Quantifying the genetic basis of antigenic variation among human influenza A viruses

by

William Thomas Harvey

A thesis submitted for the degree of

Doctor of Philosophy (Ph.D.)

Institute of Biodiversity, Animal Health and Comparative Medicine

College of Medical, Veterinary and Life Sciences

University of Glasgow

April 2016

Abstract

Influenza viruses are a major cause of morbidity and mortality worldwide, with seasonal epidemics of influenza resulting in around three to five million cases of severe illness globally each year. The evolution of influenza A viruses is characterised by rapid antigenic drift, which allows mutant viruses to evade host immunity acquired to previously circulating viruses. Antigenic variation is observed across a wide range of infectious organisms and can circumvent long-lasting immunity in hosts leading to repeated infection or non-clearance. Influenza A viruses can often be effectively combatted by the immune system and vaccines also exist to protect at-risk individuals, limiting the burden of disease. However, the effectiveness of the vaccine depends on constituents being antigenically similar to circulating viruses. Antigenic drift of influenza viruses therefore requires a global surveillance system responsible for the antigenic characterisation of circulating viruses. The identification of emerging antigenic variants is critical to the vaccine virus selection process and in addition experts must anticipate which viruses are likely to predominate in forthcoming epidemic seasons. Mutations to B-cell epitopes on the surface of haemagglutinin (HA) that facilitate escape from neutralising antibodies play a key role in influenza antigenic drift. Consequently the haemagglutination inhibition (HI) assay, which measures HA cross-reactivity, is commonly used to approximate antigenic phenotype.

In this thesis, I investigate the genetic basis of antigenic variation among human influenza A viruses through analysis of HI data collected in recent decades and associated HA gene sequence data. In Chapter 2, I use phylogenetic methods and antigenic cartography to characterise the genetic and antigenic variation among the viruses studied and evaluate the usefulness of these methods for epitope identification. In Chapter 3, I extend a model developed to investigate antigenic differences among foot-and-mouth disease (FMD) viruses to former seasonal A(H1N1) viruses. By attributing variation in HI titre to amino acid differences between viruses, while accounting for phylogenetic relationships, I identify substitutions that have driven the antigenic evolution of the virus. Reverse genetics was then used to validate model predictions experimentally. In Chapter 4, I further extend the model and investigate the genetic drivers of antigenic drift among A(H3N2) viruses, comparing model results with published HI data generated using mutant

recombinant viruses. In Chapter 5, I explore the power of the identified genetic determinants for predicting antigenic relationships among A(H1N1) and A(H3N2) viruses. Specifically I show that sequence-based models can be used to estimate the antigenicity of emerging viruses directly from their sequence and that by including substitutions of smaller antigenic impact, in addition to the high-impact substitutions that are often focused on, predictions were improved. I also demonstrate the versatility of these methods by extending this sequence-based approach to predict antigenic relationships among viruses of three serotypes of FMD virus.

Determining phenotype from genotype is a fundamental challenge for virus research. It is of particular interest in the case of the antigenic evolution of influenza viruses, given the need to continually track changes in the virus population, anticipate which viruses will predominate in future seasons, and select vaccine viruses. Collectively, the results I present demonstrate an enhanced quantitative understanding of the molecular genetic basis of the adaptive phenotype of influenza viruses. The ability to quantify the phenotypic impact of specific amino acid substitutions should help to refine methods that predict, from genotype, the fitness and evolutionary success of influenza viruses from one season to the next, strengthening the theoretical foundations for vaccine virus selection. The techniques presented also have great potential to be extended to other antigenically variable pathogens and to elucidate the genetic basis of their antigenic variation.

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank Richard Reeve for his extensive supervision and guidance throughout the course of my studentship and for encouraging me to pursue a wide range of scientific interests and to proactively seek opportunities to collaborate with others. I would like to thank John McCauley at the Worldwide Influenza Centre at the Francis Crick Institute for his supervision and sharing his expertise and extensive knowledge of the influenza field with me. Rod Daniels and Alan Hay have also been great in this respect and very willing to give up their time whenever I have visited Mill Hill for which I thank them. During my time in the lab in Mill Hill, I also received great help from Vicky Gregory, Stephen Wharton, Michael Bennett, and in particular Donald Benton who gave up his time to show me the ropes in the lab, providing training and advice.

In Glasgow, I would also like to thank Dan Haydon and Roman Biek for academic assistance. As part of my supervisory team and head of the institute, I have always found Dan's input and advice concerning research questions and strategies very useful. Roman has acted as my assessor and has been available and supportive in this role. I am also grateful for the work of Florence McGarrity and Lorna Kennedy who have always being happy to help tackle any problems, and countless others within the department for their friendship, support, and willingness to help whenever they can. For designing the format of the Latex document used to present my thesis, I am very grateful to Matt Denwood. I would like to thank all of the students with whom I have shared an office during my time in the Graham Kerr, particularly Kirstyn and Hannah who have suffered my company in Room 308 longer than any others. Together with the many other inhabitants of the building and the wider IBAHCM community, these people make the department such a sociable, stimulating and hugely enjoyable place in which to do research.

I acknowledge the useful comments and suggestions made by my thesis examiners, Olwyn Byron (University of Glasgow) and Sam Lycett (The Roslin Institute, University of Edinburgh).

HI and genetic sequence data used throughout were generated by staff at the Crick Worldwide Influenza Centre, Mill Hill, London, UK (formerly the WHO Collaborating Centre for Reference and Research on Influenza, MRC National Institute for Medical Research) over a number of years. These data are introduced in 2 and are further analysed in the following chapters. I also acknowledge the network of WHO National Influenza Centres and WHO Collaborating Centres

that comprise the WHO Global Influenza Surveillance and Response System, who provided the influenza viruses used in this study.

I specifically thank Donald Benton (Crick Worldwide Influenza Centre, UK) who generated the ΔK130 mutant in Chapter 3 and Jan de Jong (Erasmus Medical Centre, Rotterdam, Netherlands) for providing the A/Netherlands/1/93 virus used in reverse genetics experiments described in Chapter 3.

Finally, I'll always be grateful for all the support I've received during my time in Glasgow from my parents Ann and Jim, friends, family, and my girlfriend Emily who has been particularly supportive and encouraging during the final stages of my PhD.

Declaration

This thesis, and the work contained within it, was conducted from October 2011 to April 2016 by myself, unless stated otherwise. No part of this thesis has been submitted for another degree.

William Thomas Harvey

Abbreviations

293T Human embryonic kidney 293T cells

2D Two-dimensional

3D Three-dimensional

AIC Akaike information criterion

BAY Bayern

BE89 A/Beijing/353/89 (H3N2)

BE92 A/Beijing/32/92 (H3N2)

BE95 A/Beijing/262/95 (H1N1)

BEAST Bayesian Evolutionary Analysis Sampling Trees

BK79 A/Bangkok/1/79 (H3N2)

BMA Bayesian model averaging

BMDS Bayesian multidimensional scaling

BZ78 A/Brazil/11/78 (H1N1)

CA04 A/California/7/2004 (H3N2)

CC Collaborating Centre

cDNA Complementary DNA

DNA Deoxyribonucleic acid

EN72 A/England/42/72 (H3N2)

ESS Effective sample size

FEL Fixed-effects likelihood

FMD Foot-and-mouth disease

FMDV Foot-and-mouth disease virus

FU02 A/Fujian/411/2002 (H3N2)

FUBAR Fast unconstrained Bayesian approximation

GISN Global Influenza Surveillance Network

GISRS Global Influenza Surveillance and Response System

GTR General time reversible

HA Haemagglutinin

HA0 Uncleaved HA precursor

HA1 N-terminal section of cleaved HA

HA2 C-terminal section of cleaved HA

HI Haemagglutination Inhibition

HK Hong Kong

HK68 A/Hong Kong/1/68 (H3N2)

HKY85 Hasegawa, Kishino, and Yano, 1985

IFEL Internal fixed-effects likelihood

JC69 Jukes and Cantor, 1969

JGB Johannesburg

K80 Kimura, 1980

M1 Matrix 1

M2 Matrix 2 (Ion channel protein)

M42 Matrix protein variant

mAb Monoclonal antibody

MCC Maximum clade credibility

MCMC Markov chain Monte Carlo

MDCK Madin-Derby canine kidney

MDCK-SIAT MDCK cell line transfected with plasmid expressing 2,6 sialyltransferase

MDS Multidimensional scaling

MEME Mixed effects model evolution

mRNA Messenger RNA

MUSCLE Multiple sequence comparison by log-expectation

NA Neuraminidase

NC New Caledonia

NEP Nuclear export protein

Neth93 A/Netherlands/1/93 (H1N1)

Neth93 Δ130 Neth93 with a deletion at HA1 position 130

NIC National Influenza Centre

NP Nucleoprotein

NS1 Non-structural protein 1

NS2 Non-structural protein 2

P1 Protein 1 (foot-and-mouth disease virus)

PA Polymerase acidic

PA-X Polymerase acidic protein variant

PB1 Polymerase basic 1

PB1-F2 Polymerase basic 1 protein variant

PB1-N40 Polymerase basic 1 protein variant

PB2 Polymerase basic 2

PBS Phosphate-buffered saline

PDB Protein Data Bank

pdm09 2009 A(H1N1) pandemic

PE09 A/Perth/16/2009 (H3N2)

RdRp Viral ribonucleic acid dependent ribonucleic acid polymerase

RNA Ribonucleic acid

RNP Ribonucleoprotein

SAT South African Territories

SI87 A/Sichuan/2/87 (H3N2)

SN Serum neutralisation

SO06 A/Solomon Islands/3/2006 (H1N1)

SY97 A/Sydney/5/97 (H3N2)

TCID Tissue culture infective dose

TEM Transmission electron micrograph

TX77 A/Texas/1/77 (H3N2)

UK United Kingdom of Great Britain and Northern Ireland

US United States of America

UU Ulan-Ude

VI75 A/Victoria/3/75 (H3N2)

Vic B/Victoria/2/87

VN Virus neutralisation

vRNA Viral RNA

WHO World Health Organization

WU95 A/Wuhan/359/95 (H3N2)

X-31 High growth reassortant containing A/Aichi/2/68(H3N2) HA and NA with a A/Puerto Rico/8/34(H1N1) background

Yam B/Yamagata/16/88

CHAPTER 1

Introduction

Introduction

1.1 Influenza

Influenza viruses, identified as the causative agents of human epidemics of influenza in 1933 (Smith *et al.*, 1933), are one of the major causes of infectious disease in humans. Influenza viruses have a segmented, negative-sense, single-stranded RNA genome and are members of the family *Orthomyxoviridae*. Influenza viruses are classified into three genetically and antigenically distinct types (A, B, C, classified based on the nucleoprotein and matrix segments). Influenza A viruses infect a range of bird and mammal species, though migratory birds (shorebirds and waterfowls) are understood to be the natural reservoir of these viruses. Influenza B and C viruses have been isolated primarily from humans, though there have been reports of the isolation of influenza B virus from horses and seals (Ohishi *et al.*, 2002; Osterhaus *et al.*, 2000) and of influenza C virus from pigs (Kimura *et al.*, 1997; Yuanji *et al.*, 1983). Among influenza viruses, influenza A viruses are the most prevalent pathogen for both humans and animals and will form the focus throughout.

Influenza viruses are one of the most important causes of respiratory disease in humans and are a major cause of morbidity and mortality worldwide. Globally, seasonal influenza epidemics were estimated to result in three to five million cases of severe illness and about 250,000 to 500,000 deaths annually in years prior to 2009 (WHO, 2009). Influenza epidemics occur annually during autumn and winter in temperate regions. Influenza A viruses, together with influenza B viruses, circulate globally and are responsible for these seasonal epidemics, though the majority of severe illness is caused by influenza A viruses. The most striking feature of human influenza A viruses is perhaps their capacity to evade host immunity and cause recurrent annual epidemics of disease and, at infrequent intervals, major global pandemics (Hay *et al.*, 2001). Influenza pandemics occur due to the introduction of antigenically novel viruses into a human population that is immunologically naive to the new virus and resulted in tens of millions of deaths in the 20^{th} century (Hay *et al.*, 2001; Kilbourne, 2006). Among influenza viruses, only influenza A has the capacity to cause pandemics owing to its extended host range and greater antigenic diversity (Hay *et al.*, 2001).

The influenza A genome which is around 13 kb in size, consists of eight single-stranded

RNA segments and encodes at least 14 proteins. Each gene segment is encapsidated into ribonucleoproteins (RNPs) by the viral RNA polymerase and multiple copies of the viral nucleoprotein. RNPs are imported into the nucleus of an infected cell where the viral RNA is transcribed to give positive sense mRNA (Wise *et al.*, 2012). Segments are numbered one–eight in decreasing order of size. Segments one, four, five, and six each encode a single protein: polymerase basic protein two (PB2), haemagglutinin (HA), nucleoprotein (NP), and neuraminidase (NA) respectively. The primary protein products encoded by gene segments two and three are polymerase basic protein one (PB1) and polymerase acidic protein (PA) respectively, however additional protein products encoded by both. The proteins PB1-F2, PB1-N40, and PA-X are also encoded by gene segments two and three as a result of leaky ribosomal scanning and traslation termination-reinitiation in the case of segment two and +1 ribosomal frameshifting for segment three (Chen *et al.*, 2001; Jagger *et al.*, 2012; Wise *et al.*, 2011, 2009). The coding capacity of segments seven and eight is expanded by differential RNA splicing. Unspliced segment seven encodes the matrix one (M1) protein. Three spliced transcripts have been detected (mRNAs 2–4). Of these, mRNA2 has been demonstrated to encode matrix protein two (M2), while mRNA4 has been demonstrated to encode a M2-related variant, designated M42 (Lamb *et al.*, 1981; Wise *et al.*, 2012). A single spliced species encoded by segment eight has been described, producing nuclear export protein/non-structural protein two (NEP/NS2), while non-structural protein one (NS1) is the unspliced product (Inglis *et al.*, 1980; Lamb & Lai, 1980).

Influenza A is a pleomorphic, enveloped virus. Laboratory-adapted strains often form spherical virions that are ∼100 nm in diameter whereas samples isolated from the human upper respiratory tract show mainly filamentous forms that are ∼100 nm in diameter and up to 20 μm in length (Chu *et al.*, 1949; Kilbourne & Murphy, 1960). The lipid envelope, which is derived from the host cell membrane from which the virus has budded, is studded with two morphologically distinct glycoproteins: HA and NA, encoded by gene segments four and five respectively. The HA protein is a homotrimer responsible for binding to host cells, fusion, and cell entry, while NA is a homotetrameric sialidase responsible for cell cleavage enabling budding viruses to be released from infected cells (Gamblin & Skehel, 2010). HA, NA, and the ion channel protein (M2) are anchored in the in the lipid bilayer of the viral envelope. The ratio of HA to NA molecules in the viral envelope usually ranges between 4:1 and 5:1 (Subbarao & Joseph, 2007). M1 which is the most abundant structural protein in the virus particle lies beneath the virus envelope and associates with the RNP complex. Inside the M1 layer, each of the eight RNA segments forms a helical RNP complex, being encapsidated with NP and associated with a viral RNA dependent RNA polymerase (RdRp) which is comprised of the proteins PB1, PB2, and PA. The structure of the influenza A virus particle is summarised in Figure 1.1.

HA and NA are also the primary antigens of influenza A viruses, which are classified into subtypes according to their antigenic composition, being named after the combination of HA

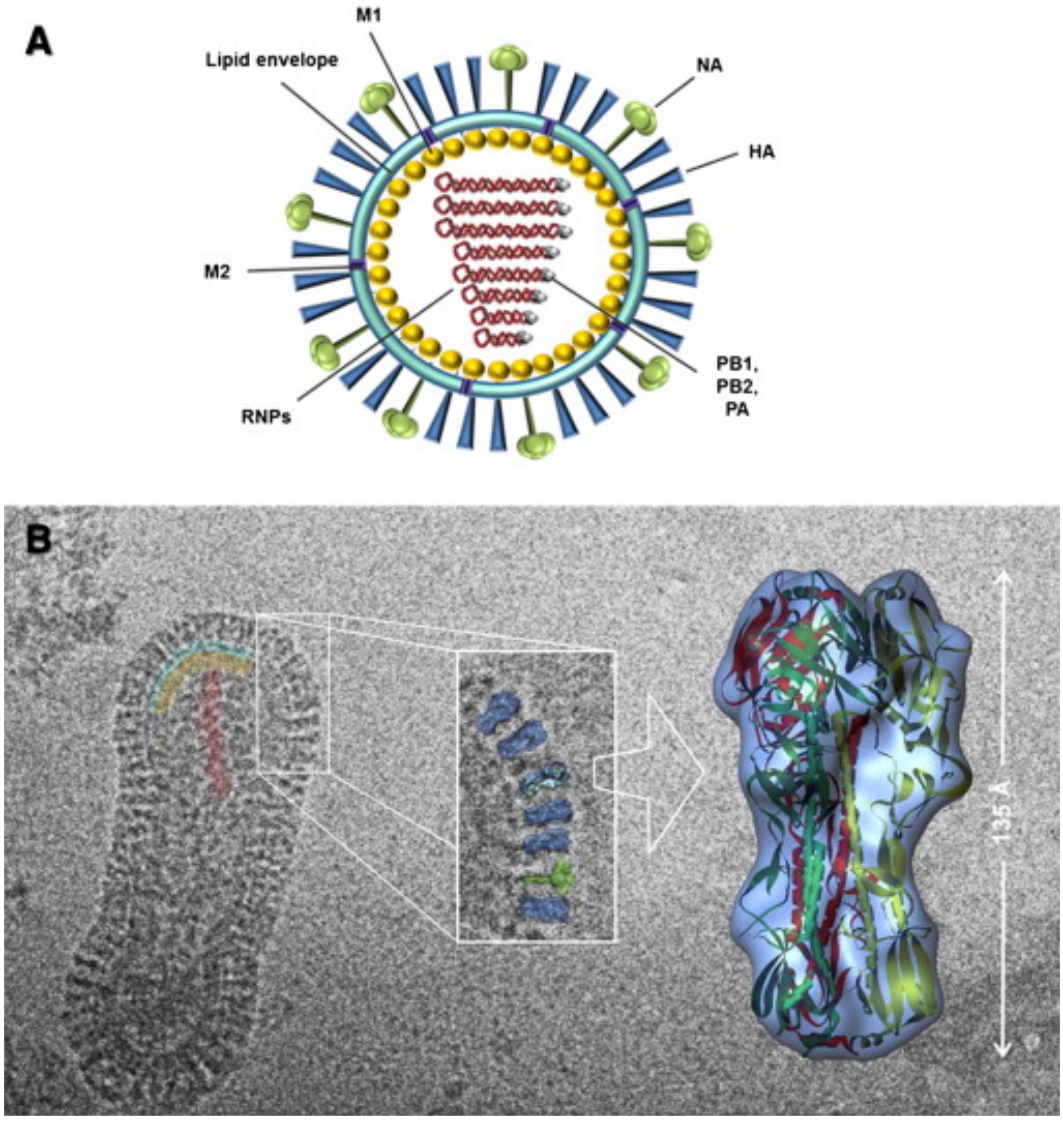

Figure 1.1: Overview of the structure of an influenza virus particle: (**A**) Schematic representation of an influenza A virus particle. The surface glycoproteins HA (blue), NA (green) and the proton channel protein M2 (violet) are embedded in the host-derived lipid envelope (turquoise). The membrane is lined on the interior with the M1 capsid protein (yellow). The viral genome consists of eight ribonucleoprotein particles (RNPs, red), with each segment formed by viral RNA (vRNA), the nucleoprotein (NP), and the viral polymerase proteins (PB1, PB2 and PA, coloured in grey). (**B**) Transmission electron micrograph (TEM) of an influenza A/X-31 virion. Regions of the lipid membrane (turquoise) and of the M1 capsid (yellow) as well as one RNP (red) are coloured as in (**A**). In the magnified section of the electron micrograph HA and NA (blue and green respectively) were overlaid with surface representations of the corresponding crystal structures filtered to an TEM-comparable resolution (attainable by 3D-TEM-reconstruction techniques) (Böttcher *et al.*, 1999). On the right side the enlarged surface representation of the HA 3D-structure is overlaid with its crystal structure (Protein Data Bank (PDB) ID: 2YPG (Lin *et al.*, 2012)), monomers in green, red, yellow. Image reproduced as appears in Figure 1 in Mair *et al.* (2014) with permission from the rightsholder.

and NA they express. To date, 18 HA subtypes and 11 NA subtypes have been identified. HA subtypes 1–16 and NA subtypes 1–9 have all been isolated from birds and reassortment among them is feasible. In recent years, novel influenza virus genomes (designated H17N10 and H18N11) have been isolated from bats (Tong *et al.*, 2012, 2013), though studies indicate that the surface proteins of these viruses have unusual structures and functions, that they are unlikely to reassort with either influenza A or B viruses, and therefore that they may be considered as belonging to a separate species or genus of the *Orthomyxoviridae* (Sun *et al.*, 2013; Zhou *et al.*, 2014; Zhu *et al.*, 2012, 2013). Antibodies specific for the M2 ion channel protein cannot neutralise infectivity but are protective *in vivo* (Webster *et al.*, 1992; Zebedee & Lamb, 1988). The cellular immune response is chiefly directed against NP, PB2 and PA (Subbarao & Joseph, 2007).

The global pandemic of 1918, caused by an influenza A(H1N1) virus, is the most severe pandemic of human disease recorded. During the 1918 "Spanish" influenza pandemic an estimated one third of the global population was infected (Taubenberger & Morens, 2006). The disease was exceptionally severe, case fatality rates are estimated to have exceeded 2.5% and total deaths have been estimated at somewhere in the region of 50 million (Johnson & Mueller, 2002; Taubenberger & Morens, 2006). Phylogenetic analysis of molecular genetic data suggest the HA and NA genes of the 1918 virus to be of avian origin (Taubenberger *et al.*, 2001). Influenza viruses of pandemic potential emerge when the genomic reassortment of virus gene segments derived from different parental strains results in the formation of new influenza virus subtypes with novel combinations of HA and NA. This process is commonly referred to as "antigenic shift".

In 1957, circulating A(H1N1) viruses that had acquired by reassortment HA, NA and polymerase basic protein 1 (PB1) segments of avian origin caused the "Asian" influenza pandemic (Scholtissek *et al.*, 1978b). These A(H2N2) viruses replaced subtype A(H1N1) viruses in the human population and circulated until their subsequent replacement by viruses of the A(H3N2) subtype that were responsible for the 1968 "Hong Kong" influenza pandemic. The pandemic virus in 1968 resulted from a reassortment event between circulating A(H2N2) viruses and avian influenza viruses in which the human strain acquired novel HA and PB1 segments (Scholtissek *et al.*, 1978b). In 1977, A(H1N1) viruses genetically and antigenically similar to those circulating two decades previously re-emerged in the human population in Russia or Northern China causing a pseudo-pandemic almost entirely restricted to persons of less than 25 years of age who did not possess protective antibodies from prior exposure (Scholtissek *et al.*, 1978a; WHO, 1977). A(H1N1) and A(H3N2) viruses have co-circulated in the human population, alongside influenza B viruses, to the present, though viruses descended from the 1977 A(H1N1) virus were replaced by a novel triple reassortant virus, A(H1N1)pdm09, that was transmitted from swine to humans causing a relatively mild pandemic in 2009 (Garten *et al.*, 2009; Smith *et al.*, 2009).

Interpandemic periods in the evolution of influenza are characterised by seasonal epidemics of varying severity (Fleming & Elliot, 2008; Hay *et al.*, 2001). Point mutations in the antigenic regions of the HA and NA glycoproteins can alter the structure of epitopes affecting antibody binding (Nelson & Holmes, 2007; Skehel & Wiley, 2000). Novel antigenic variants capable of evading the host immunity and infecting previously exposed individuals emerge by this process and may, when the selective advantage conferred is sufficient, out-compete other viruses of the same subtype, which they replace in the virus population. This process is commonly referred to as "antigenic drift". High rates of mutation and a well-connected global population with frequent migration events contribute to diversity of antigenic phenotypes exposed to selection (Bedford *et al.*, 2015), as does the mutational tolerance of the influenza structural proteins relative to those found in some other viruses (Thyagarajan & Bloom, 2014).

The antigenic evolution and epidemiology of human influenza, characterised by infrequent pandemics and annual epidemics, reflects the particular characteristics of the virus genome whereby antigenic shift (caused by genetic reassortment) and drift (caused by point mutations) can be attributed to the segmented and single-stranded RNA properties of the genome respectively. Intra-genomic recombination is not important in the evolution of influenza viruses (Strelkowa & Lässig, 2012). Studies on the HA gene have shown that high mutation rates sometimes result in the same mutation originating independently but the absence of any correlation between this occurring and distance between sequence sites indicates that any small deviations from complete association are generated by independent originations of the same mutation rather than by recombination of alleles (Strelkowa & Lässig, 2012).

The antigenic variability of influenza viruses requires a World Health Organization (WHO) coordinated Global Influenza Surveillance and Response System (GISRS) responsible for virological surveillance and the genetic and antigenic characterisation of influenza viruses (WHO, 2015a). The objectives are: 1) the early detection and characterisation of novel influenza A subtypes with the potential to cause pandemics and 2) the identification of new antigenic variants among circulating influenza A and B viruses in order to ensure that influenza vaccine components reflect the antigenic characteristics of the most prevalent viruses (Hay *et al.*, 2001).

1.2 Vaccination and antigenic drift

The influenza vaccine, first used on a population-scale in 1945 among US military personnel (Grabenstein *et al.*, 2006), remains the primary means of disease prevention and control (WHO, 2015c). It is administered biannually in anticipation of the influenza seasons that occur in winters of the Northern and Southern hemispheres. The vaccine has most commonly been trivalent comprising inactivated strains of influenza A(H1N1), A(H3N2) and B viruses

(Barr *et al.*, 2014; WHO, 2015d), although quadrivalent vaccines containing one virus from each of the two antigenically distinct lineages of influenza B viruses, B/Victoria/2/87-like (Vic) and B/Yamagata/16/88-like (Yam), have also been used in recent years (Barr *et al.*, 2014; WHO, 2015d). An intranasally administered, live attenuated vaccine accounts for a smaller proportion of influenza vaccination (Osterholm *et al.*, 2012). A meta-analysis of randomised controlled trials has suggested that influenza vaccines can provide moderate protection against virologically confirmed influenza, but that vaccine effectiveness is variable for seasonal influenza and that in some seasons protection is greatly reduced (Osterholm *et al.*, 2012). Vaccination provides minimal protection across subtypes and effectiveness within-subtype is maximised when the vaccine virus is more antigenically similar to circulating influenza viruses (Belongia *et al.*, 2009). In contrast to many other vaccines, influenza vaccines are assessed for updating twice-yearly and updates to one or more components are frequently made to provide protection against emerging antigenic variants (WHO, 2015d). When antigenic drift results in the antigenic phenotype of predominant circulating influenza viruses becoming distinct to that of the vaccine virus, the protective efficacy of the vaccine may be insufficient resulting in more serious epidemics such as observed in 1947 (Kilbourne & Smith, 2002) and 1997/98 (de Jong *et al.*, 2000). In the 2014/15 influenza season in the United Kingdom the vaccine effectiveness of the quadrivalent live attenuated influenza vaccine was reduced to 34% as a result of antigenically drifted A(H3N2) and B viruses (Pebody *et al.*, 2014).

Soon after the introduction of the first inactivated vaccines in the 1940s, it became apparent that antigenic changes occurring in the virus population reduced vaccine efficacy (Salk & Suriano, 1949). The consequences of antigenic drift for vaccine effectiveness and epidemic severity necessitate a global surveillance system which was first started following the winter of 1948/49 (Chu *et al.*, 1950). Since 1952, the WHO has coordinated monitoring of the antigenic properties of influenza viruses in humans and, more recently, in animals (WHO, 2015a). The GISRS (formerly the Global Influenza Surveillance Network, GISN) currently comprises of six WHO collaborating centres (CCs), four essential regulatory laboratories (ERLs), 142 national influenza centres (NICs) in 111 WHO member states and 13 H5 reference laboratories (WHO, 2015a). The responsibilities of the NICs include collecting samples from patients with symptoms of influenza-like illness or acute respiratory disease and reporting on influenza activity. The CCs receive clinical samples or virus isolates from NICs and carry out detailed characterisation which includes assessment of antiviral drug susceptibility, genetic sequencing and phylogenetic analysis, and antigenic analysis.

The GISRS subject matter experts then meet twice-yearly at WHO convened technical consultations (vaccine composition meetings) to review these data in the context of candidate vaccine virus selection (Barr *et al.*, 2014). Due to the logistics of the vaccine production process it is necessary for decisions on vaccine composition to be made several months in advance of the influenza season. Data on the antigenic characteristics and epidemiological features of

viruses circulating in the months up to September must be reviewed to determine whether they justify a change in the recommended composition of the vaccine ahead of the influenza season in the following year in the Southern Hemisphere, with the corresponding decision for the Northern Hemisphere being made in February (WHO, 2015d). As a result, in addition to identifying circulating antigenic variants, decision makers must attempt to anticipate which viruses will predominate in the influenza population up to a year in advance.

Given the importance of having a vaccine virus that is antigenically similar to circulating viruses, it is vitally important that the process by which influenza viruses change antigenically is well understood. The ability to accurately characterise the antigenic phenotype of circulating influenza viruses is crucial. Influenza HA and NA are both targets for activated B cells, however HA is the primary target for neutralising antibodies and principal antigenic determinant of influenza viruses (Skehel & Wiley, 2000). Because of the immunodominance of HA, decisions regarding recommendations on vaccine viruses are are made using data on HA antigenicity (Barr *et al.*, 2014). The NA in the vaccine is always from the same virus as the HA and is not independently antigenically matched to circulating strains (WHO, 2015d). For these reasons, HA has been at the focus of a majority of studies investigating the antigenicity of influenza viruses. Much of the fundamental understanding of the antigenicity of HA comes from experimental studies carried out in the 1980s (e.g. Caton *et al.*, 1982; Knossow *et al.*, 1984; Skehel *et al.*, 1984; Webster & Laver, 1980; Wiley *et al.*, 1981; Wilson *et al.*, 1981; Yewdell *et al.*, 1986) that revealed details of the structure and location of antibody-binding sites upon the surface of the HA protein. Phylogenetic studies on the influenza HA gene have helped to elucidate the mechanisms of antigenic drift of influenza A viruses in nature. HA neutralising antibodies present in the human population due to either prior infection or vaccination select for novel antigenic variants, and when the selective advantage is sufficient these antigenic variants replace circulating viruses (Fitch *et al.*, 1997).

1.3 Haemagglutinin

HA is the receptor-binding and membrane fusion glycoprotein of influenza virus and also the principal target for neutralising antibodies (Skehel & Wiley, 2000). The first study to resolve the structure of HA revealed a homotrimeric structure consisting of a stalk region extending from the virus membrane and a globular head domain containing the receptor binding site and variable antigenic regions (Wilson *et al.*, 1981). Each HA monomer contains two polypeptides (HA1 and HA2) that result from enzymatic proteolytic cleavage of a single precursor HA0 (Skehel & Wiley, 2000). The primary receptor-binding site is composed of a number of amino acid residues in a shallow depression near the head of the HA protein. The cellular receptors for influenza viruses are terminal sialic acids of glycoproteins and glycolipids (Wiley & Skehel, 1987). Neu5Ac α(2,3)-Gal and Neu5Ac α(2,6)-Gal are the two

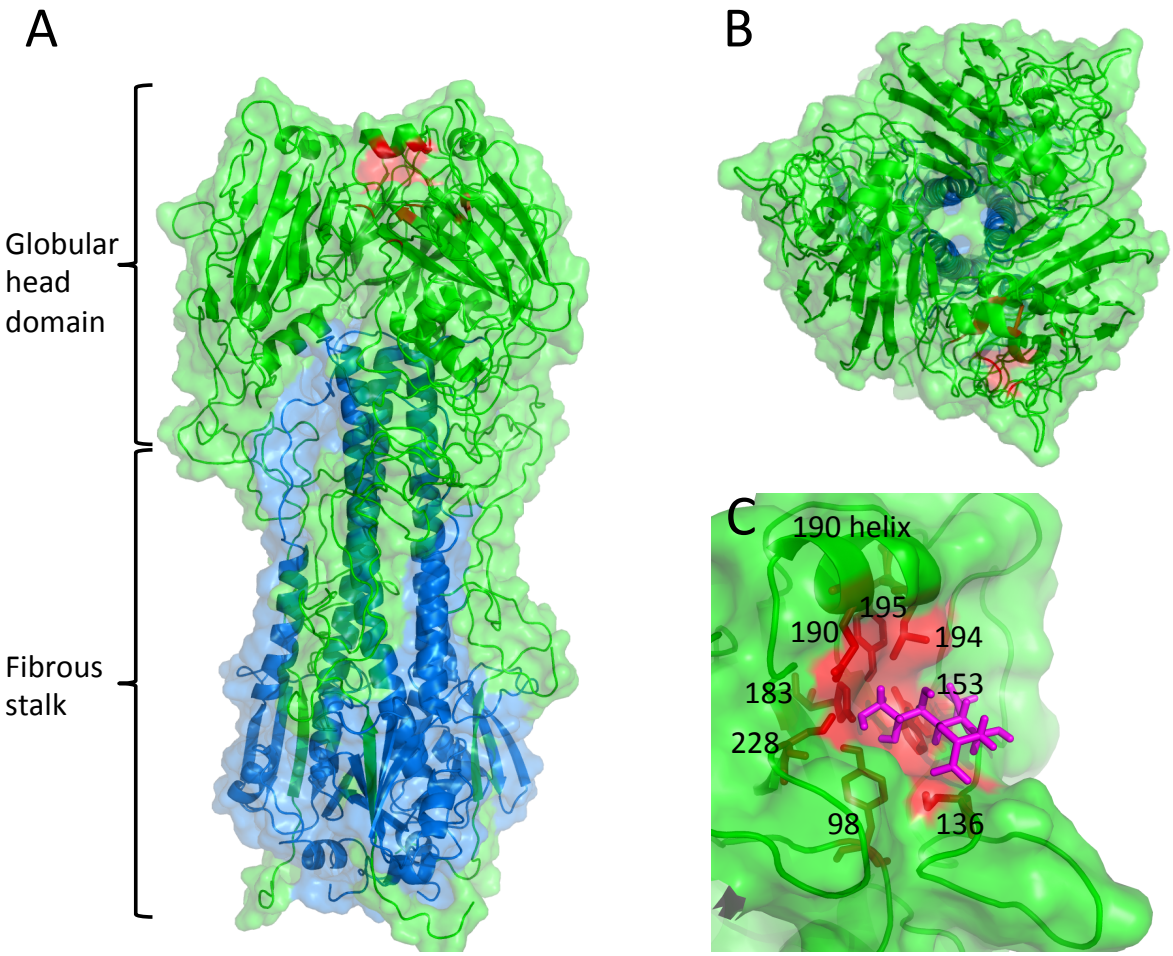

Figure 1.2: Haemagglutinin structure: Surface representation of the homotrimeric HA 3D-structure is overlaid with its crystal structure (PDB ID: 1HGG (Sauter *et al.*, 1992)). HA1 is shown in green and HA2 in blue, while the amino acids comprising the receptor-binding site of the frontmost monomer are shown in red. (**A**) The globular head domain (rich in β-pleated sheet) and fibrous stalk (rich in α-helix) are indicated. Note that the surface-exposed regions of the fibrous stalk are comprised of both HA1 and HA2 (Figure 1.3) (**B**) The homotrimeric structure of HA is more apparent when viewed from above. (**C**) Functional groups of the amino acid residues that are highly conserved among A(H3N2) viruses and have a confirmed role in binding (numbered, red) and the sialic acid receptor ($C_{11}H_{19}NO_6$, magenta) are shown. The 190 helix on the upper boundary of the binding site is also indicated.

major linkages between the sialic acid and penultimate galactose residues of carbohydrates found in nature and different HAs have different recognition specificities for these linkages (Skehel & Wiley, 2000). For example, the avian derived HA of the A(H3N2) virus responsible for the 1968 influenza pandemic changed on introduction into humans from a preference for avian receptors (α(2,3)-linked sialic acid) to a preference for human receptors (α(2,6)-linked sialic acid) (Lin *et al.*, 2012). Influenza viruses also vary in their receptor-binding avidity. Avidity refers to the accumulated strength of multiple affinities of individual non-covalent binding interactions participating in a biomolecular interaction. Receptor-binding avidity is

not necessarily the sum of the constituent individual affinities involved as other interactions occur, according to valency and structure, that influence the accumulated binding strength. The avidity of A(H3N2) viruses for the human receptor has decreased over time with a ∼4-fold reduction between 1968 and 2001, a further ∼200-fold drop during the period of 2001–2004, and a subsequent drop of such a magnitude that it could not be quantified (Lin *et al.*, 2012).

Influenza epitopes have principally been identified by a combination of structural analysis, analysis of gene sequences and the characterisation of monoclonal antibody (mAb) escape mutants, genetic variants generated under selective pressure through exposure to neutralising monoclonal antibodies (mAbs) (e.g. Caton *et al.*, 1982; Skehel *et al.*, 1984; Webster & Laver, 1980). mAbs are produced by a clonal population of immune cells derived from a unique parent cell and have affinity for a single epitope. Influenza viruses grown in the presence of a single anti-HA neutralising mAb typically acquire single amino acid substitutions that reduce antibody binding and crystallography studies have shown that these substitutions distort the structure of a single antigenic site leaving the other antigenic sites unaffected (Knossow *et al.*, 1984; Wiley *et al.*, 1981). In this context, the structure of an epitope includes the biophysical and biochemical properties of the constituent amino acid residues. Sequencing of mutants selected by exposure to neutralising mAbs can reveal amino acid positions that comprise an epitope and substitutions that allow the virus to escape the mAb binding.

mAbs that compete for binding are considered to bind to the same antigenic site, which may comprise many epitopes or antibody-binding sites which are structurally defined as the amino acids of the antigen that contact amino acids of a particular mAb. A single antibody is not likely to bind to all amino acids of a particular antigenic site but epitopes in the same antigenic site may be physically overlapping and a single amino acid substitution within that site may affect the binding of several antibodies with specificity for that site. mAb competition assays originally identified four antigenic sites (A–D) on the surface of H3 HA1 (Webster & Laver, 1980). Further mAb escape studies subsequently identified a fifth antigenic site (E) on H3 HA1 (Skehel *et al.*, 1984). Amino acid substitutions present in laboratory-selected variants derived from viruses grown in the presence of hyperimmune antisera or monoclonal antibodies have been mapped to the structure of HA and this information was used to define five topographically distinct clusters of surface-exposed amino acids on H3 HA1 that comprise antigenic sites A–E (Figure 1.3) (Wiley & Skehel, 1987; Wiley *et al.*, 1981). Similar mapping experiments have identified four comparable antigenic sites (Ca, Cb, Sa, and Sb) in similar regions close to the globular head of the HA1 of H1 HA (Figure 1.4) (Caton *et al.*, 1982).

Examination of the locations on the HA structure of amino acid substitutions in natural and monoclonal antibody-selected antigenic variants indicates that all are at surface-exposed positions and that these are predominantly on the membrane distal HA1 domain, particularly in areas surrounding the receptor-binding site (Skehel & Wiley, 2000). A notable feature of

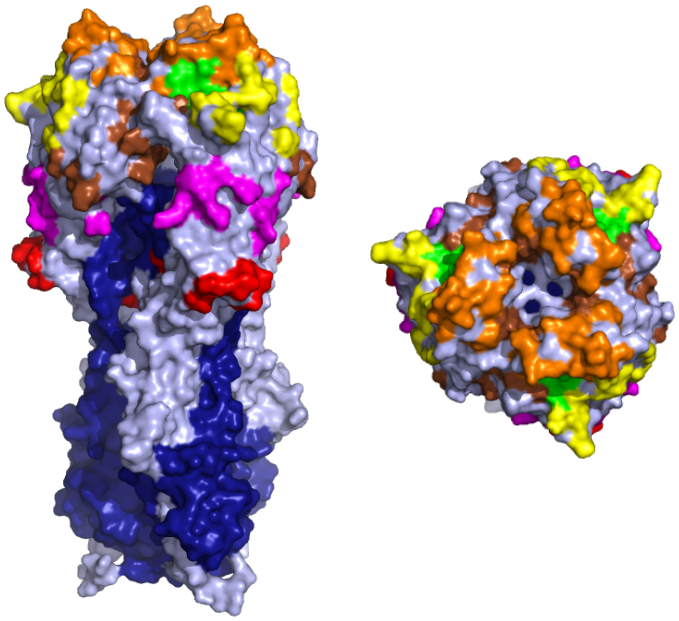

Figure 1.3: A(H3N2) HA antigenic sites: Surface representation of HA of A/Aichi/2/68 (H3N2) (Protein Data Bank ID: 1HGG (Sauter *et al.*, 1992)) with HA1 antigenic sites shown. Antigenic sites are coloured yellow (A), orange (B), red (C), brown (D) and pink (E). The primary receptor binding site is shown in green. Remaining parts of the HA1 surface are shown in light blue while HA2 is shown in dark blue.

HA antigenic sites is that they are largely composed of residues belonging to loop-like structures projecting from the HA surface. These prominent loops are often able to accept amino acid substitutions without affecting the stability or function of the HA therefore antibody-selected mutations at these sites may be tolerated with minimal associated fitness cost (Skehel & Wiley, 2000; Wilson & Cox, 1990). Constituents of antigenic sites are listed in Table 1.1.

The close proximity to the primary receptor-binding site of several residues in the HA antigenic sites results in steric hindrance of receptor association while other antibody-selected changes influence receptor-binding directly by affecting specific interactions with the sugars linked to sialic acid (Gamblin & Skehel, 2010). Examination of the structure of HA in complex with the fragment antigen-binding (Fab) of an antibody that binds at a site distant to the receptor-binding site has also been shown to neutralise infectivity by blocking receptor-binding indirectly through the bulk of either the Fab fragment or immunoglobulin molecule, indicating that disruption of receptor-binding is an important mode of neutralisation for antibodies binding across HA (Skehel & Wiley, 2000). In general the concave shape of the receptor-binding site appears to prevent the production of avid antibodies that bind directly to its principal constituents allowing these to remain conserved without the concurrent production of neutralising cross-protective antibodies (Skehel & Wiley, 2000; Wiley *et al.*, 1981). In some cases the footprint of an antibody may overlap with the receptor-binding site but these antibodies are still susceptible to escape mutations if the subset of residues that

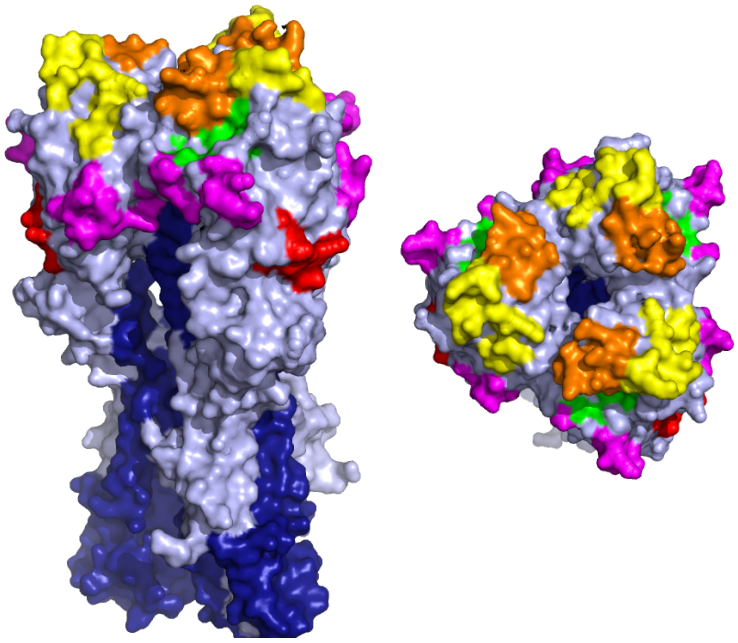

Figure 1.4: A(H1N1) HA antigenic sites: Surface representation of HA of A/Puerto Rico/8/34 (H1N1) (Protein Data Bank ID: 1RU7 (Gamblin *et al.*, 2004)) with HA1 antigenic sites shown. Antigenic sites are coloured pink (Ca), red (Cb), yellow (Sa) and orange (Sb) while the primary receptor binding site is shown in green. Remaining parts of the HA1 surface are shown in light blue and HA2 is shown in dark blue.

are important for antibody binding are not identical to those necessary for receptor-binding (Colman, 1997).

As a result of the proximity of the antibody-binding sites and the receptor-binding site, substitutions may simultaneously affect both antigenicity and receptor-binding properties including specificity and avidity (Daniels *et al.*, 1984; Hensley *et al.*, 2009; Li *et al.*, 2013). For example, two receptor-binding mutants of X-31 A(H3N2) influenza virus which differ in receptor-specificity (recognising either α(2,3)- or α(2,6)-linked sialic acid as the result of a single amino acid difference at HA1 position 226) could also be distinguished between antigenically using a panel of mAbs (Daniels *et al.*, 1984). Passage of A(H1N1) virus in immunised mice was also shown to select for single amino acid HA substitutions that increased receptor-binding avidity (Hensley *et al.*, 2009). Subsequent passage in naive mice resulted in compensatory avidity-decreasing mutations suggesting that increases to avidity may play an important role in escape from polyclonal antibodies (Hensley *et al.*, 2009).

It is not clear whether or not the assertion made by Hensley *et al.* (2009), that antigenic drift of influenza A viruses is driven by changes to receptor-binding avidity, is compatible with the observed long-term decrease in avidity of A(H3N2) viruses (Lin *et al.*, 2012), though Hensley *et al.* (2009) suggest that alternating passage between immune and naive individuals (e.g. children) could resolve this apparent contradiction. Recent work on differences in the

Table 1.1: HA1 amino acid residues in defined antigenic sites and the primary sialic acid receptor-binding site.

Subtype	Biological role		HA1 amino acid position
A(H1N1)	Antigenic site*	Ca	137,138,139,140,141,142,166,167,168,169,170,203,204, 205,221,222,235,236,237
		Cb	69,70,71,72,73,74
		Sa	124,125,153,154,155,156,157,159,160,161,162,163,164
		Sb	184,185,186,187,188,189,190,191,192,193,194,195
	Receptor binding site†		94,131,133,150,152,180,187,191,223,225
A(H3N2)	Antigenic site‡	A	**122**,**124**,**126**,130,**131**,132,**133**,**135**,**137**,138,140,**142**, 143,**144**,**145**,**146**,150,152,168
		B	128,129,**155**,**156**,**157**,**158**,**159**,**160**,**163**,**164**,165,**186**, 187,**188**,**189**,**190**,**192**,**193**,194,**196**,**197**,198
		C	44,45,46,47,48,**50**,51,**53**,**54**,273,**275**,**276**,**278**,279,280, 294,297,**299**,300,304,305,307,308,309,310,311,312
		D	96,102,103,117,**121**,167,170,171,**172**,**173**,**174**,175,176,177, 179,182,**201**,203,**207**,208,209,212,**213**,214,215,216,**217**,218, 219,**226**,**227**,228,229,230,238,240,**242**,**244**,246,247,**248**
		E	**57**,59,**62**,**63**,**67**,**75**,**78**,80,**81**,**82**,**83**,86,87,88,91,**92**,**94**, 109,**260**,261,**262**,265
	Receptor binding site§		**98**,134,135,**136**,137,138,**153**,155,**183**,**190**,**194**, **195**,224,225,226,227,**228**

* Reported by Brownlee & Fodor (2001).

† Reported by Skehel & Wiley (2000) and Soundararajan *et al.* (2011).

‡ All residues are reported by Bush *et al.* (1999a) which builds on previous studies (e.g. Wiley *et al.*, 1981; Wilson & Cox, 1990). Subset in bold reported by Shih *et al.* (2007).

§ Reported by Skehel & Wiley (2000); Weis *et al.* (1988). Bold positions are highly conserved and confirmed using crystallographic studies of complexes with receptor analogs (Skehel & Wiley, 2000).

patterns of global circulation and rates of antigenic drift of A(H1N1) and A(H3N2) viruses has revealed differences in age of infection as a potentially important factor, with A(H1N1) viruses being associated with lower ages of infection (Bedford *et al.*, 2015). If A(H1N1) viruses disproportionally affect younger, naive individuals it could be that avidity differences are more important relative to antigenic differences in this subtype.

In addition to substitutions that directly impact on antibody binding or receptor-binding, substitutions introducing glycosylation sites have been observed during HA evolution. The introduction of N-linked glycosylation sites near the antigenic sites represents another mechanism for influenza A to escape antibody neutralisation and has been described in human H3 and H1 viruses (Gambaryan *et al.*, 1998; Skehel *et al.*, 1984). Oligosaccharides bound to these sites can sterically inhibit existing antibodies from accessing their binding footprint or shield

nearby epitopes preventing these from inducing an antibody response. Human H3 viruses have acquired novel glycosylation sites at a greater rate, compared with H1 viruses, with four potential sites introduced between 1968 and 1999 (Skehel & Wiley, 2000). The introduction of glycosylation sites can also influence receptor-binding. Generally, glycosylation reduces avidity, due to either steric hindrance or an overall reduction in surface charge, preventing hyperglycosylation from being an effective means of escaping immunity (Das *et al.*, 2011).

In summary, neutralising antibodies may select for HA substitutions that change the structure or biophysical properties of epitopes at various locations on the protein, cause changes to receptor-binding avidity, and introduce or remove glycosylation sites. Human influenza A viruses can therefore evade immune pressure by a variety of mechanisms, and while the molecular basis of these mechanisms is relatively well characterised, the relative contributions of each of these mechanisms to antigenic drift, and consequently the fitness impact of specific substitutions in adaptive evolution more generally is less well understood.

1.4 Characterising antigenic phenotype

Effective surveillance and antigenic characterisation of circulating influenza viruses is critical for ensuring that vaccination provides adequate protection against circulating viruses. Ultimately we are interested in the the level of *in vivo* cross-protection conferred by the vaccine virus. We want to know what level of protection antibodies raised in response to the vaccine virus will provide against other influenza viruses of the same subtype circulating in the human population. Most aspects of phenotype (e.g. metabolic rate, bird wing span, virus receptor-binding avidity) can be measured independently for a given individual, however in the case of antigenic phenotype we are instead interested in the contrasts between individuals and the degree of antigenic similarity to other individuals (viruses). To assess the antigenic novelty of circulating viruses, antigenic similarity to existing antigenic phenotypes, especially those of the vaccine viruses, must be assessed. It is of critical importance to the vaccine virus selection process that the GISRS is able to accurately quantify cross-reactivity and the antigenic similarity of strains, and whether exposure to the vaccine virus will protect against emerging viruses. Antigenic similarity is typically assessed by measuring the degree of viral cross-reactivity. Cross-reactivity refers to the extent to which antibodies react with an antigen that differs from the immunogen, or in this context, the extent to which antibodies raised against one influenza virus react with a second virus. It is sometimes also referred to as cross-immunity or cross-protective immunity, though cross-reactivity does not necessarily equate to cross-protection.

The antigenic phenotype of human influenza viruses is most commonly characterised using the haemagglutination inhibition (HI) assay. The HI assay is used as a surrogate for *in vivo* protection and shows only an indirect relationship as each titre value reflects a ternary

reaction between antigen, erythrocyte and antibody, and as such may be influenced by a number of variables in each of these, including limitations in the accuracy of their dilution. A schematic illustration of the HI assay is shown in Figure 1.5. In the absence of influenza virus, a suspension of erythrocytes in phosphate-buffered saline (PBS) will sink and form a button in the bottom of a V-shaped well. When a sufficient amount of virus able to bind to the receptors on the erythrocyte surface is present when erythrocytes are added, haemagglutination occurs and the erythrocytes are bound in a diffuse lattice giving the well a cloudy appearance. When antiserum contains antibodies that are able to bind to the virus, haemagglutination is inhibited and the erythrocytes sink to the bottom of the well giving a button appearance. Antigenic relationships are inferred by measuring the effectiveness of polyclonal antibodies in post-infection ferret antisera, raised against a panel of reference viruses, to inhibit agglutination of erythrocytes by viruses isolated from clinical samples (Barr *et al.*, 2014). Post-infection antisera are serially diluted (usually two-fold) and a titre value is generated for each combination of virus isolate and reference virus that represents the reciprocal of the maximum dilution of antiserum raised against the reference virus that is able to inhibit HA binding to receptor analogs on the erythrocyte surface, and is used to estimate the antigenic similarity of a pair of viruses (Hirst, 1942, 1943).

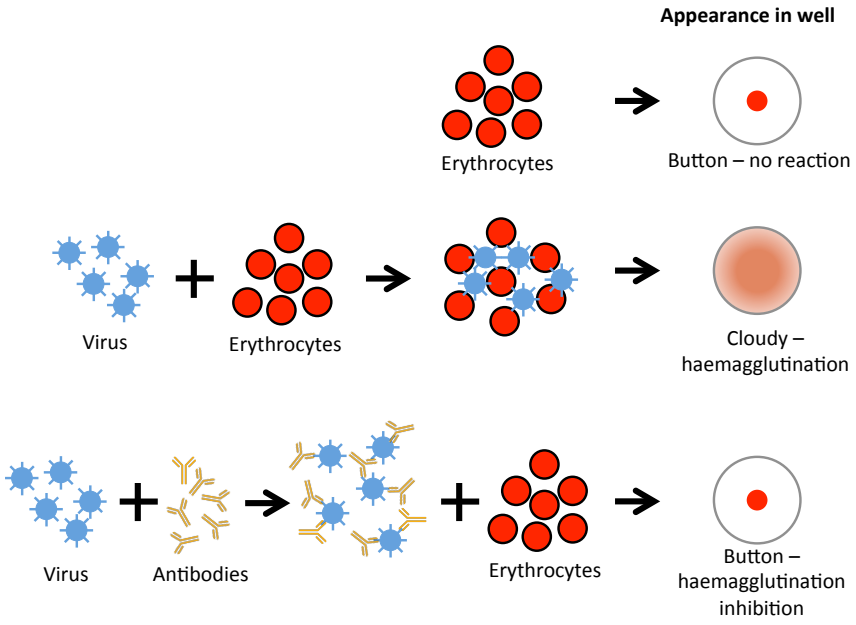

Figure 1.5: HI assay schematic: An illustration of the components of the HI assay, interactions between particles, and the resulting appearance of the well. In the top row, erythrocytes in the absence of virus particles sink to the bottom of a V-shaped well giving the button appearance. In the middle row, virus particles haemagglutinate erythrocytes resulting in a cloudy appearance. In the bottom row, antibodies bind to virus particles inhibiting haemagglutination resulting in the button appearance produced in the absence of virus particles.

The HI assay is powerful in comparison with other immunological assays for characterising

antigenic phenotype as it is relatively quick, inexpensive and simple to perform in a high-throughput manner in laboratories across the world, however it also suffers from a number of limitations. Several of these limitations are shared with other immunological binding assays while some are specific to the HI assay. Working with a human virus necessitates working with reagents from outside the natural host. Using human antiserum to characterise viruses is problematic as infection history cannot usually be determined and may be the result of multiple infections and boosts. In contrast, post infection ferret antiserum can be raised in naive animals and this has become the accepted model for this system. It is, however, possible that discrepancies in the immune response recognised by human and ferret immune cells complicate the relationship between ferret post-infection HI titres and protection in humans (Lee & Yang, 2003; Nolan *et al.*, 2003), though there is no convincing evidence. Agglutination of non-human erythrocytes (usually guinea pig, turkey or chicken) is used as a surrogate for attachment to the human receptor. The HI assay therefore measures the reaction between ferret antibodies, non-human erythrocytes and human viruses that may have acquired changes upon propagation in cells or hen's eggs.

It is recognised that HI titres can be affected by a variety of non-antigenic factors. In particular, it was quickly noted after the introduction of the HI assay that changes in viral receptor-binding avidity can impact on HI assays (Archetti & Horsfall, 1950; Hirst, 1943). Influenza viruses that bind to erythrocytes with increased avidity may require a higher concentration of antibodies to overcome this and inhibit agglutination in the HI assay. This may have important consequences for the interpretation of the HI assay. Increased avidity may reduce HI titres resulting in the virus being falsely defined as antigenically novel, while viruses with low avidity may be falsely defined as being more antigenically similar to the reference virus than they actually are (Daniels *et al.*, 1984; Hensley *et al.*, 2009; Yewdell *et al.*, 1986). Experimental studies of A(H1N1) have shown that this effect can result from a single amino acid substitution (Hensley *et al.*, 2009).

Somewhat counter-intuitively, it is also possible that increased viral receptor-binding avidity may lead to higher HI titres. Influenza samples are standardised prior to HI based on their ability to agglutinate erythrocytes using the haemagglutination assay (Ndifon, 2011). A standardised number of haemagglutinating units of virus are used in HI, usually four or eight (i.e. a sample that can agglutinate a standardised suspension of erythrocytes when diluted four- or eight-fold, but not when diluted eight- or 16-fold respectively). A virus that binds to erythrocytes with decreased avidity will require more virus particles to agglutinate the same number of erythrocytes in the haemagglutination assay resulting in a higher concentration of virus particles in the HI assay, which may require more antibodies for total inhibition of agglutination and thus result in higher HI titres. The complexities of particle interactions of this kind in the HI assay are explored in depth using biophysical models by Ndifon (2011).

Similarly, differences in the magnitudes of titres measured using different reference strains

arise, perhaps due to differences in their immunogenicity or antibody-inducing capacity, further complicating inference of results (Hirst, 1943). Antisera raised against some viruses appears to be more potent resulting in higher titres against viruses in general, regardless of antigenic differences. To account for differences of this kind HI titres are sometimes considered relative to the homologous titre recorded when antisera is used to test the reference virus used to generate it (e.g. Klimov *et al.*, 2012) or the maximum titre recorded using that antiserum (Smith *et al.*, 2004). Furthermore, differences in the immune-response of individual ferrets may influence HI titres.

In addition to the limitations described above, the behaviour of A(H3N2) viruses in haemagglutination and HI assays has changed in recent years and, as a result, interpretations have become even more complex (Barr *et al.*, 2010; Lin *et al.*, 2010). In particular, the ∼200-fold reduction in the avidity of A(H3N2) for the human receptor during the period 2001–2004 and greater subsequent drop described above, resulted in a change in reactivity against erythrocytes of various species and necessitated a switch in HI from turkey to guinea pig erythrocytes (Lin *et al.*, 2012). Participation of the virus NA in the agglutination of erythrocytes by A(H3N2) viruses has also been observed to influence HI titres and to control for this, HI assays may be performed in the presence of the neuraminidase inhibitor oseltamavir (Lin *et al.*, 2010).

Virus neutralisation (VN) assays (plaque reduction and microneutralisation) are the principal alternatives to the HI assay for characterising antigenic phenotype of influenza viruses. An advantage of VN assays is that virus neutralisation is more biologically relevant than inhibition of binding to erythrocytes and studies have shown that VN assays are more sensitive than HI assays for detecting antigenic differences (Belshe *et al.*, 2000). Nonetheless, HI and VN titres are usually well correlated and given that VN assays are relatively time-consuming and difficult to perform, HI assays remain the principal method of antigenic characterisation (Barr *et al.*, 2014; Belshe *et al.*, 2000).

The results of many HI assays, together with the results of VN assays performed using smaller numbers of viruses, are assessed in order to identify patterns of antigenic cross-reactivity among circulating influenza viruses and to identify antigenic variants (Barr *et al.*, 2014). In addition to these antigenic studies that use post-infection ferret antisera, human serum panels obtained pre- and post-vaccination with seasonal influenza vaccine formulations are used to assess the coverage afforded by current vaccines against panels of representative recently circulating viruses. These data in combination are used to inform the decision on whether or not to re-formulate the seasonal vaccine (Barr *et al.*, 2014).

In summary, the HI assay, which remains the most important assay for the antigenic characterisation of influenza viruses, reflects a complex interplay of factors including antibody binding, serum potency, and virus receptor-binding avidity and specificity. In addition to these various factors that directly impact HI titres, the assay is affected by a degree of exper-

imental variation stemming from multiple sources. The complex nature of HI data requires expert decision-makers with the ability to identify patterns in tables of HI titres that indicate the presence or absence of genuinely antigenically distinct viruses among the circulating strains. Even when antigenic differences between viruses are apparent, these factors further complicate efforts to quantify the antigenic impact of specific genetic differences. For instance, with an emerging clade of antigenically distinct viruses, the antigenic similarity of these viruses to previously circulating viruses may be under- or over-estimated if the viruses in the emerging clade also have lower or higher titres as a result of altered receptor-binding avidity. The ability to accurately characterise the influenza virus population is determined, at least to some extent, by the range of antigenic phenotypes represented by the panel of antiserum used. There are practical limitations on the number of reference viruses that can be used, so choice of reference virus becomes another important decision in the surveillance process that relies on expert knowledge of the system.

1.5 Computational methods

A variety of mathematical, statistical and computational techniques have been used to investigate antigenic drift of influenza A viruses. These methods have used various types of data, but principally have involved the analysis of either genetic sequence data, antigenic data from the HI assay, or both. These data have been used explore a diverse range of research questions relating to antigenic drift. Among these, various approaches have been used to attempt to identify which codons/amino acid residues are important in adaptive evolution of influenza viruses, to characterise the dynamics of antigenic evolution relative to the molecular evolution of the virus, and to explore possibilities for predicting the future composition of the virus population based on knowledge of the current population and historical events.

1.5.1 Phylogenetic analysis of haemagglutinin

Phylogenetic analysis is used to infer evolutionary history from genetic sequence data. Studying the shape of inferred phylogenies can allow inferences of how epidemiological, immunological and evolutionary processes act upon the genetic sequence in question. In Figure 1.6, a phylogenetic tree for the HA1 of 229 human A(H3N2) viruses isolated between 1968 and 2013 is shown. Each tip or terminal node of the phylogeny (filled black circles in Figure 1.6) represents an actual sampled virus of known sequence. Internal nodes of the phylogeny represent hypothetical ancestors and the branching pattern or tree topology describes the evolutionary relationships between nodes. The phylogeny in Figure 1.6 is rooted and the inferred position of the root, or ancestor of all sampled taxa, is indicated by a red diamond. Alternatively, phylogenetic trees may be unrooted. An unrooted tree only positions individual taxa relative

to one another without indicating the direction of the evolutionary process.

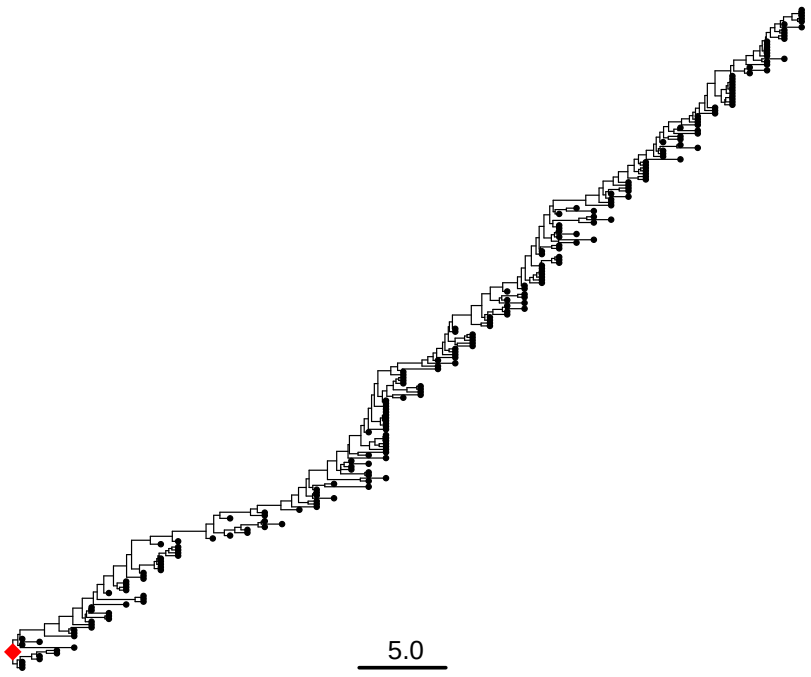

Figure 1.6: Typical A(H3N2) HA1 phylogenetic tree: The maximum clade credibility, time-resolved HA1 phylogenetic tree for 229 A(H3N2) viruses isolated between 1968 and 2013. The position of the root is indicated with a red diamond. Branch lengths are measured in evolutionary time in terms of years rather than substitutions and the scale is indicated.

In Figure 1.6 genetic change to evolutionary lineages is portrayed in the horizontal dimension. The vertical dimension in Figure 1.6 has no meaning and is simply used for visualisation purposes. Vertical lines in the branching structure therefore simply indicate the topology of the phylogenetic tree and their length is irrelevant. At each node in the phylogeny shown in Figure 1.6, the branch leading to the greater number of tips has been positioned higher in the vertical dimension. The length of a branch in the horizontal dimension represents the amount of genetic change. The units of branch length are usually nucleotide substitutions per site (i.e. the number of nucleotide changes divided by the length of the sequence). Models of nucleotide substitution provide a statistical description of the evolutionary process and are used to compute genetic distance. The simplest possible nucleotide substitution model, introduced by Jukes and Cantor in 1969 (JC69), specifies that equilibrium frequencies of the four nucleotide bases are equal and that any nucleotide has the same probability to be replaced by any other (Jukes & Cantor, 1969). The K80 model distinguishes between transitions and transversions while retaining the assumption of equal base frequencies (Kimura, 1980). Contrastingly, the F81 model is an extension of the JC69 model in which base frequencies are allowed to vary from 0.25 (Felsenstein, 1981). The HKY85 model can be thought of as combining the extensions made in the K80 and F81 models as it distinguishes between

transitions and transversions and allows unequal base frequencies (Hasegawa *et al.*, 1985). The general time reversible (GTR) model allows each of the nucleotide substitution rates between four bases and frequencies of the four bases at equilibrium to vary (Tavare, 1986; Yang, 1994). To account for differences in the rate of evolution at different sites in a DNA sequence, a gamma (Γ) distribution of rates across sites can be used and in addition a proportion of invariant sites can be estimated (Yang, 1993). In Figure 1.6 branch lengths are however shown in time rather than nucleotide substitutions per site. The phylogeny shown in Figure 1.6 is therefore said to be time-resolved and has been generated using a molecular clock model. The molecular clock is a hypothesis that mutations and substitutions occur in lineages of a tree at a particular rate (Zuckerkandl & Pauling, 1965). Therefore, if all lineages of a tree are from the same time they should all have the same genetic distance from the root. If taxa are from different times, the distance of a particular sequence from the root of the tree should be proportional to the amount of time that has accumulated from the root to the time of sampling. For a viral population sampled through time, the differences in date of isolation can be used to calibrate the rate (nucleotide substitutions per site per unit time) of the molecular clock. A time-resolved phylogeny allows for the date of divergence events in the evolutionary history of sampled taxa to be estimated. The existence of a molecular clock seems to support the neutral theory of evolution and deviations from clock-like behaviour may reveal adaptive evolution, relaxing functional constraints, or changes in effective population size (Kimura, 1968).

BEAST (Bayesian Evolutionary Analysis Sampling Trees) is a program for Bayesian analysis of molecular sequences using Markov chain Monte Carlo (MCMC) used at various points points throughout this thesis (Drummond *et al.*, 2012). BEAST is used to co-estimate phylogenies and divergence times. Phylogenies estimated by BEAST are rooted, bifurcating and time-resolved inferred using strict or relaxed molecular clock models. Evolutionary hypotheses can also be tested without conditioning on a single tree topology. BEAST uses MCMC to average over tree space, so that each tree is weighted proportional to its posterior probability. When selecting a single tree is chosen from the posterior sample, the maximum clade credibility (MCC) tree is usually identified. The maximum credibility method evaluates each of the sampled posterior trees. Each clade within the tree is scored based on the proportion of times that it appears in the set of sampled trees and the product of these scores is taken as the trees overall score. The tree with the highest score is the MCC tree.

Phylogenetic trees generated using the HA gene of human influenza A viruses are characterised by the presence of a single predominant trunk lineage, relatively short side branches, and the absence of deep bifurcations (Fitch *et al.*, 1997; Nelson & Holmes, 2007). A typical example is shown in Figure 1.6. This branching structure arises due to rapid turnover in the HA population; contemporaneous virus lineages coalesce to a common recent ancestor at a rate that exceeds the temporal span of samples used to generate the tree (Bedford *et al.*, 2011). This is in contrast to, for example, the HA gene of the measles virus which gener-

ates bushier trees with far deeper branching (Bedford *et al.*, 2011). Influenza A(H3N2) HA phylogenetic trees are, in particular, characterised by this high rate of lineage coalescence, the absence of deep bifurcations or distinct lineages that co-circulate for several years, and a relatively low effective population size (Bedford *et al.*, 2011). It is striking that the immune pressure exerted upon A(H3N2) HA results in a single predominant trunk lineage rather than antigenic diversification as is observed in some other virus proteins such as the foot-and-mouth disease virus (FMDV) capsid protein whose adaptive evolution is also shaped by immune pressure (Logan *et al.*, 1993). This predominance of a single lineage may be linked to the global circulation patterns of A(H3N2) viruses; the global population is so highly interconnected that genetic variants do not persist locally between epidemics but are instead postulated to be reseeded from South and Southeast Asia (Russell *et al.*, 2008). In contrast, global movements of A(H1N1) viruses occur less frequently, with genetic lineages sometimes persisting locally across several seasons (though still only for ∼9 months on average) (Bedford *et al.*, 2015). This co-occurs with a lower rate of coalescence, a bushier phylogenetic tree with deeper bifurcations, and a higher effective population size exhibited by A(H1N1) viruses, relative to A(H3N2) (Bedford *et al.*, 2015). Though it is not as apparent as in A(H3N2), A(H1N1) HA phylogenetic trees still exhibit a predominant trunk lineage and are relatively lacking in deep bifurcations when compared with phylogenetic trees generated for the HA of influenza B viruses or other viral genes in general (Bedford *et al.*, 2015).

1.5.2 Genotypic analysis

Various approaches utilising genetic sequence data have been used to investigate antigenic drift of influenza A viruses and to identify antigenically important substitutions. Several studies have estimated ratios of rates of synonymous (dS) and non-synonymous (dN) mutation across whole genes, parts of genes or at specific codons using reconstructed phylogenetic trees (e.g. Bush *et al.*, 1999b; Fitch *et al.*, 1997; Yang, 2000). Estimated dN/dS ratios that are significantly different from unity (one) indicate adaptive evolution. dN/dS ratios greater than one indicate positive selection, while ratios lower than one indicate purifying or negative selection. Analysing an alignment of A(H3N2) HA1 sequences, Fitch *et al.* (1997) observed there to be an excess of non-synonymous mutations in the side branches and tips of the phylogenetic tree compared with the trunk lineage and identified 25 HA1 codons with dN/dS ratios indicating positive selection and six indicating purifying selection. A further excess of non-synonymous mutations in terminal branches leading to egg-propagated viruses indicated that dN/dS ratios were affected by genetic changes not relevant to the adaptive evolution of the virus. It is possible that the observed excess of non-synonymous mutations in terminal branches is due to the presence of deleterious substitutions in the sequences of isolated viruses that are sampled once before they are removed by selection and are not therefore important in the evolution of the virus at the population level.

In response to these potential problems, a follow-up analysis calculated codon-specific dN/dS ratios for internal and terminal branches separately (Bush *et al.*, 1999b). dN/dS ratios indicating positive selection were detected at eighteen codons in the HA1 domain when only mutations in internal branches were considered, indicating positive selection at these positions at the population level. Each of these codons coded for surface-exposed amino acid residues in positions assigned to antigenic sites A (positions 124, 133, 135, 138, 142 and 145), B (156, 158, 186, 190, 193, 194 and 197), C (275), D (121, 201 and 226) or E (262) (Bush *et al.*, 1999b). These results suggested that of the five antigenic sites defined based on characterisation of mAb escape mutants, antigenic sites A and B were more important in the adaptive evolution of influenza than antigenic sites C–E. The competitive race against the human immune repertoire is expected to be the major selective pressure causing positive selection in HA, however evidence of positive selection does not necessarily indicate that a position is antigenically important. Five of the 18 positively selected codons (135, 138, 190, 194 and 226) were also among the 16 positions identified by Weis *et al.* (1988) as being involved in receptor-binding. While some positions involved in receptor-binding at the centre of the pocket which binds sialic acid are highly conserved (Weis *et al.*, 1988), it is possible that changes to receptor-binding avidity or specificity caused by substitutions at these positions are an important component of the adaptive evolution of influenza viruses, perhaps as part of the response to immune-pressure as suggested by Hensley *et al.* (2009). Another possibility is that some of the positively selected positions undergo compensatory changes that occur alongside antigenic changes which are required to remain to maintain HA function.

In contrast to these methods which investigate mutations at individual sites, Tusche *et al.* (2012) developed a dN/dS-based method that incorporated estimates of positive selection for individual sites and also information on the spatial distances between residues on the 3D protein structure in order to detect dense patches of residues showing high average positive selection. With this approach, individual sites that would not be identified by codon-specific methods may be identified if they are in close proximity to other positions with strong evidence of positive selection. Intuitively, this is a good idea, since we know antibodies bind to epitopes composed of clustered residues. When applied to A(H1N1) and A(H3N2) HA, this method identified eight and nine patches of positive selection respectively of which all but one contained antigenic residues identified by other methods (Tusche *et al.*, 2012).

dN/dS-based methods evidently have some power in their ability to detect antigenically important positions in the evolution of influenza viruses and complement mAb escape mutant studies well. mAb studies can identify specific substitutions that inhibit binding of neutralising antibodies to HA epitopes, but these substitutions are not necessarily all important in nature. For example, some antigenic substitutions may cause maladaptive effects that are not clear in the laboratory but that have consequences for intra-strain competition. The analysis of temporally sampled virus genotypes can help to reveal which positions among experimentally defined antigenic sites have also been important in nature, and can also give

indications of their relative importance. For example these methods have commonly identified more positions showing evidence of positive selection in antigenic sites A and B, relative to sites C, D and in particular E. The identification of positions under positive selection must be interpreted with caution as inferences are made in the absence of phenotypic data and so the substitutions identified provide a good set of candidate substitutions for further consideration, for instance by testing in HI assays. dN/dS-based methods also possess some disadvantages when applied to viral datasets because the relationship between the strength of selection and dN/dS ratio can become complex when applied to sequences sampled serially through time from a single population (Bhatt *et al.*, 2011). By definition, these methods are also more sensitive to recurrent selection (repeated substitutions to the same codon) than to selected mutations that occur only once (solitary selective sweeps) (Bhatt *et al.*, 2011), which may actually be very important in human influenza A viruses given that evolution is characterised by the serial replacement of variants in the trunk lineage and not antigenic diversification (Bedford *et al.*, 2011).

Alternative approaches that utilise genetic sequence data without estimating rates of synonymous and non-synonymous mutation generally rely on tracking changes in the frequency of amino acid variants. Bhatt *et al.* (2011) analysed serially sampled A(H1N1) and A(H3N2) whole genomes using a population genetics method based on an extension of the McDonald-Krietman test (McDonald & Kreitman, 1991) to quantify the rate of adaptive fixation of substitutions across genes or parts of genes through time. This analysis found that HA1 exhibited a much higher rate of adaptive evolution than HA2 and that the rate of adaptive evolution is greater for A(H3N2) than for A(H1N1). This supported previous findings, and also revealed high rates of evolution for surface-exposed regions of NA and for internal genes (Bhatt *et al.*, 2011), though this method does not detect specific codons under positive or purifying selection. Steinbrück & McHardy (2011) presented a method that explicitly investigates changes in the frequency of mutations at specific sites through time in order to identify individual alleles or sets of mutations that might be under positive selection.

Population-genetic analysis has also been used to characterise the dynamics of antigenic evolution. A punctuated pattern of evolution has been observed in antigen-antibody binding data (Smith *et al.*, 2004) and also in the temporal distribution of fixations of amino acid substitution in the virus population (Shih *et al.*, 2007). Clustering has been described by a model of episodic evolution (Koelle *et al.*, 2006) indicating that rare high-impact substitutions are important in the adaptive evolution of influenza A viruses, however other studies investigating frequency changes and rates of mutation indicate that most substitutions to positions in antigenic sites are adaptive (Shih *et al.*, 2007). Under this scenario where beneficial mutations are common, clustering can be explained by fitness interactions or epistasis between substitutions occurring within antigenic sites (Kryazhimskiy *et al.*, 2011; Shih *et al.*, 2007). Analysis of polymorphism frequency time-series data has shown that beneficial mutations are common in HA evolution and that observed clustering is the result of competition

between HA genotypes, and not a limited supply of beneficial mutations (Strelkowa & Lässig, 2012).

1.5.3 Phenotypic analysis

Multidimensional scaling (MDS) covers a range of techniques geared towards dimensionality reduction and the graphical representation of data, where algorithms are used to identify a low dimensional space, usually Euclidean, in which points represent the pattern of proximities (i.e. similarities or distances) among a set of objects (Cox & Cox, 2000). An example of an MDS analysis of distances between pairs of geographic locations is given in Figure 1.7. In this example, the proximities among a set of objects (US airports) are geographic distances measured in miles. Figure 1.7 shows that an MDS analysis of the matrix of distances between pairs of US airports (shown in A) is capable of recovering the true geographic structure, shown in map in C.

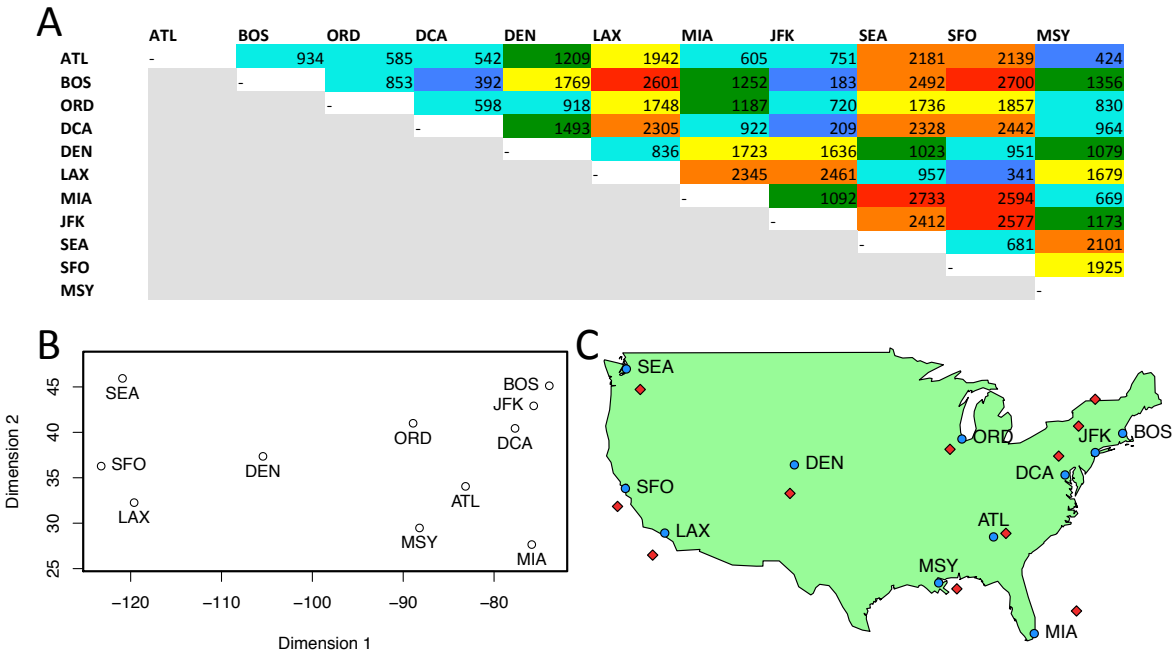

Figure 1.7: Multidimensional scaling analysis of pairwise distances between US airports: Distances in miles between 11 US airports which are indicated by their three-letter International Air Transport Association codes are shown (**A**). Distances are colour coded: red, 2501 or more miles; orange, 2001-2500 miles; yellow, 1501-2000 miles; green, 1001-1500 miles; cyan, 501-1000 miles; blue, 500 or fewer miles. An MDS analysis of these distances was used to estimate airport locations in two-dimensional space (**B**) using the R function cmdscale(). Distances in (**B**) are proportional and do not represent actual distances in miles. (**C**) Points in (**B**) were rescaled using the mean and standard deviations of true latitude and longitude coordinates and are plotted as red diamonds alongside true locations denoted by blue points.

MDS was adopted by Smith *et al.* (2004) and applied to HI data to explore antigenic similarity

of human influenza A(H3N2) viruses isolated from 1968 to 2003 and (Figure 1.8). The application of MDS to antigenic data such as HI or VN titres is commonly referred to as "antigenic cartography". Viruses are positioned on "antigenic maps" so that distances to antisera, which are positioned in the same space, best satisfy HI titres measured between them, with distance between virus points intended to be inversely related to their antigenic similarity. Antigenic maps provide a visual representation of antigenic evolution and the analysis of Smith and colleagues showed that the antigenic evolution of influenza A(H3N2) viruses was punctuated relative to molecular evolution, considered in terms of both nucleotide and amino acid substitutions (Smith *et al.*, 2004). Clusters of antigenically similar viruses on the A(H3N2) map were defined using k-means clustering analysis and, while in many cases viruses belonging to the same antigenic cluster are more distant on the map (Figure 1.8), these transitions between clusters likely represented important changes in antigenic phenotype. Antigenic cartography also allows the antigenic similarity of pairs of viruses not directly compared by HI assay to be estimated. This is useful as realistically only a small minority of viruses can be directly compared experimentally and predictions for unobserved titres can help to fill in gaps in the data providing a more complete antigenic profile of the virus population.

In Figure 1.8 the greatest amount of antigenic distance is portrayed in the vertical dimension, though the orientation of the map in antigenic spaces is free as only the relative positions of viruses and antisera can be determined. The earliest A(H3N2) viruses isolated from humans belong to purple HK68 cluster at the top of Figure 1.8 while the most recently isolated viruses belong to the orange FU02 cluster at the bottom. Broadly speaking, movement in antigenic space from top to bottom of the antigenic map in Figure 1.8 corresponds to the changing antigenic phenotype of A(H3N2) HA through time. This means that to an extent, with the orientation of map in Figure 1.8, antigenic evolution in the trunk lineage of A(H3N2) is expressed in the vertical dimension while antigenic evolution in the side branches of the A(H3N2) phylogeny is expressed in the horizontal dimension. The unusual evolutionary ecology of the influenza A HA gene that leads to distinctive phylogenies (1.6) may also result in two dimensions typically resulting in the best resolution antigenic maps generated from influenza HI data (e.g. Bedford *et al.*, 2014; Smith *et al.*, 2004) while three dimensional maps have typically been favoured for antigenic data collected for other viruses (e.g. Horton *et al.*, 2010; Ludi *et al.*, 2014) (though occasionally three dimensions have been used for influenza HI data (e.g. Sun *et al.*, 2013)).

Following the original analysis of A(H3N2) virus HI data using antigenic cartography, the technique has been used to visualise antigenic differences among other influenza viruses, including A(H1N1) (Bedford *et al.*, 2014; Sandbulte & Westgeest, 2011) (including A(H1N1)pdm09 (Garten *et al.*, 2009; Klimov *et al.*, 2012)), influenza B (Barr *et al.*, 2010, 2014; Bedford *et al.*, 2014), avian influenza A(H5N1) (Koel *et al.*, 2014), and extended to other viruses including FMDV serotype A (Ludi *et al.*, 2014), several species of lyssavirus (Horton *et al.*, 2010), and

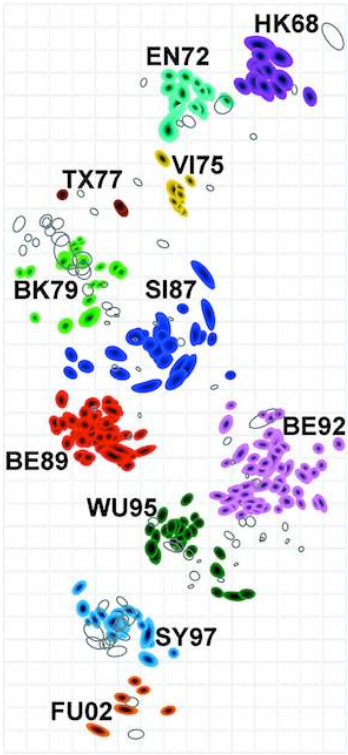

Figure 1.8: Antigenic map of influenza A(H3N2) viruses isolated from 1968 to 2003: Viruses (coloured shapes) and antisera (uncoloured open shapes) have been positioned so that distances between them represent corresponding HI data with the least error. Colours indicate antigenic clusters which were identified using a k-means clustering analysis. Clusters are named after the first vaccine strain they include with letters referring to the location of isolation (Hong Kong, England, Victoria, Texas, Bangkok, Sichuan, Beijing, Wuhan, Sydney, and Fujian) and numbers referring to the year of isolation. Vertical and horizontal gridlines both represent antigenic distance with the spacing between lines representing corresponding to a two-fold dilution of antiserum in the HI assay. As only the relative positions of antigens and antisera can be determined, the orientation of the map within these axes is free. Image reproduced as appears in Figure 1 in Smith *et al.* (2004) with permission from the rightsholder.

enterovirus 71 (Huang *et al.*, 2009). Across these studies utilising antigenic cartography, genetically similar viruses are generally placed closer together in antigenic space. For example, more genetically related species of lyssavirus tested by Horton *et al.* (2010) cluster closer together on an antigenic map generated from VN data. A notable exception to this observation is found on the antigenic map generated from VN titres measured between FMDV serotype A viruses where the three topotypes (genetically and geographically defined lineages) tested were well intermixed (Ludi *et al.*, 2014).

Antigenic cartography does not directly infer mechanistic relationships between structural changes to antigenic sites and variation in antibody binding. However, cartographic methods may be used to select amino acid substitutions for further investigation by experimentation. When clusters of virus are apparent on antigenic maps, consistent amino acid differences

between clusters represent candidate substitutions that may have caused the antigenic change. This approach has been used to determine the amino acid substitutions that have caused transitions between clusters of antigenically similar viruses on antigenic maps of A(H3N2) and A(H5N1) viruses (Koel *et al.*, 2014, 2013). Between one and thirteen amino acid substitutions were associated with transitions between clusters of antigenically similar viruses defined on the A(H3N2) antigenic map generated by Smith *et al.* (2004). Subsequent testing of all candidate substitutions by reverse genetics revealed which substitutions had caused the transitions between clusters (Koel *et al.*, 2013). Antigenic maps often lack obvious clustering of viruses (e.g. Garten *et al.*, 2009; Klimov *et al.*, 2012; Ludi *et al.*, 2014) and in these cases it is not possible to identify which substitutions may have caused the antigenic variation expressed on the map. Genetic mutations drive antigenic change and so genetically related viruses sharing a greater proportion of their evolutionary history are also likely to be antigenically similar. This results in a correlation between antigenically neutral genetic changes and antigenic distance as expressed in HI titres and on antigenic maps. The same problem is faced when trying to determine which substitutions may have caused antigenic differences represented as within-cluster variation on antigenic maps where clustering is present.

All of the substitutions identified as causing transitions between antigenic clusters of A(H3N2) and A(H5N1) viruses occurred at positions very close to the primary sialic acid-binding site (Koel *et al.*, 2014, 2013). Li *et al.* (2013) used a simple method described by Archetti & Horsfall (1950) to account for variation in HI titres caused by differences in the receptor-binding avidity of viruses. On A(H3N2) maps generated using the avidity-corrected titres the extent to which viruses formed distinct clusters was reduced, suggesting that variation in virus avidity affects positioning on antigenic maps and that substitutions that affect antigenicity are more likely to be detected as causing antigenic change using antigenic cartography if they simultaneously alter avidity (Li *et al.*, 2013). This represents an important consideration when interpreting the results of Koel *et al.* (2014, 2013), as it is possible that antigenically important substitutions that do not affect avidity are underestimated in their impact by antigenic cartography.

1.5.4 Integrating genotypic and phenotypic analyses

Various modelling approaches have been used to explore the mechanisms underlying the relationship between genotype and antigenic phenotype using models that incorporate genetic sequence data and data from antigenic analyses such as the HI assay. Lee & Chen (2004) developed a model that used the count of amino acid differences in the five H3 antigenic sites A–E to predict antigenic distances, measured by HI assay, between A(H3N2) viruses. This approach did not account for variation in the antigenic impact of substitutions within the antigenic sites that is dependent on either the position at which the substitution occurs or the amino acids involved, or for the potential antigenic impact of substitutions occurring

outside the described antigenic sites. The predictive power of Lee and Chen's model was improved by grouping amino acids into similarity classes based on properties such as polarity and charge and considering only amino acid substitutions between these classes (Liao *et al.*, 2008).

While the models described by Lee & Chen (2004) and Liao *et al.* (2008) are clearly simplistic, regression-based approaches such as these have advantages over methods that make inferences based on either genetic (i.e. dN/dS-based selection analyses) or antigenic (i.e. antigenic cartography) data only. By directly attributing antigenic variation expressed in assays such as HI to differences in virus genotype, these methods may quantify the antigenic change caused by specific genetic changes. The methods can, like positive selection analyses, identify codons linked to adaptive evolution of natural isolates, however in this case the identification is supported by phenotypic data. They can also, like antigenic cartography, be used to predict cross-reactivity of pairs of virus that have not been directly assessed experimentally on the basis of differences in genotype. However, predictions made by these models differ qualitatively from antigenic cartography in that they are based on an attempt to understand the mechanistic relationship between structural changes to epitopes and antibody binding.

Conceptually similar work has tested amino acid diversity at each HA amino acid position, as measured by Shannon entropy, as a predictor of antigenic variation in HI assay (Huang *et al.*, 2012). Applying this method to A(H1N1) data, Huang *et al.* (2012) identified as important several HA positions outwith the H1 antigenic sites identified experimentally (Caton *et al.*, 1982) but at positions in locations similar to remaining parts of A(H3N2) antigenic sites A–E, a result consistent with entropy-based analyses of H1 HA sequences that did not use HI data (Deem & Pan, 2009).

The evolutionary process creates a significant challenge for regression-based analyses that seek to attribute antigenic differences, expressed in serological assays, to genetic differences between viruses. The antigenic similarity of influenza viruses depends on amino acid differences that affect the binding of antibodies to epitopes. Antigenically important HA substitutions are expected to correlate with variation in HI titre, however non-antigenically important substitutions are also expected to correlate with reduced cross-reactivity, as a result of the shared evolutionary history of genetically and antigenically similar viruses. It has long been recognised that individuals drawn from a hierarchically structured phylogeny cannot be regarded as being independent for statistical purposes. Measures that control for shared evolutionary history when investigating genetic predictors of phenotype typically focus on properties of the individual (Felsenstein, 1985), rather than the degree of similarity or difference between individuals which is what is relevant when considering levels of antigenic cross-reactivity.

Analysing VN measurements of cross-reactivity between pairs of FMDV serotypes SAT1 and SAT2 isolates, Reeve *et al.* (2010) introduced a method for detecting antigenically important amino acid changes while controlling for the shared evolutionary history of viruses. This

method was based on the recognition that a substitution occurring at a single position in the phylogeny which correlates with reduced cross-reactivity represents a single independent piece of evidence that the substitution is antigenically important, regardless of how many different pairs of virus with reduced titres are separated by that substitution. By identifying branches of the phylogeny that correlate with reduced cross-reactivity between virus pairs and only identifying amino acid differences that improved upon this model, positions are only identified as antigenically important if they possess substitutions correlating with antigenic change in multiple branches of the phylogeny (Reeve *et al.*, 2010). This approach was used to identify surface-exposed loops of the FMDV capsid protein containing antigenically important substitutions and to make sequence-based predictions of titre. Steinbrück & McHardy (2012) applied a similar method to influenza A(H3N2) HI data and HA sequences, estimating two independent antigenic weights (up and down) for each branch of the phylogeny. Ancestral state reconstruction was then used to map substitutions to the branches of the HA phylogeny and substitutions associated with multiple branches with an average antigenic weight above a threshold value were identified. It is not clear exactly why up and down weights should differ, but it is possible that they reflect other aspects of virus phenotype such as receptor-binding avidity, something Reeve *et al.* (2010) accounted for using virus-specific corrections to titre. Methods of this kind, that account for the phylogenetic structure in cross-reactivity data, possess greater power for distinguishing antigenically important substitutions from neutral or near-neutral hitchhikers.

1.5.5 Predicting evolutionary success of genotypes

Predicting the direction of evolution of human influenza A viruses is an important aim as the vaccination process requires decision makers to anticipate which viruses will predominate in epidemics several months in advance. Early attempts at prediction showed that viral lineages with the greatest number of substitutions at positions in antigenic sites A and B or the receptor binding site showing signatures of positive selection were the progenitors of future A(H3N2) lineages in most seasons (Bush *et al.*, 1999a). Current methods for predicting the evolutionary success of circulating influenza viruses have also focused on A(H3N2) and demonstrate an ability to make informative predictions of evolutionary success in most seasons (Łuksza & Lässig, 2014; Neher *et al.*, 2014). Neher *et al.* (2014) presented a method which uses information on the shape and branching pattern of the HA phylogenetic tree to inform predictions of future increase or decrease in genotype frequency. This method therefore relies on tracking and extrapolating the frequency of genetic variants, and does not require an understanding of the mechanistic relationship between genetic changes and antibody binding. Łuksza & Lässig (2014) also used information on the frequency of genetic variants, however their predictions were also informed by an understanding of which codons are important in antigenic evolution. Their predictive model included a sequence-based estimate of virus fitness, where antigenic novelty was approximated by a count of amino acid differences in

antigenic sites reported by Shih *et al.* (2007), while substitutions outside these sites were penalised on the assumption that they are likely to be deleterious. A more sophisticated estimate of antigenic novelty could be incorporated into these models however this depends on greater understanding of the impact of specific amino acid substitutions on antigenicity, and indeed the adaptive and maladaptive fitness contributions of specific substitutions more generally.

1.6 Summary and statement of aims

Host-pathogen interactions drive the adaptive evolution of human influenza A viruses. Exposure via infection or vaccination results in lasting immunity against the virus the host is exposed to and partial immunity against infection by antigenically similar viruses. Immunity in the host population exerts a strong selective pressure on the virus population. Mutations that alter the structure of epitope regions in such a way that existing antibodies are inhibited in their ability to neutralise are positively selected and novel antigenic variants out compete other viruses and cause new epidemics. The surveillance and antigenic characterisation of circulating influenza viruses is critical to the maintenance of an effective influenza vaccine, as effectiveness depends on the antigenic similarity of the vaccine virus to circulating viruses. However, attempts to predict the antigenic novelty of genetic variants are limited by our understanding of the impact of specific genetic changes on antigenic phenotype.

The primary aim of this thesis is the development of models based on the mechanistic relationship between genetic changes and variation in the structure of antigenic sites that affects recognition by antibodies. This will primarily be done through analysis of HI data collected in recent decades and associated genetic sequence data. The HI assay measures the cross-reactivity of influenza HA and while there are other aspects of the antigenic phenotype of influenza viruses, including the targets of anti-NA antibodies and T cells, the contribution of HA epitopes to antigenicity and to the wider adaptive phenotype of influenza viruses is clear. The methods used are, however, not intended to be specific to influenza A and the HI assay but to be generalisable, capable of being extended to other antigenically variable viruses where pairwise estimates of antigenic similarity and relevant sequence data exist.

In Chapter 2, I use phylogenetic methods, analyses for the detection of positive selection, and antigenic cartography to characterise the genetic and antigenic variation among the viruses studied throughout the thesis. I evaluate various genetic measures as predictors of antigenic similarity, as expressed in HI titres and evaluate the usefulness of the aforementioned methods for epitope identification. In Chapter 3, I extend to influenza A viruses, the model used by Reeve *et al.* (2010) to investigate antigenic differences among FMDV. Reeve *et al.* (2010) developed a model to identify regions of the FMDV capsid where counts of amino acid differences between viruses reduced cross-reactivity, as expressed in VN assays. In Chapter

3, I extend this model to identify specific amino acid substitutions that have caused antigenic changes in the evolution of A(H1N1) viruses and quantify their relative impact. I then use site-directed mutagenesis and antigenic characterisation of mutant recombinant viruses produced by reverse genetics to experimentally validate model predictions. In Chapter 4, I further extend the model and investigate the antigenic drivers of antigenic drift among A(H3N2) viruses, comparing model results with published HI data generated using recombinant viruses. In Chapter 5, I explore the power that identified antigenic determinants have for predicting antigenic relationships among A(H1N1) and A(H3N2) influenza viruses and also among FMD viruses belonging to three antigenically distinct serotypes. Collectively these studies will lead to improvements in the quantification of the genetic basis of variation in antigenic phenotype with consequences for sequence-based prediction of virus fitness and evolutionary success.

CHAPTER 2

Characterisation of the molecular and antigenic evolution of human influenza A viruses

Characterisation of the molecular and antigenic evolution of human influenza A viruses

2.1 Abstract

Antigenic drift allows human influenza A viruses to circumvent immunity that exists in the human population and that conferred by vaccination, contributing to the disease burden of influenza. Increased understanding of the antigenic difference between strains is critical for surveillance and vaccine strain selection. Mutations to antigenic sites of the haemagglutinin (HA) protein result in antigenically distinct clades that co-exist for relatively short periods. When the selective advantage conferred is sufficient, novel antigenic variants out-compete other viruses and are incorporated into the trunk lineage from which all future viruses are descended. The haemagglutination inhibition (HI) assay is the principal method of antigenic characterisation. HA sequence data from influenza A(H3N2) and former-seasonal A(H1N1) isolates and reference viruses with complementary data from the HI assay were analysed. A range of genetic and phylogenetic techniques and antigenic analyses were used to characterise the molecular and antigenic evolution of the HA. Results indicate only a weak relationship between existing measures of phylogenetic and antigenic dissimilarity. Identification of known antigenic HA positions using analyses for the detection of positive selection and further antigenic analysis indicates that these methods are useful for the identification of antigenically important genetic changes. However, despite some degree of correspondence in the amino acid positions identified by these two methods, residual, unexplained variation in HI titres indicate a need for better quantitative methods to fully explore the relationship between molecular and antigenic evolution.

2.2 Introduction

A thorough understanding of antigenic evolution, and in particular the HA protein, is vital for maintenance of an effective vaccination programme. Experimental studies have been vital to the characterisation of the antigenicity of the HA protein, primarily through genetic

sequencing of monoclonal escape antibodies (e.g. Caton *et al.*, 1982; Skehel *et al.*, 1984; Webster & Laver, 1980; Wiley *et al.*, 1981). The results of these experiments and antibody competition assays have been used to map antigenic sites to the globular head region of the influenza HA glycoprotein (Caton *et al.*, 1982; Wiley *et al.*, 1981). Laboratory studies of this kind cannot however fully predict the antigenic impact of, or fitness consequences of mutations to antigenic regions of HA in nature.

Quantifying the phenotypic impact of specific mutations and related fitness consequences is a significant challenge for virus research that remains largely unresolved (Koelle & Rasmussen, 2014). Many antigenic regions are also able to influence receptor-binding avidity or specificity (Daniels *et al.*, 1984; Hensley *et al.*, 2009; Yewdell *et al.*, 1986) and the ability of monoclonal antibody escape mutants to replicate under laboratory conditions does not necessarily mean that an identified antigenic variant will be capable of competing with other viruses in natural host populations. Assessing the fitness impact of specific substitutions that affect antibody binding is complex and difficult. It is however an important goal, as it is likely that a better, quantitative understanding of the fitness consequences of specific mutations would enable better predictions of the direction of evolution among circulating viruses to be made. Recently published methods have used simple counts of non-synonymous mutations inside and outside defined antigenic sites for such purposes (Łuksza & Lässig, 2014).

The principal methods for quantifying antigenic differences among circulating influenza viruses assess the antigenic similarity of strains by evaluating the ability of a panel of post-infection antisera, raised against a set of strains representative of the known antigenic variation observed among recently circulating strains, to inhibit haemagglutination of ertythrocytes by circulating viruses (Barr *et al.*, 2014). This is primarily done using the HI assay, though virus neutralisation (VN) tests are also performed using far smaller numbers of viruses (Barr *et al.*, 2014). Variation in HI assays and various sources of non-antigenic variation complicate their interpretation and correlations between antigenic and molecular evolution make it difficult to track changes in phenotype as measured by these serological assays back to genetic differences between viruses. One solution to this problem has been to summarise these data using multidimensional scaling and to identify candidate mutations separating clusters of similar viruses on the resulting antigenic maps (Smith *et al.*, 2004) with reverse genetics being used subsequently to resolve remaining ambiguities (Koel *et al.*, 2013). Alternative methods use phylogenetic information to infer knowledge of the substitutions affecting antigenic phenotype by detecting positions under positive selection (Bush *et al.*, 1999b).

The aims of this thesis are to investigate the genetic basis of antigenic evolution by identifying the HA codons where mutations cause antigenic change, to quantify the antigenic impact of specific amino acid substitutions, and to explore the possibilities for sequence-based prediction of antigenic similarity. In this chapter, I explore the usefulness of existing methods for achieving these aims. These methods include the reconstruction of phylogenetic trees, the

detection of signatures of positive and purifying selection, and antigenic cartography. Historical A(H1N1) and A(H3N2) datasets consisting of genetic sequence data and antigenic data from the HI assay were analysed.

2.3 Materials and Methods

2.3.1 Viruses, antisera and HI assays

The genetic and antigenic data analysed in this chapter were generated by staff of the Crick Worldwide Influenza Centre, UK (formerly the WHO Collaborating Centre for Reference and Research on Influenza, MRC National Institute for Medical Research, UK). Influenza viruses were originally isolated from clinical specimens by either the WHO National Influenza Centres (NICs) or the Crick Worldwide Influenza Centre and were further propagated in either MDCK or MDCK-SIAT1 cells or, for a few viruses only, in 10-day old embryonated hen eggs. Post-infection ferret antisera were raised against cell- or egg-propagated reference viruses.

Ferrets were from a Home Office approved supplier and housed in containment level 2 in the UK Home Office approved facilities at the Mill Hill laboratories. The studies were approved by the Ethical Review Bodies of The Francis Crick Institute and the MRC National Institute for Medical Research and licensed by the UK Home Office under license numbers 80/2541, 80/2090 and previous licenses under the UK Animals (Scientific Procedures) Act 1986. Infected ferrets were monitored closely with respect to their health (e.g. lethargy, inability to feed/drink, etc.) and any that considered to be showing severe symptoms were culled by terminal anesthesia. Two weeks post-infection the ferrets were put under terminal anesthesia using a specific mixture of drugs (Vetalar, Rompun and Pentoject) dependent on the weight of the ferret and were exsanguinated to provide antiserum for use in HI studies.

HI assays were performed to compare the reactivity of cultured viruses with post-infection ferret antisera according to standard methods (WHO, 2011). Antiserum samples were treated with receptor-destroying enzyme to eliminate non-specific inhibitors present in serum, inactivated at 56°C and then two-fold serially diluted in phosphate-buffered saline (PBS) from an initial dilution of 1:10, 1:20 or 1:40. Virus samples were diluted in PBS to either four or eight haemagglutinating units. Virus dilutions were then added to the serially diluted antiserum samples and left for 30 minutes to allow antibodies to bind. Suspensions of erythrocytes in PBS were then added to antiserum:virus mixtures. Titres were recorded as reciprocals of the highest dilution of antisera that inhibited haemagglutination. HI assays for A(H3N2) were conducted using suspensions of chicken (0.75%), turkey (0.75%) or guinea pig (1.0%) erythrocytes, while turkey (0.75%) erythrocytes were used for all A(H1N1) assays. The behaviour of A(H3N2) viruses in haemagglutination and HI assays has changed during the period of data

collection potentially complicating the relationship between HI titres and antigenic change (Lin *et al.*, 2010, 2012). In the period 2005–2009 the role of neuraminidase-mediated agglutination of erythrocytes was observed to complicate the relationship between HI titres and antigenic change in A(H3N2) viruses (Lin *et al.*, 2010). HI assays for A(H3N2) in more recent years have therefore been carried out in the presence of the neuraminidase inhibitor oseltamavir.

A(H1N1) and A(H3N2) HI datasets were assembled from historical data by including all titres for which HA1 sequence for both virus and reference virus (antiserum) were known. The HI dataset assembled for A(H1N1) comprised 506 viruses including 43 reference viruses against which post-infection ferret antisera were raised, with 19,905 HI titres measured between 3,734 unique combinations of virus and antiserum, recorded on 351 dates from 1997 to 2009. Year of isolation for A(H1N1) viruses in this dataset ranged between 1978 and 2009. The larger A(H3N2) HI dataset comprised 49,766 titres for 15,855 virus-antiserum pairs measured using 1,857 viruses and antiserum raised against 194 reference viruses recorded on 732 dates. Year of isolation of A(H3N2) viruses in this dataset ranged between 1968 and 2014. Computational limitations necessitated the creation of a subset of the A(H3N2) dataset to be used in some analyses. This subset was chosen to have a high proportion of viruses used as reference viruses. It contained 7,315 titres recorded for 2,738 different virus and antiserum pairs between 229 viruses and antiserum raised against 169 different reference viruses. Both A(H3N2) datasets included HI and associated HA1 gene sequence data, primarily relating to pre-1990 viruses, published by Bedford *et al.* (2014) in addition to data collected by the Crick Worldwide Influenza Centre, UK. This additional HI dataset included titres recorded for 650 virus-antiserum pairs with 91 viruses tested using antisera raised against 36 reference strains.

2.3.2 Phylogenetic analysis

HA1 nucleotide sequences were aligned using MUSCLE alignment software (Edgar, 2004). Phylogenetic reconstruction and analysis were carried out using BEAST v1.7.5 (Drummond *et al.*, 2012) which uses a Markov chain Monte Carlo (MCMC) algorithm to explore parameter space and evaluate phylogenetic models. BEAST also uses date of isolation to calibrate a molecular clock model allowing rates of evolution along branches to be estimated. The goodness-of-fit of 88 models of nucleotide substitution were evaluated prior to analysis in BEAST using jModelTest v2.1.7 (Darriba *et al.*, 2012; Guindon & Gascuel, 2003). The preferred model from this analysis was tested against similar models in BEAST through comparison of Bayes factors (Suchard *et al.*, 2001) calculated using Tracer v1.5 (Rambaut & Drummond, 2009). The best model of nucleotide substitution was then tested with codon-position models that allow rates to differ across all three codon positions and at the third position relative to the first two. These comparisons were made using the simplest molecular clock model (strict clock) (Zuckerkandl & Pauling, 1965) and relatively uninformative priors

on all model parameters.

The strict molecular clock model, which assumes a constant rate of substitution along all branches of the phylogeny, was then tested against other molecular clock models, using the optimal model of nucleotide substitution with or without codon-partitioning. The alternative clock models tested were the random local clock model which allows substitution rate to vary across the tree, while being constant within local subtrees (Drummond & Suchard, 2010) and the relaxed (uncorrelated) molecular clock where each branch is assigned its own substitution rate which are collectively assumed to have a common probability distribution (Drummond *et al.*, 2006). The relaxed molecular clock was tested with branch substitution rates drawn from both lognormal and exponential probability distributions. Different priors on the coalescence process, based on assumptions of the demographic history of the influenza population, were also tested. A model assuming constant population size was tested against models of exponential, logistic and linear growth. Additionally the tree was reconstructed with a Bayesian skyline model where the effective population size is estimated from the data allowing more complex patterns of demographic change. Both molecular clock models and priors on demographic history were again assessed by calculation of Bayes factors in Tracer.

Autocorrelation and run length were assessed through calculation of effective sample size (ESS) using Tracer. The ESS of a parameter sampled from an MCMC is the number of effectively independent draws from the posterior distribution that the Markov chain is equivalent to. The ESS of a parameter is reduced when autocorrelation, the similarity of parameter estimates as a function of the number of samples of the Markov chain separating them, is high. High autocorrelation and low ESS indicate poor mixing which means the efficiency with which the MCMC algorithm has sampled a parameter is low and the estimate for the parameter is less likely to be accurate. The ESS of parameters was recalculated across independent MCMC simulations to assess convergence. The maximum clade credibility tree was identified from a posterior sample of 10,000 trees using the TreeAnnotator v1.7.5 program, which is distributed within the BEAST package. For A(H1N1), substitution at position 187, associated with adaptation to propagation in eggs (Gambaryan *et al.*, 1999; Raymond *et al.*, 1986; Robertson *et al.*, 1987), was assumed to be an artefact with potential to distort phylogenetic inference, so nucleotides coding for position 187 were excluded from phylogenetic analysis unless otherwise stated.

Two measures of genetic similarity were tested as correlates of HI titres. Amino acid distances were calculated as the Hamming distance between pairs of aligned HA1 amino acid sequences. Patristic distances between pairs of viruses in the HA1 phylogeny were also calculated. Patristic distances are the sum of the branch lengths (branch lengths are represented in the horizontal dimension in Figures 2.3 and 2.3 only and not the vertical dimension) separating two viruses and are therefore the genetic distance as estimated using a nucleotide substitution model. These patristic distances are referred to here as phylogenetic distances.

2.3.3 Signatures of selection

Codon-specific rates of non-synonymous (dN) and synonymous (dS) nucleotide substitutions were estimated across HA1. Estimates of dN significantly different from dS imply non-neutral evolution, with $dN/dS > 1$ indicative of positive selection and $dN/dS < 1$ indicative of purifying selection. All methods of analysis used for detecting site-by-site selection were implemented using either the HyPhy package v2.2.4 (Kosakovsky Pond *et al.*, 2005) or the associated Datamonkey webserver (http://www.datamonkey.org/) (Delport *et al.*, 2010; Kosakovsky Pond & Frost, 2005a).

At each site, dN and dS were directly estimated based on a codon-nucleotide substitution model using the fixed-effects likelihood (FEL) method (Kosakovsky Pond & Frost, 2005b). The FEL method takes into account, for each three-base sequence, the number of synonymous and non-synonymous non-stop codons accessible given a single nucleotide substitution according to the structure of the genetic code. At each codon, rates of synonymous and non-synonymous nucleotide substitution are estimated and codon-specific p-values are generated using a likelihood ratio test. These likelihood ratio tests evaluate whether the estimated pattern of mutation can be explained by a single rate of mutation (1 parameter) or if two different rates of mutation for synonymous and for non-synonymous mutations is preferable (2 parameters). The likelihood ratio test at each codon thus has one degree of freedom to test $dN \neq dS$. This analysis was repeated using the internal fixed-effects likelihood (IFEL) method, a variation of the FEL method which excludes mutations occurring in terminal branches of the phylogenetic tree when deriving dN/dS (Kosakovsky Pond *et al.*, 2006). It is useful to consider the dN/dS ratio derived using only internal branches of the phylogeny as non-synonymous mutations in terminal branches are effectively removed from the population and are therefore not important at the population level. The mixed effects model evolution (MEME) method is an extension of FEL that allows some branches to be under positive selection while others are under purifying selection (Murrell *et al.*, 2012). MEME was used to identify codons subject to episodic diversifying selection. Codons identified by FEL/IFEL/MEME methods with p-values < 0.05 are reported.

Fast unconstrained Bayesian approximation (FUBAR), an alternative method for detecting codon-specific positive and purifying selection based on dN and dS, was also used to analyse HA1 (Murrell *et al.*, 2013). FUBAR ensures robustness against model misspecification by averaging over a large number of predefined site classes, leaving the distribution of selection parameters essentially unconstrained. Codons with dN/dS ratios greater than one with posterior probabilities greater than 0.95 are reported.

The PyMOL Molecular Graphics System v1.7.7.2 (http://www.pymol.org) was used to visualise the locations of surface-exposed codons exposed to positive or purifying selection on the HA 3D structures of A/Puerto Rico/8/34 (H1N1) (Protein Data Bank (PDB) ID: 1RU7

(Gamblin *et al.*, 2004)) and A/Aichi/2/68 (H3N2) (PDB ID: 1HGG (Sauter *et al.*, 1992)). Proportions of identified codons associated with defined antigenic sites were calculated and compared with proportions calculated across the full HA1 alignments. Codons associated with A(H1N1) and A(H3N2) antigenic sites are based on those sets of 50 and 60 reported by Brownlee & Fodor (2001) and Shih *et al.* (2007) respectively (detailed in Table 1.1).

2.3.4 Antigenic cartography

Virus and antiserum locations in antigenic space were estimated using a Bayesian multidimensional scaling (BMDS) model (Bedford *et al.*, 2014). This methodology extends that introduced by Smith *et al.* (2004) by incorporating a phylogenetic diffusion process and estimates of antiserum and virus reactivity to account for variation in the immunogenicity of reference viruses and in the receptor-binding avidity of viruses. This model was used to position viruses and antisera in a two dimensional space, following Smith *et al.* (2004) and Bedford *et al.* (2014), so that distances best satisfied the HI data. After accounting for variation in titres dependent on differences in antiserum and virus reactivity, viruses and antisera were positioned so that distances between their locations on the map best reflected the inverse of observed $\log_2$ HI titres. Here, the two dimensions in antigenic space are termed the primary and secondary antigenic dimensions.

Multiple configurations of virus and serum locations in two-dimensional space may give the same likelihood, given an incomplete matrix of observed HI titres. Very antigenically distant viruses (i.e. those sampled decades apart) are rarely tested together by HI, accentuating this issue of model identifiability. Given the expectation that antigenic distance between seasonal influenza viruses increases with time, a weak prior was used to inform the expected location of viruses and antisera whereby their position in the primary antigenic dimension was informed by date of sampling (Bedford *et al.*, 2014). A weakly informative diffuse gamma prior with shape = 0.001 and rate = 0.001 was put on the annual rate of drift in the primary antigenic dimension. This assumption that temporally distant viruses are also antigenically distant accounts for the issue with model identifiability in part, producing more interpretable visualisations of the HI data.

Genetic relatedness, as estimated by the phylogenetic analysis described above, was also used to inform the locations of viruses in antigenic space. This involved estimating locations of each internal node of the phylogenetic tree, including the root, in antigenic space, in addition to the tips (viruses) by modelling antigenic phenotype as an evolutionary diffusion; a character state that evolves along branches of the phylogenetic tree according to a Brownian motion process (Lemey *et al.*, 2010). Tree topology and branch lengths were used to predict virus locations with an assumption that a node or tip location follows from the location of its parent node, with the addition of the temporal drift in the primary antigenic dimension described above

(Bedford *et al.*, 2014). Thus, prior locations of viruses were informed by date of sampling and genetic similarity to other viruses. Phylogenetic information from 50 trees sampled from the posterior of the BEAST phylogenetic analyses described above was incorporated. Trees from this set were randomly proposed and accepted following the Metropolis-Hastings algorithm (Pagel *et al.*, 2004), thus accounting for phylogenetic uncertainty (Bedford *et al.*, 2014). MCMC was used to sample from the posterior distribution of the BMDS model. This was implemented in BEAST using XML files adapted from published templates (Bedford *et al.*, 2014). MCMC samples were then analysed, and antigenic maps constructed, using R (R Core Team, 2015). Antigenic maps were constructed using the full A(H1N1) dataset and the subset of the A(H3N2) dataset.

k-medoids clustering was used to determine clusters of antigenically similar viruses according to the X and Y coordinates describing the locations in antigenic space. k-medoids clustering is a partitioning technique that clusters a set of objects (e.g. virus locations in antigenic space) into a pre-determined number of k clusters (Theodoridis & Koutroumbas, 2006). A medoid is the object of a cluster whose average dissimilarity (e.g. distance in multidimensional antigenic space) to all objects in the cluster is minimal. k-medoids clustering is conceptually similar to k-meas clustering, though it is more robust to noise and outliers because it minimises a sum of pairwise dissimilarities instead of a sum of squared Euclidean distances. k was determined using the silhouette method which determines the number of clusters by minimising the average distance between medoids and other objects assigned to the same cluster (Rousseeuw, 1987). Amino acid differences conserved between at least 0.95 of viruses in adjacent antigenic clusters were identified by comparison with HA1 alignments. The Partitioning Around Medoids (PAM) algorithm for k-medoids clustering was implemented using the R package *fpc* (Hennig, 2015).

2.4 Results

HI titres may be affected by several non-antigenic factors. Variation in titres can occur due to variability in other characteristics of viruses such as receptor-binding avidity. Similarly, different antisera may produce titres of lower or higher magnitude due to differences in the antibody-inducing capacity of the viruses used to generate them. Variation in the immune response of individual animals used to generate antiserum for use in the assay and between batches of cells may also impact titres. Other sources of experimental variation arise due to limitations in the accuracy of dilutions of reagents used and the subjective nature involved in reading assay results from the plate. The observed variation in titres recorded for any given virus-antiserum pair gives an indication of the degree of experimental variability in HI titres and is plotted for each subtype in Figure 2.1. Here, for every virus-antiserum pair with at least two observations, each $\log_2$ HI titre recorded is plotted against the mean titre recorded

for the virus-antiserum pair.

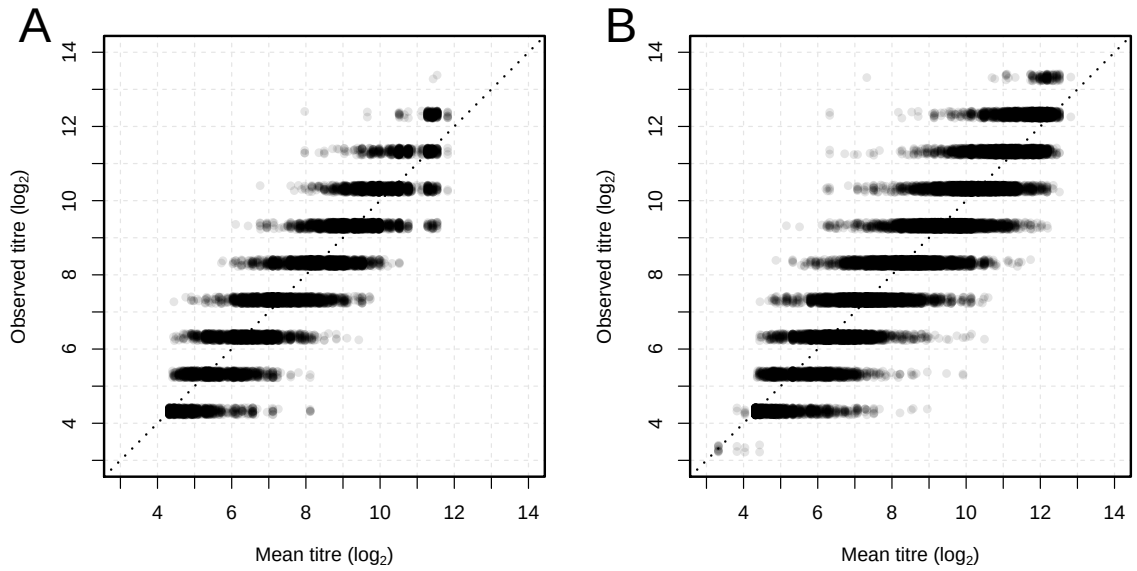

Figure 2.1: Variation in HI titres: Mean and observed $\log_2$ titres are shown for each virus-antiserum pair with at least two HI titres recorded. Discrete bands in observed titres correspond to two-fold dilutions in the HI assay and points have been randomly displaced by a small amount in this dimension to facilitate visualisation of overlapping points. 16,801 titres for 630 virus-antiserum pairs are shown for A(H1N1) (**A**) and 36,133 titres for 2,222 virus-antiserum pairs are shown for A(H3N2) (**B**). Black dotted lines with slope =1 and intercept = 0 are shown.

Notable levels of variability in the titres recorded per virus-antiserum pair are observed in both subtypes. In the A(H1N1) dataset, 16,801 titres were recorded for 630 virus-antiserum pairs with at least two titres. Observed titres could predict much of the variation in mean titres ($R^2 = 0.87$), however discrepancies between mean and observed $\log_2$ titres as great as 4.35 were observed; the mean, absolute error for a single reading (the difference between mean and observed $\log_2$ titre) was calculated as 0.55. This is roughly half of a two-fold dilution in HI assay and slightly more than double than the value of 0.25 which is the optimal expected error given the two-fold dilutions involved even when everything else occurs without error. The A(H3N2) HI dataset comprised 36,133 titres recorded for 2,222 virus-antiserum pairs with at least two titres. The relationship between mean and observed $\log_2$ titres was very similar to that recorded for A(H1N1) ($R^2 = 0.88$), however the greatest discrepancy was significantly larger (6.00) and the mean, absolute error was also slightly greater (0.59). Figure 2.1 indicates that in many instances, variation in titres measured for a given virus-antiserum pair are of a magnitude that might otherwise indicate antigenic dissimilarity. This demonstrates the necessity of experience when interpreting and recognising patterns among the results of many HI assays and shows why it is often difficult to track changes in antigenic phenotype back to specific genetic differences between viruses using HI data. Given the observed levels of variability in titres, comparisons of molecular and antigenic evolution were

made using mean $\log_2$ titres calculated for each virus-antiserum pair.

2.4.1 Phylogenetic analysis

For both A(H1N1) and A(H3N2), the GTR+I+Γ_4 model was determined to be the best-fit model of nucleotide substitution by jModelTest. This model, a general time-reversible model of nucleotide substitution with a proportion of invariant sites and a gamma distribution describing among-site rate variation with four categories estimated from the empirical data was tested against similar models in BEAST and was also supported by comparison of Bayes factors (Suchard *et al.*, 2001). Bayes factor analysis also determined that a codon-position model that allowed rates of nucleotide substitution to vary at the third codon position relative to the first and second codon positions should be used. For both subtypes, a relaxed (uncorrelated) molecular clock (Drummond *et al.*, 2006) with branch substitution rates drawn from an exponential distribution was determined to best represent the rate of evolution along branches of the phylogeny. For A(H1N1), the constant size population model outperformed other models. This is not surprising since, while influenza population sizes fluctuate regionally depending on season, global population size is not expected to have changed dramatically during the period studied. With higher-resolution data on the dates of isolation of viruses, the Bayesian skyline demographic model (Drummond *et al.*, 2005) may have been able to more accurately capture seasonal fluctuations in population size. For A(H3N2), Bayes factor analysis indicated support for both an assumption of constant population size and the Bayesian skyline model ahead of the various models of population growth tested, though these models were equally well supported (i.e. standard errors on marginal likelihoods of models were greater than differences in marginal likelihood of models). For computational reasons, the assumption of constant population size was used in subsequent analyses. The resulting maximum clade credibility phylogenetic trees generated for A(H1N1) and A(H3N2) are shown in Figures 2.2 and 2.3 respectively. Figure 2.3 shows separate trees for the full genetic sequence data set consisting of 1860 virus sequences and the subset of 229 sequences. Reference viruses used to raise antisera present in the HI dataset and viruses included in seasonal influenza vaccines are highlighted in each phylogenetic tree.

2.4.2 Comparison of antigenic and phylogenetic data

The relationship between molecular and antigenic evolution was investigated by testing amino acid and phylogenetic distances between pairs of viruses as predictors of HI titres. Amino acid distances between all viruses in the HI datasets were calculated from translated HA1 nucleotide sequences. Phylogenetic distances were calculated as the sum of branch lengths, measured in years, separating virus pairs on the HA1 phylogeny. Pairwise amino acid and phylogenetic distances are plotted against mean $\log_2$ titres for each virus-antiserum pair in

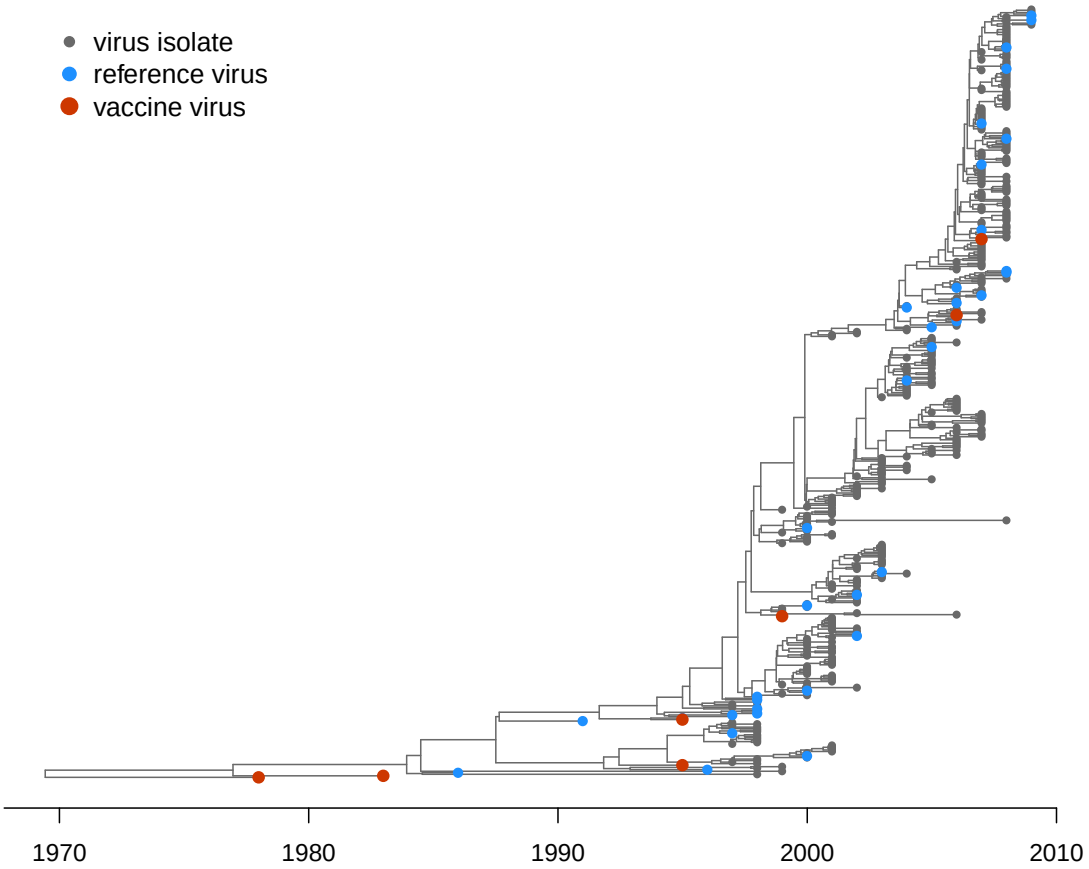

Figure 2.2: A(H1N1) phylogenetic tree: The maximum clade credibility, time-resolved HA1 phylogenetic tree for 506 A(H1N1) viruses with associated HI data. Reference viruses used to raise antisera and vaccine components are coloured blue and red respectively. Branch lengths are measured in evolutionary time in terms of years rather than substitutions.

the A(H1N1) and A(H3N2) datasets in Figures 2.4 and 2.5. These plots suggest that the antigenic similarity of strains is related to their genetic similarity but that the relationship is not adequate for sequence-based inference of antigenic phenotype. This likely represents both heterogeneity in the antigenic impact of genetic changes and the influence of other sources of variation on mean HI titres.

For A(H1N1), 183 amino acid positions in HA1 were conserved while 144 were variable. The largest amino acid distance between any pair of viruses in the HA1 alignment was 38 but among the virus-antiserum pairs in the HI data the highest distance was 26. A linear model determined that on average each amino acid difference between viruses resulted in a drop in $\log_2$ titre of 0.23 ($p < 10^{-10}$), though amino acid distance was a relatively poor predictor of HI titre ($R^2 = 0.31$). Phylogenetic distance, measured in evolutionary time, was an even weaker predictor of titre ($R^2 = 0.17$), though a significant relationship was seen, with each year of evolutionary time resulting in a drop in $\log_2$ titre of 0.14 ($p <$

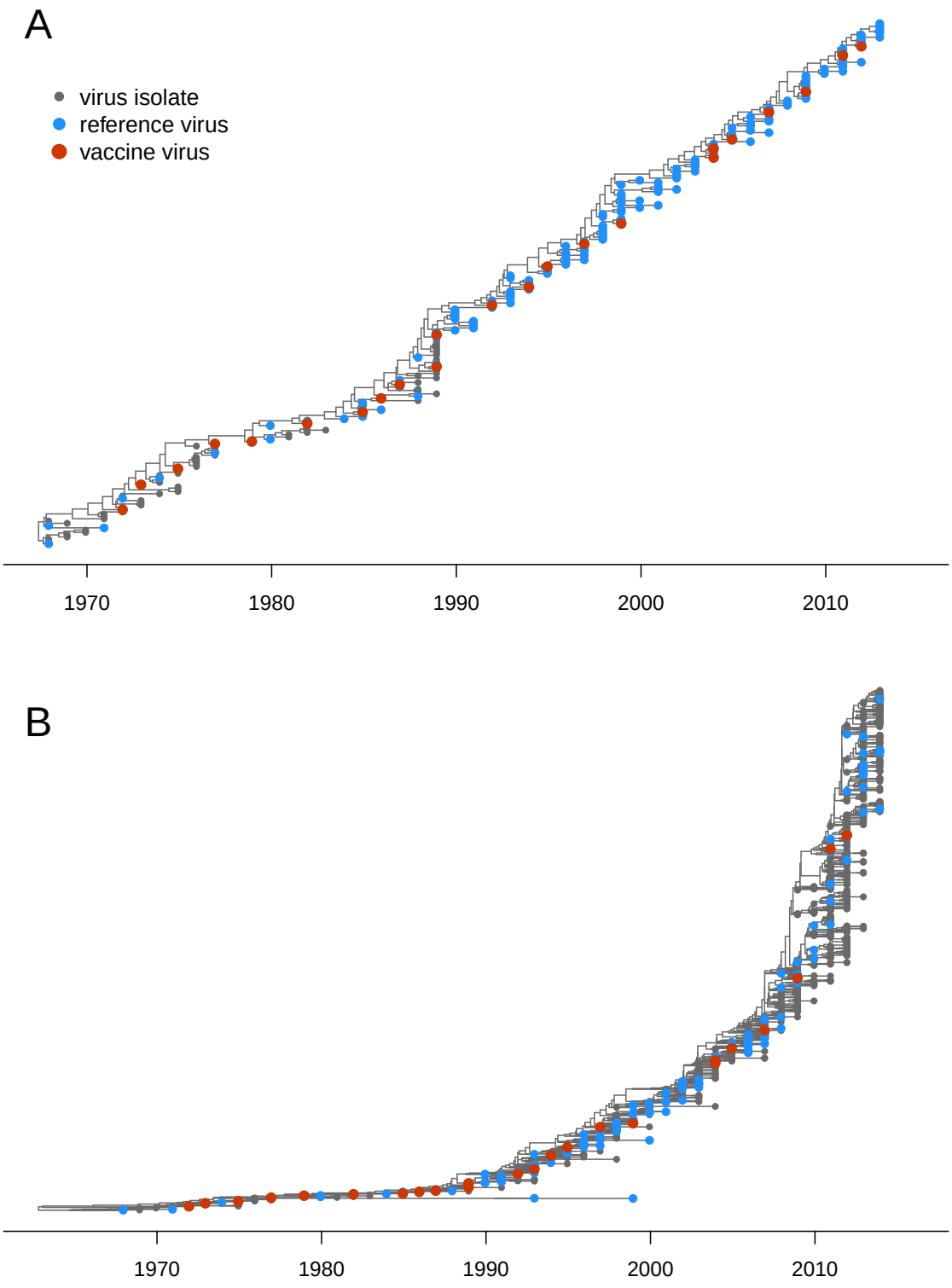

Figure 2.3: A(H3N2) phylogenetic trees: The maximum clade credibility, time-resolved HA1 phylogenetic tree for (**A**) 229 and (**B**) 1860 A(H3N2) viruses with associated HI data. Reference viruses used to raise antisera and vaccine components are coloured blue and red respectively. Branch lengths are measured in evolutionary time in terms of years rather than substitutions.

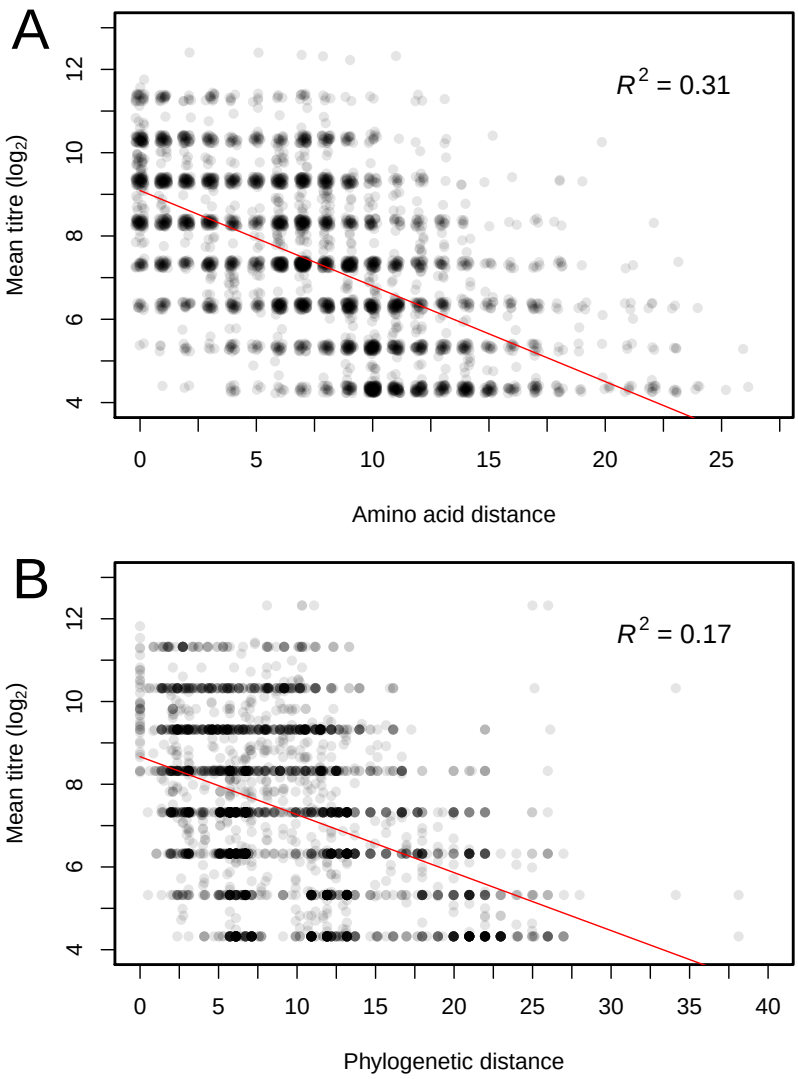

Figure 2.4: A(H1N1) pairwise amino acid and phylogenetic distances as predictors of HI titres: The mean log_2 HI titre for each virus-antiserum pair (n = 3,734) plotted against (**A**) the pairwise amino acid distance, the number of differences in the HA1 amino acid sequences of test and reference virus and (**B**) the pairwise phylogenetic distance, the sum of branch lengths separating two viruses in a time-resolved HA1 phylogenetic tree. Given the discrete nature of amino acid distances, a small amount of random displacement in both axes has been added to the position of points in (**A**) to facilitate visualisation of overlapping points. Red lines show linear model fits and R^2 values are noted.

10^{-10}). A greater proportion of HA1 amino acid positions were variable among the viruses in the extended A(H3N2) alignment (233 out of 328 HA1 positions). The largest amino acid distance between viruses in the A(H3N2) alignment was 70 while 62 was the greatest distance between a pair of viruses with associated HI data. On average each amino acid difference resulted in a smaller drop in $\log_2$ titre (0.14, $p < 10^{-10}$), relative to the A(H1N1) dataset, and amino acid distance explained an even smaller proportion of variation in titre ($R^2 = 0.18$). For A(H3N2), phylogenetic distance was a comparable predictor of $\log_2$ titre ($R^2 = 0.18$) but was still a statistically significant correlate with each year of evolutionary time resulting in an average drop in $\log_2$ titre of 0.26. Thus, genetic distances provide some power for estimating antigenic similarity of influenza A(H1N1) and A(H3N2) viruses, however the low coefficients of determination (R^2) indicate that simple measures of genetic distance are inadequate predictors.

A visual representation of the relationship between between molecular and antigenic evolution for A(H1N1) viruses is shown in Figure 2.6. Here, viruses and antisera are sorted phylogenetically according to the maximum clade credibility trees generated in BEAST and coloured cells indicate average HI titres for pairs of virus and antiserum tested. It is clear from this figure that, generally, viruses that are phylogenetically related are also antigenically similar, however there are instances where phylogenetically similar viruses are antigenically distinct. For example, the starkest antigenic change in the A(H1N1) HI dataset is represented as the red to yellow change in the columns for the antisera raised against reference viruses A/Johannesburg/82/96 and A/Bayern/7/95 to the left of the heat-map in Figure 2.6. The colour change in these columns of the heat-map corresponds to a relatively deep bifurcation in the HA1 phylogeny to the left of Figure 2.6, where one or more changes in amino acid must have caused a disproportionally large change in antigenic structure. This further indicates the level of heterogeneity in the antigenic impact of genetic changes.

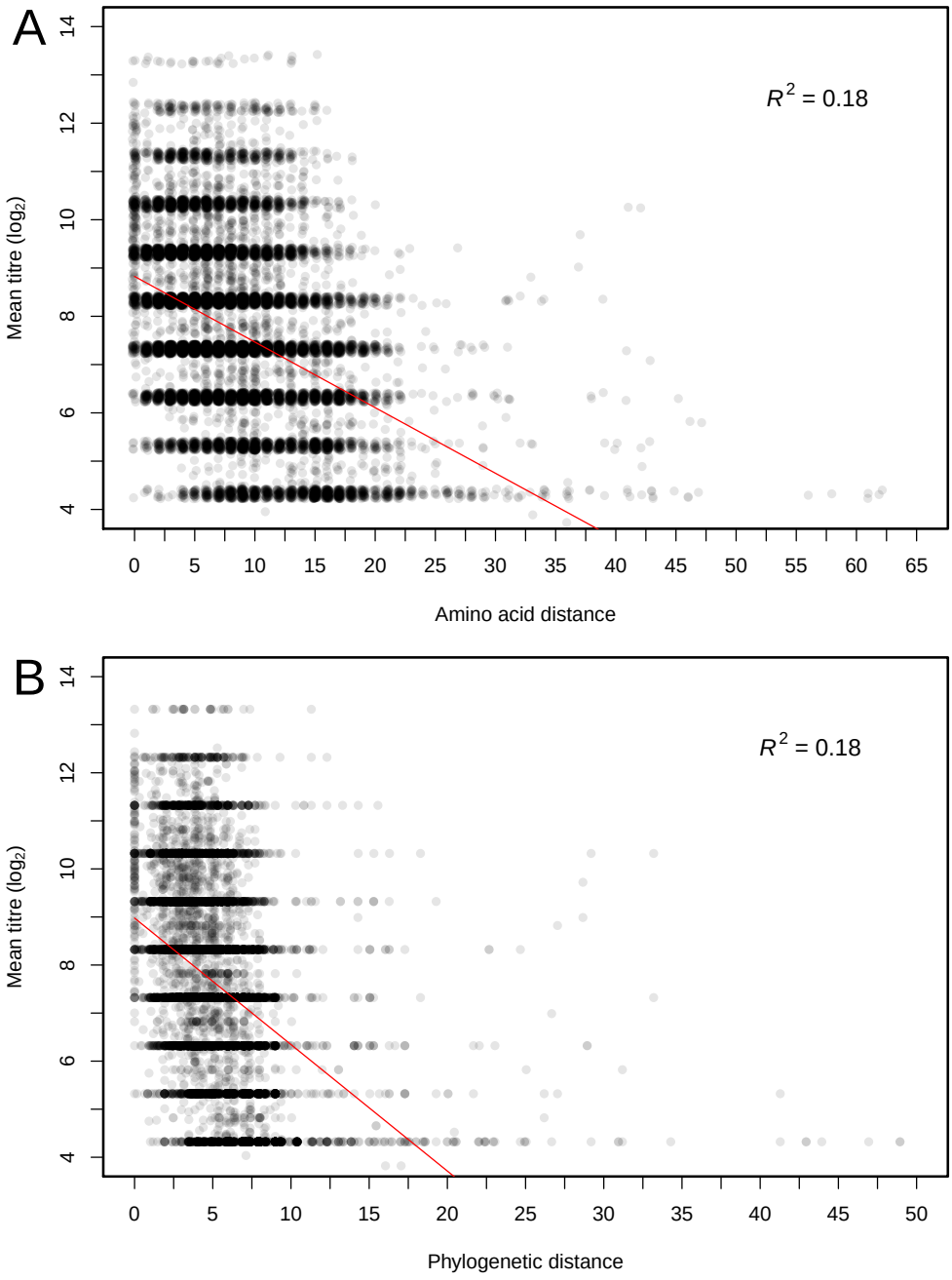

Figure 2.5: A(H3N2) pairwise amino acid and phylogenetic distances as predictors of HI titres: The mean log_2 HI titre for each virus-antiserum pair (n = 15,855) plotted against (**A**) the pairwise amino acid distance, the number of differences in the HA1 amino acid sequences of test and reference virus and (**B**) the pairwise phylogenetic distance, the sum of branch lengths separating two viruses in a time-resolved HA1 phylogenetic tree. Given the discrete nature of amino acid distances, a small amount of random displacement in both axes has been added to the position of points in (**A**) to facilitate visualisation of overlapping points. Red lines show linear model fits and R^2 values are noted.

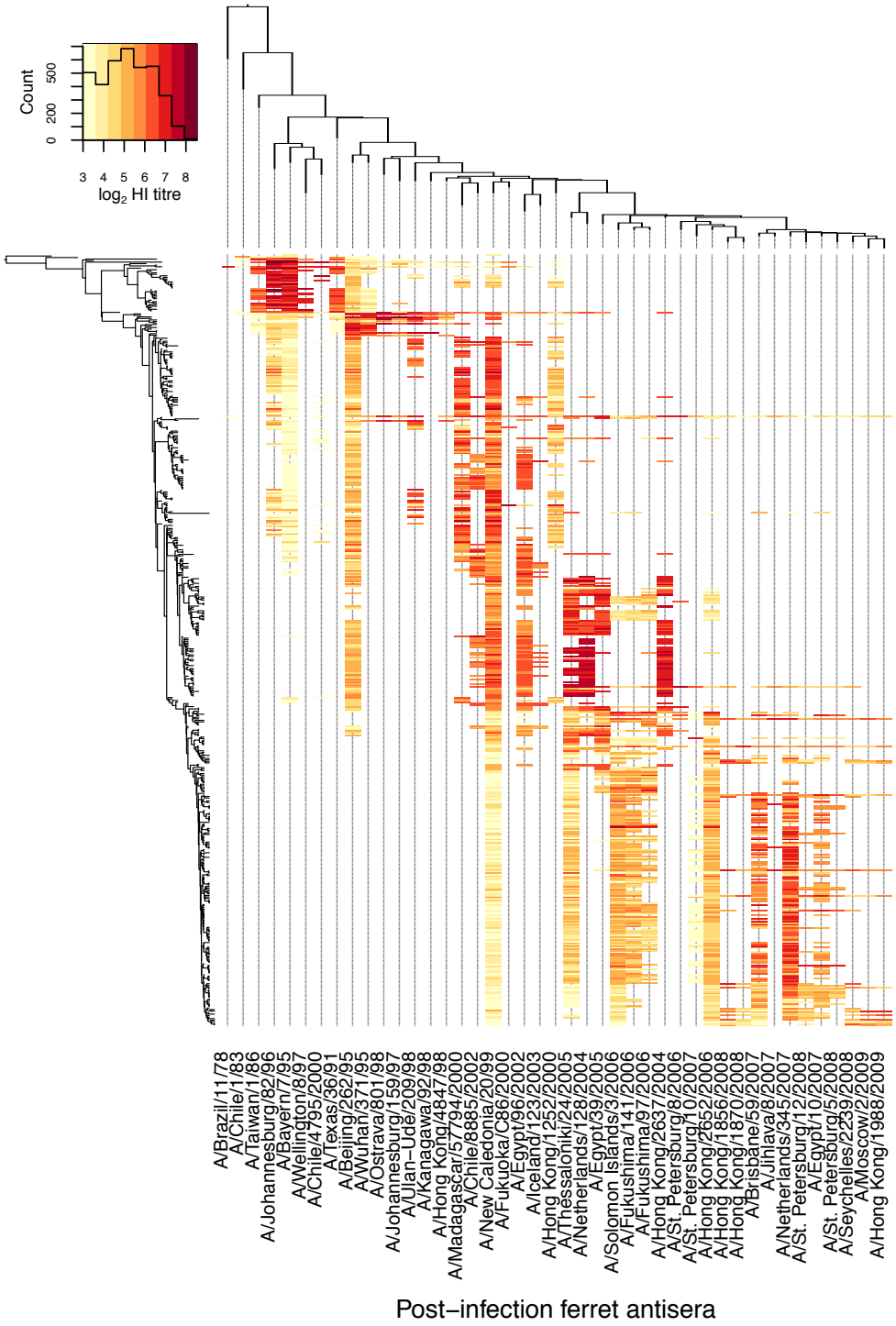

Figure 2.6: Heat-map of A(H1N1) HI titres sorted phylogenetically according to HA1 gene: Cells are coloured by mean $\log_2$ HI titre for each pairing of antiserum and test virus present in the full dataset. The colour key for HI titres is shown in the histogram at top left along with the number of assays yielding each titre. Test viruses and reference viruses used to generate post-infection ferret antisera are sorted phylogenetically on the HA1 gene along the vertical and horizontal axes respectively. Phylogenies are shown to the left for test viruses and above for reference viruses.

2.4.3 Signatures of selection

Four codon-based maximum likelihood methods, FEL, IFEL, MEME and FUBAR, were used to estimate the dN/dS ratio (also known as the K_a/K_s or ω ratio) at each codon in alignments of A(H1N1) and A(H3N2) HA1 sequences. Retaining the topology of the time-resolved, maximum clade credibility phylogenetic trees generated using BEAST, codon-specific rates of synonymous and non-synonymous mutations were estimated. The FEL and FUBAR methods were used to estimate codon-specific dN/dS ratios across the full HA1 phylogenetic trees. The IFEL method was used to identify codons subject to selective pressures at the population level (i.e. along internal branches only). The MEME method was used to estimate codon-specific dN/dS ratios that may vary across branches of the tree, potentially identifying codons exposed to episodic positive selection. These analyses revealed the proportion of HA1 codons under purifying selection to be notably higher than the proportion under positive selection in both subtypes (Table 2.1).

Table 2.1: Proportion of codons across HA1, and in antigenic sites, under positive and purifying selection

Subtype	Method	Positive selection		Purifying selection	
		HA1	Antigenic sites	HA1	Antigenic sites
A(H1N1)	FEL	0.01(2)	0.02(1)	0.35(116)	0.20(10)
	IFEL	0.01(3)	0.02(1)	0.18(59)	0.08(4)
	MEME	0.01(4)	0.04(2)	-	-
	FUBAR	0.01(4)	0.06(3)	0.45(146)	0.28(14)
A(H3N2)	FEL	0.04(13)	0.17(10)	0.24(80)	0.03(2)
	IFEL	0.03(9)	0.13(8)	0.12(40)	0.02(1)
	MEME	0.04(14)	0.18(11)	-	-
	FUBAR	0.03(9)	0.12(7)	0.27(87)	0.02(1)

Proportion (and count in parentheses) of codons with dN/dS ratios significantly greater than 1 (p-value < 0.05 using FEL/IFEL/MEME or posterior probability > 0.95 using FUBAR).

The proportions of codons coding for amino acid positions in defined antigenic sites detected as being exposed to positive or purifying selection were also calculated and these proportions were compared with HA1-wide proportions. The proportions of codons associated with antigenic sites in A(H1N1), identified by the various methods, were marginally higher than the HA1-wide proportions (Table 2.1). In A(H3N2), the differences in observed proportions were starker, with more codons in antigenic sites exposed to positive selection and fewer exposed to purifying selection, relative to the proportions calculated across the entire HA1. This possibly indicates a more significant role of immune-mediated evolution in A(H3N2), relative to A(H1N1).

Surface-exposed amino acid positions identified by any of the four methods as being exposed

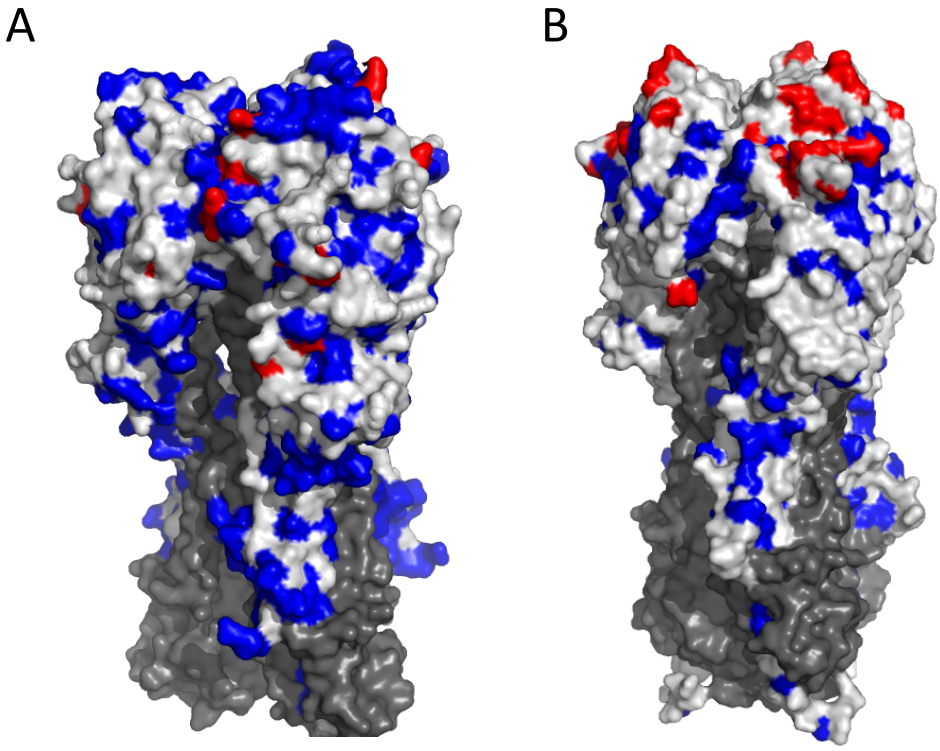

Figure 2.7: HA1 selection surfaces: Surface representations of HA of (**A**) A/Puerto Rico/8/34 (H1N1) (PDB ID: 1RU7 (Gamblin *et al.*, 2004)) and (**B**) A/Aichi/2/68 (H3N2) (PDB ID: 1HGG (Sauter *et al.*, 1992)) with positions under positive and negative selection shown. HA1 is shown in white and HA2 in grey. Amino acid positions with codon-specific dN/dS ratios indicating significant signatures of positive or negative selection as detected by at least one method are coloured red or blue respectively ($p < 0.05$ using FEL/IFEL/MEME or posterior probability > 0.95 using FUBAR).

to either positive or purifying selection are shown on the HA structure, in red and blue respectively, in Figure 2.7. No codon was identified as showing a significant signature of both positive and purifying selection using different methods. Strikingly, a large proportion of the positions identified as being under positive selection are close to the top globular head of the A(H3N2) HA. This is consistent with expectations based on knowledge of the locations of neutralising antibody binding sites (Wiley & Skehel, 1987) and the importance of substitutions at positions in the periphery of the binding site in antigenic evolution of this subtype (Koel *et al.*, 2013). Contrastingly, positively selected positions in A(H1N1) are more sparse and regions analogous to those positively selected at the head of A(H3N2) HA are, in many cases, associated with purifying selection. Root mean square deviation (RMSD) was used to evaluate the similarity of the HA1 structures of A/Puerto Rico/8/34 and A/Aichi/2/68. The RMSD weighted over 315 aligned HA1 amino acids was calculated, using the software SAP (Taylor, 1999, 2000), to be 2.69 Å.

HA1 codons in each subtype showing evidence of positive selection are shown in Table 2.2.

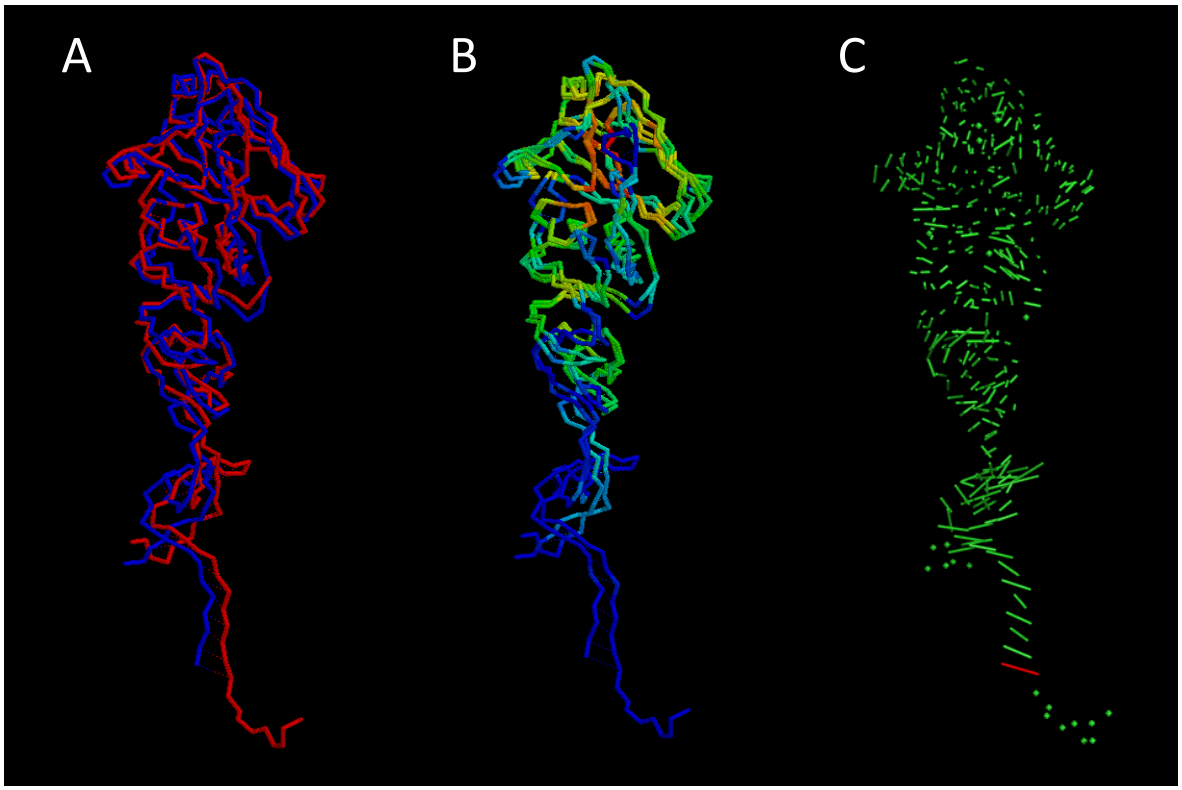

Figure 2.8: Comparison of the HA1 structures of A/Puerto Rico/8/34 (H1N1) and A/Aichi/2/68 (H3N2): In (**A**) and (**B**) aligned HA1 peptide backbone chains are shown as solid lines linking α carbons. In (**A**) A/Puerto Rico/8/34 (PDB ID: 1RU7) is coloured blue and A/Aichi/2/68 (PDB ID: 1HGG) is coloured red. In (**B**) chains are coloured according to the proximity of aligned α carbons using the temperature colouring option in RasMol v 2.7.5.2. which gives the sequence blue, cyan, green, yellow, orange, and red, where blue indicates greatest distance between α carbons, and red indicates least. (**C**) shows a visualisation of the distances between aligned α carbons. For reference, the red line at the base of the structure represents a distance of 8.21 Å. For instances where an α carbon was not aligned (i.e. at the base of the structure), no line is drawn and instead a dot is drawn at the position of the residue for which there are coordinates.

There is a high degree of correspondence between methods. All codons identified using either the FEL or FUBAR methods were identified using both methods, though the level of confidence in the identification of some codons did differ between methods. A(H1N1) codons 141, 186 and 222 and A(H3N2) codons 138 and 186 were identified using the FEL and FUBAR methods which analyse mutation rates in all branches of the phylogenetic tree, but not by the IFEL method which estimates dN and dS rates using only mutations inferred to have occured in internal branches of the tree. Evidence of positive selection at these five sites is therefore dependent on an overabundance of non-synonymous mutations occurring in terminal branches. This suggests that amino acid substitutions contributing to a signature of positive selection are transient, are removed from the population, and do not contribute to the future evolution of the virus. It can therefore be inferred that there is no evidence for the

role of these five positions in the adaptive evolution of either subtype at a population level.

Table 2.2: HA1 codons showing evidence of positive selection

Subtype	Codon	Method				Antigenic site/ receptor binding
		FEL	IFEL	MEME	FUBAR	
A(H1N1)	57	✓	✓	✓	✓	
	87			✓		
	94	∗	✓	∗		binding
	141	∗		∗	✓	Ca
	170			✓		Ca
	186	∗			✓	Sb
	189		✓			Sb
	222	✓		✓	✓	Ca/binding
A(H3N2)	128	✓	∗	✓	✓	
	133			✓		A
	135	✓	✓	✓	∗	A/binding
	137	✓	✓	✓	✓	A/binding
	138	✓		✓	∗	binding
	142	✓	∗	∗	∗	A
	145	✓	✓	✓	✓	A
	155			✓		B/binding
	157	✓	∗	✓	✓	B
	158	∗	✓	∗	∗	B
	159	✓	✓	✓	✓	B
	186	✓		✓	✓	B
	193	✓	✓	✓	✓	B
	194	✓	✓	✓	✓	binding
	226	✓	✓	✓	✓	D/binding
	262	✓	✓	✓	∗	E

✓Codon-specific dN/dS ratio significantly greater than 1 (p-value < 0.05 using FEL/IFEL/MEME or posterior probability > 0.95 using FUBAR).
∗ Positive selection identified with reduced certainty: p-value < 0.1 (FEL/IFEL/MEME) or posterior probability > 0.9 (FUBAR).

MEME, an extension of the FEL methodology that allows dN/dS to vary between sites and between lineages, was used to detect codons with dN/dS ratios indicative of positive selection in some part(s) of the phylogenetic tree but not others. Theoretically, it is possible that a codon involved in antigenic evolution could be subject to positive selection in one area of the phylogeny and then subject to purifying selection in the descendent part of the phylogeny. A particular amino acid substitution at a position within an epitope might provide a significant selective advantage and this codon may therefore have a dN/dS ratio indicative of positive selection and the substitution would be expected to increase in frequency within the influenza population. For a period, subsequent amino acid substitutions, either reversions of the positively selected substitution or alternative substitutions away from the antigenically distinct state towards other amino acids more easily bound by existing antibodies in the

human population, may be selected against. As A(H1N1) codons 87 and 170 and A(H3N2) codons 133 and 155 were identified using the MEME method but not by other methods, it can be inferred that these positions may have experienced episodic positive selection. For A(H1N1), each method of analysis was repeated with codon 187 included and in each case a highly significant signature of positive selection was detected with the highest dN/dS ratio of any codon in either subtype which is unsurprising given the tendency for amino acid substitutions to occur at this position as an adaptation to growth in eggs and cell culture.

Table 2.2 shows that several codons showing evidence of positive selection are known to code for amino acid residues belonging to antigenic sites. It is notable however that these codons are not evenly distributed across antigenic sites. In A(H1N1) positively selected codons were associated with antigenic sites Ca and Sb but not Cb or Sa, while in A(H3N2) antigenic sites A and B are represented in greater numbers than sites C–E. Evidence of positive selection was also identified at a number of codons associated with a role in receptor-binding which is consistent with observations of the importance of receptor binding in the adaptive evolution of influenza viruses (Hensley *et al.*, 2009; Lin *et al.*, 2012).

2.4.4 Antigenic cartography

Viruses and antisera of each subtype were positioned in two-dimensional antigenic maps using a BMDS model (Bedford *et al.*, 2014) that estimates reference virus immunogenicity and test virus receptor-binding avidity while also incorporating information on the year of virus isolation and the phylogenetic relatedness of strains. A(H1N1) and A(H3N2) viruses positioned in antigenic maps are shown in Figure 2.9. Viruses are represented on these maps as coloured circles coloured by year of isolation. To aid the interpretability of the maps a prior was used resulting in most variation being expressed in the primary antigenic dimension (shown in the x-axis in both maps in Figure 2.9), with a smaller amount of variation being contained in the secondary antigenic dimension. Thus the majority of antigenic drift through time is shown in the x-axis of the maps in Figure 2.9 while antigenic variation in the y-axis principally represents antigenic differences between lineages that emerge and transiently coexist.

The rate of antigenic drift, or change of location in the primary antigenic dimension, was calculated for each subtype. Viruses of the A(H1N1) subtype evolved at a mean rate of 0.51 antigenic units per year (95% HPD: 0.48–0.54), slower than the rate of 0.60 estimated for A(H3N2) (95% HPD: 0.55-0.65). This observation is consistent with the significantly slower rate of vaccine updates in A(H1N1) and also the application of this Bayesian multidimensional scaling analysis to other A(H1N1) and A(H3N2) HI datasets though the disparity in rates observed here is less stark (Bedford *et al.*, 2010). Both of these estimates of antigenic change per year are significantly lower than the figures reported above when estimating the drop in

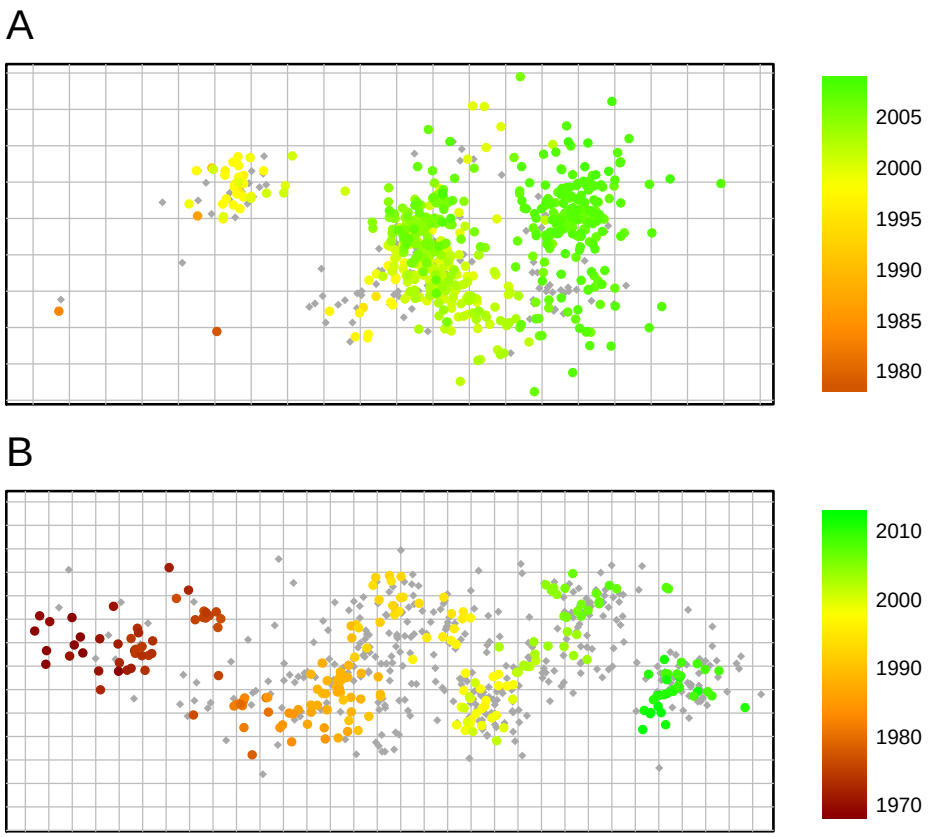

Figure 2.9: A(H1N1) and A(H3N2) antigenic maps: For each subtype, map locations are shown for a representative example from a Bayesian multidimensional scaling model that estimates virus location, antiserum location, reference virus immunogenicity and test virus receptor-binding avidity. (**A**) 506 A(H1N1) and (**B**) 229 A(H3N2) viruses are shown as circles coloured by date of isolation according to the legends at right and antisera are shown as grey diamonds. Primary and secondary antigenic dimensions are shown in the horizontal and vertical dimensions respectively. Gridlines represent single antigenic units, twofold dilutions in the HI assay.

$\log_2$ HI titre per unit of evolutionary time separating viruses in phylogenetic trees generated using HA1 (0.14 and 0.26 $\log_2$ titre per year for A(H1N1) and A(H3N2) respectively). These figures are not however contradictory as on the map antigenic change is averaged across the number of years in the period of data collection, whereas using the phylogenetic method antigenic change is averaged across evolutionary time which is summed over all branches of the phylogeny, a figure which may be significantly higher. Using either method, antigenic change per unit of time in A(H3N2) is estimated to be greater than in A(H1N1).

Distinct clusters of antigenically similar viruses, indicating punctuated antigenic evolution,

are apparent in the maps constructed for both subtypes (Figure 2.9). k-medoids clustering was used to define clusters of antigenically similar viruses based on locations estimated using BMDS. For both subtypes, antigenic clusters are shown in Figure 2.10. Clusters are named after the first vaccine virus they contain with letters and numbers indicating the location and year of virus isolation respectively. The 506 A(H1N1) viruses were identified as belonging to three distinct clusters of antigenically similar viruses, shown in red, green and blue in Figure 2.10. Of the seven vaccine viruses in the A(H1N1) dataset (red circles in the phylogenetic tree in Figure 2.2), the three earliest (A/Brazil/11/78, A/Chile/1/83 and A/Bayern/7/95) were assigned to the red cluster, though A/Brazil/11/78 and A/Chile/1/83 are the two obvious outliers, antigenically distant from all other viruses, to the lower left of the map (their colour in Figure 2.9 indicates their age relative to other viruses in the dataset). It is anticipated that the addition on higher numbers of earlier viruses would result in aditional antigenic clusters to which these two older vaccine viruses would be associated. The next two vaccine viruses occurring in the evolution of the virus (A/Beijing/262/95 and A/New Caledonia/20/99) were assigned to the central green cluster and the two most recent vaccine viruses (A/Solomon Islands/3/2006 and A/Brisbane/59/2007) were assigned to the cluster shown in blue. It is notable that, even after discounting some of the more obvious outliers at the fringes of clusters, distances between some viruses belonging to the same cluster exceed distances between some virus belonging to different clusters.

k-medoids clustering was used to define 11 clusters of antigenically similar A(H3N2) viruses isolated during the period 1968–2013 (Figure 2.10). Smith *et al.* (2004) identified 11 clusters over the shorter period 1968–2003 using a k-means clustering analysis, though the method for determining k in that study is not fully described. $k = 10$ was used, and subsequently an eleventh cluster was defined based on *a priori* knowledge of a genetic and antigenic variant (A/Texas/1/77). The identification of fewer clusters here is however consistent with observations that the number of antigenic clusters identified is reduced by accounting for differences in viral receptor-binding avidity (Li *et al.*, 2013) and by the inclusion of greater numbers of viruses (Sun *et al.*, 2013).

From left to right on the map in Figure 2.10, similar groupings of virus to clusters named by Smith *et al.* (2004) after A/Hong Kong/1/68 (HK68, green), A/England/42/1972 (EN72, orange) and A/Victoria/3/75 (VI75, magenta) are retained. The A/Texas/1/77 cluster defined manually post-analysis by Smith *et al.* (2004) is clustered together with the A/Bangkok/1/79 cluster (TX77, light blue). Viruses belonging to Smith *et al.* (2004) clusters A/Sichuan/2/87 and A/Beijing/353/89 form the black cluster (SI87). Similar groupings of viruses to Smith *et al.* (2004) clusters A/Beijing/32/92 (BE92, red), A/Wuhan/359/95 (WU95, brown), A/Sydney/5/97 (SY97, grey) and A/Fujian/411/2002 (FU02, yellow) were also identified. The seven most recent vaccine viruses in the HI dataset investigated here post-date viruses included on the map constructed by Smith *et al.* (2004). Of these, A/Wellington/1/2004 was identified in the yellow cluster alongside A/Fujian/411/2002, while A/California/7/2004, A/Wisconsin/67/2005

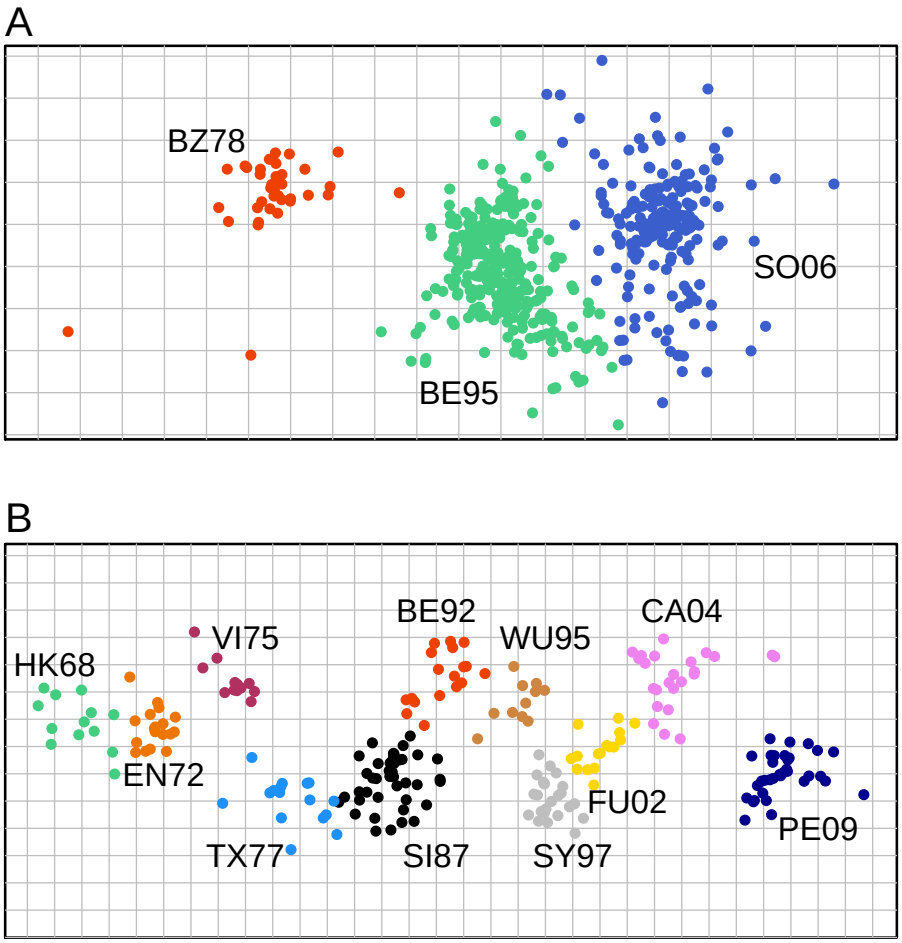

Figure 2.10: A(H1N1) and A(H3N2) antigenic clusters: A(H1N1) (**A**) and A(H3N2) (**B**) viruses positioned in antigenic space (as per Figure 2.9) represented by filled circles coloured by antigenic cluster. Clusters of antigenically similar viruses identified by k-medoids clustering analysis are named after the first vaccine virus in the cluster — letters refer to the location of isolation (Brazil, Beijing, and Solomon Islands for A(H1N1) and Hong Kong, England, Victoria, Texas, Sichuan, Beijing, Wuhan, Sydney, Fujian, California, and Perth for A(H3N2)) and numbers indicate year of isolation. Gridlines represent single antigenic units, twofold dilutions in the HI assay.

and A/Brisbane/10/2007 were among viruses forming the descendent antigenic cluster shown in pink (CA04) and the most recent antigenic cluster shown in dark blue (PE09) contained the vaccine viruses A/Perth/16/2009, A/Victoria/361/2011 and A/Texas/50/2012. Antigenic maps including these more recent viruses have shown similar clustering patterns (Bedford *et al.*, 2014).

Following Smith *et al.* (2004) and Koel *et al.* (2013), amino acid differences between ancestral and descendent antigenic clusters were identified in order to produce sets of candidate substitutions for each phenotypic change. On both antigenic maps distances between viruses

belonging to the same antigenic cluster often exceed distances between viruses in adjacent clusters. To reflect this, and the possible uncertainty in the assignment of some viruses to clusters that could arise due to this, amino acid differences between clusters were considered that were not completely correlated with the separation of viruses into adjacent clusters; 0.95 was used as a cutoff. The amino acid positions associated with transitions between antigenic clusters are detailed in Table 2.3.

Table 2.3: Amino acid substitutions associated with transitions between antigenic clusters

Subtype	Cluster transition*	Amino acid substitutions
A(H1N1)	BZ78–BE95	R43L, S69L, F71I, A80V, ΔK130, T133S, S183P, S271P, T310A
	BE95–SO06	K141E
A(H3N2)	HK68–EN72	T122N, G144D, T155Y, N188D, R207K, D275G
	EN72–VI75	L164Q, F174S, Q189K, R201K, I217V, I278S
	VI75–TX77	K50R, G158E, M260I
	TX77–SI87	Y155H, S159Y, K189R
	SI87–BE92	S133D, E190D
	BE92–WU95	I121T, G172D, R197Q, N262S, S278N
	WU95–SY97	K62E, K156Q, E158K, V196A, N276K
	SY97-FU02	
	FU02–CA04	K145N
	CA04–PE09	K158N, N189K

* Antigenic clusters identified by k-medoids clustering are shown Figure 2.10. Names refer to location and year of isolation.

Nine amino acid differences were identified as separating the ancestral BZ78 cluster in Figure 2.10 from the descended BE95 cluster shown in green: R43L, S69L, F71I, A80V, ΔK130 (a deletion of lysine at position 130), T133S, S183P, S271P, and T310A. Only A80V was totally correlated with assignment to red and green clusters. Of the positions at which these nine substitutions occur, only positions 69 and 71 belong to the defined H1 antigenic sites, both being constituents of the Cb antigenic site (Brownlee & Fodor, 2001). Despite this, ΔK130 has been previously identified as causing a large antigenic change in A(H1N1) subtype viruses and is associated with a change in phenotype from viruses antigenically similar to the vaccine virus A/Bayern/7/95 (BZ78 cluster in Figure 2.10) to viruses antigenically similar to the vaccine virus A/Beijing/262/95 (BE95 cluster in Figure 2.10) (McDonald *et al.*, 2007) (ΔK134 in cited work). K141E was identified as the substitution best correlated with assignment of viruses to either the green antigenic cluster or the descendent cluster shown in blue in Figure 2.10. Position 141 is a constituent of the Ca antigenic site and has previously been identified as a substitution separating clusters of antigenically distinct viruses on a map constructed from HI data collected using A(H1N1) viruses from the same period as those used in this

study (Koel *et al.*, 2013). Of the ten candidate positions identified using antigenic maps of A(H1N1) viruses (43, 69, 71, 80, 130, 133, 141, 183, 271, and 310), only position 141 was also associated with a signature of positive selection being identified using the FUBAR analysis and with less certainty by the FEL and MEME analyses (Table 2.2).

Position 130 has not been identified as belonging to one of the principal antigenic sites identified from monoclonal antibody escape mutant studies (Brownlee & Fodor, 2001) but presumably the deletion affected antibody binding either by inhibiting the binding of antibodies which were interacting with position 130 directly or by causing structural changes that affected antibody binding at nearby epitopes. The impact on structure of ΔK130 has been investigated by McDonald *et al.* (2007) by the prediction of molecular graphic images based on the A/Puerto Rico/8/34 (which bears a deletion of position 130) and A/swine/Iowa/30 (which bears K130) HA trimers (Gamblin *et al.*, 2004) using the Chimera package (Computer Graphics Laboratory, University of California, San Francisco, USA) (Huang *et al.*, 1996; Sanner *et al.*, 1996). K130 was predicted to lie along the edge of the receptor-binding pocket, potentially reducing the size of the pocket and impacting on the specificity or the avidity of receptor-binding. This analysis indicated that the addition of K130 into the A/Puerto Rico/8/34 HA would introduce a strong positive charge on the right side of the receptor-binding as well as slightly reducing negative charge at the bottom of the pocket (Figure 2.11). Position 130 was also predicted to have high solvent accessibility and therefore could play a role in binding of neutralising antibodies (McDonald *et al.*, 2007). The results of this structural analysis combined with those of reverse genetics experiments conducted by (McDonald *et al.*, 2007) indicate that K at position 130 acted as an immunodominant epitope and that ΔK130 resulted in escape from neutralising antibodies.

To reflect the increased uncertainty in the assignment of viruses to clusters upon the A(H3N2) antigenic map, the criteria for identifying amino acid substitutions separating viruses belonging to adjacent clusters was relaxed. If no substitutions were identified using the 0.95 cutoff, this was relaxed in intervals of 0.05 until either substitutions were identified or 0.80 was reached. In total, 33 substitutions were identified by this method. Substitutions also identified as causing transitions between antigenic clusters by reverse genetics experiments (Koel *et al.*, 2013) are highlighted in bold in Table 2.3. As many as six substitutions were found to be associated with transition between antigenic clusters, however for the SY97–FU02 transition no substitutions were identified. Of these 33 cluster-defining substitutions identified, eight (T155Y, Q189K, G158E, Y155H, S159Y, K189R, K156Q, and E158K) have been previously determined to cause transitions between clusters using reverse genetics methods to test all cluster-defining substitutions (Koel *et al.*, 2013). Koel *et al.* (2013) additionally identified K156E as causing a transition between Smith clusters A/Texas/1/77 and A/Bangkok/1/79, N145K as causing a transition to another cluster not identified here (A/Beijing/353/89), E156K as causing the SI87–BE92 transition, N145K as causing the BE92–WU95 transition, and Q156H as causing the SY97–FU02 transition.

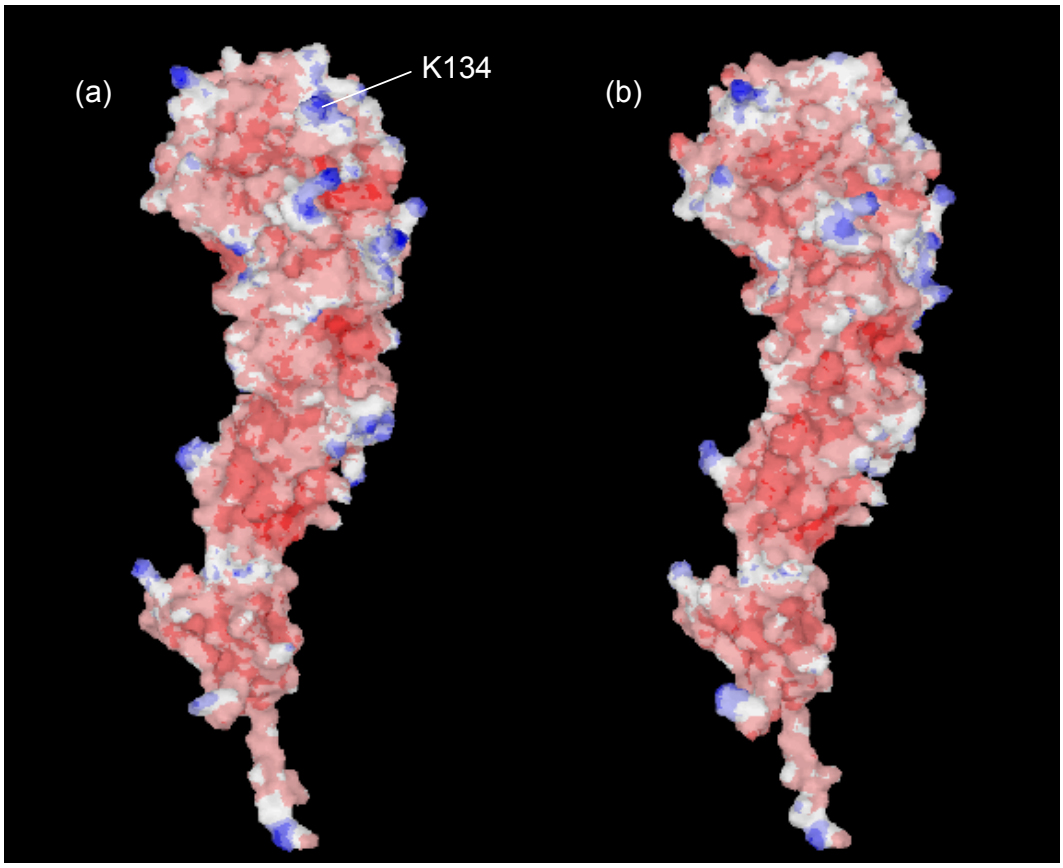

Figure 2.11: Predicted impact of ΔK130 on HA1 domain surface charge distribution: Positive and negative charges are indicated in blue and red, respectively. Neutral charges are depicted in white. (**a**) Predicted structure of A/Puerto Rico/8/34 with insertion of K130 (labelled as position 134 in original publication). (**b**) Structure of A/Puerto Rico/8/34 which bears a deletion of residue 130 (PDB ID: 1RU7 (Gamblin *et al.*, 2004)). Image reproduced as appears in Supplementary Figure S2 in McDonald *et al.* (2007) with permission from the rightsholder.

The two antigenic clusters coloured pink and dark blue to the right of Figure 2.10 post-date the viruses analysed by Koel *et al.* (2013). Clustering analysis of the virus locations estimated by BMDS identify the substitutions K145N as the candidate substitutions best correlated with the causing the transition from the FU02 to the CA04 cluster (though the substitutions Y159F and S189N were also identified when the cutoff was reduced to 0.80). The substitution N145K has previously been found to be associated with two cluster transitions in the antigenic evolution of A(H3N2) (Koel *et al.*, 2013). The CA04–PE09 transition correlated best with the substitutions K158N and N189K. Each of these three substitutions (and indeed Y159F and S189N) occur at positions within the A or B antigenic sites (Wiley *et al.*, 1981) that have been previously associated with transitions between antigenic clusters (Koel *et al.*, 2013) and therefore represent reasonable candidates. Multiple cluster-defining substitutions were identified at positions 155 (2), 158 (3), 189 (3) and 278 (2), therefore in total the 33 identified substitutions occurred at 27 positions. Each of these 27 positions are among the 60

positions assigned to antigenic sites by Shih *et al.* (2007), while 6 (0.18) were also identified as being positions exposed to positive selection by the analyses described above (positions 133, 145, 155, 158, 159 and 262). Seven of the 13 positions in antigenic sites identified by selection analyses were not identified as cluster-defining on the A(H3N2) antigenic map, potentially highlighting an antigenically adaptive role of substitutions causing smaller antigenic changes not detected using antigenic cartography.

2.5 Discussion

When investigating simple measures of pairwise genetic distance as correlates of the antigenic distance between viruses as reflected by HI titres, amino acid distance was more highly correlated with HI titres than phylogenetic distance for both serotypes. There are two potential explanations: Firstly, phylogenetic distances are influenced by synonymous mutations, which are antigenically irrelevant as they do not affect protein structure, while amino acid distances are not. Secondly, phylogenetic distances cannot reflect cases of convergent evolution or the reversal of amino acid substitution while amino acid distance can. When viruses in separate lineages of phylogeny converge on the same amino acid state at one or more positions, amino acid distance between them is reduced and intuitively these viruses are more likely to have become more antigenically similar than not, however the phylogenetic distance between them is increased by the mutations causing the convergence of amino acid states. Increased genetic distance quantified by each of these measures was significantly correlated with decreased antigenic cross-reactivity, as represented by lower HI titres. However, the low coefficients of determination associated with these correlations indicate that the relationship between molecular and antigenic evolution is more complex and that increased knowledge of the heterogeneities in the antigenic impact of genetic changes is required for accurate sequence-based prediction of antigenic similarity.

Various approaches for estimating codon specific dN/dS ratios were used to identify codons showing evidence of exposure to positive and purifying selection. In both subtypes a higher proportion of HA1 was identified as being exposed to purifying selection rather than positive selection. Despite a high degree of mutational tolerance of influenza A HA relative to proteins of some other viruses with similarly high mutation rates (Thyagarajan & Bloom, 2014), it is unsurprising that a high proportion of HA1 codons are resistant to amino acid substitution given the requirement for HA to retain the ability to bind host receptors and to initiate membrane fusion, two functions essential for viral replication (Russell *et al.*, 2004; Wiley & Skehel, 1987). The proportion of HA1 codons showing evidence of purifying selection, estimated using three methods, was noticeably higher in A(H1N1) (0.18–0.45) compared with A(H3N2) (0.12–0.27). This perhaps indicates an increased level of mutational tolerance in A(H3N2), relative to A(H1N1), which may be a factor contributing to the increased rate of antigenic

drift of A(H3N2) viruses identified by the BMDS analysis presented here and in previous studies (Bedford *et al.*, 2015, 2014). Several codons identified by these analyses belong to described antigenic sites suggesting that such methods have some power for identification of the amino acid positions at which substitutions are important in antigenic drift.

Antigenic characterisation of recombinant viruses generated by reverse genetics is a powerful method for determining the antigenic impact of specific amino acid substitutions, though in the absence of unlimited resources a method for identifying candidate substitutions is required. A BMDS model applied to HI data was used to estimate virus locations in antigenic space and a clustering analysis of these locations identified cluster-defining substitutions that could be considered as candidates for reverse genetics experiments. Koel *et al.* (2013) used reverse genetics to test 54 substitutions identified as defining antigenic clusters on maps estimated using a conceptually similar multi-dimensional scaling technique applied to a broadly similar HI dataset collected over a similar period. Their experiments identified 14 substitutions at seven positions that had caused the largest antigenic changes in the evolution of A(H3N2) over a 35 year period. However, this did require the definition of an antigenic cluster not identified by a clustering analysis of virus locations but instead based on *a priori* knowledge of a genetic and antigenic variant (A/Texas/1/77) and the 14 also included the substitution S159Y that was not identified as a cluster-defining candidate substitution in their analyses (though it was identified here — Table 2.3). There are several inconsistencies in the candidate substitutions identified here and by Koel *et al.* (2013), with only eight of 14 cluster transition substitutions being identified consistently across studies. This is due in part to the A/Texas/1/77 case however inconsistencies remain even after accounting for this. Another similar analysis of virus locations in antigenic space produced yet another set of candidate substitutions (Sun *et al.*, 2013).

This difficulty in replicating the same set of candidate substitutions raises the question of whether it would be preferable to test all substitutions estimated as having occurred in the trunk lineage of influenza using a phylogenetic analysis to reconstruct ancestral sequences. This approach would have resulted in 52 of the 54 substitutions Koel *et al.* (2013) identified as cluster-defining on A(H3N2) antigenic maps still being tested by reverse genetics. The substitutions N145K and N193S identified as defining the A/Beijing/353/89 cluster would not be identified as candidates as they lie outside the trunk lineage. The inability to identify potentially antigenically important substitutions outside the trunk lineage is a clear limitation of using this phylogenetic approach, however a strict implementation of the testing of cluster-defining substitutions would fail to identify candidate substitutions leading to and from the the A/Texas/1/77 antigenic cluster and the S159Y substitution, which each occurred in the trunk lineage.

Several amino acid positions with dN/dS ratios indicating exposure to positive selection were not identified as candidate substitutions using antigenic cartography. While it is possible that

these dN/dS ratios are false positives or that these codons have been selected to change for reasons other than antibody evasion, it is possible that this indicates a role in antigenic evolution for substitutions that do not result in major changes in antigenic evolution that are apparent on maps generated by applying BMDS to HI data. Previous work supporting the importance of smaller-impact amino acid substitutions in the evolution of influenza viruses include other studies (Bush *et al.*, 1999b) that also identify evidence of positive selection at many more codons than those identified by Koel *et al.* (2013) and the inclusion of a wider selection of codons in antigenic sites in models used to predict changes in frequencies of lineages between influenza seasons (Łuksza & Lässig, 2014). Much of the variation expressed on the antigenic maps is within-cluster rather than between-cluster, though it is not clear how to identify the amino acid substitutions responsible for this variation. Correlating amino acid substitutions against cartographic distance without distinct clusters indicating instances of phenotypic change is problematic given the shared evolutionary history of viruses which results in a strong correlation between genetic and antigenic change through time. If indeed non-cluster-defining substitutions are important in the antigenic evolution of influenza viruses, alternative methods for their identification are required.

CHAPTER 3

Statistical identification and experimental validation of genetic drivers of antigenic change in influenza A(H1N1)

Statistical identification and experimental validation of genetic drivers of antigenic change in influenza A(H1N1)

This chapter (together with some parts of Chapters 2 and 6) has been published as:

Harvey, W.T., Benton, D.J., Gregory, V., Hall, P.J., Daniels, R.S., Bedford, T., Haydon, D.T., Hay, A.J., McCauley, J.W. & Reeve, R. (2016) Identification of low- and high-impact hemagglutinin amino acid substitutions that drive antigenic drift of influenza A(H1N1) viruses. PLoS Pathogens. 12(4): e1005526. doi: 10.1371/journal.ppat.1005526.

A pre-print version of this work is available at arXiv.org:

Harvey, W.T., Gregory, V., Benton, D.J., Hall, P.J., Daniels, R.S., Bedford, T., Haydon, D.T., Hay, A.J., McCauley, J.W. & Reeve, R. (2014) Identifying the genetic basis of antigenic change in influenza A(H1N1). arXiv:1404.4197v2.

3.1 Abstract

Determining phenotype from genotype is a fundamental challenge in virus research, especially as sequence data become ever easier to generate. Identification of emerging antigenic variants among circulating influenza viruses is critical to the vaccine virus selection process, with vaccine effectiveness maximised when constituents are antigenically similar to circulating viruses. Haemagglutination inhibition (HI) assay data are commonly used to assess influenza antigenicity. Here, sequence and 3D structural information for haemagglutinin (HA) glycoproteins were analysed together with corresponding HI assay data for former seasonal influenza A(H1N1) virus isolates (1997–2009) and reference viruses. The models developed identify and quantify the impact of eighteen amino acid substitutions that affect the antigenicity of HA, two of which were responsible for major transitions in antigenic phenotype. Reverse genetics was used to demonstrate the causal effect on antigenicity for a subset of these substitutions including one instance where multiple contemporaneous substitutions made a definitive identification impossible from the evolutionary history. The ability to quantify the

phenotypic impact of specific amino acid substitutions increases the value of new HA gene sequence data for monitoring antigenic drift and phenotypic evolution and should help refine emerging techniques to predict the evolution of virus populations from one year to the next, leading to stronger theoretical foundations for selection of candidate vaccine viruses. The generality of an approach originally developed for foot-and-mouth disease virus is demonstrated, and this calls for extension to other antigenically variable pathogens.

3.2 Introduction

Evolution of human influenza A viruses is characterised by rapid antigenic drift, with structural changes in antigenic epitopes allowing the virus to escape existing immunity. Antigenic changes in circulating influenza viruses are principally assessed by the HI assay (Hirst, 1942; WHO, 2011). Results of many HI assays can be summarised using multidimensional scaling approaches, which approximate antigenic dissimilarity by Euclidean distances between viruses and antisera on a map, with antigenic evolution in influenza represented as movement between clusters of viruses (Smith *et al.*, 2004). The non-synonymous genetic mutation(s) causing transitions between antigenic clusters can be determined experimentally by reverse genetics (Koel *et al.*, 2013), though this approach is often laborious, as multiple amino acid substitutions bridge each antigenic cluster transition, and individual substitutions need to be assessed. This approach recently demonstrated that transitions between antigenic clusters of H3N2 viruses are caused predominantly by single amino acid substitutions at positions near the receptor-binding site (Koel *et al.*, 2013). However, clusters cannot always be readily identified by automated techniques (Chapter 2) and major cluster transitions may not be the only antigenically important events and an exhaustive reverse genetics analysis of all observed substitutions is not feasible due to high levels of amino acid sequence diversity in HA (e.g. at 46% of amino acid positions, in this study).

An alternative approach is to integrate matching sequence, antigenic and structural data into models that allow us to attribute the observed antigenic differences in a dataset directly to their underlying causes. Reeve *et al.* (2010) developed such a model to identify surface-exposed regions of the capsid proteins of foot-and-mouth disease virus where substitutions were correlated with antigenic change, but were unable to show definitive causal connection with specific substitutions. Various other computational approaches have similarly been used to identify antigenically important amino acid positions in influenza HA by comparison of predominant sequences of successive antigenic clusters and by comparing sequence and antigenic data (Huang *et al.*, 2012; Lee & Chen, 2004; Smith *et al.*, 2004; Steinbrück & McHardy, 2012; Sun *et al.*, 2013).

This chapter: 1. extends the modelling approach of Reeve *et al.* (2010) to former seasonal influenza A(H1N1) viruses, focusing on these first rather than A(H3N2) viruses, for which the

role of neuraminidase-mediated agglutination of erythrocytes has complicated the relationship between HI data and antigenic change (Lin *et al.*, 2010), or the distinct A(H1N1)pdm09 viruses, which have remained antigenically similar since emerging in humans in 2009 (Barr *et al.*, 2014); 2. attributes variation in HI titres to individual amino acid substitutions; 3. quantifies their antigenic impact; 4. validates the model experimentally by assessing the impact of a subset of the identified substitutions introduced by reverse genetics. This quantitative understanding of the phenotypic impact of specific HA substitutions improves our understanding of the antigenic evolution of the virus and should allow for more accurate predictions of the antigenic phenotype of emerging influenza viruses, measurement of which is critical to predicting the evolutionary success of newly emerging variants.

3.3 Materials and Methods

3.3.1 Data

The former seasonal influenza A(H1N1) dataset is introduced in detail in Chapter 2. These data were collected by staff of the of the Crick Worldwide Influenza Centre, UK (formerly the WHO Collaborating Centre for Reference and Research on Influenza, MRC National Institute for Medical Research, UK) over a number of years. Viruses were originally isolated from clinical specimens either by WHO National Influenza Centres (NICs) or by the Crick Worldwide Influenza Centre (formerly the WHO Collaborating Centre for Reference and Research on Influenza, MRC National Institute for Medical Research, UK). The antigenic dataset encompassed 506 A(H1N1) viruses for which HA gene sequence data were available, inclusive of 43 reference viruses against which post-infection ferret antisera were raised, with 19,905 HI titres measured between 3,734 unique combinations of virus and antiserum, made on 351 dates from 1997 to 2009.

3.3.2 Phylogenetic analysis

HA1 nucleotide sequences of all viruses were aligned using MUSCLE (Edgar, 2004). Phylogenies were estimated using a variety of nucleotide substitution and molecular clock models. A relaxed, uncorrelated clock and a GTR+I+Γ_4 nucleotide substitution model were determined to be most suitable through comparison of Bayes factors (Suchard *et al.*, 2001). Year of isolation was used to calibrate the molecular clock allowing rates of evolution along branches to be estimated. To incorporate phylogenetic error, a posterior sample of 9,000 trees (after removing 1,000 as burnin) was sampled across and the maximum clade credibility tree was selected. Substitution at HA position 187 is associated with adaptation to growth in eggs (Gambaryan *et al.*, 1999; Raymond *et al.*, 1986; Robertson *et al.*, 1987) and was therefore

considered to be an artefact with potential to distort phylogenetic inference, so nucleotides coding for position 187 were excluded from the phylogenetic analysis. Phylogeny construction and analysis was carried out using BEAST v1.7.4 (Drummond *et al.*, 2012) and Tracer v1.5 (Rambaut & Drummond, 2009). Ancestral amino acid state at each node in the phylogeny for each position identified by modelling was estimated using the FLU amino acid substitution model (Dang *et al.*, 2010) and unlinked strict molecular clocks for each amino acid position.

3.3.3 Mixed effects modelling and model selection

Co-variance between HI titre and the size of residuals from models using linear HI titres necessitated the use of logarithmically transformed HI titre, whose residuals were homoscedastic (i.e. their variance was equal across the observed range of HI titres), as the response variable. Base 2 was chosen (without loss of generality) for the logarithm to follow Smith *et al.* (2004) and work throughout was in terms of $\log_2$ titre (or antigenic units), so 1 antigenic unit corresponds to a two-fold dilution of antiserum in the HI assay. Goodness of fit of models including each of the following variables was assessed by likelihood ratio test: The reference virus against which the antiserum was raised, the test virus, and the date on which the assay was performed. Models were fitted using the package *lme4* (Bates *et al.*, 2012) in R (R Core Team, 2015). Parameters values for all terms were estimated using restricted maximum likelihood (REML).

Following Reeve *et al.* (2010), each branch of the HA1 phylogeny was tested in the model as a fixed effect term. Each branch term was included as a discrete indicator variable: 1 when reference virus and test virus were separated by the branch in the phylogenetic tree and 0 otherwise. Random restart hill-climbing (Russell & Norvig, 1995) was used to determine the best model as computation of all possible combinations of branch terms was not feasible. To a random consistent starting model, branch terms were added and removed stepwise at random to maximise model fit, assessed by AIC (Akaike, 1974). This was repeated while randomising their order to avoid sensitivity to the order in which the parameters are presented. This approach is conservative in that it is biased towards inclusion of parameters including possible false positives. This is desirable since it was used to determine the branches used to control for phylogenetic correlations in the data, and adding in extra unnecessary terms simply reduced the power of the analysis. In an ongoing collaboration with Vinny Davies, Richard Reeve and Dirk Husmeier (all University of Glasgow) we show that sparse Bayesian variable selection methods are more powerful for such problems, but these are not currently computationally feasible for use on such large datasets (Davies *et al.*, 2014). Conceptually similar machine learning techniques have been used on related influenza datasets (Sun *et al.*, 2013), but these did not control for phylogenetic correlation, which we consider to be critical to avoid false positives, and this remains the computationally expensive step.

PyMOL Molecular Graphics System v1.7.7.2 (http://www.pymol.org) was used to visualise and identify surface-exposed positions on the external surface of the HA 3D structure of A/Puerto Rico/8/34 (resolved to 2.3 Å, Protein Data Bank (PDB) ID: 1RU7) (Gamblin *et al.*, 2004), which had previously been identified as having significant solvent accessibility using naccess v2.1.1 (Hubbard & Thornton, 1993) (absolute solvent accessibility > 18 Å^2) . To account for potential structural changes in the HA1 of human A(H1N1) viruses since the isolation of A/Puerto Rico/8/34 in 1934, the lower resolution (3.19 Å) structure of HA of A/Solomon Islands/3/2006 (PDB ID: 3SM5) (Whittle *et al.*, 2011) was also used. The HA1 sequences of A/Puerto Rico/8/34 and A/Solomon Islands/3/2006 are identical at 86.1% of nucleotide sites. Root mean square deviation (RMSD) is often used to measure the difference between structures. The RMSD of the structures 1RU7 and 3SM5 weighted over 320 aligned HA1 positions was calculated, using the software SAP (Taylor, 1999, 2000), to be 1.35 Å. The superimposed structures of the A/Puerto Rico/8/34 and A/Solomon Islands/3/2006 HA1 peptide backbone chains are shown in Figure 3.1. For comparison, the RMSD of 1RU7 and the structure of A/Aichi/2/68 (H3N2) (PDB ID: 1HGG) (Sauter *et al.*, 1992) was calculated by the same method to be 2.69 Å (see Figure 2.8 for a visual comparison).

Amino acid positions determined to be surface-exposed according to the structure of either A/Puerto Rico/8/34 or A/Solomon Islands/3/2006 were classed as surface-exposed. Depending on the structure, 24-28 residues identified as being surface-exposed using naccess were identified as being located on the interior of the HA homotrimeric protein and were omitted from the analysis on the basis that they are not exposed on the true external surface. Amino acid dissimilarity between reference virus and test virus at each surface-exposed position that was not conserved within the dataset, was tested as a predictor of reduced HI titre ($p < 0.05$) using a Holm-Bonferroni correction to account for multiple tests (Holm, 1979). To test whether the omission of non-surface-exposed residues may have mistakingly resulted in false-negatives, substitutions at buried positions were separately tested as predictors of reduced titre. Although position 187 was excluded from phylogenetic inference for A(H1N1) (see Chapter 2), amino acid dissimilarity at this position was tested as a predictor of antigenic difference. At each HA position identified at this stage, the mean antigenic impact of specific amino acid substitutions was determined by examining the associated parameter (k_1, Equation 3.4) using data subsets consisting of viruses and reference viruses with one of the two amino acid variants being considered at that position.

3.3.4 Recombinant viruses

Viruses were generated using a protocol based on that described by Hoffmann *et al.* (2000). HA and neuraminidase cDNAs of A/Netherlands/1/93 (Neth93), which had been exclusively propagated in cell culture, were amplified using a standard RT-PCR protocol. These cDNAs were cloned into the pHW2000 vector (Lin *et al.*, 2010, 2012). Mutations were introduced

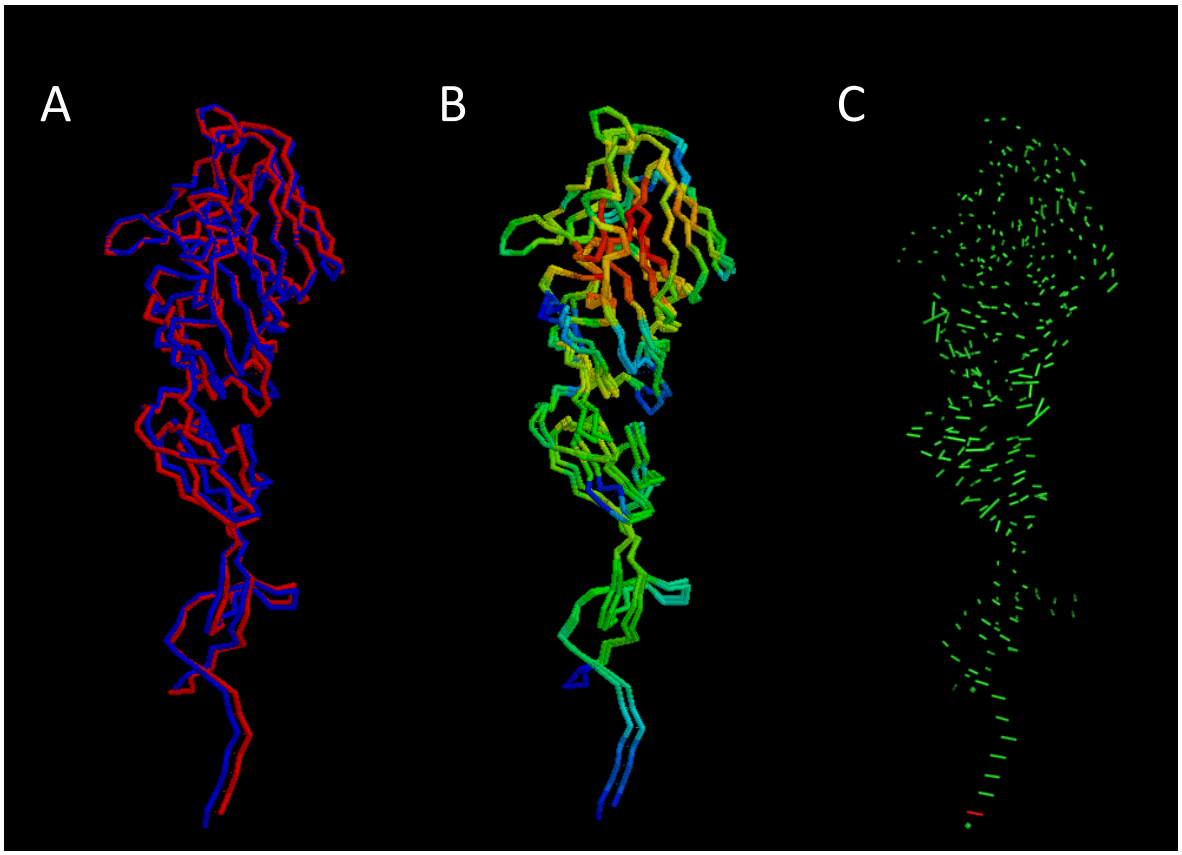

Figure 3.1: Comparison of the HA1 structures of A/Puerto Rico/8/34 and A/Solomon Islands/3/2006: In (**A**) and (**B**) aligned HA1 peptide backbone chains are shown as solid lines linking α carbons. In (**A**) A/Puerto Rico/8/34 (PDB ID: 1RU7) is coloured blue and A/Solomon Islands/3/2006 (PDB ID: 3SM5) is coloured red. In (**B**) chains are coloured according to the proximity of aligned α carbons using the temperature colouring option in RasMol v 2.7.5.2. which gives the sequence blue, cyan, green, yellow, orange, and red, where blue indicates greatest distance between α carbons, and red indicates least. (**C**) shows a visualisation of distances between aligned α carbons. For reference, the red line at the base of the structure represents a distance of 2.54 Å. For instances where an α carbon was not aligned (i.e. at the very base of the structure), no line is drawn and instead a dot is drawn at the position of the residue for which there are coordinates.

into the HA plasmid using the QuikChange lightning site-directed mutagenesis kit (Agilent Technologies). Co-cultured 293T and Madin-Darby canine kidney (MDCK) epithelial cells were co-transfected with plasmids containing HA and neuraminidase derived from Neth93 with the remaining six genes from A/Puerto Rico/8/34. MDCK cells were purchased from the American Type Culture Collection (ATCC) and maintained at Mill Hill laboratories. After 2–3 days, recombinant viruses in the supernatant of transfected cells were recovered and propagated in MDCK cells as described by Lin *et al.* (2012). Virus HA sequences were verified after passage. These methods were performed at containment level 2.

3.3.5 HI assays and analysis

HI assays were performed on recombinant viruses by standard methods (WHO, 2011). Antiserum samples were treated with receptor-destroying enzyme to eliminate non-specific inhibitors present in serum, inactivated at 56°C and then two-fold serially diluted in phosphate-buffered saline (PBS) from an initial dilution of 1:40. Virus samples were diluted in PBS to eight haemagglutinating units. Virus dilutions were then added to the serially diluted antiserum samples and left for 30 minutes to allow antibodies to bind. Suspensions of turkey (0.75%) erythrocytes in PBS were then added to antiserum:virus mixtures. Titres were recorded as reciprocals of the highest dilution of antisera that inhibited haemagglutination. The reciprocal of the highest two-fold dilution of post-infection ferret antiserum to completely inhibit haemagglutination (non-haemagglutinated blood drop reached the side of the well within 30 seconds when a plate was turned on its side) was recorded as the HI titre. Transfections were performed at containment level 2.

Post-infection ferret antisera raised against the following seven reference viruses were used to characterise recombinant viruses: A/Bayern/7/95, A/Johannesburg/82/96, A/Johannesburg /159/97, A/Ulan-Ude/209/98, A/Hong Kong/4847/98, A/New Caledonia/20/99 and A/Hong Kong/1252/2000. Antisera samples were obtained from archives of the WHO Collaborating Centre. Reference strains were chosen, in part, due to their phylogenetic proximity to the parent virus, Neth93. Where possible reference strains possessing and lacking the amino acid substitutions introduced by reverse genetics were chosen to allow changes in HI titres, associated with the introduction of substitutions, to be attributed to antigenic and non-antigenic effects of substitutions.

Average changes in $\log_2$ HI titre between parent and mutant recombinant viruses were quantified. These were partitioned into antigenic (ΔH_A) and non-antigenic (ΔH_N) effects using Equation 3.1:

$$\Delta H_A = \frac{\Delta H_2 - \Delta H_1}{2} \qquad \Delta H_N = \frac{\Delta H_1 + \Delta H_2}{2} \tag{3.1}$$

If the amino acid substitution introduced was antigenically important it was expected to cause a decrease in HI titre against parent-like virus derived antisera (ΔH_1) and a corresponding increase in HI titer against mutant-like virus derived antisera (ΔH_2). Conversely, a change in virus receptor-binding avidity is expected to cause a consistent decrease (or increase) in titre with these two groups of antisera (ΔH_1 and ΔH_2). Consequently, for each substitution the associated change in $\log_2$ HI titre, relative to the parental virus, was partitioned into antigenic (ΔH_A) and non-antigenic (ΔH_N) components. (ΔH_A) is estimated by halving the difference between (ΔH_1) and (ΔH_2) whereas (ΔH_N) is estimated by taking the mean of (ΔH_1) and (ΔH_2). Therefore if a substitution has an impact on phenotype which, on average, decreases

titres against parent-like antisera by 1 $\log_2$ titre ($\Delta H_1 = -1$) and increases titre against mutant-like antisera by 1 $\log_2$ titre ($\Delta H_2 = +1$), the substitution would be estimated to have an antigenic effect of 1 $\log_2$ titre ($(+1 - -1)/2 = 1$) and no non-antigenic effect on titre ($(+1 + -1)/2 = 0$). If a substitution impacts phenotype in such a way as to, on average, increase titres by 1 $\log_2$ titre against both parent- and mutant-like antisera ($\Delta H_1 = +1$ and $\Delta H_1 = +1$), the substitution would be estimated to have no antigenic effect ($(+1 - +1)/2 = 0$) and to increase titres by 1 $\log_2$ titre as a result of a non-antigenic effect ($(+1 + +1)/2 = 1$).

Antigenic effects were compared with predictions from modelling. Mean error in predictions across all substitutions was calculated as the average difference between the predicted mean and each measured antigenic change in HI using a specific virus dilution measured against a particular antiserum (excluding measurements restricted by the lower threshold of the HI assay). Small non-antigenic changes in HI titre (ΔH_N) between two viruses could be explained by the routine standardisation of both viruses by the haemagglutination assay prior to HI. Limitations in the accuracy of the haemagglutination assay controlling for virus concentration for both parent and mutant viruses mean that small effects on $\log_2$ HI titre (95% CI: $\pm$ 0.78 $\log_2$ titre) could be a result of test error. The haemagglutination assay was used to generate a virus sample that could agglutinate 0.50 µl of turkey erythrocytes (0.75% suspension in phosphate buffered saline) suspended in a further 0.50 µl of phosphate buffered saline when diluted 8-fold (or 3 $\log_2$ titre) but not when diluted 16-fold (or 4 $\log_2$ titre). If the exact dilutions of each virus are uniformly distributed between 3 and 4 $\log_2$ titres, the potential difference between the parent and mutant virus dilutions is between + 1 $\log_2$ titre and − 1 $\log_2$ titre, which could result in a two-fold difference in HI titre ($\pm$ 1 $\log_2$ titre). The distribution of errors is therefore the difference between two independent uniform variables, and 95% of the time this will be within $1-\sqrt{0.05} \approx 0.78$ $\log_2$ titre.

3.4 Results

3.4.1 The effect of amino acid substitutions at specific positions

Linear mixed effects models were used to identify antigenic relationships and their predictors by accounting for variation in HI titres, as described by Reeve *et al.* (2010). Initial model selection identified non-antigenic sources of variation in HI titre. It was determined that a fixed effect, A_v, for each virus, v, should be included in the model (n = 506, p < 10^{-20}), to account for consistent differences in titres between viruses, reflecting changes in receptor-binding avidity amongst other factors. A further fixed effect, I_r, was required for each reference virus, r (n = 43, p < 10^{-20}), to account for consistent differences in titres between antisera raised against different reference viruses, potentially reflecting differences in immunogenicity. Date of test (n = 351) needed to be controlled for as a random effect, ϵ_D (n

$= 351$, $p < 10^{-20}$), accounting for variability in batches of erythrocytes and dilutions of erythrocytes, antisera and viruses. These factors compensate for non-antigenic effects impacting on HI titres (Equation 3.2).

$$H_{r,v} = k_0 + I_r + A_v + k_1\alpha_1(r,v) + \epsilon_D + \epsilon_R \tag{3.2}$$

$H_{r,v}$ is the $\log_2$ HI titre for test virus v and antiserum raised against reference virus r. k_0 is a baseline, and ϵ_R is the residual measurement error not explained by the model.

Equation 3.2 includes a term, $k_1\alpha_1(r,v)$, to investigate the effect of amino acid substitutions at specific positions: α_1 represents the presence (1) or absence (0) of an amino acid mismatch between the reference virus, r, and test virus, v, at a specific position. k_1 is the associated regression coefficient. Using this model (Equation 3.2) substitutions at over 50% of non-conserved, surface-exposed positions and over 25% of non-conserved, non-surface-exposed positions were significantly correlated with reduced HI titre ($p < 0.05$) using a Holm-Bonferroni correction for multiple tests (Holm, 1979). Furthermore, the number of synonymous mutations between viruses was significantly correlated with reduced titre ($p < 10^{-15}$) because of a correlation between molecular and antigenic evolution. This demonstrates that a simple regression analysis will incorrectly identify some antigenically neutral changes as antigenically important (i.e. false positives) simply because they occur in either the same branch or in a nearby branch to one where antigenically important substitution(s) occur.

3.4.2 Incorporating phylogenetic structure

The described tendency for identification of false positives required phylogenetic structure to be incorporated in the model. Equation 7 of Reeve *et al.* (2010) was used to identify branches of the phylogeny that were correlated with lower HI titres when they separated reference virus and test virus:

$$H_{r,v} = k_0 + I_r + A_v + \sum\nolimits_i m_i\delta_i(r,v) + \epsilon_D + \epsilon_R \tag{3.3}$$

Equation 3.3 incorporates branch terms $m_i\delta_i(r, v)$ instead of the term representing substitutions at a single amino acid position: $\delta_i = 1$ when reference virus (r) and test virus (v) are separated by branch i of the phylogeny and $\delta_i = 0$ otherwise, with m_i being the associated regression coefficient from the mixed effects model. Because of the computational infeasibility of searching the 2^{1010} possible antigenically important sets of branches (in a tree containing 1010 branches), a stochastic hill-climbing approach was used. This identified 62 branches as correlating with drops in HI titre when they separated reference and test viruses, indicating

that antigenic evolution occurred in these branches. Such antigenic events were found in much higher proportion in the trunk (38.3%) than in side (4.6%) branches (χ^2 test, $p < 10^{-13}$), supporting the standard model of influenza antigenic drift, whereby substitutions altering antigenicity without loss of fitness undergo preferential fixation, thus forming the trunk lineage from which future viruses descend (Fitch *et al.*, 1997; Nelson & Holmes, 2007). With these 62 branch terms included in the model (Equation 3.3), there was no longer a significant correlation between HI titre and the number of synonymous mutations between reference and test virus ($p > 0.05$), nor with any of the non-surface-exposed positions. This shows that including branch terms accounted for the correlation between molecular and antigenic evolution and therefore reduced the false discovery rate as required.

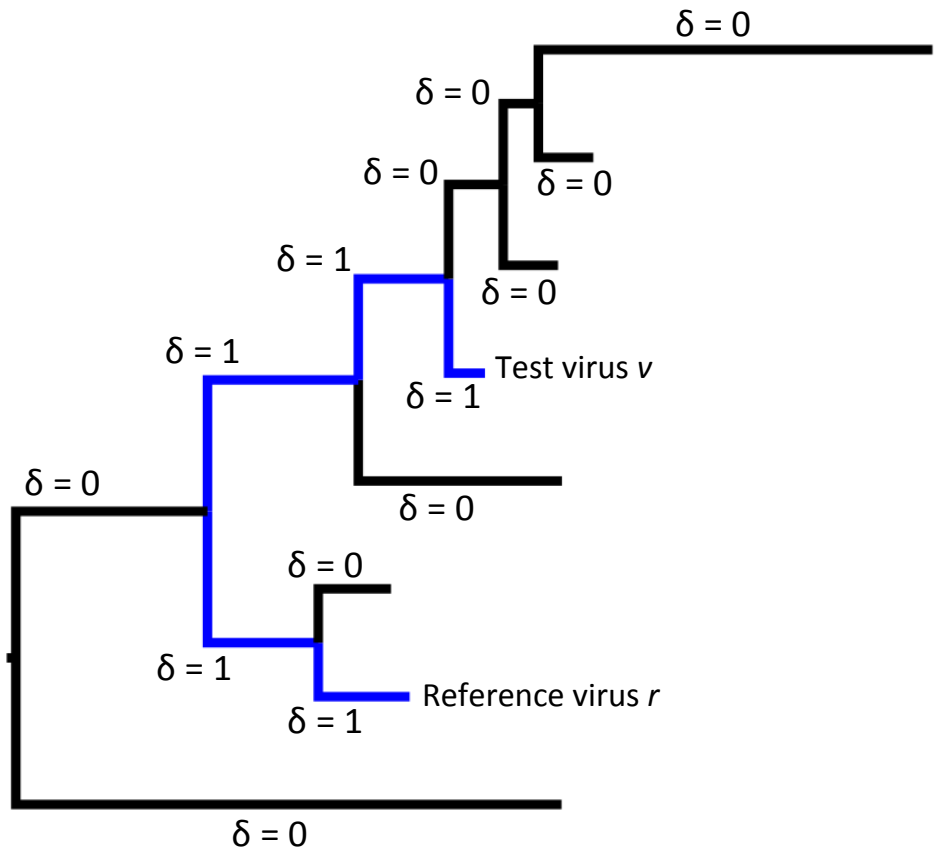

Figure 3.2: An example of a phylogeny including δ terms introduced in Equation 3.3: The assignment of values to δ terms associated with each branch, given the indicated positions of a reference virus, r and a test virus, v, is illustrated. $\delta = 1$ for each branch of the phylogeny that separates reference virus r and test virus v (these branches are also coloured blue), while $\delta = 0$ for each branch that does not occur between them.

3.4.3 Substitutions affecting antigenicity in multiple positions of the phylogeny

The model was extended by combining Equations 3.2 and 3.3 to include explicit terms for amino acid substitutions to identify antigenically important substitutions:

$$H_{r,v} = k_0 + I_r + A_v + k_1\alpha_1(r,v) + \sum_i m_i\delta_i(r,v) + \epsilon_D + \epsilon_R \tag{3.4}$$

Equation 3.4 incorporates the term $k_1\alpha_1(r,v)$ from Equation 3.2 representing specific amino acid substitutions. Each of the previously identified 62 branch terms (δ_i) were included and associated regression coefficients (m_i) were re-estimated in a model containing the $k_1\alpha_1(r,v)$ term. Because δ_i terms account for the antigenic changes inferred to occur in single specific branches of the phylogeny, any significant improvement to model fit by α_1 must be a result of the term representing amino acid substitutions at a single HA1 position being correlated with a change in the antigenicity of the virus represented in multiple branches of the phylogeny. Thus an improvement to model fit achieved by inclusion of α_1 indicates that there have been alternative, convergent or back-substitutions at the same amino acid position associated with antigenic change in at least two branches of the phylogeny.

One hundred and thirteen non-conserved, surface exposed amino acid positions of the HA1 domain were tested as predictors of variation in HI titre. At four of these (positions 141, 153, 187 and 190), the inclusion of an α_1 term representing a substitution at that site (Equation 3.4) improved model fit compared with the model containing only branch terms (Equation 3.3). Since the identified positions improve model fit in the presence of branch terms (δ_i), it can be inferred that substitutions at these positions correlate with antigenic change in more than one position of the phylogeny. Thus these substitutions are described as having *support across the phylogeny*. Each of these four amino acid positions (Figure 3.3) has previously been allocated to one of the H1 antigenic sites (Caton *et al.*, 1982). Position 187 is also a constituent of the primary sialic acid receptor-binding site and the analogous position 190 in H3-HA has been described as forming hydrogen bonds with the 9-hydroxyl group of sialic acid (Skehel & Wiley, 2000). Non-surface exposed positions were examined separately and model fit was not significantly improved by substitution at any of these positions.

Different substitutions at the same position are expected to vary in antigenic impact according to the biochemical properties of the amino acids involved. To account for this, the significance and average impact (in antigenic units, where a unit corresponds to a 2-fold dilution in the HI assay) of each substitution at HA1 positions 141, 153, 187 and 190 that was observed to have occurred between reference and test viruses in the dataset was measured. Substitutions between seven pairs of amino acids at the four positions showed significant antigenic impact with support across the phylogeny. The mean antigenic impact (k_1 in Equation 3.4) of

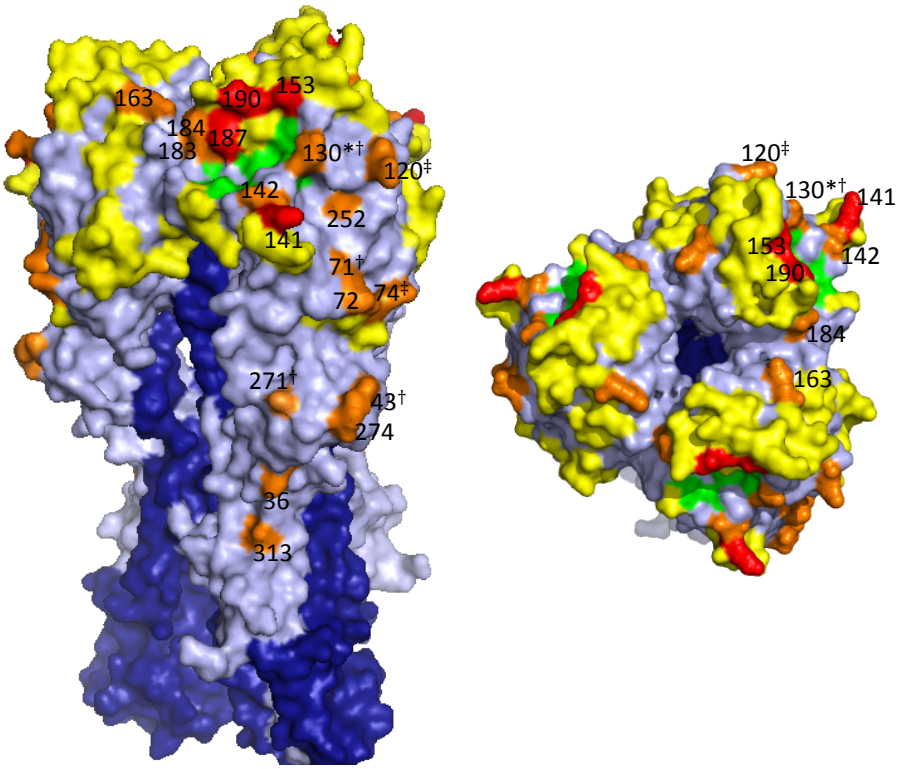

Figure 3.3: A(H1N1) HA positions implicated in antigenic change: Amino acid positions identified as affecting antigenic phenotype shown on the HA structure of A/Puerto Rico/8/34 (Protein Data Bank ID: 1RU7) (Gamblin *et al.*, 2004). HA1 is shown in light blue and HA2 in dark blue. The primary sialic acid binding site is shown in green and antigenic sites Ca, Cb, Sa and Sb are coloured yellow. Amino acid positions of substitutions identified as causing antigenic change in multiple positions of the phylogeny are shown in red. Positions of substitutions correlated with antigenic change in a single position of the phylogeny are shown in orange. *As A/Puerto Rico/8/34 possessed a deletion at 130 relative to some viruses used in the study, position 129 is coloured instead. Substitutions at positions 43, 71, 130 and 271 (†) and at positions 74 and 120 (‡) were each identified as alternative causes of the antigenic change inferred to have occurred at two specific points in the evolutionary history of the virus.

exchange between amino acids of each pair is shown in Table 3.1. Among these substitutions, exchange between lysine and glutamic acid at position 141, the most prominently protruding position on the 140-loop (antigenic site Ca) is associated with the greatest mean drop in cross-reactivity (2.37 antigenic units). The remaining three positions at which there was support across the phylogeny (153 (150-loop), 187 and 190 (190-helix)) form an exposed ridge bordering the upper lip of the depression containing residues involved in binding to host receptors (Figure 3.3).

Table 3.1: HA1 amino acid substitutions that correlate with antigenic changes in A(H1N1) viruses

Substitution(s) (H1-HA numbering)	Antigenic site H1	Antigenic site H3	Antigenic impact*
Substitutions with support across phylogeny identified using Equation 3.4†:			
K141E	Ca	A	2.37 (2.27-2.47)
E153G	Sa	B	0.20 (0.07-0.33)
E153K	Sa	B	0.66 (0.39-0.93)
G153K	Sa	B	1.50 (0.51-2.49)
D187N	Sb	B	0.33 (0.30-0.36)
D187V	Sb	B	0.88 (0.51-2.49)
A190T	Sb	B	0.24 (0.17-0.31)
Substitutions with support across phylogeny identified using Equation 3.5†:			
S36N	-	C	0.66 (0.22-1.11)
S72F	Cb	E	0.81 (0.49-1.13)
E74G, E120G‡	Cb,-	E,A	0.43 (0.29-0.57)
R43L, F71I, ΔK130‡, S271P	-,Cb,-,-	C,-,A,-	3.53 (3.44-3.62)
S142N	Ca	A	0.75 (0.58-0.92)
K163N	Sa	-	0.67 (0.62-0.73)
S183P	-	B	0.61 (0.33-0.89)
N184S	Sb	B	0.51 (0.31-0.70)
W252R	-	-	0.37 (0.32-0.43)
E274K	-	-	1.31 (0.68-1.93)
R313K	-	-	1.47 (0.84-2.10)

* k_1 in Equation 3.4 or k' in Equation 3.5. Mean and 95% CI are shown.

† Substitutions identified by likelihood ratio test using p-value of 0.05 adjusted using Bonferroni correction.

‡ Multiple substitutions in the same branch offer alternative explanations for the associated antigenic change.

3.4.4 Substitutions affecting antigenicity at single positions in the phylogeny

Next, terms k_j and α_j representing each of the the seven inferred antigenic substitutions at the four positions with support across the phylogeny (Table 3.1) were added to produce Equation 3.5. The causes of antigenic change in branches that still had large estimated antigenic impacts were investigated:

$$H_{r,v} = k_0 + I_r + A_v + k'\alpha'(r,v) + \sum_j k_j\alpha_j(r,v) + \sum_i m_i\delta_i(r,v) + \epsilon_D + \epsilon_R \qquad (3.5)$$

Terms for these seven substitutions absorbed variation in HI previously explained by branch terms that correspond to the positions in the phylogeny where those substitutions were estimated to have occurred. However, this model still included 18 branch terms representing internal branches of the phylogeny whose estimated impact on the HI assay (m_i), in the model containing terms for each of the seven substitutions, remained at least 0.25 antigenic units. Each of these 18 branch terms were excluded in turn, the model re-built with the residual branches, and each remaining amino acid position (as k' and α') was retested to determine which substitution(s) could explain the variation in HI titre associated with the excluded branch term. A substitution identified at this stage (when a branch term had been excluded) was inferred to have caused the associated antigenic change at that position in the phylogeny if it was the only substitution to be identified.

In nine cases, a single substitution was identified as explaining variation in HI titre upon exclusion of one branch. These substitutions were at positions 36, 72, 142, 163, 183, 184, 252, 274 and 313 (Table 3.1). Unique identification was not possible in two further cases, as multiple substitutions occurring in the same branch could not be discriminated. The branch associated with the greatest drop in HI titre across the phylogeny (starred in Figure 3.4) has a deletion of lysine at position 130 (ΔK130) and substitutions R43L, F71I and S271P. The antigenic significance of ΔK130 has been described (McDonald *et al.*, 2007); however each of these co-occurring substitutions have also been identified as antigenic determinants by another *in silico* technique, which did not identify them as false positives (Huang *et al.*, 2012). Each of these four changes is assigned equal weight in our model but it is identified explicitly that they offer alternative explanations for the same antigenic change and are not independent antigenic determinants. To infer their individual effects experimental investigation was required. One further instance of alternative substitutions at different positions explaining an antigenic change equally well involved positions 74 and 120 (Table 3.1).

Although the substitutions identified when branch terms were excluded correlated with antigenic change at only a single position in the phylogeny, it is notable that, among them, positions 72, 74, 142, 163 and 184 map to previously described H1 antigenic sites while positions analogous to 36, 120 and 183 are constituents of H3 antigenic sites (Wiley & Skehel, 1987). The amino acids and inferred mean antigenic impact of each of the substitutions is shown in Table 3.1 alongside the seven substitutions identified with support across the phylogeny. Locations within the phylogeny where any of the identified substitutions in Table 3.1 altered the antigenic phenotype of the virus by at least 0.5 antigenic units and the degree of correspondence with changes to the H1 vaccine component are shown in Figure 3.4.

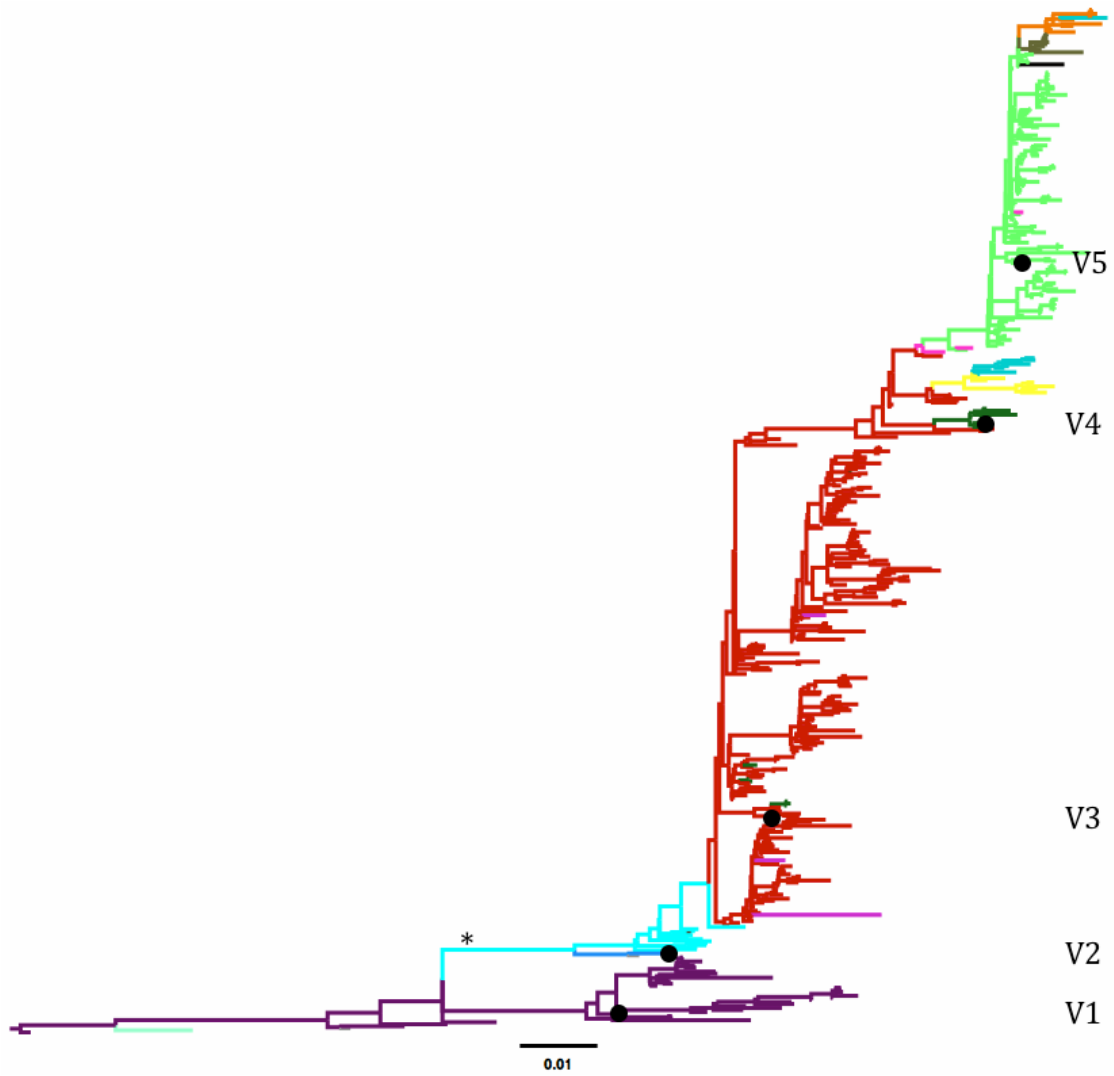

Figure 3.4: Locations of significant antigenic change in HA1 phylogeny: HA1 phylogeny showing positions of significant antigenic substitutions. Colour changes mark the locations of substitutions associated with changes in antigenic phenotype of at least 0.5 antigenic units according to Table 3.1. The position of the branch associated with the greatest drop in cross-reactivity among the sampled viruses is marked (*). Filled black circles indicate the positions of viruses included in the influenza vaccine over the period of HI data collection and are labeled: A/Bayern/7/95 (V1), A/Beijing/262/95 (V2), A/New Caledonia/20/99 (V3), A/Solomon Islands/3/2006 (V4) and A/Brisbane/59/2007 (V5). Older vaccine viruses are not indicated. Branch length indicates the estimated number of nucleotide substitutions per site.

3.4.5 Production of mutant viruses by reverse genetics

To validate the identification of substitutions affecting antigenicity and assess the accuracy of estimated antigenic effects mutant viruses containing a subset of the amino acid substitutions identified in Table 3.1 were generated by reverse genetics. The HA gene of an exclusively cell-propagated virus, A/Netherlands/1/93 (Neth93) was used. The K130 deletion (ΔK130) and the R43L substitution were introduced into the Neth93 HA independently to test whether both of these changes cause antigenic change. Consistent with the results of McDonald *et al.* (2007), a large antigenic impact associated with ΔK130 was observed. The introduction of ΔK130 therefore generated an additional, antigenically distinct HA background (Neth93 Δ130) in which to further test the effects of substitutions. The HA genes of both Neth93 and Neth93 Δ130 were used to produce viruses carrying individual substitutions of K141E, E153K, D187N and D187V. E153G and A190T were not investigated as it was considered that their antigenic impact might be too small to detect; the lower 95% confidence interval on their predicted antigenic impact was 0.07 and 0.17 $\log_2$ titre respectively (for all substitutions tested, the lower 95% confidence interval $\geq$ 0.3 $\log_2$ titre).

Table 3.2: Antisera used to characterise recombinant viruses

Substitution resulting from mutagenesis	Reference viruses against which antisera were raised						
	A/BAY /7/95	A/JBG /82/96	A/JBG /159/97	A/UU /209/98	A/HK /4847/98	A/NC /20/99	A/HK /1252/00
R43L	R	R	L	L	L	L	L
ΔK130	K	K	Δ*	Δ*	Δ*	Δ*	Δ*
K141E	K	K	K	K	K	K	E
E153K	E	E	E	K	G	G	G
D187N	D	D	D	D	N	D	D
D187V	D	D	D	D	N	D	D

Abbreviations in reference virus names: Bayern (BAY), Johannesburg (JBG), Ulan-Ude (UU), Hong Kong (HK) and New Caledonia (NC). * Δ indicates deletion of amino acid corresponding to position 130). Red indicates that the reference virus used to generate antisera lacked the introduced substitution and so was in the ancestral state (e.g. R43) and blue indicates that the reference virus shared the introduced substitution (e.g. L43). Absence of colour indicates that the reference amino acid identity at the position of substitution in the recombinant virus was different from both of the parental viruses (Neth93 and Neth93 Δ130) and from the mutant virus.

Mutant recombinant viruses were characterised by HI using a panel of post-infection ferret antisera raised against seven reference viruses. Amino acid identities of these seven reference viruses at positions at which substitutions were introduced in recombinant viruses (HA1 positions 43, 130, 141, 153 and 187) are shown in Table 3.2. Geometric mean titres for these HI assays averaged across four repeats are shown in Appendix A, Table A.1 and a visual representation of these data is given in Figure 3.5. The top row of the heatmap in

Figure 3.5 shows the average log titres for the parental virus Neth93 against each of the antisera used. High titres (dark red) indicate that, of the antisera used, Neth93 is most antigenically similar to A/Bayern/7/95 and A/Johannesburg/82/96 and low titres indicate that Neth93 is least antigenically similar to A/Ulan-Ude/209/98, A/Hong Kong/4847/99, and A/Hong Kong/1252/2000. In the second row, the colours for the mutant recombinant virus possessing the substitution R43L are the same as in the top row and the dendogram bifurcation separating these two rows at the left of the figure is almost flat indicating that R43L has had almost no impact on HI titre.

The deepest bifurcation in the dendogram at the left of the figure corresponds to the introduction of ΔK130. Titres recorded for the ΔK130 mutant are lower than Neth93 for the antisera A/Bayern/7/95 and A/Johannesburg/82/96 and higher for the antisera A/Johannesburg/159/97, A/Ulan-Ude/209/98, A/Hong Kong/4847/99, and A/New Caledonia/20/99 indicating reduced and increased antigenic similarity against these two groups of antisera respectively. Table 3.2 shows that the the viruses A/Bayern/7/95 and A/Johannesburg/82/96 possess K130 while A/Johannesburg/159/97, A/Ulan-Ude/209/98, A/Hong Kong/4847/99, and A/New Caledonia/20/99 bear a deletion of residue 130. This indicates that the ΔK130 change is also responsible for the deepest bifurcation in the antisera dendogram above the heatmap in Figure 3.5. A/Hong Kong/1252/2000 is also Δ130 but the ΔK130 change has not increased titre, relative to the parental virus Neth93, against this antiserum.

The second deepest bifurcation in the virus dendogram is associated with the introduction of E153K or K141E substitutions into the Neth93 Δ130 background. In addition to lowering titres against A/Johannesburg/159/97, A/Ulan-Ude/209/98, A/Hong Kong/4847/99, and A/New Caledonia/20/99, K141E is associated with an increased titre against the A/Hong Kong/1252/2000 antisera. The titre recorded for the Neth93 ΔK130 K141E mutant was almost as high as the homologous titre recorded for A/Hong Kong/1252/2000 (174 vs. 202, Table A.1), the only reference virus used with E141 indicating that the introduction of two amino acid changes (ΔK130 K141E) could convert the antigenic phenotype of Neth93 from A/Bayern/7/95-like to A/Hong Kong/1252/2000-like.

3.4.6 The antigenic and non-antigenic effects of introduced substitutions

It is possible that HI titres recorded, and represented in Figure 3.5, were influenced by a combination of both the antigenic and non-antigenic properties of the viruses tested. To assess the antigenic impact of each amino acid substitution introduced by mutagenesis, antisera of two types were chosen: 1) antisera raised against parent-like viruses that had amino acid identity in common with the parent virus (i.e. R43 for the substitution R43L); 2) antisera raised against mutant-like viruses that had amino acid identity in common with the recombinant virus (i.e. L43 for the substitution R43L). For each amino acid substitution

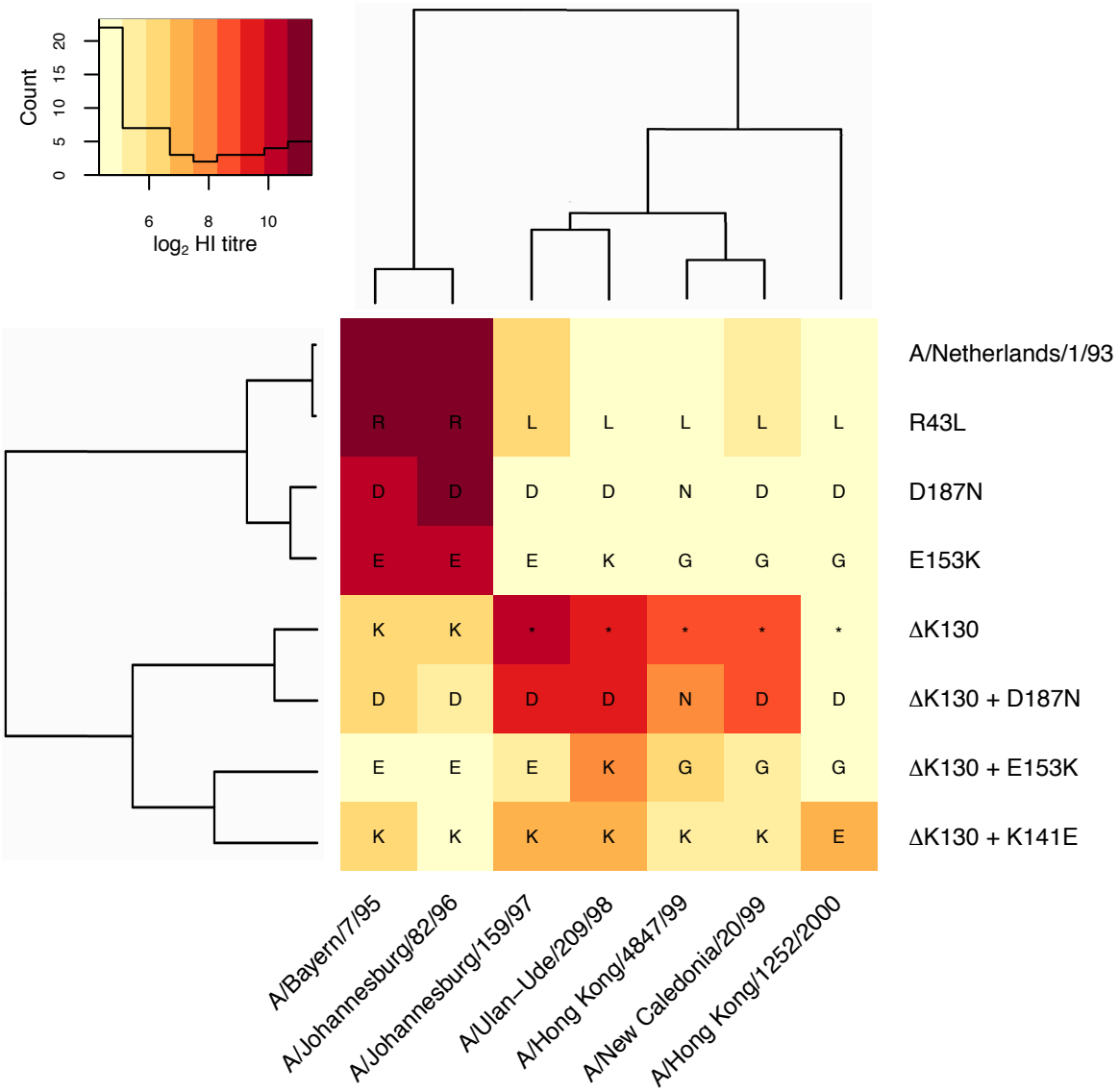

Figure 3.5: Heat-map and clustering analysis of HI titres with mutant viruses: Hierarchical clustering of the wild type Neth93 and mutant viruses generated from it by reverse genetics. Reference viruses used to generate antisera are arranged according to their ability to inhibit agglutination of turkey erythrocytes by each virus. Viruses are simultaneously clustered along the vertical axis according to their antigenic profile. Dendograms indicating antigenic relatedness are shown at the top (for antisera) and to the left (for viruses). Colouring represents $\log_2$ HI titre as indicated at top left with the histogram showing the frequency (count) for each titre. Cells are labelled to indicate reference virus amino acid identity at the position of substitution. In the bottom three rows, labels refer to the second substitution (and not ΔK130).

introduced into the recombinant viruses, the assignment of antisera to these two categories, based on reference virus amino acid sequence, is shown in red or blue cell colour respectively in Table 3.2.

For each recombinant virus, changes in $\log_2$ HI titre, relative to the parental virus Neth93 or Neth93 Δ130, were partitioned into antigenic and non-antigenic components using Equation 3.1. The predicted antigenic effect of each substitution (from Table 3.1) is shown alongside mean observed changes in $\log_2$ HI titre partitioned into antigenic (ΔH_A) and non-antigenic effects (ΔH_N) in Table 3.3. Comparing columns two and four in Table 3.3 shows the similarity in the mean antigenic impact of substitutions in models applied to the full HI dataset and in the reverse genetics experiments described above. The range of antigenic effects of K141E, ΔK130, E153K and D187N amino acid substitutions, measured against the panel of antisera shown in Table 3.2, were consistent with predictions from modelling. The predicted antigenic impact and the mean and range in antigenic impact (ΔH_A) measured experimentally using individual antisera is shown in Figure 3.6. Across all substitutions a mean error in predictions of only 0.14 antigenic units was calculated.

Table 3.3: Predicted and observed antigenic impacts of A(H1N1) HA1 amino acid substitutions

Substitution	Predicted antigenic effect*	Mutagenesis background	Observed effect† ΔH_A	Observed effect† ΔH_N
K141E‡	2.37	Neth93 Δ130	2.60	+0.27
E153K	0.66	Neth93	0.67	-0.42
		Neth93 Δ130	0.65	**-2.15**
		Averaged	0.66	**-1.28**
D187N	0.33	Neth93	0.41	-0.41
		Neth93 Δ130	-0.08	-0.54
		Averaged	0.16	-0.47
ΔK130	3.53§	Neth93	4.10	-0.78
R43L	3.53§	Neth93	0.01	-0.01

* Predicted mean antigenic impact (from Table 3.1) measured in antigenic units.

† Mean observed changes in $\log_2$ HI titre (in antigenic units) partitioned into antigenic (ΔH_A) and non-antigenic (ΔH_N) effects.

‡ HA plasmids were generated for the mutant Neth93 K141E but multiple attempts to rescue were unsuccessful.

§ ΔK130 and R43L occur in the same branch of the phylogeny and thus offer alternative explanations for the associated antigenic change.

The predicted and observed antigenic impacts, based on HI results, shown in Table 3.3 and in Figure 3.6 indicate that our model captures the mean impacts of the HA1 amino acid substitutions identified. However, non-antigenic effects (ΔH_N) of substitutions that resulted

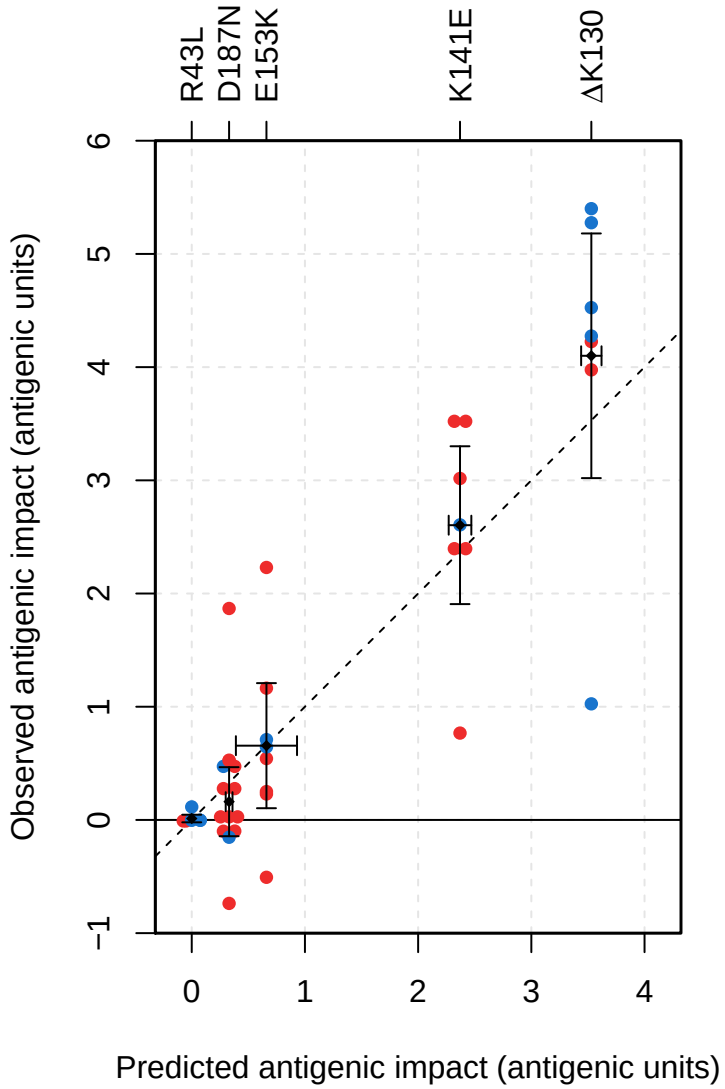

Figure 3.6: Observed and predicted antigenic impact of H1 HA1 amino acid substitutions: The mean antigenic impact of each substitution predicted from modelling (Table 3.1) plotted against the mean observed impact averaged across antisera in the panel (Table 3.2). 95% confidence intervals are shown for both. Each point shows the observed mean antigenic impact (ΔH_A, change in HI titre for a recombinant virus relative to its parent virus) of a particular amino acid substitution (labelled at top) with each antiserum in the panel. For each amino acid substitution an antiserum was classified into one of two groups dependent on whether the reference virus it was raised against lacked or shared that particular deletion/substitution. Red points indicate that the reference virus lacked the amino acid substitution, so the predicted impact of mutation is a reduction in titre; blue points indicate that the reference virus shared the substitution, so the predicted impact of mutation is an increase in titre. The number of points for each substitution is dependent on whether it was inserted into one or both (Neth93 and Neth93 Δ130) parental viruses and on the number of reference viruses sharing one of the two amino acid identities being considered. A negative observed antigenic impact indicates a change in HI titre in the opposite direction to that predicted.

in higher or lower HI titres irrespective of antigenic similarity between test virus and the reference virus against which a particular antiserum was raised were also observed. Such effects exceeding 0.78 antigenic units, shown in bold in Table 3.3, cannot be explained solely by differences in virus concentration resulting from the limited accuracy of the haemagglutination assay, used to standardise haemagglutinating units prior to HI. A relatively small antigenic impact of the E153K substitution, but a large non-antigenic effect, on HI titre was observed. This result is supported by previous work showing E153K to have a relatively small impact on monoclonal antibody binding while causing a large increase in erythrocyte binding avidity that together contributed to large reductions in HI titre indicating escape from inhibition by polyclonal antiserum (Hensley *et al.*, 2009).

Two HA plasmids were generated containing the D187V substitution; multiple attempts to rescue Neth93 D187V were unsuccessful and while Neth93 ΔK130 D187V was rescued, the absence of an available antiserum raised against a virus carrying this substitution (Table 3.2) prevented the partitioning of changes in HI, relative to the parental virus, into antigenic and non-antigenic effects. Where no reference strain exists that carries the substitution, limitations in the accuracy of the haemagglutination assay controlling for virus concentration (± 0.5 antigenic units) for both parent and mutant viruses mean that small effects (95% CI: ±0.78 antigenic units) could be a result of test error. This is because the haemagglutination inhibition assay is is used to generate a sample that can agglutinate a standardised sample of erythrocytes when diluted 8-fold (or 3 $\log_2$ titre) but not 16-fold (or 4 $\log_2$ titre). With a potential difference of 1 $\log_2$ titre between parent and mutant, 95% of the time this would be expected to be within 0.78 $\log_2$ titre as explained in Section 3.3.5. Having reference viruses that have both amino acid identities allows antigenic and non-antigenic effects to be distinguished (see Equation 3.1). A mean drop of 0.17 units measured against antisera raised against viruses lacking the D187V substitution was therefore consistent with our predicted effect of between 0.51 and 1.25, giving 95% CI of -0.61 to 0.95.

3.5 Discussion

Using a modelling approach that integrated HA sequence data and HI antigenic data for over 500 viruses, substitutions responsible for the antigenic evolution of former seasonal influenza A(H1N1) viruses over a period of more than 10 years were identified. Substitutions at 15 amino acid positions in HA1 that not only had high-impact on antigenicity (which individually may lead to a need to change vaccine virus), but also those of low-impact with many possibly not being directly observable in routine HI assays were identified. Antigenic cartography of these A(H1N1) viruses identified three antigenic clusters (Figure 2.10 in Chapter 2). In addition to the two amino acid substitutions that can explain the transitions between these clusters (ΔK130 and K141E), the antigenic determinants identified include substitu-

tions at thirteen further positions. At three of these positions (153, 187, and 190), the same substitutions co-occur with detectable antigenic change at multiple times in the evolutionary history of the virus, consistently explaining observed antigenic changes. The application of the modelling framework used here to influenza viruses demonstrates the generality of an approach used previously to build predictive models of antigenic cross-reactivity between foot-and-mouth disease viruses (Reeve *et al.*, 2010). This approach is distinguished from other regression-based approaches by the use of phylogenetic information to reduce false positive detection rates and to identify when alternative explanations for the same antigenic event exist. By incorporating phylogenetic information, the locations in the HA1 phylogeny where antigenic changes occurred were also estimated. Antigenic events were found disproportionately in the trunk lineage using Equation 3.3 (χ^2 test, $p < 10^{-13}$) supporting the standard model of antigenic drift in influenza A HA, whereby substitutions altering antigenicity without loss of fitness undergo preferential fixation, thus forming the trunk lineage from which future viruses descend (Fitch *et al.*, 1997; Nelson & Holmes, 2007).

In antigenic maps of A(H3N2) viruses, antigenic distances between viruses belonging to the same antigenic cluster often exceeded distances between viruses in adjacent clusters, demonstrating the need to assess non-cluster defining substitutions (Smith *et al.*, 2004). Further, the selection of approximately twice as many H3N2 vaccine viruses as antigenic clusters identified by Smith *et al.* (2004) during the period 1968 to 2003 is supportive of the importance of substitutions not readily identified using antigenic maps in the evolution of influenza viruses. Substitutions identified in this chapter ranged in their antigenic impact from high-impact antigenic determinants which individually may have led to a need to change the vaccine virus (ΔK130 which was previously identified as antigenically important and as being associated with the A/Bayern/7/95 to A/Beijing/262/95 vaccine change by McDonald *et al.* (2007) and the K141E substitution which was observed multiple times and associated with the A/New Caledonia/20/99 to A/Solomon Islands/3/2006 vaccine change) to smaller-impact antigenic determinants that were individually less apparent but that may have accumulated to cause antigenic drift. Inclusion of antigenic substitutions of lower-impact allowed us to trace the antigenic evolution of influenza A(H1N1) in finer detail and the identification of substitutions responsible for such smaller incremental changes in antigenicity also raises the prospect of fine-tuning vaccine viruses by mutating existing candidate vaccine viruses or their derivatives.

The model also dissects the results of HI assay into antigenic and non-antigenic effects and in the characterisation of mutant viruses generated by reverse genetics non-antigenic effects of specific substitutions on HI were identified, in addition to their antigenic effects. Changes in the receptor-binding avidity of influenza viruses are known to contribute to apparent antigenic effects as measured by HI assay (Daniels *et al.*, 1984; Hensley *et al.*, 2009). The results of the described reverse genetics experiments showed that investigated substitutions also affected HI titre as a result of non-antigenic effects (ΔH_N) in addition to the predicted antigenic effects (ΔH_A). It is clear that that understanding the impact of non-antigenic variation on

titre is important when interpreting HI data and drawing conclusions on antigenic similarity of influenza viruses. Understanding the genetic variation underlying changes in the receptor-binding avidity of influenza viruses that contribute to apparent antigenic effects as measured by HI assay is clearly a very important area for further investigation. A sequence-based method to investigate the mechanistic relationship between HA amino acid substitutions and variation in HI attributable to changes in avidity, to complement our current model of antigenic evolution, is desirable.

The ability to quantify the phenotypic impact of substitutions involved in antigenicity introduces an opportunity to investigate interactions between such substitutions. It has been proposed that epistasis is prevalent in the evolution of the influenza surface proteins, however few examples have been confirmed using phenotypic data (Kryazhimskiy *et al.*, 2011). The models presented in this chapter do not explicitly account for how the antigenic impact of substitution at one position depends on substitutions present at other HA, or neuraminidase, positions. In the computational analysis described, variation in the antigenic effects of particular substitutions across the full dataset was observed. Further, in the assessment of recombinant viruses, variation in the effect of substitutions was observed between mutagenesis backgrounds (Neth93 and Neth93 ΔK130), dependent on which antiserum was used in HI assays. I speculate that the antigenic impact of a particular HA substitution is dependent on how similar or different the HA epitope structure of a test virus is to that of the reference virus used to raise antiserum, on the basis that successive substitutions in the same antigenic site will have diminishing fitness returns in terms of decreased antibody binding. Understanding how variation in the antigenic impact of a particular substitution can be attributed to amino acids present at other positions on the antigenic surface is another area for investigation, called for by this work, that should further improve our understanding of how specific substitutions affect antigenicity.

CHAPTER 4

Identifying the genetic drivers of antigenic evolution of influenza A(H3N2) viruses

Identifying the genetic drivers of antigenic evolution of influenza A(H3N2) viruses

4.1 Abstract

Understanding how specific features of the viral genome contribute to strain fitness and evolutionary success is vital for increasing our ability to predict their evolutionary future. Influenza A(H3N2) viruses circulate globally in the human population with increased epidemic frequency and severity, relative to other seasonal influenza viruses. Current methods for predicting which influenza A(H3N2) viruses among circulating variants will seed epidemics in future seasons suggest that the fitness profile in the viral population is complex with the accumulation of several adaptive and maladaptive mutations contributing to virus fitness, though our knowledge of the phenotypic impact of genetic differences remains limited. The haemagglutination inhibition assay is commonly used to assess influenza antigenic relationships, but only summarises a complex interplay of factors that affect survival. In particular, it confounds the effects of antigenicity and receptor binding avidity, both of which are well known to play critical roles. Here a sequence-based method that has previously been shown to predict vaccine match for foot-and-mouth disease and identify drivers of antigenic evolution of A(H1N1) viruses is extended and applied to A(H3N2) HI data. By quantifying the antigenic impact of specific amino acid substitutions that have occurred during the evolution of A(H3N2) a pathway towards the inclusion of a more mechanistic, detailed understanding of the role-play between molecular and antigenic evolution in sequence-based predictions of evolutionary success is presented.

4.2 Introduction

A(H3N2) viruses have circulated globally in the human population since 1968 causing seasonal epidemics of disease. Influenza A(H1N1)pdm09 and influenza B viruses also circulate in the human population, cause similar symptoms and evolve by similar mechanisms of immune escape, however A(H3N2) viruses have a considerably higher disease burden and have been

responsible for the majority of severe illness caused by seasonal influenza in recent decades (Barr *et al.*, 2014). Influenza A viruses exhibit higher rates of nucleotide mutation and amino acid substitution, and reduced genealogical diversity relative to influenza B viruses, suggesting an increased role of adaptive evolution in A viruses (Bedford *et al.*, 2015). Though rates of molecular evolution in A(H3N2) and A(H1N1) viruses are similar, the rate of antigenic drift of A(H3N2) viruses is higher than that of both A(H1N1) and B viruses (Bedford *et al.*, 2014). Owing to a range of evolutionary and epidemiological factors, a more interconnected global population of A(H3N2) viruses has been shown to coincide with reduced rates of local persistence of genetic and antigenic variants, higher ages of infection, and larger, more frequent epidemics relative to A(H1N1) and B viruses (Bedford *et al.*, 2015).

As a result of the increased frequency and severity of seasonal A(H3N2) epidemics, relative to other human influenza viruses, this subtype has been at the focus of recently emerging techniques for predicting the evolutionary trajectory of circulating influenza viruses (Łuksza & Lässig, 2014; Neher *et al.*, 2014). The method described by Neher *et al.* (2014) uses information in the shape and branching pattern of the HA phylogenetic tree to track changes in the frequency of genetic variants and to make predictions of whether variants will increase or decrease in frequency in the future. Łuksza & Lässig (2014) also used information on the frequency of clades defined on the HA phylogenetic tree to inform predictions of future composition of the viral population, however their predictions were also informed by our understanding of which codons are involved in the antigenic evolution of A(H3N2) viruses. Their model incorporated a sequence-based estimate of viral fitness where antigenic novelty was approximated by a count of amino acid differences in antigenic sites reported by Shih *et al.* (2007), while substitutions outside these sites were penalised on the assumption that they are likely to to be deleterious (Łuksza & Lässig, 2014). A natural extension of these methods is the incorporation of a quantitative model of the impact of specific amino acid substitutions on antigenicity, and indeed other relevant facets of viral phenotype. Research in this area is restricted by limitations in our understanding of the molecular basis of antigenic evolution, the phenotypic impact of specific amino acid substitutions, and the consequences for viral fitness.

Much of our understanding of the antigenicity of A(H3N2) is derived from fundamental experimental work in the 1980s (Knossow *et al.*, 1984; Skehel *et al.*, 1984; Webster & Laver, 1980; Wiley *et al.*, 1981; Wilson *et al.*, 1981). The contributions of these studies included the solving of the 3D structure of the H3 HA glycoprotein (Wilson *et al.*, 1981) and the mapping of substitutions facilitating escape from neutralising monoclonal antibodies to antigenic sites denoted A–E (Wiley *et al.*, 1981). Various approaches have been used to investigate the mechanisms of antigenic drift (Hensley *et al.*, 2009; Koelle *et al.*, 2006) and phylogenetic-based methods have been used to identify positively selected codons that may have been important in the antigenic evolution of the virus (Bush *et al.*, 1999b; Shih *et al.*, 2007). The application of multi-dimensional scaling to HI data, termed antigenic cartography, has allowed

for visualisation of HI data and antigenic maps generated for A(H3N2) have indicated that antigenic evolution is punctuated, relative to the more continuous accumulation of genetic changes (Smith *et al.*, 2004). Substitutions causing major changes in A(H3N2) antigenic evolution, many of which are apparent as transitions between clusters on these antigenic maps, have subsequently been identified by exhaustive testing of viruses generated by reverse genetics containing cluster-difference substitutions introduced by site-directed mutagenesis (Koel *et al.*, 2013). This experimental approach provides definitive identification of the genetic determinants causing very large antigenic changes.

It is possible that the adaptive evolution of A(H3N2) viruses proceeds via rare amino acid substitutions with large phenotypic effects only, however under this scenario the fitness profile of the viral population is homogenous most of the time. That Łuksza & Lässig (2014) and Neher *et al.* (2014) were able to make informative predictions of evolutionary success in most seasons implies persistent fitness variation among circulating A(H3N2) viruses suggesting that adaptive evolution proceeds by the accumulation of a greater number of small effect mutations. Łuksza & Lässig (2014) found that while substitutions to antigenic site E did not inform predictions, substitutions in sites A–D did, which adds support to the hypothesis that a wider set of codons than those identified as causing the largest antigenic changes, which occur in sites A and B only (Koel *et al.*, 2013), are important in A(H3N2) antigenic evolution.

In this chapter, the application of the modelling approach introduced in Chapter 3 to viruses of the A(H3N2) subtype is described. In A(H1N1) it was observed that corrections to baseline titres made for each virus tested by HI co-varied with year of isolation, suggesting a potential molecular basis of variation in some non-antigenic trait affecting HI titre, possibly receptor-binding avidity. In this chapter the method of incorporating phylogenetic structure into the model is developed to allow non-antigenic differences between virus and antisera, which may be the result of changes in virus avidity and immunogenicity respectively, to be explained by branches of the phylogeny where these changes are estimated to have occurred. The aims of this chapter are to identify antigenic determinants that have caused both large and more gradual changes in antigenic phenotype of A(H3N2) viruses and to quantify their relative antigenic impact.

4.3 Materials and Methods

4.3.1 Data

The A(H3N2) dataset used in this chapter is introduced in Chapter 2. Most of the HI data and associated HA sequence data analysed were generated by staff of the Crick Worldwide Influenza Centre (formerly the WHO Collaborating Centre at the National Institute for Medical

Research, MRC National Institute for Medical Research, UK) over a number of years. These data were augmented with data from a study by Smith *et al.* (2004), made available subsequently (Bedford *et al.*, 2014), which comprised observations of 650 unique virus:reference strain combinations between 91 viruses, all isolated prior to 1990, and antisera raised against 36 reference strains. For computational reasons, a subset of the full A(H3N2) dataset described in Chapter 2 was used here. This subset consisted of 7,315 titre measurements for 2,738 unique virus:reference strain combinations between 229 viruses and antisera raised against 169 reference strains and was inclusive of the data published by Bedford *et al.* (2014).

4.3.2 Phylogenetic analysis

HA1 nucleotide sequences of all viruses were aligned using MUSCLE (Edgar, 2004). Phylogenies were estimated using a variety of nucleotide substitution and molecular clock models. A GTR+I+Γ_4 nucleotide substitution model was determined to be most suitable through comparison of Bayes factors (Suchard *et al.*, 2001). Bayes factor analysis also favoured allowing rates to differ at the third codon position relative to the first two and a relaxed, uncorrelated molecular clock model with branch rates drawn from an exponential distribution. Year of isolation was used to calibrate the molecular clock allowing rates of evolution along branches to be estimated. A coalescent demographic model with constant effective population size was used. To incorporate phylogenetic error, a posterior sample of 10,000 trees was sampled across and the maximum clade credibility tree was selected. Phylogeny construction and analysis was carried out using BEAST v1.7.5 (Drummond *et al.*, 2012) and Tracer v1.5 (Rambaut & Drummond, 2009).

4.3.3 Mixed effects modelling and model selection

As in Chapter 3, HI titres were log transformed (using base 2) to ensure homoscedasticity of residuals. A difference in $\log_2$ titre of one corresponds to a two-fold dilution of antiserum in the HI assay. Goodness of fit of models including each of the following variables were assessed by likelihood ratio test: the reference virus against which the antiserum was raised, the test virus, erythrocyte species (turkey or guinea pig), the date on which the assay was performed, and the presence or absence of the neuraminidase inhibititor oseltamavir in the assay.

In Chapter 3 each branch of the HA1 phylogeny was incorporated as a discrete indicator variable: a binary variable where 1 indicated that reference virus and test virus were separated by the branch in the phylogenetic tree and 0 otherwise with an associated antigenic impact, following Reeve *et al.* (2010). Each branch splits the tree in two and can be thought of as defining a single virus or clade of viruses separated from all other viruses in the tree. The inclusion of a branch term of this kind indicates that amino acid substitutions that occurred

in the branch changed the antigenic phenotype of the virus such that a lower than otherwise expected HI titre is observed when viruses separated by the branch in the phylogeny are tested together.

Other non-antigenic properties of viruses can also affect HI titre. In particular, substitutions altering the receptor-binding avidity may influence HI resulting in lower or higher titres against all antisera. Likewise, substitutions may alter some property of the virus so that reference viruses carrying these substitutions will produce antisera which will have higher or lower titres against all test viruses irrespective of antigenic similarity. Here, this property of a reference virus is referred to as its immunogenicity. These two properties of the virus, receptor-binding avidity and immunogenicity, were not reflected by the phylogenetic terms incorporated in the models in Chapter 3. In this chapter, the model is developed to allow changes in virus and antiserum reactivity associated with the genetic changes occurring in branches to be estimated. It was possible to do this in this chapter because the larger HI dataset was sub-sampled to give a dataset rich in reference viruses. When only a small proportion of viruses in a dataset have been used to generate antisera it is, in most cases, not possible to differentiate whether lower titres reflect antigenic differences or are due to changes in receptor-binding avidity.

In the method introduced in this chapter, each branch also had associated binary indicator variables to indicate whether the clade defined by that branch contained (1), or did not contain (0), the test virus or separately the reference strain used to generate antisera in the HI assay. The inclusion of either of these two variables indicates a receptor-binding avidity or immunogenic effect associated with the branch defining that clade respectively. Where possible, for each branch, we included these binary, indicator variables for the associated reactivity and immunogenic effects in addition to the antigenic branch effect used in Chapter 3 and described above. It was possible to estimate these three phylogenetic terms independently when the clade defined by a branch contained at least one virus tested by HI and at least one reference virus used to generate antisera used in the dataset. In the HA1 phylogeny that contained 458 branches, 335 branches defined clades containing at least one test virus and at least one reference virus allowing antigenic, avidity and immunogenicity effects to be estimated independently. Where a clade contains only test viruses (or reference viruses), the antigenic and receptor-binding avidity (or immunogenic) effects associated with the branch defining it cannot be discriminated. Therefore, branches were associated with either three binary indicator variables (allowing antigenic, avidity and immunogenicity effects to be estimated independently) or a single binary indicator (in cases where these effects could not be discriminated between). These various binary variables reflecting phylogenetic structure were tested as fixed effect terms with random restart hill-climbing (Russell & Norvig, 1995) used to determine the best model. To a random consistent starting model, branch terms were added and removed stepwise at random to maximise model fit, assessed by AIC (Akaike, 1974), as in Chapter 3. This was repeated while randomising their order to avoid sensitivity to the

order in which the parameters were presented.

The PyMOL Molecular Graphics System v1.7.7.2 (http://www.pymol.org) was used to visualise and identify surface-exposed amino acid positions on the HA1 according to the structure of the recombinant viral strain A/X-31, that carries HA from the A/Aichi/2/68 isolate (Protein Data Bank ID: 1HGG) (Sauter *et al.*, 1992). Amino acid dissimilarity between reference virus and test virus at each of 132 surface-exposed HA1 positions, not conserved within the dataset, was tested as a predictor of reduced HI titre ($p < 0.05$) using a Holm-Bonferroni correction to account for multiple tests (Holm, 1979). As in the previous chapter, HA2 positions were not tested for two reasons: First, the HA1 has been identified as the target for most anti-HA neutralising antibodies and for this reason is though to be the principal antigenic determinant for influenza A; second, for many viruses only HA1 gene sequences exist and thus opting to use the full HA gene would have greatly diminished the number of pairwise comparisons that could have been investigated. At each HA position identified at this stage, the mean antigenic impact of specific amino acid substitutions was determined by examining the associated parameter (k_1, Equation 4.3) using data subsets consisting of viruses and reference viruses with one of the two amino acid variants being considered at that position. Substitutions identified at this stage were added together into the same model along with identified phylogenetic terms. Phylogenetic terms that were still associated with drops in $\log_2$ titre exceeding 0.3 were then removed in turn. Upon removal of each of these branches, substitutions not identified as correlating with antigenic change at multiple points in the phylogeny were re-tested, again using a Holm-Bonferroni correction for multiple tests. The substitution, or substitutions, identified upon re-testing in the absence of a particular branch were only significantly correlated with antigenic change at that single point in the evolutionary history of the virus. The antigenic impact of all identified substitutions was then re-estimated in a single model lacking any of the branch terms included in models used to identify antigenically important substitutions. The estimated antigenic impacts of substitutions in this model were used for comparison with the experimentally estimated impact of substitutions introduced into recombinant viruses.

4.3.4 Analysis of HI assays using recombinant viruses

Koel *et al.* (2013) carried out an intensive reverse genetics search, introducing substitutions associated with transitions between eleven clusters of antigenically similar viruses expressed on previously generated antigenic maps of A(H3N2) viruses isolated between 1968 and 2003 (Smith *et al.*, 2004). To identify the HA substitutions responsible for transitions between antigenic clusters, viruses containing consensus amino acid sequences were constructed. Koel and colleagues then introduced each of the 54 cluster-difference substitutions within HA1 positions 109 to 301 into corresponding consensus HA. In several cases, double and triple substitution mutants were also tested and substitutions showing signs of altering antigenicity

were tested in reverse too. When cluster-difference substitutions did not bring about the anticipated change in phenotype other substitutions were also tested by the same method (Koel *et al.*, 2013). Parent and mutant recombinant viruses were tested by HI using antisera raised against a panel of genetically and antigenically variable viruses.

Koel *et al.* (2013) estimated the antigenic effect of introduced substitutions by comparing the position of mutant and parental viruses on an antigenic map generated from HI data, though there are indications that this method is also sensitive to changes in receptor-binding avidity that may be caused by the introduction of certain HA1 substitutions (Li *et al.*, 2013). Here, the HI results for mutant viruses, geometric mean titres averaged over a number of repeats, were considered relative to the titres of the parental virus. Differences in $\log_2$ HI titre associated with introduced substitutions were attributed to antigenic and non-antigenic changes in phenotype using Equation 3.1. As discussed in Chapter 3, this is done by attributing changes in titre associated with a substitution to antigenic effects if titres drop against antisera raised against viruses lacking the introduced substitution and increase titres against antisera raised against viruses sharing the introduced substitution. Changes in titre that are consistent across both these groups of antisera are considered non-antigenic and are likely to be the result of changes in receptor-binding avidity caused by the substitution. The allocation of antisera to these two categories is detailed in Appendix B, Table B.1. Titres below the lower threshold of the assay (i.e. the most concentrated sample of antisera tested did not inhibit erythrocyte agglutination) were excluded from this analysis.

4.4 Results

To identify the amino acid changes responsible for antigenic differences among A(H3N2) viruses and to quantify heterogeneities in their impact, variation in HI titres was modelled together with amino acid sequence data. The response variable throughout model development was $H_{r,v}$, the $\log_2$ titre recorded when a virus, v, was tested using antiserum raised against reference virus r. In common with A(H1N1) (Chapter 3), it was determined that terms for each virus, v, and antiserum raised against each reference virus, r should be included in the model ($p < 10^{-20}$) to account for consistent differences in titres between viruses and antisera respectively. These differences likely represent, amongst other factors, differences in virus receptor-binding avidity and reference virus immunogenicity respectively. Additionally an error term representing date of test had to be included ($p < 10^{-20}$). Day-to-day variability in titres may be due to variability in batches of erythrocytes and in dilutions of erythrocytes, viruses, and antisera which are typically made each day. Using these terms to account for non-antigenic variation in titre, amino acid difference at each variable surface-exposed HA1 position was tested as a predictor of variation in titre using

$$H_{r,v} = k_0 + I_r + A_v + k_1\alpha_1(r,v) + \epsilon_D + \epsilon_R \tag{4.1}$$

which is identical to Equation 3.2. The term $k_1\alpha_1(r,v)$ was used to test amino acid difference (or substitution) at specific amino acid positions. α_1 is an indicator variable that was 1 when reference virus, r, and test virus, v, possessed different amino acids, and zero otherwise. k_1 is the associated regression coefficient that represents the antigenic impact or expected drop in $\log_2$ titre when viruses are separated by substitution at the position being tested. In addition to terms representing each virus, A_v, antiserum raised against each reference virus I_r and date of test ϵ_D, k_0 is an intercept or baseline titre and ϵ_R represents residual measurement error not explained by the model.

Using Equation 4.1 substitutions at over 60% of variable, surface-exposed positions and 25% of non-surface-exposed were identified as being significantly correlated with antigenic change (reduced cross-reactivity in the HI assay) ($p < 0.05$) using a Holm-Bonferroni correction for multiple tests (Holm, 1979). As only surface-exposed HA1 positions have been identified as antigenically important in A(H3N2) (Skehel & Wiley, 2000), these percentages almost certainly reflect a high false positive discovery rate. Furthermore a significant correlation between the count of synonymous mutations between viruses and reduced cross-reactivity in the HI assay ($p < 10^{-10}$) caused by a strong correlation between molecular and antigenic evolution demonstrates how a regression analysis is predisposed to identify a significant number of false positives when applied to this problem.

4.4.1 Incorporating phylogenetic structure

Phylogenetic structure was incorporated in the model to reduce the described tendency for identification of false positives. Each branch was investigated as a predictor of reduced antigenic cross-reactivity apparent as a drop in HI titre when reference and test virus in HI are separated by the branch. Each branch was also tested as predictors of variation in the magnitude of HI titres associated with the use of particular reference or test viruses. These phylogenetic terms associated with branches of the phylogeny were tested using:

$$H_{r,v} = k_0 + I_r + A_v + \sum_{i \in F_{anti}} m_i\delta_i(r,v) + \sum_{i \in F_{immu}} n_i\zeta_i(r) + \sum_{i \in F_{avid}} p_i\zeta_i(v) + \sum_{i \in F_{anti'}} m'_i\delta_i(r,v) + \epsilon_D + \epsilon_R \tag{4.2}$$

which modifies Equation 4.1 by replacing the term representing substitution at a single amino acid position with various phylogenetic terms. $m_i\delta_i(r,v)$ terms are phylogenetic terms representing antigenic differences. As described in Chapter 3 and used in Equation 3.3, δ_i is an

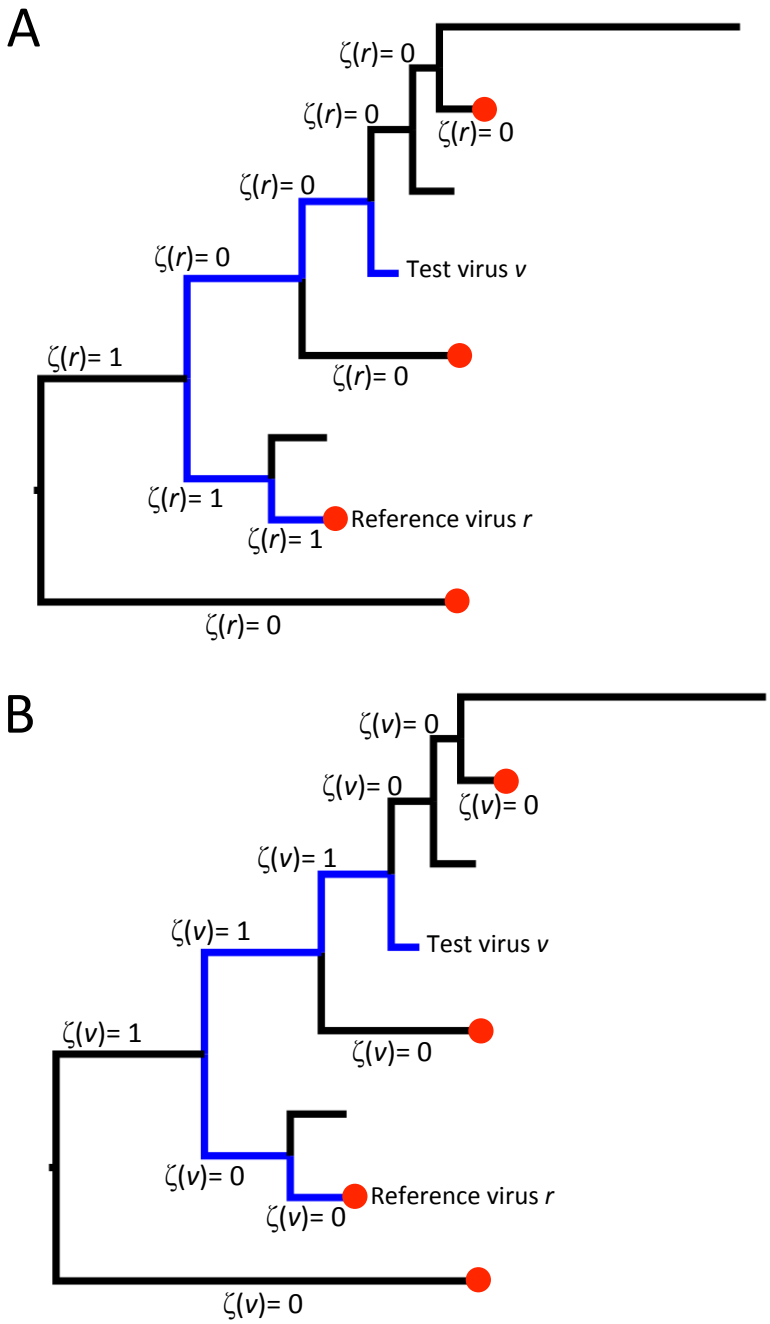

Figure 4.1: An example of a phylogeny including ζ terms introduced in Equation 4.2: The assignment of values 0 and 1 to $\zeta(r)$ (**A**) and $\zeta(v)$ (**B**) terms associated with branches of an example phylogeny, given the indicated positions of a reference virus, r and a test virus, v, is shown. Red circles indicate a set of four possible reference viruses that includes the reference virus used in this example. $\zeta()$ terms are assigned only to branches whose descendants include both antisera (reference viruses) and test viruses (i.e. branches that occur between the root of the tree and either of the four references strains). (**A**) $\zeta(r) = 1$ when the reference virus, r, is descended from a branch and $\zeta(r) = 0$ otherwise. (**B**) $\zeta(v) = 1$ when the test virus, v, is descended from a branch and $\zeta(v) = 0$ otherwise.

indicator variable and $\delta_i(r, v) = 1$ when reference virus (r) and test virus (v) are separated by branch i of the phylogeny and $\delta_i(r, v) = 0$ otherwise, with m_i being the associated regression coefficient. However, there are now terms to identify avidity or immunogenicity changes in the phylogeny. ζ_i is a second kind of indicator variable and $\zeta_i() = 1$ when a virus, either a reference virus r or a test virus v, is descended from branch i and $\zeta_i() = 0$ otherwise. $n_i\zeta_i(r)$ phylogenetic terms are associated with variation in the immunogenicity of reference viruses, or in other words the reactivity of antisera raised against them, descended from a particular branch, with n_i being the the associated regression coefficient. $p_i\zeta_i(v)$ phylogenetic terms are associated with variation in the receptor-binding avidity of test viruses descended from a particular branch, with p_i being the the associated regression coefficient. It is only possible to differentiate between these three phylogenetic effects for branches whose descendants include both antisera (reference viruses) and test viruses. For this set of branches, associated antigenic, avidity and immunogenicity effects are tested separately as F_{anti}, F_{immu} and F_{avid}. For branches where this was not true, phylogenetic terms in the set $F_{anti'}$ were tested. For branches in this set, $\delta_i(r, v)$ terms are the same as described above and m'_i is now the associated regression coefficient. In Figure 4.1, the allocation of $\zeta(r)$ and $\zeta(v)$ to branches of the same phylogeny represented in Figure 3.2 is shown. Figure 3.2 shows the assignment of values of 0 or 1 to $\delta_i(r, v)$ terms associated with each branch of this phylogeny.

In total, 281 phylogenetic terms were selected by a search that maximised AIC. 81 branches were identified as correlating with antigenic change and these become F_{anti} in Equation 4.3. 77 branches were were identified as explaining variation in HI titre explained by differences in antiserum reactivity and become F_{immu} in Equation 4.3. 109 branches were identified as explaining significant variation in HI titre caused by differences in virus reactivity and become F_{avid} in Equation 4.3. 14 branches were identified as being associated with variation in HI titre as a result of a change in either antigenicity or virus reactivity. As no viruses descended from these 14 branches were reference viruses used to generate antiserum in this dataset, it was not possible to resolve whether differences in titre observed when these viruses were tested were due to antigenic differences or the result of differences in receptor-binding avidity, relative to viruses outside the clades defined by these branches. These 14 branches become $F_{anti'}$ in Equation 4.3. Using maximisation of AIC tends towards the selection of false positives however this actually represents a conservative choice as these phylogenetic terms are being used as a control mechanism to prevent false positive identification of substitution terms.

4.4.2 Identification of substitutions affecting antigenicity

To identify antigenically important substitutions, variable, surface-exposed HA1 amino acid positions were tested alongside the 281 phylogenetic terms by the model:

$$H_{r,v} = k_0 + I_r + A_v + k_1\alpha_1(r,v) + \sum_{i \in F_{anti}} m_i\delta_i(r,v) + \sum_{i \in F_{immu}} n_i\zeta_i(r) + \sum_{i \in F_{avid}} p_i\zeta_i(v) + \sum_{i \in F_{anti'}} m'_i\delta_i(r,v) + \epsilon_D + \epsilon_R \quad (4.3)$$

which retains the 81 $m_i\delta_i(r,v)$ terms in F_{anti}, 77 $n_i\zeta_i(r)$ terms in F_{immu}, 109 $n_i\zeta_i(v)$ terms in F_{avid}, and 14 $m_i\delta_i(r,v)$ terms in $F_{anti'}$ selected using Equation 4.2. As in Equation 4.1, $k_1\alpha_1(r,v)$ terms represent amino acid substitution at a specific position. After accounting for multiple testing, amino acid differences at 20 positions (2, 62, 83, 124, 133, 138, 144, 145, 155, 156, 158, 172, 183, 189, 193, 197, 214, 217, 262 and 276) were found to be significantly correlated with variation in HI titres. As these positions were identified alongside phylogenetic terms selected using Equation 4.2, substitutions at each of these positions must be correlated with antigenic change in at least two locations in the phylogeny (there is *support across the phylogeny*). To identify specific amino acid substitutions with support across the phylogeny, $k_1\alpha_1(r,v)$ terms were re-examined using subsets of the data that included only viruses possessing the various possible combinations of two amino acid variants at each of these positions with support across the phylogeny. Specific amino acid substitutions identified as impacting on antigenic phenotype with support in multiple positions of the phylogeny are included in Table 4.1. The antigenic impact of substitutions in Table 4.1 is the absolute value of k_1 in Equation 4.3.

Table 4.1: HA1 amino acid substitutions correlated with antigenic change at multiple points in the evolution of A(H3N2) viruses

HA position	Antigenic site	Amino acid pair	Antigenic impact*
124	A	D-G	0.62
133	A	D-N	0.40
		N-S	0.51
138	A	A-S	0.29
144	A	I-N	0.55
145	A	D-N	0.45
		K-N	0.48
		N-S	0.50
156	B	E-K	0.42
158	B	E-K	0.67
		K-N	1.01
172	D	D-G	0.61
183	-	H-L	0.31
193	B	N-S	0.49
197	B	Q-R	0.69
214	D	I-T	0.36
217	D	I-V	0.56

*Estimated mean antigenic impact of substitution(s) associated with a particular branch measured in $\log_2$ HI titre (antigenic units).

Substitutions with support across the phylogeny were added to Equation 4.3 and the regression coefficients associated with phylogenetic terms were re-estimated. Antigenic phylogenetic terms ($m_i\delta_i()$ and $m'_i\delta_i()$ terms) were re-examined in this model. Any of these terms that remained associated with significant changes in antigenicity were not therefore explained, or completely explained, by substitutions with support across the phylogeny. To identify the amino acid substitutions responsible for the antigenic changes associated with these branches of the phylogeny, single phylogenetic branch terms were dropped in turn. As in Equation 3.5 in the previous chapter, to identify the most likely substitutions responsible for the antigenic change associated with the excluded branch, variable surface-exposed HA1 amino acid positions were retested as k' and α' in place of each excluded phylogenetic term using the equation:

$$H_{r,v} = k_0 + I_r + A_v + k'\alpha'(r,v) + \sum_{i \in S} k_i\alpha_i(r,v) + \sum_{i \in F_{anti}} m_i\delta_i(r,v) + \sum_{i \in F_{immu}} n_i\zeta_i(r) + \sum_{i \in F_{avid}} p_i\zeta_i(v) + \sum_{i \in F_{anti'}} m'_i\delta_i(r,v) + \epsilon_D + \epsilon_R \quad (4.4)$$

where terms k_i and $\alpha_i(r,v)$ represent each of the 17 substitutions with support across the

phylogeny (substitutions in Table 4.1). Substitutions identified using Equation 4.4 correlate with antigenic change at a single point in the evolutionary history of the virus and are also shown in Table 4.2. They are sorted by estimated antigenic impact. In several cases, multiple substitutions estimated to occur simultaneously in the evolution of the virus provide equally likely explanations for the antigenic change associated with a particular branch of the phylogeny. Among the substitutions in Table 4.2, those occurring at HA1 positions where there was support across the phylogeny for a role in antigenicity for substitutions in general (but not for the specific substitution) are marked ✓. For example, five different substitutions at position 189 were identified using Equation 4.4 and are present in Table 4.2 (N189K, K189Q, R189S, K189R and S189N). Each of these substitutions was identified as correlating with antigenic change at a single point in the phylogeny, so individually none was identified with support across the phylogeny, whereas substitution at position 189 in general was. Eighteen out of the 24 branches investigated had a least one potentially causative substitution that occurred at a position with support across the phylogeny.

The substitutions identified using Equation 4.4 included several substitutions identified by Koel *et al.* (2013) as being involved in major antigenic changes, causing transitions between clusters apparent on antigenic maps generated using HI data (T155Y, Y155H, K156Q, E158K, S159Y, K189Q and K189R — Table 4.2). This indicates that several of the most antigenically important A(H3N2) substitutions have occurred at only single points in the evolution of the virus, based on knowledge of sampled viruses. This suggests that the immune repertoire of the human population prevents the reversion to antigenic phenotypes, even when they were most recently circulating many years ago. Candidate substitutions identified for nine of the eleven points in the antigenic evolution of A(H3N2) where antigenic change is estimated to be over 1 antigenic unit included substitution(s) at at least one of the positions 155, 156, 158, 159 or 189 (all of which are in antigenic site B). This highlights the immunodominant role of amino acid positions forming an antigenic ridge bordering the upper lip of the primary receptor-binding site (Figure 4.2).

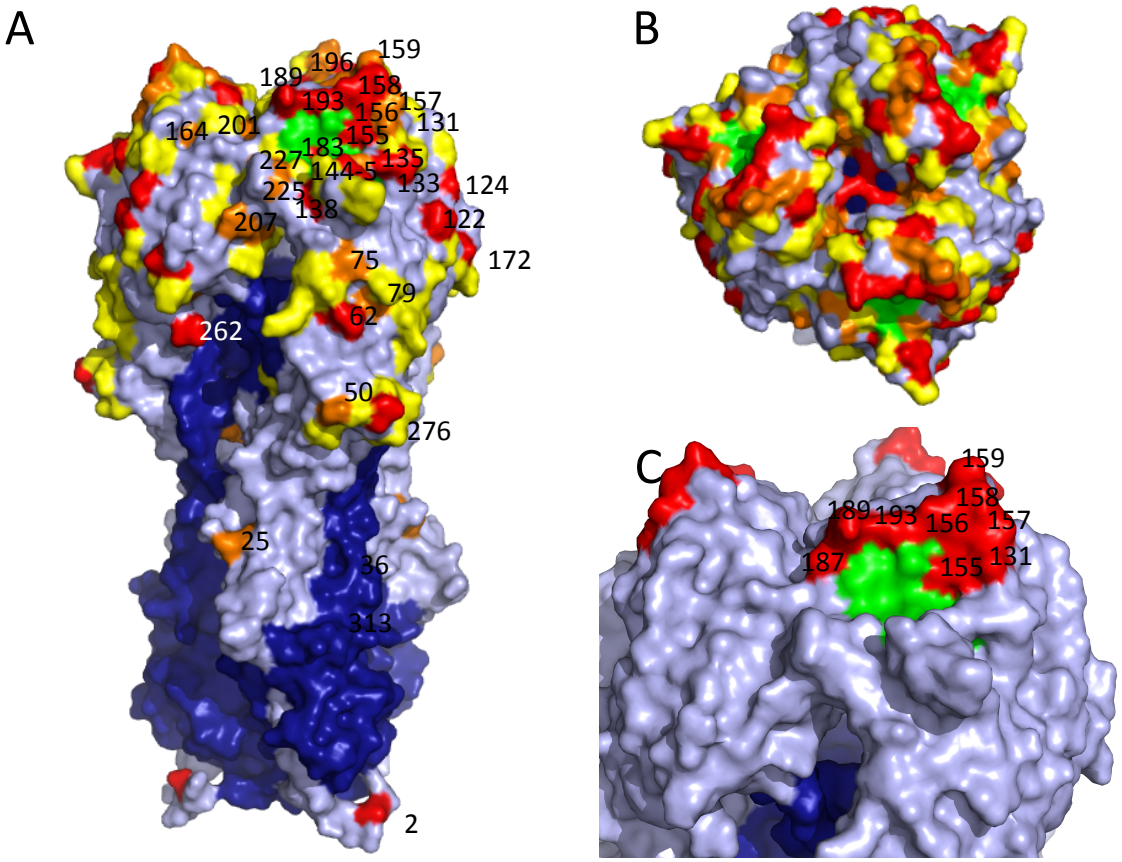

Figure 4.2: A(H3N2) HA positions implicated in antigenic change: HA1 positions where substitutions were identified as affecting antigenic phenotype shown on the HA structure of A/Aichi/2/68 (H3N2) (Protein Data Bank ID: 1HGG) (Sauter *et al.*, 1992) viewed side on (**A**) and from above (**B**). HA1 is shown in light blue and HA2 in dark blue. Amino acid positions where there is a significant correlation between substitution and antigenic change at multiple points in the evolution of the virus are shown in red. Positions where substitution is correlated with antigenic change in only a single part of the phylogeny are shown in orange. Remaining parts of the primary sialic acid binding site and antigenic sites A–E are shown in green and yellow respectively. (**C**) An exposed ridge at the upper boundary of the receptor-binding site comprised of residues at positions 131, 155, 156, 157, 158, 159, 187, 189, and 193 is shown in red.

Table 4.2: HA1 amino acid substitutions associated with antigenic change at single points in the evolutionary history of A(H3N2)

Antigenic weight associated with branch*	Candidate substitution(s)	Antigenic site	Support for position across phylogeny
2.68	L25I	-	
	H75Q	E	
	A131T	A	
	H155T	A	✓
2.55	K158N	B	✓
	N189K	B	✓
2.29	T122N	A	
	G144D	A	✓
	T155Y	B	✓
	R207K	D	
1.93	K62E	E	✓
	V144I	A	✓
	K156Q	B	✓
	V196A	B	
	N276K	C	✓
1.79	N262T	E	✓
1.77	I62M	E	✓
	F79L	-	
	S159N	B	
1.65	L164Q	B	
	R201K	D	
1.41	K189Q	B	✓
1.29	G135K	A	
1.19	R189S	B	✓
1.07	K307R	C	
	D144V	A	✓
0.88	K135T	A	
	N262S	E	✓
0.83	Y155H	B	✓
	K189R	B	✓
0.72	Y159F	B	
	S189N	B	✓
	S227P	D	
0.69	N2K	-	✓
	D144V	A	✓
0.66	D158E	B	✓
0.63	K158N	B	✓
0.62	K50R	C	
	E158G	B	✓
	M260I	E	
0.62	E62K	E	✓
	N144K	A	✓
0.62	S193F	B	✓
	D225N	-	
0.57	S159Y	B	
0.48	G135E	A	
0.43	T212A	D	
0.41	L157S	B	

*Estimated mean antigenic impact of substitution(s) associated with a particular branch measured in $\log_2$ HI titre (antigenic units).

4.4.3 Analysis of HI assays using recombinant viruses

Predictions of the antigenic impacts of substitutions identified as affecting antigenic phenotype were compared with the results of antigenic analyses of recombinant viruses possessing those identified substitutions. Koel *et al.* (2013) generated recombinant viruses possessing amino acid substitutions associated with transitions between clusters of antigenically similar viruses on antigenic maps generated from HI data collected using A(H3N2) viruses isolated over a similar period to this study. Koel *et al.* (2013) then assessed the antigenic impact of introduced substitutions by positioning recombinant viruses on antigenic maps among circulating viruses using data generated from HI assays. The estimated antigenic impacts of these experimentally introduced substitutions were compared with the predicted antigenic effects of substitutions in Tables 4.1 and 4.2 when these were added together into a single model lacking the branch terms included in models used for identification of substitutions.

Substitutions identified in Tables 4.1 and 4.2 that were also introduced into recombinant viruses by Koel *et al.* (2013) are shown in Table 4.3. To determine the antigenic and non-antigenic impact of introduced substitutions, Equation 3.1 was applied to the relevant HI data. ΔH_1 is the mean change in log titre associated with each substitution averaged across HIs measured using antisera from viruses lacking the introduced substitution. An antigenic change associated with substitution is expected to reduce cross-reactivity or antigenic similarity to these viruses so ΔH_1 is expected to be negative. Conversely, the antigenic impact of a substitution is expected to increase cross-reactivity or antigenic similarity to viruses sharing the introduced substitution so the change in $\log_2$ titre measured against these viruses, ΔH_2, is expected to be positive when a substitution affects antigenicity. A substitution affecting antigenicity can therefore decrease or increase antigenic similarity to reference viruses, and therefore HI titre, depending on whether those reference viruses lack or share the introduced substitution. The antigenic effect on $\log_2$ titre of substitutions, ΔH_A, is therefore reported without a sign in Table 4.3, except in cases where the change in titre is the opposite of what is expected when an antigenic change is observed (i.e. $\Delta H_1 > \Delta H_2$). In these instances where mutating away or towards a reference strain increases or decreases HI titre respectively, changes in titre are assumed to be the result of experimental variability. In these cases, a negative sign is shown with the estimated antigenic change (ΔH_A). A consistent decrease or increase in $\log_2$ titre associated with the introduction of a substitution that is measured using antisera raised against viruses both lacking and sharing the introduced substitution cannot be considered the result of an antigenic change and these effects are considered to be the result of a non-antigenic impact of a substitution, ΔH_N.

Table 4.3: Antigenic and non-antigenic impacts of H3 HA1 substitutions introduced into recombinant viruses

HA1 position	Introduced substitution	Mutagenesis background*	Observed effect†			
			ΔH_1	ΔH_2	ΔH_A	ΔH_N
122	T122N	HK68REC	+0.14	+0.84	0.34	+0.49
133	N133S	TX77REC	+0.04	-0.14	-0.09	-0.05
135	G135E	SI87REC	-0.46	+1.18	0.81	+0.36
144	G144D	HK68REC	-0.03	+1.58	0.80	+0.78
145	N145K	SI87REC	-1.98	+0.63	1.31	-0.68
	N145K	BE92REC	-1.92	+1.93	1.93	+0.01
	K145N	BE89REC	+1.25	+2.83	0.79	+2.04
	K145N	WU95REC	-1.08	+2.95	2.01	+0.93
	K-N mean				1.51	
155	T155Y	HK68REC	-2.95	+1.06	2.01	-0.95
	Y155T	EN72REC	-1.14	+3.20	2.17	+1.03
	T-Y mean				2.09	
	H155T	SY97REC	+1.15	+1.48	0.16	+1.32
	T155H	FU02REC	-2.24	-0.39	0.92	-1.31
	H-T mean				0.54	
	Y155H	BK79159SY189KR	-1.34	-0.15	0.59	-0.74
156	K156E	TX77REC	+0.69	+2.04	0.67	+1.36
	K156E	BE92REC	-1.10	+2.76	1.93	+0.83
	E156K	BK79REC	-1.62	-0.52	0.55	-1.07
	E156K	SI87REC	-0.11	-2.58	-1.24	-1.34
	E-K mean				0.48	
	K156Q	WU95REC	-2.10	+0.60	1.35	-0.75
	Q156K	SY97158KE	-0.52	+1.61	1.07	+0.53
	K-Q mean				1.21	
158	G158E	VI75REC	-1.52	-0.02	0.75	-0.77
	E158K	WU95156KQ	0.20	+1.00	0.40	+0.60
	K158E	SY97REC	-1.10	+2.26	1.68	+0.58
	E-K mean				1.04	
159	S159Y	TX77156KE	+1.10	+1.55	0.22	+1.32
	S159Y	BK79155YH189KR	-1.25	+1.34	1.30	+0.05
	S-Y mean				0.76	
164	L164Q	EN72REC	-0.07	+0.22	0.15	+0.07
172	G172D	TX77REC	-0.14	0.01	0.08	-0.07
189	Q189K	EN72REC	-0.58	+5.00	2.79	+2.21
	K189Q	VI75REC	-4.08	+0.04	2.06	-2.02
	K-Q mean				2.43	
	K189R	BK79155YH159SY	-4.73	-0.32	2.21	-2.52
	R189S	SI87REC	-0.55	-0.58	-0.02	-0.57
	S189R	BE92REC	-0.34	+0.53	0.44	+0.10
	R-S mean				0.21	
193	N193S	SI87REC	0.00	-0.01	0.00	0.00

HA1 position	Introduced substitution	Mutagenesis background*	Observed effect† ΔH_1	ΔH_2	ΔH_A	ΔH_N
	S193N	BE92REC	-1.51	-0.28	0.89	-0.61
	N-S mean				0.45	
196	V196A	WU95REC	-0.69	-0.61	0.04	-0.65
197	Q197R	TX77REC	0.44	-0.10	-0.27	+0.17
207	R207K	HK68REC	+0.24	+1.07	0.42	+0.66
217	I217V	EN72REC	+0.14	+0.22	0.04	+0.18
	V217I	TX77REC	+0.04	+0.22	0.09	+0.13
	I-V mean				0.07	
260	M260I	VI75REC	+0.67	+0.14	-0.27	+0.41
262	N262T	BE92REC	+0.10	+0.05	0.03	+0.08
276	N276K	WU95REC	-0.50	+0.18	0.34	-0.16

* Mutagenesis backgrounds labels are taken from Koel *et al.* (2013) and refer to antigenic clusters defined by Smith *et al.* (2004).

† Mean observed changes in $\log_2$ HI titre (in antigenic units) against antisera lacking (ΔH_1) and sharing (ΔH_2) the introduced substitution are partitioned into antigenic (ΔH_A) and non-antigenic (ΔH_N) effects.

The antigenic and non-antigenic impacts of laboratory-tested substitutions identified in Tables 4.1 and 4.2 are shown in Table 4.3. For each substitution introduced into multiple mutagenesis backgrounds, antigenic effects were averaged across mutagenesis backgrounds. These mean, observed antigenic effects are plotted against the antigenic impacts of substitutions estimated by modelling in Figure 4.3. The predicted and observed antigenic impacts of substitutions are significantly positively correlated ($p < 0.05$), however there is a considerable degree of variation around this relationship and this is reflected by a relatively low coefficient of determination ($R^2 = 0.37$). The spread of antigenic impacts around the observed means is also high. This demonstrates the degree to which the observed antigenic impact of substitution on HI titre depends on the existing genetic and antigenic differences between the parent virus and the viruses used to generate antiserum.

Equation 3.1 was also used to estimate the non-antigenic impact of substitutions (ΔH_N in Table 4.3). Smaller differences in this value could represent differences in the dilution of parental and mutant viruses prior to HI testing given the limited accuracy of the haemagglutination inhibition assay (> 0.78 — see Chapter 3, section 3.3.5). Larger values likely reflect differences to the viral receptor-binding avidity caused by the introduced substitution. Non-antigenic impacts exceeding 0.78 were found to occur as a result of substitution at positions 145, 155, 156, 158, 159 and 189 (Table 4.3). If a substitution resulted in lower HI titres as a result of a change to receptor binding-avidity, the reverse substitution would be expected to increase titres by the same amount, and vice versa. Across the amino acid differences with non-antigenic impacts exceeding 0.78 where substitutions were tested in both directions

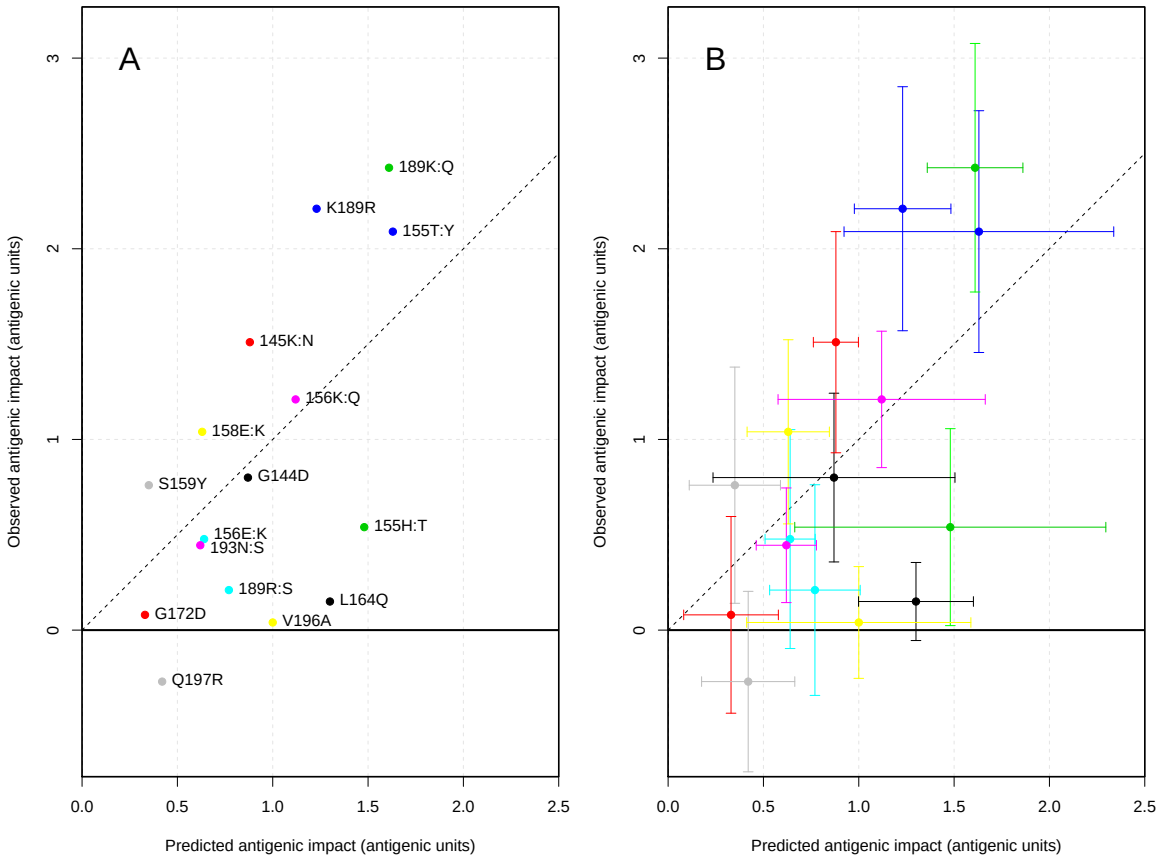

Figure 4.3: Observed and predicted antigenic impacts of H3 HA1 amino acid substitutions: The mean antigenic impact of each substitution predicted from modelling (Tables 4.1 and 4.2) plotted against the mean observed impact averaged across antisera in the panel (Table B.1) and mutagenesis backgrounds (Table 4.3). For each substitution, a point shows the observed mean antigenic impact (ΔH_A) associated with a change in HI titre (averaged across mutagenesis backgrounds and antisera) plotted against the antigenic effect predicted by modelling. A negative observed antigenic impact indicates a change in HI titre in the opposite direction to that predicted. Substitutions are labelled in (**A**) and shown with 95% confidence intervals in (**B**). Substitutions in (**A**) and (**B**) are coloured to ease comparison; colours are assigned randomly and recycled. (**A**) Labels indicate if substitutions were tested in a single direction (e.g. K189R) or in both directions (e.g. 189K:Q). (**B**) 95% confidence intervals are shown for both predicted and observed antigenic effects.

(e.g. H155T and T155H), substitution was found to increase titre when introduced in one direction and decrease titre when the reverse was introduced. The non-antigenic impacts of substitutions in each of these pairs, and in particular those at positions 155, 156 and 189, were of similar magnitudes (Table 4.3 and Figure 4.4). While the negative correlation between the estimated non-antigenic effects of K145N and N145K is not as strong, the relationship holds in that one is associated with an increase in titre while the other is associated with an drop in titre, albeit a smaller one. This negative correlation supports the hypothesis that the observed effects on HI titre are due to genuine differences in viral phenotype, as these effects would presumably be drawn randomly from some underlying distribution otherwise.

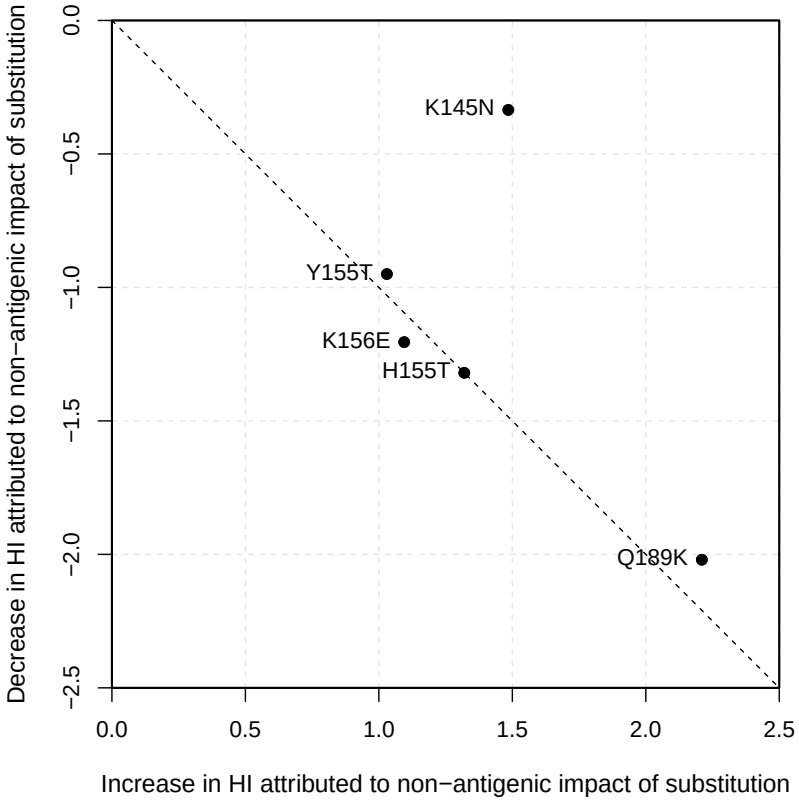

Figure 4.4: Non-antigenic impacts of paired amino acid substitutions on HI titres: Estimated impact on $\log_2$ HI titre of non-antigenic changes in viral phenotype caused by the introduction of paired substitutions (e.g. H155T and T155H). This is the consistent change in HI titre for a mutant virus relative to its parent virus, averaged across groups of antisera lacking and sharing the introduced substitution. Paired substitutions are labelled with the substitution whose non-antigenic impact resulted in an increase in titre (i.e. H155T rather than T155H).

4.5 Discussion

By attributing antigenic variation in HI titres to HA1 amino acid differences between viruses, a detailed quantitative analysis of the determinants of antigenic drift of human influenza

A(H3N2) viruses isolated during the period 1968-2014 is presented. In addition to several previously identified determinants of major antigenic changes that have been incorporated into the trunk lineage during the evolution of the virus, the described methods also identify further substitutions of varying antigenic impact by incorporating information on substitutions occurring in multiple places in the phylogeny including those occurring in side branches that may occur only at low frequency within the viral population. While there is notable uncertainty in the identification of antigenic determinants in several cases where the antigenic impact of co-occurring substitutions could not be discriminated, nearly all identified substitutions have been experimentally validated or occur at positions belonging to the main antigenic sites defined based on the results of monoclonal antibody escape mutant studies, indicating the usefulness of this approach. I am now working as part of a collaboration involving the authors of recently developed methods for predicting the evolutionary success of competing viral lineages in the A(H3N2) population (Łuksza & Lässig, 2014) to assess how these could benefit from our ability to quantify the relative antigenic importance of substitutions occurring in the HA1 domain.

Estimated antigenic impacts of substitutions identified by modelling were compared with the published results of HI assays performed using mutant recombinant viruses possessing several of the identified substitutions. Changes in HI titre associated with the introduction of substitutions were partitioned into antigenic and non-antigenic effects. Ideally the antigenic impact of an introduced substitution would be assessed using antisera raised against both the parental and mutant viruses, though this is rarely feasible. In the absence of these antisera, the antigenic impact of an introduced substitution must be estimated using HI titres measured against a panel of antisera that would preferably include antisera raised against reference viruses both lacking and sharing the introduced substitution, as I did in Chapter 3.

The antisera used by Koel *et al.* (2013) were raised against viruses spread more widely across the phylogeny. This increases the range of genetic and antigenic differences between the viruses used for comparison which is likely to result in the estimated antigenic impacts of substitutions in Table 4.3 being less reliable as other substitutions are likely to have occurred at other positions in epitopes which may affect inference. Koel *et al.* (2013) reported the antigenic distances as between the parental virus and the mutant possessing the introduced substitution on an antigenic map. While antigenic distances estimated by this method may be influenced by changes to receptor-binding avidity caused by the substitution (Li *et al.*, 2013), they may better represent antigenic impact of substitutions under these experimental conditions. These observations demonstrate how difficult it can be to quantify the antigenic impact of specific genetic changes, even when the substitution causing the change in antigenic phenotype is known.

Substitutions tested by Koel *et al.* (2013) at six positions were identified as having a non-

antigenic impact on HI titres. Residues at five of these six positions (155, 156, 158, 159 and 189) form an exposed ridge above the primary receptor-binding site (Figure 4.2). Changes in HI titre due to altered receptor-binding avidity caused by substitutions at positions in the 150-loop have been described in former seasonal A(H1N1) (Hensley *et al.*, 2009) and in A(H1N1)pdm09 (Liu *et al.*, 2010). At the final position among the set of six, 145, the negative correlation between the substitutions K145N and N145K was not as strong as between the other pairs of substitutions highlighted, however changes in HI titre caused by a change in receptor-binding avidity have previously been attributed to the substitution N145K which also influenced clustering of viruses on antigenic maps of A(H3N2) viruses (Li *et al.*, 2013). These six positions were also among the seven identified by Koel *et al.* (2013) as being substitutions causing the major changes in antigenic phenotype of A(H3N2) viruses in the period 1968–2003. That these positions are identified as having affected both binding and antigenic phenotype could be supportive of the important role of changes in receptor-binding avidity in influenza A antigenic drift (Hensley *et al.*, 2009). Alternatively, it could demonstrate the influence that changes in receptor-binding avidity have on virus clustering on antigenic maps generated using HI data. This feature of antigenic maps has been described previously (Li *et al.*, 2013) and may have implications for which "kind" of antigenic changes are likely to be identified using methods based around the construction of antigenic maps.

In this application of this approach to this A(H3N2) dataset there is increased uncertainty, relative to the A(H1N1) analysis described in Chapter 3, in the both the identification of substitutions causing antigenic drift and in the estimation of antigenic impact. This is at least in part because the viral population is sampled more sparsely here, relative to Chapter 3, with fewer viruses isolated over a longer period included. This has resulted in substitutions, that may occur in different branches of the phylogeny with denser sampling, appearing to have occurred simultaneously. Substitutions that achieved only low frequency in the virus population are also more likely to be missed with sparser sampling. Computational limitations informed the decision to use the sparser sampling here. A more CPU intensive Bayesian alternative to the methodology described here, which similarly attributes variation in pairwise measurements of antigenic cross-reactivity to sequence differences while incorporating phylogenetic information in a single step, and as a result less conservatively, is currently being developed and has successfully been applied to smaller foot-and-mouth disease virus datasets (Davies *et al.*, 2014). Ongoing work is currently speeding up this algorithm so it will be viable on larger datasets like the ones studied here. Using statistical methods, it is difficult to differentiate between the antigenic impact of co-occurring substitutions that do not show significantly different patterns of substitution elsewhere in the phylogeny. Challenges in identifying a definitive set of genetic determinants of antigenicity using statistical methods call for a method that accounts for model uncertainty when using viral sequence to predict antigenicity. This is explored using a method of Bayesian model averaging in Chapter 5.

CHAPTER 5

Quantifying and predicting antigenic relationships among influenza A and foot-and-mouth disease viruses

Quantifying and predicting antigenic relationships among influenza A and foot-and-mouth disease viruses

5.1 Abstract

Maintaining an efficacious influenza vaccine in the face of rapid antigenic drift requires a global surveillance effort. This involves the antigenic characterisation of circulating viruses and anticipation of which strains will predominate in future epidemics. Various modelling approaches have used data from antigenic assays, such as the haemagglutination inhibition (HI) assay, to characterise antigenically variable viruses, to quantify and predict antigenic similarity, and to explore the relationship between genetic and antigenic evolution (e.g. Bedford *et al.*, 2014; Sun *et al.*, 2013). In previous chapters I have described methods for the identification of genetic drivers of antigenic change in influenza A viruses. In this chapter I test the ability of sequence-based models that incorporate these genetic determinants to predict antigenic relationships, as expressed in the HI assay, among viruses of influenza subtypes A(H1N1) and A(H3N2). I compare predictive performance against antigenic cartography, showing the sequence-based approach to be similarly capable of the relatively unchallenging task of predicting unobserved HI titres and better able to predict titres for potential vaccine viruses. A further advantage of the sequence-based approach is the ability to predict antigenic relationships for uncharacterised, novel viruses from their amino acid sequence. Under this scenario, I show that substitutions identified as having low antigenic impact are a critical component of virus cross-reactivity and that by including these in addition to the high-impact substitutions often focused on, the accuracy of predicting antigenic phenotypes of emerging viruses from genotype was improved dramatically. By applying the same methods to viruses of three antigenically distinct serotypes of foot-and-mouth disease virus (FMDV) I demonstrate the versatility of the approach presented. Accurate sequence-based prediction of antigenic phenotype has the potential to inform choices of reference viruses used to generate antiserum, the targeting of antigenic assays, and predictions of the evolutionary success of different genotypes. This could ultimately help to inform the vaccine selection process.

5.2 Introduction

Antigenic drift of seasonal influenza viruses presents a significant challenge to the maintenance of effective influenza vaccines and requires ongoing antigenic characterisation of the virus population. The HI assay is the most widely used method of antigenic characterisation, however interpretation of the results of HI assays are complicated by a variety of non-antigenic factors that impact titres (Archetti & Horsfall, 1950; Ndifon, 2011). These complicating factors hamper attempts to quantify the antigenic impact of specific genetic changes and consequently, current methods for predicting the evolutionary future of genotypes within the influenza population do not account for differences in the antigenic impact of HA substitutions (Łuksza & Lässig, 2014; Neher *et al.*, 2014). In Chapters 3 and 4 I describe methods for the identification of amino acid substitutions that affect the antigenic similarity of influenza A(H1N1) and A(H3N2) viruses. In this chapter I investigate the predictive power of these genetic determinants of antigenicity. As discussed in previous chapters, uncertainties in the identification of the specific genetic changes responsible for differences in antigenic phenotype arise when substitutions occur at the same (or almost the same) position in the phylogeny. Rather than attempting to converge on an optimal set of genetic determinants, here I use a form of Bayesian model averaging (BMA) that explicitly incorporates uncertainty in the identification of causative substitutions and reflects this uncertainty in predictions of titre (Hoeting *et al.*, 1999).

5.2.1 Foot-and-mouth disease virus

To further investigate the predictive performance of the approaches presented, and to ascertain their generalisability, antigenic relationships among viruses of FMDV serotypes A, O and SAT1 were also investigated. FMDV is a single-stranded, positive-sense RNA virus that together with Equine rhinitis A virus constitutes the genus *Apthovirus* of the family *Picornaviridae* (Fry *et al.*, 2005). The $\sim$ 8.2 kb genome encodes a single polyprotein that is cleaved post-translationally by viral proteases to yield structural and non-structural proteins. The virion is roughly spherical with a diameter of $\sim$ 30 nm and is, unlike influenza, non-enveloped. The capsid comprises 60 copies of four proteins VP1, VP2, VP3, and VP4 arranged in a pseudo T=3 icosahedral lattice (Figure 5.1). VP1, VP2, and VP3 are wedge-shaped eight-stranged β-sandwiches that fit together and are exposed on the capsid surface while VP4, together with the N-termini of VP1 and VP3, is located at the interior of the capsid (Fry *et al.*, 2005).

FMDV causes a highly contagious disease that predominantly affects animals of the order *Artiodactyla*. Globally, foot-and-mouth disease (FMD) is one of the most economically important diseases of livestock with the primary hosts being cattle, water buffalo (*Bubalus bubalis*) and small ruminants although there are also significant levels of virus circulation in

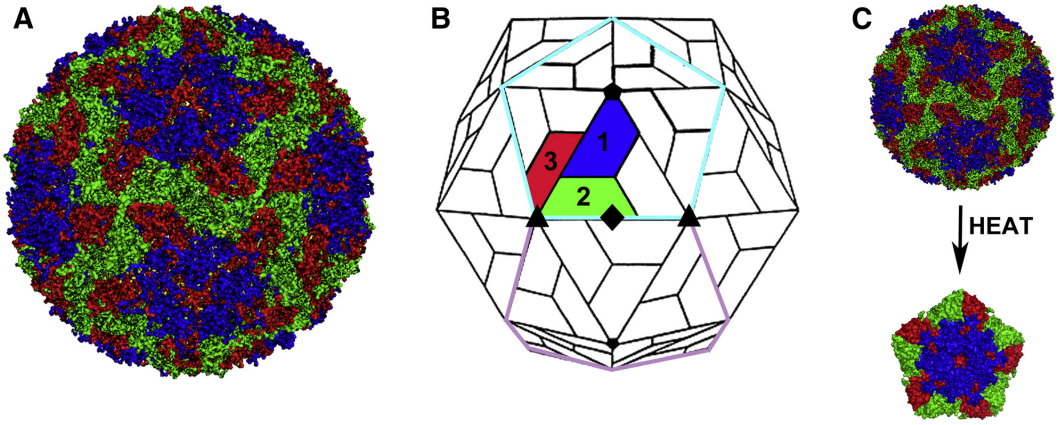

Figure 5.1: Overview of the structure of the FMD virus particle: Surface representations of the FMD capsid. VP1, VP2, and VP3 are coloured blue, green, and red respectively. VP4 is internal and is not shown. (**A**) Atomic model of the C-S8c1 crystal structure (Lea *et al.*, 1994). (**B**) Schematic of the capsid architecture with one protomer shaded and numbers to denote VP1, VP2, and VP3. Five protomers form a pentamer (delimited by cyan lines). The neighbouring pentamer below is delimited by violet lines. Axes of symmetry are also indicated (five-fold, pentamer; three-fold, triangle; two-fold, diamond). The psuedo T=3 icosahedral lattice is comprised of 60 protomers (12 pentamers). (**C**) FMD virion and one of the pentamers into which it dissociates are represented to scale. Dissociation can occur due to a variety of factors that included elevated temperatures. Image reproduced as appears in Figure 1 in Rincon *et al.* (2014) with permission from the rightsholder.

wildlife, particularly in the African buffalo (*Syncerus caffer*) (Thomson *et al.*, 2003; Vosloo *et al.*, 1996). There are seven antigenically distinct FMDV serotypes, namely A, C, O, Asia 1, and Southern African Territories (SAT) 1, 2, and 3. While individual vaccines may protect against large groups of genetically diverse viruses within some serotypes, there are still antigenically distinct groups within all serotypes with little or no antigenic cross-reactivity among them. Antigenic variability between and within serotypes prevents the production of a universally protective vaccine capable of protecting against all FMD viruses, or even all viruses of the same serotype (Paton *et al.*, 2005).

Inactivated vaccines play an important role in the control of FMD, and though the quality of the vaccine used is probably the most important factor for the effectiveness of vaccination, a sufficient degree of antigenic similarity between vaccine virus and circulating strains is also considered essential (Paton *et al.*, 2005). In common with influenza, antibodies play a critical role in the host response to FMDV infection (Alexandersen *et al.*, 2003; Juleff *et al.*, 2009) and generally, levels of neutralising antibody correlate well with protection observed *in vivo* (Pay & Hingley, 1987; Van Bekkum, 1969). Serotypes do vary in degree of antigenic variability that they exhibit, with viruses of serotype A showing particularly high levels of genetic and antigenic variability that reduce cross-reactivity and hinder control by vaccination. The antigenic similarity of FMD viruses can be assessed using the virus neutralisation (VN) test. VN titres give an *in vitro* measure of the whether the epitopes targeted by neutralising

antibodies remain sufficiently similar to cross-react.

In common with the HI assay, VN tests are affected by several non-antigenic factors (Reeve *et al.*, 2010). Variation in titres can occur due to variability in other characteristics of viruses such as receptor-binding avidity. Similarly variation in the magnitude of titres measured using antisera raised against different vaccine or reference strains is observed. Sources of experimental variation include potential differences in the immune response of individual animals used to generate antiserum for use in the assay, differences between batches of cells, and limitations in the accuracy of dilutions of reagents used. In contrast to the HI assay used to characterise human influenza viruses, we have the advantage of being able to work with antiserum drawn from the definitive host (cattle) when using VN to assess antigenic similarity of FMD viruses.

5.2.2 Prediction

There are various challenges faced by traditional serological approaches used for the antigenic characterisation of viruses such as influenza or FMDV. These are accentuated by the fact that both HI and VN titres are affected by a variety of non-antigenic factors which hinder the interpretability of the results of these assays and complicate the process of making decisions regarding vaccination. Computational methods have the potential to aid and inform these important decisions. For example, antigenic cartography is routinely used to visualise antigenic relationships among circulating influenza viruses and is incorporated into reports containing vaccine virus recommendations (Barr *et al.*, 2010, 2014; Klimov *et al.*, 2012). Emerging techniques have demonstrated accurate predictions of changes in frequency of genotypes within the A(H3N2) population between epidemic seasons (Łuksza & Lässig, 2014; Neher *et al.*, 2014), presenting an opportunity to enhance vaccine selection decisions that must be made up to a year in advance of anticipated epidemics. In this chapter I investigate the usefulness of sequence-based models for predicting antigenic relationships between viruses (using HI and VN titres as surrogate measures). Model cross-validation typically involves repeatedly dividing the data into training and test datasets chosen at random, parameterising the model using the training data and testing predictions for the test data. In this chapter I assess the ability of sequence-based predictive models to answer a variety of different, but related, questions that are relevant to the vaccine decision-making process by varying the composition of data subsets used to train and test models.

Both HI and VN assays are used to measure the antigenic similarity of pairs of viruses. Decisions regarding vaccination require the wider viral population to be antigenically characterised using available data on the pairwise similarity of sampled virus isolates. Typically only a small fraction of pairwise measurements among all relevant, isolated viruses are made making this task even more difficult. Here, by parameterising sequence-based and Bayesian

multidimensional scaling (BMDS) antigenic cartography models using training datasets lacking a proportion of virus:antiserum pairs and predicting titres for those missing pairs, the suitability of these methods for predicting unobserved antigenic relationships and aiding this process was assessed.

Due to limitations on the availability of animals for the generation of antiserum, the number of virus strains used to raise antiserum is typically a major limiting factor on the proportion of pairwise antigenic relationships that can be investigated experimentally. Given that the ability to characterise the viral population necessarily depends on the range of antigenic phenotypes represented in the panel of antisera used, the ability to predict antigenic relationships for antiserum before it is produced is useful. By predicting titres for antisera excluded from the data used to parameterise models, the power of sequence-based and antigenic cartography models for predicting antigenic coverage of new reference or vaccine viruses was also assessed.

By excluding viruses from the training datasets entirely, the accuracy of sequence-based prediction of titres for non-antigenically characterised viruses was investigated. Predictions of this kind are not possible using antigenic cartography because viruses must be characterised experimentally using some antisera before they can be positioned in antigenic space. Predicting titres for non-antigenically characterised viruses is potentially a more difficult task than those described above. This is partly because in the absence of any serological data for a virus it is not possible to estimate the impact of receptor-binding avidity on titres. The ability to predict antigenic phenotype from viral sequence data is however a desirable goal, and has implications for vaccine matching and for the informed targeting of laboratory-based antigenic assays.

If the antigenic similarity of emerging viruses to existing viruses or vaccines can be inferred from sequence, this also has the potential to inform predictions of viral fitness, epidemic severity, and the likely evolutionary success of genotypes. Specifically for influenza, we therefore examine the power of sequence-based approaches for predicting titres for emerging viruses using models trained using data collected in previous years. Under this scenario, I formed comparisons with predictions made using sequence-based models with terms for only the substitutions causing the largest changes in antigenic phenotype which are often focussed on. The ability to form sequence-based predictions of the antigenic novelty of viruses circulating in future influenza seasons has implications for the theoretical foundations of vaccine selection.

5.3 Materials and Methods

5.3.1 Influenza data

The influenza A HI and genetic sequence data were introduced in Chapter 2. The identification of amino acid positions at which substitution is associated with antigenic drift is described for subtypes A(H1N1) and A(H3N2) in Chapters 3 and 4 respectively. These analyses also identified other sources of experimental and non-antigenic variation in HI titres.

5.3.2 Foot-and-mouth disease virus data

The FMD viruses investigated were either supplied by the World Reference Laboratory for Foot-and-Mouth Disease at the Pirbright Institute, UK (serotypes A and O) or form part of the virus databank at the Transboundary Animal Diseases Programme (TADP), Onderstepoort Veterinary Institute, South Africa (serotype SAT1). Antigenic relationships among viruses of each serotype were determined by VN test. Viral sequence and VN data were generated for serotypes A and O by researchers at the Pirbright Institute and for serotype SAT1 by researchers and staff at the Onderstepoort Veterinary Institute. These data have been published by Reeve *et al.* (2010) (SAT1), Bari *et al.* (2014) (A), and Reeve *et al.* (2016) (under revision) (O, SAT1).

VN test datasets included titres for 56 A, 77 O and 42 SAT1 viruses and cattle antisera raised against seven A, five O and five SAT1 strains. For serotypes A and O, post-vaccination bovine antisera were generated by vaccinating cattle with 10 micrograms of vaccine antigen per dose after sublimating with montanide ISA 206 (Sepic, France). 21-day or 28-day post-vaccination antisera were pooled from five animals vaccinated with each virus. Two serotype O antisera (O Manisa/69 and O TUR/5/2009) were procured from Intervet, Germany. For SAT1, a group of five cattle were inoculated intradermolingually with 10^4 $TCID_{50}$ of each virus and convalescent sera were taken from cattle at 28 days post-infection. SAT1 antisera from individual animals were then used for VN tests.

Datasets consisted of 929 observations for 371 pairs of virus and antiserum for serotype A, 768 observations for 334 pairs for O and 1809 observations for 153 pairs for SAT1. Two-dimensional virus neutralisation tests were also carried out following established methodology (Rweyemamu *et al.*, 1978). Antibody titres were calculated from regression data as the $\log_{10}$ reciprocal antibody dilution required for 50% neutralisation of 100 tissue culture infective doses of virus ($\log_{10}SN_{50}$/100 $TCID_{50}$). For SAT1, all the serological tests were repeated on different dates between one and five times and carried out independently by three different individuals (operators). Date of test was also recorded for A and O and operator was recorded

for O, while all A tests were completed by the same individual.

Viral RNA was extracted from cell culture adapted isolates and reverse transcription (RT) polymerase chain reaction (PCR) was used to amplify capsid encoding genes and nucleotide sequencing was performed as previously described (Reeve *et al.*, 2010; Upadhyaya *et al.*, 2014). The P1 region of the viral genome was then sequenced and consensus nucleotide sequences were used to determine the amino acid sequence of the full capsid (VP1-4) for each reference strain and test virus present in the VN dataset. P1 nucleotide sequences aligned using ClustalW v2.1 (Thompson *et al.*, 1994) were used to generate phylogenetic trees for the viruses in each serotype using BEAST v1.7.2 (Drummond *et al.*, 2012).

Antigenic determinants were identified by attributing variation in VN test titres to amino acid differences at surface-exposed positions of the capsid using models similar to those presented in Chapters 3 and 4. These analyses are described for serotypes O and SAT1 by Reeve *et al.* (2016) (under revision) and for serotype A by Bari *et al.* (2015) and the amino acid positions identified by those studies were considered here.

5.3.3 Underlying titres

HI and VN titres are frequently affected by experimental variability and other forms of non-antigenic variability. By identifying variables that explain this variation and quantifying their effect on titres, it was possible to generate fitted underlying titres with the explained variation removed. An underlying titre, with experimental sources of variation removed, was estimated for each virus-antiserum pair present in the dataset. These underlying titres are the estimated *true* titres for each virus-antiserum pair and these, rather than the raw titres, were used to validate predictive models. For each dataset, various experimental variables were tested and slightly different combinations of these were identified as explaining significant variation in titre. The general structure of the model fitted to the data is described by:

$$H_{r,v} = I_r + A_v + D_{r,v} + \epsilon_T \tag{5.1}$$

where $H_{r,v}$ is the $\log_2$ HI or VN titre for virus v and antiserum raised against reference virus r. There are effects for reference strain immunogenicity (I_r) and test virus avidity (A_v) which reflect differences in the magnitude of titres when particular antisera or viruses are used irrespective of antigenic similarity to other strains. In addition to these two terms, there is a term for each combination of reference strain and test virus ($D_{r,v}$), which describes their antigenic relationship. This term reflects the extent to which titres for reference strain r and test virus v are lower than would otherwise be expected for these two viruses as a result of reduced antigenic similarity and can not take a positive value.

The term ϵ_T in Equation 5.1 represents the total error in the model. Some of this error was attributed to variation in titre caused by experimental variability, with the specific terms used to do so varying between datasets. For influenza A(H1N1) and A(H3N2) and for FMDV serotypes A and O, $\epsilon_T = \epsilon_d + \epsilon$ where ϵ_d is a term to absorb day-to-day variability that can take a different value for each date, d, on which titres were measured. In HI assays day-to-day variability can arise due to variability in batches of erythrocytes, and in dilutions of erythrocytes, antisera, and viruses. ϵ represents residual error not explained by any other parameter in the model. Day-to-day variability can arise in VN tests as a result of differences in cells and in dilutions of viruses and antiserum. For FMDV serotype SAT1 where tests were carried out by multiple operators on multiple dates and information on the individual animal used to raise antiserum was available, $\epsilon_T = \epsilon_r + \epsilon_s + \epsilon$ where ϵ_r is an effect to account for differences in titre attributable to run (assays completed by a specific operator on a specific date) and ϵ_s is an effect to account for differences in titre dependent on the individual animal from which serum was drawn reflecting animal-to-animal variability in the immune response. Differences in the optimal error structure between datasets likely represents differences in data collection, and not differences in the underlying biology.

The prior distribution on the effect for each reference strain was defined as a normal distribution such that $I_r \sim \mathcal{N}(\mu_I, \sigma_I^2)$ where $\mu_I \sim \mathcal{N}(0,\ 1{,}000{,}000)$ and the variance, σ_I^2, was drawn from an inverse gamma distribution such that $\sigma_I^2 \sim \mathcal{IG}(0.001, 0.001)$. μ_I is analogous to the intercept present in regression models described in previous chapters and forms a baseline for titres around which a normal distribution exists that the effects for each reference strain, I_r, are drawn. The prior distribution on the effect for each virus was defined as $A_v \sim \mathcal{N}(0, \sigma_A^2)$ where ${\sigma^2}_A \sim \mathcal{IG}(0.001,\ 0.001)$. The $D_{r,v}$ term describing the antigenic relationship between viruses was set to zero when the antisera used was raised against the same virus being tested (i.e. homologous titres). For heterologous titres, a truncated normal distribution was used where $D_{r,v} \sim \mathcal{N}(0, \sigma_D^2)|_{D_{r,v} \leq 0}$ where $\sigma_D^2 \sim \mathcal{IG}(0.001, 0.001)$. For each term, x, in $\epsilon_T, \epsilon_x \sim \mathcal{N}(0, \sigma_{\epsilon_x}^2)$ where $\sigma_{\epsilon_x}^2 \sim \mathcal{IG}(0.001,\ 0.001)$. As we assume that titres could be lower as a result of differences between viruses, but not higher, the antigenic term for heterologous virus-antiserum pairs was constrained to be negative.

Variation explained by ϵ_T was excluded from the posterior underlying titres which were described by:

$$H'_{r,v} = I_r + A_v + D_{r,v} \tag{5.2}$$

where $H'_{r,v}$ is the underlying $\log_2$ HI or VN titre for virus v and antiserum raised against reference virus r and all terms on the right-hand side of the equation are fixed from Equation 5.1. The posterior median value of $H'_{r,v}$ was calculated and used to validate predictions from antigenic cartography and sequence-based models. Models were implemented using

JAGS v3.3.0 (Plummer, 2012) through R (R Core Team, 2015) using the *runjags* package (Denwood, 2013). Convergence and autocorrelation were evaluated using Gelman & Rubin and Heidelberg & Welch diagnostic statistics and by visual inspection of MCMC plots for a posterior sample of 500 taken from two independent MCMC chains

5.3.4 Sequence-based prediction

The identification of HA amino acid positions involved in antigenic evolution of influenza A subtypes A(H1N1) and A(H3N2) is described in Chapters 3 and 4 respectively. Amino acid positions exposed on the surface of the FMDV capsid involved in the antigenic evolution of serotypes O and SAT1 were identified by Reeve *et al.* (2016) (under revision) while the identification of positions involved in antigenic evolution of serotype A is described by Bari *et al.* (2015). these identified amino acid positions were considered as candidate variables to be used for sequence-based prediction and are shown in Table 5.1. To assess the predictive performance of substitutions at these positions, cross-validation was carried out as described above. The constituents of training and test datasets is detailed in the relevant sections of the Results.

Table 5.1: Candidate amino acid positions for sequence-based approach

Virus	Protein	Amino acid positions
Influenza A(H1N1)	HA	36, 43, 71, 72, 74, 80, 120, 130, 141, 142, 153, 163, 183, 184, 187, 190, 252, 274, 313
Influenza A(H3N2)	HA	2, 25, 50, 62, 75, 79, 83, 122, 124, 131, 133, 135, 138, 144, 145, 155, 156, 157, 158, 159, 164, 172, 183, 189, 193, 196, 197, 201, 207, 212, 214, 217, 225, 227, 260, 262, 276, 307
FMDV A	VP1	83, 147, 151, 157, 172, 180
	VP2	74, 79
	VP3	135
FMDV O	VP1	142, 169, 211
	VP2	74, 130, 193
	VP3	56, 177
FMDV SAT1	VP1	102, 144, 146, 149, 163, 164, 209
	VP2	72
	VP3	72, 77, 171

* Candidate positions for influenza subtypes A(H1N1) and A(H3N2) identified by analyses described in Chapters 3 and 4 respectively. FMDV amino acid positions involved in antigenic evolution of serotypes O and SAT1 were identified by Reeve *et al.* (2016) (under revision) while the identification of positions involved in antigenic evolution of serotype A is described by Bari *et al.* (2015).

The simplest sequence-based models tested contained terms for the presence or absence of amino acid substitution between reference virus and test virus at each position in the set of candidate positions, Λ, which is given for each dataset in Table 5.1. Training datasets were used to parameterise a model described by:

$$H_{r,v} = I_r + A_v + \sum_{\lambda \in \Lambda} m_\lambda \delta_\lambda(r, v) + \epsilon_T \tag{5.3}$$

which modifies Equation 5.1 by replacing the term $D_{r,v}$ describing the antigenic relationship between reference strain and test virus (r and v) with terms representing the presence or absence of amino acid substitution at each candidate position in Λ. δ_λ(r, v) is 1 when reference virus r and test virus v are separated by any substitution(s) at a specific position in Λ and 0 otherwise. Each m_λ parameter represents the estimated drop in titre resulting from the antigenic impact of substitution at a particular amino acid position in Λ and has the truncated prior $m_\lambda \sim \mathcal{N}(0, \sigma^2_{m_\Lambda})|_{m_\lambda \leq 0}$ where the hyperparameter $\sigma^2_{m_\Lambda}$ is shared by all elements of Λ and has the prior $\sigma^2_{m_\Lambda} \sim \mathcal{IG}(0.001, 0.001)$. The terms retained from Equation 5.1 have the same priors as before.

Variation in titre ($H_{r,v}$) attributed to experimental and residual variation ϵ_T was excluded as it is irrelevant to predictions made for titres in the test dataset which were generated by:

$$H_{r,v} = I_r + A_v + \sum_{\lambda \in \Lambda} m_\lambda \delta_\lambda(r, v) \tag{5.4}$$

where all terms on the right-hand side of the equation are fixed from Equation 5.3. If a reference virus, r, or test virus, v, in the test data was not present in the training data then the relevant parameter relating to the magnitude of titres, I_r or A_v, could not be estimated by Equation 5.3 and so was given the value μ_I or 0 respectively in Equation 5.4. Assigning μ_I or 0 in this situation is approximately equivalent to assigning the average value of I_r or A_v estimated for reference or test viruses present in the training dataset. This practice was used for all subsequent sequence-based predictive models. Titres predicted for the test datasets using Equation 5.4 were compared with underlying titres extracted from the raw data using Equation 5.2 by calculating the mean, absolute error and its lower 95% credible interval.

5.3.4.1 Variable selection using binary mask models

The methods described in the cited works for the selection of the candidate positions in Table 5.1 did not consider, in depth, models containing combinations of terms representing

substitution at these amino acid positions. Amino acid positions or substitutions with statistical support were identified singly after accounting for the phylogenetic structure of the data. It is possible that some positions/substitutions explain some of the same variation in titre and statistically there is uncertainty about which terms offer the best explanation for the observed variation. Determining the optimal combination of these position/substitution variables using standard stepwise regression techniques is problematic given the sensitivity of these methods to the order in which variables are added or removed. Here, two approaches implementing binary mask models were used to address this issue and to optimise the combination of variables associated with substitution at the candidate antigenic positions. Binary mask models (e.g. Jow *et al.*, 2014) have been used to perform variable selection in a variety of contexts and are discussed together with other Bayesian alternatives, such as spike-and-slab priors, in Chapter 13 of Murphy (2012). Each variable considered using a binary mask model was associated with a 0–1 binary mask term which effectively excludes or includes the variable from the model. MCMC was used to converge on the optimal combination of included variables.

The first approach using binary mask terms was a two-step process. Candidate positions were considered using a binary mask model applied to the full dataset in order to identify an optimal set of variables representing substitutions at candidate positions. Following this, a model including the optimal set of variables was used for prediction. To implement the binary mask model, Equation 5.3 was modified:

$$H_{r,v} = I_r + A_v + \sum_{\lambda \in \Lambda} y_\lambda m_\lambda \delta_\lambda(r, v) + \epsilon_T \tag{5.5}$$

so that each position in Λ is associated with a binary mask parameter, y_λ, which takes the value 0 or 1. When y_λ is 0, the term representing the impact of substitution at a given position, $m_\lambda \delta_\lambda(r, v)$, is multiplied by 0 and therefore has no impact on titre, $H_{r,v}$, and is said to be masked. Each binary mask parameter was given the prior $y_\lambda \sim \text{Bernoulli}(p)$, where p was given the flat prior $p \sim \text{Beta}(1, 1)$ which was chosen to be minimally informative. The posterior of p represents the estimated proportion of positions in Λ that ought to be included in the optimal model.

Amino acid positions were ranked based on the how often the position parameter was not masked (the proportion of MCMC samples where the relevant binary mask term took the value 1). A higher posterior mean value of the binary mask term, y_λ, thus represents increased confidence in the inclusion of the associated amino acid position in the optimal model. The posterior median value of the inclusion parameter p was used to determine the proportion of positions from the ranked list of amino acid positions included in the optimal model. These selected amino acid positions became elements of Λ'. To assess the predictive power of substitutions at the positions in Λ', the model was rerun with only the selected positions,

so Λ' replaces Λ in Equations 5.3 and 5.4.

5.3.4.2 Bayesian model averaging (BMA)

The second approach used to implement binary mask terms involved only a single step in which predictions were made using a model including binary mask terms. By not requiring an optimal set of positions to be chosen, this approach is a form of Bayesian model averaging (BMA) that provides a coherent mechanism for accounting for model uncertainty. When substitutions occur together in the phylogeny at a point where antigenic change occurs, there is uncertainty as to which substitution(s) have caused the antigenic change. Information on the antigenic consequence of the same substitutions occurring in other branches of the phylogeny or other substitutions at the same sites can be used to resolve these ambiguities in some cases, but often there is unresolved uncertainty in terms of which substitutions are actually causing the observed antigenic change. One option is to attempt to optimise the model despite this uncertainty. BMA reflects this uncertainty in predictions that are made. Prediction is weighted towards the combination of position variables favoured by binary mask terms (i.e. those that form Λ') with less favourable combinations of position variables represented less often according to their favourability as assessed by MCMC. This approach involved applying Equation 5.5 to training data only, rather than the full dataset. Predictions of titres in the test data were then made using:

$$H_{r,v} = I_r + A_v + \sum_{\lambda \in \Lambda} y_\lambda m_\lambda \delta_\lambda(r, v) \tag{5.6}$$

where the values estimated for all terms on the right-hand side of the equation are fixed from Equation 5.5. Again, variation attributed to the error term, ϵ_T, was not used as it is irrelevant to prediction.

5.3.4.3 Allowing for variation in the impact of different substitutions

To account for potential differences in the antigenic impact of different substitutions occurring at the same position, Equation 5.3 was modified by replacing terms representing presence or absence of substitution at each position in Table 5.1 with terms representing the presence or absence of specific amino acid substitutions:

$$H_{r,v} = I_r + A_v + \sum_{\lambda \in \Lambda} \sum_{\kappa \in K_\lambda} m'_{\kappa,\lambda} \delta'_{\kappa,\lambda}(r, v) + \epsilon_T \tag{5.7}$$

While δ_λ in Equation 5.3 took the value 0 or 1 to indicate the absence or presence of amino acid dissimilarity at candidate position λ separating reference virus r and test virus v, here δ'_κ indicates the absence or presence of a specific amino acid difference, indexed by κ, among the set of all observed substitutions K_λ at each position λ in the full set of positions Λ taken from Table 5.1. Each m'_κ parameter represents the estimated drop in titre resulting from the antigenic impact on titre of a specific substitution and has the truncated prior $m'_\kappa \sim \mathcal{N}(\mu_\lambda, \sigma_K^2)|_{m'_\kappa \leq 0}$. This gives a normal distribution with a mean μ_λ specific to that position representing the average effect on titre across all substitutions at position λ and a variance shared by all substitutions $\sigma_K^2 \sim \mathcal{IG}(0.001, 0.001)$. The position-specific mean μ_λ retains the prior associated with the analogous m_λ term in Equation 5.3 so that $\mu_\lambda \sim \mathcal{N}(0, \sigma_{m_\Lambda}^2)|_{\mu_\lambda \leq 0}$ with the hyperparameter $\sigma_{m_\Lambda}^2 \sim \mathcal{IG}(0.001, 0.001)$. Thus, the estimated antigenic impact of a specific substitution depends upon both the amino acid position at which the substitution takes place and the amino acids residues involved.

Again, variation in titre explained by ϵ_T was excluded from predictions which were generated by:

$$H_{r,v} = I_r + A_v + \sum_{\lambda \in \Lambda} \sum_{\kappa \in K_\lambda} m'_{\kappa,\lambda} \delta'_{\kappa,\lambda}(r, v) \tag{5.8}$$

which was applied to test data with parameters fixed from Equation 5.6 applied to the training data.

The optimal combination of positions used for prediction was re-estimated while accounting for heterogeneity in the antigenic impact of different substitutions at the same position using binary mask models. Equation 5.7 was modified:

$$H_{r,v} = I_r + A_v + \sum_{\lambda \in \Lambda} y_\lambda \sum_{\kappa \in K_\lambda} m'_{\kappa,\lambda} \delta'_{\kappa,\lambda}(r, v) + \epsilon_T \tag{5.9}$$

to include binary mask terms, y_λ, for each position, λ, in the set of candidate positions Λ. All terms for specific substitutions, $m'_{\kappa,\lambda}\delta'_{\kappa,\lambda}(r, v)$, are multiplied by a position-specific binary mask term which takes the value 0 or 1, having the prior $y_\lambda \sim$ Bernoulli(p) where $p \sim$ Beta(1,1). Positions were again ranked based on the proportion of MCMC samples in which they were unmasked and the posterior median value of p was used to select the optimal proportion of positions from this ranked list of positions. To test the predictive performance of terms representing substitutions at the set of positions, Λ^*, selected by applying Equation 5.9 to the full datasets, Equations 5.7 and 5.8 were re-applied to training and test datasets with Λ replaced by Λ^*. Again, as an alternative to implementing the two step process of selecting an optimal set of variables using a binary mask model and then predicting titres

with the optimised set, predictions were made using a BMA approach by applying Equation 5.9 to training data and predicting titres in the test dataset using:

$$H_{r,v} = I_r + A_v + \sum_{\lambda \in \Lambda} y_\lambda \sum_{\kappa \in K_\lambda} m'_{\kappa,\lambda} \delta'_{\kappa,\lambda}(r,v) \tag{5.10}$$

where the error term, ϵ_T, is dropped from Equation 5.9 while all other terms are parameterised using the training data.

5.3.5 Antigenic cartography

Virus and antiserum locations in antigenic space were estimated as described in Chapter 2 (Section 2.3.4) using a Bayesian multidimensional scaling (BMDS) model (Bedford *et al.*, 2014). HI or VN titres are used to position antisera and viruses on the antigenic map such that distances between them are inversely proportional to $\log_2$ titres and together all distances are arranged in order to best satisfy the data. To follow previous studies, influenza and FMD viruses were positioned in two-dimensional and three-dimensional antigenic space respectively (Bedford *et al.*, 2014; Ludi *et al.*, 2014; Smith *et al.*, 2004). The BMDS method used extends the method of Smith *et al.* (2004) by incorporating estimates of antiserum and virus reactivity to account for variation in titres caused by differences in the immunogenicity of different reference viruses and in the receptor-binding avidity of viruses. These effects were estimated for every antiserum and virus respectively in each influenza and FMDV dataset. In addition, the prior locations of influenza viruses were informed by date of isolation and by phylogenetic relatedness to other viruses as described in Chapter 2. Date of isolation was not used to inform the prior locations of FMD viruses as there was no assumption of a relationship between temporal separation and antigenic distance. Phylogenetic distances were not used to inform the position of FMD viruses in antigenic space, however previous studies suggest improvements to the predictive performance achieved by incorporating phylogenetic information are minimal (0.01–0.05 $\log_2$ HI titre measured across four influenza datasets (Bedford *et al.*, 2014)). Models were implemented using BEAST (Drummond *et al.*, 2012) and posterior inference was carried out using R.

A posterior sample of locations was taken from those generated using MCMC after distances between antiserum and virus locations had converged. Predictions of either HI or VN titre are made using:

$$H_{r,v} = \frac{I_r + A_v}{2} - C_{r,v} \tag{5.11}$$

where $H_{r,v}$ is the predicted $\log_2$ HI or VN titre for virus v and antiserum raised against

reference virus r. For each titre the average of the estimated antiserum and virus reactivity terms (I_r and A_v respectively) is a baseline titre from which the Cartesian distance between the virus and antiserum ($C_{r,v}$), from the antigenic map, is subtracted. As with sequence-based modelling, predictive performance under a variety of scenarios was assessed using a variety of training and test datasets. Maps were generated using subsets of the full data, training datasets, and positions of viruses and antisera on the resultant maps were used to predict underlying titres for virus-antiserum pairs in the remaining test data. Prediction error was calculated by comparing titres predicted by Equation 5.11 with underlying titres fitted using the model described by Equation 5.2. The mean and 95% credible interval on the absolute prediction error, averaged across MCMC samples, was calculated.

5.4 Results

5.4.1 Binary mask models

Binary mask models were used to determine the optimal combinations of variables representing amino acid substitutions for predicting antigenic relationships. The influenza HA1 and FMDV capsid amino acid positions selected using these models are shown in Tables 5.2 and 5.3 respectively. Two alternative binary mask models were applied to each of the five virus datasets. Firstly, amino acid positions were tested using a model with the assumption that all substitutions occurring at a particular position have the same antigenic impact (Equation 5.5). The results of these models are shown in table columns with the heading Λ'. Secondly, amino acid positions were tested using a model which allowed different antigenic impacts for different amino acid differences at the same position to be estimated (Equation 5.9). The results of these models are shown in table columns with the heading Λ^*.

Amino acid positions in Tables 5.2 and 5.3 are ranked by their inclusion probability (IP). The inclusion probability is the proportion of the MCMC chain in which an amino acid position was included. The proportion of amino acid positions that ought to be included in the optimal model was independently estimated using each model and is shown as a dashed line in Tables 5.2 and 5.3. For each model, the proportion of amino acid positions estimated to represent the optimal combination of genetic determinants of antigenicity is therefore found above the dashed lines. These optimised sets of selected positions, Λ' and Λ^*, were used in sequence-based predictive models labelled 2 and 5 in Tables 5.4–5.6.

Table 5.2: Influenza HA1 amino acid positions selected using binary mask models.

A(H1N1)				A(H3N2)			
Λ'		Λ^*		Λ'		Λ^*	
Position	IP	Position	IP	Position	IP	Position	IP
36	1.00	36	1.00	124	1.00	2	1.00
43	1.00	43	1.00	135	1.00	50	1.00
72	1.00	71	1.00	133	1.00	62	1.00
74	1.00	72	1.00	138	1.00	133	1.00
141	1.00	74	1.00	144	1.00	135	1.00
142	1.00	141	1.00	145	1.00	138	1.00
153	1.00	142	1.00	156	1.00	145	1.00
163	1.00	153	1.00	158	1.00	156	1.00
183	1.00	163	1.00	164	1.00	157	1.00
184	1.00	183	1.00	183	1.00	158	1.00
187	1.00	184	1.00	189	1.00	159	1.00
190	1.00	187	1.00	193	1.00	172	1.00
252	1.00	190	1.00	196	1.00	183	1.00
274	1.00	252	1.00	207	1.00	189	1.00
313	1.00	274	1.00	212	1.00	193	1.00
71	0.79	130	0.99	262	1.00	212	1.00
130	0.71	120	0.97	25	0.94	262	1.00
80	0.57	313	0.95	172	0.94	131	0.99
120	0.45	80	0.89	79	0.92	155	0.99
				75	0.88	276	0.99
				214	0.85	225	0.98
				50	0.84	307	0.98
				197	0.75	25	0.97
				201	0.55	75	0.96
				122	0.50	196	0.96
				225	0.41	197	0.96
				217	0.31	207	0.95
				62	0.18	164	0.94
				227	0.10	201	0.71
				157	0.07	79	0.65
				307	0.07	214	0.58
				131	0.06	122	0.49
				159	0.06	260	0.49
				276	0.06	217	0.41
				83	0.05	124	0.26
				155	0.05	227	0.26
				260	0.05	83	0.04
				2	0.04	144	0.02

Positions are ranked by their inclusion probability (IP) - the proportion of MCMC samples in which the position is unmasked (i.e. the associated binary mask term = 1), with positions above the dashed lines being selected. Λ' indicates positions selected by a model without independent effects for different substitutions at the same position (Equation 5.5) and Λ^* with independent effects for different substitutions at the same position (Equation 5.9).

The proportion of candidate amino acid positions selected using both binary mask models was higher for A(H1N1) than A(H3N2). This reflects the greater uncertainty associated with the identification of genetic determinants of antigenic change in A(H3N2) as discussed in Chapter 4, which arises when it is not possible to discriminate between multiple substitutions that have occurred at the same point (or similar points) in the evolutionary history of the virus. For A(H1N1), this situation arose in only two instances (Table 3.1) where it was not possible to discriminate between substitutions at positions 74 and 120 in one case and between substitutions at 43, 71, 80, and 130 in the other. In the second case, experimental results indicate that the antigenically important change occurred at position 130 and not at the others (McDonald *et al.*, 2007, Chapter 3: Figure 3.6). Here, position 80 was not selected using either binary mask model, however positions 43 and 71 were selected in addition to 130 indicating that other substitutions at these positions may correlate with antigenic change at other points in the phylogeny. Position 74 was selected over 120 using the simpler binary mask model where the antigenic impact of all substitutions at a position is assumed to be the same (Equation 5.5), while both positions were selected when different amino acid substitutions at a position were allowed to vary in their impact (Equation 5.9). A lower proportion of A(H3N2) candidate positions were selected using both models, relative to A(H1N1). Two positions, 155 and 159, where substitutions have been experimentally confirmed as causing high-impact changes in HA antigenicity were not selected in Λ' but were selected in Λ^*. This possibly indicates an advantage of accounting for differences in the impact of different amino acid changes when implementing binary mask models of this kind.

Table 5.3: FMDV capsid amino acid positions selected using binary mask models.

A				O				SAT1			
Λ'		Λ^*		Λ'		Λ^*		Λ'		Λ^*	
Position	IP	Position	IP	Position	IP	Position	IP	Position	IP	Position	IP
VP1 83	1.00	VP1 83	1.00	VP1 142	1.00	VP3 177	1.00	VP2 72	1.00	VP2 72	1.00
VP1 147	1.00	VP1 147	1.00	VP3 177	1.00	VP1 211	0.99	VP1 102	1.00	VP1 144	1.00
VP1 151	1.00	VP1 151	1.00	VP2 74	0.32	VP1 142	0.82	VP1 144	1.00	VP1 149	1.00
VP1 172	1.00	VP1 157	1.00	VP2 130	0.31	VP2 130	0.60	VP1 149	1.00	VP1 163	1.00
VP2 79	0.93	VP1 172	1.00	VP2 193	0.28	VP2 74	0.51	VP1 163	1.00	VP1 209	1.00
VP1 157	0.93	VP2 74	0.97	VP1 211	0.24	VP2 193	0.37	VP3 171	1.00	VP3 72	0.99
VP1 180	0.90	VP2 79	0.88	VP3 56	0.18	VP1 169	0.31	VP1 209	1.00	VP1 102	0.99
VP2 74	0.89	VP1 180	0.78	VP1 169	0.12	VP3 56	0.29	VP3 72	0.99	VP3 171	0.91
VP3 135	0.64	VP3 135	0.58					VP3 77	0.97	VP1 146	0.85
								VP1 164	0.96	VP3 77	0.82
								VP1 146	0.88	VP1 164	0.72

Positions are ranked by their inclusion probability (IP) - the proportion of MCMC samples in which the position is unmasked (i.e. the associated binary mask term = 1), with positions above the dashed lines being selected. Λ' indicates positions selected by a model without independent effects for different substitutions at the same position (Equation 5.5) and Λ^* with independent effects for different substitutions at the same position (Equation 5.9).

5.4.2 Predicting unobserved antigenic relationships

One of the challenges faced when attempting to identify the antigenic characteristics of a virus population is that only a very small fraction of the viral population is ever isolated and among these isolated viruses it is normal for only a small proportion of pairwise antigenic relationships to be determined experimentally. To assess their ability to predict unobserved antigenic relationships, models were parameterised using training datasets composed of titres recorded between 90% of virus-antiserum pairs. Performance was assessed by calculating the mean, absolute error (or test error) of titres predicted for test datasets containing the remaining 10% of virus-antiserum pairs. The composition of test and training datasets was randomised and this was repeated 50 times (i.e. 50-fold cross-validation) The mean, absolute error for various sequence-based models described above and for antigenic cartography are shown in Table 5.4.

Table 5.4: Sequence-based prediction: Average absolute prediction error of $\log_2$ titres for test datasets consisting of missing virus:antiserum pairs

Model	Positions or Substitutions	Variable set	Mean absolute error				
			Influenza		FMDV		
			A(H1N1)	A(H3N2)	A	O	SAT1
1	Positions	All	0.63	0.69	0.74	0.55	0.35
2	Positions	Optimised	0.62	0.69	0.75	**0.55**	0.36
3	Positions	BMA	0.64	0.69	0.76	0.55	0.36
4	Substitutions	All	0.62	0.66	0.75	0.57	0.34
5	Substitutions	Optimised	**0.60**	0.68	0.74	0.56	0.34
6	Substitutions	BMA	0.61	**0.66**	**0.74**	0.57	**0.34**
7	Antigenic cartography		0.54	0.77	0.68	0.54	0.65

Positions or substitutions refers to whether or not different substitutions at the same position were able to have independent antigenic impacts. Either all variables were used, a set optimised using binary mask models were used or BMA was used to average across models. The most accurate sequence-based models are shown in bold.

Across the five datasets there is very little variability in the predictive performance of the various sequence-based approaches tested. Accounting for differences in the antigenic impact of different substitutions at the same position improves predictive performance for A(H1N1), A(H3N2) and SAT1 (mean, absolute error for model 4 < mean, absolute error for model 1) but not for FMDV serotypes A or O. A potential disadvantage of BMA is overfitting to the data, though this does not seem to have occurred here as the predictive performance of models 3 and 6 are not notably less accurate than the other sequence-based models tested. Using BMA to account for model uncertainty has not however resulted in better prediction either. As the quantitative differences between sequence-based models is small, I favour Model 6 which is qualitatively superior for two reasons: Firstly it is preferable to allow different substitutions at the same position to vary in antigenic impact as we know the antigenic impact of substitutions does in fact depend on the different biophysical properties of the

amino acids involved; and secondly because uncertainty in the best model does exist due to multiple substitutions occurring at similar points in the evolution of the virus where antigenic changes are estimated to have occurred, so it is better to reflect this in predictions.

On average, Model 6 outperforms antigenic cartography for two datasets (A(H3N2) and SAT1) while antigenic cartography produced more accurate predictions of titres for A(H1N1), and FMDV serotypes A and O. This shows that on average the sequence-based approach presented here and state of art methods for antigenic cartography are similarly capable of predicting unobserved antigenic relationships, as expressed in HI and VN tests. In addition to the average error in predictions it is important to consider upper limits in the inaccuracy of these models as this gives an indication of how risky it might be to rely on these techniques. To this end, 95% credible intervals for the absolute error of predictions made using Model 6 and antigenic cartography were calculated and these are shown in Figure 5.2. Figure 5.2 shows that 95% credible intervals for predictions made using Model 6 were more accurate than BMDS for influenza A(H3N2) and FMD serotypes A, O, and SAT1 and less accurate for influenza A(H1N1). Figure 5.2 also shows 95% of predictions made using Model 6 across influenza subtypes had a mean, absolute error less than two antigenic units, or a four-fold dilution in the HI assay, while predictions for A(H3N2) made using antigenic cartography were only marginally less accurate than this.

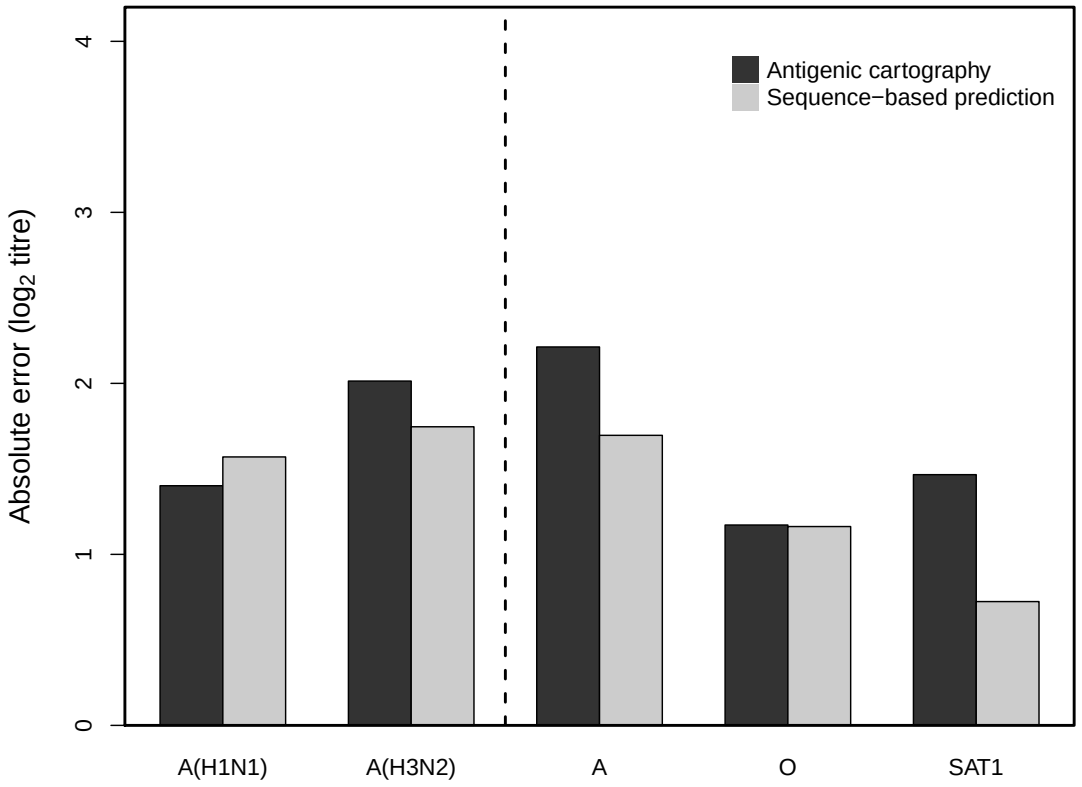

Figure 5.2: Model cross-validation: predicting titres for missing virus-antiserum pairs: Mean, absolute error and lower 95% credible intervals in predictions of $\log_2$ titres for test datasets consisting of 10% of virus-antiserum pairs. Predictive performance is shown for null, BMDS and sequence-based predictive models across datasets consisting of influenza A subtypes A(H1N1) and A(H3N2) and FMD serotypes A, O, and SAT1 viruses and antisera.

5.4.3 Predicting cross-reactivity of uncharacterised antisera

Predicting titres for unobserved virus-antiserum pairs is a useful method of model cross-validation and gives an indication of whether models are being overfitted to the data. However, viruses and antisera in the test datasets are also, in most cases, present in the training dataset being tested against different antisera/viruses and so these represent viruses and antisera that have already been antigenically characterised by HI or VN assay, and so provide limited value. The ability to predict antigenic relationships, as expressed in these assays, for either antisera or viruses that have not been antigenically characterised experimentally, and therefore do not appear at all in training datasets, is more informative. The ability to antigenically characterise a viral population depends to an extent on the range of antigenic phenotypes represented by available antisera; if all antisera in a panel are antigenically similar, it is possible to learn only whether a test virus is similar to or different from the viruses used to generate those antisera. There are practical limitations on the number of antisera that can be used so there is therefore an implicit prediction of the antigenic characteristics

of antiserum involved when a virus is chosen to raise antiserum.

To provide a more detailed understanding of the likely properties of an antiserum as part of the process of selection of reference antiserum, I tested the ability of sequence-based models to predict HI or VN titres for antisera that were entirely excluded from training data, using the genetic sequence of the associated virus. For comparison, predictions were made using the position of the associated virus in antigenic space estimated using BMDS. For A(H1N1) and A(H3N2) test datasets containing 10% of antisera were repeatedly chosen at random. Each FMDV dataset had seven or fewer antisera in total so test datasets containing each antiserum in turn were used. The mean, absolute error of predictions of titre made for uncharacterised antisera are shown in Table 5.5.

Table 5.5: Sequence-based prediction: Average absolute error for $\log_2$ titres estimated for test datasets consisting of missing antisera

Model	Positions or Substitutions	Variable set	Mean absolute error				
			Influenza		FMDV		
			A(H1N1)	A(H3N2)	A	O	SAT1
1	Positions	All	0.78	0.90	1.61	1.79	0.42
2	Positions	Optimised	0.79	0.94	**1.57**	**1.59**	0.41
3	Positions	BMA	**0.75**	0.89	1.62	1.72	0.42
4	Substitutions	All	0.85	0.88	1.77	1.76	0.36
5	Substitutions	Optimised	0.81	**0.85**	1.71	1.61	**0.35**
6	Substitutions	BMA	0.81	0.86	1.77	1.72	0.36
7	Antigenic cartography		0.75	0.98	2.02	1.96	1.87

Positions or substitutions refers to whether or not different substitutions at the same position were able to have independent antigenic impacts. Either all variables were used, a set optimised using binary mask models were used or BMA was used to average across models. The most accurate sequence-based models are shown in bold.

Again, neither of the sequence-based models consistently outperforms the others. Models 2, 3 and 5 produced the lowest test errors on average for different datasets though the magnitudes of the differences between these means indicate that all sequence-based models perform similarly well. On average, the BMA model of prediction that allows for differences in the antigenic impact of different amino acid substitutions (Model 6) outperforms antigenic cartography when predictions are made from antisera of influenza A(H3N2) and all three FMDV serotypes, while predictions for A(H1N1) are slightly more accurate on average when made using antigenic cartography. 95% credible intervals for Model 6 and antigenic cartography were calculated and are shown in Figure 5.3. When 95% credible intervals are considered the sequence-based approach outperforms antigenic cartography across all five datasets.

Predictions of titres for uncharacterised antisera made for both influenza subtypes are less accurate than predictions made for unobserved virus-antiserum pairs, as the exclusion of all observations for a particular antiserum prevents the model learning the immunogenicity parameter associated with each antiserum (I_r terms in Equations 5.1–5.10) resulting in

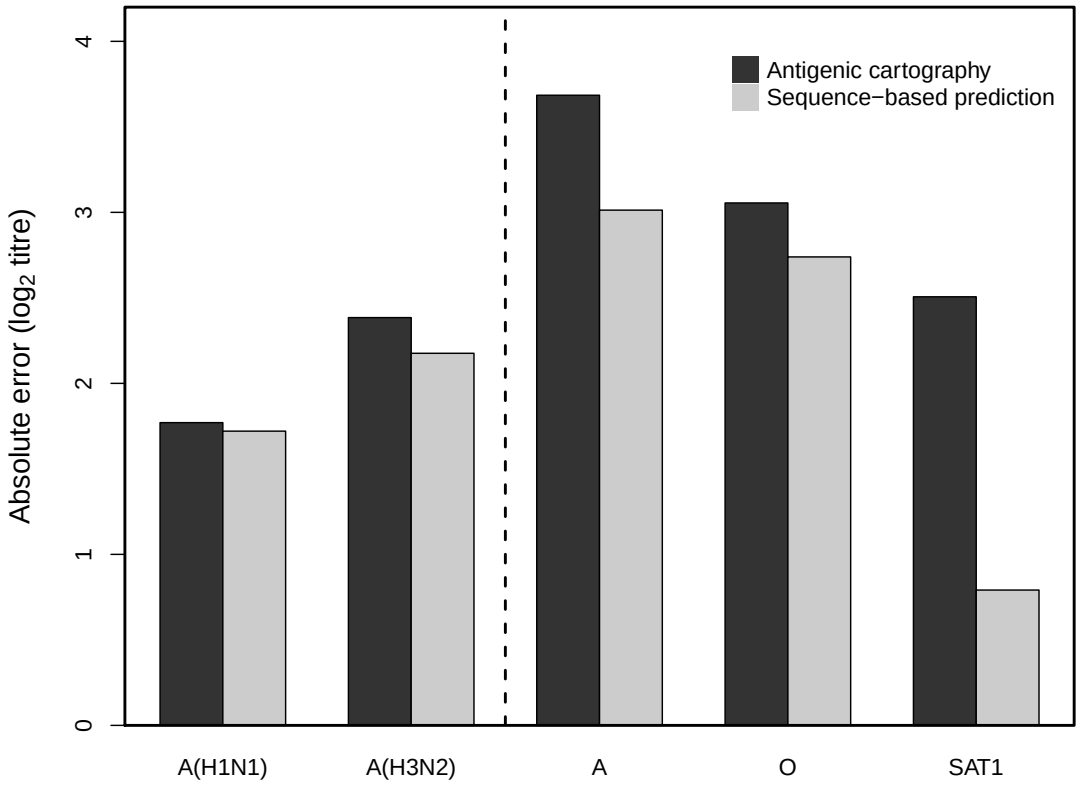

Figure 5.3: Model cross-validation: predicting titres for uncharacterised antisera: Mean, absolute error and lower 95% credible intervals in predictions of $\log_2$ titres for test datasets consisting of 10% of virus-antiserum pairs. Predictive performance is shown for null, BMDS and sequence-based predictive models across datasets consisting of influenza A subtypes A(H1N1) and A(H3N2) and foot-and-mouth disease virus serotypes A, O and SAT1 viruses and antisera.

systematic prediction errors.

5.4.4 Predicting cross-reactivity of uncharacterised viruses

One of the advantages of a sequence-based method is that antigenic relationships can be predicted for viruses that have not been tested experimentally. This is potentially the most useful application of prediction because, as genetic sequence data become easier to obtain, it will provide the ability to quickly estimate antigenic phenotype before experimentation. Test datasets consisting of all observations for 10% of the viruses in each dataset were repeatedly chosen at random in a cross-validation procedure and predictions were made based on sequence differences to the reference/vaccine strains used to generate antiserum in the training datasets. Mean, absolute error in predictions made by each of sequence-based models across each of the five datasets are shown in Table 5.6. For four of the five datasets, the most accu-

rate predictions were made for missing viruses when different amino acid substitutions were allowed to vary in their predicted antigenic impact. In two cases, for A(H1N1) and SAT1, Model 6 which also incorporates model uncertainty in predictions outperformed other models. Viruses for which no HI or VN data has been gathered cannot be positioned in antigenic space using cartographic methods so under this scenario predictions cannot be made using with antigenic cartography.

Table 5.6: Sequence-based prediction: Average absolute prediction error of $\log_2$ titres for test datasets consisting of missing viruses

Model	Positions or Substitutions	Variable set	Mean absolute error				
			Influenza		FMDV		
			A(H1N1)	A(H3N2)	A	O	SAT1
1	Positions	All	0.87	0.99	0.99	**0.64**	0.55
2	Positions	Optimised	0.90	0.98	0.95	0.65	0.55
3	Positions	BMA	0.86	0.98	1.05	0.68	0.56
4	Substitutions	All	0.86	0.99	**0.91**	0.71	0.54
5	Substitutions	Optimised	0.88	**0.93**	0.92	0.66	0.54
6	Substitutions	BMA	**0.85**	0.97	0.93	0.75	**0.53**

Positions or substitutions refers to whether or not different substitutions at the same position were able to have independent antigenic impacts. Either all variables were used, a set optimised using binary mask models were used or BMA was used to average across models. The most accurate models are shown in bold.

For both influenza subtypes the mean, absolute errors of each predictive model were less than one $\log_2$ titre (antigenic unit) so on average predictions of HI titres are within one two-fold dilution or well on the plate of an HI assay (Table 5.6). Influenza predictions for uncharacterised viruses are, however, less accurate than those made when predicting titres for unobserved virus-antiserum pairs or for uncharacterised antisera. Viruses present in the test data were totally absent from the training data, so the impact of virus receptor-binding on titre could not be learnt (A_v parameters in Equations 5.1–5.10) using the data and was therefore taken as the average of A_v effects estimated for viruses present in the training dataset instead. Titres predicted for missing viruses of each FMDV serotype were less accurate than those made for missing virus-antiserum pairs suggesting that non-antigenic aspects of virus phenotype may also be important in VN titres recorded for FMD viruses. The titres predicted for serotypes A and O are, however, much more accurate than those made for uncharacterised antisera. This indicates that the inability to train these virus reactivity parameters (A_v) is of less importance than the inability to train immunogenicity effects in these datasets.

5.4.5 Influenza: prediction through time

To further investigate the power of the sequence-based approach to predict antigenic phenotype in influenza, HI titres for viruses isolated in each year of our dataset (from 1998 to 2009

for A(H1N1) and from 1998 to 2015 for A(H3N2)) were predicted using models trained to data collected in previous years. Test datasets comprised all titres measured between viruses isolated in a given calendar year and antisera to viruses collected over previous years. These predictions are therefore directly relevant to decisions made during the vaccine selection process since they are estimating the antigenic novelty, as expressed in the HI assay, of emerging genotypes. 1998 was chosen as the first year to be predicted as there was a noticeable improvement in the number of data across both subtypes in the period 1996–98. Predictions were made using Model 6 which was was the best and second best model for predicting titres for uncharacterised viruses of subtypes A(H1N1) and A(H3N2) respectively (Table 5.6).

To investigate the relative contributions to HA cross-reactivity, I made further comparisons between predictions made using the full sequence-based model with a model that used only substitutions detectable as causing transitions between clusters of antigenically similar viruses on an antigenic map for prediction. These substitutions are, in general, those of greatest antigenic impact, which individually may lead to a need to change vaccine virus. For A(H1N1), ΔK130 and K141E were used in this model. In Chapter 2 these substitutions were found as being able to explain transitions between clusters of antigenically similar viruses (Figure 2.10) and their large antigenic impact has been demonstrated experimentally (McDonald *et al.*, 2007, Chapter 3: Figure 3.6). For A(H3N2), this model included substitutions at the seven positions 145, 155, 156, 158, 159, 189 and 193. The substitutions N145K, T155Y, Y155H, K156E, K156Q, Q156H, G158E, E158K, S159Y, Q189K, K189R, D193N have been identified using reverse genetics experiments as causing transitions between antigenic clusters apparent on maps of A(H3N2) viruses isolated between 1968 and 2003 (Koel *et al.*, 2013) and sequence analysis of more recent antigenic clusters indicates that the substitutions K145N, Y159F, S189N, K158N and N189K may have caused more recent transitions between clusters (Chapter 2, Section 2.4.4).

For A(H1N1), mean absolute prediction error when only the high-impact substitutions ΔK130 and K141E were included averaged across the twelve years was 1.81 $\log_2$ titre (SEM = 0.0139, SD = 0.48). This was reduced to 0.82 $\log_2$ titre (SEM = 0.0072, SD = 0.25) when all substitutions at each position in Table 5.1 were considered using BMA (Model 6). The mean absolute prediction error of each model in each year is shown in Figure 5.4. In all years the accuracy is improved by the inclusion of the smaller-scale antigenic determinants in addition to the two substitutions that define antigenic clusters on the map.

For A(H3N2) mean absolute prediction error was also improved by the inclusion of substitutions causing changes in antigenic phenotype of lower-impact in addition to the set of substitutions that have been identified as being responsible for the major changes in antigenic phenotype that have occurred during the evolution of H3 HA in humans. The mean, absolute error averaged across the 17 years used for prediction was 1.47 (SEM = 0.116, SD = 0.47) when the antigenic substitutions of highest impact were included and this was reduced to 1.01

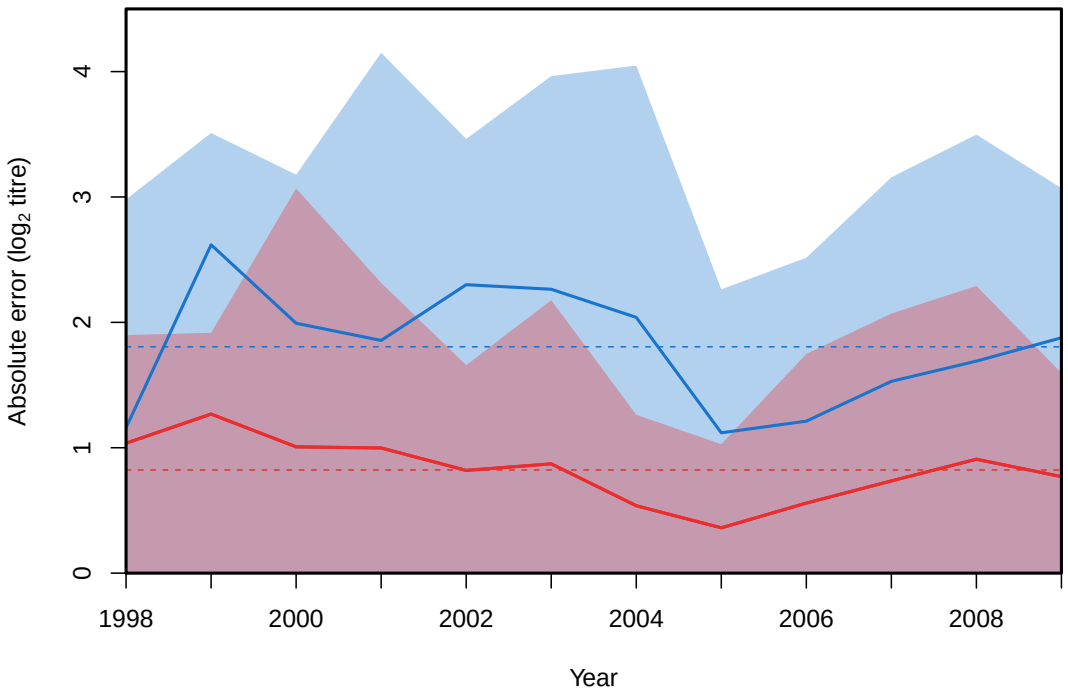

Figure 5.4: A(H1N1): Prediction error through time for models used to predict HI titres of viruses isolated in the following year: The mean, absolute difference between observed titres for viruses isolated in a given year and titres predicted using models trained to HI data collected in previous years is shown. Predictive models included terms for cluster-defining substitutions ΔK130 and K141E only (solid blue line) or used a BMA model allowing different antigenic impacts for different substitutions at each position in Table 5.1 (solid red line). For each model, shaded areas show the lower 95% credible interval on the absolute prediction error. In each year the blue 95% credible interval extends vertically on the y-axis above the red 95% credible interval and the overlapping area appears purple. The mean, absolute prediction errors averaged across the 12 years are shown as dashed lines.

(SEM = 0.080, SD = 0.32) when the full sequence-based model of prediction was used. The improvement in predictive performance achieved by a deeper understanding of the genetic changes impacting upon HI titres was therefore smaller than in A(H1N1), however the BMA model still improved prediction in each year (Figure 5.5).

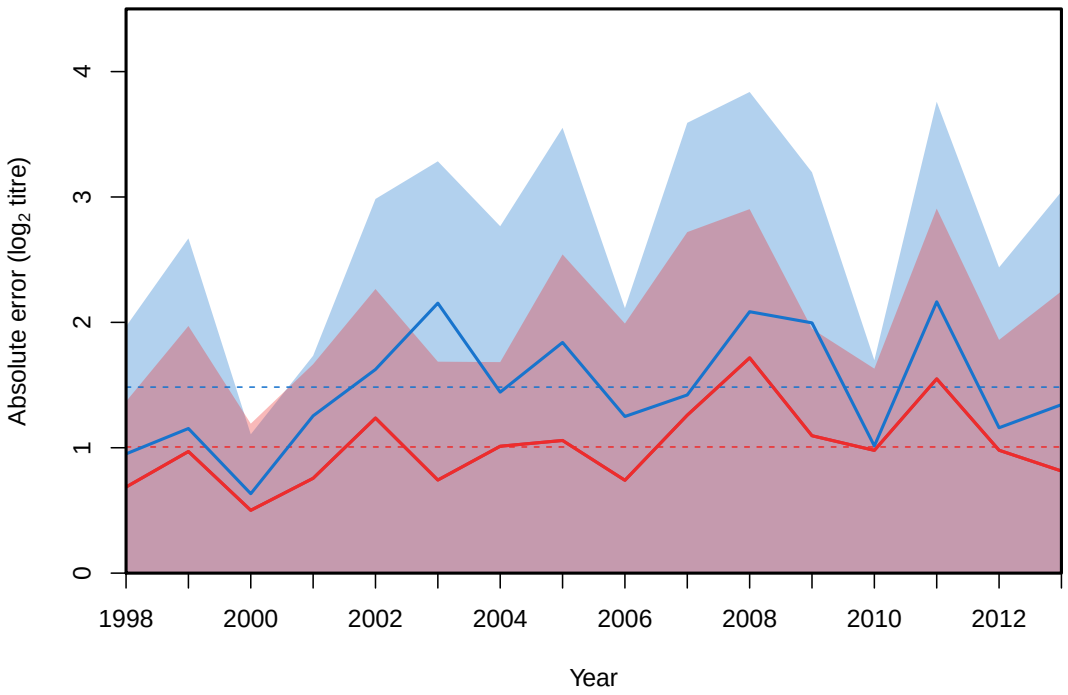

Figure 5.5: A(H3N2): Prediction error through time for models used to predict HI titres of viruses isolated in the following year: The mean, absolute difference between observed titres for viruses isolated in a given year and titres predicted using models trained to HI data collected in previous years is shown. Predictive models included terms for cluster-defining substitutions (solid blue line) or used a BMA model allowing different antigenic impacts for different substitutions at each position in Table 5.1 (solid red line). For each model, shaded areas show the lower 95% credible interval on the absolute prediction error. In each year except 2000, the blue 95% credible interval extends vertically on the y-axis above the red 95% credible interval and the overlapping area appears purple. The mean, absolute prediction errors averaged across the 16 years are shown as dashed lines. After 2003, cluster-defining substitutions are inferred from antigenic maps in Chapter 2 and are not experimentally validated.

5.5 Discussion

In this chapter, I demonstrate the ability to predict antigenic relationships, as expressed in HI or VN titres, among viruses of two influenza A subtypes and three FMDV serotypes from genetic sequence data under a variety of scenarios. Constraints on time and resources limit the number of viruses that can be investigated experimentally using antigenic assays, and HI and VN titres give only pairwise measures of cross-reactivity. Expert decision-makers must attempt to form a complete picture of the antigenic profile of a virus population based on the available set of pairwise measurements that realistically can only represent a small minority of the possible pairwise comparisons between genotypes in the population. I first examined the ability of sequence-based models to predict the antigenic cross-reactivity of pairs of virus and antisera not directly compared experimentally by forming test datasets

comprised of a proportion of virus-antiserum pairs in the full datasets. This is useful as it shows the potential for this kind of method to provide a more complete antigenic profile of the virus population. This analysis showed the presented sequence-based approach to be similarly capable of predicting titres when compared to a BMDS antigenic cartography model which has previously been demonstrated to be significantly better than other models of antigenic cartography (Table 1 in Bedford *et al.*, 2014). In addition to predicting titres more accurately than BMDS for the A(H3N2) and SAT1 datasets on average (Table 5.4), the sequence-based approach could be considered a less risky form of prediction as 95% lower credible intervals on prediction error were reduced in four out of five datasets (Figure 5.2). In addition to the quantitative argument, a sequence-based approach has a qualitative advantage in that it is based on a falsifiable, mechanistic explanation of the relationship between genetic changes and antigenic variation.

When predicting titres for uncharacterised antisera using sequence-based models, prediction accuracy was lower for each of the five datasets than when predicting missing virus-antisera pairs. Dropping reference antisera from test datasets prevents models being trained to the estimated immunogenicity of those antisera so the immunogenicity term (I_r in Equations 5.1–5.11) is instead estimated as the average of antiserum effects estimated for antisera that remain in the training data. This issue appears to have caused significant differences in predictions made for FMDV serotypes A and O, where prediction accuracy using either sequence-based models or BMDS is dramatically reduced. Particular difficulty may arise in estimating the I_r terms for excluded antisera in these two datasets, where there are so few antisera and given their immunogenicity is known to be highly variable since they include antiserum from animals vaccinated with commercial high potency vaccines and animals vaccinated with custom vaccines not tested for potency.

Less variation in the immunogenicity of SAT1 antisera meant that the problem of estimating I_r terms for excluded antisera did not seem to be problematic for that serotype, hence titres estimated using the sequence-based approach are only slightly less accurate than when predicting unobserved virus-antiserum pairs. This suggests that there are further disadvantages to using antigenic cartography to estimate titres in this dataset as titres estimated for missing SAT1 antiserum were significantly less accurate (Table 5.5 and Figure 5.3). This is potentially a consequence of using virus–virus distances from the maps to predict titres. When predicting titres for unobserved virus-antiserum pairs, and in all previously published cross-validation of antigenic cartography that I am aware of, virus-antiserum distances from maps are used. When antigenic maps are generated virus-antiserum distances are optimised to reflect titres from antigenic assays, however virus-virus distances are what is most commonly interpreted from these maps. With influenza antigenic maps, predicting HI titres from virus–virus distances taken from antigenic maps has not resulted in a loss of accuracy of a similar magnitude.

While the ability to predict titres for unobserved virus-antiserum pairs or uncharacterised antisera is useful and interesting, the ability to predict titres for new virus isolates that have not been antigenically characterised is a more desirable aim. Genetic variants are constantly emerging in each of the influenza A subtypes and FMDV serotypes studied and as the sequencing of viruses becomes less expensive and more routine, the ability to estimate the antigenic similarity of these viruses to existing viruses will become more valuable. In each of the five datasets investigated, sequence-based models proved capable of predicting titres for uncharacterised viruses, not present in training datasets, with a mean, absolute error below one two-fold dilution in either the HI or VN assay. Again, there is a loss in accuracy relative to making predictions for unobserved virus-antiserum pairs. This could be accounted for by differences in virus reactivity (A_v terms in Equations 5.1–5.11), which cannot be learnt. It has previously been shown that amino acid substitutions can allow escape from the polyclonal antibody response represented in HI assays by increasing receptor-binding avidity (Hensley *et al.*, 2009), though it not known to what extent this kind of effect is important in immune escape in nature. Identifying the amino acid differences responsible for the variation in avidity that affects HI titres and using population genetics methods to assess the fitness of those substitutions could potentially resolve this question, and additionally enable more accurate sequence-based predictions of titres when A_v effects cannot be learned from the data.

For influenza subtypes A(H1N1) and A(H3N2), predictions were also made for viruses isolated in each year from 1998 using BMA models that also account for differences in the antigenic impact of different substitutions at the same position, trained to data collected in previous years. The mean, absolute error of these predictions was roughly the same as when making predictions for test datasets comprised of 10% of viruses. Under this scenario the full sequence-based model produced significantly more accurate predictions than those achieved using only substitutions that have caused high-impact changes in antigenic cross-reactivity and that have been identified through analysis of antigenic maps generated using HI data, followed by extensive reverse genetics experiments (Chapter 3 and Koel *et al.* (2013)). This effect was greater in the A(H1N1) dataset, perhaps indicating that smaller-impact amino acid substitutions play a greater role in the antigenic evolution of A(H1N1) viruses, relative to A(H3N2). However, as there are other differences in the structure of datasets any such inference should be made cautiously. The A(H3N2) dataset also contains a much higher proportion of reference viruses. Reference viruses are selected partly because of their antigenic properties and it is possible that high impact substitutions are more likely to be important in defining antigenic relationships between reference viruses which are usually chosen to be more antigenically diverse than a similar number of non-reference viruses.

In contrast to the gathering of genetic sequence data, which is becoming more routine, the antigenic characterisation of A(H3N2) viruses, particularly those of the phylogenetic sub-clade 3C.2a, has recently become more technically difficult because of their changing properties

(WHO, 2015b). Specifically, many 3C.2a viruses had low or undetectable haemagglutination activity which has required modifications to HI and VN assays to be made in order to infer antigenic relationships by these methods. Problems of this kind faced by traditional approaches call for new modelling approaches to help them. In addition to predicting antigenic variants based on prior knowledge of the importance of a genetic change that has arisen previously in the evolution of a virus, a sequence-based approach can predict when a genetic variant, possessing substitutions not associated with antigenic change in the past, will be antigenically similar to existing viruses. In addition to the antigenic characterisation of circulating viruses, vaccine virus selection requires an anticipation of which viruses will predominate in future epidemics up to a year in advance. Recently published methods for predicting the evolutionary success of influenza viruses based on HA genotype (Łuksza & Lässig, 2014; Neher *et al.*, 2014) show that statistical methods can predict which strains will predominate in future seasons. HA cross-reactivity as expressed in the HI assay is an important aspect of antigenic phenotype and therefore of virus fitness. The ability to predict the antigenic cross-reactivity, as measured by HI, of emerging influenza viruses using models parameterised using data collected in previous years should help to refine such techniques. This is discussed further in Chapter 6.

CHAPTER 6

Discussion

Discussion

6.1 General discussion

The aims of this thesis were to investigate the genetic basis of antigenic variation among human influenza A viruses using models based on the mechanistic relationship between amino acid changes and antigenic evolution. I have attempted this using a range of models that attribute variation in haemagglutinin (HA) cross-reactivity, measured using the haemagglutination inhibition (HI) assay, to amino acid substitutions that have occurred during the evolution of A(H3N2) and former seasonal A(H1N1) viruses. By quantifying the antigenic impact of specific amino acid substitutions, our knowledge of how specific genetic changes contribute to the antigenic phenotype of influenza viruses and therefore virus fitness is enhanced. In this chapter, I will begin summarising the key findings of each chapter, and will then explore some of the unresolved issues that arose throughout the analyses described, with a particular focus on the role of changes in receptor-binding avidity on HI titres. Finally, I will describe the significance of the results in terms of how they could complement and enhance related techniques and speculate on future directions.

In Chapter 2, the A(H1N1) and A(H3N2) datasets that are explored throughout the thesis were analysed using methods that focus on either genotype or phenotype. Phylogenetic analysis indicated that the rate of genetic evolution is similar in both subtypes. Two measures of genetic distance were tested as predictors of antigenic distance as reflected by HI titres, and for both subtypes amino acid distance was a better predictor of titre than phylogenetic distance, though neither measure was well correlated with HI titre indicating that a better understanding of the heterogeneity in the antigenic impact of genetic changes is required for accurate sequence-based prediction of antigenic similarity. The ratios of rates of synonymous (dS) and non-synonymous (dN) mutations at each position in the HA1 domain of HA were estimated using various related methods. These analyses showed the proportion of HA1 codons under purifying selection to be notably higher than the proportion under positive selection in both subtypes. The proportion of positions in H1 antigenic sites with evidence of positive selection was marginally higher than the proportion detected as being under positive selection across the whole of HA1, while in H3 antigenic sites the proportion

of positions exposed to positive selection was much higher and fewer positions were exposed to purifying selection. A(H3N2) viruses also exhibited higher rates of antigenic drift than A(H1N1) viruses on antigenic maps generated from HI data, possibly indicating a more significant role of immune-mediated evolution in A(H3N2), relative to A(H1N1). Eight A(H1N1) and 16 A(H3N2) HA1 codons were detected as showing signatures of positive selection. The majority of these positions have previously been identified as having a role in either antibody binding or virus receptor binding or both but their role in adaptive evolution must be interpreted with caution as their identification is based on phenotypic data.

Substitutions between antigenic clusters on antigenic maps generated using HI data were also identified in Chapter 2, though notably the set of identified substitutions was inconsistent with previously published similar studies that were also inconsistent with each other (Koel *et al.*, 2013; Sun *et al.*, 2013). This difficulty in replicating the same set of candidate substitutions raises the question of whether it would be preferable to test all substitutions estimated to have occurred in the trunk lineage of HA phylogenetic trees. A significant degree of variation expressed on the antigenic maps shown in Chapter 2 was within-cluster rather than between-cluster and the genetic causes of this variation cannot be determined by existing methods.

In Chapter 3, a modelling approach used to identify surface-exposed regions of the capsid protein of foot-and-mouth disease virus (FMDV) where substitutions were correlated with antigenic change was extended and applied to A(H1N1). Nineteen amino acid substitutions were identified as causing antigenic variation among the viruses studied. These included seven substitutions detected as correlating with antigenic change at multiple points in the evolution of the virus and one instance where two candidate substitutions (E74G and E120G) could not be discriminated between to identify the genuine cause of antigenic change. The antigenic impact of substitutions was also quantified and substitutions ranged in their average antigenic impact from 0.20 to 3.53 $\log_2$ titre. The two identified substitutions with highest antigenic impact, K141E and ΔK130, were also identified as substitutions capable of explaining transitions between antigenic clusters in Chapter 2. Reverse genetics was used to validate the importance of four substitutions and the quantification of their antigenic impact.

In Chapter 4, I further extended the model used in Chapter 3 and used it to identify genetic drivers of antigenic change that has occurred in the evolution of human A(H3N2) viruses. Analysing a dataset with a much higher proportion of reference viruses enabled me to modify the method of accounting for the phylogenetic structure of the dataset to allow changes in virus and antiserum reactivity associated with genetic changes occurring in branches to be estimated, in addition to the antigenic changes in viruses estimated for branches in Chapter 3. Seventeen substitutions were identified as being correlated with antigenic change at multiple points in the evolution of the virus. Candidate substitutions were identified as plausible explanations of the antigenic change identified as occurring at 24 further points in the phylogenetic tree. In most of these 24 instances, co-occurring substitutions repre-

senting alternative candidate explanations for antigenic change meant there was increased uncertainty in the identification of antigenic determinants of antigenic change, when compared with the A(H1N1) analysis described in Chapter 3. Although definitive identification of the substitutions causing antigenic change was not possible in every instance, the alternative candidate substitutions identified represent good candidates for further experimental investigation. Without experimental data that can resolve these ambiguities, a model of sequence-based prediction should reflect these ambiguities.

The predictive power of the sets of genetic determinants of antigenic variation identified in Chapters 3 and 4 was tested in Chapter 5. Various models were used to predict antigenic cross-reactivity based on amino acid sequence, including a model that accounted for the variability in antigenic impact of different amino acid substitutions to the same position and uncertainty in the identification of substitutions using Bayesian Model Averaging (BMA). To test the generalisability of these methods, predictions of virus neutralisation titres that quantify antigenic relationships among viruses of three serotypes of FMDV were also made. Sequence-based predictions proved to be similarly capable of predicting titres for unobserved virus-antiserum pairs and better able on average to predict titres for uncharacterised antiserum. More importantly, I demonstrated that the sequence-based models of prediction were capable of predicting titres for uncharacterised viruses. The ability to predict titres for emerging influenza viruses using models trained to data collected in previous years was also shown. Using the preferred BMA model these predictions were made to within one $\log_2$ titre on average. Figures 5.4 and 5.5 show that, for both A(H1N1) and A(H3N2), predictions of HI titre for emerging viruses were improved when substitutions not causing high-impact antigenic changes and therefore transitions between antigenic clusters are included in addition to those substitutions that do.

To demonstrate the impact of including substitutions of lower antigenic impact in addition to those of largest effect the titres predicted by various models are plotted against observed titres for test datasets consisting of 10% of observations in Figure 6.1. In (**A**) a simple null model was used to make predictions according to the average titre for a reference antiserum with a mean, absolute error of 1.32 $\log_2$ titre. In (**B**) this was reduced to 1.02 $\log_2$ titre by including the high-impact substitutions, K141E and ΔK130, which explain transitions between antigenic clusters shown in Figure 2.10. With this model, predictions are essentially based on whether virus and antiserum are predicted to belong to the same antigenic cluster in the antigenic map or not. In (**C**) the average error of predictions is further reduced to 0.69 $\log_2$ titre by using all substitutions identified in Chapter 3 (Table 3.1) for prediction. A final gain in accuracy was achieved in (**D**) by including terms for the reactivity of each virus that account for differences in titres related to properties of the virus such as receptor-binding avidity (mean, absolute error = 0.54 $\log_2$ titre).

Understandably, much of the focus of studies of the genetic basis of antigenic variation of

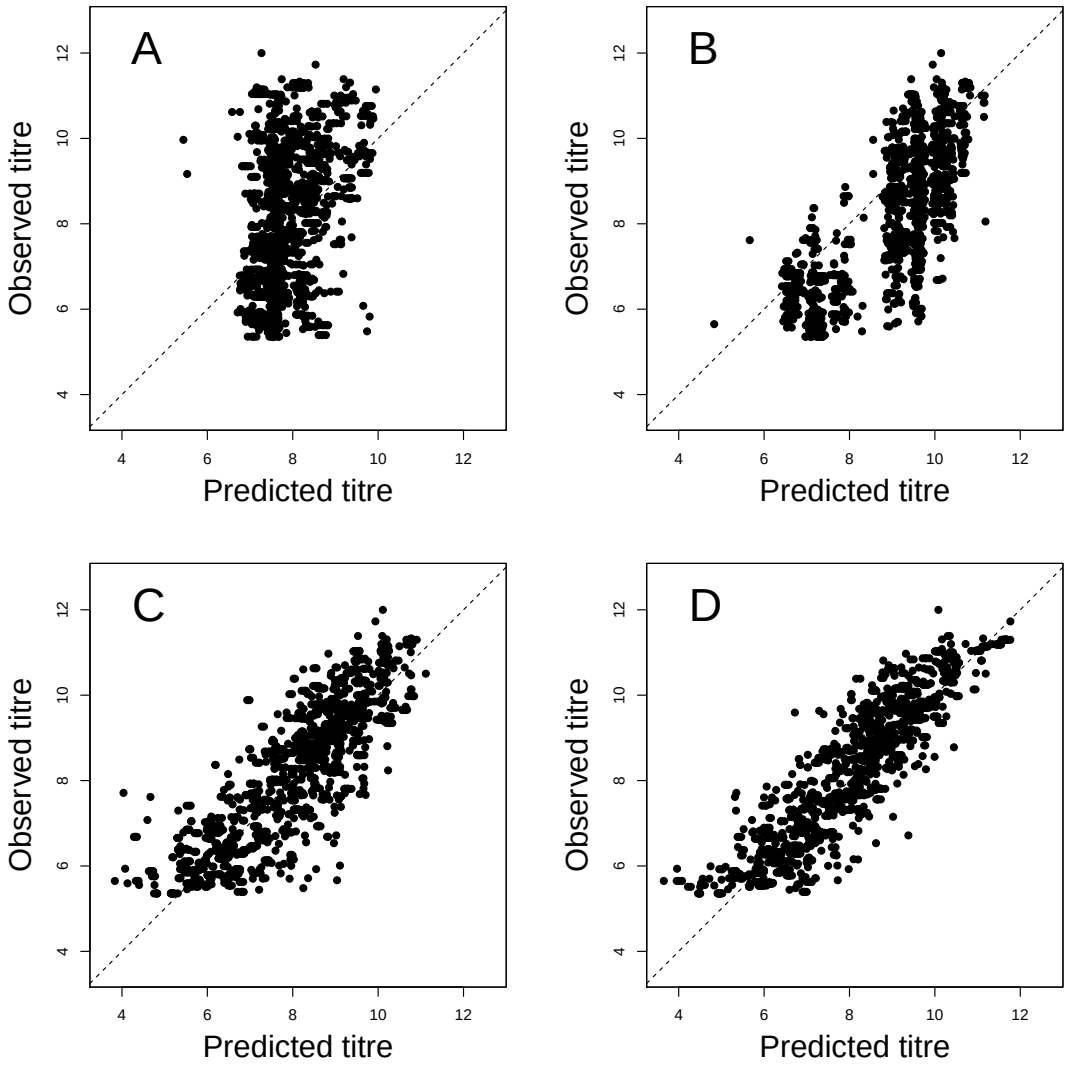

Figure 6.1: The roles of low- and high-impact HA amino acid substitutions in predicting antigenic relationships between A(H1N1) viruses: Observed and predicted HI titres plotted on log2 scale (antigenic units) using representative models trained with 90% of the data. Predictive models contained terms for (**A**) average titres for each reference virus, (**B**) antigenic cluster-defining substitutions ΔK130 and K141E, (**C**) all 18 antigenic substitution(s) shown in Table 3.1, (**D**) all 18 antigenic substitution(s) shown in Table 3.1 with additional terms that estimate differences in test virus receptor-binding avidity (non-antigenic variation in titre associated with each virus). Each model was fitted to the same training dataset comprising 90% of all observations and predictions for the remaining data are shown.

influenza viruses has been concentrated on the substitutions of highest antigenic impact (e.g. Koel *et al.*, 2013). However, the accuracies of predictions made using these alternative models demonstrate that including substitutions that cause smaller antigenic changes in addition to those that cause the largest antigenic changes significantly improves the accuracy of predictions of the HA cross-reactivity of emerging viruses. Whether or not they are also important in the wider antigenic phenotype and contribute to viral fitness remains to be tested. It is possible that by examining the accuracy of predictions of genotype survival between influenza seasons made using models that included substitutions of high-impact, low-impact, or both, questions of this nature could be resolved.

6.2 The impact of receptor-binding avidity

As stated in Chapter 1, HI titres reflect a ternary reaction between antigen, erythrocyte and antibody. The HI assay measures the ability of reference antisera to prevent binding of influenza viruses to erythrocytes and after the introduction of the HI assay, it was quickly noted that changes in viral receptor-binding avidity can impact HI titres (Archetti & Horsfall, 1950; Hirst, 1943). Understanding how substitutions that affect HA receptor-binding avidity influence HI titre is critical to the interpretation of the HI assay, and therefore the antigenic characterisation of influenza viruses. Avidity changes can potentially result in antigenic differences between viruses being accentuated or masked. There is however no standard method of accounting for the influence of receptor-binding avidity when interpreting HI data, and instead GISRS expert decision-makers rely on experience when interpreting HI data as part of the vaccine virus selection process. There is no method to correct for the impact of variation in avidity on HI titres that has been systematically tested experimentally. Archetti & Horsfall (1950) proposed a method to control for variation in receptor-binding avidity producing an antigenic distance, d between two viruses, X and Y, that is roughly independent of their receptor-binding avidities, though this measurement requires both homologous (H^{XX} and H^{YY}) and both heterologous titres (H^{XY} and H^{YX}) to be known:

$$d(X,Y) = \log_2\left(\sqrt{\frac{H^{XX}H^{YY}}{H^{XY}H^{YX}}}\right) \tag{6.1}$$

Using this Equation to control for the effect of avidity in has been shown to influence the clustering of A(H3N2) viruses using antigenic cartography (Li *et al.*, 2013). Equation 6.1 is however limited in its usefulness as it requires both homologous and both heterologous titres. As a result, Li *et al.* (2013) could only re-estimate positions in antigenic space for a small subset of viruses positioned by Smith *et al.* (2004).

In Chapters 3 and 4 model selection showed that it was necessary to account for variation

in the magnitude of HI titres measured using different viruses by including virus reactivity parameters. These reactivity parameters (A_v terms in equations used throughout chapters 3–5) may reflect differences in virus receptor-binding between viruses but could also reflect a variety of factors related to individual viruses, including their dilution if they are tested on only a single date. When plotted against year of isolation, these A_v factors did not seem to be distributed across years at random, with A(H1N1) virus A_v terms in particular showning a trend through time (Figure 6.2). This suggested that these A_v factors might reflect changes in the virus that have occurred over the sampled period of evolution.

Similar parameters reflecting differences in the reactivity of influenza viruses in HI assays were estimated by BMDS models used to construct antigenic maps for the same viruses in Chapter 2. The terms estimated by the BMDS model were strongly correlated with the A_v terms plotted in Figure 6.2 for A(H3N2) viruses ($R^2 = 0.80$). While the correlation was weaker for A(H1N1) ($R^2 = 0.54$), plotting the virus reactivity terms estimated using the BMDS model against year of isolation for the A(H1N1) viruses (Figure 6.3) revealed a very similar pattern to that shown in Figure 6.2 (the virus reactivity parameters in the BMDS model include an estimate of the baseline HI titre across all observations so appear higher on the y-axis than those in 6.2). That the virus reactivity parameters estimated by these two approaches show similar patterns indicates that the parameters are not just artefacts of either modelling approach but that they do instead reflect genuine sources of variation in the data.

To test the hypothesis that estimated A_v effects reflected some genuine aspect of HA function with an underlying genetic basis, virus reactivity parameters were modelled as a continuous trait reconstructed on the HA1 phylogeny using BEAST (Drummond *et al.*, 2012). The continuous trait model in BEAST was originally developed as a phylogeographic technique to allow geographic coordinate data to be incorporated into phylogenetic tree reconstruction, so that historical migration events could be reconstructed upon the phylogeny (Lemey *et al.*, 2010). It can also be used for ancestral reconstruction of any continuously measured phenotypic trait (e.g. five morphological traits of species of Darwin's finches including wingspan and tarsus length (Drummond *et al.*, 2012)). Under the phylogeographic scenario, a continuous trait model can indicate the relationship between geographic proximity and genetic relatedness of individuals, and allows key migration events to be mapped to branches of the phylogeny that connect hypothetical ancestors. Phylogenetic trees were generated using the same methodology described in Chapter 2 and the distribution of changes in A_v across branches of the phylogeny were drawn from a Cauchy distribution using default parameterisation.

In Figure 6.4, A_v parameters are reconstructed upon a phylogenetic tree for the 43 A(H1N1) reference viruses described in the dataset in previous chapters. This was done for reference viruses in the first instance because A_v terms estimated for these viruses are more reliable. For

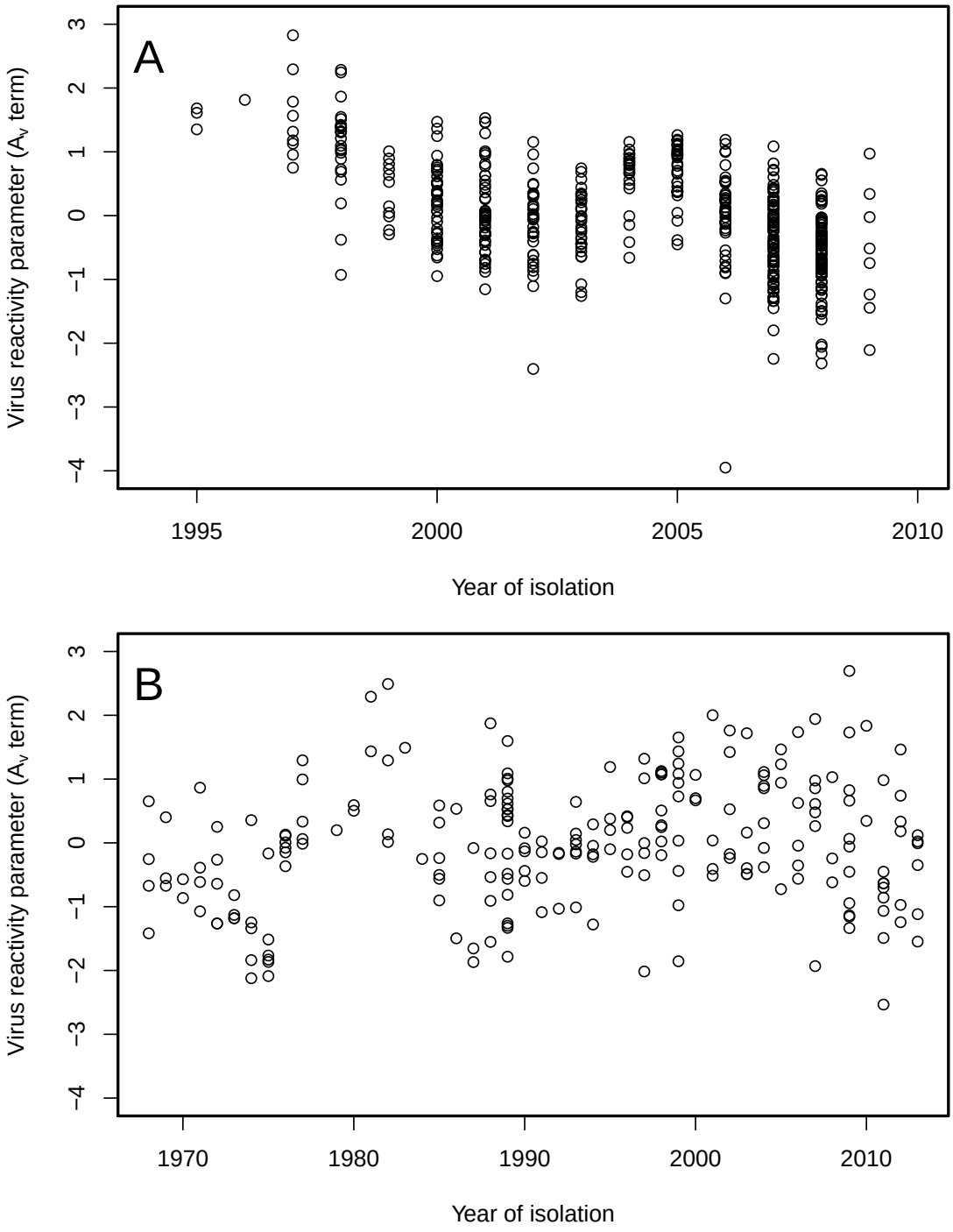

Figure 6.2: Virus reactivity parameters plotted against year of isolation: A_v parameters used to adjust titres according to the reactivity of each virus estimated using Equation 5.1 are plotted against the year of virus isolation. 502 A(H1N1) viruses are shown in (**A**) and 229 A(H3N2) viruses are shown in (**B**). A_v terms estimated for four A(H1N1) viruses isolated during the period 1978–1994 are not shown in (**A**).

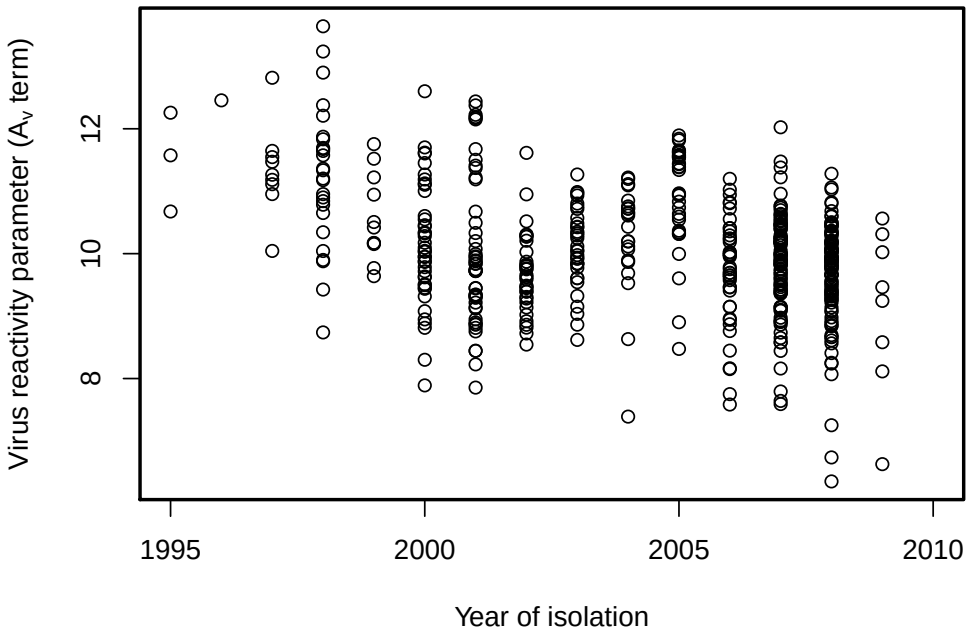

Figure 6.3: A(H1N1) virus reactivity parameters estimated by BMDS plotted against year of isolation: A_v parameters used to adjust titres according to the reactivity of each virus estimated by a Bayesian multidimensional scaling model, described in Chapter 2 are plotted against the year of virus isolation. A_v terms estimated for four A(H1N1) viruses isolated during the period 1978–1994 are not shown.

a non-reference virus in the dataset, low titres to all reference antisera that it is tested against could indicate that the virus is antigenically distinct or that it has increased avidity and that it overcomes the polyclonal antibody response as a result of this. For reference viruses with associated antiserum and known homologous titres, this ambiguity is resolved. Figure 6.4 indicates that the trend in A_v patterns through time shown in Figure 6.2 is the result of underlying genetic variation and has not arisen because of some factor (i.e. changes in experimental practice through time). For example, the A_v parameter estimated for A/Chile/4765/2000 is much more similar to those of phylogenetically related viruses in the clade close to the bottom of the tree than it is to other viruses isolated in the same year (A/Madagascar/57794/2000, A/Fukuoka/C86/2000 and A/Hong Kong/1252/2000) which are phylogenetically distant and all appear higher up in the tree as it is displayed in Figure 6.4.

Discounting A_v parameters for the oldest two viruses in the tree, which are likely to be unreliable given how antigenically and genetically distant these viruses are to each other and to all other viruses in the phylogeny, there is a decrease in A_v across the period of evolution studied. This means that generally titres are dropping as a result of non-antigenic change in virus phenotype, which could indicate an increase in avidity, which tends to result in lower titres as greater numbers of antibodies are required to overcome virus-erythrocyte binding. This could complement the model of antigenic drift proposed by Hensley *et al.*

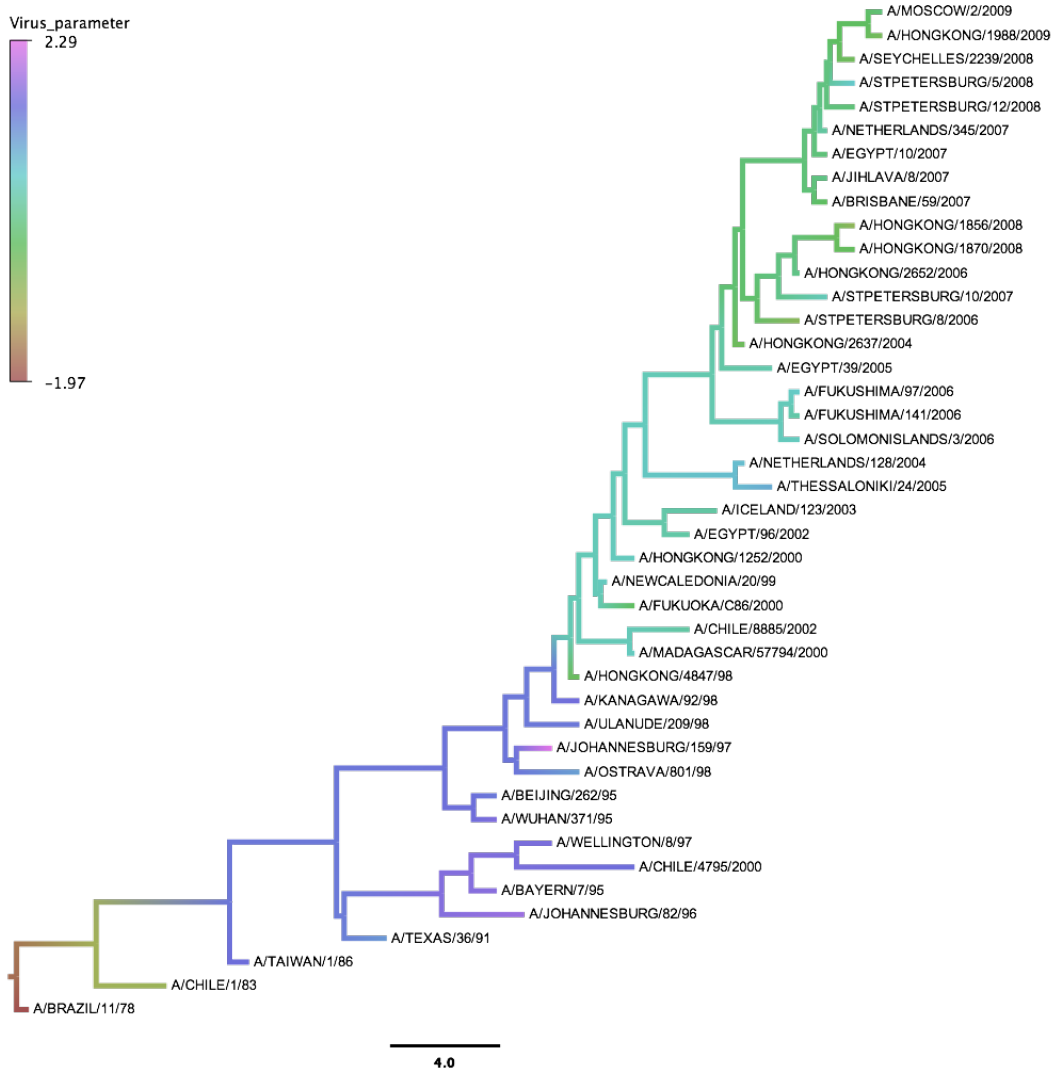

Figure 6.4: A(H1N1) virus reactivity parameters reconstructed upon HA1 phylogeny: A_v terms were estimated using Equation 5.1 for each reference virus in the A(H1N1) dataset and modelled as a continuous trait in BEAST. The colouring of branches is according to the legend at top left. This shows the A_v parameter used to adjust titres according to the reactivity of each virus and is shown as $\log_2$ titre. Branch lengths are measured in evolutionary time in terms of years rather than substitutions with scale at bottom.

(2009) introduced in Chapter 1 whereby A(H1N1) viruses escaped the polyclonal antibody response of the host population by acquiring avidity increasing substitutions. There is no obvious pattern between A_v and year of isolation in A(H3N2) and phylogenetic analyses incorporating this information are ongoing. A long term decrease in the receptor-binding avidity of A(H3N2) viruses has been observed since their emergence in the human population in 1968, though the haemagglutination assay performed prior to HI in an attempt to negate the effect of differences in receptor-binding avidity on HI titres may have succeeded in this is not reflected in A_v terms estimated.

In the same way in which historical migration events can be mapped to particular internal branches of the phylogeny when a continuous trait model is used for phylogeography, in this scenario changes in receptor-binding avidity that affect HI titres can also be. Reconstructing which substitutions that occurred in the branches in which ancestral states of A_v were estimated to change most dramatically could indicate which genetic changes are the basis of this phenotypic variation. After ignoring the oldest viruses in the phylogenetic tree, the most apparent change in A_v occurs in the trunk lineage in the area around A/Kanagawa/92/98. Substitutions mapping to the trunk lineage in the branches either side of this virus are E153G, S183P and I191L.

The development of methods to identify which amino acid substitutions have caused changes in the receptor-binding avidity of influenza, as detected in HI assays, is a potential area for future work. I have been unable here to attribute variation in receptor-binding avidity, as expressed in HI, to specific amino acid substitutions that have occurred, instead using A_v corrections for individual viruses. However the impact on HI assays of single amino acid substitutions introduced by mutagenesis was assessed in the reverse genetics experiment described in Chapter 3. In addition to quantifying the antigenic impact of introduced substitutions, differences in titres for mutant viruses, relative to parental viruses, that were not consistent with a pattern indicating antigenic change were quantified (Equation 3.1). The same method was used to estimate the antigenic and non-antigenic impact on HI titres of amino acid changes in A(H3N2) experiments carried out by Koel *et al.* (2013). This method is conceptually similar to that of Archetti & Horsfall (1950) and can identify the impact on HI titre of what appear to be changes in avidity. Substitutions with an estimated non-antigenic impact on HI titre of a magnitude that indicates a change in virus properties are shown in Table 6.1.

Table 6.1 also includes results from amino acid substitutions present in escape mutants from studies using immunised mice (Hensley *et al.*, 2009) and introduced by reverse genetics (Li *et al.*, 2013). These studies used erythrocyte binding assays in addition to HI assays to show that the increases and decreases in HI titres caused by single mutations were the result of decreases and increases in receptor-binding avidity respectively. These results validate the inferences made at H1 HA position 153 and H3 HA position 145. More generally, they

Table 6.1: HA1 amino acid substitutions identified as affecting antigenic analyses as a result of changes to receptor-binding avidity

Subtype	HA1 position*	Substitution	Charge change†	Impact on titre‡	Reference§
A(H1N1)	153 (156)	E to K	++	drop (1.28)	Chapter 3
		E to K	++	drop	Hensley *et al.* (2009)
	155 (158)	E to K	++	drop	Hensley *et al.* (2009)
	243 (246)	E to G	+	drop	Hensley *et al.* (2009)
A(H3N2)	145	K to N	−	rise (1.49)	Chapter 4
		K to N	−	rise	Li *et al.* (2013)
		N to K	+	drop (0.34)	Chapter 4
		N to K	+	drop	Li *et al.* (2013)
	155	T to Y	0	drop (0.95)	Chapter 4
		Y to T	0	rise (1.03)	Chapter 4
		H to T	−	rise (1.32)	Chapter 4
		T to H	+	drop (1.31)	Chapter 4
	156	E to K	++	drop (1.21)	Chapter 4
		K to E	−−	rise (1.10)	Chapter 4
	189	K to Q	−	drop (2.02)	Chapter 4
		Q to K	+	rise (2.21)	Chapter 4
		K to R	0	drop (2.52)	Chapter 4
	226	I to V	0	rise	Li *et al.* (2013)

* For A(H1N1), H3 HA1 numbering is shown in brackets.

† ++ (or −−) indicates substitution from a negatively charged residue to a positively charged residue (or the reverse). + (or −) indicates substitution from a negatively (positively) charged residue to a non-charged residue or from a non-charged residue to a positively (negatively) charged residue. 0 indicates no change in charge.

‡ Estimated impact on $\log_2$ titre in brackets.

§ Results from Chapter 4 are based on analysis of HI data published by Koel *et al.* (2013).

show that amino acid substitutions increasing and decreasing positive charge at the identified positions typically result in lower and higher HI titres, respectively, as a result of non-antigenic changes to virus phenotype. Changes in charge in Table 6.1 are based on side chain charge at pH 7.4. This pH reflects the conditions under which binding is initiated; the host cell membrane and the sialic acid receptor are predominantly negatively charged and therefore, generally, amino acid substitutions increasing and decreasing the positive charge of HA1 increase and reduce, respectively, the binding avidity and affinity (Kobayashi & Suzuki, 2012). After the virus has bound to sialic acid, it is then endocytosed and transported within the endosome where a declining pH (5.0–6.0) triggers a major conformational change (Mair *et al.*, 2014). The correlation between charge and receptor-binding avidity is seen across the results shown in Table 6.1 except in the case of K–Q at position 189 in A(H3N2) where a

drop in positive charge coincides with a lower HI titre and an increase in positive charge coincides with an increase in titre.

These results provide support for the role of amino acid changes that alter avidity in the escape from polyclonal antibody response, as represented in the HI assay, in both former seasonal A(H1N1) and A(H3N2) viruses. Furthermore, substitutions in the 153-157 region of A(H1N1)pdm09 viruses that influence receptor-binding specificity or avidity have been observed to contribute to apparent antigenic effects in HI (Liu *et al.*, 2010). I have shown that it is possible to quantify the antigenic and non-antigenic impacts of amino acid changes in HA and I think there is an opportunity to incorporate inference of this kind into the interpretation of HI data in a more systematic way. It is clear that changes to receptor binding avidity affect escape from polyclonal antibodies in HI assays, and laboratory studies demonstrate that changes in avidity can also allow influenza viruses to escape the immunity of vaccinated mice (Hensley *et al.*, 2009). Determining the extent to which changes in avidity contribute to escape from the polyclonal antibody response in nature is more difficult to assess.

6.3 More complex antigenic effects

There are various ways in which the complexity of the models presented could be increased to better reflect the relationships between antibodies and epitopes. A potential weakness of the models presented is an assumption that the antigenic impact of forward and reverse substitutions are the same (i.e. they are symmetric). Experimental studies have shown that this is not necessarily the case. In A(H3N2), the substitution N145K has been shown to introduce an antigenically dominant epitope. The substitution N145K does not reduce cross-reactivity to viruses possessing N145, because antibodies do not tend to bind to N145, however the substitution K145N does reduce cross-reactivity to viruses possessing K145 because viruses possessing K145 have an antigenically dominant epitope containing that residue (Li *et al.*, 2013). This asymmetric antigenic effect was seen in addition to a change in receptor binding avidity caused by substitution in both directions (Table 6.1). The asymmetric antigenic effect associated with N145K and K145N likely explains why the paired amino acid substitutions K145N and N145K appear distant from the line in Figure 4.4, while the effects of all the other pairs of forward or reverse substitutions that impact HI titres as a result of avidity changes are very well negatively correlated. The inclusion of asymmetric antigenic effects of substitutions, or any other more complex relationships of this kind, that are based on a deeper knowledge of the interactions that occur between antibodies and the epitopes which they bind to has real potential to increase the accuracy of sequence-based models such as those I present.

To my knowledge no models of the genetic basis of antigenic variation that incorporate phe-

notypic data estimate any effect of epistasis. It is possible that doing so would achieve a better understanding of the relationship between molecular and antigenic evolution. The models I have presented assume that a specific amino acid substitution occurring at a particular position will always have the same antigenic impact. Potentially models could allow for the antigenic effect of substitutions to depend on the amino acid sequence in which they arise. These kind of dependencies could also be more specific (perhaps between pairs of amino acid positions). Dependencies between amino acid positions could be made to depend on the proximity of residues on the surface of the structure of the glycoprotein or on membership of the same antigenic site defined by monoclonal antibody mapping experiments. In interpretating HI assays carried out using viruses generated by reverse genetics with single antigenically important substitutions, it was clear that the impact of a substitution on the antigenic similarity of the mutated virus to any given reference antiserum depended on existing genetic and antigenic differences between parental virus and reference virus (e.g. the variation in observed impact of substitutions in Figure 3.6). A more sophisticated model of the genetic basis of antigenic relationships between viruses could attempt to quantify how substitutions affecting antigenic similarity to specific viruses (e.g. vaccine viruses) depend on existing genetic and antigenic differences between viruses.

6.4 Future directions and final conclusions

In Chapter 5, I demonstrate that the genetic determinants of antigenicity identified in Chapters 3 and 4 can be used to predict HI titres of emerging influenza viruses from HA sequence using models trained to data collected in previous years. Antigenic cross-reactivity of the HA glycoprotein, as expressed in the HI assay, is a critical aspect of the antigenic phenotype of influenza A viruses, and consequently, an important factor contributing to the adaptive phenotype and virus fitness more generally. For this reason, I think it is possible that the approaches described in previous chapters could be used to complement and refine existing methods for predicting the evolutionary success of influenza viruses from genotype.

In previous chapters I describe the model presented by Łuksza & Lässig (2014) that predicts the frequency of HA genotypes in one season based on their frequency within the virus population in the previous season and a fitness score inferred from HA genotype. Amino acid substitutions that arise define clades of viruses on the HA phylogeny. The fitness score for a clade rewards antigenic novelty relative to other viruses in the tree, which is approximated by a count of amino acid substitutions in the sets of positions in antigenic sites A–D classified by Shih *et al.* (2007) and penalises substitutions that occur outside these antigenic sites on the assumption that other regions of HA are likely to be under negative selection so that protein structure and function is retained. Including quantitative estimates of the relative importance of positions within epitopes and of the antigenic impact of specific amino acid substitutions

should help to refine these techniques further. Together with my PhD supervisors, I am currently working with Marta Łuksza and Michael Lässig to explore these possibilities.

In addition to providing more accurate predictions of the evolutionary success of HA genotypes, combining the inferences made in previous chapters with this model of evolutionary success could enhance our knowledge of antigenic evolution in other ways. For example, the previous application of this model showed that counts of substitutions in H3 antigenic site E did not improve model predictions, while counts of substitutions in antigenic sites A–D did (Łuksza & Lässig, 2014). Testing the accuracy of predictions made using various subsets of the genetic determinants identified could help us to understand the roles of antigenic substitutions of low- and high-impact in the adaptive evolution of influenza viruses, while including terms for substitutions identified as affecting avidity or even finding a way of including A_v terms in the fitness score used for predictions could inform us about whether avidity effects apparent in HI are actually contributing to virus fitness or if they are actually only complicating factors that need to be accounted for when the results of HI assays are interpreted.

The collective results of previous chapters indicate an ability to quantify the effect of genetic changes on a complex aspect of the phenotype of influenza A viruses. In this chapter I have speculated on how these methods could be incorporated into various aspects of the process of which the ultimate outcome is the selection of vaccine viruses. It should also be stressed that the methods presented appear to be versatile. They were originally developed for FMDV and in Chapter 5, the most recently developed methods are re-applied to FMD viruses showing that is possible to predict VN titres. This generalisability calls for an extension to other antigenically variable pathogens, especially where antigenic variation creates difficulties for existing vaccines (e.g the orbiviruses, bluetongue, and African horse sickness, where serotype identification is important for vaccination strategies) or prevents the development of effective vaccines (e.g. hepatitis C). In particular, it would be interesting to apply these methods to viruses about which less is known of their antigenic determinants. For a virus about which far less is known of the constituents of epitopes and antigenic evolution more generally, there is potentially more to be gained from the epitope identification aspect of the methods presented.

APPENDIX A

Mean haemagglutination inhibition titres for recombinant viruses generated in Chapter 3

Mean haemagglutination inhibition titres for recombinant viruses generated in Chapter 3

Table A.1: Mean haemagglutination inhibition titres for reference viruses and recombinant viruses generated in Chapter 3 against antisera raised against reference viruses presented in Table 3.2

	Reference viruse against which antisera were raised						
	A/BAY /7/95	A/JBG /82/96	A/JBG /159/97	A/UU /209/98	A/HK /4847/98	A/NC /20/99	A/HK /1252/00
Reference virus:							
A/Bayern/7/95	1280	1016	32	20	20	20	20
A/Johannesburg/82/96	2560	2281	80	36	20	57	20
A/Johannesburg/159/97	40	40	1613	640	160	254	20
A/Ulan-Ude/209/98	26	20	202	453	71	113	20
A/Hong Kong/4847/98	20	20	45	127	72	57	20
A/New Caledonia/20/99	20	25	180	359	202	285	32
A/Hong Kong/1252/2000	20	20	20	22	25	32	202
Recombinant virus:							
A/Netherlands/1/93 (Neth93)	2792	2792	87	28	20	40	20
Neth93 ΔK130	87	104	1174	698	453	453	24
Neth93 R43L	2792	2792	80	28	20	40	20
Neth93 E153K	1660	1522	28	34	20	24	20
Neth93 D187N	1396	1660	698	20	20	28	20
Neth93 ΔK130 K141E	62	24	174	160	48	48	174
Neth93 ΔK130 E153K	28	20	57	247	80	44	20
Neth93 ΔK130 D187N	67	52	698	538	293	320	28
Neth93 ΔK130 D187V	67	80	830	761	415	415	28

Geometric mean HI titres are recorded as the reciprocal of the highest dilution of a particular antiserum that inhibited haemagglutination of a standardised concentration of erythrocytes by eight haemagglutinating units of each recombinant virus. Abbreviations in reference virus names: Bayern (BAY), Johannesburg (JBG), Ulan-Ude (UU), Hong Kong (HK) and New Caledonia (NC). A visual description of these data is provided in Figure 3.5.

APPENDIX B

Antisera used to characterise recombinant viruses in Chapter 4

Antisera used to characterise recombinant viruses in Chapter 4

Table B.1: Antisera used to characterise recombinant viruses are categorised as lacking or sharing substitutions introduced by mutagenesis

Substitution introduced by mutagenesis	Reference virus against which antisera were raised	
	Lacking introduced substitution	Sharing introduced substitution
S133D	A/Stockholm/10/85, A/Sichaun/2/87 A/Lyon/1149/91	A/Shandong/9/93, A/Lyon/2279/95
G135E	A/Stockholm/10/85, A/Sichaun/2/87 A/Shanghai/11/87, A/Shandong/9/93	A/Lyon/1149/91
G144D	A/Hong Kong/1/68	A/Hong Kong/107/71, A/Bilthoven/21793/72 A/Port Chalmers/1/73, A/Victoria/3/75
N145K	A/Netherlands/209/80, A/Stockholm/10/85 A/Sichaun/2/87, A/Shanghai/11/87 A/Hong Kong/34/90, A/Beijing/32/92 A/Shandong/9/93, A/Johannesburg/33/94	A/Beijing/353/89, A/Victoria/2/90 A/Lyon/1149/91, A/Netherlands/823/92 A/Paris/548/92, A/Lyon/2279/95 A/Nanchang/933/95, A/Wuhan/359/95 A/Brisbane/8/96, A/Sydney/5/97
S145N	A/Hong Kong/1/68, A/Hong Kong/107/71 A/Bilthoven/21793/72, A/Port Chalmers/1/73 A/Scotland/840/74, A/Victoria/3/75	A/Netherlands/209/80
H155T	A/Lyon/1149/91, A/Beijing/32/92 A/Shandong/9/93, A/Lyon/2279/95 A/Nanchang/933/95, A/Wuhan/359/95 A/Brisbane/8/96, A/Sydney/5/97 A/Netherlands/118/2001, A/Netherlands/88/2003	A/Fujian/411/2002, A/Netherlands/22/2003 A/Wisconsin/67/2005, A/Netherlands/42/2006
T155Y	A/Hong Kong/1/68, A/Hong Kong/107/71	A/Bilthoven/21793/72, A/Port Chalmers/1/73 A/Victoria/3/75
Y155H	A/Bilthoven/21793/72, A/Port Chalmers/1/73 A/Victoria/3/75, A/Netherlands/209/80 A/Netherlands/241/82, A/Philippines/2/82 A/Stockholm/10/85	A/Sichuan/2/87, A/Shanghai/11/87 A/Hong Kong/34/90, A/Lyon/1149/91
E156K	A/Netherlands/209/80, A/Stockholm/10/85, A/Sichuan/2/87, A/Shanghai/11/87 A/Beijing/353/89, A/Hong Kong/34/90 A/Victoria/2/90, A/Lyon/1149/91 A/Netherlands/823/92, A/Paris/548/92	A/Beijing/32/92, A/Shandong/9/93 A/Johannesburg/33/94, A/Lyon/2279/95
K156E	A/Bilthoven/21793/72, A/Port Chalmers/1/73 A/Victoria/3/75, A/Texas/1/77	A/Bangkok/1/79, A/Netherlands/241/82 A/Philippines/2/82, A/Stockholm/10/85 A/Sichuan/2/87
Q156H	A/Sydney/5/97, A/Netherlands/118/2001 A/Netherlands/88/2003, A/Wisconsin/67/2005	A/Fujian/411/2002, A/Netherlands/22/2003 A/Netherlands/42/2006
K156Q	A/Shandong/9/93, A/Lyon/2279/95 A/Brisbane/8/96	A/Sydney/5/97, A/Netherlands/118/2001
K158E	A/Sydney/5/97, A/Netherlands/118/2001	A/Lyon/1149/91, A/Shandong/9/93

Substitution introduced by mutagenesis	Reference virus against which antisera were raised	
	Lacking introduced substitution	Sharing introduced substitution
	A/Fujian/411/2002, A/Netherlands/22/2003 A/Netherlands/88/2003, A/Wisconsin/67/2005 A/Netherlands/42/2006	A/Johannesburg/33/94, A/Nanchang/933/95 A/Wuhan/359/95, A/Brisbane/8/96
G158E	A/Hong Kong/107/71, A/Bilthoven/21793/72 A/Port Chalmers/1/73, A/Victoria/3/75	A/Netherlands/209/80, A/Stockholm/10/85
S159Y	A/Bilthoven/21793/72, A/Port Chalmers/1/73 A/Victoria/3/75, A/Netherlands/209/80 A/Netherlands/241/82, A/Philippines/2/82	A/Stockholm/10/85, A/Shanghai/11/87 A/Hong Kong/34/90, A/Lyon/1149/91
G172D	A/Netherlands/209/80, A/Netherlands/241/82 A/Philippines/2/82, A/Stockholm/10/85 A/Sichaun/2/87	A/Bilthoven/21793/72, A/Port Chalmers/1/73 A/Victoria/1/77
K189R	A/Victoria/3/75, A/Netherlands/209/80 A/Netherlands/241/82, A/Philippines/2/82 A/Stockholm/10/85	A/Sichuan/2/87, A/Shanghai/11/87 A/Hong Kong/34/90, A/Lyon/1149/91
Q189K	A/Hong Kong/1/68, A/Hong Kong/107/71 A/Bilthoven/21793/72, A/Port Chalmers/1/73	A/Victoria/3/75, A/Netherlands/209/80
R189S	A/Sichaun/2/87, A/Shanghai/11/87 A/Hong Kong/34/90, A/Lyon/1149/91 A/Netherlands/823/92, Beijing/32/92 A/Paris/548/92	A/Shandong/9/93, A/Johannesburg/33/94 A/Lyon/2279/95
D193N		A/Hong Kong/107/71, A/Port Chalmers/1/73 A/Netherlands/209/80, A/Stockholm/10/85
N193S	A/Stockholm/10/85, A/Sichuan/2/87 A/Shanghai/11/87	A/Hong Kong/34/90, A/Lyon/1149/91 A/Beijing/32/92, A/Shandong/9/93 A/Johannesburg/33/94, A/Lyon/2279/95
V196A	A/Lyon/1149/91, A/Shandong/9/93 A/Lyon/2279/95, A/Brisbane/8/96	A/Sydney/5/97, A/Netherlands/118/2001
Q197R	A/Bilthoven/21793/72, A/Port Chalmers/1/73 A/Victoria/3/75	A/Netherlands/241/82, A/Philippines/2/82 A/Stockholm/10/85, A/Sichuan/2/87
I217V	A/Hong Kong/1/68, A/Hong Kong/107/71 A/Bilthoven/21793/72, A/Port Chalmers/1/73 A/Netherlands/209/80, A/Netherlands/241/82 A/Philippines/2/82, A/Stockholm/10/85 A/Sichuan/2/87	A/Victoria/3/75
T262N	A/Netherlands/209/80, A/Stockholm/10/85 A/Sichuan/2/87, A/Beijing/353/89 A/Victoria/2/90, A/Lyon/1149/91 A/Netherlands/823/92, A/Paris/548/92	A/Shanghai/11/87, A/Hong Kong/34/90 A/Beijing/32/92, A/Shandong/9/93 A/Johannesburg/33/94, A/Lyon/2279/95
N276K	A/Shandong/9/93, A/Lyon/2279/95 A/Brisbane/8/96	A/Sydney/5/97, A/Netherlands/118/2001
I278S	A/Hong Kong/1/68, A/Hong Kong/107/71 A/Bilthoven/21793/72, A/Port Chalmers/1/73	A/Victoria/3/75, A/Netherlands/209/80

BIBLIOGRAPHY

Bibliography

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**; pages 716–723. 67, 92

Alexandersen, S., Zhang, Z., Donaldson, A. I., & Garland, A. J. M. (2003). The pathogenesis and diagnosis of foot-and-mouth disease. *Journal of Comparative Pathology*, **129**; pages 1–36. 113

Archetti, I. & Horsfall, F. L. (1950). Persistent antigenic variation of influenza A viruses after incomplete neutralization *in ovo* with heterologous immune serum. *Journal of Experimental Medicine*, **92**; pages 441–462. 16, 27, 112, 145, 150

Bari, F. D., Parida, S., Asfor, A. S., Haydon, D. T., Reeve, R., Paton, D. J., & Mahapatra, M. (2015). Prediction and characterization of novel epitopes of serotype A foot-and-mouth disease viruses circulating in East Africa using site-directed mutagenesis. *Journal of General Virology*, **96**; pages 1033–1041. 117, 119

Bari, F. D., Parida, S., Tekleghiorghis, T., Dekker, A., Sangula, A., Reeve, R., Haydon, D. T., Paton, D. J., & Mahapatra, M. (2014). Genetic and antigenic characterisation of serotype A FMD viruses from East Africa to select new vaccine strains. *Vaccine*, **32**; pages 5794–5800. 116

Barr, I. G., McCauley, J., Cox, N., Daniels, R., Engelhardt, O. G., Fukuda, K., Grohmann, G., Hay, A., Kelso, A., Klimov, A., Odagiri, T., Smith, D., Russell, C., Tashiro, M., Webby, R., Wood, J., Ye, Z., & Zhang, W. (2010). Epidemiological, antigenic and genetic characteristics of seasonal influenza A(H1N1), A(H3N2) and B influenza viruses: basis for the WHO recommendation on the composition of influenza vaccines for use in the 2009-2010 Northern Hemisphere season. *Vaccine*, **28**; pages 1156–1167. 17, 25, 114

Barr, I. G., Russell, C., Besselaar, T. G., Cox, N. J., Daniels, R. S., Donis, R., Engelhardt, O. G., Grohmann, G., Itamura, S., Kelso, A., McCauley, J., Odagiri, T., Schultz-Cherry, S., Shu, Y., Smith, D., Tashiro, M., Wang, D., Webby, R., Xu, X., Ye, Z., & Zhang, W. (2014). WHO recommendations for the viruses used in the 2013-2014 Northern Hemisphere influenza vaccine: Epidemiology, antigenic and genetic characteristics of influenza A(H1N1)pdm09, A(H3N2) and B influenza viruses collected from October 2012 to January 2013. *Vaccine*, **32**; pages 4713–4725. 7, 8, 15, 17, 25, 34, 66, 89, 114

Bates, D., Maechler, M., & Bolker, B. (2012). Linear mixed-effects models using S4 classes. http://lme4.r-forge.r-project.org/. 67

Bedford, T., Cobey, S., Beerli, P., & Pascual, M. (2010). Global migration dynamics underlie evolution and persistence of human influenza A (H3N2). *PLoS Pathogens*, **6**; page e1000,918. 53

Bedford, T., Cobey, S., & Pascual, M. (2011). Strength and tempo of selection revealed in viral gene genealogies. *BMC Evolutionary Biology*, **11**; page e220. 20, 21, 23

Bedford, T., Riley, S., Barr, I. G., Broor, S., Chadha, M., Cox, N. J., Daniels, R. S., Gunasekaran, C. P., Hurt, A. C., Kelso, A., Klimov, A., Lewis, N. S., Li, X., McCauley, J. W., Odagiri, T., Potdar, V., Rambaut, A., Shu, Y., Skepner, E., Smith, D. J., Suchard, M. A., Tashiro, M., Wang, D., Xu, X., Lemey, P., & Russell, C. A. (2015). Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature*, **523**; pages 217–220. 6, 13, 21, 61, 89

Bedford, T., Suchard, M. A., Lemey, P., Dudas, G., Gregory, V., Hay, A. J., McCauley, J. W., Russell, C. A., Smith, D. J., & Rambaut, A. (2014). Integrating influenza antigenic dynamics with molecular evolution. *eLife*, **3**; page e01,914. 25, 36, 39, 40, 53, 56, 61, 89, 91, 111, 124, 137

Belongia, E. A., Kieke, B. A., Donahue, J. G., Greenlee, R. T., Balish, A., Foust, A., Lindstrom, S., & Shay, D. K. (2009). Effectiveness of inactivated influenza vaccines varied substantially with antigenic match from the 2004-2005 season to the 2006-2007 season. *The Journal of Infectious Diseases*, **199**; pages 159–167. 7

Belshe, R. B., Gruber, W. C., Mendelman, P. M., Mehta, H. B., Mahmood, K., Reisinger, K., Treanor, J., Zangwill, K., Hayden, F. G., Bernstein, D. I., Kotloff, K., King, J., Piedra, P. a., Block, S. L., Yan, L., & Wolff, M. (2000). Correlates of immune protection induced by live, attenuated, cold-adapted, trivalent, intranasal influenza virus vaccine. *The Journal of Infectious Diseases*, **181**; pages 1133–1137. 17

Bhatt, S., Holmes, E. C., & Pybus, O. G. (2011). The genomic rate of molecular adaptation of the human influenza A virus. *Molecular Biology and Evolution*, **28**; pages 2443–2451. 23

Böttcher, C., Ludwig, K., Herrmann, A., van Heel, M., & Stark, H. (1999). Structure of influenza haemagglutinin at neutral and at fusogenic pH by electron cryo-microscopy. *FEBS Letters*, **463**; pages 255–259. 4

Brownlee, G. G. & Fodor, E. (2001). The predicted antigenicity of the haemagglutinin of the 1918 Spanish influenza pandemic suggests an avian origin. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **356**; pages 1871–1876. 13, 39, 57, 58

Bush, R. M., Bender, C. A., Subbarao, K., Cox, N. J., & Fitch, W. M. (1999a). Predicting the evolution of human influenza A. *Science*, **286**; pages 1921–1925. 13, 29

Bush, R. M., Fitch, W. M., Bender, C. a., & Cox, N. J. (1999b). Positive selection on the H3 hemagglutinin gene of human influenza virus A. *Molecular Biology and Evolution*, **16**; pages 1457–1465. 21, 22, 34, 62, 89

Caton, A. J., Brownlee, G. G., Yewdell, J. W., & Gerhard, W. (1982). The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell*, **31**; pages 417–427. 8, 10, 28, 34, 74

Chen, W., Calvo, P. A., Malide, D., Gibbs, J., Schubert, U., Bacik, I., Basta, S., O'Neill, R., Schickli, J., Palese, P., Henklein, P., Bennink, J., & Yewdell, J. W. (2001). A novel influenza A virus mitochondrial protein that induces cell death. *Nature Medicine*, **7**; pages 1306–1312. 3

Chu, C. M., Andrewes, C. H., & Gledhill, A. W. (1950). Influenza in 1948-1949. *Bulletin of the World Health Organization*, **3**; pages 187–214. 7

Chu, C. M., Dawson, I. M., & Elford, W. J. (1949). Filamentous forms associated with newly isolated influenza virus. *The Lancet*, **1**; pages 602–603. 3

Colman, P. M. (1997). Virus versus antibody. *Structure*, **5**; pages 591–593. 12

Cox, T. F. & Cox, M. A. A. (2000). *Multidimensional Scaling*. Chapman and Hall/CRC, second edition. 24

Dang, C. C., Le, Q. S., Gascuel, O., & Le, V. S. (2010). FLU, an amino acid substitution model for influenza proteins. *BMC Evolutionary Biology*, **10**; page e99. 67

Daniels, R. S., Douglas, A., Skehel, J. J., Wiley, D., Naeve, W., Webster, R., Rogers, G., & Paulson, J. (1984). Antigenic analyses of influenza virus haemagglutinins with different receptor-binding specificities. *Virology*, **138**; pages 174–177. 12, 16, 34, 85

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**; pages 772–772. 36

Das, S. R., Hensley, S. E., David, A., Schmidt, L., Gibbs, J. S., Puigbò, P., Ince, W. L., Bennink, J. R., & Yewdell, J. W. (2011). Fitness costs limit influenza A virus hemagglutinin glycosylation as an immune evasion strategy. *Proceedings of the National Academy of Sciences of the United States of America*, **108**; pages E1417–E1422. 14

Davies, V., Reeve, R., Harvey, W., Maree, F. F., & Husmeier, D. (2014). Sparse Bayesian variable selection for the identication of antigenic variability in the foot-and-mouth disease virus. *Journal of Machine Learning Workshop and Conference Proceedings*, **33**; pages 149–158. 67, 109

de Jong, J. C., Beyer, W. E. P., Palache, A. M., Rimmelzwaan, G. F., & Osterhaus, A. D. (2000). Mismatch between the 1997/1998 influenza vaccine and the major epidemic A(H3N2) virus strain as the cause of an inadequate vaccine-induced antibody response to this strain in the elderly. *Journal Of Medical Virology*, **61**; pages 94–99. 7

Deem, M. W. & Pan, K. (2009). The epitope regions of H1-subtype influenza A, with application to vaccine efficacy. *Protein Engineering, Design & Selection*, **22**; pages 543–546. 28

Delport, W., Poon, A. F. Y., Frost, S. D. W., & Kosakovsky Pond, S. L. (2010). Datamonkey 2010: A suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*, **26**; pages 2455–2457. 38

Denwood, M. J. (2013). runjags: Interface utilities for MCMC models in Just Another Gibbs Sampler (JAGS) using parallel and distributed computing methods. http://cran.r-project.org/web/packages/runjags/index.html. 119

Drummond, A. J., Ho, S. Y. W., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**; page e88. 37, 42

Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22**; pages 1185–1192. 42

Drummond, A. J. & Suchard, M. A. (2010). Bayesian random local clocks, or one rate to rule them all. *BMC Biology*, **8**; page e114. 37

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**; pages 1969–1973. 20, 36, 67, 91, 117, 124, 146

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**; pages 1792–1797. 36, 66, 91

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, **17**; pages 368–376. 19

Felsenstein, J. (1985). Phylogenenies and the comparative method. *The American Naturalist*, **125**; pages 1–15. 28

Fitch, W. M., Bush, R. M., Bender, C. A., & Cox, N. J. (1997). Long term trends in the evolution of H(3) HA1 human influenza type A. *Proceedings of the National Academy of Sciences of the United States of America*, **94**; pages 7712–7718. 8, 20, 21, 73, 85

Fleming, D. M. & Elliot, A. J. (2008). Lessons from 40 years' surveillance of influenza in England and Wales. *Epidemiology and Infection*, **136**; pages 866–875. 6

Fry, E. E., Newman, J. W. I., Curry, S., Najjam, S., Jackson, T., Blakemore, W., Lea, S. M., Miller, L., Burman, A., King, A. M. Q., & Stuart, D. I. (2005). Structure of foot-and-mouth disease virus serotype A1061 alone and complexed with oligosaccharide receptor: Receptor conservation in the face of antigenic variation. *Journal of General Virology*, **86**; pages 1909–1920. 112

Gambaryan, A., Marinina, V., & Tuzikov, A. (1998). Effects of host-dependent glycosylation of hemagglutinin on receptor-binding properties of H1N1 human influenza A virus grown in MDCK cells and in embryonated eggs. *Virology*, **177**; pages 170–177. 13

Gambaryan, A. S., Robertson, J. S., & Matrosovich, M. N. (1999). Effects of egg-adaptation on the receptor-binding properties of human influenza A and B viruses. *Virology*, **258**; pages 232–239. 37, 66

Gamblin, S. J., Haire, L. F., Russell, R. J., Stevens, D. J., Xiao, B., Ha, Y., Vasisht, N., Steinhauer, D. A., Daniels, R. S., Elliot, A., Wiley, D. C., & Skehel, J. J. (2004). The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science*, **303**; pages 1838–1842. 12, 39, 50, 58, 59, 68, 75

Gamblin, S. J. & Skehel, J. J. (2010). Influenza hemagglutinin and neuraminidase membrane glycoproteins. *Journal of Biological Chemistry*, **285**; pages 28,403–28,409. 3, 11

Garten, R. J., Davis, C. T., Russell, C. A., Shu, B., Lindstrom, S., Balish, A., Sessions, W. M., Xu, X., Skepner, E., Deyde, V., Okomo-Adhiambo, M., Gubareva, L., Barnes, J., Smith, C. B., Emery, S. L., Hillman, M. J., Rivailler, P., Smagala, J., de Graaf, M., Burke, D. F., Fouchier, R. A. M., Pappas, C., Alpuche-Aranda, C. M., López-Gatell, H., Olivera, H., López, I., Myers, C. A., Faix, D., Blair, P. J., Yu, C., Keene, K. M., Dotson Jr., P. D., Boxrud, D., Sambol, A. R., Abid, S. H., St George, K., Bannerman, T., Moore, A. L., Stringer, D. J., Blevins, P., Demmler-Harrison, G. J., Ginsberg, M., Kriner, P., Waterman, S., Smole, S., Guevara, H. F., Belongia, E. A., Clark, P. A., Beatrice, S. T., Donis, R., Katz, J., Finelli, L., Bridges, C. B., Shaw, M. W., Jernigan, D. B., Uyeki, T. M., Smith, D. J., Klimov, A. I., & Cox, N. J. (2009). Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans. *Science*, **325**; pages 197–201. 5, 25, 27

Grabenstein, J. D., Pittman, P. R., Greenwood, J. T., & Engler, R. J. M. (2006). Immunization to protect the US Armed Forces: Heritage, current practice, and prospects. *Epidemiologic Reviews*, **28**; pages 3–26. 6

Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**; pages 696–704. 36

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**; pages 160–174. 20

Hay, A. J., Gregory, V., Douglas, A. R., & Lin, Y. P. (2001). The evolution of human influenza viruses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **356**; pages 1861–1870. 2, 6

Hennig, C. (2015). fpc: Flexible Procedures for Clustering. R package version 2.1-10. URL http://CRAN.R-project.org/package=fpc. 40

Hensley, S. E., Das, S. R., Bailey, A. L., Schmidt, L. M., Hickman, H. D., Jayaraman, A., Viswanathan, K., Raman, R., Sasisekharan, R., Bennink, J. R., & Yewdell, J. W. (2009). Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. *Science*, **326**; pages 734–736. 12, 16, 22, 34, 53, 84, 85, 89, 109, 138, 148, 150, 151, 152

Hirst, G. K. (1942). The quantitative determination of influenza virus and antibodies by means of red cell agglutination. *Journal of Experimental Medicine*, **75**; pages 49–64. 15, 65

Hirst, G. K. (1943). Studies of antigenic differences among strains of influenza A by means of red cell agglutination. *Journal of Experimental Medicine*, **78**; pages 407–423. 15, 16, 17, 145

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging. *Statistical Science*, **14**; pages 382–401. 112

Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G., & Webster, R. G. (2000). A DNA transfection system for generation of influenza A virus from eight plasmids. *Proceedings of the National Academy of Sciences of the United States of America*, **97**; pages 6108–6113. 68

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**; pages 65–70. 68, 72, 93, 95

Horton, D. L., McElhinney, L. M., Marston, D. A., Wood, J. L. N., Russell, C. A., Lewis, N., Kuzmin, I. V., Fouchier, R. A. M., Osterhaus, A. D. M. E., Fooks, A. R., & Smith, D. J. (2010). Quantifying antigenic relationships among the Lyssaviruses. *Journal of Virology*, **84**; pages 11,841–11,848. 25, 26

Huang, C. C., Couch, G. S., Pettersen, E. F., & Ferrin, T. E. (1996). Chimera: an extensible molecular modeling application constructed using standard components. In Hunter, L. & Klein, T. E., editors, *In Pacific Symposium on Biocomputing '96*, page 724. World Scientic Publishing, Singapore. 58

Huang, J.-W., Lin, W.-F., & Yang, J.-M. (2012). Antigenic sites of H1N1 influenza virus hemagglutinin revealed by natural isolates and inhibition assays. *Vaccine*, **30**; pages 6327–6337. 28, 65, 77

Huang, S. W., Hsu, Y. W., Smith, D. J., Kiang, D., Tsai, H. P., Lin, K. H., Wang, S. M., Liu, C. C., Su, I. J., & Wang, J. R. (2009). Reemergence of enterovirus 71 in 2008 in Taiwan: Dynamics of genetic and antigenic evolution from 1998 to 2008. *Journal of Clinical Microbiology*, **47**; pages 3653–3662. 26

Hubbard, S. & Thornton, J. (1993). NACCESS, computer program. Department of Biochemistry Molecular Biology, University College London (http://www.bioinf.manchester.ac.uk/naccess/). 68

Inglis, S. C., Gething, M.-J., & Brown, C. M. (1980). Relationship between the messenger RNAs transcribed from two overlapping genes of influenza virus. *Nucleic Acids Research*, **8**; pages 3575–3589. 3

Jagger, B. W., Wise, H. M., Kash, J. C., Walters, K.-A., Wills, N. M., Xiao, Y.-L., Dunfee, R. L., Schwartzman, L. M., Ozinsky, A., Bell, G. L., Dalton, R. M., Lo, A., Efstathiou, S., Atkins, J. F., Firth, A. E., Taubenberger, J. K., & Digard, P. (2012). An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science*, **337**; pages 199–204. 3

Johnson, N. P. A. S. & Mueller, J. (2002). Updating the accounts: Global mortality of the 1918-1920 "Spanish" influenza pandemic. *Bulletin of the History of Medicine*, **76**; pages 105–115. 5

Jow, H., Boys, R. J., & Wilkinson, D. J. (2014). Bayesian identification of protein differential expression in multi-group isobaric labelled mass spectrometry data. *Statistical Applications in Genetics and Molecular Biology*, **13**; pages 531–551. 121

Jukes, T. H. & Cantor, C. R. (1969). Mammalian protein metabolism. In Munro, H. N., editor, *Mammalian Protein Metabolism*, pages 21–123. Academic Press, New York. 19

Juleff, N., Windsor, M., Lefevre, E. a., Gubbins, S., Hamblin, P., Reid, E., McLaughlin, K., Beverley, P. C. L., Morrison, I. W., & Charleston, B. (2009). Foot-and-mouth disease virus can induce a specific and rapid CD4+ T-cell-independent neutralizing and isotype class-switched antibody response in naïve cattle. *Journal of Virology*, **83**; pages 3,626–3,636. 113

Kilbourne, E. (2006). Influenza pandemics of the 20th century. *Emerging Infectious Diseases*, **12**; pages 9–14. 2

Kilbourne, E. & Smith, C. (2002). The total influenza vaccine failure of 1947 revisited: major intrasubtypic antigenic change can explain failure of vaccine in a post-World War II epidemic. *Proceedings of the National Academy of Sciences of the United States of America*, **99**; pages 10,748–10,752. 7

Kilbourne, E. D. & Murphy, J. S. (1960). Genetic studies of influenza viruses. I. Viral morphology and growth capacity as exchangeable genetic traits. Rapid *in ovo* adaptation of early passage Asian strain isolates by combination with PR8. *The Journal of Experimental Medicine*, **111**; pages 387–406. 3

Kimura, H., Abiko, C., Peng, G., Muraki, Y., Sugawara, K., Hongo, S., Kitame, F., Mizuta, K., Numazaki, Y., Suzuki, H., & Nakamura, K. (1997). Interspecies transmission of influenza C virus between humans and pigs. *Virus Research*, **48**; pages 71–79. 2

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, **217**; pages 624–626. 20

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**; pages 111–120. 19

Klimov, A. I., Garten, R., Russell, C., Barr, I. G., Besselaar, T. G., Daniels, R., Engelhardt, O. G., Grohmann, G., Itamura, S., Kelso, A., McCauley, J., Odagiri, T., Smith, D., Tashiro, M., Xu, X., Webby, R., Wang, D., Ye, Z., Yuelong, S., Zhang, W., & Cox, N. (2012). WHO recommendations for the viruses to be used in the 2012 Southern Hemisphere influenza vaccine: Epidemiology, antigenic and genetic characteristics of influenza A(H1N1)pdm09, A(H3N2) and B influenza viruses collected from February to September 2011. *Vaccine*, **30**; pages 6461–6471. 17, 25, 27, 114

Knossow, M., Daniels, R., Douglas, A., Skehel, J. J., & Wiley, D. (1984). Three-dimensional structure of an antigenic mutant of the influenza virus haemagglutinin. *Nature*, **311**; pages 678–680. 8, 10, 89

Kobayashi, Y. & Suzuki, Y. (2012). Compensatory evolution of net-charge in influenza A virus hemagglutinin. *PLoS One*, **7**; page e40,422. 151

Koel, B., Vliet, S. V. D., & Burke, D. (2014). Antigenic variation of clade 2.1 H5N1 virus is determined by a few amino acid substitutions immediately adjacent to the receptor binding site. *mBio*, **5**; pages e01,070–14. 25, 27

Koel, B. F., Burke, D. F., Bestebroer, T. M., van der Vliet, S., Zondag, G. C. M., Vervaet, G., Skepner, E., Lewis, N. S., Spronken, M. I. J., Russell, C. A., Eropkin, M. Y., Hurt, A. C., Barr, I. G., de Jong, J. C., Rimmelzwaan, G. F., Osterhaus, A. D. M. E., Fouchier, R. A. M., & Smith, D. J. (2013). Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science*, **342**; pages 976–979. 27, 34, 50, 56, 58, 59, 61, 62, 65, 90, 93, 94, 100, 103, 105, 108, 109, 134, 138, 142, 145, 150, 151

Koelle, K., Cobey, S., Grenfell, B., & Pascual, M. (2006). Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science*, **314**; pages 1898–1903. 23, 89

Koelle, K. & Rasmussen, D. A. (2014). Prediction is worth a shot. *Nature*, **507**; pages 47–48. 34

Kosakovsky Pond, S. L. & Frost, S. D. W. (2005a). Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics*, **21**; pages 2531–2533. 38

Kosakovsky Pond, S. L. & Frost, S. D. W. (2005b). Not so different after all: A comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*, **22**; pages 1208–1222. 38

Kosakovsky Pond, S. L., Frost, S. D. W., Grossman, Z., Gravenor, M. B., Richman, D. D., & Leigh Brown, A. J. (2006). Adaptation to different human populations by HIV-1 revealed by codon-based analyses. *PLoS Computational Biology*, **2**; pages 0530–0538. 38

Kosakovsky Pond, S. L., Frost, S. D. W., & Muse, S. V. (2005). HyPhy: Hypothesis testing using phylogenies. *Bioinformatics*, **21**; pages 676–679. 38

Kryazhimskiy, S., Dushoff, J., Bazykin, G. A., & Plotkin, J. B. (2011). Prevalence of epistasis in the evolution of influenza A surface proteins. *PLoS Genetics*, **7**; page e1001,301. 23, 86

Lamb, R. A. & Lai, C.-J. (1980). Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. *Cell*, **21**; pages 475–485. 3

Lamb, R. A., Lai, C.-J., & Choppin, P. W. (1981). Sequences of mRNAs derived from genome RNA segment 7 of influenza virus : Colinear and interrupted mRNAs code for overlapping proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **78**; pages 4170–4174. 3

Lea, S., Hernandez, J., Blakemore, W., Curry, S., Domingo, E., Fry, E., Abu-Ghazaleh, R., King, A., & Newman, J. (1994). The structure and antigenicity of a type C foot-and-mouth disease virus. *Structure*, **2**; pages 123–139. 113

Lee, M.-S. & Chen, J. S.-E. (2004). Predicting antigenic variants of influenza A/H3N2 viruses. *Emerging Infectious Diseases*, **10**; pages 1385–1390. 27, 28, 65

Lee, M.-S. & Yang, C.-F. (2003). Cross-reactive H1N1 antibody responses to a live attenuated influenza vaccine in children: implication for selection of vaccine strains. *The Journal of Infectious Diseases*, **188**; pages 1362–1366. 16

Lemey, P., Rambaut, A., Welch, J. J., & Suchard, M. a. (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution*, **27**; pages 1877–1885. 39, 146

Li, Y., Bostick, D. L., Sullivan, C. B., Myers, J. L., Griesemer, S. B., StGeorge, K., Plotkin, J. B., & Hensley, S. E. (2013). Single hemagglutinin mutations that alter both antigenicity and receptor binding avidity influence influenza virus antigenic clustering. *Journal of Virology*, **87**; pages 9904–9910. 12, 27, 55, 94, 108, 109, 145, 150, 151, 152

Liao, Y.-C., Lee, M.-S., Ko, C.-Y., & Hsiung, C. A. (2008). Bioinformatics models for predicting antigenic variants of influenza A/H3N2 virus. *Bioinformatics*, **24**; pages 505–512. 28

Lin, Y. P., Gregory, V., Collins, P., Kloess, J., Wharton, S., Cattle, N., Lackenby, A., Daniels, R., & Hay, A. (2010). Neuraminidase receptor binding variants of human influenza

A(H3N2) viruses resulting from substitution of aspartic acid 151 in the catalytic site: A role in virus attachment? *Journal of Virology*, **84**; pages 6769–6781. 17, 36, 66, 68

Lin, Y. P., Xiong, X., Wharton, S. A., Martin, S. R., Coombs, P. J., Vachieri, S. G., Christodoulou, E., Walker, P. A., Liu, J., Skehel, J. J., Gamblin, S. J., Hay, A. J., Daniels, R. S., & McCauley, J. W. (2012). Evolution of the receptor binding properties of the influenza A (H3N2) hemagglutinin. *Proceedings of the National Academy of Sciences of the United States of America*, **109**; pages 21,474–21,479. 4, 9, 10, 12, 17, 36, 53, 68, 69

Liu, Y., Childs, R. A., Matrosovich, T., Wharton, S., Palma, A. S., Chai, W., Daniels, R., Gregory, V., Uhlendorff, J., Kiso, M., Klenk, H.-D., Hay, A., Feizi, T., & Matrosovich, M. (2010). Altered receptor specificity and cell tropism of D222G hemagglutinin mutants isolated from fatal cases of pandemic A(H1N1) 2009 influenza virus. *Journal of Virology*, **84**; pages 12,069–12,074. 109, 152

Logan, D., Abu-Ghazaleh, R., Blakemore, W., Curry, S., Jackson, T., King, A., Lea, S., Lewis, R., Newman, J., Parry, N., Rowlands, D., Stuart, D., & Fry, E. (1993). Structure of a major immunogenic site on foot-and-mouth disease virus. *Nature*, **362**; pages 566–568. 21

Ludi, A. B., Horton, D. L., Li, Y., Mahapatra, M., King, D. P., Knowles, N. J., Russell, C. A., Paton, D. J., Wood, J. L. N., Smith, D. J., & Hammond, J. M. (2014). Antigenic variation of foot-and-mouth disease virus serotype A. *Journal of General Virology*, **95**; pages 384–392. 25, 26, 27, 124

Łuksza, M. & Lässig, M. (2014). A predictive fitness model for influenza. *Nature*, **507**; pages 57–61. 29, 34, 62, 89, 90, 108, 112, 114, 139, 153, 154

Mair, C. M., Ludwig, K., Herrmann, A., & Sieben, C. (2014). Receptor binding and pH stability - How influenza A virus hemagglutinin affects host-specific virus infection. *Biochimica et Biophysica Acta*, **1838**; pages 1153–1168. 4, 151

McDonald, J. H. & Kreitman, M. (1991). Adaptive evolution at the *Adh* locus in *Drosophila*. *Nature*, **351**; pages 652–654. 23

McDonald, N. J., Smith, C. B., & Cox, N. J. (2007). Antigenic drift in the evolution of H1N1 influenza A viruses resulting from deletion of a single amino acid in the haemagglutinin gene. *Journal of General Virology*, **88**; pages 3209–3213. 57, 58, 59, 77, 79, 85, 127, 134

Murphy, K. P. (2012). *Machine Learning: a Probabilistic Perspective*. MIT Press, Cambridge, MA. 121

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S. L., & Scheffler, K. (2013). FUBAR: A Fast, Unconstrained Bayesian AppRoximation for inferring selection. *Molecular Biology and Evolution*, **30**; pages 1196–1205. 38

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, **8**; page e1002,764. 38

Ndifon, W. (2011). New methods for analyzing serological data with applications to influenza surveillance. *Influenza and Other Respiratory Viruses*, **5**; pages 206–212. 16, 112

Neher, R. A., Russell, C. A., & Shraiman, B. I. (2014). Predicting evolution from the shape of genealogical trees. *eLife*, **3**; page e03,568. 29, 89, 90, 112, 114, 139

Nelson, M. I. & Holmes, E. C. (2007). The evolution of epidemic influenza. *Nature Reviews. Genetics*, **8**; pages 196–205. 6, 20, 73, 85

Nolan, T., Lee, M.-S., Cordova, J. M., Cho, I., Walker, R. E., August, M. J., Larson, S., Coelingh, K. L., & Mendelman, P. M. (2003). Safety and immunogenicity of a live-attenuated influenza vaccine blended and filled at two manufacturing facilities. *Vaccine*, **21**; pages 1224–1231. 16

Ohishi, K., Ninomiya, A., & Kida, H. (2002). Serological evidence of transmission of human influenza A and B viruses to Caspian seals (*Phoca caspica*). *Microbiology and Immunology*, **46**; pages 639–644. 2

Osterhaus, A., Rimmelzwaan, G., Martina, B., Bestebroer, T. M., & Fouchier, R. A. M. (2000). Influenza B virus in seals. *Science*, **288**; pages 1051–1053. 2

Osterholm, M. T., Kelley, N. S., Sommer, A., & Belongia, E. A. (2012). Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis. *The Lancet Infectious Diseases*, **12**; pages 36–44. 7

Pagel, M., Meade, A., & Barker, D. (2004). Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology*, **53**; pages 673–684. 40

Paton, D. J., Valarcher, J. F., Bergmann, I., Matlho, O. G., Zakharov, V. M., Palma, E. L., & Thomson, G. R. (2005). Selection of foot and mouth disease vaccine strains - a review. *Revue Scientifique et Technique (International Office of Epizootics)*, **24**; pages 981–993. 113

Pay, T. & Hingley, P. (1987). Correlation of 140S antigen dose with the serum neutralizing antibody response and the level of protection induced in cattle by foot-and-mouth disease vaccines. *Vaccine*, **5**; pages 60–64. 113

Pebody, R., Warburton, F., Andrews, N., Ellis, J., von Wissmann, B., Robertson, C., Yonova, I., Cottrell, S., Gallagher, N., Green, H., Thompson, C., Galiano, M., Marques, D., Gunson, R., Reynolds, A., Moore, C., Mullett, D., Pathirannehelage, S., Donati, M., Johnston, J., de Lusignan, S., McMenamin, J., & Zambon, M. (2014). Effectiveness of trivalent seasonal influenza vaccine in preventing laboratory-confirmed influenza in primary care in the United Kingdom: 2012/13 end of season results. *Eurosurveillance*, **20**; page pii=30,013. 7

Plummer, M. (2012). Just Another Gibbs Sampler v3.3.0 (JAGS): A program for analysis of Bayesian graphical models using Gibbs sampling. http://mcmc-jags.sourceforge.net. 119

R Core Team (2015). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. 40, 67, 119

Rambaut, A. & Drummond, A. J. (2009). Tracer v1.5 http://tree.bio.ed.ac.uk/software/tracer/. 36, 67, 91

Raymond, F. L., Caton, A. J., Cox, N. J., Kendal, A. P., & Brownlee, G. G. (1986). The antigenicity and evolution of influenza H1 haemagglutinin, from 1950-1957 and 1977-1983: Two pathways from one gene. *Virology*, **148**; pages 275–287. 37, 66

Reeve, R., Blignaut, B., Esterhuysen, J. J., Opperman, P., Matthews, L., Fry, E. E., de Beer, T. A. P., Theron, J., Rieder, E., Vosloo, W., O'Neill, H. G., Haydon, D. T., & Maree, F. F. (2010). Sequence-based prediction for vaccine strain selection and identification of antigenic variability in foot-and-mouth disease virus. *PLoS Computational Biology*, **6**; page e1001,027. 28, 29, 30, 65, 67, 71, 72, 85, 91, 114, 116, 117

Reeve, R., Borley, D. W., Maree, F. F., Upadhyaya, S., Lukhwareni, A., Esterhuysen, J. J., Harvey, W. T., Fry, E. E., Parida, S., Paton, D. J., & Mahapatra, M. (2016). Tracking the antigenic evolution of foot-and-mouth disease virus. *PloS One - under review*. 116, 117, 119

Rincon, V., Rodriguez-Huete, A., Lopez-Arguello, S., Ibarra-molero, B., Sanchez-ruiz, J. M., Harmsen, M. M., & Mateu, M. G. (2014). Identification of the structural basis of thermal lability of a virus provides a rationale for improved vaccines. *Structure*, **22**; pages 1560–1570. 113

Robertson, J. S., Bootman, J. S., Newman, R., Oxford, J. S., Daniels, R. S., Webster, R. G., & Schild, G. C. (1987). Structural changes in the haemagglutinin which accompany egg adaptation of an influenza A(H1N1) virus. *Virology*, **160**; pages 31–37. 37, 66

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, **20**; pages 53–65. 40

Russell, C. A., Jones, T. C., Barr, I. G., Cox, N. J., Garten, R. J., Gregory, V., Gust, I. D., Hampson, A. W., Hay, A. J., Hurt, A. C., de Jong, J. C., Kelso, A., Klimov, A. I., Kageyama, T., Komadina, N., Lapedes, A. S., Lin, Y. P., Mosterin, A., Obuchi, M., Odagiri, T., Osterhaus, A. D. M. E., Rimmelzwaan, G. F., Shaw, M. W., Skepner, E., Stohr, K., Tashiro, M., Fouchier, R. A. M., & Smith, D. J. (2008). The global circulation of seasonal influenza A (H3N2) viruses. *Science*, **320**; pages 340–346. 21

Russell, R. J., Gamblin, S. J., Haire, L. F., Stevens, D. J., Xiao, B., Ha, Y., & Skehel, J. J. (2004). HI and H7 influenza haemagglutinin structures extend a structural classification of haemagglutinin subtypes. *Virology*, **325**; pages 287–296. 60

Russell, S. & Norvig, P. (1995). Artificial intelligence: A modern approach. In *Artificial Intelligence: A Modern Approach*, chapter 4. Pearson Education Limited. 67, 92

Rweyemamu, M. M., Booth, J. C., Head, M., & Pay, T. W. (1978). Microneutralization tests for serological typing and subtyping of foot-and-mouth disease virus strains. *The Journal of Hygiene*, **81**; pages 107–123. 116

Salk, J. E. & Suriano, P. C. (1949). Importance of antigenic composition of influenza virus vaccine in protecting against the natural disease. *American Journal of Public Health*, **39**; pages 345–355. 7

Sandbulte, M. & Westgeest, K. (2011). Discordant antigenic drift of neuraminidase and hemagglutinin in H1N1 and H3N2 influenza viruses. *Proceedings of the National Academy of Sciences of the United States of America*, **108**; pages 2–7. 25

Sanner, M. F., Olson, A. J., & Spehner, J. C. (1996). Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers*, **38**; pages 305–320. 58

Sauter, N. K., Hanson, J. E., Glick, G. D., Brown, J. H., Crowther, R. L., Park, S. J., Skehel, J. J., & Wiley, D. C. (1992). Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: analysis by proton nuclear magnetic resonance spectroscopy and X-ray crystallography. *Biochemistry*, **31**; pages 9609–9621. 9, 11, 39, 50, 68, 93, 101

Scholtissek, C., Hoyningen, V. V., & Rott, R. (1978a). Genetic relatedness between the new 1977 epidemic strains (H1N1) of influenza and human influenza strains isolated between 1947 and 1957 (H1N1). *Virology*, **89**; pages 613–617. 5

Scholtissek, C., Rohde, W., Hoyningen, V. V., & Rott, R. (1978b). On the origin of the human influenza virus subtypes H2N2 and H3N2. *Virology*, **87**; pages 13–20. 5

Shih, A. C.-C., Hsiao, T.-C., Ho, M.-S., & Li, W.-H. (2007). Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **104**; pages 6283–6288. 13, 23, 30, 39, 60, 89, 153

Skehel, J. J., Stevens, D., Daniels, R. S., Douglas, A., Knossow, M., Wilson, I. A., & Wiley, D. (1984). A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody. *Proceedings of the National Academy of Sciences of the United States of America*, **81**; pages 1779–1783. 8, 10, 13, 34, 89

Skehel, J. J. & Wiley, D. C. (2000). Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin. *Annual Review of Biochemistry*, **69**; pages 531–569. 6, 8, 9, 10, 11, 13, 14, 74, 95

Smith, D. J., Lapedes, A. S., de Jong, J. C., Bestebroer, T. M., Rimmelzwaan, G. F., Osterhaus, A. D. M. E., & Fouchier, R. A. M. (2004). Mapping the antigenic and genetic evolution of influenza virus. *Science*, **305**; pages 371–376. 17, 23, 24, 25, 26, 27, 34, 39, 55, 56, 65, 67, 85, 90, 91, 93, 105, 124, 145

Smith, G. J. D., Vijaykrishna, D., Bahl, J., Lycett, S. J., Worobey, M., Pybus, O. G., Ma, S. K., Cheung, C. L., Raghwani, J., Bhatt, S., Peiris, J. S. M., Guan, Y., & Rambaut, A. (2009). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature*, **459**; pages 1122–1125. 5

Smith, W., Andrewes, C., & Laidlaw, P. (1933). A virus obtained from infuenza patients. *The Lancet*, **255**; pages 66–68. 2

Soundararajan, V., Zheng, S., Patel, N., Warnock, K., Raman, R., Wilson, I. A., Raguram, S., Sasisekharan, V., & Sasisekharan, R. (2011). Networks link antigenic and receptor-binding sites of influenza hemagglutinin: Mechanistic insight into fitter strain propagation. *Scientific Reports*, **1**; page e200. 13

Steinbrück, L. & McHardy, A. C. (2011). Allele dynamics plots for the study of evolutionary dynamics in viral populations. *Nucleic Acids Research*, **39**; pages 1–12. 23

Steinbrück, L. & McHardy, A. C. (2012). Inference of genotype-phenotype relationships in the antigenic evolution of human influenza A (H3N2) viruses. *PLoS Computational Biology*, **8**; page e1002,492. 29, 65

Strelkowa, N. & Lässig, M. (2012). Clonal interference in the evolution of influenza. *Genetics*, **192**; pages 671–682. 6, 24

Subbarao, K. & Joseph, T. (2007). Scientific barriers to developing vaccines against avian influenza viruses. *Nature Reviews. Immunology*, **7**; pages 267–278. 3, 5

Suchard, M. A., Weiss, R. E., & Sinsheimer, J. S. (2001). Bayesian selection of continuous-time Markov chain evolutionary models. *Molecular Biology and Evolution*, **18**; pages 1001–1013. 36, 42, 66, 91

Sun, H., Yang, J., Zhang, T., Long, L.-p., Jia, K., Yang, G., Webby, R. J., & Wan, X.-f. (2013). Using sequence data to infer the antigenicity of influenza virus. *mBio*, **4**; pages e00,230–13. 5, 25, 55, 61, 65, 67, 111, 142

Taubenberger, J. K. & Morens, D. M. (2006). 1918 Influenza: The mother of all pandemics. *Emerging Infectious Diseases*, **12**; pages 15–22. 5

Taubenberger, J. K., Reid, A. H., Janczewski, T. A., & Fanning, T. G. (2001). Integrating historical, clinical and molecular genetic data in order to explain the origin and virulence of the 1918 Spanish influenza virus. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **356**; pages 1829–1839. 5

Tavare, S. (1986). Some probabilistic and statistical problems on the analysis of DNA problems. *Lecture Notes on Mathematical Modelling in the Life Sciences*, **17**; pages 57–86. 20

Taylor, W. R. (1999). Protein structure comparison using iterated double dynamic programming. *Protein Science*, **8**; pages 654–665. 50, 68

Taylor, W. R. (2000). Protein structure comparison using SAP. *Methods in Molecular Biology*, **143**; pages 19–32. 50, 68

Theodoridis, S. & Koutroumbas, K. (2006). *Pattern Recognition*. Academic Press, Orlando, FL, 3rd edition. 40

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**; pages 4673–4680. 117

Thomson, G. R., Vosloo, W., & Bastos, a. D. S. (2003). Foot and mouth disease in wildlife. *Virus Research*, **91**; pages 145–161. 113

Thyagarajan, B. & Bloom, J. D. (2014). The inherent mutational tolerance and antigenic evolvability of influenza hemagglutinin. *eLife*, **3**; page e03,300. 6, 60

Tong, S., Li, Y., Rivailler, P., Conrardy, C., Castillo, D. a. A., Chen, L.-M., Recuenco, S., Ellison, J. a., Davis, C. T., York, I. a., Turmelle, A. S., Moran, D., Rogers, S., Shi, M., Tao, Y., Weil, M. R., Tang, K., Rowe, L. a., Sammons, S., Xu, X., Frace, M., Lindblade, K. a., Cox, N. J., Anderson, L. J., Rupprecht, C. E., & Donis, R. O. (2012).

A distinct lineage of influenza A virus from bats. *Proceedings of the National Academy of Sciences of the United States of America*, **109**; pages 4269–4274. 5

Tong, S., Zhu, X., Li, Y., Shi, M., Zhang, J., Bourgeois, M., Yang, H., Chen, X., Recuenco, S., Gomez, J., Chen, L.-M., Johnson, A., Tao, Y., Dreyfus, C., Yu, W., McBride, R., Carney, P. J., Gilbert, A. T., Chang, J., Guo, Z., Davis, C. T., Paulson, J. C., Stevens, J., Rupprecht, C. E., Holmes, E. C., Wilson, I. a., & Donis, R. O. (2013). New world bats harbor diverse influenza A viruses. *PLoS Pathogens*, **9**; page e1003,657. 5

Tusche, C., Steinbrück, L., & McHardy, A. C. (2012). Detecting patches of protein sites of influenza A viruses under positive selection. *Molecular Biology and Evolution*, **29**; pages 2063–2071. 22

Upadhyaya, S., Ayelet, G., Paul, G., King, D. P., Paton, D. J., & Mahapatra, M. (2014). Genetic basis of antigenic variation in foot-and-mouth disease serotype A viruses from the Middle East. *Vaccine*, **32**; pages 631–638. 117

Van Bekkum, J. (1969). Correlation between serum antibody level and protection against challenge with FMD virus. Technical report, Presented at the Meeting of the Standing Technical Comittee of the European Comission for the control of Foot-and-Mouth Disease, Brescia, Italy, Food and Agriculture Organization of the Rome United Nations,. 113

Vosloo, W., Bastos, a. D., Kirkbride, E., Esterhuysen, J. J., Janse Van Rensburg, D., Bengis, R. G., Keet, D. W., & Thomson, G. R. (1996). Persistent infection of African buffalo (*Syncerus caffer*) with SAT-type foot-and-mouth disease viruses: Rate of fixation of mutations, antigenic change and interspecies transmission. *Journal of General Virology*, **77**; pages 1457–1467. 113

Webster, R. & Laver, W. (1980). Determination of the number of nonoverlapping antigenic areas on Hong Kong (H3N2) influenza virus hemagglutinin with monoclonal antibodies and the selection of variants with potential epidemiological significance. *Virology*, **104**; pages 139–148. 8, 10, 34, 89

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., & Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiological Reviews*, **56**; pages 152–179. 5

Weis, W., Brown, J., Cusack, S., Paulson, I., Skehel, J. J., & Wiley, D. (1988). Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature*, **333**; pages 426–431. 13, 22

Whittle, J. R. R., Zhang, R., Khurana, S., King, L. R., Manischewitz, J., Golding, H., Dormitzer, P. R., Haynes, B. F., Walter, E. B., Moody, M. A., Kepler, T. B., Liao, H.-X., & Harrison, S. C. (2011). Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin. *Proceedings of the National Academy of Sciences of the United States of America*, **108**; pages 14,216–14,221. 68

WHO (1977). Influenza. *Weekly Epidemiological Record*, **52**; pages 397–404. 5

WHO (2009). WHO Influenza fact sheet. Technical report, WHO. 2

WHO (2011). Manual for the Laboratory Diagnosis and Virological Surveillance of Influenza. 35, 65, 70

WHO (2015a). Global Influenza Surveillance and Response System (GISRS) - http://www.who.int/influenza/gisrs_laboratory/en/. 6, 7

WHO (2015b). Recommended composition of influenza virus vaccines for use in the 2015-2016 northern hemisphere influenza season. *Weekly Epidemiological Record*, **9**; pages 97–108. 139

WHO (2015c). Vaccines - http://www.who.int/entity/influenza/vaccines/en/. 6

WHO (2015d). WHO recommendations on the composition of influenza virus vaccines - http://www.who.int/influenza/vaccines/virus/recommendations/en/. 7, 8

Wiley, D. C. & Skehel, J. J. (1987). The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annual Review of Biochemistry*, **56**; pages 365–394. 8, 10, 50, 60, 77

Wiley, D. C., Wilson, I. A., & Skehel, J. J. (1981). Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature*, **289**; pages 373–378. 8, 10, 11, 13, 34, 59, 89

Wilson, I. & Cox, N. (1990). Structural basis of immune recognition of influenza virus hemagglutinin. *Annual Review of Immunology*, **8**; pages 737–771. 11, 13

Wilson, I. A., Skehel, J. J., & Wiley, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature*, **289**; pages 366–373. 8, 89

Wise, H. M., Barbezange, C., Jagger, B. W., Dalton, R. M., Gog, J. R., Curran, M. D., Taubenberger, J. K., Anderson, E. C., & Digard, P. (2011). Overlapping signals for translational regulation and packaging of influenza A virus segment 2. *Nucleic Acids Research*, **39**; pages 7775–7790. 3

Wise, H. M., Foeglein, A., Sun, J., Dalton, R. M., Patel, S., Howard, W., Anderson, E. C., Barclay, W. S., & Digard, P. (2009). A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. *Journal of Virology*, **83**; pages 8021–8031. 3

Wise, H. M., Hutchinson, E. C., Jagger, B. W., Stuart, A. D., Kang, Z. H., Robb, N., Schwartzman, L. M., Kash, J. C., Fodor, E., Firth, A. E., Gog, J. R., Taubenberger, J. K., & Digard, P. (2012). Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain. *PLoS Pathogens*, **8**; page e1002,998. 3

Yang, Z. (1993). Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Molecular Biology and Evolution*, **10**; pages 1396–1401. 20

Yang, Z. (1994). Estimating the pattern of nucleotide substitution. *Journal of Molecular Evolution*, **39**; pages 105–111. 20

Yang, Z. (2000). Maximum likelihood estimation on large phylogenies and analysis of adaptive evolution in human influenza virus A. *Journal of Molecular Evolution*, **51**; pages 423–432. 21

Yewdell, J. W., Caton, A. J., & Gerhard, W. (1986). Selection of influenza A virus adsorptive mutants by growth in the presence of a mixture of monoclonal antihemagglutinin antibodies. *Journal of Virology*, **57**; pages 623–628. 8, 16, 34

Yuanji, G., Fengen, J., Ping, W., Min, W., & Jiming, Z. (1983). Isolation of influenza C virus from pigs and experimental infection of pigs with influenza C virus. *Journal of General Virology*, **34**; pages 177–182. 2

Zebedee, S. L. & Lamb, R. A. (1988). Influenza A virus M2 protein : Monoclonal antibody restriction of virus growth and detection of M2 in virions. *Journal of Virology*, **62**; pages 2762–2772. 5

Zhou, B., Ma, J., Liu, Q., Bawa, B., Wang, W., Shabman, R. S., Duff, M., Lee, J., Lang, Y., Cao, N., Nagy, A., Lin, X., Stockwell, T. B., Richt, J. A., Wentworth, D. E., & Ma, W. (2014). Characterization of uncultivable bat influenza virus using a replicative synthetic virus. *PLoS Pathogens*, **10**; page e1004,420. 5

Zhu, X., Yang, H., Guo, Z., Yu, W., Carney, P. J., Li, Y., Chen, L.-M., Paulson, J. C., Donis, R. O., Tong, S., Stevens, J., & Wilson, I. A. (2012). Crystal structures of two subtype N10 neuraminidase-like proteins from bat influenza A viruses reveal a diverged putative active site. *Proceedings of the National Academy of Sciences of the United States of America*, **109**; pages 18,903–18,908. 5

Zhu, X., Yu, W., McBride, R., Li, Y., Chen, L.-M., Donis, R. O., Tong, S., Paulson, J. C., & Wilson, I. A. (2013). Hemagglutinin homologue from H17N10 bat influenza virus exhibits divergent receptor-binding and pH-dependent fusion activities. *Proceedings of the National Academy of Sciences of the United States of America*, **110**; pages 1458–1463. 5

Zuckerkandl, E. & Pauling, L. (1965). Evolutionary divergence and convergence. In Bryson, V. & Vogel, H. J., editors, *Evolving Genes and Proteins*, pages 97–166. Academic Press, New York. 20, 36