Viegas and Mistchenko, 2005, Zlateva et al., 2007). The number of identical sequences is indicated as x n. Branches are indicated at the right by brackets. The scale at the bottom indicates years since first isolation.

TABLE 2.8 Estimates of selection using the modified Nei-Gojobori method(Nei and Gojobori, 1986) for the different genotype BA branches. Ka/Ks calculated between the individual branches are shown below the diagonal in each column of the table, and within each branch are in bold on the diagonal. Ka/Ks>1 are shaded.

KA/KS	BA-I	BA-II	BA-III	BA-IV	BA-VI
BA-I	**0.8**				
BA-II	1.56	**1.25**			
BA-III	1.03	1.3	**1**		
BA-IV	1.02	1.15	0.85	**0.6**	
BA-VI	1.06	1.46	0.8	0.86	**2**

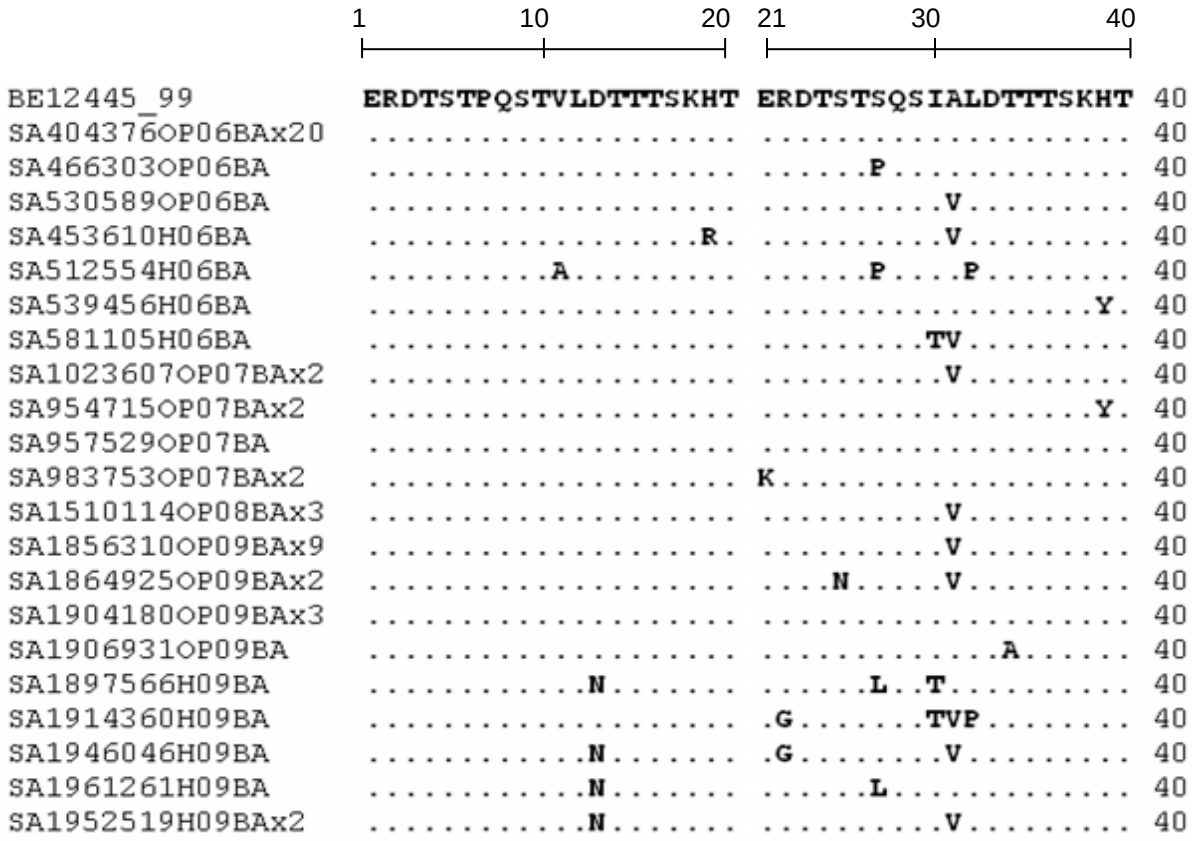


FIGURE 2.8 Amino acid alignment of the duplicated region within the BA genotype. The alignment is shown relative to the BE12445 strain isolated in Belgium in 1999.

2.3.5 Correlation of G protein variability with differing disease manifestations

Disease severity and infecting subtype or genotype were correlated by comparing out-patients with hospitalized patients. Twenty four of the specimens had co-infections with other viruses as tested with an in-house multiplex real-time RT-PCR(Lassauniére et al., 2010). These viruses include adeno virus (ADV), human metapneumo virus (hMPV), human boca virus (hBoV), Influenza A and B, para-influenza virus (PIV) 1 and 3 and human corona viruses (hCov) 229E, NL63 and OC43. These 24 specimens were excluded for this analysis.

Of the 185 specimens analysed, a 118 (64%) were hospitalized and 67 (36%) were out-patients. Both patient groups had a high percentage of subtype A infections over the four years (Table 2.9). The association between subtype and disease severity over all four years was not statistically significant, though when looking at the different years separately there was some association found in 2008 where subtype A was associated with more severe disease ($p=0.04$). It was evident that subtype A dominated over subtype B for 2006, 2007 and 2008 and therefore both patient groups had a high subtype A prevalence during these three years. During 2009, 74% of the hospitalized group had subtype B infections and only 21% were infected with subtype A while all the out-patients had subtype B infections (Table 2.9).

To investigate the association between genotype and disease severity, we compared genotypes identified in hospitalized and out-patients. The only year where a difference in disease severity was seen is 2006 where more hospitalized patients had GA5 infections (67%) relative to 45% in the out-patient group while 24% of hospitalized patients had BA infections while 45% were identified in the out-patient group ($p=0.275$).

TABLE 2.9 Disease severity associations with infecting subtype and genotype

	Subtype A	Subtype B	Dual infections	p-value	GA2	GA5	BA	BA+Other	p-value
2006:									
HP (45)	31 (69%)	11 (24%)	3 (7%)		2 (4%)	30 (67%)	11 (24%)	2 (4%)	
OP (20)	10 (50%)	9 (45%)	1 (5%)	0.234	1 (5%)	9 (45%)	9 (45%)	1 (5%)	0.275
Total (65)	**41 (63%)**	**20 (31%)**	**4 (6%)**		**3 (5%)**	**39 (60%)**	**20 (31%)**	**3 (5%)**	
2007:									
HP (34)	33 (97%)	1 (3%)	0		27 (79%)	6 (18%)	1 (3%)	0	
OP (27)	26 (96%)	0	1 (4%)	0.693	20 (74%)	6 (22%)	0	1 (4%)	0.726
Total (61)	**59 (96%)**	**1 (2%)**	**1 (2%)**		**47 (77%)**	**12 (20%)**	**1 (2%)**	**1 (2%)**	
2008:									
HP (20)	20 (100%)	0	0		16 (80%)	4 (20%)	0	0	
OP (12)	9 (75%)	3 (25%)	0	0.044*	8 (67%)	1 (8%)	3 (25%)	0	0.085
Total (32)	**29 (91%)**	**3 (9%)**	**0**		**24 (75%)**	**5 (16%)**	**3 (9%)**	**0**	
2009:									
HP (19)	4 (21%)	14 (74%)	1 (5%)		4 (21%)	0	14 (74%)	1 (5%)	
OP (8)	0	7 (88%)	1 (13%)	0.338	0	0	7 (88%)	1 (13%)	0.338
Total (27)	**4 (15%)**	**21 (78%)**	**2 (7%)**		**4 (15%)**	**0**	**21 (78%)**	**2 (7%)**	
All 4 years:									
HP (118)	88 (75%)	26 (22%)	4 (3%)		49 (42%)	40 (34%)	26 (22%)	3 (3%)	
OP (67)	45 (67%)	19 (28%)	3 (4%)	0.612	29 (43%)	16 (24%)	19 (28%)	3 (4%)	0.435
Total (185)	**133 (72%)**	**45 (24%)**	**7 (4%)**		**78 (42%)**	**56 (30%)**	**45 (24%)**	**6 (3%)**	

***Statistically significant

2.4 DISCUSSION

Our study indicated that the BA genotype has now replaced all previous subtype B genotypes identified in SA and illustrates the ability of RSV to evolve and overcome herd immunity. Such variability complicates human RSV vaccine development. Evidence exist for a common ancestor of all BA viruses with this 60 nucleotide duplication (Trento et al., 2006) and since its discovery, strains with this duplication have been found in clinical specimens from distantly related places (Kuroiwa et al., 2005, Nagai et al., 2004, Sato et al., 2005, Scott et al., 2004b, Zlateva et al., 2005b, Zlateva et al., 2004).

In the first three years studied in SA, subtype A predominated, whereas subtype B predominated in the fourth year. Two genotypes (GA2 and GA5) dominated the subtype A specimens. These genotypes were also found to be dominant in a previous study from 1997-2000 in SA(Venter et al., 2001). The SAA1 genotype was however not detected again, suggesting GA2 and GA5 are the current stable genotypes. All the subtype B specimens clustered with the BA genotype and had the characteristic 60 nucleotide duplication replacing all B genotypes previously identified in SA. Referring to Venter *et al.(Venter et al., 2002b),* identical genotypes were previously found in different regions in SA during one season, and thus strains identified in one region likely represent the strains circulating in SA. The fact that only BA subtype B strains were identified over all four years in major public sector referral hospitals in the capital of the country is also suggestive that this is representative of the rest of the country. Investigations of strains from the rest of Africa will determine if this is true for the whole continent. The BA genotype has also been detected in Kenya in 2006 but co-circulated with the SAB1 genotype in that epidemic(Scott et al., 2004a). When compared to other genotypes, SA BA strains were more closely related to the SAB3 genotype previously detected in SA. It is not certain if new genotypes in SA occurred as a result of spontaneous mutations or of importation.

Apart from the 60 nucleotide duplication, no deletions, insertions or frame shift mutations were detected in the C-terminal region of the G protein although the full G protein sequence identified a 6 nucleotide (2 amino acid) deletion at position 490 resulting in a missing Proline and Lysine. This deletion has also been detected in strains from Kenya

and Buenos Aires(Agoti et al., 2010).The dominance of this genotype indicates that the duplication might give this genotype an evolutionary advantage. Sequencing of the full gene suggests further changes in the G gene may occur. Full genome sequence for the new BA strains would elucidate changes in other proteins.

Evolutionary rates have been estimated for only a limited number of viruses. In general, values lie close to 1×10^{-3} substitutions/site/year, although considerable variation exists. This variation is partly due to their error prone replication which causes mutations to accumulate quickly(Jenkins et al., 2002). To determine the evolutionary rate of the SA strains, Bayesian evolutionary analysis was conducted on all SA subtype A and B isolates. A higher evolutionary rate within the subtype A SA strains indicate that subtype A strains evolve faster with more nucleotide substitutions occurring per site per year than the subtype B strains. This could explain why acquisition of the 60 nucleotide duplication was to the advantage of subtype B strains for overcoming herd immunity.

Further investigation of the BA genotype was conducted by Bayesian analysis of SA BA isolates as well as BA isolates from all over the world since it was first discovered in 1999. Firstly it was shown that all the subtype B strains isolated in SA clustered within sub-genotype IV. According to Trento *et al.(Trento et al., 2006),* this is the most heterogeneous branch regarding the date and the place of isolation. It is possible that this 60 nucleotide duplication changed the antigenic structure of these BA genotypes, giving them an evolutionary advantage to re-infect individuals previously exposed to subtype B. When the BA viruses were first isolated in 1999, the 60 nucleotide duplication was an exact replicate of the preceding 60 nucleotides (Trento et al., 2006). This region accumulated further nucleotide substitutions over time and certain nucleotide positions were shown to be under positive selection (Botosso et al., 2009, Melero et al., 1997, Woelk and Holmes, 2001, Zlateva et al., 2005b, Zlateva et al., 2004). These changes might enhance the fitness of these viruses which would explain why they are now dominating worldwide and replacing all other B genotypes. This region is believed to be involved in changing/replacing the antigenic epitopes which contribute to the evasion of the immune response of the population. Certain amino acid in this region have been shown to be under positive selection, as well as the change 952CAA to UAA (Q313 to STOP) that was present in the most recent lineages of the BA-IV

branch(Botosso et al., 2009, Zlateva et al., 2005b). Most of the SA BA strains (79%), all from BA-IV, had this stop codon change. However, it was at position 946 because of the 6 nucleotide deletion. Changes in stop codon usage are thought to be associated with antigenic variation that allows immune evasion, especially in the subtype B isolates that are less variable than the subtype A isolates (Martinez et al., 1999).

Positive selection was shown to occur within some of the different BA branches. The dates when the strains were isolated and their placement on the tree indicate how much this genotype changed over time, especially the SA strains (BA-IV). The strains first isolated in 1999 lie closer to the root of the tree than the strains isolated in 2009. The appearance of new sub-genotypes within the BA-IV branch may occur as drift continues with four different clusters already existing within BA-IV (Figure 2.7). The rate of evolution for the BA genotype was calculated as 3.143×10^{-3} substitutions/site/year, which correlates with findings by Trento et al.(Trento et al., 2006). According to them, the origin of the MRCA was dated between 1998 and 1999, shortly before it was discovered in June 1999. Our data also indicate that the MRCA of the SA BA sequences date back to 1997 however, when the MRCA is calculated for all SA strains together with strains from other countries it dates back to 1992. These strains may have circulated at low levels and was therefore only detected a few years later.

Several studies have attempted to correlate clinical severity with the different subtypes. Some have found no difference between subtype A and B, although others have reported that subtype A is more severe (Brandenburg et al., 2000, Cane, 2001, Savon et al., 2006). It was also shown that the two subtypes co-circulate but that subtype A was dominant in more epidemics(Cane, 2001). Martinello *et al*.(Martinello et al., 2002) found that GA3 was associated with significantly greater severity of illness compared to GA2 and GA4, while other studies have failed to show any associations. In general the subtype/genotype dominating each of the seasons in the present study was also the subtype/genotype that was associated with increased hospitalizations. The only year where a specific genotype were more associated with hospitalized cases than with out-patients was 2006 when GA5 was associated with more severe disease while BA caused milder disease. The following year it was both replaced completely with GA2. This switch in dominance may be due to herd immunity against subtype A and GA5,

because of the previous year's high prevalence. It is known that multiple genotypes within both antigenic subtypes can co-circulate in one epidemic, with one or two genotypes being replaced in successive years (Cane et al., 1994, Galiano et al., 2005, Garcia et al., 1994, Hall et al., 1990, Peret et al., 2000, Peret et al., 1998). Venter *et al*.(Venter et al., 2002a) compared community circulating strains with isolates from hospitalized infants and found that the same strains were present in cases of mild URTI as well as more severe LRTI. This is however the first time that clinical data was available from patients attending the hospitals as in and out-patients and allowed more comprehensive analysis over seasons.

This study had several limitations. We were not able to fully assess possible associations between disease severity and strain type. Although we had information on hospitalization, the number of out-patients was small and no significant associations between subtype/genotype and disease severity were found. Also, we lacked information on HIV infection and immune status for most study participants. Thus, the role HIV infection play in circulating RSV genotypes could not be further explored.

To conclude, G protein variability may play a significant role in RSV pathogenesis by allowing immune evasion. Certain substitutions or alterations may enhance the fitness of viruses as is evident with the BA strains that replaced all other B genotypes previously identified in SA. The G protein's ability to accommodate such substantial changes and facilitate immune evasion may complicate vaccine development. It remains to be seen if this BA genotype will remain dominant or if the dominance will eventually fade because of herd immunity. Our study suggests that subtype B strains may need to be updated with recent strains in experimental vaccines to re-evaluate their efficacy.

CHAPTER 3

Identification of deletion mutant Respiratory Syncytial virus strains lacking most of the G protein in immunocompromised children with pneumonia in South Africa

3.1 INTRODUCTION

Evolutionary studies of the G protein of subtype B strains have described major differences in the G protein length due to alternative termination codon usage, premature stop codons, inframe duplications, deletions and insertions in the second variable region (Martinez et al., 1999, Trento et al., 2006). G proteins of especially subtype B isolates has been identified that are up to 30 nucleotides shorter due to frame shift mutations(Melero et al., 1997). The central region remained conserved in all sequences with an absolute conserved region between amino acid positions 164-187 in all strains(Zlateva et al., 2005a).

GAGs are responsible for the majority of RSV attachment to cultured cells leading to infection (Feldman et al., 2000). RSV virions containing the F protein as sole surface protein bind to GAGs as well as another unidentified molecule suggesting the F protein may have an auxiliary role as an attachment protein (Techaarpornkul et al., 2002).

Variation in neutralizing epitopes in the hypervariable region of the G protein suggest immune selection of new variants may contribute to generation of HRSV diversity(Trento et al., 2010). A cold-passaged subtype B mutant containing large deletions spanning most of the coding sequences of the small hydrophobic (SH) and attachment (G) proteins has been isolated in-vitro(Karron et al., 1997a). This virus replicated efficiently in Vero cells, but was found to be overattenuated in RSV sero-negative infants and children. Deletion of the central conserved domain and cystine noose was shown not to affect virus growth in vitro or in vivo (in mice)(Teng and Collins, 2002), although no record exists of strains that lack this in patients with clinical disease. Augmentation of the Th2 immune response by the G protein as well as antigenic variation in G makes delta G mutants attractive vaccine candidates(Teng et al., 2001).

As part of a molecular epidemiology investigation of a nosocomial outbreak of RSV in a ward for premature infants in Kalafong hospital in Pretoria, SA in 2006, RSV positive NPAs from patients in the general paediatric ward were investigated. The HIV-1 seroprevalence among mothers to these infants was 52.6%, suggesting a high level of perinatal exposure (Visser et al., 2008). Here we describe identification of G protein deletion mutants that were significantly reduced in size in comparison to the prototype controls in children with pneumonia. Sequence analysis identified subtype B G protein deletion mutant strains that lacked most of the ectodomain, including the conserved cystine noose in two children with pneumonia. In addition, all RSV positive specimens identified over a period of four years were screened to identify further cases.

3.2 Materials and Methods

3.2.1 Study Population

All RSV positive patients over a period of 4 years were selected as described in chapter 2, paragraph 2.2.1.

3.2.2 RNA isolation

Refer to chapter 2, paragraph 2.2.3.

3.2.3 Primer selection

Refer to chapter 2, table 2.1 for more information on primers used for subtyping and full G protein amplification and sequencing. Figure 3.1 indicate a diagram of all primers used to amplify and sequence the deletion mutant strains.

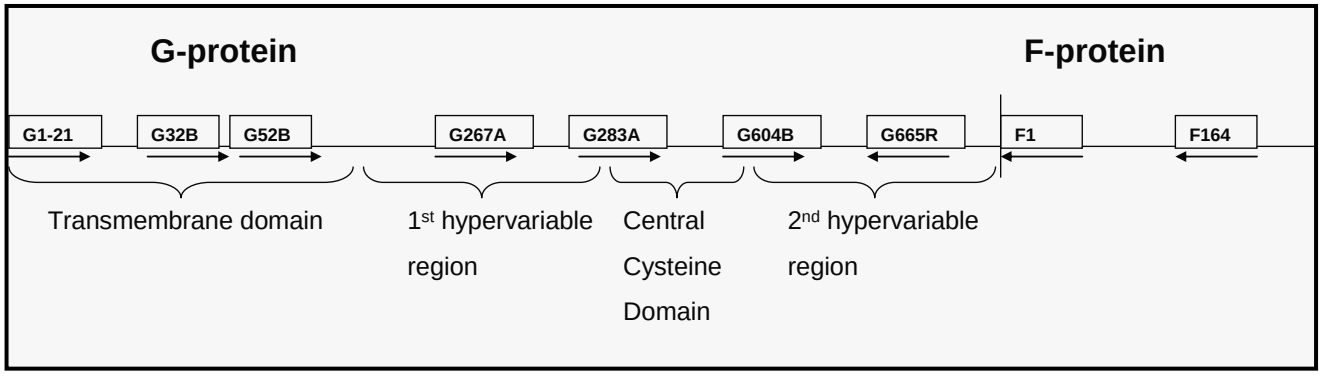


FIGURE 3.1 Diagram of primers used. G32B, G267A and F164: 1st round multiplex RT-PCR; G52B, G283A and F1: Nested multiplex PCR for subtyping. G1-21 and F164: 1st round RT-PCR; G32B and F1: Nested PCR for full length G protein amplification. G32B, G604B, F1 and G665R: Sequencing primers.

3.2.4 Full length G protein amplification

Subtyping entailed amplification of the G protein from NPAs by nested PCR with primers that distinguished subtype A and B by size(Madhi et al., 2001, Venter et al., 2002a, Venter et al., 2001) as explained in chapter 2, paragraph 2.2.5. Full G protein amplification of each of the subtype B genotypes were performed by Andronica Rakgantso as part of her BSc(Hons) project. These strains were deposited on genbank

with accession numbers JF704213-JF704219 (Appendix A). All PCR products were analyzed on a 1.5% agarose gel, using a 100bp ladder as molecular weight marker (DNA molecular marker XIV, Roche Diagnostics, Mannheim, Germany).

3.2.5 Nucleotide sequencing

Primers were as indicated in table 2.1 and figure 3.1. PCR product purification and sequencing was performed as in chapter 2, paragraph 2.2.7.

3.2.6 Multiplex Real-Time RT-PCR

A multiplex real-time RT-PCR that can detect 14 different viruses was performed on one of the deletion mutants (GTS374801) to detect any possible co-infections with other viruses(Lassauniére et al., 2010). The other specimen was not tested due to insufficient amounts of nucleic acid. The real-time RT-PCR protocol was as described in chapter 2, paragraph 2.2.8.

3.2.7 Phylogenetic analysis

Sequence alignments were performed with Clustal X version 1.81(Thompson et al., 1997). A summary alignment of the three deletion mutants and a subtype A and B sequence was generated with GeneDoc version 2.7.000(Nicholas and Nicholas, 1997). BioEdit Sequence Alignment Editor version 7.0.4.1 was used to compare nucleotide as well as amino acid differences between the three deletion mutant strains and the full length subtype B G protein gene sequences(Hall, 1999). Kyte and Doolittle Mean Hydrophobicity profiles were also created for these sequences with BioEdit. In addition, Parker antigenic plots were generated with Antheprot ‘98 for the deletion mutants and each full length subtype B genotype(Parker et al., 1986). Genbank accession numbers: HQ711840-HQ711842 (three deletion mutants); JF704213-JF704219 (full G protein sequences) (Appendix A).

3.3 RESULTS

3.3.1 Subtype identification: Multiplex RT-PCR and Nested PCR

Clinical specimens isolated over four consecutive years (2006-2009) were subtyped with a multiplex RT-PCR and nested PCR (Chapter 2). RSV positive specimens resulted in PCR fragments of 950bp for subtype A and 1200bp for subtype B in the first round and 728 and 946bp respectively for the nested PCR. In 2006, one specimen was found that was not of the correct size upon gel electrophoresis analysis. Figure 3.2 lane 1, indicate the specimen (SA367859Pt06) with a size of 450bp next to a negative control and the prototype subtype A and B positive control strains. Subsequently, screening of 209 clinical specimens identified an additional specimen with an altered profile. Figure 3.3 lane 4 indicate the second specimen (SA374801Pt06) that had three different bands present, 400-, 200- and 150bp respectively. Sequence analysis and BLAST searches confirmed these bands to be of RSV origin and belonging to subtype B isolates. An alignment summary indicated in figure 3.4 shows the similarities between the three deletion mutants and a subtype A and B strain, indicating that all three deletion mutant strains were more similar to subtype B.

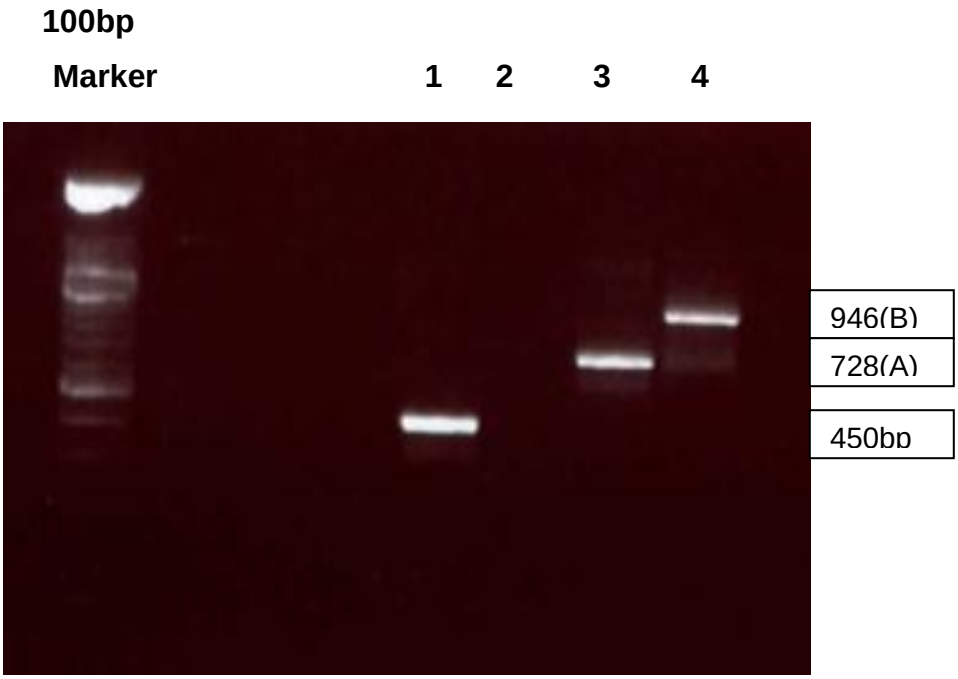


FIGURE 3.2 G protein specific subtyping multiplex nested RT-PCR for patient 1. Lane 1: Amplicon from patient one (SA367859Pt06) with a size of 450bp; Lane 2: Negative control; Lane 3: Positive control A; Lane 4: Positive control B.

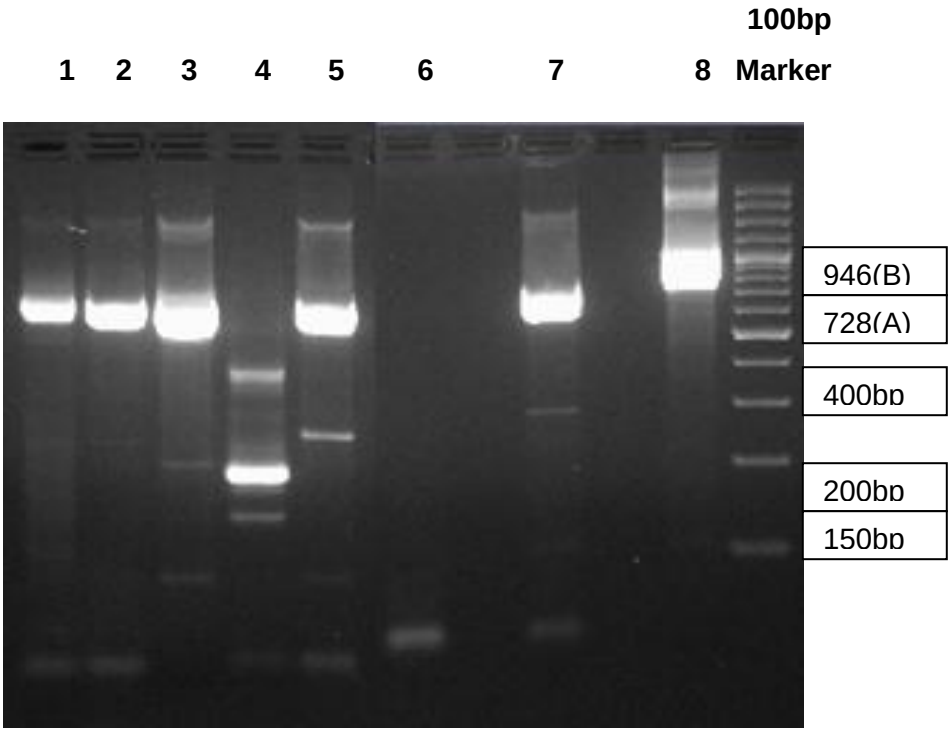


FIGURE 3.3 G protein specific subtyping multiplex nested RT-PCR for patient 2. Lane 1-5: Clinical specimens. Lane 6: Negative control; Lane 7: Positive control A; Lane 8: Positive control B; Lane 4: Amplicon from patient 2 with sizes 400-, 200- and 150bp.

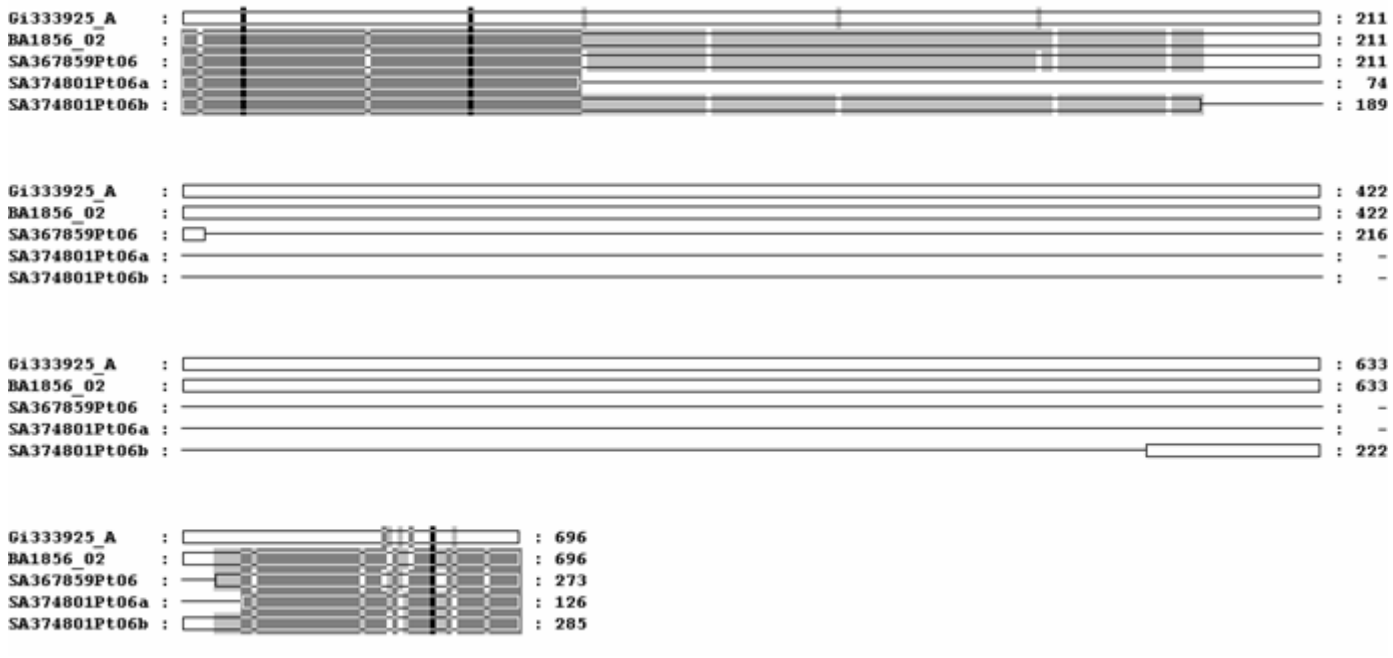


FIGURE 3.4 Alignment summary of the three deletion mutant strains with a subtype A and BA strain. Black lines indicate similarity between all five strains and grey lines indicate similarity between four strains. All three deletion mutants were more similar to subtype B isolates.

3.3.2 Patient clinical description

The first specimen identified that yielded a suspicious G protein amplicon (SA367859Pt06) was from an 8 month old HIV positive male hospitalized with pneumonia that tested positive for RSV with the rapid antigen kit. This child had symptoms of breathing distress and tachypnoea, very underweight (close to marasmic) and was on continues oxygen treatment. His CD4 count was 1038(27.5%). The percentage of CD4 cells is the best indicator of immunodeficiency in children, although guidelines only exist for infants older than 12 months. The immune system of infants <1 year of age are still immature and HIV infected infants are at serious risk of morbidity and mortality(2010).

The second specimen (SA374801Pt06) was from a 2 month, 13 day old HIV exposed male hospitalized for seven days for gastroenteritis and severe dehydration. He developed pneumonia in the ward and tested positive for RSV five days after admission. The child and mother live in an informal settlement with limited resources and the infant was extremely malnourished upon admission. HIV PCR was performed due to HIV sero-positive status of the mother as part of the antiretroviral treatment programme for prevention of mother to child transmission. Follow up testing for HIV would be carried out up to 18months as per protocol for HIV exposed children.

No amplicons were visible at the expected position for full length G proteins and no other co-infections were identified. SA374801Pt06 was negative for hMPV, hBoV, PIV 1,2 and 3, Influenza A and B, ADV and hCov- 229E, OC43, HKU1 and NL63(Lassauniére et al., 2010). SA367859Pt06 was negative for other viruses detected by the IFA but insufficient material was available for additional PCRs. Attempts to culture virus from these specimens were not successful.

3.3.3 Nucleotide and amino acid analysis

Amino acid alignment comparisons of the amplicon of the specimen from patient one (SA367859Pt06) identified a deletion mutant of which nucleotides 318 to 905 was deleted which corresponds to amino acids 106 to 302 of the G protein gene. This comprises nearly the entire G protein ectodomain corresponding to the BA1428/99B (AY333364) reference strain used(Trento et al., 2003).

Two deletion mutants were identified in patient 2 of which the smaller of the two had a deletion from nucleotide 108 to 911 (SA374801aPt06) corresponding to amino acids 36 to 304, while nucleotides 276 to 854 (amino acids 92 to 285) were deleted in the larger of the two amplicons (SA374801bPt06). The larger band (400bp) was shown to be non-specific. All three deletion mutant strains lacked the conserved cystine region.

Figure 3.5 indicates amino acid alignments of specific genotypes within subtype B as well as the three deletion mutants. The first three strains indicate the three deletion mutants. A large amount of nucleotide substitutions resulted in amino acid changes. One strain (AgJ15_99 SAB3) described by Venter *et al*.(Venter et al., 2002a) had a C to T transversion which resulted in an amino acid change from Glutamine (Gln) to a stop codon (TAA), truncating the strain by 41 amino acids. Additional in frame insertions were also identified in the C-terminal hypervariable domain of some of the SA BA strains.

3.3.4 Antigenic site prediction and hydrophobicity profiles

In order to compare the antigenic structure of the deletion mutants relative to the other subtype B genotypes, Parker antigenic plots were generated with the programme Antheprot ‘98(Parker et al., 1986)(Figure 3.6). The two hypervariable sites, which are the major antigenic sites of the G protein, were absent in the three deletion mutants. The mean hydrophobicity profile of the three deletion mutants and the different subtype B genotypes are shown in figure 3.7. The transmembrane region and cystine noose are indicated as the hydrophobic regions, whereas the two hypervariable regions are indicated as hydrophilic domains.

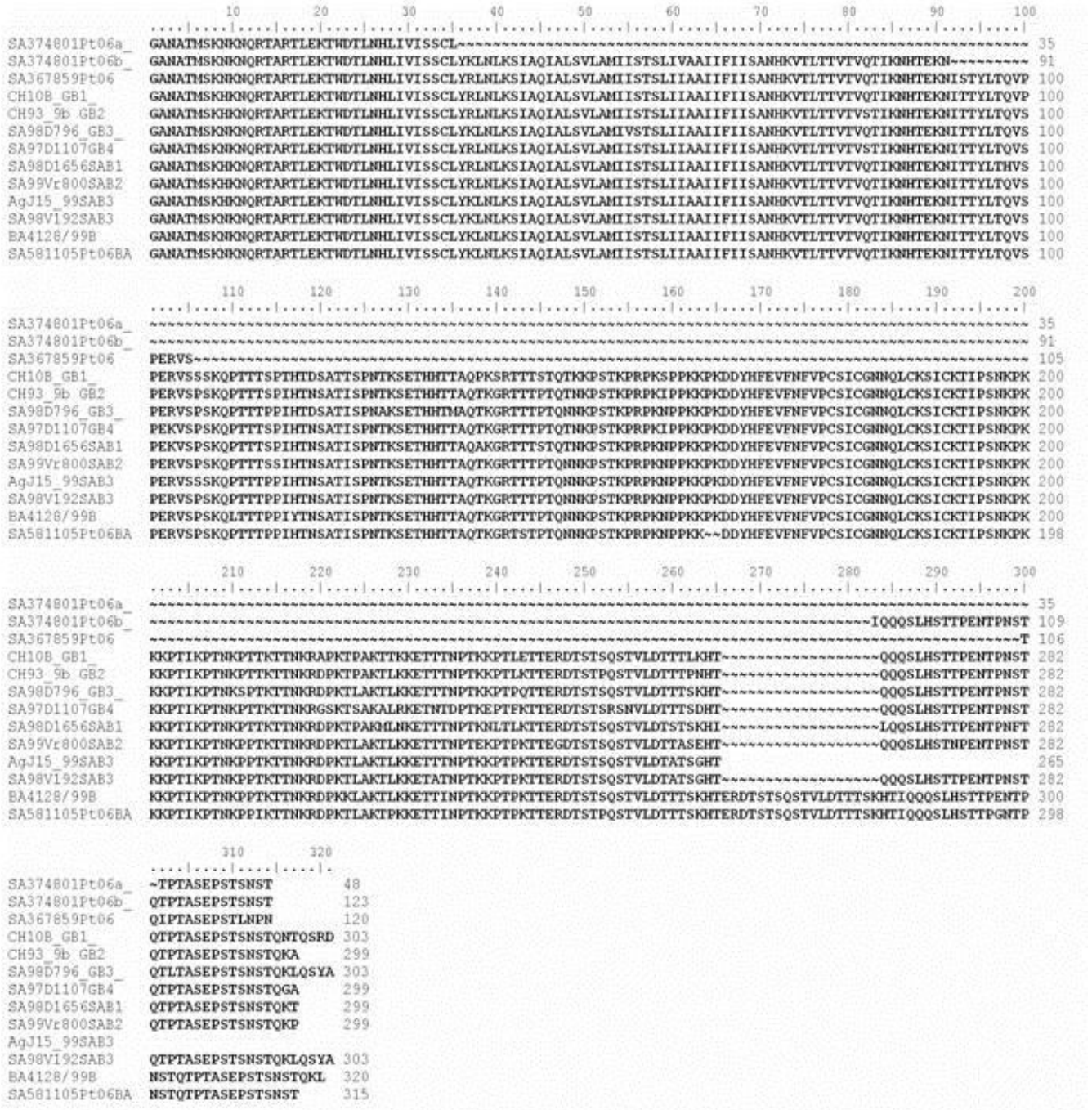


FIGURE 3.5 Complete amino acid alignment of the three deletion mutants in comparison to the full G protein subtype B sequences of each genotype. The genotype name is indicated after the strain name. Accession numbers: HQ711840-HQ711842 (Deletion mutant strains); AF065250 (GB1); AF065251 (GB2); JF704217 (GB3); JF704214 (GB4); JF704213 (SAB1); JF704218 (SAB2); JF704215 (Truncated SAB3); JF704216 (SAB3); AY333364(BA); JF704219(South African BA).

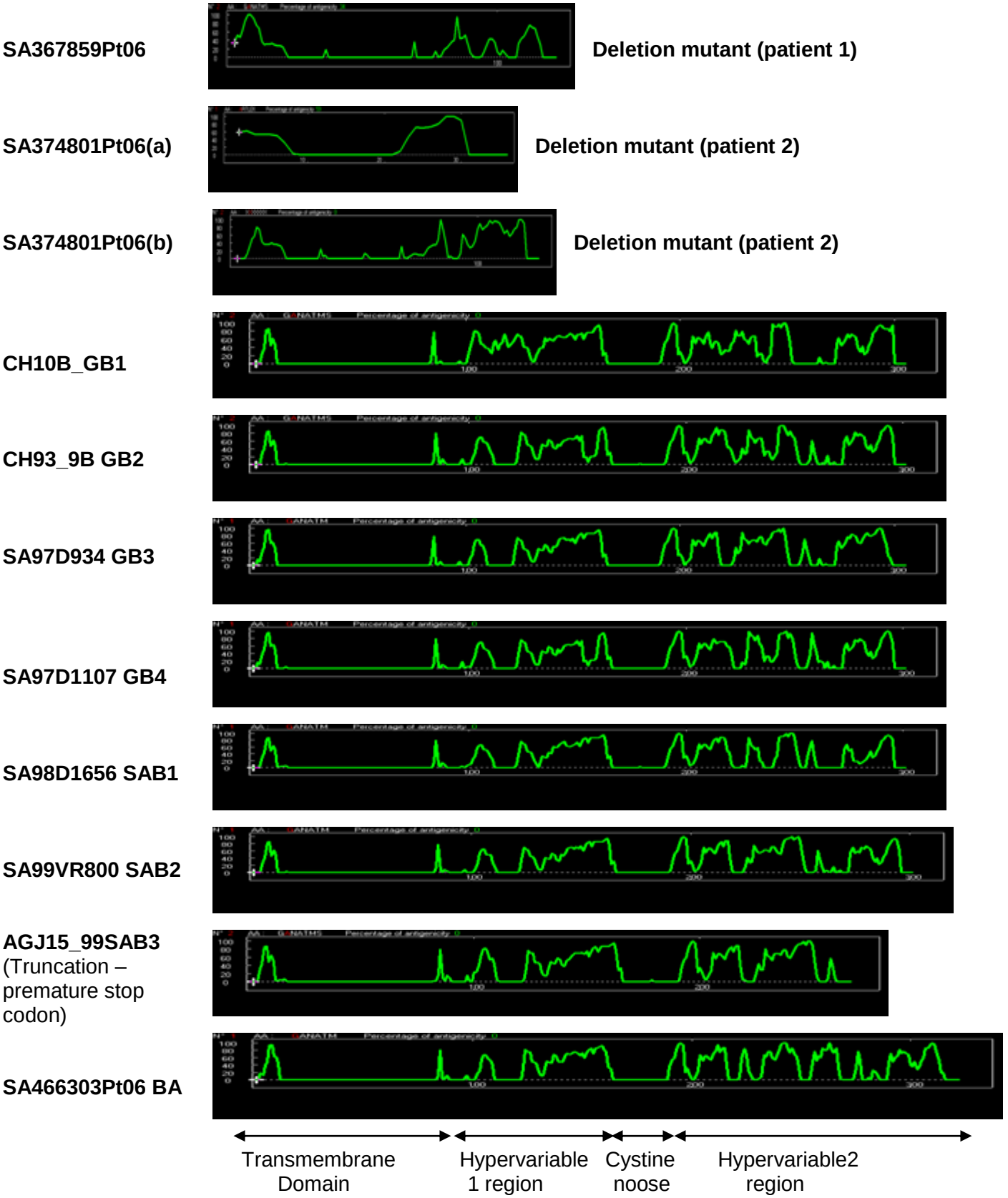


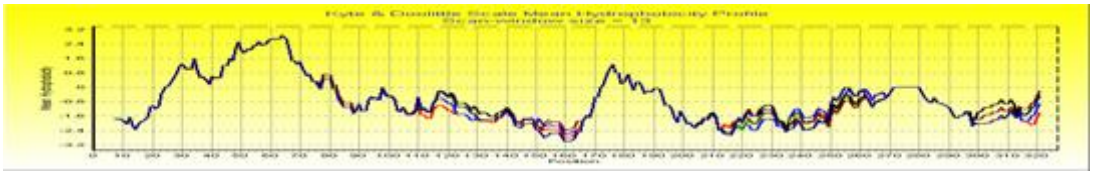
FIGURE 3.6 Parker's antigenicity prediction plots for the three deletion mutants, SAB3 (with a premature stop codon) and each of the subtype B genotypes generated with Antheprot'98(Parker et al., 1986).

SA367859Pt06 Deletion mutant (patient 1)

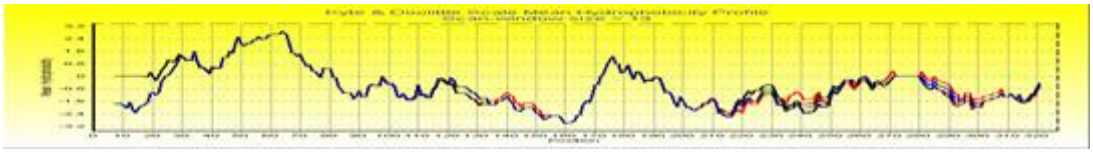
SA374801Pt06(a) Deletion mutant (patient 2)

SA374801Pt06(b) Deletion mutant (patient 2)

GB1, GB2, GB3 and GB4



SAB1, SAB2 and SAB3



BA genotypes

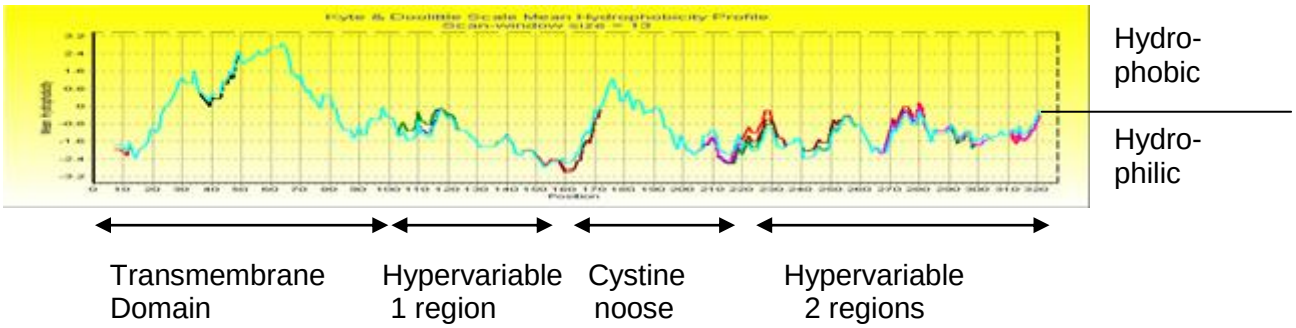


FIGURE 3.7 Complete hydropobicity profile (Kyte and Doolittle) of the three deletion mutants and all other genotypes within subtype B generated with BioEdit Sequence Alignment Editor version 7.0.4.1(Hall, 1999).

3.4 DISCUSSION

During the 2006 season, two specimens were detected that differed substantially in size from the prototype subtype A and B controls upon gel electrophoresis analysis. Deletion mutant strains were identified that had most of the G protein ectodomain missing and belonging to subtype B. To our knowledge, this is the first time that strains with such extreme deletions in the G glycoprotein have been found to cause LRTI in infants. This emphasises the flexibility of the G protein ectodomain to accommodate changes on its surface. This may suggest that major deletions can occur in the host without affecting its ability to cause LRTI.

Frame shift mutations, premature stop codons and hypermutation events that caused severe truncations in the G protein have been shown to generate neutralization resistant mutants in serial passages in-vitro but it remained to be seen whether these changes in the escape mutants are also observed among natural isolates(Garcia-Barreno et al., 1990, Martinez et al., 1997, Rueda et al., 1991). Lazar *et al*.(Lazar et al., 2006) identified premature stop codons in the G glycoprotein gene sequenced at three different time points from a patient with severe combined immune deficiency syndrome. The RSV G gene of these isolates encodes a truncated G glycoprotein lacking the 42 carboxy-terminal amino acid. It was hypothesized that these mutations developed during prolonged infection due to severe immune deficiency as well as immunologic pressure as a result of the monthly treatment with intravenous immunoglobulin.

Both patients with deletion mutant strains identified in the present study were immune suppressed and hospitalized with LRTI. There was no other obvious co-infection with any other virus that could have caused the symptoms. Attempts to culture virus from these deletion mutants were unsuccessful, probably due to low concentrations present and repeat freeze thawing of specimens that had been screened retrospectively. Amino acid analysis as well as antigenic site prediction showed that nearly the entire G protein ectodomain, together with the cystine noose were deleted in the deletion mutants. Due to the size of the deletions, comparing the antigenic plots of the deletion mutants with that of the subtype B genotypes, could not show that the deletion mutants belong to a specific genotype.

These findings suggest that the G protein might not be necessary for infection and that the F protein might be sufficient to play the role of auxiliary attachment protein. Techaarpornkul *et al*.(Techaarpornkul et al., 2002)have shown that RSV virions containing the F protein as the sole surface protein, do bind to GAGs and that part of the F protein binding represents an interaction with a molecule other than a GAG. This demonstrates that the F protein may have an auxiliary role as an attachment protein.

This study suggests that RSV clinical strains that lack most of the G protein gene may occur in immunocompromised patients with LRTI. The molecular mechanism whereby this occurs is not clear, however reduced immune pressure in these patients may allow these strains to utilise the F protein for binding and replication. Further characterization of such strains may elucidate the replication and pathogenic potential, however low viral load and storage of specimens may complicate the isolation of such strains. This finding may have implications for the utilisation of certain attenuated strains as vaccines in immunocompromised children.

CHAPTER 4

Genetic- and expression level differences in the Non-structural proteins as determinants of RSV pathogenesis

4.1 INTRODUCTION

RSV is unique among paramyxoviruses in having two non-structural (NS) proteins NS1 and NS2 of 139 and 124 amino acids, respectively(Collins et al., 1986). The proteins are encoded from promoter-proximal genes to ensure high expression levels. NS1 and NS2 show no sequence similarity except for four carboxy-terminal amino acids (Collins and Wertz, 1985, Johnson and Collins, 1989). Evans *et al*.(Evans et al., 1996) have shown that both the NS1 and NS2 proteins are synthesised early in the infectious cycle with NS1 expression appearing to be stable while NS2 disappeared rapidly. NS1 was also found to be closely associated with the M protein in infected cells(Evans et al., 1996).

Alpha and beta Interferon (INF) proteins are not only a powerful first line of defence against pathogens but also have potent immunomodulatory activities(Le Bon *et al.*, 2001). The success of viruses in establishing productive infections depends on viral expression of proteins or mechanisms for evasion of the IFN pathway to enhance virulence. It has been shown that the two NS proteins perform different and complementary functions to inhibit the host's immune response by affecting key points of the IFN pathway (Kotelkin et al., 2006, Moore et al., 2008, Spann et al., 2004, Swedan et al., 2009). Viruses use several strategies to modify antiviral effects of IFNs such as (1) minimizing IFN production; (2) blocking IFN receptors; (3) inhibiting IFN signalling; and (4) bypassing antiviral protein effects(Ploegh, 1998, Sen, 2001).

Other IFN-independent functions has also been associated with these two proteins such as suppression of dendritic cell maturation which results in decreased antigen presentation and T lymphocyte activation leading to incomplete and/or weak immune responses, suppression of premature apoptosis and a potent inhibitor of minigenome transcription and replication(Atreya et al., 1998, Bitko et al., 2007, Munir et al., 2008).

Sequence analysis of HRSV NS proteins of clinical specimens has been limited. NS protein sequences of HRSV subtype A and B have been compared to that of BRSV and amino acid identities of 69 and 68% for the NS1 protein and 84 and 83% for the NS2 protein was shown respectively(Pastey and Samal, 1995). Deplanche et al., (2007) analysed the evolution of bovine RSV and found the variable region of the G protein to be genetically stable after virus isolation and over 10 serial infections in bovine turbinate (BT) cells. The mutant spectrum of the G protein of several populations derived from one isolate were examined and it was proposed that populations of RNA viruses evolve as dynamic distributions of closely related mutant genomes that exist in equilibrium around a theoretical consensus sequence. Such mutants would provide evidence of quasispecies dynamics for RSV, implying the presence of a variant reservoir for viral adaptation. No such analysis of mutant spectra has been reported for HRSV to date.

Investigation of the sequence variation of the NS proteins in strains identified in patients with mild and severe disease may determine if specific genetic differences exist that could alter the protein function. Expression levels of these proteins may also have an effect on the IFN response and in effect the pathogenic potential in patients. To address these questions, sequence variation and expression levels of the NS proteins were investigated in paediatric patients without other predisposing factors with mild and severe disease experiencing their first RSV infection.

4.2 MATERIALS AND METHODS

4.2.1 Study population

The study population was as described in chapter 2, paragraph 2.2.1.

4.2.2 Selection of specimens to include in NS protein investigation

All specimens that tested positive for RSV by RT-PCR for the 2006 and 2007 season (Chapter 2) were used for NS1 and NS2 protein amplification and all those that could be amplified by first round RT-PCR were used for quantification analysis. Patients were specifically chosen to have mild or severe RSV infection to investigate disease severity. Mild disease was classified as outpatients not requiring hospitalization and severe disease were patients requiring hospitalization with severe LRTI or requiring intensive care treatment. No underlying risk factors such as prematurity or heart disease were present. For some patients the HIV status was known but for most of them it was unknown. All the specimens were single infections and no co-infections were present as tested with a respiratory multiplex described in chapter 2.

4.2.3 Non-structural protein amplification and sequencing

4.2.3.1 Primer selection

For the amplification of the NS1 and NS2 genes, primers NS1F44 and NR1214 were used for first round RT-PCR (Venter, unpublished) that is conserved for both subtypes. Sequencing was performed with the same set of primers as well as an additional set (NSF460 and NSR760) to cover the entire fragment (Table 4.1/Figure 4.1). All primers were designed using Primer3 software (http://www.broad.mit.edu/cgi-bin/primer/primer3_www.cgi).

TABLE 4.1 Primers and probes used for NS amplification/sequencing and quantification

Primer	Orientation	Sub-group	Target gene	PCR/Seq	Primer sequence (5' to 3')	Reference
NS1F44	Forward	A & B	NS1	RT&S	TGGGGCAAATAAGAATTTGATAA	Venter, unpublished
NR1214	Reverse	A & B	NS2	RT&S	AT(AG)GTGTATTTGCTGGATGACAG	Venter, unpublished
F460	Forward	A & B	NS1	S	TGACCAATTATATGAATCAATTATCTG	Venter, unpublished
R760	Reverse	A & B	NS2	S	GCCTGTCTTTCATCAAGTTTTCT	Venter, unpublished
NS1F17	Forward	A & B	NS1	QPCR	TGAG(CT)ATGATAAA(AG)GTTAGATTACAAA	Venter, unpublished
NS1R399	Reverse	A & B	NS1	QPCR	AAGTAA(AG)TCAGATA(AG)TTGATTCATAT	Venter, unpublished
NS2F39	Forward	A & B	NS2	QPCR	AGA(AG)TGATGATCACAGACATGAGACC	Venter, unpublished
NS2R342	Reverse	A & B	NS2	QPCR	GGAGTRTG(AG)TTTGTAGGCTTAATG	Venter, unpublished
Taqman Probes						
NS2 RSV A	Forward	A	NS2	QPCR	YAK-TGCACAAAGTGGGAAGCAACTAAATAYAAGA-BBQ	Venter, unpublished
NS2 RSV B	Forward	B	NS2	QPCR	FAM-TGCACAAAGTAGGGAGTACCAAATACAAGA-BBQ	Venter, unpublished

NS=Non-structural; RT=Reverse transcription; S=Sequencing; QPCR=Quantitative PCR

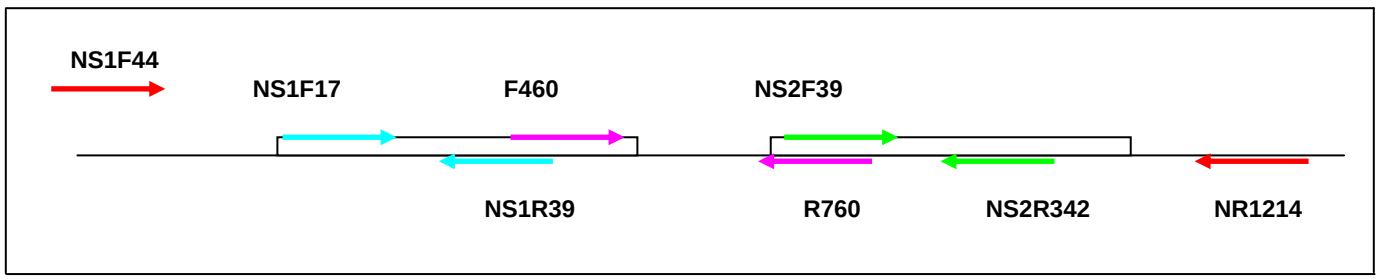


FIGURE 4.1 Primer combinations and positions for amplifying, sequencing and quantifying the NS proteins. Red= amplification of NS1/2; Pink= sequencing and quantifying antigenome; Blue= quantifying NS1; Green= quantifying NS2.

4.2.3.2 RNA isolation

Refer to chapter 2, paragraph 2.2.3.

4.2.3.3 RT-PCR for NS1/2 amplification

Each PCR reaction contained 10 µl of RNA template, 1 µl dNTP mix (10mM), 2 µl of 20 pmol primer (NS1F44 and NR1214), 2.5 µl DTT (100mM), 0.25 µl RNase Inhibitor (40U/ µl), 10 µl 5x RT- buffer and 1 µl enzyme mix in a final reaction volume of 50 µl. Reactions were incubated at 50°C for 30 min, followed by 94°C for 2 min, 35 cycles of 94°C for 10 s, 54°C for 30 s and 68°C for 1 min and finally one step of 68°C for 7 min. For specimens that could not be detected, the annealing temperature was lowered to 51˚C. Cycle elongation for an additional 5 s for each cycle was performed for specimens that could still not be detected. All PCR products were analysed on a 1% agarose gel, using a 100bp ladder as molecular weight marker (DNA molecular marker XIV, Roche Diagnostics, Mannheim, Germany).

4.2.3.4 Nucleotide sequencing

The Wizard SV Gel and PCR Clean-up system were used for PCR product purification (Promega, Madison, WI). Cycle sequencing was performed with the BigDye Terminator 3.1 Cycle Sequencing kit (Applied biosystems, Foster City, CA). Primers are shown in Table 4.1. Nucleotide sequencing was carried out on both strands and the editing was performed with Sequencher^{TM} Version 4.6, (Gene Codes Corporation, Ann Arbor, MI). NS protein sequences can be found on GenBank, Accession numbers: JQ359529-JQ359608.

4.2.3.5 Nucleotide and amino acid sequence analysis

Nucleotide sequences of subtype A and B viruses were aligned separately with Clustal X 1.81, using the multiple alignment option(Thompson et al., 1997). P-distances were calculated using Mega version 4(Tamura et al., 2007). BioEdit Sequence Alignment Editor version 7.0.4.1(Hall, 1999) and GeneDoc version 2.7.000(Nicholas and Nicholas, 1997) was used for amino acid analysis.

4.2.3.6 Phylogenetic analysis

Midpoint-rooted neighbour-joining tree were constructed with Mega version 4(Tamura et al., 2007). Kimura 2 distance parameter under 1000 bootstrap intervals was used.

4.2.3.7 Cloning of NS sequences

Sequence analysis of ten molecular clones each of four specimens that had two different peaks present during sequencing were performed using the Clone JET PCR Cloning Kit according the manufacturer's recommendations (Fermentas, Canada, USA) to identify the quasispecies distribution.

4.2.4 Quantification of the NS proteins

4.2.4.1 Primer and probe design

For quantification purposes, primers NS1F17 and NS1R399 were used for NS1 amplification of a PCR product of 382bp; primers NS2F39 and NS2R342 for NS2 amplification (303bp product) while primers F460 and R760 were used to amplify the antigenome for relative quantification of the protein's expression levels to the genome copies (332bp product). The NS1 gene and the antigenome were quantified using a SYBR Green assay whereas the NS2 gene was quantified using a taqman assay. The probes were designed by TIB MOLBIOL (Roche diagnostics, Mannheim Germany) for each subtype. More information about the primers and probes are shown in table 4.1 and figure 4.1.

4.2.4.2 Standard curve

The PCR product amplified with primers NS1F44 and NR1214 was purified using the Wizard SV Gel and PCR Clean-up system (Promega, Madison, WI) and quantified spectrophotometrically using a Nanodrop. Three different PCR amplifications (NS1, NS2 and antigenome) were performed with template dilutions ranging from 10^{-1} to 10^{-6} copies. Copy number was calculated with the following equation: DNA concentration (g/mol) = (Size of template in bp)(330Da)(2nt/bp)/Avogadro's number 6.0221415 x 10^{23}) (Whelan et al., 2003).

The C_T (Cycle Threshold) values were used to calculate and plot a linear regression line by plotting the algorithm of template concentration (X-axis) against the corresponding

threshold cycle (Y-axis). The C_T is the PCR cycle where the target amplification is first detected, so where fluorescence intensity is greater than background fluorescence(Wong and Medrano, 2005). The slope of the line was used to determine the efficiency of the target amplification (=Ex) using the equation $Ex=(10^{-1/slope})-1$(Rasmussen, 2001). A PCR efficiency of 2 is the best, and corresponds to the synthesis of one full length DNA copy per every available template after each cycle(Plumet and Gerlier, 2005). Finally, a correlation coefficient (r) of -1 showed that a linear relation is observed.

4.2.4.3 RNA isolation and cDNA synthesis

RNA isolation was as described in chapter 2, paragraph 2.2.3. For cDNA synthesis, 9.5µl of RNA were used, 1µl of 20pmol primers (random hexamers), 2µl of dNTP mix (10mM), 4µl of 5X RT buffer (100mM), 2µl DTT (10mM), 0.5µl RNase inhibitor (50U) and 1µl of Expand Reverse Transcriptase enzyme (50U/µl) to make up 20µl.

The following cycle was used for cDNA synthesis. 30ºC for 10min and 42º for 50min.

4.2.4.4 Quantitative PCR of NS1 protein and the antigenome

Quantifications were performed with the Lightcycler FastStart DNA Master SYBR Green I (Roche, Mannheim, Germany) kit according to the manufacturer's recommendations. In brief, each reaction contained 5µl of cDNA, 1µl of 10pmol primer each, 1.6µl of $MgCl_2$ stock solution and 2µl of the Lightcycler FastStart DNA Master SYBR Green I, 10x concentration in a final volume of 20µl.

The following programme was followed on a Lightcycler 1.5 system (Roche Diagnostics^{Ltd} Mannheim, Germany): Preincubation at 95ºC for 10min, [95ºC for 10s, 54ºC for 8s, 72ºC for 16s] for 45 cycles. Fluorescent data were acquired during each extension phase. After 45 cycles a melting curve was generated by heating the sample to 95ºC for 0s (slope=0.1ºC/sec), 65ºC for 15s and 95ºC for 0s (continues). Finally the samples were cooled down to 40ºC for 30s. The crossing points (CP) for each transcript was determined. To confirm amplification specificity the PCR products from each primer pair were subjected to agarose gel electrophoresis. The fluorescence from primer dimers may be eliminated by measuring the fluorescence at a temperature above their Tm but several degrees below that of the specific product(Morrison et al., 1998). Under

these conditions the primer dimers will be single stranded and therefore not bind SYBR Green I dye.

4.2.4.5 NS2 quantification

The NS2 gene was quantified using the Lightcyler TaqMan Master kit (Roche, Mannheim, Germany) according to the manufacturer's recommendations. Each reaction contained 5µl of cDNA template, 1µ of 10pmol primers each, 0.2µl of 10pmol probe (subtype A or B) and 4µl of the Lightcycler TaqMan Master mix 5x concentration to a final volume of 20µl. The following programme was followed on a Lightcycler 1.5 system: 95ºC for 15min, [95ºC for 10s, 51ºC for 40s, 72ºC for 1s (single acquisition)]X45.

4.2.4.6 Statistical analysis

Statistical analysis was performed using the Two-sample Wilcoxon rank-sum (Mann-Whitney) test and the Krustal-Wallis equality-of-populations rank test. Probability values <0.05 were considered statistically significant.

4.3 RESULTS

4.3.1 Genetic differences of the NS proteins as determinant of RSV pathogenesis

4.3.1.1 NS protein amplification

For both seasons (2006 and 2007), 142 specimens met the criteria for inclusion in this study and could be amplified with the G protein multiplex nested PCR (Chapter 2). All 74 specimens from 2006 and 68 specimens from 2007 were used for NS protein amplification. In total, a 111 of the 142 specimens were successfully amplified with a first round PCR targeting both NS proteins. Figure 4.2 indicates a 1% agarose gel of the NS RT-PCR products. NS RT-PCR resulted in products of 1100bp in size for both subtypes.

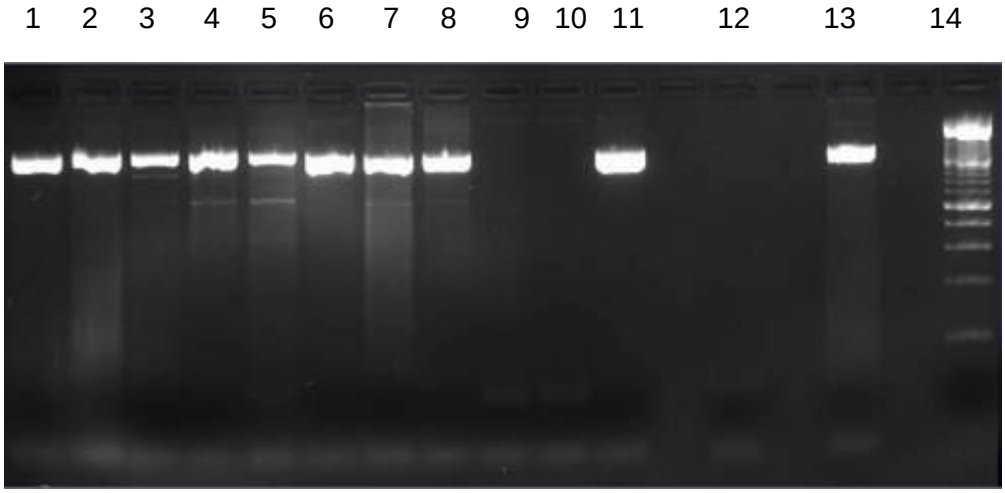


FIGURE 4.2 Agarose gel (1%) of the NS RT-PCR products. Lanes 1-11 indicate clinical specimens. Lane 12: Negative control; Lane 13: Positive control; Lane 14: 100bp molecular weight marker.

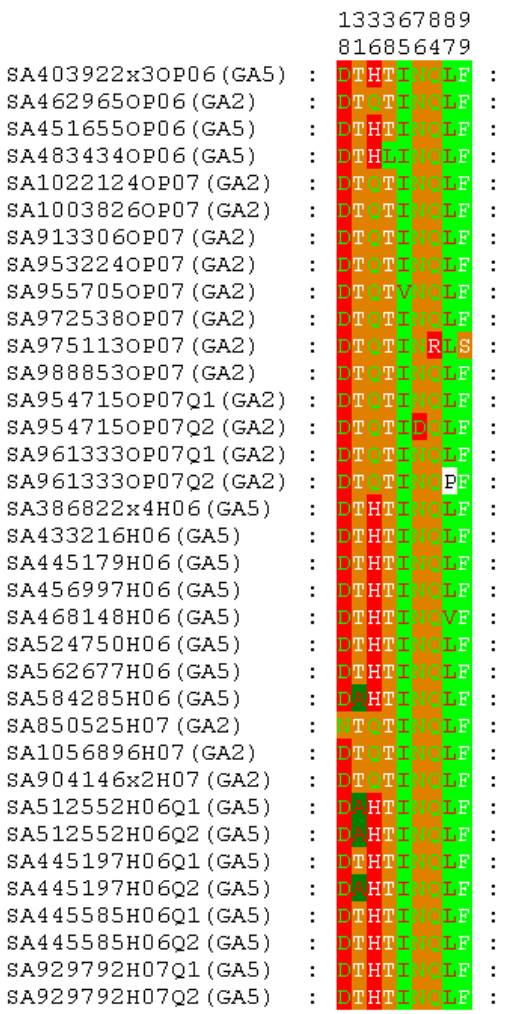
4.3.1.2 NS protein sequence- and phylogenetic analysis

A total of 57 specimens, 42 from 2006 and 15 from 2007 were selected for NS protein sequencing. Of these, 27 were specimens from outpatients and 30 were specimens from patients that were hospitalized. Two primer sets were used to cover the entire fragment. Figure 4.3 A and B indicate the variable amino acid sites of the NS1 gene subtype A and B separately while figure 4.3 C and D indicate the variable amino acid sites of the NS2 gene of subtype A and B strains.

Both proteins were highly conserved with only a few amino acid changes. The subtype A genes were generally more variable than the subtype B genes. A few amino acid changes could be attributed to a specific G protein genotype while no one specific change could be identified that was more prevalent in the hospitalized group than in the out-patient group. P-distance analysis showed that the NS1 gene had differences of between 0% and 6.4% at nucleotide level with an average difference of 3% between specimens while at amino acid level only up to 2.9% differences were seen with an average difference of 0.8%. The subtype B specimens were more conserved with only up to 1% difference between specimens at nucleotide level and 0.7% difference at amino acid level.

A range of 0% to 4% differences were seen for the NS2 gene both at nucleotide and amino acid level with an average mean of 2.1% at nucleotide level and 1.2% at amino acid level. The subtype B NS2 specimens differed by 1.6% at nucleotide level and 4% at amino acid level.

A)



B)

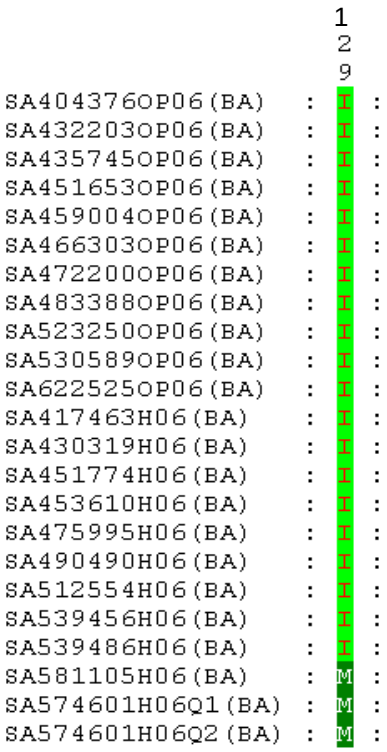


FIGURE 4.3 Variable amino acid sites of **(A)** the NS1 gene of subtype A, **(B)** NS1 gene of subtype B, **(C)** NS2 gene of subtype A and **(D)** the NS2 gene of subtype B. Specimens are ordered according the disease severity and year, where OP=Outpatient (mild) and H=Hospitalized (severe). Quasispecies are also indicated as Q1, Q2, Q3 etc.

C)

```
                          3455
                         78509
SA432598OP06(GA5)    : GRHRF :
SA451655OP06(GA5)    : GRHRL :
SA462965OP06(GA2)    : DKHRF :
SA483434OP06(GA5)    : DRHRF :
SA505872OP06(GA5)    : GRHRF :
SA953224OP07(GA2)    : DKHRF :
SA954715OP07(GA2)    : DKHRF :
SA955705OP07(GA2)    : DRHRF :
SA961333OP07(GA2)    : DKHRF :
SA972538OP07(GA2)    : DKHRF :
SA10038260P07(GA2)   : DKHRF :
SA10221240P07(GA2)   : DKHRF :
SA403922OP06Q1(GA5)  : GRHRF :
SA403922OP06Q2(GA5)  : GR RL :
SA445585H06(GA5)     : GRHRF :
SA447060H06(GA5)     : GRHRL :
SA452947H06(GA5)     : GRHRL :
SA456997H06(GA5)     : GRHRL :
SA524750H06(GA5)     : GRHRF :
SA562677H06(GA5)     : GRHRF :
SA386822H06(GA5)     : GRHRL :
SA584285H06(GA5)     : DRHRF :
SA433216H06(GA5)     : GRHRF :
SA445179H06(GA5)     : GRHRF :
SA468148H06(GA5)     : GR RL :
SA470433H06(GA5)     : GRHRL :
SA850525H07(GA2)     : DKHKF :
SA930011H07(GA2)     : DKHRF :
SA975113OP07(GA2)    : DRHRF :
SA988853OP07(GA2)    : DKHRF :
SA913306OP07(GA2)    : DRHRF :
SA1056896H07(GA2)    : DKHRF :
SA904146OH07(GA2)    : DKHRF :
SA445197H06Q1(GA5)   : GRHRF :
SA445197H06Q2(GA5)   : DRHRL :
SA512552H06Q1(GA5)   : DRHRF :
SA512552H06Q2(GA5)   : GRHRF :
SA929792H07Q1(GA5)   : GRHRL :
SA929792H07Q2(GA5)   : DRHRF :
SA929792H07Q3(GA5)   : GRHRF :
```

D)

```
                          1
                       2581
                       8537
SA404376OP06(BA)     : TRDI :
SA432203OP06(BA)     : TRDI :
SA435745OP06(BA)     : TRDI :
SA451653OP06(BA)     : TRDI :
SA459004OP06(BA)     : TRDI :
SA466303OP06(BA)     : TRDI :
SA472200OP06(BA)     : TRDI :
SA523250OP06(BA)     : TRDI :
SA483388OP06Q1(BA)   : TKDI :
SA483388OP06Q2(BA)   : TRDI :
SA530589OP06Q1(BA)   : TRDI :
SA530589OP06Q2(BA)   : IRDI :
SA622525OP06Q1(BA)   : IRKI :
SA622525OP06Q2(BA)   : TRDI :
SA622525OP06Q3(BA)   : TRKI :
SA622525OP06Q4(BA)   : TRDI :
SA417463H06(BA)      : TRDM :
SA430319H06(BA)      : TRDI :
SA451774H06(BA)      : TRDI :
SA453610H06(BA)      : TRDI :
SA475995H06(BA)      : TKDI :
SA490490H06(BA)      : TRDI :
SA512554H06(BA)      : TRDI :
SA574601H06(BA)      : TRDI :
SA581105H06(BA)      : TRDI :
SA539456H06Q1(BA)    : TRDI :
SA539456H06Q2(BA)    : TRDI :
SA539486H06Q1(BA)    : IRDI :
SA539486H06Q2(BA)    : IRKI :
```

FIGURE 4.3 (C) NS2 subtype A and (D) NS2 subtype B strains

4.3.1.3 Quasispecies in the NS proteins

Even though both NS proteins are highly conserved, a vast amount of quasispecies were detected that was not seen in the G protein of the same strains of the corresponding specimen. Two clear peaks could be distinguished when the chromatogram was analysed for all strains indicated as quasispecies in figure 4.4. Overall, seven specimens (6 subtype A and 1 subtype B) had quasispecies present within the NS1 protein and nine specimens (4 subtype A and 5 subtype B) in the NS2 protein. There was no specific relationship between the quasispecies found and the rate of hospitalization of RSV, because both the out-patient group as well as the hospitalized group had these quasispecies present in almost equal amount. Some of these quasispecies could only be seen at nucleotide level and did not change the amino acid composition while some had non-synonymous substitutions (Figure 4.5 A-D). No distinct association could be made regarding a specific polymorphism and disease severity. Both NS proteins are highly conserved with only a few polymorphisms present. Sequence analysis of ten molecular clones was performed for four specimens that had these quasispecies. Only up to two different quasispecies was seen for each specimen.

Phylogenetic analysis indicated that the specimens that clustered within a specific genotype during G protein subtyping (GA2 and GA5 for subtype A and BA for subtype B) also clustered together with the NS protein analysis, except for two specimens from 2007 within the subtype A group. SA961333OP07 and SA954715OP07 clustered within GA5 with G protein subtyping but now cluster with GA2 genotype with the NS protein analysis for both the NS1 and NS2 genes. Quasispecies were also identified in both these specimens. The NS1 subtype A phylogenetic tree is indicated (Figure 4.6).

Seq 1
Forward

Seq 1
Reverse

Seq 2
Forward

Seq 2
Reverse

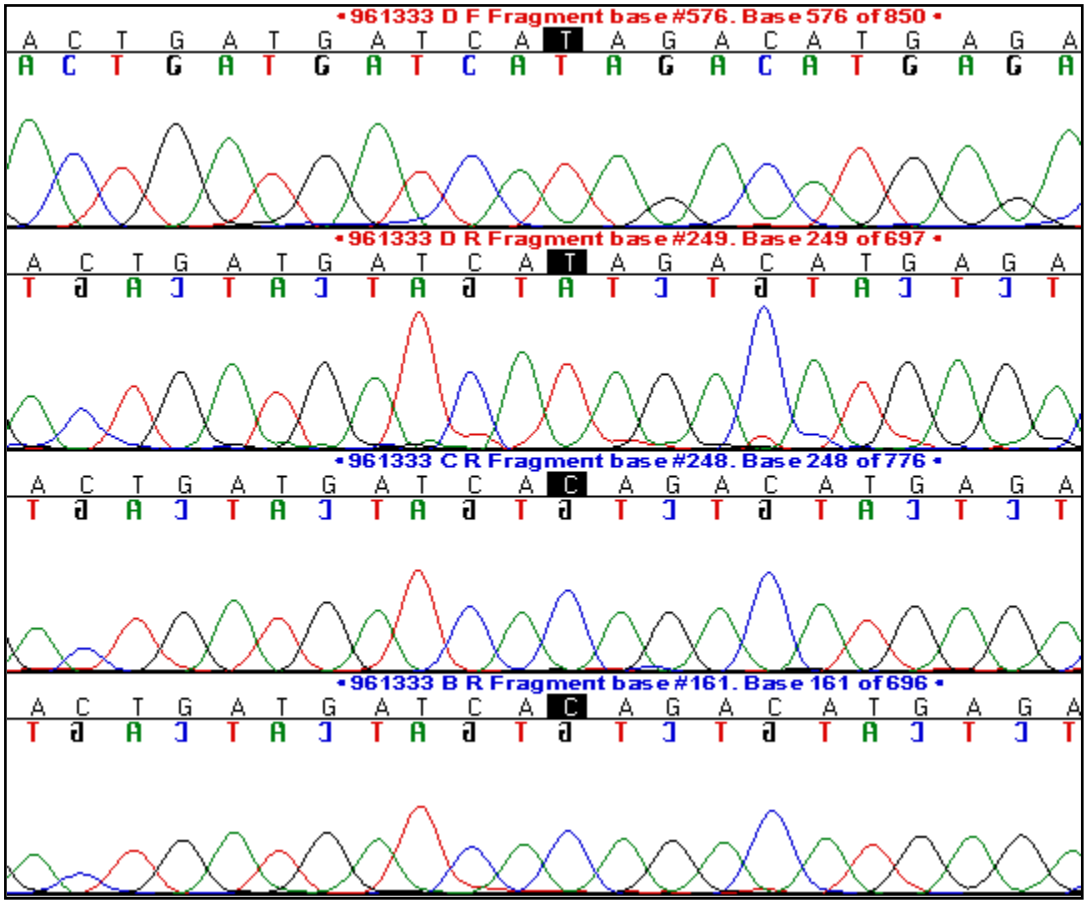


FIGURE 4.4 A chromatogram showing two sequences of one specimen that showed variability within the NS proteins indicated in black suggesting the presence of quasispecies in the NS proteins.

A)

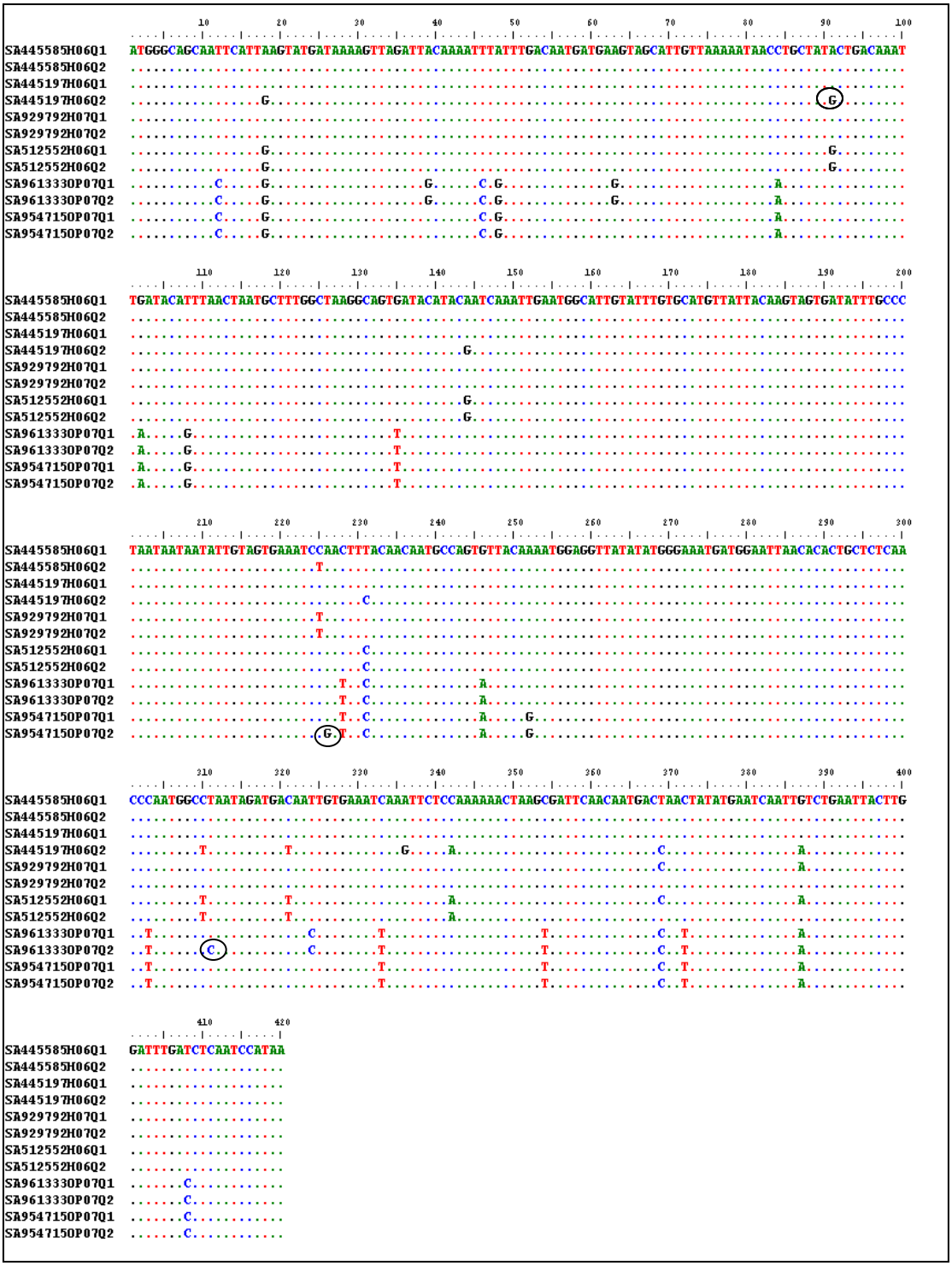
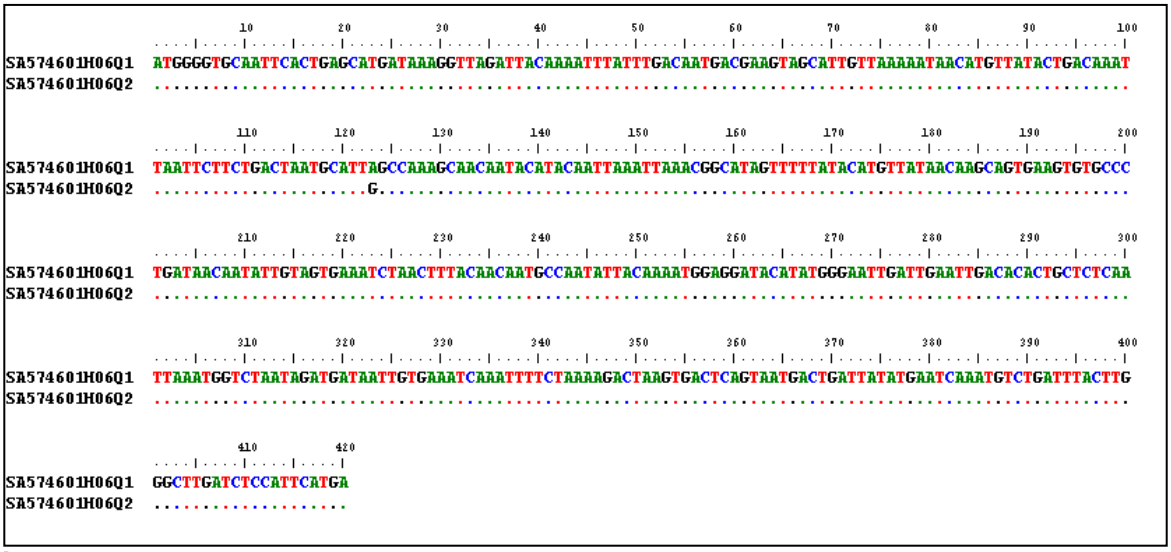


FIGURE 4.5 Nucleotide alignment of **(A)** NS1 subtype A specimens; **(B)** NS1 subtype B specimens; **(C)** NS2 subtype A specimens and **(D)** NS2 subtype B specimens where quasispecies were identified**.** The quasispecies are indicated at the back as Q1, Q2 etc. The encircled bases are non-synonymous substitutions.

B)



C)

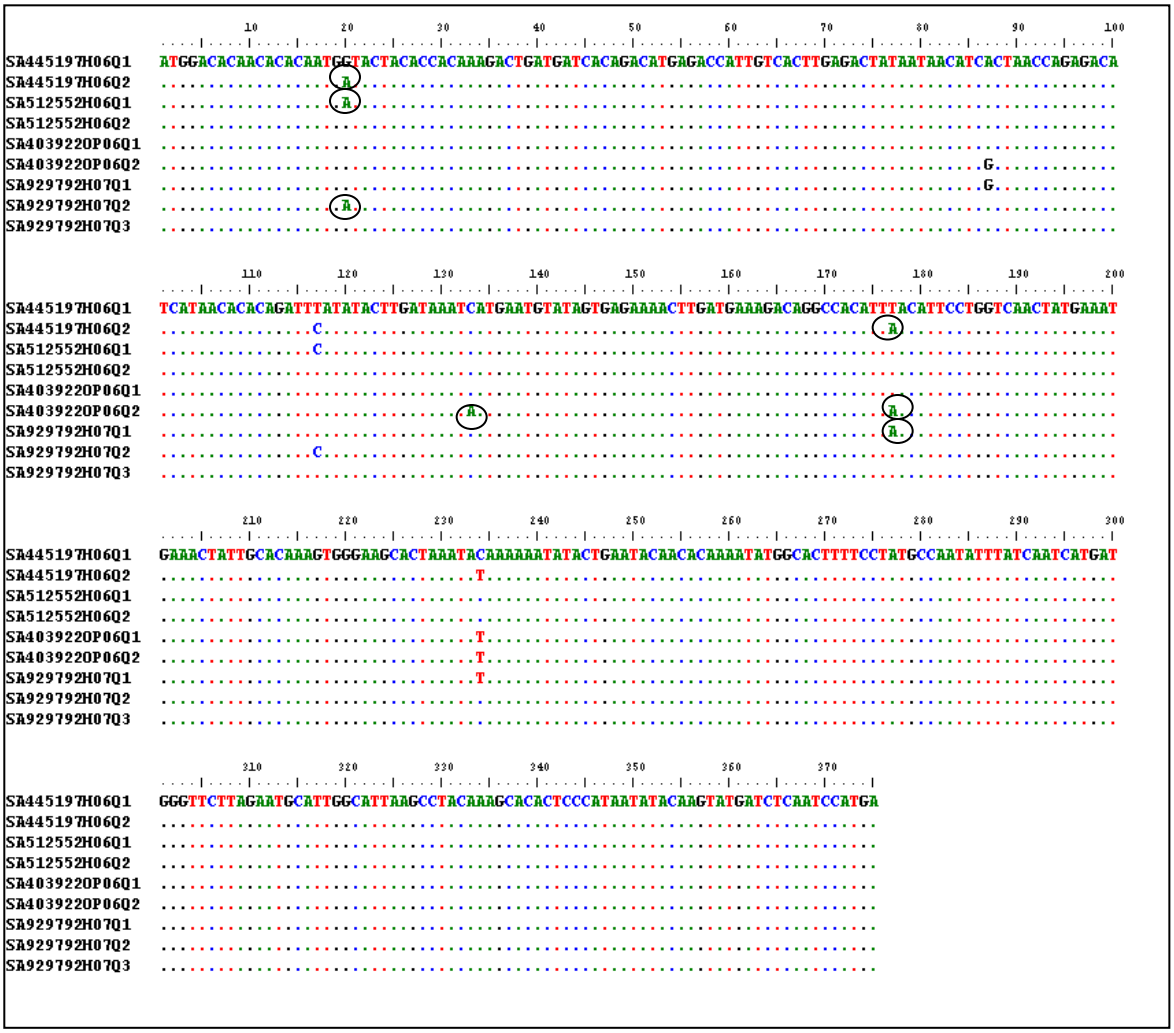


FIGURE 4.5 (B) NS1 subtype B and **(C)** NS2 subtype A specimen.

D)

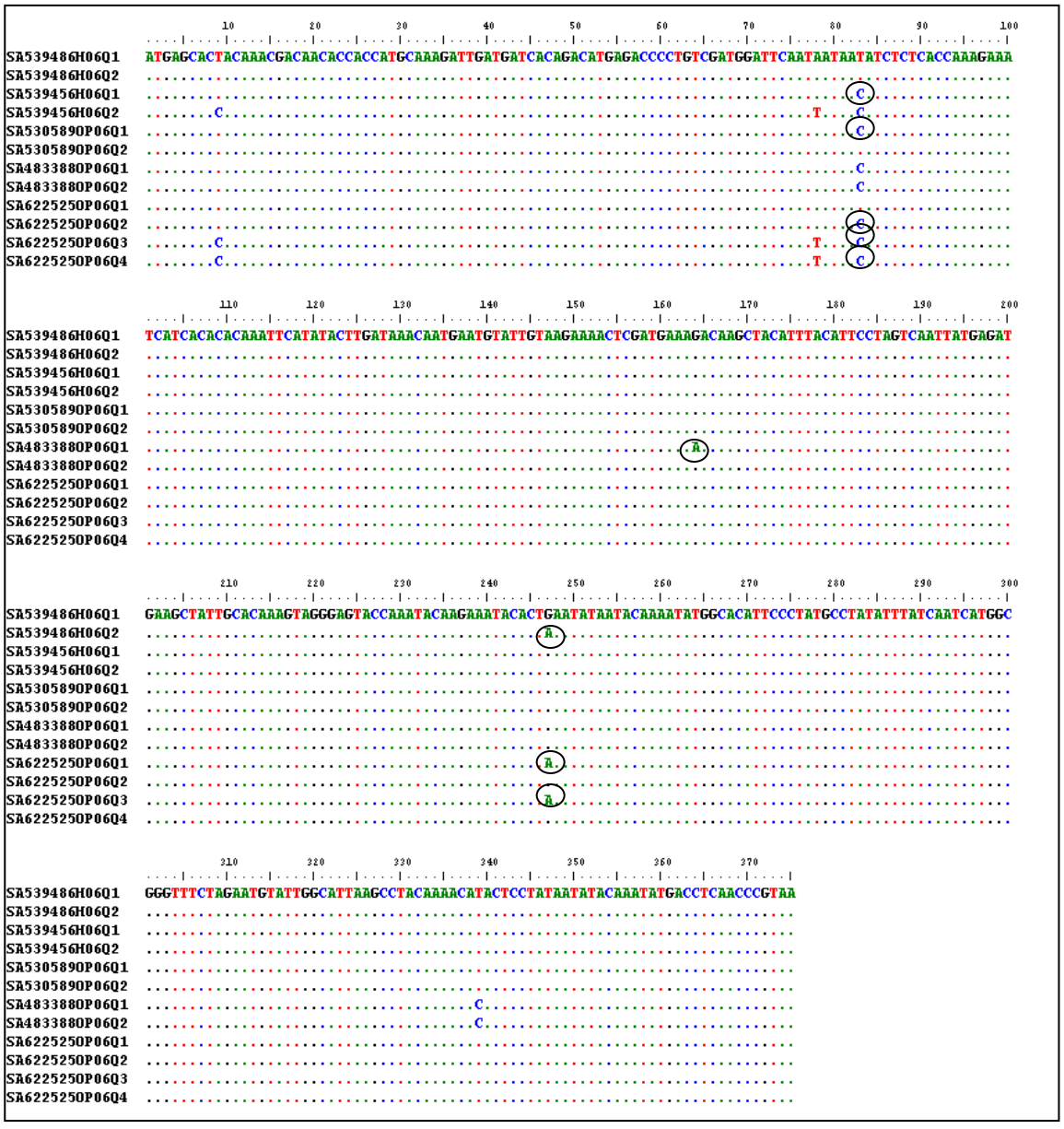


FIGURE 4.5 (D) NS2 subtype B specimens

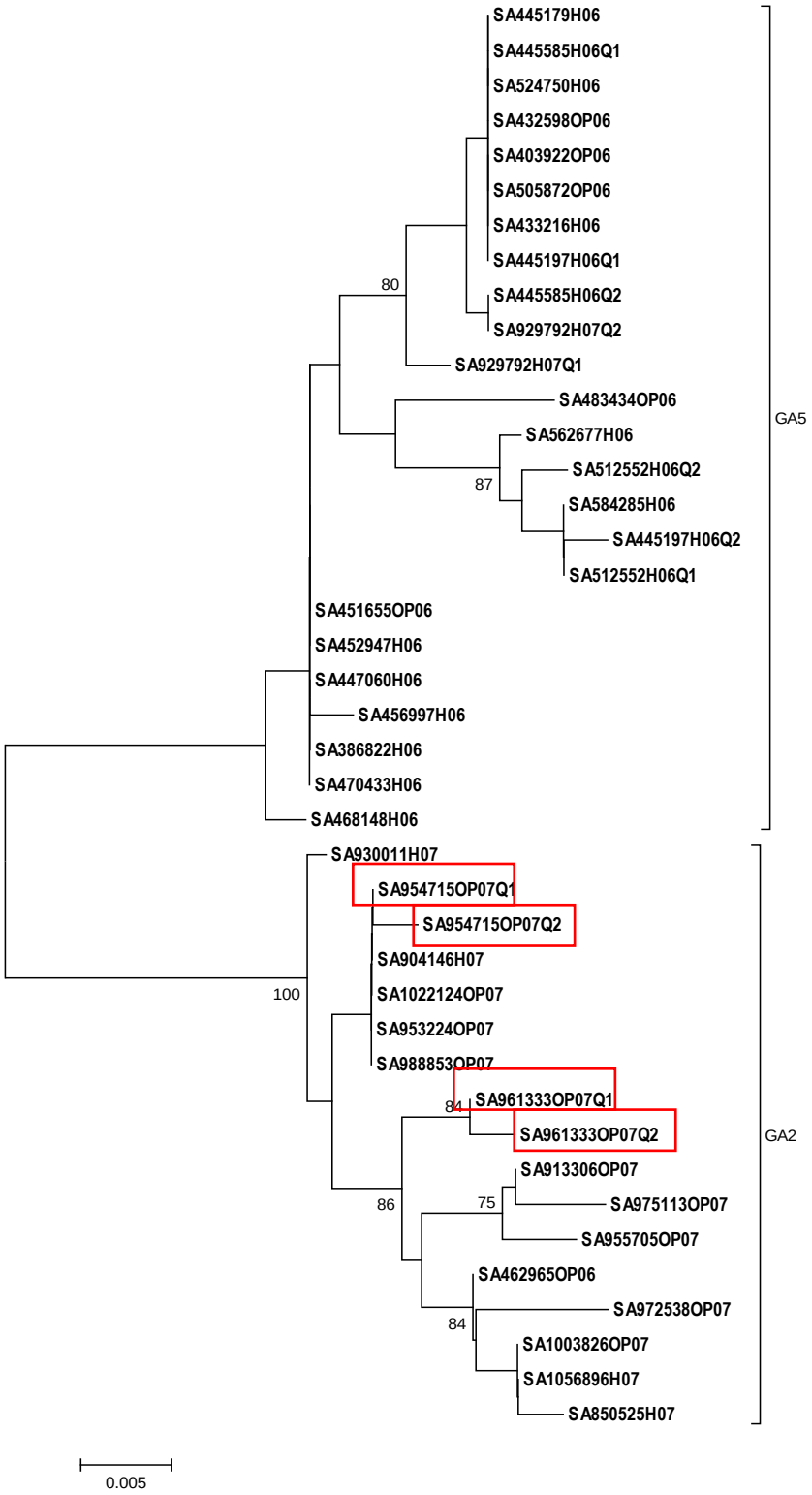


FIGURE 4.6 Midpoint rooted neighbour-joining tree for subtype A NS1 gene constructed with MEGA version 4(Tamura et al., 2007). Only bootstrap values >70 are shown. G-protein Genotypes are indicated on the right with brackets. The two specimens that clustered within genotype GA5 for the G-protein are indicated in the red blocks that now cluster within GA2 for the NS proteins.

4.3.2 Quantitative PCR for detecting expression level differences in the NS proteins

4.3.2.1 Optimization: Specific amplification

Following optimization of the three different primer pairs for NS1, NS2 and the RNA antigenome, all amplified a single PCR product without primer-dimers or any other misprimed products. Figure 4.7 A indicate an agarose gel of the PCR products for the three different regions (NS1, NS2 and antigenome) amplified with SYBR green I. A LightCycler melting curve analysis indicated specific amplification without primer-dimers. All three primer pairs worked well on control samples although upon clinical specimen amplification, the NS2 primers caused non-specific amplification that would lead to incorrect quantification. Taqman probes were further designed for the NS2 region for subtype A and B separately for use in quantification and the amplified products are shown in figure 4.7 B.

4.3.2.2 Optimization: Creating of standard curves and optimizing PCR efficiency

The amplification curves and standard curves that were used for quantification are shown in Appendix B. All three standard curves had a correlation coefficient (r) of -1 showing that a linear relation was observed. Real-time PCR efficiencies were calculated from the given slopes using LightCycler software with the following equation: $Ex=(10^{-1/slope})-1$. Rounded to one decimal, all four PCRs had a PCR efficiency of 1.6 (80%).

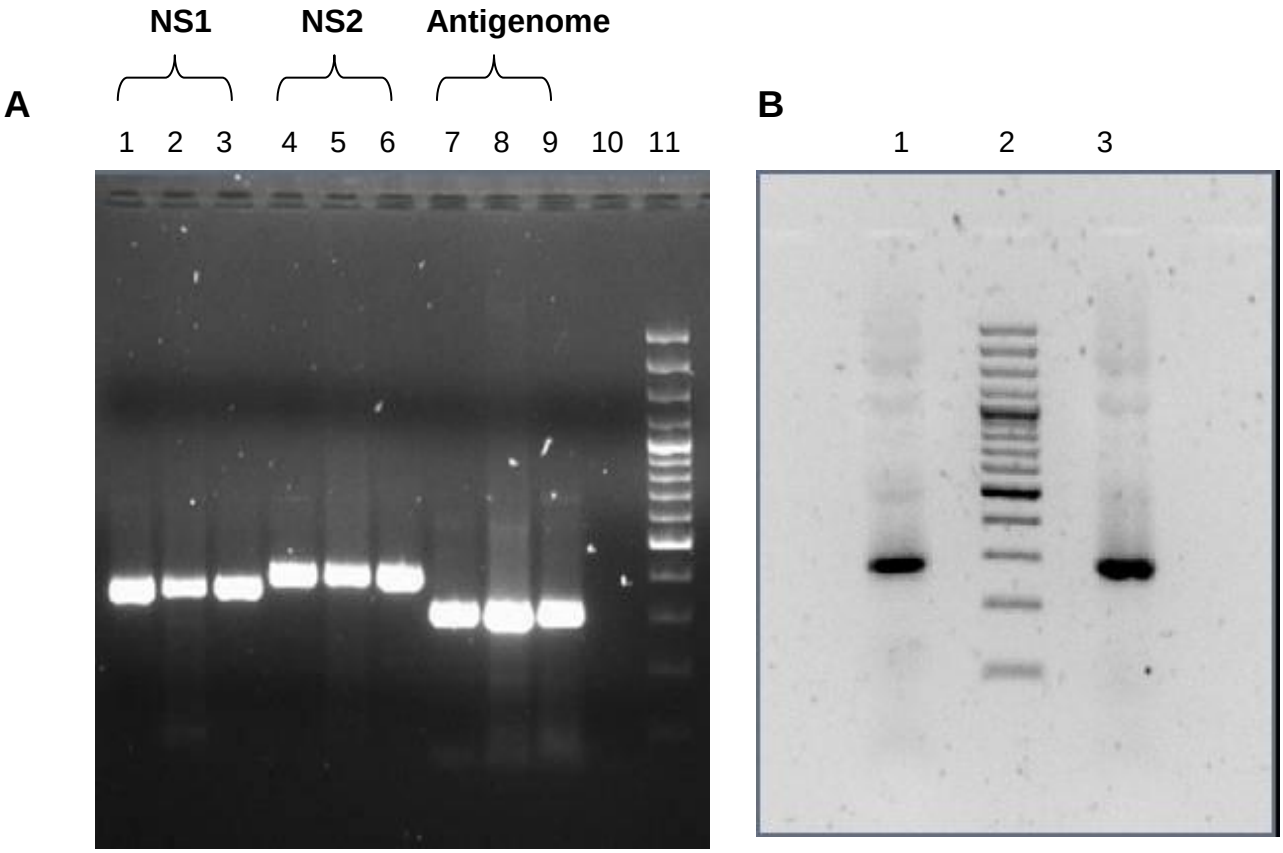


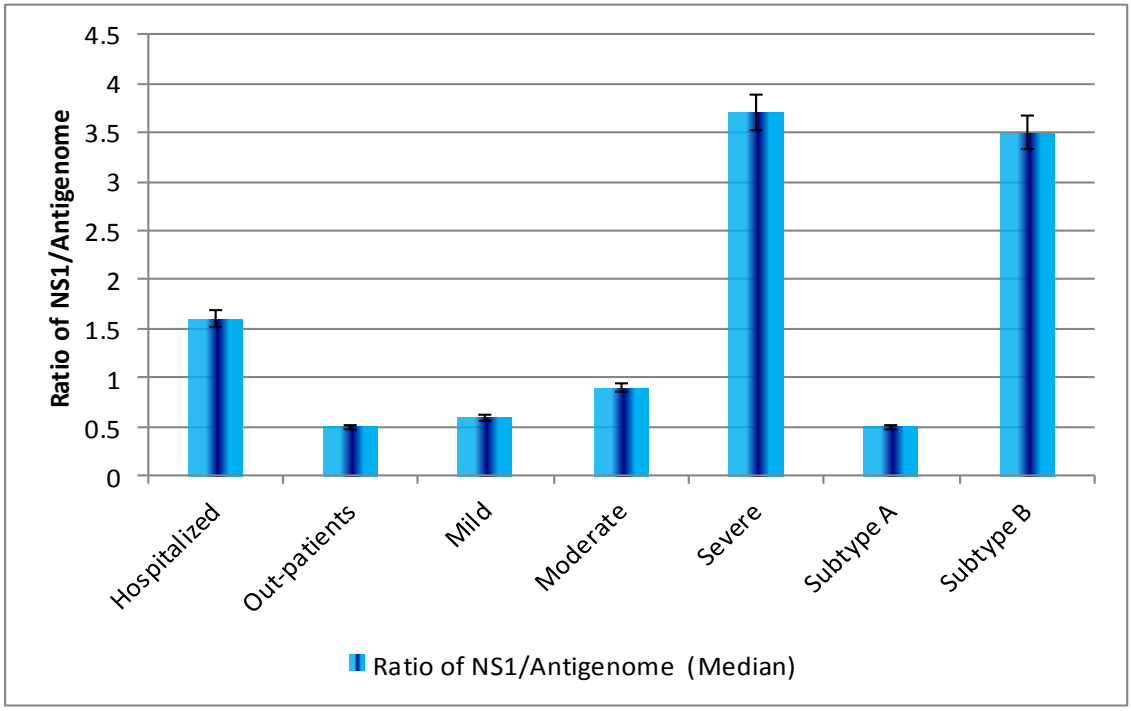
FIGURE 4.7 Agarose gel of the PCR products for **A)** the three different regions (NS1, NS2, RNA template) amplified with SYBR Green I. Lane 1-3=Control and two clinical specimens (RNA template); Lane 4-6=Control and two clinical specimens (NS1); Lane 7-9=Control and two clinical specimens (NS2); Lane 10=Negative PCR control; Lane 11=Molecular weight marker; **B)** subtype A and B gel products amplified with the Lightcyler TaqMan Master kit. Lane 1=Subtype A control(NS2); Lane 3=Subtype B control (NS2); Lane 2=Molecular weight marker.

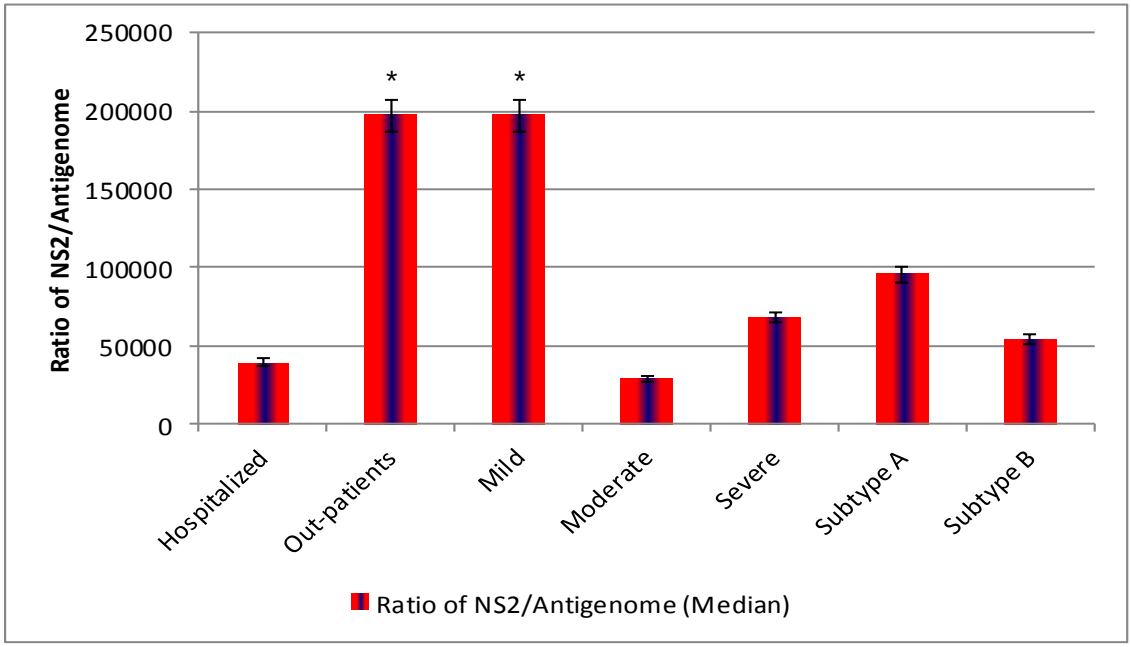
4.3.2.3 Clinical specimen quantification and analysis

A total of 58 specimens where the NS proteins could successfully be amplified in the first round were selected for quantification of the NS1 and NS2 proteins. Of these 33 (57%) were hospitalized patients and 25 (43%) were out-patients. To enable accurate analysis of the relative NS1 and NS2 gene expression between runs the quantitative values were expressed relative to the amount of antigenome present for transcription. The amount of NS1/NS2 proteins relative to the amount of antigenome is given as a ratio: NS1/antigenome and NS2/antigenome.

An assessment of NS1/2 expression level differences with disease severity and infecting subtype can be seen in figure 4.9 A and B. For the NS1 protein, three times higher expression levels were present in the hospitalized group relative to the out-patient group ($p=0.37$). When analyzing disease severity in more detail, patients with moderate infections had twice as much NS1 expression than the mild group, the severe group had four times higher expression then the moderate group, and the severely affected patients had six times higher expression levels then the mild group ($p=0.24$). It was also noted that patients infected with subtype B strains had seven times higher NS1 expression levels then the subtype A infected patients ($p=0.11$) although not statistically significant. The opposite was seen for the NS2 protein. Five times higher expression was seen in the out-patient group compared to the hospitalized group ($p=0.003$). Patients with mild infections had seven and three times higher expression levels then the moderate and severe group respectively, while the severe group had NS2 levels twice as high as the moderate group ($p=0.008$). All of these were highly significant. With respect to infecting subtype, subtype A infected patients had twice as much NS2 levels then patients infected with subtype B ($p=0.46$).

To explain the expression level differences seen in the two NS proteins, the gene-start for NS2 and gene-end for NS1 sequences were also analysed. For both proteins these sequences were 100% conserved with no genetic changes that could explain the differences (Figure 4.10 A and B).

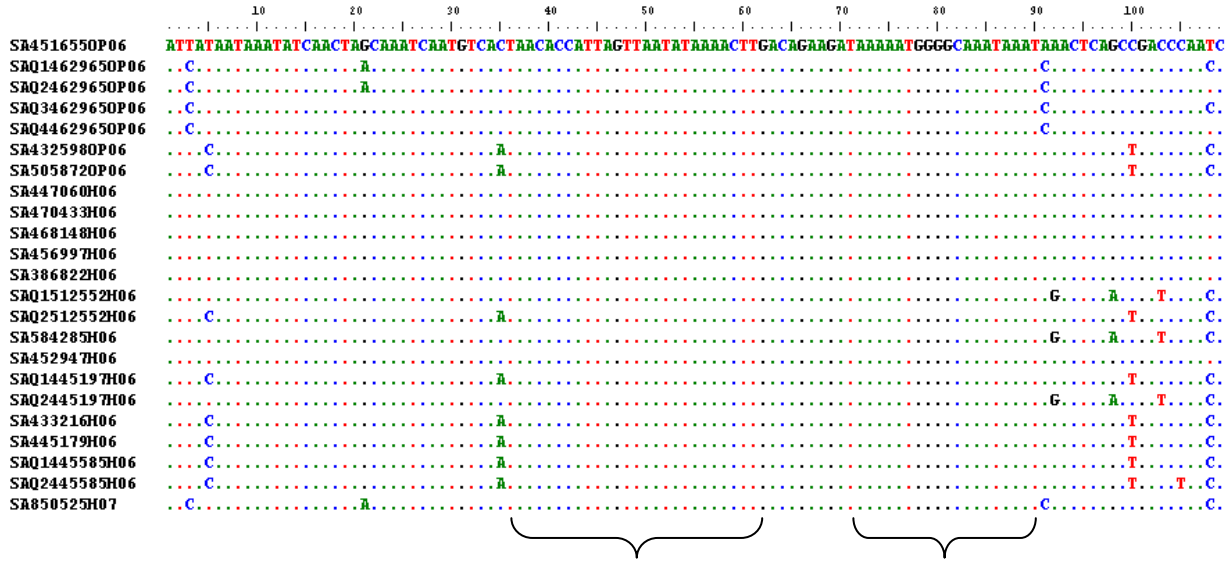




* Statistically significant

FIGURE 4.8 Assessment of NS1/2 expression level differences with disease severity and infecting subtype. Values are expressed as relative values of NS1/2 protein/antigenome.

A)



B)

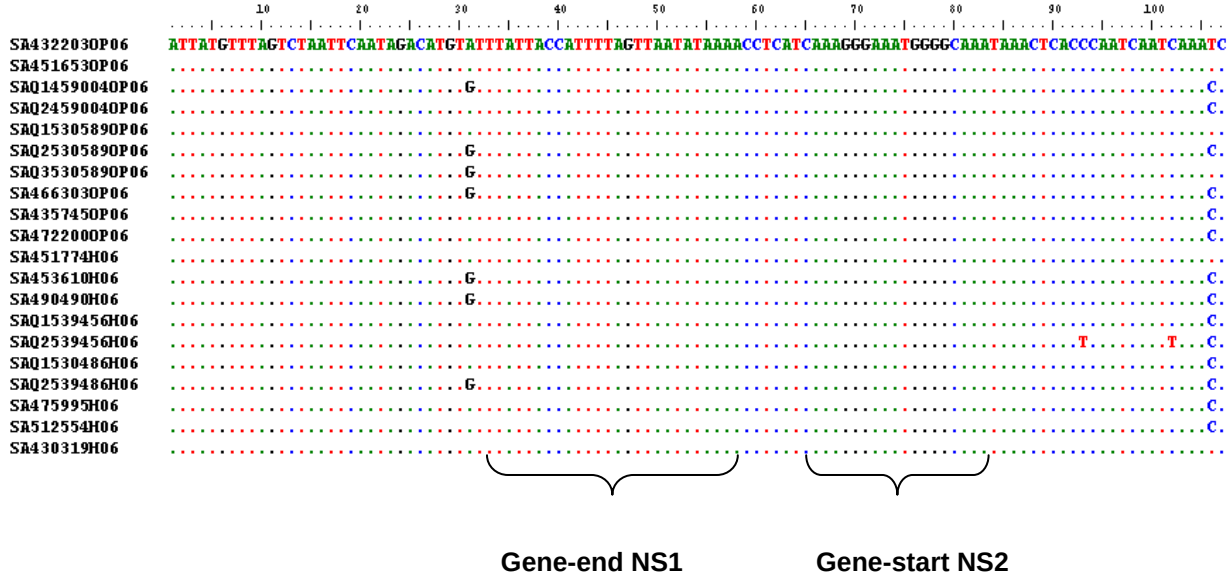


FIGURE 4.9 Nucleotide alignment of **A)** subtype A and **B)** subtype B intron between the NS1 and NS2 gene with the gene-end of NS1 and gene-start of NS2 indicated in brackets.

4.4 DISCUSSION

The two NS proteins play an important part in the infectious life cycle of RSV by antagonizing the host's INF pathway(Spann *et al.*, 2004) by affecting the steady-state levels of key proteins of the IFN pathway(Swedan *et al.*, 2009). We have shown that both proteins are highly conserved at nucleotide and amino acid level, however quasispecies were identified in some specimens. Expression level differences were also seen between specimens and between the two different proteins.

No specific polymorphisms in the NS genes were identified that was associated with more severe disease, although both genes are highly conserved with only a few polymorphisms. We did however see two strains that clustered with genotype GA5 with G protein genotyping but with GA2 with the NS protein sequences. This might indicate recombination between different strains within a specific subtype increasing the variation of the strains.

Quasispecies were identified in some of the clinical specimens and sequence analysis of molecular clones revealed several changes within one individual. Some of the changes were only at nucleotide level while others resulted in amino acid changes. Sanz-Ramos *et al*.(Sanz-Ramos et al., 2008) provided experimental evidence that viral quasispecies in the NS proteins of the Foot-and-mouth disease virus (FMDV) in vivo consists of many subpopulations that evolve independently in different tissues within the same infected host. Their hypothesis is that the quasispecies are composed of "non-pathogenic" and "pathogenic" genomes, and that interplay between the different genomes will lead to virulent or attenuated populations.

Continuous generation of mutant virus is currently regarded as a key adaptive strategy of RNA viruses where the mutations are well tolerated and may not adversely affect protein function(Deplanche et al., 2007). The ability of viruses to generate a complex mutant spectrum may allow the viral populations to adapt and survive in different environments, including selective pressures generated by the host's immune response(Domingo et al., 2006, Pfeiffer and Kirkegaard, 2005, Vignuzzi et al., 2006). It is difficult to analyze the heterogeneity of RNA genomes and no such analysis has been

reported for HRSV, however in vivo quasispecies distributions were evident in the bovine RSV genome(Deplanche et al., 2007).

The IFN system serves as a potent first line of defence against virus infection, however most viruses have responded to this antiviral system by encoding IFN antagonists such as the NS1 protein of Influenza A virus(Wang *et al.*, 2000), the NSs protein of Bunyamwera Bunyavirus(Weber *et al.*, 2002) and the NS1 and NS2 proteins of BRSV(valarcher *et al.*, 2003). This was also shown for the two NS proteins of HRSV(Spann *et al.*, 2004). Therefor we hypothesised that higher levels of both NS proteins will be seen in the hospitalized group compared to the out-patient group. This proved to be true for the NS1 protein, however the opposite was seen for the NS2 protein. Higher expression levels of NS1 was also seen in subtype B infected patients while subtype A infected patients had higher levels of NS2. The NS proteins has been shown to have a complex interaction with a regulatory interaction on each other(Swedan *et al.*, 2009) and this opposite dosing effect for the two proteins may play a role in this.

To investigate these differences, the gene-end sequence of the NS1 gene and the gene-start of the NS2 gene were also analysed to determine if genetic variability within these regions could confer the differences seen in expression levels. Kuo et al.(Kuo *et al.*, 1996) showed that the only sequences required for efficient transcription and replication were the 44-nucleotide 3’ leader region, the last 40 nucleotides of the 5’ trailer region, and the 9- to 10-nucleotide gene-start and 12- to 13-nucleotide gene-end sequences. The gene-start and gene-end sequences were shown to be absolutely conserved between all specimens analysed. Differences within the promoter area in the beginning of the genome (nucleotide 1-44) may also contribute to differences in expression levels and sequencing of the promoter area may elucidate this in future studies.

To conclude, differences in disease severity may not be attributed to specific polymorphisms within the NS genes, however genome heterogeneity (quasispecies) may allow different genomes to circulate within one patient causing the virus to evade immune pressure and enhance disease progression and severity. In addition, higher expression levels of NS1 were seen in the hospitalized group compared to the out-patient group, strengthening our hypothesis that increased levels of NS may enhance

suppression of the host's IFN pathway and thus more severe disease. However, the opposite was shown for NS2. These proteins may act at different sites in the IFN pathway and have a more complicated interplay that work on regulation of this pathway. Down regulation of the IFN response by NS2 could also result in a decreased inflammatory response and milder disease. Genetic variability within the gene-end and gene-start sequences is not responsible for the differences in expression levels observed, although variability within the promoter area may need to be investigated.

CHAPTER 5

Concluding remarks

Respiratory Syncytial virus is the leading cause of pneumonia and bronchiolitis in infants and the elderly (Glezen, 1987). Additionally RSV is an important nosocomial infectious agent and the rate of hospital acquired infection for infants and young children during a RSV season ranges from 26 to 47% in newborn units and from 20 to 40% for older children. Hospital staff members have also been shown to play an important role in the transmission and they have an infection rate of 25% to 50% during RSV epidemics(Hall et al., 1975).

Repeat infections are known to occur throughout life(Wertz and Moudy, 2004) and it is therefore of utmost importance to develop an effective vaccine, which has been unsuccessful to date. It is unclear why some children experience severe disease and others develop only milder disease. It may be due to host factors, maternal immunity or to differences in the virus itself. Characterizing the differences in the virus may potentially aid scientists in the development of a suitable vaccine and targeting therapeutics. In addition, epidemiological analysis of RSV is important for providing the necessary background information for the development of an effective vaccine.

The molecular epidemiology of RSV has been defined in previous investigations in SA in 1997-2000. The emergence of the BA genotype has prompted us to re-investigate the situation 5 years later. Since it has been shown that different genotypes co-circulate each year, it is of interest to investigate the evolution of subtype A and B in South Africa since studies conducted in 2000(Venter et al., 2001) and also to investigate the relationship between G protein variability and disease severity. It was previously shown that the same strains could cause mild or severe disease, however investigation of specimens over 4 years for which clinical data was available may aid in defining the role of protein variability in disease severity.

Bayesian analysis revealed a replacement of all subtype B genotypes previously identified in SA with the BA genotype since 2006. Certain subtype A genotypes identified in previous years are still circulating (GA2 and GA5), although others such as SAA1 was not detected in the 4 years studied. Replacement of all subtype B genotypes with the BA genotype might indicate that the 60 nucleotide duplication may enhance the fitness of these viruses and allow immune evasion in infants who has been previously exposed to subtype B. An additional 6 nucleotide deletion was also detected when the full G protein was sequenced for selected strains. Full genome sequencing of BA strains may elucidate further changes in other proteins. Even though the subtype B strains had more prominent changes, the subtype A strains were shown to evolve faster with a higher evolutionary rate. The slower evolutionary rate of the subtype B strains might explain why more drastic changes are necessary in the subtype B strains to overcome herd immunity. To investigate the evolution of BA strains in SA relative to the rest of the world, a comprehensive Bayesian analysis of SA BA isolates as well as BA isolates from all over the world since it was first discovered in 1999 were performed. It was shown that all SA BA isolates clustered with the BA-IV sub-genotype and the appearance of new sub-genotypes within this branch may occur if drift continues. The most recent common ancestor of the SA strains date back to 1997. Introduction was however only in the past 8 years since no BA strains were detected during the evaluation in 2000.

Analysis over the four years also identified G protein PCR amplicons significantly reduced in size in 2 out of 209 clinical specimens screened over 4 years. Sequence analysis revealed subtype B strains lacking nearly the entire G protein ecto-domain in one HIV positive and one HIV exposed child hospitalized with pneumonia. G protein deletion mutants replicate effectively in vitro but have not been detected in nature (Karron et al., 1997b, Teng and Collins, 2002). This study suggests that RSV clinical strains that lack most of the G protein gene may occur in immunocompormised patients with LRTI. The molecular mechanism whereby this occurs is not clear, however reduced immune pressure in these patients may allow these strains to utilise the F protein for binding and replication. Further characterization of such strains may elucidate the replication and pathogenic potential. In addition, only subtype B deletion mutant strains could be investigated because the primer position on the subtype A strains would have caused us to miss subtype A deletion mutants if they did not come up in the first round

PCR. Future investigations should attempt to design nested primers that will also pick up subtype A deletion mutants.

The two non-structural proteins were subsequently investigated. These proteins have been shown to be involved in inhibiting the host's IFN pathways(Spann et al., 2004) thus playing a vital role in RSV pathogenesis. Sequence variability as well as expression level differences were analyzed to elucidate the role of dose effect of the NS proteins in disease severity. RSV is unique among paramyxoviruses in having two non-structural proteins that play a major role in inhibiting the host's immune response. We were unable to contribute differences in disease severity to specific polymorphisms within the NS genes, however we identified genome heterogeneity (quasispecies) in certain individuals which may allow different genomes to circulate within one patient causing the virus to evade the IFN response and enhance disease progression and severity. In addition, higher expression levels of NS1 were seen in the hospitalized group compared to the out-patient group, strengthening our hypothesis that increased levels of NS protein may lead to enhanced suppression of the host's IFN pathway and thus more severe disease. However, the opposite was shown for NS2. NS1 and NS2 may interact at different stages in the IFN pathway and interplay to complement each other to regulate the innate immune response. Genetic variability within the gene-end and gene-start sequences could not be linked to the differences in expression levels observed, although variability within the promoter area may need to be investigated.

To conclude, acquisition of the 60 nucleotide duplication appeared to have improved the fitness of the BA viruses and more recent subtype B strains may need to be included in experimental vaccines to evaluate their efficacy in the current setting of evolved circulating strains. In addition, the association of clinical strains lacking most of the G protein with LRTI may have implications for the utilisation of certain attenuated strains in immunocompromised children. The two NS proteins were shown to be important during RSV infection, where higher levels of NS1 were shown to be associated with a higher hospitalization rate. Although NS1 and NS2 work together to inhibit the IFN pathway, higher levels of NS2 was present in the out-patient group. This phenomenon needs further investigation to explain to role of NS1 and NS2. The differences in expression levels cannot be attributed to changes in the gene-start and gene-end sequences

although further sequence analysis of the promoter area might be useful in future studies. This study emphasizes the role of both structural and non-structural proteins of RSV to overcome the innate and adaptive responses of the host.

REFERENCES

AGENBACH, E., TIEMESSEN, C. T. & VENTER, M. (2005) Amino acid variation within the fusion protein of respiratory syncytial virus subtype A and B strains during annual epidemics in South Africa. *Virus Genes,* 30**,** 267-78.

AGOTI, C. N., MBISA, J. L., BETT, A., MEDLEY, G. F., NOKES, D. J. & CANE, P. A. (2010) Intrapatient variation of the respiratory syncytial virus attachment protein gene. *Journal of Virology,* 84**,** 10425-10428.

ANDERSON, L. J., HIERHOLZER, J. C., TSOU, C., HENDRY, R. M., FERNIE, B. F., STONE, Y. & MCINTOSH, K. (1985) Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J Infect Dis,* 151**,** 626-33.

ASENJO, A., GONZALEZ-ARMAS, J. C. & VILLANUEVA, N. (2008) Phosphorylation of human respiratory syncytial virus P protein at serine 54 regulates viral uncoating. *Virology,* 380**,** 26-33.

ATREYA, P. L., PEEPLES, M. E. & COLLINS, P. L. (1998) The NS1 protein of human respiratory syncytial virus is a potent inhibitor of minigenome transcription and RNA replication. *J Virol,* 72**,** 1452-61.

BARIK, S., MCLEAN, T. & DUPUY, L. C. (1995) Phosphorylation of Ser232 directly regulates the transcriptional activity of the P protein of human respiratory syncytial virus: phosphorylation of Ser237 may play an accessory role. *Virology,* 213**,** 405-12.

BECKER, Y. (2006) Respiratory syncytial virus (RSV) evades the human adaptive immune system by skewing the Th1/Th2 cytokine balance toward increased levels of Th2 cytokines and IgE, markers of allergy--a review. *Virus Genes,* 33**,** 235-52.

BERMINGHAM, A. & COLLINS, P. L. (1999) The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription. *Proc Natl Acad Sci U S A,* 96**,** 11259-64.

BITKO, V., SHULYAYEVA, O., MAZUMDER, B., MUSIYENKO, A., RAMASWAMY, M., LOOK, D. C. & BARIK, S. (2007) Nonstructural proteins of respiratory syncytial virus suppress premature apoptosis by an NF-kappaB-dependent, interferon-independent mechanism and facilitate virus growth. *J Virol,* 81**,** 1786-95.

BOECK, K. D. (1996) Respiratory syncytial virus bronchiolitis: clinical aspects and epidemiology. *Monaldi Arch Chest Dis,* 51**,** 210-3.

BOTOSSO, V. F., ZANOTTO, P. M., UEDA, M., ARRUDA, E., GILIO, A. E., VIEIRA, S. E., STEWIEN, K. E., PERET, T. C., JAMAL, L. F., PARDINI, M. I., PINHO, J. R., MASSAD, E., SANT'ANNA, O. A., HOLMES, E. C. & DURIGON, E. L. (2009) Positive selection results in frequent reversible amino acid replacements in the G protein gene of human respiratory syncytial virus. *PLoS Pathog,* 5**,** e1000254.

BRANDENBURG, A. H., VAN BEEK, R., MOLL, H. A., OSTERHAUS, A. D. & CLAAS, E. C. (2000) G protein variation in respiratory syncytial virus group A does not correlate with clinical severity. *J Clin Microbiol,* 38**,** 3849-52.

BUCKINGHAM, S. C., JAFRI, H. S., BUSH, A. J., CARUBELLI, C. M., SHEERAN, P., HARDY, R. D., OTTOLINI, M. G., RAMILO, O. & DEVINCENZO, J. P. (2002) A randomized, double-blind, placebo-controlled trial of dexamethasone in severe respiratory syncytial virus (RSV) infection: effects on RSV quantity and clinical outcome. *J Infect Dis,* 185**,** 1222-8.

BUKREYEV, A., YANG, L., FRICKE, J., CHENG, L., WARD, J. M., MURPHY, B. R. & COLLINS, P. L. (2008) The secreted form of respiratory syncytial virus G glycoprotein helps the virus evade antibody-mediated restriction of replication by acting as an antigen decoy and through effects on Fc receptor-bearing leukocytes. *J Virol,* 82**,** 12191-204.

CANE, P. A. (2001) Molecular epidemiology of respiratory syncytial virus. *Rev Med Virol,* 11**,** 103-16.

CANE, P. A., MATTHEWS, D. A. & PRINGLE, C. R. (1994) Analysis of respiratory syncytial virus strain variation in successive epidemics in one city. *J Clin Microbiol,* 32**,** 1-4.

CENTERS FOR DISEASE CONTROL AND PREVENTION, retrieved 4 October 2010, www.cdc.gov/nicdod/dvrd/revb/respiratory/rsvfeat.html.

CHANOCK, R. M., ROIZMAN, B. & MYERS, R. (1957) Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). I. Isolation, properties and characterization. *American Journal of Hygiene,* 66**,** 281-290.

CHRISTENSEN, L. S., LARSEN, L. B., JOHANSEN, J., ANDERSEN, E. A., WEJSE, C., KLUG, B. & HORNSLETH, A. (1999) The fluctuating pattern of various genome types of respiratory syncytial virus in Copenhagen and some other locations in Denmark. *Apmis,* 107**,** 843-50.

COLLINS, P. L., DICKENS, L. E., BUCKLER-WHITE, A., OLMSTED, R. A., SPRIGGS, M. K., CAMARGO, E. & COELINGH, K. V. (1986) Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order. *Proc Natl Acad Sci U S A,* 83**,** 4594-8.

COLLINS, P. L. & GRAHAM, B. S. (2008) Viral and host factors in human respiratory syncytial virus pathogenesis. *J Virol,* 82**,** 2040-55.

COLLINS, P. L., HILL, M. G., CRISTINA, J. & GROSFELD, H. (1996a) Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. *Proc Natl Acad Sci U S A,* 93**,** 81-5.

COLLINS, P. L., HILL, M. G. & JOHNSON, P. R. (1990) The two open reading frames of the 22K mRNA of human respiratory syncytial virus: sequence comparison of antigenic subgroups A and B and expression in vitro. *J Gen Virol,* 71 (Pt 12)**,** 3015-20.

COLLINS, P. L. & JR CROWE, J. E. (2007) Respiratory syncytial virus and metapneumovirus. IN KNIPE, D. M. & HOWLEY, P. M. (Eds.) *Fields Virology.* 5th ed. Philidelphia, Lippincott Williams and Wilkens.

COLLINS, P. L., MCINTOSH, K. & CHANOCK, R. M. (1996b) *Respiratory Syncytial Virus,* Lippencott-Raven, Philadelphia.

COLLINS, P. L. & MURPHY, B. R. (2002) Respiratory syncytial virus: reverse genetics and vaccine strategies. *Virology,* 296**,** 204-11.

COLLINS, P. L. & MURPHY, B. R. (2005) New generation live vaccines against human respiratory syncytial virus designed by reverse genetics. *Proc Am Thorac Soc,* 2**,** 166-73.

COLLINS, P. L., OLMSTED, R. A., SPRIGGS, M. K., JOHNSON, P. R. & BUCKLER-WHITE, A. J. (1987) Gene overlap and site-specific attenuation of transcription of the viral polymerase L gene of human respiratory syncytial virus. *Proc Natl Acad Sci U S A,* 84**,** 5134-8.

COLLINS, P. L. & WERTZ, G. W. (1985) Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of human respiratory syncytial virus. *Virology,* 143**,** 442-51.

CROWE, J. E., JR. (2001) Respiratory syncytial virus vaccine development. *Vaccine,* 20 Suppl 1**,** S32-7.

CROWE, J. E., JR. & WILLIAMS, J. V. (2003) Immunology of viral respiratory tract infection in infancy. *Paediatr Respir Rev,* 4**,** 112-9.

DEPLANCHE, M., LEMAIRE, M., MIRANDETTE, C., BONNET, M., SCHELCHER, F. & MEYER, G. (2007) In vivo evidence for quasispecies distributions in the bovine respiratory syncytial virus genome. *J Gen Virol,* 88**,** 1260-5.

DEVINCENZO, J. P. (2005) Factors predicting childhood respiratory syncytial virus severity: what they indicate about pathogenesis. *Pediatr Infect Dis J,* 24**,** S177-83, discussion S182.

DICKENS, L. E., COLLINS, P. L. & WERTZ, G. W. (1984) Transcriptional mapping of human respiratory syncytial virus. *J Virol,* 52**,** 364-9.

DOMACHOWSKE, J. B. & ROSENBERG, H. F. (1999) Respiratory Syncytial Virus Infection: Immune Response, Immunopathogenesis, and Treatment. *Clinical Microbiology Reviews,* 12**,** 298-309.

DOMINGO, E., MARTIN, V., PERALES, C., GRANDE-PEREZ, A., GARCIA-ARRIAZA, J. & ARIAS, A. (2006) Viruses as quasispecies: biological implications. *Curr Top Microbiol Immunol,* 299**,** 51-82.

DONNERSTEIN, R. L., BERG, R. A., SHEHAB, Z. & OVADIA, M. (1994) Complex atrial tachycardias and respiratory syncytial virus infections in infants. *J Pediatr,* 125**,** 23-8.

DRUMMOND, A. J. & RAMBAUT, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol,* 7**,** 214.

EISENHUT, M. (2006) Extrapulmonary manifestations of severe respiratory syncytial virus infection--a systematic review. *Crit Care,* 10**,** R107.

EVANS, J. E., CANE, P. A. & PRINGLE, C. R. (1996) Expression and characterisation of the NS1 and NS2 proteins of respiratory syncytial virus. *Virus Res,* 43**,** 155-61.

FALSEY, A. R., FORMICA, M. A. & WALSH, E. E. (2002) Diagnosis of respiratory syncytial virus infection: comparison of reverse transcription-PCR to viral culture and serology in adults with respiratory illness. *J Clin Microbiol,* 40**,** 817-20.

FALSEY, A. R., HENNESSEY, P. A., FORMICA, M. A., COX, C. & WALSH, E. E. (2005) Respiratory syncytial virus infection in elderly and high-risk adults. *N Engl J Med,* 352**,** 1749-59.

FELDMAN, S. A., AUDET, S. & BEELER, J. A. (2000) The fusion glycoprotein of human respiratory syncytial virus facilitates virus attachment and infectivity via an interaction with cellular heparan sulfate. *J Virol,* 74**,** 6442-7.

FREYMUTH, F., EUGENE, G., VABRET, A., PETITJEAN, J., GENNETAY, E., BROUARD, J., DUHAMEL, J. F. & GUILLOIS, B. (1995) Detection of respiratory syncytial virus by reverse transcription-PCR and hybridization with a DNA enzyme immunoassay. *J Clin Microbiol,* 33**,** 3352-5.

FUENTES, S., TRAN, K. C., LUTHRA, P., TENG, M. N. & HE, B. (2007) Function of the respiratory syncytial virus small hydrophobic protein. *J Virol,* 81**,** 8361-6.

GALIANO, M. C., PALOMO, C., VIDELA, C. M., ARBIZA, J., MELERO, J. A. & CARBALLAL, G. (2005) Genetic and antigenic variability of human respiratory syncytial virus (groups a and b) isolated over seven consecutive seasons in Argentina (1995 to 2001). *J Clin Microbiol,* 43**,** 2266-73.

GARCIA-BARRENO, B., PORTELA, A., DELGADO, T., LOPEZ, J. A. & MELERO, J. A. (1990) Frame shift mutations as a novel mechanism for the generation of neutralization resistant mutants of human respiratory syncytial virus. *Embo J,* 9**,** 4181-7.

GARCIA, J., GARCIA-BARRENO, B., VIVO, A. & MELERO, J. A. (1993) Cytoplasmic inclusions of respiratory syncytial virus-infected cells: formation of inclusion bodies in transfected cells that coexpress the nucleoprotein, the phosphoprotein, and the 22K protein. *Virology,* 195**,** 243-7.

GARCIA, O., MARTIN, M., DOPAZO, J., ARBIZA, J., FRABASILE, S., RUSSI, J., HORTAL, M., PEREZ-BRENA, P., MARTINEZ, I., GARCIA-BARRENO, B. & ET AL. (1994) Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. *J Virol,* 68**,** 5448-59.

GERNA, G., CAMPANINI, G., ROGNONI, V., MARCHI, A., ROVIDA, F., PIRALLA, A. & PERCIVALLE, E. (2008) Correlation of viral load as determined by real-time RT-PCR and clinical characteristics of respiratory syncytial virus lower respiratory tract infections in early infancy. *J Clin Virol,* 41**,** 45-8.

GLEZEN, W. P. (1987) Incidence of Respiratory Syncytial and Para-influenza type 3 viruses in an urban setting. *Pediatric Virology,* 2**,** 1-4.

GLEZEN, W. P., TABER, L. H., FRANK, A. L. & KASEL, J. A. (1986) Risk of primary infection and reinfection with respiratory syncytial virus. *Am J Dis Child,* 140**,** 543-6.

GOETGHEBUER, T., ISLES, K., MOORE, C., THOMSON, A., KWIATKOWSKI, D. & HULL, J. (2004) Genetic predisposition to wheeze following respiratory syncytial virus bronchiolitis. *Clin Exp Allergy,* 34**,** 801-3.

GRAHAM, B. S. (1996) Immunological determinants of disease caused by respiratory syncytial virus. *Trends Microbiol,* 4**,** 290-3.

GROSFELD, H., HILL, M. G. & COLLINS, P. L. (1995) RNA replication by respiratory syncytial virus (RSV) is directed by the N, P, and L proteins; transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-length mRNA. *J Virol,* 69**,** 5677-86.

GUINDON, S., LETHIEC, F., DUROUX, P. & GASCUEL, O. (2005) PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res,* 33**,** W557-9.

HALL, C. B. (2001) Respiratory syncytial virus and parainfluenza virus. *N Engl J Med,* 344**,** 1917-28.

HALL, C. B., DOUGLAS, R. G., JR., GEIMAN, J. M. & MESSNER, M. K. (1975) Nosocomial respiratory syncytial virus infections. *N Engl J Med,* 293**,** 1343-6.

HALL, C. B., POWELL, K. R. & MACDONALD, N. E. (1986) Respiratory syncytial viral infection in children with compromised immune function. *New England Journal of Medicine,* 315**,** 77-81.

HALL, C. B., WALSH, E. E., SCHNABEL, K. C., LONG, C. E., MCCONNOCHIE, K. M., HILDRETH, S. W. & ANDERSON, L. J. (1990) Occurrence of groups A and B of respiratory syncytial virus over 15 years: associated epidemiologic and clinical characteristics in hospitalized and ambulatory children. *J Infect Dis,* 162**,** 1283-90.

HALL, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser,* 41**,** 95-98.

HANCOCK, G. E., SPEELMAN, D. J., FRENCHICK, P. J., MINEO-KUHN, M. M., BAGGS, R. B. & HAHN, D. J. (1995) Formulation of the purified fusion protein of respiratory syncytial virus with the saponin QS-21 induces protective immune responses in Balb/c mice that are similar to those generated by experimental infection. *Vaccine,* 13**,** 391-400.

HARDY, R. W. & WERTZ, G. W. (1998) The product of the respiratory syncytial virus M2 gene ORF1 enhances readthrough of intergenic junctions during viral transcription. *J Virol,* 72**,** 520-6.

HENKEL, J. H., ABERLE, S. W., KUNDI, M. & POPOW-KRAUPP, T. (1997) Improved detection of respiratory syncytial virus in nasal aspirates by seminested RT-PCR. *J Med Virol,* 53**,** 366-71.

HENRICKSON, K. J. (2004) Advances in the laboratory diagnosis of viral respiratory disease. *Pediatric Infectious diseases Journal,* 23**,** S6-S10.

HENRICKSON, K. J. & HALL, C. B. (2007) Diagnostic assays for respiratory syncytial virus disease. *Pediatr Infect Dis J,* 26**,** S36-40.

JANSSEN, R., BONT, L., SIEZEN, C. L., HODEMAEKERS, H. M., ERMERS, M. J., DOORNBOS, G., VAN 'T SLOT, R., WIJMENGA, C., GOEMAN, J. J., KIMPEN, J. L., VAN HOUWELINGEN, H. C., KIMMAN, T. G. & HOEBEE, B. (2007) Genetic susceptibility to respiratory syncytial virus bronchiolitis is predominantly associated with innate immune genes. *J Infect Dis,* 196**,** 826-34.

JENKINS, G. M., RAMBAUT, A., PYBUS, O. G. & HOLMES, E. C. (2002) Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J Mol Evol,* 54**,** 156-65.

JOHNSON, P. R. & COLLINS, P. L. (1989) The 1B (NS2), 1C (NS1) and N proteins of human respiratory syncytial virus (RSV) of antigenic subgroups A and B: sequence conservation and divergence within RSV genomic RNA. *J Gen Virol,* 70 (Pt 6)**,** 1539-47.

JOHNSON, P. R., JR., OLMSTED, R. A., PRINCE, G. A., MURPHY, B. R., ALLING, D. W., WALSH, E. E. & COLLINS, P. L. (1987a) Antigenic relatedness between glycoproteins of human respiratory syncytial virus subgroups A and B: evaluation of the contributions of F and G glycoproteins to immunity. *J Virol,* 61**,** 3163-6.

JOHNSON, P. R., SPRIGGS, M. K., OLMSTED, R. A. & COLLINS, P. L. (1987b) The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proc Natl Acad Sci U S A,* 84**,** 5625-9.

KARRON, R. A., BUONAGURIO, D. A., GEORGIU, A. F., WHITEHEAD, S. S., ADAMUS, J. E., CLEMENTS-MANN, M. L., HARRIS, D. O., RANDOLPH, V. B., UDEM, S. A., MURPHY, B. R. & SIDHU, M. S. (1997a) Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral replication in vitro: clinical evaluation and molecular characterization of a cold-passaged, attenuated RSV subgroup B mutant. *Proc Natl Acad Sci U S A,* 94**,** 13961-6.

KARRON, R. A., WRIGHT, P. F., BELSHE, R. B., THUMAR, B., CASEY, R., NEWMAN, F., POLACK, F. P., RANDOLPH, V. B., DEATLY, A., HACKELL, J., GRUBER, W., MURPHY, B. R. & COLLINS, P. L. (2005) Identification of a recombinant live attenuated respiratory syncytial virus vaccine candidate that is highly attenuated in infants. *J Infect Dis,* 191**,** 1093-104.

KARRON, R. A., WRIGHT, P. F., CROWE, J. E., JR., CLEMENTS-MANN, M. L., THOMPSON, J., MAKHENE, M., CASEY, R. & MURPHY, B. R. (1997c) Evaluation of two live, cold-passaged, temperature-sensitive respiratory syncytial virus vaccines in chimpanzees and in human adults, infants, and children. *J Infect Dis,* 176**,** 1428-36.

KHO, N., KERRIGAN, J. F., TONG, T., BROWNE, R. & KNILANS, J. (2004) Respiratory syncytial virus infection and neurologic abnormalities: retrospective cohort study. *J Child Neurol,* 19**,** 859-64.

KIM, H. W., CANCHOLA, J. G., BRANDT, C. D., PYLES, G., CHANOCK, R. M., JENSEN, K. & PARROTT, R. H. (1969) Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. *Am J Epidemiol,* 89**,** 422-34.

KIM, Y. K., CHOI, E. H. & LEE, H. J. (2007) Genetic variability of the fusion protein and circulation patterns of genotypes of the respiratory syncytial virus. *J Med Virol,* 79**,** 820-8.

KOTELKIN, A., BELYAKOV, I. M., YANG, L., BERZOFSKY, J. A., COLLINS, P. L. & BUKREYEV, A. (2006) The NS2 protein of human respiratory syncytial virus suppresses the cytotoxic T-cell response as a consequence of suppressing the type I interferon response. *J Virol,* 80**,** 5958-67.

KUMAR, M. L., SUPER, D. M., LEMBO, R. M., THOMAS, F. C. & PROKAY, S. L. (1987) Diagnostic efficacy of two rapid tests for detection of respiratory syncytial virus antigen. *J Clin Microbiol,* 25**,** 873-5.

KUO, L., GROSFELD, H., CRISTINA, J., HILL, M. G. & COLLINS, P. L. (1996) Effects of mutations in the gene-start and gene-end sequence motifs on transcription of monocistronic and dicistronic minigenomes of respiratory syncytial virus. *J Virol,* 70**,** 6892-901.

KUROIWA, Y., NAGAI, K., OKITA, L., YUI, I., KASE, T., NAKAYAMA, T. & TSUTSUMI, H. (2005) A phylogenetic study of human respiratory syncytial viruses group A and B strains isolated in two cities in Japan from 1980-2002. *J Med Virol,* 76**,** 241-7.

LASSAUNIÉRE, R., KRESFELDER, T. & VENTER, M. (2010) A novel multiplex real-time RT-PCR assay with FRET hybridization probes for the detection and quantification of 13 traditional and newly identified respiratory viruses. *Journal of Virological Methods*.

LAZAR, I., CANAAN, A., WEIBEL, C. & KAHN, J. S. (2006) Novel mutations in the respiratory syncytial virus G gene identified in viral isolates from a girl with severe combined immune deficiency treated with intravenous immune globulin. *J Clin Virol,* 37**,** 168-73.

LE BON, A., SCHIAVONI, G., D'AGOSTINO, G., GRESSER, I., BELARDELLI, F. & TOUGH, D. F. (2001) Type I interferons potently enhance humoral immunity and can promote isotype switching by stimulating dendritic cells in vivo. *Immunity,* 14**,** 461-470.

LE MERCIER, P., GARCIN, D., HAUSMANN, S. & KOLAKOFSKY, D. (2002) Ambisense Sendai viruses are inherently unstable but are useful to study viral RNA synthesis. *Journal of Virology,* 76**,** 5492-5502.

LI, D., JANS, D. A., BARDIN, P. G., MEANGER, J., MILLS, J. & GHILDYAL, R. (2008) Association of respiratory syncytial virus M protein with viral nucleocapsids is mediated by the M2-1 protein. *J Virol,* 82**,** 8863-70.

LLORENTE, M. T., GARCIA-BARRENO, B., CALERO, M., CAMAFEITA, E., LOPEZ, J. A., LONGHI, S., FERRON, F., VARELA, P. F. & MELERO, J. A. (2006) Structural analysis of the human respiratory syncytial virus phosphoprotein: characterization of an alpha-helical domain involved in oligomerization. *J Gen Virol,* 87**,** 159-69.

MADHI, S. A., KUWANDA, L., CUTLAND, C. & KLUGMAN, K. P. (2006) Five-year cohort study of hospitalization for respiratory syncytial virus associated lower respiratory tract infection in African children. *J Clin Virol,* 36**,** 215-21.

MADHI, S. A., VENTER, M., MADHI, A., PETERSEN, M. K. & KLUGMAN, K. P. (2001) Differing manifestations of respiratory syncytial virus-associated severe lower respiratory tract infections in human immunodeficiency virus type 1-infected and uninfected children. *Pediatr Infect Dis J,* 20**,** 164-70.

MAGGON, K. & BARIK, S. (2004) New drugs and treatment for respiratory syncytial virus. *Rev Med Virol,* 14**,** 149-68.

MARTINELLO, R. A., CHEN, M. D., WEIBEL, C. & KAHN, J. S. (2002) Correlation between respiratory syncytial virus genotype and severity of illness. *J Infect Dis,* 186**,** 839-42.

MARTINEZ, I., DOPAZO, J. & MELERO, J. A. (1997) Antigenic structure of the human respiratory syncytial virus G glycoprotein and relevance of hypermutation events for the generation of antigenic variants. *J Gen Virol,* 78 (Pt 10)**,** 2419-29.

MARTINEZ, I., VALDES, O., DELFRARO, A., ARBIZA, J., RUSSI, J. & MELERO, J. A. (1999) Evolutionary pattern of the G glycoprotein of human respiratory syncytial viruses from antigenic group B: the use of alternative termination codons and lineage diversification. *J Gen Virol,* 80 (Pt 1)**,** 125-30.

MCCARTHY, C. A. & HALL, C. B. (2003) Respiratory syncytial virus: concerns and control. *Pediatr Rev,* 24**,** 301-9.

MCNAMARA, P. S., FLANAGAN, B. F., HART, C. A. & SMYTH, R. L. (2005) Production of chemokines in the lungs of infants with severe respiratory syncytial virus bronchiolitis. *J Infect Dis,* 191**,** 1225-32.

MCNAMARA, P. S. & SMYTH, R. L. (2002) The pathogenesis of respiratory syncytial virus disease in childhood. *Br Med Bull,* 61**,** 13-28.

MELERO, J. A., GARCIA-BARRENO, B., MARTINEZ, I., PRINGLE, C. R. & CANE, P. A. (1997) Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. *J Gen Virol,* 78 (Pt 10)**,** 2411-8.

MOORE, E. C., BARBER, J. & TRIPP, R. A. (2008) Respiratory syncytial virus (RSV) attachment and nonstructural proteins modify the type I interferon response associated with suppressor of cytokine signaling (SOCS) proteins and IFN-stimulated gene-15 (ISG15). *Virol J,* 5**,** 116.

MOORE, M. L. & PEEBLES, R. S., JR. (2006) Respiratory syncytial virus disease mechanisms implicated by human, animal model, and in vitro data facilitate vaccine strategies and new therapeutics. *Pharmacol Ther,* 112**,** 405-24.

MORRIS, J. A., BLOUNT, R. E. & SAVAGE, R. E. (1956) Recovery of cytopahtic agent from chimpanzees with coryza. *Preceedings of the Society for Experimental Biology and Medicine,* 92**,** 544-550.

MORRISON, T. B., WEIS, J. J. & WITTWER, C. T. (1998) Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques,* 24**,** 954-8, 960, 962.

MUNIR, S., LE NOUEN, C., LUONGO, C., BUCHHOLZ, U. J., COLLINS, P. L. & BUKREYEV, A. (2008) Nonstructural proteins 1 and 2 of respiratory syncytial virus suppress maturation of human dendritic cells. *J Virol,* 82**,** 8780-96.

MURPHY, B. R., GRAHAM, B. S., PRINCE, G. A., WALSH, E. E., CHANOCK, R. M., KARZON, D. T. & WRIGHT, P. F. (1986) Serum and nasal-wash immunoglobulin G and A antibody response of infants and children to respiratory syncytial virus F and G glycoproteins following primary infection. *J Clin Microbiol,* 23**,** 1009-14.

NAGAI, K., KAMASAKI, H., KUROIWA, Y., OKITA, L. & TSUTSUMI, H. (2004) Nosocomial outbreak of respiratory syncytial virus subgroup B variants with the 60 nucleotides-duplicated G protein gene. *J Med Virol,* 74**,** 161-5.

NATIONAL DEPARTMENT OF HEALTH SOUTH AFRICA. (2010) Guidelines for the Management of HIV in Children. (2nd ed).

NEI, M. & GOJOBORI, T. (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol,* 3**,** 418-26.

NICHOLAS, K. B. & NICHOLAS, H. B. (1997) *GeneDoc: a tool for editing and annotating multiple sequence alignments*.

NOKES, D. J., OKIRO, E. A., NGAMA, M., OCHOLA, R., WHITE, L. J., SCOTT, P. D., ENGLISH, M., CANE, P. A. & MEDLEY, G. F. (2008) Respiratory syncytial virus infection and disease in infants and young children observed from birth in Kilifi District, Kenya. *Clin Infect Dis,* 46**,** 50-7.

NOYOLA, D. E. & DEMMLER, G. J. (2000) Effect of rapid diagnosis on management of influenza A infections. *Pediatr Infect Dis J,* 19**,** 303-7.

OLSON, M. R. & VARGA, S. M. (2008) Pulmonary immunity and immunopathology: lessons from respiratory syncytial virus. *Expert Rev Vaccines,* 7**,** 1239-55.

PARKER, J. M. R., GUO, D. & HODGES, R. S. (1986) New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and x-ray-derived accessible sites. *Biochemistry,* 25**,** 5425-5432.

PASTEY, M. K. & SAMAL, S. K. (1995) Nucleotide sequence analysis of the non-structural NS1 (1C) and NS2 (1B) protein genes of bovine respiratory syncytial virus. *Journal of General Virology,* 76**,** 193-197.

PERET, T. C., HALL, C. B., HAMMOND, G. W., PIEDRA, P. A., STORCH, G. A., SULLENDER, W. M., TSOU, C. & ANDERSON, L. J. (2000) Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. *J Infect Dis,* 181**,** 1891-6.

PERET, T. C., HALL, C. B., SCHNABEL, K. C., GOLUB, J. A. & ANDERSON, L. J. (1998) Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. *J Gen Virol,* 79 (Pt 9)**,** 2221-9.

PERKINS, S. M., WEBB, D. L., TORRANCE, S. A., EL SALEEBY, C., HARRISON, L. M., AITKEN, J. A., PATEL, A. & DEVINCENZO, J. P. (2005) Comparison of a real-time reverse transcriptase PCR assay and a culture technique for quantitative assessment of viral load in children naturally infected with respiratory syncytial virus. *J Clin Microbiol,* 43**,** 2356-62.

PFEIFFER, J. K. & KIRKEGAARD, K. (2005) Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. *PLoS Pathog,* 1**,** e11.

PLOEGH, H. L. (1998) Viral strategies of immune evasion. *Science,* 280**,** 248-53.

PLUMET, S. & GERLIER, D. (2005) Optimized SYBR green real-time PCR assay to quantify the absolute copy number of measles virus RNAs using gene specific primers. *J Virol Methods,* 128**,** 79-87.

POWER, U. F., NGUYEN, T. N., RIETVELD, E., DE SWART, R. L., GROEN, J., OSTERHAUS, A. D., DE GROOT, R., CORVAIA, N., BECK, A., BOUVERET-LE-CAM, N. & BONNEFOY, J. Y. (2001) Safety and immunogenicity of a novel recombinant subunit respiratory syncytial virus vaccine (BBG2Na) in healthy young adults. *J Infect Dis,* 184**,** 1456-60.

RAKUSAN, T. A., PARROTT, R. H. & SEVER, J. L. (1991) Limitations in the laboratory diagnosis of vertically acquired HIV infection. *J Acquir Immune Defic Syndr,* 4**,** 116-21.

RASMUSSEN, R. (2001) *Quantification on the lightcycler,* Heidelberg, Springer Press.

RICCETTO, A. G., SILVA, L. H., SPILKI, F. R., MORCILLO, A. M., ARNS, C. W. & BARACAT, E. C. (2009) Genotypes and clinical data of respiratory syncytial virus and metapneumovirus in brazilian infants: a new perspective. *Braz J Infect Dis,* 13**,** 35-9.

ROBERTS, S. R., LICHTENSTEIN, D., BALL, L. A. & WERTZ, G. W. (1994) The membrane-associated and secreted forms of the respiratory syncytial virus attachment glycoprotein G are synthesized from alternative initiation codons. *J Virol,* 68**,** 4538-46.

RUEDA, P., DELGADO, T., PORTELA, A., MELERO, J. A. & GARCIA-BARRENO, B. (1991) Premature stop codons in the G glycoprotein of human respiratory syncytial viruses resistant to neutralization by monoclonal antibodies. *J Virol,* 65**,** 3374-8.

SANZ-RAMOS, M., DIAZ-SAN SEGUNDO, F., ESCARMIS, C., DOMINGO, E. & SEVILLA, N. (2008) Hidden virulence determinants in a viral quasispecies in vivo. *J Virol,* 82**,** 10465-76.

SATO, M., SAITO, R., SAKAI, T., SANO, Y., NISHIKAWA, M., SASAKI, A., SHOBUGAWA, Y., GEJYO, F. & SUZUKI, H. (2005) Molecular epidemiology of respiratory syncytial virus infections among children with acute respiratory symptoms in a community over three seasons. *J Clin Microbiol,* 43**,** 36-40.

SAVON, C., GOYENECHEA, A., VALDES, O., AGUILAR, J., GONZALEZ, G., PALERM, L., GONZALEZ, G. & PEREZ BRENA, P. (2006) Respiratory syncytial virus group A and B genotypes and disease severity among Cuban children. *Arch Med Res,* 37**,** 543-7.

SCOTT, P., OCHOLA, R., NGAMA, M., OKIRO, E., NOKES, D., MEDLEY, G. & CANE, P. (2004a) Molecular epidemiology of respiratory syncytial virus in Kilifi district, Kenya. *J Clin Microbiol,* 74**,** 344-354.

SCOTT, P. D., OCHOLA, R., NGAMA, M., OKIRO, E. A., NOKES, D. J., MEDLEY, G. F. & CANE, P. A. (2004b) Molecular epidemiology of respiratory syncytial virus in Kilifi district, Kenya. *J Med Virol,* 74**,** 344-54.

SEN, G. C. (2001) Viruses and interferons. *Annu Rev Microbiol,* 55**,** 255-81.

SHAY, D. K., HOLMAN, R. C., ROOSEVELT, G. E., CLARKE, M. J. & ANDERSON, L. J. (2001) Bronchiolitis-associated mortality and estimates of respiratory syncytial virus-associated deaths among US children, 1979-1997. *J Infect Dis,* 183**,** 16-22.

SHOBUGAWA, Y., SAITO, R., SANO, Y., ZARAKET, H., SUZUKI, Y., KUMAKI, A., DAPAT, I., OGUMA, T., YAMAGUCHI, M. & SUZUKI, H. (2009) Emerging genotypes of human respiratory syncytial virus subgroup A among patients in Japan. *J Clin Microbiol,* 47**,** 2475-82.

SIBER, G. R., LEOMBRUNO, D., LESZCZYNSKI, J., MCIVER, J., BODKIN, D., GONIN, R., THOMPSON, C. M., WALSH, E. E., PIEDRA, P. A., HEMMING, V. G. & ET AL. (1994) Comparison of antibody concentrations and protective activity of respiratory syncytial virus immune globulin and conventional immune globulin. *J Infect Dis,* 169**,** 1368-73.

SIGURS, N., GUSTAFSSON, P. M., BJARNASON, R., LUNDBERG, F., SCHMIDT, S., SIGURBERGSSON, F. & KJELLMAN, B. (2005) Severe respiratory syncytial virus bronchiolitis in infancy and asthma and allergy at age 13. *Am J Respir Crit Care Med,* 171**,** 137-41.

SIMON, A., KHURANA, K., WILKESMANN, A., MULLER, A., ENGELHART, S., EXNER, M., SCHILDGEN, O., EIS-HUBINGER, A. M., GROOTHUIS, J. R. & BODE, U. (2006) Nosocomial respiratory syncytial virus infection: impact of prospective surveillance and targeted infection control. *Int J Hyg Environ Health,* 209**,** 317-24.

SIMPSON, D. A., FEENEY, S., BOYLE, C. & STITT, A. W. (2000) Retinal VEGF mRNA measured by SYBR green I fluorescence: A versatile approach to quantitative PCR. *Mol Vis,* 6**,** 178-83.

SINGLETON, R., ETCHART, N., HOU, S. & HYLAND, L. (2003) Inability to evoke a long-lasting protective immune response to respiratory syncytial virus infection in mice correlates with ineffective nasal antibody responses. *Journal of Virology,* 77**,** 11303-11311.

SPANN, K. M., TRAN, K. C., CHI, B., RABIN, R. L. & COLLINS, P. L. (2004) Suppression of the induction of alpha, beta, and lambda interferons by the NS1 and NS2 proteins of human respiratory syncytial virus in human epithelial cells and macrophages [corrected]. *J Virol,* 78**,** 4363-9.

SRIKIATKHACHORN, A., CHANG, W. & BRACIALE, T. J. (1999) Induction of Th-1 and Th-2 responses by respiratory syncytial virus attachment glycoprotein is epitope and major histocompatibility complex independent. *J Virol,* 73**,** 6590-7.

STEC, D. S., HILL, M. G., 3RD & COLLINS, P. L. (1991) Sequence analysis of the polymerase L gene of human respiratory syncytial virus and predicted phylogeny of nonsegmented negative-strand viruses. *Virology,* 183**,** 273-87.

STEVENS, W. W., FALSEY, A. R. & BRACIALE, T. J. (2008) RSV 2007: recent advances in respiratory syncytial virus research. *Viral Immunol,* 21**,** 133-40.

SULLENDER, W. M. (2000) Respiratory syncytial virus genetic and antigenic diversity. *Clin Microbiol Rev,* 13**,** 1-15, table of contents.

SULLENDER, W. M., SUN, L. & ANDERSON, L. J. (1993) Analysis of respiratory syncytial virus genetic variability with amplified cDNAs. *J Clin Microbiol,* 31**,** 1224-31.

SULLENDER, W. M. & WERTZ, G. W. (1991) Synthetic oligonucleotide probes differentiate respiratory syncytial virus subgroups in a nucleic acid hybridization assay. *J Clin Microbiol,* 29**,** 1255-7.

SWEDAN, S., MUSIYENKO, A. & BARIK, S. (2009) Respiratory syncytial virus nonstructural proteins decrease levels of multiple members of the cellular interferon pathways. *J Virol,* 83**,** 9682-93.

TAMURA, K., DUDLEY, J., NEI, M. & KUMAR, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol,* 24**,** 1596-9.

TANG, R. S., NGUYEN, N., ZHOU, H. & JIN, H. (2002) Clustered charge-to-alanine mutagenesis of human respiratory syncytial virus L polymerase generates temperature-sensitive viruses. *Virology,* 302**,** 207-16.

TECHAARPORNKUL, S., BARRETTO, N. & PEEPLES, M. E. (2001) Functional analysis of recombinant respiratory syncytial virus deletion mutants lacking the small hydrophobic and/or attachment glycoprotein gene. *J Virol,* 75**,** 6825-34.

TECHAARPORNKUL, S., COLLINS, P. L. & PEEPLES, M. E. (2002) Respiratory syncytial virus with the fusion protein as its only viral glycoprotein is less dependent on cellular glycosaminoglycans for attachment than complete virus. *Virology,* 294**,** 296-304.

TENG, M. N. & COLLINS, P. L. (2002) The central conserved cystine noose of the attachment G protein of human respiratory syncytial virus is not required for efficient viral infection in vitro or in vivo. *J Virol,* 76**,** 6164-71.

TENG, M. N., WHITEHEAD, S. S. & COLLINS, P. L. (2001) Contribution of the respiratory syncytial virus G glycoprotein and its secreted and membrane-bound forms to virus replication in vitro and in vivo. *Virology,* 289**,** 283-96.

THOMPSON, J. D., GIBSON, T. J., PLEWNIAK, F., JEANMOUGIN, F. & HIGGINS, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res,* 25**,** 4876-82.

TOMOCHIKA, K., ICHIYAMA, T., SHIMOGORI, H., SUGAHARA, K., YAMASHITA, H. & FURUKAWA, S. (2009) Clinical characteristics of respiratory syncytial virus infection-associated acute otitis media. *Pediatrics International,* 51**,** 484-487.

TRAN, T. L., CASTAGNE, N., DUBOSCLARD, V., NOINVILLE, S., KOCH, E., MOUDJOU, M., HENRY, C., BERNARD, J., YEO, R. P. & ELEOUET, J. F. (2009) The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner. *J Virol,* 83**,** 6363-74.

TRENTO, A., CASAS, I., CALDERON, A., GARCIA-GARCIA, M. L., CALVO, C., PEREZ-BRENA, P. & MELERO, J. A. (2010) Ten years of global evolution of the human respiratory syncytial virus BA genotype with a 60-nucleotide duplication in the G protein gene. *J Virol,* 84**,** 7500-12.

TRENTO, A., GALIANO, M., VIDELA, C., CARBALLAL, G., GARCIA-BARRENO, B., MELERO, J. A. & PALOMO, C. (2003) Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. *J Gen Virol,* 84**,** 3115-20.

TRENTO, A., VIEGAS, M., GALIANO, M., VIDELA, C., CARBALLAL, G., MISTCHENKO, A. S. & MELERO, J. A. (2006) Natural history of human respiratory syncytial virus inferred from phylogenetic analysis of the attachment (G) glycoprotein with a 60-nucleotide duplication. *J Virol,* 80**,** 975-84.

TRIPP, R. A. (2004) Pathogenesis of respiratory syncytial virus infection. *Viral Immunol,* 17**,** 165-81.

TRIPP, R. A., JONES, L. P., HAYNES, L. M., ZHENG, H., MURPHY, P. M. & ANDERSON, L. J. (2001) CX3C chemokine mimicry by respiratory syncytial virus G glycoprotein. *Nat Immunol,* 2**,** 732-8.

TRIPP, R. A., MOORE, D., JONES, L., SULLENDER, W., WINTER, J. & ANDERSON, L. J. (1999) Respiratory syncytial virus G and/or SH protein alters Th1 cytokines, natural killer cells, and neutrophils responding to pulmonary infection in BALB/c mice. *J Virol,* 73**,** 7099-107.

VALARCHER, J. F., FURZE, J., WYLD, S., COOK, R., CONZELMANN, K. K. & TAYLOR, G. (2003) Role of Alpha/Beta Interferons in the Attenuation and Immunocenicity of Recombinant Bovine Respiratory Syncytial Viruses Lacking NS Proteins. *Journal of Virology,* 77**,** 8426-8439.

VALOSKY, J., HISHIKI, H., ZAOUTIS, T. E. & COFFIN, S. E. (2005) Induction of mucosal B-cell memory by intranasal immunization of mice with respiratory syncytial virus. *Clinical Diagnostic Laboratory Immunology,* 12**,** 171-179.

VAN DRUNEN LITTEL-VAN DEN HURK, S., MAPLETOFT, J. W., ARSIC, N. & KOVACS-NOLAN, J. (2007) Immunopathology of RSV infection: prospects for developing vaccines without this complication. *Rev Med Virol,* 17**,** 5-34.

VENTER, M., COLLINSON, M. & SCHOUB, B. D. (2002a) Molecular epidemiological analysis of community circulating respiratory syncytial virus in rural South Africa: Comparison of viruses and genotypes responsible for different disease manifestations. *J Med Virol,* 68**,** 452-61.

VENTER, M., MADHI, S. A., TIEMESSEN, C. T. & SCHOUB, B. D. (2001) Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. *J Gen Virol,* 82**,** 2117-24.

VENTER, M., TIEMESSEN, C., COLLINSON, M. & SCHOUB, B. (2002b) Molecular epidemiological analysis of community circulating respiratory syncytial virus in rural South Africa: comparison of viruses and genotypes causing different disease manifestations. *J. Med. Virol,* 68**,** 452-61.

VENTRE, K. & RANDOLPH, A. (2004) Ribavirin for respiratory syncytial virus infection of the lower respiratory tract in infants and young children. *Cochrane Database Syst Rev***,** CD000181.

VIDAL, S. & KALAKOFSKY, D. (1989) Modified model for the switch from Sendai virus transcription to replication. *Journal of Virology,* 63**,** 1951-1958.

VIEGAS, M. & MISTCHENKO, A. S. (2005) Molecular epidemiology of human respiratory syncytial virus subgroup A over a six-year period (1999-2004) in Argentina. *J Med Virol,* 77**,** 302-10.

VIGNUZZI, M., STONE, J. K., ARNOLD, J. J., CAMERON, C. E. & ANDINO, R. (2006) Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature,* 439**,** 344-8.

VISSER, A., DELPORT, S. & VENTER, M. (2008) Molecular epidemiological analysis of a nosocomial outbreak of respiratory syncytial virus associated pneumonia in a kangaroo mother care unit in South Africa. *J Med Virol,* 80**,** 724-32.

VOGEL, A. M., LENNON, D. R., BROADBENT, R., BYRNES, C. A., GRIMWOOD, K., MILDENHALL, L., RICHARDSON, V. & ROWLEY, S. (2002) Palivizumab prophylaxis of respiratory syncytial virus infection in high-risk infants. *Journal of Paediatrics and Child Health,* 38**,** 550-554.

WALSH, E. E. & HRUSKA, J. (1983) Monoclonal antibodies to respiratory syncytial virus proteins: identification of the fusion protein. *J Virol,* 47**,** 171-7.

WALSH, E. E., MCCONNOCHIE, K. M., LONG, C. E. & HALL, C. B. (1997) Severity of respiratory syncytial virus infection is related to virus strain. *J Infect Dis,* 175**,** 814-20.

WANG, X., LI, M., ZHENG, H., MUSTER, T., PALESE, P., BEG, A. A. & GARCIA-SASTRE, A. (2000) Influenza A Virus NS1 Protein Prevents Activation of NF-$_{\kappa}$B ανδ Ινδυχτιον οφ Αλπηα/Βετα Ιντερφερον. *Journal of Virology,* 74**,** 11566-11573.

WEBER, F., BRIDGEN, A., FAZAKERLEY, K., STREITENFELD, H., KESSLER, N., RANDALL, R. E. & ELLIOT, R. M. (2002) Bunyamwera Bunyavirus Nonstructural Protein NSs Counteracts the Induction of Alpha/Beta Interferon. *Journal of Virology,* 76**,** 7949-7955.

WELLIVER, R. C. (2003) Review of epidemiology and clinical risk factors for severe respiratory syncytial virus (RSV) infection. *J Pediatr,* 143**,** S112-7.

WERTZ, G. W., COLLINS, P. L., HUANG, Y., GRUBER, C., LEVINE, S. & BALL, L. A. (1985) Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein. *Proc Natl Acad Sci U S A,* 82**,** 4075-9.

WERTZ, G. W. & MOUDY, R. M. (2004) Antigenic and genetic variation in human respiratory syncytial virus. *Pediatr Infect Dis J,* 23**,** S19-24.

WHELAN, J. A., RUSSELL, N. B. & WHELAN, M. A. (2003) A method for the absolute quantification of cDNA using real-time PCR. *Journal of Immunological Methods,* 278**,** 261-269.

WHO (2009) Initiative for Vaccine Research (IVR). Repiratory syncytial virus and parainfluenza viruses. Disease burden. www.who.int/vaccine_research/diseases/ari/en/indexz.html.

WOELK, C. H. & HOLMES, E. C. (2001) Variable immune-driven natural selection in the attachment (G) glycoprotein of respiratory syncytial virus (RSV). *J Mol Evol,* 52**,** 182-92.

WONG, M. L. & MEDRANO, J. F. (2005) Real-time PCR for mRNA quantification. *Biotechniques,* 39**,** 75-85.

WRIGHT, P. F., KARRON, R. A., BELSHE, R. B., THOMPSON, J., CROWE, J. E., JR., BOYCE, T. G., HALBURNT, L. L., REED, G. W., WHITEHEAD, S. S., ANDERSON, E. L., WITTEK, A. E., CASEY, R., EICHELBERGER, M., THUMAR, B., RANDOLPH, V. B., UDEM, S. A., CHANOCK, R. M. & MURPHY, B. R. (2000) Evaluation of a live, cold-passaged, temperature-sensitive, respiratory syncytial virus vaccine candidate in infancy. *J Infect Dis,* 182**,** 1331-42.

ZHANG, W. & TRIPP, R. A. (2008) RNA Interference Inhibits Respiratory Syncytial Virus Replication and Disease Pathogenesis without Inhibiting Priming of the Memory Immune Response. *Journal of Virology,* 82**,** 12221-12231.

ZLATEVA, K., LEMEY, P., MOES, E., VANDAMME, A. & RANST, M. V. (2005a) Genetic variability and molecular evolution of the human Respiratory Syncytial Virus subgroup B attachment G protein. *J Viro,* 79**,** 9157-67.

ZLATEVA, K. T., LEMEY, P., MOES, E., VANDAMME, A. M. & VAN RANST, M. (2005b) Genetic variability and molecular evolution of the human respiratory syncytial virus subgroup B attachment G protein. *J Virol,* 79**,** 9157-67.

ZLATEVA, K. T., LEMEY, P., VANDAMME, A. M. & VAN RANST, M. (2004) Molecular evolution and circulation patterns of human respiratory syncytial virus subgroup a: positively selected sites in the attachment g glycoprotein. *J Virol,* 78**,** 4675-83.

ZLATEVA, K. T., VIJGEN, L., DEKEERSMAEKER, N., NARANJO, C. & VAN RANST, M. (2007) Subgroup prevalence and genotype circulation patterns of human respiratory syncytial virus in Belgium during ten successive epidemic seasons. *J Clin Microbiol,* 45**,** 3022-30.

APPENDIX A

SEQUENCES WITH ACCESSION NUMBERS USED IN THIS STUDY

SPECIMEN NAME	ACCESSION NUMBER	SPECIMEN NAME	ACCESSION NUMBER
PARTIAL G PROTEIN GENE SEQUENCES			
SA462965OP06	HQ711628	SA1073000OP07	HQ711734
SA972538OP07	HQ711629	SA1041485H07	HQ711735
SA975346OP07	HQ711630	SA1017503H07	HQ711736
SA850525H07	HQ711631	SA972744OP07	HQ711737
SA1056896H07	HQ711632	SA419352OP06	HQ711738
SA463070H06	HQ711633	SA483434OP06	HQ711739
SA967456OP07	HQ711634	SA505872OP06	HQ711740
SAQ11003826OP07	HQ711635	SA510933OP06	HQ711741
SA1008168H07	HQ711636	SA522892OP06	HQ711742
SA1020317OP07	HQ711637	SA530589OP06A	HQ711743
SAQ21003826OP07	HQ711638	SA456997H06	HQ711744
SA1000391OP07	HQ711639	SA541471H06	HQ711745
SA1315234H08	HQ711640	SA440007H06	HQ711746
SA1425925H08	HQ711641	SA463036H06	HQ711747
SA953470H07	HQ711642	SA463045H06	HQ711748
SA1372749H08	HQ711643	SA441801OP06	HQ711749
SAQ11040558H07	HQ711644	SA451655OP06	HQ711750
SAQ21040558H07	HQ711645	SA386822H06	HQ711751
SA535057H06	HQ711646	SA430263H06	HQ711752
SA1421083OP08	HQ711647	SA441802H06	HQ711753
SA1044903H07	HQ711648	SA447937H06	HQ711754
SA1377350H08	HQ711649	SA468148H06	HQ711755
SA1415396OP08	HQ711650	SA468917H06	HQ711756
SA1380743OP08	HQ711651	SA470433H06	HQ711757
SA975113OP07	HQ711652	SA477847H06	HQ711758
SA913306OP07	HQ711653	SA576338H06	HQ711759
SA1044904OP07	HQ711654	SA584285H06	HQ711760
SAQ11031251OP07	HQ711655	SA452947H06	HQ711761
SAQ21031251OP07	HQ711656	SA954715OP07A	HQ711762
SA1045506OP07	HQ711657	SA439979OP06	HQ711763
SA1031281H07	HQ711658	SA405170H06	HQ711764
SA1040552H07	HQ711659	SA1395352OP08	HQ711765
SA1013624H07	HQ711660	SA1484890H08	HQ711766
SA955705OP07	HQ711661	SA1491542H08	HQ711767
SA1576527OP08	HQ711662	SA1536976H08	HQ711768
SA997237H07	HQ711663	SA1562390H08	HQ711769
SA983883OP07	HQ711664	SA5201630OP06	HQ711770
SA998662H07	HQ711665	SA533624H06	HQ711771
SA1095437H07	HQ711666	SA562677H06	HQ711772
SA1376979H08	HQ711667	SA576319H06	HQ711773

SAQ1976837H07	HQ711668	SA447060H06	HQ711774
SAQ2976827H07	HQ711669	SA929792H07	HQ711775
SA1911792H09	HQ711670	SA512552H06	HQ711776
SA1875019H09	HQ711671	SA645573H06	HQ711777
SA1378978OP08	HQ711672	SA445197H06	HQ711778
SA1916598H09	HQ711673	SA1122432OP07	HQ711779
SA1404700H08	HQ711674	SA1601297H08	HQ711780
SA1405122H08	HQ711675	SA481582OP06	HQ711781
SA1376978H08	HQ711676	SA1897937H09	HQ711782
SA1597150H08	HQ711677	SA459004OP06	HQ711783
SA1377345H08	HQ711678	SA1904180OP09	HQ711784
SA1405123OP08	HQ711679	SA622525OP06	HQ711785
SA998657H07	HQ711680	SA1906931OP09	HQ711786
SA984954H07	HQ711681	SA1898386H09	HQ711787
SA1040547OP07	HQ711682	SA475995H06	HQ711788
SA996747H07	HQ711683	SA957529OP07	HQ711789
SA1037536H07	HQ711684	SA430310OP06	HQ711790
SA1005653H07	HQ711685	SA432203OP06	HQ711791
SA988853OP07	HQ711686	SA451653OP06	HQ711792
SA1473284H08	HQ711687	SA523250OP06	HQ711793
SA1389148H08	HQ711688	SA430319H06	HQ711794
SA1022133OP07	HQ711689	SA451774H06	HQ711795
SA1010538H07	HQ711690	SA435745OP06	HQ711796
SA994216OP07	HQ711691	SA490490H06	HQ711797
SA974731H07	HQ711692	SA983753OP07B	HQ711798
SA1349873OP08	HQ711693	SA1138929H07	HQ711799
SA1540738H08	HQ711694	SA472200OP06	HQ711800
SAQ1983753OP07	HQ711695	SA539486H06	HQ711801
SAQ2983753OP07	HQ711696	SA433941OP06	HQ711802
SA445984H06	HQ711697	SA555134H06	HQ711803
SA1031469H07	HQ711698	SA466313H06B	HQ711804
SA967452H07	HQ711699	SA539456H06B	HQ711805
SA1893782H09	HQ711700	SA954715OP07B	HQ711806
SA1013623H07	HQ711701	SA921110H07	HQ711807
SA980164OP07	HQ711702	SA404376OP06	HQ711808
SA1046284H07	HQ711703	SA1961261H09	HQ711809
SA1104553H07	HQ711704	SA1903163H09	HQ711810
SA1022124OP07	HQ711705	SA1897566H09	HQ711811
SA953224OP07	HQ711706	SA1946046H09	HQ711812
SA930011H07	HQ711707	SA1914306H09	HQ711813
SA1377240H08	HQ711708	SA483388OP06	HQ711814
SA974725H07	HQ711709	SA574601H06	HQ711815
SA9998477H07	HQ711710	SA1507096OP08	HQ711816
SA904146H07	HQ711711	SA1913535H09	HQ711817
SA1008161H07	HQ711712	SA1559601OP08	HQ711818
SA971424H07	HQ711713	SA1856310OP09	HQ711819
SA1040546OP07	HQ711714	SA581105H06	HQ711820
SA1115939H07	HQ711715	SA1952519H09	HQ711821
SA1389133OP08	HQ711716	SA1903864H09	HQ711822
SA1208873OP07	HQ711717	SA1966323H09	HQ711823
SA1574707H08	HQ711718	SA1945943H09	HQ711824

SA963960OP07	HQ711719	SA1856312OP09	HQ711825
SA403922OP06	HQ711720	SA1954125H09	HQ711826
SA470498OP06	HQ711721	SA1901413H09	HQ711827
SA972413H07	HQ711722	SA1023607OP07	HQ711828
SA947759H07	HQ711723	SA1510114OP08	HQ711829
SA432598OP06	HQ711724	SA1911347OP09	HQ711830
SA524750H06	HQ711725	SA1037074H07	HQ711831
SA524904H06	HQ711726	SA1864925OP09	HQ711832
SA433216H06	HQ711727	SA1960816H09	HQ711833
SA445585H06	HQ711728	SA466303OP06	HQ711834
SA961333OP07	HQ711729	SA530589OP06B	HQ711835
SA445179H06	HQ711730	SA1918530OP09	HQ711836
SA969294OP07	HQ711731	SA453610H06	HQ711837
SA972743H07	HQ711732	SA512554H06	HQ711838
SA984829OP07	HQ711733	SA417463H06	HQ711839
FULL G PROTEIN GENE SEQUENCES			
SA98D1656	JF704213	SA957529Pt07	JF704224
SA97D1107	JF704214	SA983753Pt07	JF704225
AgJ15_99	JF704215	SA1023607Pt07	JF704226
SA98V192	JF704216	SA1138929Pt08	JF704227
SA98D796	JF704217	SA1510114Pt08	JF704228
SA99Vr800	JF704218	SA1904180Pt09	JF704229
SA581105Pt06	JF704219	SA1906931Pt09	JF704230
SA466303Pt06	JF704220	SA1918530Pt09	JF704231
SA486540Pt06	JF704221	SA1917319Pt09	JF704232
SA539456PT06	JF704222	SA1954125Pt09	JF704233
SA555134PT06	JF704223	SA1913535Pt09	JF704234
DELETION MUTANT G PROTEIN GENE SEQUENCES			
SA374801Pt06a	HQ711840		
SA374801Pt06b	HQ711841		
SA367859Pt06	HQ711842		
FULL NS1 AND NS2 PROTEIN GENE SEQUENCES			
SA1003826OP07	JQ359529	SA954715OP07	JQ359569
SA1022124OP 07	JQ359530	SA955705OP07	JQ359570
SA1056896H07	JQ359531	SA961333OP07	JQ359571
SA386822H06	JQ359532	SA972538OP07	JQ359572
SA403922Q1OP06	JQ359533	SA975119OP07	JQ359573
SA403922Q2OP06	JQ359534	SA988853OP07	JQ359574
SA432598OP06	JQ359535	SA468917H06	JQ359575
SA433216H06	JQ359536	SA404376OP06	JQ359576
SA445179H06	JQ359537	SA417463H06	JQ359577
SA445197Q1H06	JQ359538	SA430319H06	JQ359578
SA445197Q2H06	JQ359539	SA432203OP06	JQ359579
SA445585Q1H06	JQ359540	SA435745OP06	JQ359580
SA445585Q2H06	JQ359541	SA451653OP06	JQ359581
SA445585Q3H06	JQ359542	SA451774H06	JQ359582
SA445585Q4H06	JQ359543	SA453610H06	JQ359583
SA447060H06	JQ359544	SA459004Q1OP06	JQ359584
SA451655OP06	JQ359545	SA459004Q2OP06	JQ359585
SA452947H06	JQ359546	SA466303OP06	JQ359586
SA456997H06	JQ359547	SA472200OP06	JQ359587

SA462965Q1OP06	JQ359548	SA475995H06	JQ359588
SA462965Q2OP06	JQ359549	SA483388Q1OP06	JQ359589
SA462965Q3OP06	JQ359550	SA483388Q2OP06	JQ359590
SA462965Q4OP06	JQ359551	SA490490H06	JQ359591
SA468148H06	JQ359552	SA512554H06	JQ359592
SA470433H06	JQ359553	SA523250OP06	JQ359593
SA483434OP06	JQ359554	SA530589Q1OP06	JQ359594
SA505872OP06	JQ359555	SA530589Q2OP06	JQ359595
SA512552Q1H06	JQ359556	SA530589Q3OP06	JQ359596
SA512552Q2H06	JQ359557	SA530589Q4OP06	JQ359597
SA524750H06	JQ359558	SA530589Q5OP06	JQ359598
SA562677H06	JQ359559	SA530589Q6OP06	JQ359599
SA584285H06	JQ359560	SA539456Q1H06	JQ359600
SA850525H07	JQ359561	SA539456Q2B-H06	JQ359601
SA904146H07	JQ359562	SA539486Q1H06	JQ359602
SA913306OP07	JQ359563	SA539486Q2H06	JQ359603
SA929792Q1H07	JQ359564	SA581105H06	JQ359604
SA929792Q2H07	JQ359565	SA622525Q1OP06	JQ359605
SA929792Q3H07	JQ359566	SA622525Q2OP06	JQ359606
SA930011H07	JQ359567	SA622656Q3OP06	JQ359607
SA953224OP07	JQ359568	SA622525Q4OP06	JQ359608

APPENDIX B

CREATING OF STANDARD CURVES AND OPTIMIZING PCR EFFICIENCY

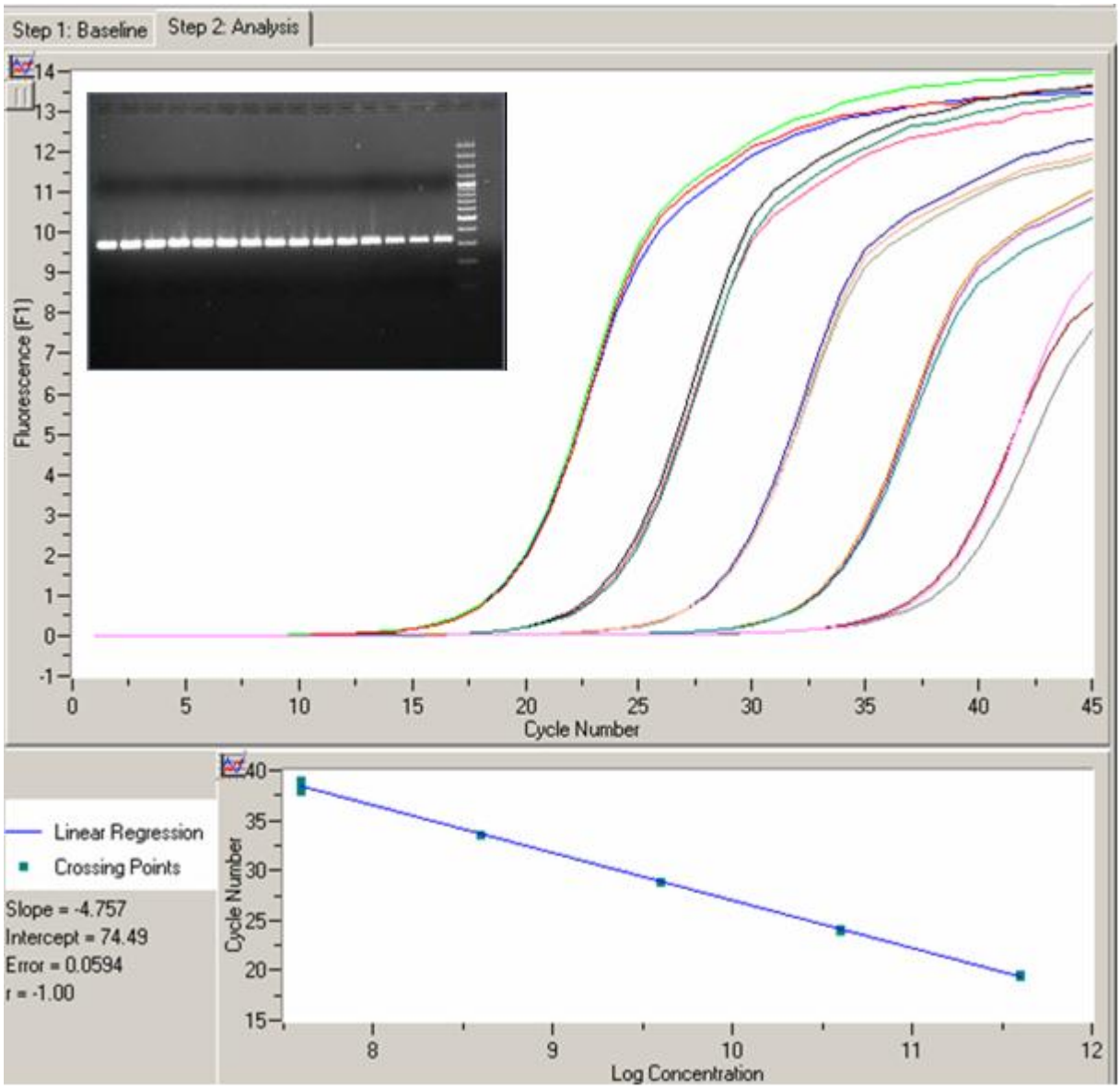


FIGURE A Standard curve with amplification curves and agarose gel electrophoresis for the antigenome using SYBR Green I dye. All dilutions were performed in triplicate.

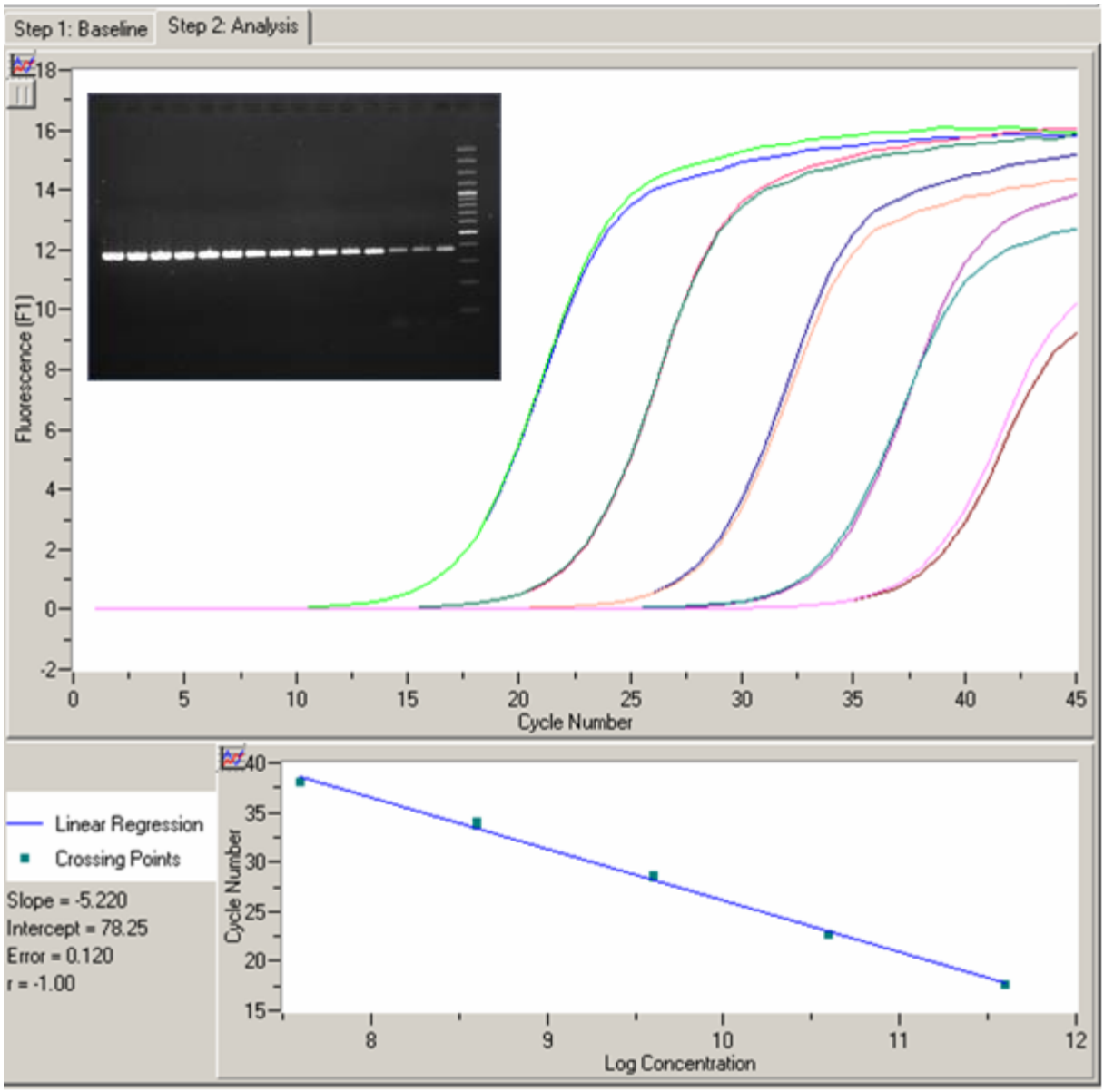


FIGURE B Standard curve with amplification curves and agarose gel electrophoresis for the NS1 region using SYBR Green I dye. All dilutions were performed in triplicate.

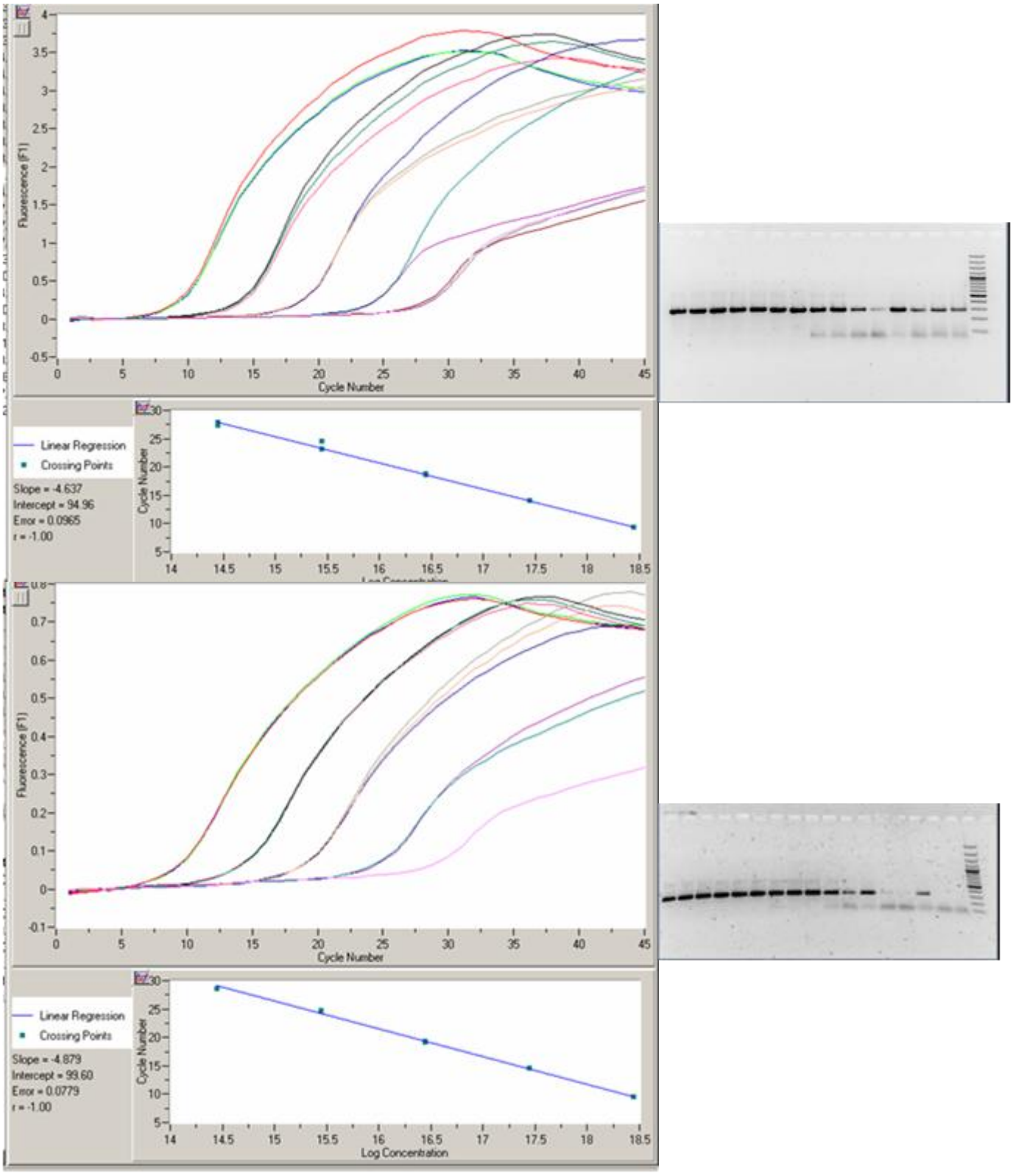


FIGURE C Standard curve with amplification curves and agarose gel electrophoresis for the NS2 region amplified with the Lightcyler TaqMan Master kit for **A)** subtype A **B)** subtype B control specimens. All dilutions were performed in triplicate.