A study of virulence mechanisms by highly pathogenic avian influenza H5N1 viruses in the avian host

Division of Infectious Diseases
Faculty of Medicine

Jason Stephen Long

Supervised by Professor Wendy Barclay
(Imperial College London)
Dr Jill Banks and Dr Wendy Howard
(AHVLA)

Doctor of Philosophy
2014
Author’s Declaration

I confirm that all work presented here is my own and that the use of all material from other sources has been properly and fully acknowledged.

Jason Stephen Long

Copyright declaration

‘The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work’
Acknowledgements

First of all I would like to thank my supervisors Professor Wendy Barclay, Dr Jill Banks and Dr Wendy Howard, not only for this opportunity but for the experience over the last four years. Wendy B has been a continual source of support, insight and understanding, I will be forever grateful. Adapting to different working environments has its challenges and I am extremely grateful to Jill for helping the transition and to Wendy H for being a constant source of help and happiness, making the long trek to the AHVLA worthwhile.

I am exceptionally thankful for the friendly advice offered from members of the Barclay and AHVLA influenza labs, past and present. I will always treasure the friendships made in my time spent here and am lucky to have been part of such an amazing group of people.

My dear friends, be them Scientists or ‘Muggles’, I apologise for my recent neglect, but your understanding has been a comfort that has helped to no end, thank you.

To my family, your capacity to love and wish me happiness astounds me, I’m so fortunate to have you and I hope I have done you proud.

Of course this study would not have been possible without the funding provided by the Seedcorn funding mechanism by the Department for Environment, Food and Rural affairs (Defra), and I am grateful to Her Majesty’s Government for this initiative.

The journey I have undertaken both professionally and personally has been a surprise at every corner. Most importantly I met my partner, Alex. Here words are not enough; you know what your support has meant.
“There is only one good, knowledge, and one evil, ignorance”

-Socrates-

Greek Athenian philosopher, 469-399 BC
(As quoted in Diogenes Laërtius’s Lives of Eminent Philosophers)
Abstract

The study of highly pathogenic avian influenza (HPAI) seeks to understand the pathogenic nature of these viruses. Belonging to the H5 or H7 subtype, HPAI causes significant economic loss in poultry and in recent years has also led to die offs in wild birds. In addition, the zoonotic nature of H5N1 has caused human infection associated with high mortality, although these viruses have not yet evolved to cause a pandemic.

Mechanisms of virulence by HPAI have previously been attributed to a multi-basic HA cleavage site (MBCS) which facilitates systemic dissemination of the virus. Yet more recently, other factors have been described as necessary for a highly pathogenic phenotype. This study utilises reverse genetics to investigate the importance of three mutations in the HA of A/turkey/England/50-92/91 (50-92) H5N1 virus which are crucial for virulence. A novel mutation in the fusion peptide appeared to alter the pH of fusion by the HA. The importance of HA fusion for pathogenesis and host adaptation is discussed.

Eurasian lineage H5N1 highly pathogenic avian influenza viruses (HPAIV), were first detected in Qinghai Lake, China in 2005, and subsequently spread through Asia, Europe and Africa. Importantly, these viruses carried a lysine at position 627 of the PB2 protein (PB2 627K), a known mammalian adaptation motif. Previous avian influenza isolates have carried glutamic acid in this position (PB2 627E), commonly described to restrict virus polymerase function in the mammalian host. We sought to examine the effect of PB2 627K on viral maintenance in the avian reservoir. Reverse genetics viruses were engineered to contain converse PB2 627K/E mutations in a Eurasian H5N1 virus (A/turkey/Turkey/5/2005; Ty/05) and, for comparison, a historical pre-Asian HPAI H5N1 virus that naturally bears PB2 627E (A/turkey/England/50-92/1991; 50-92). Our observations suggest PB2 627K is supported in Eurasian lineage viruses; in contrast PB2 627K carries a significant fitness cost in the historical pre-Asian 50-92 virus. These data have important implications for understanding the restriction of HPAI for humans and for the maintenance of a key genetic motif associated with host range in the avian reservoir.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declaration</td>
<td>1-</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>2-</td>
</tr>
<tr>
<td>Abstract</td>
<td>4-</td>
</tr>
<tr>
<td>Contents</td>
<td>5-</td>
</tr>
<tr>
<td>Chapter 1 Introduction</td>
<td>14-</td>
</tr>
<tr>
<td>1.1 Virion structure and genome organisation</td>
<td>15-</td>
</tr>
<tr>
<td>1.2 Virus life cycle</td>
<td>17-</td>
</tr>
<tr>
<td>1.2.1 Attachment and entry</td>
<td>18-</td>
</tr>
<tr>
<td>1.2.2 Replication</td>
<td>18-</td>
</tr>
<tr>
<td>1.2.3 Evading the host innate immune response</td>
<td>21-</td>
</tr>
<tr>
<td>1.2.4 Packaging and release</td>
<td>21-</td>
</tr>
<tr>
<td>1.3 A focus on HA fusion</td>
<td>23-</td>
</tr>
<tr>
<td>1.4 Influenza virus reverse genetics</td>
<td>26-</td>
</tr>
<tr>
<td>1.5 Ecology and Evolution</td>
<td>28-</td>
</tr>
<tr>
<td>1.6 History of highly pathogenic avian influenza</td>
<td>31-</td>
</tr>
<tr>
<td>1.6.1 Epidemiology and emergence of H5N1</td>
<td>33-</td>
</tr>
<tr>
<td>1.6.2 Pathobiology and disease</td>
<td>35-</td>
</tr>
<tr>
<td>1.6.3 Mechanisms of pathogenesis/ increased virulence by HPAI</td>
<td>36-</td>
</tr>
<tr>
<td>1.6.4 A case study of avirulent and virulent clones from HPAI virus A/turkey/England/50-92/1991</td>
<td>45-</td>
</tr>
<tr>
<td>1.7 Host adaptation by influenza polymerase</td>
<td>48-</td>
</tr>
<tr>
<td>1.7.1 Role of PB2 in host range adaptation</td>
<td>51-</td>
</tr>
<tr>
<td>1.8 Thesis aims</td>
<td>58-</td>
</tr>
<tr>
<td>Chapter 2 Mechanisms of increased virulence by influenza H5N1 HA protein</td>
<td>60-</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>60-</td>
</tr>
<tr>
<td>2.1.1 Role of glycosylation 158 for influenza virus</td>
<td>61-</td>
</tr>
<tr>
<td>2.1.2 Role of residue 193 in receptor binding of influenza HA</td>
<td>62-</td>
</tr>
<tr>
<td>2.1.3 Mutations in and beyond the fusion peptide that alter the ability of HA to fuse with the host membrane</td>
<td>62-</td>
</tr>
<tr>
<td>2.2 Results</td>
<td>66-</td>
</tr>
<tr>
<td>2.2.1 Location of residues 160, 193 and HA2 4 on the H5 haemagglutinin</td>
<td>67-</td>
</tr>
<tr>
<td>2.2.2 Restoration of the highly pathogenic phenotype of 50-92 virus</td>
<td></td>
</tr>
</tbody>
</table>
2.2.3 HA mutations A160T, E193K and HA2 G4R restore the highly pathogenic phenotype of 50-92 virus

2.2.4 50-92-HP replicates more efficiently than 50-92-LP in avian cells

2.2.5 HA mutation HA2 G4R is required for restoration of the HP phenotype in vivo

2.2.6 Investigation of virus tropism by in ovo influenza infection

2.2.7 Investigation of growth kinetics of the 50-92 HA mutants in vitro

2.2.8 Mutation HA2 G4R alters the pH of fusion of 50-92 HA protein

2.2.9 Bioinformatic analysis of HA 158 glycosylation, 193 and HA2 4 for avian influenza isolates

2.3 Conclusions and Discussion

Chapter 3 The emergence of PB2 627K in the avian influenza virus reservoir

3.1 Introduction

3.1.1 Molecular basis of PB2 627E restriction in mammalian cells

3.1.2 The biological consequence of the 627K mutation

3.1.3 The implication of the 627K mutation for the avian host

3.2 Results

3.2.1 Prevalence of the PB2 mutation 627K in the avian population

3.2.2 Generation of the PB2 627K mutants

3.2.3 spontaneous selection of additional mutations upon passage of pre-Asian H5N1 50-92 PB2 627K

3.2.4 H5N1 50-92/gKR virus is attenuated in vivo

3.2.5 PB2 627K does not carry a fitness cost for Eurasian lineage H5N1 Ty/05

3.2.6 Growth kinetics in avian cells

3.2.7 Differences in polymerase activity

3.3 Conclusions and Discussion

Chapter 4 Investigating the underlying mechanism of the PB2 627K associated fitness cost in non-Asian HPAI H5N1 virus

4.1 Introduction

4.2 Results

4.2.1 Investigation of avian influenza virus sequence in the PB2 627 region

4.2.2 PB2 628 has a prominent location on the 627 linker region

4.2.3 PB2 628R does not provide a fitness advantage for 50-92 reassortant virus in vitro

4.2.4 PB2 627K confers a fitness gain for Ty/05 reassortant virus
4.2.5 Altering the genetic constellation of 50-92 prevented reversion of the 627K genotype

4.3 Conclusions and Discussion

Chapter 5 Discussion

Chapter 6 Materials and Methods

6.1 Materials

6.2 Methods

6.2.1 Molecular Biology

6.2.2 Cell lines and transfection

6.2.3 Infectious studies

6.2.4 Bioinformatic analysis

Chapter 7 Appendices

Chapter 8 References
<table>
<thead>
<tr>
<th>Acronym/Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHVLA</td>
<td>Animal Health and Veterinary Laboratories Agency</td>
</tr>
<tr>
<td>AI</td>
<td>Avian influenza</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>cRNA</td>
<td>Complementary RNA</td>
</tr>
<tr>
<td>cRNP</td>
<td>Complementary Ribonucleoprotein</td>
</tr>
<tr>
<td>FCS</td>
<td>Feotal Calf Serum</td>
</tr>
<tr>
<td>GFP</td>
<td>Green Fluorescent Protein</td>
</tr>
<tr>
<td>GsGd/96</td>
<td>A/goose/Guangdong/1/1996</td>
</tr>
<tr>
<td>H&amp;E</td>
<td>Hematoxylin &amp; Eosin</td>
</tr>
<tr>
<td>HA</td>
<td>Haemagglutinin</td>
</tr>
<tr>
<td>HK/97-like</td>
<td>HongKong97-like</td>
</tr>
<tr>
<td>HP</td>
<td>Highly pathogenic</td>
</tr>
<tr>
<td>HPAI</td>
<td>Highly Pathogenic Avian Influenza</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>IHC</td>
<td>Immunohistochemistry</td>
</tr>
<tr>
<td>IVPI</td>
<td>Intravenous Pathogenicity Index</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>LP</td>
<td>Low pathogenic</td>
</tr>
<tr>
<td>LPAI</td>
<td>Low Pathogenic Avian Influenza</td>
</tr>
<tr>
<td>lsNA</td>
<td>Long stalk NA</td>
</tr>
<tr>
<td>M1</td>
<td>Matrix protein 1</td>
</tr>
<tr>
<td>M2</td>
<td>Matrix protein 2</td>
</tr>
<tr>
<td>MBCS</td>
<td>Multi-basic cleavage site</td>
</tr>
<tr>
<td>MDT</td>
<td>Mean Death Time</td>
</tr>
<tr>
<td>MOI</td>
<td>Multiplicity of infection</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>Mx</td>
<td>Myxovirus-resistance protein</td>
</tr>
<tr>
<td>NA</td>
<td>Neuraminidase</td>
</tr>
<tr>
<td>NDV</td>
<td>Newcastle Disease Virus</td>
</tr>
<tr>
<td>NEP</td>
<td>Nuclear Export Protein</td>
</tr>
<tr>
<td>NS</td>
<td>Non-structural</td>
</tr>
<tr>
<td>nt</td>
<td>Nucleotide</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>OIE</td>
<td>Organisation for Animal Health</td>
</tr>
<tr>
<td>PA</td>
<td>Polymerase acidic protein</td>
</tr>
<tr>
<td>PB1</td>
<td>Polymerase basic 1</td>
</tr>
<tr>
<td>PB2</td>
<td>Polymerase basic 2</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>pfu</td>
<td>plaque forming unit</td>
</tr>
<tr>
<td>pH1N1</td>
<td>Pandemic H1N1 (2009)</td>
</tr>
<tr>
<td>POL</td>
<td>Polymerase</td>
</tr>
<tr>
<td>PR8</td>
<td>A/Puerto Rico/8/34 (H1N1)</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative reverse transcriptase Polymerase chain reaction</td>
</tr>
<tr>
<td>RBC</td>
<td>Red blood cell</td>
</tr>
<tr>
<td>RBS</td>
<td>Receptor Binding Site</td>
</tr>
<tr>
<td>REU</td>
<td>Relative Egg infectious dose&lt;sub&gt;50&lt;/sub&gt; (EID&lt;sub&gt;50&lt;/sub&gt;) Unit</td>
</tr>
<tr>
<td>RG</td>
<td>Reverse Genetics</td>
</tr>
<tr>
<td>RLU</td>
<td>Relative Light Units</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>RNP</td>
<td>Ribonucleoprotein</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>SA</td>
<td>Sialic acid</td>
</tr>
<tr>
<td>siRNA</td>
<td>Short interfering RNA</td>
</tr>
<tr>
<td>ssNA</td>
<td>Short stalk NA</td>
</tr>
<tr>
<td>svRNA</td>
<td>Small viral RNA</td>
</tr>
<tr>
<td>Ty/05</td>
<td>A/turkey/Turkey/05/2005</td>
</tr>
<tr>
<td>UTR</td>
<td>Un-translated region</td>
</tr>
<tr>
<td>VIC</td>
<td>A/Victoria/3/1975 (H3N2)</td>
</tr>
<tr>
<td>vRNA</td>
<td>Viral RNA</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>WT</td>
<td>Wild Type</td>
</tr>
<tr>
<td>X-gal</td>
<td>5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside</td>
</tr>
</tbody>
</table>
### Table of figures

**Chapter 1.**

**Figure 1.** Virus structure and genome organisation of influenza A virus - 16-

**Figure 2.** The cycle of transcription and replication - 20-

**Figure 3.** Crystal structures and phylogenetic organisation of influenza HAs - 24-

**Figure 4.** Model of membrane fusion by influenza HA - 25-

**Figure 5.** Influenza Reverse Genetics: 12 plasmid system - 27-

**Figure 6.** Ecology of influenza A viruses - 31-

**Figure 7.** Timeline of major events of Asian H5N1 evolution - 35-

**Figure 8.** Three-dimensional structure of influenza RNP and polymerase - 49-

**Figure 9.** Mutations mapped to increase polymerase activity in mammalian cells - 50-

**Figure 10.** Effect of the K627E and 590S/591R mutation on the electrostatic surface of the 627-domain - 54-

**Figure 11.** Polymerase reconstitution assay - 55-

**Chapter 2.**

**Figure 12.** Conservation of the influenza fusion peptide across 17 different antigenic subtypes - 63-

**Figure 13.** Fusion peptide sequences of the fusion glycoproteins from influenza viruses, paramyxoviruses, retroviruses and filoviruses - 63-

**Figure 14.** HA structure of H5N1 Vietnam/04 annotated with the amino acids of interest - 66-

**Figure 15.** 5092-HP replicates more efficiently than 5092-LP in Chicken Embryonic Fibroblasts - 69-

**Figure 16.** HA2 mutation G4R is necessary for restoration of the HP phenotype in vivo - 71-

**Figure 17.** H&E longitudinal section of a non-infected 9-day old chicken embryo - 72-

**Figure 18.** Dissemination of influenza virus 50-92 in the chicken embryo detected by NP staining - 74-

**Figure 19 A&B.** Exploration of virus tropism in the chicken embryo - 76-

**Figure 20.** Replication of 50-92 HA mutants in CEF cells - 79-

**Figure 21.** Plaque phenotype of 50-92 viruses in MDCK cells - 80-

**Figure 22.** Fusion peptide mutation HA2 4R compensates for the decreased fusogenicity by 160T and 193K, measured by virus haemolysis - 81-

**Figure 23.** Fusion peptide mutation HA2 4R compensates for the decreased fusogenicity by 160T and 193K, measured by polykaryon formation - 83-

**Figure 24.** HA mutations 160T, 193K and HA2 4R did not alter cell expression - 84-

**Figure 25.** Quantitative analysis of a predicted N-linked Glycosylation at HA residue 158, identity at 193 and HA2 4 for H5N1 influenza viruses isolated from 1956 to 2013 - 86-
Chapter 3.

**Figure 26.** Phylogenetic tree of birds

**Figure 27.** Phylogenetic tree demonstrating the prevalence of the 627K mutation in the H5 clades

**Figure 28.** Mutation of recombinant non-Asian H5N1 50-92 PB2 627K upon virus passage

**Figure 29.** Virulence in chickens demonstrated by IVPI

**Figure 30.** PB2 627K carries a fitness cost for non-Asian H5N1 50-92

**Figure 31.** Reversion of PB2 K627E and R628Q in mutant non-Asian lineage H5N1 50-92/gKR upon replication in chickens

**Figure 32.** Quantification of Ty/05 virus replication in chickens after intravenous inoculation

**Figure 33.** PB2 627K is not a fitness cost for Eurasian lineage H5N1 Ty/05

**Figure 34.** H5N1 Ty/05/K and Ty/05/E virus antigen distribution of tissues in experimentally infected ducks

**Figure 35.** Growth kinetics in avian cells

**Figure 36.** Differences in polymerase activity

Chapter 4.

**Figure 37.** Summary of the Ty/05 and 50-92 virus genotypes

**Figure 38.** Location and structure of the 627-Domain and amino acid 628

**Figure 39.** Generation of reassortant viruses by reverse genetics

**Figure 40.** PB2 628R confers a fitness cost for 50-92 reassortant virus in avian cells

**Figure 41.** PB2 627K confers a fitness gain for Ty/05 reassortant virus

**Figure 42.** Altering the genetic constellation of 50-92 prevented reversion of the 627K genotype during egg passage

**Figure 43.** Altering the genetic constellation of 50-92 prevented reversion of the 627K genotype in MDCK cells

**Figure 44.** A hypothetical model explaining the 627K phenotype observed for 50-92 virus

Chapter 5.

**Figure 45.** Duck cells demonstrate higher levels of DNA fragmentation following 50-92 infection

**Figure 46.** Virus growth kinetics of 50-92 and Ty/05 NS reassortants in CEF and DEF cells

**Figure 47.** The world production of poultry livestock between 1990 and 2010

Chapter 6.

**Figure 48.** An example of Flow cytometry data detecting 50-92 HA expressed on the cell surface of transfected HeLa cells
List of tables

Chapter 1.
Table 1. Influenza A virus encoded proteins and their function, adapted from Knipe et al. (2007)

Table 2. A summary of amino acid changes in the HA gene from plaque clones compared to the 50-92 isolate sequence together with an IVPI score of pathogenicity and Mean Death Time (MDT) in chicken eggs.

Table 3. A summary of the IVPI scores from work carried out previously, and HA genotype for the original 50-92 isolate; plaque picked clones and RG generated virus with cognate HA and NA or mis-matched H7 and N1 from HPAI A/Chicken/Italy/99 RG system.

Chapter 2.
Table 4. 50-92 viruses were generated by reverse genetics and pathogenicity determined by the IVPI test in ten 6-week old chickens.

Table 5. IVPI score of 50-92 HA mutant viruses.

Table 6. Quantitative analysis of HA2 sequence 4-7 from avian influenza viruses isolates from 1956 to 2013.

Chapter 3.
Table 7. Quantitative analysis of the PB2 mutation 627K in avian isolates from 1956 to 2013.

Table 8. Prevalence of the PB2 627K mutation in HPAI H5N1 virus isolates.

Table 9. HPAI H5N1 model viruses used for this study.

Table 10. Intravenous Pathogenicity Index of reverse genetics generated viruses.

Chapter 4.
Table 11. Quantitative analysis of PB2 627 codon sequence, amino acid PB2 628 and nt 1869 from avian influenza viruses isolated from 1956 to 2013.

Table 12. Summary of data from Chapters 4 and 5 comparing 627E and 627K phenotypes for Ty/05 and 50-92 virus.

Table 13. Describes the nucleotide identity of the gene segments between 50-92 and Ty/05. Included are the amino acid differences of the polymerase complex and NP, examined for their investigation in the literature.

Chapter 5.
Table 14. The mutations required for H5N1 respiratory droplet transmission in the ferret model, compared to the genotype of clade 2.2 virus, Ty/05.

Chapter 6.
Table 15. Cell lines used in this study.

Table 16. A list of animal species used in this study, together with age and experimental use.
Table 17. List of oligonucleotide primers designed in this study

Table 18. Reverse genetics systems of PR8 (A/Puerto Rico/8/1934), Ty/05 (A/Turkey/Turkey/1/2005) and 50-92 (A/Turkey/England/50-92/91). Helper plasmids expressing polymerase and NP of A/Victoria/3/75. Ty/05, 50-92 and E195 polymerase and NP expressing plasmids

Table 19. Reporter constructs used in this study

Table 20. Antibodies used in this study

Table 21. List of buffers and culture media used in this study

Table 22. Poultry record sheet for clinical observations

Chapter 7.

Table 23. Clinical signs observed in ducks infected with Ty/05/K or Ty/05/E HSN1 virus

Table 24. Quantification of viral RNA isolated from the tissues of ducks infected with either Ty/05/K or Ty/05/E by qRT-PCR

Table 25. Immunohistochemical distribution of influenza A nucleoprotein from 12-week old ducks 2 days post infection with either Ty/05/K or Ty/05/E

Table 26. List of RG viruses generated in the 50-92, Ty/05 and PR8 genetic backgrounds
Chapter 1. Introduction

The earliest estimate dates the origin of avian influenza (AI) to around 200,000 years ago, calculated by molecular dating techniques (1). Certainly the Orthomyxovirus family established itself a long time ago, forming five genera over the years: Influenza virus A, B and C, Thogotovirus and Infectious Salmon Anaemia virus. Influenza A viruses have evolved a diverse array of subtypes based on the structural proteins; Haemagglutinin (HA) and Neuraminidase (NA). To date, there are seventeen known HA subtypes (H1-17) and ten NA subtypes (N1-10). Only in avian species have subtypes combining H1-16 and N1-9 been identified, making influenza A primarily a virus of avian origin.

Retrospectively there is evidence of disease resembling avian and human influenza dating back to the classical literatures of ancient civilizations. Reports such as: “liquid sputa, sneezing, coughs, ardent fever” (Hippocrates 400 BC), “This year happened that great destruction among the fowls” (The Anglo-Saxon Chronicle 671 AD), “desolation and despair among the barnyard pets” and “swelled heads of the thanksgiving gobblers” (New York Herald, 1872 AD (2)), are dotted throughout history. The first possible evidence of influenza pandemics may be considered as 927, 1173 and 1386 AD. More convincingly there is evidence of 3 pandemics in the 16th century, after which we can reliably describe a further 11 pandemics, making a total of 14 between 1510 and 2009 AD (3, 4).

The first definitive record of avian influenza was by Perroncito, 1878 who described a disease with high flock mortality, named ‘Epizoozia tifoide nei gallinacei’ (fowl typhoid), affecting chickens in Italy (5). Such reports of severe disease likely refer to highly pathogenic avian influenza (HPAI) a term previously known as ‘fowl plague’. HPAI was first recognised as the ‘filterable’ infective agent to cause ‘fowl plague’ in 1901 by Centanni and Savonuzzi, and later characterised as influenza virus by Schäfer 1955 (6).

No matter their origins or nomenclature, influenza viruses are pathogens of serious concern. Capable of inducing severe disease in both humans and numerous animal species, the social and economic impact of influenza has been and continues to be substantial. Despite several licensed antiviral drugs and successful vaccination programs, the virus
continues to evade our efforts of disease control by evolving resistance or novel antigenicity. Clearly, research of this fascinating virus is vital, allowing us to prepare against and combat influenza by furthering our knowledge.

1.1 Virion structure and genome organisation

Influenza A virus particles are pleomorphic, that is to say they may be spherical (-100nm) or filamentous (up to 20µm), depending on the virus strain, cell origin, passage history and Matrix protein 1 (M1) (7). A lipid membrane, derived from the host cell, encases the virion and harbours three transmembrane proteins: Haemagglutinin (HA), Neuraminidase (NA) and Matrix protein 2 (M2) (Figure 1A). HA and NA are glycoproteins, present as a homotrimer and homotetramer, respectively, involved in cell attachment and release. M2 acts as an ion channel facilitating genome release upon virus entry to the cell. Surrounded by this envelope is a layer of M1 protein, providing a structural platform for the virion, which in turn protects the virus genome.

The virus genome is present in the virion as 8 negative-sense RNA segments (Figure 1B). Each segment exists as a ribonucleoprotein (RNP) complex, consisting of the gene segment together with nucleoprotein (NP) and the polymerase complex (Polymerase Basic 2 [PB2], Polymerase Basic 1 [PB1] and Polymerase Acidic [PA] protein). Each gene segment is flanked by untranslated regions (UTR) at the 5’ and 3’ ends. Influenza utilises a number of genetic tricks to encode multiple protein products from a gene segment. So far, the PB1, PA, M and NS (non-structural) gene segments have been found to encode more than one protein product (Figure 1C). The PB1 gene encodes proteins PB1-F2 and PB1-40 by an alternative open reading frame (8, 9). Similarly the PA gene also contains alternative initiation codons to generate two N-terminal truncated forms of PA, PA-N155 and PA-N182, together with PA-X which carries a C-terminus in a different reading frame as a result of a ribosomal frameshift (10, 11). As well as coding for M1 and NS1 (Non-structural protein 1) the M and NS gene segments use alternative splicing as a mechanism to generate M2 and M42 and NEP/NS2 (Nuclear Export Protein/Non-structural protein 2) respectively (12, 13).

However, it is not clear whether all influenza A viruses encode the recently discovered protein products. For example, human isolates often lack the PB1-F2 protein, and therefore the products of each genome must be considered strain by strain.
Figure 1. Virus structure and genome organisation of influenza A virus

A. A cartoon representation of an influenza A virion, displaying the HA, NA and M2 proteins in the lipid membrane, surrounding the M1 matrix and RNP s. Adapted from CDC free resources (14). B. Electron micrograph of budding influenza virions, with a magnified image of a single virion (circled) with two transections (1 and 2), showing the distribution of the 8 gene segments. Adapted from Gavazzi et al. (2013) (15). C. The genome organisation of the 8 gene segments, displaying the encoded proteins in colour and the UTRs in grey.
1.2 Virus life cycle

The influenza life cycle has been well documented and although many techniques influenza utilises to manipulate the host cell have been identified (summarised Table 1), mysteries remain such as the function of newly discovered proteins PA-N182 and PA-N155 (10, 11).

Table 1. Influenza A virus encoded proteins and their function, adapted from Knipe et al. (2007) (16)

<table>
<thead>
<tr>
<th>vRNA segment</th>
<th>Proteins encoded</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Polymerase Basic protein 2 (PB2)</td>
<td>Component of RNA polymerase; cap-binding</td>
</tr>
<tr>
<td></td>
<td>Polymerase Basic protein 1 (PB1)</td>
<td>RNA-dependent RNA polymerase</td>
</tr>
<tr>
<td></td>
<td>PB1-F2</td>
<td>Pro-apoptotic activity; virulence factor</td>
</tr>
<tr>
<td></td>
<td>PB1-N40</td>
<td>Interaction with polymerase complex; role unclear</td>
</tr>
<tr>
<td>2</td>
<td>Polymerase Acidic protein (PA)</td>
<td>Component of RNA polymerase; Endonuclease</td>
</tr>
<tr>
<td></td>
<td>PA-N182</td>
<td>Role unknown</td>
</tr>
<tr>
<td></td>
<td>PA-N155</td>
<td>Role unknown</td>
</tr>
<tr>
<td></td>
<td>PA-X</td>
<td>Modulation of host response</td>
</tr>
<tr>
<td>3</td>
<td>Haemagglutinin (HA)</td>
<td>Surface glycoprotein; binds to cell receptor, sialic acid; membrane fusion</td>
</tr>
<tr>
<td>4</td>
<td>Nucleoprotein (NP)</td>
<td>Binds vRNA; RNA synthesis</td>
</tr>
<tr>
<td>5</td>
<td>Neuraminidase (NA)</td>
<td>Surface glycoprotein; neuraminidase activity cleaves sialic acid from cell surface</td>
</tr>
<tr>
<td>6</td>
<td>Matrix protein 1 (M1)</td>
<td>Virion structure; virus budding</td>
</tr>
<tr>
<td></td>
<td>Matrix protein 2 (M2)</td>
<td>Membrane protein; ion channel; virus assembly and budding</td>
</tr>
<tr>
<td>7</td>
<td>Non-structural protein 1 (NS1)</td>
<td>Multi-functional protein; viral IFN antagonist; inhibit cellular mRNA export; virulence factor</td>
</tr>
<tr>
<td></td>
<td>Non-structural protein 2 / Nuclear export protein (NS2/NEP)</td>
<td>Nuclear export of vRNPs; polymerase co-factor</td>
</tr>
</tbody>
</table>


1.2.1 Attachment and Entry

If the journey of transmission from an infected host to the susceptible tissues of the naive host is to be successful, influenza must then bind to suitable cells to initiate infection. Influenza HA is responsible for this attachment by binding to sialic acids present on oligosaccharides at the cell surface. Human influenza viruses preferentially bind to sialic acid attached to the penultimate galactose sugar by an α2,6 linkage (SAα2,6Gal), whereas avian viruses bind to sialic acid with an α2,3 linkage (SAα2,3Gal) (17, 18). Once attached, influenza enters the cell by endocytosis. Both clathrin-mediated endocytosis and non-clathrin, non-caveolae internalisation mechanisms have been described (19, 20).

Like many other enveloped viruses, influenza requires a low pH to fuse its viral membrane with the cell endosomal membrane. Before fusion occurs, the HA precursor (HA0) must first have been cleaved into two subunits, HA1 and HA2. The low pH of the endosome triggers a conformational change, exposing the fusion peptide and facilitating the fusion event (21). Meanwhile the M2 ion channels present in the virus membrane allow the influx of H+ ions which disrupt the virion structure. Together, the result is the release of the RNP complexes into the cytoplasm (22). Not all enveloped viruses undergo pH-dependent fusion, for example, paramyxoviruses fuse at the cell surface by a pH-independent mechanism by means of a dedicated fusion protein (23).

1.2.2 Replication

The RNPs rapidly transport to the nucleus, by means of nuclear localisation signals on NP protein and its interaction with importin α (24, 25). Once in the nucleus, primary transcription can occur. The polymerase complex (PB1, PB2 and PA) is attached to the UTRs of the vRNA, a prerequisite due to the lack of an RNA-dependent RNA polymerase in the host cell (Figure 2). The UTRs contain the highly conserved vRNA promoter sequence (13 nucleotides of the 5’ end and 12 nucleotides of the 3’ end) present in all influenza viruses. The promoter sequence forms a secondary structure, mediated by the partial complementarity of the 5’ and 3’ termini (26, 27). Several promoter structures have been suggested, with a corkscrew configuration favoured as both a vRNA and cRNA promoter (28–30).
**Viral mRNA synthesis**

The anti-sense vRNA is first transcribed into positive-sense mRNAs. Influenza mRNA synthesis requires a 5’ capped primer, gained by a ‘cap-snatching’ mechanism from the host cell pre-mRNAs. ‘Cap-snatching’ is performed by the binding of PB2 to the 5’ cap of host pre-mRNAs (31–33) which are then cleaved by the endonuclease activity of PA (34, 35), generating 9-17 nucleotide long capped RNA products which are used to prime mRNA transcription from the vRNA template. The PB1 protein contains the conserved motifs required by RNA polymerases to carry out chain elongation. Messenger RNA synthesis is initiated by the addition of a guanosine residue to the stolen primer, which base-pairs with the penultimate cytosine residue at the 3’ end of the vRNA (36).

Throughout the transcription the polymerase remains bound to the 5’ end, this is thought to cause steric hindrance when the transcription process reaches a poly-U stretch near the 5’ end of the vRNA causing stuttering and the generation of a poly(A) tail at the 3’ end of the newly transcribed mRNA (37). The resultant product is now treated as a cellular mRNA since the influenza polymerase complex has successfully added a methylated 5’ cap and a poly(A) tail.

**cRNA synthesis**

The generation of cRNA is distinct from that of mRNA. Unlike mRNA transcription, cRNA synthesis occurs without a capped primer and full-length positive copies of the vRNA template are made. The mechanism distinguishing mRNA synthesis from cRNA synthesis is unclear. The cRNA products are encapsidated by NP protein, and it has been suggested that the presence of NP protein prevents termination at the U tract near the 5’ end of the vRNA template (38).

**vRNA synthesis**

The cRNA anti-genome serves as a template for the synthesis of the vRNA. The initiation of vRNA synthesis by the cRNA promoter differs; transcription initiates a few nucleotides inside the 3’ cRNA promoter and is then realigned to the 3’ end where elongation continues to the end of the cRNA template (39). The generation of 3 RNA species is described in Figure 2.
The switch from transcription to replication is a subject of great deliberation. The current model is that primary transcription occurs in cis by the polymerase bound to the vRNA, and replication in trans by a newly generated soluble polymerase (40). The accumulation of NP protein or of polymerase subunits later in infection stabilises the cRNA template which is otherwise degraded and thus permits more replication (41, 42). More recently, there is evidence that NEP protein is able to regulate transcription and replication, independently of its nuclear export function (43, 44). In addition, newly discovered small viral RNAs (svRNA), derived from the ends of the vRNA template, may play a role in driving replication (45, 46).

Once generated the vRNPs must make their way into the cytoplasm. M1, produced late during infection, associates with the vRNPs in the nucleus. The NEP protein then interacts with the C-terminus of M1 together will the cell nuclear export proteins, directing the vRNP though the nuclear pore and into the cytoplasm (reviewed, 44).

![Diagram of vRNP cycle](image)

**Figure 2. The cycle of transcription and replication**

Transcription and replication of the anti-sense vRNA by the viral polymerase complex (PB1, PB2 and PA). The vRNP is displayed as a helical structure, with the polymerase attached to the 5’ and 3’ extremities. The vRNP image was taken from Manz et al. (2013) (48).
1.2.3 Evading the host innate immune response

Arguably the most industrious influenza protein, NS1 is truly multifunctional and has been well studied (Reviewed, 46). Two domains have been defined in its structure, which exists as a dimer; the so called ‘RNA-binding’ domain at the N-terminus and the ‘Effector’ domain of the C-terminus. Many functions in the virus life cycle have been ascribed to the NS1 protein. Upon infection, the host cell has evolved ways to detect pathogens and induce an anti-viral state. Thus, NS1 primarily acts an antagonist to the IFN response, both pre-transcriptionally by binding pattern recognition receptors, (e.g. RIG-I and PKR), which detect pathogen associated molecular patterns. The dsRNA-dependent protein kinase R (PKR) is bound by NS1 in a dsRNA-independent manner preventing dimerization and therefore IFNβ and apoptosis induction, and inhibition of translation (50, 51). NS1 is able to bind and block the IFNβ inducing activity of retinoic acid-inducible gene 1 (RIG-I) otherwise induced by viral 5’-triphosphorylated ssRNAs (52). It is not clear if NS1 is able to bind RIG-I directly, and some data suggest the presence of 5’-triphosphorylated ssRNAs are able to enhance the interaction, indicating a complex formation. The activity of IFNβ is also blocked post-transcriptionally by preventing cell mRNA processing by binding CPSF30, a key cleavage and polyadenylation specificity factor or PAF1 (53). Despite preventing cell mRNA translation, NS1 is able to promote the translation of viral mRNAs by interacting with the cell translation complex (54). Both pro- and anti-apoptotic properties have been demonstrated by NS1 (55–60); possibly suggesting a temporal control of apoptosis by this protein. Such mechanisms require further understanding, since NS1 is a key marker of pathogenesis in terms of disease outcome in the infected host.

1.2.4 Packaging and release

Replication of the virus must conclude with the exit of progeny virions, a complex process that requires each of the eight gene segments to be selectively incorporated into each virion, surrounded by matrix protein and coated with a lipid membrane derived from the cell membrane.

The successful selection of the 8 gene segments requires RNA/RNA interactions, or so called ‘packaging signals’. Packaging signals have been mapped to the ends of all 8 gene segments, spanning the non-coding region and into the terminal coding regions (~100nts
from the vRNA ends) (Reviewed, 47). The mapping of packaging signals in the coding regions was first discovered by replacing stretches of a NS sequence with a reporter gene and recovering virus by superinfection of an influenza helper virus (62); and later by use of reverse genetics (63, 64). Studies have identified all eight segments in virions, with specific segment/segment interactions, often describing HA vRNP as having a central position (15, 65, 66) (Figure 1B). Fournier et al. (2011) suggest the formation of a supramolecular RNP complex, interconnected by RNA/RNA interactions (67). Analysis of codon conservation has also identified key residues that can drastically affect packaging levels, further supporting a selective segment incorporation model based on vRNA interactions (68). The role of a ‘master segment’ directing the assembly of the genome complex is a generally favoured hypothesis of packaging at present, with the ‘master’ so far suggested to be segment 1 or 7 (PB2 and M) (61, 67, 69–72). In contrast, other RNA segmented viruses have evolved a different mechanism altogether. For example, it appears the segmented genome of dsRNA Bacteriophage ϕ6 is packaged by each segment being selectively recognised by packaging signals on their 5’ends and bound by the procapsid (73).

As the gene segments are gathered, the M1 protein is thought to associate the vRNPs with the cell membrane (74). This is mediated by the HA and NA proteins which are present as a lipid raft on the cell membrane (75). M1 polymerisation is thought to drive the generation of the virion structure, securing the incorporation of the vRNPs, as it binds to the cytoplasmic tails of HA and NA. M2 protein is thought to stabilise this structure and enable membrane scission at the base of the budding virion by altering the curvature of the extruding membranes and thus resulting in a separate virion (Reviewed, 58).

The two glycoproteins present on the virion, HA and NA, have opposing roles. Since HA binds to SA present on the cell membrane, the virion must somehow be released from its originating cell and prevent reattachment. The NA protein fulfils this role as it also recognises SA, but cleaves SA from the cell surface and virions by its enzymatic activity, preventing binding by HA (77). It has been recognised that NA also has a role to enable virion egress, or indeed ingress, through mucus present on respiratory epithelium by cleaving SA present on sialated glycans (78, 79).
After the journey through the micro-environment of the cell is complete, the virus is released into the extracellular environment. The successful transmission of the influenza virus to the next suitable host is dependent on its survival in this environment and also its initial delivery, e.g. the expulsion of a sufficient infective dose from the originating host in the form of respiratory droplets as well as its inherent infectiousness in the recipient host.

1.3 A focus on HA fusion

As previously mentioned, the release of the vRNPs into the cytoplasm from the endosome is made possible by the fusion of the viral and endosomal membranes by HA protein. The X-ray crystal structure of the HA protein was first resolved in 1981 by Wilson et al. (80). Since these investigations, various HA protein structures have been solved, to increasing resolution. The most significant observations from these works are the high level of conservation in the protein structure, despite significant variation of the amino acid sequence between HA subtypes (Reviewed, 101) (Figure 3). Influenza HA proteins can be divided into two phylogenetic groups, group 1 and group 2, shown in Figure 3.

The mechanism of fusion has been well studied, permitted by the cleavage of the HA0 precursor into the HA1 and HA2 segments. Fusion of the viral and cell endosomal membranes is a pH mediated process. As a class I fusion protein, influenza HA has been extensively characterised (81–85).
Figure 3. Crystal structures and phylogenetic organisation of influenza HAs

Trimetric complex of a H1, H2 and H3 human HA proteins, viewed from the side (top) and from above (below). The similarity of the protein structures is clearly visible. The phylogenetic grouping of 16 HA subtypes (bottom). Taken from Gamblin et al. (2010) (81).

A. The HA protein is anchored by the transmembrane domain of the C-terminus in the viral membrane present at neutral pH where the fusion peptide is hidden in the protein structure (Figure 4).

B. As the pH of the endosome lowers, an acid-induced conformational change occurs that leads to the release of the fusion peptide and insertion into the endosome membrane. The conformational change, where the N-terminal HA2 domain transitions into a helical structure, exposes the hydrophobic fusion peptide and positions it antiparallel to the transmembrane domain creating an extended HA2 intermediate. During this conformational...
change, the HA1 de-trimerises and remains attached but removed from the HA2 structure (not shown in Figure 4).

**C.** The intermediate confirmation collapses as the C-terminal and N-terminal domains of the HA2 draw together. The intermediate helical structure of the N-terminus transform to a loop, reducing the distance of the C-terminal domain and fusion peptide to the transmembrane anchor.

**D.** The resulting structure causes the two membranes to be pulled together, forming a hemifusion stalk and allowing lipid mixing.

**E.** Formation of the fusion pore is stabilised by the interaction of the fusion peptide and the transmembrane domain. The contents of the virus particle are released into the cytoplasm.

![Figure 4. Model of membrane fusion by influenza HA](image)

Two HA trimers are presented, the HA2 N-terminal portion (red) and C-terminal portion (blue, with fusion peptide (green). Taken from Floyd et al. (2008), originally from Harrison (2008) (83, 86).

Although this classic mechanism has been well established, the data is based on a limited array of subtypes. More recently it has been recognised that fusion of HA has an important role in the adaptation of influenza to different species. For example, the pH of fusion differs between avian and human viruses, the latter fusing at a lower pH (87–94). This emerging area of research seeks to understand the role that HA fusion plays on host adaptation, transmission and pathogenesis and will be discussed further in Chapters 3 and 6.
1.4 Influenza virus reverse genetics

Many studies that have explored and deciphered the life cycle of influenza virus were performed using influenza reverse genetics (RG). The technique of generating influenza viruses from cDNA (virus rescue) has revolutionised the study of influenza.

The first influenza virus rescue was performed by Enami et al. (1990) whereby a recombinant NA gene cloned and mutated in plasmid DNA was transcribed \textit{in vitro} and by super-infection incorporated into a helper influenza virus lacking an NA gene (95).

Work by Pleschka et al. (1996) developed a technique to generate vRNA segments in human cells from cDNA, i.e. a negative sense RNA not capped and without a poly(A) tail (96). By flanking the NA gene sequence of an influenza virus with the human RNA polymerase I promoter and a ribozyme sequence (from Hepatitis Delta virus), transcription was initiated at the beginning of the vRNA sequence and terminated by the cleavage of the ribozyme. Using a helper virus, Pleschka et al. confirmed successful generation of a virus carrying the intended NA sequence (96).

The same technique was used to generate an influenza virus entirely from cDNA using plasmids for each gene segment. Fodor et al. (1999) successfully generated influenza virus utilizing this method with the ribozyme sequence. Under the control of the poll promoter, vRNAs for each gene segment were produced after which the influenza polymerase complex and NP (expressed in the cell from protein expression plasmids) carried out virus transcription and replication, resulting in progeny virions (97). This 12 plasmid transfection was essentially the same process as that used by Neumann et al. (1999), except their technique used the mouse poll terminator sequence instead of the hepatitis D ribozyme and included plasmids expressing any or all of the virus structural proteins as well as the polymerase in the transfected cells (98). An outline of the methods is described in Figure 5. Influenza reverse genetics have been routinely used by the influenza research community, allowing a revolutionary platform by which to generate specific point mutations in the influenza genome and investigate their properties. The triumph of this technique has extended to influenza B virus in 2002 (99, 100) and influenza C virus in 2007 (101).
Several adaptations to the original methods published have been used, including a bi-directional system with both polI and polII promoters resulting in the expression of influenza proteins as well as each vRNA segment (102, 103), and the generation of an influenza virus from a single plasmid encoding all 8 gene segments together (104). However the 12-plasmid transfection system remains the most frequently used since it allows greater flexibility, for example, in the generation of reassortant viruses from 2 or more different strains simply by mixing different plasmids together.

Figure 5. Influenza Reverse Genetics: 12 plasmid system

The 8 gene segments are transcribed from polI promoter with either a polI terminator or ribozyme (red) (a magnified image is provided on the left in black). Subsequent virus transcription and replication of the newly generated vRNA gene segments is carried out by the influenza polymerase complex and NP expressed by the polII plasmids (blue). Adapted from Neumann et al. (1999) (98).
1.5 Ecology and evolution

The natural reservoir

Wild aquatic birds are described as the primordial source of influenza viruses (105, 106). Since the first isolation of avian influenza from a wild bird in 1961 (A/tern/South Africa/61 (H5N3)) (107), it has been established that AI viruses have a global distribution in wild birds and thus wild birds, in particular waterfowl, are considered the natural reservoir of influenza A viruses (106, 108–110). Up to 16 different HA subtypes and 9 NA subtypes can all be found in the wild bird reservoir (105). Of the wild birds, two orders are most frequently found to be infected with influenza viruses: the Anseriformes (ducks, geese and swans) and the Charadriiformes (gulls, terns and shorebirds). Furthermore, the evolutionary rates of influenza viruses (based on non-synonymous vs. synonymous mutations) from wild birds have been described as slow, suggesting a mutual existence between virus and host (111, 112). Harmony between virus and host likely represents a long-standing relationship through evolution over time. Indeed, it is generally considered that wild birds typically do not present clinical signs when infected by the majority of influenza viruses. The natural site of replication is the intestinal tract of wild birds and high titres of virus can be shed (113). Transmission is by the faecal/oral route (114). However, this dogma may be challenged in the future as experimental procedures advance. It is difficult to measure the natural route of infection, but recently intracloacal inoculation was described as a successful route of infection in experimentally infected ducks (115). It is thought that influenza viruses can persist in certain environments for a significant length of time, and virus has been successfully isolated from lake surface waters (116, 117).

Domestic birds

Domestic birds such as chickens, turkeys (both Galliformes) and ducks have reportedly been infected with subtypes H1-H13, with H1, H4-5, H7 and H9 isolated most frequently (105). The introduction of influenza viruses into the agricultural environment can have huge economic consequences. Galliformes are generally more susceptible to influenza infection suffering a respiratory disease and/or egg drop and may present clinical signs, depending on the virus subtype or strain (118). Adaptations of influenza viruses may arise upon introduction into poultry hosts, characterised by an increase in the evolutionary rate and in some subtypes, e.g. H5N1, an increase in HA glycosylation and a shortening in the
The stalk of NA protein (119–121). Interestingly, serological studies have failed to detect avian influenza in wild gallinaceous birds, suggesting Galliformes are not the natural reservoir of influenza (122). Wild and domesticated ducks and other waterfowl are thought to play an important role in the transmission of avian influenza from the natural reservoir into poultry, especially in outdoor and backyard farm systems (105, 123, 124). Human cases of avian influenza, mainly of the H5 and H7 subtype, have been attributed to exposure to poultry. For example, a recent epidemiological study in China found poultry exposure was linked to 75% of human infections by H7N9 and 71% for H5N1 (125).

**Mechanisms of evolution**

Since influenza has an RNA genome which is transcribed by an RNA-dependent RNA polymerase with no proof reading capabilities, the rate of mutation is fairly high, calculated to be approximately $0.6-3.1 \times 10^{-5}$ (mutations/site) (126–129). The appearance of point mutations, when advantageous, enables positive selection. The antigenic sites of the HA and NA proteins are particularly variable due to the immunological pressures of the host population, an interactive process between host and virus resulting in antigenic drift of these proteins.

Reassortment of the 8 gene segments is another method that enables evolution of influenza viruses. Theoretically, co-infection of a cell with two (or more) influenza viruses could result in the reassortment of gene segments and progeny virions carrying a mixture of genes derived from the two parental strains. If the HA or NA antigens acquired are novel to the human population, this mechanism is known as antigenic shift and can lead to new pandemics. Extensive reassortment in wild bird populations has been reported (117, 130–132). Charadriiformes and other migratory species such as geese may play a key role in the spread and evolution of influenza viruses, since certain species infected with influenza in the absence of symptoms may be able to spread virus over huge distances during migrations. The level of reassortment has also been investigated in swine revealing significant rates of gene swapping (133).

An interesting study recently highlighted the role of heterosubtypic immunity against AI infecting ducks. By acquiring a large number of samples from a single landing site for Mallard ducks in Sweden and analysing for patterns of immunity, the authors find...
reinfection by the same subtype is uncommon as well as demonstrating heterosubtypic immunity by closely related HA subtypes (134). It is suggested that this observation of immune pressure may explain HA subtype diversity in the AI reservoir.

Recombination can also occur between gene segments. The swapping of genetic material in this way has been detected in the NA gene derived from the PB1, PB2 or NP genes increasing the stalk length of the protein (135), as well as between two NP genes (136). Although rare, recombination can have huge consequences for virulence, such as basic amino acid insertions into the cleavage site of HA, giving rise to highly pathogenic viruses e.g. Mexico H7N3 (2012) inserted host 28S ribosomal RNA (rRNA) (137–139).

**Consequence for host adaptation**

Birds are not the only host of influenza viruses and influenza may be described as a zoonotic disease. Species specific lineages are formed by some influenza strains in humans. Currently, the H3N2 and H1N1 (2009 pandemic) influenza A virus lineages circulate in humans along with influenza B and C viruses that likely diverged from the common ancestor millennia ago.

Reassortment is responsible for the generation of pandemic influenza viruses, such as the triple reassortant H1N1 pandemic virus of 2009 derived from circulating swine viruses in North America (140–142). Novel reassortant viruses can overcome species restriction more efficiently than the accumulation of point mutations by utilising gene segments already adapted to the new host. Descendants of pandemic viruses continue to circulate and evolve in the human population causing seasonal epidemics.

Influenza viruses have also been described in a number of other animals including swine, equine, deer, big cats, dogs, ferrets, seals, whales and bats (105) (Figure 6). Similarly to humans, influenza viruses that have crossed the species barrier maintain preferred subtypes in certain hosts. The H1N1, H1N2 and H3N2 subtypes are endemic in swine (143, 144) and the H3N8 subtype is commonly found in horses and dogs (145, 146).

The ecological relationship and consequence of influenza viruses present in unusual hosts such as marine animals and bats with those currently circulating in birds and humans is unclear. Analysis of bat H17N10 virus reveals the HA to be quite distinct to all other
influenza HA proteins (147). Whether these hosts can be described as ‘dead-end’ or represent a new dawn of influenza evolution are interesting postulates.

**Figure 6. Ecology of influenza A viruses**

Wild birds are the natural reservoir of influenza A viruses where the H1-16 and N1-9 subtypes can be found in nearly all combinations. Other hosts select preferential subtypes and may require specific adaptations. More recently a unique lineage, H17N10, was discovered in bats and the interaction here is unclear. Picture taken from Manz et al. (2013) (48).

**1.6 History of highly pathogenic avian influenza**

Some influenza viruses may be described as Highly Pathogenic (HP). The phenotype of low pathogenic (LP) and HP avian influenza (LPAI and HPAI) viruses is officially defined by their effect in chickens. The standard test for pathogenicity recommended by the World Organisation for Animal Health (OIE) is the intravenous pathogenicity index (IVPI), carried out in ten 4 to 8 week old chickens intravenously injected with virus and scored for disease over 10 days (see Chapters 2 and 6) (148).

As previously mentioned, cleavage of the HA is required for infectivity, and the nature of the cleavage site is where LP and HPAI differ. Typically for LP influenza, cleavage is carried out by specific extracellular proteases that recognise the penultimate Arg or Lys
residue (monobasic) of the HA1 sequence. So far, several proteases have been described: human air-way trysin-like protease (HAT), transmembrane protease serine 2 (TMPRSS2) (149), and matripase (150) (human epithelial cell examples); rat bronchial epithelium trypase Clara (151); as well as Factor Xa-like protease in embryonated eggs (152, 153). Influenza with a monobasic cleavage site is restricted to sites of replication that express these proteases i.e. in the respiratory or intestinal tract of birds. Poultry infected with LPAI typically exhibit mild clinical signs and low instances of mortality, if any, although increases in mortality can be seen in some instances such as co-infection with other pathogens, e.g. H9N2 and Ornithobacterium rhinotracheale (154).

HP HA proteins carry a multi-basic cleavage site (MBCS) and may be cleaved by ubiquitous intracellular furin-like serine proteases such as furin and PC6 (155–160). The previous restriction to the respiratory or intestinal tract no longer applies and a systemic, infection is permitted, increasing the virulence of the virus and, depending on the host, increasing pathogenesis (161). These viruses cause high levels of morbidity and mortality in chickens and other poultry (5, 162).

The acquisition of a MBCS appears to occur by a number of mechanisms. Since the RNA sequence preceding the encoded cleavage site is rich in purines, polymerase stuttering in this region is postulated to add additional purines which are translated as basic amino acids (Arg and Lys). Also reported are recombination events with either the NP or M genes which introduces additional amino acids before the cleavage site (Reviewed, 5, 80). Occasionally, HPAI viruses may increase pathogenicity by particularly large insertions in the cleavage site derived from host 28S rRNA (163).

Since the use of phylogenetic analysis, it has generally been agreed that HP viruses evolve from a LP progenitor virus. Although, attempts to observe this in a HPAI outbreak or recapitulate it in an experimental setting have proved difficult. The HP H7N1 outbreak in Italy (1999-2000) was preceded by the circulation of LP H7N1 viruses (164). So far only two LP avian influenza subtypes, H5 and H7, have been found to naturally acquire a highly pathogenic phenotype by the generation of a MBCS in poultry (165–169). Attempts have been made to introduce a MBCS into other subtypes (170–173). Interestingly, not all
Chapter 1 | Introduction

subtypes or strains have supported the pathogenic phenotype, suggesting that the HP phenotype may depend on other virus characteristics (see section 1.6.3).

1.6.1 Epidemiology and emergence of H5N1

Historically, HPAI H5N1 viruses have caused sporadic outbreaks of severe disease restricted to gallinaceous poultry with the exception of one wild bird isolate in 1961, A/tern/South Africa/61 (5, 107, 174). The pathogenic nature of these viruses in poultry is attributed to the presence of the MBCS site in the HA protein (175–177). Between 1959 and 1995 there were an estimated 15 HPAI outbreaks recorded in poultry worldwide (of both H5 and H7 subtypes). During this time the historic sporadic outbreaks in poultry caused no known human infections and were self-limiting, stamped out by culling or controlled by H5 and H7 vaccination of poultry (5). Since the emergence of the H5N1 HPAI in Asia at the end of the 20th century, the epidemiology of HPAI altered substantially such that it reached panzootic proportions (118). The foundation of this significant epidemiological change can be traced back to the 1997 H5N1 poultry outbreaks in Hong Kong, which coincided with the first human cases (178, 179). To this day, resultant Asian lineage H5N1 HPAI viruses have caused ongoing outbreaks of severe disease in poultry, humans and other hosts, and have undergone continual evolution and reassortment leading to the emergence of the genotype Z H5N1 viruses (180). Subsequently, an unprecedented outbreak of H5N1 HPAIV in wild aquatic birds occurred in 2005 at Qinghai Lake, China (181–183).

The progenitors of the Hong Kong97-like viruses (HK/97-like) can be traced to a H5N1 HPAI outbreak that occurred in commercial geese in Guandong Province, China, late in 1996 (A/goose/Guangdong/1/1996 [GsGd/96]) (184, 185). Several important details should be recollected from the 1996/7 outbreaks and their subsequent evolution: i) the outbreak was unprecedented with millions of birds dead or culled in the following months, ii) the viruses replicated in geese and ducks as well as poultry, iii) human infections and mortalities by HPAI were reported for the first time in Hong Kong, iv) the newly emerged Asian H5N1 lineage viruses have produced more frequent outbreaks than before with unprecedented geographical spread.
A timeline of the evolution of Asian H5N1 is provided in Figure 7. The ancestral GsGD/96 viruses gave rise to the Asian H5N1 lineage through reassortment events with LPAI. The resultant HK/97-like viruses, caused devastation in poultry and domestic waterfowl, as well as human infections in 1997 (A/HongKong/156/97) (184, 186, 187). After this initial epizootic, these viruses became endemic in geese and domestic ducks in China (188, 189). During this time, a remarkable array of diverse genotypes evolved with outbreaks repeatedly detected in poultry (190–192). The genetic diversity of H5N1 viruses is so complex that a nomenclature system was developed specifically for this lineage (193). In 2002 the first wild bird outbreaks were confirmed in Hong Kong, described as genotype Z and having increased virulence in wild birds (190, 194–196). The genotype Z viruses continued to circulate and expanded to other South East Asian countries (180, 191, 197). The detection of this virus may have begun in 2002, but the virus may well have established itself earlier, remaining undetected. It is interesting that the reports of die offs in wild birds coincide with the spread of the viruses to other countries and continents.

In 2005 an outbreak in wild birds of previously unrecognised proportions occurred in Qinghai lake, China (181, 182). The Qinghai-lineage viruses, of HA clade 2.2, were the result of a reassortment event between genotype V (Clade 1) and genotype Z (Clade 2.2) H5N1 viruses (181, 198). This outbreak is thought to have seeded the geographical expansion of H5N1 outside of Asia for the first time. The clade 2.2 viruses expanded across Eurasia the Middle East and Africa, likely by migratory flyways or transport of poultry, causing extensive outbreaks in wild birds and poultry and sporadic outbreaks in humans (199, 200). Interestingly clade 2.2 viruses maintained endemicity in Egypt with a significant number of human infections (200–202). A most striking feature of the clade 2.2 H5N1 viruses is that they carry certain human adaptation markers, such as PB2 627K, a key interest of this thesis, addressed in Chapters 3 and 4.

More recently, the advent of the clade 2.3.2.1 viruses in wild birds (191, 200, 203) caused concern because of its wide geographical spread, including incursion into Europe in 2010, which transmitted to European poultry (200, 204).
Figure 7. Timeline of major events of Asian H5N1 evolution

The major ecological events of the Asian H5N1 virus are displayed in chronological order, separated by geographical distribution. Blue lines represent aquatic poultry hosts, green lines terrestrial poultry hosts, and red lines wild bird hosts. Solid lines depict stable interactions between virus and host and dashed lines signify a transient interaction. Adapted from Sonnberg et al. (2013) (205).

Since 2003 there have been 360 human fatalities as a result of 610 infections of H5N1 avian influenza (206). This high mortality rate of 59% has gained worldwide attention, prompting a significant amount of research into H5N1 HPAI.

1.6.2 Pathobiology and disease

In comparison the death toll for the poultry industry has been tremendous; since 2003 an estimated 400 million domestic poultry have died by disease or culling (207). The number of wild birds that have died as a result of H5N1 infection is not known, but it is undoubtedly a great number.

The pathobiology of H5N1 HPAI infection has been well described. Severe damage of endothelial cells and parenchymal organs is typical, although pathobiology may vary and include: swelling of the head, hemorrhages, cyanosis, edema, congestion and organ necrosis and inflammation. Interestingly, infections in waterfowl (i.e. ducks and geese) by Asian H5N1 demonstrate more neurologically associated pathobiology revealed by clinical signs such as ruffled feathers (inability to groom), torticollis (twisted neck) and loss of balance. The overall mortality for waterfowl such as domestic ducks is generally lower than for chickens which are particularly susceptible, although in recent years the virulence of certain lineages of Asian H5N1 has increased in waterfowl and wild birds. Previous to Avian H5N1
HPAI in 1996, HPAI typically did not cause illness or death in ducks and geese. (Reviewed, 164).

1.6.3 Mechanisms of pathogenesis/ increased virulence by HPAI

Pathogenicity or virulence?

The two terms are often used interchangeably. However, pathogenicity refers to the ability of a pathogen to cause disease as a result of the relationship between the virus and host. Virulence refers to the intensity of pathology caused as a result of the virus genetic makeup, and often correlates with the level of replication, a factor which it is also often used to describe.

Early on in the investigations of HPAI, Alexander et al. (1986) showed differences in pathogenesis by H5 viruses in different avian hosts with various degrees of virulence as well as a range of abilities to transmit by contact to naive birds (209). In particular, ducks were resistant to infection by this assortment of HPAI H5 viruses. One virus, A/chicken/Scotland/59 (H5N1) was pathogenic for chickens but not turkeys, whereas A/turkey/Ontario/7732/66 (H5N9) was pathogenic in turkeys but not chickens.

Regarding host factors, wild birds are generally more resistant to HPAI than are poultry. Barber et al. (2010) suggested that RIG-I accounted for decreased AI virulence in ducks, compared to chickens which are highly susceptible and lack RIG-I. However the universality of this hypothesis is unclear since chicken cells have an intracellular sensor capable of detecting flu, MDA5, which initiates the same signalling cascades as RIG-I in avian cells (210–214). Indeed, IFN was first discovered using chicken eggs infected with influenza by Isaacs and Lindemann (1957), so clearly there are chicken sensors capable of mounting an IFN response (215). Even with a competent RIG-I, different species of ducks demonstrate varying levels of susceptibility to disease by H5N1 (216). Similarly, unravelling of the chicken genome has revealed genetic differences between chicken populations which may translate as increased or decreased resistance to AI. Myxovirus-resistance (Mx) proteins induced by type-I IFN are capable of limiting influenza infection. Differences in the Mx1 gene of chickens was thought to be responsible for variation in the susceptibility of HPAI (217). However, these findings were in disagreement with another study which found no difference between the Mx1 genotype and AI susceptibility (218). Hunt et al. (2010) found
that the chickens major histocompatibility complex (MHC) haplotype determined the level of resistance to HPAI, although this study recognised other background genetic differences between the different chicken breeds were also responsible for influencing resistance (219). Questions remain over avian species specific resistance to HP avian influenza. Age is also a factor with young birds suffering a great deal more, possibly due simply to their size and the risk of having fewer cells that are overwhelmed by virus infection (220, 221). However a study by Pantin-Jackwood et al (2012) suggested the difference could be due to the innate immune response. Comparison of infected 5-week old and 2-week old ducks revealed that 5-week old ducks had a higher temperature and increased expression of several innate immune-related genes, compared to 2-week old ducks that were more susceptible to the disease (220). Microarray analysis may also highlight differences in the host response, with the development of chicken microarrays and chicken gene probes, use of this technology is increasing. Microarray analysis of chicken lungs infected with Al has already shed light on chicken genes that may alter survivability, namely: CD274, RNF19B, OASL, ZC3HAV1, PLA2G6, GCH1, and USP1 (222). However functional analysis is still required to understand if these candidates actually play a role in the host response to AI infection. With the recent increase of virulence by Asian H5N1 in both poultry, waterfowl and wild birds, these questions are more pertinent than ever (208).

Dose and route of inoculation can also impact pathogenesis. Decreasing dose of HPAI in chickens has been shown to delay the onset of clinical signs and reduce virus shedding, although not affecting overall mortality (223–225). The route of inoculation has been reported to dictate pathogenesis of AI for both HP and LPAI, and this can also depend on the species of bird being inoculated. Intravenous inoculation often increases mortality and shedding over intranasal inoculation of HPAI in chickens (209, 226, 227). The intranasal route increased shedding of a HPAI H5 in chickens and ducks compared to inoculation by the alimentary route, either directly or by consumption of infected chicken meat (227). In contrast, a more recent study showed infection of mallards with LPAI was exacerbated by intra-inguinal (alimentary) inoculation compared to intranasal, intratracheal, intraocular or intracloacal inoculation (115).

In terms of viral factors, the multi-basic cleavage site is the fundamental to the pathogenesis of HPAI in birds. Systemic replication is a clear requisite of HPAI, yet the
mechanism of virus dissemination is unclear. Clearly HPAI is able to travel through the blood and infect different organs, as demonstrated by the routine IVPI assay, yet how influenza makes this journey via the natural route of infection is unclear. A further dilemma is how the virus crosses the blood brain barrier, one theory could be entry via infected migrating immune cells (228).

1.6.3.1 HPAI virulence beyond the multi-basic cleavage site

**HA**

Despite being an overriding trait, viral factors other than a MBCS may contribute to pathogenesis. A recent study used reassortant viruses from two HP H5N1 viruses to distinguish the genes responsible for differing pathogenicities in mallard ducks. Exchange of HA alone in single gene reassortant viruses reversed the phenotypes of this viruses which differ only by 4 amino acids in the HA protein, all located outside the MBCS. The mechanism behind the difference was not elucidated (229). In a similar study that compared a pathogenic H5 virus from 2005 to a HK/97-like virus, and generated single gene reassortants, the HA of the modern H5 was shown to contribute significantly to mortality in ducks (230). It would be interesting to understand whether the increase virulence in these circumstances is due to altered receptor binding or by an alternative HA function such as fusion.

The pH of fusion by HA has been found to correlate with H5N1 pathogenicity in chickens and ducks (93). This will be discussed further in Chapter 2.

**NA**

The stalk length of NA for HPAI viruses has repeatedly been described to adapt influenza virus to poultry (120, 121, 231, 232). Hoffman et al. (2012) compared the effect of a short-stalk and long-stalk NA of HP H7 virus in chickens and ducks. A short-stalk NA was detrimental for the virus in ducks, whereas a long-stalk reduced replication and mortality in chickens (233). Whether stalk deletions directly increase pathogenicity in chickens or if it is a matter of contributing to successful replication of the virus is unclear.

Glycosylation of the NA head has been reported to increase pathogenicity in chickens (234), the authors suggest this may be via the increased activation of host
protease, such as NA cleavage of SA obscuring the cleavage site as observed by others (235), yet no direct evidence of this was provided.

**PB1, PB2, PA and NP**

The polymerase genes of influenza A viruses have been shown to contribute to virulence in both mammalian and avian hosts. Several studies have suggested that increased pathogenicity is associated with an increase in the replicative ability of the virus.

To understand the basis of increased pathogenicity of recent H5N1 viruses, a comparison between a HPAI H5N1 (A/chicken/Yamaguchi/7/2004) and a less virulent HPAI H5N1 (A/duck/Yokohama/aq10/2003) was performed in chickens. This study demonstrated that increased pathogenicity was associated with rapid replication of the virus. The more pathogenic virus caused death in chickens before fever was induced, whereas the less pathogenic induced fever and time to death was longer. Cytokines were measured by RT-PCR and a positive correlation between increased early replication and induction of the innate immune response was shown (236).

*In vitro* analysis of polymerase genes has also suggested increased polymerase activity correlated with increased pathogenicity for two closely related H5N1 duck viruses, A/duck/Fujian/01/2002 (HPAI) and A/duck/Guangxi/53/2002 (LPAI), although no further analysis was performed to confirm this *in vivo* (237). By exchanging polymerase components between the two viruses, it was suggested that differences in PA and PB2 were responsible for the increased polymerase activity of the highly pathogenic virus.

Several studies have implicated PA as a virulence factor for HPAI (229, 230, 238). Song et al. (2010) found two amino acid changes in PA increased virulence in ducks, S224P and N383D, the latter shown to increase polymerase activity *in vitro* (238). Recently, Hu et al. (2013) described PA residues 101G and 237E as increasing virulence in ducks (229). Polymerase activity was also increased in duck cells, which corresponded with an increase in PA nuclear accumulation (likely by aa101G present in the NLS domain) and an enhanced innate immune response in the brain. Interestingly PA 237E is located in the PA-X ORF the function of which for avian cells is unknown.
On the other hand, another study mapped virulence determinants to the PA and PB1 genes of a HPAI H5N1, yet was not able to link this to increased levels of polymerase activity (239). Using reverse genetics, Hulse-Post et al. (2007) assigned pathogenicity of A/Vietnam/1203/04 (H5N1) in mallards to the PA (T515A) and PB1 proteins (K207R and Y436H), yet in vitro and in vivo measurements of replication showed no difference between the mutant viruses. This suggests mechanisms may exist for the polymerase genes to increase pathogenesis other than by increasing virus replication. It is interesting to speculate if the roles of the recently discovered PA-X, PA-N155 and PA-N182 could alter pathogenesis in the avian host (10, 11). So far, in the investigations of PA-X function, a PA-X deficient virus was found to cause greater disease in the mouse model, hinting at roles of PA-X in the modulation of the host immune response, likely due to its ability to cause host cell shutoff (10), a trait recently confirmed by a second study (240).

Influenza NP plays an important role in polymerase function, through its oligomerization and RNA binding properties (241). A study comparing two H5N1 viruses with differing pathogenicity found the NP of the more HP virus increased virulence of the less pathogenic virus in chickens; the authors suggest this was due to an upregulation of IFN induced genes, likely caused by an observed increase in replication (242). This was later mapped to a single residue change, NP 184K (243). A small increase in virulence was also gained by swapping the HA, NS and M genes into the less pathogenic background. Interestingly, in this instance, the PB1 and PB2 genes from the HP virus decreased replication and ultimately virulence in the LP virus background.

However, other functions of NP protein have been described, and their role in virulence has been examined. Tada et al. (2011) assigned the PB2 and NP genes of a pathogenic chicken H5N1 virus as responsible for increased virulence over a duck H5N1 virus in chickens (244). Both the NP and PB2 genes increased pathogenicity in chickens when introduced into the duck virus background by reverse genetics. A single point mutation NP 105V, was discovered that converted the phenotype which the authors suggest could be a motif for adaptation of duck viruses to chickens. NP has also been associated with increased pathogenicity of an H7N1 virus following intracerebral inoculation. Although a duck H7N1 virus had a LP phenotype by IVPI assay, intracerebral inoculation increased the pathogenicity in chickens. This was mapped to NP residues 50N and 98K by reverse genetics,
when compared to a H7N1 virus that did not acquire a HP phenotype after intracerebral inoculation (245). In mammalian systems, mutations in NP that are host adapting or affect virulence have been mapped to a domain thought to be the target of the antiviral Mx protein. The role of influenza NP in modulating the innate immune response by Mx has been described (246). In addition, the regions around the residues described above has been implicated in the Mx adaptation of avian viruses to humans, namely at 100 and 49, demonstrated by Manz et al. (2013) (247). It is possible that influenza viruses require Mx adaptation between certain avian species, increasing virulence as they do so. However, since chicken Mx reportedly lacks anti-influenza virus activity, it is not clear that this would explain the difference in that particular avian host.

PB1-F2 has been identified as a key virulence factor, and since its discovery in 2001 (8) many functions have been attributed to it, all of which may have the potential to alter virulence. A recent review by Chakrabarti et al. (2013) describes the primary role of PB1-F2, to induce apoptosis in the cell by depolarization of the mitochondrial membrane, along with other functions that have been reported including deregulation of the innate immune response and inducing inflammation, regulating polymerase activity (in a strain specific manner) and increasing the risk of secondary bacterial infections (248). H5N1 PB1-F2 contributed to the HP phenotype in mallard ducks. Mutation at three residues 51, 56 and 87 reduced lethality in mallards and correlated with increased polymerase activity in chicken cells (249). A single mutation in PB1-F2 of H5N1 virus, N66S, was found to have a small increase of pathogenicity in ducks (measured by IVPI test), compared to mice where the mutation significantly increased replication and pathogenesis, particularly in the brain. However, a virus engineered to lack PB1-F2 significantly reduced virulence in the duck by IVPI test, with a delay in clinical signs and systemic spread but not overall mortality rates following a natural route of infection (250). PB1-F2 N66S has previously been associated with increased antagonism with the mammalian RIG-I/MAVS protein complex, it would be interesting to examine the effect in regard to duck RIG-I or chicken MAVS (251). Interestingly, PB1-F2 motif associated with increased cytokine release and inflammation in mammalian influenza models (including 62L, 75R, 79R, and 82L), have been found to be present in a high proportion of H5N1 avian isolates (252). For now, the mechanism behind the observed increase in pathogenesis by PB1-F2 in vivo in ducks remains unclear and the
effect of PB1-F2 for poultry hosts has not yet been explored. Furthermore, since some influenza strains do not encode for functional PB1-F2 proteins, understanding the role of PB1-F2 is complicated further (Reviewed, 196)

**NS1**

Being the main IFN antagonist protein of influenza, the role of NS1 and virulence may seem obvious. Certainly *in vitro* investigations of NS1 from avian influenza viruses have been conducted. The first links to the IFN antagonist function of NS1 was by the replication of NS1 deletant viruses in IFN deficient cells (253, 254). Since these initial observations NS1 anti-IFN abilities have been described as strain specific, at least in mammalian cells (255), and the mechanism by which NS1 modulates the IFN response in mammalian cells has been thoroughly investigated (49). *In vitro* investigations in avian cells using IFN reporter plasmids driven by the chicken IFN-β promoter have confirmed this function against avian hosts (210, 256).

However, knowledge of the described functions *in vivo* in avian species is limited. Indeed, work by Penski et al. (2011) proved that the ability of NS1 to limit IFN induction was not indicative of its function during infection of chickens, at least for HPAI. Although attenuated, viruses with mutated NS1 proteins did not induce greater IFN in chickens than WT. Furthermore pre-treatment of chickens with IFN-α failed to protect them from HPAI (257). However, several studies have made a link between the ability of NS1 to control IFN and pathogenesis in avian species. One study compared two H5N1 viruses of different phenotypes, *A/goose/Guangdong/1/96* (highly pathogenic in chickens) and *A/goose/Guangdong/2/96* (unable to replicate in chickens). The phenotypic difference was mapped to a single mutation in NS1, V149A, which converted GsGd/2/96 to a pathogenic infection in chickens, likely by means of its ability to counter the avian IFN response as demonstrated *in vitro* (258).

It has been noted that since 2000, many HPAI H5N1 viruses lacked 15 nucleotides at positions 263-277 from the NS1 gene (aa80-84, N-terminus). Reverse genetics was used to determine that this deletion together with a resultant amino acid change, D92E, contributed to an increased pathogenesis in chickens, although the mechanism behind this observation was not examined (259). A later study by Zhou et al. (2010) showed a 5aa deletion (NS1 85-
adjacent to the domain described above actually attenuated HP H5N1 virus in chickens (260). This was due to loss of binding to cellular eIF4GI, an important function of NS1 which enables preferential translation of vRNAs. This may explain the D92E shift reported by Long et al. (2008) as a means to readapt the region since eIF4GI binding is important for the virus. Interestingly, a 5aa deletion between residues 191-195 (C-terminus) was reported to severely attenuate a H5N1 virus (derived from swine) in chickens, and failed to prevent IFN induction in chicken cells (261).

Comparisons of two viruses with different pathogenic phenotypes can provide interesting results and allude to virulence mechanisms, but interpretations from these studies must be made with caution. An example of this has been the investigation of NS reassortant viruses. By swapping the NS segment for two HPAI viruses, one very virulent in ducks and one moderately virulent, Sarmento et al. (2010) did not detect any changes in pathogenesis or regulation of immune-related genes (262). Yet in a separate study, swapping the NS genes of a HP H7 and H5 virus increased the virulence of H7 in poultry (263). These findings may demonstrate the issue with strain specific mechanisms which are difficult to untangle, or highlight alternative mechanisms such as NEP (also expressed from NS segment) which can affect replication. Furthermore, single gene reassortants may give false impressions on the mechanisms of pathogenesis. Liniger et al. (2012) demonstrated a gene constellation from a HP H5 virus was necessary for controlling IFN, after comparative analysis with an attenuated H5 virus. The PB1, PB2 and PA genes were required together with the NS for optimal control of the IFN response. Crucially this study also demonstrated that PB2 could partially inhibit the IFN response in chicken cells; exogenous expression of either PB2 protein in chicken cells inhibited IFN activation stimulated by co-expression of CARDIF or MDA5, cofactors of the IFN stimulatory pathway, with the PB2 from the virulent virus achieving this more efficiently (264). The role of influenza proteins other than NS1 in altering virulence by modulation of the innate immune pathway is less understood and could highlight a further mechanism of increased virulence by HPAI.

The NS1 protein of influenza viruses carries a PDZ domain on its C-terminus ((postsynaptic density protein 95, Drosophila disc large tumor suppressor, and zonula occludens 1 protein) domain). This domain occurs on many cellular proteins and is thought to be involved in a number of processes such as cellular polarity and cell signalling (265).
The role of this domain in influenza infection is not yet clearly described. To determine whether this domain contributed to virulence in avian hosts, truncation mutants of HPAI H5N1 A/Vietnam/1203/04 were generated. Infection in 6-week old chickens showed no significant difference in virulence (266). Another study using a LP H7N1 virus showed deletion of the ESEV domain did not affect replication in both ducks and chickens (267). However, this virus carried the ESEV motif, whereas the most recent Eurasian lineages of H5N1 viruses carry ESKV, and typical human adapted seasonal viruses RSKV. Interestingly, the human RSKV motif was shown to increase virulence for a H7N1 virus in pekin ducks with increased replication and Mx induction, although in comparison WT H7N1 with its ESEV motif was shed for a longer time, which the authors speculate could affect transmission of the virus and may explain the maintenance of ESEV in nature (268). Further study may be needed to determine the role of this PDZ domain in the avian host. Investigations in the effects of the PDZ domain in mammalian hosts may shed light on this. A study of human influenza viruses encoding the ESEV sequence showed the disruption of cellular tight junction by binding to Scribble and Dlg1 in human cells (269). The authors suggest this may be a mechanism by which the virus can disseminate and could explain the increased pathogenesis of human H5N1 infections by viruses carrying the ESEV motif. It may be that this PDZ domain of typical AI is not a contributor to virulence in the avian host but enables efficient transmission. It would be interesting to examine the effect of the Eurasian H5N1 ESKV motif for avian species, and if this able to increase virulence as seen by the human RSKV sequence.

In conclusion it appears there are many routes by which HPAI can increase virulence, and reassortment of gene segments may be an efficient way to do this. However, many studies have revealed increases of pathogenicity by single or multiple gene reassortments with HPAI viruses. Strain specificity, gene constellation, specific protein functions (gain or loss) and host adaptation all have a role, which presents a challenge if we wish to understand HPAI virulence and use this knowledge to better prevent and control disease outbreaks.
1.6.4 A case study of avirulent and virulent clones from HPAI virus A/turkey/England/50-92/1991

In 1991 there was a HPAI H5N1 virus outbreak in Norfolk, England causing the death of 7129 18-week-old turkeys accommodated in a single house (270). The virus was isolated and termed A/turkey/England/50-92/1991 (50-92). Virus pathogenicity was measured by intravenous pathogenicity index (IVPI) in 6-week-old chickens and gave the maximum score of 3.00 (270), confirming this virus to be HPAI.

Interestingly, viruses of both low and highly pathogenic phenotype were cloned from the original virus isolate by plaque picking in MDCK cells, at the Animal Health and Veterinary Laboratories Agency (AHVLA), Weybridge. The virulence of the clones was measured by means of the IVPI test in chickens. Viruses were sequenced and comparison to the original isolate revealed several mutations in the HA gene, described in Table 2 (271). Plaque clones with a lower level of pathogenicity carried the same genetic sequence in the HA gene as the consensus sequence of the original isolate at positions 160, 193 (H3 numbering) and HA$_{24}$ (H5 aa residue 348), including the same multi-basic cleavage site. Plaque clones that had the highest level of pathogenicity contained one or all of 3 mutations in the HA gene, Table 2. In addition, the Mean Death Time (MDT) of embryonated chicken eggs inoculated with each plaque clone correlated to the pathogenicity identified by IVPI test (unpublished data by Banks et al. AHVLA).
Table 2. A summary of amino acid changes in the HA gene from plaque clones compared to the 50-92 isolate sequence together with an IVPI score of pathogenicity and Mean Death Time (MDT) in chicken eggs

<table>
<thead>
<tr>
<th>Virus</th>
<th>HA amino acid</th>
<th>IVPI(^b)</th>
<th>MDT(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>160(^a)</td>
<td>193</td>
<td>HA(_2)</td>
</tr>
<tr>
<td>50-92 isolate</td>
<td>A</td>
<td>E</td>
<td>G</td>
</tr>
<tr>
<td>Pathogenicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5L</td>
<td>A</td>
<td>E</td>
<td>G</td>
</tr>
<tr>
<td>7S</td>
<td>A</td>
<td>E</td>
<td>G</td>
</tr>
<tr>
<td>4S</td>
<td>A</td>
<td>E</td>
<td>G</td>
</tr>
<tr>
<td>2L</td>
<td>A</td>
<td>E</td>
<td>G</td>
</tr>
<tr>
<td>13L</td>
<td>A</td>
<td>E</td>
<td>R</td>
</tr>
<tr>
<td>3L</td>
<td>T</td>
<td>K</td>
<td>R</td>
</tr>
<tr>
<td>11L</td>
<td>T</td>
<td>K</td>
<td>R</td>
</tr>
</tbody>
</table>

\(^a\) 160T permits an N-linked Glycosylation at aa158.
\(^b\) IVPI score in 6-week old chickens, ≥1.20= HP. IVPI data from Wood et al. 1994 (271).
\(^c\) Mean Death Time (hrs) in 9 day old embryonated chicken eggs. Carried out by Banks et al. (data not published).

Taken together this work suggests that the original 5092 isolate was a quasispecies, meaning a group of genetically diverse viruses closely related to one another, containing at least some highly pathogenic virus but also harbouring genotypes that in themselves were not HPAI. It should be reiterated that all 50-92 clones contained identical MBCS sequences in HA.

A reverse genetics system was developed for the rescue of recombinant 50-92 based on the vRNA sequence of the 2L clone, by Howard et al. (2007) (272). Virus was generated using this system and an IVPI test performed. The RG 50-92 did not score as highly as expected and was classified as a low pathogenic virus (Table 3). In addition a reassortant 50-92 virus with the H7 and N1 genes of HPAI A/Chicken/Italy/13474/99 was generated and pathogenicity deduced by the IVPI test. This virus was highly pathogenic (unpublished data performed by Howard et al., AHVLA) suggesting that the internal 50-92 genes rescued in the RG viruses were able to support a HPAI phenotype. Thus it seemed likely that the HA
sequence cloned for the 50-92 RG system did not support a HP phenotype, or that there were other sequence differences in the generated RG virus that were not revealed as attenuating when combined with the H7N1 HA NA gene pair. This study seeks to determine the restriction of the 50-92 RG system and restore pathogenesis (see section 1.8).

**Table 3. A summary of the IVPI scores from work carried out previously, and HA genotype for the original 50-92 isolate; plaque picked clones and RG generated virus with cognate HA and NA or mis-matched H7 and N1 from HPAI A/Chicken/Italy/99 RG system**

<table>
<thead>
<tr>
<th>Virus</th>
<th>HA genotype</th>
<th>IVPI score</th>
<th>Carried out by:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original isolate</td>
<td>A160, E193, HA₂G₄</td>
<td>3.00</td>
<td>Wood et al. (1994)</td>
</tr>
<tr>
<td>Clones IVPI&lt;2.00</td>
<td>A160, E193, HA₂G₄</td>
<td>0.00-1.77</td>
<td>Wood et al. (1994)</td>
</tr>
<tr>
<td>Clones IVPI=3.00</td>
<td>160T, 193K, HA₂₄R</td>
<td>3.00</td>
<td>Wood et al. (1994)</td>
</tr>
<tr>
<td>RG 50-92</td>
<td>A160, E193, HA₂G₄</td>
<td>0.20</td>
<td>Howard et al. (unpublished data)</td>
</tr>
<tr>
<td>RG 50-92 (internal genes only) *</td>
<td>H7 gene</td>
<td>3.00</td>
<td>Howard et al. (unpublished data)</td>
</tr>
</tbody>
</table>

*A/Chicken/Italy/13474/99 HA and NA (HPAI H7N1) (273) with 6 internal genes from 50-92.
1.7. Host adaptation by influenza polymerase

An important area of on-going research seeks to understand the host restriction of influenza viruses. A great deal of research has addressed the role of HA in host adaptation, and the mutations that mediate the switch in preference of HA to bind to avian \( \text{SA}\alpha_2,3\text{Gal} \) or human \( \text{SA}\alpha_2,6\text{Gal} \) receptors. Several domains and mutations have been mapped for numerous HA subtypes that enable host adaptation, particularly from the perspective of avian to human adaptation (Reviewed, 262). However, influenza polymerase also plays an important role in the adaptation of influenza viruses to different hosts.

Often an important route to understanding the mechanisms by which mutations achieve changes in protein function is to resolve the protein structure of the concerned constituent. The influenza polymerase structure has proved difficult to resolve, but recent years have shed more light on the structures with advances in technology. Partial crystal structures for PB1 and PB2 are available, and complete crystal structure for NP with separate N and C-terminal structures for PA (35, 275–279). The crystal structures together with Cryo-EM analysis have provided useful information, supporting proposed models such as the trans/cis model for replication/transcription (40, 280). A summary of recent data by Arranz et al. (2012) and Moeller et al. (2012) is described in Figure 8. These data confirmed the double helix formation of the vRNA with NP, and confirmed the location and length of RNA per NP (20-24nt/NP) (281, 282).
Figure 8. Three-dimensional structure of influenza RNP and polymerase

A. Cryo-EM reconstruction of a double-helical RNP by Arranz et al. (2012) (281). Terminal NP loop (yellow), polymerase (green and brown). B. Model showing the localisation of vRNA (yellow) with NP crystal structure modelled into place showing surface potential (Left [Red and Blue]) and crystal structure (Right [red]) (281). C. Cryo-EM reconstruction of the polymerase bound to RNP, PB1 and PB2 (orange), PA (red) and NP (green) adapted from Moeller et al. (2012) (282). D. Model showing crystal structure of PA, with demonstration of observed arm rotation (282).

Other techniques are used together with structural data to identify mutations and mechanism that confer host adaption. Approaches include sequence analysis of viral genomes from isolates or viruses passaged through foreign hosts; bioinformatic analysis of viral sequence data available in large public databases, for example, gene alignment; and functional and biochemical analysis of virus proteins or virus organisms (87, 88, 283–285). A large number of domains and residues have been mapped using such techniques for each protein involved in genome replication, outlined in Figure 9. The majority of these data relate to the adaptation of avian origin polymerases to mammals, since avian polymerases do not function in mammalian cells. Little work has been done to understand the role these mutations have for avian viruses in avian hosts and why they may not naturally occur in the avian reservoir.
Figure 9. Mutations mapped that increase polymerase activity in mammalian cells

Examples of mutations and domains linked to increase of polymerase activity from published data, derived either experimentally or by bioinformatics. Functional domains: with viral proteins (green), nuclear localization (red), nuclear export (purple), with cellular proteins (orange), MxA resistance (yellow) and RNA binding (blue). Taken from Manz et al. (2013) (48).
1.7.1 Role of PB2 in host range adaptation

PB2 has long been identified as a gene that requires adaptation to enable avian virus replication in mammals. The first report of this nature was made by J. W. Almond in 1977. PB2 was identified as the restrictive protein of an avian influenza strain, Rostock, which failed to form plaques in mammalian cells, compared to a virus that could Dobson, (286). These two viruses were further characterised by Yao et al. (2001) and specific residues in the carboxy terminus of the PB2 protein were identified as being host adapting, able to increase polymerase activity in mammalian cells (287).

In research over the following years, several mutations in PB2 were implicated in host range adaptation. The most distinctive mutation identified to adapt avian origin polymerase to mammals is a mutation from a Glu to Lys at residue 627 of PB2. Human influenza viruses typically bear 627K on their PB2 and avian viruses carry PB2 627E.

The first complete sequence of PB2 was produced in 1983, and comparison of three human strains revealed all had PB2 627K, typical of human isolates, although the importance of this was not then known (288). The first publications that revealed the difference between human and avian PB2 sequences were in 1988 and 1990. Avian PB2 carried 627E and human viruses 627K, although these studies also did not identify PB2 627 as a host determining residue (289, 290). Clements et al. (1992) reassorted a human and avian virus and tested each one’s ability to replicate in squirrel monkeys and human volunteers (291). This work again identified avian PB2 as a restrictive factor in mammals. It was not until 1993 that PB2 627 was genetically identified as a host determining residue. Work in the lab of Brian Murphy conducted by Kanta Subbarrao passaged these reassortant viruses through mammalian MDCK cells and found the restricted virus acquired a replicative advantage by mutating at PB2 E627K (292).

The first human viruses for which we have sequence are from the pandemic of 1918, the sequences of which were painstakingly amplified from paraffin embedded lung pathology slides of soldiers who were victims of the 1918 pandemic. These were found have PB2 627K, although their phylogeny of the entire PB2 gene segment suggested a close relationship with avian virus sequences, suggesting that the amino acid change at residue 627 might have been recently acquired or selected for (112). Descendants of the pandemic
virus established themselves in swine, which continued circulation as ‘classical’ swine and retained the 627K mutation (293). Interestingly, it has recently been shown that introducing a reversion to PB2 627E did not alter the pathogenicity of the virus in pigs, a striking contrast to the observation in mice which strongly correlates 627K with increased virulence (293–295), and surprising bearing in mind that Moncorgé et al. showed that the 627K mutation did indeed enhance activity of avian polymerases in pig cells. Subsequent human seasonal H1N1 influenza also derived from the pandemic maintained the 627K motif, as did the later H2N2 and H3N2 pandemic viruses and their derivatives. pH1N1 2009 proved an exception to this rule and circulates in humans without PB2 627K. However, work from both the Kawaoka groups and Mehle and Doudna has shown that this virus overcomes the polymerase host range restriction by a pair of mutations 590S and 591R, which enable polymerase activity in human cells (296). Similarly to the 627K mutation, 591R is a positive charge and lies very close to the 627K domain (297) Moreover, restriction is overcome by either an R or K at 591, and the addition of 627K to the 590/591SR mutation did not significantly increase activity further in human cells, suggesting a common mechanism (296, 298–300).

1.7.2.1 Bioinformatic analysis

Bioinformatics can use powerful tools to process large amounts of data and analyse mutations present in genes using defined parameters. In addition the evolution of viral genes has been revolutionised by the use of computational phylogenetic analysis. Yet the power of these tools relies on the collection of vital data from the lab and the field to obtain the samples in the first place in order to understand the significance of sequence variation within the biological context. Global surveillance for human influenza has operated globally since 1952, through WHO’s Global Influenza Surveillance and Response System (GISRS). This information is used to recommend strains for inclusion in seasonal vaccines. Monitoring and surveillance of influenza in animal species is led by the OIE. Information is shared between organizations to help inform vaccine strain recommendations for pre-pandemic vaccines e.g. for H5 and H9 subtypes. Much virus sequence data is deposited into public databases through these networks.
Bioinformatic studies have utilised these data and describe host specific residues, deduced by statistical probabilities. PB2 627 has repeatedly been highlighted by this method confirming the presence of 627K in humans isolates and 627E in avian isolates (301–309). A summary of these data for PB2 and other viral proteins involved in replication is shown in Figure 9.

Bioinformatics has also enabled the processing of complex experimental data, such as that from large scale RNAi screens which seek interacting host factors (310–312). Such systems biology approaches produce a great deal of information, but often this is hard to understand and apply to a biological context, such that follow up experiments are required to confirm the initial findings. Such a study has been performed comparing host factors that interact with PB2 627E or PB2 627K, and revealed several proteins to be differentially involved, in particular host protein DDX17 (313).

1.7.2.2 Structural analysis

The structure of the domain of PB2 that contains residue 627 has been determined by X-ray crystallography for A/Victoria/3/1975(H3N2) (VIC) (residues 538-693) and A/Puerto Rico/8/34 (H1N1) (residues 535-759) (297, 314). A prominent 627-domain was observed where residue 627 was surface exposed. The E627K mutation altered the electrostatic charge of the exposed surface, resulting in a highly basic surface (297, 314). These findings supported the hypothesis that host restriction was the result of a loss of a host specific interaction with a cellular and/or viral partner. Moreover, mapping of the pH1N1 590S and 591R mutations onto an avian polymerase highlight their ability to disrupt the negative charged surface of the 627 domain (shown in red) similarly to 627K (Figure 10) (296), supporting the common mechanism between these residues and 627K.
Figure 10. Effect of the K627E and 590S/591R mutation on the electrostatic surface of the 627-domain

Electrostatic surface potentials were mapped onto models of A/green-winged teal/OH/175/1983 (S009), including WT and the 590S/591R and 627K mutants. The 627 residue is highlighted. Electrostatic charge is coloured: -4 kT/e (red) to 4 kT/e (blue). Figure adapted from Mehle et al. (2009) (296).

1.7.2.3 Functional Analysis

As well as measuring the replication of a virus in vitro or in vivo, the effect of PB2 627 has often been analysed using the polymerase assay. This assay allows the measurement of polymerase activity via a ‘minigenome’ reporter which encodes a reporter gene such as CAT, firefly luciferase or GFP flanked by the 5’ and 3’ non-coding regions which contain the viral promoters. The minigenome is delivered into the cell in a species-specific poll vector which transcribes a virus-like RNA, with authentic 5’ and 3’ termini. Co-expression of the influenza polymerase and NP genes leads to the transcription and replication of the in situ generated viral like RNA, the mRNA species of which is subsequently translated and measured accordingly (96). An illustration of the assay principle is shown in Figure 11.

This technique drastically reflects the restriction of avian origin polymerases carrying PB2 627E in human cells (292, 315–318). A single switch to PB2 627K permits the polymerase to function in the unfamiliar cell type (317, 319). This is presented in Figure 11, the 627K mutation permits polymerase activity in human cells (293T) whereas 627E is restricted; The effect was minimal in avian cells (DF-1) in the assay shown with this particular mouse-adapted polymerase constellation, A/WSN/33 (H1N1) (320).
Figure 11. Polymerase reconstitution assay

Co-transfection of plasmids expressing the influenza polymerase, NP and a minigenome reporter results in viral transcription and replication of the virus-like RNA, typically encoding luciferase or GFP. The mRNA is translated into protein and can be measured by the appropriate assay system. This diagrammatic representation was adapted from Manz et al. (2013) (48). An example is included which demonstrates the restriction of influenza polymerase with 627E in human 293T cells, which is not true for chicken DF-1 cells, measured by a luciferase minigenome reporter. Taken from Hudjetz et al. (2012) (320).

The 627K mutation has also been observed to be cold adapting (316, 321, 322). In mammalian cells, viruses carrying 627K had a significant fitness advantage over those with 627E that was only pronounced at the cooler temp of 33°C. This difference was not apparent in avian cells, indicating the 627K mutation confers little or no advantage in avian cells compared to human cells, and/or that PB2 activity is less affected by temperature in avian cells (316, 321). The effects of temperature have been mapped to the enzyme kinetics of influenza polymerase. Biochemical investigation of polymerase reconstituted from a baculovirus expression system showed PB2 627K had significantly reduced activity at the higher temperatures of 37 and 42°C compared to 627E, but 627K was more catalytically active at 34°C, suggesting that the 627K mutation alters enzyme catalysis, rather than by RNA binding (323).
Initial experiments suggested PB2 627K stabilised the assembly of vRNPs by increased binding to NP in human cells (319, 324–326), supported by the fact that the origin of the NP protein has been implicated in the selection of PB2 627K (327). However, despite this correlation, recent work by Cauldwell et al. (2012) demonstrated that the observations of PB2:NP stability mediated by 627K was an indirect observation accounted for by amplification of viral RNA which results in an increased PB2-NP co-precipitation (328).

627K is not the only polymerase host determining residue on PB2. Other mutations can also increase activity of avian polymerase in mammalian cells, such as PB2 D701N (322). The C-terminus of PB2 has been shown to interact with importin-alpha proteins, and enable the nuclear import of PB2 (279, 329). The 701N mutation causes a switch of avian PB2 dependence on importin-α 3 to human importin-α 7 which increase virus replication and has been shown to increase virulence in mice (329, 330). PB2 701N has also been implicated in altering cap-binding by the PB2 protein, possibly providing further evidence of how this mutation increase replication and virulence in mammalian hosts (331). Recently, it was shown that PB2 627K also has a host-specific role in the interaction of PB2 with importin alphas. The polymerase complex containing PB2 627K was dependent on importin-α 1 and 7, whereas the 627E mutant was not. The link with the 701N mechanism may explain why 701N is able to compensate the lack of polymerase activity by 627E in mammalian cells (322). These recent studies have demonstrated surprising overlaps in the function of residues 627 and 701.

It has also been shown that the 627 region of PB2 is involved in RNA binding (314). In the absence of cell co-factors the ability of PB2 to bind to RNA oligonucleotides was measured. PB2 with 627K had significantly higher degree of binding than PB2 with 627E. However, it is hard to consider the relevance of this with host restriction, unless the binding to RNA involves a species-specific host co-factor, since PB2 627E is perfectly capable of functioning in avian hosts. Furthermore, an alternative study has shown that the effect of 627K is enzymatic rather than dependent on the RNA template (323). Further work is needed to understand the role of PB2 627K and RNA binding.

Further insight into the mechanism behind the restriction of avian virus polymerase such as those containing PB2 627E in human cells was revealed by Manz et al. 2012 (43).
Previous to this work, it was unknown at which point in replication avian origin polymerases were restricted in human cells; during vRNA, cRNA or mRNA synthesis, or indeed all three. Using PB2 complementation assays with a PB2 protein modified to prevent mRNA synthesis, it was shown avian origin polymerase could perform mRNA synthesis when vRNA template was provided *in trans*, since co-expression of the avian PB2 with the transcription deficient human PB2 resulted in restoration of mRNA synthesis. This was further demonstrated in the context of virus by the infection of human cells with either avian virus (627E) or adapted virus (627K) in the presence of cycloheximide. Since translation was blocked by cycloheximide, the demonstration of equal levels of primary transcription by both viruses concluded avian influenza polymerases are not restricted for mRNA synthesis in mammalian cells. Furthermore when polymerase with either 627E or 627K were expressed in cells and treated with cycloheximide (preventing further protein synthesis) and then cells were infected with either 627E or 627K virus, results showed vRNA synthesis did not occur (43). This was perhaps due to defective cRNA templates produced by the avian origin polymerase (627E). This observation was overcome by cRNAs generated by a polymerase containing PB2 with the 627K mutation, yet whether the defective cRNPs were such because they lacked a cellular co-factor remains unclear. Interestingly, an alternative mechanism by which restriction to avian polymerase in human cells could be overcome was revealed by Manz et al. showing that mutations in NEP protein could override the requirement for 627K (43).

Despite a large effort in this area from many labs, the underlying mechanism of 627K and host adaptation remains unknown. The role of DDX17 is not convincing, discussed further in Chapter 4, and studies highlighting the role of importin alpha subunits in host restriction have largely been performed with RNAi which might affect the accumulation of several different host factors in the nucleus in addition to the importins themselves. The demonstration of defective vRNA transcription by avian origin polymerase and the novel role of NEP have brought fresh insight into the field.
1.8 Thesis aims and objectives

There has been an increasing amount of evidence gathered that described virulence mechanisms of HPAI beyond their MBCS sequence (Section 1.6). The predominant role that the MBCS plays in systemic replication and ultimate pathology cannot be ignored, yet it appears other viral genetic features contribute to pathogenesis. It is important to understand the mechanisms by which AI increases virulence due to the pathogenic nature of these viruses and the disease and economical costs involved. Section 1.6.4 describes a set of virus clones from a H5N1 virus, 50-92, that varied in pathogenesis despite all carrying a MBCS. An attempt to generate a HP H5N1 virus based on the existing RG system will be attempted, and following successful generation of this virus the difference in pathogenesis between the virus clones will be examined. The recent increase of virulence by Asian H5N1 viruses is of particular concern. As is their apparent spread and increased pathogenesis in wild bird species (Section 1.6).

A recent clade of H5N1 HPAI viruses carried the human adaptation mutation, PB2 627K (Section 1.6). This increases the pandemic threat of HPAI H5N1, which have caused hundreds of human infections in the past 16 years but have failed to initiate a pandemic. The role of PB2 627K has rarely been studied in the context of the avian host, since in vitro analysis has demonstrated the ability of polymerase bearing 627K to function in avian cells (Section 1.7). It is important to understand the effects of this mutation in the context of virus infection in the avian host, since the vast majority of AI viruses isolated carry PB2 627E and not 627K.

**Aim 1: Restoration of the highly pathogenic 50-92 H5N1 reverse genetics system**

Objective 1. Introduce the three HA mutations discussed in section 1.6.4 present in the HP virus clones into the 50-92 RG system by site-directed mutagenesis

Objective 2. Generate corrected version by Reverse Genetics virus rescue in high containment and assess the pathogenicity by IVPI index in chickens.

**Aim 2: Elucidate the mechanism behind the differing pathogenicity of the 50-92 virus clones**
Objective 1. Generate alternative combinations of the HA mutations in the 50-92 genetic background by site-directed mutagenesis and generate viruses by reverse genetics virus rescue.

Objective 2. Examine the effect on pathogenesis and virus replication of the different virus by IVPI index in chickens.

Objective 3. Examine structural location of these mutations and known functions. Use appropriate *in vitro* assays to elucidate the mechanism of increased pathogenesis.

**Aim 3: Examine the role of PB2 627K for the avian host**

Objective 1. Generate the converse PB2 627 mutants for a modern HPAI H5 virus (K627E) and a historic HPAI H5 virus (E627K) by reverse genetics.

Objective 2. Examine the effect of PB2 627 for both virus backgrounds in the appropriate host *in vivo*. Analyse the effect on polymerase activity and virus growth *in vitro*.

The overall aim of this study was to understand the mechanisms underlying pathogenesis of HPAI H5N1 for the avian host and the observation of maintained human adaptation motifs in the avian reservoir by recent Asian HP H5N1. This was to be achieved using a reverse genetics approach whereby specific mutations could be engineered into the avian influenza virus genome and their effects in the avian hosts monitored by *in vivo* studies. Towards this end a novel reverse genetics system for a traditional H5N1 avian influenza virus had to be generated and validated. A reverse genetics system for a modern Eurasian lineage H5N1 virus was obtained from another laboratory and again it had to be validated. Then using these two systems the effects of the mammalian adapting mutation in PB2 E627K on virus replication and transmission *in vivo* was tested.
Chapter 2. Mechanisms of increased virulence by influenza H5N1 HA protein

2.1 Introduction

As discussed in Chapter 1, the presence of a MBCS in the HA protein of influenza is the most predominant determinant of the HP phenotype of H5 and H7 influenza viruses, permitting systemic dissemination of the virus. However, research has shown that presence of this motif may not be solely responsible for increased pathogenesis. There are several examples of different virus subtypes engineered by reverse genetics to artificially contain a MBCS. The H2, H4, H6, H8, H9 and H14 subtypes acquired high pathogenicity after the introduction of a MBCS (170, 172, 173), yet the H1 and H3 subtypes did not (170, 171). This indicates that other factors may also be necessary for the HP phenotype. Such viral genetic factors could reside on internal gene segments, on the NA or on the HA gene itself.

A classic example of the influence of the HA protein sequence on the presentation of a HP phenotype from a virus with a MBCS was discovered for plaque variants of a HP H5N2 virus, A/Chicken/Pennsylvania/1983. Analysis revealed glycosylation of the HA protein at residue 11 was masking the MBCS from access of proteases and preventing efficient cleavage and therefore reducing virulence (332–334). Interestingly, Ohuchi et al. (1989) demonstrated the inhibition by the glycosylation site was possibly due to the unusually short multi-basic cleavage site of Pennsylvania/1983 (PQKKKR*GLF) which prevented access of ubiquitous proteases and that this could be circumvented by altering the positively charged sequence of the MBCS (335). Similarly, deletions in the length of the MBCS of A/Turkey/Ireland/1378/85 (H5N8) virus was shown to prevent cleavage, and it was postulated this was also due to the glycosylation at residue 11 (336). As suggested by others, it is likely the length and/or sequence of the MBCS dictates the extent to which the glycosylation at residue 11 can affect cleavage of a MBCS in H5 viruses (158, 337).

The mechanism described for Pennsylvania/1983 provides an example of how pathogenesis can be altered by few genetic changes, but analysis of HA gene sequence confirms that these exact mechanisms do not apply to HPAI H5N1 virus 50-92 (Section
1.6.4). The MBCS for 50-92 is PQRKRKTR*GLF, which is not particularly short for H5 viruses (338–340); moreover the glycosylation consensus at aa11 was identical in both the LP and HP 50-92 viruses and both phenotypic groups carried the same MBCS.

Section 1.6.4 describes 3 mutations in the HA gene of 50-92 that correlated with increased pathogenicity of 50-92 in the IVPI assay; the addition of a glycosylation site at 158N by the mutation of A160T, and two further changes at E193K in the receptor binding site and HA24R in the fusion peptide.

2.1.1 Role of glycosylation 158 for influenza virus

The role of glycosylation of proteins in eukaryotic cells has been well studied; the addition of N-Glycans to a newly translated protein in the endoplasmic reticulum can increase stability, prevent protein mis-folding and degradation (341). There have been several demonstrations of the important consequence of glycosylation for influenza HA glycoprotein. One of the effects of glycosylation of HA has been to alter pathogenesis. Glycosylation on the globular head of the HA for a H7N7 virus was shown to recover the HP phenotype in chickens as opposed to the LP variant which had the glycosylation site deleted (342). Another study showed the positive selection of HA glycosylation at 123 and/or 149 in H7N1 viruses isolated from a poultry outbreak in Italy, 1999, together with a deletion in the NA stalk (120). In contrast, pathogenicity phenotype in chickens for H5N1 Vietnam/04 virus was indifferent to the presence or absence of glycosylation at 158N. Vietnam/04 was mutated to lose the glycosylation at 158 plus an additional mutation in the RBS at S227N and this virus did not affect lethality for chickens by IVPI assay, scoring 2.98 compared to the WT 3.00.

Recent studies have shown that loss of glycosylation at HA 158N of H5N1 viruses was one of the mutations supporting more efficient respiratory droplet transmission between ferrets (87, 88). In these independent studies, with different H5 HA genes from H5N1 viruses, the deletion of glycosylation in addition to mutations in the receptor binding pocket to confer the α2,6 receptor switch and an additional mutation in HA enabled transmission. Similarly a live attenuated vaccine of H5N1 Vietnam/04 required loss of glycosylation at 158N and RBS mutations for efficient replication in the upper respiratory tract of ferrets (343).
Residue 158 has also demonstrated the ability to alter binding affinity in H1 influenza viruses. Although glycosylation was not present in this HA subtype at aa158, a Gly to Glu mutation was shown to reduce binding to both α2,3 and α2,6 SA receptors, as well as enable antibody escape (344, 345), which suggests the proximal distance of the 150-loop to the RBS is a mechanism whereby influenza may alter receptor binding affinity in a number of ways and protect against antibody recognition.

2.1.2 Role of residue 193 in receptor binding of influenza HA

Residue 193 forms part of the 190-helix found in the influenza HA RBS. Mutations in this region have been associated with receptor binding switching from avian α2,3 to human α2,6 SA. In particular for H1 and H3 subtypes, the E190D mutation has contributed to receptor switching, E190 is the typical avian signature, whereas D190 is predominant in human strains (17, 346–348). X-ray analysis of crystal structures of a H5 HA protein reveals the presence of the 190-helix, and highlights interaction of these residues with α2,3 SA (349). As well as switching receptor preference, mutations to this region have been shown to increase binding to SA regardless of receptor preference. Introduction of a positive charge, either Lys or Arg at 193 was found to increase receptor binding for human and equine H3’s to α2,6 or α2,3 SA respectively. Mutation K193S of the equine H3 decreased its ability to bind to its preferred α2,3 SA receptor but did not alter the receptor binding preference (350).

2.1.3. Mutations in and beyond the fusion peptide that alter the ability of HA to fuse with the host membrane

HA2 mutation G4R, present in the HP virus clones, at first seems an unlikely candidate to increase pathogenesis, being such a dramatic amino acid substitution to a predominantly hydrophobic region. Indeed, the influenza fusion peptide has been described as highly conserved (Figure 12) (82). The first 11 amino acids of HA2 are highly conserved amongst the 17 subtypes, including HA2 4G. Surprisingly, this region in the recent H17 HA isolated from bats was also conserved, despite the divergent properties exhibited by H17 HA, including its resistance to low pH compared to other HA subtypes (351, 352).
Figure 12. Conservation of the influenza fusion peptide across 17 different antigenic subtypes

Representative fusion peptide sequences of the 17 subtypes, compared to H1. HA2 residue 4 is highlighted in red. Adapted from Cross et al. (2009) to include the recent H17 subtype (82).

Fusion peptides also demonstrate conservation across different virus families, described in Figure 13 (353). There is a distinct lack of positively charged amino acids in each of the virus families described. In addition there is a similar distribution in the abundant Glycine residues, including HA2 4G.

Figure 13. Fusion peptide sequences of the fusion glycoproteins from influenza viruses, paramyxoviruses, retroviruses and filoviruses

Representative fusion peptide sequences of viruses belonging to different families. The conservation of distributed Glycine residues is highlighted in red. Adapted from Skehel et al. (2001) (353).

The distribution of Gly residues in Figure 13 (highlighted in red) are recognisable as a GxxxG motif for transmembrane helix-helix association, together with the residues Ala, Val, Ile and Leu situated between the Gly residues (354). This suggests a structural function of Gly residues at HA2 4 and 8. Mutagenesis studies have revealed the role of glycine residues in the fusion peptide for influenza. The initial Gly 1 residue of HA2 has been well studied and
alterations to this amino acid often reduce fusion or prevent it completely. Qiao et al. (1999) substituted G1 for Ala, Ser, Val, Glu, Gln or Lys and measured the effect on fusion (355). Only Ala was able to support fusion and Ser enabled penetration but not mixing of the membranes. Interestingly, Val, a hydrophobic residue, was unable to initiate fusion, suggesting the N-term amino acid of HA2 is required to be small and apolar. Further studies of the G1S mutant have revealed the importance of shape in enabling fusion, Ser allowed penetration due to a kink in the structure, whereas G1V was too linear to allow penetration (356). Gething et al. (1986) revealed a G1E substitution prevented fusion activity with red blood cells, compared to a G4E mutant that resulted in an increase in the pH of fusion, although the efficiency was decreased, likely by the instability. Further studies of the G4E mutant by Kantchev et al (2004) showed the Glu mutation caused the fusion peptide to penetrate the lipid bilayer less deeply. However influenza does require polar residues at some positions in the fusion peptide; such as Glu at position 11, since Ala and Val mutants have been shown to inhibit fusion (357). This study also showed that a G8E substitution was able to increase the pH of fusion.

As well as the fusion peptide, mutations in the fusion pocket are commonly described. Reed et al. (2009) generated 19 HA proteins with point mutations in the peptide pocket of H5N1 A/Chicken/Vietnam/C58/04. Five mutations were found to alter the pH of fusion, Y23H and H24Q of HA1 and E105K, N114K and K58I of HA2. Interestingly this study compared their findings to fusion pocket mutations in other HA subtypes, and revealed such mutations were generally subtype specific, indicating the conformational mechanism behind such changes due to the reliance on the structure of each HA subtype (358).

It has recently been established that the pH of fusion is a key host-specific barrier of influenza virus and can also determine virulence. The recent studies that describe the successful adaptation of H5 for respiratory droplet transmission in ferrets contained a HA mutation other than loss of glycosylation and receptor binding switching. Imai et al. (2012) showed the role of T318I to reduce pH of fusion and stabilise the HA from temperature degradation. Herfst et al. (2012) reported the H103Y mutation, and although they did not present a mechanism to justify its selection, its location in the fusion pocket suggests a similar mechanism to that described by Imai et al (2012) (87, 88). In addition Shelton et al. (2013) presented a study using the ferret model and found H5 H24Q mutation decreased
the pH of fusion and in turn enabled more efficient transmission when coupled with receptor binding changes and loss of Glycosylation at 158N (89). A similar study was subsequently published by Zaraket et al. (2013) whereby H5 virus Vietnam/04 was transmitted efficiently together with a mutation that reduced pH of fusion, K58I, together with receptor binding mutations and linked the improved transmissibility with a shift in replication from the lung to the upper respiratory tract (91). The pH of fusion has also been implicated in the adaptation and pathogenesis of mice. Together with receptor binding changes Leu78 and Glu354 adapted PR8 virus to mice and increased virulence (359). Analysis of fusion activity revealed these mutations increased the pH of fusion. A/Chicken/Vietnam/58/04 virus containing K58I was also associated with increased pathogenesis in mice, although this virus was attenuated in ducks (90).

Zaraket et al. (2013) determined the higher pH of fusion (around pH 6.0) was preferred for a H5N1 virus in ducks, and that adaptation to mammals required a drop to pH 5.4 (90, 92). Reed et al. (2010) were the first to demonstrate the role of pH of fusion in pathogenesis and transmissibility in ducks (92). In this study, WT HA of A/Chicken/Vietnam/58/04 fused at pH of 5.9 and a N114K mutant that increased the pH further rendered the virus attenuated in replication and decreased its environmental stability. Similarly, mutations Y23H and K58I that increased and decreased the pH of fusion by pH +0.4 and -0.5, respectively, were associated with reduced weight loss, indicating pH 5.9 is the optimum pH for this virus in mallard ducks (92). Dubois et al. (2011), showed an increase in pH of activation of HA correlated with greater pathogenicity in the chicken using A/chicken/HonKongYU562/2001 and A/goose/HongKong/437-10/1999 H5N1 viruses, modelled on residues 104 and 115 (93). The pH of fusion of this virus panel ranged from pH 5.2 to 6.0 and the higher pH resulted in the greatest pathogenicity in the chicken.

We hypothesised that the three mutations present in the HP clones in the HA gene were required for efficient HA binding, fusion and release in the chicken host. We aimed to generate a 50-92 virus with all three changes and measure its pathogenicity compared to the LP 50-92 virus generated by Howard et al. lacking the hypothesised HP motif. Subsequently we aimed to elucidate the mechanism by which 50-92 HA can alter virulence in the chicken host, based on these three residues in the HA gene.
2.2 Results

2.2.1 Location of residues 160, 193 and HA2 4 on the H5 haemagglutinin

To identify the regions associated with the 160, 193 and HA2 4 mutations we located the residues on the protein structure of H5N1 Vietnam/04 solved by Stevens et al. (2006). We found the residues responsible for the conversion from LP to HP of 50-92 virus are in areas of known importance in the HA protein. As described in Figure 14, 160T is adjacent to the RBS on the prominent 150-loop and permits an N-linked glycosylation at Asn158. Mutation 193K is located on the 190 α-helix, directly in the HA binding pocket. Residue HA2 4 is the 4th amino acid in the fusion peptide which wraps around the HA stalk.

A. H5 Haemagglutinin structure

![HA1 head domain and HA2 with Lys193, Thr160, and Gly4 annotated]

B. Receptor Binding Site

![150-loop, 190-Helix, 130-loop, 220-loop with Lys193 and Thr160 annotated]

Figure 14. HA structure of H5N1 Vietnam/04 annotated with the amino acids of interest

The structure of Vietnam/05 HA was used to annotate the residues of interest for 50-92 virus. The amino acid identity, written in red, is that of Vietnam/04. Protein secondary structure is colour coded: pink is
α-helix, yellow is β-sheet. Residues of interest are labelled in red. (Vietnam/04= 160T, 193K, HA\textsubscript{2} 4G. 50-92-HP= 160T, 193K, HA\textsubscript{2} 4R). A. HA structure of H5N1 Vietnam/04 including the HA1 and HA2 domains. B. Receptor Binding site of HA1 globular head: including the 190-helix; 130-, 150- and 220-loops. The structure of Vietnam/04 HA was solved by Stevens et al. (2006) (360) and images generated and annotated using Geneious R6.

2.2.2 Restoration of the highly pathogenic phenotype of 50-92 H5N1 virus

In order to transform the LP phenotype of the 50-92 RG system to HP, we engineered the HA gene to carry three additional mutations, A160T, E193K and HA\textsubscript{2} G4R, previously observed in the HP plaque picked clones (Table 2). All the HA constructs carried a MBCS and differed only at the 3 HA residues described. The remaining segments of the 50-92 RG system were fully sequenced, which revealed a mutation in the NS gene causing a NS1 P85H substitution, not present in the original virus isolate (data not shown) or displayed on the 50-92 sequences present of the NCBI Influenza Database. Mutations to alter the HA and correct the NS genes were engineered in the poll RG plasmids by site directed mutagenesis.

2.2.3 HA mutations A160T, E193K and HA\textsubscript{2} G4R restore the highly pathogenic phenotype of 50-92 virus

We based the generation of the HP phenotype 50-92 RG virus on the RG system designed by Howard et al. 2007, which was previously shown to give an IVPI score of 0.20 (unpublished data). Firstly a new RG virus was made that was identical to that made by Howard et al. except the NS plasmid which now included the correct 85P residue in the NS1 protein. The newly generated 50-92 virus scored a similar IVPI to the previous attempt at 0.02, a LP score, suggesting that the NS sequence changes had not significantly affected the pathogenicity of the RG virus. In addition we generated a virus based on the HP plaque clone genotype carrying HA mutations 160T, 193K, HA\textsubscript{2} 4R (Table 4) together with NS1 H85P. This virus gave an IVPI score of 2.88, similar to the original isolate, thus defining this virus as highly pathogenic.

The ability of the P85H mutation to alter NS1 function was examined \textit{in vitro}. 50-92 NS1 protein with either mutation and two NS1 proteins of known characteristics (PR8 H1N1 and VIC H3N2) were exogenously expressed in human 293T cells and were able to prevent IFN induction under a Newcastle Disease Virus or Poly:(IC) stimulation, measured by a
luciferase reporter driven by an IFNβ promoter (data not shown). In addition, chimeric PR8 H1N1 viruses were generated carrying only the NS segment of 50-92. Both the P85 and H85 virus grew to high titres and neither induced significantly different levels of IFN in infected human A549 or chicken DF-1 cells, measured by the IFNβ luciferase reporter (data not shown).

Taken together it is unlikely the mutation discovered in the NS1 protein was responsible for the LP phenotype of the original RG virus generated by Howard et al. Only the addition of the 3 HA mutations produced a significant change in pathogenicity caused by the virus.

Table 4. 50-92 viruses were generated by reverse genetics and pathogenicity determined by the IVPI test in ten 6-week old chickens

<table>
<thead>
<tr>
<th>RG virus</th>
<th>HA genotype</th>
<th>IVPI*</th>
<th>New name</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-92</td>
<td>A160, E193, HA2G4</td>
<td>0.02</td>
<td>50-92-LP</td>
</tr>
<tr>
<td>50-92160T,193K,4R</td>
<td>160T, 193K, HA24R</td>
<td>2.88</td>
<td>50-92-HP</td>
</tr>
</tbody>
</table>

*A high path virus has an IVPI score ≥1.20.

From this point on, 50-92 (A160, E193, HA2G4) and 50-92160T,193K,4R are termed 50-92-LP and 50-92-HP, respectively, to reflect their phenotype defined by IVPI.

2.2.4 50-92-HP replicates more efficiently than 50-92-LP in avian cells

During egg passage it was observed that 50-92-LP virus had a slower death rate in embryonated chicken eggs compared to 50-92-HP virus, similar to observations by Banks et al. (Table 2) (data not shown). This suggested an increased lethality of 50-92-HP in ovo. To further examine this observation, we performed multi-cycle virus growth kinetics experiments on primary chicken embryonic fibroblast (CEF) cells. Although both viruses grew well and caused cytopathic effect, 50-92-HP demonstrated higher rates of replication than 50-92-LP, statistically significant at 24 hpi (Figure 15).
Figure 15. 5092-HP replicates more efficiently than 5092-LP in Chicken Embryonic Fibroblasts

Chicken Embryonic fibroblasts (CEF) cells were infected with either 5092-HP or 5092-LP (MOI 0.0001 by pfu/ml). vRNA was extracted from the cell supernatant and quantified by Q RT-PCR using M gene specific primers. (REU = Relative EID\textsubscript{50} unit). Two tailed t-test (Holm-Sidak method) * p<0.01.

2.2.5 HA mutation HA\textsubscript{2} G4R is required for restoration of the HP phenotype \textit{in vivo}

In addition to generating 50-92-LP and 50-92-HP we also attempted to generate other combinations of the HA mutations, utilising the engineered HA plasmid panel. We failed to generate the single point virus mutants 50-92\textsubscript{160T} and 50-92\textsubscript{193K}, despite multiple attempts (Table 5). vRNA was extracted from egg allanotic fluid of each virus genotype and HA mutations confirmed by sequencing.
Table 5. IVPI score of 50-92 HA mutant viruses

<table>
<thead>
<tr>
<th>Virus</th>
<th>HA mutations</th>
<th>Successful generation?</th>
<th>IVPI Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-92-LP</td>
<td>A160 E193 HA2 G4</td>
<td>✓</td>
<td>0.02</td>
</tr>
<tr>
<td>50-92_{160T}</td>
<td>160T E193 HA2 G4</td>
<td>× (n=2)</td>
<td>ND</td>
</tr>
<tr>
<td>50-92_{193K}</td>
<td>A160 193K HA2 G4</td>
<td>× (n=2)</td>
<td>ND</td>
</tr>
<tr>
<td>50-92_{4R}</td>
<td>A160 E193 HA2 4R</td>
<td>✓</td>
<td>ND</td>
</tr>
<tr>
<td>50-92_{160T,193K}</td>
<td>160T 193K HA2 G4</td>
<td>✓</td>
<td>0.68</td>
</tr>
<tr>
<td>50-92_{160T,4R}</td>
<td>160T E193 HA2 4R</td>
<td>✓</td>
<td>ND</td>
</tr>
<tr>
<td>50-92_{193K,4R}</td>
<td>A160 193K HA2 4R</td>
<td>✓</td>
<td>ND</td>
</tr>
<tr>
<td>50-92-HP</td>
<td>160T 193K HA2 4R</td>
<td>✓</td>
<td>2.88</td>
</tr>
</tbody>
</table>

ND= Not Done  
HPAI= IVPI≥1.20

Viruses were generated by the standard RG system method with one passage in MDCK cells and IVPI assay performed for 50-92-LP, 50-92_{160T,193K} and 50-92-HP. This table includes the results previously shown in Chapter 2 Table 4 (IVPI scores for 50-92-LP and 50-92-HP).

50-92-LP and 50-92-HP were tested for their pathogenicity in chickens by the IVPI test and scored 0.02 and 2.88, respectfully. In order to determine which of the three amino acids were required to induce the HP phenotype, we also intravenously infected chickens with 50-92_{160T,193K} virus. Interestingly, although some clinical signs were observed, the virus retained its LP phenotype scoring 0.68 with the greatest onset of clinical signs at 6 dpi (Figure 16A and Table 5). None of the chickens infected with 50-92_{160T,193K} or 50-92-LP succumbed to infection. Virus replication was measured by quantifying vRNA present in oropharyngeal and cloacal swabs. In addition to increased clinical signs, greater levels of vRNA for 50-92_{160T,193K} virus were observed compared to 50-92-LP, for which negligible vRNA was detected from oropharyngeal or cloacal swabs at 1dpi or 11dpi (Figure 16 B&C). However, significantly higher levels of vRNA were detected for 50-92-HP over both 50-92-LP and 50-92_{160T,193K}, with all birds succumbing to infection by 2dpi with severe clinical signs. Taken together the data suggested HA2 4R was necessary for the HP phenotype in chickens. The results did not answer whether HA2 4R is capable of inducing a HP phenotype without mutations 160T and 193K, which themselves did not significantly increase pathogenicity. However the increase in virus titre in cloacal swabs at 1dpi for the double mutant virus and
the intermediate IVPI score indicate that these two mutations may contribute to the HP phenotype.

![Image of a graph showing IVPI scores over time for 50-92-HP, 50-92_{160T,193K}, and 50-92-LP viruses.

Figure 16. HA\(_2\) mutation G4R is necessary for restoration of the HP phenotype \textit{in vivo}

The IVPI assay was performed for 50-92-LP, 50-92-HP and 50-92_{160T,193K} virus, whereby ten 6-week old SPF chickens were intravenously injected with >16HA units of allantoic fluid and scored for pathogenicity over 10 days (0 if normal, 1 if sick, 2 if severely sick, 3 if dead). Oropharyngeal and cloacal swabs were taken from 5 randomly selected birds 1dpi and 11 dpi (unless dead). Viral RNA was extracted and quantified by qRT-PCR of the M gene, displayed as log10 Relative EID\(_{50}\) units (REU). A. Plot of clinical score over 10 days. B and C. Oropharyngeal and Cloacal REU titres respectively. Error bars displayed as SEM. † all birds infected with 50-92-HP were dead by 2dpi (shown in red). Statistical analysis was by two-way ANOVA with Bonferroni's multiple comparisons test. * p>0.05.
2.2.6 Investigation of virus tropism by *in ovo* influenza infection

Influenza viruses are frequently propagated using embryonated chicken eggs. LP viruses (with a single basic HA cleavage site) are inoculated in the allantoic fluid and replicate in the chorio-allantoic membrane. Viruses with a MBCS however, are able to traverse the allanotic membranes and infect tissues throughout the embryo (361). Figure 17 displays a H&E section of an un-infected 9-day old chicken embryo, with the differentiated tissues clearly visible.

![Figure 17. H&E longitudinal section of a non-infected 9-day old chicken embryo](image)

Embryos were decapitated and fixed in formalin for a minimum of 5 days. H&E stain sections were generated by the Pathology team at AHVLA, Weybridge. Images were captured using an electronic scanner.

We sought to use the *in ovo* model of influenza infection to detect possible cell tropism differences between 50-92-LP, 50-92-HP and 50-92*\textsubscript{160T,193K}*, which may contribute to the IVPI phenotypic differences. We hypothesised mutations 160T and 193K may alter tissue tropism due to their location in the HA RBS, and because together they slightly increased
virus shedding detected 1dpi (Figure 16 B&C). 9-day old chicken embryos were inoculated via the allantoic route with $10^2$ pfu of either virus (n=2). Embryos were analysed for evidence of infection by the detection of influenza NP protein at 24 and 48hpi. As previously observed, embryos infected with 50-92-LP outlived those infected with either 50-92-HP or 50-92$_{160T,193K}$ virus, although by 48hpi embryos infected with 50-92-LP demonstrated vascular constriction on the interior of the egg shell, typical of sick and dying embryos infected with HP influenza virus (data not shown). Embryos infected with either 50-92-HP or 50-92$_{160T,193K}$ were dead at the point of harvest. During the dissection allantoic fluid samples were taken and vRNA was quantified by qRT-PCR. At 24hpi all three viruses replicated to high titres and mean REU values were comparable (1.02x10$^7$ for 50-92-LP, 8.98x10$^6$ for 50-92-HP and 3.28x10$^7$ for 50-92$_{160T,193K}$ (n=3)). Values at 48hpi could not be readily obtained, due to the deterioration of the allantoic fluid caused by virus infection.

Images of the whole embryo demonstrated the widespread dissemination of all three viruses by 48hpi with significant levels of immunolabelling and no gross differences between the viruses (Figure 18).
Figure 18. Dissemination of influenza virus 50-92 in the chicken embryo detected by NP staining

9-day old embryonated chicken eggs infected with $10^2$ pfu of either 50-92-LP, 50-92-HP or 50-92160T,193K virus, and mock (virus diluents only) (n=2). Embryos were culled (if surviving) by decapitation and fixed in formalin before H&E sections being made. Immunohistochemical demonstration of virus NP antigen
was examined. Influenza NP stain (brown). Mayer’s Haematoxylin counter-stain (blue). Brown pigment (melanin) is present in the retinal pigment epithelium of the eye, clearly distinguishable from true NP staining due to its dark appearance and exclusive cytoplasmic location. Positive and negative controls were as expected. Sections were prepared by the Pathology team at AHVLA, Weybridge.

More detailed analysis was carried out by microscopy and the cell tropism of each virus was explored. At 24hpi, both 50-92_{160T,193K} and 50-92-HP virus antigen was found in all of the tissues examined, whereas 50-92-LP was only detected in the allantoic and amniotic membranes and the skin epithelium, despite reaching high viral titres in the allantoic fluid (Figure 19A). This difference may account for the survival of the chick embryos infected with LP virus. By 48hpi, all three viruses were detected in abundance throughout the embryo. Virus antigen was present in the ectoderm, endoderm and surrounding the blood vessels of the allantoic membrane for all viruses. Allantoic membrane was not recovered for 50-92-LP virus at 48hpi, however, NP positive cells can be seen attached to the allantoic matrix (Figure 19A&B). The allantoamniotic membrane was positive for NP on both the allantoic and amniotic epithelium, suggesting a route of possible spread from the allantoic to the amniotic cavity. At 24hpi NP staining of the organs by 50-92-HP and 50-92_{160T,193K} was predominantly of the blood vessel endothelium. By 48hpi virus antigen from all 50-92 viruses was widespread in endothelial cells in multiple organs and also observed in parenchymal cells of organs. There was no observable difference in tissue tropism between the three viruses in the different organs. At 48hpi immunolabelling was observed in lymphohistiocytic cells and endothelial cells. In particular the spleen, a major lymphoid organ of avian species, was heavily stained, suggesting a strong preference of 50-92 virus for immune cells. In addition to virus antigen, inflammation and/or necrosis of tissues could be observed, for example the liver and spleen at 24hpi is clearly inflamed for 50-92-HP and 50-92_{160T,193K} compared to 50-92-LP where no antigen was detected. Deterioration of embryo tissues was clearly identified by 48hpi suggesting necrosis as a result of virus infection and/or autolysis after death (Figure 18 and 19B). Mock infected tissues did not demonstrate tissue inflammation or necrosis (Data not shown).

The embryo model did not detect any differences in cell tropism between the different 50-92 viruses. Nonetheless these data suggested the dissemination of 50-92-LP virus through the embryo was reduced. Despite this, 50-92-LP was able to reach comparable levels of infection by 48hpi.
Figure 19A. Exploration of virus tropism in the chicken embryo at 24hpi
Figure 19B. Exploration of virus tropism in the chicken embryo at 48hpi
2.2.7 Investigation of virus replication of the 50-92 HA mutants *in vitro*

To further examine the effect of the HA mutations on 50-92 virus replication, virus growth curves were performed with the entire panel of 6 RG viruses on CEF cells at 41°C, representative of the avian body temperature. Viral RNA was extracted from supernatant and quantified by qRT-PCR. As previously shown in figure 15, 50-92-HP displayed a replication advantage over 50-92-LP in CEF cells (Figure 15). Indeed 50-92-HP replicated at statistically significant higher levels than any other virus HA genotype at 20hpi, particularly 50-92-LP and 50-92_{193K,4R} (Figure 20). In addition, 50-92-LP and 50-92_{193K,4R} replicated at significantly lower rates than 50-92_{160T,193K} and 50-92_{4R}.

![Graph showing virus replication comparison](image)
Figure 20 (Previous page). Replication of 50-92 HA mutants in CEF cells

CEF cells were infected with either 5092-HP or 5092-LP (MOI 0.0001 by pfu/ml). vRNA was extracted from the cell supernatant and quantified by Q RT-PCR using M gene specific primers. (REU= Relative EID$_{50}$ unit). Two way ANOVA, multiple comparisons Bonferroni’s post-test. Statistical analysis revealed significance only at 20hpi, displayed in the tables above. * p<0.01, ** p<0.001, **** p<0.00001.

Interestingly, the small differences observed by multi-cycle growth curves were somewhat reflected by the analysis of plaque size in MDCK cells. 50-92-LP appeared to have a slightly smaller plaque phenotype than 50-92-HP, 50-92$_{160T,4R}$ and 50-92$_{160T,193K}$ (Figure 21A+B). 50-92$_{193K,4R}$ virus had the smallest plaque phenotype. Comparison of the different plaque sizes does not highlight one HA mutation alone as capable of increasing or decreasing size, since 193K decreased plaque size in combination with HA$_2$ 4R but not 160T, whereas HA2 4R alone did not alter plaque size. Taken together the plaque phenotypes suggested a mechanism whereby the mutations act together to alter the replication of the virus shown by growth kinetics and plaque size. Lack of all three mutations or the combination of 193K and HA$_2$ 4R are the most detrimental to 50-92 virus.
Figure 21. Plaque phenotype of 50-92 viruses in MDCK cells

MDCK cells were infected with a dilution series of either 50-92-HP, 50-92-LP, 50-92_{160T,4R}, 50-92_{193K,4R}, 50-92_{193K}, 50-92_{4R} in the absence of trypsin with avicel plaque overlay media in 12-well plate format. After 3 days incubation at 37°C, cells were stained with crystal violet.

2.2.8 Mutation HA\textsubscript{2} G4R alters the pH of fusion of 50-92 HA protein

Due to the location of the HA\textsubscript{2} G4R mutation in the fusion peptide we hypothesised that the pH of fusion was altered and that this could account for the increased pathogenesis in the chicken. We investigated the pH at which the HA protein triggered membrane fusion by a haemolysis assay. 64HA units of each 50-92 RG virus was incubated with chicken red blood cells (cRBCs) at 4°C. After binding of virus particles by HA to SA on cRBCs, we incubated the mixture with a low pH buffer to trigger the HA conformation change and induce fusion. Successful fusion was measured by the release of haemoglobin from the erythrocytes. 50-92-HP and 50-92-LP were able to fuse more readily at pH 4.75 than 50-92_{160T,193K} (Figure 22). This suggests mutations 160T and 193K reduced the pH of fusion which was compensated by the HA\textsubscript{2} 4R mutation present in 50-92-HP virus. The levels of
fusion were lower than expected at the higher pH values (> pH 5.0), possibly due to the effect of the low pH on mock infected cRBCs which made it difficult to detect virus specific lysis.

![Graph showing pH vs % Haemolysis](image)

**Table 1**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-92-HP vs. 50-92_{160T,193K}</td>
<td>***</td>
</tr>
<tr>
<td>50-92-LP vs. 50-92_{160T,193K}</td>
<td>**</td>
</tr>
</tbody>
</table>

**Figure 22. Fusion peptide mutation HA₂ 4R compensates for the decreased fusogenicity by 160T and 193K, measured by virus haemolysis**

64 HA units of each virus was mixed with cRBCs. After incubation for 30mins at 4°C, the mixtures were pelleted and washed in PBS. This was replaced by MES buffer at varying pH. After 30mins incubation at 37°C, fusion was measured by absorbance at 450nm. Values were normalised to those obtained following exposure of unagglutinated red cells to each pH (mock adjusted) and presented as mean % Haemolysis (100% = pH 4.5). Error bars are SEM. Statistical analysis was by two-way ANOVA with Bonferroni’s post-test. ** p< 0.01, ***p< 0.001.

These data with whole 50-92 virus demonstrated significant variability, likely due to the technical constraints working in a high containment laboratory. To this end, we sought to use transient expression of HA as a means to further examine the effect of the HA mutations on pH of fusion. The 50-92 HA constructs were sub-cloned into an expression plasmid, PCAGGS. Human 293T cells were transfected to express the HA proteins and after
24 hours, cell monolayers were washed at varying pH to trigger HA fusion, visible by the formation of polykaryon cells. 50-92-LP HA was able to induce fusion up to pH 5.5 (Figure 23). The introduction of 160T or both 160T and 193K reduced the pH of fusion to pH 5.0. Interestingly, as seen in the haemolysis assay (Figure 22), HA 4R mutation was able to compensate this reduction and increase the pH of fusion to pH 5.5 for 50-92-HP. 50-924R did not increase the pH of fusion above that of 50-92-HP, despite lacking 160T and 193T. Unexpectedly, pH 5.25 failed to trigger fusion with any HA construct. The pH of this buffer was repeatedly measured so the significance of this observation remains unclear.

The expression of each HA protein from the pCAGGS plasmid was explored by flow cytometry and IHC to determine if the mutations in the HA were affecting cell surface expression. HeLa cells were transfected with each HA construct and after 24 hours fixed in paraformaldehyde and stained with anti-HA antibodies. Flow cytometry analysis revealed equivalent cell surface expression between the HA mutants, suggesting none of the mutations altered the expression of HA protein. Figure 24A presents data from two independent experiments; on both occasions the proportion of HA positive cells was low, suggesting inefficient transfection of HeLa cells. In addition to flow cytometry analysis, 50-92-LP, 50-92-HP, 50-92160T and 50-92160T,193K HA expression was analysed by IHC in 293T cells. Protein expression was detected for all the constructs (Figure 24B).
<table>
<thead>
<tr>
<th>pH</th>
<th>4.5</th>
<th>4.75</th>
<th>5.0</th>
<th>5.25</th>
<th>5.5</th>
<th>5.75</th>
<th>6.0</th>
<th>HA$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-92-LP</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>160</td>
</tr>
<tr>
<td>50-92$_{160T}$</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>193</td>
</tr>
<tr>
<td>50-92$_{160T,193K}$</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>50-92$_{4R}$</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>A</td>
</tr>
<tr>
<td>50-92-HP</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>E</td>
</tr>
<tr>
<td>Empty Vector</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>E</td>
<td>G</td>
<td>T</td>
<td>K</td>
<td>G</td>
<td>T</td>
<td>K</td>
</tr>
</tbody>
</table>
Figure 23 (Previous page). Fusion peptide mutation HA2 4R compensates for the decreased fusogenicity by 160T and 193K, measured by polykaryon formation

Human 293T cells in 12-well format were transfected with 500ng of either HA plasmid. After 24hours incubation, media was replaced with pre-warmed MES buffer at varying pH and incubated at 37°C for 5 mins. After this time, cells were washed in PBS and media replaced. Cells were incubated at 37°C for 2.5 hours when polykaryon formation occurred. Carried out in triplicate, representative images were selected.

A.

![Graph showing FITC Median and % Positive Cells for different conditions.]

B.

Figure 24. HA mutations 160T, 193K and HA2 4R did not alter cell expression

A. HeLa cells were transfected with 500ng of either HA expression plasmid in 12-well plate format, plus empty vector mock. After 24hrs cells were washed and probed with chicken polyclonal anti-50-92-HA sera (1:100) (AHVLA, Weybridge). HA was indirectly detected by AlexaFlour 488 FITC antibody (1:500). Cells were fixed in 4% paraformaldehyde before flow cytometry was performed. Data from two independent experiments is displayed here as median fluorescence with percentage FITC positive cells. Empty vector, and controls
excluding 1° or 2° antibody gave expected results. Error bars are SEM. Statistical analysis by two-way ANOVA revealed no significant differences. Flow cytometry analysis examples are provided (Chapter 6). B. 293T cells were transfected with 500ng of either HA expression plasmid in 12-well plate format on glass coverslips, including empty vector control. A. 50-92-LP, B. 50-92-HP, C. 50-92_{160T,193K}. D. Mock (empty plasmid). After 24hrs cells were fixed in paraformaldehyde. Cells were incubated with sheep α-HA antibody (1:300) (Vietnam/04, NIBSC) and HA indirectly detected by goat α-sheep antibody conjugated to FITC (1:500) (Merck Millipore).

2.2.9 Bioinformatic analysis of HA 158 glycosylation, 193 and HA_{2} 4 for avian influenza isolates

Since N-linked glycosylation at residue 158 has previously been reported in the literature, we sought to examine the prevalence of the predicted glycosylation at HA residue 158 in avian H5N1 influenza viruses isolated from 1956 to 2013. Overall 42.45% of 2052 isolates were predicted to carry the glycosylation site at 158 (Figure 25A). Almost half the isolates were from chickens and these had a higher rate of glycosylation than viruses from other hosts, 49.35 vs 35.98% respectively. Since viruses belonging to clade 2.2, which emerged in 2005, have reportedly lacked the 158 glycosylation site we separated the data based on the year isolated. It was found that isolates from 1956 to 2004 had a greater proportion of glycosylation at 158 than those between 2005 and 2013, 52.81 and 40.67% respectively. In particular, during 1956 to 2004 chicken viruses had a higher incidence of glycosylation, 64.96% compared to 43.33% for non-chicken derived isolates. Taken together it suggests H5N1 viruses may favour 158 glycosylation in the chicken host, although it is clearly not essential to all H5N1 viruses, particularly in more recent years.
A. Sequences are segregated based on the year of isolation and whether isolated from a chicken host.

B. Sequence identity of aa193 (1956-2013).

C. Sequence analysis of aa HA$_2$ 4 (1956 to 2013). The analysis was performed using full length HA sequences of avian H5N1 viruses from NCBI Influenza Virus Resource database, May 2013. Data are percentages. Sequences were aligned using ClustalOmega EBI resource (284), and analysed in Geneious R6.

Figure 25. Quantitative analysis of a predicted N-linked Glycosylation at HA residue 158, identity at 193 and HA$_2$ 4 for H5N1 influenza viruses isolated 1956 to 2013.
In addition to the glycosylation at 158, we examined the frequency of different amino acids at HA 193 and HA_2 4 in the H5N1 isolates. We found HA 193 was highly conserved with a preference for a positively charged amino acid, Lys (62.13%) or Arg (35.82%), suggesting an important function for H5N1 HA protein (Figure 25B). Arg was more predominantly observed in recent H5N1 isolates (data not shown). We identified almost total conservation of glycine at HA_2 4, with only one example of alanine at this residue (Figure 25C). Further analysis of the HA_2 sequence (aa4-7) was conducted in all influenza subtypes isolated from 1956 to 2013. Again, the region was highly conserved, with a minority of variations identified (Table 6). A small number of isolates carried a positively charged amino acid in this region with a −GIKA- motif, demonstrating a positively charged residue may be tolerated in this region in nature. Interestingly, only one isolate was found to carry the HA_2 4R mutation, a non-virulent variant of a H6N1 virus (362).

**Table 6. Quantitative analysis of HA_2 sequence 4-7 from avian influenza viruses isolates from 1956 to 2013**

<table>
<thead>
<tr>
<th>HA2 sequence (4-7)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>−GAIA−</td>
<td>99.59 (9498)</td>
</tr>
<tr>
<td>−GIKA−</td>
<td>0.34^a (32)</td>
</tr>
<tr>
<td>−VAIA−</td>
<td>0.03 (3)</td>
</tr>
<tr>
<td>−SAIA−</td>
<td>0.02 (2)</td>
</tr>
<tr>
<td>−GDIA−</td>
<td>0.01 (1)</td>
</tr>
<tr>
<td>−RAIA−</td>
<td>0.01^b (1)</td>
</tr>
</tbody>
</table>

A total of 9537 avian influenza isolates were downloaded from the NCBI Influenza Virus Resource database, May 2013, aligned using ClustalOmega EBI resource (284) and analysed using Geneious R6.

^a A mixture of influenza subtypes carried this genetic signature

^b A/Chicken/Taiwan/2838N/00 (H6N1) (non-virulent) (362)

HA2 residue 1 and any charged amino acids are written in red.
3.3 Conclusions and Discussion

**Restoration of the highly pathogenic 50-92 H5N1 reverse genetics system**

H5N1 50-92 RG virus was successfully generated and demonstrated high pathogenicity in the chicken host in contrast to the previously rescued 50-92 RG LP virus. Three mutations in the HA gene were responsible for the HP phenotypic switch: A160T, E193K and HA\textsubscript{2}G4R which also increased replication of 50-92 virus *in vitro* in CEF cells.

The original outbreak was highly pathogenic and caused high rates of mortality in turkeys, yet previous work by Woods et al. (1994) showed the sequence of the original isolate was the same as the lower pathogenicity clones that had IVPI <2.00 (271). However, since Sanger sequencing was used, we speculate the HP genotype was present as a quasispecies and undetected during this analysis. This minority population may also have been inoculated into the chickens when tested by IVPI and have been responsible for the maximum HP score of 3.00. Indeed, the IVPI test for statutory diagnostics aims in part to capture the presence of such minor HP viral populations. The original 50-92 virus was isolated from turkeys and may have had an advantage to carry the “LP” genotype in the turkey host, perhaps enriching for those mutations. Our analysis has been performed in chickens and it is possible the 3 HA mutations that conferred pathogenicity in the chicken by IVPI may not carry the same phenotype in turkeys. Pathogenic markers between different avian species are poorly understood. A virus such as 50-92 could prove a useful tool in deciphering these differences.

50-92 (derived from original isolate) causes systemic infection and death in chickens, even by intranasal and intraoral inoculation (363). Histological examination by Kobayashi et al. (1996) showed typical HPAI infection with pathogenesis of the heart and brain in particular. This suggests that the high pathogenesis shown by the RG virus in the IVPI test might also be apparent in chickens infected by a more natural route, and the histological analysis supports the severe clinical signs observed by RG 50-92-HP virus.

A second isolate from the 1991 outbreak was taken from the brain of a turkey carcass A/turkey/England/87-92/1991 (87-92). Interestingly this isolate carried the same 193 and HA\textsubscript{2}4 residues as the HP plaque clones of 50-92 but scored an IVPI of 0.00 (271). A
clone derived from 87-92 selected by three limiting dilution passages also scored 0.00. Banks and Plowright (2003) serially passaged this virus in embryonated eggs, selecting a glycosylation at HA residue 160 by the fourth passage, although this virus was still found to have a LP phenotype by IVPI test (364). Subsequent work by Londt et al. (2007) successfully increased the virulence of 87-92 virus by serial intracerebral passage in chickens, selecting additional changes in the HA, PB2, PB1, NP and PA, that restored the sequence of those polymerase genes and NP to that seen in the 50-92-HP virus used in this project. This suggests that mutations throughout the genome may have been present in the quasispecies during the outbreak and many of these likely altered virulence (365).

The finding that the 50-92-LP RG virus displayed a LP phenotype in the IVPI assay has implications for the OIE guidelines on the identification and control of highly pathogenic avian influenza viruses. Currently, both 50-92-LP and 50-92-HP are considered notifiable highly pathogenic viruses (i.e. subject to statutory control) due to the presence of a MBCS alone, despite having distinctly different disease outcomes in the chicken. For example, an outbreak of 50-92-LP virus may go unnoticed, despite carrying a MBCS and clearly having the potential to become HP on mutation. This highlights the necessity to sequence influenza viruses by surveillance to assess their pathogenic potential as well as confirming an already HP outbreak. Understanding the mechanisms of pathogenicity and the biological significance of genetic signatures may help refine such guidelines and risk assessments leading to a proportionate response. The mechanisms behind the generation of HPAI viruses from LP precursors are poorly understood. The mutations described by this work highlight a route to increased pathogenicity by mutations accumulating in HA.

The successful generation of a functioning highly pathogenic avian influenza RG system for 50-92 may be a useful tool for the field. 50-92 has been used as a model avian H5N1 virus in numerous publications by different laboratories (39, 271, 272, 317, 318, 328, 363, 365–370). In addition to understanding the variable pathogenesis of the 50-92 clones, the system may now be used in future research projects.

Limitations and future work

A limitation of this work includes the use of the IVPI index as a measure of pathogenicity. This assay is a certified method by which to define AI as HP or LP, but its use
to determine more subtle differences in pathogenicity are limited. However, the clinical
signs observed in birds infected by RG 50-92-HP (IVPI 2.88) were similar to those expected
and previously seen in chickens infected with WT virus (IVPI 3.00) derived from the clinical
isolate (personal communication with AHVLA staff). The 50-92 RG virus by Howard et al.
based on the 2L genotype scored an IVPI of 0.20 and the 50-92-LP virus generated in this
study with the same HA genotype scored 0.02. These small differences are difficult to
define, although both are LP. Such a difference may be explained by differences in the input
of virus which is roughly required as >8HA units. The original 2L clone scored 1.20 (a HP
score), this difference may be explained by the lack of full genome sequencing conducted on
this virus and the possible presence of mutations elsewhere in the genome affecting
pathogenesis.

Elucidation of the mechanism behind the differing pathogenicity of the 50-92 virus clones

We sought to examine which of the three HA amino acid differences between 50-92-
LP and 50-92-HP were responsible for the pathogenic phenotype switch. We found all three
HA mutations were necessary for HP phenotype and increased virus shedding in the
chicken. 50-92-LP was severely attenuated with negligible shedding detected \textit{in vivo}. The
combination of 160T and 193K alone failed to confer the HP phenotype in chickens,
although there was some evidence that they might have contributed to increased
replication \textit{in vivo}. Mutation HA2 4R was necessary to complete the transformation into a HP
virus. We did not determine whether HA2 4R mutation was capable of the HP phenotype
alone, and whether mutations 160T and 193K were indeed necessary.

Bioinformatic analysis revealed that glycosylation at 158N is often positively selected
in the chicken (Figure 25), as described by others (88, 121, 308, 371). 50-92 virus carries a
short-stalk (ss) NA gene due to a 22aa deletion, and research suggests this is linked with the
acquisition of HA glycosylation at 158N. Baigent and McCauley (2001) illustrated how a HA
lacking glycosylation at 158 restricted growth when combined with a ssNA, due to poor NA
activity, demonstrated by failure to elute bound red blood cells (371). This suggests
glycosylation at the tip of HA is able to improve release from its receptor, especially since
when combined with a long-stalk (Is) NA excessive NA activity was observed which was
detrimental to virus growth. The optimal conditions were described as glycosylation at 158 with a ssNA and lack of glycosylation with a lsNA, but which drives which in terms of evolution in poultry is not clear (371). Imai et al. (2012) suggest glycosylation at 158N for a H5 virus can sterically block receptor binding due to its close proximity (87). HK/97 HSN1 viruses also demonstrated reduced binding affinity to SA and increased NA elution in presence of the 158 glycosylation, which likely compensated for the ssNA of these viruses (121). Similar research has been conducted for H7 virus, where removal of a glycosylation on the HA head restricted growth and resulted in a smaller plaque size. The HA:NA balance was also affected and NA reassortant viruses showed this was subtype specific; N1 reduced elution from bound red blood cells to a greater extent than N2 when combined with the mutant HA, suggesting N2 was partly able to recover the detrimental high-affinity binding by the HA mutant (372). Taken together this suggests the NA short stalk of 50-92 selected for during passage in poultry upset the HA:NA balance which was then restored by mutation 160T in the 50-92-HP virus. Questions remain as to why poultry select a ssNA; what is unique about the presentation of SA receptors in the chicken compared to duck or human?

Despite the presumed effect of glycosylation to decrease binding to sialic acid, preliminary results comparing HAU and pfu/ml titres of 50-92-HP and 50-92-LP have suggested greater binding of 50-92-HP as less pfu were required to match the HA titre of 50-92-LP (data not shown). Although this preliminary finding does not elucidate which of the 3 amino acids may be responsible, we might reasonably speculate mutation 193K increases SA receptor binding due to its location in the 190-helix of the RBS, although we have not demonstrated this experimentally. Bioinformatic analysis showed the vast majority of HSN1 viruses have a positive charged amino acid at residue 193, suggesting a strong preference (Figure 25).

193K enabled the increased binding of equine H3 to α2,3 SA, discussed previously in the introduction (350). Thus 193K and gly158 may be jointly involved in receptor binding, potentially interacting in opposition. Intriguingly, for 50-92 virus, the 193K mutation was detrimental for 50-92 virus in combination with HA2 4R when in absence of 160, resulting in reduced growth and a smaller plaque phenotype than 50-92-LP (Figures 20&21). We speculate this may be due to reduced virus release, since 193K may increase binding, and the lack of a glycosylation site to sterically block SA binding with a ssNA results in low NA
activity and reduced virus release. In addition, despite multiple attempts, virus carrying the 193K genotype alone could not be generated. 50-92_4R virus was able to grow efficiently and had an increased plaque size, suggesting 193K was the restrictive factor for 50-92_193K,4R virus.

It is unlikely that 193K and 160T switched receptor binding preference from avian α2,3 to human α2,6. Interestingly, chicken and other gallinaceous species express more α2,6 linked SA compared to ducks (373–375). No differences in cell tropism were observed in embryos infected with either virus which may be expected if receptor binding had altered (Figure 18). Taken together this may suggest it is the overall ability to bind that has altered rather than a switching of receptor preference.

Receptor binding mutations have previously been described to alter the stability of the HA. The H5 HA of Vietnam/04 was adapted for ferret respiratory droplet transmission by Imai et al (2012) (87). Interestingly, the introduction of receptor binding mutations N224K/Q226L increased the pH of fusion compared to WT HA, shown by polykaryon formation by expression of HA in HeLa cells. Removal of a glycosylation site, N158D, did not appear to alter pH of fusion further. To compensate this change in pH a fourth mutation was selected, T318I which stabilized the HA, and decreased pH of fusion either alone or in combination with the receptor binding mutations. Since human viruses have been shown to require a lower pH of fusion than avian influenza virus, the T318I mutation was required to correct the negative effect of the receptor switching mutations. A similar set of mutations was described by Herfst et al. (2012) which included a HA mutation likely acted to alter pH of fusion, H103Y, although this was not confirmed in the study (88). Clearly all three sets of mutations were required together to adapt 50-92 HA for increased pathogenicity in poultry, and the function described by each of these mutations is remarkably similar to the mutations described for the two H5 viruses that were adapted to ferrets (87, 88).

Introducing either 160T or 193K resulted in a reduced pH of fusion compared to 50-92-LP. As we have discussed, it is likely these mutations are required for efficient binding and release in chicken cells. The pH of fusion was recovered by the HA_2 G4R mutation, to a level more usual for avian viruses, around pH 5.5 (Figure 23) (92–94). The increased fusion measured by haemolysis by whole virus bound to red blood cells required a lower pH, 4.75 (Figure 22), although this assay was technically challenging in the high containment
environment and presence of NA (on the virus as opposed to using an assay where HA was expressed alone from a plasmid) has been shown affect the measured pH of fusion of avian HA viruses (92).

In a previous study, a plaque clone (13L) that carried only HA$_2$ 4R in its HA was highly pathogenic and scored an IVPI of 2.36, which was lower than the clones with all three mutations (3L and 11L) scoring 3.00 (Chapter 1, Table 2). This suggests mutations 160T and 193K can contribute to 50-92-HP virulence, but a worthy experiment in the future would be to perform the \textit{in vivo} experiment with a reverse genetics generated 50-92$_{4R}$ virus to confirm the findings by Wood et al. (1994) (271). Thus we speculate all three mutations were vital for efficient replication and the combination determined the increased virulence in the chicken.

We are not aware of previous descriptions of the effects of the HA$_2$ 4R mutation in the literature, yet Steinhauer et al. (1995) demonstrated the effects of a G4A mutant (376). G4A increased the pH of fusion by 0.4 pH units although this was much less efficient at fusion overall. In addition Korte et al. (2001) revealed that the removal of negative charge in the fusion peptide resulted in deeper penetration into the host cell membrane (377). An observation such as this may suggest the introduction of a positive charge could counteract the negative charge of the peptide and result in deeper penetration, allowing fusion to occur at a higher pH. For HA$_2$ 4R destabilising the HA in this manner may have counteracted the stabilising 160T and 193K mutations.

We speculate that 193K was required for efficient binding together with glycosylation at 158, a common poultry adaptation, and these mutations increased HA stability which was disadvantageous for the virus. Mutation HA$_2$ 4R was required to correct the balance and restore the threshold of pH to that necessary for avian viruses. Since we have not seen 4R in nature, we describe it as a novel virulence marker and highlight the importance of 193K and glycosylation at 158 for H5 viruses in the chicken.

It seems there are many residues that can alter pH of fusion by the HA, which may explain why we have not seen this particular mutation occur in wild birds and other HPAI outbreaks. Our pathogenesis model has not measured transmission. The HA$_2$ 4R mutation may render the HA unstable in the environment, preventing or reducing transmission of the
virus. Such mutations might be selected for in densely populated poultry houses where distance between animals is very short. An in vivo transmission model of chickens infected with 50-92-HP could address this hypothesis.

These data provide useful information for understanding virulence mechanisms of H5N1 viruses for poultry and contribute to the concept of pH as a virulence factor and interspecies host barrier for influenza virus.

**Limitations and future work**

Although the plaque sizes of several of the 50-92 RG viruses differed, additional plaques could be measured and compared to ensure a reliable comparison between all viruses.

Expression of HA examined by FACS analysis in HeLa cells was low. In order to reliably test the expression of each HA construct, this should be repeated with a more transfectable cell line such as 293T cells.

The bioinformatic analysis carried out in this study relies on the previously acknowledged association between increased glycosylation of H5 HPAI and the chicken host. In order to independently prove an association between glycosylation and chicken isolates, a Monte Carlo analysis on randomly selected amino acids on the HA would enable statistical significance to be applied.

Examination of the pH of HA fusion faced technical limitations in high containment and exogenously expressed HA in absence of other influenza proteins limits the interpretation of the data. The generation of single basic HA 50-92 viruses may be a useful tool to examine pH of fusion at a level of containment less straining on technical abilities. This was attempted during the study and interestingly the HA engineered to be SB failed to generate virus by RG virus rescue (data not shown). The engineered cleavage sequence may require further mutations to be viable.

In order to evaluate these findings further, a more comprehensive analysis of the effect of fusion should be carried out. By collaboration pseudotyped lentiviral vectors carrying the different 50-92 HA proteins will be generated and used to measure the fusion
of pH in combination with cellular pH altering drugs. If possible, the pathogenicity of the RG virus with the HA$_2$4R genotype will be determined by IVPI index. This may indicate whether the 160T and 193K mutations are ultimately required for increased pathogenicity or if HA$_2$4R is capable of generating a HP phenotype alone.
Chapter 3. The emergence of PB2 627K in the avian influenza virus reservoir

3.1 Introduction

An unprecedented outbreak of H5N1 HPAIV in wild aquatic birds occurred in 2005 at Qinghai Lake, China (181–183). Descendants of these Qinghai-like viruses spread into Europe and Africa thus establishing the Eurasian lineage Z viruses (clade 2.2) (118, 378–381). Strikingly, the Qinghai lake isolates and descendants possessed Lys 627 on the PB2 protein, a motif previously associated with adaptation to mammals (182, 183, 199). In contrast, H5N1 viruses isolated from avian hosts prior to 2005 typically possessed a Glu at position 627, emblematic of the majority of other avian influenza virus subtypes. Bioinformatic and biological studies have highlighted PB2 627 as a key adaptation marker of avian viruses to human hosts (301–308).

3.1.1 Molecular basis of PB2 627E restriction in mammalian cells

Avian derived influenza polymerases bearing PB2 627E are restricted in mammalian cells including cells of human, murine, primate and porcine origin (292, 315–318). Despite recent developments, the exact mechanism for this host-specific restriction by PB2 627 remains unclear.

In 2008 it was suggested an inhibitory factor in human cells restricted 627E polymerase activity by preventing binding of the polymerase to NP, preventing efficient vRNP assembly (319). However, more recent work concluded the restriction of 627E polymerase activity in human cells was due to the lack of a positive co-factor and that this factor was present in avian cells (317). In both these studies the approach was to fuse human and avian cells and measured polymerase activity of polymerases bearing either 627E or 627K by species-specific polymerase reporters. The latter study from our laboratory showed no evidence of a mammalian inhibitory activity, as when human cells were fused with avian cells, 627E activity was not reduced. Conversely, 627E activity was increased in human cells fused with avian cells. Explanations for the discrepancy between the two studies might be the exact strain of virus used to generate the polymerase complexes, or
difference in the origins and cell passage history of the cells used that might affect the relative expression of different host factors. In addition, the methods of heterokaryon fusion used were different and this may have up or down regulated the expression of various host factors that then influenced the polymerase assay. It remains likely that several host factors could impact on the influenza polymerase function and affect the observed host range restriction. The PB2-binding cellular factor, DEAD-box RNA-helicase DDX17/p72 has been shown to increase viral mRNA and vRNA synthesis in a 627-specific fashion (313). In human cells, exogenous expression of DDX17 increased the activity of the polymerase harbouring 627K whereas activity of the 627E mutant remained unchanged. In addition, using short interfering RNA (siRNA), knockdown of the chicken DDX17 equivalent was shown to decrease avian 627E polymerase activity in chicken cells 3-fold. A 2-fold reduction was observed when the polymerase carried 627K (data not shown in paper), which suggests DDX17 acts independently of PB2 627 in avian cells since both residues at 627 had a similar outcome. Despite the identification of this factor, the mechanism by which availability of DDX17 affects the polymerase has yet to be shown.

3.1.2 The biological consequence of the 627K mutation

As discussed in Chapter 1, the effect of 627K can be cold-adapting. The temperature of the upper respiratory tract in humans is 33-35°C, as compared to the avian intestinal tract temperature around 38-41°C. The ability to replicate at these sites may have a significant advantage in terms of transmission, indeed it has been shown that 627K increased virus transmission between guinea pigs (322). PB2 627K was one of the few mutations necessary to support H5N1 respiratory droplet transmission in the ferret host (88). The sensitivity to temperature may simply represent the biochemical basis of the 627K mechanisms, yet it highlights the biological importance of this mutation to overcome the physiological and anatomical constraints provided by a foreign host.

With regard to H5N1 HPAIV infection, the acquisition of PB2 627K by avian H5N1 viruses has been shown to be important for infection of the human and mouse host (295, 327, 382–387). An advantage conferred by 627K has been seen in other influenza subtypes both in the field and experimentally (292, 388, 389).
The relationship of PB2 627K and adaptation in original Asian H5N1 outbreak of 1997 is less clear. HK/97-like viruses replicated significantly better with higher mortality in mice when they carried the 627K mutation (294, 295, 382) and so 627K permitted a more virulent virus. This correlation was not so clear in human cases since the 627K mutation was not always selected and was not always necessary for virulence (294). Descendants of these viruses infected and killed big cats: leopards were infected with viruses bearing 627K, whereas tigers were not, the virus carried the WT 627E signature. In a later outbreak tigers were killed by a H5N1 virus that selected 627K. This ambiguity may indicate an unknown factor restricting the selection of 627K in some particular virus strains.

Interestingly, viruses from the 2009 H1N1 pandemic had PB2 627E but carried other PB2 mutations, 590-91SR and 271A, that appeared to compensate for PB2 627E and adapt them for efficient replication in mammals (296, 307, 390). In addition, not all mammals tend to preferentially select the 627K motif. Viruses circulating in horses and swine often to not carry PB2 627K, but retain the avian 627E signature (391). Indeed, comparisons between swine viruses containing the cognate swine PB2 or foreign avian PB2 bearing either 627K or 627E in the pig model, revealed 627E decreased pathogenesis in the context of a swine-adapted PB2, but not for a virus that retained an avian PB2 (293).

3.1.3 The implication of the 627K mutation for the avian host

The implication of the PB2 627K mutation for the avian host is little understood. Previously it has widely been considered irrelevant following in vitro analysis. Polymerase function and virus growth of PB2 627E/K viruses in avian derived cells have demonstrated no host-specificity (321, 327, 382). However, recent work by Schat et al. 2012 suggested a small fitness cost associated with H5N1 PB2 627K in the avian host (392). The biological significance of this observation remains unclear as this single piece of work used a human H5N1 isolate that might contain other genetic adaptations and the effect was only apparent in a genetically modified virus background from which the multi-basic cleavage site in HA was removed. In a separate study Bogs et al (2011) studied the reversion of PB2 627E to the K genotype using a Eurasian lineage clade 2.2 virus in mammalian cells. They found that in MDCK cells this occurred frequently but only in the presence of the cognate NP protein.
contrast, reversion from E627 to K was not observed for any genotype constellation in avian cells (327).

Similarly to mammals, certain avian species have also been shown to select the PB2 627K mutation. These species belong to the Ratitae; namely the ostrich, emu and rhea. Viruses isolated from these species have often carried the PB2 627K or 701N mutation or alternatively the 590/591 mutations that may act in a similar manner, and positive selection was recorded in experimentally infected ostriches (393). It has been suggested that these avian species share genetic similarities with their mammalian cousins. Therefore, the relevance of this finding to the emergence and maintenance of PB2 627K in the avian influenza reservoir is questionable, since Ratitae are genetically distinct from other avian taxa (394). Figure 26 describes the genetic distance of Ratitae from Galloanserae, which contains examples of species known to contribute to the maintenance of the avian influenza reservoir, such as wild ducks (395).

![Figure 26. Phylogenetic tree of birds](image)

The phylogenetic tree is based on documented literature. Modern birds are classified into three large taxa (Ratitae, Galloanserae and Neoaves), here each is labelled with a common example. Adapted from Suzuki et al. 2013 (394).
We hypothesised that the PB2 627K mutation of a Eurasian H5N1 virus would have little effect in the avian host, based on previous work suggesting there is no PB2 627 host restriction demonstrated in avian cells, except those of the Ratitae. To this end, we planned to generate viruses carrying either the 627E or K mutations and measure their effects in vivo.

### 3.2 Results

To further our understanding of the effect of the PB2 627K mutation in the avian host, we used reverse genetics to generate viruses bearing either PB2 627K or 627E. As an example of a historical pre-Asian lineage HPAI H5N1 bearing the typical avian signature PB2 627E, we used 50-92 (A/turkey/England/50-92/1991), reportedly restricted to Galliformes (270, 366). We compared this to a modern Eurasian lineage virus (HA clade 2.2) Ty/05 (A/turkey/Turkey/05/2005) naturally bearing PB2 627K, known to infect a variety of avian hosts including Anseriformes (396). The converse mutations were engineered resulting in an isogenic pair of viruses for each lineage differing only at position 627 in the PB2 protein.

### 3.2.1 Prevalence of the PB2 mutation 627K in the avian population.

To quantify the frequency of different amino acids at PB2 627 for avian influenza isolates, all full length PB2 sequences from 1956 to 2013 were downloaded from the NCBI Influenza Virus Resource database and analysed. As expected, the majority of isolates (95.2%) carried 627E (Table 7). Surprisingly, 0.78% carried 627V, previously reported for several H9N2 avian viruses isolated in Israel (397), along with two swine H1N1 isolates (307). 4% of isolates carried 627K, and the majority of these were identified as belonging to the H5N1 subtype. In addition, 627K was identified in several other influenza subtypes (H1N1, H5N2, H6N8, H7N1, H7N3 and H9N2), including those from ostrich, emu and rhea, hosts known to select the 627K mutation (Table 7).

We next sought to further examine residue 627 for the HPAI H5N1 subtype, given the higher proportion of H5N1 isolates carrying 627K (Table 8). Of the two available PB2 sequences of HPAI H5N1 prior to 1996 (A/chicken/Scotland/1959 and A/turkey/England/50-92/1991), neither harboured the PB2 627K mutation.
Table 7. Quantitative analysis of the PB2 mutation 627K in avian isolates from 1956 to 2013

<table>
<thead>
<tr>
<th>Avian Influenza A virus isolates 1956-2013</th>
<th>Position 627 %b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total isolates (multiple subtypes) (6698)</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>95.2 (6375)</td>
</tr>
<tr>
<td>K</td>
<td>4.0 (268)</td>
</tr>
<tr>
<td>Other</td>
<td>0.8 (55)</td>
</tr>
<tr>
<td>Total H5N1 subtype (1041)</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>74.8 (795)</td>
</tr>
<tr>
<td>K</td>
<td>23.1 (246)</td>
</tr>
</tbody>
</table>

Non-H5N1 subtypes carrying PB2 627K (22):

- A/rhea/TX/39923/1993(H5N2)*
- A/emu/TX/39924/1993(H5N2)*
- A/emu/TX/25414/1995(H7N2)*
- A/ostrich/South Africa/9508103/95(H9N2)*
- A/ostrich/Italy/2332/00(H7N1)*
- A/ostrich/South Africa/Al2114/2011(H5N2)*
- A/ostrich/South Africa/Al2512/2011(H5N2)*

*host known to positively select PB2 627K

- A/duck/Hong Kong/702/1979-quail adapted(H9N2)
- A/turkey/Kansas/4880/1980 (H1N1)
- A/turkey/NC/17026/1988(H1N1)
- A/turkey/NC/19762/1988(H1N1)
- A/laughing gull/New Jersey/276/1989(H6N8)
- A/laughing gull/New Jersey/276/1989(H6N8)
- A/chicken/PA/35154/1991(H1N1)
- A/chicken/NY/21665-73/1998(H1N1)
- A/chicken/Italy/5093/99(H7N1)
- A/chicken/Jiangsu/7/2002(H9N2)
- A/aquatic bird/Korea/MA81K/2007(H5N2)
- A/aquatic bird/Korea/MA44/2007(H7N3)
- A/chicken/Guangdong/V/2008(H9N2)
- A/northern pintail/Interior Alaska/10BM08540R0/2010(H4N6)

---

a Influenza A full-length PB2 sequences of avian origin were downloaded from the NCBI Influenza Virus Resource database, May 2013. The position of PB2 627 was analysed using Excel and Geneious5.6.
b Data are percentages. The number of sequences is shown in brackets.
c Other = (A=1, G=2, V=52)

To determine the prevalence of the PB2 627K mutation in HPAI H5N1 viruses isolated between 1996 and 2013 from avian and human hosts in Asia, Europe and Africa, viruses were first assigned to the WHO H5 nomenclature by building a H5N1 HA neighbour joining tree and then identifying the corresponding PB2 mutations (Figure 27). This analysis demonstrated a high prevalence of the PB2 627K mutation in clade 2.2 viruses compared to other H5N1 isolates (Figure 27 and Table 8). This observation is in accordance with the paper from Russell et al. (2012) (308). Human isolates displayed a higher proportion of PB2
627K compared to avian isolates, indicative of the potential for positive selection of this residue in the human host. However, the majority of the human isolates retained the 627E motif (Table 8).

![Phylogenetic tree](image)

**Figure 27. Phylogenetic tree demonstrating the prevalence of the 627K mutation in the H5 clades**

1325 H5N1 PB2 and 2392 HA sequences were downloaded from the NCBI Influenza Virus Resource April 2013 from avian, human and other hosts. 59 additional human H5N1 genomes and one avian genome were also obtained through GISAID. The nucleotide sequences were aligned by codon using Muscle in MEGA 5.05 and manually adjusted. HP isolates were identified by multi-basic cleavage site on the HA protein sequences. A neighbour-joining tree was used to identify the virus clades in HA according to the WHO H5 nomenclature [http://www.who.int/influenza/gisrs_laboratory/h5n1_nomenclature/en/]. The tree was generated in MEGA from the downloaded HP HA nucleotide sequences, together with the sequences used in the WHO tree [http://www.who.int/influenza/gisrs_laboratory/201101_h5smalltreerealignment.txt] using the Tamura-Nei model. The available PB2 sequences corresponding to the labelled isolates (final number after matching and de-duplication = 963) were analysed for PB2 E627K mutations by host type and clade using custom R scripts. *Analysis carried out by Samantha Lycett, Edinburgh University.*
Table 8. Prevalence of the PB2 627K mutation in HPAI H5N1 virus isolates

<table>
<thead>
<tr>
<th>Asian HPAI H5N1 isolates</th>
<th>Sequence at PB2 position 627&lt;sup&gt;b&lt;/sup&gt;</th>
<th>E</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Avian</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hong-Kong 97-like</td>
<td>100 (13)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1996-2004</td>
<td>98.9 (173)</td>
<td>1.1 (2)</td>
<td></td>
</tr>
<tr>
<td>2005-present (exc. Clade 2.2)</td>
<td>97.8 (396)</td>
<td>2.2 (9)</td>
<td></td>
</tr>
<tr>
<td>2005-present (Clade 2.2)</td>
<td>11.3 (24)</td>
<td>88.7 (188)</td>
<td></td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hong-Kong 97-like</td>
<td>60 (3)</td>
<td>40 (2)</td>
<td></td>
</tr>
<tr>
<td>1996-2004</td>
<td>68.8 (11)</td>
<td>31.2 (5)</td>
<td></td>
</tr>
<tr>
<td>2005-present (exc. Clade 2.2)</td>
<td>80.2 (93)</td>
<td>19.8 (23)</td>
<td></td>
</tr>
<tr>
<td>2005-present (Clade 2.2)</td>
<td>0</td>
<td>100 (21)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> The analysis was performed using full length PB2 sequences of avian and human HPAI H5N1 viruses from NCBI Influenza Virus Resource database, and human HPAI H5N1 viruses from GISAID, April 2013. **Analysis carried out by Samantha Lycett, Edinburgh University.**

<sup>b</sup> Data are percentages. The number of sequences analysed is shown in brackets.

3.2.2 Generation of H5N1 viruses with mutations at position 627 in PB2

To examine the effect of the 627K genotype in a H5N1 avian influenza virus that circulated before 1996, the mutation was engineered into the PB2 gene of a non-Asian lineage H5N1 virus, 50-92, isolated in 1991. The corresponding virus pair, bearing either E or K at this position, were rescued by reverse genetics, and termed 50-92/aEQ (WT) and 50-92/aKQ (Refer to Figure 28A for sequence). Similarly, an isogenic pair of recombinant Eurasian lineage H5N1 viruses that differed only by E or K at PB2 residue 627 were rescued, and termed Ty/05/K (WT) and Ty/05/E (outlined in Table 9). The 50-92 PB2 627K mutation was introduced by a single nucleotide change, g1881a; Ty/05 PB2 627E was engineered by a double nucleotide change at codon 627, aag to gaa.
### Table 9. HPAI H5N1 model viruses used for this study

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Year</th>
<th>Lineage</th>
<th>PB2 aa 627</th>
<th>PB2 nt 1869</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutant:</td>
<td></td>
<td></td>
<td>627K</td>
<td></td>
<td>50-92/aKQ</td>
</tr>
<tr>
<td>A/turkey/Turkey/5/2005 (H5N1)</td>
<td>2005</td>
<td>Eurasian lineage (HA clade 2.2)</td>
<td>627K</td>
<td>a</td>
<td>Ty/05/K</td>
</tr>
<tr>
<td>Mutant:</td>
<td></td>
<td></td>
<td>627E</td>
<td></td>
<td>Ty/05/E</td>
</tr>
</tbody>
</table>

#### 3.2.3 Rescue of recombinant viruses altered in PB2 at residue 627 and spontaneous selection of additional mutations upon passage of pre-Asian H5N1 50-92 PB2 627K

Viruses generated by reverse genetics were passaged in 9 or 10-day old embryonated chicken’s eggs and sequenced through the PB2 627 region at the time of rescue and at each egg passage. No mutations were observed in Ty/05/K, Ty/05/E or 50-92/aEQ virus (Figure 28A, B&C). In contrast, 50-92/aKQ readily mutated, either acquiring a synonymous mutation a1869g or reverting to 627E. Moreover, viral RNA could not be amplified and sequenced from all 50-92/aKQ infected eggs due to insufficient virus titre. A recombinant virus carrying the synonymous a1869g mutation was subsequently engineered, termed 50-92/gKQ (Figure 28E). Carried out in duplicate, one 50-92/gKQ virus displayed genetic instability at the point of generation, resulting in change at residue 628 from Gln to Arg, whilst retaining the engineered 627K mutation as well as the synonymous 1869g during MDCK passage. The resultant 50-92/gKR virus was maintained during egg passage (Figure 28F). The second recombinant 50-92/gKQ virus also selected the 628R mutation, albeit less readily in 1/7 eggs by the second egg passage (EP2) (Figure 28F). Indeed, mutation of Q628R was observed in the context of E627K on multiple separate occasions during egg or MDCK passage of 50-92 recombinant viruses (data not shown). It was not possible to generate a pure 50-92/aKQ virus of sufficient titre for further experimentation therefore subsequent in vivo experiments were carried out with 50-92/gKR virus as it appeared most stable during egg passage.
Figure 28. Mutation of recombinant non-Asian H5N1 50-92 PB2 627K upon virus passage

Reverse genetics (RG) viruses were generated by one passage in MDCK cells. Viral RNA of RG virus was extracted, amplified by RT-PCR and sequenced in the PB2 627 region. A. DNA sequencing chromatograms of 1867-1884 nt sequence of PB2 gene of RG viruses (Ty/05/E, Ty/05/K, 50-92/aEQ, 50-92/aKQ) and/or RG passage mutants (50-92/gKQ, 50-92/gKR). Nucleotide changes are shaded in grey. Mutations introduced by
site-directed mutagenesis are underlined by a solid line. Mutations detected upon virus passage are underlined by a dashed line. B-F. RG viruses were blind passaged in embryonated chicken eggs up to egg passage (EP) 3 and the PB2 627 region identified by Sanger sequencing (n=7) i.e. each result is the consensus sequence of an egg isolate. Eggs were inoculated with <4HA units determined by HA assay using turkey red blood cells and incubated at 37°C. E&F. 50-92/aKQ and 50-92/gKQ were twice repeated using separately generated RG viruses. Fi. 50-92/gKQ input has an unknown mix of 50-92/gKQ and 50-92/gKR virus after generation through one MDCK passage.

3.2.4 H5N1 50-92/gKR virus is attenuated in vivo

To assess the effect of PB2 627K mutation on pathogenicity in the chicken host, the IVPI assay was used to assign a pathogenicity index (148). For each virus, ten 6-week old chickens were intravenously inoculated with virus and clinical signs measured daily for 10 days. Both Ty/05 viruses scored a maximum 3.00 regardless of PB2 residue 627. On the other hand, 50-92/aEQ with the typical avian residue scored a higher pathogenicity index than 50-92/gKR, 2.88 and 1.68 respectively (Table 10 and Figure 29). In addition, sequencing of samples taken from surviving birds at the experiment endpoint revealed that the 50-92/gKR virus had undergone PB2 K627E reversion but retained the 628R mutation during the experiment (data not shown). MDCK origin 50-92 viruses were also assigned a pathogenicity index, 50-92/aEQ had an identical score as the egg grown virus, whereas 50-92/gKR was slightly less pathogenic than the equivalent virus grown in eggs (Figure 29).

To further examine the lower pathogenicity in chickens associated with the 627K mutation in 50-92, seven 10-week old White Leghorn chickens were inoculated oculonasally with $10^6$ EID$_{50}$ units of either 50-92/aEQ or 50-92/gKR (egg origin) and virus shedding was quantified in both groups. Oropharyngeal virus shedding was significantly higher at 36 h.p.i. for 50-92/aEQ (2.75 REUlog$_{10}$) than 50-92/gKR (0.49 REUlog$_{10}$) (Figure 30A). Cloacal virus shedding was significantly higher at 36 h.p.i. for 50-92/aEQ (4.03 REUlog$_{10}$) than 50-92/gKR (1.55 REUlog$_{10}$) (Figure 30B). Area under the curve analysis was performed on oropharyngeal and cloacal virus shedding in the first 36 hours of infection, values were higher for 50-92/aEQ than 50-92/gKR, and this difference was statistically significant for cloacal virus shedding (Figure 30C). Viral titres in the basal calamus of feathers plucked from birds at 36 h.p.i. were significantly higher for 50-92/aEQ (3.94 REUlog$_{10}$) than 50-92/gKR (1.62 REUlog$_{10}$), suggesting 50-92/gKR virus disseminated through the chicken host but replicated to lower titres (Figure 30D). In addition to higher shedding, 50-92/aEQ caused
greater mortality in the first 72 h.p.i. Mean survival time across the experiment was less for chickens infected with 50-92/aEQ (3 days) than 50-92/gKR (5 days) (Figure 30E), although this difference was not statistically significant. The number of clinical signs was observed and recorded for both groups. Chickens inoculated with 50-92/aEQ virus demonstrated earlier onset of clinical signs, coinciding with mortality. Once onset, clinical signs for both groups were typical of HPAI and included dropped wings, oedema, cyanosis, conjunctivitis, loss of balance and tremors. Taken together, the data show that 50-92/gKR replicated less efficiently in chickens than wild type 50-92/aEQ virus.

Table 10. Intravenous Pathogenicity Index of reverse genetics generated viruses

<table>
<thead>
<tr>
<th>Virus</th>
<th>Origin</th>
<th>IVPI score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ty/05/K</td>
<td>Egg</td>
<td>3.00</td>
</tr>
<tr>
<td>Ty/05/E</td>
<td>Egg</td>
<td>3.00</td>
</tr>
<tr>
<td>50-92/aEQ</td>
<td>Egg</td>
<td>2.88</td>
</tr>
<tr>
<td>50-92/aEQ</td>
<td>MDCK</td>
<td>2.88</td>
</tr>
<tr>
<td>50-92/gKR*</td>
<td>Egg</td>
<td>1.59</td>
</tr>
<tr>
<td>50-92/gKR*</td>
<td>MDCK</td>
<td>1.32</td>
</tr>
</tbody>
</table>

*unknown gKR/gKQ mix, passage mutant
Figure 29. Virulence in chickens demonstrated by IVPI

Ten 6-week old white leghorn chickens were intravenously infected with either virus, in accordance with OIE guidelines (148), and scored daily for clinical signs over 10 days. 0=healthy, 1=ill, 2=severely ill, 3=dead. The daily sum of clinical signs was calculated and divided by the total number of birds per group. **A.** Ty/05/K and Ty/05/E scored identical. **B.** Egg origin 50-92 virus. **C.** MDCK origin 50-92 virus. 50-92/gKR virus was an unknown gKR/gKQ mix, passage mutant.
Figure 30. PB2 627K carries a fitness cost for non-Asian H5N1 50-92

Seven 10-week old chickens were oculonasally inoculated with $10^6$ EID$_{50}$ units of 50-92/aEQ or 50-92/gKR. Viral RNA was extracted from oropharyngeal, cloacal or feather samples and virus titre measured by qRT-PCR of the matrix gene, displayed as mean log$_{10}$ Relative EID$_{50}$ Unit (REU). A. Oropharyngeal virus shedding. B. Cloacal virus shedding. C. Area Under Curve (AUC) analysis of shed virus up to 36 h.p.i.. D. Virus titre of the basal calamus of feather samples taken from the brow of chickens at 36 h.p.i.. E. Percent survival of chickens. F. Average number of clinical signs observed per bird at chosen time points. (A, B, C, F) log values (plus an arbitrary 0.01 to allow log transformation of 0 values) were used for suitable statistical analysis by two-way ANOVA corrected by Bonferroni post test. (D) two-tailed paired t-test. Error bars are displayed as SEM. * p<0.05, ** p<0.01.
### 3.2.4.1 Reversion of PB2 K627E and R628Q in mutant non-Asian lineage H5N1 50-92/gKR upon replication in chickens

The PB2 gene of virus isolates taken at various time points from both groups of infected chickens was sequenced. Viruses were sequenced from oropharyngeal or cloacal samples that were positive for virus by qRT-PCR to determine the PB2 627-628 amino acid sequence. Five out of seven birds infected with the 50-92/gKR virus showed reversion of the 627K mutation to 627E during the course of the experiment. In 4 of these 5 birds, reversion of the second site mutation 682R back to Q was detected in some samples. The K627E reversion was first detected at 60 h.p.i. in both oropharyngeal and cloacal samples (Figure 31). After this time, the birds inoculated with 50-92/gKR virus showed increased average shedding (Figure 30A+B) and increased clinical signs (Figure 30F). Interestingly, a cloacal sample from 1 bird carried the aKQ signature at 132 h.p.i., the biological significance of this is unclear as the same bird revealed the gER signature in oropharyngeal shed virus at a later time-point. In contrast to the genetic instability in birds infected with gKR virus, there was no sequence change in the PB2 gene detected in birds inoculated with 50-92/aEQ virus (data not shown).

**Figure 31. Reversion of PB2 K627E and R628Q in mutant non-Asian lineage H5N1 50-92/gKR upon replication in chickens**

At chosen time points viral RNA, from virus positive oropharyngeal and cloacal samples from chickens infected with 50-92/gKR (Figure 30), was extracted and amplified by RT-PCR and sequenced in the PB2 627 region. Graphs show where 627E was detected; calculated percentages include undetectable shedding and dead birds. **A.** oropharyngeal samples. **B.** Cloacal samples.
3.2.5 PB2 627K does not carry a fitness cost for Eurasian lineage H5N1 Ty/05

Ty/05 virus was highly virulent for intravenously infected chickens (Table 10 and Figure 30). Viral RNA was quantified from oropharyngeal and cloacal swabs from five dead birds infected with either Ty/05/K or Ty/05/E (Figure 32). A small increase in replication was observed with Ty/05/E virus over 627K. Since poultry were exceptionally susceptible to Ty/05 infection, viral fitness differences between Ty/05/K and Ty/05/E were not readily measured in this host.

Figure 32. Quantification of Ty/05 virus replication in chickens after intravenous inoculation

Ten 6-week old chickens were intravenously inoculated with either Ty/05/K or Ty/05/E virus, as described in Table 10. Oropharyngeal and cloacal swabs were taken from five randomly selected dead birds 1dpi. vRNA was extracted and quantified by Q-RT-PCR of the Haemaglutinin gene, displayed as mean log_{10} Relative EID_{50} Unit (REU). Statistical analysis was by One-way ANOVA, * p<0.01. Error bars are displayed as SEM.

Domesticated ducks are a known avian host of modern HPAI H5N1 viruses. Ty/05/K WT virus was previously shown to readily infect 12 week old pekin ducks and was not lethal in birds of this age, although this virus was 100% lethal to 4 week old pekin ducks (221, 396). Therefore, twelve 12-week old pekin ducks (Anas platyrhynchos) were infected with 10^6 EID_{50} units of either Ty/05/K or Ty/05/E and five 4-week old naive contact sentinel pekin ducks were co-housed 1 d.p.i. to measure contact transmission efficiency. Daily oropharyngeal and cloacal virus shedding was measured in all birds by qRT-PCR. Cloacal virus shedding was absent or extremely limited in 12-week old ducks (data not shown). There was no significant difference in oropharyngeal virus shedding between the two 12-
week old duck groups (Figure 33A). Similarly, no significant differences were observed in oropharyngeal and cloacal virus shedding of 4-week old contact ducks, all of whom contracted infection following exposure to directly infected animals (Figure 33B). Feathers were taken from dead or euthanized 4-week old ducks and virus titre of the basal calamus was measured by qRT-PCR. There was no significant difference in mean virus titre of the feathers (Figure 33C). The mean survival time was 7 days for both groups of 4-week old ducks, both viruses causing 80% mortality (Figure 33D). Similar clinical signs were observed in both groups, predominantly those of a neurological nature, and comparable to those previously observed by Londt et al. 2008 (396) (Appendix i, Table 22).

Figure 33. PB2 627K is not a fitness cost for Eurasian lineage H5N1 Ty/05

Twelve 12-week old pekin ducks were oculonasally inoculated with $10^6$ EID$_{50}$ units of Ty/05/K or Ty/05/E. Five 4-week old naive contact pekin ducks were co-housed 1 d.p.i.. Five 12-week old ducks were euthanized for post mortem analysis 2 d.p.i.. Virus RNA was extracted from oropharyngeal, cloacal or feather samples and virus titre measured by qRT-PCR of the HA gene, displayed as mean log$_{10}$ REU. A. Daily buccal virus shedding of 12-week old ducks. B. Daily buccal and cloacal virus shedding of 4-week old contact ducks. C. Virus titre of the basal calamus of feather samples taken from the breast of dead or euthanized 4-week old ducks. D. Percent survival of 4-week old contact ducks. Error bars are displayed as SEM.
Five randomly selected 12-week old ducks from each infected group were euthanized 2 d.p.i., viral loads in tissues were quantified by qRT-PCR and viral tropism was analysed by IHC. 1-3 out of 5 birds in each group showed positive qRT-PCR in some or all tissues but loads varied widely and no significant differences were observed (Appendix ii, Table 23). No significant gross or histopathological changes were observed in ducks challenged with either viral isolates at 2 dpi. In tissues that were positive by IHC from birds infected with either 627E or 627K virus, virus distribution was similar (Figure 34). Immunolabelling was observed in lymphohistiocytic cells and endothelial cells in lamina propria in nasal turbinates; epithelial cells in air capillaries and air sacs; medullary cells (lymphocytes, macrophages and reticular cells) of bursal follicles; splenic red pulp and stromal cells in ovary (Figure 34). Overall, distribution of viral antigen by immunohistochemistry was sparse, with 1 bird infected with Ty/05/E scoring positive and 2 birds from the Ty/05/K group (Appendix ii, Table 24). Although less clear, there was evidence of epithelial staining of the nasal turbinates by both Ty/05/K and Ty/05/E, as well as on cervical trachea of a Ty/05/K infected duck (Figure 34ii). Virus antigen from both groups was detected in air capillaries of the lung. In addition, Ty/05/K antigen was found in the brain and liver, whereas Ty/05/E antigen was present in the kidney, jejunum and thymus (Figure 34ii). In tissues infected by Ty/05/E but not Ty/05/K, antigen distribution appeared similar than those reported from ducks previously challenged by Ty/05 isolate (WT, carries 627K) (221, 396). Post-mortem analysis was performed 2dpi as we sought to detect subtle differences between the virus groups; however it appeared that this time-point was unsuitable for thorough IHC analysis, and a later time point may have detected more virus infection.

All remaining 12-week old ducks survived the infection, haemagglutination inhibition of duck sera demonstrated all surviving ducks sero-converted (data not shown). No mutations were observed in the 627 region of PB2 in oropharyngeal virus samples from any surviving birds at 9 d.p.i. (data not shown). After detailed analysis, these results did not demonstrate any fitness cost or advantage of PB2 627K for Ty/05 virus in ducks.
Figure 34. H5N1 Ty/05/K and Ty/05/E virus antigen distribution of tissues in experimentally infected ducks

Immunohistochemical demonstration of influenza A nucleoprotein (brown) in tissues of ducks challenged with either Ty/05/K (left) or Ty/05/E (right) at 2dpi.


3.2.6 Growth kinetics in avian cells.

To further examine the effect of PB2 627K in vitro in the two different virus backbones, multi-cycle growth curves were performed on primary chicken embryonic fibroblast (CEF) cells at both 37 and 41°C, the latter being more representative of avian body temperature. Viral RNA was extracted from supernatant and quantified by qRT-PCR. At the final time point of growth curves conducted at 37°C (28 h.p.i.) viral RNA was used to sequence the PB2 gene. The aKQ signature conferred a significant cost to viral replication at both 37 and 41°C in the 50-92 background and mutation to gKQ was detected at the 28 h.p.i. time point. Surprisingly, viral RNA titres were slightly higher for 50-92/gKQ than 50-92/aEQ at 8 h.p.i., 37°C but this pattern was not consistent at later time points (Figure 35A&B). Ty/05/K and Ty/05/E both replicated to high titres in CEF cells and there was little difference in the quantity of viral RNA generated by the two viruses (Figure 35C&D).
Figure 35 (Including previous page). Growth kinetics in avian cells

Primary chicken embryonic fibroblast (CEF) cells were inoculated with 50-92/aEQ, 50-92/gKR, 50-92/gKQ, or 50-92/aKQ at a MOI of 10⁻⁵. Cells were incubated at 37°C (A) or 41°C (B). (A+B) At 28 hours post-infection viral RNA was extracted, amplified by RT-PCR and sequenced in the PB2 627 region. Reversion from 50-92/aKQ to 50-92/gKQ was detected at 28 hours post-infection (black arrow). In addition, CEF cells were inoculated with Ty/05/K and Ty/05/E at a MOI of 10⁻⁴. Cells were incubated at 37°C (C) or 41°C (D). Virus RNA was extracted from supernatant and virus titre measured by qRT-PCR of the matrix (50-92) or HA (Ty/05) gene, displayed as mean Log₁₀ REU, infections were carried out in triplicate. These trends were reproducible. Error bars displayed as SEM. Statistical analysis was by either two-way ANOVA corrected by Bonferroni post-test (50-92) or two-tailed t-test (Holm-Sidak method) (Ty/05). * p<0.05, ** p<0.01, **** p<0.0001.

3.2.7 Differences in polymerase activity.

To understand whether effects on polymerase activity underlie the fitness cost associated with 50-92 PB2 627K mutation, we performed an in vitro polymerase assay using plasmids that expressed WT PB2 (50-92 PB2 627E or Ty/05 PB2 627K) or corresponding PB2 mutants with their cognate PB1, PA and NP expression plasmids, together with the minigenome firefly reporter plasmid pChicken-Poll-Firefly or pHuman-Poll-Firefly in avian DF-1 or human 293T cells, respectively. These reporter constructs use the cellular RNA polymerase I to generate a virus-like mini-genome comprised of the firefly gene flanked by the NS gene segment’s NCRs. For comparison, the polymerase complex and NP protein of A/England/195/2009 (pH1N1) was also reconstituted in human cells. In this assay, polymerase containing Ty/05/K consistently directed a higher luciferase signal than Ty/05/E polymerase at 37, 39 and 41°C in chicken DF-1 cells, although this was only statistically significant at 37°C (Figure 36A). Likewise, 50-92 polymerase activity was also increased by the 627K mutation in PB2 at 37°C (Figure 36A). As expected, polymerases bearing PB2 627E were severely restricted in their ability to support minigenome amplification and expression.
in human cells whereas polymerases of both 50-92 and Ty/05 replicated efficiently when carrying the PB2 627K signature (Figure 36C).

To determine the effect of the 50-92 628R mutation that had emerged during rescue of the recombinant virus, plasmids bearing PB2 628R paired with either 627K or E (50-92/KR and 50-92/ER) were combined with the other 50-92 polymerase constructs and tested for activity in DF-1 cells at 37 and 41°C. Little difference in activity was observed between polymerase constellations with this PB2 change; 50-92/KR caused a small decrease in activity at 37°C compared to 50-92/K (Figure 36B). These data do not explain the selection of the 628R mutation in the context of the rescued virus.
Figure 36 (previous page). Differences in polymerase activity

Polymerase complex and NP expression plasmids, together with the mini-genome firefly reporter plasmid pChicken-PolI-Firefly or pHuman-PolI-Firefly were transfected into avian DF-1 or human 293T cells, respectively. Cells were lysed after 20 hours and the level of firefly activity measured. A. Log_{10} firefly activity of Ty/05/K, Ty/05/E, 50-92/EQ and 50-92/KQ with cognate polymerase complex and NP conducted at 37, 39 and 41°C in DF-1 cells. B. Log_{10} firefly activity of 50-92/EQ or 50-92/KQ together with the 628R mutation, conducted at 37 and 41°C in DF-1 cells. C. Log_{10} firefly activity of Ty/05 and 50-92 PB2 E/K constructs together with pH1N1 human control in 293T cells at 37°C. Mock is transfection of firefly reporter only. Values were carried out in triplicate. These trends were reproducible. Error bars displayed as SEM. Statistical analysis was by two-way ANOVA corrected by Bonferroni post-test. ** p<0.01, **** p<0.00001.

3.3 Conclusions and Discussion

The aim of this chapter was to determine the effect of the PB2 627K mutation for the avian host. Being such a well demonstrated adaptation marker of successful human influenza viruses, the recent emergence of this mutation in the avian host must be understood. Our observations suggest modern HPAI H5N1 viruses tolerate either genotype at PB2 627 with no obvious fitness cost or advantage, in contrast to 50-92 which was intolerant of the PB2 627K mutation.

Since 2005, the unprecedented spread of the Eurasian lineage H5N1 HPAI viruses from Asia to Europe, the Middle East and Africa has posed a continual threat to domesticated and wild bird populations. In addition, cumulative human cases have been reported in these areas, underscoring the potential for these viruses to give rise to a pandemic. Indeed, although H5N1 has not naturally evolved the ability to spread from person to person, under experimental conditions a clade 2.1 H5N1 virus was shown to acquire mutations that supported transmission between ferrets by the respiratory droplet route, an indicator of pandemic potential (88).

One of the key mutations required for airborne H5N1 transmission in ferrets was the well known human adaptation motif, PB2 627K. Here we confirmed the high prevalence of this mutation in clade 2.2 viruses isolated from both wild and domestic birds compared to other H5N1 clades and indeed other influenza subtypes (Table 8 and Figure 27), in agreement with Russell et al. 2012 (308).

Using a reverse genetic approach we examined the consequence of this mutation for virus replication and transmission in the avian host. We concluded that in the Eurasian
lineage clade 2.2 virus, A/turkey/Turkey/05/2005, PB2 627K provided neither significant fitness cost nor gain and thus did not affect pathogenesis or transmission in ducks or IVPI score in chickens. Interestingly, Ty/05/E replicated to slightly higher titres over Ty/05/K in chickens from the IVPI assay (Figure 32). However, since the IVPI is a crude assay whereby chickens are inoculated by an artificial route with an undefined quantity of virus, it is difficult to establish the importance of this small observation. Alternatively, if this observation described a real difference between ducks and chickens and PB2 627K, this could represent a difference in a host cell factor between the two species. Indeed, this might support the observations of Schat et al. who also found an in vivo advantage in chickens for a virus with 627E over K, but only in a virus engineered to be of low pathogenicity. However, we have not observed a negative impact of the 627K mutation on polymerase activity in chicken cells, which is quite apparent with PB2 627E in human cells which lack an efficient cell co-factor.

The finding that this mutation is neutral in ducks may help refine risk assessments, such as those described by Russell et al. (2012) and indicates the high probability that this mammalian adaptation motif will be maintained in the avian reservoir (308). Similarly, Schat et al. (2011) concluded that neither disease outcome nor transmission of a HP H5N1 virus in ducks or chickens were altered by PB2 627 E or K. Their study used a human H5N1 virus isolate of clade 1 (A/Vietnam/1203/04) which bears 627K. Within the clade 1 virus group some viruses isolated from human and feline cases contained the 627K mutation but the majority of natural avian isolates retained 627E (Table 8) (398, 399). Interestingly, Schat and co-workers also performed in vivo studies with viruses engineered to be LPAI by removal of the multi-basic cleavage site in the HA protein. In chickens the LP virus with PB2 627E induced more profound clinical signs, was detected in more tissues by IHC and virus isolated from more infected bird tissues than the 627K virus. However, the mechanism by which 627E increases replication in chickens in a LP context was not elucidated. Our own data did not conclude any advantage of 627E for polymerase function (Figure 36) or virus replication (Figure 35) in chicken cells in either genetic backbone. Schat et al. did not investigate transmission of the LP viruses between chickens or ducks, so the biological relevance of the observations in chickens for the maintenance of 627E genotype in LP or HP avian influenza viruses in a natural reservoir is not clear. Another way to measure modest difference in
fitness between the two viruses would be to perform competition assays where mixtures of viruses would be used in various ratios to infect cells or animals. We have not performed this style of experiment with pairs of viruses that differ in PB2 627 E or K in the present study but this might be a useful avenue to pursue in the future.

Clade 2.2 viruses are unusual as they are shed via the respiratory route in domestic waterfowl, in contrast to the majority of avian influenza viruses that are shed cloacally (396, 400). The mechanism for this switch in tropism remains unclear. PB2 627K has been shown to confer cold adaptation in mammalian hosts, and it is possible that this mutation provided a small advantage in the cooler temperature of the avian respiratory tract (316). Indeed, we observed an increase in polymerase activity of Ty/05/K as compared to Ty/05/E at the cooler temperature of 37°C (Figure 36). However, our in vivo experiments showed that Ty/05/E was also shed via the respiratory route suggesting that others genetic changes may also contribute to the shift in tropism of the clade 2.2 H5N1 viruses.

Indeed the clade 2.2 Ty/05 HA also naturally carries the mutations, T158A, which results in the loss of a glycosylation site, as well as R193 in the receptor binding site; both of these mutations can increase human α2,6SA binding, at least when combined with the human receptor changing mutations Q226/G228 (350, 401). Several poultry species have been shown to contain α2,6SA receptors in the respiratory tract (374, 402) which may have selected for the HA mutations and simultaneously supported the 627K genotype in Ty/05 and other clade 2.2 viruses (401). On the other hand, it has also been suggested that such genetic changes have been selected via an intermediate mammalian host (e.g. Pika) or avian species of the Ratitae (e.g. Ostrich), before returning to the avian influenza reservoir (327, 393). These mutations combined with the cold adapting PB2 627K switch, were amongst those recently shown to confer airborne transmission of HPAI H5N1 virus between ferrets (87, 88). Therefore our data underscore that in a natural setting the PB2 627K mutation has arisen and is likely to be maintained in a virus whose other genetics support the evolution of a human transmissible phenotype.

Intriguingly, and in contrast to the results with the modern H5N1 virus, our experiments with a historical H5N1 virus, A/turkey/England/50-92/91, revealed a significant fitness cost resulting from genetic instability associated with the nucleotide changes
required to generate PB2 627K. Since 50-92 is genetically distinct from the Asian H5N1 viruses, the significance of this finding to currently circulating Asian lineage H5N1 viruses is difficult to determine, but may be more pertinent to the observation that most but not all avian isolates, naturally maintain a glutamic acid at PB2 627. Table 8 demonstrated that the majority of human H5N1 infections retained the 627E motif, whether these viruses were forced to select other adaption markers such as PB2 701N is not known, and may be worthy of further investigation. The effect of the 627K mutation in the 50-92 virus background could not be examined in the duck host, since ducks are refractive to infection by 50-92 virus (270, 366). Yet, since 50-92 627K virus selected compensatory mutations even in mammalian MDCK cells, it suggests a more universal mechanism behind our observations, independent of host origin.

Bogs et al. (2011), provided evidence that other viral genes affect the selection of the 627K mutation in PB2. They generated two PB2 gene reassortant viruses with PB2 627E; a Eurasian lineage (clade 2.2) virus that carried the Hong Kong 1997 (clade 0) H5N1 virus PB2 gene, and a Hong Kong 1997 H5N1 virus that carried the Eurasian lineage PB2 gene. When passaged through MDCK cells, neither of these viruses selected the PB2 627K mutation. In contrast, a Eurasian lineage virus engineered to bear PB2 627E reverted to 627K by the fifth passage. Further reassortant viruses demonstrated a requirement for the Eurasian lineage NP protein to support the 627K reversion (327). Surveillance carried out in Chinese pikas (a small mammalian species) near Qinghai Lake revealed two distinct H5N1 virus lineages present in the pika population, a Vietnam-like lineage and a Qinghai-like lineage (403). These virus lineages differed at PB2 627. Only the Qinghai-like viruses carried the PB2 627K mutation and were pathogenic in mice, whereas Vietnam-like viruses retained PB2 627E (327, 403). Since the Vietnam-like lineage virus did not select for 627K, it may suggest that not all avian influenza viruses favour the 627K genotype, even under natural selective conditions i.e. in mammals. On the other hand, several reports have described the selection of avian influenza viruses with the PB2 627K mutation particularly during passage in mice, including H9N2 (389, 404) and H7N7 (391). There are also several reports of viruses that have been stably engineered to possess the 627K mutation (293, 322, 382, 392, 405). Indeed in the recent outbreak of a novel H7N9 virus in China, 2013, the PB2 627K mutation is often selected for upon human infection (406, 407).
The mechanism driving the observed fitness cost by PB2 627K in the 50-92 genetic background used here remains unclear. Exploration of polymerase activity did not account for *in vitro* and *in vivo* observations, suggesting the 627K coding change had an effect on the PB2 gene independent of its role in the polymerase complex. A recent study has highlighted the role of NEP in regulating polymerase activity. Host adaptation of NEP was able to compensate the restriction of avian polymerase in human cells (43). The exact mechanism of NEP and how it affects polymerase activity remains unclear. NEP was not expressed in the polymerase assays conducted in this study, however, work in our laboratory suggests NEP does not greatly affect polymerase activity in avian cells (data unpublished) and since avian cells did not demonstrate host-specificity by PB2 627 it is unlikely that NEP would have altered polymerase activity of our constructs and accounted for our *in vivo* observations.

We engineered the E627K mutation into 50-92 PB2 by a single nucleotide change g1879a. In response to this change we observed the selection of two adenine to guanine transitions, a1869g and a1884g, the later coding a Glu to Arg change at aa 628. It is unclear why these mutations occurred when a single nucleotid transition a1879g was sufficient to revert PB2 to K627E. When the a1869g mutation was engineered into the PB2 gene, the 628R mutation was also selected, to varying degrees during independent experiments (Figure 28). Although there are several examples of the 627K mutation being selected for during passage of virus with avian PB2 in cells or mice, the co-selection of further mutations synonymous or otherwise in PB2 has not been previously been reported (292). The a1869g synonymous change appeared to support the 627K phenotype *in vitro* as replication of 50-92/gKQ and 50-92/gKR viruses were comparable to 50-92/aEQ in CEF cells (Figure 35). However, there was still a significant fitness cost for 50-92/gKR *in vivo*. The mis-matching phenotype of *in vitro* growth kinetics and polymerase activity to *in vivo* chicken experiments poses an interesting speculation, “How does the animal host present a more challenging environment for the virus to conquer?”

Viruses of the Qinghai lake outbreak of 2005 carried an important mutation, PB2 627K, significantly increasing the threat of HP H5N1 viruses to the human population. We have demonstrated no fitness cost associated with this mutation in the clade 2.2. viruses from the Eurasian lineage, offering an explanation as to why this human adapting mutation was maintained in the avian reservoir. Such occurrences increase the chance of further
mutations that can lead to zoonoses or human transmission. Furthermore, the converse
finding in an historic H5N1 may suggest an explanation for the paucity of the 627K mutation
in other influenza viruses in the avian reservoir.

Limitations and future work

Although observations suggest that Ty/05 virus has a neutral acceptance of either
PB2 627K or E, the attempt to compare tissue tropism and virus load in adult pekin ducks
infected with either virus would have been more reliable if post-mortem was performed at
the peak day of shedding, i.e. day 3 post-infection. In addition virus load and shedding from
all birds was measured by qRT-PCR of vRNA. Although this is a good measure of virus
amplification and a practical method to conduct in high containment, it does not directly
represent infectious virus. Any additional work could include the measure of infectious virus
by titration of virus in cell tissue culture. Pekin ducks are an appropriate model in which to
examine the fitness cost of H5N1 viruses, especially of the Eurasian lineage which are known
to have an extended host range and pathogenesis in the avian reservoir. Given this, the
effect of PB2 627K on different wild bird species may be worth studying in the future, do
avian species other than the Ratitae preferentially select the 627K mutation?
Chapter 4. Investigating the underlying mechanism of the PB2 627K associated fitness cost in non-Asian HPAI H5N1 virus

4.1 Introduction

The inherent genetic instability of 50-92 after the introduction of PB2 627K was not explained by polymerase activity, Chapter 3. A phenotype such as this has not been described previously, but another H5N1 virus, A/HongKong/156/97 (HK156), isolated from the first human outbreak of H5N1 in Hong Kong has been used in studies on host range by several different laboratories (408) and demonstrated a failure to select the 627K mutation upon passage in mammalian cells or in vivo host. As previously discussed, Bogs et al. (2011) passaged HK156 reassortant viruses in mammalian cells and demonstrated its failure to select 627K (passage of WT HK156 was not described), no compensatory mutations were described (327). Hiromoto et al. (2000) sequenced HK156 derived from MDCK cells and infected mice, neither selected the PB2 627K mutation, however, viruses derived from MDCK cells carried the PB2 701N mutation and were more pathogenic in mice (409).

The study by Bogs et al. (2011) described the requirement for the cognate NP protein for the selection of PB2 627K by a Eurasian lineage clade 2.2 H5N1 virus (327). The activity of PB2 627K has been previously reported to be linked to the NP protein, where reduced polymerase activity in human cells of avian origin polymerase (627E) abrogated the formation of vRNPs by reduced binding of PB2 to NP, a phenotype rescued by the 627K mutation (319, 324–326). However, recent work by Cauldwell et al. (2013) demonstrated these observations were due to an indirect measurement of the 627K mutation, where increased polymerase activity resulted in more PB2-NP co-precipitation (328).

Since Ty/05 was observed to support the 627K mutation, in dramatic contrast to 50-92, we sought to investigate the mechanism behind our observations by mis-matching gene segments of the two viruses. Since polymerase activity did not account for these differences we hypothesised the fitness cost may occur at the level of RNA on the PB2 gene segment.
4.2 Results

4.2.1 Investigation of avian influenza virus sequence in the PB2 627 region

Figure 37 summarises the sequence of the PB2 627 region for our virus panel. Since our results suggest an RNA incompatibility between gene segments and the PB2 627 region, we investigated the identity of the 627 codon, as well as nt 1869 and amino acid 628 for full length PB2 sequences of avian influenza viruses (all subtypes) isolated from 1956 to 2013, downloaded from the NCBI Influenza Virus Resource database, May 2013 (Table 11).

![Figure 37. Summary of the Ty/05 and 50-92 virus genotypes](image)

mRNA and amino acid sequence of the 627 region of PB2 for Ty/05 and 50-92 and mutants. Light blue= WT sequence. Coloured areas= mutated nucleotide or amino acid.

Of the natural virus isolates encoding 627K, 93.3% of these used the aag Lys codon, the alternative aaa codon was found less extensively, 6.7%. Ty/05 naturally evolved to encode Lys by aag, whereas 50-92 was engineered to carry aaa since its natural codon for 627E was gaa and the switch to aaa required only a single nt change.

The majority of isolates carried a Glutamine at PB2 628, 99.9% (PB2E) and 99.3% (PB2K). Sequence gazing (that is to say looking at a protein sequence in an alignment of sequences to find features that may interpret function) revealed few other examples with a
positive charged aa at residue 628 that were excluded from the previous analysis as only partial sequencing of PB2 was available, these included genotype tER (A/shorebird/DelawareBay/13/2004(H6N8)) and aEK (A/goose/Guangdong/72/2004 (H5N1)). In conclusion, 628R is rarely observed in nature.

Most isolates carried either thymine or adenine at PB2 nt 1869. A minority of 627E isolates carried cytosine or guanine at nt 1869. In addition to these data, sequence gazing of isolates with partial PB2 sequences identified examples of other H5N1 isolates that encoded 1869g; including 9 isolates from Nigerian chickens in 2006 with the gKQ genotype and one gEQ (A/duck/France/05066b/2005(H5N1)). Although less common, 1869g is observed in nature, including viruses of the H5N1 subtype and in conjunction with 627K.

Table 11. Quantitative analysis of PB2 627 codon sequence, amino acid PB2 628 and nt 1869 from avian influenza viruses isolated from 1956 to 2013

<table>
<thead>
<tr>
<th>PB2 627 Genotype</th>
<th>E (6375)</th>
<th>K (268)</th>
</tr>
</thead>
<tbody>
<tr>
<td>627 codon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gaa</td>
<td>68.6 (4372)</td>
<td>93.3 (250)</td>
</tr>
<tr>
<td>gag</td>
<td>31.4 (2003)</td>
<td>6.7 (18)</td>
</tr>
<tr>
<td>628aa</td>
<td>Q (99.9 (6424))</td>
<td>Q (99.3 (266))</td>
</tr>
<tr>
<td>Other</td>
<td>0.1 (8)</td>
<td>0.7 (2)</td>
</tr>
<tr>
<td>1869 nucleotide</td>
<td>t (61.2 (3935))</td>
<td>t (2.2 (6))</td>
</tr>
<tr>
<td>a</td>
<td>35.7 (2295))</td>
<td>a (97.8 (262))</td>
</tr>
<tr>
<td>c</td>
<td>1.9 (122)</td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>1.2 (78)</td>
<td></td>
</tr>
</tbody>
</table>

The analysis was performed using full length PB2 sequences of avian and human HPAI H5N1 viruses from NCBI Influenza Virus Resource database, May 2013. Analysis carried out using Microsoft Excel and Geneious. Data are percentages. The number of sequences analysed is shown in brackets.

a Analysis of PB2E includes 55 isolates carrying 627A,G or V (i.e. excluding PB2K).

b Other includes (H=1, K=1, L=2, P=2, R=2).

4.2.2 PB2 628 has a prominent location on the 627 linker region

To our knowledge the 628R mutation has not been previously described in the literature, and bioinformatic analysis revealed it to be a rare occurrence in nature (Table
11). Figure 38 demonstrates the location of 627 and 628 on the PB2 protein structure, on an extended peptide region which links the alpha- and beta- halves of the 627 domain (297). Tarendeau et al. 2008 observed no significant structural change to the 627-domain by altering K627E, but mainly a reversal of electrostatic potential. The close proximity of 628 suggests a change in electrostatic charge by the Gln to Arg switch is likely to alter the interactions made by 627, although the position of the side chain is orientated away from that of 627 (Figure 38). Secondary protein structure was predicted using the Protein Homology/analogY Recognition Engine 2.0 (Phyre²) (410) for both genotypes KQ and KR. Models were predicted to the template sequence identified as that from Kuzuhara et al. 2009 for which the crystal structure was obtained (314), and had 96.2% identity similarity with 50-92 PB2 627-Domain. Both linker regions from each genotype were predicted to remain disordered, as observed by the resolved crystal structure (297, 314) (data not shown).

Figure 38. Location and structure of the 627-Domain and amino acid 628
A ribbon diagram of the 627-domain (aa537-693) from A/Victoria/3/1975(H3N2) PB2 showing secondary structure and the position of 627E and 628Q, α-helices in pink, beta-strands in yellow, as defined by Tarendeau et al. 2008. Figure generated using Geneious, Glu627 in red, Gln 628 in light blue.

4.2.3 PB2 628R confers a fitness cost for 50-92 reassortant virus in vitro

In addition to generating highly pathogenic viruses of the H5 subtype, reassortant viruses were generated by reverse genetics combining the five internal gene segments of 50-92 virus together with the HA, NA and M gene segments of PR8 (a lab adapted H1N1
strains). Interestingly, RG viruses with 50-92 PB2 627E or 627K were not rescued on multiple occasions suggesting that 50-92 PB2 gene was incompatible with just the PR8 HA and NA gene segments (data not shown). However, virus could be generated to sufficient titres with the addition of the PR8 M gene to make a 5:3 constellation, as well as with Ty/05 HA, NA and M. Reassortant viruses were also generated in the Ty/05 background, in contrast to 50-92, these viruses were successfully generated with just PR8 HA and NA (6:2 constellation) (Figure 39), or all three PR8 segments (5:3) (not shown in figure).

![Diagram](image)

**Figure 39. Generation of reassortant viruses by reverse genetics**

A mixture of the eight Poll reverse genetics plasmids were transfected into cells by the method described in Chapter 1. PR8 HA, NA and M genes in red, generate a virion composed of PR8 proteins. 50-92 or Ty/05 genes, in blue and green, respectively, encode the polymerase, NP, NS1 and NEP proteins.

To investigate the effect of the PB2 628R mutation that arose from the generated aKQ genotype, 50-92 viruses were generated in the PR8 HA, NA and M backbone with either PB2 627K or 627E (WT) and either the 628R or 628Q (WT) mutation. Viruses that carried 627K were also engineered to carry the 1869g genotype. Growth kinetics in CEF cells were assessed for the virus panel at 37 and 41°C. PB2 628R caused a significant fitness cost for both the 627K and 627E virus (Figure 40A&B), despite this mutation being selected for by egg passage in the context of whole H5 virus (Figure 28). It was previously shown (Figure 36) that 628R decreased polymerase activity for the 627K polymerase but not the 627E polymerase, and that the difference was only significant at 37°C not at 41°C. Together these
data suggested the 628R fitness cost observed for the 5:3 virus constellation was likely due to incompatibility with a PR8 gene segment and not at the protein level of the polymerase.

Figure 40. PB2 628R confers a fitness cost for 50-92 reassortant virus in avian cells

Primary chicken embryonic fibroblast (CEF) cells were inoculated with 50-92/aEQ/PR8HA,NA,M, 50-92/aER/PR8HA,NA,M, 50-92/gKQ/PR8HA,NA,M, 50-92/gKR/PR8HA,NA,M at a MOI of 10^-3. Cells were incubated at 37°C A. or 41°C B. in the presence of TPCK trypsin. Cell supernatant was removed at 12, 20 and 48 hours post-infection and plaqued on MDCK cells to determine virus titre in pfu/ml. Data displayed as mean Log_{10} pfu/ml and were carried out in triplicate. These trends were reproducible. Error bars displayed as SEM. Statistical analysis was by two-way ANOVA corrected by Bonferroni post-test. * p<0.05, ** p<0.01, *** p<0.0001.

4.2.4 PB2 627K confers a fitness gain for Ty/05 reassortant virus

Reassortant Ty/05 viruses were generated with PR8 HA and NA gene segments, with either PB2 627K or 627E. Both viruses grew well in MDCK cells and induced rapid cytopathic effect (data not shown). Virus growth kinetics were measured by infecting CEF and duck embryonic fibroblast (DEF) cells. In contrast to whole H5 virus (Figure 35), PB2 627K provided a statistically significant fitness advantage in both CEF and DEF cells at either 37 or 39°C (Figure 41A-D). Virus titres were consistently 1-1.5 log_{10} pfu/ml higher for Ty/05/K. Despite this difference, both viruses grew rapidly to high titres, displaying similar growth kinetics; viruses grew faster at the higher temperature (Figure 41B&D). A similar pattern of results was observed for both CEF and DEF cells, suggesting each virus was equally competent in either avian cell type.
Figure 41. PB2 627K confers a fitness gain for Ty/05 reassortant virus

Primary CEF and DEF cells were inoculated with Ty/05/K/PR8 or Ty/05/E/PR8 at a MOI of 10^-3. Cells were incubated at 37°C. (A&C) or 39°C (B&D) in the presence of TPCK trypsin. Cell supernatant was removed at 12, 24 and 48 hours post-infection and plaqued on MDCK cells to determine virus titre in pfu/ml. Data displayed as mean Log_{10} pfu/ml and were carried out in triplicate. These trends were reproducible using both MDCK and egg origin virus. Error bars displayed as SEM. Statistical analysis was by two-way ANOVA corrected by Bonferroni post-test. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001.

4.2.5 Altering the genetic constellation of 50-92 prevented reversion of the 627K genotype

Since polymerase activity did not account for the decreased replication in vitro of the 50-92 whole recombinant virus with the aKQ PB2 genotype and in vivo with the gKR PB2 genotype, we hypothesised that the selection of the synonymous mutation at PB2 1869 during passage of this virus indicated an RNA-dependent mechanism. Furthermore, investigation of the effect of PB2 628R by virus growth kinetics did not present an advantage for 50-92 in combination with PB2 627E or K in the PR8 reassortant background. To address this, the reassortant viruses, including a 50-92 virus generated to carry the HA, NA and M gene segments of Ty/05 virus were passaged in embryonated chicken eggs. The PR8 viruses were generated with the aKQ, aEQ and gKQ PB2 sequences and the Ty/05 virus carried the
aKQ sequence. After three blind passages in embryonated chicken eggs, no changes were detected in the PB2 gene sequence in any of these viruses (Figure 42). The introduction of PR8 or Ty/05 HA, NA and M genes to the virus with 50-92 PB2 aKQ motif prevented reversion to gKQ or aEQ, in contrast to 50-92/aKQ virus with its cognate HA, NA and M gene segments (Figure 28).

![Diagram](image)

Figure 42. Altering the genetic constellation of 50-92 prevented reversion of the 627K genotype during egg passage

RG viruses were generated by one passage in MDCK cells, then blind passaged in embryonated chicken eggs up to EP3, vRNA extracted from cell supernatant or egg fluid and the PB2 627 region amplified by RT-PCR and identified by sanger sequencing. Eggs were inoculated with <4HA units determined by HA assay using turkey red blood cells and incubated at 37°C (n=5). A. 50-92/aKQ/PR8HA,NA,M. B. 50-92/gKQ/PR8HA,NA,M. C. 50-92/aEQ/PR8HA,NA,M. D. 50-92/aKQ/Ty/05HA,NA,M.

Next we passaged the recombinant virus panel five times through MDCK cells, a mammalian cell line known to select for the 627K mutation (292, 327) (Figure 43). The sequence of the 627K region of Ty/05/K virus remained the same through each MDCK passage. In contrast, Ty/05/E reverted to PB2 627K by passage five. Remarkably, the lysine residue was selected by a double nt change (gaa to aag), despite a single nt being sufficient to switch to the lysine amino acid (gaa to aaa). Bioinformatic analysis revealed the majority of viruses with 627K carried the aag codon (Table 11). All the 50-92 recombinant viruses that had cognate HA, NA and M genes acquired mutations that resulted in the genotype gKQ during passage in MDCK cells, regardless of their starting genotype. Thus 50-92/aEQ virus mutated to 627K by passage 3, and also acquired the synonymous substitution a1869g
at this time. 50-92/aKQ rescued virus had already gained the synonymous mutation at nucleotide 1869 to become gKQ by the first passage, 50-92/gKR mutated to gKQ at the fourth passage and 50-92/gKQ remained stable during MDCK passage. This may be indicative of the fitness cost of PB2 628R observed in the 50-92:PR8 5:3 growth curve kinetics. Taken together these data provide further evidence of the genetic instability in 50-92 virus caused by the introduction of the PB2 627K mutation. The introduction of PR8 HA, NA and M into the 50-92 virus background abrogated this genetic instability and all three of these reassortant viruses remained genetically stable throughout 5 passages in MDCK cells, even aEQ. These observations suggest an RNA incompatibility between the 627K mutation in PB2 and the HA, NA and/or M of 50-92.

Figure 43. Altering the genetic constellation of 50-92 prevented reversion of the 627K genotype in MDCK cells
RG viruses were blind passaged in MDCK cells in T25 flasks (n=1). Cell supernatant was harvested at each passage, viral RNA extracted and the PB2 627 region amplified by RT-PCR and sequenced by Sanger sequencing. Chromatogram trace files for each virus and passage of PB2 nt 1867-9 and 1879-84 are shown. Solid arrows represent a nucleotide change. Dashed arrows show a possible minor population in the trace file.

4.3 Conclusions and Discussion

Taken together chapters 3 and 4 demonstrate the indifference of the clade 2.2 H5N1 virus to the humanizing PB2 627K mutation, in contrast to a non-Asian lineage classic H5N1 virus that does not tolerate the change. A summary of the data is provided in Table 12 which depicts the consequence of PB2 627K/E across the range of experiments used during the study. Of particular note is the altered phenotype that occurs for both virus backgrounds when gene segments of a H1N1 virus are introduced into the genetic constellation.

| Table 12. Summary of data from Chapters 3 and 4 comparing 627E and 627K phenotypes for Ty/05 and 50-92 virus |
|---|---|---|---|
| Polymerase | Ty/05 | 50-92 |
| 37°C | WT | +PR8HA,NA | WT | +PR8HA,NA,M |
| K>E | - | K>E | - |
| 41°C | K=E | - | K>E | - |

<table>
<thead>
<tr>
<th>Growth Kinetics</th>
<th>CEF</th>
<th>Ty/05</th>
<th>50-92</th>
</tr>
</thead>
<tbody>
<tr>
<td>37°C</td>
<td>K=E</td>
<td>-</td>
<td>aEQ&gt;&gt;aKQ</td>
</tr>
<tr>
<td>41°C</td>
<td>K&gt;E</td>
<td>-</td>
<td>aEQ=gKQ</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Virus Passage</th>
<th>EGG</th>
<th>Ty/05</th>
<th>50-92</th>
</tr>
</thead>
<tbody>
<tr>
<td>K and E stable</td>
<td>-</td>
<td>627E stable</td>
<td></td>
</tr>
<tr>
<td>K selected for over E</td>
<td>-</td>
<td>627K unstable</td>
<td></td>
</tr>
<tr>
<td>gKQ selected over aEQ</td>
<td>gKQ selected over aEQ, aKQ and gKR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>In vivo</th>
<th>Chicken</th>
<th>Ty/05</th>
<th>50-92</th>
</tr>
</thead>
<tbody>
<tr>
<td>K=E</td>
<td>-</td>
<td>aEQ=gKQ</td>
<td></td>
</tr>
<tr>
<td>Duck</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Table summarises data from Chapter 3 and 4. Coloured font reflects observed genotype advantage. (K=red, E=blue). a aag Lys codon selected instead of aaa, despite requiring an extra nt change. b IVPI score identical, Ty/05/E virus replication, measured by virus shedding, demonstrated small advantage over Ty/05/K. c observed gKQ slight advantage over aEQ at early timepoint. d 627K selected after passage through MDCK cells together with PB2 mutation a1869g. e chickens inoculated with unknown mixture gKQ and gKR, gKR was selected, gKQ virus was not detected throughout the infection. IVPI for aEQ was 2.88. IVPI for gKR was 1.32-1.59. f aEQ vs aKQ PR8 reassortant virus growth kinetic comparison not done. g aKQ genotype also stable when coupled with Ty/05 HA, NA and M.

In this chapter I have attempted to examine the mechanism behind the observations in Chapter 3. Although differences in polymerase activity in avian cells with either 627K or 627E were small, the 627K mutation consistently increased activity, particularly at the cooler
temperature of 37°C (Figure 36), and this phenotype has repeatedly been observed in the literature (316, 317, 327). However, this phenotype remains strikingly less dramatic than that observed in human cells. The effect may be purely biochemical such as altering enzyme kinetics, and not due to compatibility with other polymerase subunits or host co-factor (323). Alternatively, this in vitro effect may be an artefact of cultured cell physiology, offering altered expression of a co-factor such as chicken DDX17, which may account for the small increase in activity observed with 627K. In addition the 627K mutation has been associated with altered RNA binding in human cells where 627K appeared to decrease RNA binding (314). If this observation were also true for avian cells, it is unclear how this may relate to our observations.

Clearly, unlike the striking adaptation enabled by the 627K mutation to human cells, the same is not true for avian cells, suggesting host factor present in avian cells are not restrictive to either 627E or 627K. As previously mentioned, DDX17 has been proposed as the PB2 627-specific mammalian host restriction factor by Aggarwal et al. (2011) (313). However, a criticism can be made of the data by which the authors concluded their findings; the effect on polymerase activity by altering expression of DDX17, said to be 627K specific, was only modest and does not fit with the striking phenotype often observed in the field. The DDX17 study did not deduce the mechanism by which DDX17 affects polymerase or demonstrate a direct interaction of PB2 627K and DDX17. In regard to chicken DDX17, it was mentioned that a 3 fold and 2-fold reduction in the activity of polymerase carrying 627E or 627K, respectively, was caused by siRNA knockdown of chicken DDX17 in avian cells, suggesting chicken DDX17 did not show a preference for E or K in chicken cells. The suggestion by Aggarwal et al. that human DDX17 acts to antagonise avian origin polymerase does not fit with observations by Moncorgé et al. (2010) (317). Here, fusion of human cells to avian cells demonstrated the presence of an avian-cofactor to increase avian polymerase activity in human cells; conversely, fusion of human cells to avian cells containing 627E polymerase did not result in the inhibition of polymerase activity.

We did not demonstrate if the introduction of PB2 627K into the 50-92 polymerase complex caused an increased error rate, which may have provided an explanation for the multiple mutations observed during virus passage. However, we found no great alteration in polymerase activity, which might be required for the accumulation of errors. Moreover, by
generating reassortant viruses we demonstrated that mutation was prevented by the HA, NA or M gene segments, which are not known to function in the polymerase complex.

The profound genetically instability of the engineered 5092 PB2 627K gene was circumvented when the internal gene segments of 50-92 were reassorted with the HA, NA and M of PR8 (H1N1) or of Ty/05. In these mosaic genotype viruses, 50-92/aKQ no longer reverted to PB2 1869g. This suggests an interaction between the PB2 gene segment and other gene segments in these viruses that can be disrupted by mutations around the 627 coding region. The significance of the a1869g mutation was highlighted when 50-92/aEQ virus was passaged though MDCK cells and it was selected together with the nonsynonymous mutation that encoded 627K (Figure 43). Interestingly, passage of Ty/05/E virus (Glu codon gaa) through MDCK cells selected for the 627K change as expected but favoured a double nucleotide change (Lys codon aag) rather than a single nt change (Lys codon aaa). This may further demonstrate the importance of the RNA sequence in this region. Virus sequence databases revealed the majority of viruses encoded 627K by codon aag, as well as natural examples of the 1869g mutation (Table 11). We engineered 50-92 627K by a single nucleotide change from gaa to aaa. A worthy experiment for the future may be to engineer a 50-92 627K virus utilising a aag Lys codon, although this codon sequence was not detected in our passaged virus that selected other compensatory mutations.

The only described cis-acting RNA functions in influenza segments thus far are packaging signals. These have been mapped for individual gene segments in various viruses (reviewed 61) and some work supports the idea that some RNA segments are more dominant than others in directing packaging (15, 67, 411). Work by Gog et al. (2007) demonstrates how single nucleotide mutations can substantially reduce incorporation of the gene segment during virus particle assembly (68). However, PB2 codons 623, 627 and 628 were not mentioned in this study and packaging signals of the non-coding and codon regions for segment PB2 have not been mapped to the 627 region (412). Disruption of RNA-RNA interactions between segments affects packaging of the viral RNPs (15, 66, 67). In particular, PB2 vRNA has been shown to interact with NA vRNA using a H5N2 virus model (15). We did not attempt to model RNA secondary structure changes by our mutations. Such a computational approach would unlikely explain our findings, as influenza vRNA is known
to be bound to NP (413) and this knowledge has recently been confirmed by solving the vRNP structure to such an unprecedented resolution that it indicates the vRNA is exposed on the external surface of the NP monomers (282, 414). Comparison of the RNA sequences of 50-92 and Ty/05 reveals a significant lack of homology, particularly for the NS and PB2 sequence (Table 13). Given this observation, diversification of packaging sequences during virus evolution may not be entirely surprising. One approach that may reveal conserved packaging signals could be a sliding window analysis, ignoring the difference in gross homology, but possibly highlighting smaller areas of conservation.

Table 13. Describes the RNA sequence homology of the gene segments between 50-92 and Ty/05

<table>
<thead>
<tr>
<th>Gene segment</th>
<th>Sequence homology (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PB2</td>
<td>87.1</td>
</tr>
<tr>
<td>PB1</td>
<td>92.3</td>
</tr>
<tr>
<td>PA</td>
<td>90.1</td>
</tr>
<tr>
<td>HA</td>
<td>91.0</td>
</tr>
<tr>
<td>NP</td>
<td>93.8</td>
</tr>
<tr>
<td>NA</td>
<td>91.9</td>
</tr>
<tr>
<td>M</td>
<td>93.1</td>
</tr>
<tr>
<td>NS</td>
<td>70.4</td>
</tr>
</tbody>
</table>

We speculate that one or all of the cognate HA, NA and M vRNA segments of 50-92 may have interfered with the packaging of the vRNA carrying the PB2 627K mutation. However, it remains uncertain whether the rescued phenotype observed by the addition of PR8 or Ty/05 HA, NA and M is a result of vRNA interactions or another unknown mechanism increasing virus fitness. A hypothetical model of how RNA interactions of the aKQ genotype may affect packaging is demonstrated in Figure 44.

Since the PB2 628R mutation was selected for during egg passage and upon infection of chickens, we sought to examine the structural consequence of this mutation for the PB2 protein. Although it provides a change in electrostatic charge, this change may be tolerated in the linker region of the 627-domain, which may explain its positive selection. Interestingly, when engineered into a reassortant 50-92 virus carrying PR8 HA, NA and M, PB2 628R was associated with a small fitness loss (Figure 40), a phenotype not displayed in conjunction with the 50-92 equivalent gene segments (Figure 35). Since 628R was selected for during whole virus 50-92/gKQ passage, the fact that the amino acid change may actually
cause a fitness cost, supports the hypothesis that the nucleotide change at the 628 codon outweighs this cost by compensating an RNA-dependent mechanism. Alternatively the fitness cost observed with the PR8 segments may describe a segment mis-match between 628R and PR8 HA and/or NA. In addition, the observation that Ty/05 with 627K gained fitness from reassortment with PR8 HA and NA (Figure 41), may reinforce the finding that the effect of the 627K sequence is dependent on the origin of the HA and NA gene segments. It is unclear why the compensatory mutations occurred on the same gene segment, although since the HA and NA genes were not sequenced after virus passaging, it cannot be ruled out that these gene segments may also harbour mutations to prevent mis-matching.

Similarly, genetic manipulation of a HP H5N1 revealed a fitness cost by 627K that was previously masked. Schat et al. (2012) showed a small fitness cost associated with Vietnam/04 H5N1 and PB2 627K, but only with a LP recombinant version where the MBCS of HA was removed. It would be interesting to see if the same phenotype was demonstrated by Ty/05. If Ty/05 HA was engineered to carry a single-basic cleavage site, would the fitness of the virus be reduced by such a degree that a fitness cost associated with 627K could now be observed?

As mentioned in the introduction, H5N1 HK156 virus has been extensively used as a model H5N1 virus and has demonstrated a failure to select the 627K mutation and favoured the PB2 701N mammalian adaptation motif instead. Since this is an early H5N1 virus, a pertinent question to ask would be: Would HK156 suffer the same fitness cost caused by PB2 627K as observed for 50-92? Work by Nakazono et al. (2012) demonstrated the ability of 627K to increase polymerase activity of HK156 in human cells, but this work was only conducted with reconstituted polymerase, not in the context of virus (415). Bogs et al. (2011) demonstrated the failure of HK156 virus to select PB2 627K in mammalian cells, and suggested this was due to an incompatibility with the NP protein, since HK156 NP prevented a clade 2.2 virus to select 627K. The presence of PB2 701N was not mentioned in this study, which may have greatly influenced the likelihood of HK156 to select PB2 627K. A useful experiment would be to lock the 701D codon and reassort HK156 with the clade 2.2. NP to observe if the PB2 627K mutation may now be selected. At this time alternative mechanisms such a packaging signals, as hypothesised for 50-92 virus, cannot be excluded for HK156. A
closely related HK/97-like virus, A/HongKong/483/97, was found to carry the PB2 627K mutation. Isolated from a human case and able to cause lethal infection in mice, the virus was clearly able to replicate effectively despite carrying the PB2 627K motif (285).

Given the observation from Bogs et al. (2011) we sought to examine whether the identity of NP affected polymerase with 50-92 and Ty/05 PB2. We postulated that the clade 2.2 NP could be necessary to support polymerase activity when PB2 had 627K and may explain our observations for the fitness cost of 50-92 virus bearing PB2 627K. However, our results suggest that 50-92 NP did not detrimentally affect the polymerase activity of 50-92/aKQ. In fact, 50-92 NP slightly increased polymerase activity when combined with the Ty/05 polymerase complex in avian cells, and this increase was independent of PB2 residue 627 (data not shown). Moreover, subsequent investigations revealed the role of RNA sequences, beyond the protein level at the polymerase.

Taken together these data may have important implications for the surveillance of emerging influenza virus infections and for risk assessments of pandemic potential. Pandemic H1N1 virus of 2009 did not carry the 627K mutation, instead it adapted to humans via the PB2 590/91 SR mutation. It is important to understand why viruses may take different routes of evolution. Studies have attempted to introduce the 627K mutation into the pH1N1 background. Several laboratories have shown no fitness cost or increase by introducing PB2 627K into pH1N1 containing 590/91 SR (296, 298, 299) and one study demonstrated a small fitness cost (300). As mentioned previously, ostriches can select mammalian adaptation motifs in PB2. Analysis from Yamada et al. (2010) revealed ostriches infected with an Avian H5N1 isolate had examples of 591K as well as the 627K mutation (298). This suggests that certain H5N1 viruses may also be able to tolerate a 590/91 change as opposed to 627K or 701N upon mammalian adaptation.
Figure 44. A hypothetical model explaining the 627K phenotype observed for 50-92 virus

Viral RNPs are attracted by complementary RNA interactions on exposed vRNA bound to NP, such as 1869g. These interactions may enable gene segment incorporation into the virion. Insufficient RNA interactions may result in poor accumulation of the 8 gene segments and prevent budding and/or viable infectious particles.

Limitations and future work

This work highlights that the codon sequence rather than the amino acid affects virus fitness. However, the study has not investigated the effect of different codons in this region. For example, a 50-92 virus could be generated with the aag Lys codon and the effect on virus fitness measured. We speculate that the sequence changes in this region affect vRNA interactions. A worthwhile experiment to investigate this in the future could be to measure the dimerization of different vRNA segments expressed in vitro, such experiments have been demonstrated by Fournier et al. (15, 66, 67).
Chapter 5. Discussion

Summary of future work

Chapter 2

- The pathogenic phenotype of the 50-92-HA₄ 4R virus to be determined *in vivo* by IVPI index.
- The generation of viable single basic HA 50-92 viruses may be a useful tool to examine pH of fusion by haemolysis assay.
- Analysis of expression of the HA constructs should be conducted again in a suitable cell line.
- Pseudotyped lentiviral vectors carrying the different 50-92 HA proteins will be generated and used to measure the fusion of pH in combination with cellular pH altering drugs.
- Monte Carlo analysis on randomly selected amino acids on the HA would enable statistical significance to be applied and significance of the three HA mutations to be estimated.

Chapters 3 and 4

- The effect of PB2 627K on different wild bird species may be worth studying in the future, such as migrating species or birds of prey.
- Engineer a 50-92 627K virus with an aag codon and compare its fitness to that of the virus with the aaa Lys codon.
- Measure vRNA interactions by the level of dimerization by different vRNA segments expressed *in vitro*.
- Perform sliding window analysis on the different gene segments to identify possible conserved areas that may act as packaging signals.
Virulence motifs for H5N1 avian influenza

We sought to generate a HP H5N1 virus from a previously established reverse genetics system based on a clone of virus isolate, 50-92. Surprisingly the RG virus recovered from the original system had not led to an HP phenotype by IVPI. By the introduction of three mutations in the HA gene, described in the sequence of a different clone from the same 50-92 isolate, the HP phenotype was recovered. The increase in pathogenesis was not attributed to the MBCS of the HA, since both the LP phenotype RG virus and HP RG virus had an identical MBCS.

We suggest that two of the mutations affected viral binding to sialic acid and also altered the pH stability of the HA, and that the third compensatory mutation was required to enable efficient HA fusion. Without this crucial compensatory mutation, virulence was significantly reduced and the virus lost its pathogenic phenotype. A similar concept has been described by two recent studies that successfully adapted H5N1 to the ferret host. Mutations that altered sialic acid binding affected HA stability and were compensated by mutations in the HA fusion pocket, shown by Imai et al. (2012) and Herfst et al. (2012), although the latter study did not demonstrate this mechanism directly (87, 88).

Since the precursor virus was never identified, we can only speculate at the virus origins of the 50-92 outbreak in 1991; likely an introduction from a wild bird. AI introduction from the avian reservoir appears to require and select for adaptation to poultry. The selection of a MBCS from a single-basic precursor by the H5 and H7 subtypes in poultry has often been observed, and more recently, HP H5N1 viruses carrying a MBCS are found in the AI reservoir. Studies have also highlighted that a shortening of the NA stalk is often selected for in poultry, although not always. The glycosylation of the binding site appears to be related to this shortening, but little is understood about receptor binding changes between different avian species (135, 371).

The diagram below describes the effects of AI adaptation to poultry and raises a pertinent question on whether these changes have implications for other influenza host species including humans and indeed re-entry into the avian reservoir. Understanding the adaptation of HA and NA to poultry require further investigation. With poultry reportedly expressing more SAα2,6Gal than other avian species, gaining insight into poultry adaptation...
may have implications for understanding AI virulence but also how that might predispose for human adaptation (374, 375).

The route to successful adaptation to a new host appears to be complicated, involving many different interconnecting factors. A further complication may be the environment and crucially transmission efficiency. Pressures external to the host can also influence influenza evolution, such as persistence in freshwater, dependent on the stability of the virion proteins and envelope.

**HA fusion is an emerging host range restriction factor and virulence mechanism**

The role of HA fusion as a host range restriction and virulence factor has recently become apparent (416). Here we present a mutation in the fusion peptide of a H5N1 virus that was crucial to the pathogenicity of the chicken host, HA₂ G4R. Previous studies have correlated an increase in HA fusion with increasing pathogenicity in chickens (93). The mutation described in this thesis is unusual as it resides directly in the fusion peptide, in contrast to previous studies which describe mutations in the fusion pocket and HA trimer interface (87, 91, 358). Understanding which pH of fusion is required for specific hosts and the effect on pathogenesis is important since it can help guide anti-viral and vaccine development as well as point out high risk mutation motifs. Anti-viral compounds that prevent HA fusion have been described, although none have yet proved suitable for
commercial license (417). An exception is a commonly used drug in Russia and China that claims to treat influenza infection by blocking fusion, Arbidol, although this is not licensed in the West (418). The discovery of fusion related mutations has also led to their introduction into vaccines, improving infectivity and immunogenicity (419).

**Other virulence mechanisms beyond the MBCS**

Several attempts have recently been made to insert a MBCS into different HA subtypes and measure their ability to support a HP phenotype. It is interesting to see that some HA subtypes supported the HP phenotype, H2, H4, H6, H8, H9 and H14, whereas H1 and H3 did not (170, 172, 173)(170, 171). It highlights the fact that pathogenesis can be caused by a number of factors. Whether the HA subtypes that failed to transform into HP viruses lack virulence factors on other gene segments, or the HA protein itself is incompatible with the MBCS is unclear.

Systemic replication has even been described by viruses without a MBCS. LP avian influenza viruses, characterised by a monobasic cleavage site, have on occasion been isolated beyond the respiratory or intestinal tract. LP H5 or H7 viruses with a single-basic cleavage site have been found in the pancreas, kidney and oviduct in infected chickens (420), however this may not be surprising since these subtypes are known to support systemic replication. In the case of H9 viruses, the evidence is more convincing. Several reports have described H9N2 virus replicating in the kidney of infected poultry (421–423), where matriprase is expressed at high levels and able to cleave the H9 HA protein (422). Whether replicating in the kidney provides influenza with any advantage is unclear, but the role of increased tropism in the recent endemic spread of H9N2 influenza warrants further investigation. Acquiring a better understanding of the binding profile and function of the HA in avian species may elucidate virulence mechanisms and the potential of avian viruses to infect humans.

**Host restriction between avian species**

Little is understood in the field about the adaptation of influenza viruses to different avian species and the role this plays in the ecology of AI. An obvious point of restriction may well be in attachment. Little is known about the differences in receptors used by the virus in
different avian species. Although variation in the distribution of SAα2,6Gal has been reported (374, 402), knowledge of the exact composition of these sugars is lacking for avian species. Beyond attachment, differences may also lie between the host cell environments of avian species. Indeed, it has been shown that chickens lack some components of the innate immune cellular pathways, such as RIG-I, present in, ducks (210). Such differences may also account for their differing susceptibilities to HPAI.

One incompletely answered question pertinent to the viruses used in this study is why ducks fail to support replication of H5N1 50-92, shown by Wood et al. (1995) by intranasal inoculation (366), but permit Ty/05 replication (396). Kuchipudi et al. (2011) demonstrated an increased level of DNA fragmentation in duck cells infected with 50-92 compared to chicken cells, a phenotype not demonstrated by Ty/05 (Figure 45). It was thought this was due to apoptosis induction. Since work in this thesis had generated two reverse genetics systems for the relevant viruses on which this work was based, by collaboration we attempted to investigate this hypothesis and map the responsible gene segment.

![Figure 45. Duck cells demonstrate higher levels of DNA fragmentation following 50-92 infection (369)](image)

Chicken cells (white bars) or duck cells (black bars) were infected with either 50-92 or Ty/05 H5N1 viruses. The level of DNA fragmentation was measured by propidium iodide staining of ethanol fixed cells by flow cytometry. Taken from Kuchipudi et al. (2011) (369).

Several RG reassortant viruses were generated between 50-92, Ty/05 and PR8; in particular, NS single gene reassortant viruses were generated in the H5N1 background (Appendix iv, Table 25). Virus growth curves of NS reassortant 50-92 and Ty/05 viruses revealed 50-92 NS provided a small fitness advantage for Ty/05 virus in chicken cells, but not
in duck cells (Figure 46), possibly highlighting a reduced efficiency of the 50-92 NS1 or NEP function in duck cells. Furthermore, preliminary results using the MTT assay which measures cell proliferation, suggested greater cell death was induced in duck cells infected with a PR8 virus carrying the NS segment of 50-92 compared to that of Ty/05 NS (data not shown). However, the MTT assay failed to duplicate these results in the HP H5 background; Ty/05 caused exceptionally high readings of cell proliferation regardless of NS origin and cell type, possibly due to cell membrane disruption preventing the assay from providing reliable comparisons as this may actually have been a measure of cell death. Disruption of the cell membrane by modern H5 viruses has been demonstrated previously, advocating a role in apoptosis in duck cells, in contrast to the hypothesis of Kuchipudi et al. (424). Furthermore, attempts to measure caspase induction, suggested Ty/05 induced greater apoptosis in duck cells than 50-92 (data not shown). However, these results are inconclusive due to technical difficulties within the high containment environment. Further work using suitable assays may provide more conclusive evidence about the restriction of 50-92 virus in ducks. In addition, the reassortant viruses generated in this study could be used to identify other virulence factors carried by modern H5 viruses such as Ty/05. Caution is required in this approach since it may be difficult to decipher virulence mechanisms due to incompatibility of virus genes. Indeed, certain combinations of 50-92 and Ty/05 appeared to be attenuated (Appendix iv, Table 25). Interestingly the M gene of 50-92 was restrictive for RG viruses when combined with PR8 HA and NA, a phenotype overcome by Ty/05 M.

![Figure 46. Virus growth kinetics of 50-92 and Ty/05 NS reassortants in CEF and DEF cells](image)

CEF cells (A) and DEF cells (B) were infected with 50-92 and Ty/05 viruses together with their NS reassortants at an MOI 0.001. Virus replication was measure by qRT-PCR of the M gene (50-92) or HA gene
(Ty/05) vRNA in the cell supernatant. Statistical analysis was carried out by Two-way ANOVA, adjusted for Bonferroni post-test. **** P<0.0001.

Role of quasispecies and pathogenicity

An original hypothesis addressing the failure of the 50-92 (2L) RG system to generate a HP virus asked whether a reduced quasispecies as a result of virus generation from cDNA caused a reduced pathogenesis. This does not appear to have been the case, and correction of the HA has proved the suitability of RG influenza viruses for use in experimental systems replicating observations from clinical isolates. However, there may still be some merit behind this speculation. Increased fidelity of the poliovirus polymerase by a single amino acid change, induced by passaging in the presence of ribavirin, resulted in reduced virulence in vivo (425–427). To our knowledge such an examination has not been adopted for influenza and may be worthy of investigation, especially since influenza is susceptible to the effects of ribavirin suggesting this approach may be fruitful. An influenza virus attenuated in this way could prove useful for vaccine development and fundamental understanding of virulence mechanisms. On the other hand, a polymerase that is too error prone and reaches error catastrophe is clearly detrimental to virus. This is likely the basis for the new antiviral drug favipiravir (428). If there is a link between error rates and pathogenicity, it might be important to consider during the development and roll out of such drugs in the clinic.

Why we need to control avian influenza?

The current world population of 7.2 billion is projected to increase by almost one billion people within the next twelve years, reaching 8.1 billion in 2025 and 9.6 billion in 2050, according to a recent United Nations report (429).

This, together with the effects of climate change and decline in global economic growth presents a major challenge for the world in the future. The current challenge of feeding the world has been met with some success, with the Millennium Development Goals reporting an expected 50% reduction in the number of people who suffer from hunger by 2015, since 1990 (430). However, there is still a great deal of work to achieve Goal 1: ‘Eradicate extreme poverty and hunger’, with an estimated 870million people currently undernourished.
Since 1990 the use of poultry to feed the world has increased at a greater rate than any other livestock. The production of chicken meat has doubled in 10 years to a gigantean 99.1 million tonnes in 2010. The same is true for egg production, reaching 69.1 million tonnes (431) (Figure 47). This highlights the rising dependence of poultry as a food source for the world and raises the concern to control the effects of avian influenza.

LPAI may present itself in poultry as a mild respiratory disease, and cause egg drop for layers, whereas HPAI can cause high mortality. Successful prevention against AI in the poultry industry is primarily by the implementation of biosecurity, preventing contamination and spread by disinfection, confinement and control of movement. Secondly vaccination can be applied when possible. Vaccination against AI in poultry has increased significantly in the past 10 years, although it is not practised in the UK. Most commonly, vaccination is employed against HPAI of the H5 and H7 subtypes, and LPAI H9N2 (432). Vaccine use varies globally, and is most often reactive to recurrent HPAI outbreaks in the related geographical regions, particularly in countries where the virus is endemic. Problems associated with vaccines as a control strategy include their efficacy and effectiveness, and the requirement for multiple doses in the life of a layer chicken to be successful. In addition the nature of influenza evolution dictates the requirement to redevelop vaccines for different strains and over time. HPAI outbreaks may be controlled by attempts at culling the population. Both the disposal of infected carcasses, active culling of nearby poultry populations ‘at risk’, as well as vaccination in surrounding areas can help prevent spread to a wider area.
Figure 47. The world production of poultry livestock between 1990 and 2010

A. Type of meat production in millions of tonnes worldwide. B. Recorded egg production in millions of tonnes worldwide. Data taken from FAO Statistical Yearbook 2013, Part 3 ‘Feeding the World’ (431)

As mentioned in Chapter 1, there is a great deal of surveillance of influenza viruses. This can aid in the planning and decision making of biosecurity and to inform local risk assessment. Biosecurity can be increased if surveillance detects viruses of concern. Here the investigation and identification of virulence motifs or markers of pandemic potential can directly inform such decisions. Well informed judgement is essential to maintain the viability of the industry due to the cost of biosecurity and to maintain sufficient production of poultry meat and eggs.
Designer chickens?

Another approach to prevent AI in poultry other than by disease control is to alter the poultry themselves. We discussed in Chapter 1 the differing susceptibilities of avian species to AI. Chickens, and turkeys in particular, are highly susceptible to AI especially HPAI.

After the publication of the first draft of the chicken genome in 2004, its analysis has led to a greater understanding of chicken genetics (433). Moreover using transgenic technology, it is possible to generate chickens with altered genomes (434, 435). Researchers at the universities of Cambridge and Edinburgh have developed a transgenic chicken that carries a virus decoy. Chicken cells were engineered to express decoy RNA hairpins designed to interfere with the influenza polymerase and attenuate the virus. *In vivo* challenge of these birds by H5N1 virus showed the transgene offered limited resistance to disease but did prevent transmission to other birds, (436). Prevention of spread to other chickens and indeed any other host not only prevents loss of poultry but could limit the pandemic potential of AI derived from poultry.

The proof of concept behind transgenic chickens and increasing the resistance to AI or controlling AI spread is an exciting prospect. A better understanding of AI interaction with chicken cell factors and other avian species, such as Mx variants and RIG-I, may further development of a transgenic chicken less susceptible to AI infection. Would a chicken expressing RIG-I from a duck be more resistant to HPAI? The discovery of AI resistant polymorphisms in the chicken genome such as residue S631N of chicken Mx could also be applied by transgenic technology (437–439), although the efficacy of chicken Mx in any of its isoforms against influenza has been questioned (438). Interestingly, expression of mouse Mx in chicken cells has been shown to increase resistance to influenza infection (440). These concepts could also be applied to transgenic swine, as some breeds contain deficient Mx alleles (441). Indeed, the generation of a transgenic pig line expressing mouse Mx1 has been attempted, although the method used proved problematic for the organism and gene recombination prevented protein synthesis (442).

Eradication of influenza virus from the globe remains impossible. Although within domesticated species it could be a possibility. In the not so distant future, cross-reactive vaccines could provide a viable method of influenza control both in humans and in
domesticated animals, although generation of appropriate immunogens has yet to be achieved and cost and logistics would likely be problematic (443).

As a commercial venture in more economically developed countries, society may not be ready for genetically modified meat, but there may be a time in the future when circumstances and opinions change, and the scientific community and poultry industry should be ready for this.

Is H5N1 still a pandemic threat?

PB2 627K is a mammalian adaptation mutation and a virulence factor. Our results assessed the effect of PB2 627K for avian influenza H5N1 in the avian host. We showed there was neither a positive or negative fitness associated with this mutation within a clade 2.2 background, a result supported by its maintenance in the current avian influenza reservoir. This was in contrast to the detrimental effect in an ‘ancient’ non-Asian H5N1 background, which caused a strong attenuation that we suggest is due disruption of RNA interactions between gene segments. These findings are presented in the next diagram, highlighting their broader consequence.

Understanding the maintenance of human adaptation motifs in the avian reservoir is critical, particularly in the case of HPAI H5N1. It is worthy to ask if avian viruses other than 50-92 have the same inability to accept the 627K mutation. In addition it could be interesting to observe whether other PB2 adapting mutations such as 701N are supported by 50-92 virus. The PB2 701N mutation increases polymerase activity of 50-92 in human and pig cells, although this has not been explored in the context of virus infection (318). Unravelling this observation may help distinguish which avian subtypes pose the greatest pandemic risk. The fact this 627K restricted phenotype has not been observed experimentally for other strains so far may indicate that this is rare. On the other hand, the lack of the 627K mutation in the avian reservoir could demonstrate that the 627 domain is often restricted in evolution due to cis-acting vRNA interactions.
Worryingly, the clade 2.2 viruses typically carry additional human adaptation motifs other than PB2 627K, for example in the HA. Russell et al. (2012) have shown the large number of viruses isolated that carry PB2 627K and have lost the 158 glycosylation site in the HA protein, mainly those belonging to clade 2.2 (308). An observation confirmed by our own analysis in collaboration with Dr Sam Lycett, Edinburgh University (data not shown). This demonstrates the accumulation of additional mutations and the increase in the pandemic potential of these viruses. Although H5N1 viruses have the ability to infect and cause disease in humans, so far they have not caused a pandemic due to their lack of transmission between humans. Table 14 describes the predicted mutations necessary for H5N1 human to human transmission, based on the ferret model by Imai et al. (2012) and Herfst et al. (2012). The similarities between these two independent experiments are striking, as is their comparison to Ty/05, a typical clade 2.2 virus. Others in the field have commented on the very few mutations required by some H5N1 viruses, particularly in Egypt which has experienced a significant number of human infections in recent years (444). Interestingly, Ty/05 carries a second mammalian adaptation motif present on the NP protein, V33I. This has been previously identified by bioinformatic analysis of mammalian and avian influenza virus isolates (305, 306, 445).
The role of HA fusion in host restriction is little understood. It appears human influenza viruses favour a more stable HA fusing at a lower pH (87–89, 94). The ability to map these mutations and make predictions may be more difficult, as so far it appears there could be many different mutations whereby the virus could alter pH stability, including our own example of HA₂ G4R by 50-92, as well as several mutations described by Russell and co-workers (93).

Table 14. The mutations required for H5N1 respiratory droplet transmission in the ferret model, compared to the genotype of a typical clade 2.2 virus, Ty/05

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N154D</td>
<td>T156A</td>
<td>Loss of glycosylation</td>
<td>✓ 156A</td>
</tr>
<tr>
<td>Q222L</td>
<td>Q222L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N220K</td>
<td>G224S</td>
<td>Receptor binding switch</td>
<td>✓ 220K</td>
</tr>
<tr>
<td>T314I</td>
<td>H103Y</td>
<td>Fusion of pH?</td>
<td>? *</td>
</tr>
<tr>
<td>PB2 590/91 (H1N1)</td>
<td>PB2 E627K</td>
<td>Overcome polymerase restriction</td>
<td>✓ PB2 627K</td>
</tr>
</tbody>
</table>

* the nature of these mutations appear to be diverse and therefore could prove difficult to map.

Despite the appearance of H5N1 viruses teetering on the edge of human adaptation, the fact that these viruses have circulated since 2005 and have not yet caused a pandemic could imply yet further restrictions down the human adaptation path that have not yet been described in the ferret transmission experiments.

Recently, Blumenkrantz et al. (2013) demonstrated the limitation of influenza transmission in the ferret model conferred by a short stalk NA (79). 96% of all H5N1 isolated viruses carry a short stalk NA (79), including Ty/05, and this may restrict their efficiency at human to human transmission. Either the stalk must be lengthened by insertion of genetic material (by recombination for example) or a reassortment event must take place which introduces an adapted human NA. In addition, results by Herfst et al. (2012) suggest that the pathogenicity caused by the transmissible H5 virus is reduced in the ferret compared to WT. This is an important consideration when preparing for such an event.

An interesting consideration is how the clade 2.2 viruses are maintaining the loss of a glycosylation site together with the PB2 627K mutation? Another interesting feature of
these viruses is their replication in the respiratory tract as opposed to the gut in ducks. One hypothesis is that the tropism shift from gut of ducks to respiratory tract may have been driven by the increased replication of the virus carrying the cold adapting 627K mutation, since the respiratory track is a cooler environment than the gut. This may then have selected mutations in the HA due different expression of receptors on different cell types in the duck. Our experiments did not support this hypothesis in ducks infected with Ty/05 627E or K since no significant difference in tropism was observed. However, the HA between these viruses was identical and may have dictated the tropism in a way that would not allow this hypothesis to be adequately tested. An interesting experiment of the future may be to map this change in tropism. Comparison of the Ty/05 HA sequence with a H5 virus that has not altered tropism may highlight potential mutations that have enabled the change. These residues could engineered into Ty/05 HA and their effect measured. If the tropism is successfully altered, a comparison of viruses bearing PB2 627K or E could highlight fitness differences in either shedding location.

During our bioinformatic analysis of PB2 627 we discovered several avian viruses encoding valine at this position, of the H9N2 and H1N1 subtype (Table 7). Several H9N2 viruses isolated from Israel have been reported to carry this motif (397). The PB2 627V sequence has also previously been reported for swine H1N1 isolates (307). It may be worthwhile to investigate whether the 627V is able to overcome the restriction of avian polymerase in mammalian cells, it is possible that this introduction disrupts the negative charge of the 627 region, in a similar manner to the 590/91SR mutations (296). A study by Mehle and Doubna (2008) demonstrated the introduction of alanine, another small neutral residue, at PB2 627 partially permitted polymerase activity in human cells, compared to the 627E mutant that was severely restricted (319).

**What other avian viruses currently pose a significant pandemic threat?**

**H7N9**

H7N9 viruses were first detected in humans in February 2013, China, and have so far caused 135 reported cases, including 44 deaths (446, 447). The considerably high fatality
rate is characterised by respiratory distress, virus present in stool and secondary bacterial pneumonia (448). The human population is naive to the H7 subtype. Of concern are reported human to human transmission events, although the virus has not demonstrated widespread transmission (449, 450). Indeed, the number of human cases has decreased significantly since April, and authorities will be watching cautiously in the start of the influenza season in the coming months.

Human isolates of H7N9 have been tested in the ferret influenza model by several laboratories and the virus has demonstrated various levels of respiratory droplet transmission between ferrets, although generally limited (451–456). The surprising finding that there is any transmission of H7N9 between ferrets may be due to the fact that the virus carries certain human adaptive mutations. The HA carries the 226L mutation that humanizes receptor preference, and PB2 has been found with 627K or 701N in human isolates (454, 457). The selection of the PB2 627K mutation suggests there is no restriction, similar to our observations with Ty/05 and in contrast to the situation with 50-92. The internal gene segments of the H7N9 virus are reportedly derived from H9N2 strains, and H9N2 viruses have previously been described to select the PB2 627K mutation in mammals and generally support a high level of reassortment with other virus subtypes (389, 446, 458–463). It appears the virus has been transmitted to humans by poultry which have demonstrated asymptomatic infection (464), and so far, viruses isolated from chickens appear to maintain the PB2 627E mutation (464).

The zoonotic behaviour of these novel H7N9 viruses is reminiscent of the H5N1 situation in 1997. An interesting question is whether the H7N9 viruses will maintain the HA and PB2 mutations in the avian reservoir. No studies have yet tested whether the 627K mutation confers a fitness costs in the avian host for this virus. Sequences on the NCBI influenza database reveal all H7N9 avian isolates carry the 627E mutations, together with a mix of HA 226Q (Avian) or 226L (Human) (NCBI influenza database resource, accessed August 2013). The latter may well provide evidence of re-introduction from human infection back into birds, although some H9N2 viruses have reportedly carried this mutation for a number of years (465).
**H9N2**

Along with H5 and H7, H9 viruses have also caused sporadic infection in humans (465–467). Moreover, LPAI H9N2 has become endemic in poultry in Asia and the Middle East (118). The H9N2 viruses are of particular concern as a potential pandemic since recent human and avian isolates of H9N2 carry the same human adaptation motif as H7N9, HA 226L. This mutation has been shown to increase SAα2,6Gal binding and replication in ferret and human airway epithelial cells (462, 468, 469). Viruses of this subtype have also successfully selected the PB2 627K mutation upon replication in mice (404). Together this knowledge warns us of the high risk these viruses pose for a pandemic. With many poultry infections the pandemic risk increases, as a matter of chance due to the greater number of H9N2 infections. Yet clearly H9N2 requires further adaptation to cause a pandemic. Reassortant events with human H1N1 and H3N2 viruses combined with some adaptive mutations in HA, that map both to receptor binding changes and HA stability, has permitted aerosol transmission of the H9 HA in the ferret model (465, 466, 470). These studies suggest further restrictions of the pandemic potential of the H9N2 viruses lie in the NA and/or internal genes. Indeed, both N2 and N9 genes of the two highlighted poultry viruses discussed here harbour NA stalk truncations although neither as large as that for the N1 NA from H5N1.

**H2N2**

The pandemic of 1957 was caused by a H2N2 virus and descendants of this continued to circulate in the human population until 1968, where it is believed they were displaced by the H3N2 pandemic (471). Structural studies have examined its ability to bind human receptors and note the important of the 226L mutation to enable efficient SAα2,6Gal binding (472). H2N2 viruses originating from the 1957 pandemic still continue to circulate in avian and swine populations and structural and binding analysis reveals these viruses still maintain a degree of human receptor binding specificity (473). It has been argued that the waning immunity of the elderly population is disappearing, and the younger population is vulnerable to the H2 subtype (474). Taken together it seems likely that H2
viruses could generate another pandemic in the future. There has been a surprising lack of H2N2 influenza research in recent years.

Should ‘gain of function’ experiments be permitted?

The experiments presented in this thesis, where the PB2 627K mutation was engineered into the H5N1 virus 50-92 may be regarded as ‘gain of function’. The term simply describes the forward adaptation of the H5N1 virus to humans. For many years scientists sought to map the mutations necessary by H5N1 to transmit between humans, using the ferret model. In 2012 two teams were successful and attempted to publish their results. The work was met with surprise by US government regulators, media and other research scientists. Concern was raised over the possible ‘dual use’ of this work. Dual use research is defined as work that provides benefit for society but can also lead to a potential for misuse, e.g. bioterrorism (475). Debate ensued and publication of the results was delayed, although completed sometime later. Some doubt the merit of such research whereas others argue the clear benefits of the knowledge gained, such as better knowledge of vaccine development and anti-viral treatment by modelling the predicted mutations. The restriction of the results to prevent this information being misused is unfavourable and likely impossible. The very nature of science is the sharing of knowledge and restricting this could be extremely detrimental to future development. Restricting such information to certain individuals is unfair as it could lead to monopolies and prevent other worthy researchers developing their ideas. Both sides of the argument agree the requirement for appropriate biocontainment and biosafety.

Appropriate risk assessment was of course in place for the work in this thesis. 50-92 virus is a Non-Asian lineage H5N1, and there is no evidence of human infection with such H5N1 viruses. Analysis of the HA shows it lacks human receptor binding motifs. In fact, the introduction of the glycosylation during the use of the 50-92-HP RG system in this study has led the virus further away from adaptation to humans, as shown by Herfst and Imai et al. (2012). The fact we are able to deduce this information and make informed predictions from the virus gene sequences is reliant upon the ‘gain of knowledge’ from previous studies in the field.
Recently, species-specific RNAs have been demonstrated as a tool to strengthen biosecurity: Langlois et al. (2013) incorporated target sites of mircoRNA miR-192 into the HA gene of influenza virus (476). This miRNA was found to be expressed at significantly higher levels in human and mice epithelial cells, compared to ferret epithelial cells. The engineered virus was proven to replicate efficiently and transmit in the ferret in vivo model but was attenuated in human cells and in mice due to the miR-192 target. This concept could well be applied to other influenza viruses in a bid to further prevent their accidental escape into the environment. So far chicken microRNA targets have not been well described but in the future RG viruses that contain them might be generated in order to increase safety of experiments with HPAI and alleviate concerns about risks to poultry industry from this type of experimentation.

An influenza pandemic caused by an influenza virus containing a MBCS such as H5N1, which is associated with such high mortality, is a real cause for concern. Current antiviral treatments may help in some cases, but delivery and virus resistance can prevent their effectiveness. Our greatest weapon is vaccination. Successful immunisation against H5N1 has been demonstrated in the poultry industry. The problem with vaccination is the generation time and ability to produce sufficient quantities and distribute to the international community. Several seed strains currently exist and these can be screened for their ability to grow, be developed and effectiveness against other H5 strains. Knowledge of the mutations which enable H5N1 transmission may assist the development of these vaccines. For example, the loss of a glycosylation site may significantly alter the antigenic profile, and this may now be tested and expected should a H5N1 pandemic strike. Moreover increased ability to survive in the slightly acidic pH of the human nose may increase LAIV vaccine take rate (89).

It is crucial that the influenza science community promote informed discussion about influenza and experimentation with the media, government officials and other infectious disease scientists. Society is still damaged today after the scandalous 1998 MMR vaccine controversy. Such a repeat in history should never be allowed.
Chapter 6. Materials and Methods

6.1 Materials

Cell lines

Table 15. Cell lines used in this study

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>293T</td>
<td>Human kidney cells expressing large T antigen of SV-40 ATCC</td>
</tr>
<tr>
<td>DF-1</td>
<td>Chicken fibroblast cell line ATCC</td>
</tr>
<tr>
<td>MDCK</td>
<td>Madin Darby canine kidney cell Glaxo Smithkline or AHVLA</td>
</tr>
<tr>
<td>HeLa</td>
<td>Immortalised cervical cancer cell line from Henrietta Lacks AHVLA</td>
</tr>
<tr>
<td>A549</td>
<td>Human lung cell line ATCC</td>
</tr>
<tr>
<td>A549 IFN-luc</td>
<td>Human lung cell line stably expressing Firefly luciferase reporter gene under the control of IFN-β promoter Hayman et al., 2006 (255)</td>
</tr>
<tr>
<td>CEF</td>
<td>Primary Chicken Embryonic Fibroblasts By Jason Long or AHVLA (Eggs AHVLA)</td>
</tr>
<tr>
<td>DEF</td>
<td>Primary Duck Embryonic Fibroblasts By Jason Long (Eggs AHVLA)</td>
</tr>
</tbody>
</table>

Animals

Table 16. A list of animal species used in this study, together with age and experimental use

<table>
<thead>
<tr>
<th>Species</th>
<th>Age</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>White leg-horn chicken (Gallus gallus)</td>
<td>Embryonic, 9-10 days</td>
<td>Virus amplification. Primary cell generation.</td>
</tr>
<tr>
<td></td>
<td>6-weeks</td>
<td>IVPI.</td>
</tr>
<tr>
<td></td>
<td>10-weeks</td>
<td>Virus infection.</td>
</tr>
<tr>
<td>Pekin Duck (Anas platyrhynchos domestica)</td>
<td>Embryonic, 14-15 days</td>
<td>Virus amplification. Primary cell generation.</td>
</tr>
<tr>
<td></td>
<td>5-weeks</td>
<td>Virus infection and transmission.</td>
</tr>
<tr>
<td></td>
<td>12-weeks</td>
<td>Virus infection.</td>
</tr>
</tbody>
</table>
## Oligonucleotides

Table 17. List of oligonucleotide primers designed in this study

Primers were generated by MWG Eurofins and diluted in sterile dH₂O to a 100pmol/µl stock concentration.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence, 5' to 3'</th>
<th>Primer function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ty/05PB1NotfW</td>
<td>GCGGCCGCGCCACCATGAGTTGATCTACCCGAC</td>
<td>pCAGG cloning (introduces NotI/MluI sites into coding sequence)</td>
</tr>
<tr>
<td>Ty/05PB1MluRev</td>
<td>ACCGGCTCTATTTGGGCCCTGAGGCTCTTCAATGGG</td>
<td></td>
</tr>
<tr>
<td>Ty/05PB2NotfW</td>
<td>GCGGCCGCGCCACCATGAGTTGATCTACCCGAC</td>
<td></td>
</tr>
<tr>
<td>Ty/05PB2MluRev</td>
<td>ACCGGCTCTATTTGGGCCCTGAGGCTCTTCAATGGG</td>
<td></td>
</tr>
<tr>
<td>Ty/05PANotfW</td>
<td>GCGGCCGCGCCACCATGAGTTGATCTACCCGAC</td>
<td></td>
</tr>
<tr>
<td>Ty/05PA MluRev</td>
<td>ACCGGCTCTATTTGGGCCCTGAGGCTCTTCAATGGG</td>
<td></td>
</tr>
<tr>
<td>50-92 NS1 NotfW</td>
<td>CTCTTTGGGAAAAGCGCGGCGCCACCATGAGTTG</td>
<td></td>
</tr>
<tr>
<td>50-92 NS1 MluRev</td>
<td>ACCGGTAAATTTGTCATCTCCTGAGGCTCTTCAATGGG</td>
<td></td>
</tr>
<tr>
<td>pCAGGS(4715)fw</td>
<td>GCTAACCATGTTCATGCCTTC</td>
<td>pCAGG sequencing primer</td>
</tr>
<tr>
<td>pCAGGS(85)rev</td>
<td>GTATTTTTGGGCCACCATGAGTTGATCTACCCGAC</td>
<td></td>
</tr>
<tr>
<td>50-92 HA A160Tfw</td>
<td>GGTTATGGCTTATGGAAGAGAGGACTATTTGG</td>
<td>A160T mutagenesis</td>
</tr>
<tr>
<td>50-92 HA A160Trv</td>
<td>GGTTATGGCTTATGGAAGAGGACTATTTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 HA E193Kfw</td>
<td>CTTTTTTGGAAAAGCGCGGCGCCACCATGAGTTG</td>
<td>E193K mutagenesis</td>
</tr>
<tr>
<td>50-92 HA E193Krev</td>
<td>CTTTTTTGGAAAAGCGCGGCGCCACCATGAGTTG</td>
<td></td>
</tr>
<tr>
<td>50-92 HA G348Rfw</td>
<td>AGGAAAAGAAAACAAAAGAGAGCAGTATTTGG</td>
<td>G348R mutagenesis</td>
</tr>
<tr>
<td>50-92 HA G348Rrev</td>
<td>AGGAAAAGAAAACAAAAGAGAGCAGTATTTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 HA sb Mutfw</td>
<td>ACCCGAGAAGCTCCCTCAAAAGAGGACTATTTGG</td>
<td>removal of Multi-basic cleavage site</td>
</tr>
<tr>
<td>50-92 HA sb Mutrev</td>
<td>ACCCGAGAAGCTCCCTCAAAAGAGGACTATTTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 HA Mutfw</td>
<td>CTTAACAGATTTCAATGGGGAAGCGGCTGTTCAATGGG</td>
<td>removal of Multi-basic cleavage site (HA 4)</td>
</tr>
<tr>
<td>50-92 HA Mutfrev</td>
<td>CTTAACAGATTTCAATGGGGAAGCGGCTGTTCAATGGG</td>
<td></td>
</tr>
<tr>
<td>50-92 acc NS1fw</td>
<td>CTTAACCTCCATGCGGGGGGCGCCACCATGAGTTG</td>
<td>removal of NEP splice acceptor site</td>
</tr>
<tr>
<td>50-92 acc NS1rev</td>
<td>CTTAACCTCCATGCGGGGGGCGCCACCATGAGTTG</td>
<td></td>
</tr>
<tr>
<td>50-92 NS1 H85Pfw</td>
<td>GCCATTGCTCTCCTGAGGTGTTAGGTGGG</td>
<td>H85P mutagenesis</td>
</tr>
<tr>
<td>50-92 NS1 H85Prev</td>
<td>GCCATTGCTCTCCTGAGGTGTTAGGTGGG</td>
<td></td>
</tr>
<tr>
<td>Ty/05 PB2 K627Efw</td>
<td>CATTGCAGCAGCCCCACCGGAAAGAGGACTATTTGG</td>
<td>K627E mutagenesis</td>
</tr>
<tr>
<td>Ty/05 PB2 K627Erev</td>
<td>CATTGCAGCAGCCCCACCGGAAAGAGGACTATTTGG</td>
<td></td>
</tr>
<tr>
<td>Ty/05 PB2 701Nfw</td>
<td>CTTGCTATTTAGGAGGGCATTTGAGGCGGCTGTTCAATGGG</td>
<td>D701N mutagenesis</td>
</tr>
<tr>
<td>Ty/05 PB2 701Nrev</td>
<td>CTTGCTATTTAGGAGGGCATTTGAGGCGGCTGTTCAATGGG</td>
<td></td>
</tr>
<tr>
<td>PR8 HA 300bp rev</td>
<td>GACCATGATCCTACCTGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 HA 300bp rev</td>
<td>GACCATGATCCTACCTGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 PB1 300bp rev</td>
<td>GACCATGATCCTACCTGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 PB2 627 fw 300bp</td>
<td>GAGGATGGCGGCCATGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 PB2 627 rev 300bp</td>
<td>GAGGATGGCGGCCATGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 pa 300bp rev</td>
<td>GACCATGATCCTACCTGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 NP 300bp rev</td>
<td>GACCATGATCCTACCTGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 M 300bp rev</td>
<td>GACCATGATCCTACCTGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>50-92 NS 300bp rev</td>
<td>GACCATGATCCTACCTGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>ty HA 300bp rev</td>
<td>CTTGATTTTGATGGACCATGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>ty PB1 300bp rev</td>
<td>CTTGATTTTGATGGACCATGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>ty PB2 627 fw 300bp</td>
<td>GGTGATTTTGATGGACCATGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>ty PB2 627 rev 300bp</td>
<td>GGTGATTTTGATGGACCATGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>ty PA 300bp rev</td>
<td>CTTGATTTTGATGGACCATGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>ty NP 300bp rev</td>
<td>CTTGATTTTGATGGACCATGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>ty M 300bp rev</td>
<td>CTTGATTTTGATGGACCATGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>ty NS 300bp rev</td>
<td>CTTGATTTTGATGGACCATGAGGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>Matrix primer1</td>
<td>AGATGAGTCTTACACTCCAGGAGAGATAGAGTGG</td>
<td></td>
</tr>
<tr>
<td>Matrix primer2</td>
<td>TGAAACACCTTACACTCCAGGAGAGATAGAGTGG</td>
<td></td>
</tr>
</tbody>
</table>

Virus sequencing (amplifies ~300bp product combined with Hoffmann universal primers)
Table 18. Reverse genetics systems of PR8 (A/Puerto Rico/8/1934), Ty/05 (A/Turkey/Turkey/1/2005) and 50-92 (A/Turkey/England/50-92/91). Helper plasmids expressing polymerase and NP of A/Victoria/3/75. Ty/05, 50-92 and E195 polymerase and NP expressing plasmids and 50-92 HA expressing plasmids

<table>
<thead>
<tr>
<th>Name</th>
<th>Plasmid description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pPol I PR8 PB2</td>
<td>pPol I vector containing PR8 Segment 1 cDNA FluPAN backbone</td>
<td>Whitley et al. (2007) (273)</td>
</tr>
<tr>
<td>pPol I PR8 PB1</td>
<td>pPol I vector containing PR8 Segment 2 cDNA FluPAN backbone</td>
<td></td>
</tr>
<tr>
<td>pPol I PR8 PA</td>
<td>pPol I vector containing PR8 Segment 3 cDNA FluPAN backbone</td>
<td></td>
</tr>
<tr>
<td>pPol I PR8 HA</td>
<td>pPol I vector containing PR8 Segment 4 cDNA FluPAN backbone</td>
<td></td>
</tr>
<tr>
<td>pPol I PR8 NP</td>
<td>pPol I vector containing PR8 Segment 5 cDNA FluPAN backbone</td>
<td></td>
</tr>
<tr>
<td>pPol I PR8 NA</td>
<td>pPol I vector containing PR8 Segment 6 cDNA FluPAN backbone</td>
<td>Dr Ron Fouchier, Erasmus (370)</td>
</tr>
<tr>
<td>pPol I PR8 M</td>
<td>pPol I vector containing PR8 Segment 7 cDNA FluPAN backbone</td>
<td></td>
</tr>
<tr>
<td>pPol I PR8 NS</td>
<td>pPol I vector containing PR8 Segment 8 cDNA FluPAN backbone</td>
<td></td>
</tr>
<tr>
<td>pPol I Ty/05 PB2</td>
<td>pPol I vector containing A/turkey/Turkey/05 segment 1 cDNA</td>
<td>Howard et al. (2007) (272)</td>
</tr>
<tr>
<td>pPol I Ty/05 PB1</td>
<td>pPol I vector containing A/turkey/Turkey/05 segment 2 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I Ty/05 PA</td>
<td>pPol I vector containing A/turkey/Turkey/05 segment 3 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I Ty/05 HA</td>
<td>pPol I vector containing A/turkey/Turkey/05 segment 4 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I Ty/05 NP</td>
<td>pPol I vector containing A/turkey/Turkey/05 segment 5 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I Ty/05 NA</td>
<td>pPol I vector containing A/turkey/Turkey/05 segment 6 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I Ty/05 M</td>
<td>pPol I vector containing A/turkey/Turkey/05 segment 7 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I Ty/05 NS</td>
<td>pPol I vector containing A/turkey/Turkey/05 segment 8 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I 50-92 PB2</td>
<td>pPol I vector containing A/Turkey/England/50-92/91 segment 1 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I 50-92 PB1</td>
<td>pPol I vector containing A/Turkey/England/50-92/91 segment 2 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I 50-92 PA</td>
<td>pPol I vector containing A/Turkey/England/50-92/91 segment 3 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I 50-92 HA</td>
<td>pPol I vector containing A/Turkey/England/50-92/91 segment 4 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I 50-92 NP</td>
<td>pPol I vector containing A/Turkey/England/50-92/91 segment 5 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I 50-92 NA</td>
<td>pPol I vector containing A/Turkey/England/50-92/91 segment 6 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I 50-92 M</td>
<td>pPol I vector containing A/Turkey/England/50-92/91 segment 7 cDNA</td>
<td></td>
</tr>
<tr>
<td>pPol I 50-92 NS</td>
<td>pPol I vector containing A/Turkey/England/50-92/91 segment 8 cDNA</td>
<td></td>
</tr>
<tr>
<td>pCMV-Victoria-NP</td>
<td>Express human influenza Victoria NP protein</td>
<td>Barclay Lab (7)</td>
</tr>
<tr>
<td>pCMV-Victoria-PB1</td>
<td>Express human influenza Victoria PB1 protein</td>
<td></td>
</tr>
<tr>
<td>pCMV-Victoria-PB2</td>
<td>Express human influenza Victoria PB2 protein</td>
<td></td>
</tr>
<tr>
<td>pCMV-Victoria-PB2 627E</td>
<td>Express human influenza Victoria PB2 protein mutated K627E</td>
<td></td>
</tr>
<tr>
<td>pCMV-Victoria-PA</td>
<td>Express human influenza Victoria PA protein</td>
<td></td>
</tr>
<tr>
<td>pCAGGS 50-92 PB1</td>
<td>pCAGGS vector expressing 50-92 PB1 protein</td>
<td>Manuela Mura, Imperial College London (IC) (370)</td>
</tr>
<tr>
<td>pCAGGS 50-92 PB2</td>
<td>pCAGGS vector expressing 50-92 PB2 protein</td>
<td></td>
</tr>
<tr>
<td>50-92 PB2 627K</td>
<td>pPoll &amp; pCAGGS vector expressing 50-92 PB2 627K</td>
<td></td>
</tr>
<tr>
<td>50-92 PB2 701N</td>
<td>pPoll &amp; pCAGGS vector expressing 50-92 PB2 701N</td>
<td></td>
</tr>
<tr>
<td>pCAGGS 50-92 PA</td>
<td>pCAGGS vector expressing 50-92 PA protein</td>
<td></td>
</tr>
<tr>
<td>pCAGGS 50-92 NP</td>
<td>pCAGGS vector expressing 50-92 NP protein</td>
<td></td>
</tr>
<tr>
<td>pCAGGS 50-92 NS</td>
<td>pCAGGS vector expressing 50-92 NS protein</td>
<td>Matt Smith (IC) [477]</td>
</tr>
<tr>
<td>50-92 NS PB5H</td>
<td>pPoll &amp; pCAGGS vector expressing 50-92 NS PB5H</td>
<td></td>
</tr>
<tr>
<td>pCAGGS 50-92-LP</td>
<td>pCAGGS vector expressing 50-92 HA protein</td>
<td></td>
</tr>
<tr>
<td>pPoll &amp; pCAGGS 50-92-160T</td>
<td>pPoll &amp; pCAGGS vector expressing 50-92 HA 160T</td>
<td>This study (478)</td>
</tr>
</tbody>
</table>
### Reporter constructs

**Table 19. Reporter constructs used in this study**

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFN-β luc</td>
<td>Fire fly luciferase reporter gene under the control of IFN-β promoter</td>
<td>Prof Steve Goodbourn, St. Georges, London</td>
</tr>
<tr>
<td>β-gal</td>
<td>β-galactosidase reporter with a rat actin promoter and a 3’ SV40 T Ag splice/polyadenylation site</td>
<td>Prof Steve Goodbourn, St. Georges, London</td>
</tr>
<tr>
<td>pHuman-Poll-Firefly</td>
<td>Human minigenome luciferase reporter, with mouse poll terminator</td>
<td>Dr. Olivier Moncorgé (IC) (317)</td>
</tr>
<tr>
<td>pHuman-Poll-GFP</td>
<td>Human minigenome GFP reporter</td>
<td>Dr. Olivier Moncorgé (IC)</td>
</tr>
<tr>
<td>pChicken-Poll-Firefly</td>
<td>Chicken minigenome luciferase reporter</td>
<td>Dr. Laurence Tiley, Cambridge University</td>
</tr>
</tbody>
</table>

### Antibodies

**Table 20. Antibodies used in this study**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-vinculin</td>
<td>Goat polyclonal antibody raised against vinculin</td>
<td>Santa Cruz</td>
</tr>
<tr>
<td>Goat α-mouse HRP</td>
<td>Horse radish peroxidase conjugated antibody</td>
<td>AbD Serotec</td>
</tr>
<tr>
<td>donkey α-sheep/goat HRP</td>
<td>Horse radish peroxidase conjugated antibody</td>
<td>AbD Serotec</td>
</tr>
<tr>
<td>α-SV5</td>
<td>Mouse monoclonal raised against SV5 tag</td>
<td>AbD Serotec</td>
</tr>
<tr>
<td>α-NP</td>
<td>Mouse monoclonal antibody raised against Influenza A H7 NP clone 2F6.C9</td>
<td>Dr. Maria Zambon, HPA</td>
</tr>
<tr>
<td>Goat α-mouse IgG- β-galactosidase</td>
<td>Secondary antibody conjugated to β-galactosidase</td>
<td>AbD Serotec</td>
</tr>
<tr>
<td>α-HS HA</td>
<td>sheep α-HA antibody (Vietnam/04)</td>
<td>NIBSC</td>
</tr>
<tr>
<td>Goat α-sheep FITC</td>
<td>Goat α-sheep antibody conjugated to FITC</td>
<td>Merck Millipore</td>
</tr>
</tbody>
</table>
Chapter 6 | Materials and Methods

α-H5 50-92 | Polyclonal chicken sera from chickens infected with 50-92 virus | AHVLA, Weybridge
Goat α-chicken 488 | Polyclonal goat α-chicken antibody conjugated to DyLight® 488 | Abcam

Buffers and culture media

Table 21. List of buffers and culture media used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Recipe</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cell culture media</strong></td>
<td>Dulbecco’s Modified Eagle Medium (DMEM) with pyruvate and L-Glutamine (Gibco)</td>
<td>Maintenance of cell lines</td>
</tr>
<tr>
<td></td>
<td>Heat inactivated Fetal Calf Serum (FCS) 10% (v/v) (Biosera)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-essential amino acids (NEAA) 1% (v/v) (Sigma)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>With/without Penicillin and Streptomycin (P/S) 1% (v/v) (Gibco)</td>
<td></td>
</tr>
<tr>
<td><strong>Virus infection media</strong></td>
<td>DMEM (Gibco)</td>
<td>Virus growth media</td>
</tr>
<tr>
<td></td>
<td>With/without FCS 2% (v/v) (Biosera)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>With/without 1mg/ml trypsin (Worthington)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NEAA 1% (v/v) (Sigma)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P/S 1% (v/v) (Gibco)</td>
<td></td>
</tr>
<tr>
<td><strong>Cell trypsin</strong></td>
<td>2x solution 200ml PBS (Gibco)</td>
<td>Maintenance of cell lines</td>
</tr>
<tr>
<td></td>
<td>4ml 2.5% (10x) Trypsin solution (Gibco)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5mM EDTA</td>
<td></td>
</tr>
<tr>
<td><strong>Virus plaque assay overlay</strong></td>
<td>100 ml 10 x Earle’s minimal essential medium (EMEM)</td>
<td>Virus plaque assays</td>
</tr>
<tr>
<td></td>
<td>28 ml 7.5% BSA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1% glutamine (200mM)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20 ml 7.5% NaHCO3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 ml 1M HEPES</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 ml 1% DEAE Dextran (Sigma)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1% penicillin-streptomycin (5000IU/ml)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2% Agarose (Oxoid)</td>
<td></td>
</tr>
<tr>
<td><strong>10x CN buffer</strong></td>
<td>30mM Ferricyanide</td>
<td>Blue cell assays</td>
</tr>
<tr>
<td></td>
<td>30mM Ferrocyanide</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10mM MgCl2</td>
<td></td>
</tr>
<tr>
<td><strong>X-Gal substrate</strong></td>
<td>Prepared to 2% in Dimethylformamide (DMF)</td>
<td>Blue cell assays</td>
</tr>
<tr>
<td><strong>LacZ Buffer</strong></td>
<td>60mM Na2HPO4.7H2O</td>
<td>Developing β-galactosidase assays</td>
</tr>
<tr>
<td></td>
<td>40mM NaH2PO4.H2O</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10mM KCl</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1mM MgSO4</td>
<td></td>
</tr>
<tr>
<td><strong>LB Agar</strong></td>
<td>LB + 1.5% (w/v) Difco Agar</td>
<td>Culturing transformed bacterial cells</td>
</tr>
<tr>
<td></td>
<td>Supplemented with 1% Ampicillin</td>
<td></td>
</tr>
<tr>
<td><strong>Lysogeny Broth (LB)</strong></td>
<td>1% Oxoid tryptone</td>
<td>Culturing transformed bacterial cells</td>
</tr>
<tr>
<td></td>
<td>0.5% Oxoid yeast extract</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5% NaCl</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1% glucose</td>
<td></td>
</tr>
<tr>
<td><strong>TAE Buffer</strong></td>
<td>40mM Tris-acetate pH 8</td>
<td>DNA gel electrophoresis</td>
</tr>
<tr>
<td></td>
<td>1mM EDTA</td>
<td></td>
</tr>
<tr>
<td><strong>Phosphate Buffered Saline (PBS)</strong></td>
<td>7.4pH (Gibco)</td>
<td>Used for cell washes</td>
</tr>
<tr>
<td><strong>Crystal violet solution</strong></td>
<td>300 ml Crystal violet stock solution</td>
<td>Cell monolayer staining</td>
</tr>
<tr>
<td></td>
<td>300 ml ethanol</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6 l water</td>
<td></td>
</tr>
</tbody>
</table>
6.2 Methods

6.2.1 Molecular Biology

PCR amplification

PCR was performed with K.O.D. Taq polymerase (Novagen) or GoTaq® green master mix (Promega). For K.O.D. Taq polymerase PCR a 50μl PCR mix was made consisting of 5 μl 10X Buffer and 3μl MgSO4 for KOD DNA Polymerase, 0.15μl (100pmol/μl) forward and reverse primers, 5 μl dNTPs 2 mM, 1μl KOD polymerase (2.5 U/μl), 10-100ng of DNA template and nuclease free water to 50μl. For Go-Taq polymerase PCR of bacterial colony screening, the following was used: 7.5μl 2X GoTaq® Green Master Mix, 0.25μM forward and reverse primers, bacteria swab and nuclease free water to 15μl. Cycling conditions were performed to the manufacturer’s guidelines and varied according to the construct length.

Site directed mutagenesis of the constructs

Site directed mutagenesis was carried out on both the 50-92 and Ty/05 rescue plasmids using the QuikChange® Lightning Site-Directed Mutagenesis Kit to the manufacturer’s guidelines (Agilent Technologies). Where QuikChange mutagenesis failed, overlapping PCR was performed using K.O.D polymerase. The first round of PCR was performed (described in section 8.2.1.1), generating DNA fragments with the engineered mutations and an overlapping region. For the second round of PCR, equal amounts of the purified DNA fragments were used as templates for the DNA reaction. 10 cycles of PCR were first performed without primers, to allow annealing of the DNA templates before completing the PCR reaction.

Agarose gel electrophoresis

DNA fragments were separated on 1% or 2% agarose gels diluted with 0.5XTAE buffer and supplemented with 1x gel red (Cambridge Bioscience). DNA was loaded with 5x Gel loading dye (Qiagen). Gels were run in 0.5XTAE buffer. Samples were run alongside either 1kb or 1kb plus DNA size markers (Invitrogen) at 80-120V until the bands had sufficiently separated. DNA was visualised using a UV trans-illuminator.

DNA purification

DNA fragments were cut from the agarose gel under UV light, and DNA extracted using a QIAquick Gel extraction kit (Qiagen) following manufacturer’s instructions. Similarly,
digestion products or PCR product (directly) were purified using the Gel extraction kit. DNA fragments were eluted with 35-50 μl sterile water.

**Double restriction enzyme digest**

In order to clone genes into the pCAGGS vector, the NotI and MluI restriction enzyme sites engineered into the vector were utilised. For all double digests, using NEB reagents, the following quantities were used. For example a 10μl reaction volume: 1μl NEBuffer 3, 0.16μl of each Enzyme, 0.33μl BSA, ~250ng purified plasmid, 8.33μl nuclease free water. Digests were left for the acquired amount of time.

**DNA ligation**

Constructs were ligated into the pCAGGS vector using Clonables™ (Novagen®) to the manufacturer’s guidelines. Typically, 0.5µl of vector was incubated at RT for 20mins with 4.5µl construct and 5µl Clonables™.

**Transforming competent bacterial cells**

1 to 5ul of the DNA ligation (or ~50ng of plasmid) was mixed with 50µl OneShot®TOP10 chemically competent *E.coli.* (Invitrogen) and incubated on ice for 30 minutes. Cells were heat shocked for 30 seconds at 42°C and replaced on ice for 2 minutes. 250µl of pre-warmed S.O.C. media (Invitrogen) was added to the mix and incubated for 1 hour at 37°C in a shaking incubator, after which a suitable volume was spread on to pre-warmed LB agar plates containing 1% Ampicillin and incubated overnight at 37C for colony formation. Colonies were screened by PCR, and grown in 5ml LB supplemented with Ampicillin (100µg/ml).

**Plasmid purification**

**Small scale.** Bacterial cells grown in 5ml LB (100µg/ml Ampicillin) were pelleted by centrifugation at 3000xg for 5 minutes. The supernatant was discarded and DNA purified by the QIAprep Spin Miniprep kit (QIAGEN) following the manufacturers guidelines. DNA was eluted with 50µl warm sterile water and stored at 20°C.

**Large scale.** 250ml of LB supplemented with 100µg/ml Ampicillin was inoculated with 200µl bacterial cells and incubated at 37°C for 16 hours. Cells were pelleted by centrifugation at 3000xg for 5 minutes. The supernatant was discarded and DNA purified by
the QIAfilter Plasmid Maxi kit (QIAGEN) following the manufacturers guidelines. The DNA pellet was resuspended with 1ml warm sterile water and stored at 20°C.

The concentration and quality of DNA was measure using a NANODROP spectrophotometer. Good quality DNA was considered such with a $A_{260}/A_{280}$ ratio of 1.7-2.0.

**Plasmid constructs**

Poll and pCAGGS plasmids for A/Puerto Rico/8/1934 (PR8), A/Turkey/England/50-92/91 (50-92), A/Turkey/Turkey/1/2005 (Ty/05) (supplied by R. Fouchier, Erasmus University, Rotterdam, Netherlands) and A/England/195/09 (pH1N1) have been described previously (272, 318). 50-92, Ty/05 and pH1N1 PB1, PB2, PA, and NP genes were subcloned into the pCAGGS expression plasmid as described previously (317, 318). Briefly, mutagenesis of PB2 genes was carried out using the QuikChange® Lightning Site-Directed Mutagenesis Kit to the manufacturer’s guidelines (Agilent Technologies). Polymerase assays utilised the mini-genome firefly reporter plasmid pChicken-Poll-Firefly or pHuman-Poll-Firefly as described previously (317).

Mutagenesis of the 50-92 HA gene was carried out sequentially to introduce each of the three HA mutations using the QuikChange® Lightning Site-Directed Mutagenesis Kit. Conventional overlapping PCR was used to introduce the H85P mutation into 50-92 NS1 gene. 50-92 NS1 P85H and the 50-92-LP, 50-922160T 50-922160T,193K, 50-924r and 50-92-HP HA genes were successfully inserted into the pCAGGS expression vector via the introduction of the NotI and MluI restriction enzyme sites on the 5’ and 3’ ends by PCR. The stop codon of the 50-92 NS1 P85H gene was removed to allow expression of the V5 tag on the C-terminus.

**Sequencing of constructs**

**Plasmid constructs.** All plasmid constructs were verified by DNA sequencing. Sequencing of constructs was conducted by the MRC DNA Core Genomics Laboratory, Hammersmith Hospital, using the primer diluted to 3.2pmol/μl and >500ng DNA in a total reaction volume of 10μl.

**Virus gene sequencing.** All virus mutations were verified by sequencing the mutated region of the viral gene of successfully generated RG virus. RT-PCR was carried out on viral RNA using QIAgen OneStep RT-PCR kit (QIAGEN) according to the manufacturer’s instructions. The product was isolated using QIAquick Gel extraction kit (Qiagen). Chain termination PCR was performed using the BigDye Terminator x3.1 Cycle Sequencing Kit.
Chapter 6 | Materials and Methods

(Applied Biosystems). Briefly, 4µl reaction mix was added to 1.6µl DNA primer (1pmol) and 4.4µl of eluted DNA product and PCR performed according to manufacturer’s instructions. Precipitation of the reaction product was performed by adding 5µl 125mM EDTA and 60µl 100% ethanol, left at RT for 15 minutes and centrifuged at 20,000xg for 30 minutes. Supernatant was removed and the remaining pellet was washed in 150µl 70% ethanol and centrifuged for a further 10 minutes. Ethanol was removed and the pellet left to dry before adding 12.5µl Hi-Di Formamide (Applied Biosystems). The products were run on the 3130 or 3730 capillary sequencing machines by the AHVLA Central Sequencing Unit.

All sequencing results were analysed using Geneious R6 software.

6.2.2. Cell lines and transfection

Cell culture

Primary chicken embryo fibroblast (CEF) cells and duck embryo fibroblast (DEF) cells were generated from 9-10 day old chicken embryos or 14-15 day pekin duck embryos using 0.25% trypsin (Invitrogen). In short, using sterile technique, embryos were culled by decapitation, limbs were removed and the bodies were processed through a 10ml syringe. Tissues were washed in PBS supplemented in 1% Penicillin-Streptomycin (PS) (Invitrogen), allowed to settle, supernatant removed and trypsinised at 37°C by aid of a magnetic stirrer. Once trypsinised, cells were filtered through sterile muslin and resuspended in Fetal Calf Serum (FCS) before being centrifuged at 800xg for 5 minutes in a graduated tube. The 0.1ml of the cell pellet was resuspended in 50ml growth media.

CEF, DEF, Human embryonic kidney (293T), chicken fibroblast (DF-1), human alveolar epithelial cells (A549) and Madin-Darby canine kidney (MDCK) cells were maintained in Dulbecco’s modified Eagle’s medium (DMEM; Invitrogen) supplemented with 10% fetal calf serum (FCS) (Biosera) and 1% PS. The cell lines were maintained at 37°C in a 5% CO₂ atmosphere. DF-1 cells were maintained at 39°C. The A549-luc cell line was maintained by G418 drug selection (SigmaAldrich). Cell lines were trypsinised and passaged twice weekly and monitored for signs of fatigue.

Cell transfections

Transfections were performed using Lipofectamine 2000 reagent (Invitrogen) following manufacturer’s instructions. DNA:reagent ratio volumes were scaled appropriately.
for cell transfections. DNA and Lipofectamine were diluted in the appropriate volumes using OptiMEM (Invitrogen). Transfection of RG plasmids for RG virus generation was carried out with Fugene HD (Promega), described section 8.2.3.

**Luciferase transgene assay**

To measure the level of cell protein expression altered by NS1, 0.5μg pCAGG-NS1, 0.125μg pCAGGS-luciferase plasmid DNA was transfected into 293T cells in 12-well format as described previously (477). Cells were incubated at 37°C. Cells were harvested after 48 hours.

**IFN-β luciferase reporter assay**

To measure the ability of NS1 protein to inhibit IFN-β induction, 0.375μg pCAGG-NS1, 0.25μg luciferase reporter (human-IFN-β or chicken-IFN-β) and 0.125μg β-gal plasmid DNA was transfected into 293T and DF-1 cells in 12-well format as previously described (477). After 24 hours, stimulation of IFN induction by influenza virus infection (MOI 3.0), Poly(I:C) or NDV was conducted as previously described (477). Cells were incubated at 37°C and harvested after a further 24 hours.

**A549-luc assay**

To measure IFN-β induction of A549-luc cells, cells in 24-well format were inoculated with influenza virus (MOI 3.0), as previously described (477). Cells were harvested after 24 hours.

**Minigenome assay** NP: 320 ng, PB1 and PB2: 160 ng, PA: 40 ng

To measure polymerase activity, 160ng PB1, 160ng PB2, 40ng PA and 320ng NP pCAGGS expression plasmids, together with 160ng of the mini-genome firefly reporter plasmid pChicken-Poll-Firefly or pHuman-Poll-Firefly were transfected into avian DF-1 or human 293T cells, respectively, as previously described (318). Double the amount of DNA was used for DF-1 cells. Cells were incubated at 37, 39 or 41°C. Cells were harvested after 20 hours.

**Firefly luciferase activity assay**

To measure firefly activity, 20 or 24 hours after transfection, cells were lysed with 300μl (12-well) or 200μl (24-well) of passive lysis buffer (Promega), and firefly luciferase
activity measured using Luciferase Assay System reagent (Promega) on a FLUOstar Omega plate reader (BMG Labtech).

**β-galactosidase assay**

Cell protein expression or transfection efficiency was measured by β-gal activity. Cells were lysed with 300µl (12-well) or 200µl (24-well) of passive lysis buffer (Promega). 10µl of lysate, 80µl lacZ buffer and 20µl ortho-Nitrophenyl-β-galactosidase (ONPG) (4mg/ml) was added to a well of a 96well plate and incubated at room temperature until a yellow colour change was observed. Absorbance was measured at 450nm using a FLUOstar Omega plate reader.

**Western blot of NS1 protein**

* Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Cells lysates were diluted 1:1 in Protein Loading Buffer (20µl 1M 1,4-Dithiothreitol per 1ml buffer) and denatured by heating at 100°C for 2 minutes. 10µl of each sample was loaded and then separated on a 15% SDS-PAGE at 160V, until the pre-stained marker had sufficiently separated (Novex).

* Immuno blotting. Proteins were then transferred onto a methanol activated Polyvinylidene fluoride membrane (PVDF) (VWR), at 15v for 1 hour. Membranes were blocked for 1 hour in 5% fat free milk powder (Fluka) solution, diluted in 0.5% TWEEN-PBS at 4°C. Membranes were washed 4 times in 0.1% TWEEN-PBS and probed by anti-V5 mouse monoclonal antibody (ABD Serotec) and anti-viniculin goat monoclonal antibody (Santa Cruz) both at 1:1000 dilution in milk solution for 1 hour. Four 5 minute washes in 0.1% TWEEN-PBS followed before 2° antibodies were then added at 1:1000 dilution, (including: Horse radish peroxidise (HRP)-conjugated to Goat anti-mouse or donkey anti-goat), for 1 hour. Washing followed as before. HRP-conjugated antibodies were detected by incubation of the membrane for 3 minutes with ECL Plus Western Blotting Detection System (Amersham) as manufacturer’s instructions. The resulting chemiluminescence was captured on ECL-Hyperfilm (GE Lifesciences) and developed in an automated Fuji X-ray film developer RGII.
Influenza HA detection

**IHC.** 293T cells were transfected with 500ng 50-92 HA pCAGGS plasmid on glass coverslips and incubated for 24 hours. Cells were fixed in 4% paraformaldehyde and probed with sheep α-HA antibody (1:300) (Vietnam/04, NIBSC) and HA indirectly detected by goat α-sheep antibody conjugated to FITC (1:500) (Merck Millipore). Pictures of fluorescent cells were acquired with an Axiovert 200M microscope (Zeiss) with a Axiocam HRC camera (Zeiss) and were edited with AxioVision Rel 4.7 program.

**Flow cytometry.** HeLa cells in 12-well plate format were transfected with 500ng of pCAGGS plasmid expressing 50-92 HA, and incubated for 24 hours. After 24 hours incubation, cells were trypsinised and washed in ice-cold PBS then probed with chicken polyclonal anti-50-92-HA sera (1:100) (AHVLA, Weybridge). After HA was indirectly detected by AlexaFlour 488 FITC antibody (1:500). Prior to flow cytometry analysis, cells were fixed with 1% paraformaldehyde for 10 minutes on ice. Cells were analysed using a MACSQuant flow cytometer (Miltenyi biotec) and data analysed using FACS Quant software. Figure 48 is an example of the FACS analysis conducted. Per sample, a double SSC plot-gated area was generated (P1) then selected for analysis; FSC vs SSC plot and FITC vs FSC-gate (P3) which contains positive cells. Controls included 50-92-LP HA with/without secondary or primary antibody and a mock (empty pCAGG vector) transfected cell sample.
Figure 48. An example of the FACS analysis data detecting 50-92 HA expressed on the cell surface of transfected HeLa cells


**Polykaryon fusion assay**

293T cells were transfected with 500ng of 50-92 HA pCAGGS plasmid. After 24 hours, cells were washed in PBS and treated with pre-warmed MES buffer at varying pH for 5 minutes. Growth media was replaced and the cells incubated at 37°C until polykaryon formation occurred (approximately 3 hours). Pictures were acquired using the Axiovert 200M microscope described previously.
Chapter 6 | Materials and Methods

6.2.3 Infectious studies

6.2.3.1 In vitro

Safety/ Biosecurity

All infectious experiments with HPAI were performed in Biosafety Level 3 + (BSL-3 +) [UK Department of Environmental, Food and Rural Affairs (Defra) Specific Animal Pathogen Order (SAPO) 4; Advisory Committee on Dangerous Pathogens (ACDP) 3] facilities at AHVLA, Weybridge. Infectious experiments conducted with PR8 H1N1 virus carrying the internal gene segments of HPAI viruses were conducted at BSL-2 at Imperial College London, St Mary’s Campus.

Generation of recombinant viruses

Viruses were generated by reverse genetics using a 12-plasmid system as described previously (318). Briefly, all 12 plasmids were transfected into 293T cells using in 12-well plates using Fugene HD (Promega). After 24 hours, the transfected cells were removed from the wells, mixed with MDCK cells, and co-cultured in 25-ml flasks using DMEM containing 10% serum. 6 to 8 hours later the cells were washed briefly in serum-free DMEM and the medium was replaced with serum-free DMEM. For generation of viruses with a single-basic HA cleavage site, serum-free DMEM was supplemented with 2.5 µg/ml TPCK trypsin (Worthington). Reverse transcription-PCR (RT-PCR) was carried out on the cell supernatant of the recovered viruses, the PB2 627K, NS1 P85H or HA 160, 193, HA24R region amplified and cDNAs were sequenced to confirm their genotype. Virus stocks were propagated in 9 day old embryonated chicken eggs incubated at 37°C. Infectious titres were determined by plaque assay on MDCK cells or by Egg Infectious Dose50 (EID50).

Influenza virus titration by plaque assay

12 well plates seeded to form a confluent monolayer to infect the next day. The media was removed, and cells washed in PBS. Virus was diluted in SF DMEM in a 10-fold dilution series. 200µl of virus was added to each well and incubated at 37°C. After 1 hour, 1ml of flu overlay/avicel (Sigma Aldrich) mix containing 0.4µl/ml of TPCK trypsin, was added to each well. After 3 days incubation at 37°C, the overlay was removed and cells fixed/stained using Crystal Violet stain for at least 30 minutes. Monolayers were washed before counting the plaques.
Haemagglutination (HA) assay

Virus was diluted in a two-fold serial dilution using PBS in 96 well V-bottomed plates. An equal volume of 0.5% chicken or turkey red blood cells (diluted in PBS) was then added to each well. After at least 1 hour incubation on ice, the resulting HA titre was determined, as the well prior to the first well displaying a blood pellet.

Viral RNA isolation

Viral RNA was isolated from oropharyngeal, cloacal, feather samples, egg allantoic fluid and cell supernatant using the automated QIAquick RNA extraction method (Qiagen) as previously described (396). Sample sizes of approximately 12 or less were isolated using the manual QIAamp Viral RNA mini extraction method (Qiagen) following the manufacturer’s protocol. Samples were stored at -70°C.

Quantification of Viral RNA

50-92 and Ty/05 viral RNA was quantified by influenza Matrix and H5 gene qRT-PCR respectively, as previously described (396). Relative EID$_{50}$ units (REU) were calculated by quantitative standards of five 10-fold dilutions of extracted RNA from $10^6$ EID$_{50}$ Ty/05 or 50-92. Not detected (negative) samples were deemed so if they held a ct value of more than 35-36, judged depending on the performance of the negative and positive controls included in the qRT-PCR run. For each qRT-PCR the negative control included the water used to dilute the PCR reaction, and the positive control included vRNA at a known concentration that was extracted at the same time as the RNA isolation of the samples to be tested.

In vitro growth kinetics

CEF and DEF cells in 6-well plate format were inoculated at a multiplicity of infection of $10^3$ (50-92) or $10^{-4}$ (Ty/05), unless otherwise stated, washed once in PBS and incubated in 2ml DMEM supplemented with 2% FCS and 1% P/S at 37 or 41°C. At the times indicated, cell supernatant was collected, vRNA isolated and viral titre determined by qRT-PCR.

Haemolysis assay

64 HA units of virus was mixed with chicken red blood cells. After incubation for 30 minutes at 4°C, the mixtures were pelleted and washed in PBS. Centrifugation was repeated and the pellet was resuspended in MES buffer at varying pH. After 30 minutes incubation at 37°C, cells were centrifuged briefly and fusion was measured by absorbance of the
supernatant at 450nm. Values were normalised to mock and presented as mean % Haemolysis.

**Haemagglutination inhibition assay**

To detect antibody to influenza HA, the haemagglutination assay was used on animal blood serum prior to and post-infection with influenza virus. Freeze dried inactivated antigen of Ty/05 or 50-92 HA (AHVLA, Weybridge) was reconstituted in PBS and titrated to 4HAU by HA assay. 25µl serum sample, control serum (AHVLA, Weybridge) and negative control (PBS) were added to 25µl PBS and titrated 2-fold in a 96-well V-bottomed plate. 25ul of the 4HAU antigen stock was added to all wells and incubated for 30 minutes at RT. 25µl of 1% chicken red blood cells were added and mixed. Results were determined after 30-40 minutes incubation at RT.

**6.2.3.2 In vivo and in ovo infection studies**

**Ethics statement**

The animal experiments were evaluated and approved by the AHVLA ethics committee for animal studies and was carried out in accordance with the UK 1986 Animal Scientific Procedure Act and AHVLA code of practise for performance of scientific studies using animals (available upon request). In all in vivo experiments, animals were routinely monitored and animals deemed moribund following H5N1 infection, were killed humanely.

**Chicken experiments**

Oropharyngeal, cloacal and blood samples were taken from all chickens prior to infection to test for previous influenza infection. Samples revealed no previous exposure to H5 infection by M gene qRT-PCR or by haemagglutination inhibition tests using A/turkey/England/50-92/91 antigen (AHVLA, Weybridge).

**IVPI.** A standard intravenous pathogenicity index test on Ty/05/K, Ty/05/E, 50-92/gKR, 50-92/aEQ (Also 50-92-HP), 50-92-LP and 50-92_{360T,193K} was performed, according to OIE guidelines (148). Briefly, ten 6-week old chickens were intravenously infected with a minimum of 16HAU units of virus diluted 1 in 10 in PBS. Clinical signs were recorded daily for 10 days.
Chapter 6 | Materials and Methods

Chapter 2, 50-92 HA comparison. Five randomly selected chickens inoculated for IVPI with either 50-92-LP, 50-92-HP or 50-92\textsubscript{160T,193K} were swabbed by cloacal and oropharyngeal means after 24 hours, and 11 dpi. Chickens infected with 50-92-HP were swabbed at death at 2 dpi.

Chapters 3 and 4, PB2 627 comparison. Seven 10-week old White Leghorn specific-pathogen-free chickens (Cherry Valley Farms Ltd, UK) were infected oculonasally with 10\textsuperscript{6} EID\textsubscript{50} units of either 50-92/aEQ or 50-92/gKR. Virus shedding was measured by oropharyngeal, cloacal, and feather samples from the brow. Clinical signs were recorded for each bird at multiple times daily (Table 22).

Table 22. Poultry record sheet for clinical observations

Where possible, answers were given semi quantitatively out of 5, where 1/5 is mild and 5/5 is bad/advanced.

<table>
<thead>
<tr>
<th>Clinical sign</th>
<th>7am Bird No.</th>
<th>1pm Bird No.</th>
<th>7pm Bird No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alertness: group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alertness: individual</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evidence of eating?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nervous signs?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torticollis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking in circles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wind drop</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limping</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss of balance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tremors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paralysis/moribund</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ruffled feathers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huddling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diarrhoea</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyes shut</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infra-orbital oedema</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oedema of the head</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swollen eyes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conjunctivitis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pus-like discharge from eyes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watery discharge from eyes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pus-like discharge from nares</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyanosis of extremities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10p sized sub-cutaneous haemorrhages</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pin head sub-cutaneous haemorrhages</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coughing/snicking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other observations</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Duck experiments

High-health status Pekin ducks were acquired from Cherry Valley Farms Ltd, UK. Oropharyngeal, cloacal and blood samples were taken from all ducks to test for previous influenza infection. Samples revealed no previous exposure to H5 infection by H5 gene quantitative reverse-transcriptase PCR (qRT-PCR) or by haemagglutination inhibition tests using A/turkey/Turkey/1/05 antigen (AHVLA, Weybridge). Twelve 12-week old Pekin ducks were each oculonasally inoculated with $10^6$ EID$_{50}$ units of either Ty/05/K or Ty/05/E. Five 4-week old naive contact Pekin ducks were co-housed 1 day post-infection (d.p.i.). Five 12-week old ducks were selected randomly and culled humanely 2 d.p.i. for post-mortem examination. Oropharyngeal and cloacal swabs were taken daily from all birds Clinical signs were observed daily and recorded as negative or positive (Appendix i, Table 23).

Embryo infection

For purposes of virus amplification. Embryos were candled in the dark and a point of inoculation which allows injection into the allantoic fluid away from the embryo and any major veins was marked. Holes were drilled in the egg shell at this point. Virus was diluted in SF DMEM containing 1% PS to an appropriate concentration and 100-200µl injected into the egg using a 25 gauge 5/8 inch needle and 0.1ml syringe. Eggs were monitored daily by candeling. Eggs infected with HPAI were chilled at death or ideally at the point where the egg veins were decreasing and embryo dropping. LPAI infected eggs were chilled after 2 to 3 dpi. Allantoic was harvested using sterile technique and stored at -70°C.

Influenza virus titration by Egg Infectious Dose$_{50}$ (EID$_{50}$). A ten-fold dilution series was made by diluting virus in SF DMEM supplemented with 1% PS. 100µl of each dilution was injected into the allantoic fluid of five 9-day old chicken eggs. Eggs were monitored and chilled at death or after 3 days incubation. Allantoic fluid was harvested and tested by HA assay in chicken or turkey red blood cells and recorded as negative or positive. EID$_{50}$ was calculated by the Reed and Muench mathematical technique.

Chapter 2. Embryo infection for 50-92 HA comparison. Four 9-day old embryonated chicken eggs were infected with $10^2$ pfu/ml of either 50-92-LP, 50-92-HP or 50-92$_{160T,199K}$ virus. After 24 hours (n=2) or 48 hours (n=2) embryos were culled by decapitation and fixed in formalin for at least 5 days before H&E sections being made. IHC demonstration of virus NP antigen was examined. Images were captured using the Axiovert 200M microscope.
described previously. Images representative of both infected embryos were selected. Sections were prepared by the Pathology team at AHVLA, Weybridge.

**Immunohistochemistry**

**In vivo.** At post mortem tissue samples were harvested (including nasal turbinates, air sac, bursa, spleen and gonad) and fixed in 10% buffered formalin.

**In ovo.** The egg membranes were harvested and chick embryos were decapitated and submerged in 10% buffered formalin for a minimum of 5 days.

**Both.** Four µm thick sections were stained with haematoxylin and eosin or used for immunohistochemical detection of influenza A nucleoprotein as previously described (221). Briefly, sections for immunohistochemistry (IHC) were dewaxed in xylene, and passed through graded alcohols to Tris Buffered Saline solution (TBS) (0.005M Tris, pH 7.6, 0.85% w/v NaCl). Endogenous peroxidase activity was quenched with a methanol/hydrogen peroxide block (BDH) and treated with Protease XXIV. Samples were incubated at room temperature with an anti-influenza A nucleoprotein primary antibody (Statens Serum Institute, Denmark) 1/4000 in TBS for 1 hour, and Dako ENVISION™ polymer for 30 minutes. Sections were washed three times with TBS between incubations. The immunohistochemical signal was visualised using 3,3 diaminobenzidine (Sigma-Aldrich), and sections were counterstained in Mayer's haematoxylin (Surgipath, UK).

Harvesting of tissues was carried out by Jason Long, Alex Nunez (AHVLA) and AHVLA staff. IHC was performed by the Pathology team, AHVLA, Weybridge. IHC scoring of duck tissues was performed blind by Alex Nunez. IHC analysis of chick embryos was performed by Jason Long, and reviewed, blind, by Professor Wendy Barclay.

**6.2.4 Bioinformatic analysis**

**Sequence analysis**

AI influenza sequences were downloaded from the NCBI Influenza Virus Resource (285) and aligned using the using Clustal Omega EBI resource (284). Sequence alignments were analysed using Geneious R6 software.
Phylogenetic analysis

To investigate the prevalence of the PB2 627K mutation in H5N1 viruses, Samantha Lycett (Edinburgh University) downloaded 1325 H5N1 PB2 and 2392 HA sequences from the NCBI Influenza Virus Resource (285), April 2013 from avian, human and other hosts. 59 additional human H5N1 genomes and one avian genome were also obtained through GISAID. The nucleotide sequences were aligned by codon using Muscle in MEGA 5.05 and manually adjusted. HP isolates were identified by multi-basic cleavage site on the HA protein sequences. A neighbour-joining tree was used to identify the virus clades in HA according to the WHO H5 nomenclature [http://www.who.int/influenza/gisrs_laboratory/h5n1-nomenclature/en/b]. The tree was generated in MEGA from the downloaded HP HA nucleotide sequences, together with the sequences used in the WHO tree [http://www.who.int/influenza/gisrs_laboratory/201101_h5smalltreerealignment.txt] using the Tamura-Nei model allowing for site and branch heterogeneity. HP isolates from the Hong Kong 97-like clade, Clade 2.2, and the time periods 1996-2004 and 2005-Present (excluding isolates in HK/97-like and Clade 2.2) were labelled manually using FigTree, and duplicate isolates removed. The available PB2 sequences corresponding to the labelled isolates (final number after matching and de-duplication = 963) were analysed for PB2 E627K mutations by host type and clade using custom R scripts.

Statistical analysis

All statistical analyses were performed using Prism 6 software (GraphPad). Statistical calculations included two-tailed paired t-test, two-way ANOVA with Bonferroni’s post-test, one-way ANOVA and Area Under the Curve (AUC) analysis. P-values less than 0.05 were considered significant. When performing multiple comparisons, the Bonferroni correction ensures that the 5% probability applies to the entire family of comparisons, and not only to each individual comparison.
Chapter 7. Appendices

Appendix i

Table 23. Clinical signs observed in ducks infected with Ty/05/K or Ty/05/E H5N1 virus

<table>
<thead>
<tr>
<th>Clinical sign</th>
<th>Ty/05/K</th>
<th>Ty/05/E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Juvenile duck</td>
<td>Adult duck</td>
</tr>
<tr>
<td>Wing drop</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Ruffled feathers</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Huddling</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Lethargy</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Eyes shut</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Conjunctivitis</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Laboured breathing</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Loss of balance*</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Walking in circles*</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Tremors*</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Paralysis*</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Fitting*</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Torticollis*</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>

Ticks were given if the clinical sign was observed in any of the 10 adult ducks or 5 juvenile ducks for either group during infection. * Indication of neuropathogenesis.

Appendix ii

Table 24. Quantification of viral RNA isolated from the tissues of ducks infected with either Ty/05/K or Ty/05/E by qRT-PCR

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Ty/05/K</th>
<th>Ty/05/E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n°</td>
<td>Titre²</td>
</tr>
<tr>
<td>Nasal turbinate</td>
<td>2/5</td>
<td>4.20</td>
</tr>
<tr>
<td>Lower trachea</td>
<td>2/5</td>
<td>5.93</td>
</tr>
<tr>
<td>Upper trachea</td>
<td>2/5</td>
<td>4.52</td>
</tr>
<tr>
<td>Tracheal swab</td>
<td>2/5</td>
<td>3.73</td>
</tr>
<tr>
<td>Lung</td>
<td>2/5</td>
<td>4.32</td>
</tr>
<tr>
<td>Air sac</td>
<td>3/5</td>
<td>4.98</td>
</tr>
<tr>
<td>Liver</td>
<td>2/5</td>
<td>5.35</td>
</tr>
</tbody>
</table>
Heart 2/5 2.60 2.04 1/5 4.05
Kidney 2/5 3.18 2.52 1/5 5.46
Brain 2/5 6.73 6.88 1/5 5.24
Spleen 2/5 5.80 5.94 2/5 4.11
Pancreas 1/5 2.57 1/5 2.41
Thymus 2/5 4.26 3.99 1/5 5.61
Bursa 2/5 4.36 4.47 1/5 4.19
Duodenum 2/5 4.11 4.23 1/5 3.82
Caecum 2/5 4.16 4.28 1/5 4.98
Breast muscle 1/5 2.29 1/5 2.78
Thigh muscle 1/5 2.29 1/5 2.47

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Ty/05/K</th>
<th>Ty/05/E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>Skin</td>
<td>- - - - -</td>
<td>- - - + -</td>
</tr>
<tr>
<td>Feather follicle</td>
<td>- - - - -</td>
<td>- - - + -</td>
</tr>
<tr>
<td>Flight Feathers</td>
<td>- - - - -</td>
<td>- - - + -</td>
</tr>
<tr>
<td>Heart</td>
<td>- - - - -</td>
<td>- - - + -</td>
</tr>
<tr>
<td>Breast muscle</td>
<td>- - - - -</td>
<td>- - - + -</td>
</tr>
<tr>
<td>Thigh muscle</td>
<td>- - - - -</td>
<td>- - - + -</td>
</tr>
<tr>
<td>Air sac</td>
<td>- - - + -</td>
<td>- - - ++ -</td>
</tr>
<tr>
<td>Epithelial cells</td>
<td>- - - + -</td>
<td>- - - ++ -</td>
</tr>
<tr>
<td>Lymphoid cells</td>
<td>- - - +/-</td>
<td>- - - + -</td>
</tr>
<tr>
<td>Blood vessels</td>
<td>- - - - -</td>
<td>- - - +/-</td>
</tr>
<tr>
<td>Lung</td>
<td>- - - +/-</td>
<td>- - - +/-</td>
</tr>
<tr>
<td>Air capillaries</td>
<td>- - - +/-</td>
<td>- - - +/-</td>
</tr>
<tr>
<td>Blood vessels</td>
<td>- - - - -</td>
<td>- - - - -</td>
</tr>
<tr>
<td>Bronchi</td>
<td>- - - - -</td>
<td>- - - - -</td>
</tr>
<tr>
<td>Trachea</td>
<td>- - - + -</td>
<td>- - - * -</td>
</tr>
<tr>
<td>Epithelium</td>
<td>- - - + -</td>
<td>- - - - -</td>
</tr>
</tbody>
</table>

\footnote{a}{number of animals positive by qRT-PCR out of 5.}
\footnote{b}{Virus RNA was extracted from oropharyngeal, cloacal or feather samples and virus titre measured by qRT-PCR of the HA gene, displayed as mean log_{10} REU.}
\footnote{c}{Standard deviation of the mean, displayed as mean log_{10} REU.}

Appendix iii

Table 25. Immunohistochemical distribution of Influenza A nucleoprotein by IHC from 12-week old ducks 2 days post infection with either Ty/05/K or Ty/05/E
IHC carried out by Pathology department, AHVLA, Weybridge. IHC scoring carried out by Alejandro Núñez (AHVLA) and Jason Long. IHC score was an overall indication of the extent of virus antigen in the whole tissue. Positive and negative controls showed appropriate and expected labelling (data not shown). Data for birds 1-5 in each group.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>IHC Score</th>
<th>Examples:</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>+/-</td>
<td>Weak</td>
<td>Spleen (Figure 34iC)</td>
</tr>
<tr>
<td>+</td>
<td>Moderate</td>
<td>Feather (Figure 34iiD)</td>
</tr>
<tr>
<td>++</td>
<td>Strong</td>
<td>Air Sac (Figure 34iB Ty/05/E)</td>
</tr>
<tr>
<td>+++</td>
<td>Very strong</td>
<td>Brain (Figure 34iiE)</td>
</tr>
</tbody>
</table>
## Appendix iv

### Table 25. List of RG viruses generated in the 50-92, Ty/05 and PR8 genetic backgrounds

<table>
<thead>
<tr>
<th>Viral background</th>
<th>Segment ratio</th>
<th>Mutation panel</th>
<th>Genotype</th>
<th>HA</th>
<th>NA</th>
<th>PB1</th>
<th>PB2</th>
<th>PA</th>
<th>NP</th>
<th>M</th>
<th>NS</th>
<th>Viral titre</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>50-92</strong></td>
<td>8</td>
<td>HA, PB2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>6:2</td>
<td>PB2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>5:3</td>
<td>PB2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>4:1:3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>4:1:3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>4:1:3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>4:1:3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>5:1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>5:1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>5:1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>5:1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>5:1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td><strong>Ty/05</strong></td>
<td>8</td>
<td>PB2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>7:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>6:2</td>
<td>PB2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>5:1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>5:1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>5:1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>5:1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td><strong>PR8</strong></td>
<td>7:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>7:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td><strong>5092</strong></td>
<td>5:3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td><strong>TY/05</strong></td>
<td>5:3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
</tbody>
</table>

**Key:**

- **50-92**
- **Ty/05**
- **PR8**

Viral titres recovered upon virus generation are described approximately as low (L) (<1x10^6 pfu/ml) or high (H) (>1x10^6 pfu/ml).

* A range of viral titres were recovered, depending on the PB2 genotype.
Chapter 8. References


13. **Lamb R, Choppin P.** 1979. Segment 8 of the influenza virus genome is unique in coding for two polypeptides. Proc. Natl. ... **76**:4908–4912.


16. **David M. Knipe, PhD; Peter M. Howley, MD; Diane E. Griffin, MD, PhD; Robert A. Lamb, PhD, ScD; Malcolm A. Martin, MD; Bernard Roizman, ScD; Stephen E. Straus M.** 2007. Field’s Virology, 5th ed. Lippincott Williams & Wilkins.


46. Umbach JL, Yen H-L, Poon LLM, Cullen BR. 2010. Influenza A virus expresses high levels of an unusual class of small viral leader RNAs in infected cells. MBio 1.


89. **Shelton H, Roberts KL, Molesti E, Temperton N, Barclay WS.** 2013. Mutations in hemagglutinin that affect receptor binding and pH stability increase replication of a PR8 influenza virus with H5 HA in the upper respiratory tract of ferrets and may contribute to transmissibility. J. Gen. Virol.


139. **Maurer-Stroh S, Lee RTC, Gunalan V, Eisenhaber F.** 2013. The highly pathogenic H7N3 avian influenza strain from July 2012 in Mexico acquired an extended cleavage site through recombination with host 28S rRNA. Virol. J. **10**:139.


248. **Chakrabarti AK, Pasricha G.** 2013. An insight into the PB1F2 protein and its multifunctional role in enhancing the pathogenicity of the influenza A viruses. Virology **440**:97–104.


364. Banks J, Plowright L. 2003. Additional glycosylation at the receptor binding site of the hemagglutinin (HA) for H5 and H7 viruses may be an adaptation to poultry hosts, but does it influence pathogenicity? Avian Dis. 47:942–50.


412. **Dos Santos Afonso E, Escriou N, Leclercq I, van der Werf S, Naffakh N.** 2005. The generation of recombinant influenza A viruses expressing a PB2 fusion protein requires the conservation of a packaging signal overlapping the coding and noncoding regions at the 5’ end of the PB2 segment. Virology **341**:34–46.


Chapter 8 | References


