Genetic and Molecular Characterisation of Non-Tissue Culture Adapted Influenza B Viruses detected in Singapore between 2004-2009

MUHAMMAD RAIHAN JUMAT
SCHOOL OF BIOLOGICAL SCIENCES
2013
Genetic and Molecular Characterisation of Non-Tissue Culture Adapted Influenza B Viruses detected in Singapore between 2004-2009

MUHAMMAD RAIHAN JUMAT

SCHOOL OF BIOLOGICAL SCIENCES

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (Ph.D)

2013
Acknowledgements

I would like to thank my supervisor, Associate Professor Richard J Sugrue for his guidance and mentorship over the period of my candidature. I would also like to thank our collaborator, Assistant Professor Tan Boon Huan, for her relentless support and advice for this project.

I would like to thank the members of my Teaching Advisory Committee, Associate Professor Tan Suet Mein, Associate Professor Julien Lescar and Assistant Professor Tang Kai for reviewing my work and for their invaluable inputs.

My heartfelt gratitude goes out to my mentors and lab mates, (Liat Hui, Laxmi, Patricia, Dawn, Jasper, Chen Hui, Fu Yi, Debbie, Elizabeth, Puisan, Peijun, Lifang, Shirley, and Richard) for teaching me the ropes to my first scientific endeavor and for providing a conducive environment for research.

I would also like to thank Professor Thorsten Wolff for the gift of anti-influenza B NS1 antibody, without which most of the confirmatory experiments could not be carried out. I would like to thank the hosts of This Week in Virology (TWiV) podcast for the lively discussions about the latest in this interesting field.

I would like to thank Nanyang Technological University for providing me with a research scholarship and Muslimin Trust Fund Association (MTFA) for bursary awards received throughout my undergraduate as well as post graduate studies.

I thank Allah (swt), for his Benevolence, especially in granting me a family which had never ceased in supporting me especially through the last five years, despite my negligence towards them. I thank my siblings (Muhsin, Azizah, Salwaa and Nooreen) for giving me a reason to laugh even when times are hard. I thank my in-laws (Mak and Ayah) for their support and encouragement.

My wife (Dahlia), to whom I owe the deepest of gratitude, I am extremely thankful to her accepting me whilst still a student and for undergoing the trials of this candidature together, providing me strength and for reminding me of the greater purpose. I am extremely thankful to my son (Hasan) for providing me reason to get up (often, several times) in the mornings.

Lastly I would like to thank my parents (Mummy and Abah), for instilling in me the love of science and the teaching me the importance of patience and diligence. Words cannot describe how much I am eternally grateful to you. This work, albeit humble in its nature, is dedicated to the two of you.
Chapter 1: Introduction

1.1 Classification and Nomenclature

1.2 Symptoms and Epidemiology of Influenza B

1.3 Influenza B Tropism

1.4 Influenza B Lineages

1.5 Influenza B Evolution

1.6 Influenza B Vaccination

1.7 Global Prevalence of Influenza B

1.7.1 Influenza in Singapore

1.7.1.1 Influenza B Prevalence in Singapore

1.7.1.2 National Service in Singapore

1.8 Influenza Virus B Structure and Morphology

1.9 Influenza B Genome Structure

1.10 Influenza B Proteins

1.10.1 Influenza B Hemagglutinin (HA) Protein

1.10.2 Influenza B Neuraminidase (NA) Protein

1.10.3 Influenza B NB Protein

1.10.4 Influenza B BM2 Protein

1.10.5 Influenza B Non Structural Proteins 1 (NS1)

1.10.6 Influenza B Nuclear Export Protein (NEP/NS2)

1.11 Study Outline & Objectives

Chapter 2: Materials & Methods

2.1 General Reagents

2.1.1 Cells

2.1.2 Viruses

2.1.3 Tissue Culture Reagents

2.1.4 Antibodies

2.1.5 Immunofluorescence and Immunostaining Reagents

2.1.6 Commercially Available Kits
2.1.7 Tissue Culture Propagation & Infection Media and Reagents ..................44
2.1.8 Bacterial Culture Media and Solution ..............................................45
2.1.9 DNA Analysis..................................................................................45
2.1.10 Protein Analysis by SDS-PAGE gel ..............................................46
2.1.11 Western Blotting Reagents ............................................................46
2.1.12 Coomassie Blue Staining Reagents ..............................................46
2.1.13 SYPRO Ruby Staining Reagents ....................................................47
2.1.14 Reagents for Immunoprecipitation ...............................................47
2.2 Clinical Specimen Sequencing ............................................................47
2.2.1 Reference Strain ..............................................................................47
2.2.2 Clinical Specimens .........................................................................47
2.2.3 Tube Cultures of Viral isolates .......................................................47
2.2.4 Shell Vial Assays .............................................................................48
2.2.5 Egg Passaging ..................................................................................48
2.2.6 Diagnostic Immunofluorescent Assays ...........................................48
2.2.7 Extraction of Influenza B Virus vRNA ............................................49
2.2.8 Sequencing Primers .........................................................................49
2.2.9 Real Time PCR detection of Influenza B Virus ...............................50
2.2.10 cDNA Synthesis ............................................................................50
2.2.11 Hemagglutinin (HA) Gene Sequencing Strategy ................................50
2.2.12 Neuraminidase (NA) Gene Sequencing Strategy ...........................51
2.2.13 Non Structural (NS) Gene Sequencing Strategy ..............................51
2.2.14 Gel Electrophoresis and PCR Product Extraction .........................52
2.2.15 Sequencing and Bioinformatics ......................................................52
2.3 Virus Infection in Tissue Culture .........................................................53
2.3.1 Mammalian Cell Maintenance .........................................................53
2.3.2 Harvesting and Culturing of Chicken Embryonic Fibroblasts (CEF) ....53
2.3.3 Virus Infection on Cell Culture .......................................................53
2.3.4 Conventional Plaque Assay .............................................................54
2.3.5 End Point Assay ...............................................................................54
2.3.6 Conventional Coverslip Immunofluorescent Assay ..........................55
2.3.7 Immunofluorescent Microplate Assay .............................................55
2.3.8 Drug Treatment ..............................................................................56
2.4 Molecular Cloning of NS1, BM2 and NB genes .................................................. 56
  2.4.1 Restriction Endonuclease Selection ............................................................... 56
  2.4.2 PCR amplification of Genes for Cloning ....................................................... 56
  2.4.3 Ligation into pDRIVE Cloning Vector .......................................................... 57
  2.4.4 Bacterial Transformation of pDRIVE Ligation .............................................. 57
  2.4.5 Restriction Endonuclease Treatment ............................................................ 60
  2.4.6 Ligation into pCAGGS vector ........................................................................ 60
  2.4.7 Bacterial Transformation of Insert-pCAGGS .................................................. 60
  2.4.8 Site-Directed Mutagenesis .............................................................................. 60
  2.4.9 Plasmid Transfection into Mammalian Cells .................................................. 63
  2.4.10 Endoglycosidase Treatment ......................................................................... 63
  2.4.11 Chemical cross linking of proteins ............................................................... 64
  2.4.12 Immunoprecipitation .................................................................................... 64
  2.5 Protein Analysis by SDS-PAGE ......................................................................... 64
    2.5.1 Western Blotting ........................................................................................... 65
    2.5.2 Coomassie Brilliant Blue Staining ................................................................. 65
    2.5.3 SYPRO Ruby Staining .................................................................................. 65

Chapter 3: Cultivation and passage of clinical specimens and B/Lee/40............ 67
  3.1 Introduction ........................................................................................................ 67
  3.2 Results ............................................................................................................... 69
    3.2.1 Cultivation of Clinical Specimens ............................................................... 69
    3.2.2 B/Lee/40 Passage and Cultivation ............................................................... 72
      3.2.2.1 B/Lee/40 Infection Characterisation by Immunofluorescence Assay (IFA) ................................................................. 72
      3.2.2.2 Purification of B/Lee/40 .......................................................................... 74
      3.2.2.3 Transmission Electron Microscopy of Purified B/Lee/40 ..................... 76
    3.2.3 Growth Kinetics of B/Lee/40 ...................................................................... 76
    3.2.4 Immunofluorescent Microplate Assay ......................................................... 77
  3.3 Discussion ......................................................................................................... 81

Chapter 4: Genetic and phylogenetic analysis of gene segment 4 encoding for the hemagglutinin (HA) protein................................................................. 85
  4.1 Introduction ...................................................................................................... 85
  4.2 Results ............................................................................................................. 86
    4.2.1 Sequencing of gene segment 4 .................................................................... 86
4.2.2 Nucleotide sequence and phylogenetic analysis of gene segment 4 ........ 87
4.2.3 Amino acid sequence analysis of gene segment 4, HA0 precursor protein ... 88
4.3 Discussion ........................................................................................................... 93

Chapter 5: Genetic and phylogenetic analysis of gene segment 6 encoding for the neuraminidase (NA) and NB proteins ................................................................. 96
5.1 Introduction ....................................................................................................... 96
5.2 Results ............................................................................................................ 98
  5.2.1 Sequencing of gene segment 6 ................................................................. 98
  5.2.2 Nucleotide sequence and phylogenetic analysis of gene segment 6 ...... 99
  5.2.3 Amino acid sequence analysis ................................................................. 102
    5.2.3.1 Amino acid sequence analysis of the NA protein ....................... 102
    5.2.3.2 Amino acid sequence analysis of the NB protein ....................... 103
  5.2.4 Molecular characterisation of the NB protein isolated from non-tissue culture adapted influenza B clinical specimens ............................................. 104
    5.2.4.1 Molecular Cloning of the NB protein and Expression by Plasmid DNA Transfection ................................................................. 104
    5.2.4.2 Endoglycosidase treatment of the NB protein ......................... 108
    5.2.4.3 Cross-linking of NB protein ......................................................... 109
5.3 Discussion ........................................................................................................ 111

Chapter 6: Genetic and phylogenetic analysis of gene segment 8 encoding for the Non-Structural 1 Protein (NS1) and Nuclear Export Protein (NEP) ............ 117
6.1 Introduction ....................................................................................................... 117
6.2 Results ............................................................................................................ 121
  6.2.1 Sequencing of gene segment 8 ................................................................. 121
  6.2.2 Nucleotide sequence and phylogenetic analysis of gene segment 8 ...... 121
  6.2.3 Amino Acid sequence analysis ............................................................... 122
    6.2.3.1 Amino acid sequence analysis of the NS1 protein .................. 125
    6.2.3.2 Amino acid sequence analysis of the NEP .............................. 127
  6.2.4 Molecular Characterisation of the NS1 protein from non-tissue culture adapted clinical specimens ................................................................. 128
    6.2.4.1 Cellular localisation of NS1 protein during influenza B infection ... 128
    6.2.4.2 Molecular Cloning of NS1 and Expression by Plasmid DNA Transfection .......................................................... 133
      6.2.4.2.1 Comparison of Influenza A and Influenza B NS1 localisation 138
6.2.4.2 Expression of Individual Influenza B NS1 Domains ...........................................139
6.2.4.3 Effects of Mutations on Influenza B NS1 Protein Localisation.................................142
6.2.4.3 SDS-PAGE analysis of influenza B NS1 ...............................................................146
   6.2.4.3.1 SDS-PAGE analysis of NS1 in B/Lee/40 infection and Lee-NS1 transfection ................146
   6.2.4.3.2 Time course expression of Influenza B NS1 protein .........................................148
   6.2.4.3.3 Expression of individual domains of B/Lee/40 NS1 protein .................................149
   6.2.4.3.4 Expression of influenza B NS1 of the representative clinical specimens ..........151
   6.2.4.3.5 Cytoplasmic and nuclear fractionation of transfected/infected HEK 293T cells .................................................................153
   6.2.4.3.6 Mutational analysis of influenza B NS1 protein ................................................155
   6.2.4.3.7 Effect of protease inhibition on the expression of Lee-NS1 .................................160
6.3 Discussion ..................................................................................................................162

Chapter 7: Molecular characterisation of influenza B BM2 protein cloned from non-tissue culture adapted clinical specimens .........................................................175
7.1 Introduction .................................................................................................................175
7.2 Results .........................................................................................................................176
   7.2.1 Molecular Cloning of the BM2 protein and Expression by Plasmid DNA Transfection .................................................................176
   7.2.2 Chemical cross-linking of the Lee-BM2 protein .....................................................180
7.3 Discussion ....................................................................................................................182

Chapter 8: Conclusions and Future Direction ..................................................................186
Protein Alignments ...........................................................................................................196
References .......................................................................................................................219
List of Figures

Fig. 1.1 Genetic reassortment between 2 different influenza strains 6
Fig. 1.2 Number of global influenza specimens collected in 2012 (GISRS) 10
Fig. 1.3 Percentage of influenza virus specimen collected in 2012 (GISRS) 11
Fig. 1.4 Number of SEA influenza specimens collected in 2012 (GISRS) 12
Fig. 1.5 Distribution of influenza viruses in Singapore from June 2009- July 2012 (CDC, MOH) 14
Fig. 1.6 Weekly distribution of influenza A and B viruses in Singapore form May 2009 to Aug 2010 15
Fig. 1.7 Influenza distribution in Singapore, Mexico and USA 2009-2010 (GISRS) 16
Fig. 1.8 Electron microscopy of influenza A and B viruses 18
Fig. 1.9 Schematic diagram of influenza B structure 19
Fig. 1.10 Radiolabelled genomes of influenza A and B 20
Fig. 1.11 Gene segments of influenza B viruses 21
Fig. 1.12 General schematic of influenza virus replication 23
Fig. 1.13 Alignment of B/Lee/40 HA against A/PR/8/34 HA 24
Fig. 1.14 HA, pre and post fusion 26
Fig. 1.15 Fusion of viral and host cell membrane 28
Fig. 1.16 Schematic diagram of the mechanism of NAI inhibition 30
Fig. 1.17 The first 350 nucleotides of gene segment 6 31
Fig. 1.18 Genome structure of segment 7 of influenza A and B viruses 32
Fig. 1.19 Structure of influenza B BM2 protein 33
Fig. 1.20 Amino acid alignment of the TM domain of the M2 protein form influenza A and B viruses 34
Fig. 1.21 Schematic representation of influenza M2 proton channel 34
Fig. 1.22 Schematic representation of influenza B NS1 protein 35
Fig. 1.23 Localisation of NS1 protein in nuclear speckle domains 36
Fig. 1.24 Gene arrangements of NS1 and NEP in gene segment 7 of influenza A and B viruses 37
Fig. 2.1 Schematic of strategy used to sequence HA/NA gene segments 49
Fig. 2.2 Schematic of strategy used to sequence NS gene segment 50
Fig. 3.1 Diagnostic IFA of clinical specimens 68
Fig. 3.2 MDCK cells infected with supernatants harvested from tube cultures 69
Fig. 3.3 Shell vial infection assays 70
Fig. 3.4 Time course infection of egg-passaged B/Lee/40 on human alveolar A549 cells 70
Fig. 3.5 B/Lee/40 virus purification through a 30%-60% sucrose gradient 71
Fig. 3.6 30% and 40% virus bands from the sucrose gradient analysed on 12% SDS PAGE
Fig. 3.7 TEM image of purified B/Lee/40 virus from the 40% virus band
Fig. 3.8 CPE caused by B/Lee/40 infection on MDCK cells
Fig. 3.9 Titration of B/Lee/40 virus
Fig. 3.10 Titration of A/WSN/33 (H1N1) virus
Fig. 4.1 Schematic representation of the influenza B HA protein
Fig. 4.2 Sizes of the overlapping HA ORF fragments
Fig. 4.3 HA phylogenetic tree
Fig. 5.1 Schematic representation of the influenza B NA and NB proteins
Fig. 5.2 Sizes of overlapping NA ORF fragments of B/Lee/40
Fig. 5.3 NA phylogenetic tree
Fig. 5.4 Cloning of NB gene
Fig. 5.5 SDS-PAGE analysis of NB
Fig. 5.6 Time course expression of 70-NB and 147-NB in HEK 293T cells
Fig. 5.7 MDCK cells were transfected with NB
Fig. 5.8 A549 cells were transfected with NB
Fig. 5.9 Endoglycosidation of HEK 293T cells transfected with NB
Fig. 5.10 Cross linking of Lee-NB
Fig. 6.1 NoLS of influenza A NS1
Fig. 6.2 Speckled localisation of influenza B NS1 in early infection
Fig. 6.3 Schematic representation of NS1 and NEP proteins
Fig. 6.4 Size of the coding sequence of gene segment 8 of B/Lee/40
Fig. 6.5 NS phylogenetic tree
Fig. 6.6 Time course infection of B/Lee/40 on A549 cells (NS1 and NP staining)
Fig. 6.7 Time course infection of B/Lee/40 on A549 cells (NS1 and Lamin A/C staining)
Fig. 6.8 Time course infection of B/Lee/40 on A549 cells (NS1 and NPC staining)
Fig. 6.9 Time course infection of B/Lee/40 on A549 cells (NS1 and NPM staining)
Fig. 6.10 Time course infection of B/Lee/40 on A549 cells (NS1 and Coillin staining)
Fig. 6.11 Time course infection of B/Lee/40 on A549 cells (NS1 and SC35 staining)
Fig. 6.12 Time course infection of B/Lee/40 on MDCK cells (NS1 and Lamin A/C staining)
Fig. 6.13 Time course infection of B/Lee/40 on MDCK cells (NS1 and NPC staining)
Fig. 6.14 Time course infection of B/Lee/40 on MDCK cells (NS1 and SC35 staining)
Fig. 6.15  Time course infection of B/Lee/40 on MDCK cells (NS1 and NPM staining)  129
Fig. 6.16  Cloning of NS1 gene  131
Fig. 6.17  Transfection of NS1-pCAGGS constructs in MDCK cells (FLAG and SC35 staining)  132
Fig. 6.18  Transfection of NS1-pCAGGS constructs in MDCK cells (FLAG and NPM staining)  133
Fig. 6.19  Transfection of NS1-pCAGGS constructs in MDCK, A549 and CEF cells (FLAG staining)  134
Fig. 6.20  Cellular localisation of influenza A NS1  136
Fig. 6.21  Alignment of influenza A NS1 analysed in this study  137
Fig. 6.22  Individual domains of influenza B NS1 which were cloned for transient expression  137
Fig. 6.23  Expression of individual domains of Lee-NS1 in MDCK  138
Fig. 6.24  Expression of individual domains of 132-NS1 in MDCK  139
Fig. 6.25  Transfection of wild-type Lee-NS1 and LeeNS1 SSS107-109  140
Fig. 6.26  Transfection of Lee-NS1, 132-NS1, 132-NS1 ASS107-109, 132-NS1 AAS107-109 and 132-NS1 AAA107-109  141
Fig. 6.27  Transfection of 132-NS1, 132-NS1 S69N, 132-NS1 S123P and 132-NS1 S215N in MDCK cells  143
Fig. 6.28  Cellular localisation of wt 117-NS1 and 117-NS1 P94S in MDCK cells  143
Fig. 6.29  SDS-PAGE analysis of influenza B NS1 protein  144
Fig. 6.30  SDS-PAGE analysis of influenza A and B NS1 protein  145
Fig. 6.31  Time course analysis of B/Lee/40 NS1 upon HEK 293T infection  146
Fig. 6.32  SDS-PAGE analysis of individual NS1 domains  147
Fig. 6.33  Transfection of the 6-NS1 constructs on HEK 293T cells  149
Fig. 6.34  Cellular fractionation of NS1-transfected cells  150
Fig. 6.35  Cellular fractionation of double-tagged NS1-transfected cells  151
Fig. 6.36  Western blot analysis of Lee-NS1, Lee-NS1 D92N and 132-NS1  152
Fig. 6.37  Time course analysis of Lee-NS1 and Lee-NS1 D92N  153
Fig. 6.38  SDS-PAGE analysis of Lee-NS1, Lee-NS1 D92N, 132-NS1 and 132-NS1 N92D  154
Fig. 6.39  SDS-PAGE analysis of wt Lee-NS1, Lee-NS1 D92A, M91A, P93A, S94P, wt 117-NS1 and 117-NS1 P94S  155
Fig. 6.40  SDS-PAGE analysis of wt Lee-NS1 and Lee-NS1 SSS  156
Fig. 6.41  SDS-PAGE analysis of wt 132-NS1 and 132-NS1 S69N, S123P, S215N, ASS107-109, AAS107-109, AAA107-109  157
Fig. 6.42  VAD treatment of Lee-NS1 expression  158
Fig. 6.43  Doxycycline treatment of Lee-NS1 expression  158
Fig. 6.44  Gene alignment of influenza B NS1 and NEP ORF  166
Fig. 6.45  Proportion of specimens displaying N92 or D92 over the period of 1940-2011  169
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 6.46</td>
<td>3D structure of influenza B NS1 protein</td>
</tr>
<tr>
<td>Fig. 7.1</td>
<td>Size of BM2 ORF amplified form B/Lee/40</td>
</tr>
<tr>
<td>Fig. 7.2</td>
<td>Alignment of BM2 genes cloned in this study</td>
</tr>
<tr>
<td>Fig. 7.3</td>
<td>SDS-PAGE analysis of cloned Bm2 proteins</td>
</tr>
<tr>
<td>Fig. 7.4</td>
<td>IFA analysis of BM2-pCAGGS in MDCK cells (FLAG and GM130 staining)</td>
</tr>
<tr>
<td>Fig. 7.5</td>
<td>IFA analysis of BM2-pCAGGS in A549 cells (FLAG and GM130 staining)</td>
</tr>
<tr>
<td>Fig. 7.6</td>
<td>SDS-PAGE analysis of BM2 protein treated with cross-linker</td>
</tr>
<tr>
<td>Fig. 7.7</td>
<td>IFA of BM2 in HeLa cells</td>
</tr>
</tbody>
</table>
List of Tables

Table 1.1  Influenza B component of the TIV from 2009-2013  
Table 1.2  Lineages of the circulating influenza B strains in US compared to the vaccine strains 1999-2008  
Table 1.3  Respiratory viruses isolated from a SAF military barracks from 2006-2007  
Table 1.4  Influenza B proteins and their respective functions  
Table 1.5  Cleavage recognition site of the Ha0 from 3 different influenza strains  
Table 1.6  Primers used for sequencing in this study  
Table 1.7  Touchdown PCR thermocycling conditions  
Table 1.8  Accession numbers of the genes of the reference and ancestor strains used in this study  
Table 1.9  Primers used in amplifying NS1, NB, BM2 and ISG15 genes for molecular cloning  
Table 1.10  PCR thermocycling conditions used to amplify gene segments for molecular cloning  
Table 1.11  PCR thermocycling conditions for colony screening  
Table 1.12  Primers used for mutagenesis of the cloned specimens in this study  
Table 1.13  Thermocycling conditions recommended for site-directed mutagenesis  
Table 1.14  Reagents used in endoglycosidation of proteins in this study  
Table 1.15  Sizes of influenza B proteins in KiloDaltons (kDa)  
Table 1.16  Titres of egg-passaged B/lee/40 by end-point assay  
Table 1.17  Comparison of various titration assays for titering egg-passaged B/Lee/40  
Table 2.1  Accession numbers of the top hits of BLAST search results (HA)  
Table 2.2  Amino acid substitutions of the Victoria 'Tip' of the HA protein  
Table 2.3  Amino acid substitutions of the Yamagata 'Loop' of the HA protein  
Table 2.4  Amino acid substitutions of the glycosylation site 211 of the HA protein  
Table 2.5  Accession numbers of the top hits of the BLAST search results (NA)  
Table 2.6  Amino acid substitutions of the calcium binding domains of the NA protein  
Table 2.7  Amino acid substitutions of the NS protein  
Table 2.8  NB protein sequences grouped according to their similarity and their representative strains  
Table 2.9  Accession numbers of the top hits of the BLAST search results (NS)  
Table 2.10  Amino acid substitution of the NS1 NLS  
Table 2.11  Amino acid substitution of ISG15 binding site  
Table 2.12  Amino acid substitution of NEP  
Table 2.13  NS1 protein sequences grouped according to their similarity and their respective strains  
Table 2.14  Cloned influenza A NS1 genes  
Table 2.15  Ratio of p23: Full Length NS1 band intensity  
Table 2.16  Ratio of p23: Full Length NS1 band intensity of Lee-NS1 and 117-NS1 wt and mutants  
Table 2.17  Alignment of 388 influenza B NS1 protein, highlighting amino acids 87-87  
Table 2.18  List of specimens and their sequenced genes  

xvi
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-Terminus</td>
<td>amino terminus</td>
</tr>
<tr>
<td>α</td>
<td>anti-</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>BLAST</td>
<td>basic local alignment search tool</td>
</tr>
<tr>
<td>C-Terminus</td>
<td>carboxyl terminus</td>
</tr>
<tr>
<td>CDC</td>
<td>Centres for Disease Control</td>
</tr>
<tr>
<td>CEF</td>
<td>chicken embryonic fibroblast</td>
</tr>
<tr>
<td>CDC (Singapore)</td>
<td>Communicable Disease Centre</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary deoxyribonucleotide</td>
</tr>
<tr>
<td>CPE</td>
<td>cytopathic effect</td>
</tr>
<tr>
<td>Cyto</td>
<td>cytoplasmic</td>
</tr>
<tr>
<td>CT</td>
<td>cytoplasmic tail</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleotide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulphoxide</td>
</tr>
<tr>
<td>DSP</td>
<td>dithiobis[succinimidyl propionate]</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>DOX</td>
<td>doxycycline</td>
</tr>
<tr>
<td>ED</td>
<td>ecto domain</td>
</tr>
<tr>
<td>ED</td>
<td>effector domain</td>
</tr>
<tr>
<td>EM</td>
<td>electron microscopy</td>
</tr>
<tr>
<td>Endo H</td>
<td>endoglycosidase H</td>
</tr>
<tr>
<td>FBS</td>
<td>fetal bovine serum</td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
</tr>
<tr>
<td>GISRS</td>
<td>Global Influenza Surveillance Response System</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescence protein</td>
</tr>
<tr>
<td>HA</td>
<td>hemagglutinin</td>
</tr>
<tr>
<td>hr</td>
<td>hour(s)</td>
</tr>
<tr>
<td>hpi</td>
<td>hours post infection</td>
</tr>
<tr>
<td>HAT</td>
<td>human airway trypsin-like protease</td>
</tr>
<tr>
<td>HEK</td>
<td>human embryonic kidney</td>
</tr>
<tr>
<td>IFA</td>
<td>immunofluorescence assay</td>
</tr>
<tr>
<td>ITZ</td>
<td>influenza transmission Zone</td>
</tr>
<tr>
<td>kDa</td>
<td>kilo Daltons</td>
</tr>
<tr>
<td>LED</td>
<td>linker + effector domain</td>
</tr>
<tr>
<td>MDCK</td>
<td>Madin Darby canine kidney</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectroscopy</td>
</tr>
<tr>
<td>M1</td>
<td>matrix 1</td>
</tr>
<tr>
<td>MMP</td>
<td>matrix metalloprotease</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleotide</td>
</tr>
<tr>
<td>MOH</td>
<td>Ministry of Health</td>
</tr>
<tr>
<td>min</td>
<td>minute(s)</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>MCS</td>
<td>multi cloning site</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>MOI</td>
<td>multiplicity of infection</td>
</tr>
<tr>
<td>NEA</td>
<td>National Environmental Agency</td>
</tr>
<tr>
<td>NA</td>
<td>neuraminidase</td>
</tr>
<tr>
<td>NAI</td>
<td>neuraminidase inhibitor</td>
</tr>
<tr>
<td>NS1</td>
<td>non-structural 1</td>
</tr>
<tr>
<td>Nuc</td>
<td>nuclear</td>
</tr>
<tr>
<td>NEP</td>
<td>nuclear export protein</td>
</tr>
<tr>
<td>NLS</td>
<td>nuclear localisation signal</td>
</tr>
<tr>
<td>NPC</td>
<td>nuclear pore complex</td>
</tr>
<tr>
<td>NoLS</td>
<td>nucleolar localisation signal</td>
</tr>
<tr>
<td>NPM</td>
<td>nucleophosmin</td>
</tr>
<tr>
<td>NP</td>
<td>nucleoprotein</td>
</tr>
<tr>
<td>NISN</td>
<td>Neuraminidase Inhibitor Susceptibility Network</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>pC</td>
<td>pCAGGS</td>
</tr>
<tr>
<td>PNGase F</td>
<td>peptide-N-glycosidase F</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethanesulfonylfluoride</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>pfu</td>
<td>plaque forming unit(s)</td>
</tr>
<tr>
<td>PA</td>
<td>polyacidic</td>
</tr>
<tr>
<td>PB1</td>
<td>polybasic 1</td>
</tr>
<tr>
<td>PB2</td>
<td>polybasic 2</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>RS</td>
<td>arginine/serine-rich domain</td>
</tr>
<tr>
<td>QIV</td>
<td>quadrivalent inactivated vaccine</td>
</tr>
<tr>
<td>RIPA</td>
<td>radioimmunoprecipitation assay</td>
</tr>
<tr>
<td>Rf</td>
<td>relative migration distance</td>
</tr>
<tr>
<td>RE</td>
<td>restriction endonuclease</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleotide</td>
</tr>
<tr>
<td>RBD</td>
<td>RNA binding domain</td>
</tr>
<tr>
<td>RBL</td>
<td>RNA binding domain + linker</td>
</tr>
<tr>
<td>RRM</td>
<td>RNA recognition motif</td>
</tr>
<tr>
<td>SAF</td>
<td>Singapore Armed Forces</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SEA</td>
<td>South East Asia</td>
</tr>
<tr>
<td>TCID50</td>
<td>tissue culture infectious dose 50</td>
</tr>
<tr>
<td>TM</td>
<td>transmembrane</td>
</tr>
<tr>
<td>TEM</td>
<td>transmission electron microscopy</td>
</tr>
<tr>
<td>TIV</td>
<td>trivalent inactivated vaccine</td>
</tr>
<tr>
<td>TPCK</td>
<td>L-(tosylamido-2-phenyl) ethyl chloromethyl ketone</td>
</tr>
<tr>
<td>vRNP</td>
<td>viral ribonucleoprotein</td>
</tr>
<tr>
<td>vRNA</td>
<td>viral ribonucleotide</td>
</tr>
<tr>
<td>VTM</td>
<td>viral transport media</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>VAD</td>
<td>Z-VAD-FMK (caspase inhibitor)</td>
</tr>
</tbody>
</table>
Abstract

In Singapore, both influenza A and B viruses circulate in significant proportions throughout the year, with bi-annual peaks corresponding to the monsoon seasons. This statistic is not shared in most temperate countries where influenza is only prevalent during winter. Exploiting this unique situation in Singapore, we studied the characteristics of non-tissue culture adapted influenza B clinical specimens collected from SAF servicemen between the years of 2004 and 2009. A total of 81 clinical specimens were received and 66 tested positive for influenza B upon diagnostic PCR. Growth kinetics of these clinical specimens could not be studied as none of these specimens could replicate in tissue culture. Forty-six of these clinical specimens yielded sequence information for further analysis.

Lineage identity of these specimens were determined through phylogenetic analysis, followed by a comparison of the amino acid sequences between the HA, NA, NB, NS1, NEP and BM2 proteins. Molecular cloning of the NB, NS1 and BM2 genes of representative strains were carried out to determine if differences in sequence would translate into phenotypic differences.

Phylogenetic analysis show that majority of the specimens were reassortants with the HA belonging to the Victoria lineage and the NA from the Yamagata lineage. Only one specimen had both these gene segments from the Yamagata lineage. This finding has resulted in us determining several vaccine mismatches over the six years in focus. The NS1 phylogenetic tree showed that seven of the specimens isolated in 2004 clustered closely to B/Lee/40 than the rest. This finding provides an interesting look into the separation of the influenza B lineages.

Analysis of the amino acid sequences reveal that majority of the clinical specimens harboured a glycosylation site at position 211 of the HA protein, which was previously not dominant. No mutations conferring drug resistance were noticed among the HA and NA proteins analysed. The 7 specimens clustering with B/Lee/40 in the NS phylogenetic tree had almost identical amino acid sequences to B/Lee/40 NS1 and NEP proteins, differing from the other clinical specimens.

We found an inability of these clinical specimens to replicate well in tissue culture, a trait shared with the reference strain B/Lee/40. In an attempt to exploit the differences in amino acid sequences within the clinical specimens, the NB, BM2 and NS1 genes of representative sequences were cloned into a mammalian expression
vector. Basic molecular characterisation indicates that the NB protein, which has an unknown role, localised primarily in the cis-Golgi and is present throughout the Golgi. This is similar to the staining pattern displayed by the BM2 protein, suggesting a similarity in the function. SDS-PAGE analysis reveals that both these proteins are modified by the cell and further elucidation is required to determine the kinetics and the function of these modifications.

The bulk of the work presented in this thesis revolves around the molecular characterisation of the influenza B NS1 protein. Two main ‘phenotypes’ were observed in respect to cellular localisation; speckled and un-speckled. Two of the strains did not display localisation with nuclear speckles while three did. The introduction of mutations as well as expression of the different domains of the protein did not clarify which residues were involved in the dynamics of this nuclear speckle interaction.

Four of the specimens analysed presented a second, smaller peptide on SDS-PAGE analysis. Site-directed mutagenesis revealed that an aspartic acid residue at position 92, instead of an asparagine seems to be responsible for the presence of this smaller peptide. Even so, the mechanism yielding this peptide is still unknown since caspase and MMP inhibitors did not have an effect in its expression. Further elucidation has to be carried out to determine both the mechanism of expression as well as its function in viral replication.

The inability of the clinical specimens collected in this study prompts a detailed exploration into the choice of cell-culture systems in passaging influenza B clinical specimens. This may also reflect the lack of understanding of influenza biology, which has been focused primarily on lab-adapted strains instead of naturally circulating strains. The sequence information of the analysed specimens reveal that majority of the clinical specimens sequenced in this study were re-assortants with their HA from the Victoria lineage and the NA from the Yamagata lineage. Only 1 specimen proved to have both these glycoproteins from the Yamagata lineage. Sequence analysis of the HA protein show that a glycosylation site at position 211 is now prevalent amongst circulating strains and no neuraminidase inhibitor mutations were observed in the HA and NA genes.

Cloning of the NS1 gene showed that some of the specimens had a nuclear speckled localisation while other did not. The exact domain within the protein responsible for this was not determined in this study and further probing has to be
carried out. A smaller peptide, p23, was expressed in 4 of the 6 cloned NS1 genes. This is the first time the expression of p23 has been documented. The amino acids at the interface of the RNA-binding domain and the linker were crucial for its expression or silencing. The exact mechanism of its expression is yet to be elucidated.
Chapter 1: Introduction

1.1 Classification and Nomenclature

Influenza B viruses belong to the Orthomyxoviridae family of virus. This family is grouped under Class V of the Baltimore Classification of Viruses, together with other negative single-stranded RNA viruses (Baltimore, 1971). In addition, Orthomyxoviridae viruses have segmented genomes and are enveloped (Palese, 2007).

There are four other genera of virus in this family; Influenza Virus A, Influenza Virus C, Thogotovirus and Isavirus (Fauquet and Fargette, 2005). The three influenza viruses within this family are differentiated by the antigenic properties of their nucleoprotein and matrix proteins (Webster et al., 1982). While co-infection of Influenza B viruses with other Orthomyxo viruses can occur in vitro, it’s occurrence is rare and no inter-typic re-assortment of these viruses have been observed (Eshaghi et al., 2009; Falchi et al., 2008; Palese, 2007).

Studies on the RNA-dependent RNA polymerase complex of influenza A and B have revealed that the contact residues between the subunits (PB1, PB2 and PA) of each polymerase complex are not conserved between the two genera. Binding of influenza A PB1 to influenza B PA is only allowed when threonine-6 of influenza A PB1 is substituted to tyrosine, which naturally occurs in influenza B PB1 (Wunderlich et al., 2009).

Influenza viruses are named according to the following format; Genus/Species isolated from/Location of Isolation/Isolate Number/Year of Isolation. An influenza B virus strain isolated in from a human in Singapore, in 1999 with an isolate number of 1234 would be named; B/Human/Singapore/1234/1999. The species name is usually omitted if the virus was isolated from a human (Palese, 2007).

1.2 Symptoms and Epidemiology of Influenza B

In general the symptoms presented following an influenza B virus infection are similar to that of influenza A virus infection. Simply by observation, it is impossible to diagnose between the two viral infections due to the overlapping symptoms which include a sudden onset of high fever, sore throat, runny nose, dry cough, muscle and joint pains and severe malaise (Derlet, 2010; WHO, 2009; Wright, 2006).

Even though it has been generally accepted that influenza B viruses are less pathogenic than influenza A, influenza B viruses have been reported to cause severe
symptoms which include rhabdomyolysis, onset of toxemia of pregnancy, thyroid dysfunction, Stevens-Johnson syndrome, neurologic disorders, Reye’s syndrome and otitis media. Symptoms such as myositis and gastrointestinal complications are also closely associated with influenza B rather than influenza A infection (Derlet, 2010; Kim et al., 2009; Michael et al., 1980; William et al., 1980; Wright et al., 1980; Wu et al., 2009). Only recently, a 14-year old Japanese boy had developed Kleine-Levin Syndrome (KLS), a disease of hypersomnia and odd child-like behavior, just 6 days after being diagnosed with influenza B infection. Even though the causative agent of this disease is still unknown, influenza B (Kodaira and Yamamoto, 2012) as well as influenza A (Garland et al., 1965) viruses now are shortlisted as possible candidates.

By looking at the severe symptoms associated with influenza B infection, William et al have actually suggested that even though the transmission of influenza B virus is not comparable to that of influenza A, the pathogenicity and the severity of disease are (William et al., 1980). However this is still to be confirmed.

All of the mentioned symptoms and complications are more common among children and the elderly infected with influenza B rather than among adults. This could be due to the slower rate of evolution of influenza B, allowing adults to have fully primed immune responses against influenza B but not having a similar response against influenza A (Wright et al., 1980). Typically individuals between the age of 5-9 years present with the highest rate of infection and disease due to influenza A and B virus infection. This can be attributed to the ‘unprimed’ immune response to influenza coupled to the nature of communal pre/early-school conditions. However the greatest morbidity and mortality caused by both these viruses are observed in the elderly, those younger than 2 years of age, as well as the immune-compromised (WHO, 2005).

1.3 Influenza B Tropism

While Influenza A viruses infect a large range of mammalian and avian hosts and Influenza C viruses infects both humans and pigs, influenza B viruses primarily infect humans only (Murphy, 1990). There have been, however, a few reports indicating that influenza B viruses are able to replicate naturally in wild animals. Antibodies isolated from the sera of domesticated pigs proved to be specific to a circulating human strain of influenza B virus (Takatsy et al., 1967). The same group then found that domesticated piglets could be experimentally infected with influenza B virus which
could then be spread through contact. This experimental infection was believed to cause mild illness in the piglets (Takatsy et al., 1969).

In 1971, a strain of influenza B virus was isolated from a Dachshund dog in Taipei city, Taiwan. This strain of virus had hemagglutination inhibition properties similar to B/Victoria/98926/70, a human strain isolated from the previous year (Chang, 1976). Interestingly, when infected intranasally with a strain of influenza B virus, stray mongrel dogs were not susceptible while cats (Felis Catus L.) and monkeys (Macaca radiata) were (Paniker and Nair, 1972).

To date, data from studies of marine animals have indicated that seals are susceptible to influenza B viruses. In 1999, Osterhaus et al had isolated an influenza B virus from a seal (Phoca vitulina) which was believed to have been naturally infected. This isolated virus (B/Seal/Netherlands/1/99) is highly related to virus strains which were present in humans 5 years earlier. The authors also determined that 2% of the seals they studied had antibodies specific to influenza B viruses. This finding had prompted the authors to suggest that seals serve as an animal reservoir of influenza B (Fouchier et al., 2001; Osterhaus et al., 2000). While no conclusive evidences have been found to support this postulation, recent findings have confirmed the ability of influenza B to infect seals (Bodewes et al., 2013).

Recently, a study has found that in an experimental setting, circulating strains of influenza B viruses were able to infect and replicate to high titres in the upper respiratory tracts of guinea pigs. These viruses also had the ability to transmit from the inoculated to a naïve guinea pig, leading to authors to propose the guinea pig as a viable animal model for the study of influenza B viruses (Pica et al., 2012). Further elucidation is needed to determine if guinea pigs are naturally infected with influenza B in the wild, validating its use as a representative animal model.

The non-structural protein, NS1, of influenza B viruses has been implicated in the host-range specificity of influenza B viruses. Influenza B, but not influenza A, NS1 protein binds to an ubiquitin-like ISG15 protein of the infected host cell (Yuan and Krug, 2001). ISG15 is one of the most abundantly upregulated proteins upon interferon α/β activation in virus infected cells (Der et al., 1998; Farrell et al., 1979; Liu et al., 2012). Following its upregulation, ISG15 is conjugated to a large range of host cell proteins (Giannakopoulos et al., 2005; Zhao et al., 2010). Even though the mechanism of which has not been elucidated, the upregulation of ISG15 and/or its conjugation is
believed to be responsible for the inhibition of influenza B viruses. This was further confirmed by the observation that ISG15 knockout mice were susceptible to influenza B infection (Lenschow et al., 2007). The ability of influenza B NS1 to bind to ISG15 is therefore believed to prevent ISG15 in inhibiting virus replication.

Influenza B NS1 has been shown to bind to human and primate ISG15 but not mouse or canine ISG15. The specificity of NS1 to human and primate ISG15 is attributed to the hinge region of the ISG15 protein which is variable across different species but conserved between humans and primates (Guan et al., 2011a; Sridharan et al., 2010). This specificity of influenza B NS1 could explain the virus’ inability to replicate hosts other than humans.

Besides the NS1 protein, a single amino acid change in the C-terminus of the matrix (M1) protein of a strain of influenza B virus allows for a virulent adaptation in mice, suggesting a role for the matrix protein in the host-range selectivity of influenza B viruses (McCullers et al., 2005).

1.4. Influenza B Lineages

Ever since the early 80s, 2 lineages of influenza B viruses have circulated; the Yamagata lineage and the Victoria lineage (named after their first representatives; B/Yamagata/16/88 and B/Victoria/2/87) (Rota et al., 1990). It is postulated that the Victoria lineage emerged from China in the late 60s or early 70s (Chen J, 2002; Luo et al., 1999; Shaw et al., 2002).

The Victoria lineage was more dominantly circulating in the 80s (Rota et al., 1992; Rota et al., 1990). The Yamagata lineage became prominent in the early 90s with the viruses from the Victoria lineage being isolated infrequently in Eastern Asia (CDC, 2001; Nakagawa, 2000; Nerome et al., 1998; WHO, 1999a, 2002). Between the years 1994-1997, the Victoria lineage started emerging in South China and Japan (McCullers et al., 1999) and by 2002, viruses from the Victoria lineage were detected throughout Asia, Europe and North America (Shaw et al., 2002). Ever since the turn of the century, both lineages have caused global epidemics at similar frequencies (Chi, 2005; Chi et al., 2003).

A study characterizing 205 influenza B clinical specimens collected in China between March 2009 and October of 2010 showed 2 distinct epidemics. The first in early 2009 contained both lineages circulating while the second in early 2010 had the 80% of the strains belonging to the Victoria lineage (Tan et al., 2013). A recent study
sequenced the genome of 6 influenza B viruses circulating in India in 2010 and found that 3 specimens had their HA belonging to the Yamagata lineage and 3 to the Victoria lineage. The NA gene segments of all 6 belonged to the Yamagata lineage. This proves that multiple genotypes of influenza B are circulating concurrently (Patil et al., 2013). In 2012, 92 circulating influenza B specimens in Australia were analysed through serological methods to show that 70 belonged to the Victoria lineage (Fielding et al., 2013).

A retrospective study in Brazil, another tropical country, showed that in influenza B viruses analysed between 1996 and 2001 were exclusively of the Yamagata lineage. The Victoria lineage emerged in 2002 and circulated together with the Yamagata lineage, this was the case in 2005, 2006 and 2008. The Yamagata lineage was predominant in 2010 and 2011 with the Victoria lineage circulating in the intervening years (Paiva et al., 2013).

These 2 lineages differ antigenically and genetically (Kanegae et al., 1990). Ferrets previously infected by a strain from one lineage do not develop immunity to the other (Rota et al., 1992). Exploiting the genetic difference between the lineages, a real-time PCR method has been develop to determine which lineage a particular strain is from, by using the hemagglutinin gene as template (Biere et al., 2010).

1.5 Influenza B Evolution.

It has generally been accepted that influenza B viruses evolve at a much slower rate than influenza A (Air et al., 1990; Berton et al., 1984; Berton and Webster, 1985; Hay et al., 2001; Krystal, 1983 ; Nerome et al., 1998; Nobusawa and Sato, 2006; Rota et al., 1990) but faster than influenza C (Yamashita et al., 1988). Various reasons have been put forward to explain this. Due to the lack of an animal reservoir, influenza B viruses does not rely heavily on antigenic shift, as influenza A viruses, but only antigenic drift as a means of evolution (Cox and Bender, 1995; Nerome et al., 1998; Weber et al., 1997). It has also been suggested that the polymerase complex of influenza B viruses are more accurate than influenza A viruses, resulting in a lower nucleotide mutation rate (Nobusawa and Sato, 2006).

Air et al have compared the rates of nucleotide and amino acid mutations in influenza A and B viruses and have found that in influenza A, half of the nucleotide changes cause a corresponding change in the amino acid sequence while only 30% of the nucleotide mutations in influenza B results in an amino acid change. This larger
variation at the nucleotide level in influenza B has caused the authors to believe that there exists greater negative pressure on nucleotide variants for influenza B than influenza A viruses (Air et al., 1990).

Reassortment of the gene segments between the Yamagata and Victoria lineages have been postulated as a main mechanism of influenza B viruses (Chen and Holmes, 2008; McCullers et al., 1999; Xu et al., 2004). This occurs when 2 viruses of different lineages infect a single cell; the resulting progeny virus would then contain gene segments from both lineages, resulting in it being a reassortant (Domingo, 2006; Wright, 2006). Studies (including data from this thesis) have suggested that most of the circulating influenza B viruses presently are reassortants, having the HA and NA genes from differing lineages (Luo et al., 1999; McCullers et al., 2004). The genetic reassortment of influenza B viruses is believed to be driven by immune selection on the HA, NA and NS1 proteins (Chen and Holmes, 2008). Figure 1.1 displays a schematic representation of re-assortment between a ‘black’ virus and a ‘grey’ virus. Color coding in this example represents the 2 different lineages for influenza B.

Figure 1.1 Schematic of genetic reassortment between 2 different influenza strains. 2 different strains of influenza B virus (grey and black) infect a single host cell, resulting in the generation ‘hybrid’ reassortant progeny, displaying different combinations of the gene segments. Image taken from Fields Virology 5th Edition (Wright, 2006).
1.6 Influenza B Vaccination

The surface antigens of influenza viruses (both A and B) undergo a substantial amount of antigenic drift between influenza seasons. This would result in the generation of an antigenically new strain of influenza virus during each influenza season. Therefore a new influenza vaccine has to be developed for each new season (WHO, 2005). Biannually, the World Health Organization (WHO) would recommend three different influenza strains to make up the influenza vaccine for the winters for the Northern and Southern Hemispheres (WHO, 2011, 2012). These recommendations are based on worldwide influenza surveillance by the WHO. The selection of the influenza B component of the vaccine recommendation is based on assaying the primary isolates collected by the WHO through Hemagglutination Inhibition (HI) tests with post-infection ferret antisera (Global Influenza Programme, 2002). Typically, the vaccine recommendation is trivalent, composing of 2 influenza A strains (H1N1 and H3N2) and 1 influenza B strain. In accordance to the HI test, the WHO has traditionally classified the influenza B component as either belonging to the Yamagata or the Victoria lineage. The components of the 2012-2013 influenza season for the northern hemisphere are:

- an A/California/7/2009 (H1N1) pandemic2009–like virus
- an A/Victoria/361/2011 (H3N2)-like virus
- a B/Wisconsin/1/2010-like virus

Table 1.1 below shows the influenza B virus component of the trivalent inactivated vaccine (TIV) from 2004 to 2013. These candidate viruses are also listed with their respective lineage, as indicated by the HI-test.

The TIV is delivered by intramuscular injection, typically the deltoid muscle for adults and the antero-lateral aspect of the thigh for infants (WHO, 2005). Initial studies on influenza vaccination, seem to suggest that vaccination can yield up to 80% efficacy (Patriarca et al., 1985), however recent scrutiny of the vaccine have suggested that modifications need to be made with regards to the influenza B component of the vaccine.

The HI test employed by the WHO to determine the antigenicity of the primary isolates has been shown to be less accurate than antibody-neutralisation assays. It has been suggested that the HI-positive sera do not necessarily inhibit the virus successfully (Cheng et al., 2012). This lack of coherence between the two assays can be attributed to the fact that the HI test only focuses on the antigenicity of the hemagglutinin (HA)
protein of the virus while the antibody-neutralisation test measures the antigenicity of the virus as a whole. Besides the HA protein, the neuraminidase (NA) protein is also present on the exterior of the virus. Albeit being stoichiometrically less abundant than the HA protein, the NA protein, when expressed in a virus-like particle, has been shown to elicit an immune response, strong enough to confer resistance to antigenetically similar and distinct viruses (Quan et al., 2012). This finding shows that the combination of both spike proteins are an important determinant in the immune response elicited upon infection.

<table>
<thead>
<tr>
<th>Year</th>
<th>Hemisphere</th>
<th>Strain</th>
<th>Lineage (by HI-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012-2013</td>
<td>Northern</td>
<td>B/Wisconsin/1/2010-like virus</td>
<td>Yamagata</td>
</tr>
<tr>
<td>2012</td>
<td>Southern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-2012</td>
<td>Northern</td>
<td>B/Brisbane/60/2008-like virus</td>
<td>Victoria</td>
</tr>
<tr>
<td>2011</td>
<td>Southern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-2011</td>
<td>Northern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>Southern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009-2010</td>
<td>Northern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>Southern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>Southern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007-2008</td>
<td>Northern</td>
<td>B/Malaysia/2506/2004-like virus</td>
<td>Victoria</td>
</tr>
<tr>
<td>2007</td>
<td>Southern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006-2007</td>
<td>Northern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>Southern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005-2006</td>
<td>Northern</td>
<td>B/Jiangsu/10/2003-like virus</td>
<td>Yamagata</td>
</tr>
<tr>
<td>2005</td>
<td>Southern</td>
<td>B/Shanghai/361/2002-like virus</td>
<td>Yamagata</td>
</tr>
<tr>
<td>2004-2005</td>
<td>Northern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>Southern</td>
<td>B/Hong Kong/330/2001-like virus</td>
<td>Victoria</td>
</tr>
</tbody>
</table>

As explained in sections 1.4 and 1.5, influenza B viruses belong to either the Yamagata or the Victoria lineage (Rota et al., 1990). Children and young infants who have been vaccinated with a virus from the Yamagata lineage are not immunologically primed against the Victoria lineage (Levandowski et al., 1991b). Since, influenza B viruses rely heavily on genetic re-assortment, it is highly possible to have reassortant viruses with their HA and NA proteins belonging to the different lineages (McCullers et al., 1999; Xu et al., 2004). Hence, solely recommending a vaccine strain according to the antigenicity of the HA protein should be deemed insufficient.
Only 1 representative of the 2 lineages of influenza B viruses is recommended to be in the biannual trivalent vaccine. This creates the almost impossible task of predicting which lineage will predominate in the next influenza season. As an example, more than 95% of the influenza B viruses circulating in the 2007-2008 season belonged to the Yamagata lineage while the vaccine strain, B/Malaysia/2506/2004 belonged to the Victoria lineage (CDC, 2008). The task of predicting the next lineage is made more arduous as it seems that both lineages are now equally circulating during each season. Table 1.2, lists the lineage of the circulating strains in the United States compared to that of the vaccine strain from 1999 to 2008 (Belshe et al., 2010).

The data from table 1.2 show that only in 5 out of the 10 seasons listed did the vaccine strain match the circulating strain. The spread of influenza is also specific to geography. An example can be found in the choice for influenza B component of the 2012-2013 Northern Hemisphere vaccine. B/Wisconsin/1/2010 (Table 1.1) belongs to the Yamagata lineage. This was chosen as in the WHO report; viruses of the Yamagata lineage were of the predominant lineage worldwide. It was also reported that this was not the case in China, where viruses of the Victoria lineage were predominant (WHO, 2012). Influenza outbreaks during mass gatherings may also cause a shift in the distribution of influenza spread. As seen in World Youth Day 2008 celebration in Australia, the circulating influenza B strain in Australia before the event was B/Florida/4/2006-like (Yamagata lineage), which matched the vaccine component of the 2008 TIV for the Southern Hemisphere (Table 1.1). However, after the event, B/Malaysia/2506/2004-viruses (Victoria Lineage) were circulating as predominantly as B/Florida/4/2006 (Blyth et al., 2010). Not only has it been shown that influenza incidences were significantly lowered when the vaccine matches the circulating strain, but influenza stricken patients spend 1.36 times more during years where the influenza vaccine does not match (Karve et al., 2013).

Findings such as these have called for the implementation of a quadrivalent influenza vaccine (QIV), where both lineages of influenza B are represented in the biannual vaccine (Ambrose and Levin, 2012; Belshe, 2010; Reed et al., 2012). Initial analysis shows that the addition of a fourth component to the vaccine does not impair the immunological response elicited by the other three components. The addition of a second influenza B strain also results in broader immunological protection against influenza in general (Beran et al., 2013).
Table 1. Lineages of the circulating influenza B Strains in the US compared to the recommended vaccine strains from 1999 to 2008. Taken from Belshe et al 2010.

<table>
<thead>
<tr>
<th>Year</th>
<th>Lineage</th>
<th>Proportion of circulating influenza B viruses by HA lineage (%)</th>
<th>Match of predominant epidemic strain to vaccine strain?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006-2009</td>
<td>Yamagata</td>
<td>87, 13</td>
<td>No</td>
</tr>
<tr>
<td>2007-2008</td>
<td>Victoria</td>
<td>3, 97</td>
<td>No</td>
</tr>
<tr>
<td>2006-2007</td>
<td>Victoria</td>
<td>72, 28</td>
<td>Yes</td>
</tr>
<tr>
<td>2005-2006</td>
<td>Yamagata</td>
<td>81, 19</td>
<td>No</td>
</tr>
<tr>
<td>2004-2005</td>
<td>Yamagata</td>
<td>26, 74</td>
<td>Yes</td>
</tr>
<tr>
<td>2003-2004</td>
<td>Victoria</td>
<td>7, 93</td>
<td>No</td>
</tr>
<tr>
<td>2002-2003</td>
<td>Victoria</td>
<td>100, –</td>
<td>Yes</td>
</tr>
<tr>
<td>2001-2002</td>
<td>Yamagata</td>
<td>77, 23</td>
<td>No</td>
</tr>
<tr>
<td>2000-2001</td>
<td>Yamagata</td>
<td>–, 100</td>
<td>Yes</td>
</tr>
<tr>
<td>1999-2000</td>
<td>Yamagata</td>
<td>–, 100</td>
<td>Yes</td>
</tr>
</tbody>
</table>

* Influenza seasons in which influenza B strain accounted for at least 20% of all isolates (source: B, Couch, Feb 28, 2007 VRBPAC meeting) [13].  
* Although the lineage was matched, there were significant amino acid differences in clinical isolates of the Yamagata lineage compared with the vaccine strain.

1.7 Global Prevalence of Influenza B

The WHO has set up the Global Influenza Surveillance Response System (GISRS) (http://www.who.int/influenza/gisrs_laboratory/en/) for the purpose of surveilling influenza activity globally as well in specific regions and countries.

![Global circulation of influenza viruses](http://www.who.int/influenza/gisrs_laboratory/en/)

Figure 1.2 Number of global influenza specimens collected by GISRS over the period of a year up to week 41 of 2012. Colour coding of the bar indicate the genus, sub-type and lineage of specimens. (Taken from http://www.who.int/influenza/gisrs_laboratory/en/ Oct, 2012)

Figure 1.2 shows the breakdown of influenza specimens collected for a year up to week 41 of 2012 (30th September - 6th October). This data indicates that in the past year, the main influenza virus circulating belonged to the influenza A genus,
specifically the H3 subtype. The peak of number of specimens collected was observed within January and March. This increase in specimens collected also saw an increase in influenza B virus specimens. Less than half of the influenza B viruses had their lineages determined, but out of those determined it can be said that globally viruses of the Victoria lineage predominated. This observation is obvious between weeks 1-13 (GISRS, 2012).

Despite the overall predominance of influenza A on the global scale, when we observe the breakdown of influenza specimens collected within week 39 (23rd-29th September), we can see that certain regions have a higher prevalence of influenza B than A as indicated in Figure 1.3.

Each pie chart indicated in Figure 1.3 is representative of 1 influenza transmission zone (ITZ), as designated by the WHO. Eight out of the 15 ITZs displayed show more influenza B viruses collected than influenza A, suggesting that the global surveillance data in Figure 1.2 does not discriminate between the different ITZs. Figure 1.4 below shows the composition of influenza specimens collected over the past year, specifically for the South East Asia (SEA) ITZ.

The distribution of influenza specimens collected from SEA is not similar to the distribution of influenza virus specimens collected globally. Influenza B specimens range from 25.9% (week 41) to 81.3% (week 25). The peak of influenza activity in SEA (weeks 27-41) does not correlate with the peak in the Figure 1.2. These data suggests that circulation of influenza viruses are region specific and that SEA is a good region to study influenza B transmission.

### 1.7.1 Influenza in Singapore

Singapore lies in the tropics having a latitude and longitude of 1.5°N and 104°E, respectively. On average its climate ranges from a minimum range of 23.9°C to 25.4°C and a maximum of 30.3°C to 32.3°C and experiences a relative humidity of 84%. Two monsoon seasons; in April and December, gives Singapore its maximum rainfall (Chow, 2006; NEA, 2010). Singapore is also a densely populated country with a population of 4.99 million as of June 2009 occupying a land space of 710.3 square kilometers (Singstat, 2010). Its population size and geographic location allows it to serve as an international hub of trade, commerce and tourism (Doraisingham, 1987).
Figure 1.3 Percentage of influenza virus specimens collected within week 39 by the GISRS. Genus and sub-type are color coded as indicated in the diagram. (Taken from http://www.who.int/influenza/gisrs_laboratory/en/ Oct, 2012)
All these stated reasons are believed to contribute to the unique influenza viral dissemination in Singapore, compared to that in temperate countries. In non-tropical countries, influenza viruses circulate predominantly during winter, whereas in Singapore, influenza viruses circulate all year round with biannual peaks between April-July and November-January (Chew et al., 1998; Chow, 2006; Doraisingham, 1986, 1987, 1988a, b; Hampson, 1999). Understanding Singapore’s position in the global community, the state of influenza circulation locally would mirror that of the international community.

It has been estimated that out of 100000 influenza infections, 14.8 deaths are incurred in tropical countries. The exact breakdown of influenza A and influenza B related deaths are unknown. Between the years 1996 to 2003 out of a number of suspected respiratory specimens isolated in Singapore, the percentage of influenza A positives ranged between 2.6%–9.1% while that for influenza B ranged between 0.4%–1.6% (Chow, 2006). In a recent study, 1352 nose swabs of soldiers from a Singaporean military camp suspected of having respiratory infections were tested for influenza
viruses by diagnostic PCR methods. 24% of these samples were positive for influenza A while 12% were positive for influenza B (Seah et al., 2010a). With Singapore’s increasing population and the current trend in increasing resistance to treatment (Hatakeyama et al., 2007; Moscona and McKimm-Breschkin, 2007), the infection rate of Influenza B is anticipated to increase.

All of these reasons stated make Singapore an ideal country to study influenza virus transmission, circulation and evolution. However, not much attention has been devoted to the study of influenza B in Singapore. The earliest record of such a study was in 1965 where an influenza B virus outbreak was observed in the National University of Singapore in December 1963-February 1964. This study claims to report the novel introduction of Influenza B virus to Singapore (Yin-Coggrave, 1965). The viruses isolated in this study were antigenically related to an influenza B strain from Johannesburg rather than one from Malaya (current day Malaysia and Singapore) which was isolated in the 50s (Godon Smith, 1956). The virus could have been introduced to this community from a foreign visitor to the university and spread within the closely-knit university population.

Between the years of 1972-1999, 25 influenza epidemics have been reported in Singapore. Only 4 of these epidemics were caused by influenza B. However, 2 of the 4 epidemics also had influenza A strains co-circulating with them (Lee et al., 2009).

The study by Yin-Coggrave, 1965, and subsequent studies of influenza B viruses in Singapore have merely focused on the antigenic and epidemiologic properties of the virus and not on sequence analysis of the genes and proteins. To the best of our knowledge, this is the first in depth sequence analysis of influenza B viruses in Singapore, giving us greater insight into the evolution of the virus and individual gene segments.

1.7.1.1 Influenza B Prevalence in Singapore

Figure 1.5, taken from the Communicable Diseases Centre, Ministry of Health (MOH) of Singapore shows the distribution of influenza viruses in Singapore from June 2009 to July 2012 (Communicable Diseases Division, 2012).
Figure 1.5 Distribution of Influenza Viruses in Singapore from June 2009 to July 2012. Taken from CDC, MOH, Singapore. ■ Influenza B, □ Intermediate (low viral titre), ■ Influenza A (H1N1 pandemic 2009), ■ Influenza A (H3N2) (Aug, 2012 Taken from http://www.moh.gov.sg/content/moh_web/home/statistics/infectiousDiseasesStatistics/weekly_infectiousdiseasebulletin.html)

Figure 1.5 shows that since early 2010, influenza B viruses (blue bar) have been dominantly circulating in Singapore. The distribution of influenza B seems to be almost reciprocal to that of influenza A (H3N2) (grey bar). The spread of influenza B peaked twice, once in September 2010 and another in February 2012, causing up to 80% of all influenza infections in Singapore. Correspondingly, when the prevalence of influenza A (H3N2) peaked in June to August 2011, influenza B was reduced to a 10-20%. However, it is interesting to note that the prevalence of both seasonal influenza A (H3N2) and influenza B is dependent on the prevalence of the pandemic influenza A H1N1, represented by the red bar. At the peak of the H1N1 pandemic, between August 2009 to February 2010 minimal to little influenza B or influenza A (H3N2) was observed. Figure 1.6 shows a similar chart of the influenza distribution during the period of the H1N1 pandemic from May 2009 to Aug 2010.

In February 2009, an outbreak of influenza A H1N1 was reported in a Mexican town of La Gloria, Veracruz and by end of April, this virus was circulating worldwide. Genetic analysis of this virus showed that it was a triple-reassortant with segments belonging to human, avian as well as swine influenza viruses (Neumann et al., 2009). Not surprisingly, the spread of this novel influenza A virus reached Singapore, through imported cases (Mukherjee et al., 2010). The spread of this pandemic H1N1 virus has impacted on the spread of seasonal influenza A and B as seen in figures 1.5 and 1.6. Also shown in both these figures is the appearance of influenza B virus upon the decline of the pandemic H1N1 virus.
Figure 1.6 Weekly Prevalence of influenza A and influenza B virus in Singapore from May 2009 to Aug 2010. Blue Bars: Seasonal Influenza A strains (%), Pink Bars: Pandemic H1N1 Influenza A (%), Gray Bars: Influenza B (%). Taken from CDC, MOH Singapore. (Aug. 2012)
To further understand the relationship between the spread of pandemic H1N1 and influenza B, the following charts were compiled from the Global Influenza Surveillance and Response System (GISRS), with regards to the prevalence of influenza viruses in Singapore, Mexico and the United States of America (Figure 1.7) (GISRS).

Data from all three countries show a spike in the prevalence of the pandemic H1N1 towards the second quarter of 2009. At the peak of the pandemic in all countries, influenza B was reduced. However, at the end of the pandemic (second quarter of 2010), the re-emergence of influenza B viruses in Singapore occur instantly but not in Mexico or in USA. From week 10 of 2010, minimal cases of influenza were reported in both Mexico and USA, while Singapore had a significant number of cases of both pandemic H1N1 and influenza B. Interestingly, when the pandemic H1N1 dwindled in week 30 of 2010, majority of the influenza cases reported in Singapore were of influenza B. This was not seen in Mexico or in USA. The lack of data between weeks 38-47 in Singapore could be attributed to the cessation of surveillance due to the national response to the pandemic (Fig 1.7).

The data from figures 1.5-1.7 could shed an important aspect about influenza B prevalence worldwide. After the end of the pandemic, most of the population could possibly be immunologically primed against influenza A in general. Therefore it would be easier for influenza B to secure a foothold in such a population.

1.7.1.2 National Service in Singapore

Singaporean and second generation permanent resident males, who have reached the age of 18, are enrolled into the military in a conscription exercise known as National Service. Since a large percentage of Singaporeans are serving in the Singapore Armed Forces (SAF), a sample of the viruses circulating within the army barracks would mirror the general population (Seah et al., 2010a).

The close proximity of the living quarters and often unsanitary training conditions have been attributed to the reason why military barracks have uniquely high rate of transmission for respiratory infections (Gray et al., 1999; Kak, 2007; Russell, 2006). Early studies suggest a positive correlation to the incidence of respiratory infections to the number of soldiers in training (Miller, 2007). Interestingly, the first few reports of the 1918 influenza A (H1N1) pandemic was reported from a military unit (Anonymous, 1918; Cole, 1918).
Figure 1.7 Data regarding the distribution of influenza cases in Singapore (top panel), Mexico (middle panel) and USA (bottom panel) throughout 2009 and 2010. (Taken from http://www.who.int/influenza/gisrs_laboratory/en/ Aug, 2012)

Most studies of influenza transmission in a military setting are based in the USA. Most of these findings suggest that influenza is not the most common cause of respiratory disease in barracks. This could be due to the mandatory annual vaccination
of US soldiers (Kak, 2007; Russell, 2006). Despite being mandatory, the vaccine does not have 100% efficacy (Eick-Cost et al., 2012; Strickler, 2007), as demonstrated by how a strain of influenza A imported from Puerto Rico managed to cause an outbreak in Lackland Air Force Base in San Antonio, Texas during the summer of 1999 (Laurel et al., 2001).

Influenza vaccination was not compulsory in the SAF, which could be the reason as to why influenza was found to be the most common cause of respiratory illness in a study done from 1 SAF military camp, causing 36% of the respiratory viruses screened (Table 1.3) (Seah et al., 2010a). Interestingly, adenovirus which is the most common respiratory virus in US military barracks (Hilleman et al., 1957; Russell, 2006), only attributed to 0.4% of respiratory viruses screened in the Singapore military camp, lower than both influenza A and B viruses, rhinovirus and human metapneumovirus (Table 1.3).

Table 1.3 Respiratory viruses isolated from a SAF military barrack from 2006-2007. Taken from Seah, 2010

<table>
<thead>
<tr>
<th>Viral agents tested</th>
<th>Number(a) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of specimens collected</td>
<td>1354</td>
</tr>
<tr>
<td>Adenovirus</td>
<td>5 (0.4)</td>
</tr>
<tr>
<td>Influenza A virus</td>
<td>326 (24)</td>
</tr>
<tr>
<td>Influenza B virus</td>
<td>159 (12)</td>
</tr>
<tr>
<td>Influenza C virus</td>
<td>4 (0.3)</td>
</tr>
<tr>
<td>Parainfluenza type 1</td>
<td>1</td>
</tr>
<tr>
<td>Parainfluenza type 2</td>
<td>0</td>
</tr>
<tr>
<td>Parainfluenza type 3</td>
<td>4 (0.3)</td>
</tr>
<tr>
<td>Parainfluenza type 4</td>
<td>0</td>
</tr>
<tr>
<td>Rhinovirus</td>
<td>15 (1.1)</td>
</tr>
<tr>
<td>Coronavirus, OC43</td>
<td>0</td>
</tr>
<tr>
<td>Coronavirus, 229E</td>
<td>1</td>
</tr>
<tr>
<td>Respiratory syncytial virus</td>
<td>0</td>
</tr>
<tr>
<td>Human metapneumovirus</td>
<td>9 (0.7)</td>
</tr>
<tr>
<td>Total number tested positive</td>
<td>524 (38.7)</td>
</tr>
</tbody>
</table>

\(a\) The number of clinical specimens that tested PCR positive for respiratory viruses in Camp X.

1.8 Influenza Virus B Structure and Morphology

The structure of influenza B viruses has been elucidated through electron microscopy, hydrodynamical measurements and by small-angle neutron scattering. Influenza B viruses are practically indistinguishable from influenza A viruses through electron microscopy (Palese, 2007) (Figure 1.8). On average, Influenza B viruses have a molecular mass of 170-200 x10\(^6\) Daltons. Virions are mostly spherical, however filamentous and kidney shaped virions have been observed through electron microscopy (Mellema et al., 1981; Ruigrok et al., 1984; Servidei et al., 1987; Waterson et al., 1963).
Influenza B viruses are enveloped, attaining the lipid bilayer by budding through plasma membrane of an infected cell (Harrison, 2007). This lipid bilayer has an internal radius of 425Å and an external radius of 580Å. This lipid bilayer has an estimated molecular weight of $60 \times 10^6$ Daltons (Mellema et al., 1981).

![Figure 1.8 Electron Microscopy of influenza A and B](image)

**Figure 1.8 Electron Microscopy of influenza A and B**

- b. Electron micrograph of B/Lee/40. Showing clearly the envelope/spike proteins. The dark regions in both virions are believed to be depressions on the surface. Adapted from Waterson, Hurrell et al. 1963.

The influenza B virion has 4 proteins inserted into the envelope; HA, NA, NB, and BM2 (Betakova et al., 1996; Brassard et al., 1996; Odagiri et al., 1999; Zebedee and Lamb, 1988). These proteins form a shell of 10-14nm around the virion. Gene segments are encapsidated with viral nucleocapsid proteins (NP) and are coupled with proteins of the polymerase complex, (PA, PB1 and PB2). These gene segment and protein conjugates are known as ribonucleoprotein complexes (RNPs) (Harrison, 2007). The egress of influenza B RNP into the cytoplasm has been shown to activate cellular RNA-dependent Protein Kinase (PKR). This is believed to be brought about by the panhandle structure of the viral RNA Other proteins which are inside the virion are NS2/NEP and M1 (Imai et al., 2003). A schematic diagram of an influenza B virus is shown in figure 1.9.

### 1.9 Influenza B Genome Structure

The genome of influenza B viruses is composed of 8 negative stranded, RNA segments (Palese et al., 1980; Ritchey, 1976). The number of gene segments has been elucidated by Palese *et al.*, 1980, by visualizing them on a polyacrylamide gel and comparing them to the gene segments of influenza A viruses (figure 1.10). These
individual segments are later transcribed to messenger RNA molecules which are regarded as positive polarity (Baltimore, 1971)

These gene segments are numbered according to their sizes; segment 1 being the longest segment and segment 8 being the shortest. Gene segments 1-5 are monocistronic while gene segments 6-8 code for 2 proteins, each with a different mechanism as shown in figure 1.11 (Palese, 2007)


**Figure 1.9** Schematic diagram of Influenza B structure. Taken from [http://www.expasy.ch/viralzone/all_by_species/80.html](http://www.expasy.ch/viralzone/all_by_species/80.html) Aug, 2012

![Radio labeled genomes](image)

**Figure 1.10** Radio labeled genomes of influenza A and B viruses (lanes 1 and 2) visualized on a polyacrylamide gel, showing 8 differently sized segments each. The gene segments in this figure are named after the main protein that it transcribes. Taken from (Palese et al., 1980)
The 5’ and 3’ non-coding regions of each gene segment had been found to show partial and inverted complementarity (Desselberger et al., 1980; Lee and Seong, 1998). This partial complementarity results in the formation of panhandle structure as shown in figure 1.9. This panhandle structure mimics double stranded RNA and has been postulated to be the reason why RNA-dependent Protein Kinase (PKR) is activated in infected host cells upon influenza B RNP nuclear egress (Dauber and Wolff, 2009).

The sequences flanking the coding regions are also unique to each genus, allowing for the genus to be identified by the first 5 nucleotides of each gene segment. Interestingly the first 9 nucleotides of each gene segment of influenza B viruses are conserved, AGCAGAAGA. Variability in the sequence only sets in at nucleotide position 10 onwards for each gene segment. With the exception of nucleotide 4, the first 12 nucleotides of each gene segment of influenza A viruses are conserved, AGCRAAAGCAGG, where R=G/A (Stoeckle et al., 1987). These sequences have been shown to play a crucial role in the initiation of transcription in influenza A studies (Engelhardt et al., 2005; Fodor, 2013).

![Gene Segment Diagram](image)

**Figure 1.11** Gene Segments of Influenza B viruses. The segments are arranged according to their segment numbers; PB2 is coded by segment 1 while NS1/NS2 by segment 8. The numbers flanking each gene segment denote the length of the entire segment in nucleotides while the numbers within each segment denote the length of the protein in amino acids. Segments 6, 7 and 8 each code for 2 different proteins. Taken from Fields Virology 5th edition
1.10 Influenza B Proteins

Table 1.4 lists the proteins of influenza B and their respective functions in the replication cycle. All of the proteins listed in table 1.4 are structural proteins, except for NS1. Despite being closely related, there exist certain proteins which are unique to influenza A or B viruses. Firstly, gene segment 2, which encodes the PB1 protein of both influenza A and B viruses, also encodes for PB1-F2 and PB1-N40 in influenza A viruses, which have not been detected in influenza B viruses. The PB1-F2 protein is coded by a gene in an alternate reading frame (+1) near the 5’ end of the PB1 gene. This protein has been characterized as pro-apoptotic and binds to result in the destabilization of the mitochondrial membrane of infected cells (Chanturiya et al., 2004; Chen et al., 2001; Palese, 2007; Zamarin et al., 2006). A recent study has shown an affinity of PB1-F2 to host cell IKKβ and also its ability to block NF-κB from binding to DNA, which results in an inhibition of its downstream processes (Reis and McCauley, 2013). The start codon of the PB1-N40 gene lies at codon 40 of the PB1 gene. While it has been classified as a nonessential protein, its over expression has been shown to be detrimental to viral growth in culture. The PB1-N40 protein has been shown to bind to the PB2 protein of the viral polymerase complex (Tauber et al., 2012; Wise et al., 2009).

In addition to encoding for the PA protein, the third gene segment of influenza A viruses also encode the PA-X protein which is a product of frame shifting in the +1 ORF (Shi et al., 2012). This 30kDa protein, functions to quell pathogenicity, almost antagonistic to the function of PB1-F2 (Jagger et al., 2012; Yewdell and Ince, 2012).

A spliced product in gene segment 7 of influenza A viruses, denoted as M42, encodes for a protein which has a transmembrane domain and functions to rescue viruses with defective M2 proteins (Wise et al., 2012). A third spliced mRNA yielding form gene segment 7 of influenza A viruses has been identified, but this mRNA has not been found to encode for a protein (Inglis and Brown, 1981; Lamb et al., 1981; Shih et al., 1998). These findings have not been made in influenza B viruses as yet.

The next difference lies in the NB protein of influenza B viruses, which lies in the -1 open reading frame of the NA protein (Shaw et al., 1983). Despite not having its function fully elucidated, there has not been the identification of an NB cognate in influenza A or C viruses (Palese, 2007).
Table 1.4 Influenza B proteins and their respective functions.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Function</th>
</tr>
</thead>
</table>
| PB2     | Component RDRP Complex  
         | Involved in mRNA cap recognition  
         | Initiates viral transcription |
| PB1     | Component RDRP Complex  
         | Has endonuclease activity  
         | Serves as a catalyst in transcription initiation  
         | Involved in RNA elongation |
| PA      | Component RDRP Complex  
         | Unknown function |
| HA      | Receptor Binding  
         | Fusion Protein |
| NP      | Involved in RNA Binding  
         | Involved in RNA synthesis |
| NA      | Cleaves glycosidic bonds of terminal sialic acid moieties |
| NB      | Unknown function  
         | Membrane Protein |
| M1      | Matrix Protein  
         | Interacts with viral RNP and viral surface proteins |
| BM2     | Proton channel |
| NS1     | Multifunctional Protein  
         | Interferon antagonist  
         | Inhibits PKR Kinase |
| NEP/NS2 | Exports vRNPS out of nucleus |

The functions of each of these proteins work throughout the replication cycle of the virus (Fig 1.12). The adsorption process is chiefly brought about by the HA protein’s ability to recognize and bind to sialic acid moieties present on the glycoproteins of the host cell. Binding of the HA proteins to the receptor triggers endocytosis, encapsulating the virus in an endosome. The endosome, the BM2 protein, present on the viral surface, facilitates the influx of protons into the virion from the endosomal lumen. This influx of protons triggers the HA protein to expose its fusion peptide, resulting in the fusion of the viral and endosomal membranes (Fusion & Uncoating Fig1.12). The vRNPS now enter the nucleus, facilitated by the nuclear localization signals present in the NP and the RDRP. The genome of the virus is both transcribed to mRNA and replicated to cRNA by the functions of the RDRP and cellular RNA polymerase II. Viral mRNA leaves the nucleus to be translated by ribosomes in the cytoplasm of the infected cell. Glycoproteins such as HA and NA get inserted in the membranes of the secretory pathway, undergoes post-translational
modification while being transported to the cellular membrane. Proteins such as the NP, PA, PB1, PB2, NS1 and NEP contain nuclear localization signals (NLS) and enter the nucleus to further amplify the generation of cRNA and mRNA. The NS1 functions to inhibit cellular responses which would be deleterious to the replication of the virus. The cRNA gets encapsidated with the NP protein and conjugated to the RDRP, forming the RNP complex. The matrix (M1) protein is believed to play an important role in the assembly of the virus. M1 binds to the vRNPs facilitating its dissociation with the nucleoplasm. This newly formed vRNP then leaves the nucleus via the assistance of the NEP, which contains a functional nuclear localization signal. The vRNP-M1 complex now assembles with the viral glycoproteins just prior to budding. The NA protein cleaves the glycosidic bonds of sialic acid which may be bound to the HA proteins, allowing for the HA protein to function in the proceeding infection (Palese, 2007). The characteristics and functions of the proteins studied in this thesis (HA, NA, NB, BM2, NS1, NEP) will be described further in the proceeding sections (Sections 1.10.1-1.10.6).

Figure 1.12 General schematic of influenza virus replication. Taken from (Palese, 2007).
1.10.1 Influenza B Hemagglutinin (HA) Protein

The influenza B hemagglutinin (HA) protein is coded for by gene segment 4, the gene and protein lengths of HA are listed in figure 1.11. The influenza B HA gene is 100 nucleotides longer than the influenza A HA gene (Desselberger, 1978; Krystal et al., 1982; Palese et al., 1980). The molecular weight of the HA protein is approximately 84kDa (Racaniello and Palese, 1979) and only shares 25% homology to its influenza A orthologue (Krystal, 1983; Krystal et al., 1982). The HA protein is the most abundant protein in the external surface of the virion. It is a type I integral membrane protein (N\textsubscript{out} C\textsubscript{in}) (Palese, 2007). The HA protein is translated as a precursor denoted as HA\textsubscript{0}. The influenza B HA\textsubscript{0} protein has 2 cleavage sites, the signal peptide is cleaved off from the first 15 amino acids from its amino terminus and the remainder of the protein is cleaved at the 360\textsuperscript{th} amino acid, maturing the HA\textsubscript{0} precursor protein into HA\textsubscript{1} and HA\textsubscript{2} fragments (Krystal et al., 1982). Figure 1.13 shows an alignment of the HA\textsubscript{0} precursor protein, highlighting the signal peptide and the 2 cleavage sites which yield the HA\textsubscript{1} and HA\textsubscript{2} fragments.

![Figure 1.13 Alignment of B/Lee/40 HA against A/PR/8/34 HA showing the conserved cleavage sites yielding the two functional peptides: HA\textsubscript{1} and HA\textsubscript{2}.](image)
Even though the HA proteins of influenza A and B viruses only share 25% homology, the structural features between them are rather conserved; their signal peptides, amino and carboxyl termini are hydrophobic and their maturational cleavage sites centre on a key arginine residue (Krystal et al., 1982) (Fig 1.13). The maturational cleavage sites of influenza B compared to influenza A (H1 and H3) are shown in table 1.5 below (Bianchi et al., 2005).

Table 1.5 Cleavage Recognition Site of the HA<sub>0</sub> from 3 different influenza strains.

<table>
<thead>
<tr>
<th>Virus/subtype</th>
<th>Strain</th>
<th>Sequence&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/H3/HA&lt;sub&gt;0&lt;/sub&gt;</td>
<td>Consensus</td>
<td>NVPEKQTR ↓ GIFGAIAGFIE</td>
</tr>
<tr>
<td>A/H1/HA&lt;sub&gt;0&lt;/sub&gt;</td>
<td>Consensus</td>
<td>NIPSIQS↓ GFLGAIAGFIE</td>
</tr>
<tr>
<td>B/HA&lt;sub&gt;0&lt;/sub&gt;</td>
<td>Consensus&lt;sup&gt;b&lt;/sup&gt;</td>
<td>PAKLKKER ↓ GFFGAIAGFIE</td>
</tr>
</tbody>
</table>

<sup>a</sup> The position of cleavage between HA<sub>1</sub> and HA<sub>2</sub> is indicated by the arrow.

<sup>b</sup> The consensus is the same for both the Victoria and Yamagata lineages.

Adapted from Bianchi et al., 2005

The signal peptide (Fig 1.13) targets the HA<sub>0</sub> precursor to intracellular membranous structures and the HA<sub>0</sub> precursor is cleaved by host cell proteases into its functional HA<sub>1</sub> and HA<sub>2</sub> fragments. Even though the protease which cleaves the HA<sub>0</sub> precursor has been determined in other animal systems (Goto and Kawaoka, 1998; Gotoh et al., 1992; Kido, 2007.; Lazarowitz, 1973; LeBouder, 2008), the human counterpart of this protease has only been identified recently; HAT (human airway trypsin-like protease) and TMPRSS2 (Bottcher-Friebertshauser et al., 2010; Böttcher et al., 2009; Bottcher et al., 2006; Chaipan et al., 2009). The studies identifying this protease were carried out on influenza A HA<sub>0</sub> precursor and it still needs to be confirmed if the same protease cleaves the influenza B HA<sub>0</sub> precursor.

The products of the maturational cleavage, HA<sub>1</sub> and HA<sub>2</sub>, are linked to each other by disulphide bonds forming a heterodimer. On the exterior of the virion, the HA exists as a trimer of these heterodimers. The HA<sub>1</sub> contains the receptor binding domain, the HA<sub>2</sub> contains the fusion peptide and the transmembrane domain at its carboxyl terminus (Earp et al., 2005). Figure 1.14 shows a schematic of the mature HA in its native state (Fig 1.14A) and at low-pH (Fig 1.14B).

The structure of influenza B HA protein has only recently been determined in 2007 and 2008 (Wang et al., 2008; Wang et al., 2007), thus not much is known about the
influenza B HA from a structural standpoint, compared to influenza A HA (Tung et al., 2004).

The main functions of the influenza HA protein include receptor binding and fusion of viral and host-cell membranes (Palese, 2007). The cellular receptors for influenza A and B viruses are terminal sialic acid residues of glycoproteins and glycolipids (Gottschalk, 1959; Palese, 2007; Wiley and Skehel, 1987). The HA protein of influenza A viruses bind to sialic acid moieties with either α(2,3) linkages or α(2,6). The HA proteins isolated from different hosts have different binding affinities to either of these linkages, influenza A viruses which infect humans have a higher affinity to α(2,6)-linked sialic acid receptors, influenza A viruses infecting avian and equine hosts bind stronger to α(2,3)-linked receptors while influenza A viruses infecting swine have strong affinities to both (Connor et al., 1994; Gambaryan et al., 2004; Gambaryan et al., 1997; Matrosovich et al., 1999; Matrosovich et al., 1997; Rogers, 1989; Rogers and Paulson, 1983; Skehel and Wiley, 2003; Wiley and Skehel, 1987). Specific glycan arrays which use various sialosides of both linkages as ‘bait’ have been employed to determine the binding preference of viruses. These arrays can also be used as a tool to differentiate between the different influenza A subtypes (Liao et al., 2010).

**Figure 1.14 HA pre and post fusion.**

This figure shows a general schematic of the structure of the mature HA protein of both influenza A and B viruses. The HA1 fragment is shown in gray. The HA2 fragment is composed of the blue coiled-coil, the green C-terminal helix, the red fusion peptide and the yellow transmembrane carboxyl domain. C and N refer to the carboxyl and amino termini of the HA2 fragment respectively.

The native conformation of the mature HA. The fusion peptide is hidden in this conformation, being clamped down buy the globular HA1 heads.

Once exposed to low pH, the globular HA1 separate and the fusion peptides are now exposed.

Adapted from Earp et al., 2005
Understanding that the influenza B viruses infect predominantly humans (Murphy, 1990; Sridharan et al., 2010), it would be intuitive to assume that the HA of influenza B viruses would have a stronger binding affinity to α(2,6)-linked sialic acid receptors, however reports are divided on this issue, suggesting that the HA protein of influenza B viruses might have a substantial binding affinity to the α(2,3)-linked sialic acid receptors as well (Gambaryan and Matrosovich, 1992; Gambaryan et al., 1995; Gambaryan et al., 1999; Gambaryan et al., 1997; Matrosovich et al., 1993; Suzuki, 1992; Xu et al., 1994).

A recent study has found that post egg adaptation, the HA protein of influenza B viruses would lose an N-linked glycosylation motif at position 196 or 197. Viruses which have this glycosylation motif can be passaged through MDCK cells well but not in eggs and its loss gives the HA protein greater affinity to α(2,3)-linked sialic acid receptors which are predominant in chicken embryo chorio-allantoic membrane cells (Chen et al., 2008b). This loss/gain of glycosylation also results in a change in the virus antigenicity. This is something that should be studied in depth as influenza B virus vaccines are generated in chicken eggs (Bright et al., 2007).

The second important function of the HA protein in influenza viruses is membrane fusion, which is pH activated. When the pH decreases, the HA protein adopts the conformation shown in figure 1.14B. The fusion peptide is now exposed and can insert into the endosomal membrane, while the transmembrane domain is still anchored in the virus envelope. This is then followed by hemifusion of the membrane and pore formation, allowing escape to the vRNPs into the cytoplasm as shown in figure 1.15 (Palese, 2007).

To understand which residues in the mature HA proteins of influenza B viruses are critical for hemifusion and pore formation, various mutational studies have been carried out. When the most carboxylic residue of the HA2 fragment, a leucine was substituted to a more hydrophilic residue, the ability to form syncytia was impaired. However, when this leucine was substituted to a more hydrophobic residue, syncytia formation was not impaired. However, when the leucine residue was deleted, full fusion of the membranes was observed. These observations led the authors to conclude that leucine at the end most position of the HA of influenza B viruses was not essential for syncytia and pore formation, but rather the hydrophobicity of this residue plays a crucial role in the formation (Ujike et al., 2006).
Acylation of specific cysteine residues of the HA protein is required for fusion to occur (Ujike et al., 2004, 2005; Wagner et al., 2005). However, the pattern and dependence of acylation for fusion differs among the different sub-types of influenza viruses (Veit, 2006) The HA protein from influenza B viruses are acylated at cysteine residues present in the cytoplasmic tail. The acylated cysteine residues were found to be dispensable in fusion but implicated in pore dilation (Ujike et al., 2004).

Several differences lie in the HA proteins of the 2 main lineages. Firstly, the main antibody neutralizing epitope of the Yamagata lineage lies in the ‘loop’ region (amino acids: 156-164) while the Victoria lineage’s neutralizing epitope lies in the ‘tip’ of the protein (amino acids: 178-185) (Nakagawa et al., 2005; Nakagawa et al., 2006).
The HA protein of the Yamagata lineage also contain a deletion of an asparagine at position 178 and ever since 2000, strains of the Victoria lineage have shown a new glycosylation site at position 211 (McCullers et al., 2004; Nakagawa et al., 2004).

1.10.2 Influenza B Neuraminidase (NA) Protein

The influenza B neuraminidase (NA) protein is the second most abundant protein in the virion shell. Unlike the HA protein, the NA protein is a type II integral membrane protein (Colman, 1994; Palese, 2007). The gene segment coding for NA is 1557 nucleotides long and the NA protein is composed of 466 amino acids (figure 1.11). The protein functions in a homo-tetramer complex, with a thin stalk inserted into the viral membrane supporting a box-like head (Air and Laver, 1989).

Even though the NA protein of influenza B viruses only share 25% amino acid sequences similarity with influenza A NA protein, the secondary structure of these two proteins are conserved (Air and Laver, 1989). Similarly, the active-site residues are both spatially and chemically conserved, suggesting strong functional similarity (Bossart-Whitaker et al., 1993). Interestingly, using a reverse genetics system, the NA from a strain of influenza B virus could replace the function of influenza A NA protein, suggesting that the NA function of both genuses are conserved (Ghate and Air, 1999).

NA proteins from both influenza A and B exist as a homo-tetramer and consists of a box-like head with a stalk which is anchored in the virus envelope (Air et al., 1989). By studying the structure of the influenza B NA, it was discovered that the NA protein of all influenza viruses not only contain a sialic acid binding site, but also 2 calcium ion (Ca\(^{2+}\)) binding sites. This Ca\(^{2+}\) binding site is believed to have a structural role as it pins an antigenic site, 316-350, into the active site of the protein (Burmeister et al., 1992).

The NA protein serves to cleave the \(\alpha-(2,3)\) and \(\alpha-(2-6)\) glycosidic links between the terminal sialic acid moieties of glycoproteins or lipids. This function of neuraminidase is crucial in ensuring successful viral release from infected cells and also to prevent self-aggregation of viral particles. Cells infected with temperature sensitive mutants with defects in their NA show, by electron microscopy, large aggregates of intact viruses at the cell membrane (Palese, 2007; Palese et al., 1974). This process of desialidation may occur at the trans-Golgi or the cell surface (Hanessian et al., 2002; Luo et al., 1999).
Neuraminidase Inhibitors (NAIs) are a class of widely used drugs for influenza. These inhibitors are sialic acid analogues which competitively inhibits binding to the neuraminidase active site. Inhibitor-bound neuraminidase cannot cleave the cellular-receptor bound to HA during viral release. This prevents successful viral shedding, discontinuing viral replication (figure 1.16).

![Figure 1.16 Schematic diagram showing the mechanism of NAI inhibition](http://www.flutrackers.com/forum/showthread.php?p=318949 Dec, 2013)

Mutations in the HA and NA genes can generate strains which are resistant to neuraminidase inhibitors (McKimm-Breschkin, 2000). In 1999, the Neuraminidase Inhibitor Susceptibility Network (NISN) was set up to monitor the use of Neuraminidase Inhibitors. NISN monitors the spread of resistant strains, possibly selected by the use of such drugs (Zambon and Hayden, 2001).

Neuraminidase Inhibitor (NAI) resistance is a real and palpable problem. Hatakeyama et al (2007) has reported the presence of resistant strains in individuals who have not been treated by any antiviral, suggesting that resistant strains may be transmitted from a treated individual to an untreated individual. But there have been cases of resistant strains isolated from patients in Australia, where the use of neuraminidase inhibitors are virtually close to none. This suggests spontaneous mutations are generating these strains (Hatakeyama et al., 2007; Moscona and McKimm-Breschkin, 2007).
1.10.3 Influenza B NB Protein

Gene segment 6 of influenza B viruses is bicistronic, coding for 2 proteins; neuraminidase (NA) and NB, the sizes of the genes and proteins are reflected in figure 1.8. The NB gene starts in a -1 open reading frame, 7 nucleotides 5’ of the start codon of the NA gene, as shown in figure 1.17 (Shaw et al., 1983).

The NB protein is unique to influenza B which has no determined function yet (Shaw et al., 1983). This protein is a type III integral membrane protein and is found abundant on the surface of virus infected cell as well as in the packaged virion (Betakova et al., 1996; Brassard et al., 1996). Increasing evidence seems to suggest that the NB protein is an ion channel that functions similar to the M2 protein of influenza A viruses (Burmeister et al., 1993; Fischer et al., 2000b; Pinto and Lamb, 2006). When placed into an artificial membrane, the NB protein forms ion channels which selectively allow passage to cations (Sunstrom et al., 1996). This activity was inhibited by the same drug inhibiting the M2 ion channel of influenza A viruses (Fischer et al., 2001).

It has been found that the mutants lacking the NB protein can replicate sufficiently well as wild type viruses, but are less pathogenic in mice (Hatta and Kawaoka, 2003). This further adds to the mystery of the function of the NB protein. It is unsure if it serves as an accessory or vital protein in the replication and pathogenicity of the virus.

1.10.4 Influenza B BM2 Protein

Gene segment 7 of influenza A and B viruses is bicistronic, coding for the Matrix (M1) protein as well as the M2 proton channel of both viruses (Fig 1.18).
Figure 1.18 Genome structure of segment 7 (black lines) from influenza A and B virus and the proteins (coloured boxes) encoded respectively. The numbers flanking the black lines represent the length of the respective genes. Numbers above the coloured boxes represent the number of amino acids making up each of the proteins. (A) genome structure of influenza A segment 7. (B) Genome structure of influenza B segment 7, highlighting the Stop-Start Pentanucleotide.

The start codon of the BM2 gene overlaps with the stop codon of the M1 gene in influenza B viruses (Horvath, 1990). This is different from influenza A, where the A/M2 protein is a product of splicing of the mRNA transcript of segment 7 (Fields et al., 2006). Despite having a different method of translation than that of influenza A viruses, this stop-start pentanucleotide sequence is also seen in other viruses, such as in the expression of VP2 in feline calicivirus (Luttermann and Meyers, 2007). Studies have shown that the last 45 nucleotides of the B/M1 gene are responsible for promoting the expression of the BM2 protein. This is displayed by inserting any gene downstream of the last 45 nucleotides + the stop-start sequence resulting in its successful expression (Hatta et al., 2009; Powell et al., 2008).

The BM2 protein is a Type III membrane protein of 109 amino acids, with the N terminus on the outside of the membrane. 5 amino acids are present on the ectodomain, 20 amino acids span the transmembrane region and the last 84 amino acids of the C-terminus are intracellular (Cross, 2009; Wang et al., 2009) (Fig 1.19). The BM2 protein
is a structural protein and expressed at the cell surface of infected cells (Paterson et al., 2003).

Structural and functional analyses of the BM2 protein have revealed that it functions as a tetramer, similar to A/M2 (Paterson et al., 2003; Pinto and Lamb, 2006). However, while the A/M2 exists as a tetramer of 2 di-sulphide linked homodimers held together by non-covalent interactions (Sugrue and Hay, 1991), no evidence of di-sulphide bridging has been found in the BM2 tetramer. Instead, the transmembrane region has been found responsible for the oligomerisation of the BM2 monomers (Balannik et al., 2008). Further elucidation has to be carried out to determine how the BM2 oligomer is stabilized.

Viruses with the BM2 gene knocked out could not replicate well in cell culture and viruses could only replicate when infected in cells stably expressing BM2 protein, suggesting that the BM2 protein is crucial for the successful replication of influenza B viruses (Imai and Odagiri1, 2004). Sedimentation experiments also prove that the BM2 protein is a structural protein, existing as a tetramer in the virus (Holsinger et al., 1994; Odagiri et al., 1999).

Just like the influenza A M2, BM2 functions as a proton channel, facilitating the influx of protons at the onset of uncoating in early infection (Palese, 2007). The
functional motif responsible for proton conductance is HXXXW which is conserved in both influenza A and B (Fig 1.20) (Ma et al., 2008; Pinto and Lamb, 2006).

![Amino acid alignment of the transmembrane domain of the M2 protein from influenza A and B. Red box shows the conserved functional motif HXXXW. Taken from (Ma et al., 2008)](image)

Both the histidine and tryptophan residues have their side chains facing the aqueous pore of the protein. The side chain of the tryptophan pore lies perpendicular to the length of the protein. This creates a closed ‘gate’, preventing the entry of protons. When the histidine residue gets protonated, the protein undergoes a conformational change, resulting in a rotation of the side chain of tryptophans becoming parallel to the ion pore. This now creates the open ‘gate’ conformation, allowing the influx of protons (reviewed in (Pinto and Lamb, 2006)) (Fig 1.21)

![Schematic representations of influenza M2 oligomers. For clarity, the HXXXW motifs of only 2 monomers are shown. Red: Histidine, Blue: Tryptophan. Top panel: when in high pH, histidine is not protonated, the tryptophan side chains block proton influx. Bottom panel: when in low pH, histidine gets protonated and tryptophan side chains rotate, allowing proton influx. Taken from (Pinto and Lamb, 2006)](image)

Amantadine, a M2 inhibitor is a common drug used to treat influenza infections. However, it has been noted that amantadine does not treat influenza B infection (Davies et al., 1964) as it does not inhibit the BM2 protein. It has been suggested that the nitrogen atom of amantadine forms a hydrogen bond with the unprotonated imidazole
histidine. This prevents the histidine from being protonated when in low pH (Mould et al., 2003a). By comparing the residues lining the pore, it can be observed that BM2 has 2 more serines than influenza A. It is possible that the more polar environment of BM2 reduces the affinity of amantadine to form a hydrogen bond with the histidine (Paterson et al., 2003; Pinto and Lamb, 2006) (Fig 1.20). This suggestion is backed up with the finding that influenza A M2 bearing the A30T mutation is resistant to amantadine (Holsinger et al., 1994).

1.10.5 Influenza B Non Structural Proteins 1 (NS1)

The NS1 protein is coded for by gene segment 8. Its gene is 843 nucleotides long, coding for a protein of 281 amino acids (Fig 1.11) (Palese, 2007). The NS1 protein is the only non-structural protein in influenza B. The different domains of the NS1 protein are shown in figure 1.22.

![Figure 1.22 Schematic representation of influenza B NS1 protein.](image)

The protein is divided into 3 domains: N-terminal RNA binding domain (RBD) of 90 amino acids, the C-terminal effector domain (ED) of 161 amino acids and a linker domain between the two of 30 amino acids (Guan et al., 2011a) (Fig 1.22).

Various reports have noted that NS1 localises to the nucleus in early infection and migrates to the cytoplasm in late infection (Norton et al., 1987; Schneider et al., 2009). The influenza B NS1 protein has 1 known nuclear localization signal (NLS); strongly basic residues between positions 46-56 (Schneider et al., 2009). A second NLS has been identified in influenza A NS1 proteins at the C-terminal end of certain strains. This second NLS has been denoted as the Nucleolar Localisation Signal (NoLS), targeting the influenza A NS1 to the nucleolus in early infection (Melen et al., 2007; Volmer et al., 2010). No known NoLS has been determined for influenza B NS1 as of now.

The NS1 protein functions in a homodimer complex and has been found to bind to single and double stranded RNAs. The RBD has been found to be crucial for this dimerization of influenza B NS1 (Wang and Krug, 1996). The structure of the first 120 amino acids of the influenza B NS1 protein has been solved and it reveals that the RBD
is composed of three alpha helixes in tandem (Li et al., 2011). The ability of the NS1 protein to bind to RNA has been attributed to the arginines and lysines present in the second helix of the RBD (Donelan et al., 2004).

An interesting observation of NS1 nuclear localisation in early infection has been reported by Schneider et al. In their study, it was found that within 6 hours of infection, the NS1 protein localises in nuclear speckle domains, this localisation is lost in late infections (Fig 1.23). This phenomenon is also not seen in influenza A NS1. It was found that NS1 protein co-localises with host cell protein SC35 which is an important player in the nuclear export mechanism (Fig 1.23).

![Figure 1.23 Localization of NS1 protein in nuclear speckle domains in early infection, co-localizing with host cell SC35. Taken from (Schneider et al., 2009)](image)

This protein has been shown to inhibit type I interferons (IFNs) and PKR kinase. Both these functions are carried out by the RBD (Dauber et al., 2004; Donelan et al., 2004). These functions are also displayed by the influenza A NS1, even though the 2 proteins share less than 25% amino acid similarity (Bergmann et al., 2000; Mibayashi et al., 2007; Min et al., 2007). Interestingly both the RBD and effector domains (ED) can individually inhibit the translocation of IRF3 into the nucleus upon infection (Donelan et al., 2004).

Even though the NS1 proteins of both influenza A and B viruses are interferon antagonists (Palese, 2007; Talon et al., 2000), they have different mechanisms in meeting this function. Unlike the NS1 of influenza A, the NS1 of influenza B binds to ubiquitin-like ISG15 proteins of the host cell and prevents this protein from conjugating with its target proteins, resulting in an inhibition of interferon release (Sridharan et al., 2010; Yuan and Krug, 2001). In influenza A infection, little ISG15 protein is present and no evidence of its binding to influenza A NS1 has been reported.
Influenza B NS1 mutants have attenuated pathogenicity and cannot replicate in infected cells (Hai et al., 2008).

A recent publication has suggested that the NS1 protein might play a role in the strict host specificity of influenza B viruses. NS1 proteins of both influenza A and B bind to ISG15, a host cell protein which has the ability to trigger an immune response against the virus. It was found that influenza B NS1 binds only to the ISG 15 protein of human and primate origin. Influenza B NS1 cannot bind to mouse or canine ISG15. This suggests that ISG15 may be active in influenza B infection in mice and canines, resulting in an immune eradication of the infection (Sridharan et al., 2010).

**1.10.6 Influenza B Nuclear Export Protein (NEP/NS2)**

The eighth gene segment of the influenza B genome codes for both NS1 and NEP proteins (figure 1.11). The NEP protein is translated from a spliced mRNA of segment 8. The spliced mRNA of the NEP protein contain the same 75 nucleotides present in the 5’ end of the NS1 mRNA but the remaining 350 nucleotides of the NEP mRNA can be translated in the +1 reading frame. This finding is rather similar with the mechanism for NS1 and NEP expression in influenza A (Briedis and Lamb, 1982) (Fig 1.24).

![Figure 1.24](image)

**Figure 1.24** Gene arrangements of NS1 and NEP in gene segment 7 of influenza A and B viruses. Taken from (Briedis and Lamb, 1982)

Despite being initially named non-structural, the NEP protein has been found in complex with the matrix protein in the packaged virion, (Richardson and Akkina, 1991;
These findings have prompted the name of the NS2 protein be changed to reflect its function; Nuclear Export Protein (NEP). This protein has been named as such because of its role played in ensuring successful egress of vRNPs from the nucleus to cytoplasm (Gomez-Puertas et al., 2000; Neumann et al., 2000; Neumann et al., 1994; O'Neill et al., 1998; Yasuda et al., 1993). Upon influenza B infection, the NEP protein can be detected in the nucleus of the infected cell through immunofluorescence, as early as 5 hours post infection (hpi) in the nucleus, at 9hpi the signal is strong in the nucleus and NEP can also be detected in the cytoplasm. And at 14hpi NEP can be detected throughout the entire cell at equal intensity. This reflects the function of the NEP which is to export vRNPs out of the nucleus to the cytoplasm for assembly (Imai et al., 2003).

Studies done in detergent-disrupted virions show that NEP was associated with the vRNPs and the M1 protein. This finding is different in influenza A virions where the NEP is found in close association with M1 protein only. This finding suggests that the NEP protein of influenza A binds only to M1, which is in association with the vRNPs, while the influenza B NEP is in direct association to both the vRNPs and the M1 protein (Imai et al., 2003).

1.11. Study Outline & Objectives

Throat swabs of SAF personnel which have been tested positive for influenza B upon collection were inoculated onto MDCK cells to test for viral viability. vRNA was then extracted from the infected cells to serve as template for cDNA synthesis. Individual gene segments were then PCR amplified from the cDNA, and were sequenced. The consensus sequences generated were aligned with vaccine strains, ancestor strains and reference strains for sequence and phylogenetic analysis. After attaining the sequences from the clinical specimens, representative sequences are cloned into a mammalian expression vector and these proteins are expressed in cell culture. By transfecting mammalian cells with these sequences, the cellular localization can be observed through immunological staining. Western blot analysis of these sequences can reveal the sizes of these proteins. Differences in the cellular localization and migration patterns of these proteins can be attributed to the differences in the sequence of these proteins, allowing greater insight into the functions of individual domains of these proteins.
This study aims to identify the lineages of the influenza B strains circulating locally and compare them with the lineage of regionally and globally circulating strains. Phylogenetic analysis of gene segments 4, 6 and 8 were carried out to determine their respective lineages. Knowing the lineages these segments would allow us to determine the reassortment pattern of circulating influenza B viruses over the period of study. The lineage identity of these viruses when compared to those circulating globally would allow us to determine the effectiveness of the vaccine strains put forth for the period in question. The proportions of each lineage circulating may also give us a look into the evolutionary kinetics of influenza B viruses, allowing better preparedness for the following influenza pandemic.

Sequence alignment, at both the gene and protein levels would allow detailed inferences of the evolutionary mechanisms employed by influenza B viruses. Mutations in the NA and HA genes which have been implicated in Neuraminidase Inhibitor Resistance can be identified in these locally circulating strains, giving us greater insights in the transmission of such strains. Amino acid substitutions in specific domains were compared to other locally and globally circulating strains to ascertain the prevalence of these mutations, ultimately shedding light on the function of these amino acids. Domains examined included known glycosylation sites, antibody recognition sites, enzyme active sites of the HA and NA, the transmembrane domain of the NB and the BM2 as well as RNA-binding of the NS1 proteins.

The cloned NB protein was examined for its cellular localisation, its post translational modification pattern as well as its oligomerisation status. Similar analysis of the BM2 protein was made. The NS1 protein was analysed extensively to determine the kinetics involved in its cellular distribution and if any differences in sequence would result in a redistribution of cellular localisation. A novel peptide was expressed in 4 out of the 6 cloned NS1 constructs. Mutational analysis and individual domain expression of the NS1 protein was then carried out to determine the characteristics of this previously undiscovered peptide.
Chapter 2: Materials & Methods

2.1 General Reagents

Stated in this section are lists of general reagents used for the experiments carried out in the following chapters. For work involving sterility, reagents are either autoclaved at 121°C for 20 min or passed through a 0.22 µm filter (Nalgene).

2.1.1 Cells

Mammalian Cells

Madin Darby Canine Kidney (MDCK) ATCC
Human Adenocarcinomic Lung Cells (A549) ATCC
Human Embryonic Kidney 293T (HEK 293T) ATCC
Embryonated Chicken Eggs Chew’s Chicken Farm
Chicken Embryo Fibroblasts Harvested from Embryonated Chicken Eggs

Bacterial Cells

EZ Competent Cells QIAGEN Ltd.
XL1-Blue Supercompetent Cells Stratagene®

2.1.2 Viruses

A/WSN/33 (H1N1) ATCC
A/Puerto Rico/8/34 (H1N1) ATCC
A/Duck/Malaysia/01 (H9N2) AVA
B/Lee/40 ATCC
A/Singapore/471/2009 (H1N1) DSO National Laboratories

2.1.3 Tissue Culture Reagents

DMEM+GlutaMAX Gibco
Fetal Bovine Serum (FBS) Gibco
2.1.4 Antibodies

mAb 8257 Mouse anti-influenza A Nucleoprotein  Chemicon
mAb 8259 Mouse anti-influenza B Nucleoprotein  Chemicon
Ab8984 Mouse anti-Lamin A + C  Abcam
Ab10530 Mouse anti-Nucleophosmin  Abcam
Ab11826 Mouse anti-SC35  Abcam
Ab24609 Mouse anti-Nuclear Pore Complex Proteins  Abcam
Sc-5621 Mouse anti-PML  Santa Cruz
Sc-1616 Rabbit anti-Actin  Santa Cruz
F7425 Rabbit anti-FLAG epitope  Sigma
Ab18230 Mouse anti-DDDK epitope (FLAG)  Abcam
mAb 9B11 Mouse anti-Myc epitope  Cell Signaling
Ab9106 Rabbit anti-Myc epitope  Abcam
IgG2a Mouse Anti-His  Amersham Biosciences
Rabbit Anti-B/NS1  Gift From Prof Thorsten Wolff (Dauber et al., 2004)
AP124 F Goat Anti-Mouse Fluorescein Conjugated  Chemicon International
AP307F Goat Anti-Rabbit Fluorescein Conjugated  Chemicon International
2.1.5 Immunofluorescence and Immunostaining Reagents

Paraformaldehyde  
Sigma Aldrich

10X Phosphate Buffered Saline pH 7.2  
1st Base

Triton X-100  
Chemicon

2.1.6 Commercially Available Kits

RNAeasy Mini Kit  
Qiagen

Gel Extraction Kit  
Qiagen

QIAprep Spin Miniprep Kit  
Qiagen

QIAGEN Plasmid Midi Kit  
Qiagen

QIAGEN PCR Cloning Plus Kit  
Qiagen

SuperscriptTM First Strand cDNA synthesis  
Invitrogen

QuikChange Site-Directed Mutagenesis  
Stratagene

NE-PER Nuclear and Cytoplasmic Extraction  
Thermo Scientific

2.1.7 Tissue Culture Propagation & Infection Media and Reagents

Cell Propagation Medium  
DMEM+GlutaMAX (Gibco) supplemented with 10% v/v Fetal Calf Serum (Gibco) and Penicillin and Streptomycin (Gibco)

Infection Medium for single infection cycle  
DMEM+GlutaMAX (Gibco) supplemented with 2% v/v Fetal Calf Serum (Gibco) and Penicillin and Streptomycin (Gibco)

Infection Medium for multiple infection cycle  
DMEM+GlutaMAX (Gibco) supplemented with 0.21% BSA (Gibco), 1μg/ml trypsin-TPCK (Worthington, USA) and Penicillin and Streptomycin (Gibco)
Fixing Solution
4% Paraformaldehyde in PBS

Permeabilising Solution
0.5% Triton-X in PBS

Conventional Plaque Overlay
DMEM (Gibco) and 1% Low Temperature Agarose (Invitrogen), 0.21% BSA (Gibco) and 1μg/ml trypsin-TPCK (Worthington, USA)

2.1.8 Bacterial Culture Media and Solution

100mg/ml Ampicillin
100mg in 1ml of ddH2O, filtered (0.22μm) and stored in -20°C

0.1M IPTG
23.83mg of IPTG (Fermentas) dissolved in 1ml of ddH2O

20mg/ml X-Gal
20mg of X-Gal (Fermentas) dissolved in 1ml of DMSO

LB-Amp Broth
LB medium (Difco™), 100μg/ml ampicillin

LB-Amp Agar Plates
LB Agar (Difco™), 100μg/ml ampicillin

LB-Amp Agar Plates (blue/white screening)
LB Agar (Difco™), 100μg/ml ampicillin, 40ul of 20mg/ml X-Gal, 40ul of 0.1M IPTG

NZY® Broth
10g of NZ amine, 5g of yeast extract, 5g of NaCl, dH2O added to 1 litre, pH adjusted to 7.5 with NaOH.
Autoclave
Post-Autoclave, add:
12.5ml of 1M MgCl₂, 12.5ml of 1M MgSO₄, 20ml of 20% (w/v) glucose

Tris-NaCl (TN) Buffer
50mM Tris, 50mM NaCl

Bacterial Lysis Buffer
0.2mM PMSF, 200μg/ml Lysozyme, in TN Buffer

2.1.9 DNA Analysis

1xTBE
100ml of 10X TBE dissolved in 900ml of ddH₂O to make 1 X TBE stock
0.1% Ethidium Bromide  
10mg of Ethidium Bromide dissolved in 10ml of ddH₂O

1% Agarose Gels  
0.5g of Agarose in 50ml of 1XTBE

**2.1.10 Protein Analysis by SDS-PAGE gel**

10% Resolving Gel  
1.33ml of 30% Bis-Acrylamide, 1.583ml ddH₂O, 1ml 1.5M Tris-HCl pH 8.8, 40µl 10% SDS, 40µl of 10% APS

12% Resolving Gel  
1.6ml of 30% Bis-Acrylamide, 1.4ml ddH₂O, 1ml 1.5M Tris-HCl pH 8.8, 40µl 10% SDS, 40µl of 10% APS

15% Resolving Gel  
2ml of 30% Bis-Acrylamide, 0.94ml ddH₂O, 1ml 1.5M Tris-HCl pH 8.8, 40µl 10% SDS, 40µl of 10% APS

4% Stacking Gel  
0.2ml of 30% Bis-Acrylamide, 0.9ml ddH₂O, 0.378ml 0.5M Tris-HCl pH 6.8, 15µl 10% SDS, 15µl of 10% APS

5X Laemmli Buffer  
31.25ml 1 M Tris-HCl pH 6.8, 10g SDS, 25ml Glycerol, 0.75ml Bromophenol Blue, 2% ethanol, 5ml 2-mercaptoethanol to 100ml of ddH₂O

1X Laemmli Buffer  
Dissolving 5X Laemmli Buffer with ddH₂O in 1:5 ratio

SDS-PAGE Running Buffer  
57.6g Glycine, 12g Tris base, 4g SDS to 4 litres of ddH₂O

**2.1.11 Western Blotting Reagents**

Transfer Buffer  
3.03g Tris base, 14.41g Glycine, 200ml Methanol to 1 litre of ddH₂O

PBST  
1x PBS, 0,05% Tween

5% Skim Milk Blocking Solution  
1g Skim Milk powder mixed in 20ml of 1ml of PBST

**2.1.12 Coomassie Blue Staining Reagents**

Coomassie Blue Stain  
0.1% Coomassie® Brilliant Blue R-250 (AppliChem), 20% Methanol, 10% Acetic Acid

Destaining Buffer  
45% Methanol, 10% Acetic Acid.
2.1.13 SYPRO Ruby Staining Reagents
SYPRO Ruby Fixing Solution  50% Methanol, 7% Acetic Acid
SYPRO Ruby Wash Solution  10% Methanol, 7% Acetic Acid
SYPRO Ruby Protein Stain  Molecular Probes

2.1.14 Reagents for Immunoprecipitation
RIPA Buffer  20mM Tris-HCl (pH 7.5), 1% NP40, 0.1% SDS, 150mM NaCl, 1mM EDTA, 2mM Lysine, 2mM PMSF (added prior to experiment)

Binding Buffer  2.5ml of NP40, 1ml of 0.5M EDTA made up to 50ml of PBS

Low Salt (Wash) Buffer  1mM EDTA, 1% Triton-X-100, in 1X PBS.

2.2 Clinical Specimen Sequencing

2.2.1 Reference Strain
B/Lee/40 virus was purchased from American Type Culture Collection (ATCC), catalogue number VR101. The virus was stored in 1ml aliquots in -80°C. The infection and propagation of this virus was carried out according to manufacturer’s instructions. This strain was used in this study as a reference strain.

2.2.2 Clinical Specimens
Virus samples were collected from medical centres of various SAF camps between 2004 and 2009. Only individuals presenting a fever of greater than 38°C (oral temperature) and acute respiratory symptoms; cough and/or sore throat were included in this study. Oropharyngeal swabs of these patients were obtained and resuspended in viral transport medium (VTM) (Seah et al., 2010a). These specimen are named according to the following format; DSO_(numerical code)_(year of collection).

2.2.3 Tube Cultures of Viral isolates
MDCK cells were grown in 37°C, in flat-sided tube cultures (Nunclon™Δ Surface, Cat. No. 156758) till confluent. VTM were inoculated onto confluent MDCK monolayers for 1 hour in 33°C to allow for viral absorption. Cultures were then incubated in 33°C in DMEM supplemented with 0.25% BSA, 2µg/ml Trypsin TPCK
and penicillin/streptomycin. Cultures were observed daily, cells and supernatants were harvested once extensive cytopathic effect (CPE) or after 14 days of incubation.

Once CPE was observed, or after 14 days of infection, the cells were scraped in the supernatant spun at 2500rpm for 15 minutes. The supernatants were stored in 3x500µl aliquots at -80°C. Cells were resuspended in 1ml of sterile Phosphate Buffered Saline (PBS) and stored in 2x500µl aliquots and stored in -80°C.

2.2.4 Shell Vial Assays

MDCK cells were grown on 12mm glass coverslips in 24 well-plates till 60-80% confluency in 37°C. These cells were washed in sterile PBS and were inoculated with 200 µl of the supernatants harvested from the tube cultures described above. The plates were spun at 2000rpm at 25°C for 45 minutes, using a Beckman Coulter centrifuge. Cultures were then incubated in 33°C in 1ml of DMEM supplemented with 0.25% BSA, 2µg/ml Trypsin TPCK and penicillin/streptomycin.

Cells were observed daily and were fixed once CPE was observed. If no CPE was observed, the cells were fixed and stained after 6 days of infection, similar to the method employed for conventional coverslip immunofluorescence.

2.2.5 Egg Passaging

Twelve day old embryonated eggs were collected from Chew’s Agricultural Farm (Singapore). Each egg was inoculated with 200µl of supernatant harvested from the viral isolation process and incubated in 37°C for 2 days. Harvested allantoic fluid were clarified at 2500rpm for 15 minutes at 4°C and stored in 1ml aliquots in -80°C. Virus infectivity was assayed using the shell vial method described above.

2.2.6 Diagnostic Immunofluorescent Assays

The resuspended cells were spotted onto multi-well microscopy slides and air dried. The cells were fixed and permeabilized in acetone and stained for 1 hour using conjugated antibodies against influenza B (LIGHT DIAGNOSTICS™ Respiratory Virus Panel DFA, 3110). Cells were washed 3 times with PBS and viewed under a fluorescence microscope.
2.2.7 Extraction of Influenza B Virus vRNA

Viral RNA (vRNA) was extracted from the harvested cells using the RNeasy minikit (Qiagen, Inc., Valencia, CA, USA) according to manufacturer’s instructions. The extracted vRNA was stored in -80°C until real time PCR analysis.

2.2.8 Sequencing Primers

Table 2.1 lists all the primers used in this study. In house designed primers were designed using an alignment of the HA, NA and NS gene segments of influenza B isolated in Singapore previously. Primers were designed using BioEdit Software.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>5’ Sequence 3’</th>
<th>Size</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diagnostic Primers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FluB 970F*</td>
<td>AAATACGGTGATTAACAAAGACAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FluB 1131R*</td>
<td>AGCTCCGAAGAAACCCCTTTC</td>
<td>161 bp</td>
<td>HA</td>
</tr>
<tr>
<td>FluB 1004 Probe*</td>
<td>CACCATATTGGGCAATTTCCATG GC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sequencing Primers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UniB`</td>
<td>AGC AGA AGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA 30F`</td>
<td>CTA CTC ATG GTA GTA ACA TCC</td>
<td>709bp</td>
<td></td>
</tr>
<tr>
<td>HA 749R`</td>
<td>YGG GAA GCC ACC AAT CTG AGA AAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA 471 F`</td>
<td>ACC TCA GGA TCT TGC CCT AAC G</td>
<td>689bp</td>
<td></td>
</tr>
<tr>
<td>HA1169R`</td>
<td>TGT GTA TCC GTG CCA ACC TGC AAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA 999F`</td>
<td>AAA GCC ATA GGA AAT TGCCCA</td>
<td>841bp</td>
<td>HA</td>
</tr>
<tr>
<td>HA 1840R`</td>
<td>TCA ATA ACG TTT CTT TGT AAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HA F§</td>
<td>TAT TCG TCT CAG GGA GCA GAA GCA GAG CAT TTT CTA ATA TC</td>
<td>1840bp</td>
<td></td>
</tr>
<tr>
<td>HA R§</td>
<td>ATA TCG CTC CGT ATT AGT AGT AAC AAG AGC ATT TTT C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA 21FΔ</td>
<td>GCT ACC TTC AAC TAT ACA AAC G</td>
<td>547bp</td>
<td></td>
</tr>
<tr>
<td>NA 568R½</td>
<td>TAC CAT CAT GCC ATG CGG A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA 361F½</td>
<td>GCT CCC TTG ATA ATA AGG GAA CC</td>
<td>477bp</td>
<td></td>
</tr>
<tr>
<td>NA 838R½</td>
<td>ATG TTC TAC TCT TCC TGT TGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA 716F½</td>
<td>GGG GGA RAT TGT TAT CTT ATG</td>
<td>789bp</td>
<td>NA</td>
</tr>
<tr>
<td>NA 1505R½</td>
<td>TTT CAG AAA CAA TTA AKT TCA GTA AGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA F·</td>
<td>TAT TCG TCT CAG GGA GCA GAA GCA GAG CA</td>
<td>1505bp</td>
<td></td>
</tr>
<tr>
<td>NA R·</td>
<td>ATA TCG CTC CGT ATT AGT AGT AAC AAG AGC ATT TT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NS 376R½</td>
<td>CTG GTG TTG AAG GGT AAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NS 700R½</td>
<td>AT CTT CTT CAT CCT CCA CTG TAA</td>
<td></td>
<td>NS</td>
</tr>
<tr>
<td>NS F*</td>
<td>TAT TCG TCT CAG GGA GCA GAA GCA GAG GAT TTT ACG C</td>
<td>1053bp</td>
<td></td>
</tr>
<tr>
<td>NS R*</td>
<td>ATA TCG TCT CGT ATT AGT AGT AAC AAG AGG ATT TTT AT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1 Primers used for sequencing in this study. bp: base pairs

* Taken from Zou et al, 1997
½ Taken from Hoffmann et al, 2000
□ In-house designed primers
Δ Taken from Tsai, 2006
* Taken from Krafft et al, 2005
2.2.9 Real Time PCR detection of Influenza B Virus.

The extracted vRNA served as template for diagnostic real time PCR as described by Krafft et al. (Krafft et al., 2005). Primers used in this diagnostic PCR are listed in table 2. The reactions were carried out in a Lightcycler 1.5 (Roche Diagnostics, Manheim, Germany).

2.2.10 cDNA Synthesis

Complementary DNA synthesis was carried out for the specimens which were positive in either or both of the diagnostic tests mentioned above, using the kits; Transcriptor First-Strand cDNA synthesis System for RT-PCR (Roche) with the reverse transcription step carried out at 65°C for 30 minutes by Invitrogen’s SuperScript II First-Strand cDNA Synthesis Kit. UniB primer (table 2) (Zou, 1997) with the reverse transcription step carried out at 42°C for 1 hour. Other than the modifications mentioned, all other steps were carried out according to manufacturer’s instructions.

2.2.11 Hemagglutinin (HA) Gene Sequencing Strategy

Using Invitrogen’s Platinum Taq Polymerase High Fidelity kit (Cat. No 11304-029) the full HA gene segment was amplified out using primers HA F and HA R (Table 2.2) (Hoffmann et al., 2002). The cDNA synthesized previously served as template for this step. Touchdown PCR thermocycling conditions from table 2.2 were used to amplify the HA gene segment. All the PCR conditions in this study were touchdown PCR, with adjustments to the annealing temperatures, and extension times.

<table>
<thead>
<tr>
<th>Initial Denaturation</th>
<th>94°C</th>
<th>2 min</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>94°C</td>
<td>30 sec</td>
</tr>
</tbody>
</table>

**Touchdown**

<table>
<thead>
<tr>
<th>11cycles</th>
<th>55°C</th>
<th>30 sec</th>
<th>-1°C/ cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>68°C</td>
<td>2 min</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Denaturation</th>
<th>94°C</th>
<th>30 sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annealing</td>
<td>48°C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Extension</td>
<td>68°C</td>
<td>2 min</td>
</tr>
<tr>
<td>Final Extension</td>
<td>68°C</td>
<td>10 min</td>
</tr>
<tr>
<td>Holding</td>
<td>12°C</td>
<td>forever</td>
</tr>
</tbody>
</table>

*Table 2.2 Touchdown PCR thermocycling conditions*

The HA gene segment was then further amplified into 3 overlapping fragments as shown in figure 2.1. The primers used fragments 1, 2 and 3 were; HA30F/HA749R,
HA471F/HA1169R and HA999/HA1840R (table 2.1) respectively. The thermocycling conditions for the amplification of all three of these fragments were similar to that in table 3 but with an extension step of 1min. Invitrogen’s Platinum Taq Polymerase High Fidelity kit (Cat. No 11304-029) was used for this step as well.

HA/NA

![Diagram of strategy used to sequence HA/NA gene segments](image)

Figure 2.1 Schematic of strategy used to sequence HA/NA gene segments

### 2.2.12 Neuraminidase (NA) Gene Sequencing Strategy

Using Invitrogen’s Platinum Taq Polymerase High Fidelity kit (Cat. No 11304-029) the full NA gene segment was amplified out using primers NA F and NA R (Table 2.1) (Hoffmann et al., 2002). The cDNA synthesized previously served as template for this step. Touchdown PCR thermocycling conditions from table 2.2 were used to amplify the NA gene segment.

The NA gene segment was then further amplified into 3 overlapping fragments as shown in figure 2.1. The primers used for fragments 1, 2 and 3 were; NA21F/NA568R, NA361F/NA838R and NA716F/NA1505R (Table 2.1) respectively. The thermocycling conditions for the amplification of NAF1, NAF2 and NAF3 fragments required annealing temperatures 45°C, 58°C and 50°C respectively and with an elongation time of 1min.

### 2.2.13 Non Structural (NS) Gene Sequencing Strategy

The NS gene segment was amplified from the synthesized cDNA by using primers NS F and NS R (Table 2.1) (Hoffmann et al., 2002). The thermocycling conditions required for NS gene segment amplification required an annealing temperature of 62°C and an elongation temperature of 72°C for 10 minutes.

The NS gene segment was sequenced directly from the amplified gene segment using primers NSF, NSR as well as NS700R and NS376R (Table 2.1). A schematic diagram of the sequencing plan of the NS gene segment is shown in figure 2.2.
2.2.14 Gel Electrophoresis and PCR Product Extraction.

Five microlitres of PCR product were initially analysed on a 1% agarose gel stained with ethidium bromide followed by UV illumination. If the sample showed a positive result, the remainder of the product was then loaded onto a preparative 1% agarose gel stained with ethidium bromide. Gel extraction was carried out using QIAGEN’s Gel Extraction Kit. PCR products were eluted in 30µl-50µl of RNase free water, depending on the intensity of the band. 5µl of the eluted product was loaded onto a 1% analytical agarose gel to confirm that the extraction was carried out successfully. The concentrations of the eluted PCR products were determined by NanoDrop Spectrophotometer (ND-1000) prior to storage and sequencing. All PCR products were stored in -20°C until sending for sequencing.

2.2.15 Sequencing and Bioinformatics.

Eluted PCR products were sent to 1st BASE Singapore for sequencing. The sequences of the fragments were aligned using SeqMan program of the DNASTAR’s Lasergene software package. Aligned contigs were saved as the consensus sequence for each of the clinical specimens. EditSeq software was used to locate the positions of the correct open reading frames (ORFs) and the amino acid sequences of each of the clinical specimen were generated as well.

Each nucleotide consensus sequence generated was ran through a BLAST search. The sequences which were highly similar to the clinical specimen were included in the analysis, together with the sequences of the reference strain and the 2 influenza B ancestor strains (B/Yamagata/16/1988 and B/Victoria/02/1987) (Table 2.3).

<table>
<thead>
<tr>
<th>Accession Numbers</th>
<th>Reference Strain</th>
<th>Ancestor Strains</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA</td>
<td>NC_002207.1</td>
<td>CY018765.1</td>
</tr>
<tr>
<td>NA</td>
<td>NC_002209.1</td>
<td>CY018767.1</td>
</tr>
<tr>
<td>NS</td>
<td>J02096.1</td>
<td>CY018769.1</td>
</tr>
<tr>
<td>B Lee 40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Yamagata 16 1988</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Victoria 02 1987</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.3 Accession Numbers of the genes of the reference and ancestor strains used in this study
Both the nucleotide and the amino acid consensus sequences were aligned through the Clustal W application of the MegAlign software of the DNASTAR software package and the phylogenetic trees were generated using this alignment.

2.3 Virus Infection in Tissue Culture

2.3.1 Mammalian Cell Maintenance

Mammalian cells, such as MDCK and A549 were grown and maintained in T75 flasks in 37°C in Dulbecco’s modified Eagle’s medium (DMEM) (GIBCO, Rockville, MD) supplemented with 10% heat inactivated fetal bovine serum (GIBCO, Rockville, MD) and antibiotics (penicillin and streptomycin). Cells were trypsinized once confluent and were continually passaged.

2.3.2 Harvesting and Culturing of Chicken Embryonic Fibroblasts (CEF)

Ten-day old embryonated chicken eggs were collected from Chew’s Farm (Singapore). The exterior of the egg shells were wiped clean with 70% ethanol prior to harvesting. The egg shells were pecked by the use of sterilized forceps, just above the air sac. Upon removal of the allantoic membrane, the chick embryo was removed by forceps. The head, limbs and internal organs were removed, leaving only the ‘torso’ of the chick. This body part was then minced in sterile PBS and subsequently washed in sterile PBS as well. The wash solution was decanted to remove away any remaining blood or yolk. The tissue was then trypsinized with 0.21% trypsin-EDTA in PBS for 20 minutes. To inactivate the trypsin, 35ml of Cell Propagation Media (Section 2.1.7) was added to the tissue. Trypsinized cells were separated from the tissue by the use of a strainer and subsequently spun down at 400g for 10 minutes. The post-spin supernatant was discarded and the pelleted cells were re-suspended in 50mls of cell propagation media. Cells were counted by the use of a hemacytometer and appropriate numbers were seeded in flasks and/or dishes at 37°C, 5% CO₂.

2.3.3 Virus Infection on Cell Culture

Mammalian cells were seeded overnight in cell propagation media at 37°C, 5% CO₂ till desired confluency. The cell propagation media was aspirated and cells were washed in sterile PBS. Upon removal of the sterile PBS, an appropriate volume of the virus inoculum was dispensed directly onto the cells. This volume of inoculum was
calculated to achieve the desired MOI (Multiplicity of Infection). Inoculum was dissolved in DMEM. Cells immersed in the inoculum were placed in 37°C, 5% CO₂ for 1 hour with intermittent rocking. If only 1 cycle of virus infection was desired, the inoculum was aspirated and replaced with an appropriate volume of DMEM+GlutaMAX supplemented with 2% Fetal Calf Serum and Penicillin and Streptomycin. To ensure that multiple cycles of virus infection occurred, the inoculum was replaced with DMEM+GlutaMAX supplemented with 0.21% BSA, 1µg/ml of Trypsin-TPCK and Penicillin and Streptomycin.

2.3.4 Conventional Plaque Assay

To carry out the conventional plaque assay, MDCK cells were seeded overnight at 1.2x10⁶ cells per well to confluency in 6-well plates. Prior to infection the cells were washed in PBS and subsequently inoculated with 200µl of the virus in 10 fold dilution. Cells were incubated at 37°C in the presence of 5% CO₂ with gentle rocking throughout. The inoculum was then replaced with a 2 ml overlay consisting of 1X DMEM, 1% low melting point agarose, 37.5 mg/ml of sodium bicarbonate, supplemented with 0.21% BSA and 1 µg/ml of TPCK-treated trypsin. The plates were placed at 4°C for 15 minutes for the overlay to solidify before incubation in a humid chamber in 37°C in the presence of 5% CO₂, till plaques were observed (3-5 days).

2.3.5 End Point Assay

Viruses were serially diluted ten folds in Phosphate-Buffered Saline (PBS). Fifty microlitres of each dilution were inoculated onto 10 adjacent wells of MDCK cells grown to confluency in a 96-well plate for 1 hour at 37°C in the presence of 5% CO₂. Cells were seeded at a density of 2x10⁴ cells per well 16 hrs prior to infection. The inoculum was aspirated and replaced with 100 µl DMEM (Invitrogen, USA) supplemented with 0.21% bovine serum albumin (BSA) (Sigma-Aldrich, USA) and 1 µg/ml trypsin treated with L-(tosylamido-2-phenyl) ethyl chloromethyl ketone (TPCK) (Biochemical Corporation, USA). The virus-infected cells were incubated at 37°C in the presence of 5% CO₂ for 5 days, until no further changes were observed in the number of wells showing CPE. Wells with virus-infected cells showing CPE were scored as positive and the value for the TCID₅₀ was calculated (Reed and Muench, 1938). This value was multiplied by 1000/50 to obtain the TCID₅₀/ml and the final virus
titre was obtained by multiplying the TCID$_{50}$/ml with the constant 0.69, based on the Poisson distribution (Murhammer, 2007; Reed and Muench, 1938).

2.3.6 Conventional Coverslip Immunofluorescent Assay

MDCK cells were grown on 12mm glass coverslips in 24 well-plates till 60-80% confluency in 37°C. These cells were washed in sterile PBS and were inoculated with 200 µl of the supernatants harvested from the tube cultures described above. The plates were incubated in 1 hour at 33°C. Each well was then supplemented with 500µl of DMEM supplemented with 0.25% BSA, 2µg/ml Trypsin TPCK and penicillin/streptomycin and observed daily for CPE.

Cells were observed daily and were fixed once CPE was observed. If no CPE was observed, the cells were fixed after 14 days of infection in 4% paraformaldehyde (PFA) and permeabilized with 0.1% Triton-X. Cells were then stained with influenza-B specific antibody (mab8259, Chemicon (Millipore)) in a 1:500 dilution and a duplicate was stained with influenza-A specific antibody (mab8257, Chemicon (Millipore)) in a 1:200 dilution, followed by a secondary antibody conjugated with FITC in a 1:100 dilution. Both primary and secondary antibodies were incubated for an hour in room temperature and the coverslips were washed with PBS in between incubations. Prior to mounting, the coverslips were washed twice with PBS and mounted on a glass slide over 2µl of mounting fluid and sealed with nail polish. Slides were observed by an immunofluorescence microscope and images were recorded.

2.3.7 Immunofluorescent Microplate Assay

Using the same viral dilutions as in the end-point assay, 250µl of each dilution was inoculated onto confluent MDCK monolayers seeded on glass coverslips in a 24-well plate for 1 hr at 37°C in the presence of 5% CO$_2$. Cells were seeded at a density of 8x10$^4$ cells per well 16 hrs prior to infection. The inoculum was then aspirated and replaced with 500µl of DMEM supplemented with 0.21% BSA with or without 1µg/ml TPCK-treated trypsin. After 20 hrs post-infection (hpi), the virus-infected cells were fixed with 4% paraformaldehyde and subsequently permeabilized with 0.5% Triton-X for 15 minutes. Infected cells were labeled with mouse anti-influenza B NP (Chemicon, USA) and subsequently with goat anti-mouse fluorescein isothiocyanate (FITC)-conjugated antibody (Millipore, USA) and counterstained with Evan’s Blue (Sigma-
Aldrich, USA). The virus-infected cells were incubated with the respective antibodies at room temperature for 1hr each. The coverslips were washed with PBS, mounted onto glass slides and infected cells were visualized and counted under a fluorescence microscope (NIKON Eclipse 80i). The calculation of the virus titre is similar to that for a conventional plaque assay, but has to be adjusted to accommodate the difference in area between the glass coverslip and the well. For this adjustment, the observed number of infected cells were divided by the surface area of the coverslip, \( \pi(6\text{mm})^2 \), and multiplied by the area of the well, \( \pi(9\text{mm})^2 \).

2.3.8 Drug Treatment

Drugs were introduced into cell cultures as specified times as indicated in each experiment. Unless otherwise stated, the drugs were diluted to the expected concentration in DMEM+GlutaMAX in the absence of any serum. Cells were treated to the drug in the medium for the specified incubation periods.

2.4 Molecular Cloning of NS1, BM2 and NB genes

The genes coding for the NS1, BM2 and NB proteins of 6 different specimens were cloned into cloning and expression vectors. The following sections describe the methods involved.

2.4.1 Restriction Endonuclease Selection

Prior to PCR amplification, the nucleotide sequences of each of the genes were submitted to NEB Cutter (http://tools.neb.com/NEBcutter2/) (Vincze et al., 2003). Restriction endonucleases which did not contain a recognition sequence within the nucleotide sequence were considered in the cloning strategy for the specified gene.

2.4.2 PCR amplification of Genes for Cloning

Each gene to be cloned was amplified from the cDNA synthesized from the vRNA extract as described in sections 2.3.7 and 2.3.10. Primers were specifically designed for each gene to contain a restriction endonuclease site as well as an epitope tag. Table 2.4 lists the primers used in the cloning of the full length of each of the genes. In addition to the full length of the gene, individual domains of the NS1 gene was amplified and cloned accordingly.
Five microlitres of the cDNA were used as template in the PCR amplification. The PCR master mix contained the following reagents from Invitrogen; 5μl of 10X High Fidelity Buffer, 5μl of 2mM dNTP, 2.5μl of 50mM MgSO$_4$, 2μl of 10mM Forward Primer, 2μl of 10mM Reverse Primer 32.8μl of water and 0.5μl of Invitrogen High Fidelity Platinum Taq Polymerase. The PCR was carried out using the thermo-cycling conditions listed in table 2.5.

PCR products were analysed on an ethidium bromide stained agarose gel. Once the DNA band corresponding to the correct size of the intended was observed, it was extracted from the gel (Section 2.3.14). The extracted gene segment was eluted in water and stored in -20°C till further use.

2.4.3 Ligation into pDRIVE Cloning Vector

The concentration of the eluted gene segment was measured were determined by NanoDrop Spectrophotometer (ND-1000) prior to ligation. Ligation into pDRIVE (Invitrogen) was carried out as in manufacturer’s instructions. Briefly, 50ng of the pDRIVE plasmid was incubated with an appropriate amount of gene segment to achieve 1:10 molar ratio. The DNA mixture was made up to 5μl and mixed with 5μl of the 2X Ligation Master Mix (Invitrogen). This ligation was then placed in 4°C for a minimum of an overnight incubation. In the event of negative results, the ligation was repeated for a longer incubation time.

2.4.4 Bacterial Transformation of pDRIVE Ligations.

Twenty five microlitres of Invitrogen’s EZ competent cells were thawed on ice. Once thawed, 4μl of the ligation mixture were introduced to the EZ cells. This mixture was then incubated for 5 minutes on ice, followed by a heat shock at 42°C for 30 seconds. Cells were then placed on ice again for 2 minutes before the addition of 400 μl of pre-warmed SOC media (Invitrogen). Cells were then plated out on LB-amp plates (+IPTG & X-Gal) for an overnight incubation in 37°C.
<table>
<thead>
<tr>
<th>Primer Name</th>
<th>5' Sequence 3'</th>
<th>Size</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NS1-For (Sac1):</strong></td>
<td>CGAGCTCGAAGTAG GCG RACAAACATGAYC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NS1-For Myc:</strong></td>
<td>CGAGCTCGAA ATGGA AA AAACTCATCTCAGAGAGGAT CTG GCGRACAACATGAYC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NS1-Rev FLAG (Xho1):</strong></td>
<td>CCGGCTCGAGCGGCTATTTACGTCACTTGCAATTTGTAATCAACCTCTTGACCTAAATTTTCC</td>
<td>750 bp</td>
<td>NS1</td>
</tr>
<tr>
<td><strong>NS1-For (Xho1) pRSETC:</strong></td>
<td>TAGGATCCACCTCGAGGAAATGGCAATAAAGGCAGAGGAT CTG GAG AAAACATGAYC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NS1-Rev FLAG (HindIII) pRSETC:</strong></td>
<td>CCGGCTCGAGGGGTACCCACGAGATCCTCTTCTGAGATGAGTTTTGTTTCAGATCTGGCTCGGAGGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PR8 NS1-For (EcoRI):</strong></td>
<td>CCGGAATTCCGGGATGGATCCAAACACTGTCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PR8 NS1-Rev FLAG (Kpn1):</strong></td>
<td>CCGGTTACCCGTCATTATATTGCTATTTGTTTAATCAACCTCTTGACCTAAATTTTCC</td>
<td>850 bp</td>
<td>Flu A</td>
</tr>
<tr>
<td><strong>471 NS1-For (EcoRI):</strong></td>
<td>CCGGAATTCCGGGATGGACTCCAACACCATGTCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>471 NS1-Rev FLAG (Kpn1):</strong></td>
<td>CCGGTTACCCGTCATTATATTGCTATTTGTTTAATCAACCTCTTGACCTAAATTTTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NB-For (EcoRI):</strong></td>
<td>CCGGAATTCCGGGACATGACTCCCTAATACAAACAGTT</td>
<td>300 bp</td>
<td>NB</td>
</tr>
<tr>
<td><strong>NB-Rev Myc (Xho1):</strong></td>
<td>CCGGCTCGAGGGGTACCCACGAGATCCTCTTCTGAGATGAGTTTTGTTTCAGATCTGGCTCGGAGGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BM2-For:</strong></td>
<td>GAAATGTCCGAACCAYTGTCATTCGATCTTCTGAGATGAGTTTTGTTTCAGATCTGGCTCGGAGGTA</td>
<td>350 bp</td>
<td>BM2</td>
</tr>
<tr>
<td><strong>BM2-Rev FLAG:</strong></td>
<td>TTATTTATCTGCTATTTGTTTAATCGATTGATTTGTTTCAGATCTGGCTCGGAGGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ISG15-For (Xho1):</strong></td>
<td>CCGGTTACCCGTAATGGGCTGGGAACCTGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ISG15-Rev Myc (Xho1):</strong></td>
<td>CCGGCTCGAGGGGTACCCACGAGATCCTCTTCTGAGATGAGTTTTGTTTCAGATCTGGCTCGGAGGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RB-Rev (Xho1):</strong></td>
<td>CCGGCTCGAGGGGTACCCACGAGATCCTCTTCTGAGATGAGTTTTGTTTCAGATCTGGCTCGGAGGTA</td>
<td>270 bp</td>
<td>N51</td>
</tr>
<tr>
<td><strong>RBL-Rev (Xho1):</strong></td>
<td>CCGGCTCGAGGGGTACCCACGAGATCCTCTTCTGAGATGAGTTTTGTTTCAGATCTGGCTCGGAGGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ED-For (Sac1):</strong></td>
<td>CCGGCTCGAGGGGTACCCACGAGATCCTCTTCTGAGATGAGTTTTGTTTCAGATCTGGCTCGGAGGTA</td>
<td>575 bp</td>
<td>N51</td>
</tr>
<tr>
<td><strong>LED-For Myc (Sac1):</strong></td>
<td>CCGGCTCGAGGGGTACCCACGAGATCCTCTTCTGAGATGAGTTTTGTTTCAGATCTGGCTCGGAGGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LED-For (Sac1):</strong></td>
<td>CCGGCTCGAGGGGTACCCACGAGATCCTCTTCTGAGATGAGTTTTGTTTCAGATCTGGCTCGGAGGTA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.4 Primers used in amplifying NS1, NB, BM2 and ISG15 genes for molecular cloning. Primers contain an epitope tag, either FLAG or Myc. The name of restriction endonuclease sequence designed in the primer is indicated in brackets. RB: RNA Binding Domain. RBL: RNA Binding Domain + Linker. ED: Effector Domain. LED: Linker + Effector Domain.
White colonies were picked and inoculated in 20μl of water. Five microlitres of this inoculation were then used as template for colony screening. The PCR master mix for this colony screen included 1μl of 10X High Fidelity Platinum Taq Buffer, 1μl of 2mM dNTP, 0.5μl of 50mM MgSO₄, 0.4μl of 10mM Forward Primer, 0.4μl of 10mM Reverse Primer and 1.6 μl of water and 0.1μl of Invitrogen High Fidelity Platinum Taq Polymerase. Primers used were specific to M13 sites present on the flanking regions of the pDRIVE cloning vector. The thermo-cycling conditions for colony screening were as listed in table 2.6.

<table>
<thead>
<tr>
<th>Initial Denaturation</th>
<th>94°C</th>
<th>15 min</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>94°C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Touchdown</td>
<td>70°C</td>
<td>30 sec</td>
</tr>
<tr>
<td></td>
<td>72°C</td>
<td>3 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94°C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>62°C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>3 min</td>
</tr>
<tr>
<td>Final Extension</td>
<td>72°C</td>
<td>10 min</td>
</tr>
<tr>
<td>Holding</td>
<td>4°C</td>
<td>forever</td>
</tr>
</tbody>
</table>

**Table 2.5 PCR Thermocycling conditions used to amplify gene segments for molecular cloning.**

<table>
<thead>
<tr>
<th>Initial Denaturation</th>
<th>94°C</th>
<th>5 min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denaturation</td>
<td>94°C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>55°C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Extension</td>
<td>68°C</td>
<td>1 min 30 sec*</td>
</tr>
<tr>
<td>Final Extension</td>
<td>68°C</td>
<td>5 min</td>
</tr>
<tr>
<td>Holding</td>
<td>4°C</td>
<td>forever</td>
</tr>
</tbody>
</table>

**Table 2.6 PCR Thermocycling conditions for colony screening.** Extension times were adjusted according to the length of the insert.

Positive colonies would display a band corresponding to the size of the insert + 200 bp. A colony not containing an insert would only display a band of 200bp. Positive colonies were inoculated in 5mls of LB-amp broth overnight. Minipreps of the positive clones were made and the concentrations of the eluted plasmids were quantified and stored in -20°C. The plasmids were sent to 1st BASE, Singapore for sequencing, using the primers specific to M13.
2.4.5 Restriction Endonuclease Treatment

Once the sequence of the insert had been determined correct, the insert was ‘cut’ out of the vector through the use of restriction endonucleases which are specific to each gene.

Five thousand nanograms of the vector was incubated with 1μl of each restriction enzyme in its appropriate buffer in the presence of BSA. Restriction digest took place at 37˚C overnight, as indicated by the manufacturer. A similar amount of pCAGGS/MCS vector was also subjected to the appropriate restriction digestion (Niwa et al., 1991). pCAGGS vector was a gift obtained from Professor C. Broder (Patch et al., 2007). To separate the insert from the cut vector, the vector was subjected to DNA gel electrophoresis on a 1% ethidium bromide stained, agarose gel. Once a band corresponding to the correct size of the insert was identified, it was extracted from the gel as in section 2.3.14. The pCAGGS vector was purified through similar methods. The cut pCAGGS vector was then subjected to Shrimp Alkaline Phosphatase (SAP, Promega) treatment at 37˚C for 30 min, followed by deactivation at 65˚C for 15min.

2.4.6 Ligation into pCAGGS vector

Post restriction digest as well as SAP treatment, the concentrations of both the pCAGGS vector and the insert were determined by the use of the Nanodrop (Section 2.3.14). Appropriate volumes of both the insert and the cut vector were mixed to achieve a 1:10 molar ratio (Vector:Insert). 1 microlitre of T4 Ligase as well as 10X T4 ligase buffer (Roche) was added to the mixture and incubated for a minimum of 16 hours in 4˚C.

2.4.7 Bacterial Transformation of Insert-pCAGGS

The ligations were then transformed into Qiagen’s EZ competent cells as in section 2.5.3. These transformed bacteria were plated out on LB-Amp plates overnight. Colonies were picked and screened using primers specific to the regions flanking the multiple cloning site of pCAGGS. Positive clones were grown up overnight in LB-Amp broth and mini-prep of the clones were obtained.

2.4.8 Site-Directed Mutagenesis

DNA substitutions were introduced into the cloned gene segments by use of the QuikChange Site-Directed Mutagenesis (Stratagene). Primers were designed with the
DNA substitutions incorporated, through the use of the manufacturer’s online primer design tool (http://labtools.stratagene.com/QC). Table 2.7 lists the sequence of all of the primers used for mutagenesis experiments in this study.

The pDRIVE vector containing the gene to be mutated was used as template in the PCR. Ten nanograms of template DNA was mixed with 5μl of 10X reaction buffer, 125ng of forward primer, 125ng of reverse primer, 1μl of dNTP mix, 1μl of *Pfu Turbo* DNA polymerase and ddH₂O was added to a final volume of 50μl. All the mentioned reagents were from the QuikChange® Site-Directed Mutagenesis kit from Stratagene. The thermo-cycling conditions for the site-directed mutagenesis PCR, as recommended by the manufacturer, are listed in table 2.8.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>5' Sequence 3'</th>
<th>Specimen</th>
<th>Substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee-NS1 M91A For:</td>
<td>GGGTAAATGTAAGTGTTCTTCTGT TTGCGG ATCCCTCTGCTG</td>
<td>Lee-NS1</td>
<td>M91A</td>
</tr>
<tr>
<td>Lee-NS1 M91A Rev:</td>
<td>CCAGCAGAGGGAGTCGCAAAACAGAAGCACT TTCATCATTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee-NS1 D92N For:</td>
<td>GAAAGTGCTTCTGTTGTATGAAATCCCTCTGCTGGAATTGA</td>
<td>Lee-NS1</td>
<td>D92N</td>
</tr>
<tr>
<td>Lee-NS1 D92N Rev:</td>
<td>CCTTCAATTCAGCAGGGGATTTCAAACAGAAGCACTTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee-NS1 D92A For:</td>
<td>GAAAGTGCTTCTGTTTTATGGATCCCTCTGCTGGAATTGA</td>
<td>Lee-NS1</td>
<td>D92A</td>
</tr>
<tr>
<td>Lee-NS1 D92A Rev:</td>
<td>CCTTCAATTCAGCAGGGGATTTCAAACAGAAGCACTTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee-NS1 P93A For:</td>
<td>GAAAGTGCTTCTGTTATGGATGCCTCTGCTGGAATTGA</td>
<td>Lee-NS1</td>
<td>P93A</td>
</tr>
<tr>
<td>Lee-NS1 P93A For:</td>
<td>AATTCAGCAGGGCATCCATAACAGAAGCACTTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee-NS1S94P For:</td>
<td>CTTCCTTCTTCTGAGCCTCCCTCTGCTGGAATGGAT</td>
<td>Lee-NS1</td>
<td>S94P</td>
</tr>
<tr>
<td>Lee-NS1S94P Rev:</td>
<td>ATCCTTCAATTCAGCAGGGGATTTCAAACAGAAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee-NS1 NPL107SSS For:</td>
<td>AATTTGAAGGATCTGCAGCTCTGCTGGAATTGAAGGG</td>
<td>Lee-NS1</td>
<td>NPL107SSS</td>
</tr>
<tr>
<td>Lee-NS1 NPL107SSS For:</td>
<td>TTGGCATTGGCTGTTACGAGGCTTTTCAGCATGACTTGGCT CAAATCTTCTAATT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>132-NS1 N92D For:</td>
<td>GAAAGATCTCTTATTATGGATCCCTCTGCTGGAATTGA</td>
<td>132-NS1</td>
<td>N92D</td>
</tr>
<tr>
<td>132-NS1 N92D Rev:</td>
<td>CCCTCCTAATTCAGCAGGGATTTCAAACAGAAGGACTTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>132-NS1 S69N For:</td>
<td>CAACAAAAGTGAGCCTGAATAAAAGGATGTCCCTTGAAG</td>
<td>132-NS1</td>
<td>S69N</td>
</tr>
<tr>
<td>132-NS1 S69N Rev:</td>
<td>CCTCAAGGAGCATTTTTTTTTTACGGCTCACTTGG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.7 Primers used for mutagenesis of the cloned specimens in this study

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>132-NS1 S123N For:</td>
<td>GGACCGATTACCCTCAACACCAGGGAGG</td>
</tr>
<tr>
<td>132-NS1 S123N Rev:</td>
<td>CTCCTCGGTGGTTGAGGGGTAAATCGGTCC</td>
</tr>
<tr>
<td>132-NS1 S215N For:</td>
<td>TGCAATAGATGAAATGACCAGAATGGAGGTTTGG</td>
</tr>
<tr>
<td>132-NS1 S215N Rev:</td>
<td>CAACAGAGGCTTCCATTTGCTCATATGCATTCAATTATGCA</td>
</tr>
<tr>
<td>132-NS1 SSS107ASS For:</td>
<td>GGAGCTCCTCAAAATAGCAACTGTACGG</td>
</tr>
<tr>
<td>132-NS1 SSS107AAS For:</td>
<td>GAGCTCCTCAAAATAGCAACTGTACGG</td>
</tr>
<tr>
<td>132-NS1 SSS107AAAFor:</td>
<td>GAGCTCCTCAAAATAGCAACTGTACGG</td>
</tr>
<tr>
<td>117-NS1 P94S For:</td>
<td>TCCTATTATGGATCGGTCTGGAATTGAAGGG</td>
</tr>
<tr>
<td>117-NS1 P94S Rev:</td>
<td>CCGTACAGTTCATTTCAGGAAGGCTTTCATACAGTATGGCTCAAAACC</td>
</tr>
</tbody>
</table>

Table 2.8 Thermocycling conditions recommended for site-directed mutagenesis PCR.

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>95°C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95°C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Annealing</td>
<td>55°C</td>
<td>1 min</td>
</tr>
<tr>
<td>Extension</td>
<td>68°C</td>
<td>1 min /kb of plasmid length</td>
</tr>
<tr>
<td>Holding</td>
<td>4°C</td>
<td>forever</td>
</tr>
</tbody>
</table>

The number of cycles for the PCR shown in table 2.8 is dependent on the nature of the mutation to be introduced. Briefly, a point mutation requires 12 cycles, 2-3 mutations resulting in a single amino acid change requires 16 cycles, while multiple amino acid changes as well as insertions or deletions require 18 cycles.
then heat shocked for 45 seconds at 42°C. Cells were then placed back on ice for 2 minutes and 500μl of NZY+ broth (Section 2.1.8) were added. NZY+ broth was preheated to 42°C. The transformations were then incubated for 1 hour at 37°C with shaking at 225rpm.

Cells were then plated onto LB-Amp plates and incubated overnight at 37°C. Colonies were picked and grown overnight in 5ml of LB-Amp broth and minipreps of each cloned were prepared. Using M13 primers, the mutated plasmids were sent for sequencing and the presence of the desired mutations were checked. Plasmids containing the desired mutations were subsequently sub-cloned into pCAGGS as described in sections 2.5.5-2.5.6.

2.4.9 Plasmid Transfection into Mammalian Cells

Cells were seeded to desired cell number and confluency. Prior to transfection, the cell propagation medium was replaced with antibiotic-free cell propagation medium. The plasmids to be transfected were dissolved in Opti-MEM® (Invitrogen) and to Lipofectamine 2000 (Invitrogen) to the appropriate volumes specified by the manufacturer. After 20 minute incubation, the DNA-Lipofectamine complex was added directly to the cells. Four to six hours post transfection, the media was replaced to DMEM+GlutaMAX supplemented with 2% FCS.

2.4.10 Endoglycosidase Treatment

Cells expressing the protein of interested were washed in sterile PBS once and scraped in Glycoprotein Denaturing Buffer (DEN Buffer) (New England Biolabs, NEB). Cells in this suspension were heated at 100°C for 20 minutes, sonicated briefly and heated at 100°C for another 10 minutes. Table 2.9 lists the different reagents used in treating the cells to 2 different endoglycosidases: Endo H and PNGase F (NEB). For negative control reactions, 1μl of water was added instead of enzyme.

<table>
<thead>
<tr>
<th></th>
<th>Endo H Treatment</th>
<th>PNGase F Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Lysate in DEN Buffer</td>
<td>15μl</td>
<td>15μl</td>
</tr>
<tr>
<td>10x G7 Buffer</td>
<td>-</td>
<td>2.5μl</td>
</tr>
<tr>
<td>10x G5 Buffer</td>
<td>2.5μl</td>
<td>-</td>
</tr>
<tr>
<td>NP40</td>
<td>-</td>
<td>2.5μl</td>
</tr>
<tr>
<td>Enzyme</td>
<td>1μl</td>
<td>1μl</td>
</tr>
<tr>
<td>dH2O</td>
<td>6.5μl</td>
<td>4μl</td>
</tr>
<tr>
<td>Total</td>
<td>25μl</td>
<td>25μl</td>
</tr>
</tbody>
</table>

Table 2.9 Reagents involved in endoglycosidation of proteins in this study.
2.4.11 Chemical cross linking of proteins

Cells were transfected with the intended gene for a specified incubation preceding a wash with PBS pH 8. Four microgrammes of chemical cross linker Dithiobis[succinimidyl propionate] (DSP) Pierce Biotechnology) was dissolved in 100μl of dimethyl sulfoxide (DMSO) to reach a stock concentration of 100mM. Cells were treated with the varying concentrations of DMSO dissolved in PBS pH 8.0 for 1 hour at room temperature. Post treatment, cells were washed with PBS pH 8.0 with 2mM lysine for quenching. Cells were then harvested in an appropriate buffer for further analysis.

2.4.12 Immunoprecipitation

Cells expressing the desired protein to be immunoprecipitated were washed in sterile PBS once and were treated to RIPA buffer for 20 minutes at 4°C with intermittent shaking. Two hundred microlitres of RIPA buffer were added to a 35mm dish of cells. Cells were then scraped in the RIPA buffer and spun down at 14000rpm for 20 minutes at 4°C. A small volume of the supernatant was removed and boiled in Laemmli Buffer for protein analysis. The supernatant was then added to Binding buffer, supplemented with 0.25% BSA and 0.5μl of antibody at 4°C for an overnight incubation. Thirty microlitres of Protein A Sepharose beads (Sigma-Aldrich) were added directly to this mix and the tubes were incubated at 4°C for 90 minutes with agitation. The tubes were spun down at 14000rpm for 5 minutes to pellet the Protein A Sepharose beads. The beads were then washed in 500μl of Low Salt Buffer 3 times. Beads were finally resuspended in 50μl of 1X Laemmli buffer for Protein Analysis. The components of the buffers used in immunoprecipitation are listed in section 2.1.13.

2.5 Protein Analysis by SDS-PAGE

Samples requiring SDS-PAGE analysis were heated in Laemmli Buffer at 95°C for 5 minutes. If the sample were total cell lysate, then cells were scraped directly in 1X Laemmli buffer after washing with PBS (Section 2.1.10). Samples presenting with viscosity, such as total cell/bacterial lysates, even after heating were sonicated till viscosity was not observed. Boiled and sonicated samples were stored in -80°C.

PAGE gels were casted to their appropriate acrylamide concentrations by the use of the Mini-PROTEAN® Tetra Cell (BIORAD). Samples were loaded into the gels
bathed in SDS-PAGE Running buffer. The proteins were separated by a potential difference of 200V till the dye front of the samples reached the base of the gel. The gels were then carefully removed from the glass plates and processed for protein visualization (Sections 2.6.1-2.6.3).

2.5.1 Western Blotting

Polyvinylidene Fluoride (PVDF) membrane (PALL Whatman Laboratories) was activated by soaking in methanol. The transfer unit cassette of the Mini Trans-Blot® Cell rig was assembled in the following order: Red Electrode, Clear plate, Pad, 3MM filter paper, PVDF membrane, SDS-PAGE gel, 3MM Filter Paper, Pad, Black Electrode. The cassette was loaded into the tank and bathed in transfer buffer with a block of ice and ran for 1 hour at 100V. The PVDF membrane was removed and blocked in 5% skim milk for a minimum of 1 hour with gentle agitation. The PVDF membrane was then washed in PBST 3 times and then incubated with the appropriate antibody diluted in PBST for 1 hour at room temperature with gentle agitation. The PVDF membrane was then washed in PBST 3 times and incubated with the appropriate secondary antibody conjugated to Horse Radish Peroxide (HRP) for 1 hour at room temperature with gentle agitation. The secondary antibody was washed with PBST 3 times and the protein bands treated with Enhanced Chemiluminescence (ECL, Amersham) briefly and exposed to KODAK O-MAT X-Ray film in the dark. The bands were visualized in Kodak Developer Machine after an appropriate exposure time (Section 2.1.11).

2.5.2 Coomassie Brilliant Blue Staining

The gel was immersed in a volume of coomassie blue stain for 1 hour with gentle agitation. The coomassie blue stain was removed and the gel was washed in water. The gel was then treated with destaining buffer till the protein bands were discernible from the background (Section 2.1.12).

2.5.3 SYPRO Ruby Staining

Gel was immersed in SYPRO Ruby Fixing Solution for 30 minutes at room temperature with gentle agitation. The fixing solution was removed and fresh fixing solution was added for another 3 minutes. The gel was then washed in dH₂O twice for 5 minutes each. The gel was then immersed in SYPRO Ruby Protein Stain in the dark
overnight at room temperature with gentle agitation (40rpm). The gel was then washed with SYPR Ruby Washing Solution for 30 minutes with gentle agitation. The gel was then washed twice in dH₂O for 5 minutes each. The gel was then scanned using the Typhoon Trio scanner (GE Amersham) to visualize the protein bands (Section 2.1.13).
Chapter 3: Cultivation and passage of clinical specimens and B/Lee/40

3.1 Introduction

The clinical specimens analysed in this study were oropharyngeal samples collected from SAF personnel displaying symptoms characteristic of influenza infection such as high temperature (≥ 38°C) with a sore throat and/or cough. These swabs were then re-suspended in viral transport medium (VTM) (Seah et al., 2010b). These samples were tested by DSO staff for the presence of influenza B virus prior to the work in this chapter. This VTM was then treated as the main inoculum in the primary cultivation of these clinical specimens. The characteristics of a viral strain can be characterized by cell culture techniques. These techniques allow the measure of Cytopathic Effect (CPE) incurred by the virus on the cells. CPE is defined as a set of morphological changes displayed by cells upon virus infection. CPE caused by infection is a gauge of both how pathogenic the strain is and/or is an indication of the viral load in the inoculum (Palese, 2007). When analysed by immunofluorescence, cellular localisation of the viral proteins can be determined, lending greater insights into the kinetics of viral proteins during infection.

Besides the initial diagnostic infection, these specimens were passaged through cell culture as well as embryonated chicken eggs, to allow us to study the growth kinetics of clinical specimen influenza B viruses. The severity of these infections was measured using the immunofluorescence techniques.

Basic characterisation of B/Lee/40 was carried out prior to the experiments in this study. The virus was inoculated into embryonated chicken eggs and the harvested allantoic fluid was used as inoculum throughout this study. The allantoic fluid was assayed for its infectivity by infection on cell culture.

To further analyse the virus, the virus was purified through centrifugation on a continuous sucrose gradient and subjected to transmission electron microscopy (TEM). Studies similar to this have been done to characterize influenza A viruses (Arora et al., 1985; Shaw et al., 2008), however, similar experiments for influenza B viruses seem to be lacking. The purified virus was then subjected to SDS PAGE analysis to characterise the structural protein components of the virus.

An important characteristic of a virus stock is its titre. The titre of a virus stock is applicable to both research and diagnostics. To ascertain the titre, most assays would quantify a biological property of the viruses, reflecting its infectivity (Condit, 2007).
The conventional plaque assay is the most widely used (Babiker and Rott, 1968; Dulbecco and Vogt, 1953). In this assay, the viruses are serially diluted and infected on cell culture monolayers. At an appropriate dilution, a single infectious particle would give rise to a single and observable plaque, define as a region of cell death, visible by light microscopy. The numbers of plaques observed can then be used to calculate the titre of the original virus stock.

Another commonly used method is the end-point assay. Similar to the plaque assay, viruses are serially diluted and infected on specific hosts. However, these infections have to be carried out in replicates. The end-point assay is specifically for viruses which do not produce plaques in tissue culture but their pathology is clearly displayed upon infection. In this assay, the host can include cells, animals or embryonated eggs. Statistical calculations are employed to determine the value for the TCID\textsubscript{50} of the virus, which corresponds to the number of viruses which would yield 50% pathology in the hosts (Condit, 2007; Reed and Muench, 1938).

A common drawback for both these assays is that they are applicable only to viruses which replicate well in cell culture. Certain viruses do not replicate well in cell culture. These viruses may be able to cause infection in cells but might not be able to produce viable progeny. Plaques and CPE cannot be produced by a single round of infection. Clinical specimens are adapted to humans and may not replicate well in cell culture, making the use of these 2 assays impractical.

Another disadvantage of these assays is the time required before significant CPE or plaques are observed. Multiple rounds of virus infection are required before observable pathology is established. In the case of influenza, trypsin has to be added to the cell culture media to facilitate successful viral spread (Gaush and Smith, 1968; Klenk et al., 1975; Lazarowitz and Choppin, 1975).

Even though the drawbacks highlighted for both these assays are relevant to clinical specimens, the egg-passaged B/Lee/40 generated in this study could not produce plaques in cell culture. The reason for this is unknown but to circumvent this issue, an immunofluorescent microplate assay was developed to measure the titre of the virus. This assay was validated against titres obtained for 2 other influenza A viruses which have also been titred through the conventional plaque and end-point assays.
3.2 Results

3.2.1 Cultivation of Clinical Specimens

A total 81 clinical specimens were tested in this study. These specimens were collected between the years 2004-2009 by the method describes by Seah, et al and were tested positive for influenza B virus by diagnostic real time PCR (RT PCR) (Seah et al., 2010a). Twenty-two of these samples were tested to be positive for influenza B by DSO staff previously and have had their vRNA extracted and stored in -80°C. These vRNA were not cultured and used directly as template for sequencing. The remaining 59 specimens were received in Viral Transport Media (VTM) and were inoculated onto MDCK tube cultures, as described in methods. Out of the 59 specimens inoculated, 48 specimens showed significant CPE within two weeks of infection. These cells were scraped and a diagnostic immunofluorescence assay (IFA) test was carried out on them. Diagnostic IFA involves staining permeabilized cells with anti influenza A and B antibodies to check for the possibility of cross contamination. Only 44 out of the 48 specimens showing CPE, stained positively for influenza B. vRNA from the 44 showing positive signal on IFA were extracted from the clinical specimens and used as template for sequencing. Figure 3.1 shows the breakdown of clinical specimens received.

<table>
<thead>
<tr>
<th>Total</th>
<th>Received as</th>
<th>Culture</th>
<th>FluB Diagnostic IFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>81</td>
<td>59 VTM</td>
<td>48 +ve CPE</td>
<td>44 positive</td>
</tr>
<tr>
<td>Clinical Specimens</td>
<td>11 -ve CPE</td>
<td>4 negative</td>
<td></td>
</tr>
<tr>
<td></td>
<td>22 vRNA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.1 Breakdown of the 81 clinical specimens received for this study.

Figure 3.2 shows representative images of the results of the diagnostic IFA (Figure 3.2). Specimens DSO_050526_2005 and DSO_050528_2005 are representative of the influenza B-positive specimens. As seen in figure 3.2, specimen DSO_040152_2006 is positive for influenza A rather than influenza B and no influenza B sequence information could be attained from this specimen. DSO_040200_2006 represents a specimen which is negative for both viruses’ (Figure 3.2). Specimen
DSO_040203_2006 did not show any CPE upon inoculation but showed a positive signal when stained with the diagnostic antibody.

The supernatants of the samples showing positive signals were stored in 1ml aliquots and stored in -80°C. The supernatants of 20 of these samples were used to infect MDCK cell grown on glass coverslips and were stained for influenza B and influenza A infection after the observation of CPE or 14 days of incubation. Out of the 20, 14 specimens showed strong positive signal for influenza B infection, 5 showed negative signal while 1 sample showed partial signal (unable to differentiate between positive and background). Figure 3.3 shows representative images of the specimens analysed this way. Specimens DSO_020132_2007 and DSO_020147_2007 showed the strongest influenza B positive signal upon infection. Infection with specimen DSO_020114_2006 resulted in changes in cell morphology and positive staining on the periphery of the cells. Specimen DSO_020113_2007 only resulted in 2 cells being infected while DSO_070214_2006 had a very faint staining. Specimen DSO_040152_2006 which showed positive influenza A staining in figure 3.2 showed very faint influenza A staining and no influenza B staining. Specimen DSO_050320_2006 showed similar staining as the negative control (Figure 3.3).

Out of the 44 specimens showing positives, 3 specimens showing the strongest signals were chosen to be inoculated into eggs. These specimens were; DSO_020147_2007, DSO_070214_2006 and DSO_040174_2006. The harvested allantoic fluid was used to infect MDCK cells and was stained for influenza A and B infection after observable CPE. The results of a conventional infection did not show any positive staining. To circumvent this problem, a shell vial infection assay was carried out. The cells were fixed and stained after 3 days of infection. Only samples DSO_020147_2007 and DSO_040174_2006 showed low positive signals. The supernatant of specimen DSO_020147_2007 only infected a small portion of the monolayer and DSO_040174_2006 only infected 2 cells in the monolayer (Figure 3.4).
Figure 3.2 Diagnostic IFA of clinical specimens. MDCK cells inoculated with clinical specimens are scraped after CPE is observed, spotted onto multiwell glass slides and stained for influenza B (top panels) and influenza A (bottom panels). Green staining shows positive staining for the intended antigen while red staining (Evans Blue) serves as a control, to prove the presence of cells. Images taken at 20X magnification.
3.2.2 B/Lee/40 Passage and Cultivation

3.2.2.1 B/Lee/40 Infection Characterisation by Immunofluorescence Assay (IFA)

B/Lee/40 virus purchased from ATCC, were inoculated in 12-day old chicken eggs for 48hrs and the clarified allantoic fluid were used as inoculum in the course of this study. To study the characteristics of influenza B infection, human alveolar A549 cells were infected with egg-passaged B/Lee/40 and immunologically-labeled with anti influenza B Nucleoprotein (NP) antibodies at 8, 12 and 16 hours post infection (hpi) (Figure 3.5).
Figure 3.4 Shell vial infection assay for allantoic fluids harvested from eggs infected with the supernatants of specimens DSO_020147_2007, DSO_040174_2006 and DSO_070214_2006. 4.3a, MDCK cells infected with egg passaged DSO_020147_2007. 4.3b, MDCK cells infected with egg passaged DSO_040174_2006. 4.3c, MDCK cells infected with egg passaged DSO_070214_2006. (10x magnification)

Figure 3.5 Time course infection of egg-passaged B/Lee/40 on human alveolar A549 cells. Cells were stained for both influenza B NP and Lamin A/C. Rex box highlights uninfected cells (negative control). Images were taken under oil immersion (100X magnification).

Cells in figure 3.5 were co-stained with anti-Lamin A/C as a nuclear marker. Early in infection, NP is located primarily in the nucleus. The intensity of staining increases at 12hrs and at 16hrs NP can be found in the cytoplasm as well. The region highlighted in the red box in figure 3.5 highlights uninfected cells which show Lamin
A/C staining but not NP staining. These were cells which remained uninfected despite being exposed to virus.

### 3.2.2.2 Purification of B/Lee/40

To further analyse the properties of B/Lee/40, the virus from the allantoic fluid was purified through a 30%–60% sucrose gradient as described in Methods (Chapter 2). Interestingly, 2 opalescent bands were observed in the gradient at 30% and at 40% (Figure 3.6) when observed with the aid of a light box.

![Figure 3.6 B/Lee/40 virus purification through a 30%–60% sucrose gradient.](image)

The virus bands observed were harvested and pelleted. The virus pellets were analysed on a polyacrylamide gel stained with Sypro Ruby (Figure 3.7). The 40% virus pellet displayed more intense bands than the 30% pellet on 12% SDS-PAGE (Figure 3.7a). For greater resolution of the individual viral protein bands, the 40% pellet was ran on 7.5% and 15% SDS-PAGE (Figure 3.7b and c).
Bands which could potentially be influenza B viral proteins are identified in Figure 3.7 and their possible identities are suggested in table 3.1, which lists the molecular weights of influenza B proteins (structural proteins are listed in bold). It is important to note that this labeling is not definitive and more confirmatory experiments need to be carried out to determine the identity of these proteins.

**Influenza B Viral Protein Sizes**

<table>
<thead>
<tr>
<th>Protein</th>
<th>PB1</th>
<th>PB2</th>
<th>PA</th>
<th>HA</th>
<th>NP</th>
<th>NA</th>
<th>NB</th>
<th>M1</th>
<th>BM2</th>
<th>NS1</th>
<th>NEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (kDa)</td>
<td><strong>84</strong></td>
<td><strong>88</strong></td>
<td><strong>83</strong></td>
<td><strong>80</strong></td>
<td><strong>70</strong></td>
<td><strong>51</strong></td>
<td><strong>15</strong></td>
<td><strong>32</strong></td>
<td><strong>15</strong></td>
<td><strong>35</strong></td>
<td><strong>14</strong></td>
</tr>
<tr>
<td>p1</td>
<td>p1</td>
<td>p2</td>
<td>p3</td>
<td>p3</td>
<td>p5</td>
<td>p4</td>
<td>p5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1 Sizes of Influenza B proteins in KiloDaltons (kDa). Structural proteins are listed in bold. Suggestions to the peptides observed in figure 3.7 are listed in the bottom row.
3.2.2.3 Transmission Electron Microscopy of Purified B/Lee/40

The 40% virus band observed in the sucrose gradient was pelleted and resuspended in Phosphate Buffered Saline (PBS) and was analysed by Transmission Electron Microscopy (TEM) (Figure 3.8).

Figure 3.8 TEM image of purified B/Lee/40 virus from the 40% virus band. Both images show the same field at different magnifications, 150000X (a) and 250000X (b). Inset shows the magnified region of the spike proteins of a virion (b).

TEM images show circular structures of similar size to an influenza virion, 50-100nm. These structures are mostly spherical in shape and contain a dense shell of 5-10nm around their periphery (Figure 3.8b inset). This coincides with the spike proteins on the exterior of the virus, mostly made up of the HA and NA proteins.

3.2.3 Growth Kinetics of B/Lee/40

To study the replication kinetics of B/Lee/40, Madin Darby Canine Kidney (MDCK) cells were infected with the virus and the supernatant was harvested every 12hrs for 4 days to assay for the presence of virus. Figure 3.9 shows the CPE displayed by the monolayer over the course of 4 days.

Prior to infection (day 0), cells formed a monolayer but as the infection progressed more CPE was observed. Few cells remained on the surface by the 3rd day of infection and by the 4th day, all cells had lifted off, suggesting 100% cell death (Figure 3.9a). To determine if this effect was a product of virus infection, cells were collected at 24hpi and spotted on glass slides. These cells showed positive signal (Figure 3.9b) when labeled with anti influenza B NP antibody.

After proving that the cells were infected with influenza B, the harvested supernatant from each of the time points, indicated in figure 3.9a, were used as
inoculum to infect a fresh set of MDCK cells. Interestingly, this inoculum did not result in infection of this new set of cells after 20 hours of infection.

![Figure 3.9 CPE caused by B/Lee/40 infection on MDCK cells.](image)

3.2.4 Immunofluorescent Microplate Assay

To determine the titre of egg-passaged B/Lee/40, 4 sets of end-point assays were carried out. The titres produced were of similar orders of magnitude and had an average of $5.89 \pm 1.1 \times 10^4$ pfu/ml (Table 3.2).

Using the same dilutions as the end point assay, B/Lee/40 was inoculated on MDCK cells for the immunofluorescent microplate assay. A virus-infected cell shows prominent nuclear fluorescence when fluorescently labeled with anti NP, as highlighted in dilution $10^{-2}$ (Figure 3.10, inset) and $10^{-3}$ (figure 3.11, inset). Countable infected cells were observed from the viral dilution $10^{-2}$ onwards (Figure 3.10). This assay was repeated 4 times and yielded an average titre of $9.90 \pm 3.5 \times 10^4$ pfu/ml, similar to that obtained from the end-point assay (Tables 3.2 and 3.3). To test the genetic stability of this assay, A/WSN/33 (H1N1) and A/Duck/Malaysia/1/01 (H9N2) were titred as well, yielding $2.4 \times 10^8$ pfu/ml and $8.0 \times 10^6$ pfu/ml respectively (Table 3.3)
## End-Point Assay of B/Lee/40

<table>
<thead>
<tr>
<th>Dilution</th>
<th>aPlate 1</th>
<th>aPlate 2</th>
<th>aPlate 3</th>
<th>aPlate 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^0$</td>
<td>10/10</td>
<td>10/10</td>
<td>10/10</td>
<td>10/10</td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td>10/10</td>
<td>10/10</td>
<td>10/10</td>
<td>10/10</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>10/10</td>
<td>10/10</td>
<td>10/10</td>
<td>10/10</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>8/10</td>
<td>10/10</td>
<td>8/10</td>
<td>9/10</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>3/10</td>
<td>3/10</td>
<td>0/10</td>
<td>1/10</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
<td>0/10</td>
</tr>
</tbody>
</table>

| TCID$_{50}$ | 7.96x10$^4$ | 1.04x10$^5$ | 6.31x10$^4$ | 9.48x10$^4$ |
| Titre (pfu/ml) | 5.49x10$^4$ | 7.18x10$^5$ | 4.36x10$^4$ | 6.54x10$^5$ |

Average Titre $5.89 \pm 1.1 \times 10^4$

### Table 3.2 Titres of egg passaged B/Lee/40 by end-point assay

- a 4 replicates of the end point assay were carried out, denoted by Plate 1, Plate 2, Plate 3 and Plate 4
- b Virus inoculum was serially diluted 10 fold from $10^0$ to $10^{-5}$ in PBS.
- c TCID$_{50}$ value calculated by Reed and Muench Method, 1938
- d Titre in plaque forming units/ml (pfu/ml), estimated by multiplying the TCID$_{50}$ by the constant 0.69

### Comparison of Various Titration Assays

<table>
<thead>
<tr>
<th>Method</th>
<th>B/Lee/40</th>
<th>H1N1</th>
<th>WSN</th>
<th>H9N2</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-Point Assay</td>
<td>5.89x10$^4$</td>
<td>3.27x10$^8$</td>
<td>4.36x10$^7$</td>
<td></td>
</tr>
<tr>
<td>Conventional Plaque Assay</td>
<td>b -</td>
<td>2.0x10$^8$</td>
<td>1.2x10$^7$</td>
<td></td>
</tr>
<tr>
<td>Immunofluorescent Microplate Assay</td>
<td>9.90x10$^4$</td>
<td>2.4x10$^8$</td>
<td>8.0x10$^6$</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3.3 Comparison of various titration assays for titering egg-passaged B/Lee/40

- a Titre in plaque forming units/ml (pfu/ml). b – denotes that no plaque formation was observed.
Figure 3.10 Titration of influenza B/Lee/40 virus. Virus inoculum was serially diluted ($10^0$-$10^3$) in phosphate buffer saline and used to infect confluent MDCK monolayers on glass coverslips in the absence (-Trypsin) and presence (+Trypsin) of TPCK-treated Trypsin. The inoculum was removed after 1hr incubation at 37°C and replaced with media supplemented with serum. Cells were fixed 20hrs post-infection and stained as described in methods. Fluorescent cells were viewed and counted under a fluorescence microscope. Images were taken at 20x magnification. Inset: magnified view of an infected cell observed at $10^{-2}$ dilution (boxed). * indicates zones of clearings in the monolayer of the virus-infected cells. Mock, represents uninfected cells. Circled: background fluorescence due to non-specific staining of antibody.
Figure 3.1 Titration of influenza A/WSN/33 (H1N1) virus. Virus inoculum was serially diluted ($10^0$-$10^3$) in phosphate buffer saline and used to infect confluent MDCK monolayers on glass coverslips in the absence (-Trypsin) and presence (+Trypsin) of TPCK-treated trypsin. The inoculum was removed after 1hr incubation at 37˚C and replaced with media supplemented with serum. Cells were fixed 20hrs post-infection and stained as described in methods. Fluorescent cells were viewed and counted under a fluorescence microscope. Images were taken at 20x magnification. Inset: magnified view of an infected cell observed at $10^3$ dilution (boxed). Mock, represents uninfected cells. Circled: background fluorescence due to non-specific staining of antibody.
3.3 Discussion

MDCK cells were chosen for primary cultivation of VTMs due to its inherent deficient interferon-induced antiviral response (Frensing et al., 2011). While the study by Frensing et al., showed that influenza B virus infection as well as propagation in MDCK cells is highly efficient, it is important to note that the studies done were on a lab-adapted strain B/Malaysia/2506/2004. The results obtained by infection with lab-adapted strains cannot be directly applied to clinical specimens. As it can be seen from section 3.2.1, only 14 out of positively-tested 20 specimens could be passaged. Moreover, the 3 specimens which could be passaged in cell culture could not propagate well in embryonated chicken eggs. 2 of these egg-passaged clinical specimens required the shell-vial method to establish infection. Even then, only a small percentage of cells were infected (Figure 3.4).

These findings suggest that cell culture may not necessarily be the best way to characterize clinical specimens. These viruses have adapted to human hosts and could have gained certain mutations which do not allow them to propagate well in cell culture. Currently, PCR-based assays are the main technique used in diagnostic laboratories. PCR based techniques can allow us to determine the viral load of a sample and also determine the specific sub-type or lineage of the specimen (Biere et al., 2010; Hernes et al., 2012). However, certain characteristics of viruses are not reflected in PCR-based assays. These include growth kinetics and susceptibility of the virus to drugs. Culture-based assays are still the only method which can determine such characteristics. It has been found that culture-based assays are less sensitive than PCR-based assays (Cheng et al., 2004). The results of our current study agree with this data, showing that specimens detected as positive through RT-PCR may not be able to be passaged through cell-culture. More studies need to be carried out to improve culturing of clinical specimens.

Unlike the clinical specimens, B/Lee/40 could be successfully propagated and amplified in embryonated chicken eggs (Figure 3.5). A549 cells were chosen to characterize the infection of egg-passaged B/Lee/40 as they are adenocarcinomic human alveolar basal epithelial cells, similar to respiratory cells which influenza B virus infects naturally (Giard et al., 1973). Unlike MDCK cells, the A549 cells do not have a deficient interferon-induced antiviral response (Sutejo et al., 2012), this could
account for the small proportion of uninfected cells even when infected at a high titre of B/Lee/40 (Figure 3.5, red box).

When separated on a sucrose gradient, egg-passaged B/Lee/40 displayed 2 opalescent bands, migrating at 30% and 40% (Figure 3.6). The 30% band was thinner than the 40% band (Figure 3.6) and also displayed less intense proteins on an SDS PAGE gel compared to the 40% band (Figure 3.7). Infectious viruses could be isolated from both bands, but the titre of the viruses from the 30% band was ten times lower than the viruses isolated from the 40% (data not shown). Since viruses can be isolated from the 30% band, it could be possible that the 30% band contains a pleomorphic form of influenza B. Since the viruses isolated from the 40% band were spherical (Figure 3.8), the viruses migrating at 30% could be filamentous or kidney shaped, as described in Chapter 1 (Mellema et al., 1981; Ruigrok et al., 1984; Servidei et al., 1987; Waterson et al., 1963). Further analysis, especially electron microscopy should be carried out to determine the morphology of the viruses in the 30% band.

The naming of the protein bands in figure 4.6 was done by comparing the sizes of the prominent bands to the known sizes of the influenza B proteins. To confirm the identity of these protein bands, mass spectrometry (MS) analysis should be carried out. However, these bands were only visible under Sypro-Ruby staining. This makes it difficult to determine the exact location of the bands needed to be excised for MS.

Despite being able to infect MDCK cells, egg-passaged B/Lee/40 cannot be recovered from the supernatant of infected MDCK cells. Even the original inoculum of B/Lee/40 purchased from ATCC could infect cells but no infectious particles could be recovered from the supernatant of infected cells. This inability to be passaged in cell culture is very similar to the clinical specimens. A possible explanation for this inability could be due to the lack of activation of the HA0 precursor protein. In mammalian hosts, the HA0 precursor protein is cleaved by endogenous proteases such as HAT (human airway trypsin-like protease) and TMPRSS2 (Bottcher-Friebertshauser et al., 2010; Böttcher et al., 2009; Bottcher et al., 2006; Chaipan et al., 2009). These proteases are absent in cell culture and have to be introduced exogenously post infection. The studies about the nature of this activation were done in influenza A viruses, and even though the cleavage site of influenza A and B HA0 is conserved, a different protease or a different mechanism could be involved. Hence, the addition of trypsin to cell culture may not be sufficient to activate influenza B HA0 precursor. Experiments involving
changing the concentration of trypsin or the type of protease added should be carried out to determine the ideal conditions needed to proteolytically activate influenza B HA$_0$ precursor protein.

The incompatibility of trypsin in cell culture could be the reason why egg-passaged B/Lee/40 could not produce plaques in the conventional plaque assay. However, we have decided to circumvent this issue by developing and validating the immunofluorescent microplate assay. In this assay, the number of infectious particles present in the original inoculum is titred without the need of multiple rounds on viral replication. Therefore trypsin is not needed to be added to the culture. This ensures that the numbers of infected cells observed are directly proportional to the titre of the inoculum. The effect of addition of trypsin can be seen in figures 3.9 and 3.10 where more infected cells were observed when trypsin was added to the culture.

Even though the omission of trypsin is vital for this assay to be accurate, it will still not be enough for the titration of Highly Pathogenic Avian Influenza viruses (HPAI). These viruses contain a polybasic cleavage site in their HA$_0$ precursor protein. This would allow the HA protein to undergo proteolytic activation by endogenous host cell furin and furin-like proteases, even in the absence of trypsin. This would result in an increase in viral progeny activation, causing a greater spread of viruses across the monolayer, preventing the accurate measure of the virus titre (Stieneke-Grober et al., 1992).

The titres obtained through the immunofluorescent microplate assay prove to be similar to the titres obtained through the conventional and end point assays. This assay has the following advantages against the other two assays; it is much more rapid in titre acquisition, the titre can be estimated within 12 hours of infection. The other assays require the virus to replicate multiple times before significant pathology or CPE is observed. This may take up to a week of incubation. Since fluorescence is employed in this assay, the kinetics and cellular localisation of the viral proteins can be observed, shedding greater input in the replication of the virus. Since this assay does not require the virus to replicate multiple times, it can be used to measure the titres of clinical specimens which are not adapted to cell culture. This shorter incubation time can allow this assay to be developed as a rapid and high throughput diagnostic assay, proving especially useful in times of an epidemic. This assay can be easily adapted to the use of a 96-well plate, reducing the consumption of laboratory reagents and consumables.
Another titration assay had already been developed for influenza viruses which exploit antigen presentation (Matrosovich et al., 2006). The unique feature of this assay is that it employs the use of a low viscosity overlay, allowing for faster viral spread within plaques, resulting in faster titre acquisition. Despite its obvious advantages, this assay still requires multiple rounds of viral replication. This means, viral strains (such as clinical specimens) which are not adapted to cell culture cannot be titred by this assay. This assay is also limited by the time required for a plaque to grow to an observable size, whereas the immunofluorescent microplate assay simply requires the time for a single round of virus infection.

Lastly, even though fluorescence-based titration assays are not new, we believe that this study is a first in comparing the titres acquired through this assay to more ‘conventional’ assays. We find this assay to be as accurate as the other assays and its advantages make it suitable in titering non-laboratory adapted influenza specimens with a shorter incubation time.
Chapter 4: Genetic and phylogenetic analysis of gene segment 4 encoding for the hemagglutinin (HA) protein

4.1 Introduction

Viruses belonging to the Orthomyxoviridae family have single stranded, negative stranded RNA genomes. These genomes are replicated by the RNA dependent RNA polymerase (RDRP) complex, encoded by the viral genome (Palese, 2007). The RDRP has no proof reading capability resulting in an ‘error-prone’ replication of the RNA genome. On average, the RDRP causes an inclusion of 1 mismatch in every 10000 nucleotides (Drake, 1993). The genome of influenza B is composed of 14538 nucleotides (Fig 1.11). This means that each time influenza B virus replicates, an average of 1 mismatch occurs. These mismatches may include nucleotide substitutions, deletions or even insertions. These changes would accumulate over time and would result in a change in the protein sequence. Such changes may get selected for by the immune pressures of the host resulting in fixation. Such a phenomenon is known as antigenic drift (Hampson, 2002). Even though the antigenic drift in influenza B is slower than influenza A, it is still a main driving force in influenza B evolution as the other mechanism of viral evolution; antigenic shift is not observed in influenza B viruses (Chen and Holmes, 2008; Hay et al., 2001).

As specified in Chapter 1, another mechanism of influenza B evolution is through genetic reassortment. This means that circulating viruses may have different combinations of gene segments from different lineages (McCullers et al., 2004; McCullers et al., 1999; Nerome et al., 1998). A virus that contain gene segments from different lineages is known as a reassortant, (Figure 1.1).

The main glycoprotein of influenza B viruses is hemagglutinin (HA) (Palese, 2007). This protein is abundantly displayed on the exterior of the virus and is the main antigen recognised by the host immune system. It is therefore important that the lineages of a circulating virus are compatible with that of the most recent recommended vaccine strain as individuals vaccinated with one lineage are not immunized against the other (Levandowski et al., 1991b).

The coding region of gene segment 4 was sequenced from influenza B viruses isolated from military personnel of the SAF between the years of 2004-2009. Phylogenetic analysis of these sequences was carried out to allow us to observe the clustering pattern of these sequences, allowing inference of their respective lineages, in
comparison to the bi-annual recommended vaccine strains. This would give us an assessment of the vaccine efficacy. Added in the phylogenetic tree were the HA sequences of viruses circulating within the same region as Singapore as well as previously sequenced Singaporean strains.

The predicted amino acid sequence of the HA protein was inferred by translating the sequenced gene segment in silico. These sequences were aligned with the HA protein of other circulating strains as well as previously sequenced Singaporean influenza B viruses. Several domains and motifs of the HA protein have been highlighted in the literature and are analysed in this chapter.

Figure 4.1 displays a schematic representation of the HA protein. The numbering applied refers to the amino acid sequence of the HA protein of B/Lee/40 (The UniProt Consortium, 2012).

![Figure 4.1 Schematic representation of the influenza B HA protein.](image)

584 represents the number of amino acids which make up the length of this protein.

4.2 Results

4.2.1 Sequencing of gene segment 4

The reference strain (B/Lee/40) was used in optimizing the PCR thermocycling conditions used to amplify the desired PCR products. Touchdown PCR was chosen as it reduces amplifying nonspecific sequences within the template. Figure 4.2 shows the
sizes of each of 3 overlapping fragments amplified by the thermocycling conditions using the primers described in the methods section.

**Figure 4.2** Sizes of the overlapping HA ORF fragments generated in this study analysed on an ethidium bromide stained 1% agarose gel.

Out of a total of 46 specimens isolated, the full gene sequence of the HA gene could be attained from 41. To confirm that the sequences generated were of influenza B origin, the consensus sequences were entered into the query field of the BLAST search tool. Table 4.1 lists the top hits for the HA gene sequences ([http://blast.ncbi.nlm.nih.gov/Blast.cgi](http://blast.ncbi.nlm.nih.gov/Blast.cgi)).

<table>
<thead>
<tr>
<th>Accession No</th>
<th>Description</th>
<th>Max Score</th>
<th>Total Score</th>
<th>Query Coverage</th>
<th>E Value</th>
<th>Max Ident</th>
</tr>
</thead>
<tbody>
<tr>
<td>CY033804.1</td>
<td>Influenza B virus (B/Taiwan/94/2005) segment 4, complete sequence</td>
<td>3221</td>
<td>3221</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
<tr>
<td>CY040449.1</td>
<td>Influenza B virus (B/Malaysia/2506/2004) segment 4, complete sequence</td>
<td>3217</td>
<td>3217</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
<tr>
<td>CY037391.1</td>
<td>Influenza B virus (B/Tennessee/UR06-0406/2007) segment 4, complete sequence</td>
<td>3190</td>
<td>3190</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
<tr>
<td>CY018365.1</td>
<td>Influenza B virus (B/Florida/02/2006) segment 4, complete sequence</td>
<td>3190</td>
<td>3190</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
</tbody>
</table>

**4.2.2 Nucleotide sequence and phylogenetic analysis of gene segment 4.**

The HA sequences of the locally circulating influenza B strains from years 2004-2009 shared a sequence similarity between 70.1%-100%. However if the alignment of these sequences did not include the specimens isolated in 2009, the similarity between them would be between 95%-100%. When extended to include the HA sequences of the highly related strains (downloaded from the BLAST search), the sequences share a similarity between 68.2%-100%. Again, if the 2009 sequences were omitted, the sequence similarity would be between 89%-100%. The HA gene segment
of the 4 2009 specimens share a sequence similarity between 99.8%-100%, whilst having a sequence similarity of 68.2%-73.6% with the rest of the sequences in this study.

The phylogenetic tree generated from the alignment of the HA sequences is displayed in figure 4.3. The HA phylogenetic tree shows 2 distinct clusters corresponding to the Victoria87 and Yamagata88 lineages. Each of these lineages are shown separately for ease of viewing; Victoria87 (Figure 4.3B) and Yamagata88 (Figure 4.3C).

All except for one of the specimens sequenced in this study clustered within the Victoria87 lineage (Figure 4.3B). The sequence of the HA gene from DSO_010147_2007 clustered in the Yamagata88 lineage and shares a direct node with the HA of the vaccine strain of 20012-2013 B/Wisconsin/01/2010 (Figure 4.3C). All of the Singaporean strains sequenced prior to the year 2000 as well as a minority of Singaporean strains sequenced in 2010 and 2011 belonged to the Yamagata lineage (Figure 4.3C).

All of the strains in the Victoria cluster were isolated post 2000. The locally circulating clinical specimens sequenced in this study share high homology with other specimens isolated within the same year. All the specimens isolated in 2009 (●) clustered amongst themselves, while most of the specimens isolated in years 2005 (●) and 2006 (●) clustered together (Figure 4.3B).

4.2.3 Amino acid sequence analysis of gene segment 4, HA0 precursor protein

The HA gene segment, is monocistronic, coding for only the HA0 precursor protein. This protein is later cleaved into its functional peptides HA1 and HA2 (Krystal et al., 1982; Lazarowitz et al., 1973; Wiley and Skehel, 1987). An alignment of the HA amino acid sequences of the clinical specimens and the downloaded sequences showed that the sequences shared a sequence similarity of 89.6%-100% (Supp. Fig.1). Amongst the clinical specimen, the sequence similarity ranged between 92.1%-100%.

The main neutralising epitope of HA from the Victoria lineage lies between residues 178-185 and this is known as the ‘tip’ (Table 4.2) (Nakagawa et al., 2006). The amino acid sequence of the specimens sequenced in this study as well as other Singaporean strains of Victoria lineage is identical to B/Victoria/87. Only 2 substitutions were observed in the vaccine strains: B/Hong Kong/330/2001 (E179D) and B/Brisbane/60/2008 (N180K) (Table 4.2).
A.

HA Phylogenetic Tree
B. 

HA Victoria87 Lineage
Figure 4.3 A. HA phylogenetic tree. Sequences, both local and global, clustered into either the Yamagata88 or Victoria87 lineages. B. Magnified view of the Victoria87 lineage. C. Magnified view of the Yamagata88 lineage. Specimens sequenced in this study isolated in 2004. Specimens sequenced in this study isolated in 2005. Specimens sequenced in this study isolated in 2006. Specimens sequenced in this study isolated in 2007. Specimens sequenced in this study isolated in 2009. Singaporean specimens not sequenced in this study. Vaccine strains.

Table 4.3 lists the variability displayed at the main neutralising epitope of the Yamagata-like strains known as the ‘loop’, between residues 156-164 (Nakagawa et al., 2005). DSO_010147_2007 has 2 substitutions in this epitope: V161A and R164K (Table 4.3).

All of the specimens sequenced in this study displayed a glycosylation site at position 211 except for DSO_050629_2005 (NEI) (Nakagawa et al., 2004) (Table 4.4). The same glycosylation site was observed even in Singaporean specimens isolated in 2010-2011. B/Singapore/222/1979 and B/Singapore/1964 did not have this
glycosylation site (TET and NEI), while, 3 of the 5 Singaporean specimens isolated in
the 90s contained the glycosylation site (NKT) sequences.

Table 4.2 Amino acid substitution of the Victoria ‘Tip’ of the HA protein. Specimens belonging to the Yamagata
lineage are listed under Yamagata 88. Specimens belonging to the Victoria lineage are listed under Victoria 87.
Similar residues are labeled by the dot (.). The dash (-) represents a deletion. DSO Victoria: Specimens sequenced in
this study belonging to the Victoria lineage. (Yam) Singapore 2010-2011: Singaporean strains isolated in 2010-2011
belonging to the Yamagata lineage. (Vic) Singapore 2010-2011: Singaporean strains belonging to the Victoria
lineage.

<table>
<thead>
<tr>
<th>Yamagata 88</th>
<th>D</th>
<th>N</th>
<th>K</th>
<th>T</th>
<th>A</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSO 010147 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Singapore 04 1991</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Singapore 11 1994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Singapore 22 1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Singapore 31 1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Singapore 35 1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Yam) Singapore 2010-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacc0405N 05S B Jilin 2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacc0506N B Jiangsu 10 2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacc0809NS B Florida 4 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacc1213N B Wisconsin 01 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Victoria 87</td>
<td>N</td>
<td>D</td>
<td>N</td>
<td>N</td>
<td>K</td>
<td>T</td>
</tr>
<tr>
<td>DSO Victoria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Vic) Singapore 2010-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacc04S B Hong Kong 330 2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacc0608NS B Malaysia 2506 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacc0912NS B Brisbane 60 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3 Amino acid substitution of the Yamagata ‘Loop’ of the HA protein. Specimens belonging to the
Yamagata lineage are listed under Yamagata 88. Specimens belonging to the Victoria lineage are listed under
Victoria 87. Similar residues are labeled by the dot (.). DSO Victoria: Specimens sequenced in this study belonging
to the Victoria lineage. (Yam) Singapore 2010-2011: Singaporean strains isolated in 2010-2011 belonging to the

<table>
<thead>
<tr>
<th>B Yamagata 16 88</th>
<th>G</th>
<th>S</th>
<th>C</th>
<th>P</th>
<th>N</th>
<th>V</th>
<th>T</th>
<th>S</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSO 010147 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
</tr>
<tr>
<td>Vacc0405N 05S B Jilin 2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
</tr>
<tr>
<td>Vacc0506N B Jiangsu 10 2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
</tr>
<tr>
<td>Vacc0809NS B Florida 4 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
</tr>
<tr>
<td>Vacc1213N B Wisconsin 01 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
</tr>
<tr>
<td>(Yam) Singapore 2010-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
</tr>
<tr>
<td>B Singapore 04 1991</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>B Singapore 11 1994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Singapore 22 1998</td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
</tr>
<tr>
<td>B Singapore 31 1998</td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>K</td>
</tr>
<tr>
<td>B Singapore 35 1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>B Singapore 222 1979</td>
<td></td>
<td></td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G</td>
</tr>
<tr>
<td>B Singapore 1964</td>
<td></td>
<td></td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G</td>
</tr>
<tr>
<td>B Victoria 2 87</td>
<td>G</td>
<td>S</td>
<td>C</td>
<td>P</td>
<td>N</td>
<td>V</td>
<td>T</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>DSO Victoria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacc04S B Hong Kong 330 2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacc0608NS B Malaysia 2506 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>Vacc0912NS B Brisbane 60 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>(Vic) Singapore 2010-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

92
### Table 4.4 Amino acid substitution of the glycosylation site 211 of the HA protein.
Specimens belonging to the Yamagata lineage are listed under Yamagata 88. Specimens belonging to the Victoria lineage are listed under Victoria 87. Similar residues are labeled by the dot (.). The dash (-) represents a deletion. X: unknown amino acid residue.

<table>
<thead>
<tr>
<th>B Yamagata 16 88</th>
<th>211</th>
<th>213</th>
</tr>
</thead>
<tbody>
<tr>
<td>B Yamagata 16 88</td>
<td>D</td>
<td>K</td>
</tr>
<tr>
<td>DSO 010147 2007</td>
<td>N</td>
<td>.</td>
</tr>
<tr>
<td>Vacc0405N 05S B Jiliin 2003</td>
<td>N</td>
<td>.</td>
</tr>
<tr>
<td>Vacc0506N B Jiangsu 10 2003</td>
<td>N</td>
<td>.</td>
</tr>
<tr>
<td>Vacc0809NS B Florida 4 2006</td>
<td>N</td>
<td>.</td>
</tr>
<tr>
<td>Vacc1213N B Wisconsin 01 2010</td>
<td>N</td>
<td>.</td>
</tr>
<tr>
<td>(Yam) Singapore 2010-2011</td>
<td>N</td>
<td>.</td>
</tr>
<tr>
<td>B Singapore 04 1991</td>
<td>N</td>
<td>.</td>
</tr>
<tr>
<td>B Singapore 22 1998</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>B Singapore 222 1979</td>
<td>T</td>
<td>E</td>
</tr>
<tr>
<td>B Singapore 1964</td>
<td>N</td>
<td>E</td>
</tr>
<tr>
<td>B Victoria 2 87</td>
<td>N</td>
<td>E</td>
</tr>
<tr>
<td>DSO Victoria</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>DSO 050629 2005</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>Vacc04S B Hong Kong 330 2001</td>
<td>S</td>
<td>T</td>
</tr>
<tr>
<td>Vacc0608NS B Malaysia 2506 2004</td>
<td>.</td>
<td>X</td>
</tr>
<tr>
<td>Vacc0912NS B Brisbane 60 2008</td>
<td>.</td>
<td>T</td>
</tr>
<tr>
<td>(Vic) Singapore 2010-2011</td>
<td>D</td>
<td>K</td>
</tr>
<tr>
<td>Lee40</td>
<td>K</td>
<td>T</td>
</tr>
</tbody>
</table>

#### 4.3 Discussion

The full length of the HA gene was selected in this study instead of only the HA₁ gene as it has been suggested that the evolution of Influenza B viruses is not subjected to antibody selection (Air et al., 1990), therefore phylogenetic analyses of Influenza B viruses should not be based on the most ‘exposed’ surfaces of the protein but rather the full length of the gene (McCullers et al., 1999). However, our finding that 33 out of the 42 amino acid mutations occur within the HA₁ peptide seems to challenge the suggestion that the entire protein is mutating at a uniform rate. Moreover, the clustering pattern of the phylogenetic tree using the HA₁ gene only did not differ compared to the phylogenetic tree using the full length of the HA gene segment.

Amongst the 41 specimens analysed in figure 4.3, only the HA of specimen DSO_010147_2007 belonged to the Yamagata lineage (Supp. Fig 1). Similarly, 3 out of the 11 Singaporean specimens isolated in 2010-2011 also belonged to the Yamagata lineage. While the majority of the specimens isolated post 2000 had their HA gene segment belonging to the Victoria lineage, it is interesting to note that all the Singaporean specimens isolated before 2000 were of the Yamagata lineage. The Singaporean specimens which were sequenced and deposited in GenBank (including from this study) may not necessarily be representative of all of the circulating viruses in
Singapore as sampling bias would be significant in small sample sizes. However, the changes in the proportions of the 2 different lineages do suggest that several reassortment events have occurred over the past 50 years in Singapore. The finding that 1 out of 41 specimens isolated in 2004-2009 and 3 out of 11 isolated in 2010-2011 belonged to the Yamagata lineage, alludes to the re-emergence of the Yamagata lineage. Wider surveillance and scrutiny over the circulating viruses in Singapore, as well as regionally, may allow us to study the dynamics involved in this purported re-emergence.

The close clustering of specimens isolated within the same year follows the model of viral evolution where lesser variation exists within one influenza season as compared to between influenza seasons, being a product of changes in herd immunity between the seasons (Chen and Holmes, 2008).

The deletion of asparagine at position 178 of the HA protein is a signature of the Yamagata lineage (McCullers et al., 1999). As expected, DSO_010147_2007, belonging to the Yamagata lineage has a deletion at this location amongst the specimens sequenced in this study (Table 4.2).

The HA ‘tip’ is the main neutralising epitope present in the strains of the Victoria lineage (Nakagawa et al., 2006). The specimens sequenced in this study belonging to the Victoria lineage have 100% identity to the HA of B/Victoria/2/87. Singaporean specimens sequenced in 2010-2011 have a N180K substitution which is seen the vaccine strain B/Brisbane/60/2008 (Table 4.2).

The HA ‘tip’ is not conserved amongst the Yamagata lineage. DSO_010147_2007 was the only specimen belonging to the Yamagata lineage harboring N181Y substitution which was only seen in B/Wisconsin/01/2010. This suggests that Y181 has been integrated into the influenza B HA protein only recently (Table 4.2).

The ‘loop’ of the influenza B HA protein has been found to the main antibody neutralising epitope amongst the Yamagata lineage (Nakagawa et al., 2005). The two substitutions shown by DSO_0101147_2007 in table 4.3 are conserved in all the other specimens of this lineage post 1991 (Table 4.3). Strains harboring K164 have been shown to have the same antibody reactivity as strains displaying R164 (Nakagawa et al., 2005). Further functional analyses need to be carried out to determine if the V161A substitution would bear any antigenic significance. Two new substitutions were
observed in the alignment. It has been reported that R\textsubscript{156} and G\textsubscript{164} would render the epitope unrecognizable. While none of the sequenced specimens had these substitutions, the vaccine strain B/Jiangsu/10/2003 displayed R\textsubscript{156}, while B/Singapore/222/1979 and B/Singapore/1964 had G\textsubscript{164}.

A glycosylation site at position 211(NET) was observed in most of the specimens isolated in Japan post 2002 (Nakagawa et al., 2004) (Table 4.4). Prior to 2002, position 211 displayed a variety of sequences (NEA, KET, NEN and NET). Interestingly, the HA gene segment of DSO\textsubscript{050629}\textsubscript{2005} belonged to the Victoria lineage and this glycosylation site was initially found only in Japanese strains belonging to the same lineage. This suggests that position 211 of the HA protein is probably exposed, and found located on the exterior of the globular protein. The folding of the protein at this position is also independent of its lineage. All the vaccine strains except for 2 (B/Florida/4/2006 and B/Hong Kong/330/2001) did not contain this potential glycosylation site. The amino acid at position 213 for B/Malaysia/2506/2004 is unknown, as the codon at that position has the sequence AYC (where Y= C or T) therefore it is not possible to determine if this protein has the glycosylation site (Table 4.4). Further analysis of the specimens used to construct the HA phylogenetic tree revealed that all of the regional specimens contained this potential glycosylation site, except for B/Taiwan/72068/2004, B/Taiwan/91061/2005, B/Taiwan/2894/2006, B/Taiwan/71523/2007, B/Malaysia/1749642/2007 and B/Hong Kong/310/2004 (KKT, SET, DET, SKT, SET and KET respectively). It is possible that this glycosylation site originated in Japan prior to its spread worldwide; however since the glycosylation site at position 211 was also seen in Singapore as early as 1994, it suggests that the origin of this glycosylation site occurred independent to circulating Japanese specimens. Studies in influenza A viruses have led to the suggestion that an introduction of a glycosylation site may mask an antibody-binding epitope, leading to an antigenic change (Schulze, 1997). This further suggests that a similar masking may take place at this glycosylation position for the influenza B virus HA protein.
Chapter 5: Genetic and phylogenetic analysis of gene segment 6 encoding for the neuraminidase (NA) and NB proteins.

5.1 Introduction

Gene segment 6 is 1544 nucleotides long and is bicistronic (Palese, 2007). The main protein encoded by gene segment 6 is neuraminidase (NA). This protein is 466 amino acids long and serves as a sialidase, cleaving sialic acid upon viral release (Palese et al., 1974). This protein is the target of neuraminidase inhibitors (NAI), a class of drugs which act as sialic acid analogs, binding to NA, preventing successful sialidase activity (Figure 1.16). Mutations in both the HA and NA proteins have been found within circulating strains that result in resistance to NAIs. Increasing evidence in the literature suggests that influenza B viruses gain resistance to NAIs faster than influenza A viruses.

While, not as dominant as the HA protein, the NA protein is also present on the exterior of the virion as a glycoprotein. The amount of NA protein expressed in the exterior causes it to induce significant antigenicity, such that vaccination with just the NA protein alone is enough to confer immunization to live viruses (Levandowski et al., 1991b; Quan et al., 2012). This, together with the reasons listed in the previous paragraph, strongly calls for the need for in-depth surveillance of influenza viruses, an effort which this body of work hopes to contribute to.

The NB protein is the second protein encoded for by this gene segment. The start codon for the NB gene lies in the +1 open reading frame (ORF), 7 nucleotides 5’ to the start codon of the NA start codon (Figure 1.17). This protein was discovered in 1983 and has no homolog in influenza A viruses (Shaw et al., 1983). It has been found that the NB is a structural protein, has ion channel activity when placed in an artificial membrane, which can be inhibited by amantadine, and is post-translationally modified at N3 and N7 (Fischer et al., 2001; Sunstrom et al., 1996; Williams and Lamb, 1988). This protein is a type III membrane protein and a putative transmembrane domain lies between amino acids 18-40 (Betakova et al., 1996). Mutant viruses which do not express NB can replicate to similar titers in tissue culture but are heavily attenuated in animal models, adding to the mystery of the NB function (Hatta and Kawaoka, 2003).

The ORF of the NA gene was sequenced as described in section 2.3.13 (Figure 2.1). The gene segments were aligned with other influenza B strains and analysed phylogenetically. The nucleotide sequences were then translated in silico and the amino
acid sequences of the 2 proteins were analysed as well. Owing to the fact that the primers used were initially intended to sequence the NA protein, only partial information of the NB protein was attained. Figure 5.1 shows a schematic representation of the NA and NB proteins. The amino acid positions shown refer to the sequence of the NA and NB proteins of B/Lee/40 (The UniProt Consortium, 2012).

Figure 5.1 Schematic representation of the influenza B NA and NB proteins. 466 and 100 represents the number of amino acids which make up the proteins.
5.2 Results

5.2.1 Sequencing of gene segment 6

Similar thermocycling conditions described in section 4.2.1 were used to PCR amplify the 3 overlapping fragments of the NA ORF. Figure 5.2 shows the sizes of these fragments. Viable sequence information could be attained from 44 out of the 46 clinical specimens. To test if the sequences generated in this study were correct, select sequences were queried through the BLAST search online tool. Table 5.1 lists the top hits for the NA genes sequenced in this study.

![Image of gel with NAF1, NAF2, and NAF3 labels](Image)

**Figure 5.2** Sizes of the overlapping NA ORF fragments of B/Lee/40 generated in this study analysed on an ethidium bromide stained 1% agarose gel. bp. base pairs.

<table>
<thead>
<tr>
<th>Accession No</th>
<th>Description</th>
<th>Max Score</th>
<th>Total Score</th>
<th>Query Coverage</th>
<th>E Value</th>
<th>Max Ident</th>
</tr>
</thead>
<tbody>
<tr>
<td>CY040451.1</td>
<td>Influenza B virus (B/Malaysia/2506/2004) segment 6, complete sequence</td>
<td>2606</td>
<td>2606</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
<tr>
<td>CY033870.1</td>
<td>Influenza B virus (B/NYMC BX-17(Lee/1940-Malaysia/2506/2004)) segment 6, complete sequence</td>
<td>2606</td>
<td>2606</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
<tr>
<td>CY018735.1</td>
<td>Influenza B virus (B/Hong Kong/310/2004) segment 6, complete sequence</td>
<td>2606</td>
<td>2606</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
<tr>
<td>CY015389.1</td>
<td>Influenza B virus (B/New Jersey/1/2006) segment 6 sequence</td>
<td>2540</td>
<td>2540</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
</tbody>
</table>
5.2.2 Nucleotide sequence and phylogenetic analysis of gene segment 6.

The 44 clinical specimens yielding sequence information for the NA gene segment shared a sequence similarity between 85.6%-100% at the nucleotide level. Unlike the HA gene segments, the sequence similarity does not change upon the removal of the NA sequences of the 4 2009 specimens. When the rest of the downloaded NA sequences are included in the alignment, the sequence similarity shared is between 80%-100%.

The 2 main lineages of influenza B viruses are seen as 2 distinct clusters in the NA phylogenetic tree. However, only the NA sequences from strains; B/Hong Kong/330/2001 and B/Singapore/222/1979 clustered in the Victoria cluster. All of the NA sequences from locally isolated specimens after the nineties till the present were present in the Yamagata cluster, including specimen DSO_010147_2007. For ease of analysis, two clusters within the Yamagata88 are highlighted as B and C and are expressed individually in figures 5.3B and 5.3C (Figure 5.3).

The NA gene from specimen DSO_010147_2007 clustered together with the NA gene from the vaccine strain of 2012-2013, B/Wisconsin/01/2010. This is similar to that of it’s HA gene. None of the other clinical specimens clustered together with B/Wisconsin/01/2010 (Figures 4.3 & 5.3).

All of the NA sequences of the specimens isolated in 2009 (●) clustered together, just as in the HA phylogenetic tree (Figure 4.3). However, unlike their HA genes, the NA genes did not share 100% similarity between them. All of the NA genes of the specimens isolated in 2004 clustered together and most of the NA genes of specimens isolated in years 2006 and 2007 clustered together as well (Figure 5.3). A strong clustering pattern was observed between specimens isolated in 2005 and 2006 (●, ●), all 6 of the specimens isolated in 2004 (●) clustered under 1 node (Figure 5.3).

The NA genes of specimens; DSO_020132_2007, DSO_020147_2007, DSO_070214_2006 and DSO_020114_2007 clustered together with the vaccine strain of 2009; B Brisbane 60 2008. The NA gene of vaccine strain from 2007; B Malaysia 2506 2004 clustered together with majority of the specimens isolated between years 2004-2009 (Figure 5.3).
A. NA Phylogenetic Tree
B.

NA Cluster B
C. NA Cluster C

Figure 5.3A. NA phylogenetic tree. B. Magnified view of cluster B (bracketed in A.). C. Magnified view of cluster C (bracketed in A.). Sequences, both local and global, clustered into either the Yamagata88 or Victoria87 lineages.

5.2.3 Amino acid sequence analysis

Gene segment 6 of the influenza B virus is bicistronic; coding for both the NA and NB proteins (Betakova et al., 1996; Brassard et al., 1996). The alignment of the amino acid sequences of the NA protein of clinical specimens and downloaded sequences showed sequence similarity between 82.2%-100%. Amongst the clinical specimens only the NA proteins showed 85.9%-100% amino acid sequence similarity. The NB protein alignment of all the sequences in this study showed 57%-100% sequence similarity. Within the clinical specimen sequenced, the sequence similarity shared remains unchanged.

5.2.3.1 Amino acid sequence analysis of the NA protein

The NA protein of the B Lee strain is 466 amino acids long and the mutations occur uniformly throughout the protein unlike the HA protein where most of the mutations occur in the first 374 amino acids (Supp. Fig 2). The calcium binding site of the NA lies between residues 318 and 350 (Burmeister et al., 1992). Substitutions in this domain are listed in table 5.2. All of the substitutions in this domain are conservative substitutions.
5.2.3.2 Amino acid sequence analysis of the NB protein

Table 5.3 lists the amino acid substitutions observed in the NB protein. All of the amino acid substitutions within residues 19-40, which is postulated to be the transmembrane region of the protein, are conservative. P67S and L92I are only observed in Singaporean strains isolated in 2009. These substitutions are linked. As can be observed, the amino acid profile of DSO_010147_2007 NB was the most unique of all the specimens listed (Supp Fig 3, Table 5.3).

Table 5.2 Amino acid substitution of the calcium binding domain of the NA protein. Specimens belonging to the Yamagata lineage are listed under B Yamagata 16 88. Similar residues are labeled by the dot (.).

<table>
<thead>
<tr>
<th>Specimen</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>B Yamagata 16 1988</td>
<td>E . . R</td>
</tr>
<tr>
<td>(Yam) Rest of DSO Specimens</td>
<td>. . . S</td>
</tr>
<tr>
<td>DSO 050528 2005</td>
<td>D . . S</td>
</tr>
<tr>
<td>DSO 070214 2006</td>
<td>D . . S</td>
</tr>
<tr>
<td>DSO 020114 2007</td>
<td>D . . S</td>
</tr>
<tr>
<td>DSO 020132 2007</td>
<td>D . . S</td>
</tr>
<tr>
<td>DSO 020147 2007</td>
<td>D . . S</td>
</tr>
<tr>
<td>DSO 010147 2007</td>
<td>D N . S</td>
</tr>
<tr>
<td>B Singapore 1H 2010</td>
<td>D . . S</td>
</tr>
<tr>
<td>B Singapore 32C 2010</td>
<td>D . . S</td>
</tr>
<tr>
<td>B Singapore 17H 2010</td>
<td>D . . S</td>
</tr>
<tr>
<td>B Singapore 57H 2010</td>
<td>D . . S</td>
</tr>
<tr>
<td>B Singapore 65H 2010</td>
<td>D . . S</td>
</tr>
<tr>
<td>B Singapore 68C 2010</td>
<td>. N E S</td>
</tr>
<tr>
<td>B Singapore 77C 2011</td>
<td>D . . S</td>
</tr>
<tr>
<td>B Singapore 71H 2011</td>
<td>D . . S</td>
</tr>
<tr>
<td>B Singapore 90C 2011</td>
<td>N . . S</td>
</tr>
<tr>
<td>B Singapore 83C 2011</td>
<td>. N E S</td>
</tr>
<tr>
<td>B Singapore 87H 2011</td>
<td>. N E S</td>
</tr>
<tr>
<td>B Singapore 11 1994</td>
<td>. . . S</td>
</tr>
<tr>
<td>B Singapore 22 1998</td>
<td>. . . S</td>
</tr>
<tr>
<td>B Singapore 31 1998</td>
<td>. . . S</td>
</tr>
<tr>
<td>B Singapore 35 1991</td>
<td>. . . S</td>
</tr>
<tr>
<td>B Singapore 222 1979</td>
<td>. . . L</td>
</tr>
<tr>
<td>Vacc04S B Hong Kong 330 2001</td>
<td>. . . S</td>
</tr>
<tr>
<td>Vacc0405N 055 B Jilin 20 2003</td>
<td>. . . S</td>
</tr>
<tr>
<td>Vacc0405N 055 B Shanghai 361 2002</td>
<td>. . . S</td>
</tr>
<tr>
<td>Vacc0608NS B Malaysia 2506 2004</td>
<td>. . . S</td>
</tr>
<tr>
<td>Vacc0506N B Jiangsu 10 2003</td>
<td>. . . S</td>
</tr>
<tr>
<td>Vacc0809NS B Florida 4 2006</td>
<td>. N . S</td>
</tr>
<tr>
<td>Vacc0912NS B Brisbane 60 2008</td>
<td>D N . S</td>
</tr>
<tr>
<td>B Lee 40</td>
<td>K D K L</td>
</tr>
<tr>
<td>B Victoria 02 1987</td>
<td>E D K R</td>
</tr>
</tbody>
</table>
5.2.4 Molecular characterisation of the NB protein isolated from non-tissue culture adapted influenza B clinical specimens

5.2.4.1 Molecular Cloning of the NB protein and Expression by Plasmid DNA Transfection

The amino acid sequences of the NB protein analysed in section 5.2.3.2 (Table 5.5) can be grouped according to their year of isolation. Strains representing each year were chosen for cloning of the NB protein. The strains names are listed in table 5.4. Owing to the unique sequence of specimen DSO_010147_2007, this strain was also cloned despite of specimen DSO_020132_2007 being the representative strain for 2007. Table 5.4 also lists the code by which these specimens will be referred to throughout this section.

Table 5.3 Amino Acid Substitutions of NB protein. Specimens belonging to the Yamagata lineage are listed under B Yamagata 1688. Similar residues are labeled by the dot (.)

<table>
<thead>
<tr>
<th>Specimen</th>
<th>Transmembrane Domain</th>
<th>Ecto Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>B Yamagata 16 1988</td>
<td>V I T F Y V T V F N P L</td>
<td></td>
</tr>
<tr>
<td>Rest of DSO Specimens</td>
<td>I . . . I . . . P Y . .</td>
<td></td>
</tr>
<tr>
<td>DSO 070214 2006</td>
<td>I . . . I . I L . . .</td>
<td></td>
</tr>
<tr>
<td>DSO 050091 2006</td>
<td>I . I . I I L . . .</td>
<td></td>
</tr>
<tr>
<td>DSO 050528 2005</td>
<td>I . . . I . . P Y . F</td>
<td></td>
</tr>
<tr>
<td>DSO 0003 2009</td>
<td>I . . . I . . Y S I</td>
<td></td>
</tr>
<tr>
<td>DSO 0005 2009</td>
<td>I . . . I . . Y S I</td>
<td></td>
</tr>
<tr>
<td>DSO 0070 2009</td>
<td>I . . . I . . Y S I</td>
<td></td>
</tr>
<tr>
<td>DSO 0100 2009</td>
<td>I . . . I . . Y S I</td>
<td></td>
</tr>
<tr>
<td>Vacc04S B Hong Kong 330 2001</td>
<td>. . . L . . . . . F</td>
<td></td>
</tr>
<tr>
<td>Vacc0405N 05S B Jilin 20 2003</td>
<td>. . . . . . . I . .</td>
<td></td>
</tr>
<tr>
<td>Vacc0405N 05S B Shanghai 361 2002</td>
<td>. . . . . . . I . .</td>
<td></td>
</tr>
<tr>
<td>Vacc0809NS B Florida 4 2006</td>
<td>. . . . . . I . . .</td>
<td></td>
</tr>
<tr>
<td>Vacc1213N B Wisconsin 01 2010</td>
<td>. . . . . . . I . . .</td>
<td></td>
</tr>
<tr>
<td>B Lee 40</td>
<td>I I T L T V I V F N P L</td>
<td></td>
</tr>
<tr>
<td>B Victoria 02 1987</td>
<td>V I T L I V T V F N P L</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.4 shows an alignment of the NB proteins which were cloned into a mammalian expression vector, pCAGGS, between the EcoRI and XhoI restriction sites. This vector expressed genes under the control of a chicken β actin promoter (Niwa et
al., 1991; Patch et al., 2007). Each of these constructs was cloned with a myc epitope tag (EQKLISEEDL) at the C terminus.

Table 5.4 NB protein sequences were grouped according to their similarity and a representative strain is chosen for each group. * Code refers to the abbreviation by which these specimens will be referred to.

<table>
<thead>
<tr>
<th>Year of Isolation</th>
<th>Specimen Name</th>
<th>Code*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940</td>
<td>B/Lee/40</td>
<td>Lee</td>
</tr>
<tr>
<td>2004</td>
<td>DSO_090136_2004</td>
<td>136</td>
</tr>
<tr>
<td>2005/2006</td>
<td>DSO_040117_2006</td>
<td>117</td>
</tr>
<tr>
<td>2007</td>
<td>DSO_020132_2007</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>DSO_010147_2007</td>
<td>147</td>
</tr>
<tr>
<td>2009</td>
<td>DSO_0070_2009</td>
<td>70</td>
</tr>
</tbody>
</table>

To validate the expression of the constructs, Human Embryonic Kidney (HEK) 293T cells were transfected with these plasmids for 20 hours and the cell lysates were analysed through western blotting by probing with an anti-myc antibody (Figure 5.5).

Each of the NB constructs display 4 bands each. Specimens Lee-NB, 136-NB, 117NB, and 70-NB display bands labeled 1, 2, 3 and 5 while specimens 132-NB and 147-NB display bands 1, 2, 4 and 5. The predicted size of the unmodified influenza B NB protein is 11kDa, which corresponds to band 1. If the NB protein is only glycosylated at 1 of the 2 glycosylation sites, it will migrate to 15kDa, which could correspond to bands 2 and/or 3. An NB protein which has both of its sites glycosylated would migrate to 18kDa, band 4. Each of the NB constructs display a smearing from 16kDa to 37kDa corresponding to the glycans on NB protein being modified by the addition of multiple units of lactosamine (Williams and Lamb, 1986; Williams and Lamb, 1988) (Figure 5.5).

Specimens 70-NB and 147-NB were transfected into HEK 293T cells and cell lysates were harvested at 4, 8 and 20 hours post transfection. The cell lysates were analysed by western blotting (Figure 5.6). The band at 11kDa was not visible at all time points but the all the other bands listed in figure 5.5 were visible by 8 hours post transfection. The smearing was not seen at 8 hours post but clearly visible at 20 hours (Figure 5.6).

The cellular localisation properties of each construct was analysed through immunofluorescence. MDCK and A549 cells were transfected with these constructs for 10 hours and cells were then stained with anti-myc and anti GM130 antibodies (Figures 5.7 and 5.8).
Figure 5.4 Cloning of NB gene A. Alignment of the NB representative sequences listed in Table 5.4. CT: Cytoplasmic Tail, TM: Transmembrane, ED: Ectodomain, G: Glycosylation. The TM and the myc epitope tag are highlighted. B. Schematic representation of the cloning strategy employed for these sequences. RE: Restriction Endonuclease. pCAGGS schematic, courtesy of Dr. Loo Liat Hui.

The NB proteins of all the constructs show similar cellular localisation patterns. Strong perinuclear without any intra-nuclear staining is observed. The intensity of the staining decreases away from the nucleus and the staining is most intense at a particular region around the nucleus (represented by solid white arrow, Figures 5.7 & 5.8). This region of intense staining is seen in all the constructs and in both cell types and corresponds with GM130 which is a cis-Golgi marker (Hollow white arrow Figure 5.7 & 5.8). It can be observed that cells expressing the NB protein display filamentous projections on several locations on the periphery of their cell membranes (Inset, Figure 5.8).
Figure 5.5 SDS-PAGE analysis of NB. A. 15% SDS PAGE analysis of HEK 293T cells transfected with Lee-NB, 136-NB, 117-NB, 132-NB, 70-NB and 147-NB. Mock Trans: Cells transfected with empty pCAGGS (pC) vector. Cells were transfected for 20hrs and boiled in Laemlli buffer. Similar volumes were loaded onto the gel and protein bands were detected by probing with anti-myc antibody. Bands are labeled 1, 2, 3, 4 and 5. B. Log MW vs. Rf chart plotted by comparing the Rf and the molecular weights of the molecular weight standards. C. Inferences of the sizes of the 5 bands observed in (A) through the use of the equation of the slope obtained in (B).

<table>
<thead>
<tr>
<th>Band</th>
<th>Rf</th>
<th>LogMW</th>
<th>MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.831461</td>
<td>1.023166</td>
<td>10.54791</td>
</tr>
<tr>
<td>2</td>
<td>0.730337</td>
<td>1.129346</td>
<td>13.46933</td>
</tr>
<tr>
<td>3</td>
<td>0.662921</td>
<td>1.200133</td>
<td>15.85377</td>
</tr>
<tr>
<td>4</td>
<td>0.617978</td>
<td>1.247324</td>
<td>17.67354</td>
</tr>
<tr>
<td>5</td>
<td>0.561798</td>
<td>1.306312</td>
<td>20.24475</td>
</tr>
</tbody>
</table>

Figure 5.6 Time course expression of 70-NB and 147-NB in HEK 293T cells. Cell lysates were harvested at 4, 8 and 20 hours post transfection and boiled in Laemlli buffer. Samples were analysed on 15% SDS-PAGE and bands were visualized by probing with anti myc antibody.
5.2.4.2 Endoglycosidase treatment of the NB protein

To determine the nature of the glycosylation undergone by the NB protein, HEK 293T cells were transfected with Lee-NB, 117-NB and 147-NB. These cell lysates were harvested and were subjected to treatment with Endoglycosidase H (Endo H) and Peptide-N-Glycosidase F (PNGase F) (Figure 5.9).

Untreated samples in figure 5.9 show similar migration to figure 5.5. Treatment with Endo H removes the smearing and all other bands save 15kDa and 11kDa. The extensive smearing was not removed when treated with PNGase F, but similar to Endo H treatment, only the 15kDa and 11kDa bands were visible. The band at 15kDa seems to be enriched in both Endo H and PNGase F treatments. Specimen 147-NB did not display an obvious band at 11kDa, however since this band was observed in figure 5.5, it can be assumed that this is due to low protein expression or unequal loading.

**Figure 5.7** MDCK cells were transfected with each of the 6 NB constructs. Cells were allowed to express the proteins for 10 hours before fixing and staining with anti-myc (upper panel) and anti-GM130 (lower panel) antibodies as described in section 2.4.6. Both upper and lower panels show the same field of view. Solid white arrow: strong perinuclear staining of Lee-NB protein. Hollow white arrow: same region indicated by the solid white arrow but when stained for GM130.
Figure 5.8 A549 cells were transfected with each of the 6 NB constructs. Cells were allowed to express the proteins for 10 hours before fixing and staining with anti-myc (upper panel) and anti-GM130 (lower panel) antibodies as described in section 2.4.6. Both upper and lower panels show the same field of view. Solid white arrow: strong perinuclear staining of Lee-NB protein. Hollow white arrow: same region indicated by the solid white arrow but when stained for GM130. Inset: region of the cell membrane displaying filamentous projections.

5.2.4.3 Cross-linking of NB protein

HEK 293T cells were transfected with Lee-NB and treated with increasing concentrations of a membrane permeable cross-linker, dithiobis(succinimidyl propionate) (DSP). Cell lysates were harvested in non-reducing Laemmli buffer. Cell lysates untreated with DSP, display an additional 3 bands, band F: ~30kDa, band G: ~38kDa and band H: ~58kDa (Figure 5.10). These bands could correspond to a dimer (band F), a trimer (band G) and a tetramer (band H) of the NB protein.
Figure 5.9 Endoglycosidation of HEK 293T cells transfected with Lee-NB, 70-NB and 147-NB. U: Untreated, P: PNGase F Treated, E: Endo H treated. Post treatment, lysates were boiled in Laemlli buffer and analysed on 15% SDS-PAGE. Bands were visualized by western blotting, probed with anti-myc antibody.

Figure 5.10 Cross linking of Lee-NB  
A. HEK 293T cells expressing Lee-NB for 20 hours were treated with 0mM, 0.1mM, 0.5mM, 1mM, 2mM and 5mM of DSP and boiled in non-reducing Laemmli buffer. Bands were visualized by probing with anti-myc antibody. B. Log MW vs. Rf chart plotted by comparing the Rf and the molecular weights of the molecular weight standards in A. C. Inferences of the sizes of bands f, g and h observed in (A) through the use of the equation of the slope obtained in (B).
5.3 Discussion

Figures 5.3 and 4.3 indicate that all of the clinical specimens sequenced in this study are re-assortants with their HA gene segments belonging to the Victoria lineage and their NA gene segments belonging to the Yamagata lineage, except for specimen DSO_010147_2007 which has both of its HA and NA gene segments belonging to the Yamagata lineage. This finding is in agreement with the increased frequency of Victoria-like strains since 1997 globally (Richard et al., 2010). Singaporean strains isolated in the 90s have both gene segments belonging to the Yamagata lineage while B Singapore 222 1979 has its HA gene from the Yamagata lineage and its NA gene from the Victoria lineage (Figures 4.3 & 5.3). This observation prompts the notion that reassortment of influenza B viruses have occurred even as early the late 70s which questions the choice of B/Yamagata/16/1988 and B/Victoria/02/1987 as true representatives of the 2 different lineages.

The change in the lineage patterns of the locally circulating influenza B viruses is an interesting phenomenon. The data could suggest 2 main explanations for this outcome; firstly, 2 main reassortment events occurred in Singapore, where the viruses in the late 70s have swapped their Victoria-like gene segments of its NA to the Yamagata lineage. And the second reassortment event occurred prior to the year 2000 where the HA Yamagata-like gene segment is swapped with the Victoria-like gene segment. Secondly, different lineages of influenza B viruses were introduced in Singapore over the time course 1979-present day.

The observation that DSO_010147_2007 is highly related to the vaccine strain of 2008 while the rest of the clinical specimens from that year are not seems to suggest that the vaccine strain chosen for 2008 is not a good representative of the strains circulating in 2007. Each year, the WHO suggests a vaccine strain for the upcoming year based on the composition of the current year’s circulating strains (WHO, 2008). In its report for the composition of the 2008 vaccine, the WHO had statistics showing that there were two predominant strains of influenza B viruses in circulation worldwide, one hemagglutinin inhibition tests (HI) one strain showed close relation to B/Malaysia/2506/2004 (Victoria lineage) while the other strain was highly related to B/Florida/4/2006 (Yamagata lineage). The vaccine strain, B/Florida/4/2006, was chosen as it was predicted to be the more dominant one in the coming year.
Figures 5.3 and 4.3 show that even though DSO_010147_2007 was closely related to B/Florida/2006, the rest of the clinical specimens, even those isolated in 2007 and 2009, were more closely related to B/Malaysia/2506/2004, of the Victoria lineage. Seeing that the vaccine strain of 2008 is of the Yamagata lineage it would only confer immunity to that lineage only (Levodowski et al., 1991b). The prevalence of the Victoria lineage in Singapore is evident since 2000 till present day; hence a vaccine of the Yamagata lineage would confer very little protection to the local population. The suggestion by Levandowski et al, that the WHO incorporates both lineages into its vaccine seems logical since both lineages seem to be circulating worldwide with almost equal prevalence since the start of the century. (Chi et al., 2003; Chi et al., 2005; Levandowski et al., 1991a).

An interesting observation can be made on the classification of the viruses by the WHO. The WHO carries out HI tests on the viruses collected and compares this with HI tests carried out on the ancestor strains. The WHO then determines the lineages of the collected viruses by the results of this test (WHO, 1999a, 2002, 2007, 2008, 2009). By this method, the WHO has classified both B/Malaysia/2506/2004 and B/Brisbane/60/2008 as belonging to the Victoria lineage, and B/Florida/4/2006 as from the Yamagata lineage.

Data from this study however shows that the NA gene segment from all of the vaccine strains are of the Yamagata lineage, except for B/Hong Kong/330/2001 (Figure 5.3) indicating that the B/Malaysia/2506/2004 and B/Brisbane/60/2008, 2 vaccine strains considered from the Victoria lineage, were actually reassortants. This method of classification of influenza B might give us an insight to the biochemical properties of the HA molecule, but is insufficient to classify the virus. HI tests only consider the the ability of the HA molecule to agglutinate red blood cells in the presence of antibodies which might inhibit the HA activity (Potter and Oxford, 1979). This test does not consider the immunogenicity of the HA as well as the NA molecule which is the second abundant protein in the virion exterior (Palese, 2007). It is also postulated that the NB protein is a structural protein which may be present in the virus exterior, which may mont a significant immune response (Betakova et al., 1996; Brassard et al., 1996; Sunstrom et al., 1996). For a more comprehensive classification, the antigenic properties of the NA molecule should also be considered. A virus can only be considered of a certain lineage if it has all the gene segments of the particular ancestor
strain. Even if the HAs of the vaccine strain and the circulating virus are of the same lineage, it is unclear if the vaccine can confer immunity to a virus which does not share lineage with its other structural proteins. This is especially important to note since reassortment of the gene segments is considered as a major mechanism of influenza B evolution (Hiromoto et al., 2000; Lindstrom et al., 1999; McCullers et al., 1999; Shaw et al., 2002).

The conservative nature of the substitutions observed in the calcium binding site suggests a strong positive selection for this epitope. For example, E320D is seen in 5 other specimens, B/Myanmar/M107/2007, B/Myanmar/M254/2007, B/Taiwan/14/2007, B/Taiwan/2894/2006 and B/Hong Kong/259/2010. D329 is observed in B/Malaysia/27127/2004, B/Malaysia/1899839/2007, B/Malaysia/1919534/2008, B/Taiwan/2171/2004 and B/Taiwan/71523/2007. K343E is unseen in any of the other specimens, but K343R is observed in B/Taiwan/2894/2006 and B/Taiwan/14/2007. R345S substitution is seen in the rest of the specimens analysed in the NA phylogenetic tree except for B/Bangkok/141/1994 which harboured L345. Each of the amino acids listed are hydrophilic and are substituted with another hydrophilic residue, except for R345L in B/Bangkok/141/1994 and this strongly suggests a conservation of phenotype. However, functional and structural assays are needed in order to determine if these substitutions do have an actual effect on calcium binding (Table 5.2, Supp. Fig 2).

None of the previously cited amino acid changes associated to neuraminidase inhibitor (NAI) resistance were found in this study in either the HA or NA proteins (Barnett et al., 1999; Cheam et al., 2004a; Cheam et al., 2004b; Gubareva, 2004; Gubareva et al., 1998; Hatakeyama et al., 2007; Hurt et al., 2006; McKimm-Breschkin et al., 2003; McKimm-Breschkin, 2000; Mishin et al., 2005; Staschke et al., 1995).

The lack of amino acid mutations which are linked to NAI resistance does not fully exclude the possibility of resistant strains circulating in Singapore. More biochemical based assays have to be carried out on these specimens to determine their NAI susceptibility (Supp Fig 1 & 2).

The function of the NB protein, encoded for by gene segment 4, is still undetermined, even though it has been found to be expressed abundantly in infected cells and in the virion (Betakova et al., 1996; Brassard et al., 1996; Shaw et al., 1983). The putative transmembrane region of this protein is believed to be between amino acids 19-40; IRGSIIIICVSLIVILIVFGYIAKIFI (Fischer et al., 2000a), and by using
computational analysis it has been found that amino acids S20, T24 and S28 face the pore and create a partially hydrophilic environment within the channel wall (Fischer et al., 2000b; Smart et al., 1997). When S20 was mutated to an alanine, ion-channel gating was abrogated, causing the channel to be open all the time. No notable changes were observed when T24 and S28 were replaced with alanine (Premkumar et al., 2004).

It was observed that S20 and S28 were absolutely conserved in the NB protein of all the sequences in this study but T24 was mutated to an isoleucine in specimens DSO_050091_2006, DSO_050214_2006, DSO_020113_2007, DSO_020114_2007, DSO_020132_2007, DSO_020147_2007, and DSO_010147_2007 (Table 5.3). Another notable mutation occurring in the putative transmembrane region is V31I seen in the exact same specimens as I24T, however in this instance, B/Taiwan/45/2007, B/Myanmar/M017/2007, B/Brisbane/60/2008 and B/Washington/07/2009 also shared this mutation (Table 5.3, Supp Fig 3).

Since the role of NB in the replication cycle of the virus is still unknown, there is no empirical assay that has been developed to determine if these 2 mutations might cause a change in the protein’s function. It is also interesting to note all of the substitutions in the putative transmembrane domain are conservative mutations, maintaining a hydrophobic amino acid at the same positions (Betakova et al., 1996; Fischer et al., 2000c; Hatta and Kawaoka, 2003).

Immunofluorescence analysis of cells expressing the NB protein show that the NB protein is present throughout the cytoplasm and is enriched in cis-Golgi region, as indicated by the co-localisation with GM130 (Figures 5.7 & 5.8). Proteins which are undergoing glycosylation are found in this region of the cell and this corresponds with the finding that the NB protein is found abundantly in the membranes of infected cells (Brassard et al., 1996). The observation that cells expressing the NB protein display filamentous cell-membrane projections suggests that the NB protein may have a role in viral budding (Figures 5.7 & 5.8). Confocal as well as electron microscopy studies may be required to determine the nature of these projections and their possible role in viral replication.

Post translational modification of the NB protein has been previously ascribed to N-linked glycosylation at N3 and N7 of the NB protein (Figure 5.4). The addition of these N-linked oligosaccharides to a mannose core (GlcNac)2 which are then further modified by the repeated addition of lactosamine (Galβ1→4-GlcNAc β1→3) units. The
unmodified NB protein has been shown in endoglycosidation studies to migrate to 11.5kDa while NB proteins which contain a single and double mannose core migrate to 15kDa and 18kDa respectively. The addition of polylactosaminoglycan results in the heterogeneous form of NB which migrates as a smear between 30kDa-55kDa on SDS-PAGE (Fukuda, 1985; Williams and Lamb, 1986; Williams and Lamb, 1988).

The studies mentioned were mainly carried out in non-human cells: CV1, CHO, CEF, L and MDCK. While the data from expressing the NB protein in these cells were consistent in their glycosylation patterns, it is important to note that the glycosylation of proteins would differ from cell type to cell type. This difference has been attributed to the difference in specificities of the glycosyl transferases in each cell type (Etchison and Holland, 1974; Nakamura and Compans, 1979). More importantly, influenza B viruses have been found to primarily infect humans in nature, hence determination of NB glycosylation in human cells would be better representative of native influenza B infection. (Murphy, 1990).

To the best of our knowledge, this study presents the first glycosylation analysis of the NB protein in a human cell line, HEK 293T. 6 different forms of the NB protein were found on SDS-PAGE analysis, bands 1, 2,3,4,5 and a smear observed between 20kDa-37kDa (Figure 5.5). Band 1, may correspond to the unmodified NB protein and band 2 and/or 3 may be the singly glycosylated NB protein. Specimens 132-NB, 147-NB do not have band 3, but display band 4 which is unseen in the other specimens. Bands 3/4 may be the doubly glycosylated NB. The smear observed should be the heterogeneous form of the NB protein which is polylactosaminoglycan modified. Band 5 within this smear could correspond to a major form of this heavily modified protein.

Specimens 132-NB and 147-NB do not seem to share any unique amino acid sequences which may cause the absence of band 3 and presence of band 4 (Figure 5.4 & 5.5). A more detailed analysis of this difference in banding should be carried out to determine the nature of these bands.

Time course analysis shows that the singly and doubly glycosylated forms of specimen 70-NB and 147-NB can be detected at 8 hours post transfection, more intensely than the unmodified form of NB, suggesting that the NB protein gets glycosylated upon translation. The NB protein which is modified by polylactosaminoglycan can only be detected at 20 hours, suggesting that this modification does not occur co-translationally (Figure 5.6).
Interestingly, treatment of lysates with PNGase F, does not remove the glycosylation observed in the singly glycosylated form of NB (15kDa). PNGase F has the ability to remove all N-linked glycosylations. No enrichment in the unmodified form of NB was observed, instead an enrichment of the 15kDa band was observed in all the specimens analysed in figure 5.9. Further analysis of this 15kDa band should be done to determine its resistance to PNGase F digestion. The smearing observed in the lysates was resistant to Endo H treatment, suggesting that these are mature glycosylation modifications which are added only in the medial and trans-Golgi network (Figure 5.9).

Treatment of the NB protein with DSP shows the presence of bands F, G and H, which were not obvious when cell lysates were run reduced (Figures 5.5 & 5.10). The NB protein of B/Lee/40 cloned and expressed in this study has a total of 5 cysteine residues in the cytoplasmic domain (Figure 5.4). If 4 of these cysteines are involved in intramolecular disulphide bond formation, this still leaves one cysteine for intermolecular disulphide bond formation. The presence of bands F and H have been reported previously (Williams and Lamb, 1986), with the authors suggesting that band F corresponds to a possible dimer of the NB protein and band H as an aberration of the isolation procedure. This conclusion was made due to band H being less intense than band F, an observation shared in this study as well. Moreover, increasing the concentration of DSP, does not increase the intensity of band H (Figure 5.10) However, if it were to hold true that the NB is an ion channel, forming a dimer may not suffice in reaching this target. The proton channel of influenza viruses, M2 (BM2 in influenza B), functions as a tetramer (Paterson et al., 2003; Sugrue and Hay, 1991). The pore of most reported ion channels are formed by the multimerisation of 3,4,5 or 6 subunits (Cai, 2008). An NB-dimer may not be able to serve as an ion channel. Band G which has not been previously reported migrates to 40kDa, which may possibly correspond to an NB-trimer complex. The intensity of band G increases between the concentrations of 0mM-1mM (Figure 5.4). A more detailed analysis of bands H and G needs to be carried out to determine if the NB is truly an ion channel.
Chapter 6: Genetic and phylogenetic analysis of gene segment 8 encoding for the Non-Structural 1 Protein (NS1) and Nuclear Export Protein (NEP)

6.1 Introduction

The eighth gene segment is the smallest within the influenza B genome, comprising 1096 nucleotides (Figure 1.11). This gene segment encodes for the NS1 and NEP proteins, similar to gene segment 8 of influenza A viruses (Palese, 2007). While the gene for the NS1 protein is an uninterrupted ORF between nucleotides 23-868, the NEP gene is a product of 2 separate exons located between nucleotides 23-55 and 711-1046 of the gene segment (Figure 1.24). This makes the first 11 amino acids of both proteins identical. The second exon of the NEP is in the +1 frame in relation to the NS1 gene (Briedis and Lamb, 1982).

The NS1 protein of influenza B viruses is the only non-structural protein in its genome (Fields et al., 2007). This protein has a myriad of functions and has a main role in countering the host’s antiviral responses. Knowledge of the cellular localisation of viral proteins shed light of the protein’s function during infection and also the specific protein’s interactome. The protein has been shown to have 3 separate domains (Chapter 1, Figure 1.22) (Guan et al., 2011a), similar to the domains of influenza A NS1 (Hale et al., 2008). The first 90 amino acids from the N-terminal make up the RNA Binding Domain (RB), and also contain the Nuclear Localisation Signal (NLS) between positions 46-56 (Dauber et al., 2006; Schneider et al., 2009). It is important to note that while the RNA binding domain of influenza B NS1 has been shown to bind to RNA, no known function of the effector domain has been identified (Donelan et al., 2004; Wang and Krug, 1996).

As of now, only 1 NLS has been found in influenza B NS1 proteins, while 2 have been found in the NS1 proteins of certain influenza A subtypes (Greenspan et al., 1988). This second NLS targets the protein to the nucleolus of mammalian cells, hence being dubbed the Nucleolar Localisation Signal (NoLS). The active residues of the influenza A NS1 NoLS has been mapped to the arginines and lysines of the protein’s C-terminus, between positions 219-237 (Melen et al., 2007). Figure 6.1 (taken from Volmer et al.) shows the localisation the NS1 of 2 different influenza A subtypes: A/Turkey/Italy/977/V99 (H7N1) and A/Udorn/72(H3N2). The C-terminal sequence of Udorn virus NS1 is 219-KRMARTARSKVRRDAD-237, with the underlined residues contributing to the NoLS. The C-terminal sequence of 977 virus NS1 is 219-
Not only is this protein 7 amino acids shorter than the NS1 protein of the Udorn virus, it also lacks most of the basic amino acids which makes up the NoLS. The nucleolus is indicated by a nucleolar marker; nucleophosmin (NPM). NS1 from Udorn virus co-localises with NPM, especially during early infection, while no such co-localisation is observed in the NS1 of 977 virus (Figure 6.1)(Volmer et al., 2010).

Figure 6.1 NoLS of influenza A NS1. Duck Embryonic Fibroblast cells infected with A/Turkey/Italy/977/V99 (H7N1) (-NoLS) and A/Udorn/72(H3N2) (+NoLS). The nucleolus is indicated by nucleophosmin (NPM) labeling. Asterisks indicate the NPM of uninfected cells. White arrows show the changes in NPM localisation upon infection. Image taken from (Volmer et al., 2010).

A unique characteristic of influenza B, not A, NS1 is its speckle organization in the nucleus of infected cells early in infection. This speckle distribution is lost later in infection as the NS1 protein migrates out of the nucleus to the cytoplasm. Figure 6.2, (taken from Schneider et al) displays the mentioned speckle distribution of influenza B NS1 in early infection as compared to influenza A NS1 (Fortes et al., 1995; Schneider et al., 2009; Wolff et al., 1998).
Figure 6.2 Speckled localisation of influenza B NS1 in early infection. Confocal Microscopy of A549 cells infected with the following viruses: B/Lee/40, B/Yamagata/1/73, B/Thuringen/2/06, B/Berlin/37/06 and A/PR/8/34. Early: 4-6 hours of infection. Late: 16hrs infection. Scale bar = 10µm (Taken from Schneider et al, 2009).

These speckles are found to be sites of host cell SC35 protein locations within the nucleus (Chapter 1, Figure 1.23). It was found that not only does influenza B NS1 co-localise with SC35 in early infection, but cells infected with influenza B display a more organized arrangement of SC35 even in late infection as compared to uninfected cells (Schneider et al., 2009).

The authors of the mentioned study have identified the first 90 amino acids as being the shortest possible fragment required for this co-localisation with SC35. Using mutational analysis, it was found that the arginines and lysines between residues 47-54 contribute to this co-localisation (Schneider et al., 2009). SC35 domains contain components required for pre-mRNA splicing. The SC35 protein is defined as a serine/arginine (SR)-rich splicing factor. Such domains are located in a speckled organization within the nucleus (Spector et al., 1991; Spector and Lamond, 2010; Wansink et al., 1993).

The 281 amino acids that comprise the NS1 protein of B/Lee/40 give it a theoretical molecular weight of 32.067 kilo Daltons (kDa) (The UniProt Consortium, 2012). Experimental evidence involving the expression of the NS1 from infection by influenza B virus shows that the NS1 protein is estimated to be 34.77kDa (Yuan and Krug, 2001), suggesting the possibility of post translational modification yet to be determined. The influenza A NS1 protein has been found to be phosphorylated at threonine 215, as well at serines 42 and 48 (Hale et al., 2009; Hsiang et al., 2012). In addition the influenza A NS1 is also SUMOlyted and modified by ISG15 (Pal et al., 2010; Tang et al., 2010; Xu et al., 2011; Zhao et al., 2010). As of late, no post-translational modifications have been reported for the influenza B NS1 protein. Evidence has shown that influenza B NS1 is successful in inhibiting the cellular
function of the interferon-induced ISG15, unlike influenza A NS1 (Sridharan et al., 2010; Yuan and Krug, 2001).

The NEP was initially named Non-Structural protein 2 (NS2), however, recent reports have found this protein to be included in the packaged virion, in association to the matrix protein (Richardson and Akkina, 1991; Yasuda et al., 1993). This protein serves a pivotal role in ensuring the nuclear export of vRNPs (Gomez-Puertas et al., 2000; Neumann et al., 2000; Neumann et al., 1994; O'Neill et al., 1998; Yasuda et al., 1993).

Similar to what has been previously presented in the last 2 chapters, the work compiled in this chapter serves to understand the evolution of gene segment 8 in relation to the vaccine and other circulating influenza B strains. Alignment of the amino acid sequences of the NS1 and NEP with these strains may give us greater insights into the selective pressures exerted on individual domains/motifs, allowing for functional inferences.

In addition to analyses of the sequences, further molecular characterisation of the NS1 protein was carried out. This work compiled in this chapter is dedicated to understanding the kinetics of NS1 localisation upon infection, as well as to determine the possibility of influenza B NS1 containing an NoLS similar to influenza A NS1.

In an effort to determine if the NS1 protein was post-translationally modified, the presence of a second peptide p23 was found upon western blot analysis of the NS1 protein. This peptide was found to be expressed in B/Lee/40 infection and in some of the clinical specimens studied. Two clinical specimens did not express p23 and their difference in sequence was exploited to determine the nature of p23 expression. Figure 6.3 shows a schematic representation of the domains NS1 and NEP proteins.

![Figure 6.3 Schematic representation of the NS1 and NEP proteins. The numbers to the right of the protein indicates the number of amino acids in the protein.](image_url)
6.2 Results

6.2.1 Sequencing of gene segment 8

Using the thermocycling conditions optimized for the amplification of the gene segments 4 and 6, the entire ORF for gene segment 8 was amplified yielding a PCR product of approximately 1200 base pairs. To test if the sequences generated in this study were correct, generated sequences were ran through the BLAST search online tool. Table 6.1 lists the top hits for the NS genes sequenced in this study.

![Figure 6.4](image)

**Figure 6.4** Size of the coding sequence of gene segment 8 of B/Lee/40 generated in this study analysed on an ethidium bromide stained 1% agarose gel.

<table>
<thead>
<tr>
<th>Accession No</th>
<th>Description</th>
<th>Max Score</th>
<th>Total Score</th>
<th>Query Coverage</th>
<th>E Value</th>
<th>Max Ident</th>
</tr>
</thead>
<tbody>
<tr>
<td>CY040453.1</td>
<td>Influenza B virus (B/Malaysia/2506/2004) segment 8, complete sequence</td>
<td>1829</td>
<td>1829</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
<tr>
<td>CY033872.1</td>
<td>Influenza B virus (B/NYMC BX-17(Lee/1940-Malaysia/2506/2004)) segment 8, complete sequence</td>
<td>1847</td>
<td>1847</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
<tr>
<td>DQ792901.1</td>
<td>Influenza B virus (B/Lee/40) NEP protein (NEP) and NS1 protein (NS1) genes, complete cds</td>
<td>1871</td>
<td>1871</td>
<td>100%</td>
<td>0.0</td>
<td>99%</td>
</tr>
</tbody>
</table>

6.2.2 Nucleotide sequence and phylogenetic analysis of gene segment 8.

The sequence similarity observed between the 36 clinical specimens yielding the coding sequence of the NS gene was 91.1%-100%. This range did not change even when the downloaded sequences were included in the alignment (data not shown).
Unlike both HA and NA phylogenetic trees, the specimens within the NS phylogenetic tree do not fall into the Yamagata or Victoria clusters. Instead, all the sequences fell into three clusters. The NS gene from both ancestor strains fell into cluster II whilst the NS genes from the clinical specimens sequenced in this study fell into clusters I and III. The specimens within cluster II were isolated between 1964 and 1994 while the sequences within cluster III were isolated between 1984 and 2009. All of the vaccine strains clustered within cluster III with the majority of the specimens in the phylogenetic tree. Clusters III and I are displayed separately in figures 6.5B and C respectively (Figure 6.5).

All of the NS genes of specimens isolated in 2004 (●), except for DSO_090117_2004, fell in cluster I together with B/Lee/40 (Figure 6.5C). The NS gene from specimen DSO_090117_2004 clustered surprisingly with the specimens isolated in 2009 within cluster III (Figure 6.5B).

All of the clinical specimens sequenced after 2004 (except for DSO_090117_2004) clustered within cluster III, together with the NS genes of all the vaccine strains from 2004-2013. The NS genes from these vaccine strains showed higher homology to the clinical specimens sequenced in this study than the other downloaded sequences within cluster III (Figure 6.5).

A strong clustering was observed between most of the specimens isolated in 2005 and 2006 (●, ●). Similarly, the specimens isolated in 2009 (●) clustered together. The only specimen from 2004, not to be found in cluster I, DSO_090117_2004 was found closely associated with the 2009 specimens (Figure 6.5).

Unlike what was observed in the HA and NA phylogenetic trees, the NS gene from specimen DSO_010147_2007 did not cluster with B/Florida/4/2006. Instead DSO_010147_2007 shared a node with the vaccine strains of the northern hemisphere for 2012-2013: B/Wisconsin/1/2010. DSO_010147_2007 still clustered away from the other specimens sequenced in this study (Figure 6.5).

### 6.2.3 Amino Acid sequence analysis

The NS1 protein alignment of all the sequences in this study showed 87.5%-100% sequence similarity. Within the clinical specimen sequenced, the sequence similarity shared remains unchanged. The NEP protein alignment of all the sequences in this study showed 91.8%-100% sequence similarity. Within the clinical specimen sequenced, the sequence similarity shared remains unchanged (Supp Fig 4).
A.

NS Phylogenetic Tree
B.

NS Cluster III
6.2.3.1 Amino acid sequence analysis of the NS1 protein

The NS1 protein of the 7 2004 specimens sequenced in this study showed high sequence similarity to B/Lee/40 unlike the rest of the specimens listed in figure 6.5. Two main motifs of the NS1 protein are highlighted in this study: the Nuclear Localisation Signal (NLS) (Table 6.2) and the ISG15 Binding Domain (Table 6.3).

| Table 6.2 Amino acid substitution of the NS1 NLS. I, II III: Specimens belonging to clusters I, II and III as indicated in Figure 6.5. Sequences of specimens are compared to in reference to B/Lee/40 NS1. Similar residues are labeled by the dot (.) |
|---|---|---|---|---|---|---|---|---|
| B Lee 40 | D | R | L | H | R | L | N | R | K | L | E |
| DSO 090114 2004 | . | . | . | . | . | . | . | . | . | . | . |
| DSO 090124 2004 | . | . | . | . | . | . | . | . | . | . | . |
| DSO 090131 2004 | . | . | . | . | . | . | . | . | . | . | . |
| DSO 090133 2004 | . | . | . | . | . | . | . | . | . | . | . |
| DSO 090134 2004 | . | . | . | . | . | . | . | . | . | . | . |
| DSO 090136 2004 | . | . | . | . | . | . | . | . | . | . | . |
| DSO 090138 2004 | . | . | . | . | . | . | . | . | . | . | . |
The NLS of the NS1 is located between amino acids 45 and 56 (Schneider et al., 2009). The specimens within cluster I have the exact same amino acid sequence of B/Lee/40’s NLS. The rest of the sequences analysed harbor the H49N and R53K amino acid substitution. Both of these are conservative substitutions (Figure 6.5 & Table 6.2).

The influenza B NS1 has been shown to have a unique interaction with host cell ISG15 (Yuan and Krug, 2001). Nineteen out the first 101 amino acids of the NS1 protein have been identified as directly associating with ISG15 (Guan et al., 2011a). Three out of these 19 residues have shown variability amongst the specimens analysed in this study (Table 6.3).

Positions 34, 97 and 101 of the specimens analysed have shown variability in sequence (Table 6.3). Specimens of cluster I have the exact sequence as B/Lee/40 NS1. Specimens belonging to clusters II and III display F34L, while I97V was observed in three out of the four specimens of cluster II and E101G was seen in 2 of the specimens sequenced in this study. All of these substitutions were conservative (Table 6.3).

Table 6.3 Amino acid substitution of ISG15 binding site. I, II, III: Specimens belonging to clusters I, II and III as indicated in Figure 6.5. Sequences of specimens are compared to in reference to B/Lee/40 NS1. Similar residues are labeled by the dot (.).

<table>
<thead>
<tr>
<th>Specimen</th>
<th>34</th>
<th>97</th>
<th>101</th>
</tr>
</thead>
<tbody>
<tr>
<td>B Lee 40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSO 090114 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090124 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090131 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090133 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090134 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090136 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090138 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>B Yamagata 16 1988</td>
<td>L</td>
<td>V</td>
<td>.</td>
</tr>
<tr>
<td>B Victoria 02 1987</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>B Singapore 04 1991</td>
<td>L</td>
<td>V</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090117 2004</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 050143 2005</td>
<td>L</td>
<td>.</td>
<td>G</td>
</tr>
<tr>
<td>DSO 050540 2005</td>
<td>L</td>
<td>.</td>
<td>G</td>
</tr>
<tr>
<td>Rest of DSO Specimens</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>B Singapore 22 1998</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Vacc04S B Hong Kong 330 2001</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Vacc0405N 05S B Jilin 20 2003</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Vacc0405N 05S B Shanghai 361 2002</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Vacc0608NS B Malaysia 2506 2004</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Vacc0506N B Jiangsu 10 2003</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Vacc0809NS B Florida 4 2006</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Vacc0912NS B Brisbane 60 2008</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Vacc1213N B Wisconsin 01 2010</td>
<td>L</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>
Unlike the HA, NA and NB proteins, the NS1 amino acid profile of specimen DSO_010147_2007 did not differ significantly from the rest of the sequences in the alignment (Tables 6.2 & 6.3, Supp Fig 4).

6.2.3.2 Amino acid sequence analysis of the NEP

The NEP protein is 122 amino acids long and 4 positions have been found to display variability (Table 6.4). All of the specimens of cluster I had the same amino acid sequence to B/Lee/40. A66 and S80 were unique to specimens of cluster I, while specimens of cluster II and III had the substitution I98V. Uniquely specimens DSO_090117_2004, DSO_0003_2009, DSO_0005_2009 and DSO_0070_2009 harbored S35N substitution, unseen in any of the specimens analysed (Table 6.4 and Figure 6.5, Supp Fig 5).

Table 6.4 Amino acid substitution of NEP, I, II, III: Specimens belonging to clusters I, II and III as indicated in Figure 6.5. Sequences of specimens are compared to in reference to B/Lee/40 NEP. Similar residues are labeled by the dot (.)

<table>
<thead>
<tr>
<th>B Lee 40</th>
<th>35</th>
<th>66</th>
<th>80</th>
<th>98</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSO 090114 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090124 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090131 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090133 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090134 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090136 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090138 2004</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>B Yamagata 16 1988</td>
<td>V</td>
<td>N</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>B Victoria 02 1987</td>
<td>V</td>
<td>N</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSO 090117 2004</td>
<td>N</td>
<td>V</td>
<td>N</td>
<td>V</td>
</tr>
<tr>
<td>DSO 0003 2009</td>
<td>N</td>
<td>V</td>
<td>N</td>
<td>V</td>
</tr>
<tr>
<td>DSO 0005 2009</td>
<td>N</td>
<td>V</td>
<td>N</td>
<td>V</td>
</tr>
<tr>
<td>DSO 0070 2009</td>
<td>N</td>
<td>V</td>
<td>N</td>
<td>V</td>
</tr>
<tr>
<td>Rest of DSO Seq</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>B Singapore 22 1998</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>B Singapore 31 1998</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>B Singapore 35 1998</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>B Singapore 04 1991</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>Vacc04S B Hong Kong 330 2001</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>Vacc0405N 05S B Jilin 20 2003</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>Vacc0405N 05S B Shanghai 361 2002</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>Vacc0608NS B Malaysia 2506 2004</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>Vacc0506N B Jiangsu 10 2003</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>Vacc0809NS B Florida 4 2006</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>Vacc0912NS B Brisbane 60 2008</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>Vacc1213N B Wisconsin 01 2010</td>
<td>V</td>
<td>N</td>
<td>V</td>
<td></td>
</tr>
</tbody>
</table>
6.2.4 Molecular Characterisation of the NS1 protein from non-tissue culture adapted clinical specimens

6.2.4.1 Cellular localisation of NS1 protein during influenza B infection

To analyse the localisation of NS1 protein during infection, A549 cells were infected with egg-passaged B/Lee/40 at MOI 0.5 and cells were fixed at different time points to be subsequently stained for NS1 and NP proteins with appropriate cell markers. NP is a commonly used marker of infection, which is expressed in the nucleus early in infection and is found in both the nucleus and cytoplasm in late infection (Chapter 3, Figure 3.5).

Figure 6.6 shows that NS1 protein is expressed only in cells which are expressing NP. The NP protein can be detected as early as 8 hours post infection (hpi) whereas NS1 is only obvious at 12hpi. However, the staining pattern of both proteins seem to be similar, in that they both start off concentrated in the nucleus in early infection and at later in infection, the proteins seem to be equally present in the nucleus and cytoplasm (Figure 6.6). Unlike the NP protein, at 12 hpi, the NS1 protein is found concentrated in nuclear speckles. This localisation is not observed at 16hpi (Figure 6.6, inset).

![Influenza B infection on A549](image.png)

Figure 6.6 Time course infection of egg-passaged B/Lee/40 on human alveolar A549 cells. Cells were stained for both influenza B NP and NS1. Inset: highlighting the speckled localisation of the influenza B NS1 protein at 12hpi. Images taken under oil immersion (100X magnification)
To further characterize this apparent karyophilic nature of the influenza B NS1, cells infected with egg-passaged influenza B virus were stained for both NS1 and Lamin A/C, a nuclear marker for mammalian cells (Koch, 2003) (Figure 6.7).

At 8hpi infected cells show faint but distinct NS1 speckles in the nucleus (Figure 6.7 inset). The intensity of NS1 expression increases by 12 and 16hpi, at these time points, the NS1 protein is prominently expressed in the cytoplasm (Figure 6.7). To further characterise the NS1 localisation during infection, the same experiment as above is repeated with different cell markers: Nuclear Pore Complex (NPC) (Figure 6.8), Nucleophosmin (NPM) (Figure 6.9), Coilin (Figure 6.10) and SC35 (Figure 6.11).
Figure 6.9 Time course infection of egg-passaged B/Lee/40 on human alveolar A549 cells. Cells were stained for both influenza B NS1 and Nucleophosmin (NPM). Inset: magnified view of a nucleus of influenza B infected cell, showing the organization of NS1 and NPM at 12hpi. Images taken under oil immersion (100X magnification).

Figure 6.10 Time course infection of egg-passaged B/Lee/40 on human alveolar A549 cells. Cells were stained for both influenza B NS1 and Coilin. Images taken under oil immersion (100X magnification).

Figure 6.11 Time course infection of egg-passaged B/Lee/40 on human alveolar A549 cells. Cells were stained for both influenza B NS1 and SC35. Inset 8hpi: magnified view of a nucleus of influenza B infected cell, showing the speckled organization of NS1 in early infection as similar to that of SC35 organization. Inset 12hpi: magnified view of NS1 and SC35 at 12hpi. Images taken under oil immersion (100X magnification).
The same experiments were also repeated in MDCK cells with Lamin A/C (Figure 6.12), NPC (Figure 6.13), SC35 (Figure 6.14) and NPM (Figure 6.15), as MDCKs are a common model cell line in the study of influenza.

An observation that can be made with regards to NS1 expression is the unique speckle organization within the nucleus that is seen in early infection (between 8-12hpi). At later time points, the speckles are less obvious as NS1 is expressed throughout the nucleoplasm (Figures 6.7-6.15). These speckles do not localise with coilin, which presents as speckles in the nucleus as well (Figures 6.10) but with SC35 in both A549 and MDCK cells (Figure 6.11 and 6.14).

![Figure 6.12 Time course infection of egg-passaged B/Lee/40 on MDCK cells. Cells were stained for both influenza B NS1 and Lamin A/C. Inset: magnified view of a nucleus of influenza B infected cell, showing the speckled organization of NS1 at 12hpi. Dotted Box: highlighting the same nucleus magnified in the inset but stained for Lamin A/C. Images taken under oil immersion (100X magnification) ](image)

![Figure 6.13 Time course infection of egg-passaged B/Lee/40 on MDCK cells. Cells were stained for both influenza B NS1 and Nuclear Pore Complex (NPC). Images taken under oil immersion (100X magnification) ](image)
Figure 6.14 Time course infection of egg-passaged B/Lee/40 on MDCK cells. Cells were stained for both influenza B NS1 and SC35. Inset, magnified view of a nucleus of influenza B infected cell, showing the speckled organization of NS1 at 12hpi as similar to that of SC35 organization. Images taken under oil immersion (100X magnification).

Figure 6.15 Time course infection of egg-passaged B/Lee/40 on MDCK cells. Cells were stained for both influenza B NS1 and Nucleophosmin (NPM). Inset, magnified view of a nucleus of influenza B infected cell, showing the organization of NS1 at 16hpi as compared to that of NPM organization. Images taken under oil immersion (100X magnification).

The pattern of SC35 mirrors that of NS1 in the nucleus even at 16hpi, when the expression of NS1 is no longer confined to speckles but spread throughout the nucleoplasm in both A549 and MDCK cells (Figures 6.11 and 6.14). Interestingly, at 12-16hpi when NS1 expression is present throughout the nucleoplasm, there exist regions of NS1 exclusivity. These regions appear as round structures within the nucleus. Even SC35 is not present in these NS1-exclusive regions (Figures 6.11 and 6.14). Staining with NPM shows that these regions are rich in NPM. This reciprocal expression of NS1 and NPM is not seen in early infection but only in 12hpi and even more obvious at 16hpi (Figures 6.9 and 6.15).
6.2.4.2 Molecular Cloning of NS1 and Expression by Plasmid DNA Transfection

The NS1 protein sequences analysed in Section 6.2.3.1 can be grouped according to the year of isolation. Specimens bearing similar sequences are grouped together and a representative strain is chosen for each group (Table 6.5). 2 specimens were chosen to represent 2007 specimens, owing to the unique sequence of specimen DSO_010147_2007. Table 6.5 also lists the abbreviated code by which these specimens will be referred to throughout this thesis.

Table 6.5 NS1 protein sequences were grouped according to their similarity and a representative strain is chosen for each group. * Code refers to the abbreviation by which these specimens will be referred to.

<table>
<thead>
<tr>
<th>Year of Isolation</th>
<th>Specimen Name</th>
<th>Code*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940</td>
<td>B/Lee/40</td>
<td>Lee</td>
</tr>
<tr>
<td>2004</td>
<td>DSO_090136_2004</td>
<td>136</td>
</tr>
<tr>
<td>2005/2006</td>
<td>DSO_040117_2006</td>
<td>117</td>
</tr>
<tr>
<td>2007</td>
<td>DSO_020132_2007</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>DSO_010147_2007</td>
<td>147</td>
</tr>
<tr>
<td>2009</td>
<td>DSO_0070_2009</td>
<td>70</td>
</tr>
</tbody>
</table>

An alignment of these representative sequences is shown in figure 6.16. The representative strains of each group were then cloned into a mammalian expression vector, pCAGGS, which expresses genes under the control a chicken β actin promoter. Each of the sequences was cloned with a FLAG epitope tag (DYDDDDK) at the N-terminus (Figure 6.16).

These plasmids containing the various NS1 genes were transfected into MDCK cells and 10 hours post transfection, the cells were immunologically labeled with anti FLAG antibody. Cells were co-stained with both, anti SC35 and anti NPM as markers (Figures 6.17 and 6.18 respectively).
Figure 6.16 Cloning of NS1 gene. A. Alignment of the representative sequences listed in Table 5.1. RNABD: RNA Binding Domain. ED: Effector Domain. The linker and the FLAG tag epitope are highlighted. B. Schematic representation of the cloning strategy employed for these sequences. RE: Restriction Endonuclease. pCAGGS schematic, courtesy of Dr. Loo Liat Hui.
In both figures 6.17 and 6.18, the NS1 of Lee and specimen 136 do not show obvious speckles, unlike the rest of the specimens analysed. However, the staining pattern of SC35 is similar to that of Lee-NS1 (Figure 6.17). For the case of Lee-NS1 and 136-NS1, the staining pattern of SC35 in a transfected cell is significantly altered compared to the staining pattern of an untransfected cell (box, Figure 6.17). The rest of the specimens analysed display obvious speckles which correlated with the spatial expression of SC35 (Figure 6.17).
Figure 6.18 Transfection of NS1-pCAGGS constructs in MDCK cells. Cells were fixed and stained 10 hours post transfection. Upper panels: Cells were stained with anti-FLAG. Lower panels: cells were co-stained with anti-Nucleophosmin (NPM). Inset: Magnified view of the nucleus of Lee-NS1 transfected cells showing the regions of Lee-NS1 exclusions which are coincident with NPM. White Arrow: Nucleus of 117-NS1 transfected cell showing large NS1 conglomerates. White hollow arrow: Localisation of NPM in nucleus of 117-NS1 transfected cell. Images taken under oil immersion (100X magnification).

Similar to the staining observed in infected cells, Lee-NS1 and NPM seem to have an antagonistic expression (Figures 6.9, 6.15 & 6.18). 117-NS1 has the most unique intra-nuclear organization compared to the rest of the specimens. 117-NS1 transfected cells display large ‘blobs’ within the nucleus while the other specimens either show discrete, punctate speckles or no speckles at all. Interestingly, the intra-nuclear structures displayed by 117-NS1 seem to share a similar staining pattern to SC35 and NPM (Figures 6.17 & 6.18).
Figure 6.19 Transfection of NS1-pCAGGS constructs in MDCK, human alveolar A549 and CEF cells. Cells were fixed and stained 10 hours post transfection. Images taken under oil immersion (100X magnification).
To further characterise the cellular localisation of the different NS1 constructs, similar transfections were carried out on human alveolar A549 cells as well as chicken embryonic fibroblasts (CEF) (Figure 6.19). The same patterns of localisations were seen even in the human and avian cells lines: Lee and 136, showing general nuclear and cytoplasmic staining, 117 showing large nuclear inclusions and 132, 147 and 70 showing fine speckles throughout the nucleoplasm (Figure 6.19).

6.2.4.2.1 Comparison of Influenza A and Influenza B NS1 localisation

To further characterize the staining pattern of the NS1 protein, the NS1 genes of 2 strains of influenza A (H1N1) were cloned into the pCAGGS expression vector and their transfected into MDCK cells (Table 6.6). Their cellular localisations were compared to Lee-NS1 and 132-NS1 (Figure 6.20).

The influenza A NS1 proteins did not show speckles like that seen in 132-NS1. However PR8-NS1 displays a similar NPM exclusion, like Lee-NS1 (Figure 6.20, inset). This exclusion can also be observed in 132-NS1, albeit being masked by the speckles. Interestingly specimen 471-NS1, instead of displaying NPM exclusion shows a co-localisation of NS1 and NPM (Figure 6.20). The alignment of these 2 influenza A NS1 proteins is shown in figure 6.20 below, as being compared to the NS1 protein sequence of A/Udorn/72(H3N2) which contains a known NoLS sequence, highlighted in red (Figure 6.21).

PR8-NS1 only shares 2 residues out of the 6 which compose the NoLS of Udorn-NS1, K\textsubscript{219} and R\textsubscript{220}. The sequence of 471-NS1 however, terminates after residue K\textsubscript{219}, thus not containing the rest of the residues required for a functional NoLS (Figure 6.21)

<table>
<thead>
<tr>
<th>Year of Isolation</th>
<th>Name of Specimen</th>
<th>Seq ID</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1934</td>
<td>A/Puerto Rico/8/1934 (H1N1)</td>
<td>JX120148.1</td>
<td>PR8</td>
</tr>
<tr>
<td>2009</td>
<td>A/Singapore/471/2009 (H1N1)</td>
<td>CY069630.1</td>
<td>471</td>
</tr>
</tbody>
</table>
**Figure 6.2** Cellular Localisation of influenza A NS1. MDCK cells were transfected with 471-NS1-pCAGGGS, PR8-NS1-pCAGGGS, Lee-NS1-pCAGGGS and 132-NS1-pCAGGGS. 10hrs post transfection cells were fixed, permeabilised and stained with anti-FLAG and anti-NPM antibodies. 471 inset: highlighting the nuclear inclusions of 471-NS1. PR8 inset: highlighting the nuclear exclusions of PR8-NS1. Images taken under oil immersion (100X magnification)

### 6.2.4.2.2 Expression of Individual Influenza B NS1 Domains

Individual domains of the Lee-NS1 and 132-NS1 proteins were cloned into the pCAGGS expression vector and transfected into cells to observe their respective nuclear localisations. Figure 6.22 shows the different domains cloned. Abbreviations, NS1: full length NS1 protein, RB: RNA Binding Domain, RBL: RNA Binding Domain
+ Linker, ED: Effector Domain, LED: Linker + Effector Domain (Figure 6.22). These constructs were transfected for 10 hours and 20 hours (Figures 6.23 & 6.24).

Figure 6.21 Alignment of influenza A NS1 proteins analysed in this study as compared to A/Udorn/1972 (H3N2). Residues known to contribute to the NoLS of A/Udorn/1972 are underlined in red.

![Influenza A NS1 Protein Alignment](image)

Figure 6.22 Individual domains of Influenza B NS1 which were cloned into pCAGGS vector for transient expression. NS1: full length NS1 protein, RB: RNA Binding Domain, RBL: RNA Binding Domain + Linker, ED: Effector Domain, LED: Linker + Effector Domain. MYC and FLAG epitopes are used as epitope tags to assist in immune labeling.

At 10 hours, Lee-NS1 displays a general nuclear and cytoplasmic localisation, without any speckles or nuclear inclusions. At 20 hours Lee-NS1 is found predominantly in the cytoplasm rather than in the nucleus. Lee-RB and Lee-RBL shows general nuclear and cytoplasmic staining, this pattern retains even at 20 hours post transfection. Lee-RB and Lee-RBL also displays regions of nuclear exclusion as shown by Lee-NS1 in figures 6.9 and 6.18. The nuclear exclusions of Lee-RBL coincide with
NPM staining. At 10 hours Lee-ED shows strong cytoplasmic staining, this patterning remains the same even at 20 hours but faint nuclear inclusion can be seen which seem to coincide with NPM staining. Lee-LED however shows nuclear-cytoplasmic staining at 10 hours but at 20 hours, obvious speckles can be seen in the nucleus (Figure 6.23).

132-NS1 shows speckles in the nucleus after 10 hours post transfection as also seen in figure 6.17. These speckles increase in size at 20 hours post transfection and the cytoplasmic staining also increases in intensity. 132-RB shows a slightly stronger presence in the nucleus than the cytoplasm. At 20 hours, very faint speckles seem to be present in the nucleus. 132-RBL seems to show a similar staining pattern as 132-RB except that no speckles were seen at 20 hours. 132-ED shows strong cytoplasmic staining at 10 hours but at 20 hours, 132-ED is seen with equal prominence in the nucleus and the cytoplasm. 132 LED shows general nuclear and cytoplasmic staining at 10 hours.

![Figure 6.23 Expression of the individual domains of Lee-NS1 protein. MDCK cells were transfected with the individual constructs for 10 or 20 hours, fixed and stained appropriately. Images taken under oil immersion (100X magnification)]
Figure 6.24 Expression of the individual domains of 132-NS1 protein. MDCK cells were transfected with the individual constructs for 10 or 20 hours, fixed and stained appropriately. Images taken under oil immersion (100X magnification)

6.2.4.2.3 Effects of Mutations on Influenza B NS1 Protein Localisation

Lee-NS1 has an amino acid sequence of NPL between positions 107-109. This sequence is shared with 136-NS1 but specimens 117-NS1, 132-NS1, 147-NS1 and 70-NS1 contain the sequence SSS107-109 (Figure 6.17). Since Lee-NS1 and 136-NS1 are the only 2 which do not display any observable nuclear inclusions on immunofluorescence (Figures 6.17, 6.18 & 6.19), it is possible that amino acids 107-109 contribute to this variability. Lee-NS1 was mutated to replace NPL107-109 to SSS107-109. The mutated construct was expressed in MDCK cells as well as the wild type Lee-NS1 for 10 and 20 hours and observed through immunofluorescence (Figure 6.25). No obvious speckles were seen in Lee-NS1 SSS107-109. However, Lee-NS1 SSS107-109 still localises as wild type Lee-NS1, bearing the same co-localisation with SC35.

132-NS1 was mutated to produce 132-NS1 ASS107-109, 132-NS1 AAS107-109 and 132-NS1 AAA107-109. These mutants were transfected into MDCK cells as well and observed under immunofluorescence (Figure 6.27). The mutants analysed still showed speckles in the nucleus. Interestingly, these speckles seemed to increase in number and size for each of the three mutations. The ‘mutant speckles’ strongly co-localised with
SC35. Even the speckles which were larger than ‘wild-type’ speckles co-localised with SC35 strongly, this is clearly seen in mutant 132-NS1 ASS$_{107-109}$ (Figure 6.27).

Figure 6.25 Transfection of wild type Lee-NS1 and Lee-NS1 SSS107-109 for 10 and 20 hours. Upper panel, cells stained with anti FLAG. Lower panels: cells co-stained for SC35. Images taken under oil immersion (100X magnification)
As proteins found to be associated with nuclear speckles are often phosphorylated (Bryantsev et al., 2007; Prasanth et al., 2003; Spector and Lamond, 2010), the serines unique to speckle-inducing NS1 constructs (11-NS1, 132-NS1, 147-NS1 and 70-NS1) were mutated to the residues seen in non-speckle-inducing NS1 constructs (Lee-NS1 and 136-NS1). Figure 6.27 shows the cellular localisation of wild type 132-NS1, 132-NS1 S69N, 132-NS1 S123P and 132-NS1 S215N. All the mutants still produced speckles which are similar in size and number to wild type 132-NS1.

Specimen 117-NS1 has 3 amino acids which are unique to the other constructs analysed in this study, P94, I151 and G154 (Figure 6.16). These unique residues could contribute to 117-NS1’s unique co-localisation to SC35 as well as NPM, unseen in any of the other constructs (Figures 6.17-6.19). While 117-NS1 has a proline at position 94, the other constructs contain a serine. 117-NS1 was mutated to contain a serine at position 94 and its cellular localisation was compared to wild type 117-NS1 (Figure 6.28).
Figure 6.27 Transfection of 132-NS1, 132-NS1 S69N, 132-NS1 S123P and 132 NS1 S215N into MDCK cells for 10 hours. Cells were fixed and stained with anti FLAG (upper panel) and co stained with anti SC35 (lower panel). Images taken under oil immersion (100X magnification).

Wild type 117-NS1 and 117-NS1 P94S appears to have similar cellular localisations. Cells transfected with both constructs display large nuclear inclusions which co-localise with SC35. The pattern of SC35 in transfected cells is significantly altered compared to SC35 in untransfected cells. These nuclear transfections seem to be larger in size at 20 hours as compared to 10 hours in both the wild-type and mutant (Figure 6.28a). The RB, RBL and ED domains of 117-NS1 were successfully cloned in the pCAGGS expression vector. None of these domains showed the distinct nuclear inclusions as seen with the full length 117-NS1 protein (Figure 6.28b).
6.2.4.3 SDS-PAGE analysis of influenza B NS1

6.2.4.3.1 SDS-PAGE analysis of NS1 in B/Lee/40 infection and Lee-NS1 transfection

Human Embryonic Kidney (HEK) 293T cells were both infected with B/Lee/40 virus and transfected with Lee-NS1 pCAGGS. After overnight expression, cells were harvested in Laemmli buffer and equal volumes of each sample were loaded and analysed on a 12% SDS-PAGE gel. The presence of the NS1 protein was carried out by western blotting with the aid of an antibody specific to influenza B NS1 protein. As controls HEK 293T cells were mock infected with Phosphate Buffered Saline (PBS) or transfected with empty pCAGGS vector (Figure 6.29).
In both the infection and the transfection, 2 bands can be observed (Figure 6.29A). To estimate the molecular sizes of these bands, a Log MW vs. Rf chart was plotted using the molecular standards used on the gel. An equation was generated upon plotting the best fit line (Figure 6.29B). Rf values of each of the bands were calculated and the molecular weights of each of those bands labeled 1-4 were estimated as listed in Figure 6.29C. Upon infection 2 bands were observed, one at 34.2kDa and another at 22.8kDa. A similar observation is made in the transfections where one band migrated to 35.2kDa while another band was detected at 24.2kDa. No protein bands were observed in both the mock infection and mock transfection. Both bands observed in the transfections migrate slower than their infection counterpart. The difference between band 3 to band 1 and band 4 to band 2 is roughly 1kDa. This difference could be attributed to the presence of the FLAG epitope tag added on the C-terminus of the NS1 protein. Bands 1 and 3 correspond to the molecular weight of the influenza B NS1 protein (Yuan and Krug, 2001). The presence of the second (smaller band) has not been reported, to the best of our knowledge. The rest of the work in this chapter will be
dedicated in understanding the nature of this smaller product, which will be denoted as p23 throughout this study. For the convenience of annotation, p23 will be used to refer to the smaller product observed in transfections as well.

To test if the influenza A NS1 has an expressed product similar to p23, 2 influenza A NS1 genes were cloned into the pCAGGS vector with a FLAG epitope tag at the C-terminus as well. A/Puerto Rico/8/1934 (H1N1), denoted as PR8-NS1, as well as A/Singapore/471/2009 (H1N1), denoted as 471-NS1 were similarly expressed in HEK 293T cells and as a negative control cells were transfected with an empty pCAGGS vector (pC) (Figure 6.30).

Only single bands, corresponding to the influenza A NS1 protein were detected. No smaller product was observed, even after longer exposure of the film to the membrane while Lee-NS1 clearly showed 2 intense bands as seen in figure 6.30.

6.2.4.3.2 Time course expression of Influenza B NS1 protein

To determine kinetics of p23 expression in relation to full length NS1 expression, a time course experiment was set up for both infection with B/Lee/40 and transfection with Lee-NS1. HEK 293T cells were subjected to the infections and transfections and were harvested in Laemmli buffer at 4, 8, 12 and 16 hours post infection/transfection (Figure 6.31).

Figure 6.31A shows that the NS1 protein is detectable in low amounts at 8 hours post infection (hpi), while the NS1 protein is detectable only very faintly at 8 hours post
transfection in figure 6.31B, suggesting that more protein is generated in an infection at this M.O.I than transfection. In both infections as well as transfections, p23 is observed as early as the full length NS1 protein.

![Figure 6.31 Time course infection of B/Lee/40 on HEK 293T cells. Cells were harvested at 4, 8, 12 and 16 hours post infection. Similar volumes of sample were loaded onto a 12% SDS-PAGE gel. B. Time course transfection of Lee-NS1 on HEK 293T cells. Cells were harvested at 4, 8, 12 and 16 hours post transfection. Similar volumes of sample were loaded onto a 12% SDS-PAGE gel. In both A. and B. NS1 protein bands were detected by western blotting with the aid of an antibody specific to the influenza B NS1 protein. Black Arrow: Protein band corresponding to the full length of the influenza B NS1 protein: Black Arrow head: Protein band corresponding to p23.]

6.2.4.3.3 Expression of individual domains of B/Lee/40 NS1 protein

To determine which domains make up p23, each domain was cloned into the pCAGGS expression vector with an appropriate epitope tag and transfected into HEK 293T cells. Figure 6.32A shows the different combinations of the domains cloned. Constructs NS1mf, LED, and RBL have 2 tags, the myc-tag at the N-terminus and the FLAG-tag at the C-terminus (Figure 6.32A). The log MW vs. Rf plot was according to the migration of the protein standards in the 15% SDS-PAGE gel and the sizes of each domain were accordingly estimated (Figure 6.32D & 6.32E).

Transfection with Lee-NS1 yielded 2 bands (Labeled as 1 and 2 in figure 6.32B) of sizes 36.1kDa and 25.2kDa (Figure 6.32E), which were similar to bands 3 and 4 of figure 6.2 (35.2kDa and 24.2kDa)

When NS1mf is probed with anti-FLAG, 2 bands were observed, 1 band similar to p23 and another band which migrated slower than the full length of NS1. This should be due to the added myc-tag at the C-terminus which adds 10 amino acids to the full length of the protein. However, when probed with anti-myc, NS1mf does not show p23; instead band 6 was observed which has an estimated size of 14.7kDa. This band will be referred throughout this study as p15.
Figure 6.32 SDS-PAGE analysis of individual NS1 domains. A. Schematic representation of the separate domains of Influenza B/Lee/40 NS1 which were cloned into pCAGGS vector for transient expression. NS1: full length NS1 protein, NS1mf: full length NS1 protein doubly tagged with myc at C-terminus and FLAG at N-terminus, RB: RNA Binding Domain, RBL: RNA Binding Domain + Linker, ED: Effector Domain, LED: Linker + Effector Domain. MYC and FLAG epitopes are used as epitope tags to assist in immuno-labeling. B. 16 hour expression of constructs in HEK 293T cells. Cells were harvested in Laemmli buffer and equal volumes of each sample were loaded on a 15% SDS-PAGE gel. Samples were detected by western blotting with anti-myc and anti-FLAG antibodies. C. Western blot analysis of RBL on 15% SDS-PAGE gel probed by both anti-myc and anti-FLAG antibodies. kDa: KiloDaltons Black Arrow: Protein band corresponding to the full length of the influenza B NS1 protein: Black Arrow head: Protein band corresponding to p23. Asterisks (*): Protein band corresponding to p15. D. Log MW vs. Rf chart plotted by comparing the Rf and the molecular weights of the molecular weight standards of gel in C. E. Estimated sizes of bands labeled in B through the use of the equation of the slope obtained in D.
The effector domain by itself has an estimated molecular weight of 21.9kDa, while the Linker + Effector Domain (LED) has a molecular weight of 29.2kDa. When LED is probed with anti-FLAG a band of size 25.9kDa is observed, this band is not observed when probed with anti-myc. (Figure 6.32)

The RNA Binding Domain (RB) has a molecular weight of 12.7kDa (Band 9). The RBL (RB + Linker) displays 2 bands when probed with anti myc (Band 7: 17.7kDa & Band 8: 14.0kDa) but only 1 with anti-FLAG, which corresponds to the same size as band 7 (Figure 6.32B & 6.32C). Due to their similarity in molecular weight, it is highly possible that bands 6 and band 8 correspond to the same peptide, p15.

Bands 2 and 4 seem to be similar in their migration on the gel, just as bands 6 and 8. The data from this gel shows that p23 is larger than the ED but smaller than the LED. Expression of the LED alone also yield p23, suggesting that p23 is derived from the sequence of the LED alone. Constructs containing the linker sequence show 2 bands on gel electrophoresis, suggesting that the linker sequence is responsible for this phenomenon.

6.2.4.3.4 Expression of influenza B NS1 of the representative clinical specimens

Each of the constructs of Figure 6.16 were transfected in HEK 293T cells for 16 hours, cells were harvested directly in Laemmli buffer and equal volumes of sample were loaded onto a 12% SDS-PAGE gel. As before, the NS1 protein was detected by western blotting with the aid of an antibody specific to the influenza B NS1 protein as well as the FLAG epitope (Figure 6.33).

Probing with both antibodies showed similar banding patterns. Out of the 6 different constructs analysed in figure 6.6, 4 displayed the presence of p23 (Lee-NS1, 136-NS1, 117-NS1 and 70-NS1), while 132-NS1 and 147-NS1 did not. Secondly, specimen 117-NS1 displayed a fainter band corresponding to the full length of the NS1 protein while a more prominent p23 band (Figure 6.33). To determine if the results obtained in figure 6.33 were reproducible, the experiment was repeated another 2 times. The intensity of the bands were measured by the use of a densitometer and a ratio corresponding to the ratio of the intensity of p23 to the intensity of the full length protein band was calculated, averaged and used to characterise each of the constructs (Table 6.7)
Figure 6.33 Transfection of the 6-NS1 constructs listed in Figure 6.1 on HEK 293T cells. Cells were 16 hours post transfection in Laemmli buffer. Similar volumes of sample were loaded onto a 12% SDS-PAGE gel and analysed by western blotting. Top Panel: Samples were probed with anti FLAG antibody. Bottom Panel: Samples were probed with anti influenza B NS1 antibody. kDa: KiloDaltons Black Arrow: Protein band corresponding to the full length of the influenza B NS1 protein: Black Arrow head: Protein band corresponding to p23.

Constructs 132-NS1 and 147-NS1 had the lowest p23/full length ratios amongst all of the constructs analysed. Interestingly construct 117-NS1 was the only specimen which consistently displayed a ratio >1, suggesting that more of p23 was present than the full length protein (Table 6.7).

Table 6.7 Ratio of p23: Full Length NS1 band intensity. 3 separate transfections were carried out for all 6 constructs. Samples were harvested 16 hours post transfection and analysed by western blotting by an antibody specific for the FLAG epitope and another specific to the influenza B NS1 protein. The intensities of each band were measured by a densitometer and the averages of the three ratios are listed in the table above. An average of ratios obtained from anti FLAG and anti NS1 is also listed.

<table>
<thead>
<tr>
<th>Ratio of intensity p23:Full Length NS1</th>
<th>Anti FLAG</th>
<th>Anti NS1</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee-NS1</td>
<td>0.54</td>
<td>0.59</td>
<td>0.57</td>
</tr>
<tr>
<td>136-NS1</td>
<td>0.46</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>117-NS1</td>
<td>3.06</td>
<td>1.02</td>
<td>2.0</td>
</tr>
<tr>
<td>132-NS1</td>
<td>0.11</td>
<td>0.34</td>
<td>0.22</td>
</tr>
<tr>
<td>70-NS1</td>
<td>1.25</td>
<td>0.41</td>
<td>0.83</td>
</tr>
<tr>
<td>147-NS1</td>
<td>0.22</td>
<td>0.55</td>
<td>0.39</td>
</tr>
</tbody>
</table>
6.2.4.3.5 Cytoplasmic and nuclear fractionation of transfected/infected HEK 293T cells

To determine the cellular localisation of the NS1 protein as well as p23, HEK 293T cells were transfected with all 6 constructs. 20 hours post transfection; cells were harvested and fractionated to yield the cytoplasmic and nuclear fractions. The resulting fractions were then subjected to western blotting with the aid of the anti-FLAG antibody. To confirm the purity of the fractions, anti-Lamin A/C and anti-β-Tubulin were used as nuclear and cytoplasmic markers respectively (Figure 6.34A). An infection with B/Lee/40 was also carried out on HEK 293T cells. Infected cells were also fractionated as mentioned (Figure 6.34B).

**Figure 6.34 Cellular fractionation of NS1-transfected cells**

A. HEK 293T cells were transfected with each of the 6 representative strains followed by cellular fractionation. Each fraction was then subjected to 12% SDS-PAGE analysis and protein bands were detected by western blotting with the aid of anti-FLAG antibody (Bottom Panels). Top Panels; membrane was probed by anti-Lamin A/C (69kDa & 62kDa) and anti-β Tubulin (55kDa) as nuclear and cytoplasmic markers. B. HEK 293T cells were infected with B/Lee/40 and fractionated as in A. Total, prior to fractionation a small aliquot of cells were harvested directly in Laemmli Buffer. Cyto, Cytoplasmic Fraction. Nuc: Nuclear Fraction. Protein bands were detected by western blotting by probing with anti-influenza B NS1 antibody. kDa: KiloDaltons Black Arrow: Protein band corresponding to the full length of the influenza B NS1 protein: Black Arrow head: Protein band corresponding to p23.

β-Tubulin was detected strongly in the cytoplasmic fraction and very faintly in the nuclear fraction while Lamin A/C was detected only in the nuclear fraction, proving that the fractions were pure. The full length protein of each construct was detected at
similar intensities in both the cytoplasmic and nuclear fractions. p23 of specimens Lee-NS1, 136-NS1, 117-NS1 and 70-NS1 were only strongly detected in the cytoplasmic fraction. The p23 band corresponding to those specimens were not detected as intense in the nuclear fraction (Figure 6.34A). Similar to the observations seen in the transfections, when B/Lee/40-infected HEK 293T cells were fractionated, p23 was only observed in the cytoplasmic fraction and not the nuclear. Full length NS1 protein was detected in both fractions (Figure 6.34B).

To further understand the nature of this phenomenon, the doubly-tagged Lee-NS1 myc-flag construct was transfected onto 293T cells and subjected to cellular fractionation. The total, nuclear and cytoplasmic fractions were then probed with anti-Flag and anti-myc antibodies to determine the cellular locations of p23 and p15 (Figure 6.35).

![Figure 6.35](image_url)

**Figure 6.35** HEK 293T cells were transfected with Lee NS1 myc-flag for 20 hours and subjected to cellular fractionation. Samples were then ran on a 12% (top 2 panels) and 15% (Bottom Panel) SDS-PAGE gel followed by western blotting. Lysates were probed for Lamin A/C (top panel), flag (middle panel) and myc (bottom panel). Total: cell lysates prior to fractionation. Nuc: Nuclear fraction. Cyto: Cytoplasmic fraction. kDa: KiloDaltons Black Arrow: Protein band corresponding to the full length of the influenza B NS1 protein; Black Arrow head: Protein band corresponding to p23, Asterisk (*): Protein band corresponding to p15.
Just as in figure 6.34, the samples were probed with anti-Lamin A/C, a nuclear marker to assess the purity of the fractions obtained. The full length of the Lee-NS1mf is observed in the total, cytoplasmic and the nuclear fractions; however, p15 is seen in only the total and cytoplasmic fractions (Figure 6.35). This observation is similar to that of p23 which is not seen in the nuclear fraction (Figure 6.34).

6.2.4.3.6 Mutational analysis of influenza B NS1 protein.

The difference in the amino acid sequence of the representative strains could be exploited in understanding this difference in ‘phenotype’ between the specimens. Specimens 132-NS and 147-NS1 (not expressing p23) differ from the rest of the specimens only at 1 position, 92 (Figure 6.16). While the rest of the specimens contain a D$_{92}$, 132-NS1 and 147-NS1 harbor N$_{92}$. Hence to determine if this amino acid is responsible for the presence of p23, the gene of Lee-NS1 was mutated to code for an asparagine at position 92. This mutant is denoted as Lee-NS1 D92N. The expression of this mutant is compared to that of Lee-NS1 and 132-NS1 (Figure 6.36). Lee-NS1 D92N expressed only the full length NS1 protein and no p23 was detected, similar to that of 132-NS1 (Figure 6.36).

![Figure 6.36 Western blot analysis of Lee-NS1, Lee-NS1 D92N, and 132-NS1.](image)

Figure 6.36 Western blot analysis of Lee-NS1, Lee-NS1 D92N, and 132-NS1. HEK 293T cells were transfected with the respective constructs. 16 hours post transfection cells were harvested in Laemmli buffer and similar amounts of sample were loaded on to a 12% SDS-PAGE gel. Protein bands were subsequently detected by probing with anti-FLAG antibody. kDa: KiloDaltons Black Arrow: Protein band corresponding to the full length of the influenza B NS1 protein: Black Arrow head: Protein band corresponding to p23.
To determine if Lee-NS1 D92N was not expressing p23 at all rather than expressing p23 at a much slower rate than Lee-NS1, a time course experiment was set up. HEK 293T cells were transfected with Lee-NS1 and Lee-NS1 D92N for 12, 24 and 36 hours (Figure 6.37).

![Figure 6.37 Time course analysis of Lee-NS1 and Lee-NS1 D92N](image)

Figure 6.37 Time course analysis of Lee-NS1 and Lee-NS1 D92N. Western blot analysis of Lee-NS1 and Lee-NS1 D92N in HEK 293T for 12, 24 and 36 hours. At each time point cells were harvested in Laemmli buffer and similar amounts of sample were loaded on to a 12% SDS-PAGE gel. Protein bands were subsequently detected by probing with anti-FLAG antibody. pC: Cells transfected with empty pCAGGS vector for 36 hours. kDa: KiloDaltons Black Arrow: Protein band corresponding to the full length of the influenza B NS1 protein: Black Arrow head: Protein band corresponding to p23.

The intensity of both the full length and p23 increased with time but even at 36 hours post transfection, no band corresponding to p23 was detected in cells transfected with Lee-NS1 D92N (Figure 6.37).

The reverse mutation was then introduced in 132-NS1, asparagine was replaced with aspartic acid at position 92. This mutant is denoted as 132-NS1 N92D. The phenotype of 132-NS1 N92D was compared to that of Lee-NS1, Lee-NS1 D92N and 132-NS1 (Figure 6.38A). As a negative control, cells transfected with empty pCAGGS vector were also loaded onto the gel. Substituting asparagine for an aspartic acid at position 92 resulted in 132-NS1 expressing p23 as well as the full length NS1 protein (Figure 6.38).

Since amino acid 92 lies at the interface between the RB and the linker, several other amino acid substitutions were introduced at this region to assess their role in the expression of p23 (Figure 6.39A). The S94P substitution was carried out as specimen 117-NS1 has a proline at position 94 unlike the rest of the clinical specimens. 117-NS1 was also mutated to code for a serine at position 94. This mutant is denoted at 117-NS1 P94S (Figure 6.39B).
Figure 6.38 SDS-PAGE analysis of Lee-NS1, Lee-NS1 D92N, 132-NS1 and 132-NS1 N92D

A. Alignment of Lee-NS1, Lee-NS1 D92N, 132-NS1 and 132-NS1 N92D, amino acids 81-102. Identical amino acids are represented by a dot (.).

B. Western blot analysis of 132-NS1, 132-NS1 N92D, Lee-NS1 and Lee-NS1 D92N in HEK 293T. 16 hours post transfection cells were harvested in Laemmli buffer and similar amounts of sample were loaded on to a 12% SDS-PAGE gel. Protein bands were subsequently detected by probing with anti-FLAG antibody. pC: Cells transfected with empty pCAGGS vector for 16 hours. kDa: KiloDaltons Black Arrow: Protein band corresponding to the full length of the influenza B NS1 protein: Black Arrow head: Protein band corresponding to p23.

Substitutions D92A and P93A resulted in the same phenotype as D92N; a single band corresponding to the full length of the NS1 protein and no p23. No change was observed when Lee-NS1 was mutated to code for an alanine at position 91. When a proline was introduced in position 94 of Lee-NS1, the intensity of p23 detected was higher than wild type Lee-NS1, similar to that of 117-NS1. Conversely when P\textsubscript{94} of 117-NS1 was substituted for a serine, the band corresponding to p23 was less intense than that seen in 117-NS1. Table 6.8 clearly show that residues bearing a proline at position 94 have p23/full length ratios >1 while those with serines at this positions have ratios <1, strongly implicating the amino acid at position 94 in the expression of p23.
Another mutant analysed by western blotting was Lee-NS1 SSS$_{107-109}$. The expression of this mutant was compared to wild-type Lee-NS1 (Figure 6.40). Densitometer readings of the bands in figure 6.39 D and E were recorded and the ratios of band intensity of p23: full length NS1 was calculated and listed in table 6.8.

Table 6.8 p23:Full length ratio of Lee-NS1, Lee-NS1 M91A, Lee-NS1 P93A, Lee-NS1 P93A, Lee-NS1 S94P, 117-NS1 and 117-NS1 P94S. Densitometer readings of bands corresponding to full length NS1 and p23 were measured and the subsequent ratios were calculated. Bands from figure 6.39D & E were analysed in this table.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Sample Name</th>
<th>p23/Full Length Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.39D</td>
<td>Lee-NS1</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>Lee-NS1 M91A</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Lee-NS1 P93A</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Lee-NS1 S94P</td>
<td>1.80</td>
</tr>
<tr>
<td>6.39E</td>
<td>Lee-NS1</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>117-NS1</td>
<td>2.62</td>
</tr>
<tr>
<td></td>
<td>117-NS1 P94S</td>
<td>0.45</td>
</tr>
</tbody>
</table>
Lee-NS1 SSS$_{107-109}$ still expresses 2 bands just like Lee-NS1, this is not surprising as specimens 70-NS1 and 117-NS1 also contain this triple serine motif at position 107 and still express p23. However, the 2 bands expressed by Lee-NS1 SSS$_{107-109}$ do seem to migrate slower on the SDS-PAGE gel than the 2 bands expressed by Lee-NS1. To quantify this ‘upward shift’, a log MW vs. $R_f$ plot was generated and the equation of the slope was used to estimate the sizes of each band observed in figure 6.39A. There exists a 1.19kDa size difference between bands 3 and 1 while bands 4 and 2 have a 0.86kDa difference. This suggests that p23 contains this triple serine motif and that this site might be modified to cause this retardation in migration (Figure 6.40).

The mutations introduced in specimen 132-NS1 were analysed by western blotting to determine if any of those amino acids substituted bear an impact on p23 expression (Figure 6.41).

None of the mutations substituting the unique serines of specimen 132-NS1 resulted in the expression of p23 (Figure 6.42).
6.2.4.3.7 Effect of protease inhibition on the expression of Lee-NS1

A possible explanation for the results presented up to this point is that specimens containing D92 and P93 are cleaved by a cellular protease. To test this hypothesis we expressed influenza B NS1 through transfection and infection of B/Lee/40 and inhibited the activity of caspases and matrix metalloproteinase (MMPs).

Caspases are cysteine proteases which get activated during periods of cellular stress, such as during infection. The activated caspases then work to cleave specific targets. While 12 different caspases exists in the human proteome and each recognizes a set of different proteins, a large majority of them cleave a specific D-X bond in their targets (Alnemri et al., 1996; Luthi and Martin, 2007). The chemical, Z-VAD-FMK, is a pan-caspase inhibitor, was used to inhibit any possible caspase activity (Figure 6.42). 2 experiments were set up to double confirm the results obtained. In the first set, cells were transfected for a total of 10 hours while in another cells were transfected for 20 hours. In both instances the caspase inhibitor was added 5 hours post transfection. No difference was observed between the Z-VAD-FMK treated and untreated samples (Figure 6.42).

MMPs are a class of host cell proteases which are proteolytically activated to cleave cellular proteins comprising the extra cellular matrix (Nagase and Woessner, 1999; Spinale, 2002). The activity of certain MMPs has been shown to be inhibited by the antibiotic, doxycycline (DOX) (Chang et al., 2010; Hanemaatijer et al., 1998). Figure 6.43 shows the effect of DOX on HEK 293T cells transfected with Lee-NS1. Cells were transfected for 20 hours and DOX at a range of concentrations (0-200μg/ml)
4 hours post transfection. As a control, the membrane was probed with anti-actin to assess if DOX had any cytotoxic effect on the cells (Figure 6.43).

No change was observed between the range of 0-50μg/ml of DOX. However at 100μg/ml and especially 200μg/ml the intensity of both bands decreased. The intensity of actin is also decreased at these 2 concentrations. So while Lee-NS1 treated with 200μg/ml of DOX only shows 1 band, it is not conclusive to say that DOX inhibits Lee-NS1 cleavage (Figure 6.43).

**Figure 6.42 VAD treatment on Lee-NS1 expression.** HEK 293T cells were transfected with empty pCAGGS vector and Lee-NS1mf. Lanes 2&3 Cells were transfected with Lee-NS1mf for 10 hours. Lanes 3&4, cells were transfected with Lee-NS1mf for 20 hours. In both cases Z-VAD-FMK was introduced 5 hours post transfection. kDa: KiloDaltons. Black Arrow: Protein band corresponding to the full length of the influenza B NS1 protein. Black Arrow head: Protein band corresponding to p23.

**Figure 6.43 Doxycycline treatment of Lee-NS1 expression.** HEK 293T cells were transfected with empty pCAGGS vector (pC) or Lee-NS1 for 20 hours. 4 hours post transfection doxycycline (DOX) at different concentrations. Cells were harvested directly in Laemmli buffer and similar amounts of sample were loaded onto a 12% SDS-PAGE gel for western blotting. Protein bands were detected by probing with anti FLAG antibody and as a control, the sample was probed with anti actin (lower panel).
6.3 Discussion

The HA, NA and NB amino acid profiles of DSO_010147_2007 are unique among the clinical specimens sequenced. What is even more puzzling is that its NS gene clusters closely with the rest of the clinical specimens in cluster III (Figure 6.5). And the amino acid profiles of its NS1 and NEP proteins are not significantly different from the rest (Tables 6.2, 6.3 and 6.4, Supp Fig 1,2,3 & 4). It is plausible to assume that specimen DSO_010147_2007 arose from an exogenous virus source, re-assorting with local strains carrying the NS gene segment of cluster III.

The phylogenetic tree of the NS gene segment indicates that the NS gene cannot be classified under either of the ancestor strains as of the HA or NA gene segments (Figure 6.5). It has been found that the both ancestor strains share a common NS gene segment (Abed et al., 2003; Luo et al., 1999; Matsuzaki et al., 2004; McCullers et al., 2004). This evidence is further proof that the 2 ancestor strains are inadequate to classify all the gene segments of influenza B viruses.

McCullers et al, have reported that the NS gene does not follow a linear evolutionary pattern as exhibited by the other gene segments of influenza B viruses (McCullers et al., 2004). This phenomenon is seen in figure 6.5, where all of the clinical specimens have their NS gene segments clustering in cluster III, except for 7 of the specimens isolated in 2004 which were found in cluster I, which contained B/Lee/40 as the only other member (Figure 6.5). The NS1 and NEP proteins of these seven specimens share similar amino acid profiles to B/Lee/40 (Tables 6.2, 6.3 & 6.4, Supp Fig 4 & 5). Why there is this reversion of the NS gene segment of the 2004 specimens to the NS gene segment is puzzling. This is the only instance in this study where a gene segment clusters with B Lee 40 (Figures 4.3, 5.3 and 6.5).

The specimens of cluster I have 2 substitutions within the NLS, which are not seen in the rest of the specimens analysed, however, these are conservative substitutions which may not have an effect on the nuclear localisation of these proteins (Table 6.2) (Schneider et al., 2009). Similar observations can be made to the amino acid substitutions at positions 34, 97 and 101 constituting three of the residues responsible for ISG15 binding (Table 6.3) (Guan et al., 2011a; Yuan and Krug, 2001). Further functional analysis have to be carried out to determine if the substitutions listed in the NS1 and NEP proteins result in any differences in protein activity.
The localisation of influenza B NS1 protein throughout infection in this study was similar to that which has been reported by Schneider et al, 2009, where in early infection influenza B NS1 localises to nuclear speckles and in later infection and this co-localisation is lost in later infection (Figures 6.2, 6.7 & 6.11) (Schneider et al., 2009). The localisation of NS1 during transfection is different, the nuclear inclusions increase in size over time as has been reported by Schneider et al, 2009 (Figures 6.24& 6.28a). This disparity is possibly due to the absence of the other viral proteins, especially the Nuclear Export Protein (NEP) during transfection. Interactions with other viral proteins may bring about this export of NS1 from nuclear speckles. The NEP protein carries out the function of exporting VRNPs out of the nucleus and may have a similar role in NS1’s export from nuclear speckles (Paragas et al., 2001).

A conclusion that can be made from this series of infections and transfections is that the influenza B NS1 rearranges the spatial organization of SC35 in an infected/transfected cell. Even late in infection, when NS1 no longer is found in nuclear speckles, the pattern of SC35 staining mirrors that of NS1. The difference is obvious when comparing the staining pattern of SC35 of infected/transfected cells to non-infected/transfected cells (Figures 6.17 & 6.28a). This suggests a possible functional correlation between the 2 proteins.

SC35 is a splicing factor which is required in the first step of splicing in mammalian cells (Fu and Maniatis, 1992a, b; Spector et al., 1991). SC35 contains 2 domains, the RNA recognition motif (RRM) as well as the arginine-serine-rich (RS) domain (Graveley, 2000). Hence this co-localisation between NS1 and SC35 could be brought about by their binding to RNA mediated by the RRM domain of SC35 and the RB domain of NS1. It has been postulated that the arginines and lysines between residues 47-54 of influenza B NS1 play a crucial role in RNA binding (Schneider et al., 2009). Variability corresponding to speckled arrangement of NS1 exists at position 52, where Lee-NS1 and 136-NS1 contain an asparagine, while speckle-associating constructs (117-NS1, 132-NS1, 70-NS1 and 147-NS1) contained a lysine. However, mutating Lee-NS1 to contain K52 did not increase the propensity of Lee-NS1 to co-localise with SC35. This suggests that the co-localisation with SC35 may not be facilitated by RNA binding.

While Schneider et al, 2009 found that the first 90 amino acids of the NS1 protein was the shortest possible domain required for speckle association, the results in
our study is not in agreement with this postulation. The expression of the RB (1-90) or RBL (1-120) of 132-NS1 or 117-NS1 in MDCK cells do not correspond to intra-nuclear speckles. A few explanations can be offered for this; firstly, the association with speckles may be a function of the full length NS1 protein and not just an individual domain. Secondly, individual domains may require a longer time for this apparent speckle association, hence transfection times of >20 hours would be required to observe such localisations. Thirdly, the incorporation of the FLAG/MYC epitopes in our study may interfere with the localisation of the various NS1 domains. The study by Schneider et al, 2009 was carried out by fusing the domains with GFP. GFP is a 27kDa protein which is larger than the size of the first 90 amino acids of NS1 (RB) (Prasher et al., 1992), conjugation of the RB to GFP may disrupt the function and localisation of the RB. FLAG and Myc epitopes are composed of 8 and 10 amino acids respectively (Evan et al., 1985; Thomas et al., 1988). Hence, fusing the individual domains to an epitope tag such as FLAG or Myc may serve as a better marker for monitoring localisation of the individual domains.

Splicing factors have been found to accumulate in enlarged rounded intra-nuclear speckles when transcription is inhibited (Melcak et al., 2000; Spector, 1993; Spector et al., 1991). This finding does suggest that the transfection of NS1 into cells may inhibit the process of transcription. Judging from the distinct speckles of SC35 in cells transfected with 117-NS1, 132-NS1, 70-NS1 and 147-NS1 it could be assumed that these constructs inhibit transcription activity to a greater extent than Lee-NS1 and 136-NS1 (Figures 6.17-6.19). If this line of reasoning holds true, the mutants: 132-NS1 ASS107-109, 132-NS1 AAS107-109 and 132-NS1 AAA107-109 may inhibit the process of transcription more than the wild type 132-NS1 as evidenced by the larger and numerous speckles seen in the mutants in figure 5.23. Mutating Lee-NS1 to Lee-NS1 SSS107-109 did not cause any observable changes in the absence of co-localisation to nuclear speckles. To further characterise this, 132-NS1 should be mutated to carry the sequence of Lee-NS1 between positions 107-109: 132-NS1 NPL107-109. Also Lee-NS1 should be mutated to replace NPL107-109 with alanines, such as in the case of 132-NS1.

From the data presented in this study, it seems likely that both 471-NS1 and 117-NS1 contain a NoLS. Both of these constructs present with large nuclear inclusions which co-localise with NPM. This is a hallmark of nucleolar localisation (Melen et al., 2007; Volmer et al., 2010). Even though the NoLS of influenza A NS1 has been
identified to the lysines and arginines between the positions 219-237 (highlighted in red in Figure 6.21), 471-NS1 only shares K$_{219}$. The sequence of 471-NS1 terminates at position 220. K$_{219}$ is also shared by PR8-NS1 which does not co-localise to the nucleolus (Figure 6.20), therefore it is probable that just harboring K$_{219}$ is not sufficient to target the NS1 protein to the nucleolus. Therefore, 471-NS1 may contain an undiscovered NoLS elsewhere in its sequence. Further mutational analyses coupled with confocal microscopy should be carried out to determine this.

117-NS1 is unique amongst all the influenza B NS1 proteins analysed in this study; it is the only one which co-localises with NPM, suggesting that it is targeted to the nucleolus of the cell. It localises to large nuclear inclusions which also co-localises with SC35 which is unique because SC35 usually presents itself in round speckles (Schneider et al., 2009). Lee-NS1 displays NPM exclusion in both late transfection and infection (Figures 6.15 & 6.18), suggesting that Lee-NS1 does not have an NoLS while 117-NS1 may. 117-NS1 only has three unique amino acids, not observed by any of the other constructs in this study; P$_{94}$, I$_{151}$ and G$_{154}$. Only P$_{94}$ was chosen to be substituted with serine as, P$_{94}$ lies in the linker region right adjacent to the RNA Binding Domain, which would be presumed to play a role in intra-nuclear targeting (Figure 6.16). However, no change in cellular localisation was observed in 117-NS1 P94S, even at 20 hours post transfection (Figure 6.28a). This leaves the possibility of I$_{151}$ and/or G$_{154}$ being responsible for this nucleolar localisation. It does seem unlikely that this nucleolar localisation is brought about by a single amino acid change as the NoLS of influenza A NS1 is composed of 6 basic residues. However, further mutational analysis might be required to confirm this. With the data gained in this study, it is also possible to presume that the co-localisation of NS1 to the NPM/SC35 is due to the difference in folding of the different NS1 proteins and not due to single amino acid changes.

The lack of changes observed in 132-NS1 S69N, 132-NS1 S123P and 132-NS1 S215N as compared to wild type 132-NS1 might suggest that these residues are not involved in SC35 targeting but does not rule out the possibility of these residues being phosphorylation sites (Figure 6.27). Other sites on 132-NS1 which could be potentially phosphorylated are S$_{123}$, T$_{143}$, Y$_{200}$ and S$_{215}$. Mutational analysis followed by phosphorylation detection should be carried out to determine if these sites are phosphorylated and could possibly target the NS1 protein to nuclear speckles (Bryantsev et al., 2007; Prasanth et al., 2003; Spector and Lamond, 2010).
The work in section 6.2.4.4 was predicated by the discovery of p23 in cells infected with B/Lee/40. p23 was present as early as the full length of the protein was detected during an infection. This suggests that p23 is not an aberrant breakdown product nor is it an artifact observed only during DNA transfection which is itself an artificial process (Figures 6.29 & 6.31A). The expression of p23 is clearly not a random process as it is sequence specific and genetic manipulations of the NS1 gene causes changes in the expression of p23.

Certain concrete conclusions can be made from the work presented in this chapter. Firstly, p23 is a product which contains the C-terminus of the NS1 protein. When the NS1 protein is singly tagged with a FLAG epitope at the C-terminus, p23 is detected when probed with anti-FLAG antibody. However, when the NS1 is doubly tagged (myc at the N-terminus and FLAG at the C-terminus), p23 is not detected by probing with anti myc antibody (Figure 6.32). This suggests that p23 does not contain the N-terminus of the NS1 protein. In fact when NS1mf is probed with anti-myc, a smaller product of 14.7kDA is observed (Band 6 of Figure 6.32). This smaller product (which shall be annotated as p15) contains the N-terminus of the NS1 protein.

Interestingly, p23 was also detected by the anti-influenza B NS1 antibody. p15 was not detected when the samples were probed with anti influenza B NS1, suggesting that recognition epitope of this antibody lies closer to the C rather than the N-terminus (Figures 6.29, 6.31, 6.33 & 6.34).

Figure 6.44 shows an alignment of the NS1 gene against the NEP gene and their corresponding amino acid sequence. The ability of p23 to be detected by anti-FLAG indicates that the open reading frame for p23 is in frame with the open reading frame of the full length of NS1 protein. It is important to state this because the gene segment 8 also codes for the NEP. The NEP is expressed as a spliced product of the mRNA transcribed from gene segment 8. While the NEP shares the first 11 amino acids with NS1, the rest of the NEP ORF is in the +1 frame with respect to NS1 (Briedis and Lamb, 1982) (Figure 6.44). This makes it impossible for p23 to be the NEP. It is also not possible for p15 to be or constitute part of the NEP even though p15 can be detected by the anti myc antibody. p15 is detected when the RBL is expressed. The RBL is coded for by the first 360 amino acids of the NS1 ORF. The NEP gene is comprised of 2 exons flanking an intron which is between nucleotides 28 and 688 (Figure 6.44). Since the RBL is coded for by nucleotides 1-360, the second exon of the NEP is not
present in the RBL ORF. Since splicing requires the 5’ and 3’ terminus of the first and second exons to be intact, it is highly impossible for the RBL (or LED) ORF to be spliced (Rogers and Wall, 1980).

When the RB or ED is expressed alone, only 1 protein band is detected when probed with anti myc or FLAG respectively. However, when these domains were fused to the linker (RB+Linker and Linker+ED) two bands were observed; 1 at the expected size of the product and another smaller product. This suggests that presence of the linker is promoting the expression of the smaller product. Interestingly when LED is expressed, the smaller product seen is the same size as p23, strongly suggesting that p23 is a truncated form of LED (Figure 6.32). The smaller product detected when RBL is probed with anti-myc is of the same size of p15. This then suggests that p15 is a truncation of the RBL (Figure 6.32).

When Lee-NS1 was mutated to Lee-NS1 SSS107-109, p23 was still expressed at similar intensity as wild type Lee-NS1. However, when the sizes were estimated, both products were approximately 1kDa heavier than their wild-type counterparts (Figure 6.39). Since p23 of this mutant also undergoes this size shift, it can be concluded that p23 contains the sequence 107-109. It is possible that these 3 serines in tandem are subject for post translational modification (likely phosphorylation) and more confirmatory experiments need to be carried out to determine this.

Another suggestion that can be placed forward is that p23 is expressed due to an internal initiation site within the NS1 linker region. To explore this possibility, the only start codon (ATG) in the linker was mutated to GCG. This mutation resulted in a substitution of M91A. p23 was still observed in this mutant, allowing a conclusion that p23 is not produced due to an internal initiation site (Figure 6.37).

Cellular fractionation of infected as well as transfected cells reveal that p23 and p15 exist primarily in the cytoplasm and not the nucleus, while the full length NS1 is present in both the cytoplasm and nucleus of the cell (Figures 6.34 & 6.35). As discussed in chapter 5, the NLS of influenza B NS1 exists between amino acids 46-56 (Dauber et al., 2006; Schneider et al., 2009). Western blot analysis of cellular fractionations of infections and transfections show that both p23 and p15 were present in the cytoplasm, even at a ‘late’ time-point of 20 hours. The presence of both p23 and p15 in the cytoplasm, suggests that the full length NS1 protein is processed to yield p23 and p15 in the cytoplasm. While the full length protein was able to enter the nucleus,
neither of the 2 smaller fractions were. p23 was predicted not to enter the nucleus as the
data till this point seems to suggest that amino acids 46-56 are not in p23. 2 different
explanations could explain the absence of p15 from the nucleus. Firstly, p15 does not
contain the NLS. Secondly, p15 does contain the NLS but is withheld from entering the
nucleus by unknown cellular (or viral) processes. Since western blot analysis of the
RBL domain shows a band corresponding to the same size as p15, it is likely that it
does contain the NLS and that the latter explanation holds true (Figure 6.32B).

The results presented thus far seem to strongly point to the possibility that the
NS1 protein is cleaved within the linker region yielding p23 and p15. This proteolysis is
affected by amino acids between positions 92-94. The sequence DPS_{92-94} allow for this
proposed proteolysis and DPP_{92-94} promotes more of it. APS_{92-94} and DAS_{92-94} blocks
the expression of p23.

To analyse the amino acid sequence of the region between the RB and the
linker of influenza B NS1 protein, 388 sequences were downloaded from the Influenza
2008). These specimens were isolated between the years 1940-2012. An alignment of
these specimens, highlighting the amino acid sequence 87-97 is shown in table 6.9. The
specimens are arranged in order of year of isolation and are compared against the
sequence of B/Lee/40 NS1.

The amino acid sequence in the boundary between the RB and the linker is
generally conserved throughout the specimens aligned in table 6.9. The only position
that displays variability is position 92, where either amino acid D or N is tolerated. Out
of the 388 specimens, 307 harbored N_{92}, 80 harbored D_{92} and 1 specimen had ambiguity
in the identity at position 92 (B/Malaysia/1749642/2007 contained B_{92}, where B: D or
N). Interestingly 3 specimens displayed variability at position 94, while the rest
displayed S_{94}, specimens B/Hong Kong/22/1989, B/Temple/B18/2003 and B/Tennessee
UR06-0310/2007 had F_{94}. No specimen analysed in table 6.9 had a proline at position
94. It would be interesting to analyse the characteristic of the proteins containing F_{94} as
phenylalanine is a hydrophobic amino acid while serine is hydrophilic.
Alignment of Influenza B NS1 and NEP

The translated amino acid sequence is located underneath the gene sequence. D92 and its corresponding codon is highlighted in red.
Table 6.9 Alignment of 388 Influenza B NS1 protein, highlighting amino acids 87-97. Year of isolation is reflected in the strain name. Specimens are arranged according to year of isolation. The sequence of B/Lee/40 NS1 is used as the reference strain. Identical residues are indicated by (.) ‘X’ :lack of sequence information at that position. ‘B’: Asparagine or Aspartic Acid

<table>
<thead>
<tr>
<th>Year</th>
<th>Strain</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>B Taiwan 407</td>
<td>N..</td>
</tr>
<tr>
<td>2006</td>
<td>B Alabama 02</td>
<td>N..</td>
</tr>
<tr>
<td>2007</td>
<td>B Washington 03</td>
<td>N..</td>
</tr>
<tr>
<td>2008</td>
<td>B Ohio 01</td>
<td>N..</td>
</tr>
<tr>
<td>2009</td>
<td>B Hong Kong 15</td>
<td>15</td>
</tr>
<tr>
<td>2010</td>
<td>B Illinois 03</td>
<td>N..</td>
</tr>
<tr>
<td>2011</td>
<td>B Washington 08</td>
<td>N..</td>
</tr>
<tr>
<td>2012</td>
<td>B Nevada 01</td>
<td>N..</td>
</tr>
<tr>
<td>2013</td>
<td>B Vermont 01</td>
<td>N..</td>
</tr>
<tr>
<td>2014</td>
<td>B Connecticut 02</td>
<td>N..</td>
</tr>
<tr>
<td>2015</td>
<td>B New Mexico 01</td>
<td>N..</td>
</tr>
<tr>
<td>2016</td>
<td>B Ohio 01</td>
<td>N..</td>
</tr>
<tr>
<td>2017</td>
<td>B Nevada 01</td>
<td>N..</td>
</tr>
<tr>
<td>2018</td>
<td>B Georgia 01</td>
<td>N..</td>
</tr>
<tr>
<td>2019</td>
<td>B Taiwan 01</td>
<td>N..</td>
</tr>
<tr>
<td>2020</td>
<td>B Washington 08</td>
<td>N..</td>
</tr>
<tr>
<td>2021</td>
<td>B Ohio 01</td>
<td>N..</td>
</tr>
<tr>
<td>2022</td>
<td>B Taiwan 01</td>
<td>N..</td>
</tr>
<tr>
<td>2023</td>
<td>B New Mexico 01</td>
<td>N..</td>
</tr>
<tr>
<td>2024</td>
<td>B Ohio 01</td>
<td>N..</td>
</tr>
<tr>
<td>2025</td>
<td>B Nevada 01</td>
<td>N..</td>
</tr>
<tr>
<td>2026</td>
<td>B Georgia 01</td>
<td>N..</td>
</tr>
<tr>
<td>2027</td>
<td>B Taiwan 01</td>
<td>N..</td>
</tr>
<tr>
<td>2028</td>
<td>B New Mexico 01</td>
<td>N..</td>
</tr>
<tr>
<td>2029</td>
<td>B Ohio 01</td>
<td>N..</td>
</tr>
<tr>
<td>2030</td>
<td>B Nevada 01</td>
<td>N..</td>
</tr>
</tbody>
</table>

*Table continued...*
Using the data from table 6.9, a bar chart displaying the proportions of the specimens harboring N$_{92}$ or D$_{92}$ was generated (Figure 6.45). This retrospective analysis of influenza B NS1 protein shows that while specimens harboring D$_{92}$ were a minority, their prevalence were contained to 1940, 1954 and the period between 1988-2007. No specimens containing D$_{92}$ were isolated in the years post 2008, despite the larger number of sequenced specimens deposited into the database. The specimens sequenced and analysed in chapter 2 belonged to 2004-2009, in which 19.3% of specimens contained D$_{92}$.

This data suggests that the NS1 protein of majority of the specimens would display as a single band on SDS-PAGE analysis, not displaying p23 or p15. The resurgence and disappearance of specimens harboring D$_{92}$ does seem to suggest that this amino acid substitution plays a role in the evolutionary fitness of the virus. More data, especially pertaining to the clinical pathology as well as antiviral counter activity needs to be collected to determine any correlation.

![Amino Acid Identity at Position 92](1940-2012)

**Figure 6.45** Proportion of specimens displaying N$_{92}$ (blue) or D$_{92}$ (red) over the period of 1940-2011. Y-axis values represent numbers of specimen, while the X-axis displays the year of specimen isolation.

Caspase activity was suspected due to the dependence of amino acid D$_{92}$ for the expression of p23 as (most) caspases cleave a specific D-X bond (Alnemri et al., 1996; Luthi and Martin, 2007). This suggestion was also strengthened by the finding that a large number of viral proteins are cleaved by caspases during infection (Richard and Tulasne, 2012). Even influenza A nucleocapsid (NP) and M2 have been proven to be
cleaved by a caspase during infection (Zhirmov et al., 1999; Zhirmov et al., 2002). However, inhibition of caspase activity did not yield a reduction in p23 expression. In this experiment, caspase inhibitor was added to the culture 5 hours post transfection (Figure 6.42). This may not be adequate in inhibiting caspase activity. This experiment should be repeated and the cells should be pre-treated with the anti-caspase drug.

However, it is important to note that caspases are usually not in their active forms when cells are not underdoig apoptosis (Pop and Salvesen, 2009). Simply transfecting a DNA plasmid into a cell does not activate the apoptosis pathway, these can be seen in the immunofluorescence images of cells transfected with NS1 protein for long periods of time (Section 6.2.4.1), no blebbing of the membrane or shrinkage of cells were observed (Saraste and Pulkki, 2000). The published report has determined a link between influenza B NS1 protein with apoptosis. However, the NS1 protein of PR8 (Influenza A/Puerto Rico/8/1934 H1N1) has been shown to play an active role in preventing infected cells from undergoing apoptosis (Zhirmov et al., 2002). Conversely the NS1 protein of an avian influenza A virus (A/Chicken/Jilin/2003) has been linked to the promotion of caspase-induced apoptosis in human cells (Zhang et al., 2010). The role influenza B NS1 plays in relation to apoptosis must be verified before concluding if p23 is a product of caspase cleavage.

Similarly, cells should also be pre-treated with DOX prior to transfection, to ensure that all MMPs are inhibited before reaching the conclusion that influenza B NS1 is not cleaved by MMP (Figure 6.44).

As of late no full crystal structure of the influenza B NS1 has been solved. The only crystal structure of influenza B/Lee/40 NS1 involves a dimer of just the first 103 amino acids bound to 2 molecules of ISG15 (MMDB ID: 92209 PDB ID: 3SDL) (Guan et al., 2011b; Madej et al., 2012). The structure shows the RB comprising 3 alpha helices, followed by a stretch of 13 amino acids of undetermined secondary structure (Figure 6.46).

The amino acids which demonstrate a correlation with p23 expression are highlighted in figure 6.46. All 13 amino acids of the linker shown in the crystal structure above are not shown to be present in any fixed secondary structure. This may suggest that amino acids 92-94 are present on the exterior of the protein, susceptible to host cell proteases. However, further structural data, especially pertaining to the structure of the mutant (D92N, P93A & S94P) needs to be ascertained to determine this.
Figure 6.46 3D structure of influenza B NS1  
B. Close up view of a single molecule of the first 103 amino acids of B/Lee/40. D$_{92}$ and S$_{94}$ are highlighted in yellow. Structures viewed with 3D structure viewer Cn3D ver 4.3

If no known proteases are shown to be the cause of this cleavage, it would also be possible to hypothesize that p23 is a product of an autocatalytic cleavage process, where the amino acid residues within the linker play an active role in this process.
Chapter 7: Molecular characterisation of influenza B BM2 protein cloned from non-tissue culture adapted clinical specimens

7.1 Introduction

Gene segment 7 is bicistronic in both influenza A and influenza B viruses (Fields et al., 2007). However, unlike in influenza A viruses, the BM2 gene is not a product of gene splicing (Lamb and Choppin, 1981). The stop codon of the influenza B M1 ORF overlaps with the start codon of the BM2 ORF in a motif known as the stop-start pentanucleotide, UAAUG. This couples the end of translation of the M1 protein to the initiation of BM2 translation (Figure 1.18) (Hatta et al., 2009; Horvath, 1990; Powell et al., 2008).

The BM2 protein is a structural protein which is expressed late in infection. In the membrane, the BM2 protein has its N-terminus in the ectodomain and it’s C-terminus in the cytoplasmic domain, similar to the orientation of the influenza A M2 protein. The absence of a signal peptide within its sequence classifies it as a type III integral protein (Odagiri et al., 1999; Paterson et al., 2003; Rouse et al., 2009).

The BM2 protein functions primarily in the uncoating step after receptor mediated endocytosis. 4 subunits of the BM2 form a proton channel which facilitates the entry of protons from the endosomal lumen into the virion. This increase in pH promotes the dissociation of vRNPs from the viral matrix (M1) allowing for vRNP to enter the host cell nucleus (Chizhmakov et al., 1996; Chizhmakov et al., 2003; Mould et al., 2003b; Paterson et al., 2003).

While the function of BM2 is similar to that of influenza A M2, there exist several differences in the structure and function of the proteins. Firstly, the 2 proteins display very low sequence identity, with the exception of a HXXXW motif in the transmembrane region, believed to be responsible for ion channel gating (Figures 1.20 & 1.21) (Ma et al., 2008; Pinto and Lamb, 2006). Secondly, while the influenza A M2 tetramer is formed by non-covalent interactions of 2 di-sulphide linked dimers, no cysteines exist in the BM2 cytoplasmic domain, removing the possibility of di-sulphide formation (Balannik et al., 2008; Sugrue and Hay, 1991). Lastly, the proton channel function of the influenza A M2 is inhibited by amantadine while amantadine does not have a palpable effect on BM2 function (Pinto and Lamb, 2006; Wang et al., 2009).
Initial analysis with the use of radioactive phosphorous ($^{32}$P), suggests that the BM2 is phosphorylated, however a recent study could not identify any residues on the protein which have been phosphorylated (Hutchinson et al., 2012; Odagiri et al., 1999).

The work in this chapter is dedicated to the molecular characterisation of the BM2 proteins expressed by the clinical specimens isolated in this study. Two of the specimens cloned displayed I14V substitution within the transmembrane domain which sparked and investigation as to the localisation of these proteins. Other amino acid substitutions observed were exploited to study the possibility of domains involved in protein targeting and modification.

7.2 Results

7.2.1 Molecular Cloning of the BM2 protein and Expression by Plasmid DNA Transfection

Using the cDNA generated from the representative clinical specimens as template, the BM2 gene was PCR amplified using primers BM2-For and BM2-Rev FLAG (Chapter 2, Table 2.4). Thermocycling conditions used were similar to those used in amplifying the NS1 and NB genes. A PCR product of about 350 was observed on an ethidium bromide stained agarose gel (Figure 7.1).

![Figure 7.1 BM2 ORF amplified from B/Lee/40 visualized on an ethidium bromide stained 1% agarose gel. bp: Base pairs.](image)

The ORF from the Lee as well as the 5 representative strains were cloned into the EcoRI restriction site on the pCAGGS mammalian expression vector. To confirm the genes were successfully cloned, the plasmids generated were sent for sequencing to determine if the gene sequence is in frame with the FLAG epitope tag and is in the forward orientation. The alignment of the strains is shown in Figure 7.2.
Sequence analysis show that the BM2 protein displays 20 positions of variability. Specimen 70-BM2 is identical to Lee-BM2 except at position 41 where 70-BM2 displays a methionine instead of an isoleucine. Specimens 136-BM2 and 117-BM2 are 100% identical. Specimens 132-BM2 and 147-BM2 are identical except at position 5 where 147-BM2 has a phenylalanine, while 132-BM2 harbors a leucine. The transmembrane domain is highly conserved with only 132-BM2 and 147-BM2 displaying a valine at position 14 while the other specimens have an isoleucine. The HXXXW motif (positions 19-23), responsible for pH sensing and channel gating is conserved in all the specimens analysed in this study (Figure 7.2) (Ma et al., 2008; Pinto and Lamb, 2006).

![Alignment of the BM2 genes cloned in this study. ED: Ectodomain, TM: Transmembrane Domain, CT: Cytoplasmic Tail. The Cytoplasmic Tail and the FLAG tag are highlighted in the schematic. B. Schematic representation of the cloning strategy employed for these sequences. RE: Restriction Endonuclease. pCAGGS schematic, courtesy of Dr. Loo Liat Hui](image-url)
These plasmids containing the BM2 genes from representative strains were transfected into HEK 293T cells for 20 hours. The harvested cell lysates were analysed through western blotting by probing with anti FLAG antibody (Figure 7.3).

5 protein bands were observed upon western blot analysis of the BM2 proteins. Bands 2 (14kDa), 3 (22kDa) and 4 (32kDa) were observed in all specimens but band 1 (12kDa) was only observed in specimens 136-BM2 and 147-BM2. Band 1 corresponds with the unmodified form of the BM2 protein, 12.5kDa (The UniProt Consortium, 2012), while bands 2 and 3 may correspond the phosphorylated form of the BM2, which was previously identified as a band migrating at 17kDa (Odagiri et al., 1999). Band 4 may be an oligomeric form of the BM2 protein (Figure 7.3).

Figure 7.3 SDS-PAGE analysis of cloned BM2 protein. A. 15% SDS PAGE analysis of HEK 293T cells transfected with Lee-BM2, 136-BM2, 117-BM2, 132-BM2, 70-BM2 and 147-BM2. Mock: Cells transfected with empty pCAGSS (pC) vector. Cells were transfected for 20 hours and harvested in Laemmli buffer. Similar volumes were loaded and protein bands were detected by probing with anti-FLAG antibody. Bands observed are labeled 1, 2, 3 and 4. kDa: Kilo Daltons. B. Log MW vs Rf chart plotted by comparing the Rf values of the molecular standards against the log of their molecular weights (MW). C. Inferences of the sizes of the 5 bands observed in (A) through the use of the equation obtained in (B).
To analyse the cellular localisation of the BM2 protein, these constructs were transfected into MDCK as well as A549 cells for 10 hours. Cells were fixed and permeabilized prior to staining with anti-FLAG antibody. Cells were also co-stained with anti GM130 antibody (Figures 7.4 & 7.5).

Figure 7.4 Transfections of BM2-pCAGGSin MDCK cells. Cells were fixed, permeabilised and stained 10 hours post transfection. Upper panels: Cells were stained with anti-FLAG. Lower panels: cells were co-stained with anti-GM130. Solid white arrow: strong perinuclear enhancement of the BM2 protein of Lee-BM2. Hollow white arrow: Same region indicated by the solid white arrow but when stained for GM130. Inset: cell membrane projections observed in cell transfected with 147-BM2.

In both cell types, the BM2 proteins of all specimens were observed to be present in the cytoplasm. The BM2 protein is enriched at a particular region in the cell which co-localises with GM130, a cis-Golgi marker (arrows in Figures 7.4 & 7.5). Cells
showing FLAG staining also show perturbations in the integrity of their cell membranes. BM2 expressing cells seem to have filamentous projections of the cell membrane (insets, Figures 7.4 & 7.5).

![Figure 7.5 Transfections of BM2-pCAGGSin A549 cells. Cells were fixed, permeabilised and stained 10 hours post transfection. Upper panels: Cells were stained with anti-FLAG. Lower panels: cells were co-stained with anti-GM130. Solid white arrow: strong perinuclear enhancement of the BM2 protein of Lee-BM2. Hollow white arrow: Same region indicated by the solid white arrow but when stained for GM130. Inset: cell membrane projections observed in cell transfected with 117-BM2.](image)

### 7.2.2 Chemical cross-linking of the Lee-BM2 protein.

To analyse the oligomeric form of the BM2 protein, HEK 293T cells were transfected with Lee-BM2 protein for 20 hours. Cells were treated with increasing concentrations of DSP and lysates were analysed by western blotting (Figure 7.6).
Cross-linking analysis shows the presence of 5 bands, the sizes of which are listed in figure 7.6C. Band 1 of Figure 7.6 migrates to a similar size of band 1 of figure 7.3, suggesting that this is the same form of the BM2 protein. Bands 2 (~35kDa) and 4 (~60kDa) may correspond to a dimer, trimer and tetramer of the BM2 protein. The closeness of band 3 to band 2 may suggest that band 3 is a modified form of the BM2-dimer, or could possibly be a BM2-trimer. While it is unlikely, band 5 (~80kDa) may possibly correspond to a 5 or 6 subunit oligomer of BM2 (Figure 7.6).

Figure 7.6 Western blot analysis of BM2 protein treated with cross-linker, DSP. A. HEK 293T cells expressing Lee-BM2 for 20 hours were treated with 0mM, 0.1mM, 0.5mM, 1mM, 2mM and 5mM of DSP and boiled in non-reducing Laemmli buffer. Bands were visualized by probing with anti-myc antibody. B. Log MW vs. Rf chart plotted by comparing the Rf and the molecular weights of the molecular weight standards in A. C. Inferences of the sizes of bands f, g and h observed in (A) through the use of the equation of the slope obtained in (B).
7.3 Discussion

Sequence analysis of the BM2 protein is relatively conserved in all the specimens analysed, especially amongst the clinical specimens. The HXXXW motif, which is the only motif conserved between influenza A and B viruses in all the specimens in this study. Interestingly, the TM domain is highly conserved except at position 5. However, the I5V substitution seen in 132-BM2 and 147-BM2 is a conservative substitution and may not have much phenotypic effect. The three serines present in the TM domain: S_9 S_{12} S_{16} are totally conserved (Figure 7.2). These serines are believed to be responsible for the amantadine resistance seen in influenza B viruses (Pinto and Lamb, 2006; Wang et al., 2009).

The BM2 protein has been previously shown to be a membrane protein that displays perinuclear Golgi as well as ER staining. Figure 7.7, adapted from Paterson et al., displays the strong enrichment of BM2 at a perinuclear location (arrow) as well as the membrane projections (boxed) which are seen in figures 7.4 and 7.5 (Imai et al., 2008; Paterson et al., 2003) (Figure 7.7).

![Image of immunofluorescence of BM2 in HeLa cells](image)

**Figure 7.7 Immunofluorescence of BM2 in HeLa cells.** M2gBM2: BM2 fused with M2 tag at the N-terminus. BM2-Flag: BM2 fused with a Flag epitope tag at the C-terminus. Surface: Staining done without permeabilisation. Perm: staining done post permeabilisation with 0.1% saponin. (Image adapted from Paterson et al.).
In our study, we have shown that the strong BM2 perinuclear enrichment coincides with GM130, a cis-Golgi marker. The Golgi apparatus is an important site in protein modification, where processes such as glycosylation as well as phosphorylation are initiated (Alberts B, 1994). No known glycosylation sites of the BM2 protein has been identified, especially since the ectodomain (MVEPLQI) of the protein does not have any N- or O-linked glycosylation sites (The UniProt Consortium, 2012). This therefore suggests that the accumulation of the BM2 at the cis-Golgi may be sites of phosphorylation. Further analysis involving phosphatase treatment or mutation of the BM2 may reveal more regarding its phosphorylation status, which until recently has become dubious (Hutchinson et al., 2012).

Cells which have been transfected with the BM2 display filamentous projections on their cell membrane (Inset, figures 7.4 & 7.5. Boxed in figure 7.7). This finding was observed in cells transfected with all 6 of the BM2-pCAGGS constructs generated in this study. The M2 protein of both influenza A and B has been shown in several studies to have an indispensable role in viral budding (Chen et al., 2008a; Imai et al., 2008). The influenza A M2 protein has been shown to bind cholesterol in lipid-rafts as well as caveolin-1 in the host cell membranes, sites of viral assembly and budding (Rossman et al., 2010b; Schroeder et al., 2005; Sun et al., 2010; Zou et al., 2009). These associations of the M2 protein with the membrane promote membrane scission which promotes viral budding (Rossman et al., 2010a; Rossman and Lamb, 2011). The data presented in this study seems to suggest that the BM2 is capable of altering the host cell membrane in the absence of other viral proteins. Further investigation on the BM2 protein needs to be carried out to determine its ability to associate with specific membrane components of the host cell.

No main differences were observed in the cellular localisations of the different specimens cloned in this study suggesting that the amino acid substitutions listed in figure 7.2 do not have an effect in BM2 localisation (Figures 7.4 & 7.5).

Size estimation of the different BM2 constructs reveal that the main species of each construct migrates to 14kDa (Band 2, figure 7.3). Two more species are seen in all the different constructs, 22kDa and 32kDa (Bands 3 and 4, Figure 7.3). Band 1 (12kDa) is observed only in specimens 136-BM2 and 147-BM2. This band could be correspond to the unmodified BM2 protein, which has been estimated to be 12.5kDa by The UniProt Consortium (The UniProt Consortium, 2012). Bands 2 and 3 could correspond
to the phosphorylated form of the BM2 protein. Interestingly, specimens 136-BM2 and 147-BM2 are not identical. 136-BM2 and 117-BM2 share 100% protein identity but 117-BM2 does not display band 1. The lack of band 1 in constructs Lee-BM2, 117-BM2, 132-BM2 and 70-BM2 might suggest that the protein is phosphorylated cotranslationally. Further time course analysis may shed light on the phosphorylation of these proteins.

Initial studies have suggested that the BM2 protein is phosphorylated to yield a protein band at 17kDa as early as 2 hours post infection. The native protein in this study was found to migrate to 12kDa. In their study, the 12kDa band is more intense than the 17kDa band, suggesting that there exist more of the unphosphorylated form of the protein in infection (Odagiri et al., 1999). The studies by Odagiri et al., were done in MDCK cells. While MDCK cells are generally accepted as a model host cell system for influenza studies, it is extremely important to note that influenza B does not infect non-human hosts. While input from MDCK infection is invaluable and should be considered, it has to be compared against data generated in human cell lines (Murphy, 1990). In our study, HEK 293T cells, of human origin, were transfected with the individual pCAGGS constructs. The pattern of post translational modification may differ between cell types (Brunton et al., 1994), resulting in the difference in migration pattern observed in our study and Odagiri’s. Labeling with radio-isotope $^{32}$P would allow us to determine which of the species observed in figure 7.3 are phosphorylated (Tolkovsky and Wyttenbach, 2009).

Protein phosphorylation prediction software estimates that S91 of BM2 from B/Brisbane/60/2008 is phosphorylated (Blom et al., 1999; Hutchinson et al., 2012). The residue is conserved in all of the specimens analysed in this chapter, lending stronger credence for this residue to be phosphorylated (Figure 7.2). Mutation of the codon encoding this residue may help us determine its role in BM2 phosphorylation.

Band 4 observed in figure 7.3, ~32kDa, may correspond to the dimeric form of the BM2 protein. This observation that the BM2-dimer can be detected on a reducing SDS-PAGE gel, albeit in a less intense amount, indicates that this dimer is not stabilised by disulphide bonds (Figure 7.3). This postulation is backed by the fact that the BM2 protein does not have any cysteines in its ectodomain (Figure 7.2) (Mould et al., 2003b). This is unlike the influenza A M2 proton channel which has been determined to be a tetrameric complex formed by 2 non-covalently linked disulphide-bonded M2-
dimers (Holsinger and Lamb, 1991; Holsinger et al., 1994; Sugrue and Hay, 1991). The data from the cross linking experiment carried out in this study agrees with a previous cross linking experiment, but with dithiobis[sulfosuccinimidylpropionate] (DTSSP) instead of DSP (this study) (Figure 7.6), where a dimer, trimer and tetramer of BM2 were detected with increasing concentrations of DTSSP (Balannik et al., 2008).

Unlike the previous study involving BM2 oligomerisation, we have detected a faint band corresponding to 79kDa at DSP concentration of 0.5mM and 1mM (Band 5, figure 7.6). This band may correspond to a BM2 complex made up of 5 or 6 subunits, or may be an artifact during the isolation procedure. Proper analysis, involving immunoprecipitation of DSP treated BM2 may shed light whether this band truly reflects an oligomer of BM2 or is an experimental aberration (Figure 7.6).
Chapter 8: Conclusions and Future Direction

B/Lee/40 was used in this study as a reference strain in optimizing the protocols for virus cultivation and sequencing of viral genes. MDCK and A549 cells infected with B/Lee/40 showed the characteristic nuclear staining pattern of the influenza NP in early infection, which was then detected in the cytoplasm in late infection. This same inoculum was able to yield infectious particles when inoculated in embryonated chicken eggs. Interestingly, the harvested allantoic fluid showed 2 opalescent bands at 30% and 40% when spun on a sucrose gradient. Both bands showed potential viral proteins on SDS-PAGE analysis, with the band at 40% yielding more intense bands. When harvested and analysed for electron microscopy, the 40% band showed particles which resembled spherical influenza particles.

B/Lee/40 infection on MDCK cells also showed prominent CPE throughout a 4 day infection. Despite so, the harvested supernatant could cause infection when inoculated onto MDCK cells. The reference strain could not produce plaques in a conventional plaque assay, which may suggest a deficiency in viral transmission from cell to cell in culture. To circumvent this, we employed the use of an immunofluorescent microplate assay. This assay involved the infection of MDCK cells in non-virus propagating conditions followed by immunofluorescence staining of the NP protein. Titres attained through this method were found to be similar to titres attained by the end-point assay of the same inoculum. This assay has an added advantage of rapid data acquisition. We propose that this assay be a suitable substitute for the conventional plaque assay, especially for non-tissue culture adapted viral strains such as primary isolates.

A total of 81 clinical specimens were received for this study. These specimens were tested for influenza B by real time PCR and/or diagnostic IFA. 22 of these specimens were previously tested for influenza B through diagnostic PCR. The remaining 59 specimens were inoculated onto MDCK cells and 48 showed significant CPE and only 44 of these were positive when tested on diagnostic IFA. 3 specimens which showed the strongest IFA signals were passaged through embryonated eggs. The harvested allantoic fluid did not show any significant IFA signal upon MDCK infection. A shell vial assay was carried using these 3 egg-passaged viruses and only 2 of these viruses showed positive staining.
The inability of these clinical specimens to be passaged through tissue culture is reminiscent of B/Lee/40, suggesting the need for a proper tissue culture model for influenza B virus propagation. Since influenza B viruses are not known to infect non-human hosts, it would be succinct to propose that human cell lines be used as a cell culture model to propagate influenza B clinical specimens, instead of MDCK cells. Studying which cell lines allow influenza B propagation may shed light on the mechanism involved in the strict host cell tropism exhibited by influenza B.

Using the vRNA extracted from these clinical specimens the HA, NA and NS1 gene segments were sequenced. Table 8.1 lists all the specimens which yielded full sequence information of any of the three genes. 33 specimens were able to provide sequence information for all 3 genes (labeled with an asterisk *, Table 8.1).

Phylogenetic analysis of HA and NA genes reveal that most of the clinical specimens are re-assortants of the HA-Victoria and NA-Yamagata lineages. The only clinical specimen not having this reassortment order was DSO_010147_2007 which had both gene segments yielding from the Yamagata lineage. The circulation of both reassortant viruses is proof enough that a single strain of influenza B in the bi-annual vaccine is enough to confer immunity to the general population. The data from this study supports the latest injunction to include both lineages of influenza B in the bi-annual vaccine.

While, certain conclusions can be made from the data analysed in this study, it is important to note that increased surveillance would allow greater understanding of evolutionary dynamics of influenza viruses. For example, specimen DSO_010147_2007 was the only specimen amongst the 41 isolated between 2004 and 2009 with both its HA and NA from the Yamagata lineage, this proportion then increased to 3 out of 11 Singaporean specimens isolated between 2010 and 2011. Having a detailed understanding of the viruses currently circulating would not only allow us to predict the next dominant strain, but gives us the ability to choose the best strains to be included in the vaccine.
Table 8.1 List of specimens and their sequenced genes. Total of 46 clinical specimens included in this study and their genes which have been sequenced. * Specimens which gave sequence information for all three genes

<table>
<thead>
<tr>
<th>Specimen Name</th>
<th>Gene Segments Sequenced</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSO_090114_2004*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_090117_2004</td>
<td></td>
</tr>
<tr>
<td>DSO_090124_2004</td>
<td></td>
</tr>
<tr>
<td>DSO_090131_2004*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_090133_2004*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_090134_2004*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_090136_2004*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_090138_2004</td>
<td></td>
</tr>
<tr>
<td>DSO_090143_2005*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_090151_2005</td>
<td></td>
</tr>
<tr>
<td>DSO_050526_2005*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_050528_2005*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_050540_2005*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_050541_2005</td>
<td></td>
</tr>
<tr>
<td>DSO_050544_2005*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_050549_2005*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_050566_2005*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_050578_2005*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_050599_2005</td>
<td></td>
</tr>
<tr>
<td>DSO_050600_2005*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_050619_2005</td>
<td></td>
</tr>
<tr>
<td>DSO_050622_2005</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_050625_2005*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_050629_2005</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_050633_2005</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_040105_2006*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_040117_2006*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_040122_2006*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_040131_2006*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_040133_2006*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_040136_2006</td>
<td>NA NS</td>
</tr>
<tr>
<td>DSO_040174_2006*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_040203_2006</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_050091_2006</td>
<td></td>
</tr>
<tr>
<td>DSO_050199_2006*</td>
<td>HA NA NS</td>
</tr>
<tr>
<td>DSO_050230_2006*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_070214_2006*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_020113_2007</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_020114_2007*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_020132_2007*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_020147_2007*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_010147_2007*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_0003_2009*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_0005_2009*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_0070_2009*</td>
<td>HA NA</td>
</tr>
<tr>
<td>DSO_0100_2009*</td>
<td>HA NA</td>
</tr>
</tbody>
</table>

While increased surveillance would provide the mentioned advantages, it is also imperative that the full genome of each of the specimens be analysed. This would give us the advantage of knowing which of the lineages each of the gene segments belong to.
In doing so, the reassortment pattern of circulating strains would be known, lending a greater understanding of the mechanisms involved in influenza B reassortment. Comparing the reassortment pattern of locally, regional as well as globally circulating strains may also allow the ability to predict the predominant circulating strain in upcoming influenza B pandemics. With knowledge of the sequence of the full genomes of the locally circulating strains, genetic and molecular characterisation of all the gene segments and their translated proteins can be studied on top of those analysed in this body of work.

Most of the amino acid substitutions observed in the HA protein were located in the HA1 domain, suggesting that this domain is under a stronger selective pressure than the rest of the protein. The ‘tip’ of the HA which is the most antigenic site amongst Victoria-like specimens is totally conserved in all the specimens sequenced in this study. Interestingly, DSO_010147_2007 had displayed V161A and R164K in the Yamagata ‘loop’. While the V161A has not been reported previously, specimens displaying K164 has been shown to have the same reactivity as R164 carrying specimens.

We found that all of the clinical specimens sequenced in this study contained a potential glycosylation site, except for specimen DSO_050629_2005. Representative Singaporean sequences isolated in 2010 and 2011 all had this glycosylation site. This glycosylation site was not prevalent worldwide prior to 2000. Further functional analysis has to be carried out to determine if this glycosylation site is truly glycosylated and escapes antibody recognition.

Amino acid substitutions were observed uniformly throughout the NA protein. All of the substitutions seen in the NA calcium binding site were conservative in nature, suggesting conservation of phenotype.

None of the listed substitutions, in both the HA and the NA have been linked to any known drug resistance in influenza B. Similar to the observations in the HA and NA protein, the pattern of amino acid substitutions seen in NB protein are specific to the year of isolation. The NB protein is generally conserved throughout all the different specimens. The HA, NA and NB proteins of DSO_010147_2007 were the most unique of all the other specimens sequenced in this study.

The NB protein of the specimens sequenced in this study displayed 5 conservative amino acid substitutions within the transmembrane domain, lending
evidence that this domain needs to retain its hydrophobicity, for the NB to carry out its undiscovered function.

The two ancestral strains of influenza B do not represent the different clades observed in the NS gene phylogenetic tree. In fact, all the clinical specimens sequenced in this study clustered away from the ancestral strains. Interestingly 7 of the clinical specimens isolated in 2004 clustered with B/Lee/40, unlike the other strains included in the phylogenetic tree. Uniquely, other influenza B specimens isolated in 2004 included in the NS phylogenetic tree did not cluster in cluster I. More sequence information regarding the NS gene of influenza B viruses circulating in 2004 should be ascertained to determine if this finding is accurate or an experimental aberration. If found to be true, it is interesting to note that the NS gene segment related to B/Lee/40 is presently in circulation.


Substitutions in the NS1 NLS were conservative in nature, but further functional analysis needs to be carried out to determine if H49N and R53K do have an effect on nuclear localisation.

We have discovered variability in 3 residues of the NS1 protein which are involved in ISG15 binding. Several confirmatory experiments can be carried out to determine if these substitutions have an effect on ISG15 binding or in the function of NS1 in general. Firstly, Yeast 2-Hybrid and pull down assays may be carried out to determine if these substitutions affect NS1-ISG15 binding. The kinetics of NS1-ISG15 binding may also be assayed by Fluorescence (Förster) Resonance Energy Transfer (FRET). Since the influenza B NS1 protein inhibits the ability of ubiquitin to conjugate with ISG15, an assay can be set up to determine if the different sequences have an effect on this process.

The NS1 and NEP amino acid sequences of these 7 specimens also mirrored the protein sequence of B/Lee/40. The NS1 and NEP protein profiles could generally be grouped according to the year of specimen isolation.

The protein sequences of these specimens were grouped according to their similarity and it was found that protein sequences of specimens isolated from the same year were highly similar. The genes coding for the NB, BM2 and NS1 proteins from
representative strains from each year of isolation were cloned into a mammalian expression vector for further molecular analysis.

Both the NB and BM2 proteins show similar cellular localisations when expressed in A549 and MDCK cells. These proteins show cytoplasmic staining with prominent cis-Golgi accumulation upon immunofluorescent staining. Cells which are expressing the NB or BM2 proteins show filamentous projections on their respective cell membranes. This similarity in localisation may shed light on the function of the NB, which still remains elusive. No differences in localisation patterns were observed between the different representative strains.

When analysed on SDS-PAGE, the NB protein shows five differently migrating species and a smear between 16kDa-37kDa. The main species observed migrates to ~15kDa, but this band is absent in specimens 132-NB and 147-NB. These 2 specimens express a protein of 18kDa which was not present in the other specimens. No obvious correlation with the sequence of these specimens could explain for this difference. It is possible that the 15kDa and the 18kDa represent NB proteins which are differentially modified.

Time course analysis of the NB protein show that the main 3 species were observed as early as 8 hours post transfection while the smear was generated between 8-20 hours post transfection, suggesting that glycosylation occurs co-translationally, and that further modifications of these glycosylation, possibly addition of polylactosaminoglycan, was done post-translationally.

To determine the nature of the NB glycosylation, the NB proteins can be labeled with a pulse of tritium (³H) followed by different periods of incubations (chase). This pulse-chase can allow us to determine which of the bands observed products of early or late glycosylation processes are.

The singly glycosylated form of the NB protein (15kDa) is resistant to PNGase F treatment, while NB modified with polylactosaminoglycan is resistant to EndoH treatment. The singly glycosylated form of NB may be modified such that it gains resistance to PNGase F treatment, a notion which needs to be studied in greater detail.

Cross-linking studies of the NB protein show the possibility of a dimer, a trimer and a tetramer form of the protein. This may lend proof that the NB forms oligomers of itself by disulphide bridging as suggested by the odd number of cysteine residues in the ectodomain which are conserved in all the representative strains analysed in this study.
Site directed mutagenesis substituting these cysteines may allow us to confirm this suggestion.

The BM2 protein shows 4 differently migrating species. The main species which is seen in all the strains migrates to 14kDa. The main species of BM2 correlates with the size of the phosphorylated form of BM2. The unmodified form of the BM2 (~12kDa) was only observed in specimens 136-BM2 and 147-BM2. No correlation between the presence of this band and the amino acid sequence of 136-BM2 and 147-BM2 was noted, suggesting that the appearance of this band could be an experimental aberration. A time course analysis of the BM2 expression may be required to determine the nature BM2 modification.

While the influenza A of the M2 proton channel has been shown to be a homotetramer formed by disulphide-linked dimers stabilised by non-covalent interactions, no cysteines are present in the BM2 protein for a similar mechanism of oligomerisation. Cross-linking analysis of BM2 shows a possible tetramer and possibly maybe even a pentameric form. The tetrameric form of the BM2 has already been published previously but further analysis needs to be carried out to determine the mechanism of oligomerisation.

B/Lee/40 infection on MDCK and A549 cells reveal that the NS1 protein is present in the nucleus in early infection and can be detected in the cytoplasm in late infection. In early infection the NS1 protein can be found in nuclear speckles which colocalise with SC35, a host cell splicing factor. This colocalisation does not persist throughout the infection. In later infection, certain regions of NS1 exclusions can be seen in the nucleus correlating to the location of NPM, a nucleolus marker.

To study this phenomenon better, the NS1 gene of representative strains were cloned into the same mammalian expression vector and transfected into MDCK and A549 cells. The NS1 protein of Lee-NS1 and 136-NS1 did not localise in speckles unlike specimens 132-NS1, 147-NS1 and 70-NS1. Cells expressing the NS1 protein appear to have an altered arrangement of their SC35 protein in the nucleus. While Lee-NS1 and 136-NS1 did not localise to speckles, the staining pattern of SC35 mirrored that of NS1 staining.

Expression of individual domains of the NS1 protein did not show any speckle localisation, suggesting that the localisation with SC35 might require the globular
structure of the NS1 protein. Mutation analysis of the NS1 did not yield any definitive conclusions about which residues are involved in this colocalisation.

To determine the nature of this localisation, confocal microscopy may determine any spatial resemblance. Yeast 2-Hybrid and pull down assays may also allow us to determine if NS1 is directly interacting with SC35 and NPM.

Unlike the rest of the specimens, specimen 117-NS1 localised to large ‘blobs’ in the nucleus. These ‘blobs’ localised with both SC35 and NPM. This colocalisation with NPM is reminiscent of influenza A NS1 proteins which contain a nucleolus localisation signal (NoLS). However, unlike influenza A NS1, 117-NS1 does not have an extra seven amino acids on the C-terminus. This may suggest that the NoLS of influenza B NS1 is not located at the C-terminus.

The lysates of B/Lee/40 infected cells were analysed on SDS-PAGE and 2 differently migrating protein species were noted. The full length NS1 migrated to 35kDa and a smaller band of 23kDa (p23) was observed. This band was also observed in HEK 293T cells transfected with Lee-NS1. Comparisons with the different domains of the NS1 place p23 to be a peptide containing the effector domain and part of the linker.

Analyses of the different representative strains show that specimens 132-NS1 and 147-NS1 did not express p23. p23 expression was found to be present only in specimens expressing D92. Constructs 132-NS1 and 147-NS1 harbored N92. Mutation of D92N eliminates expression of p23 in Lee-NS1 and N92D induces p23 expression in 132-NS1. The presence or absence of p23 expression did not correlate with the speckled expression of NS1 in the nucleus of transfected cells. Instead, cellular fractionation shows that the p23 (and p15) was found in the cytoplasm instead of the nucleus.

The p23 band of 117-NS1 is more intense than the band corresponding to the full length of NS1. 117-NS1 harbored a serine at position 94, while the other specimens had a proline. When Lee-NS1 harbored P94, the same phenotype was seen in 117-NS1. The reverse was true when 117-NS1 was mutated to harbor S94. Mutational analyses of the stretch of amino acids between 91 and 94 reveal that p23 expression is observed in sequences containing MDPS, MDPP or ADPS. p23 expression was silenced in specimens containing the sequences MNPS or MDAS. Expression of p23 is unaffected when M91A was observed, indicating that it is not a product of internal initiation within the NS1 ORF.
We tested if the expression of p23 was possibly due to a host cell protease activity. HEK 293T cells transfected with Lee-NS1 protein were treated with anti-caspase and anti-MMP drugs. Both treatments did not abolish p23 expression. Optimisation of the protocols may be required to reach the conclusion that NS1 is not cleaved by caspases or MMPs. *In vitro* translation may allow us to determine if the expression of p23 is a product of host cell factors, while recombinant expression of NS1 in bacteria may allow us to consider or rule out the processes which are shared between bacterial and eukaryotic cells.

While SDS-PAGE analysis did allow us to estimate that p23 is a C-terminal fragment of the NS1 protein containing the effector domain and part of the linker, its identity is still unclear. Mass spectroscopy of the peptide band of p23 would allow us to determine the exact amino acid sequence of this peptide.

The main impetus in compiling this body of work is the fact the most of the study on virology is centered upon the characterisation of lab-adapted viruses. These viruses have often been passaged through cells or animal models a large number of times and have lost many of the characteristics that made them virulent in humans to start with. So while they represent a model in virological studies, certain native and subtle features may go unnoticed.

Working with non-lab adapted primary isolates come with a large array of complications. Harvesting and isolating these viruses is not a truly efficient process. The chances of these samples of containing fungal and bacterial contaminants are also significant. Several rounds of optimisation had to be carried out to be able to sequence the genes of these viruses and not all of the specimens could yield sequence data for all three of the genes in focus. The sequence information extracted by these specimens provided the basis of the work compiled in this thesis.

Since these viruses could not propagate in cell culture, further molecular characterisation required the cloning of representative genes with an epitope tag for detection. The variation in sequences between these strains was exploited to study the differences in the properties of selected influenza B proteins. Two unique findings were discovered with regards to the influenza B NS1 proteins of these representative strains. Immunofluorescence analyses reveal that 3 of the influenza B NS1 proteins display speckled localisation within the nucleus while 2 did not. Interestingly, 1 specimen showed strong nucleolar localisation, suggesting the presence of a NoLS.
harboring D$_{92}$ instead of N$_{92}$ also express p23, which comprises the effector domain and part of the linker. To the best of our knowledge, this is the first report of influenza B NS1 having a NoLS as well as expressing p23. Further elucidation may divulge more on the characteristics of these discoveries, revealing more into the mystery that is influenza B, the oft-neglected sibling of influenza A.
Protein Alignments

Supplementary Figure 1  HA Protein Alignment  Pages 198-205
Supplementary Figure 2  NA Protein Alignment  Pages 206-211
Supplementary Figure 3  NB Protein Alignment  Pages 212-213
Supplementary Figure 4  NS1 Protein Alignment  Pages 214-217
Supplementary Figure 5  NEP Protein Alignment  Pages 218-219

Supplementary Figures 1-6. Protein alignments of HA, NA, NB, NS1 and NEP from specimens analysed in this study. Sequence identity is compared to that of B/Lee/40. Similar residues are labeled by the dot (.). The dash (-) at the ends and beginnings of the individual protein represents undetermined sequence while when in the middle represents a deletion or a lack of an insertion. X: unknown amino acid, The Asterisk (*) represents a stop codon. Sequence numbering is based on the full alignment and not on B/Lee.
References


Böttcher, E., Greiner, C., Steinmetzer, T., Klenk, H.-D., Garten, W., 2009. MDCK cells that express proteases TMPRSS2 and HAT provide a cell system to propagate influenza viruses in the absence of trypsin and to study cleavage of HA and its inhibition. Vaccine 27, 6324-6329.


Fu, X.D., Maniatis, T., 1992a. The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3' splice site. Proc Natl Acad Sci U S A 89, 1725-1729.


human ISG15 by the NS1 protein of influenza B virus. Proceedings of the National Academy of Sciences.


Imai, S.W., 2 Ai Ninomiya, 1 Masatsugu Obuchi, 1, Odagiri1, a.T., 2004. Influenza B Virus BM2 Protein Is a Crucial Component for Incorporation of Viral Ribonucleoprotein Complex into Virions during Virus Assembly. JOURNAL OF VIROLOGY, 78, 11007-11015.


Mibayashi, M., Martinez-Sobrido, L., Loo, Y.-M., Cardenas, W.B., Gale, M., Jr., Garcia-Sastre, A., 2007. Inhibition of Retinoic Acid-Inducible Gene I-Mediated Induction of Beta Interferon by the NS1 Protein of Influenza A Virus. J. Virol. 81, 514-524.


Wright, P.F., Bryant, J.D., Karzon, D.T., 1980. Comparison of Influenza B/Hong Kong Virus Infections among Infants, Children, and Young Adults. The Journal of Infectious Diseases 141, 430-435.