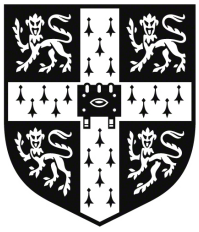
Visualisation and Analysis of Patterns in Serological Data Using "Antibody Landscapes"

Samuel H. Wilks, MA, MB BChir.

Pembroke College



This dissertation is submitted for the degree of Doctor of Philosophy.

University of Cambridge, September 2017

Supervisor:

Professor Derek J. Smith

Examiners:

Dr Simon Frost & Dr Othmar Engelhardt

Visualisation and Analysis of Patterns in Serological Data Using "Antibody Landscapes"

Samuel H. Wilks

Summary

In this thesis I develop and implement "antibody landscapes", a method to profile immunity against a pathogen as a function of antigenic differences between a range of strains. Theoretically applicable to any antigenically variable pathogen and measurement of immunity, the work here focusses on antibody-mediated immunity against the A/H3N2 influenza subtype.

Applying the methodology to study annual serum samples from individuals monitored for influenza infection over a period of six years, patterns of influenza immunity were found to be remarkably distinct and maintained almost unchanged over time in the absence of influenza exposure. Upon infection, the initial response is strikingly antigenically broad, including responses against viruses far beyond the extent of cross-reactivity observed after a primary infection.

Analysis of two vaccination cohorts, one receiving an antigenically advanced vaccine strain and one a more typical vaccine strain choice, revealed many of the same patterns of response as seen with infection. Antigenically advanced vaccination generated greater responses against later strains but surprisingly, due to equivalent boosting of prior immunity, this came at no cost to responses generated against contemporary or older strains.

Exploring in more detail the development of immunity over time, analysis of a cohort of children demonstrated that - in contrast to adults with diverse exposure histories - antibody responses to a first infections were remarkably similar in pattern and magnitude. Interestingly, for second infections, although post-infection antibody titres against circulating strains were comparable to those after first infections, overall cross-reactivity of the response against future antigenic variants appeared to be diminished.

The findings here underline the significant role prior-immunity plays in affecting the response to new exposures and the importance of understanding it. An important conclusion is that by failing to account for it, current approaches to influenza vaccine strain selection may be suboptimal and pre-emptive vaccine strain updates may improve overall vaccine efficacy where immunity to current strains already exists in the population. Building on the work presented here should help to optimise strain choice and vaccine efficacy even further.

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. This thesis totals 252 pages (including bibliography and appendices) and does not exceed the prescribed word limit.

Cambridge, September 2017

Sam Wilks

Publications

Some ideas and figures have appeared previously in the following publications:

Fonville, J. M.*, Wilks, S. H.*, James, S. L., Fox, A., Ventresca, M., Aban, M., et al. (2014). *Antibody landscapes after influenza virus infection or vaccination.* Science, **346**(6212), 996-1000.

Gordon, D. L., Sajkov, D., Honda-Okubo, Y., Wilks, S. H., Aban, M., Barr, I. G., & Petrovsky, N. (2016). *Human Phase 1 trial of low-dose inactivated seasonal influenza vaccine formulated with Advax^{TM} delta inulin adjuvant.* Vaccine, **34**(33), 3780-3786.

Fonville, J. M., Fraaij, P. L. A., de Mutsert, G., Wilks, S. H., van Beek, R., Fouchier, R. A. M., & Rimmelzwaan, G. F. (2015). *Antigenic maps of influenza A/H3N2 virus produced with human antisera obtained after primary infection.* Journal of Infectious Diseases, **213**(1), 31-38.

* Equal contribution of authors

Acknowledgements

The analyses presented in this thesis were performed independently but I am very grateful to my supervisor Derek Smith, who came up with concept of antibody landscapes, and to colleagues Judy Fonville, Mario Ventresca and Sarah James with whom I have at various times collaborated in the analysis of the infection and vaccination data in chapters 3 & 4, and in writing up results for publication. Judy Fonville in particular did a lot of work to extend the H3N2 antigenic map, upon which the antibody landscapes analyses are based, and came up with the idea of taking a 2D summary path to better visualise the landscapes, a key step.

I have been incredibly fortunate to have had the opportunity to work with some extraordinary sources of data (and some extraordinary people) and for that I give special thanks to all the collaborating groups. The ferret HI data was provided by Erasmus Medical Center, University of Rotterdam, and all the human HI data was produced by collaborators at the Victorian Infectious Disease Reference Laboratory, Melbourne. I owe a particular debt of thanks to the technicians Malet Aban and Lumin Xue for the many hours they spent patiently titrating thousands of samples, without their work, this thesis could not exist. The serum samples and PCR data analysed in chapter 3 were collected by collaborators at the Oxford University Clinical Research Unit based in Hanam, Vietnam, and the serum samples, PCR data and genetic data in chapter 5, by collaborators at the Sustainable Sciences Institute based in Managua, Nicaragua and the University of Michigan. Aubree Gordon, Amy Brandt, Brenda Lopez, Raquel Burger-Calderon, Mike Binder, Ian Barr, Aeron Hurt, Annette Marsh, Baz Nadjm and Ana Bonell have been especially generous with their time and with hosting me on visits to come and meet them. To the many many more people involved in running the various cohorts, and to all the participants who have offered their time (and blood!), a sincere thank you.

To my friends and the other people I have worked with, David Labonte, Terry Jones, Barbara Muehlemann, David Pattinson, Leah Katzelnick, David Burke, Michael Gabel, Pete Kumberger, Neha Thakre, Paola Carrillo-Bustamente, Peter Dudfield, Christian Guyader, Jack Tavener & Ed Monk, I feel lucky to have had you there for all the support, coffee breaks, cake breaks and discussions both personal and professional, you've given me insights in many different directions and made the past few years very happy ones. To my thesis advisors Eoin McKinney and Walter Federle, you have been simply excellent and your feedback was always very helpful. To Frederik Graw especially, thank you for including me so fully, for the introduction into mathematical modelling, for your considered input into my work and for generally being an all-round amazingly patient and good-natured person.

Thanks also to my examiners Othmar Engelhardt and Simon Frost for engaging so fully with the work. I'm hugely appreciative of the time they invested and the helpful and insightful comments than they were able to offer between them, this thesis has certainly improved as a result of their input.

During the course of my phD I have received funding from the Medical Research Council for which I am very thankful, while thanks also go to the University of Cambridge and Heidelberg University for the infrastructure they have provided. My college, Pembroke, has offered the perfect mixture of support and freedom that has allowed me to grow as a person and take advantage of the opportunities I have had. The same is true of my family, Philip, Sally and Tom Wilks, who I feel ever luckier to have.

Finally, I would like to thank again my supervisor Derek Smith on a personal note. From the first time I worked with him as a summer student, he has shown me incredible faith and given unwavering support in all the directions I have taken in my work and life generally. He is a fantastic source of inspiration, a terrible role model, a constant reminder of the privilege and fun of research, and a demonstration of what can be achieved when rigorous scientific approaches are applied to the right questions. And he is the reason I am a scientist.

Contents

List of Figures

1

Introduction

1.1 Chapter summary

The work in this thesis touches on aspects of many areas including the drivers and effect of influenza antigenic evolution, the role and benefits of influenza vaccination, the nature of the serological immune response and the effect of exposure history on future encounters with antigenically variable pathogens. In this chapter I provide a brief introduction to these topics, covering in particular the aspects that will become relevant in later discussions throughout this thesis. Where additional coverage of a particular subject is desired, I have tried to point in the direction of sources covering these areas in more detail.

1.2 Influenza as a study organism

A substantial proportion of infectious disease today is caused by antigenically variable pathogens, which escape immunity induced by prior infection or vaccination by changing the molecular structure of antigens targeted by the immune system. Of the many examples, the human influenza virus is amongst the most studied, yet despite decades of concerted effort influenza remains a huge burden on the global population. Annual seasonal influenza epidemics are estimated to be the cause of over 500 thousand deaths annually, alongside far greater associated rates of morbidity and huge economic costs [Kassebaum et al., 2016; Molinari et al., 2007].

In the context of such a prevalent disease, the benefits to be gained from a more detailed understanding of the immune response to influenza virus infection are clear. Additionally, the fact that influenza has been so heavily researched means that it acts as an ideal study organism for the development of techniques and investigation of immunological phenomena that are relevant to a multitude of other antigenically variable pathogens about which less may be understood. Indeed, through the work of countless individuals, we now understand much about the antigenic evolution of the influenza virus, the molecular basis for antigenic change, key determinants of influenza virus immunity and ways to enhance it through vaccination [Koel et al., 2013; Lambert and Fauci, 2010; Russell et al., 2008; Smith et al., 2004; Webster et al., 1992]. It is only because of this accumulated understanding in the particular field of influenza that the work in this thesis was possible and I hope I can add to it as part of the longer term goal to deepen understanding of our own immune system and in the process to uncover more successful ways to protect the global population against such pathogens.

1.3 The influenza virus

Influenza viruses are negative-sense-single-strand RNA viruses belonging to the family Orthomyxoviridae. They are grouped into four types, A-D [CDC, 2017], but of these types only A and B are responsible for significant disease in humans. Unusually for a virus, the genome is segmented - types A and B for example contain eight pieces of RNA, encoding for 11 proteins and 17 proteins respectively. Table 1 briefly describes the functions and separation of proteins between gene segments in influenza A, while figure 1.1 shows an overview of the viral replication lifecycle, highlighting the immunological targets that have been characterised [Lambert and Fauci, 2010].

Gene segment	Protein(s)	Protein function
Polymerase B1 (PB1)	Polymerase basic protein 1 (PB1)	Internal protein, virus replication
	PB1-F2	Mitochondrial targeting and apoptosis
	PB1-N40	Function unknown
Polymerase B2 (PB2)	Polymerase basic protein 2 (PB2)	Internal protein, virus replication
Polymerase A (PA)	Polymerase acidic protein (PA)	Internal protein, virus replication
	PA-X	Degradation of host mRNA and suppression of protein synthesis
	PA-N155, PA-N182	Function unknown
Haemagglutinin (HA)	Haemagglutinin (HA)	Surface glycoprotein, viral attachment
Nucleoprotein (NP)	Nucleoprotein (NP)	Nucleocapsid protein, RNA coating, nuclear targeting, RNA transcription
Neuraminidase (NA)	Neuraminidase (NA)	Surface glycoprotein, viral release from host cells
Matrix (M)	Matrix protein 1 (M1)	Membrane protein stability
	Matrix protein 2 (M2)	Membrane protein, viral uncoating
	Matrix protein 42 (M42)	M2 proton channel replacement
Non-structural (NS)	Non-structural protein 1 (NS1)	Internal proteins
	Non-structural protein 2 (NS2)	Regulation of virus life cycle, especially mRNA transcription and localisation of viral ribonucleic proteins
	Non-structural protein 3 (NS3)	Adaptation of avian influenza virus to new hosts

Table 1: List of Influenza A virus gene segments and encoded proteins and their functions. (see Webster et al. [1992] for an in-depth introduction and Yamayoshi et al. [2016] for an overview of more recently discovered viral accessory proteins)

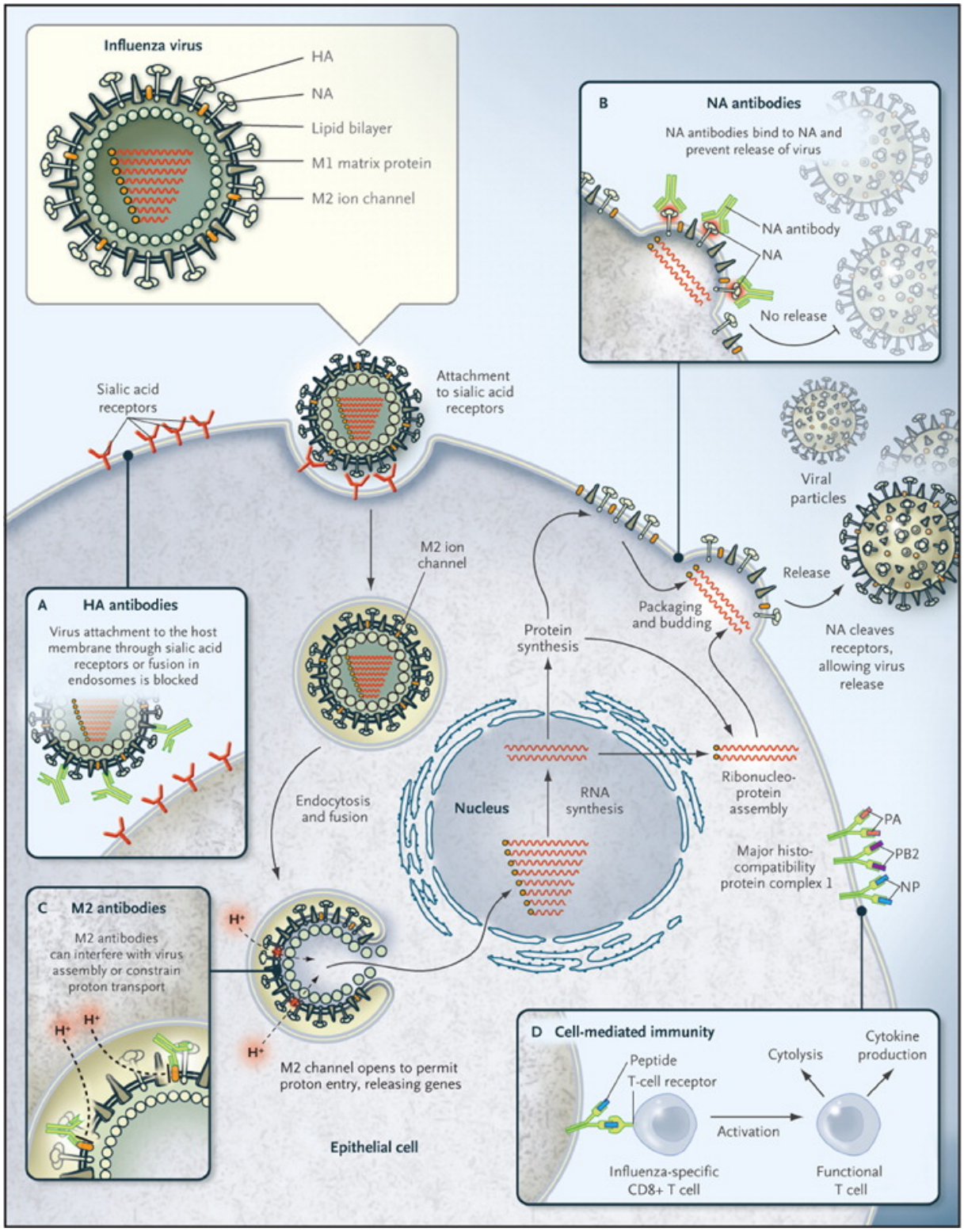

Influenza virus
HA
NA
Lipid bilayer
M1 matrix protein
M2 ion channel
B
NA antibodies
NA antibodies bind to NA and prevent release of virus
NA antibody
NA
No release
Sialic acid receptors
Attachment to sialic acid receptors
Viral particles
M2 ion channel
Release
Packaging and budding
Protein synthesis
NA cleaves receptors, allowing virus release
A
HA antibodies
Virus attachment to the host membrane through sialic acid receptors or fusion in endosomes is blocked
Endocytosis and fusion
Nucleus
RNA synthesis
Ribonucleo-protein assembly
PA
PB2
NP
Major histo-compatibility protein complex 1
C
M2 antibodies
M2 antibodies can interfere with virus assembly or constrain proton transport
H+
M2 channel opens to permit proton entry, releasing genes
Epithelial cell
D
Cell-mediated immunity
Peptide
T-cell receptor
Activation
Cytolysis
Cytokine production
Influenza-specific CD8+ T cell
Functional T cell

Figure 1.1: (facing page) Structure and replication cycle of influenza A virus and well-characterised adaptive immune responses. Eight gene segments code for 11 proteins, including hemagglutinin (HA) and neuraminidase (NA), which account for most of the known antigenic determinants. The portion of the matrix 2 (M2) protein outside the viral envelope is also antigenic. Adaptive immune responses are shown in Panels A through D. In Panel A, the influenza HA protein mediates attachment of the virus to its host cell receptor. Antibodies directed against the HA protein block attachment of the virus to the host cell receptor or block fusion of the virus and host membrane. Antibodies generated against HA are correlated with vaccine protection against infection. In Panel B, antibodies generated against the NA protein do not prevent infection but limit the release of virus from infected cells. Antibodies directed against NA have been correlated with a reduction in disease severity. In Panel C, antibodies generated against the highly conserved external domain of the M2 protein epitope interfere with virus assembly or constrain proton transport and are highly cross-reactive across virus subtypes. In Panel D, CD8+ T-cell responses to conserved influenza virus components have been correlated with enhanced clearance of virally infected cells; however, the exact degree to which they contribute to a reduction in illness remains uncertain. NP denotes nucleoprotein, PA polymerase acidic protein, and PB2 polymerase basic protein 2. Figure and caption reproduced with permission from Lambert and Fauci [2010], Copyright Massachusetts Medical Society (caption modified).

1.3.1 *Influenza A*

Influenza A viruses are further classified into subtypes (serotypes) based on antigenic characteristics of the two viral surface proteins haemagglutinin (HA) and neuraminidase (NA). There are currently 18 known HA subtypes (H1 to H18) and 11 known NA subtypes (N1 to N11). The 18 HA subtypes are then further classified phylogenetically into group 1 (containing H1-2, H5-6, H8-9, H11-13 and H16-18) and group 2 (containing H3-4, H7, H10, and H14-15). The primary reservoir for all these subtypes is aquatic birds, but poultry and other animals such as pigs, horses, dogs, seals, minks and bats[1] can also become infected [Parrish et al., 2015].

Human influenza pandemics

Significantly, influenza A has been the cause of all recorded influenza pandemics, occurring when a virus from a novel subtype not currently in the human population (or with antigenic characteristics grossly dissimilar to viruses currently circulating) accumulates the mutations necessary to transmit effectively in humans [Wilks et al., 2012]. In these cases, the population is particularly vulnerable to severe infection by viruses against which it is typically considered to be immunologically naive. Only viruses from three subtypes are known to have achieved this, namely H1N1 causing the devastating 1918 Spanish flu, followed by the less severe 1957 Asian flu (H2N2), 1968 Hong Kong flu (H3N2) and 2009 Mexican flu (H1N1 again) [Kilbourne, 2006]. Each successive pandemic displaced strains from the previous subtype but the picture was complicated in 1977 when H1N1 re-emerged in the human population after 20 years in what proved to be a relatively benign epidemic. Viruses of this type then continued

[1] Although it is not clear if bats can become infected from birds and it is now clear that bats themselves are an ancient reservoir of highly diverse influenza viruses (see for example Tong et al. [2013]).

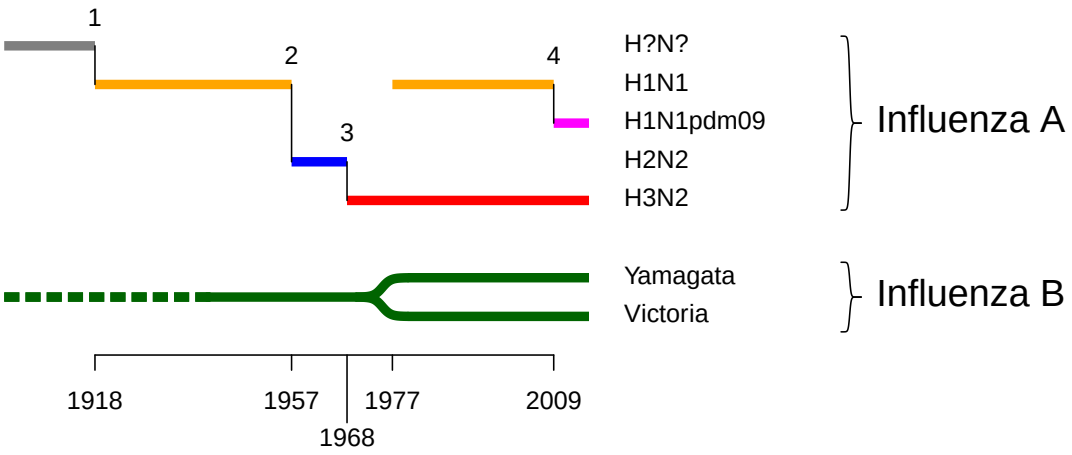


Figure 1.2: Timeline of influenza circulation in the human population. Numbers mark influenza A pandemics and the emergence of new subtypes in the human population as follows: 1 - 1918 Spanish flu (H1N1), 2 - 1957 Asian flu (H2N2), 3 - 1968 Hong Kong (H3N2), 4 - 2009 Mexican flu (H1N1pdm09). The bifurcation of influenza B in the 1970s represents the divergence of influenza B viruses into the currently circulating Yamagata and Victoria lineages. The break in the timeline for H1N1 represents the initial displacement of this subtype by the appearance of H2N2 in 1957, followed by its reappearance in the human population in 1977 before finally being displaced again by pandemic swine flu (H1N1pdm09) in 2009.

to circulate as a seasonal virus alongside H3N2 until 2009, when the pandemic swine flu (commonly designated H1N1pdm09) emerged, taking its place (see figure 1.2).

Why do influenza subtypes die out?

The reason why previous influenza subtypes have often historically been displaced upon emergence of new ones remains a source of some speculation. Sometimes it is argued that triggering of broadly cross-reactive innate immune responses during pandemic spread of a new subtype puts an immunological pressure upon older variants that, when added to the already substantial adaptive immunity that will have built up in the population, causes them to die out. Problematically however, this fails to account for the exceptions in which this did not occur, and why influenza B has been able to co-circulate undisturbed in the human population. With this in mind, it has been hypothesised that the presence or absence of shared epitopes in either the haemagglutinin or neuraminidase components determine whether previous subtypes will continue to circulate or be displaced [Palese and Wang, 2011]. Under this formulation, H2N2 displaced H1N1 in 1957 due to shared stalk epitopes between the two group 1 HAs, H3N2 displaced H2N2 due to shared NA components, and H1N1pdm09, although antigenically distinct, displaced the previous seasonal H1N1 strains due to shared epitopes in both HA and NA. Conversely, H3N2 continued to circulate after both the re-emergence of H1N1 in 1977 and the

H1N1pdm09 pandemic because neither HA group epitopes nor neuraminidase epitopes were in common. As we shall revisit later, aside from neatly accounting for the patterns of displacement and co-circulation seen historically, such an explanation also ties in well with increasing evidence of physiologically relevant levels of protection even between influenza viruses classically considered to be antigenically too highly distinct to generate much cross-reactivity. It remains to be seen, however, whether future subtypes will continue to follow such a pattern.

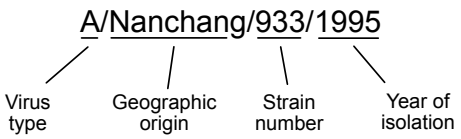
1.3.2 *Influenza B*

Unlike Influenza A, influenza B viruses almost exclusively infect humans and have not historically caused the pandemics associated with the zoonotic infections of influenza A [Hay et al., 2001]. Influenza B viruses are not divided into subtypes but rather broken down into two genetically and antigenically distinct lineages based on HA and named after their first representatives - B/Yamagata/16/88 (the B/Yamagata lineage) and B/Victoria/2/87 (the B/Victoria lineage) [Rota et al., 1990]. These lineages are thought to have diverged at some point in the 1970s when the B/Victoria lineage emerged in China before making a global appearance in the mid-1980s [Chen et al., 2007]. The earliest influenza B virus isolate originates from the 1940s but they are estimated to have circulated in the human population since well before this time. The infectious burden of influenza B is typically estimated to be less than that of influenza A viruses but they are still an important cause of morbidity and mortality, particularly in children and young adults [Sandt et al., 2015].

1.4 The antigenic evolution of influenza

Soon after isolation of the influenza virus in 1933 [Smith et al., 1933] the changeability of antigenic characteristics of the virus became clear [Magill and Francis Jr, 1936]. As is typical of RNA viruses, the low-fidelity of influenza virus polymerases result in high mutation rates, creating genetic and antigenic diversity amongst viral strains [Parvin et al., 1986]. Selection pressure from developing immunity in the population then drives selection of novel variants but there appear to be structural constraints that limit the speed and extent to which influenza viruses evolve antigenically (see Wu and Wilson [2017] for a recent review). We now understand that the majority of mutations are antigenically neutral however, and changes in certain key sites appear to have disproportionately large antigenic effects [Koel et al., 2013]. As might be expected, the result is that although the degree of genetic homology between different viral strains tends to be broadly linked with antigenic similarity, it is not a perfect correlation [Smith et al., 2004].

A note on nomenclature. Upon isolation, influenza viruses are assigned a unique identifier, based on virus type, geographic origin, strain number, year of isolation and (optionally) virus subtype, each separated by a forward slash. An example is as follows:



Note that this differs slightly from the official WHO nomenclature which uses a 2-digit year format for strains isolated before 2000. There are some other additions for viruses from the same seed strain but with different passaging histories or viruses produced using reverse genetics, as described in the strain list in appendix A.

The focus on haemagglutinin as an antigenic determinant

Antigenically speaking, the most important components of the influenza virus are haemagglutinin and neuraminidase, being the two dominant surface proteins in the viral envelope and therefore directly accessible to serum and mucosal antibodies (figure 1.1). Of these two, haemagglutinin forms the main component of influenza vaccines (see section 1.5) and has been the main focus of research into the antigenic evolution of the influenza virus for a number of reasons. First, it is both the more dominant surface protein of the two, with haemagglutinin antigenic sites estimated to be 2 to 4 times more prevalent than those of neuraminidase [Webster et al., 1968]. Second, its role in initial viral binding and membrane fusion means that antibodies raised against haemagglutinin have the capability to block viral binding entirely and prevent initial infection [Hobson et al., 1972], whereas anti-neuraminidase antibodies are typically only effective in reducing viral reproduction rates [Kilbourne et al., 1968]. Third, passive transfer of anti-haemagglutinin antibodies has been shown to be highly protective compared to antibodies raised against any of the other viral components [Virelizier, 1975]. Fourth, mutation rates in HA are higher than those found in NA, the inference being that this is a result of greater effective immune pressure in the human population [Kilbourne et al., 1990; Nobusawa and Sato, 2006].

1.4.1 *Measuring antigenic change*

The use of animal antisera to infer antigenic relatedness

Tracking the antigenic evolution of influenza viruses has been the focus of some considerable effort over the past decades and the mainstay of antigenic assessment remains the measurement of cross-reactivity of animal antisera raised against a given influenza strain [Maher and DeStefano, 2004]. Simplistically put, if an antiserum raised against a given strain A shows a high degree of cross-reactivity against strain B but little cross-reactivity against strain C, the inference is made that A and B are antigenically similar, while A and C are not. By measuring the reactivity of multiple antisera against viruses in this

way a picture can be built up of antigenic relatedness between strains - this is obviously highly relevant when trying to determine whether new influenza virus variants are likely to represent antigenic escape mutants that will successfully evade current population immunity.

The HI assay

Since it's first use in the 1940s, the HI assay has been the favoured approach for measuring reactivity of serum antibodies against given viruses [Hirst, 1943]. It is a binding assay that infers antibody-reactivity based on the ability of a given serum sample to block agglutination of red blood cells by a test-virus strain. The assay itself is discussed in more detail in appendix B, but in brief, a virus-red-blood-cell mixture is incubated with progressive dilutions of sera and the point measured at which the serum ceases to inhibit agglutination. Its advantages are its relative simplicity and repeatability and the fact that by measuring binding inhibition, only antibodies capable of blocking viral fusion will be measured. As already discussed, such anti-HA antibodies were shown to be the most protective in passive transfer studies and indeed HI titres have been consistently shown to correlate well with protection against influenza virus infection [Coudeville et al., 2010a; Fox et al., 2015; Hobson et al., 1972; Jong et al., 2003; Wikramaratna and Rambaut, 2015].

Neutralisation assays & other approaches

Although currently the most common approach, in addition to the HI assay several neutralisation assays exist that aim to account for the effect of antibodies directed against other viral components that may inhibit viral replication but not binding [Rowe et al., 1999a]. Compared to the HI assay however, these assays are considerably more labour intensive since cells must be incubated with different concentrations of virus-serum mixture and the number of viral foci compared. Since such neutralisation assays theoretically account for a wider range of protective antibodies, it may be expected that titres measured in this way would be significantly higher for a given serum sample and indeed, this has been found to be the case [Holzmann et al., 1996; Rowe et al., 1999b]. In practise, the relationship depends greatly upon the exact protocol being followed and the nature of the viruses and sera being tested. Generally speaking however, titres appear to correlate broadly well with those from the HI assay when compared [Jong et al., 2003]. Of course all these serological assays measure only immunity related to inhibitory antibodies and do not account for the cellular components, i.e. the effect of antibodies that may mediate effects through viral opsonisation, or the T-cell response (see section 1.6).

1.4.2 *The advent of antigenic cartography*

Having established an assay for measuring antigenic change, for a long time the approach evolved little further, ferret antisera would be raised against a selection of viruses and a select panel of these antisera then titrated against

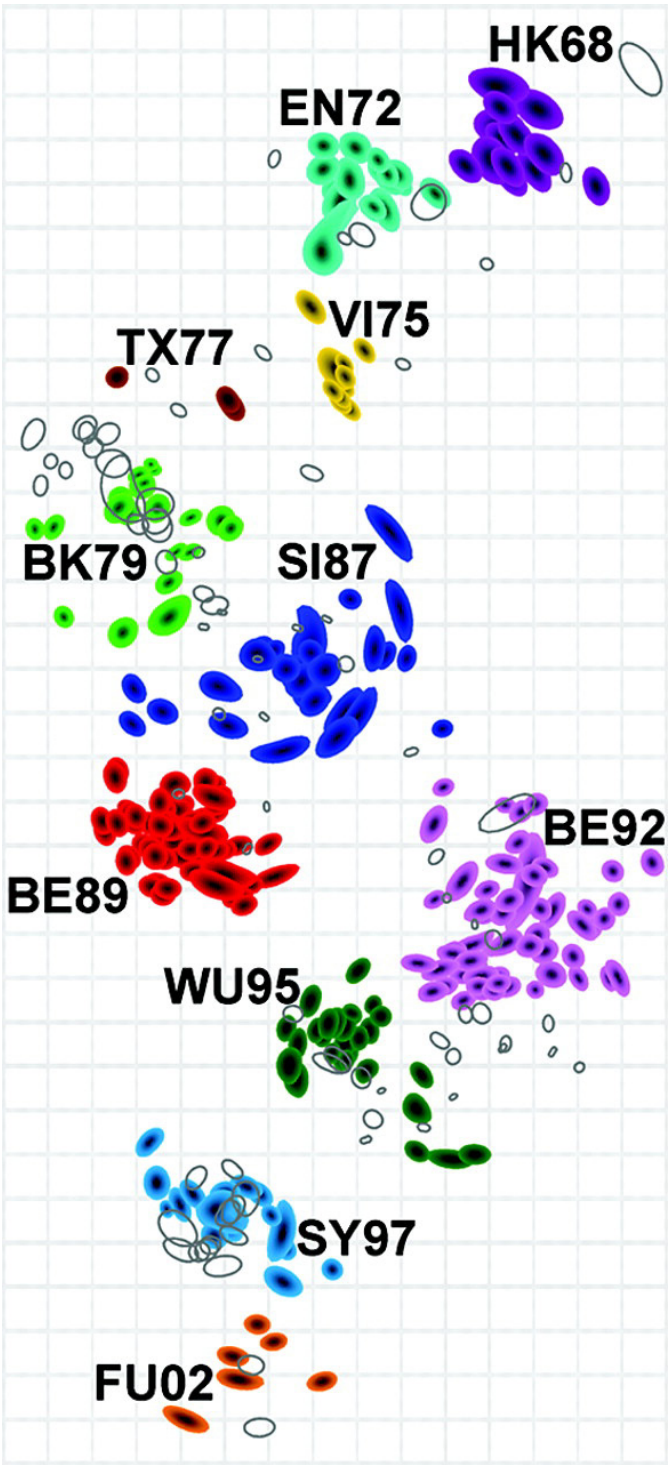


Figure 1.3: Antigenic map of H3N2. As published in Smith & Lapedes *et al.* in 2004 [Smith et al., 2004]. Coloured circles represent the antigenic position of each influenza strain and empty circles represent positions of ferret antisera raised against different viral strains. Both axes represent "antigenic distance"[a]. From Smith et al., 2004. Reprinted with permission from AAAS.

[a] A frequent question seems to be why the H3N2 antigenic map has two axes, both of which represent antigenic distance, rather than simply one. As eluded to in the text it is a question of dimensionality and how constrained the positions of points are in relation to each other. Just as two dimensions, both of which represent distance are required to represent the relative geographic positions of points on a map, so two dimensions, both of which represent *antigenic* distance are required to represent the antigenic position of strains.

Antigenic cluster	Abbreviation	Years of circulation
Hong Kong 1968	(HK68)	1968-1971
England 1972	(EN72)	1972-1974
Victoria 1975	(VI75)	1975-1976
Texas 1977	(TX77)	1975, 1977
Bangkok 1979	(BK79)	1979-1986, 1988
Sichuan 1987	(SI87)	1986-1991
Beijing 1989	(BE89)	1989-1992
Beijing 1992	(BE92)	1991-1996
Wuhan 1995	(WU95)	1993-1997
Sydney 1997	(SY97)	1997-2004
Fujian 2002	(FU02)	2002-2005
California 2004	(CA04)	2005-2006
Wisconsin 2005	(WI05)	2006-2009
Perth 2009	(PE09)	2009-2012

Table 2: The years of circulation for the different influenza H3N2 antigenic clusters. The 14 antigenic clusters from the first circulation of H3N2 in humans in 1968 until 2012 are shown.

newly emerging influenza strains. These results would be arranged in large "HI tables" and people became expert in reading the patterns of titres indicative of the emergence of new antigenic variants (in its simplest form, a reduction in titres measured against antisera raised against current strains).

The methodology of antigenic cartography

In 2004, building on the work of Lapedes and Farber [Lapedes and Farber, 2001], Smith et al. showed that HI titres could be combined computationally to quantify antigenic distances between H3N2 strains [Smith et al., 2004]. HI titres are used to infer antigenic distances between strains and antisera, creating a distance matrix that is then reduced in dimensionality through a process of multidimensional scaling. Surprisingly, they found that antigenic distances between H3N2 influenza strains could be accurately described using as few as two dimensions, allowing for visualisation of the antigenic evolution of H3N2 in terms of an antigenic map. Figure 1.3 reproduces the original H3N2 antigenic map published in their paper, just like a geographical map, distance is represented on two axes and the antigenic distance between strains is simply represented as the straight line distance between them, antigenically similar strains appearing close to each other and more distinct strains appearing being further apart.

Antigenic evolution in H3N2

Visualised as an antigenic map, the progressive and step-wise antigenic evolution of influenza viruses becomes clear, as different strains evolve to occupy

new regions of antigenic space, forming new antigenic clusters over time. Considered up until 2012, there has been circulation of 14 antigenically distinct H3N2 clusters, as shown in Table 2.

Antigenic evolution in other subtypes

Of all the influenza subtypes, the antigenic evolution of H3N2 remains the best defined. Few isolates of the H1N1 virus introduced in 1918 exist but strains from the reintroduction in 1977 also showed a similar antigenic progression before its displacement in 2009 by the new H1N1 strains of the Mexican flu [Nelson et al., 2008; Rambaut et al., 2008]. Antigenic evolution in influenza B also occurs, but it does not appear to be so dramatic as H3N2 [Air et al., 1990]. In line with this estimated amino acid fixation rates for influenza B are lower than in other subtypes [Carrat and Flahault, 2007; Rota et al., 1992; Yamashita et al., 1988]. The 2009 H1N1 pandemic strains appear to have changed relatively little antigenically speaking since their introduction [Arriola et al., 2014; Linderman et al., 2014], but again not much has been published in this regard. In addition to the better understanding of H3N2 antigenic evolution, the relatively long history of continual circulation in humans of the H3N2 subtype also makes it a particularly suitable subtype to study and work covered in this thesis focusses largely on the H3N2 subtype for these reasons.

1.5 Influenza vaccination

Vaccination represents one of the great triumphs of public health intervention, yet to date our remarkable successes have come largely against acute infections caused by antigenically invariant pathogens [Plotkin, 2005]. Influenza stands in stark contrast to this and it has required a monumental global effort to achieve protective vaccination against such a moving target - a process that still contains much room for improvement. Indeed, keeping up with the changing antigenic characteristics of the influenza virus has been the big challenge of influenza vaccination since it was first realised in 1940s that vaccine strains would have to be periodically updated to match newly circulating antigenic variants and maintain effectiveness [Hannoun, 2014]. Since 1973, the World Health Organisation (WHO) has published vaccine recommendations based on global influenza virus surveillance data. Initially this was on an annual basis but since 1998 this was extended to two vaccine recommendations relating to vaccines for use during the following winters in the northern and southern hemispheres.

1.5.1 *The vaccine strain selection process*

The Global Influenza Surveillance Network

To monitor global circulation and the antigenic evolution of the influenza virus, the WHO established the Global Influenza Surveillance Network in 1952, later

renamed as the Global Influenza Surveillance and Response System. The backbone of this system is more than 110 National Influenza Centres that collect virus specimens in their respective countries, perform preliminary analysis and ship representative clinical specimens and isolated viruses to WHO collaborating centres for further analysis. Antigenic analysis of isolated strains, alongside sequence and epidemiological data is then used to perform risk assessments and forms the basis for vaccine strain choices. The aim of this process is to ensure that vaccine viruses have antigenic properties that most closely match those of circulating strains in the upcoming season [Stöhr et al., 2012].

The vaccine manufacture process

Once vaccine strain recommendations are made the process of vaccine manufacture begins, a process that typically takes a minimum of 6 months (see figure 1.4). Currently, the vast majority of influenza vaccines produced are egg-based leading to further potential problems and delays due to egg-adaptations. In the worst cases, antigenic changes in the seed strain that occur during the egg-passaging process can significantly worsen the match against infecting viruses if not addressed by use of an alternative [Skowronski et al., 2014]. To allow time for such eventualities and for administration of vaccine doses before the onset of the influenza season, strain selections for the Northern and Southern hemisphere are currently made in February and September, a total of around 9 months in advance depending upon exactly when influenza virus circulation peaks in a given year.

1.5.2 *Vaccine effectiveness and antigenic mismatch*

The coordination of influenza virus surveillance and vaccination represents one of the great success stories of a global collaborative scientific and public health effort but it is by no means perfect. The effectiveness of the influenza vaccine is notoriously difficult to estimate and is a complex, difficult and controversial (although hugely important) research area in itself. Meta-analyses have put average vaccination effectiveness at around 59% although this is highly variable depending upon the year and method of estimation [Osterholm et al., 2012].

Antigenic mismatch and the 1997-1998 influenza season

The reasons for poor influenza vaccine effectiveness are multiple and not well understood, but a factor that has been consistently identified as a cause of particularly poor vaccine effectiveness in certain years is antigenic mismatch of the virus and circulating strains [Belongia et al., 2009]. Given the length of time between vaccine-strain selection and vaccine administration, new antigenic variants can begin to dominate in the global population, with the result that vaccine strains are no longer similar to circulating strains. In the 1997-1998 season, antigenically-drifted A/Sydney/5/1997-like viruses caused severe disease outbreaks in Europe and the US and the vaccine antigen drawn from the previous Wuhan 1995 antigenic cluster (A/Nanchang/933/1995 or A/Wuhan/

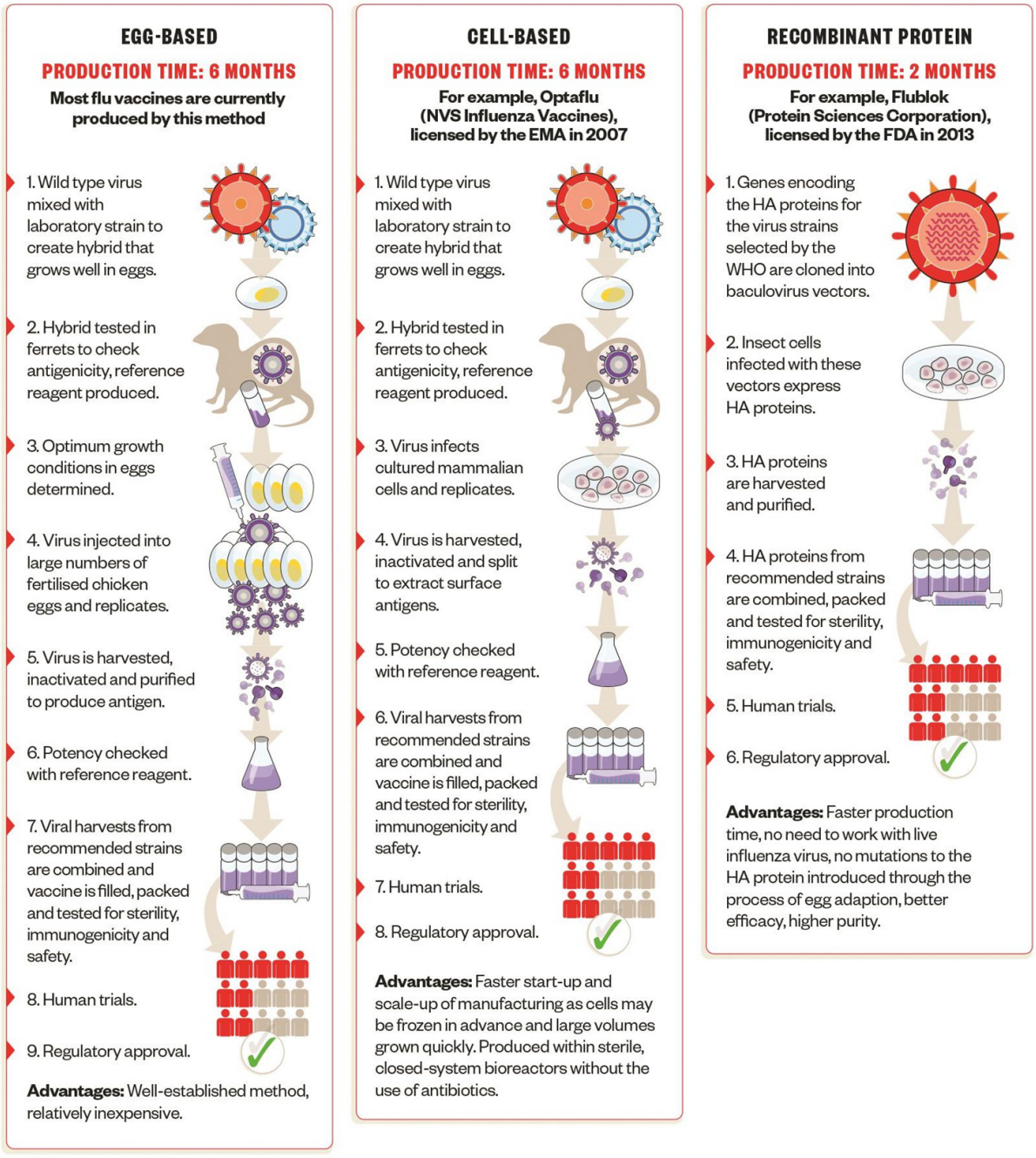


Figure 1.4: Annual production cycle for influenza vaccines. The amount of time required for sufficient vaccine manufacture means that choices for which strains to use in the vaccine must occur at least 6 months in advance of the influenza season, often longer. Modified after Connelly [2015].

359/1995) was poorly matched [Aymard et al., 1999; Klimov et al., 1999; Smith et al., 2004]. As would be expected, several studies have noted the detrimental impact that this had on vaccine effectiveness in the 1997-1998 season, with some even finding that recipients of the vaccine in that season fared no better than placebo [Bridges et al., 2000; Centers for Disease Control and Prevention (CDC), 1998; Jong et al., 2000; Nordin et al., 2001].

Although it seems to be biologically clear from the accumulated evidence that antigenic mismatch plays an important role in vaccination effectiveness, accurately estimating the effect of antigenic mismatch through vaccine effectiveness studies is a real challenge - it is not uncommon for confidence intervals of vaccine effectiveness estimates to range from below 0 to above 70% [Ohmit et al., 2009; Tricco et al., 2013].

1.6 The serological immune response

Having both extracellular and intracellular parts to its lifecycle, immunity to influenza involves multiple components of the immune system. As with all pathogens, passive and innate components of the immune system play a crucial role in protection, while the adaptive immune system allows targeting of specific antigenic features and (through the development of immunological memory) the capability to respond rapidly to re-exposure to the pathogens with shared antigenic characteristics (see again figure 1.1).

The studies in this thesis focus primarily on measuring serological immunity to influenza, i.e. the quantity and specificity of secreted antibody. High levels of serum antibody specific to an infecting virus have been consistently shown to be strongly associated with immunological protection [Coudeville et al., 2010a; Fox et al., 2015; Hobson et al., 1972; Jong et al., 2003; Virelizier, 1975; Wikramaratna and Rambaut, 2015], while they also form the basis for regulatory assessment of vaccine efficacy [Gerdil, 2003; Wood, 2003]. In this section I aim to give a brief overview of what is understood about the immunological processes behind the production and maintenance of antibody.

1.6.1 *The naive B-cell response*

In the B cell system, mature naive B cells residing in secondary lymphoid organs capture antigen specific to their B-cell receptor and begin the process of activation. For influenza, this activation is a T-cell dependent process [Mitchcll et al., 1985] - captured antigen is internalised and degraded before the resulting peptide fragments are loaded onto MHC class II molecules for presentation to CD4+ T-helper cells that have themselves been primed through T-cell-receptor-

binding-dependent interactions with antigen-presenting dendritic cells[2] [Paul, 2013].

The different progeny of activated B cells

Once formed, the antigen-dependent B cell-T cell interaction then starts a sequence of rapid B-cell proliferation and differentiation along one of two main pathways (figure 1.5). The first pathway mediates a fast initial response to infection whereby activated B-cells with unmutated BCR genes divide to produce short-lived plasma cells. In the second pathway, other activated B-cell progeny go on to form germinal centres where they interact with T follicular helper cells and antigen trapped on follicular dendritic cells, undergoing somatic hypermutation and leading to the generation of long-lived plasma cells [Allen et al., 2007]. These long-lived plasma cells, taking more time to produce as part of the germinal centre reaction, go on to constitutively produce high affinity class-switched IgG antibody, maintaining long-term levels of serum and mucosal antibody [Nutt et al., 2015].

The different types of memory B cell

The other crucial output of this process is the production of memory B cells - long-lived cells that do not produce antibody but maintain the capacity to respond to further infection, allowing for rapid recall of antibody production upon re-exposure. It was initially thought that production of memory B cells was an output solely of the germinal centre reaction, but it has more recently become clear that memory B-cells can also be produced independently of the germinal centre and that several types exist with potentially different functional roles. Generally speaking, the germinal-centre-dependent pathway is thought to produce class-switched IgG memory B cells, whereas the germinal-centre-independent pathway produces memory B cells with unmutated IgM receptors. Since memory B cells produced independently of the germinal centre have not undergone the process of affinity maturation, affinity of the BCR is generally lower but a greater diversity of binding-specificity is potentially maintained. Particularly relevant for antigenically variable pathogens such as influenza, it has been postulated that these cells play an important role in providing protection against novel antigenic variants that share some similarity with a previously encountered strain [Kurosaki et al., 2015].

1.6.2 *The B-cell response to re-exposure*

High levels of baseline serum antibody produced by long-lived plasma cells provide an important first line of defence against pathogen re-exposure (indeed

[2] How exactly the B cells themselves encounter antigen when residing in secondary lymphoid organs is still a matter of some study. Antigen-presenting cells (particularly macrophages in the subcapsular sinus of the lymph node) play an important role, while it has been shown that smaller soluble antigens can gain access to the B-cell follicles and bind B-cell receptors directly (see Yuseff et al. [2013] and Pape et al. [2007]).

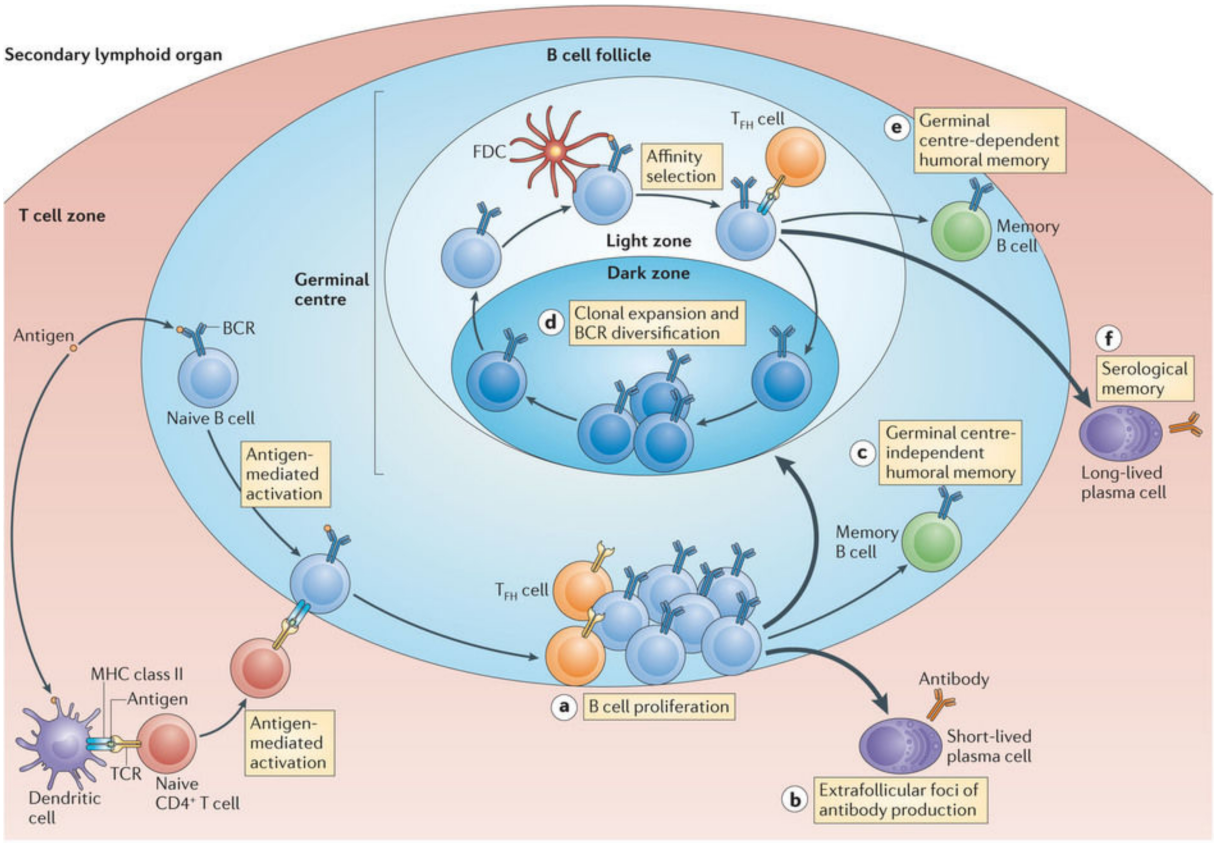


Figure 1.5: Schematic of the process of B-cell differentiation. Antigen-activated B cells and T cells migrate towards the borders of the B cell follicles and the T cell zones of secondary lymphoid organs, respectively, which leads to them establishing stable B cell-T cell interactions and enables B cells to receive helper signals from cognate CD4+ T cells. Activated B cells and T cells then migrate to the outer follicles, where B cells undergo proliferation (part a). Some of the proliferating B cells differentiate into short-lived plasma cells (part b), which give rise to the extrafollicular foci, and some develop into memory B cells (part c; germinal centre-independent memory B cells). Alternatively, the activated B cells can return to the follicle and can undergo rapid proliferation to form the germinal centre (part d). In the dark zone of the germinal centre, the clonal expansion of antigen-specific B cells is accompanied by B cell receptor (BCR) diversification through somatic hypermutation. The B cells that exit the cell cycle relocate to the light zone, where affinity selection takes place through interaction with immune complex-coated follicular dendritic cells (FDCs) and antigen-specific T follicular helper cells (TFH cells). The affinity-matured germinal centre B cells can re-enter the germinal centre cycle. Alternatively, these germinal centre B cells exit the germinal centre, either as memory B cells (part e; germinal centre-dependent memory B cells) or as long-lived plasma cells (part f) that contribute to serological memory. The strength of signals that B cells receive is likely to determine their fate; stronger signals (indicated by bold arrows) favour development into plasma cells or germinal centre B cells, whereas weaker signals (indicated by narrow arrows) determine memory B cell differentiation. TCR, T cell receptor. (Figure and caption reproduced from Kurosaki et al. [2015])

passive transfer of serum antibody alone has been shown to be sufficient for effective protection against influenza virus infection [Virelizier, 1975][3]). However, when this secreted antibody fails to contain an infection, memory B cells can become stimulated and re-activated by antigen. The primary outcome of this re-stimulation of memory B cells upon a further antigen exposure is proliferation and differentiation into short-lived antigen-secreting plasma cells, generating a rapid transient increase in the concentration of circulating antibody, alongside the production of further memory cells and long-lived plasma cells [Weill et al., 2013]. Again, this process appears to be a largely T-cell dependent one, with helper T-cells recapitulating their role in helping to boost memory B cell expansion and differentiation.

Evidence has also now accumulated that memory B cells are capable of re-entering the germinal centre reaction and undergoing further affinity maturation in the presence of new antigen although the relative importance of this process in the case of secondary exposures remains controversial [Weill et al., 2013]. Again the distinction between memory B cell types appears to be relevant here, with IgM memory cells rather than IgG memory cells shown to reinitiate a germinal centre reaction upon re-encounter of antigen.

In contrast to memory cells, plasma cells are terminally differentiated and are not thought to maintain the capacity for further division. The traditional view is that they act simply as factories for the consistent production of high baseline levels of antibody and play no further role in modulation of the immune response, having lost the ability to sense or respond to antigen through down-regulation of their BCR [Radbruch et al., 2006].

1.6.3 *The maintenance of immunity*

Long-term memory is maintained through the persistence of long-lived plasma cells and memory B and T cells. Indeed, antibodies against many vaccine antigens are still detectable in the sera of individuals many decades after their last exposure [Amanna et al., 2007; Slifka and Ahmed, 1996]. Regarding the influenza virus, antibodies against strains from 1918 H1N1 pandemic were still measurable in samples taken from subjects nearly 90 years later [Yu et al., 2008].

[3]Interestingly, in related experiments Virelizier et al. found that although memory cells should in theory be able to respond quickly to infection, passive transfer of memory B cells alone in mice failed to provide protection, only noting a benefit when these memory cells were first stimulated to differentiate to produce plasma cells and antibody. This presumably explains why the body goes to such lengths to maintain background levels of antibody even in the absence of infection.

Although incapable of further replication, long-lived plasma cells have been shown to have the capability of surviving almost indefinitely in survival niches present inside the bone marrow [Manz and Radbruch, 2002; Manz et al., 1997]. Survival factors present in these locations seem to be key to these cells' longevity however and removal from their survival niche leads to rapid cell death [Radbruch et al., 2006]. This, alongside the observation that total levels of serum IgG tend to remain constant and not increase over time, has led to the hypothesis of competition for places in a limited number of survival niches, with newly produced plasma cells potentially capable of displacing older ones [Odendahl et al., 2005].

In contrast to plasma cells, memory cells (being not yet terminally differentiated) maintain the capacity to divide and self-replenish. Studies have shown that in this way memory cells can also survive for many decades, maintaining the ability of the immune system to respond dynamically to threats upon later re-encounters [Ahmed and Gray, 1996].

1.7 The effect of previous exposure

The potential to be exposed to many antigenically variable but related strains during a lifetime creates particular immunological challenges and phenomena, unique to antigenically variable pathogens.

Generally speaking, immunity against invariant pathogens may be partial in the sense that the immunological memory present is not sufficient to provide long-lasting protection, but in such cases the problem is one of how to boost the duration or potency of an initial immune response. In diseases such as influenza however, the situation can arise that an individual has partial immunity and protection not due to a weak immune response to an initial exposure, but due to previous exposure to an antigenically similar strain. Being directed at a different constellation of epitopes, simple boosting of this immunological memory may not be enough to provide protection against infection, but cross-reactivity means that it still plays a role in the immune response against it. The immune response against the influenza virus therefore depends not only on the type of strain encountered or strength of immune stimulus, but also on the exposure history of the recipient to previous antigenic variants.

1.7.1 *Original antigenic sin*

This role of prior immunity and exposure history in affecting the responses upon infection with a new strain was first recorded as early as 1947 when Francis et al. [1947] compared antibody responses in recently infected individuals against both "Rhodes" (a newly isolated antigenic variant) and PR8 (a previ-

ously circulating virus). They observed that upon infection, while acute and convalescent antibody titres were similar against the new virus Rhodes, titres against PR8 were much higher in patients who had received a vaccine containing PR8 in the previous year - the inference being that in patients with recent PR8 exposure, antibodies with additional cross-reactivity against PR8 viruses had also become involved in the response.

Also noted in the 1947 study by Francis and Salk was the fact that although fold-change increases in HI titre were greatest against the infecting Rhodes strain, final titres and the absolute increase in titre was highest against the previously encountered PR8 strain. A study published 6 years later by Davenport et al. related to these observations, finding that titres in populations of different age groups tended to be highest against strains that circulated soon after they were born [Davenport et al., 1953]. The authors went on to hypothesise that the first infection creates a kind of imprinting effect on the immune system - subsequent exposures progressively reinforce pre-existing antibodies from the first infection that also react against antigens common to the new virus:

> *"From these data the following immunologic thesis is formulated. The antibody which is acquired during the initial infections of childhood is of limited scope and reflects the dominant antigens of the prevailing strains. The immunity conferred by the initial experiences with influenza is also limited. Successive experiences later in life with viruses of related but differing antigenic make-up result in a composite of antibody which is oriented toward a larger number of the common antigens which comprise influenza virus. These experiences confer a broader immunity which limits infection with, and antibody response to, the more recently encountered strains. The antibody-forming mechanisms appear to be oriented by the initial infections of childhood so that exposures later in life to antigenically related strains result in a progressive reinforcement of the primary antibody."*
>
> - Davenport et al. 1953

Over the following years, several studies were performed to investigate this hypothesis and the mechanistic basis behind it. Additional experiments in both humans and ferrets confirmed that upon exposure to a secondary antigenic stimulus, antibody titres against previously encountered strains were strongly boosted, even if the second strain were so antigenically dissimilar a primary exposure to it would normally generate no detectable cross-reactivity [Davenport

and Hennessy, 1956; Davenport and Hennessy, 1957; Davenport et al., 1955; Hennessy et al., 1955].

Crucially it was shown that the smaller level of antibody reactivity that was generated against a second strain was not simply produced at lower levels in addition to it, but was instead intrinsically linked to the initial response generated against the first. For example, when ferrets were serially infected with three antigenic variants, absorption of the resulting sera with the first antigen removed *all* measurable reactivity against the second and third strains [Jensen et al., 1956]. In the reverse situation where the resulting sera were absorbed with the third antigen, all reactivity was removed against the third strain, but there was little effect on titres against the first or second. The surprising conclusion from this was that the majority of the antibodies that had generated reactivity against the second and third strain, also had maintained cross-reactivity with the first. Equivalent, although not always so clear cut, results were found in humans, using instead different age groups to compare the effect of different exposure histories [Francis, 1960].

Several further studies confirmed and extended the results of Jenson et al., showing that secondary responses were characteristic of the production of antibodies reactive to epitopes shared between the new strain and those previously encountered. Fazekas de St Groth and Webster in particular went to great lengths in a series of studies published in 1966 [St Groth and Webster, 1966a; St Groth and Webster, 1966b] using X-ray irradiation and measurements of antibody binding affinity to demonstrate that antibodies produced against a new strain following prior exposure with an antigenically related one were characteristic of a secondary recall of pre-existing immunity and not an additional primary response (as was seen when the new strain was encountered instead as a first infection).[4]

Results from studies such as those described above became amalgamated into the concept of "original antigenic sin" a term that has pervaded the scientific literature ever since. Strictly speaking however, three distinct phenomena were described:

1. That static titres measured in a population tend to be highest against viruses corresponding to those circulating soon after their birth.

[4]They noted however with experiments in rabbits that the size of the first and second inoculum played a role in the types of responses seen. When a small inoculum was used as the first exposure, only the antibody produced early on in the response to a second infection with an antigenic variant was entirely cross-reactive against both strains, replaced within 14 days by a primary-type antibody response to the boosting antigen. When a very large dose of inoculum was used as the second infection, cross-reactive and primary antibody to the second antigen was produced in parallel.

2. Responses against new antigenic variants cause a boosting of antibody reactivity to previously encountered antigenic variants and that this can result in the highest post-exposure titres being maintained against these previously encountered strains rather than the new one.

3. That antibodies produced against new strains appear to be mostly ones that maintain cross-reactivity with previously encountered ones, rather than a new population of antibodies only reactive against the new strain.

These phenomena are potentially related but not necessarily so, and I would argue that the various application of the term "original antigenic sin" to refer to any one of these observations has created confusion in the scientific literature.

Timing post-infection at which samples were taken seems to play a role in the extent to which different phenomena are witnessed in different studies. Indeed this was already noted by Webster, who used sera from the same ferrets as in the original paper from Jenson et al. but examined samples taken at 21 days post infection rather than 14, and already found evidence for a more heterogenous response with both primary and secondary components [Jensen et al., 1956; Webster, 1966].

The cellular mechanism behind boosting of previous titres and skewed production of antibodies that maintain shared cross-reactivity with previous strains has never been conclusively determined but it seems likely to be a consequence of competition between memory and naive B cell populations. This concept was first eluded to with the hypothesis of "antigen trapping" whereby pre-existing immune cell populations were hypothesised through some mechanism to reduce the quantity of antigen available for stimulating a primary response [St Groth and Webster, 1966a].

The antigen trapping hypothesis was formulated before we had generated such a clear picture of the immune system as today. We can now envisage how such a situation may arise mechanistically, whereby a limited amount of antigen enters the lymph node either on macrophages or via simple diffusion and memory cells outcompete naive cells for activation. Memory cell populations have several advantages over naive B cells in such a scenario even if binding less strongly to antigen epitopes:

- They are present in larger numbers. [Rajewsky and Schittek, 1990]
- They appear to have a lower activation threshold. [Liu et al., 1995]
- They proliferate more quickly than naive B cells in response to a stimulus. [Tangye et al., 2003]

1.7.2 *The potential cost of immunological memory*

Is a memory recall response less efficacious?

Whether the apparent domination of memory cell responses comes at a cost to the effectiveness of the overall immune response is still unclear. Normally one would expect a memory response to be superior due to a more rapid and potentially larger production of antibodies with high binding specificity, but when these antibodies were produced in response to a previous antigenic variant and potentially have inferior binding properties than naive B cells that may otherwise have been more robustly stimulated, the theoretical argument becomes muddier.

Looking back at the original studies investigating the effects of different exposure history a common theme is that, while responses may have additionally been directed towards previously encountered strains, presence or absence of prior immunity had relatively little effect on the final post-infection titre to the vaccine/infection strain [Davenport and Hennessy, 1956; Francis et al., 1947; St Groth and Webster, 1966b]. Indeed, given that influenza virus infections seem to have disproportionately high morbidity and mortality in the absence of any prior immunity (for example in children or a novel influenza A subtype entering the human population and causing a pandemic) it would make sense to assume that prior immunity, even if only partially cross-reactive with a new strain, does in fact play a protective role compared to having to generate an entirely novel response.

The potential costs of relying on memory recall

Although prior immunity does not appear to affect the overall strength of response against a new antigen, evidence has mounted that a reduced diversification of the antibody repertoire can create costs in the future whereby a preferential memory response against conserved epitopes in a new strain leads to inferior protection against later antigenic variants when these epitopes mutate. Such a potential situation was noted in practice with regard to the 2009 H1N1 influenza pandemic when the pandemic H1N1 viruses began causing unusually high levels of disease in middle-aged adults during the 2013-2014 influenza season [Dávila et al., 2014]. An explanation ultimately offered was that these individuals, previously protected against the new H1N1 pandemic strains due to an epitope around position 166 of the haemagglutinin molecule that was shared with older seasonal H1N1 viruses circulating around the time of their birth suddenly became more vulnerable when this epitope mutated in the 2013-2014 season [Linderman et al., 2014]. Significantly, since novel responses against previous H1N1 2009 pandemic variants did not appear to be primarily directed against the 166 region, the mutation caused little effect in the susceptibility of other age groups and was not detected as a significant antigenic difference in tests using first infection ferret antisera.

Interestingly, although different subtypes of influenza are classically thought of as serologically distinct (i.e. there is little or no cross-reactivity of antibody responses between them) it has even been suggested that exposure history may influence patterns of immunity and susceptibility between subtypes. Recently, a study in 2016 by Gostic et al. found that the incidence and severity of H7N9 and H5N1 cases in different age groups appeared to be dependent on the year of birth of the individuals rather than their actual ages [Gostic et al., 2016]. The critical inflexion point was around the birth year 1968, when the H3N2 subtype first emerged in the human population, replacing H2N2 - those born before 1968 appeared to have better protection against H5N1 but those born after appeared to have better protection against H7N9. Since both H2 and H5 belong to HA group 1 while H7 and H3 belong to group 2, the inference was that having a first infection with a particular subtype conferred better lifelong protection against later exposure to novel subtypes from the same HA group.

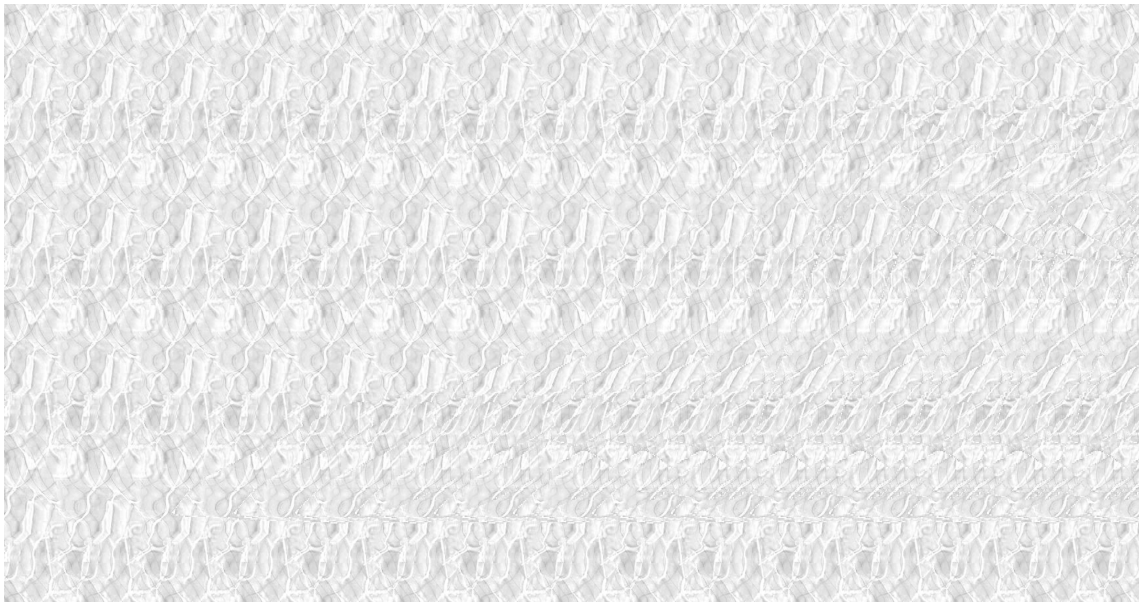
1.8 This thesis in context

Although space here is limited, I hope this introduction gives a useful outline of many of the topics that will be revisited in this thesis. In this chapter I have covered the structure and function of the influenza virus and how antigenic diversity has been increasingly appreciated and more accurately measured. In chapter 2, expanding on the work of antigenic cartography, I will explore with the methodology of antibody landscapes how meaningful understanding of serological immunity can be extended to even the individual level when this variation is accounted for and measurements against a wider range of strains combined. I have also discussed current understanding of the cellular mechanisms behind the serological immune response to influenza virus exposure. In chapter 3, antibody responses following natural influenza virus infections are analysed in detail and the phenomena that result from these processes are examined. I have considered the current approach to influenza vaccination and the successes and short-comings of the vaccination strain selection process. In chapter 4, the immunological consequences of scenarios such as vaccine-strain mismatch are investigated directly and implications for approaches to improve vaccination are discussed. Finally, I have reviewed current understanding and hypotheses relating to the question of the effect of prior immunity in influencing future immunological responses, to which work in many of the chapters relates. The work in chapter 5 however, directly addresses many of the questions regarding how infection history influences later responses compared to naive ones, investigating first and second responses to influenza virus infections and the consequences this has for our understanding of the processes that underly patterns of influenza immunity observed later in life.

Despite over 70 years of research, relevant new findings are still being uncovered about surprisingly fundamental aspects of the immune response and the best way to protect against diseases such as influenza. I hope that the work presented in the next chapters can usefully build further in some of these important areas.

2

The Antibody Landscapes Methodology



A stereogram of an antibody landscape.

2.1 Chapter summary

In this chapter I develop and test antibody landscapes, a method to help analyse complex serological data by visualising antibody-mediated immunity as a function of antigenic differences between a range of different influenza strains.

The test data set is a total of 468 human serum samples titrated against a selection of 70 strains spanning 43 years of influenza H3N2 evolution. Antigenic cartography had already been used to quantify the antigenic relationships between these viruses as a two-dimensional map and I explore the best way to use HI measurements to model serological immunity for each sample as a surface in a third dimension, representing antibody reactivity throughout antigenic space.

To aid visual comparison of multiple landscapes, 2-dimensional summaries of these landscapes are additionally created by taking a profile of the height of the full landscape along a path on the antigenic map that passes through each H3N2 antigenic cluster in chronological order.

Methodological improvements focus on the optimal way to treat non-detectable titres when modelling the surface, the most accurate way to calculate changes in antibody reactivity and how to appropriately estimate confidence intervals for the resulting fits.

2.2 Introduction

2.2.1 *An introduction to antibody landscapes*

"Antibody landscapes" is a methodology first proposed by Derek Smith as a way of better understanding patterns in serological data through analysis of serum reactivity against a wide range of influenza strains in combination, while accounting for antigenic differences between them. In essence, the aim is to create a profile of an individual's antibody-mediated immunity as it varies across antigenic space. In the case of the influenza H3N2 subtype, where antigenic differences can be approximated in two-dimensions, this allows for the immunological profile to be visualised as a surface in a third dimension with peaks representing regions of antigenic space against which an individual shows high immunity and troughs the regions of lower immunity - when antibody reactivity is being measured, an "antibody landscape".

Instead of analysis of antibody reactivity to a single strain in isolation, antigenic cartography, and the extension of an antibody landscapes methodology, provide a framework in which to infer immunity from measurements against a range of viral strains, not simply by taking averages, but by also accounting for variation that may be expected due to antigenic differences. At its best, such an immunological profile theoretically provides not only a simple visualisation of complex serological data, but also a quantification of serological immunity that is more accurate than any single measurement taken alone. Although influenza has to date provided the main study organism, an analogous approach could be applied to any antigenically variable pathogen such as for example dengue virus, malaria or HIV. In a basic form, the method involves the following steps:

1. Select a range of antigenically variant pathogen strains.
2. Create an antigenic map, quantifying their relative antigenic differences.
3. Measure antibody reactivity of a given serum sample against these strains.
4. Model the serum antibody reactivity as a function of the antigenic differences between the strains.
5. Visualise the resulting profile of serological immunity.

In this chapter I will go through this process as applied to HI measurements made using human serum samples against a range of H3N2 influenza strains. Work on the first three steps having been already completed by others, I first explore the most appropriate way to model serological immunity as a function of antigenic differences between strains. Next, I go on to explore the most effective ways to visualise the antibody landscapes created and to share them between researchers. Finally, I extend the methodology to better deal with non-detectable titres, determining how confidence intervals of the fits can be inferred and applying some diagnostic tests to check for biases and goodness of fit.

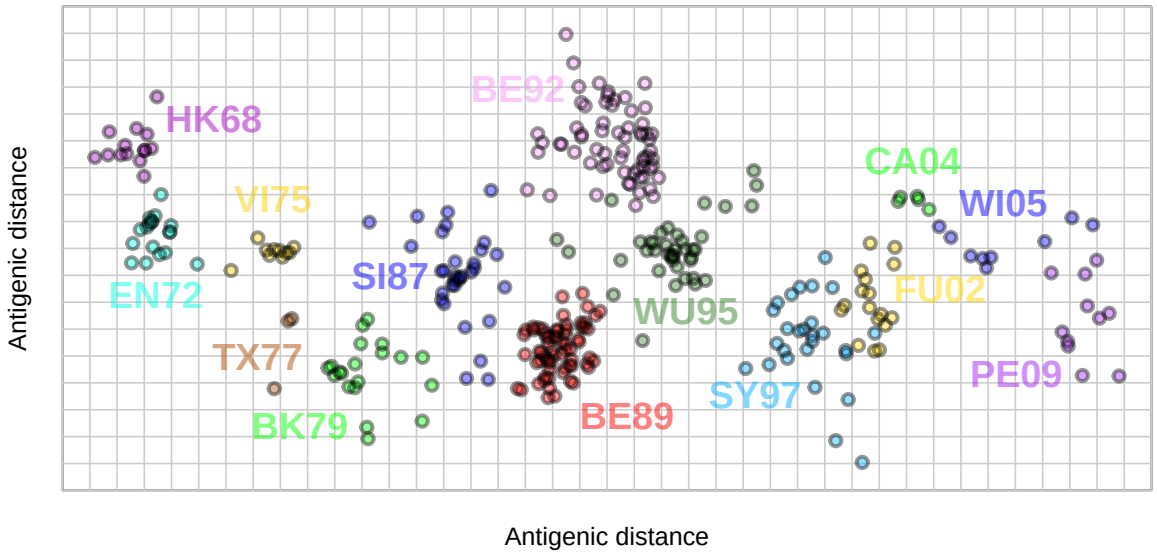


Figure 2.1: Extended H3N2 antigenic map. Antigenic map of the H3N2 evolution extended from the original H3N2 map published by Smith et al. [2004]. Both axes represent antigenic distance, and the spacing between grid lines is 1 antigenic unit, corresponding to a twofold dilution of antiserum in the HI assay. Circles represent the antigenic location of different virus strains, colour-coded by antigenic cluster (cluster name indicated). Sera were omitted for clarity. HK68: Hong Kong 1968; EN72: England 1972; VI75: Victoria 1975; TX77: Texas 1977; BK79: Bangkok 1979; SI87: Sichuan 1987; BE89: Beijing 1989; BE92: Beijing 1992; WU95: Wuhan 1995; SY97: Sydney 1997; FU02: Fujian 2002; CA04: California 2004; WI05: Wisconsin 2005; PE09: Perth 2009.

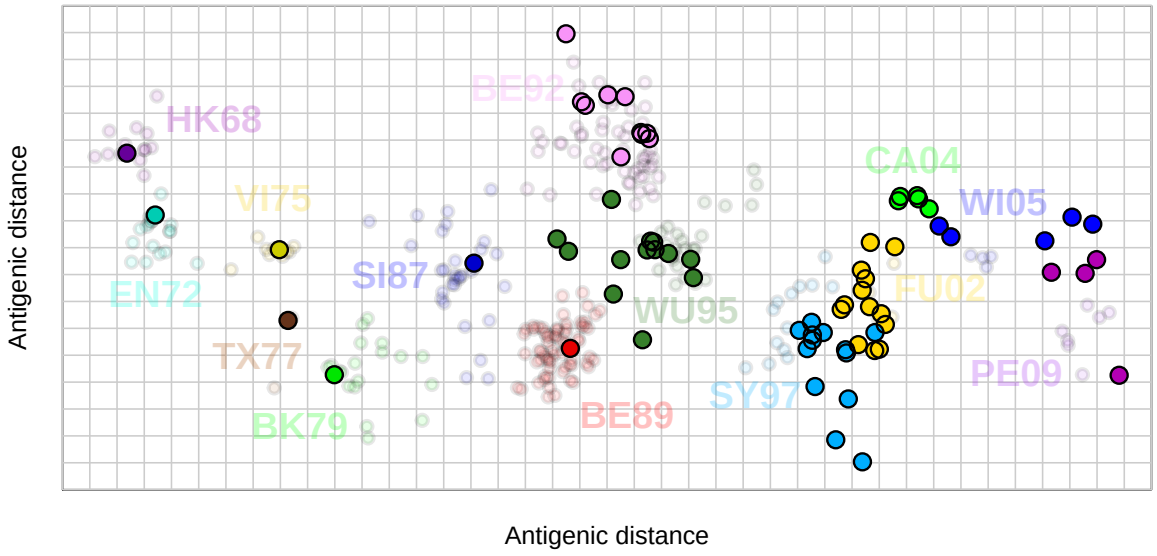


Figure 2.2: H3N2 antigenic map with test strains highlighted. Caption as figure 2.1 but now with the 70 strains against which test sera were titrated highlighted.

2.2.2 *The influenza H3N2 antigenic map*

The influenza H3N2 antigenic map which determines the relative antigenic similarities of the H3N2 strains titrated was constructed by colleagues at the Center for Pathogen Evolution, University of Cambridge, primarily through work by Judy Fonville, Mario Ventresca, Sarah James and Derek Smith. Briefly, the original H3N2 antigenic map presented by Smith et al. in 2004 (which characterised viruses only up until the Fujian 2002 antigenic cluster [Smith et al., 2004]) was extended using an additional set of ferret antisera and HI titrations against 74 viruses, to include an extra 65 H3N2 strains covering the subsequent California 2004, Wisconsin 2005 and Perth 2009 antigenic clusters. The positions of the 74 viruses were found by minimising the mapping error and then an "overlay-merge" was made with the original Smith et al. antigenic map, which minimises the squared error of positioning of viruses in common between the two datasets (9 viruses). The resulting antigenic map is shown in figure 2.1. This extended antigenic map was then used for the development of the antibody landscapes methodology and subsequent construction of all antibody landscapes in this thesis.

2.2.3 *Test data set*

The test dataset used in the analyses performed in this chapter is a series of HI titrations of a range of antigenically distinct viruses measured against pre and post vaccination sera from two vaccination trials performed in 1997 and 1998 (and studied further in chapter 4). Combined, these datasets make a total of 234 pre-vaccination and 234 post-vaccination serum samples with the majority titrated against a full selection of 70 influenza H3N2 strains. These strains are highlighted on the antigenic base-map in figure 2.2 and span from the Hong Kong 1968 antigenic cluster up until Perth 2009, but with higher sampling density against more recent viruses. The serum donors span a range of ages from 18 to 75 years old.

2.2.4 *Collection of serum and measurement of antibody reactivity*

Each serum sample had been titrated against a selection of the H3N2 influenza viral strains as detailed in appendix A using the HI assay and protocol described in appendix B.

Serum was measured for inhibition of virus red blood cell agglutination at a series of increasing 2-fold dilutions. Titres are therefore often first converted onto a log scale before quantitive analysis and in this case the relationship used was:

$$z = log_2(\frac{Z}{10}) \tag{2.1}$$

where Z is the raw HI titre. Here z represents the number of 2-fold dilutions at which inhibition of haemagglutination is still present, starting from a maximum concentration of 1:10. Where the maximum concentration of serum (1:10) still failed to inhibit haemagglutination, z was treated as -1 or a value somewhere in the range of $\leq$-0.5 depending on the methodological application. This log transformation is typical of the analysis of HI data since both titre variation and the HI measurement error is distributed on the log scale. Note however that in some cases raw titres are divided by 5 rather than 10, making the baseline 0 (for a <10 titre when this is counted as a titre of 5) [Beyer et al., 2004].

2.3 Modelling the antibody landscape surface

2.3.1 *HI titre similarity against antigenic distance*

A basic assumption of the antibody landscapes methodology is that antigens close to each other in antigenic space should return similar antibody titres when titrated against the same serum sample due to the fact that antibodies targeted against one antigen should display cross-reactivity against other antigens with shared or similar antigenic characteristics. As a first test of this principle, for each pair combination of titrated antigens, the correlation of titres measured was compared against the antigenic distance between them. The result is shown in figure 2.3 where it can be seen how correlation between antigen titres, calculated as the Spearman's rank correlation coefficient ρ, increases with diminished antigenic distance.

2.3.2 *Locally-weighted regression*

Colleagues first developing the antibody landscapes methodology had previously compared different fitting methods, identifying the R "loess" function as an appropriate method of modelling the antibody landscape surface [Fonville et al., 2013]. This function is an implementation of a form of local linear regression, often used to model underlying structure in datasets about which little or nothing about the nature of the relationship between variables is known [Cleveland and Devlin, 1988]. In addition to the fact that preliminary cross-validation results indicated a local regression approach had the greatest predictive value compared to other standard fitting approaches, the principal of fitting to locally weighted data marries nicely with the theory of serum reactivity changing throughout antigenic space - the main assumption being that titres against a given antigen will be most closely related to titres against other antigenically similar strains.

2.3.3 *Variable and fixed bandwidth local regression*

The "loess" function in the R package is an implementation of the more general concept of local regression analysis and investigation of other possible implementations

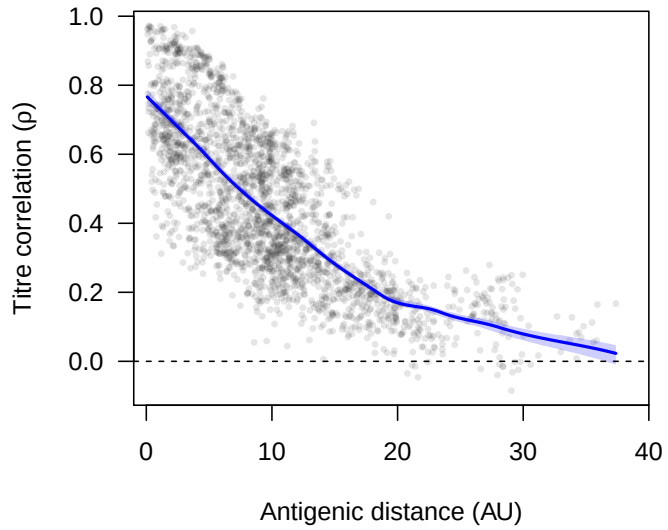


Figure 2.3: Correlation of titres vs. antigenic distance between viruses. Antigenic distance between all virus pairs plotted against the Spearman's rank correlation coefficient (*rho*) of titres measured against them. Individual points are shown in grey and the blue line represents a loess fit through the data (span=0.5). Light blue regions represent the 95% confidence interval of the loess fit, estimated from 500 bootstrap repeats (often not wider than the main blue line).

seemed appropriate. A fundamental concept is the "bandwidth", or the number of surrounding data points included when performing a local regression to calculate the loess fit for a given point. In general, the larger the bandwidth, the smoother and more constrained the fit will be.

The span variable included as part of the original R "loess" function implemented defines a variable bandwidth, by including a given percentage of the closest surrounding points when performing the regression. A span of 0.5 will thus include 50% of the surrounding points while a span of 0.8 will include 80%. This flexible definition works well when variables are homogeneously distributed and has the advantage of being scale-free, but in situations with more heterogeneous distributions (such as the distribution of virus strains sampled on the antigenic map) it can lead to very inconsistent bandwidth sizes depending upon the location being considered. This principle is illustrated in figure 2.4.

Given that it is expected to be largely the extent of cross-reactivity between different viral strains that governs the degree of correlation between titres, an intuitive alternative to variable bandwidth local regression is to use a fixed bandwidth that reflects the limit of this cross-reactivity. Regardless of the density of sampling we would expect a fixed limit of cross-reactivity, beyond which titre values should not be correlated and therefore these values should not be included in a regression analysis.

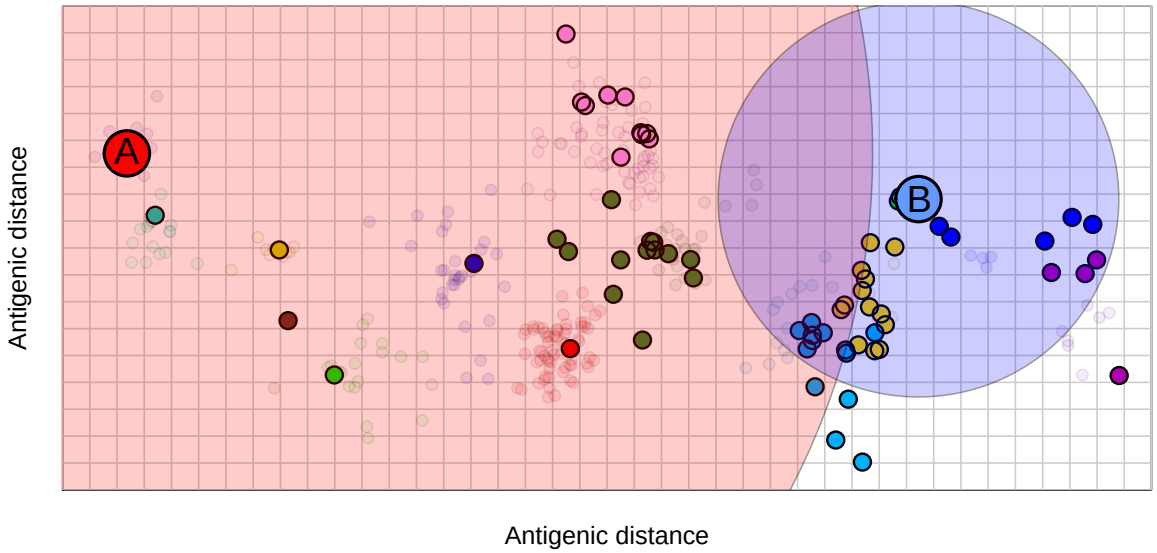


Figure 2.4: Inconsistent bandwidths with local regression using a fixed span. The antigenic map from figure 2.2 is again shown, with titrated strains highlighted. Additionally shown is the limit to which points would receive a positive weighting when performing local regression with a span of 0.5. For point A, being in a region of sparsely sampled antigenic space, encompassing the closest 50% of neighbouring points leads to a bandwidth as large as 28.6 antigenic units (the red shaded area). For point B, the bandwidth needed to encompass the nearest 50% of points is only 10.0 antigenic units. This inconsistent bandwidth can therefore be a problem when using a span rather than a bandwidth to fit datasets where points have a heterogenous distribution.

2.3.4 *The locally-weighted regression model*

To allow more flexibility a method of fixed-bandwidth analysis was implemented independently of the "loess" function originally used. It can be described as follows where each position on the antigenic map is associated with three variables: the x coordinate associated with the position in the antigenic map, x_i; the y coordinate associated with the position in the antigenic map, y_i; and the log-converted HI titre, z_i.

For a given point p_i (with associated antigenic coordinates x_i and y_i), the predicted landscape height $\hat{z}_i$ (in log units) is calculated as follows. First, Euclidean distances are calculated to obtain the antigenic distance, a_{ij}, of each antigen titrated against a serum with respect to the point in question, such that

$$a_{ij} = \sqrt{((x_j - x_i)^2 + (y_j - y_i)^2)} \tag{2.2}$$

where x_j and y_j are the x and y coordinates of the antigen. The distance of each antigen from the point p_i is then used to determine its weight when performing the linear regression, following a tricubic weighting function as follows:

$$w_{ij} = \begin{cases} (1 - (A/a_{ij})^3)^3 & \text{for } a_{ij} \leq A \\ 0 & \text{for } a_{ij} > A \end{cases} \tag{2.3}$$

Here w_{ij} is the weight of titre z_i with respect to point p_j and A is a variable that determines the maximum antigenic distance from p_j to which titrations made against viruses will receive a non-zero weighting when calculating the landscape height (such that $2A$ effectively represents the bandwidth of the local regression). Multiple linear regression is then applied to characterise the relation between the measured titres of an antigen and its coordinate variables x_j and y_j, with respect to the point p_i, such that for each antigen titrated:

$$z_j = \beta_1(x_j - x_i) + \beta_2(y_j - y_i) + \beta_3 + \mathcal{E}_j \tag{2.4}$$

where the regression coefficients β_1, β_2 and β_3, are resolved to minimise the weighted sum of squares of the errors, S, given by:

$$S = \sum_{j=1}^{n} w_{ij} {\mathcal{E}_j}^2 \tag{2.5}$$

Finally, since the regression is performed on the coordinates of antigens relative to the point being modelled, predicted values for p_i are then simply given by the intercept parameter β_3, such that:

$$\hat{z}_i = \beta_3 \tag{2.6}$$

Note that the variables β_1, β_2 and β_3 do not have a fixed value but are calculated separately for each location in the antigenic map. This is why locally-weighted-linear regression is often referred to as a "non-parametric", although of course many parameters such as A in equation 2.3 will be fixed and play an important role in the resulting fit.

The parameters β_1 and β_2 essentially describe the gradient of the plane tangent to the surface of the antibody landscape at the point in antigenic space being fitted. To avoid overfitting in more sparsely sampled antigenic regions, the decision was therefore made to constrain the parameters β_1 and β_2 such that they could not be greater than 1 or less than -1 (since a slope of 1 is assumed to be the maximum rate

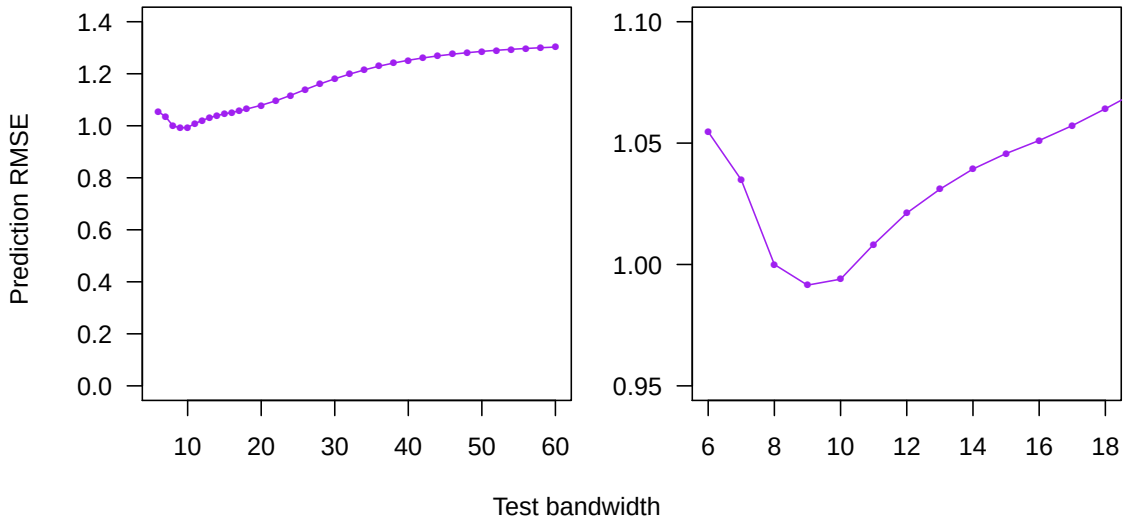


Figure 2.5: Cross-validation results for least-squares fits of different bandwidths. Fits were trained on random subsets of 90% of the data and tested on the remaining 10%, repeated 10 times with the entire dataset. The RMSE of the fit predictions when using different fitting bandwidths is shown twice, in the left panel with a y-axis scale starting from 0, and in the right panel, using a smaller x and y-axis scale to emphasise differences between the test bandwidths.

at which antibody reactivity will diminish with increasing antigenic distance from a stimulus [Smith et al., 2004]).

2.3.5 *Cross-validation of different fitting parameters*

As previously discussed in section 2.3.3 a key parameter of local regression is the bandwidth used to weight the data (A in equation 2.3). A small bandwidth will lead to a model fit that closely fits the training data but can be liable to over-fitting, while a larger bandwidth may create a fit that is not flexible enough to capture the true underlying relationship between variables. Cross-validation of models using different test bandwidths was therefore performed to try and ascertain the value that best predicted missing data when it was excluded from the training set.

For each of the serum samples in the test dataset (a total of 468 pre and post vaccination samples when combined), a random subset of 90% of the antigen titres were used as a training set used to infer the antibody landscape for that sample and predict titres for the remaining 10% of the data. This was repeated 10 times for the entire dataset, giving a total of 4680 cross-validation trials.

When comparing different bandwidths, that same training dataset was used for each value tested and the total root-mean-squared-error of the fit was calculated as,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(z_i - \hat{z}_i)^2}{n}} \tag{2.7}$$

where $\hat{z}_i$ is the value predicted when trained on the training set of titres and z_i is the actual measured titre that was excluded from the training data. Test antigens that fell outside the convex hull bounded by the training set of antigens were excluded since the intention is not that antibody landscapes should be extrapolated to infer values outside the region of antigenic space bounded by the antigens in the dataset. Predicted and measured titres were compared as per the fitting procedure on the $\log_2$ scale. Non-detectable titres were simplistically treated as -1 on the log scale while landscape predictions returning a value <-1 were also treated as -1. The result is that test antigens where the measured titre was non-detectable and the predicted value was <-1 were treated as having a residual of 0.

The results from the cross-validation testing are shown in figure 2.5. The RMSE of missing data predictions reached a minimum at a model bandwidth of between 9 or 10 antigenic units indicating that a value around this level should give a fit most representative of the data, at least for this dataset.

2.4 Data visualisation

A large part of the value of the antibody landscapes technique is to provide a simple, intuitive summary of large amounts of complex data, allowing patterns between and within individuals across large datasets to be visualised and compared. It was therefore an important part of the project to achieve an effective visualisation of antibody landscapes that could be easily generated and shared.

2.4.1 *Visualising a landscape in 3-dimensions*

As introduced at the start of this thesis, variation of H3N2 influenza strains across antigenic space is typically summarised in terms of a 2-dimensional antigenic map while the antibody landscape exists as a surface in the third dimension. To model a 3-dimensional surface, the fitting procedure described in section 2.3 can be applied grid-wise to predict the landscape height across 2-dimensional antigenic space and a surface interpolated between these points.

Although technically straight forward to model once the locally-weighted regression model has been built, effective visualisation of data in 3-dimensions is notoriously difficult but the R programming environment provides good tools for visualisation of

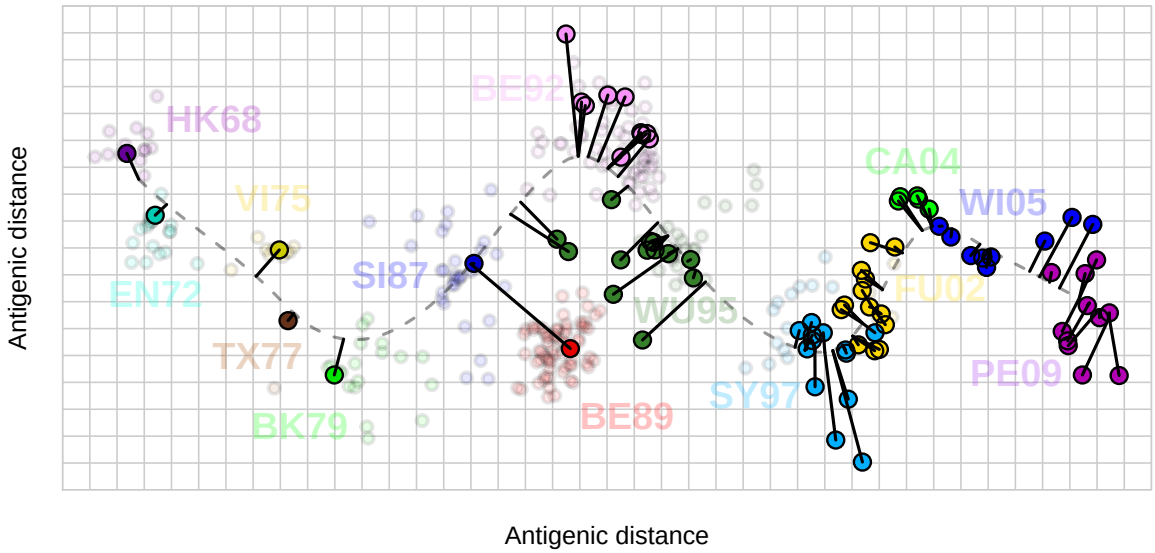


Figure 2.6: The 2-dimensional summary path. The antigenic map from figure 2.2 is again shown, with titrated antigens highlighted. The dashed line shows a smoothing spline fit through the titrated viruses, and the solid lines show corresponding projections of the viruses titrated onto this path.

such types of data. In particular the ability to add interactivity to a 3D plot can be a powerful way of allowing a better visualisation of the underlying data structure and this capability can be achieved through the package *rgl* [Adler et al., 2017], which was ultimately used as a basis for antibody landscape visualisation.

Figure 2.7 shows the end result of a stepwise visualisation of the construction of an antibody landscape, beginning first with the underlying antigenic map before showing titres as impulses in a third dimension and finally an antibody landscape surface fit to the data. Links to interactive variants of these plots and other 3D plots in this thesis are included in appendix C.5.

2.4.2 *Visualising antibody landscapes in two dimensions*

Due to the difficulties inherent when trying to visualise change between two 3D surfaces, it was suggested that the roughly linear evolution of H3N2 strains through antigenic space could be exploited to allow a summary "slice" to be taken through each landscape. Following the initial methodology applied by Judy Fonville but with some minor modifications, a smoothing spline was fit through the 81 antigens that form the total set of antigens titrated against the various datasets included in this thesis and listed in full in appendix A. The R function *smooth.spline* from the base package *stats* [R Core Team, 2016] was used to perform the fitting with the *spar*

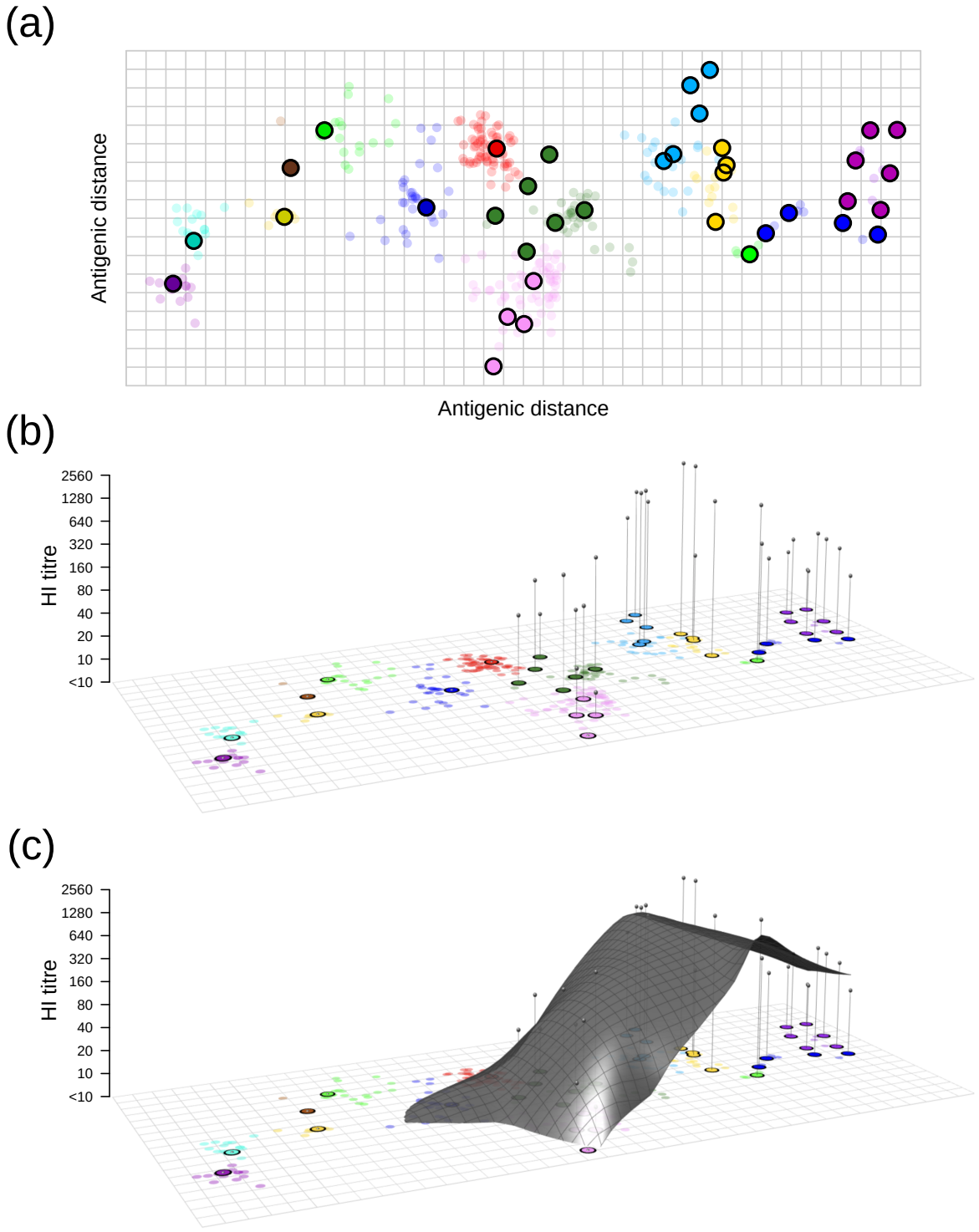


Figure 2.7: Stepwise visualisation of the construction of an antibody landscape. **(a)** Antigenic map of H3N2 from figure 2.2, showing virus strains colour-coded by antigenic cluster. **(b)** An additional dimension indicates the measured antibody titres as vertical impulses. **(c)** A smooth surface is fitted using locally-weighted multiple linear regression to create the antibody landscape within the convex hull bounded by the viruses titrated.

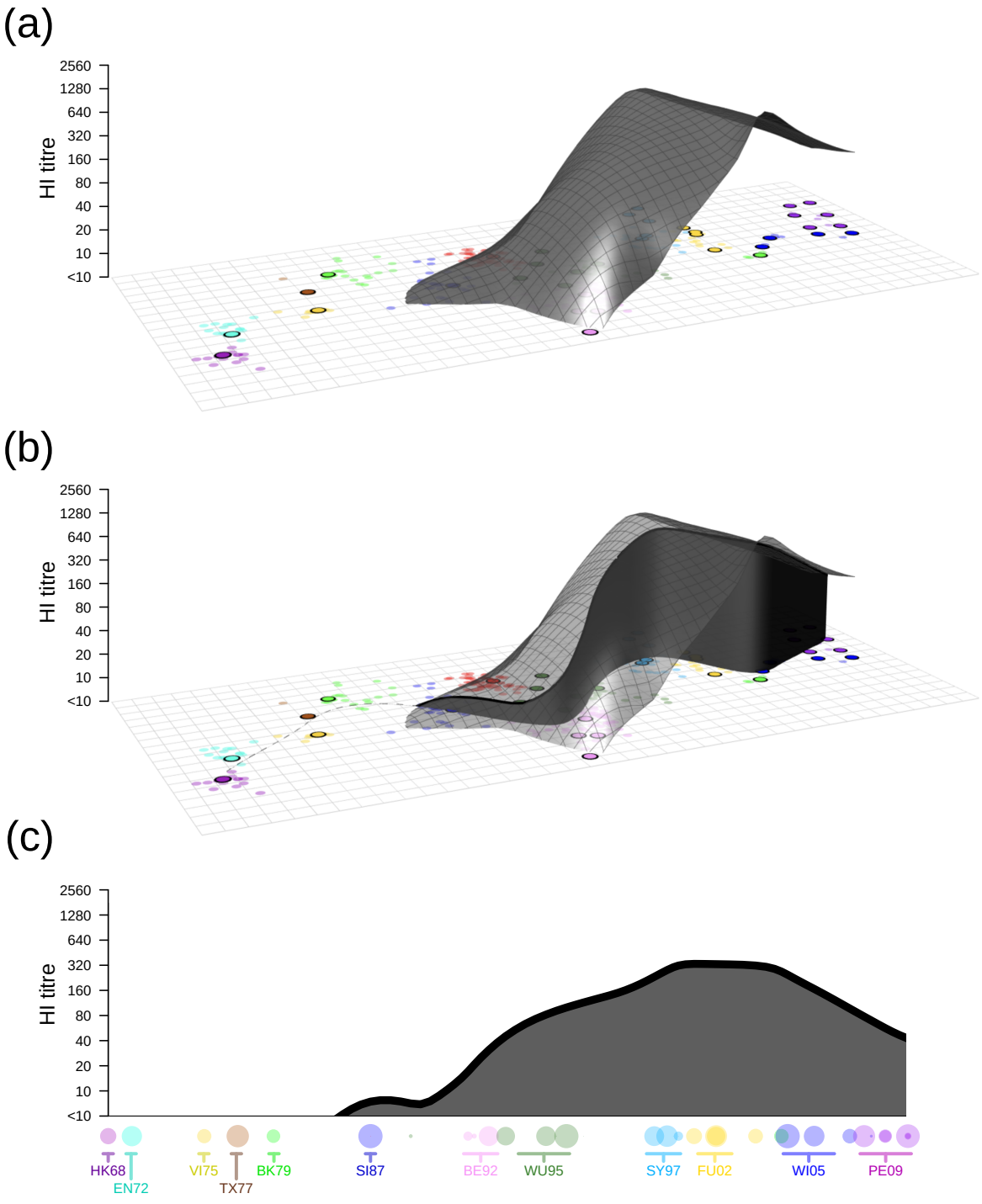


Figure 2.8: Stepwise construction of a 2-dimensional summary antibody landscape. (a) The antibody landscape from figure 2.7 is shown. **(b)** A slice from the landscape is taken along the path following the antigenic evolution of H3N2 as defined in figure 2.6. **(c)** The height of the landscape at points following the summary path is plotted against distance along it to create a two-dimensional summary. Titrated virus strains are shown along the x axis scaled inversely to the antigenic distance from the summary path; symbol colour indicates antigenic cluster.

argument set to 0.8. The resulting path through H3N2 antigenic space is shown in Figure 2.6.

Figure 2.8 shows how the corresponding values of an antibody landscape along the summary path defined can then be used to create a two-dimensional summary of the three-dimensional landscape. This 2D summary traces the height of the antibody landscape along the path through antigenic space, similar to the elevation profiles provided for race routes or hiking trails. When evaluating the summary representations along the path, it should be kept in mind that the axis of this path does not equal antigenic distance, because the path curves, and therefore two points along the path may be further apart than the direct distance between them. That being said, it reflects a good approximation of distance between strains in most cases.

To help orientate the rough antigenic location of each position along the antigenic summary path, the positions of any viruses against which the sample was titrated are shown beneath the x axis at the position of their projection onto the summary path, scaled negatively by distance from it. Due to the roughly linear nature of H3N2 evolution, the path also acts as a proxy for the different times at which the test viruses circulated and labels in figure 2.8 also give an indication of where strains from the different antigenic clusters sit along it.

2.4.3 *Adding further interactivity and portability through WebGL*

The *rgl* package [Adler et al., 2017] for R by default generates 3-dimensional plots viewed with basic interactivity using a native X11 viewer but it also provides functionality to output plots to webpages, encoding them as WebGL images. This latter approach was pursued and developed to allow plots to be viewed interactively in a web browser and easily shared as html files, without the need to download the code, data and packages needed to generate the plots locally. With some customisation, functionality to toggle different components of the antibody landscape plots on and off was also included, making it possible to focus on certain aspects of the data, either the raw titres, 3-dimensional landscape fit, or how the summary path fits through a particular landscape. This approach was used to generate the online interactive versions of the 3D plots in this thesis listed in appendix C.5.

2.5 Handling non-detectable titres

Since locally weighted multiple linear regression is used to fit the antibody landscape surface, there is always a question of how best to estimate the regression slopes and intercept used to predict the antibody landscape height at any given antigenic point for a sample where non-detectable HI titres have been recorded.

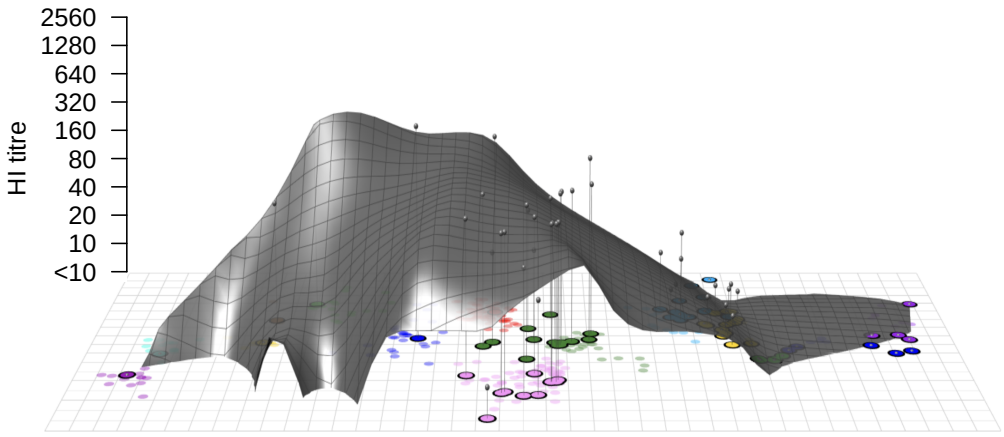


Figure 2.9: Example landscape fit to non-detectable titres using the least-squares approach. An antibody landscape is represented as in figure 2.7. The antigenic base-map is shown with titrated strains highlighted and corresponding antibody titres represented as vertical impulses. The antibody landscape surface is shown in grey. Note the plateauing of the surface fit when approaching the region of non-detectable titres on the right.

A common approach is simply to treat non-detectable titres of <10 as equal to 5, the next lowest detectable value following a two-fold dilution protocol. This makes these values equal to -1 on the $\log_2$(HI titre/10) scale used to model antibody landscapes. Although easily implemented, this inevitably leads to overestimates when fitting data since non-detectable titres are always assumed to reflect the highest possible value with which they could be associated. Another approach sometimes implemented is to exclude non-detectable titres but this too biases fits by selectively removing those measurements that have returned a low value.

Although treatment of <10 titres as equal to 5 may introduce only a negligible bias when such titres comprise only a small proportion of a whole dataset, it leads to misleading results when a larger number of non-detectable measurements are present. In the case of antibody landscapes, figure 2.9 shows how the effect is an artificial plateauing of a surface as it tends towards an area of undetectable titres.

2.5.1 *Optimising undetectable titres*

An initial solution to the treatment of non-detectable titres when fitting antibody landscapes was to allow these titres to vary and assume any value below the limit of detection in the HI assay ($\leq$-0.5 on the log scale). These values were then optimised so as to minimise the overall root-mean-square error of the fitted antibody landscape

titres, including of course the detectable titres, which remained fixed at their measured values [Fonville et al., 2014].

Although this approach solved the problem of fits plateauing as they reached antigenic regions of non-detectable titres, it is an approach with some limitations. Especially where multiple non-detectable titres are present, it is relatively computationally intensive and a non-trivial optimisation process, with solutions dependent upon initial starting values for the non-detectable values used.

A more subtle problem is that the fit fails to account for any variation due to measurement error in the HI assay which can actually give useful information about how low underlying antibody reactivity is likely to be in a given antigenic region. For example a region of exclusively non-detectable titres likely reflects underlying antibody reactivity that is not just below, but significantly below the threshold of detection since even with positive measurement error noise, which would be expected in the case of some antigens, no detectable values were recorded.

To attempt to address these shortcomings, further work was performed to improve the procedure to deal with non-detectable titres as described in the next section.

2.5.2 *Maximum-likelihood estimation*

Maximum-likelihood estimation provides an alternative approach to least-squares when fitting a model to data such as the local multiple linear regressions performed to construct the antibody landscapes. In the simple case of linear regression to data with normally distributed errors, least-squares and the maximum-likelihood estimator give the same fit, however maximum-likelihood estimation also gives the flexibility to specify different degrees of uncertainty in different measurements [Hald, 1999].

A first step in maximum-likelihood estimation is to specify the error function for your model - in the case of antibody landscapes, the amount measured values would be expected to deviate from an ideal fit. Error in the case of antibody landscapes comes from a number of potential sources:

- Inaccuracies in the positioning of viruses in the antigenic map.
- Differences between viruses and their levels of reactivity in the HI assay.
- Measurement error of the HI assay itself.

As an estimate of the total effect of all these errors, figure 2.10 shows the distribution of the absolute difference of titres between all pairs of antigens in the test dataset less than 0.2 antigenic units apart. As seen in section 2.3.1, figure 2.3, titres between pairs of antigens are expected to become less related with increasing antigenic distance but at such small distances this effect should be negligible, with most of the difference in titres attributable to the factors listed above. It can be seen that the absolute

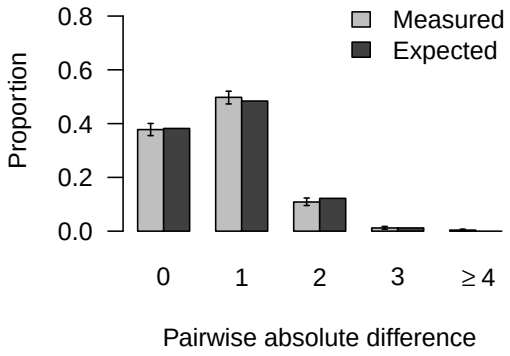


Figure 2.10: Barplot of pairwise titre differences between antigenically similar viruses. The proportions of different absolute pairwise differences between all viruses less than 0.2 antigenic units apart are shown in light grey with error bars showing the 95% confidence interval based on 500 bootstrap repeats. Dark grey bars show the proportions of pair-wise differences expected from normally distributed errors with a standard deviation of 1. Titre pairs including a non-detectable value were excluded.

measured differences between these antigen pairs are fit well by the values expected from a normal distribution with a mean of 0 and standard deviation of 1, folded about 0 (since absolute differences were compared) and rounded to the nearest integer (to account for the discrete nature of the log-transformed HI titres). Note that in the data shown, titre pairs that included a non-detectable titre were excluded.

Assuming the error distribution described above, an approach akin to Tobit regression [Tobin, 1958] was used to fit the data and account for the additional uncertainty associated with non-detectable titres. Formally, the error function $F(n, \mu)$ was defined as,

$$F(n, \mu) = \Phi\left(\frac{n - \mu}{\sigma}\right) \tag{2.8}$$

where μ is the mean of the distribution, σ is the standard deviation (set to 1), and Φ is the cumulative distribution function (CDF) of the standard normal distribution.

Next, instead of estimating the coefficients of the multiple linear regression described in section 2.3.4, equation 2.4 by minimising the weighted sum of the squared residual error as described in equation 2.5, coefficients were found to maximise the weighted log-likelihood of titre measurements given the model predictions. The likelihood of obtaining a given measurement z_i given a multiple linear regression model prediction of $\hat{z}_i$ was defined as,

$$P(z_i|\hat{z}_i) = F(n = n_i^+, \mu = \hat{z}_i) - F(n = n_i^-, \mu = \hat{z}_i) \tag{2.9}$$

where F is the error function defined in equation 2.8 and n^+ and n^- are upper and lower thresholds of the measured value. Since measurable values of z may take only integer values in the detectable range and non-detectable measurements represent values somewhere below the measurable limit of the assay, n^+ and n^- can be defined

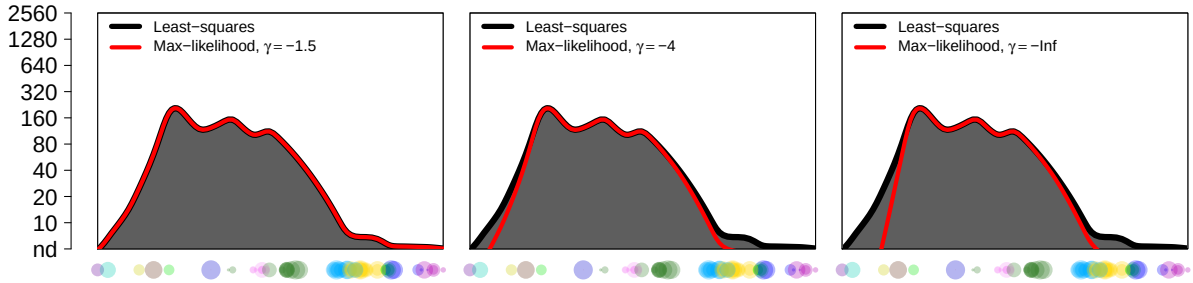


Figure 2.11: Example effects of assuming different minimum-thresholds for non-detectable titres. 2-dimensional landscape summaries are shown for an example serum comparing use of the least-squares fitting approach to (black line) to maximum-likelihood approaches assuming a different minimum threshold (γ) for non-detectable titres. Note that the maximum-likelihood fit exactly follows the least-squares fit when $\gamma = -1.5$ (left panel). Rest of figure caption as for figure 2.8 panel c.

as,

$$n_i^+ = \begin{cases} z_i + 0.5 & \text{if } z_i \notin \text{nd} \\ -0.5 & \text{if } z_i \in \text{nd} \end{cases} \qquad n_i^- = \begin{cases} z_i - 0.5 & \text{if } z_i \notin \text{nd} \\ \gamma & \text{if } z_i \in \text{nd} \end{cases} \tag{2.10}$$

where γ is an additional parameter that can be thought of as the lowest likely value on the log scale attributable to a non-detectable titre. Finally, the weighted log likelihood I to be maximised is then defined as,

$$I = \sum_{i=1}^{k} \log\left(P(z_i|\hat{z}_i)\right) w_i \tag{2.11}$$

where, as previously, w_i the weight for that titre measurement scaled negatively with distance from the antigenic point at which the landscape height is being calculated.

As mentioned with regard to equation 2.9, γ is an additional parameter to be decided when modelling the antibody landscape using this approach. In the extremes, a value of $-\infty$ would imply no constraint on the lower limit of what could count as an undetectable titre while a value of -1.5 is like treating non-detectable titres as a detectable log titre of -1 as was the initial approach with least-squares. Figure 2.11 shows the effect of different values for γ when fitting an example landscape, it can be seen that a value of -1.5 leads to a fit that exactly matches that of the least squares method, while a value of $-\infty$ fits regions of non-detectable immunity the most aggressively. A value of γ=-4 leads to an intermediate fit that still correctly reflects the region at which titres become largely undetectable but is not so heavily weighted to fit these regions.

As figure 2.11 also demonstrates, the antibody landscapes calculated using the maximum-likelihood methodology differ only slightly when compared to the original approach

described, in the antigenic regions of non-detectable titres. When cross-validation results from the maximum-likelihood fit were compared to the least-squares approach, the results were very similar, although interestingly the performance of the maximum-likelihood approach was very slightly worse overall (see figure 2.12).

The reason for a reduced performance under cross-validation when using a maximum-likelihood fit is interesting and likely partly due to the way the RMSE was calculated. As described in section 2.3.5, non-detectable titres were simplistically treated as -1 on the log scale while landscape predictions returning a value <-1 were also treated as -1. The result of this cropping of both non-detectable titres and predictions to -1 is that the residuals associated with non-detectable values tend to be lower, and improved fits to these titres contribute less to the overall RMSE of the fit than fitting to detectable titres.

Figure 2.13 shows the prediction RMSE for both approaches broken down by fits to detectable and non-detectable titres. Looked at in this way, we can see that there is indeed better fitting of non-detectable titres under maximum-likelihood approach but this comes at a cost of worse fitting to detectable titres, perhaps due to overfitting of non-detectable titres in some of the cross-validation training and test sets.

Ultimately, the predictive capacity of each fitting approach and the way in which fits to non-detectable vs detectable titres should be weighted is a matter of judgement, but the clearest benefit is seen in terms of the effect on data visualisation. Figure 2.14 reproduces the landscape modelled in figure 2.9 but now using the maximum-likelihood approach. The shape of the overall landscape is largely maintained but, where previously a misleading plateaux was present as the landscape approached but did not cross the non-detectable titre boundary, now there is a clear delineation of the regions of detectable and non-detectable reactivity in antigenic space.

2.5.3 *Estimating fold-change of serum reactivity*

Sometimes it is useful to estimate the fold-change in serum reactivity that has occurred from one sample to the next, for example pre and post vaccination or infection. The maximum-likelihood approach also provides a natural framework in which to deal with changes between measured titres that may also be censored due to non-detectable measurements. In these cases, fitting can be performed as already described but measurements are now those of fold-change difference. Specifically, in the case of a fold-change calculation, n^+ and n^- from equation 2.10 become limits of the measured fold change. Rather than taken from a single log-transformed HI measurement against a given antigen (previously z_i), the fold change is calculated from

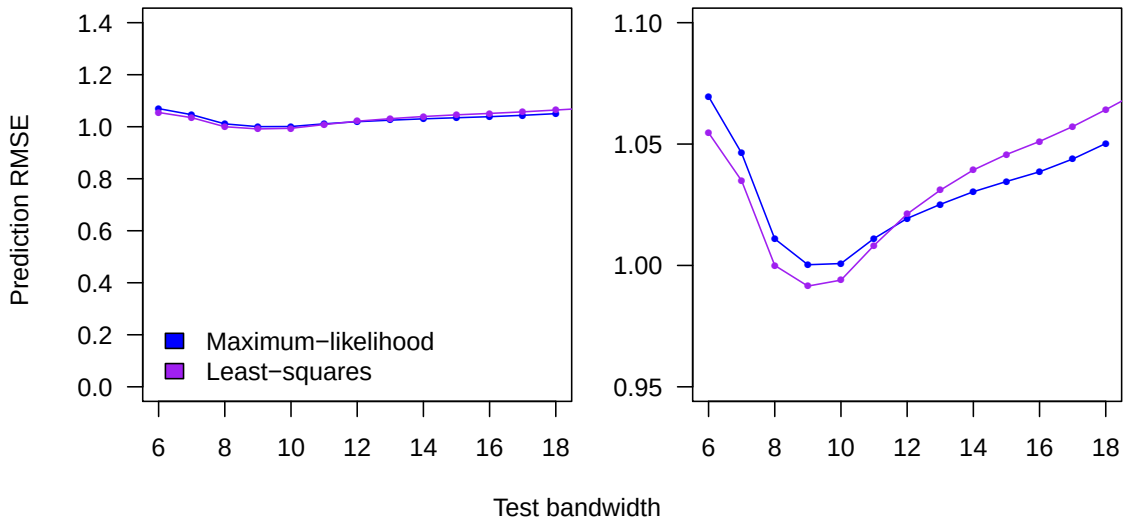


Figure 2.12: Cross-validation of max-likelihood and least-squares fitting approaches compared. Results from the least-squares fit shown in figure 2.5 compared to the results using the maximum-likelihood fitting approach to the same training and testing sets. The left panel shows the results on a y axis scale stretching to 0, while the right panel shows the same results on a scale chosen to better visualise differences in the results.

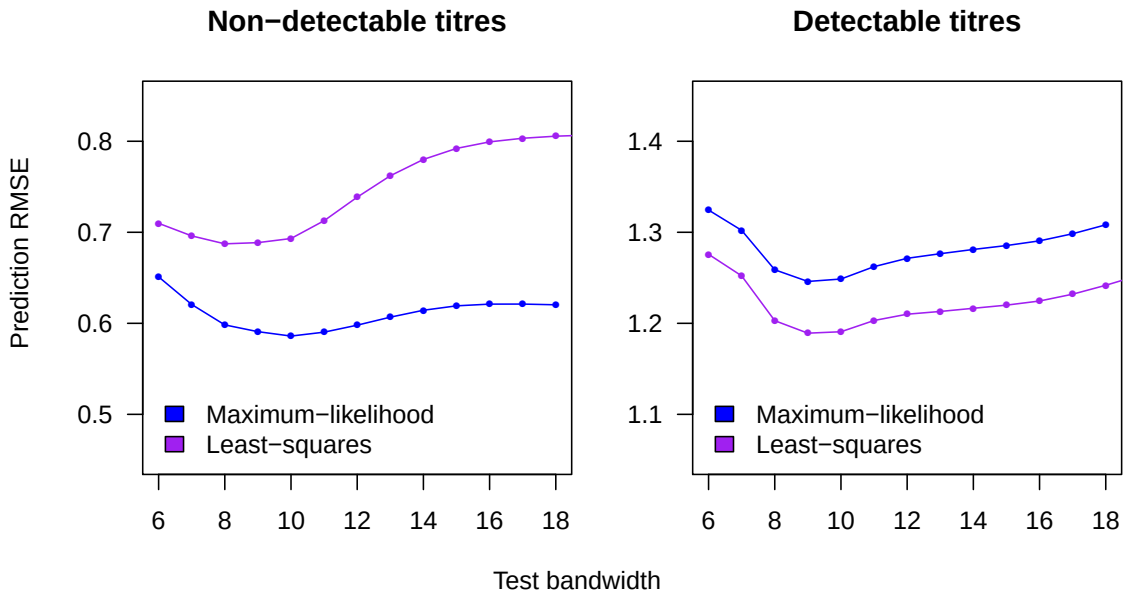


Figure 2.13: Cross validation results broken down by detectable and non-detectable titres. Results from the same cross-validation test shown in figure 2.12 broken down by predictions made against detectable and non-detectable titres.

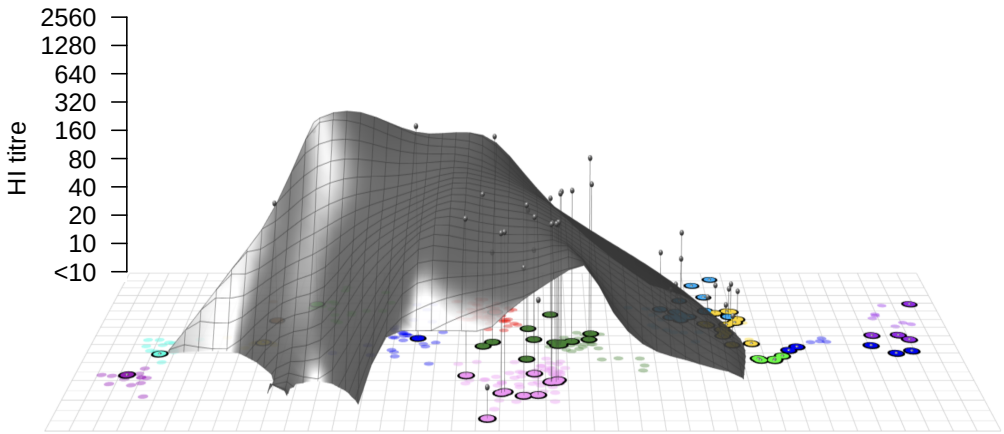


Figure 2.14: Example landscape fit to non-detectable titres using the maximum-likelihood approach. An antibody landscape fit to the same data as that in figure 2.9 but this time using the maximum-likelihood approach. Note how the landscape fit now disappears below the detection threshold as it approaches the region on non-detectable titres on the right of the plot, more accurately reflecting the overall pattern of antibody reactivity present in this serum.

one measurement to another, here represented as z_{i1} and z_{i2}. Similarly to before, upper (z^+) and lower (z^-) thresholds for either measurement are defined such that,

$$z^+ = \begin{cases} z + 0.5 & \text{if } z \notin \text{nd} \\ -0.5 & \text{if } z \in \text{nd} \end{cases} \qquad z^- = \begin{cases} z - 0.5 & \text{if } z \notin \text{nd} \\ \gamma & \text{if } z \in \text{nd} \end{cases} \tag{2.12}$$

while equation 2.10 now becomes,

$$n_i^+ = z_{i2}^+ - z_{i1}^-, \qquad n_i^- = z_{i2}^- - z_{i1}^+ \tag{2.13}$$

2.5.4 *Prediction of immunity below the HI detection threshold*

In addition to more accurate modelling of the boundary regions in antigenic space between detectable and non-detectable titres, more appropriate statistical treatment of non-detectable titres also potentially allows us to infer landscapes below the detectable threshold to a certain degree. Ideally, it would be most informative to see how predictive the height of antibody landscape fits below the detection threshold are when antigens in these regions are titrated against serum concentrations of higher than the starting dilution of 1:10. Without access to such data however, a surrogate

approach is to first fit a landscape to all the data and then compare how the fit deviates when the data is artificially censored at higher and higher levels.

The results of a landscape fit to such artificially censored data are shown in figure 2.15 where an example landscape fit is repeated when censoring the data at 6 different levels. The maximum-likelihood approach manages to infer the landscape below the detection threshold remarkably well, with virtually no effect on the landscape fit when censoring the data even up to a level where the lowest titre measurement is <160 and 67.1% of the data is non-detectable compared to only 45.7% in the uncensored dataset. Only at a censoring level of <320 does the fit begin to deviate more substantially where 94.3% of the titres are now below the detection threshold, before finally, in the most extreme case of no detectable titres using a threshold of < 1280, a flat landscape is returned.

Figure 2.16 shows the fits not of a single landscape but instead of an average population landscape, constructed by fitting the data from all serum samples of a group of individuals simultaneously - in this case all post-vaccination serum samples from individuals in the 1998 vaccine trial. Here, with more data from which to infer the landscape, the quality of fits below the detection threshold are even more impressive. Remarkably, there is still almost no change in the overall fit even when titres are censored to a level of <320, a degree to which 88.1% of the data is non-detectable and the average population landscape is always below the detectable threshold.

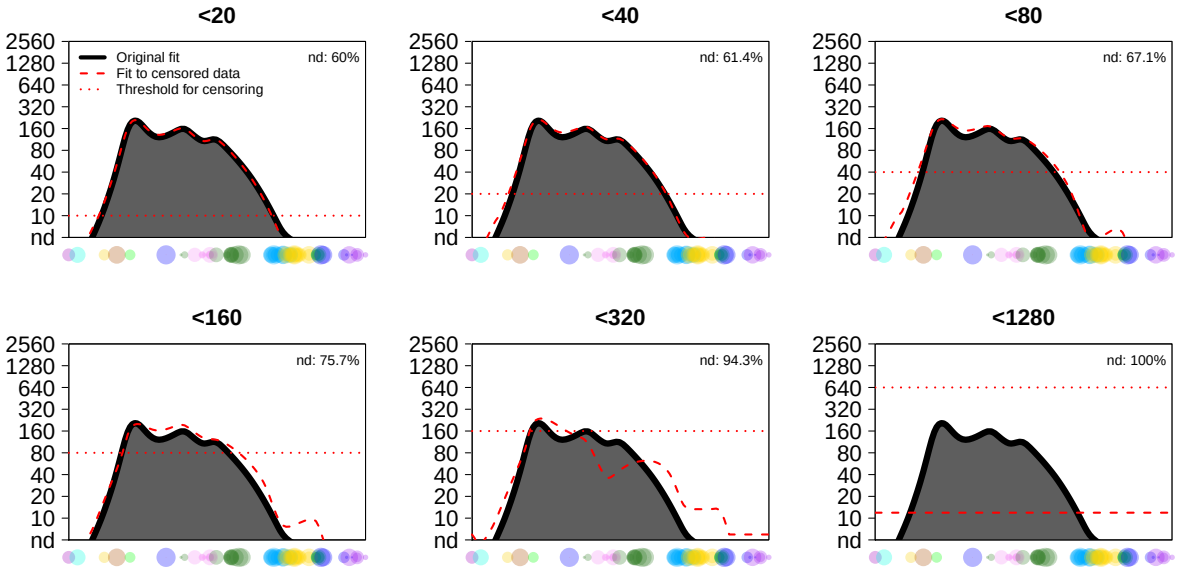


Figure 2.15: Fits to artificially censored data for a single landscape. In each panel 2-dimensional summary antibody landscape fits are shown after artificially censoring the data to different levels (red dashed line) compared to a fit made to the original data (solid black line). The censoring level used for each panel is given in the title and marked on the y-axis by a dotted red line. For example in the top left panel (<20), all titres of 10 or <10 were numerically treated as <20. The proportion of non-detectable (nd) titres below each censoring level is shown in the top-right corner. For reference, 45.7% of data was non-detectable in the uncensored titration set.

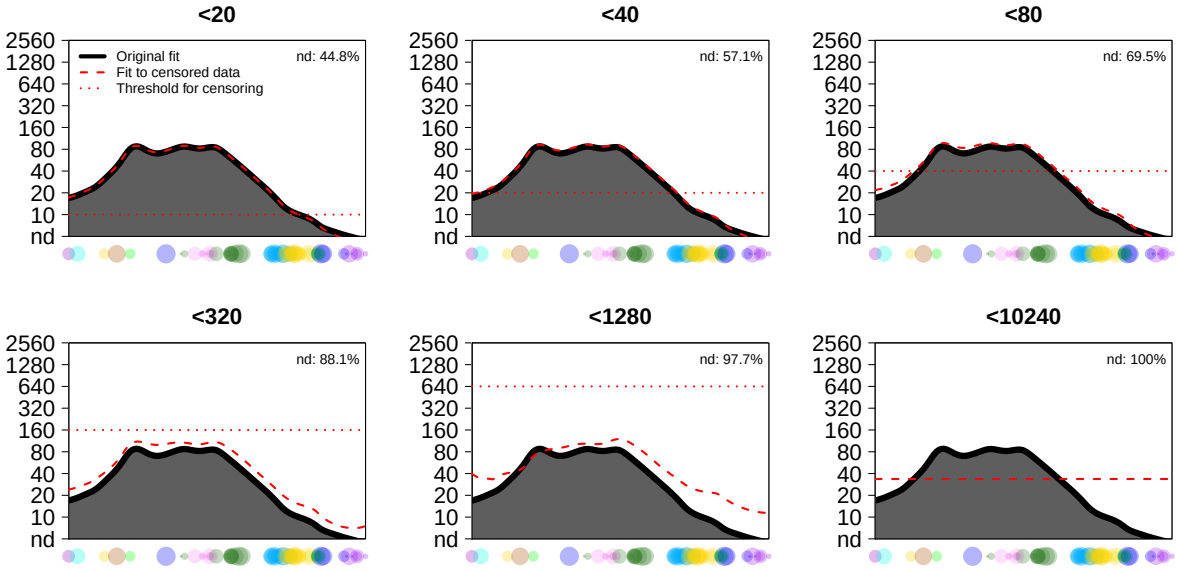


Figure 2.16: Fits to artificially censored data for a population landscape. Caption as for figure 2.15 but now comparing the effect of censoring on an average landscape fit to all post-vaccination serum samples from individuals in the 1998 vaccine trial (128 samples). For reference, 29.3% of data was non-detectable in the uncensored titration set.

2.6 Estimating confidence intervals

2.6.1 *Confidence intervals for an individual landscape fit*

An important aspect of assessing antibody reactivity effectively is to determine the confidence to which individual patterns and differences between sera can be interpreted. In the case of antibody landscapes we are interested in how affected the landscape fit can be by variation in titres due to HI error and the specific set of antigens selected for titration.

One approach to infer uncertainty is to bootstrap the data, resampling with replacement the antigens and titres used to construct the antibody landscape and examining how much variation this creates in the resulting landscape fits. Although a well-recognised method of calculating uncertainty in a wide-range of applications, the drawbacks of bootstrapping are that it is computationally intensive (since the fit must be repeated many times on different subsets of the data) and that it requires a relatively large amount of data, since it is assumed that the variation in the data sample is reflective of the variation in the "population" of possible measurements as a whole.

Maximum-likelihood fitting also gives an alternative to estimating uncertainty by calculating confidence intervals based on the profile-likelihood of relevant parameters [Raue et al., 2014]. The relevant parameter in the case of the antibody landscapes is β_3 from equation 2.4 in section 2.3.4, the intercept of the multiple linear regression, since this parameter directly determines the antibody landscape height at any given point.

Explicitly, assuming a point estimate for the parameter β_3 with a maximum-likelihood of P_{max}, two-sided 95% confidence intervals can be calculated by finding the values of β_3 that satisfy:

$$P(\beta_3) = P_{max}.F_{\chi^2}\left(\frac{1-0.95}{2}\right) \tag{2.14}$$

where F_{χ^2} is the cumulative probability function for a chi-squared distribution with 1 degree of freedom.

Computationally, solutions can be approximated for the upper/lower bound of the confidence interval by increasing/decreasing the value of the β_3 by small increments until the likelihood of the fit begins to exceed the threshold likelihood defined in equation 2.14.

Figure 2.17 compares the 95% confidence intervals for a landscape fit as calculated either by the profile-likelihood approach described above, or by bootstrapping the data. It can be seen that at the points where numerous surrounding antigens were

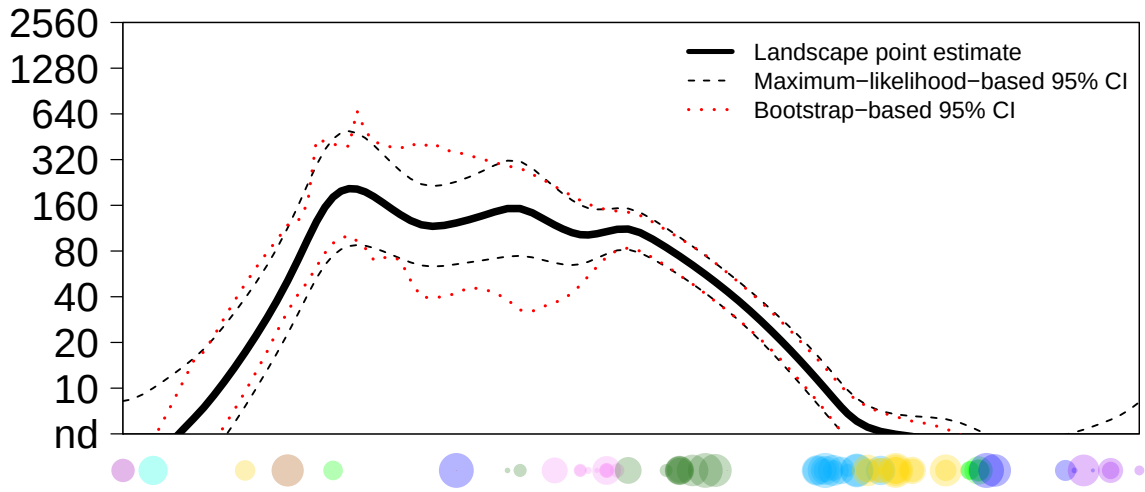


Figure 2.17: Example confidence intervals for a single landscape. The 2-dimensional antibody landscape summary fit is shown for an example serum (black solid line) alongside 95% confidence intervals for the fit based either on either 1000 bootstrap repeats (red dotted line) or a maximum-likelihood approach (black dashed line) as described in the main text.

titrated (and there is enough data to expect that bootstrapping provides a good estimate of fit confidence), the maximum-likelihood and bootstrapping approaches provide almost identical estimates. In regions further back on the summary path, confidence intervals are broadly similar but the bootstrap-estimates become more erratic - expected given the sparsity of data in these regions of the antigenic map.

In the case of bootstrapping the data, 1000 bootstraps repeats were performed, re-fitting the antibody landscape to a random sample (with replacement) of antigens and their associated titres. Confidence intervals were then inferred by taking the 2.5th and 97.5th percentile of the fits at each location along the summary path. Points where the landscape fell outside the convex hull bounded by the antigens included in a particular bootstrap sample were excluded.

2.6.2 *Confidence intervals on the population scale*

Estimating confidence intervals for an average antibody landscape on the population scale represents a slightly different challenge to individual confidence intervals. In these cases the measurements are not wholly independent since there is both per-strain variation, relevant since the same strains will have been measured across multiple samples, and per-sample variation, relevant since each sample will show variation around a population average.

One approach to accounting for such different sources of error is a mixed effects model, where the effect of the particular strain used and the particular sample being measured are modelled as additional random effects. In this case, equation 2.4, describing how the titre z measured against a strain is expected to be related to it's antigenic coordinates relative to a point i, becomes:

$$z_j = \beta_1(x_k - x_i) + \beta_2(y_k - y_i) + \beta_3 + u_n + b_k + \mathcal{E}_j \tag{2.15}$$

where u_n is a random effect due to the sample n against which the measurement was performed and b_k is a random effect due to the particular antigen k being measured. In this form x_j and y_j are better represented as x_k and y_k to reflect the fact that the antigenic coordinates x and y are associated with the particular virus measured, k, rather than the particular measurement taken, j. u_n and b_k are expected to come from a normal distribution with mean 0 and standard deviations of σ_n and σ_k respectively.

Calculating and fitting such mixed effects models can be difficult, especially in the case of censored data, but the R package "survival" [Terry M. Therneau and Patricia M. Grambsch, 2000; Therneau, 2015] provides routines for this kind of approach through use of the *Surv* and *survreg* functions. Normally distributed mixed effects can then be added to a basic weighted linear model through use of a frailty term with the function *frailty.gaussian* (see Therneau et al., 2012 and Ripatti and Palmgren, 2000). To aid fitting, based on figure 2.24 σ_k was set at a value of 0.7, while σ_n (which would be expected to vary depending on the antigenic location and population tested) was fit to the data.

Figure 2.18 shows the confidence intervals calculated for a population landscape from a sample of either 10 or 100 using the mixed-effects-model approach. In each case, bootstrapping was performed as described for the individual landscape but additionally, the individual samples themselves were sampled with replacement to capture inter-individual variation. Similarly to figure 2.17, the bootstrap confidence intervals closely match those estimated from the original data. The shape of the confidence intervals are also what would intuitively be expected. In the case of estimating the population landscape from 10 samples, the confidence intervals remain mostly constant throughout antigenic space, reflecting the fact that most uncertainty is coming from the limited number of individual samples being used to infer the overall landscape. In the case of 100 samples, confidence intervals widen more perceptibly in the region of antigenic space where titration density has become more limited, reflecting the fact that in these regions, additional uncertainty due to the limited number of titrated antigens plays a proportionally larger role.

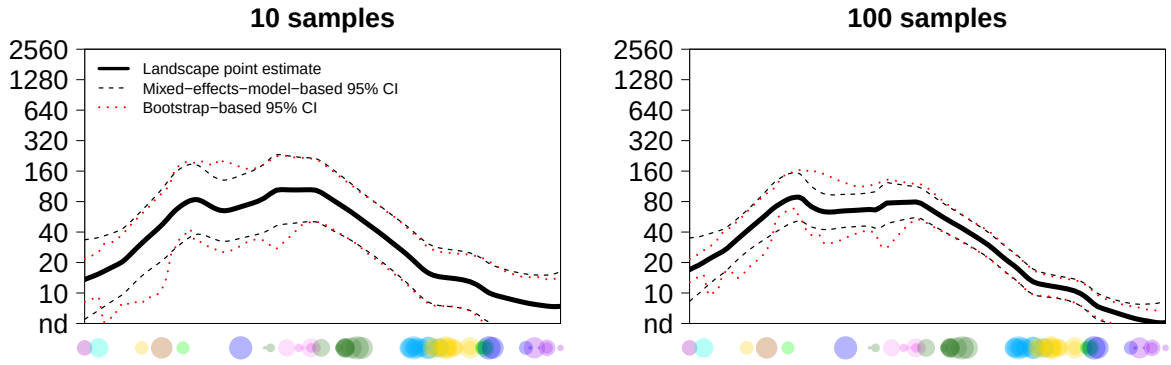


Figure 2.18: Example confidence intervals for a population landscape. The 2-dimensional antibody landscape summary fit is shown for different average population landscapes (black solid line) alongside 95% confidence intervals for the fit based either on either 1000 bootstrap repeats, bootstrapping both sera and antigens, (red dotted line) or based on the mixed-effects-model (black dashed line). The left panel shows the estimated fit and confidence intervals for 10 sera randomly selected from all post-vaccination serum from the 1998 vaccine trial and the right panel shows the same for a random selection of 100 sera.

2.6.3 *Estimated coverage from simulated data*

As an additional test, simulated data was used to estimate the coverage of the confidence intervals when fit to different random samples. In the case of fitting to a single sample, "true" values for the HI titre at each antigenic point were assigned based on the height of an example landscape at the antigenic location of the measured viruses. Samples were then repeatedly simulated by adding normally distributed error to both individual viruses and individual measurements (each assuming a standard deviation = 0.7). The subset of viruses used to fit the map with each simulation was also altered by taking a random selection of viruses from the full set of 338 shown in figure 2.1, but matching the heterogeneity of sampling throughout antigenic space by ensuring that the number of viruses drawn from each antigenic cluster was still the same as in the original sample. To simulate fitting to a population landscape, true values were assigned based on the landscape fit to post-vaccination serum samples from individuals in the 1998 vaccine trial. For a given number of simulated samples, in addition to the per-measurement and per-virus error, variation was also added on a per-sample basis. This individual variation included both general variation, simulated by adding a fixed normally-distributed amount to all measurements from a given sample, and some "shape" variation. This shape variation was modelled by taking a random walk and assigning the results as titres for each antigenic cluster centroid in chronological order. A landscape was then fit to these cluster-centroid titres and the resulting landscape added to the baseline one - an approach that was found to give variation in individual landscape shape that was visually similar to that seen in the actual data.

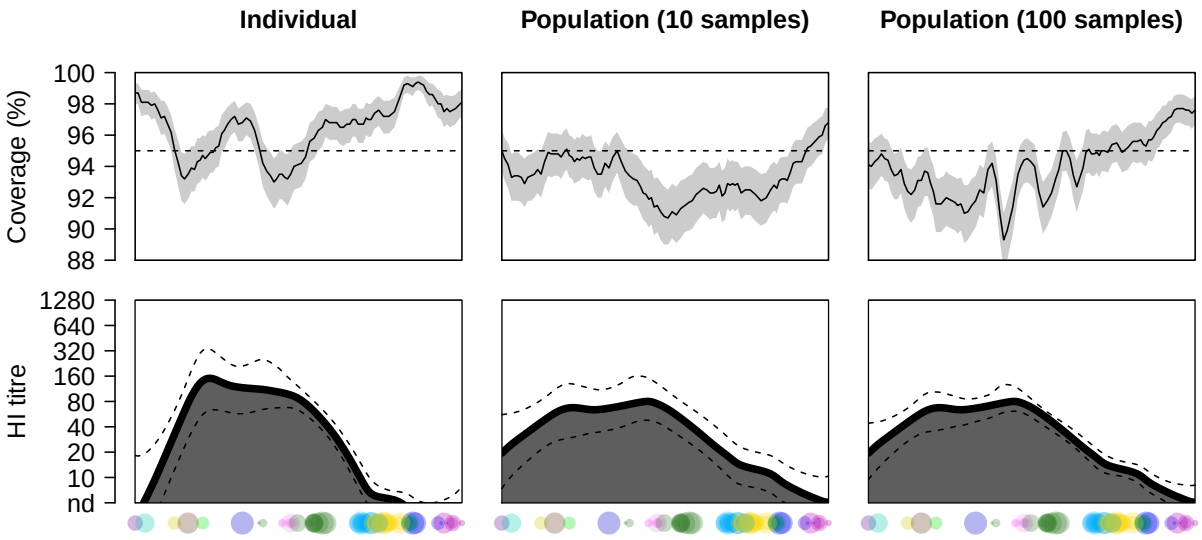


Figure 2.19: Simulated confidence interval coverage. Using the baseline antibody landscapes shown in the lower panel, noise was repeatedly added to the assumed "true" titres to simulate multiple sampling repeats. 95% confidence intervals of antibody landscapes fit to the simulated data were then compared to the original landscape fit to estimate actual confidence interval coverage under such a scheme. The upper panels show the percentage of times that confidence intervals spanned the "true" fit from 1000 simulations, across the different regions of antigenic space. The black line is the actual percentage coverage measured, with grey bands showing the 95% confidence interval for the coverage estimate based on 1000 bootstrap repeats. The horizontal dashed line in the upper panels shows the target coverage of 95%, while the dashed lines around the landscapes in the lower panel show the average lower and upper bounds of the confidence intervals from the 1000 simulation runs.

Figure 2.19 shows the results of the data simulations. The top panels show the estimated coverage from 1000 simulations, counted as the percentage of times that the antibody landscape fit to the true titres fell within the upper and lower bounds of the 95% confidence interval that was calculated when fit to the simulated data. The bottom panels show the landscape fit to the true titres as a black line, with the mean upper and lower confidence bounds of fits from the 1000 simulations. In each case, from left to right, the plots refer to estimating confidence intervals for an individual landscape, confidence intervals for a population average based on 10 samples, and confidence intervals for a population average based on 100 samples.

Encouragingly, as with comparisons to the bootstrap-based confidence intervals, coverage based on the simulated data appears to fall approximately in the target range of 95%. As with most analytical approaches I would still be hesitant to draw firm conclusions based solely on confidence intervals calculated in this way, but the methods used here appear to give a largely fair indication of the true uncertainty in the data.

2.7 Diagnostic tests

Although the antibody landscapes methodology was developed and tested on a defined dataset, the intention is that it of course be applied to datasets outside of this training set. To help assess the validity of antibody landscape fits when applied in novel situations a number of diagnostic tests were therefore developed.

2.7.1 *Fitting to randomly shuffled data*

A first check of model predictions, although not intended as a test when fitting to every dataset, is to perform cross-validation of the antibody landscapes fits when modelling data where variables and measurements have been randomly assigned. In the case of antibody landscapes the variables are x and y antigenic coordinates and the measurements are of course the HI titres. The idea is that by shuffling the data you are left with a baseline value, the size of which is determined by the amount of general variation of measurement values in the datasets, against which to compare model predictions. For example, a small model prediction error may not be particular impressive when fit to a dataset that anyway has very low variation in measurements and this test would be one way to highlight such situations.

The RMSE of cross-validation predictions when antibody landscapes were fit to data where HI titres were randomly shuffled between all antigens is shown in figure 2.20. As would be expected, it is clear that fits to the original dataset perform much better than when measurements are randomly assigned, confirming that the landscapes truly model underlying relationships between viral antigenic characteristics and serum antibody reactivity. Also shown are the results when landscapes were fit to data where the HI titres had been randomly shuffled between H3N2 strains in the same antigenic cluster. Here the differences between fits to the full dataset are far smaller, indicating that mostly it is differences in HI titre due to antigenic differences between antigenic clusters that are being modelled. However, prediction RMSEs are still larger for the "cluster-shuffled" fit, meaning that the relative positions of antigens within an antigenic cluster are still a predictor of serum reactivity to some extent.

2.7.2 *Predicted vs measured titres*

As another basic test, to check for systematic deviation of model fits from the data, fitted titres can be compared against measured titres. Figure 2.21 shows an example where model fits for all of the landscapes of pre-vaccination samples in the 1997 vaccination trial are plotted against the measured titres. Larger light blue circles show the median measured titre for fitted titres rounded to the nearest whole titre and the light blue lines mark the interquartile range. Reassuringly the medians for each group of fitted titres lie on line of equality, suggesting minimal systematic deviation of the measured data from the model in this case. Overall, there is perhaps

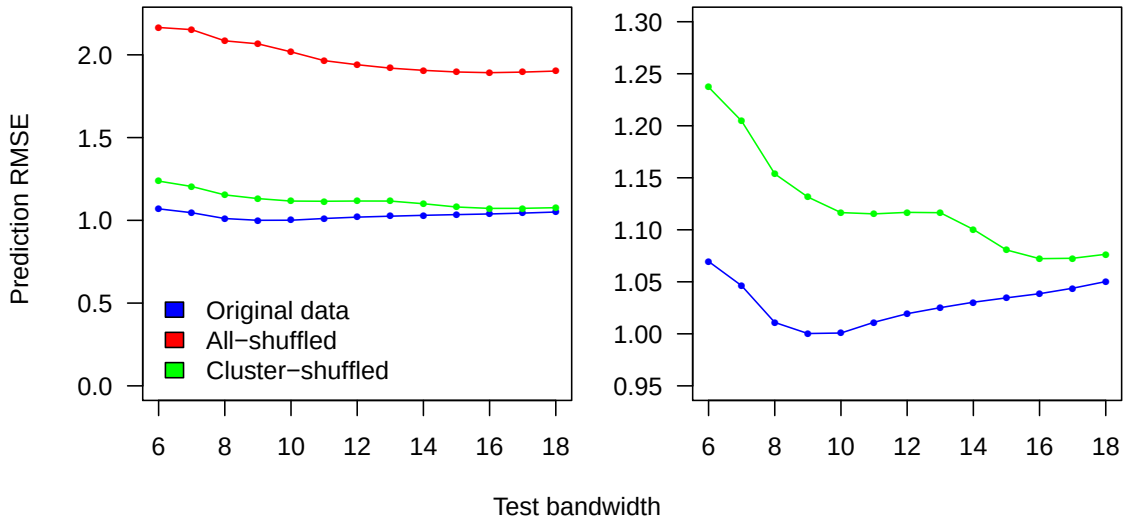


Figure 2.20: Cross-validation results compared to shuffled data. The cross-validation results from the maximum-likelihood fit shown in figure 2.12 are reproduced alongside the results when applying the maximum-likelihood fitting approach to fit data where HI titres were either randomly associated with a viruses in the base-map (red), or randomly associated with viruses in the same antigenic cluster (green). The left panel shows the results on a y axis scale stretching to 0, while the right panel shows the same results on a scale chosen to better visualise differences in the results between the original and cluster-shuffled data.

evidence for a slight bias towards negative residuals for higher titres, indicating that in some cases the fits may be over-smoothing peaks in the data, but the effect does not appear to be large.

2.7.3 *Assessing goodness of fit*

Aside from systematic positive or negative deviations from the data, another important metric is how closely the model fits the data in terms of the magnitude of these deviations. A common metric for assessing the goodness of fit is the root-mean-square error (RMSE) as used in section 2.3.5 to assess goodness of landscape predictions under cross-validation and defined in equation 2.7. In the context of landscape fits, too high an RMSE indicates a poor fit to the data while an RMSE that is lower than would be expected from measurement error indicates likely over-fitting of model to data.

Although the RMSE is a simple metric, a difficulty (as before) is how to deal with non-detectable values since it is not possible to calculate an exact residual from fits to such data. As a work-around, here again non-detectable titres were treated as -1 on the log scale while landscape predictions returning a value <-1 were also treated as -1.

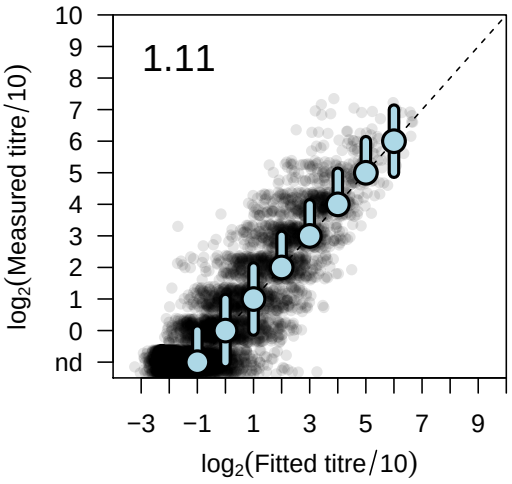


Figure 2.21: Diagnostic plot of predicted vs. measured titres. Predicted vs measured titres for antibody landscapes fit to pre-vaccination serum from the 1997 vaccination trial. Larger light blue circles show the median measured titre for fitted titres rounded to the nearest whole titre and the light blue lines mark the interquartile ranges. Individual data points are shown in transparent grey and since the measured titres are all discrete values a small amount of jitter has been added in the y axis to better visualise the spread of data. The dashed black line shows the line of equality and the number in the top left of the plot gives the RMSE of fits to measurable titres.

The left panel of figure 2.22 shows a quantile-quantile plot of the RMSEs of each landscape fit (again to pre-vaccination titres from the 1997 vaccination trial), against expected quantiles from a normal distribution with a standard deviation and mean equal to the standard deviation and mean of the RMSE distribution. In this case, although the RMSEs appear to be approximately normally distributed, the higher and lower quantiles fall lower than would be expected given a normal distribution. However, it is hard to interpret whether these fits represent outliers however since the RMSE of certain fits will be skewed to a greater or lesser degree depending upon the number of non-detectable titres measured. One approach would be to exclude residuals associated with non-detectable values but this removes a lot of the data that was used to fit the landscapes.

A solution to the issue of non-detectable titres is to numerically approximate the expected error for each landscape fit taking the effect of non-detectable titres into account and to compare this to the actual measured error. In a simple scenario with continuous measurements, non-censored data and normally distributed errors, the expected RMSE of each fit should on average be equal to the standard deviation of the error function. Although the situation with HI titres is more complicated, the expected error when fitting the antibody landscapes was already estimated in section 2.5.2 and found to be well described by a normal distribution with a standard deviation of 1, and it is still possible to compare the RMSE of fits against that expected from the error function alone.

The procedure used here is first to take the landscape fits for each antigen and simulate 10000 measurements normally distributed around the fitted value. These simulated measurements are then rounded and the RMSE calculated by treating both predic-

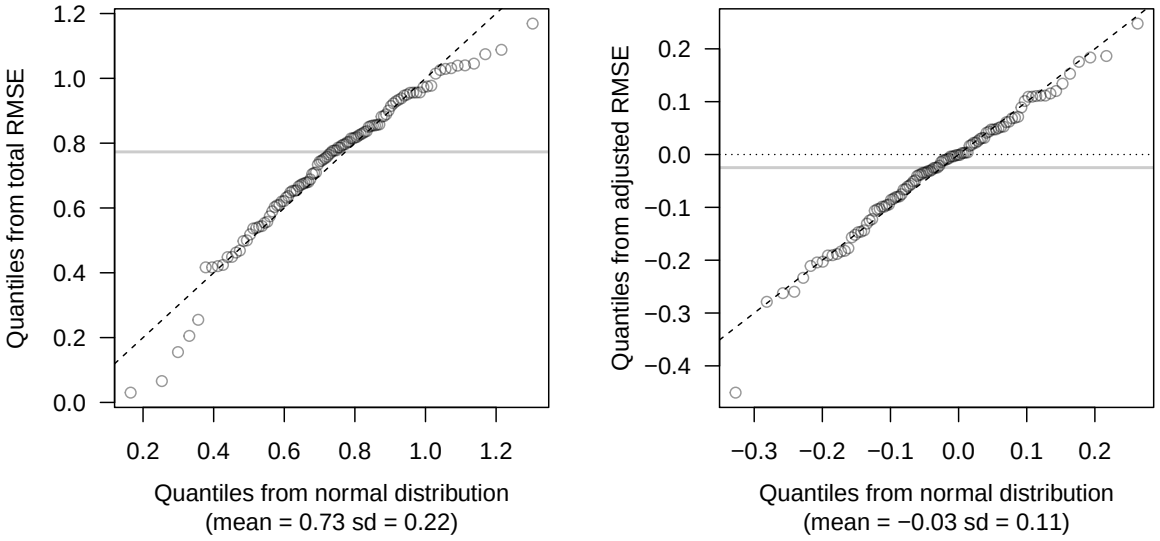


Figure 2.22: Quantile-quantile plots to assess distribution of antibody landscape fit RMSEs. The left panel shows the distribution of the RMSE of antibody landscapes fit to each of the pre-vaccination sera from the 1997 pre-vaccination trial when measured titres that are <10 and fitted titres that are <-1 are both treated as -1 on the log scale. The right panel shows instead the distribution of the RMSEs compared to the expected average RMSE for each fit taking into account the number of non-detectable titres. In each case individual quantiles (each representing an individual antibody landscape fit) are shown as grey circles and the median RMSE or adjusted RMSE is marked by the solid horizontal grey line. The dashed-black line shows the line of equality - if points are normally distributed you would expect them to lie on or close to this line.

tions and measurements less than -1 as equal to -1, exactly as when calculating the RMSE of the measured data. The RMSE of all this simulated data is then compared to the RMSE against the RMSE of the actual measured data to give an "adjusted" RMSE, defined simply as:

$$\text{Adjusted RMSE} = \text{Measured RMSE} - \text{Simulated RMSE} \tag{2.16}$$

The right panel of figure 2.22 shows the quantile-quantile plot now of the adjusted RMSEs of the same landscape fits, similarly to before, compared against expected quantiles from a normal distribution with a standard deviation and mean equal to the standard deviation and mean of the adjusted RMSE distribution. Most of the outliers from the normal distribution seen in the left panel now appear to simply follow the expected normal distribution when converted to adjusted RMSE with only one landscape still returning an unexpectedly low value, perhaps indicative of overfitting due to a paucity of detectable titres. An additional advantage of the adjusted RMSE metric is that when landscapes are fit well (neither over nor under-fitting) the average adjusted RMSE should be near to 0. In this case the median adjusted RMSE (marked

on the plot as a horizontal grey line) is indeed close to zero indicating a plausible fit to the data.

2.7.4 *Virus-specific error*

Error on an individual virus level was also examined and the question of whether there were systematic deviations from the model fit per virus. Examined across the fits to pre-vaccination samples from the 1997 vaccination trial in figure 2.23, there are clear patterns in terms of the error of the fit to each virus, with some viruses being associated with measurements that are consistently higher or lower than the predicted value.

Figure 2.24 compares the average residual for each antigen to fits to pre-vaccination sera from the 1997 vaccine trial against fits to those from the 1998 vaccine trial. The fact that there is a clear relationship between virus residuals across the two trial serum sets emphasises that many of these virus errors are systematic rather than random variation. The cause of these difference between viruses, however, is hard to ascertain. One explanation could be antigenic differences not captured by the antigenic map, either due to error in placement or patterns of cross-reactivity not captured in 2 dimensions. Equally, there could be differences in viral reactivity to serum that are not related to antigenic differences between haemagglutinin molecules, for example haemagglutinin density on the viral surface. Another well-recognised effect is that certain viral strains will be better or worse at agglutinating red blood cells used in the HI assay, as demonstrated recently in the extreme by a complete loss of the ability to agglutinate red blood cells through HA binding for some recent H3N2 viral strains [Mögling et al., 2017]. Interestingly, such virus-specific deviations have similarly been noted when using HI results in other models [Lessler et al., 2012].

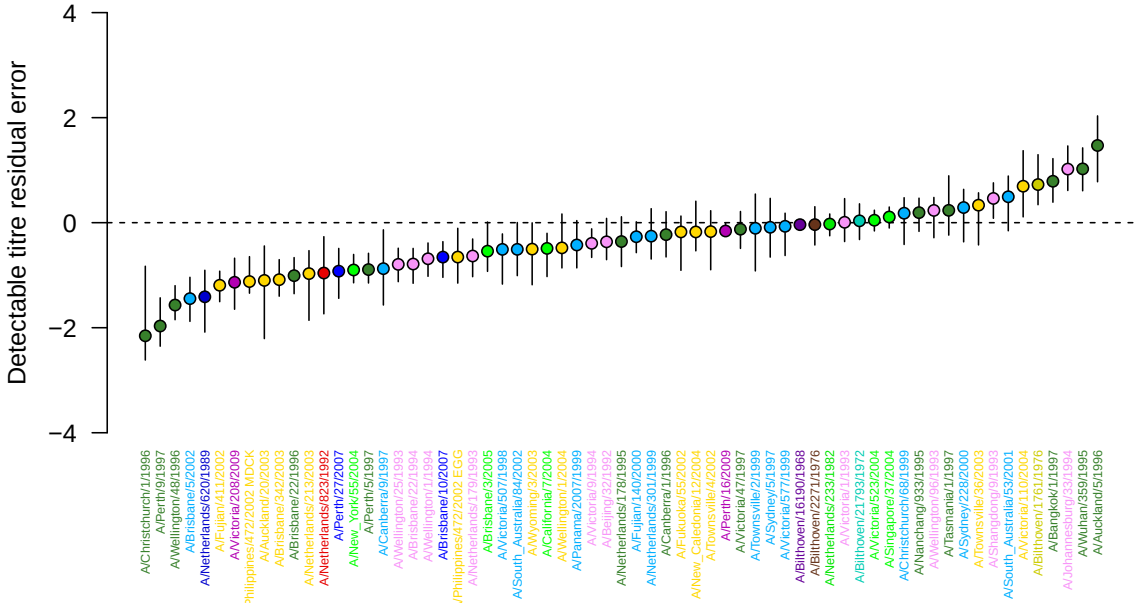


Figure 2.23: Distribution of residual error to antibody landscape fits split by virus. Viruses were ordered by the median residual error of detectable titres compared against antibody landscape fits to pre-vaccination sera from the 1997 vaccine trial. Circles colour coded by the virus antigenic cluster represent these median values while the vertical black lines mark the inter-quartile range. The dashed horizontal black line marks a residual of 0, which would be expected if detectable titres were on average no higher or lower than the value predicted by the landscape fits.

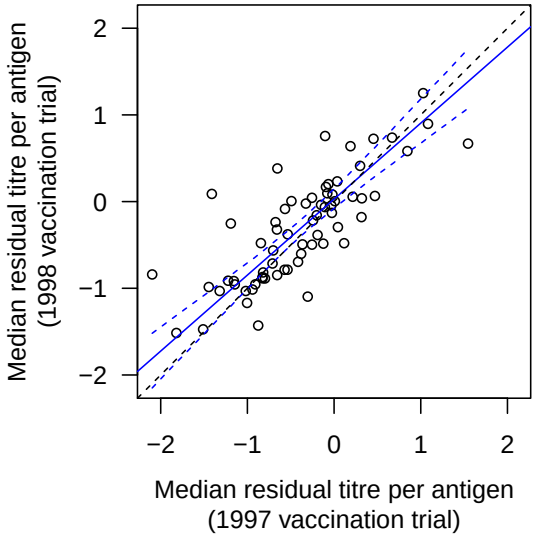


Figure 2.24: Residuals per antigen compared between the 1997 and 1998 vaccine trials. The median residual of measured titres compared to antibody landscape fits was calculated per virus and compared between fits to samples from the 1997 and 1998 vaccination trials. Each circle represents an individual virus and the dashed black line shows the line of equality. The solid blue line estimates the underlying relationship between the x and y variables through a principal component analysis. The dashed blue lines show 95% confidence intervals for this relationship based on 1000 bootstrap repeats.

2.8 Conclusions

In this chapter we have spent some time developing and examining optimal ways to model and visualise serum antibody reactivity. I hope that already the potential of the antibody landscapes technique is becoming clear but I would like to conclude with some particular observations on the strengths and caveats of this novel approach to serological analysis.

2.8.1 *Why use an antibody landscapes approach?*

When seeking to understand differences in serological data, undoubtedly one of the main advantages of an approach such as that presented here is that, since antibody reactivity is inferred from measurements against multiple different viruses, results are not so dependent upon the particular set of strains chosen for an analysis. The variability in the behaviour of different viral strains in the HI assay due to potentially confounding factors such as those discussed in section 2.7.4 is a problem that often seems under-appreciated in the literature and I would argue that the conclusions drawn from many studies using titrations against only one strain could look otherwise had a different strain with different reactivity in the HI assay been chosen instead. Modelling based on a combination of titres from different strains also allows antibody reactivity to be inferred more accurately than from single titrations alone. The fact that HI titres are limited to discrete values normally drawn from two-fold dilutions means that this can be particularly advantageous, and we shall see in later chapters how the ability to more sensitively assess serological change can aid in the interpretation of serological responses on the individual level.

The benefits of an analysis less dependent on particular strain choice and more accurate estimates from pooling data are of course not necessarily uniquely achieved with an antibody-landscapes type approach. For example, simply comparing average HI titre per antigenic cluster would already leverage both these advantages to some extent. Indeed, for many purposes this would already be sufficient to solve some of the major problems associated with single-strain analysis. However, antigenic clustering is not always distinct and antigenic changes appear also to occur within clusters to a greater or lesser degree. Similarly, antigenic differences between clusters can be highly variable, for example the antigenic change between the Wuhan 1995 and Sydney 1997 clusters seems to have been particularly large. When comparing titres or responses between different strains and clusters it is therefore often still useful to have a measurement of the relative antigenic differences between the strains used. Antibody landscapes provides a methodology to quantitatively allow for such variation while still combining information from the multiple different titres in a statistically powerful way.

Finally, a very particular advantage of antibody landscapes over simple averaging approaches based on rough antigenic groupings is that by expressing antibody reactivity

in terms of antigenic space it allows for a quantification of immunological features such as response breadth or the extent to which antibody cross-reactivity appears to diminish with antigenic distance. These quantifications can then be used to draw conclusions about the immunological processes behind them and make comparisons when such features are of interest. Excitingly, it opens the door to development of a predictive understanding of how antibody responses depend upon factors such as the distribution of prior immunity in antigenic space and the different potential antigenic characteristics of a new strain encountered. As we shall see throughout this thesis, such a predictive understanding of immune responses, supported by insight gained from detailed analysis and quantification of actual human infection and vaccination responses, makes possible the goal of vaccine strain selection that does not simply match antigenic characteristics to circulating strains, but optimises them in terms of quantitative antigenic characteristics to generate the best expected immune response in a given population or individual.

2.8.2 *Caveats and limitations*

Although a quantitative expression of immunity throughout antigenic space has powerful advantages, it should in no way be expected that an approach such as that taken here represents a perfect description of the full complexity of serological data. Antigenic cartography is at best an approximation of the true antigenic relationship between pathogen strains and the HI assay upon which the antibody landscapes in this thesis are based is a fairly course measurement of serological immunity. Two dimensions appears to be a reasonable description of antigenic relationships between strains of the H3N2 influenza subtype primarily studied in this thesis but the true dimensionality of the antigenic space they inhabit is undoubtedly much higher and more complex. Even constrained to two dimensions, the position of different viral strains is variable and prone to error.

For all the reasons mentioned, tests such as cross-validation of antibody fits and calculation of the residual errors are an important validation of the technique. The fact that missing values could be predicted from antigenic coordinates significantly better than by chance shows that the antigenic map used here, although not expected to be precise, does indeed provide useful information when inferring serological immunity in human data. Such validations should always be performed when using a new measurement of antigenic differences such as a different antigenic map.

Fitting human serological data based on antigenic relationships defined by the H3N2 antigenic map also provides a validation of the antigenic cartography approach itself. It supports the assumption that antigenic relationships as defined by ferrets also reflect differences in human serological reactivity and also gives us an indication of the scales at which antigenic differences are most relevant. The fact that shuffling of strains within clusters still diminishes the predictive value of the antigenic coor-

dinates, indicates that even on this scale the antigenic differences measured are at least somewhat meaningfully related to the underlying variation in human antibody reactivity. However, it is clear that the between-cluster antigenic differences are by far the most significant predictor of differences in antibody reactivity. On the very largest antigenic scales, where single exposure ferret sera cannot be used to directly measure antigenic distance because HI titres fall to non-detectable levels, the analyses later in this thesis provide an interesting separate indication that these inferred distances may still be approximately correct. The fact that the strong "backboost" phenomenon described in chapter 3 appears to diminish linearly with antigenic distance even across the full range of the H3N2 antigenic map is one such example, going against claims that antigenic distances that go beyond those directly measured may be misleading (see Recker et al., 2007).

A caveat to validation through prediction of missing measurements is that since antigenic cartography and the antibody landscapes in this thesis are both built upon HI titres, it may be that to some extent the predictive value (perhaps indeed that seen at the within cluster level) comes simply from describing how the strains tend to behave in the HI assay itself. If, unrelated to their actual antigenic characteristics, a set of strains tended to return low titres in the HI assay, this would have the effect of pushing them to the edges of clusters on the antigenic map, a region where antibody landscapes may indeed correctly fit a region of low reactivity. Potential problems such as this could only be addressed by the use of a different assay for the human sera in an attempt to decouple assay-specific differences that would affect both the map and antibody landscape. This would bring its own difficulties but would be a powerful test of the extent to which antigenic cartography really represents underlying antigenic differences between strains on different scales.

More generally, there are some features of the antibody landscapes technique developed here that should be remembered when it is used. For example, the variability caused by sampling of alternative viral strain combinations, leads to the conclusion that while broad patterns apparent in antibody landscapes may be confidently interpreted, one should be more cautious when it comes to inferring much from finer details in an individual landscape. In this regard, the construction of confidence intervals helps to give an idea of what should be counted as a significant feature. Additionally, where details are to be interpreted it should be remembered that although the 2D visualisations give an easily comparable summary, re-examining for any patterns in 3-dimensions will ensure that conclusions drawn are not an artefact of the specific path through the landscapes chosen.

A methodological point should also be made with regard to the calculation of confidence intervals for the antibody landscape fits. At any given antigenic point, the confidence intervals as calculated in this chapter provide a good representation of the expected range in which the underlying pattern of serological reactivity in a given

sample lies. However, when comparing antibody landscapes across a large antigenic range, care should be taken in the interpretation. When comparing to a baseline antibody landscape for example, the greater the antigenic distance over which differences are measured, the greater the probability of a region occurring whereby confidence intervals lie outside of the baseline landscape by chance variation. Compensating for this effect is statistically challenging since the probability of such chance differences occurring depends both upon the additional antigenic locations considered and upon how correlated you would expect results to be between them (i.e. are they close in antigenic space to other locations compared). A conservative approach taken in future analyses in this thesis has generally been to test for evidence of significant landscape differences at a limited number of antigenic points, allowing for any multiple testing with a Bonferonni correction. This being said, where larger differences in antibody landscapes are seen that fall well outside confidence bounds, or are seen across a wide antigenic range, significant changes in serological response can likely be safely concluded.

Finally, although an antibody landscapes approach holds advantages, there is still the potential for artefacts appearing due to the smoothing nature of the fit. For these reasons, in the analyses that follow in the next chapters, care has been taken not to over-interpret the patterns observed and to remember the limitations of the methodology. Where important conclusions have been drawn from the data, underlying patterns in the raw titrations have also been demonstrated.

2.8.3 *Areas for further development*

As a concluding note for this chapter, although much work has gone into developing and validating the technique as it stands, there is certainly room for further improvement. Local regression seems to represent a good fit both theoretically and quantitatively to the task of modelling serum reactivity throughout antigenic space, but not all potential methods and parameterisations were exhaustively tested. It is possible that other strategies may better capture variation in the data, for example the slight negative bias seen in the residuals of figure 2.21 may indicate that more flexible fits could better capture local peaks of serum reactivity in antigenic space but they would need to simultaneously guard against over-fitting of random variation in the data. In this regard, the estimate of pairwise titre differences between antigenically highly similar viruses shown in figure 2.10 gives some indication of the theoretical maximum limit of prediction accuracy using only data from the antigenic map as the predictor, namely an RMSE of around 1. Given that the RMSE when using local regression to predict missing titres in the detectable range was around 1.2 (figure 2.13), it is therefore unlikely that other approaches would substantially alter the nature of the fit.

To achieve further improvements beyond adjustments to the fitting approach, either more data would be necessary to model antigenic differences between strains or a more sensitive assay. In the case of more data, genetic variation would be one potential source but gross genetic differences have been shown to be a relatively poor predictor of antigenic characteristics in the case of influenza [Koel et al., 2013; Smith et al., 2004]. This means that without a large training set or an accurate model of how genetic change maps to antigenic change, including such data would likely simply introduce more noise and the possibility for overfitting. It is likely however, that a more accurate representation of the antigenic differences between strains could be achieved than the 2-dimensional antigenic map assumed here and it would be interesting to explore how human serological data and approaches such as antibody landscapes could even be used to better infer antigenic differences between influenza viruses.

For those interested in utilising the antibody landscapes methodology or developing it further themselves, an R package containing the functions required to build and test antibody landscapes is available on request.

3

The Response to Influenza Infection

The Hà Nam province in Vietnam, home of the Hà Nam study cohort.

3.1 Chapter summary

In this chapter the antibody landscapes methodology developed in chapter 2 is used to study antibody reactivity in samples taken annually between 2007 and 2012 from 69 individuals in the Hà Nam household cohort study in Vietnam.

Individual antibody landscapes demonstrate significant heterogeneity but show a remarkable degree of stability from one year to the next. In many cases the individual features were so distinctive that likely sample swaps could be identified. Titres against viruses circulating long before an individual was born are not found, but there are often measurable titres to clusters shortly preceding the earliest strains that were feasibly encountered, giving an indication of the extent of antibody cross-reactivity. Generally, titres against strains circulating at the time of sample collection were lower than those circulating earlier in an individual's lifetime.

Infections with H3N2 were associated with a strikingly broad antibody response, far in excess of the cross-reactivity associated with a first infection. This response often extended to include all viruses back to the very earliest antigenic variants from the Hong Kong 1968 antigenic cluster, but it was characteristically limited by the extent of an individual's prior immunity. PCR-confirmed infections with other subtypes (influenza B, H1N1 and H1N1p) caused no significant change in antibody titres, indicating that boosting of previous titres upon infection is a subtype-specific phenomenon.

Broad initial responses were followed by a period of titre decay during which antibody titres settled to a new consistent level typically sometime within the course of a year. Unlike the antigenically broad short-term boosting, the long-term response post infection was substantially narrower. It did not usually include a sustained increase in titres to very antigenically dissimilar strains that were previously encountered but still spanned multiple antigenic clusters.

3.2 Introduction

The immune response in general contains multiple components that work together in concert to deal with infection but as discussed in chapter 1, the antibody response in particular is regarded as a key component of effective protection [Coudeville et al., 2010a; Fox et al., 2015; Hobson et al., 1972; Jong et al., 2003; Virelizier, 1975; Wikramaratna and Rambaut, 2015]. Identification of common features of the response to influenza infection and how immunity is generated at the individual and population level opens up potential avenues to explore better ways to stimulate effective protection artificially, through vaccination.

In practice however, gaining any kind of understanding of generalisable patterns in the human immune response to influenza infection has proven incredibly challenging. Partly this is because of limited available studies where samples exist both before and after confirmed influenza infection. Of the few that exist, due to the difficulty and expense of active influenza monitoring, studies have typically relied upon comparison of samples taken before and after an influenza season, using seroconversion as a proxy for detecting influenza infection [Miller et al., 2013; Pica et al., 2012], or upon inferences from cross-sectional data [Kucharski et al., 2015; Lessler et al., 2012]. Exact dates of infection are therefore usually unknown unknown while by defining infection based on seroconversion - an aspect of the very response being investigated - samples are necessarily biased when detecting samples that represent actual infections, with both false positives and negatives being prevalent [Cauchemez et al., 2012].

Where human antibody responses to influenza have been investigated, a huge heterogeneity of responses has been noted and few useful patterns identified [Beyer et al., 2004; Koopmans et al., 2012]. This is perhaps unsurprising given the large array of potential variables that differ between any given infection, including the antigenic characteristics of a particular infecting strain, the time since infection, the genetic background, age and infection history of the individual infected. To exacerbate these difficulties, antibody responses are typically measured against only one or a limited number of influenza strains. The precise choice of strains used, their reactivity in the HI assay, and the extent to which they represent a good antigenic match to the actual infecting viruses also act as additional confounding variables that will have significant effects on the magnitude of effects measured despite having no bearing on the actual immune response.

Due to the difficulties of controlling for so many unknown variables in the case of human immune responses to influenza infection, much work has focussed on the studies in experimental animals, typically ferrets or mice [Angelova and Shvartsman, 1982; Li et al., 2013; St Groth and Webster, 1966b; Webster, 1966]. In such cases, the infection histories can be controlled, the timing of samples before and after infection can be standardised and measurements of the immune response to infection made against the exact strains that caused the infection. Despite these advantages, the un-

derstanding that can be gained from studies in such experimental animals can only take us so far. Differences in the immune response between humans and ferrets or mice are not fully understood, while truly recreating the extensive infection history of adult humans is impractical, if not impossible given the logistics of keeping study animals for so long and their more limited lifespan.

In this chapter, I attempt to avoid some of the short-comings of serological analysis through analysis of antibody titres and application of the antibody landscape's methodology to annual serum samples taken between 2007 and 2012 from the Hà Nam cohort study in Vietnam [Horby et al., 2012]. Unusually, in this dataset, longitudinal follow up of study participants over a number of years allows a comparison of how antibody titres in a single individual change over time, controlling for individual variation, while continuous monitoring for influenza-like-illness (ILI) and PCR confirmation of influenza infection allow for accurate information about both the timing and type of influenza infections occurring. Combined with the increased resolution of antibody analysis achievable through antibody landscapes, the aims are to describe in more detail the common features and individual differences measurable in the human antibody response to influenza infection.

3.3 Materials and Methods

3.3.1 *Sample selection*

A subset of 69 individuals were selected from the Ha Nam Household Cohort Study [Horby et al., 2012] for detailed antibody-landscapes-based analysis of serum samples collected between the study years 2007-2012. This subset included all individuals from six households that had a high sampling compliancy and a household size of five or larger (n=36), all individuals that had a PCR-confirmed infection during this time (n=10), a random subset of participants with serological conversion defined as a $\geq$4-fold change in HI titre against a single sentinel strain (n=14) and a random subset of control individuals without reported serological change (n=9). 37 of the 69 individuals had available serum samples from the full 6-year time series and a total of 324 samples were analysed.

None of the 69 subjects had ever received an influenza vaccination. Only subjects that were 5 years old or older were enrolled into the Ha Nam cohort, but otherwise no age restrictions had been imposed. The stem and leaf plot in table 1 gives the subject ages of the 69 individuals at the end of the study period in 2012.

Each subject had been informed of the aims, methods, anticipated benefits and potential risks of the study by a Vietnamese-speaking member of the research team conducting the study in a face-to-face meeting. Participants were provided with age appropriate information sheets in Vietnamese and asked to sign a consent form wit-

0	7889
1	0022223455556779
2	0122235677889
3	22358
4	0223345788999
5	000012222235
6	38
7	13
8	
9	15

Table 1: Ages of the individuals of the Ha Nam cohort in 2012. A stem-and-leaf plot representing the ages of the individuals of the Ha Nam cohort in 2012. A stem-and-leaf plot is similar to a histogram in that the length of the rows indicates the number in each group, and additionally indicates the data within each group. The row 9 | 15, for example, indicates two individuals: one aged 91, the other 95.

nessed by a member of the study team. All participants provided written informed consent.

3.3.2 *HI titrations*

HI measurements were performed by collaborators at the Victorian Infectious Disease Reference Laboratory in Melbourne following the protocol described in appendix B. One of two overlapping sets of H3N2 strains were used as the titration set as shown in figure 3.1. Sufficient serum volume permitting, samples taken in 2007-2011 from the 36 individuals from the six selected households with high sampling compliance and household size of five or larger were titrated with the HI assay against 38 H3N2 viruses (isolated between 1993 and 2011, figure 3.1 upper panel). Samples from the remaining 33 individuals and all samples taken in 2012 were titrated against a more antigenically extensive set of 30 H3N2 viruses (isolated between 1968 and 2011, figure 3.1 lower panel). The full list of strains used in each of these titration sets is given in appendix A.

The maximum dilution measured was 640, if agglutination was still inhibited at this dilution, the entry was listed as having an end-point of $\geq$1280, since 1280 represents the next two-fold dilution that would have been measured under the protocol. When modelling antibody landscapes, and for all other analyses, these $\geq$1280 measurements were treated as simply equal to a titration of 1280 (this was applicable however to only 93 out of 11597 titrations).

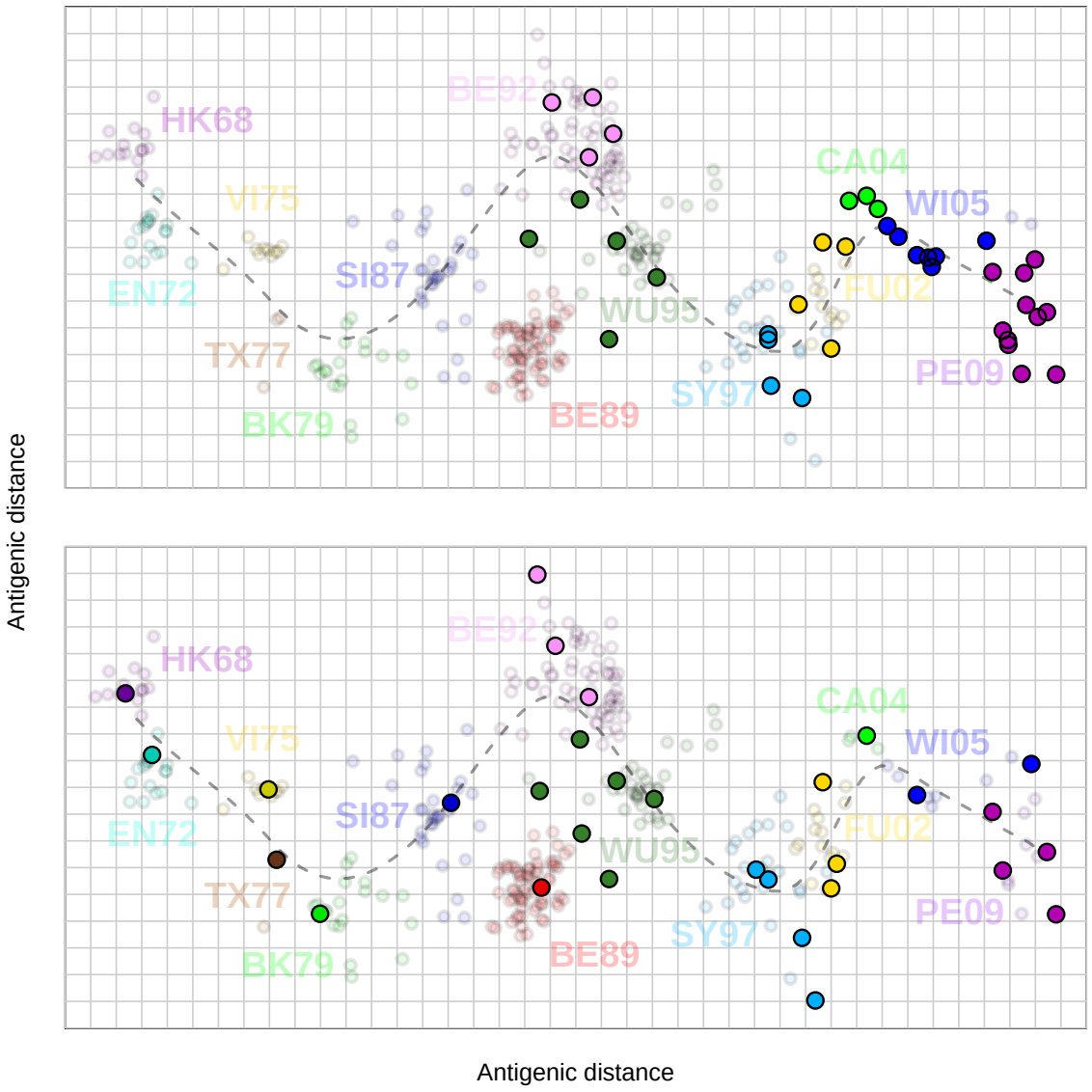


Figure 3.1: Antigenic maps highlighting the H3N2 strains used in the Hà Nam cohort study. The antigenic map and summary path defined in chapter 2 are shown, different H3N2 strains are represented by circles colour-coded by antigenic cluster. The upper panel highlights the 38 H3N2 strains (isolated between 1993 and 2011) that were used as the test set for samples between 2007 and 2011 from subjects in the six selected households with high sampling compliance and household size of five or larger. The lower panel highlights the more antigenically extensive range of 30 H3N2 viruses (isolated between 1968 and 2011) that were used as the test set for samples from the remaining 33 individuals and all of the 2012 samples. Full lists of the strains used are given in appendix A.

Year	Sampling date	Cluster(s) circulating	Inclusion test strain
2007	December 2007	WI05	-
2008	December 2008	WI05	A/Vietnam/EL140/2008
2009	June 2009	WI05, PE09	A/Vietnam/TX265/2009
2010	April 2010	PE09	A/Vietnam/TX265/2009
2011	July 2011	PE09	A/Vietnam/TX265/2009
2012	May 2012	PE09	A/Vietnam/TX265/2009

Table 2: Representative antigenic strains used for seroconversion testing in the Hà Nam cohort. Cohort year, sampling date, circulating antigenic clusters and representative antigenic strains used for seroconversion testing (WI05: Wisconsin 2005; PE09: Perth 2009).

3.3.3 *Monitoring of influenza infections*

Individuals in the cohort were closely monitored for signs of influenza like illness (ILI), participants were encouraged to report any onset of symptoms so that nose and throat swabs could be taken (typically taken the next day, or within two days if reported over a weekend) and stored for later PCR analysis and subtyping [Horby et al., 2012]. Active monitoring of participants and symptoms also took place in the form of weekly visits and a questionnaire detailing the extent and duration of any symptoms present in household members during the preceding week.

Based on antigenic analysis of circulating strains, during the cohort study period two H3N2 antigenic clusters circulated: first WI05, and later PE09, as detailed in table 2. In the following data analyses, seroconversion was defined as a $\geq$4-fold increase in HI titre against A/Vietnam/015/EL134/2008 in years with WI05 circulation, and against A/Perth/16/2009 for PE09 cluster circulation - viruses that reflected the antigenic nature of strains circulating in the previous year and against which all samples were titrated.

3.3.4 *Modelling the antibody landscapes*

Antibody landscapes were constructed using the maximum-likelihood approach as described in chapter 1, using a fitting bandwidth of 10 antigenic units assuming a minimum titre for non-detectable measurements of -4. Diagnostic plots for the antibody landscape fits are shown in appendix C.4.

In the 2-dimensional summary landscape plot, to indicate the antigenic nature of potentially infecting viruses, all viruses belonging to the cluster that circulated in the year prior to each sample's collection are highlighted. For example, in samples collected in 2008 this includes all WI05-type viruses, in 2010 it includes all PE09-

type viruses, while in 2009 it includes both WI05 and PE09 type viruses, since both clusters circulated that year (see table 2).

3.4 Results

3.4.1 *Individual patterns of influenza immunity*

Individual antibody landscape fits are shown in full in appendix C.1.2 and looking through these visualisations, several patterns became evident. Firstly and most strikingly, although there was substantial heterogeneity among the antibody landscapes of different individuals, each individual's landscape shape is often consistent from one year to the next and has distinctive individual features.

Figure 3.2 reproduces the antibody landscapes from the 9 individuals chosen as controls (from whom no sign of overt H3 influenza infection was reported between first and last serum sample collection). Here it can be seen that although the landscapes of each individual sample appear very distinct from one another, many are almost identical when comparing antibody landscapes constructed from the sample taken in 2007, to the sample taken 5 years later in 2012. Some differences are evident, as visualised by the red-coloured regions (areas where the landscape increased compared to the previous sample) and beige-coloured regions (areas where the landscape decreased compared to the previous sample). However, it is often only a part of the antibody landscape that has altered with the overall features of the landscape remaining mostly unchanged.

Changes of landscape over time will be discussed in more detail later but generally, in the absence of infection, antibody titres and consequently antibody landscapes remain remarkably consistent over time.

3.4.1.1 Potential sample swaps

The characteristic nature of the antibody landscapes was in many cases so distinct and conserved that it even acted like a fingerprint, making it possible to identify where likely sample swaps had been made and serum samples incorrectly assigned to different household members. Such errors can be difficult, if not impossible, to trace but figure 3.3 gives an example of where such a potential situation was identified.

Note that to aid interpretation and analysis in total 4 likely sample swaps (described in full in appendix C.1.1) were switched back or excluded. In the individual landscape plots, these samples are indicated with an asterisk but this switching back of likely sample swaps was also performed before all other analyses took place.

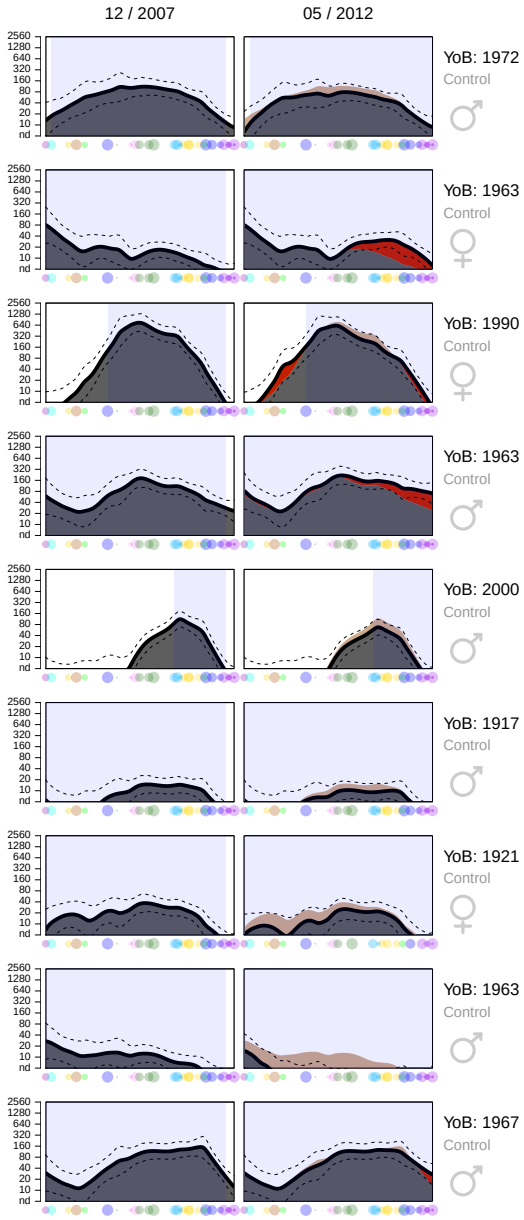


Figure 3.2: Antibody landscapes for all individuals in the Hà Nam cohort chosen as controls. Each row contains 2D summary antibody landscapes corresponding to samples taken from a single individual and the information to the right gives their date of birth, gender and reason for inclusion in the subset of test individuals. Each column contains landscapes corresponding to samples taken in a given year, with the year and month of the sample collections shown above. As described in chapter 2, the black lines represent the landscape height for each position on the antigenic summary path taken through the antigenic clusters (and shown again on the antigenic maps in figure 3.1). Dotted black lines show 95% confidence intervals for the landscape fit. The antibody landscape measured in the first sample taken is shown shaded in grey, in subsequent samples (in this case only the sample from 2012), red shading indicates increases, and beige decreases, compared with the previously sampled year. Additionally shown are blue-shaded rectangles that indicate antigenic clusters that had circulated during an individual's life span until sample collection. As in the previous chapter, dots along the x axis indicate the antigenic positions along the summary path of the subset of viruses used to generate each landscape, scaled negatively by distance from it. It can be noted in this particular figure how the shapes of the antibody landscapes are individually very distinct, yet remarkably consistent over time.

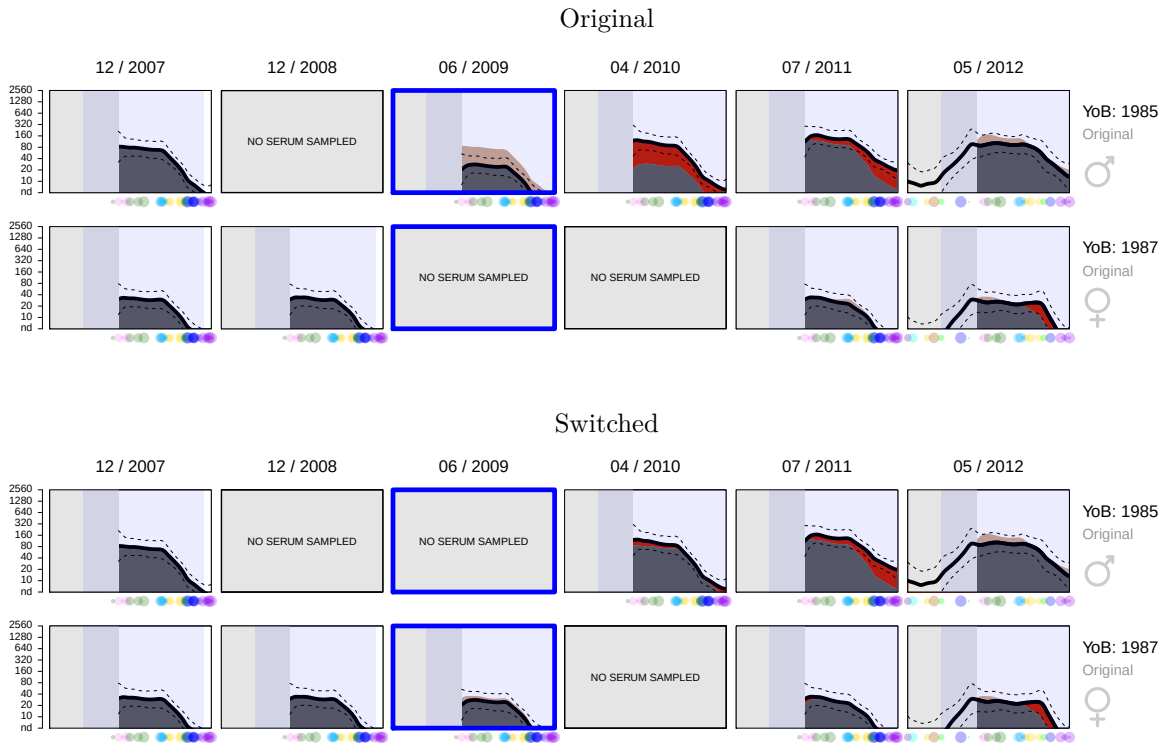


Figure 3.3: A potential sample swap in 2009 for two individuals within the same household. Caption primarily as described for figure 3.2, but additionally in this figure the samples taken in the year of the likely sample swap (2009) are outlined with a blue box. The top plots under "Original" show the sets of landscapes for two individuals as they were originally assigned in the dataset. The bottom plots under "Switched" show the sets of landscapes for the same two individuals after the decision was made to reverse a likely sample swap by switching the samples taken in 2009. Note that in some of these plots (and later ones) the antibody landscapes are cropped on the left side to the antigenic limit of the strains titrated. The blue box still however extends to the appropriate antigenic point, even if this is outside the antigenic region of strains titrated. In the final panel, where a more antigenically extensive range of strains was used for these samples, the full landscape is shown with the black line but only coloured in the antigenic region that overlaps with the coverage of the strains titrated against the previous sample.

3.4.1.2 Titres against circulating and historical strains

As has been historically observed, despite individual differences, in general antibody titres against newly circulating viruses tended to be lower than against strains that had circulated earlier in an individual's lifetime. This is shown in Figure 3.4 where titres have been grouped by either "current" antigenic cluster circulating in the season preceding sample collection, "previous" antigenic clusters not currently circulating but having circulated within an individual's lifetime, and "future" antigenic clusters that were yet to circulate at the time of sample collection.

In general, again corroborating findings from other serological studies, the antibody titres in this cohort tended to be highest against viruses that were isolated when an

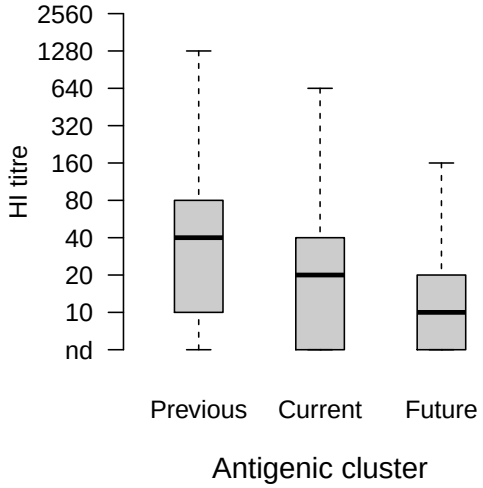


Figure 3.4: Distribution of titres to previous, current and future antigenic clusters. Box plots show the distribution of titre values to different viruses when grouped as described in the main text. Whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box.

individual was approximately 2-6 years old. This is shown in Figure 3.5 where the ages of all individuals in the year of isolation of the virus against which they had the highest HI titre have been plotted as a barplot. Where more than one virus shared the same highest titre, the earlier isolation year was counted. Individuals were excluded if there were no viruses that were isolated prior to the individual's year of birth against which titrations were performed. Where more than one sample had been taken from an individual, the first sample taken was the one used.

Measurable titres were often still present against strains that circulated only before an individual was born, consistent with antibodies produced against a given cluster having some "backwards" cross-reactivity to earlier antigenic variants [Smith et al., 2004]. Figure 3.6 shows number of antigenic clusters relative to the antigenic cluster that circulated at the birth of each individual extending up to a maximum of 2 in 5 of the 69 individuals. Note that to control for isolated titrations that recorded some reactivity for a virus (where signal could simply be measurement error) and to avoid spurious measurements due to any possible sample swaps for a given year, the lowest titre across all samples from an individual was counted as the titre for each virus. For many individuals, the number of clusters prior to birth with cross-reactivity was 0 simply because they were already alive during circulation of the first strains titrated (for example they were born before 1968).

3.4.2 *Antibody landscape responses upon infection*

In total, over the course of 5 years of follow up there were 13 PCR confirmed H3N2 infections, 5 A/H1N1 infections, 3 A/H1N1 (pandemic) infections and 2 influenza B

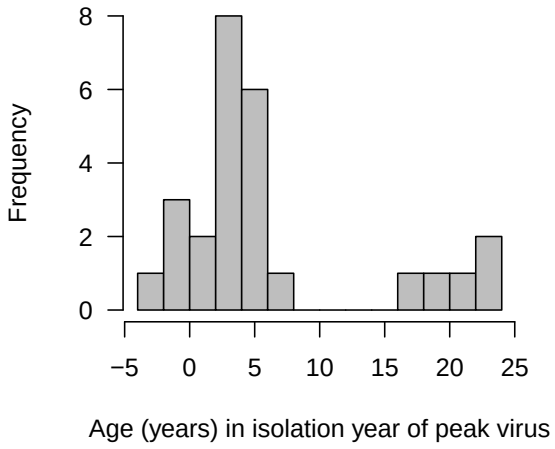


Figure 3.5: Age at the time of isolation of peak response virus. Histogram showing the age of individuals in the Hà Nam cohort in the year of isolation of the influenza virus to which they had the highest HI titre. Where more than one virus shared the same highest titre, the earlier isolation year was counted.

infections recorded in the 69 individuals selected from the cohort, while in 58 cases an ILI was reported and swabs taken but no influenza virus was detected.

3.4.2.1 Antibody response following infection with H3N2

The H3N2 antibody responses following PCR-confirmed infection with an H3N2 strain were strikingly broad, with a large part of the initial antibody response to infection dominated by increases in titres against historical strains.

Figure 3.7 shows the antibody landscapes for 2007-2012, where available, for the individuals in the Hà Nam study for whom a PCR-confirmed H3N2 infection was detected and there were samples available from both the year prior to infection and two years following, such that the change of antibody landscape over time can be fully observed. The sample taken the season after a PCR confirmed infection is indicated with a red box, while the number of days between onset of the ILI and sample collection is indicated also in red. During the cohort study period, two H3N2 antigenic clusters circulated: first WI05, and later PE09, as discussed in section 3.3.3. Viruses from the cluster(s) circulating in season prior to each sample collection are indicated with a red line below each plot.

It can be seen in figure 3.7 that although the circulating viruses highlighted are antigenically some of the most advanced against which titrations were made, infection with H3N2 was associated with a response that commonly extended to even the oldest and most antigenically distant strains titrated from the HK68 and EN72 clusters, although this breadth was variable.

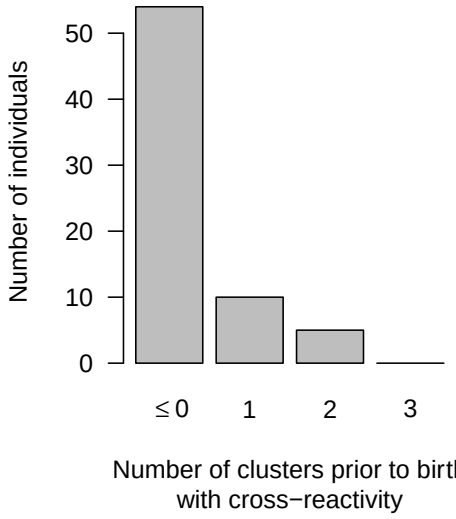


Figure 3.6: Antigenic extent of cross-reactivity to strains circulating before birth. The number of antigenic clusters relative to the antigenic cluster circulating at birth, for which antibody levels were consistently measured for an individual across all samples collected. For many individuals, the number of clusters prior to birth with cross-reactivity was 0 simply because they were already alive during circulation of the first strains titrated (for example they were born before 1968).

Typically, the broad initial response appeared to be followed by a period of decay during which antibody titres stabilised to form an altered antibody landscape that was then subsequently maintained. Comparison of the antibody landscapes of 2007 and 2012 (rightmost panel in figure 3.7) shows that the antigenic region for which increased titres were maintained long-term appears substantially narrower than that of the initial response to infection.

In addition to PCR-confirmed cases of infection with H3N2 influenza, 36 individuals met the criteria defined in section 3.3.3 for seroconversion against a sentinel H3N2 strain. 10 of these individuals had also had a PCR-confirmed influenza infection accounting for their serological response but the remaining 26 showed seroconversion in the absence of a confirmed infection.

Figure 3.8 shows each of the individual's antibody landscape post-seroconversion compared the pre-seroconversion landscape. As with the PCR confirmed H3N2 infections, the same features of a very antigenically broad response are evident in many of the cases of seroconversion. In these cases it seems highly likely that an infection has occurred to cause such dramatic antibody responses but that it was either unreported due to mildness of symptoms or non-compliance, or genuinely asymptomatic. Some seroconverters had reported an ILI in the preceding year but the sample taken was PCR-negative for all influenza subtypes, raising the possibility that a symptomatic H3N2 infection had occurred but that the swab taken was in some way inadequate to detect it.

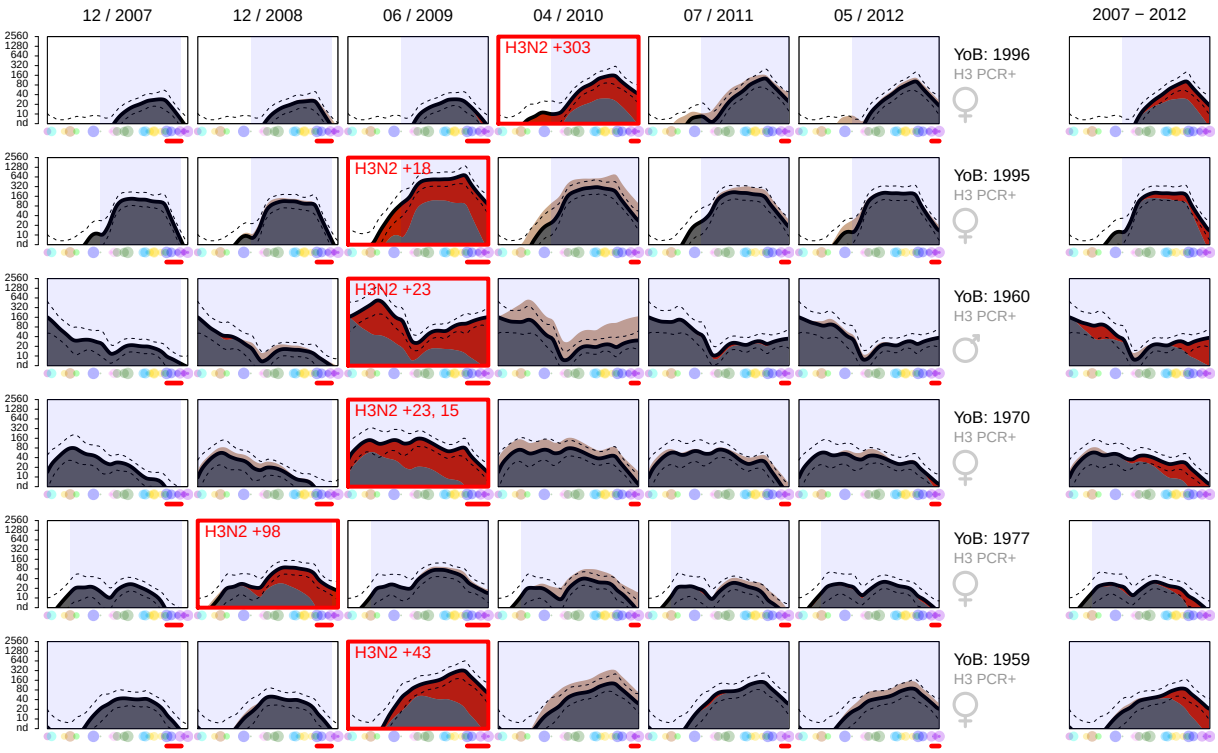


Figure 3.7: Responses to PCR-confirmed H3N2 infection. Antibody landscapes for all individuals in the Hà Nam cohort from whom samples were available at least 1 year prior and 2 years post a PCR-confirmed H3N2 infection. Legend as for figure 3.2 but additionally the first sample taken after a PCR-confirmed H3N2 influenza virus infection is marked with a red box - the red number gives the number of days prior to collection of that serum sample that the PCR-positive swab was taken. Red lines below the x axis also show the antigenic extent of viruses thought to be circulating in the season before sample collection. The extra panels on the right compare the overall change in antibody landscapes from the samples taken in 2007 to those taken in 2012.

In other cases shown in figure 3.8, although individuals seroconverted against the sentinel strain, the serological response when assessed more extensively is ambiguous. The most extreme example is subject number 4 whereby the overall picture is actually one of a decay in antibody titres rather than an increase. In this case it seems highly likely that random variation due to HI measurement error accounted for the seroconversion rather than a genuine serological response to infection. For other individuals the overall antibody response is more muted but the post-seroconversion landscape is still significantly higher than the pre-seroconversion landscape at the antigenic location of the likely infecting strains indicating that despite the smaller increase in titres compared to other individuals, a significant serological response has still occurred.

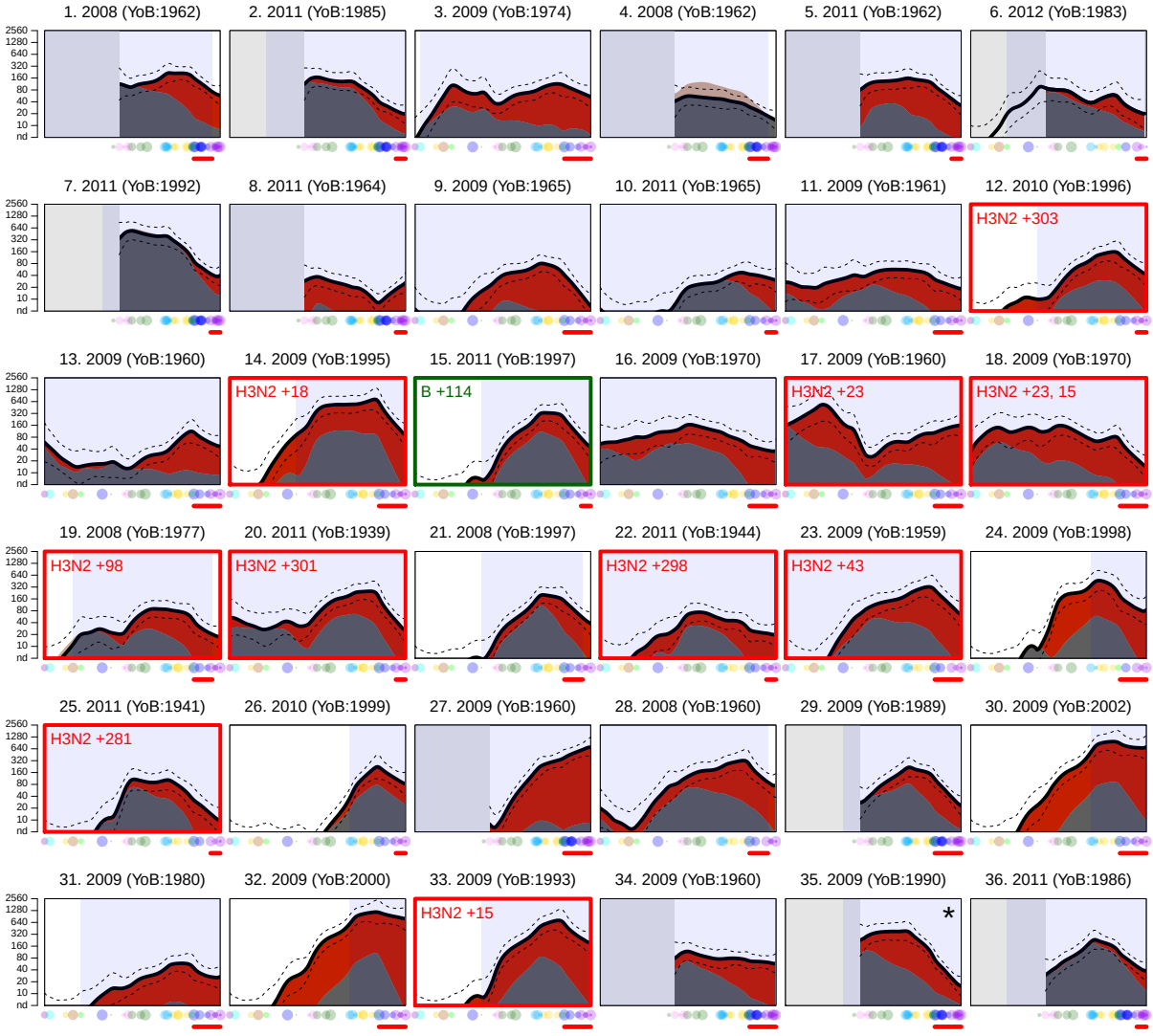


Figure 3.8: Responses following seroconversion. Antibody landscapes are shown for all samples meeting the criterion for seroconversion in the Hà Nam cohort (see section 3.3.3). Unlike previous figures, each panel represents the response in a different individual, showing 36 in total. A reference number, the date of sample collection, and the year of birth of the corresponding individual are given above each plot. The black line represents the 2-dimensional summary landscape fit for the sample taken post-seroconversion in the year given above each panel. The red shading indicates increases, beige decreases, compared to the sample taken prior to seroconversion. As with figure 3.7 samples that were also associated with a PCR-confirmed H3N2 infection are also highlighted with red boxes and show the number of days since the PCR-positive swab was taken and red lines below the x axis also show the antigenic extent of viruses thought to be circulating in the season before sample collection. One sample was taken following a PCR-confirmed influenza B infection and is highlighted in green. As with previous figures, the antigenic positions of titrated strains when projected onto the summary path are shown by circles along the x axis, coloured by antigenic cluster and scaled by distance from the path.

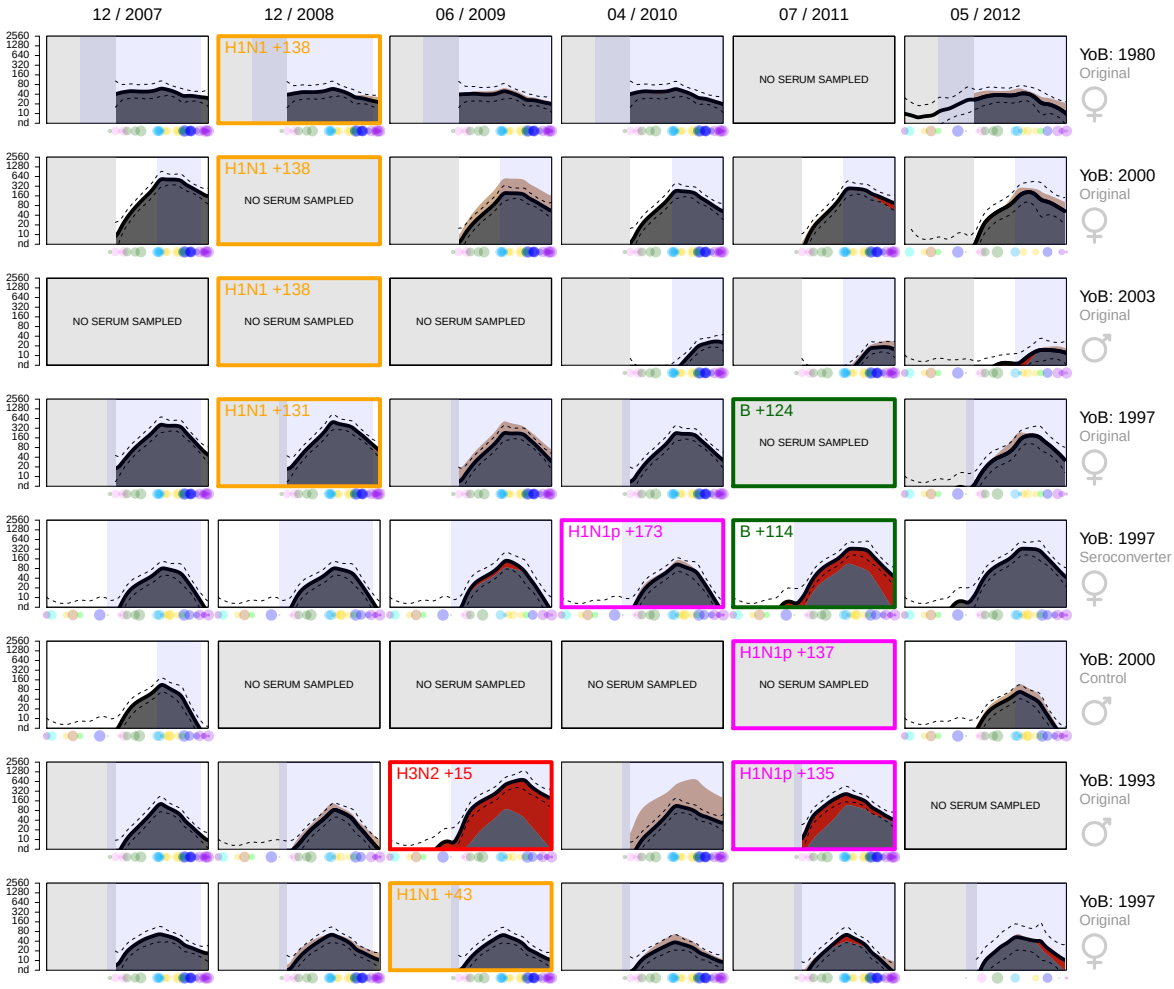


Figure 3.9: Responses to infections with other subtypes. Antibody landscapes for all individuals with a PCR-confirmed influenza A/H1N1(pdm), A/H1N1 or B virus infection. Legend as for figure 3.2 but, as with figure 3.7, the first sample after a PCR-confirmed influenza virus infection is marked with a coloured box and the number of days post collection of the PCR-positive swab shown. This time with different colours denote the different subtypes: red = H3N2, orange = H1N1, pink = H1N1p, green = influenza B.

3.4.2.2 Antibody response following infection with other subtypes

Figure 3.9 shows the series of antibody landscapes from 2007-2012 for the 8 individuals who had PCR-confirmed infection against A/H1N1, A/H1N1 post 2009 pandemic (denoted H1N1pdm) or influenza B. As for the H3N2 infections, the first sample taken after the infection is highlighted and the number of days post onset of symptoms are again shown.

In general from figure 3.9 it can be seen that most landscapes show no evidence of a response in terms of increasing H3N2 antibody titres following infection with another subtype. The two exceptions come following a B infection in the 5th individual shown

and an H1N1p infection in individual 7. Since the other examples following a B or H1N1p infection show no evidence of an associated antibody response against H3N2 it seems rather than a heterosubtypic response, the landscape changes in these two cases may instead have been due to an additional H3N2 infection that went undetected.

3.4.3 *Characterising the antibody response*

Although cellular data is not available, attempts were made to characterise the nature of the sometimes extensively broad antibody responses witnessed in response to the H3N2 infections, comparing them with expected nature of a primary response and how they related to prior immunity in the individuals.

3.4.3.1 Comparison with primary immune responses

Some cross-reactivity in terms of antibody responses to influenza exposure is expected and well recognised following a first influenza exposure due to the antigenic similarity of different influenza viruses. To put the extensive breadth of antibody response witnessed following many of the H3N2 infections in the Hà Nam cohort into context it is therefore useful to compare it to the extent of cross-reactivity that may be expected in such a primary immune response.

Figure 3.10 shows antibody landscapes constructed from 7 ferret antisera raised against clusters from the WI05 and PE09 antigenic clusters - the types of strain circulating in the Hà Nam cohort during the study period. The responses show cross-reactivity and measurable antibody responses against not only the infecting virus but additionally for viruses from antigenically similar clusters such as FU02 and CA04. However, this cross-reactivity is clearly on a different scale to many of the responses measured in the Hà Nam cohort. While detectable cross-reactivity in these 7 ferrets extended to between 3 and 5 antigenic clusters before the cluster of the infecting strain, this number is eclipsed by that of the human responses, which extend up to the full range of 14 antigenic clusters in some individuals.

To give a separate indication of the extent of cross-reactivity expected in a primary human immune response, we can also refer back to figure 3.6. Here, the maximum detectable cross-reactivity in the samples measured extended up to only 2 antigenic clusters. Ferret titres were measured at around the expected peak of the antibody response perhaps accounting for the additional cross-reactivity witnessed in these antisera. However, both ferret and the surrogate human estimates set the breadth of response witnessed in the Hà Nam cohort aside from the typical cross-reactivity generated as part of the antibody responses to a primary influenza exposure.

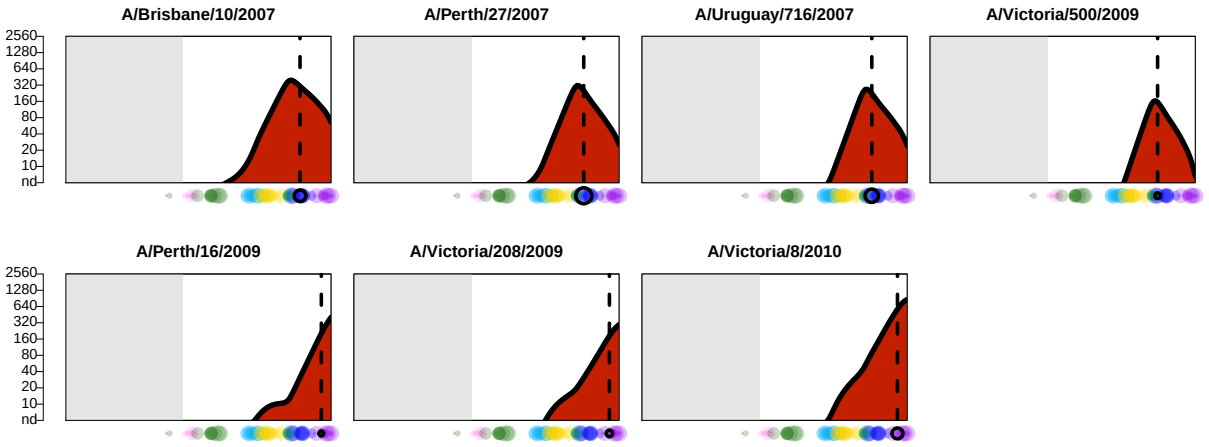


Figure 3.10: First infection ferret responses. Antibody landscapes representing primary ferret responses to exposure to viruses from the Perth 2009 and Wisconsin 2005 antigenic clusters. The solid black line represents the post-inoculation landscape and the landscape is shaded in red to represent the fact that there was no prior immunity before infection. The dashed vertical black line represents the location of the strain the ferret was inoculated with, in relation to the summary path described in chapter 2 and shown in figure 3.1. As with previous figures, the antigenic positions of titrated strains when projected onto the summary path are shown by circles along the x axis, coloured by antigenic cluster and scaled by distance from the path.

3.4.3.2 Response in relation to pre-existing antibody reactivity

Apparent from both figure 3.7 and figure 3.8 is that although post-infection responses were often broad, they were limited by the extent of the pre-exposure antibody landscape and did not tend to extend far beyond the antigenic region against which there was already some detectable levels of immunity. This is more formally examined in figure 3.11, which shows the proportion of individuals with a detectable response, and the proportion of individuals with detectable pre-exposure antibody levels, along the antigenic summary path. Detectable titres were defined as a landscape value >-1, and a detectable response as an increase in landscape value >0.5. The proportions decline similarly against older antigenic clusters further suggesting that the extent of antibody response is indeed related to the extent of detectable pre-infection antibody reactivity. Notably, the same pattern is not observed for the more recent antigenic clusters (e.g. WI05, PE09), perhaps because increases in antibody reactivity to viruses in this antigenic region are partly a result of the generation of novel antibodies directly against the infecting virus, rather than due to stimulation of prior immunity.

Significantly, although the antibody response following an H3N2 infection (as defined by either PCR or seroconversion) was antigenically very broad and often included a substantial increase in titres against older viruses, responses in terms of the fold-change were still greatest on average against strains from the likely infecting WI05

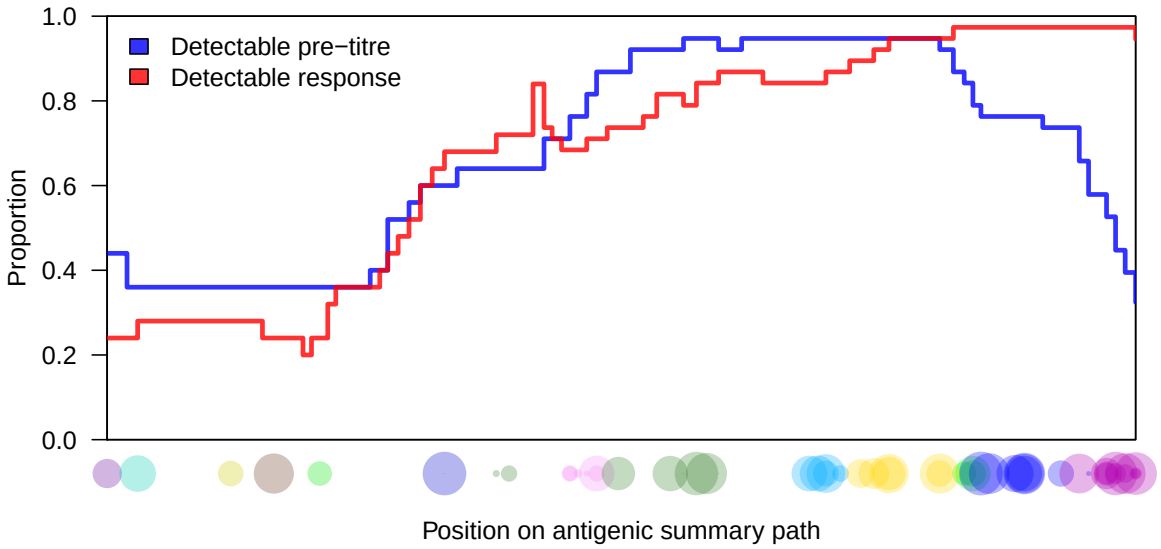


Figure 3.11: The relationship of antibody responses to prior immunity. The proportion of individuals from the Hà Nam cohort with detectable titres prior to either PCR-confirmed infection with H3N2 or seroconversion (blue line) is compared to the proportion of these individuals that show detectable increases in antibody landscapes (red line).

or PE09 clusters. This can be seen in figure 3.12 which shows the average increase in antibody landscape on the $\log_2$ scale along the length of the summary path in the sample directly following an H3N2 infection. The responses of individuals with PCR-confirmed infection (dashed lines) and those without PCR-confirmation, only seroconverting (dotted lines) are both included and follow a similar pattern, with the highest average responses against later antigenic clusters.

Also shown in figure 3.12 are the average responses maintained longer term when comparing the sample taken a season after the H3N2 infection to the one taken before the H3N2 infection (where such samples were available). As noted in figure 3.7, there is a clear overall decay of the initial post-infection response. In many individuals this decrease reflects a return to pre-infection titres against older viruses over time, while the response to more contemporary strains is maintained long-term. In a few individuals however, increased titres even to the antigenically older regions are somewhat retained - for this reason, the average long-term responses approach but do not reach 0 in these earlier regions.

3.4.3.3 Short vs long-term antibody responses

As already noted, the initial very broad response to H3N2 infection was followed by a period over which antibody titres appeared to decay before reaching new stable levels. Since samples were taken only annually for any given individual, estimating

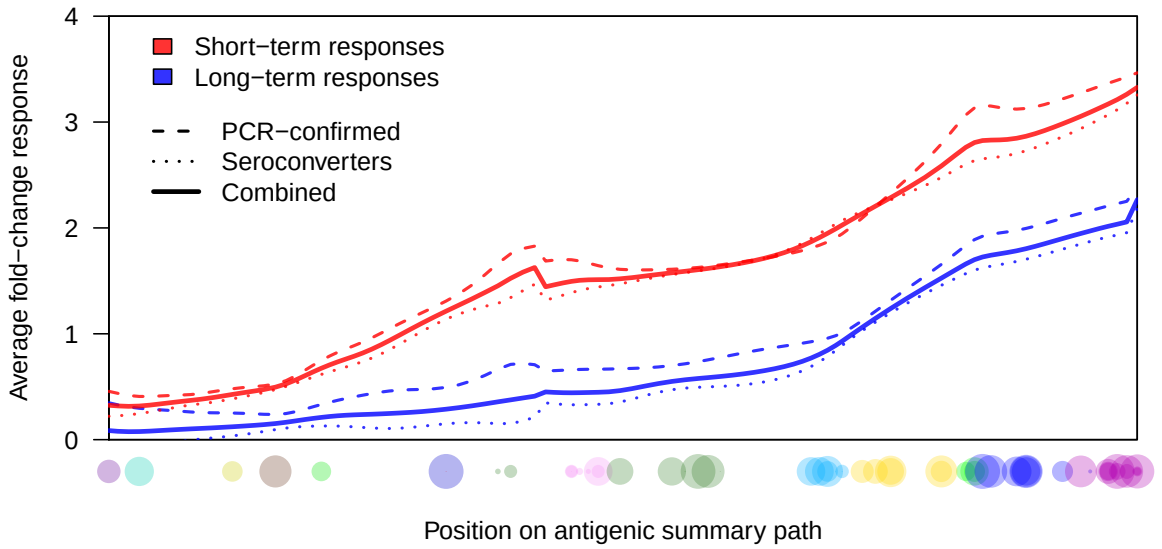


Figure 3.12: Antigenic breadth of short and long-term responses. For each position along the summary path, the mean increase in titre was calculated based on a subset of individuals from the Hà Nam cohort that had a detectable titre at that point in the pre-infection landscape (i.e. pre-infection landscape titre was 10 or higher). This was done based on PCR-confirmed infected individuals (dashed lines), seroconverters without PCR-confirmation (dotted lines), and the combination of these two subsets (solid lines). The red lines compare the sample before and immediately after infection, while the blue lines indicate the titre increase between the last available sample (often from 2012) and the antibody titres in the year prior to infection (provided this was $\geq$1 year following the infection).

the dynamics of this process is challenging. However, since the exact date of the onset of ILI symptoms is known for each of the individuals for whom there was a PCR-confirmed infection, it is possible to examine how time since infection appears to influence the pattern of response seen in the antibody landscapes.

Figure 3.13 shows the antibody landscape responses for the 12 subjects with a PCR-confirmed H3N2 infection arranged by time since onset of symptoms and how these responses later go on to decay in subsequent samples. It can be seen that the large and broad responses present in the samples from the first 5 individuals (taken 14-42 days after infection) clearly go on to further decay before settling to form a new landscape. The picture from the remaining individuals where samples were taken only 100 or more days after onset of symptoms is more mixed. The response in many cases already appears to be a little smaller in magnitude and less broad, often showing no signs of decay in further samples perhaps because the decay process has already taken place at the time of the first post-infection sample. Confusingly, the responses in the 6th and 7th individuals show no signs of decay between the 1st and 2nd samples (between 100/108 to 273/281 days post infection) but then appear later to go on to decay between the 2nd and 3rd samples (between 273/281 to 577/585 days post infection). In general however it appears true to say that landscape changes

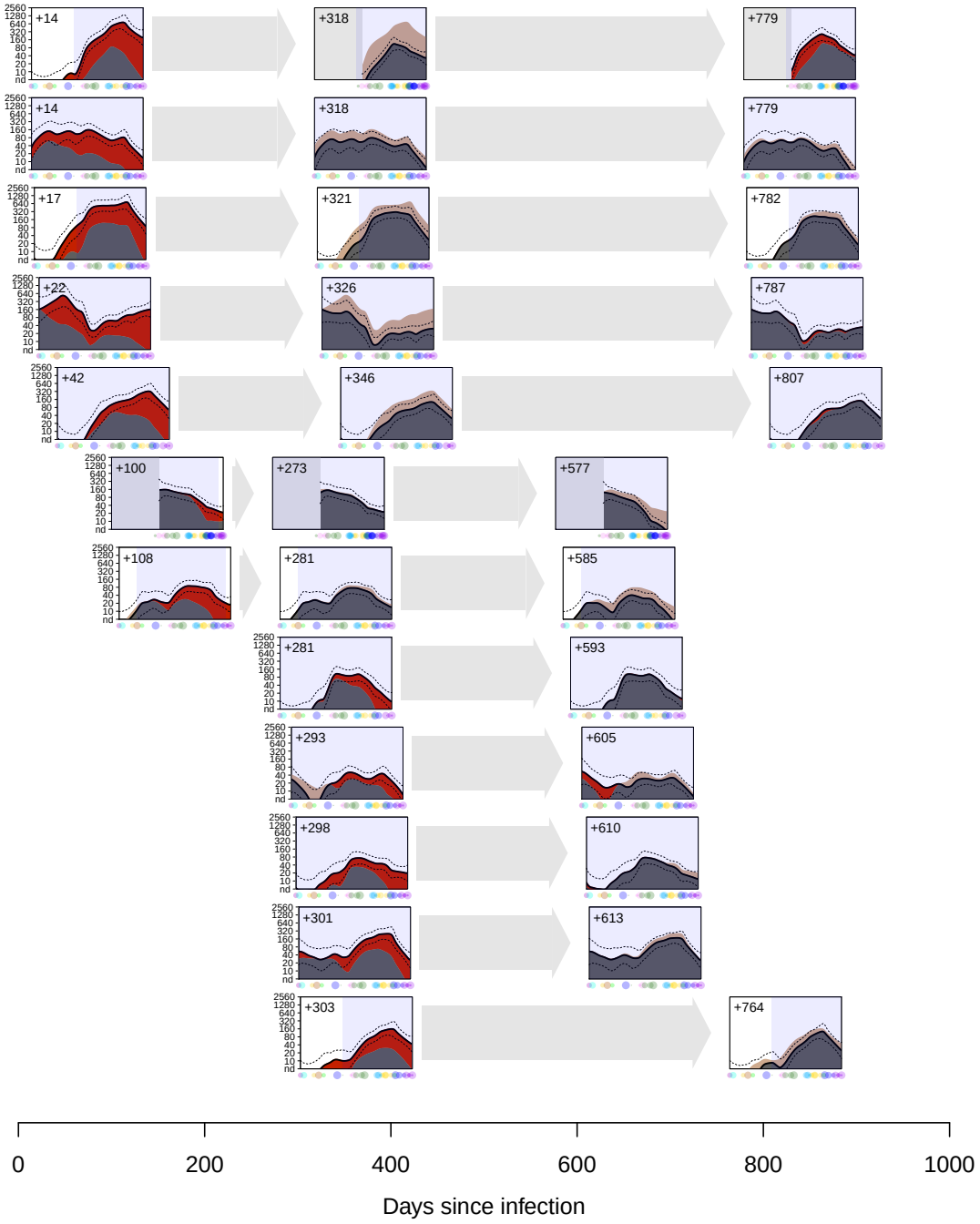


Figure 3.13: Antibody landscape responses to infection over time. Antibody landscapes are shown based on the time since the date of the last swab PCR-positive for H3N2. The left hand side of the box is aligned with the corresponding time since infection on the x axis and the exact number of days shown in the top left of each plot. Each subsequent box compares the next sample taken from that individual to the previous one in the same row, to evaluate decay of antibody responses over time since the first post-infection landscape. The individual landscapes are otherwise shown as described in figure 3.2.

that were present at least 300 days post symptom onset were typically maintained for the remainder of the available samples - a timescale in line with findings from other studies [Horsfall and Rickard, 1941].

3.4.4 *Seroepidemiological comparisons*

3.4.4.1 Estimating rates of seroconversion and asymptomatic infection

As demonstrated above in section 3.4.2.1, figure 3.8, meeting the criteria for seroconversion of a 4-fold or greater increase in titre against a single sentinel strain did not always correspond to antibody responses indicative of a true influenza infection when titre changes were assessed against a wider range of strains. Indeed, assessing seroconversion as fold-change against a single strain is well recognised as an imperfect predictor of actual influenza exposure and serological response [Cauchemez et al., 2012]. The minimum detectable difference of a 2-fold increase in titre when assessing responses against a single strain may already be too coarse to detect smaller serological responses while measurement error in the HI assay itself has the potential to cause both false positive and false negative results. Being based on titrations against many more strains, antibody landscapes therefore have the potential to provide a more robust assessment of seroconversion in response to influenza infection.

One approach to detect significant evidence of seroconversion is to look at the fold-change antibody landscape response at an antigenic point of interest from one year to the next and determine whether the lower bound of the one-sided 95% confidence interval lies above a fold-change of 0. In such circumstances it would still be possible to miss smaller antibody responses but it would be more sensitive than a single four-fold change while making it unlikely that antibody responses that pass this criteria were due simply to HI measurement error and random variation. Following this method, fold-change landscape responses were calculated for each individual from one sample to the next, as described in section 2.5.3. To avoid multiple testing, only the fold-change at the centroid of the antigenic cluster that dominated circulation between each sample pair tested was considered (Wisconsin 2005 for years 2007-2009 and Perth 2009 for years 2010-2012). In total, 45 samples from the cohort meet this criteria for seroconversion compared with 36 samples when using a four-fold or greater increase against a single sentinel strain. This would imply that using a four-fold change against a single strain to define seroconversion did indeed underestimated the actual number of infections that occurred in the cohort. This being said, it is of course not possible to independently confirm whether the extra 35 responses detected as significant using the antibody landscapes approach correspond to a true antibody response against an influenza exposure.

Figure 3.14 uses a Venn diagram to summarise the results when classifying likely influenza infections in the Hà Nam cohort based on the three different criteria, namely seroconversion against a single strain, seroconversion using the antibody landscapes

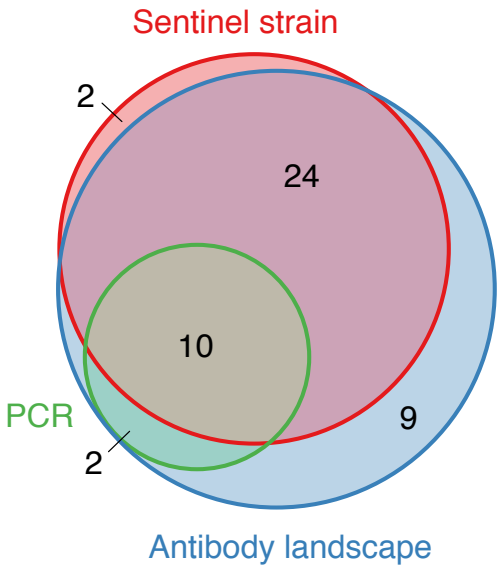


Figure 3.14: A comparison of different criteria for detecting influenza infection. Venn diagram showing the number of landscapes meeting the different criteria for an influenza infection as defined by either seroconversion against a sentinel strain (n=36), PCR-confirmation (n=12), or a significant change in the antibody landscape (n=45). Numbers in the diagram show the number of samples falling into each overlapping region, for example 10 samples met all three criteria. The areas of each region are scaled approximately according the number of samples included in them.

approach or a swab PCR positive for H3N2. It can be seen that while seroconversion against a single sentinel strain detected a serological response in 10 from a total of 12 samples following PCR confirmed H3N2 infections, the antibody landscapes method detected a significant response in all of them. This is a small piece of evidence that the antibody landscapes approach used here provides a more sensitive tool for detecting influenza infections, but without some other form of validation it is impossible to say whether this comes at the price of the detection of extra false positive serological responses.

3.4.4.2 Pre-season antibody titres in relation to infection

Figure 3.15 compares how different measurements of pre-season antibody titre relate to individuals who either went on to have no detectable influenza infection (controls), seroconversion against a sentinel strain or a symptomatic PCR-confirmed infection. In the left-most panel the distribution of pre-titres against the sentinel strain for the upcoming season are compared, in the middle panel the average pre-season titre was taken from all strains in the antigenic cluster of the upcoming season (assuming non-detectable titres to be -1 on the log scale), finally in the right panel the height of the pre-season antibody landscape at the antigenic position of the centroid of all strains in the upcoming cluster was used. Each approach shows the same overall pattern, namely that those going on to have a PCR-confirmed infection in the upcoming season or seroconversion tended to have lower antibody titres in their pre-season sample than those who did not show any evidence for H3N2 infection in the following year.

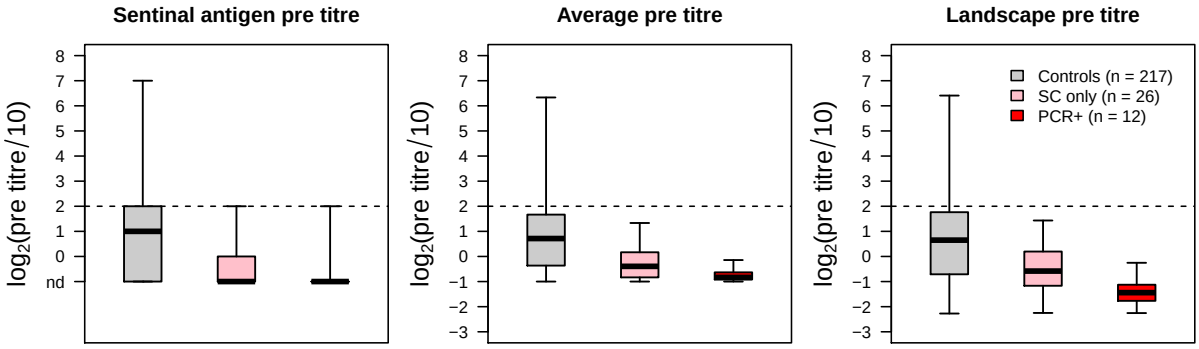


Figure 3.15: Distribution of pre-season titres for infected individuals and non-infected controls. Box-plots show the distribution of titres taken in the year prior to either a symptomatic PCR-confirmed H3N2 infection (red), seroconversion against a sentinel strain without associated PCR-confirmation (pink), or controls who meet neither of the previous two criteria. Left panel: pre-season titres as defined by the HI titre against the sentinel antigen (see section 3.3.3). Middle panel: pre-season titres as defined by the average HI titre against strains from the cluster circulating in the next season. Right panel: pre-season titres as defined by the height of the landscape fit at the centroid location of the strains circulating in the next season.

An HI titre of 40 is often associated with a 50% reduction in infection risk [Hannoun, 2004] and this is additionally marked on the plots as a dotted line. Interestingly, no individual with a titre of greater than 40 appeared to go on to be infected in the next season, supporting the idea that this is on the cusp of a protective threshold. When pre-season immunity was measured using an antibody landscapes approach, this threshold appeared to be even lower, with no participants going on to have a symptomatic PCR-confirmed infection if their antibody landscapes pre-titre at the upcoming cluster centroid was greater than around 10.

The average pre-titre and "protective" threshold was a little higher for those subjects going on to seroconvert without PCR-confirmation of illness, which may have been due to intermediate levels of antibody reactivity causing absent or reduced symptoms when subjects were exposed, but not sterile immunity and an entire absence of antibody response. Controls - being those samples that were not associated with evidence of infection in the following season - had pre-titres that spanned a wide-range. A likely interpretation here is that individuals with higher pre-titres enjoyed protection against influenza infection in the next season, while individuals with lower pre-titres may have been vulnerable but lucky enough to avoid exposure.

3.4.4.3 Generation of a protection landscape

Building on the ideas of using antibody landscapes as a more sensitive predictor of antibody-mediated immunity and associated infection risk, it is theoretically possible to extend the antibody landscapes technique to creation of a "protection landscape".

As an example, figure 3.16 shows the antibody landscape from one individual in the Hà Nam cohort converted into a protection landscape based on the quantitative relationship between HI titre and clinical protection against influenza estimated by Coudeville et al. [2010a]. Using a meta-analytical approach, they estimate the functional form of relationship between HI titre and protection (the HI protection curve) as follows:

$$\text{Probability of protection} = 1 - \frac{1}{1 + e^{\beta(log(T_j) - \alpha}} \tag{3.1}$$

where T_j represents the HI titre and α and β are parameters associated with location and steepness of the HI curve that they estimate to be 2.844 and 1.299 respectively. Although it is not possible to validate this approach, the resulting landscape in figure 3.16c gives an indication of the regions of antigenic space against which the individual may still be vulnerable.

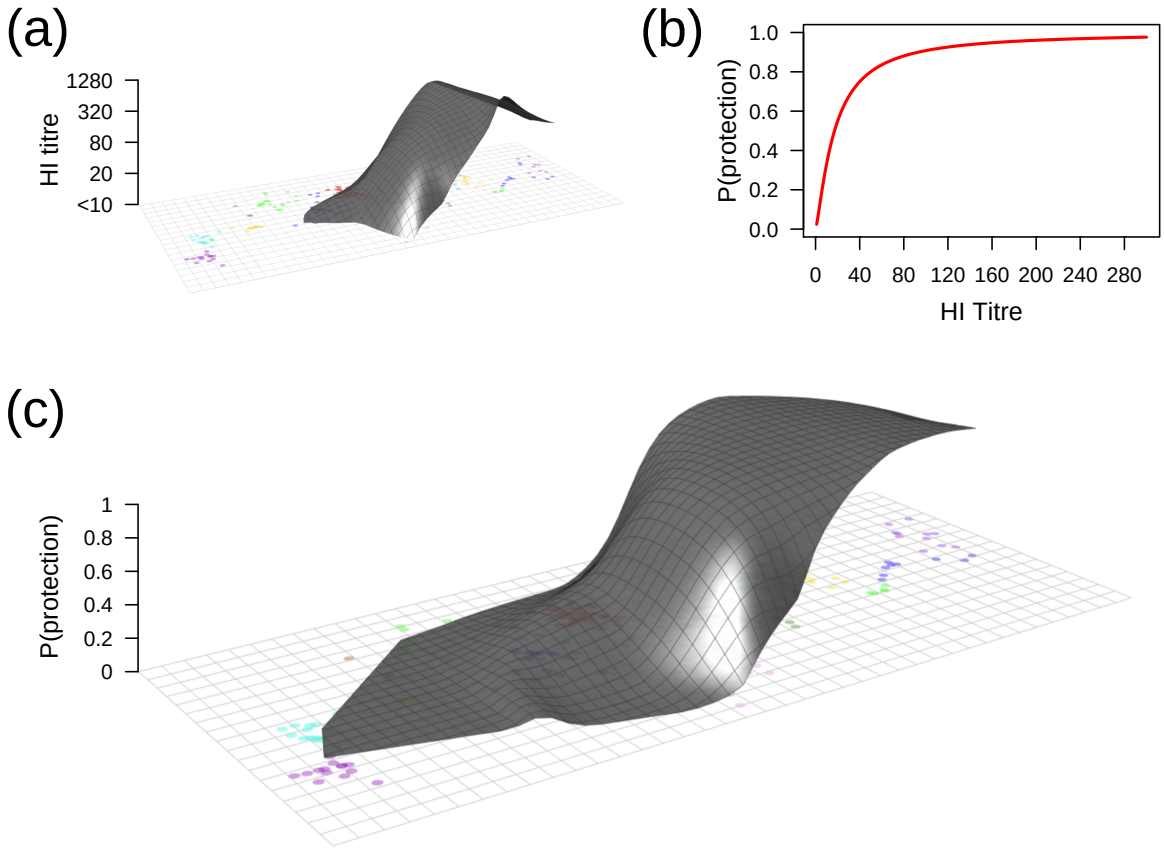


Figure 3.16: Constructing a protection landscape.(a) An example antibody landscape can be converted via **(b)** the relationship between HI titre and clinical protection against influenza defined by Coudeville et al. [2010a] to **(c)** a protection landscape, indicating the probability of protection.

3.5 Discussion

The results presented in this chapter show the value of examining the antibody response against a wide range of antigenically different strains, rather than simply examining responses to a limited number of viruses or a single sentinel strain. Whereas the coarse nature and measurement variation make interpretation of a limited set of individual titrations difficult, when performed over a range of viruses and combined to visualise immunological profiles, patterns of responses that may be otherwise obscured become clearer.

3.5.1 *Interpretation of the "backboost"*

Although recall of prior immunity on exposure to a novel influenza strain is well reported and indeed one of key observations leading to the formulation of the concept of "original antigenic sin" (see chapter 1), the extreme antigenic breadth and strength of some of the antibody responses witnessed here which frequently spanned the full antigenic range of H3N2 viruses is a novel and unexpected finding. To put it in context, there have been 13 antigenic cluster transitions from the first Hong Kong 1968 type viruses until the Perth 2009 antigenic cluster, spanning an antigenic distance that corresponds to a 24-fold dilution of antiserum in the HI assay. These antigenic changes have necessitated over 20 vaccine strain updates, and are the result of changes in 69 of the 346 amino acid positions in the HA1 domain of the haemagglutinin gene between Hong Kong 1968 and the Perth 2009 vaccine strain, including substitutions in all of the seven amino acid positions identified as key antigenic determinants by Koel et al. [2013].

Mechanistically, the explanation for the back-boosting effect described here is unknown and the immunological nature of the response was not directly investigated, however much can be inferred. As noted in the seminal studies performed investigating secondary antibody responses to influenza exposure, the response is most characteristic of a recall of immunological memory rather than an especially broad novel response, being limited by the extent of prior immunity and not universally broad in all individuals [St Groth and Webster, 1966b]. Harder to determine is whether this process of antibody recall is part of a B cell receptor (BCR) binding-dependent or binding-independent mechanism. In terms of BCR binding-independent mechanisms, memory B-cell activation as part of a generalised inflammatory response to infection has indeed previously been described [Bernasconi et al., 2002] and such a phenomenon is well recognised in terms in T-cells [Tough et al., 1996]. This mechanism would explain how infection with a PE09-type strain could induce recall of antibodies originally produced in response to a HK68 type strain, which is ordinarily considered too antigenically distinct to be cross-reactive. On the other hand, a binding-independent mechanism of memory stimulation would also imply similarly recalled responses when infected with a strain of another subtype and this was typ-

ically not witnessed. Alternatively, a BCR binding-dependent mechanism may also explain broad patterns of antibody recall if memory B cells originally produced in response to infection with a HK68 type strains were re-stimulated by low level binding to a later PE09-type virus. Naive B-cell populations have been reported to require a higher binding affinity to stimulate clonal expansion than existing memory cells [Liu et al., 1995], which could explain why a response to HK68-type strains is not ordinarily part of the first response of the immune system upon a PE09-type virus, but a memory response may be.

3.5.2 *Results with regard to "original antigenic sin"*

According to the original formulation of the concept of original antigenic sin and also to later variants such as antigenic seniority [Francis, 1960; Lessler et al., 2012], the reactivity gained against new viral strains following each infection should come primarily as a result of stimulation of antibodies produced against previous infections that have cross-reactivity with the new strains. An implication of this is that reactivity against previously encountered strains should also increase - as witnessed in the original investigations [Davenport and Hennessy, 1956; Francis et al., 1947; St Groth and Webster, 1966b; Webster, 1966].

In responses measured in the short term, the extensive boosting of prior immunity certainly recapitulates findings from previous studies into antibody recall but in the longer-term samples measured a year or later post infection, it can be seen that this boosting to older titres is not maintained, and the response that remains tends to be considerably more focussed around the infecting strain. Since previous studies investigating antibody responses in terms of original antigenic sin have focussed typically on samples taken between 2 and 4 weeks post influenza exposure, we can now see how this locates the conclusions drawn firmly in the timescale of the shorter term dynamic antibody response phase [Davenport and Hennessy, 1956; Francis et al., 1947; St Groth and Webster, 1966a; St Groth and Webster, 1966b; Virelizier et al., 1974; Webster, 1966].

It is interesting to consider what these studies would have concluded had they focussed instead upon these longer term responses a year or so post infection. As discussed in chapter 1, Jenson et al., examining samples taken from the same ferrets as those studied by Webster et al. but taken at 21 days post infection rather than 14, already found evidence for a response with a greater primary component, i.e. more antibodies thought to have arisen from the naive rather than memory B cell population [Jensen et al., 1956; Webster, 1966]. This raises the possibility that while shorter term antibody responses are indeed dominated by production of antibody by pre-existing cross-reactive memory B cells, longer term responses may include much more of a novel component and the effect of original antigenic sin type phenomenon in influencing

long-term responses and protection may have been over-estimated due to the generally shorter-term nature of the post-infection samples historically studied.

One of the tenets of the original antigenic sin dogma is that when measured in a population cross-section, antibody titres as measured by HI are generally highest against the first strains encountered because B cells producing antibodies that are cross-reactive against the first infecting strain are further stimulated upon later exposures, thereby maintaining these antibodies at the highest level compared to the strains later encountered [Francis, 1960]. Again, looking at the nature of the short-term antibody responses taken in the same time window as most previous studies, one can easily imagine how this conclusion would be drawn since at this stage, boosting of old responses is a dominant feature. However, since it is the long-term antibody responses and not short-term dynamic processes that shape the antibody repertoire over time, the fact that the long-term responses measured here did not generally sustain an increase in antibody titres to very early strains throws this conclusion into some doubt. It seems a more plausible explanation in terms of the patterns seen here is that titres against the first infecting strain are highest not because of continual boosting throughout life, but rather due to some other process that establishes higher titres earlier on in life. It is likely that initial long-term boosting of the early antibody titres may occur following second or perhaps third infections when the infecting strain is usually not so antigenically distinct but not indefinitely, as has been previously suggested [Lessler et al., 2012; Miller et al., 2013].

Whether in the short-term, antibody responses can be said to favour the first infecting strain is a matter of interpretation and to some degree, semantics. Certainly, in terms of fold-change in antibody titre, the findings here indicate that responses against earlier strains are not favoured. As perhaps might be expected, fold-change responses diminish with antigenic distance from the infecting strain. However, given that we again witness the phenomenon that pre-exposure antibody titres tend to be higher against strains encountered early in life, a consequence can be that despite smaller fold-change responses the highest titres and the highest absolute increase in antibody titre are still against older strains found in many cases. This fits with the interpretation that while antibodies produced as a result of memory B cells formed against earlier exposures may have weaker binding to novel strains and receive less stimulation, they may still dominate the response due to an initial numerical advantage and the exponential nature of most cellular growth kinetics [Tangye et al., 2003].

3.5.3 *Implications for seroepidemiology*

The observation that landscapes could be so distinct between individuals yet in many cases remain almost completely unchanged from one year to the next implies that the technique provides a robust summary of antibody-mediated immunity even in

the presence of measurement variation in the HI assay. As demonstrated, a result is that changes in antibody titre of a magnitude that is well below the sensitivity threshold of a single HI assay can be detected. This sensitivity to changes in titre makes antibody landscapes well suited for assessing seroconversion, while the fact that responses against multiple different antigens are interpolated across antigenic space means that results are not dependent upon the extent to which a given sentinel strain may match the antigenic characteristics of those actually causing infection in a given season - a common shortfall of the traditional approach.

Using antibody landscapes to make serological estimates of the number of influenza infections, the conclusions here mirror that of others in finding that the typical criteria of a 4-fold or greater increase in HI titre against a single sentinel influenza strain underestimates true attack rates [Cauchemez et al., 2012]. Exactly how accurate an antibody landscapes-based approach is when estimating influenza attack rates is hard to answer since the only gold standard for confirming influenza infection is a PCR-positive virus sample from the patient. The fact that all PCR-positive cases showed detectable serological change as measured by antibody landscapes while 2 from 12 failed to meet the criteria of a four-fold change suggests that it provides a more sensitive test but it is difficult to comment on the specificity. In the cases where serological evidence for infection was found in the absence of PCR-confirmation it is still of course possible that infections did indeed occur but were simply asymptomatic or associated with symptoms that were so minor as to go unreported [Carrat et al., 2008; Hsieh et al., 2014; Huang et al., 2011].

Whether antibody landscapes provide a better correlate of protection than a single HI measurement is difficult to assess with the relatively low number of PCR-confirmed infections in this study however the protection landscape produced as an example in section 3.4.4.3 opens potentially interesting avenues for further study. Certainly, the extensive presence of high titres of antibodies cross-reactive against historical strains provides more evidence as to why influenza does not regress but has instead continually evolved in antigenic directions that are consistent with ongoing avoidance of persistent immunological pressure [Hennessy et al., 1955; Smith et al., 2004]. As a concluding thought, if such a protection landscape could be made representing the extent of protection across antigenic space in the population as a whole rather than an individual, this would quantify a major component of viral fitness, potentially helping to identify the most probable future directions of antigenic evolution and the likelihood of persistence of new strains.

4

The Response to Influenza Vaccination

Skyline of Melbourne, home of the Victorian Infectious Diseases Reference Laboratory.

4.1 Chapter summary

In this chapter I investigate the implications that the broad antibody recall seen following infections studied in chapter 3 has for vaccination against influenza, and how it can be improved.

I focus primarily upon retrospective analysis of sera from two annual vaccine re-registration trials, one conducted in 1997 and one in 1998. Between these two years, the H3N2 component of the vaccine was antigenically updated but, significantly, there was little circulation of H3N2 in the study population. As a result, both groups had therefore received antigenically different vaccines, yet there was no significant difference in average pre-vaccination immunity, allowing a fair comparison of responses.

In the 1997 study, individuals (n=102) received a vaccine containing A/Nanchang/933/1995 from the Wuhan 1995 (WU95) antigenic cluster, against which there was substantial prior immunity. In the 1998 study, individuals (n=123) received a vaccine containing A/Sydney/5/1997 from the antigenically advanced Sydney 1997 (SY97) cluster, against which there was substantially less prior immunity, therefore mimicking a pre-emptive vaccine update.

Antibody responses 4 weeks post-vaccination showed similar patterns of response to those seen post-infection with broad boosting of pre-existing titres in many cases, however there was a large amount of individual heterogeneity. As expected, antibody responses against later antigenic clusters were on average larger in the 1998 vaccine group receiving the updated vaccine strain. Surprisingly however, boosting of prior responses was such that this came at no cost to average responses against any of the earlier clusters, even Wuhan 1995 to which the 1997 vaccine strain was a better antigenic match. A second study of the WI05-PE09 cluster transition also demonstrated a similarly strong boosting of previous titres upon vaccination.

These findings indicate that by taking the effect of prior immunity into account, pre-emptive vaccine strain updates could improve overall vaccine efficacy, especially when antigenically similar strains have already circulated for a number of seasons.

4.2 Introduction

Here, the antibody landscapes methodology developed in chapter 1 is applied to study antibody responses against vaccination. Sera collected as part of historical vaccine re-registration trials performed annually in Australia provided an opportunity to examine the antibody response against both historical strains circulating before the time of vaccination but additionally, depending upon the vaccine year, against strains that would circulate in later years but had not yet evolved at the time pre and post-vaccination serum were collected. By comparing consecutive years between which the vaccine strain had been updated, there is also the opportunity to examine how the switch in vaccine strain from one year to the next effects the magnitude and breadth of the antibody response stimulated.

A potential problem with annual vaccination trial comparisons is that a shifting baseline of pre-vaccination immunity in the population from one year to the next can make direct comparison of responses challenging since additional circulation of the same or novel strains in the population generates a different profile of prior influenza exposures in each of the trial groups [Beyer et al., 2004]. It was possible however to identify a trial year in 1998 where there was little circulation of H3N2 viruses from a new antigenic cluster in the trial population before the novel vaccine strain was first tested. Both trial groups in 1998 and the preceding year of 1997 has therefore received antigenically different vaccines yet with little additional H3N2 exposure in between to bias baseline immunity [World Health Organization, 1997].

Significantly however, because there had been little prior exposure to the new antigenic cluster in 1998 before the antigenically updated vaccine strain was given to the trial population, this year also acts as a surrogate study for the effect of pre-emptively updating the vaccine before circulation of new antigenic variants. This is a particularly relevant comparison since in the vaccine strain selection process the vaccine strain must be chosen around 6 months in advance at a time when it may be unclear what strains will dominate the upcoming season [Stöhr et al., 2012]. If antigenically novel variants are isolated in small numbers the dilemma can be whether to update the vaccine strain pre-emptively, with the risk that current variants remain dominant, or to leave the strain unchanged, with the risk that these novel variants take over and the vaccine lags behind. A comparison of responses in these years therefore offers that chance to investigate this question further and the relative benefits and risks of the two strategies.

Although the 1997 and 1998 vaccine trials formed the main focus of this part of the investigation, to explore the generalisability of conclusions drawn from this dataset, additional samples were analysed from another pair of vaccine trials taking place in 2009 and 2010 and spanning a further vaccine strain update from a strain from the Wisconsin 2005 antigenic cluster to one from the Perth 2009 antigenic cluster. Unlike the period between 1997 and 1998 however, there was significant circulation of viruses

from the new PE09 cluster in the intervening period [World Health Organization, 2009].

4.3 Methods

4.3.1 *HI measurements*

HI measurements were performed by collaborators at the Victorian Infectious Disease Reference Laboratory in Melbourne following the protocol described in appendix B. A total of 468 samples were tested, 106 pre and 4-week post vaccination samples from the 1997 vaccine re-registration trial and 128 pre and post vaccination samples from the 1998 vaccine re-registration trial. Where sufficient serum volume was available, samples were tested against 70 influenza H3N2 strains ranging across H3N2 evolution. Where serum volume was limited, titrations were made against as many strains as possible, selected so as to best cover the range of H3N2 antigenic space. For the 2009 and 2010 vaccine trials, pre and 3-week post vaccination sera were selected from a random subset of 80 individuals and titrated against a more limited range of 20 H3N2 strains.

Figure 4.1 shows the H3N2 base map with the antigenic locations of the 70 strains titrated in the 1997 and 1998 studies and the 20 strains titrated in the 2009 and 2010 studies. The full list of strains is given in appendix A.

4.3.2 *Antibody landscapes*

The antibody landscapes methodology was applied as described in chapter 1 to analyse the HI results. The pre and 4-week-post vaccination landscapes for each individual were plotted and compared to examine the magnitude and breadth of the measured antibody responses. As with other landscapes in this thesis a fitting bandwidth of 10 antigenic units was used and the minimum titre for a non-detectable HI measurement assumed to range down to a limit of -4 on the $\log_2$(HI titre/10) scale. Landscape diagnostic plots for the fits are shown in full in appendix C.4. Of note is that the median adjusted RMSEs of the landscape fits for samples from the 2009 and 2010 trials were considerably lower than 0 indicating that with the more limited titration set of 20 antigens per sample, some overfitting of the data has likely occurred in these cases. Greater caution should therefore be taken when interpreting any finer patterns in these fits, although the greater uncertainty is also reflected in the larger confidence intervals for these years.

4.3.3 *Neutralisation assay measurements*

Virus neutralisation assays were additionally performed by collaborators at the Victorian Infectious Disease Reference Laboratory in Melbourne to measure sera with

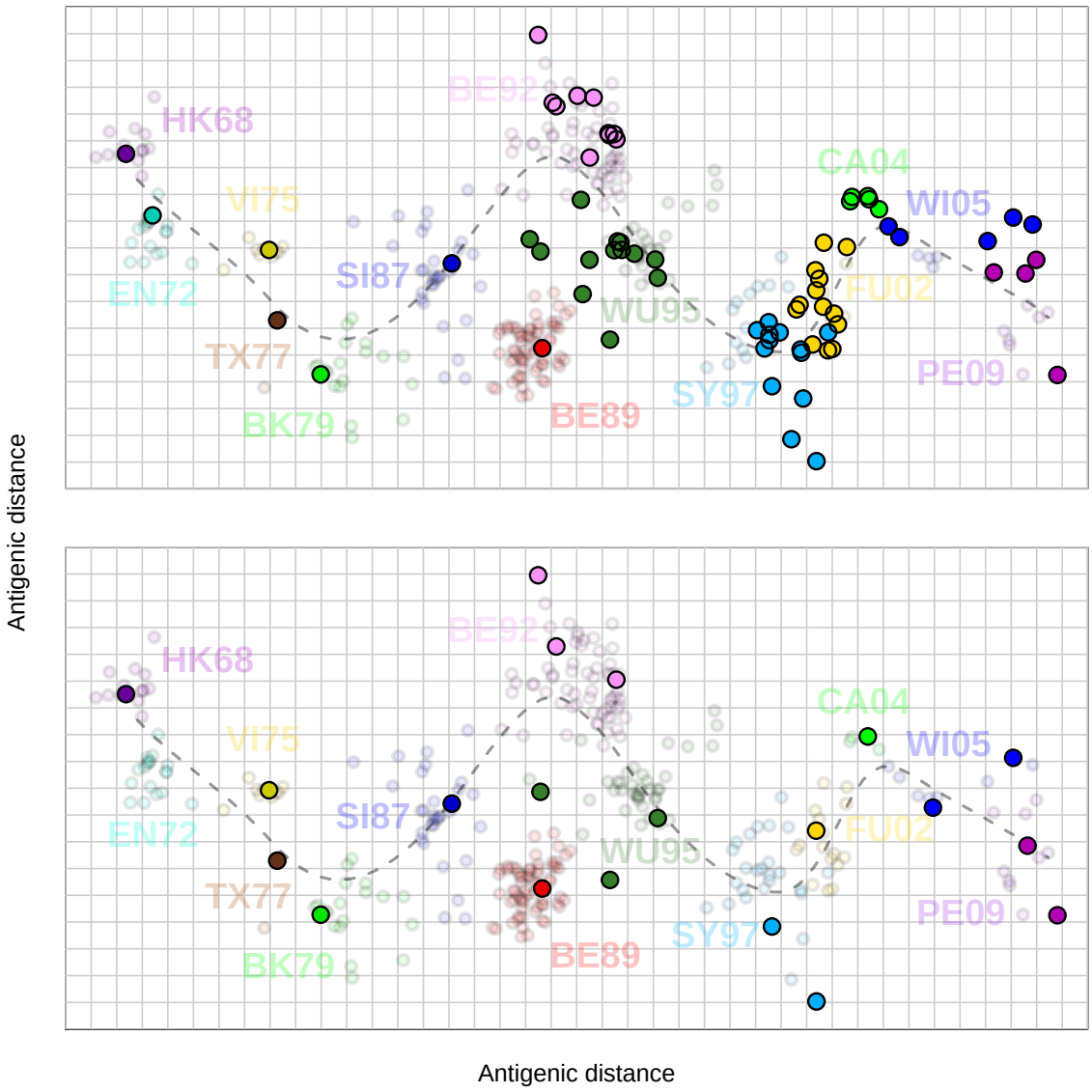


Figure 4.1: Antigenic maps highlighting the H3N2 strains used in the vaccination studies. The antigenic map and summary path defined in chapter 2 are shown, different H3N2 strains are represented by circles colour-coded by antigenic cluster. The upper panel highlights the 70 H3N2 strains that were used as the test set for samples pre and post-vaccination samples from the 1997 and 1998 vaccination trials. The lower panel highlights the less extensive range of 20 H3N2 viruses that were used as the test set for samples from the 2009 and 2010 vaccine trials. Full lists of the strains used are given in appendix A.

sufficient available volume after HI titrations against the H3N2 vaccine strains used in the two trials, A/Nanchang/933/1995 and A/Sydney/5/1997. The assays were performed by heating antisera for 30 minutes at 56°C. Twofold serial dilutions of the antisera starting at a 1:10 dilution were mixed 1:1 with 100 tissue culture infectious dose50 (TCID50) of the virus stocks. After incubation at 35°C for 2h in a 5% CO2 humidified incubator, the antiserum virus mixture was transferred to 96-wells plates containing MDCK cells, which were washed twice with PBS prior to inoculation. Plates were incubated for 2h at 37°C, and inoculum was replaced by 200μL infection medium. After four days end-point dilutions were read by haemagglutination assay. Two repeats were performed with each sample and the lower of the two titres taken as the value for that serum.

4.4 Results

4.4.1 *Individual responses to vaccination*

The complete series of landscapes constructed from all pre and post-vaccination samples are shown in full in appendix C.2 and additionally accessible as 3D interactive landscapes online as listed in appendix C.5.

In keeping with findings from other vaccination studies, the magnitude of the response was highly variable among different individuals. This can be seen in figure 4.2 where the maximum titre fold-increase following vaccination is shown as a histogram. The maximum fold-titre increase was taken across all viruses titrated since it is known that in some cases, counter-intuitively, the maximum increase following an influenza exposure as measured by HI can be against a virus other than that encountered [Smith et al., 2004]. From figure 4.2, and the full set of antibody landscapes shown in appendix C.2 it can be seen that for some individuals there is very little or no change following vaccination whereas in others a substantial change is apparent.

As with infection, the response to vaccination often generated an antibody response that was strikingly antigenically broad but mostly limited by the extent of pre-existing immunity. In this regard too however, great heterogeneity was apparent with several qualitatively different types of response apparent. Some individuals generated responses that were more specific to the vaccine strain and antigenically similar variants, while others generated responses that were very broad with boosting of titres evident across their entire pre-vaccination landscape. These differences in response were not always explicable by differences in prior immunity as illustrated in figure 4.3 where four individuals with almost identical antibody landscapes prior to vaccination go on to have very different responses.

As described in chapter 2, the black lines represent the landscape height for each position on the antigenic summary path taken through the antigenic clusters (and

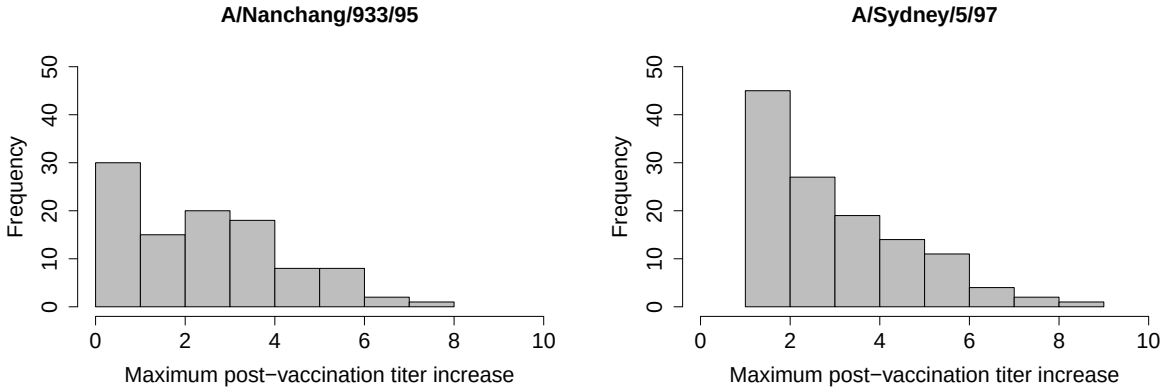


Figure 4.2: Distribution of maximum vaccine responses. Histogram displaying the maximum titre fold-increase (based on HI titres) upon vaccination against any virus for the two vaccination studies.

shown again on the antigenic maps in figure 4.1). The antibody landscape measured in the first sample taken is shown shaded in grey, in subsequent samples (in this case only the sample from 2012), coloured shading (green, light blue, purple or dark blue dependent on trial year, rather than red as in chapter 3) indicates increases, and beige decreases, compared with the previously sampled year. Additionally shown are blue-shaded rectangles that indicate antigenic clusters that had circulated during an individual's life span until sample collection. As in the previous chapter, dots along the x axis indicate the antigenic positions along the summary path of the subset of viruses used to generate each landscape, scaled negatively by distance from it. It can be noted in this particular figure how the shapes of the antibody landscapes are individually very distinct, yet remarkably consistent over time.

Interestingly there were also a minority of individuals who went on the have antibody responses so broad as to span the entire range of H3N2 antigenic space. This is particularly surprising since some of these strains were yet to circulate for another 12 years and were substantially antigenically different from either the vaccine strains or any of the viruses that could have been previously encountered by individuals in these trial populations. Figure 4.4 gives four examples of such extensive cross-reactive responses. It should be noted that although of course cross-reactivity extending to strains in future antigenic clusters would be normally be expected, the breadth seen here, extending up to 5 antigenic clusters ahead in some cases, is beyond what would normally be generated as part of a typical cross-reactive antibody response (see chapter 3, section 3.4.3.1).

At the opposite end of the scale, vaccine non-responders could also be clearly distinguished. It is already well recognised that many subjects will show no change in titre post-vaccination [Beyer et al., 2004], however since usual practice is only to titrate

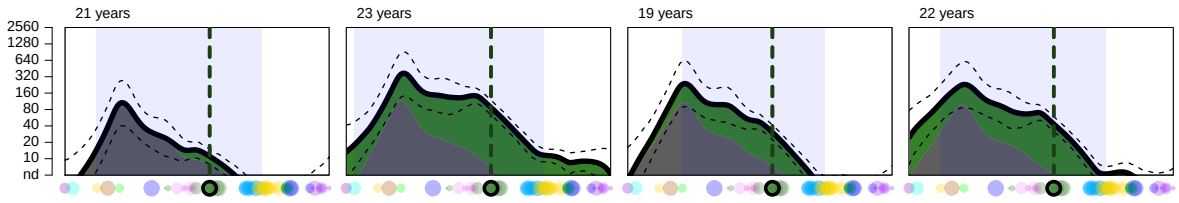


Figure 4.3: Variable vaccine responses despite similar pre-vaccination immunity. Four examples of individuals of similar ages who went on to have very different antibody responses to vaccination. Rest of legend as given below.

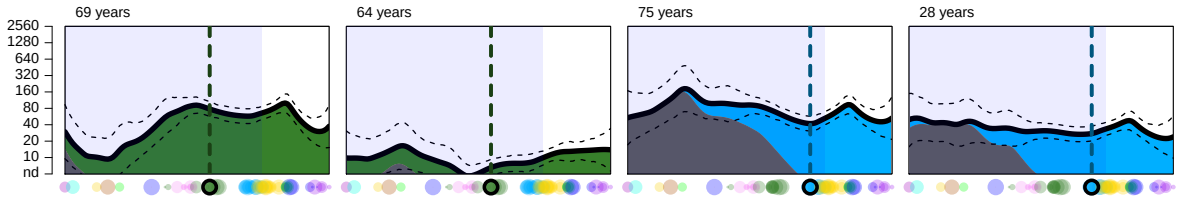


Figure 4.4: Antibody responses with very antigenically broad cross-reactivity. Four individuals who responded against the full antigenic range of H3N2 strains tested, in some cases extending up to 5 antigenic clusters ahead of strains circulating at the time of sample collection. Rest of legend as given below.

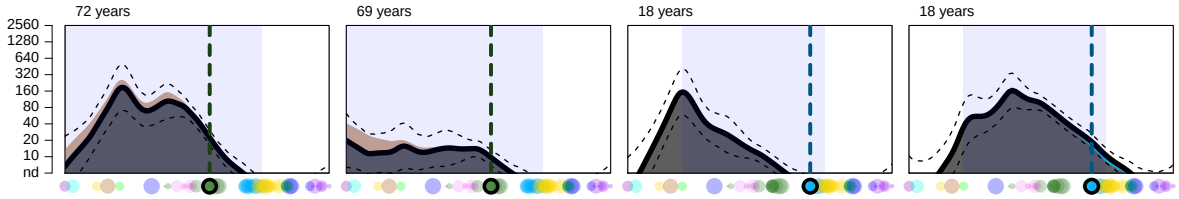


Figure 4.5: Individuals with no measurable antibody response to vaccination. Four examples of individuals who had no measurable increase in titres post-vaccination against strains from any of the antigenic clusters tested, or even a decrease (second from left). Rest of legend as given below.

Shared legend: Unlike plots in previous chapters, each panel shows the antibody response for a different individual, comparing the post-vaccination antibody landscape to the one measured pre-vaccination. For each plot, the solid black line represents the landscape height along the antigenic summary path defined in chapter 2 (and shown again on the antigenic maps in figure 4.1). Dashed black lines show the 95% confidence interval for the landscape fit. Green (for the 1997 vaccine trial) or blue (for the 1998 vaccine trial) regions show increases in antibody reactivity compared to pre-vaccination, while any beige regions represent a decrease. Dark grey regions represent areas where pre and post-vaccination antibody landscapes overlap. Additionally shown are blue-shaded rectangles that indicate antigenic clusters that had circulated during an individual's life span until sample collection. Dots along the x axis indicate the antigenic positions along the summary path of the subset of viruses used to generate each landscape, scaled negatively by distance from it. Finally, in each panel the antigenic location of the H3N2 vaccine strain (1997 = A/Nanchang/933/1995, 1998 = A/Sydney/5/1997) that the individual received is highlighted on the x-axis and marked with a vertical dotted line.

against the vaccine strain (or at most a handful of similar strains) it is difficult to draw firm conclusions as to why no rise in titres was detected. Did these subjects responded to a degree below the sensitivity of the HI assay, or perhaps responded instead against older strains, or did they truly have no overall antibody response to vaccination? With the more complete picture and higher resolution to changes in antibody reactivity afforded by the antibody landscapes analysis, it is apparent that indeed, many individuals had no detectable response to vaccination across the full range of viruses tested or indeed showed an overall decrease in HI titres post vaccination. Figure 4.5 shows some example landscapes but many more are apparent in the complete dataset shown in appendix C.2.

As can already be seen from figure 4.2, there was no clear cut-off between responders and non-responders in terms of maximum titre increase recorded with responses ranging continuously from nothing to very large titre increases. The same can be said of the vaccination responses as measured by antibody landscapes, so a distinction between a vaccine responder and non-responder will always be somewhat arbitrary. As with seroconversion in chapter 3, section 3.4.4.1, one approach is to look at the fold-change antibody landscape response to vaccination at an antigenic point of interest (in this case the antigenic location of the vaccine strain) and determine whether the lower bound of the one-sided 95% confidence interval lies above a fold-change of 0. Figure 4.6 compares the proportion of responders and non-responders in the 1997 and 1998 vaccine trials as measured using this approach and also as measured simply by a criteria of a four-fold or greater increase in titre against the vaccine strain (sera that were not titrated against the vaccine strain were excluded from this analysis).

Interestingly, there was a higher percentage of subjects classified as non-responders in the 1997 (33%) compared to 1998 (21%) vaccine trial, a difference of 12 percentage points more non-responders in 1997 (95% CI 0%,25%). In addition, as with infection, it can also be seen that a reasonable proportion of individuals classified as responders based on a significant antibody response did not reach the threshold of a four-fold or greater response to the vaccine strain alone (grey portions of the leftmost two bars in figure 4.6).

Despite the heterogeneity of individual responses, some general patterns could be identified. As discussed above, pre-vaccination immunity was clearly not accountable for a lot of the heterogeneity in terms of breadth or magnitude of individual responses but an overall relationship between prior immunity and response magnitude as measured by fold-change in HI titre was identifiable. This is in keeping with previous studies that have long recognised a negative relationship between pre-vaccination titre and magnitude of the fold-change vaccination response [Beyer et al., 2004]. Figure 4.7 shows the HI titre against the respective vaccine strain pre and post vaccination. To account for the censored nature of the data, the function “clikcorr” [Li et al., 2016] from the R package of the same name was used to estimate correlation

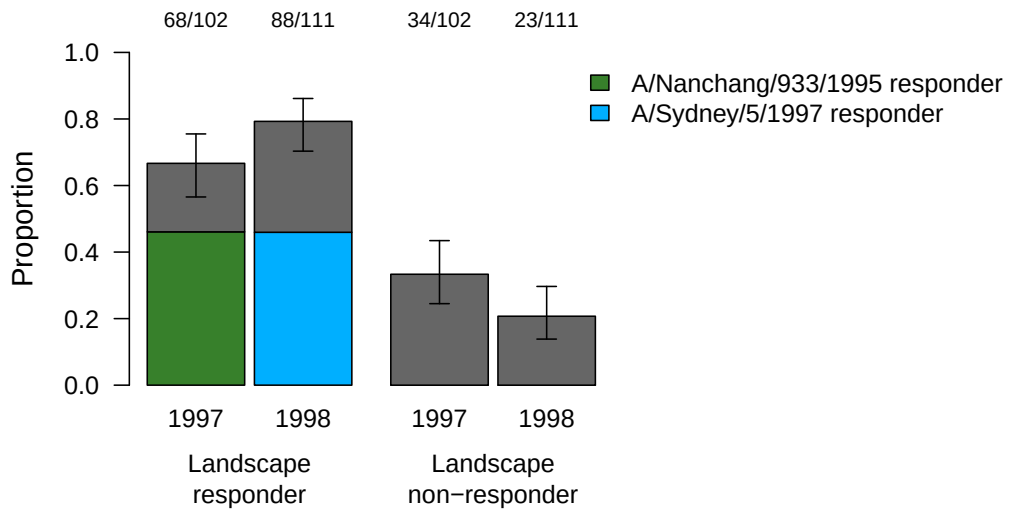


Figure 4.6: Comparison of the proportion of vaccine non-responders in the 1997 and 1998 trials. Barplots show the proportion of responders and non-responders in each vaccine trial, assessed by antibody landscapes as described in the main text. 95% confidence intervals based on a test of proportionality are shown alongside the raw numbers above. Coloured regions of the bars show the number of individuals in each category who showed a response against the vaccine strain itself. Note that many individuals did not reach the threshold of a four-fold or greater increase in HI titre to the vaccine strain yet responded significantly as measured by antibody landscapes.

between the log pre-vaccination titre and the size of the fold-change response. In both trials, the estimated correlation was significantly negative being -0.68 for 1997 (95% CI [-0.8,-0.52]) and -0.31 for 1998 (95% CI [-0.56,-0.02]).

4.4.2 *Comparison of prior immunity*

A main aim of this study was to compare the overall response generated by use of either a regular or antigenically advanced vaccine. As described in the introduction to this chapter, very little circulation of H3N2 viruses occurred between the two vaccination trials conducted in 1997 and 1998, but it was still important to check for any significant bias in terms of the extent of pre-existing antibody reactivity present against H3N2 strains in the two populations. This is especially important given the relationship between pre-existing antibody reactivity and antibody response upon influenza exposure observed.

Figure 4.8 shows the mean HI values for each virus when compared with a two-tailed t-test, treating non-detectable titres as -1. Although some viruses showed a significant difference by this measure, there appeared indeed to be no systematic bias favouring increased pre-vaccination antibody reactivity generally in one cohort compared to the other, confirming that the two vaccine trial populations started from equivalent

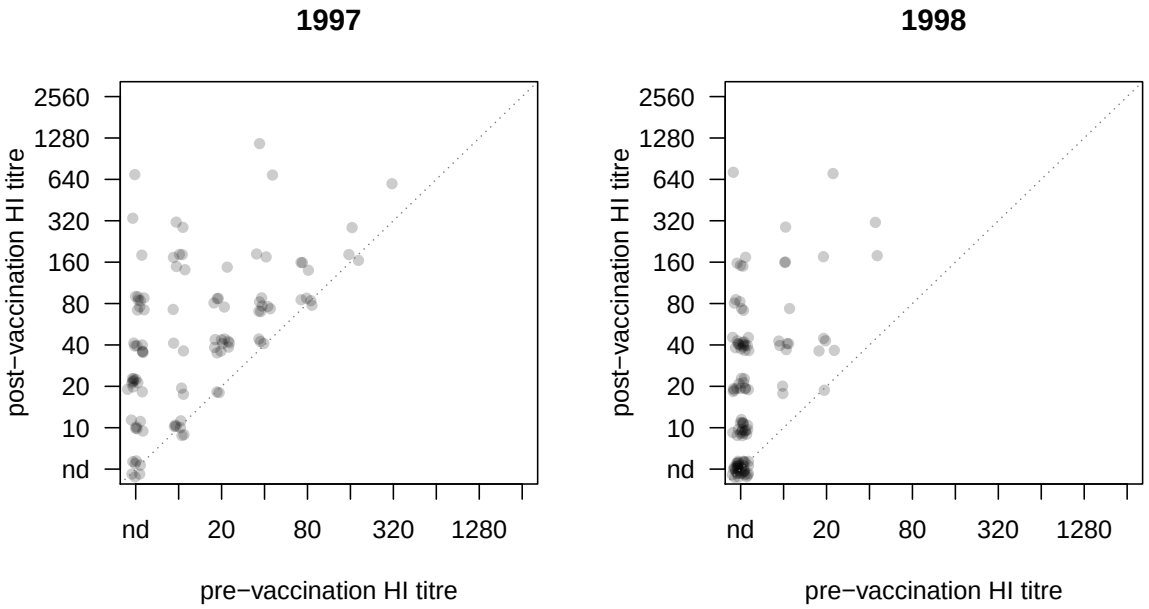


Figure 4.7: Pre vs post-vaccination HI titres. Pre vs post-vaccination HI titre to the vaccine strain. Each grey circle represents a different individual in the trial and a small amount of jitter has been applied to both axes to better visualise the density of data points.

baselines of antibody reactivity not just against the vaccine strain but across H3N2 antigenic space in general.

4.4.3 *Average vaccination responses*

In order to compare the average vaccination responses to the full range of H3N2 strains tested, an average "population landscape" was produced by calculating the mean landscape height across all individuals in each cohort along each point of the summary path. The pre and post vaccination population landscapes are shown in Figure 4.9, left panels. It can be seen that although the individual antibody landscapes and responses to vaccination were very heterogeneous, taken as a population average, the patterns of antibody reactivity are remarkably similar.

Reflecting the individual patterns noted previously, following administration of either vaccine, average antibody reactivity is increased across the full antigenic spectrum of H3N2 viruses against which there was already some prior population immunity. Again, although an average increase is detected even as far back as to the HK68 cluster, responses were on average greatest to the strains most similar to the vaccine strain administered (the antigenic location of which is marked as a dashed line).

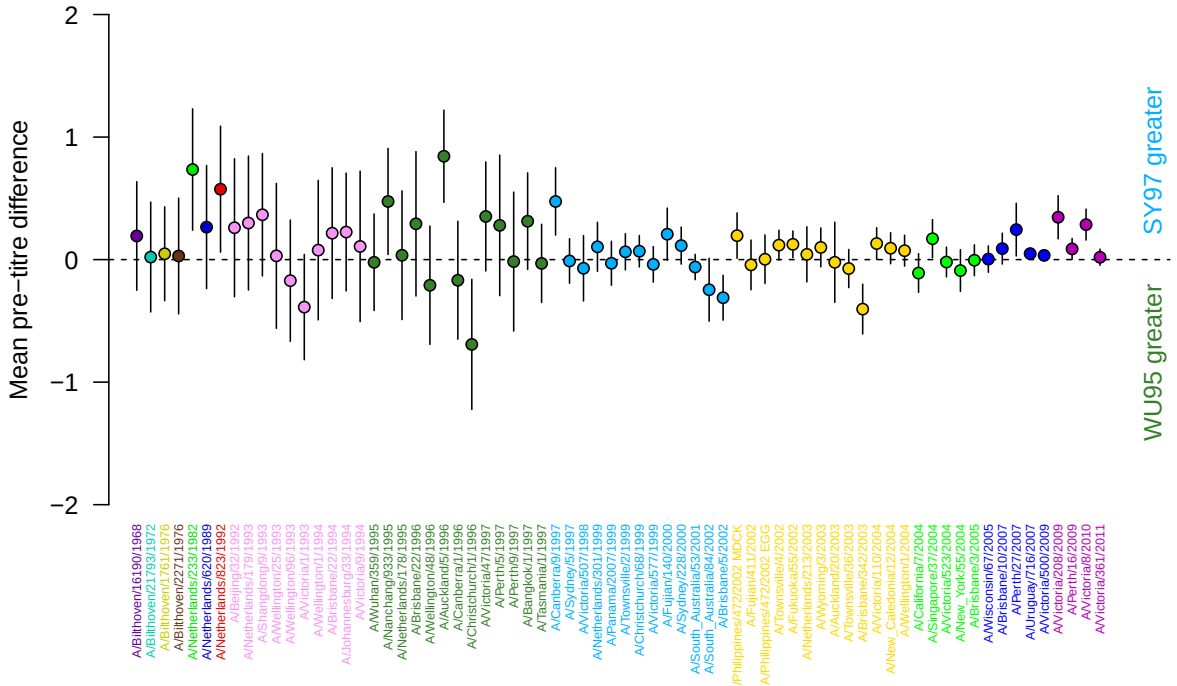


Figure 4.8: Comparison of pre-vaccination titres per antigen in the 1997 and 1998 vaccine trials. The difference between the sample means for each strain titrated against samples from the 1998 and 1997 studies is shown as a point, coloured by corresponding antigenic cluster. The bars indicate the 95% confidence interval of a two-tailed t-test for each antigen. Titres of <10 were treated as -1 on the log scale.

As would be expected, Figure 4.9 shows that average vaccination antibody responses were significantly greater against antigenic clusters from SY97 onward following vaccination with the antigenically advanced SY97-type vaccine, compared to that following vaccination with the WU95-type strain. Strikingly however, the breadth of antibody response to previous antigenic clusters was such that the SY97-type vaccine also produced antibody responses that were equally great or greater compared to previous clusters as well, even including WU95-type viruses against which it is a poorer antigenic match.

The difference in antibody response following administration of either the SY97-type or WU95-type vaccine is examined in detail in figure 4.9, right panel. This comparison of the vaccination response in the two vaccination studies uses paired pre- and post-vaccination data where for each individual, the response to vaccination was calculated as post-vaccination minus pre-vaccination antibody landscape height on the $\log_2$(HI titre/10) scale. These responses are then compared for the two vaccination studies by showing the difference between the two means: a positive titre difference indicates greater benefit from the A/Sydney/5/1997 vaccination and negative indi-

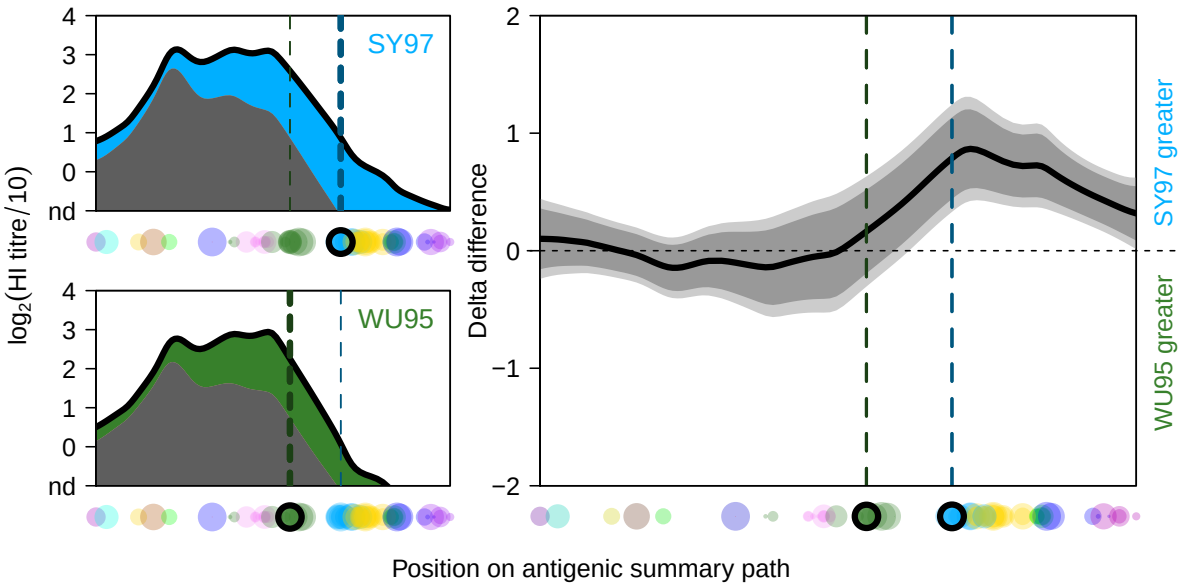


Figure 4.9: Average antibody landscape responses in the 1997 and 1998 vaccination trials compared. (top-left panel) The mean pre-vaccination landscape (grey) and landscape after vaccination with A/Sydney/5/1997 (blue) in the 1998 study (123 individuals) or **(bottom-left panel)** with A/Nanchang/933/1995 (green) in the 1997 study (102 individuals) for each position on the antigenic summary path defined in chapter 2 and shown again in figure 4.1. Dots along the x axes indicate the subset of 70 viruses used to generate these landscapes. The vertical dotted lines indicate the position of the SY97 (A/Sydney/5/1997- blue) and WU95 (A/Nanchang/933/1995- green) wild-type vaccine viruses. **(right panel)** Comparison of titre increase after vaccination with A/Nanchang/933/1995 or A/Sydney/5/1997 for each position along the antigenic summary path. When above the horizontal midpoint, the black line indicates a higher response in the group vaccinated with A/Sydney/5/1997; when below the midpoint, it denotes a higher response in the group vaccinated with A/Nanchang/933/1995. Data were calculated from the average titre increase between each individual's paired pre and post-vaccination titre, with 95% (dark grey) and 99% (light grey) t test–based confidence intervals.

cates a greater benefit from the A/Nanchang/933/1995 vaccination. The associated confidence intervals were based on t-tests comparing the responses of the individuals in the two studies.

The surprising finding of superiority of antibody responses in the group receiving the antigenically advanced SY97 also held when examining HI titres more directly. Figure 4.10 compares the mean antibody responses for each strain separately between the 1997 WU95 and 1998 SY97 trials. Antibody responses were again calculated as the $\log_2$(post titre) - $\log_2$(pre titre) with non-detectable titres treated as -1. Again it can be seen that antibody responses were significantly greater against strains from the SY97 cluster onwards while still being at least as good against strains from all previous antigenic clusters.

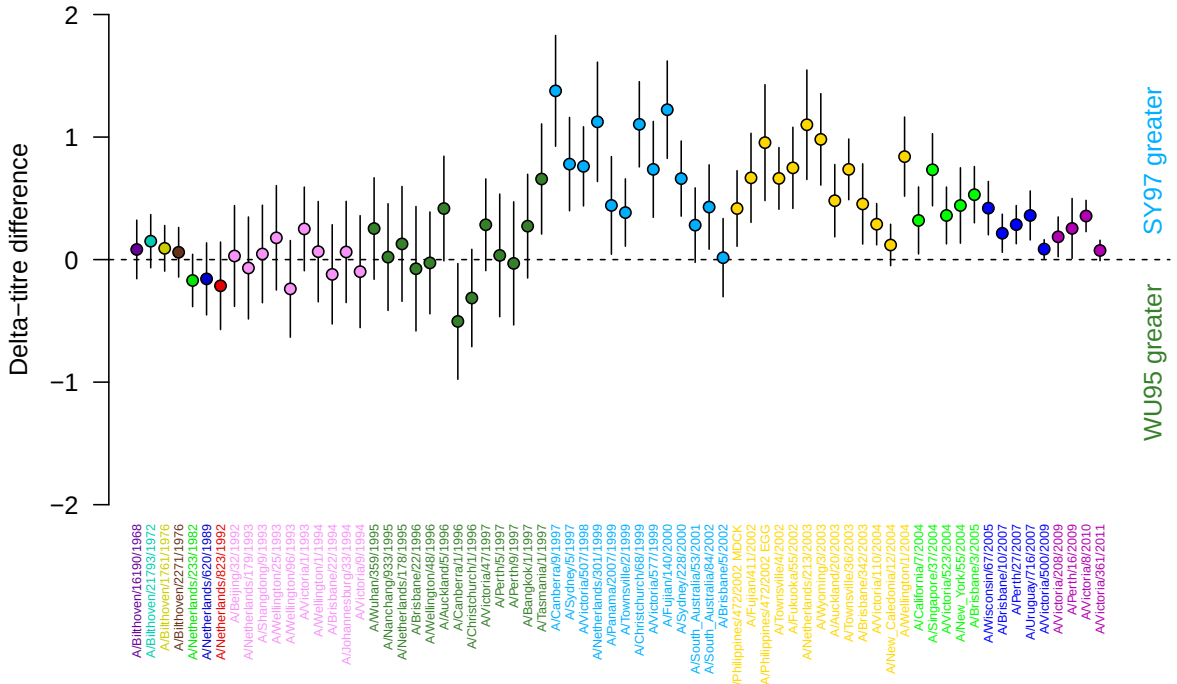


Figure 4.10: Comparison of 1997 and 1998 vaccine trial responses by strain. The difference between the sample means of the response to vaccination in the 1998 and 1997 studies (based on the post-vaccination minus pre-vaccination titre for individuals in both studies) is shown as a point for each antigen, coloured by antigenic cluster. Each vertical solid line indicates the 95% Bonferroni-corrected confidence interval of a two-tailed t-test.

4.4.3.1 Responses across the full landscape

As an additional check to assess whether the conclusions on antigenically advancing the vaccine held for all antigenic locations of the map, the difference in vaccination response was examined for each position of the landscape (not only along the summary path). Figure 4.11 shows this difference in colour-code, and the absence of green regions, and prominent presence of blue regions clearly indicates that across all regions of antigenic space, responses were either equivalent or greater following vaccination with A/Sydney/5/1997.

4.4.4 *Estimated "protection" rates*

Mean antibody titre responses in vaccine trial group are a strong indicator of the relative beneficial effects of different vaccines but it was also important to check the effect of vaccination on an individual level. Although technically the same fold increase in titre, on the individual scale we may not be so interested in use of a vaccine that increases an individual's titre from 320 to 2560 as on that increases their titre

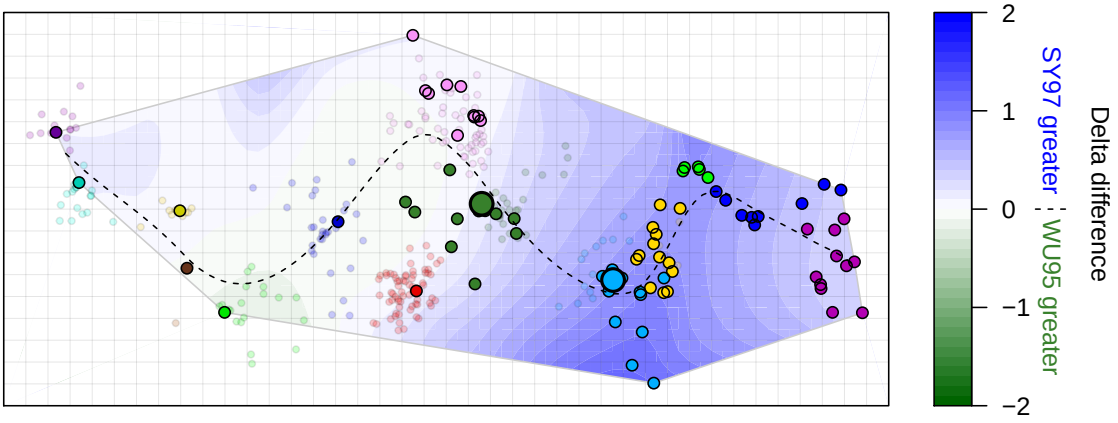


Figure 4.11: Average 1997 and 1998 vaccination trial responses compared across the full antigenic map. The antigenic map shows the antigenic position of viruses for which titres were determined in the vaccination studies, represented by dots colour-coded by antigenic cluster; the two larger symbols indicate the two vaccine strains. The shading represents the difference between the mean increase in landscape height at that antigenic position for individuals vaccinated with A/Sydney/5/1997 compared to those vaccinated with A/Nanchang/933/1995 (scale given on right-hand side). Grey represents antigenic regions outside of the convex hull bounded by the viruses titrated, outside of which the antibody landscape would need to be obtained by extrapolation. The summary path along which the 2-dimensional antibody landscapes are compared is shown by a dashed line.

from 10 to 80 since the former individual is unlikely to gain any extra protection following an increase of an already high titre, while for the latter it may translate to a great benefit. The notion of a "protective" HI titre is highly contentious but historically a titre of 40 has often been used as a cut-off for an HI titre that translates to a reasonable level of protection [Coudeville et al., 2010b]. The proportion of individuals with an antibody landscape height of $\geq$40 on the raw titre scale was therefore analysed for the 1997 and 1998 vaccine trial groups both before and after infection, with the results shown in figure 4.12.

4.4.5 *Responses in "at risk" individuals*

To check if the apparent benefits of the antigenically advanced vaccine used in 1998 still held when examining individuals typically considered at most risk of influenza infection, the analysis was repeated on two subsets of the data, elderly individuals ($\geq$60 years old) and individuals with a pre-vaccination titre of $<$40 against the 1997 vaccine strain A/Nanchang/933/1995, a titre as discussed above that is often considered to be non-protective.

Figure 4.13a shows the equivalent results from figure 4.9 when considering only individuals $\geq$60 years old. The findings of the full study are again mirrored in this subset,

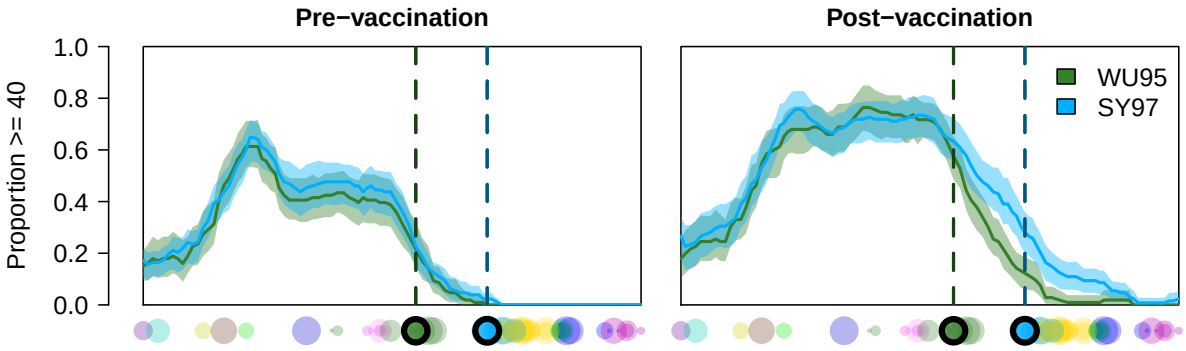


Figure 4.12: Estimated protection rates pre and post vaccination in the 1997 and 1998 trials. The proportion of individuals with a landscape value ≥40 (2 HI units) is shown for the 1997 study on A/Nanchang/933/1995 (WU95 - green) and the 1998 study on A/Sydney/5/1997 (SY97 - blue). This is calculated in both the pre-vaccination landscapes (left panel), and the post-vaccination landscapes (right panel). Lighter shaded regions represent the estimated standard error of the proportion protected in each population at each position.

with similar pre-vaccination titres and an overall antibody response in the recipients of the antigenically advanced SY97-like vaccine virus that is always equivalent to or greater than that of the WU95 group.

The subset of individuals with a pre-vaccination titre of <40 against the 1997 vaccine strain A/Nanchang/933/1995 is analysed in figure 4.13b and again shows results consistent with those drawn from the full population. Additionally, given that the back-boost appeared to be dependent upon prior immunity, figure 4.13c additionally shows the results for the subset of individuals with a non-detectable (<10) pre-vaccination titre against A/Nanchang/933/1995. There were only 41 individuals in the 1997 vaccination cohort and 38 individuals in the 1998 vaccination cohort with a <10 titre against this strain, and the confidence intervals are relatively large when response magnitudes are compared. Both groups still however, appear to have a roughly equivalent response against the WU95 cluster and thus vaccination with the SY97 strain does not appeared to have compromised the antibody response against these strains even where pre-vaccination immunity was so low. Interestingly, of the 79 individuals with a <10 HI titre against A/Nanchang/933/1995, only 8 in the 1997 and 17 in the 1998 cohort have <10 HI titres for all of the WU95 viruses. The low, but present antibodies reactive against some WU95 strains in the other individuals, may therefore still be providing a substrate for a small back-boosting effect.

4.4.6 *Neutralisation assay results*

The HI assay has been historically favoured as a relatively quick, cheap and effective assay for measuring binding of antibodies that functionally inhibit binding of influenza viruses but it does not necessarily reflect the whole serum antibody reper-

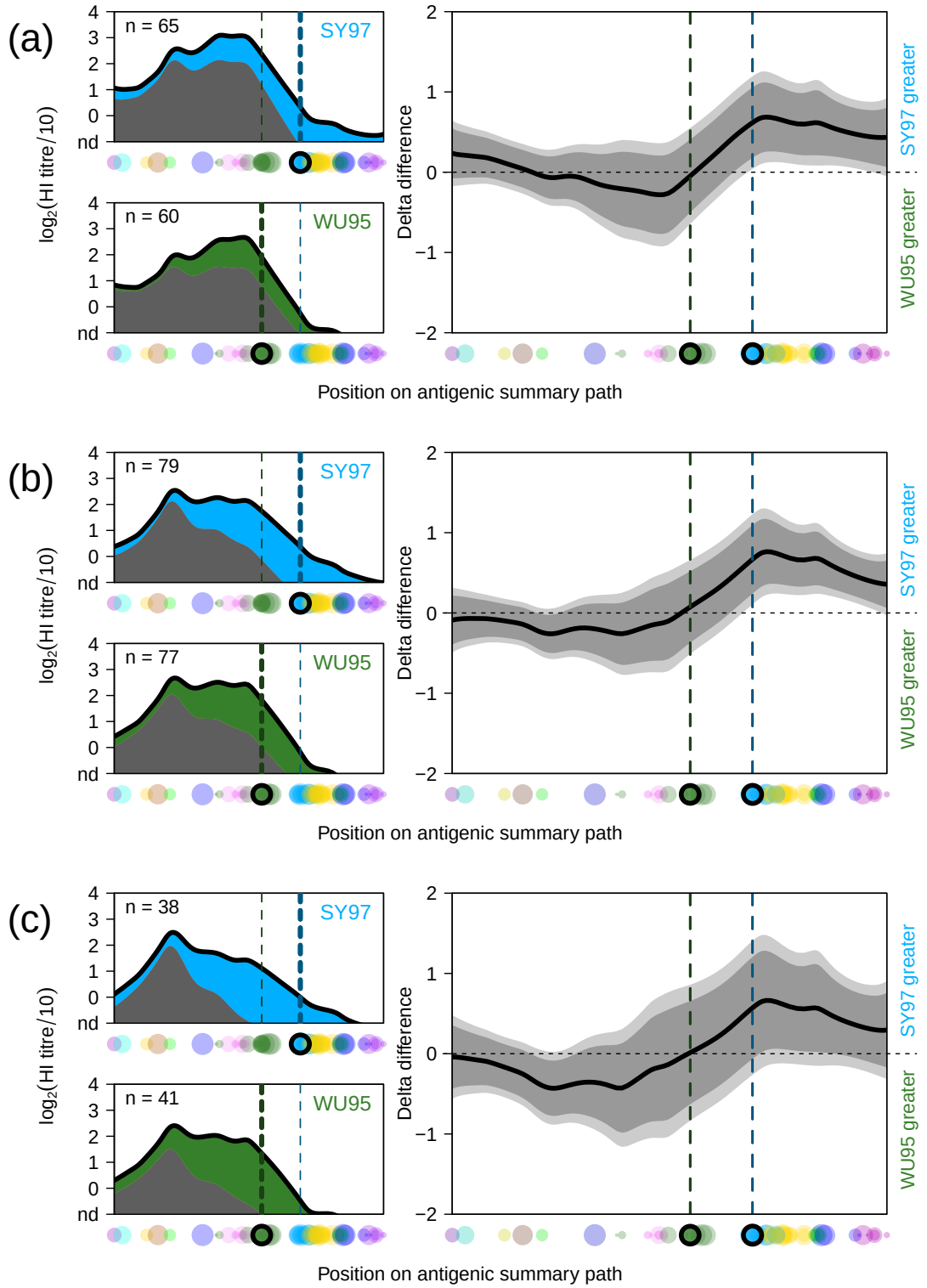


Figure 4.13: 1997 and 1998 vaccine trial responses for different subsets of at-risk individuals. (a) ≥60 years old, **(b)** A/Nanchang/933/1995 pre-vaccination titre <40, **(c)** A/Nanchang/933/1995 pre-vaccination titre <10. Legend otherwise as for figure 4.9 but the numbers in each vaccine subgroup are also shown in each plot.

toire capable of providing protection against influenza. Neutralisation assays are an alternative approach with their own limitations but provide an opportunity to look at the effectiveness of antibodies more generally to inhibit viral growth and these tests had additionally been performed by collaborators as described in the methods against both vaccine strains (A/Nanchang/933/1995 and A/Sydney/5/1997) for pre- and post- vaccination sera from the 1997 and 1998 vaccination cohorts.

Unfortunately some samples had insufficient volume for the additional viral neutralisation tests meaning that only a subset were tested. Typically the sera that had to be excluded were from vaccine high-responders as measured by HI that were previously used for other studies. As a result there was a sample selection bias favouring low-responders in each cohort. This was particularly present for samples from the 1998 cohort where only 69 from 123 samples could be tested, in comparison to 80 from 102 in the 1997 cohort.

Figure 4.14 compares raw titres and fold-change responses as measured by either the HI or viral neutralisation assay. As before in other comparisons, fold-change responses were calculated as the $\log_2$ post-titre minus the $\log_2$ pre-titre, with non-detectable titres treated as 5 on the raw titre scale.

Note that the line shown is an orthogonal regression through the data points rather than a linear regression since there is expected to be significant error in both x and y variables. In particular, orthogonal regression is a form of Deming regression, where the standard deviation of the error is assumed to be the same in both axes - a simplifying assumption made in this case (it is of course debatable whether the measurement error in the HI assay and neutralisation assay is equivalent but it is perhaps not unreasonable to assume they would be of at least similar magnitude). Multiple imputation was used to account for non-detectable values, assuming interval censoring for HI titres and fold-change differences as described in chapter 2, sections 2.5.2 and 2.5.3. Values were repeatedly imputed by drawing randomly from a uniform distribution bounded by the relevant interval for a given measurement, and orthogonal regression was then performed on the resulting data using the "deming" function from the R package of the same name [Therneau, 2014]. Median values of the orthogonal regression slope and intercept from 10000 imputations were then used to represent the most likely fit to the data. Finally, 10000 imputations with additional bootstrap resampling of the data were used to estimate the 95% confidence intervals shown.

With the caveats above in mind it can be seen in figure 4.14 that both in terms of raw titres and also fold change in titres, the fits indicate that the neutralisation assay tended to return significantly higher values. In general however, the titres were largely comparable and correlated.

Figure 4.15 goes on to compare the results of the HI and neutralisation assay titres per vaccination cohort. As already noted, neutralisation titres were higher on average

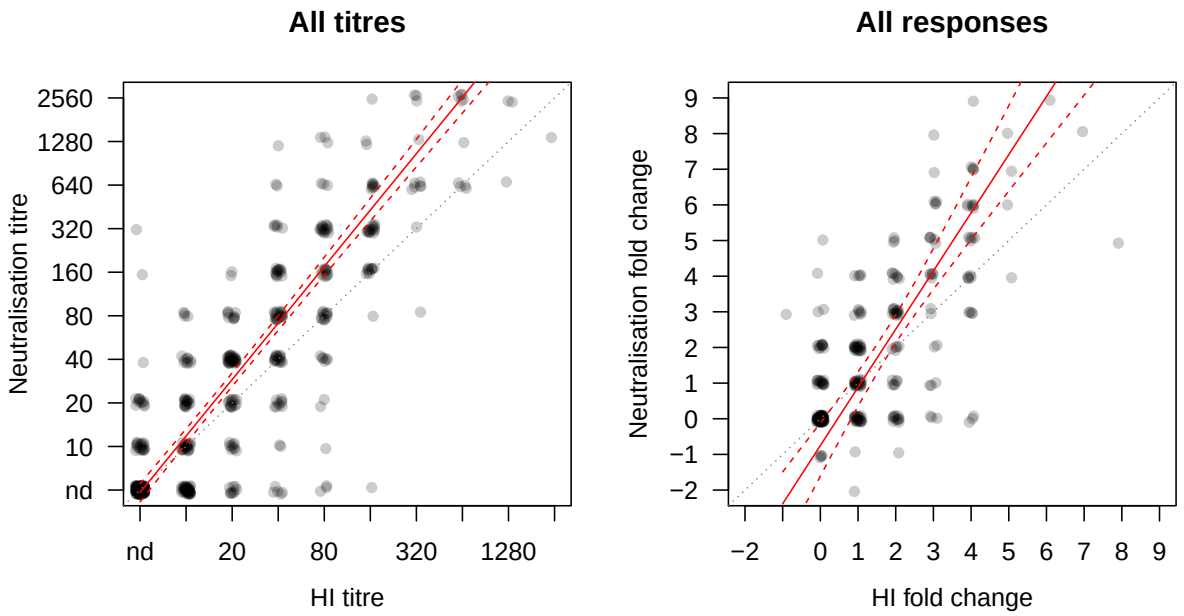


Figure 4.14: HI assay vs. neutralisation assay titres. Scatter plots with the results shown from a principal component analysis comparing HI titres with neutralisation titres. Left panel: analysis of raw titres, plotting pre- and post-vaccination titres as measured by the two assays. Right panel: analysis of the vaccination response (post-vaccination - pre-vaccination titres) as measured by HI and virus neutralisation assay. Individual points are shown as circles and a small amount of jitter has been applied to the points in both the x and y axis to better visualise the spread of data. The solid red line shows a principal component fit through the data with dashed red lines showing the 95% confidence intervals of the fit based on 500 bootstrap repeats.

than HI but the relative response patterns for the different cohorts remains similar. Most importantly, there was not a significant difference in mean post-vaccination titres against the WU95 vaccine strain A/Nanchang/933/1995 between the groups whether measured by HI or virus neutralisation. This supports further the surprising conclusion that use of the antigenically advanced vaccine strain in 1998 did not come at the cost of a reduced antibody response against the previous antigenic cluster.

It was also noted in figure 4.14 that as part of the variation between the two assays, sera from some individuals returned measurable titres in the HI assay but non-detectable titres in the virus neutralisation assay. It was therefore theoretically possible that more of these lower responders could be present in the 1997 study introducing a source of subtle bias. Reassuringly, when the proportion of these individuals with lower titres as measured by virus neutralisation vs HI against the A/Nanchang/933/1995 strain were compared between the two groups, no significant difference was found, being 24% in the 1997 cohort, and 28% in the 1998 cohort (p-value = 0.47).

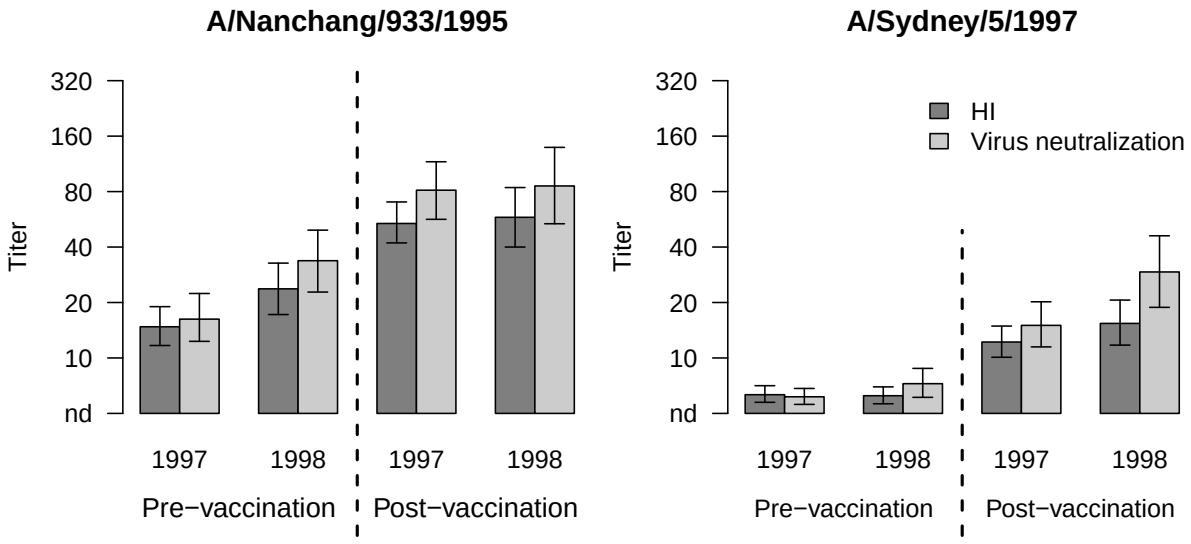


Figure 4.15: Titres for the 1997 and 1998 vaccine trials by HI or neutralisation assay. Mean pre- and post-vaccination titres in the subset of the 1997 and 1998 vaccination cohorts measured using both HI and viral neutralisation. Error bars show the standard error based on calculations of the mean from 500 bootstrap repeats.

4.4.7 *Responses in the 2009 and 2010 vaccination trials*

In general, results from the 2009 and 2010 vaccination trials corroborated findings from the 1997 and 1998 studies. Similar heterogeneity of individual responses is evident, and again the strikingly broad "back-boost" response in many individuals. Indeed, even though the vaccine strains used in the 2009 and 2010 studies were even more antigenically distinct from the earliest strains, some individuals still recorded responses against the strains tested from the earliest antigenic clusters. Individual landscapes from each vaccine trial are shown in full in appendix C.2.

With fewer individuals tested and fewer strains titrated, drawing firm conclusions about any overall difference in response to vaccination in these years is more challenging. Regarding prior-immunity, unlike the period between the 1997 and 1998 vaccination studies, there was significant circulation of the PE09-like viruses between the dates of the two studies in 2009 and 2010 [World Health Organization, 2009]. This is reflected to some extent in figure 4.16a where there was significantly increased pre-vaccination titres against A/Perth/16/2009, although there was no difference against the only other strain from the PE09 antigenic cluster that was tested, A/Hanoi/53/2010. Overall however, this translated to a significant difference in the pre-vaccination landscapes in the antigenic region of the PE09 cluster as shown in figure 4.16b.

Where average population antibody landscapes and responses are compared in figure 4.17, unlike the 1997 and 1998 trials, there was on average pre-vaccination antibody

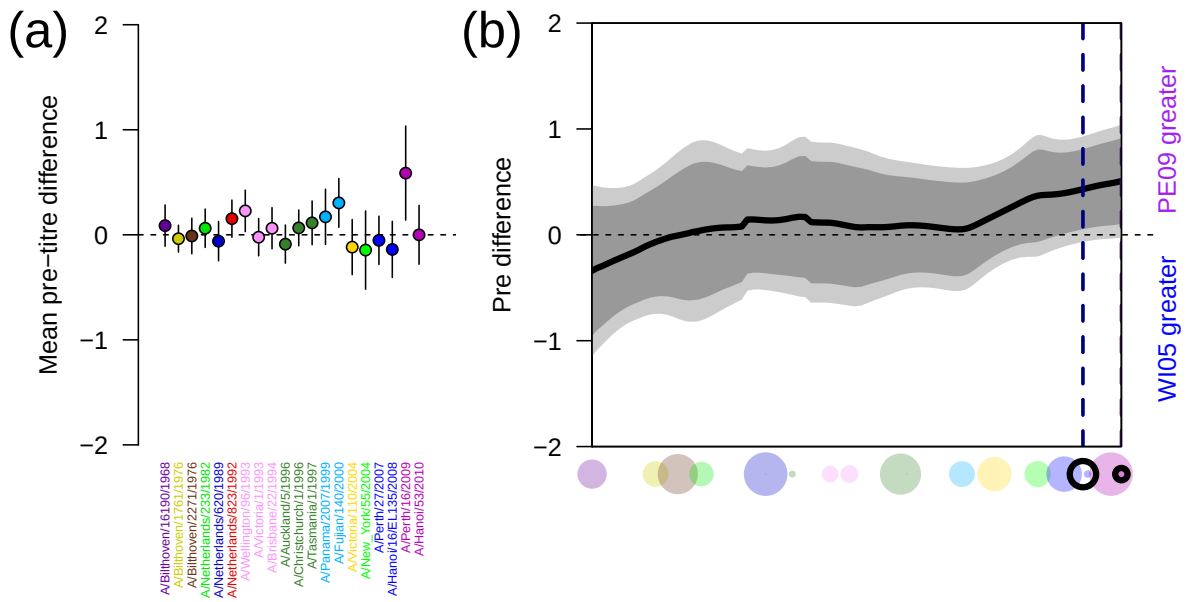


Figure 4.16: Comparison of average pre titres for the 2009 and 2010 vaccine trials. (a) Comparison of pre-vaccination titres per antigen in the 2009 and 2010 vaccine trials. The difference between the sample means for each strain titrated against samples from the 2010 and 2009 studies is shown as a point, coloured by corresponding antigenic cluster. The bars indicate the 95% confidence interval of a two-tailed t-test for each antigen. Titres of <10 were treated as -1 on the log scale. **(b)** Differences in average pre-vaccination landscape along the summary path for individuals in the 2009 and 2010 vaccine trials. When above the horizontal midpoint, the black line indicates more pre-vaccination antibody reactivity in the group vaccinated with A/Perth/16/2009; when below the midpoint, it denotes more pre-vaccination antibody reactivity in the group vaccinated with A/Brisbane/10/2007. The black line was calculated as the mean pre-vaccination antibody landscape height from those in the 2010 trial minus mean height from those in the 2009 trial, with 95% (dark grey) and 99% (light grey) t test–based confidence intervals. The vertical dotted lines indicate the position of the PE09 (A/Perth/16/2009-magenta) and WI05 (A/Brisbane/10/2007- dark blue) wild-type vaccine viruses.

reactivity well into the detectable range against both vaccine strains before they were given. A direct titre comparison cannot be made since the 2009 vaccine virus A/ Brisbane/10/2007 was not included in the titration set, however the average pre-vaccination antibody landscape height on the log scale at the location of the vaccine virus was in 0.42 in 2009 and -0.03 in 2010, compared with for example 0.77 in 1997 and -1.12 in 1998.

Differences in response did not reach the 95% level of significance at any point along the antigenic summary path although in general the overall pattern was more as expected for antigenically different vaccines with slightly larger responses against WI05 type viruses in the 2009 group that received the WI05 type vaccine virus A/ Brisbane/10/2007 and slightly larger responses against PE09 type viruses in the 2010 group that received the PE09 type vaccine virus A/Perth/16/2009.

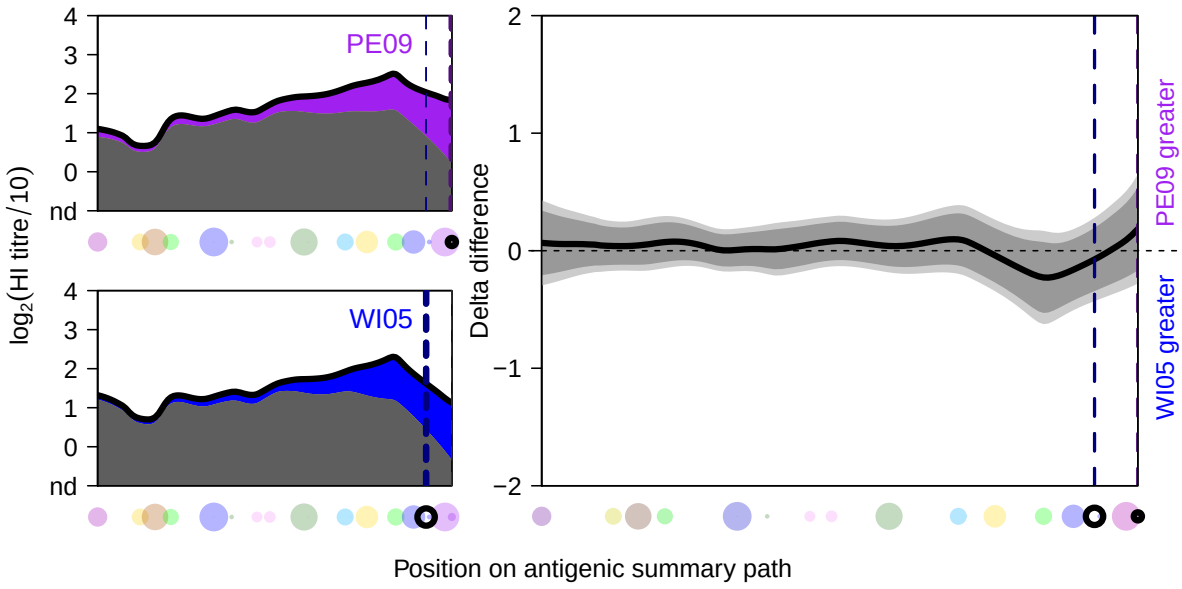


Figure 4.17: Average antibody landscape responses in the 2009 and 2010 vaccination trials. (top-left panel) The mean pre-vaccination landscape (grey) and landscape after vaccination with A/Perth/16/2009 (magenta) in the 2010 study (80 individuals) or **(bottom-left panel)** with A/Brisbane/10/2007 (dark blue) in the 2009 study (80 individuals) for each position on the antigenic summary path defined in chapter 2 and shown again in figure 4.1. Dots along the x axes indicate the subset of 20 viruses used to generate these landscapes. The vertical dotted lines indicate the position of the PE09 (A/Perth/16/2009- magenta) and WI05 (A/Brisbane/10/2007- dark blue) wild-type vaccine viruses. **(right panel)** Comparison of titre increase after vaccination with A/Brisbane/10/2007 or A/Perth/16/2009 for each position along the antigenic summary path. When above the horizontal midpoint, the black line indicates a higher response in the group vaccinated with A/Perth/16/2009; when below the midpoint, it denotes a higher response in the group vaccinated with A/Brisbane/10/2007. Data were calculated from the average titre increase between each individual's paired pre and post-vaccination titre, with 95% (dark grey) and 99% (light grey) t test–based confidence intervals.

4.5 Discussion

The vast majority of previous investigations into the antibody response to vaccination have included only titrations against the vaccine strains used but again we see here the value of using a range of titrations to assess antibody responses more broadly. The findings here have important implications for vaccine strain selection and the best approach to optimise antibody responses to vaccination.

4.5.1 *The response to antigenically advanced vaccination*

Although not entirely unexpected given the strength of antibody recall seen in response to infection, the finding that an antigenically advanced vaccine strain was able to providing better responses against later strains while still boosting prior immunity to the same extent as a theoretically antigenically better matched strain goes against the classical understanding of the antibody response to vaccination. Previous thinking, indeed the theory behind the current approach to vaccine strain selection,

would predict that while use of an antigenically advanced vaccine strain may of course stimulate responses against previously encountered antigenic variants, this would not be effective to the same extent as the use of a more appropriately matched vaccine strain.

Why antibody responses against Wuhan 1995 and earlier strains were equally stimulated by a A/Sydney/5/1997 vaccine that was theoretically less antigenically well matched to these strains than the alternate A/Nanchang/933/1995 strain requires some speculation. One interpretation is that this finding is a result of antibody recall responses being more readily stimulated than the generation of primary ones [Tangye et al., 2003]. Thus, the relatively poorer antigenic match of the A/Nanchang/933/1995 strain to Sydney 1997-type viruses meant that it generated a paucity of additional "novel" antibody reactivity against these viruses. In the reverse situation, since stimulation of the memory B cell population is less heavily dependent on binding strength of BCR-antigen interactions [Liu et al., 1995], the A/Sydney/5/1997 strain was still able to stimulate recall of antibody against old strains as adequately as A/Nanchang/933/1995. This hypothesis could explain the patterns seen to some extent but although strength of antibody recall appears to be less dependent on antigenic characteristics of the stimulating strain than a novel response, we still witness a dependence upon antigenic distance. The logical implication is that although differences in stimulation of prior immunity between the A/Nanchang/933/1995 and A/Sydney/5/1997 would be more subtle than differences in a primary response, A/Nanchang/933/1995 should still stimulate slightly greater antigenic recall, being antigenically less distant to the older strains. In actual fact we see no evidence for any differences at all in either the breadth or magnitude of antibody recall between the two strains.

An alternative explanation for the strength of antibody stimulation upon receipt of A/Sydney/5/1997 relates to the antigen trapping hypothesis introduced in chapter 1 [St Groth and Webster, 1966a]. Following a similar line of thought one could imagine that upon vaccination, a certain quantity of antigen is neutralised unproductively by free serum antibody, reducing the amount ultimately available for stimulation of both memory and naive B cell division in peripheral lymph nodes. Indeed, exactly this effect has been described in passive antibody transfer experiments [Angeletti et al., 2017]. Following this line of thought, the smaller level of prior antibody reactivity against the A/Sydney/5/1997 strain could result in a greater effective antigen dose reaching the lymph nodes, increasing both novel and recall B cell and antibody responses that compensate for the otherwise greater antigenic mismatch between A/Sydney/5/1997 and previous H3N2 strains. An interesting consequence of this hypothesis is that, depending upon the dynamics of antigen removal by free serum antibody and B cell stimulation, it may even be possible to stimulate not just equal but perhaps *better* antibody responses against current viruses through use of a mis-

matched strain, providing novel antibodies are still stimulated in sufficient quantities and maintain some cross-reactivity against the current variants.

It is also possible that the patterns seen in these studies are simply an effect of the A/Sydney/5/1997 vaccine being somehow more immunogenic for reasons unrelated to prior immunity. Given that the difference in antibody response pertained only to later strains and was not generally greater against all antigens this seems unlikely but it cannot be discounted. Perhaps more convincingly, although the potential significance does not appear to have been fully appreciated at the time, similar patterns of equivalent antibody recall upon use of an antigenically more distinct strain can also in fact be additionally seen in many of the seminal papers published on the subject of secondary antibody responses, suggesting that the effect is indeed consistently present and not simply the result of confounding variables in this particular pair of trials [Davenport and Hennessy, 1956; St Groth and Webster, 1966b].

4.5.2 *Variation in individual responses*

In many regards, the patterns of antibody response witnessed in this chapter in response to vaccination mirror those witnessed in the previous chapter in response to infection. As with infection, many of the responses are very antigenically broad, again stimulating an increase in titres against strains right back to the Hong Kong 1968 antigenic cluster in some cases. However, given that the responses here are to a much more controlled stimulus, with the same strains given and samples taken at approximately the same time post-exposure, the heterogeneity of responses becomes more obvious. Differences in age and prior immunity clearly affect antibody responses [Beyer et al., 2004; Goodwin et al., 2006], but, as highlighted in figure 4.3, individuals of similar ages and patterns of prior immunity can go on to have strikingly different responses.

It is perhaps not surprising that vaccine responses could include so much heterogeneity given the potential number of factors that could effect outcomes, including amongst surely many others, effectiveness and location of vaccine delivery [Bachmann and Jennings, 2010], stochasticity of the antibody response and individual genetic variation in the components of the immune response [Srivastava et al., 2009]. However, despite these source of variation, the difference seen not simply in the magnitude but also the breadth and overall pattern of antibody response to the same vaccination is still something I find striking. Some individuals appear to have responses that are very focussed around the vaccine strain, seemingly more characteristic of a novel-type antibody response, as discussed, others show a broader recall response similar to infection, in yet others the response is focussed but focussed around a previously encountered strain rather than the vaccine antigen and seemingly everything in between.

Arguably the most interesting individual responses are the few individuals such as those highlighted in figure 4.4 who show uniformly broad antibody responses against

all H3N2 strains measured, even those that would not go on to circulate for a further 12 years. It is of course tempting to conclude that these responses must be due to production of antibodies against some highly conserved epitope. A typical candidate in this regard would be the haemagglutinin stalk but given that such antibodies would not classically be expected to interfere with haemagglutinin binding directly, it is unlikely they would be detected by HI as is the case here [Krammer and Palese, 2013; Okuno et al., 1993]. Given that such broad antibody reactivity is only present in the samples taken post vaccination and is not seen in pre-vaccination samples, the conclusion is reached that whatever the source of these highly cross-reactive titres, they are not typically maintained long-term. Here, the same questions arise as those regarding the production of anti-HA-stalk antibodies, namely if such broadly reactive antibodies exist in an individual's repertoire, why are they not continually boosted to higher levels such that continual cross-protection is provided. Again, given that these antibodies appear to be blocking HA-binding directly, the argument applied frequently with regard to the HA-stalk - namely that it is a subdominant epitope - does not appear so readily to apply in this case [Andrews et al., 2015; Angeletti et al., 2017].

A final category of interest is that of the total non-responders to influenza vaccination. It is a well recognised phenomenon that many individuals fail to show an antibody response to vaccination but this is typically as judged by responses to a single strain (the vaccine strain) [Beyer et al., 2004; Goodeve et al., 1983]. In such cases it is unclear whether the response was simply too small to detect by HI, or was cancelled out by measurement error, or if indeed there was no response. Looking at the antibody landscape responses it now becomes clear that many individuals really appear to fail to mount any response detectable with HI to vaccination, even small.

Certainly much remains to be understood with regard to the individual responses to vaccination and techniques such as antibody landscapes could provide a useful tool for the identification of individual responses worthy of more detailed experimental analysis. It would be useful for example to investigate the specific antibody basis for the patterns of broad reactivity seen in some individuals and certainly informative, if possible, to track whether these broad responses do indeed for some reason fail to be maintained long-term as is inferred here. Identifying the reasons for such absent antibody responses in some individuals may also be a key to further unlocking improved vaccine efficacy.

4.5.3 *Implications for vaccine strain selection*

As introduced in chapter 1, the vaccine strain is selected at least 6 months in advance of delivery to the population and situations can arise whereby it is unclear whether the current strains will remain in circulation or strains with new antigenic characteristics will instead begin to dominate in the upcoming season [Stöhr et al., 2012].

Is it best to maintain the current vaccine strain based on knowledge of the strains circulating at the time of vaccine strain selection or to update the strain to match newer antigenic variants in anticipation that these variants will soon go on to replace those currently circulating? Both choices hold a risk of antigenic mismatch between vaccine and circulating strains, either that we maintain the same vaccine strain and new antigenic variants dominate next season or that we update the vaccine strain pre-emptively and the same strains do indeed continue to circulate for another season. In the context of primary immune responses, similar impacts upon overall vaccine efficacy may be expected as a result of either scenario but the results here demonstrate important differences between them in populations with prior immunity. Following a mismatch due to a delayed vaccine update where the vaccine strain lags behind antigenic evolution of the virus, neither pre-existing immunity or newly produced antibodies provide protection against the novel strains [Belongia et al., 2009]. However, following a mismatch due to an incorrectly timed pre-emptive antigenic update it appears that the extensive boosting of pre-existing antibodies would still ensure that equivalent antibody titres against viral strains are induced from the current cluster.

Although conclusions drawn here would be expected to apply for populations similar to those studied in the A/Nanchang/933/1995 and A/Sydney/5/1997 vaccine trials we should probably be cautious extrapolating this to the wider global population of vaccine recipients. None of the analyses of subpopulations typically considered most at risk from influenza infection suggested differences were present in these groups, but the study populations included only healthy adults with no severe co-morbidities. It is possible that while pre-emptive vaccine updates may improve immune responses to vaccination in the majority of individuals, they may have reduced efficacy in precisely the most vulnerable populations that require protection against influenza the most. Additionally, in the two main trial years considered here there was already significant levels of prior immunity against Wuhan 1995-like strains, this antigenic cluster having already circulated for 3 seasons at the time of the 1997 vaccine trial. Whether this would still hold in a population with less prior-immunity to the current antigenic cluster should also be an important focus of future research. Finally, given the difference seen between short and longer-term antibody responses seen in chapter 3, it will be important to determine how these antibody responses go on to later evolve throughout a potential influenza season.

5

The Development of Immunity to Influenza

A street in Managua, Nicaragau, home of the Managua cohort studies.

5.1 Chapter summary

In this chapter I explore the early stages of the development of immunity to influenza over time, focussing on antibody responses to the very first and second infections from a cohort of children studied in Managua, Nicaragua.

In contrast to adults with diverse exposure histories, antibody responses to first infections occurring within the same season were mostly remarkably similar in both pattern and magnitude but there were interesting exceptions. Season 2 responses for example were dichotomous with approximately half of the children responding strongly against strains from the Wisconsin 2005 antigenic cluster despite the fact that all circulating strains isolated in that season were from the Perth 2009 antigenic cluster. A potential explanation was that immune responses were shaped by underlying immunity from a previous infection that had not resulted in serum antibody titres detectable via HI.

Further regarding the potential for non-detectable responses to first infections, several children had PCR-confirmed H3N2 infections that resulted in no detectable HI titres against any of the strains tested against a serum sample taken post-infection. In one case, a sample taken 32 days after a first infection showed development of significant HI titres, but these fell again to undetectable levels in later samples, indicating that serum antibody may not always be reflective of the underlying immune repertoire.

Interestingly, for second infections, although post-infection HI titres against strains antigenically similar to the infecting virus were comparable to those after first infections, measurements of broader cross-reactivity showed that responses were indeed skewed overall towards prior immunity. The result was that post-infection titres after a second infection were actually smaller against future antigenic variants than those generated following first infections in the same season, providing evidence for a potential cost in terms of future protection.

Long-term antibody titres following a first infection were lower than those typically seen in adults against strains encountered early in life, but were already boosted to these later levels after only a second infection. Analysis of data from previous chapters showed that the those individuals first infected with a subtype other than H3N2 did not go on to show higher than average titres against early H3N2 strains, suggesting that the pattern of high titres against strains encountered early in life is a special feature of first infection responses and not a result of continual boosting throughout life.

5.2 Introduction

In previous chapters, serological responses to infection and vaccination have been examined in some detail, but although the antigenic nature of viruses causing infection or vaccination in these cases were known or could be inferred, all individuals began with some initial immunity as a result of prior exposures occurring before the period of study. From these later responses we can infer in broad terms how antibody repertoires are shaped over time but more detailed analysis is difficult when a large part of the infection history of these subjects remains unknown. Similarly, where heterogeneity is seen in the antibody responses to antigenically similar (or in the case of vaccination, identical) viruses it is impossible to conclude how much should be attributed to differing prior immunity shaping the overall antibody response, and how much is attributable to other factors such as genetic differences between individuals or stochasticity in the nature of antibody repertoire generated in response to antigen exposure [Duffy et al., 2012; Srivastava et al., 2009].

In particular, the importance of immunological exposures occurring early in life and how this can potentially shape a lifetime of future responses has long been appreciated but, due to the difficulty of obtaining pre and post infection sera from children at such a young age, has mostly been inferred through the use of controlled exposures in animal models or inference from adult serology [Davenport and Hennessy, 1956; Davenport et al., 1953; Jensen et al., 1956; St Groth and Webster, 1966a; St Groth and Webster, 1966b]. For example, as discussed in chapter 1, a key tenet of the long-held dogma of original antigenic sin [Francis, 1960] and it's later formulations [Lessler et al., 2012] is that repeated boosting of antibodies produced in response to the first strain encountered explains the phenomenon that antibody titres measured later in life are typically highest on average against strains circulating at around the time most individuals receive their first influenza infection. Although hinted at in some longitudinal studies [Miller et al., 2013], this effect has never been observed directly in humans and more detailed previous analysis of long-term serological responses sheds some doubt on the extent to which this actually occurs - a competing hypothesis being that due to factors such as a more robust primary antibody response as a result of greater viral replication, first infection antibody responses are the largest to begin with and this is simply maintained over time (see chapter 3, section 3.5.2).

Better understanding the processes that guide initial formation of the serological antibody repertoire is also highly relevant for vaccination strategies in children where there are fears that stimulation of narrower, misdirected or inadequate antibody responses through vaccination may lead to poor protection or even be damaging in the longer term, compared to antibody responses produced in response to natural infections [Cobey and Hensley, 2017; Gostic et al., 2016]. Is there, for example, any evidence that future responses can become diminished by the presence of prior immu-

nity directed against certain antigens as has been previously suggested with regard to some previous studies [Linderman et al., 2014; Smith et al., 1999].

In this chapter I focus on analysis of serum samples taken annually from a cohort of children participating in a range of studies coordinated by the Sustainable Sciences Institute in Managua, Nicaragua [Gordon et al., 2015]. By selecting samples from young children, it was possible in many cases to capture the dynamics of the antibody response to the very first influenza infection, examining the variability of antibody responses in the absence of prior immunity. In some cases, second infections were also captured, allowing a comparison of the effect pre-existing immunity has upon the antibody response compared to a first infection. Detailed analysis of such serum samples in children represents a rare opportunity to gain insight into a crucial phase of immunological development.

5.3 Methods

5.3.1 *Sample selection*

From the wider cohort of over 6000 individuals monitored in trials run by the Sustainable Sciences Institute in Managua, Nicaragua, samples were selected for detailed antibody-landscapes-based analysis from 94 children that met criteria designed to maximise chances of capturing immunity before and after a first H3N2 influenza infection while avoiding detectable maternal immunity:

- A first serum sample was taken after the age of 6 months and before the age of two.
- A serum sample was taken before and after a PCR-confirmed H3N2 influenza infection.
- The first recorded infection was against an H3N2 influenza subtype.

The children selected came from different households in Managua and were not closely related. Their year of birth ranged from 2003 to 2012.

5.3.2 *HI titrations*

HI measurements were performed by collaborators at the Victorian Infectious Disease Reference Laboratory in Melbourne following the protocol described in appendix B. All available sera from the subset of 94 children chosen for antibody landscape analysis were titrated against 32 H3N2 strains isolated between 1982 and 2011 and highlighted in figure 5.1. The full list of strains used is given in appendix A.

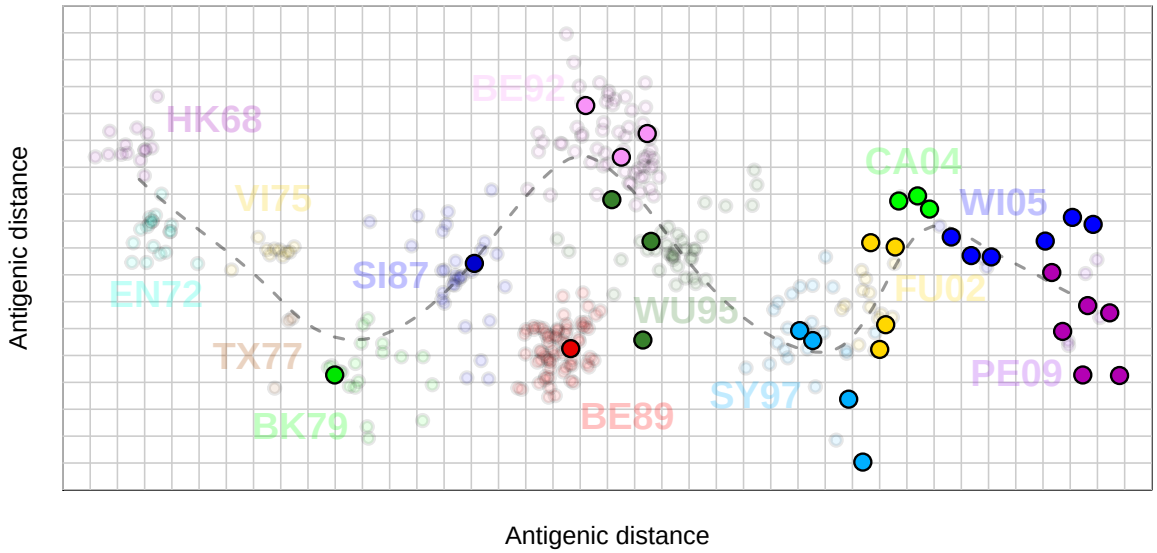


Figure 5.1: Antigenic map highlighting the H3N2 strains used in the Managua cohort study. The antigenic map and summary path defined in chapter 2 are shown, different H3N2 strains are represented by circles colour-coded by antigenic cluster. The 32 H3N2 strains that were used as the test set for samples taken from children in the Managua cohort are highlighted. Full lists of the strains used are given in appendix A.

5.3.3 *Monitoring of influenza infections*

The subset of 94 children selected were enrolled in different, often overlapping, study cohorts at different times during their lives, some including active monitoring of influenza-like illness and some including only serum sample collection. During periods of active monitoring, participants were encouraged to report any onset of symptoms so that nose and throat swabs could be taken (typically taken the next day) and stored for later PCR analysis and subtyping [Gordon et al., 2015]. Where changes in serological immunity over time are later examined in the context of recorded infections, periods of active monitoring are indicated.

5.3.4 *Analysis of sequence data*

As part of general surveillance the genomes of 245 viruses isolated from the larger cohort of study individuals were analysed. The nucleotide sequences from the HA gene segment were aligned using the "seqaln" function from the R package "bio3d"[B.J. et al., 2006] and trimmed to the coding region. The trimmed and aligned nucleotide sequences were then translated to amino acid sequences using the "translate" function from the R package "Biostrings"[Pagas et al., 2016]. Amino acid sequence identity was determined using the "seqidentity" function from the bio3d package.

Phylogenetic analysis was performed using the R packages "ape"[Paradis et al., 2004] and "phangorn"[Schliep et al., 2017; Schliep, 2011]. The tree was constructed using the functions "dist.ml" and "NJ", and topology and edge length were optimised under the GTR evolutionary model using the functions "pml" and "optim.pml".

5.3.5 *Construction of antibody landscapes*

Antibody landscapes were constructed using the maximum-likelihood approach as described in chapter 1, using a fitting bandwidth of 10 antigenic units assuming a minimum titre for non-detectable measurements of -4. Diagnostic plots for the antibody landscape fits are shown in appendix C.4.

5.4 Results

5.4.1 *Circulation of influenza within the cohort*

Figure 5.2 shows the total number of PCR positive influenza cases detected each week from subjects in the cohort between 2007 and 2016. The timing of peak influenza cases varied each year as did the number of positive cases, although the changing size of the total cohort over time should be taken into account. The infections that occurred in the antibody landscapes subset of 94 children were primarily from 5 approximate seasons of H3N2 circulation in 2007, 2010, 2011, 2012 and 2013, numbered in the figure.

5.4.2 *Genetic analysis of infecting antigens*

Figure 5.3 shows the results of a comparison of the haemagglutinin amino acid sequences of the 245 viruses that were isolated from the cohort and sequenced. The isolation dates again cluster into 5 approximate H3N2 seasons corresponding to the seasons identified in figure 5.2. Regarding HA amino acid sequence identity, the viruses cluster along largely the same lines as the seasons in which they circulated, although the differences between viruses in seasons 3 and 4 are smaller than between the other seasons. Unexpectedly, the viruses isolated in season 5 were more similar to season 2 viruses than to seasons 3 or 4, this is also reflected in the phylogenetic tree of samples shown in figure 5.4 but it is unclear why this should be the case. In general however there was very high amino acid sequence identity within a given season (in particular season 2) and many viruses with identical HA amino acid sequences were isolated.

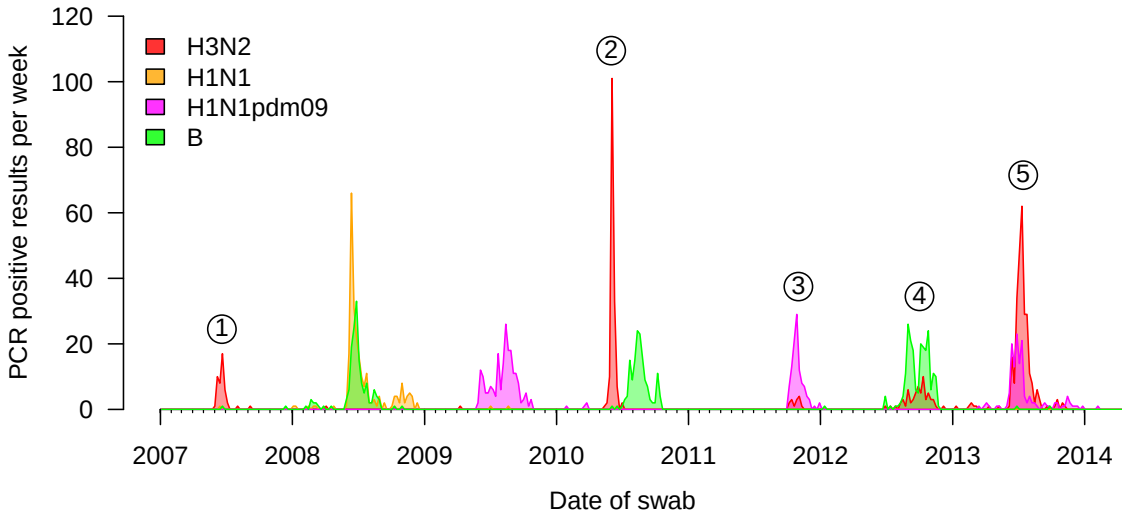


Figure 5.2: Circulation of influenza subtypes in the Managua cohort between 2007 and 2016. The number of swabs taken from individuals in study cohorts in Managua run by the Sustainable Sciences Institute that were PCR positive for different influenza subtypes per week. Note that since the total number of individuals monitored for infection in the cohorts varied over time, the numbers are not expected to represent the magnitude of different influenza epidemics in Managua, only the timing. Circled numbers mark 5 different peaks of H3N2 infection referred back to in later figures.

5.4.3 *Estimation of virus antigenic characteristics*

The available genetic data was additionally used to try and independently infer information about the antigenic characteristics of the H3N2 viruses circulating at different times in the cohort.

A first approach was to perform a nucleotide blast comparing the sequence similarity between the 245 Managua isolates and 81 viruses in the antigenic base-map for which sequences were also available. Each of the 245 Managua strains was then mapped to it's closest genetic match. Figure 5.5 shows the results of this analysis, where the frequency of different base-map antigens appearing as "best matches" to a strain from the cohort is shown over time. As can be seen all viruses in season 1 had a Wisconsin 2005-type strain as a best match (predominantly A/Perth/27/2007), while all viruses from subsequent seasons had a best match with a Perth 2009-type strain (A/Victoria/208/2009 in seasons 2, 3 and 4 and A/Victoria/361/2011 in season 5).

Since genetic change does not necessarily correlate closely with antigenic differences and small genetic differences can have a large antigenic effect [Koel et al., 2013; Smith et al., 2004], genetic data from the isolated Managua viruses was also analysed for specific amino acid mutations associated with antigenic change. Preliminary data (unpublished) indicates that the Perth 2009 cluster transition was defined by a change

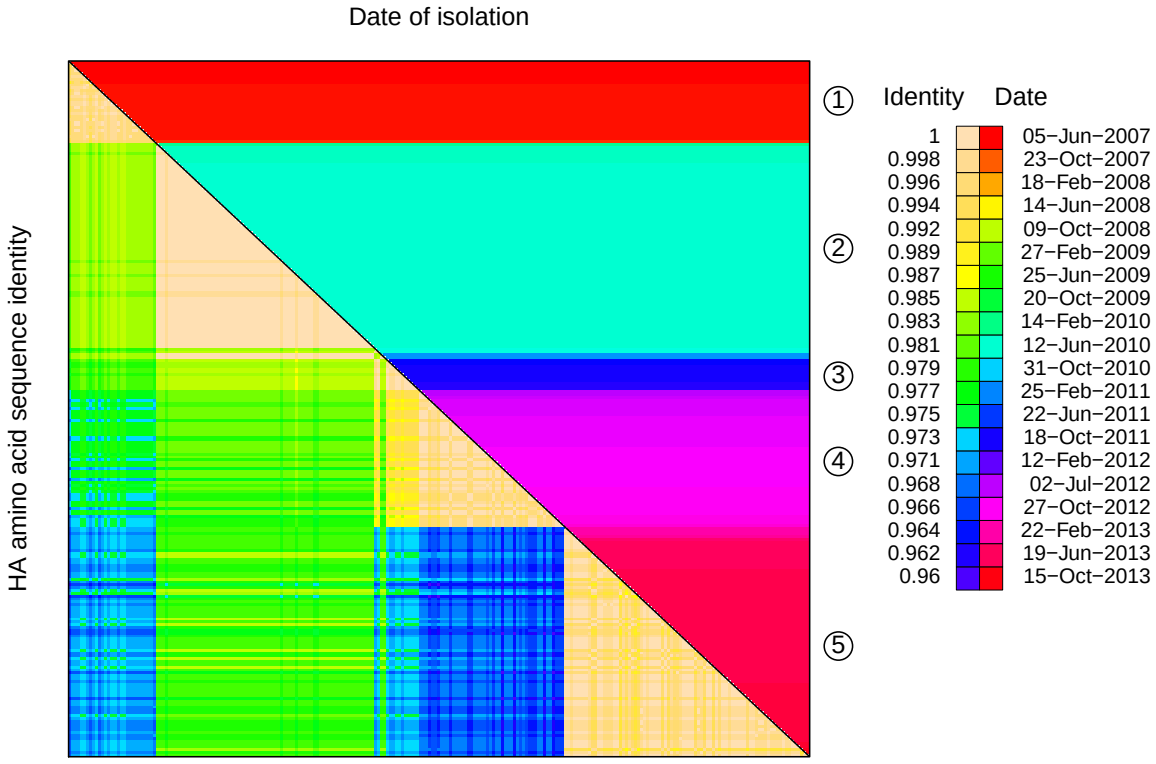


Figure 5.3: Genetic similarity of H3N2 viruses circulating in Managua by season. 245 H3N2 viruses isolated in Managua are compared pairwise, ordered by the date of isolation. The bottom left half of the matrix is coloured coded by the HA amino-acid sequence identity of each corresponding virus pair. The top right panel simply colours each row by the isolation date of the virus to the left of the row to illustrate seasonality. The circled numbers correspond back to the 5 H3N2 seasons labelled in figure 5.2.

from lysine to asparagine at position 158 and an asparagine to a lysine at position 189. In keeping with the results from the blast data, all viruses isolated in 2007 contained a lysine at position 158 and asparagine at position 189, while all viruses isolated in the subsequent seasons contained the switches at these positions in keeping with a transition to the Perth 2009 antigenic cluster.

On balance of the evidence from both overall genetic sequence similarity and specific amino-acid mutations relevant to antigenic change it seems likely that season 1 was a season dominated by viruses from the Wisconsin 2005 antigenic cluster while all later seasons were defined by the circulation of viruses from the Perth 2009 cluster.

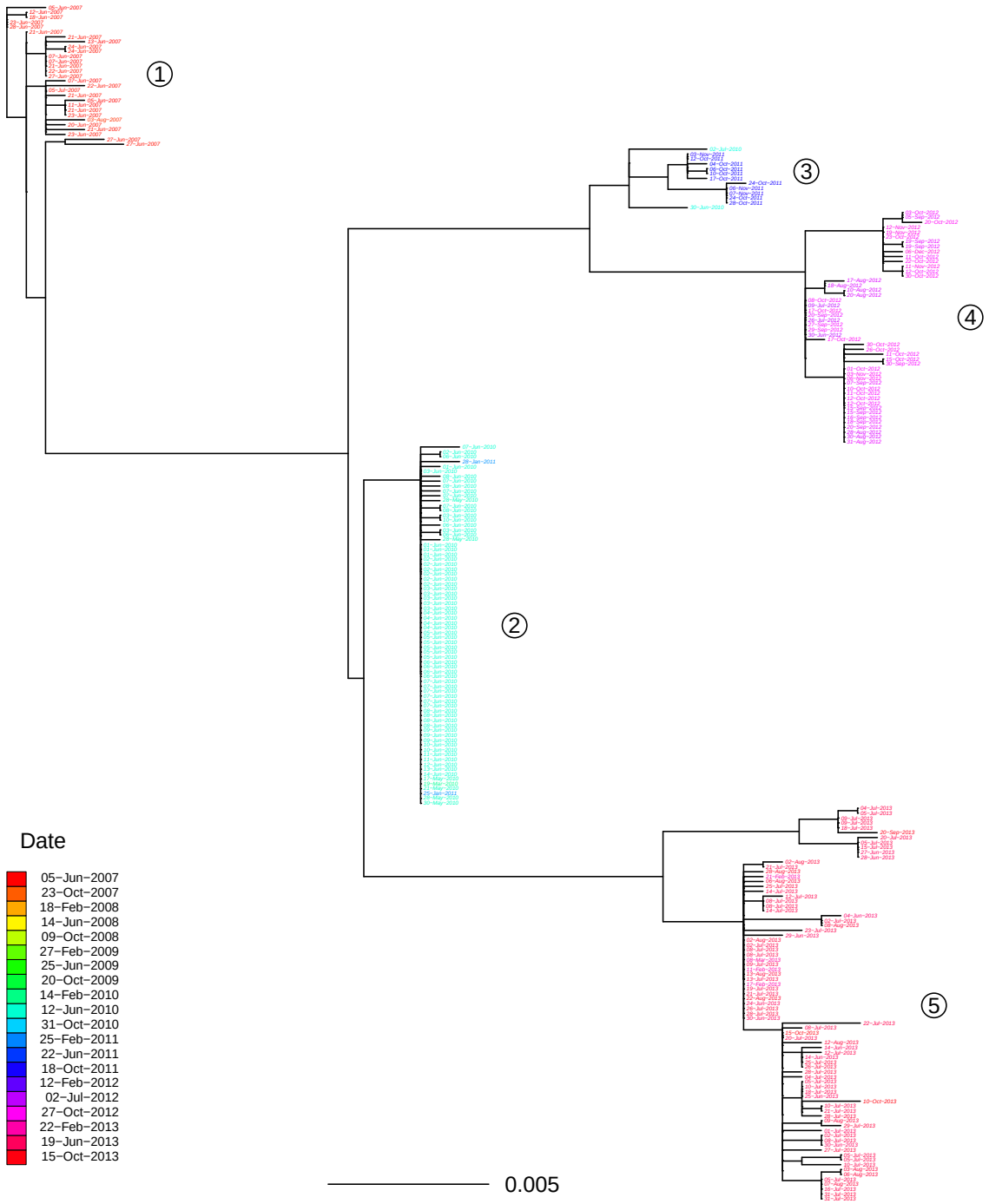


Figure 5.4: HA sequence phylogenetic tree of H3N2 viruses circulating in Managua. Tips are labelled with the isolation date of the corresponding strain and color-coded according to the legend shown. The strains form clades that closely correspond to the 5 H3N2 seasons labelled in figure 5.2, labelled here again with circled numbers. The scale bar indicates a distance of 0.005 under the GTR evolutionary model.

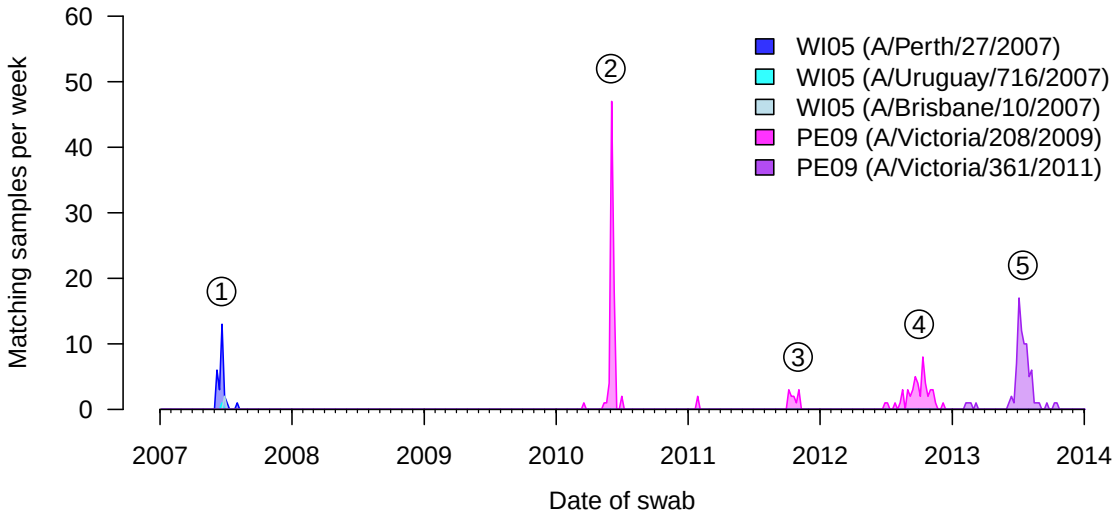


Figure 5.5: Best-matching base-map strains to viruses circulating in Managua. The number of times 81 different reference strains (for which genetic sequences were available) from the antigenic base-map were found to have the closest HA amino acid sequence identity to a virus that was isolated from a swab taken from the Managua cohorts. Only 5 strains appeared as a best match to isolated viruses (shown in the legend), and there was a clear antigenic distinction, with strains isolated in 2007 most closely matching base-map strains from the Wisconsin 2005 antigenic cluster and later strains most closely matching base-map strains from the Perth 2009 antigenic cluster. The circled numbers correspond back to the 5 H3N2 seasons labelled in figure 5.2.

5.4.4 *Visualisation of infection responses*

The antibody landscapes created for each serum sample were combined and compared to generate timelines to visualise the development of antibody titres over time for each child - an example is shown in figure 5.6.

As before with the antibody landscapes analysed from the Hà Nam cohort in chapter 2, the landscape for a given sample is marked by a black line and where possible is shown compared to the landscape from the previous sample. Red regions mark an increase in antibody reactivity and beige regions mark a decrease. The locations along the antigenic summary path of the 32 H3N2 strains titrated are again marked along the x-axis, scaled negatively by distance from it. Due to the greater complexity and variability of the sampling and influenza monitoring regimens applied for different individuals, an individual timeline is additionally shown above the landscapes. A larger black dot to the left marks the date of birth for a particular individual and the dates of serum sample collection are marked by numbered circles along it, corresponding to the antibody landscapes below. The dates of nose or throat swabs that were PCR-positive for an influenza infection are additionally marked as circles on the timeline, colour-coded by subtype as shown in the legend. Periods where the subject

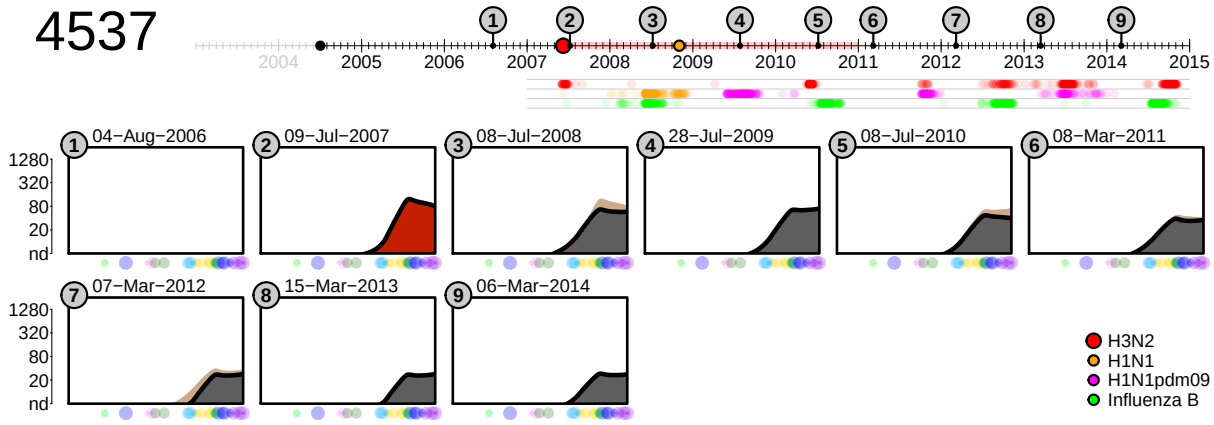


Figure 5.6: An example timeline of changes in antibody reactivity. At the top-left of the plot the subject ID is given for the child who's antibody responses are visualised. The top panel then shows a timeline of the child's life, with the date of birth marked as a large black circle to the left of the plot. Dates of serum sample collection are marked with a numbered circle and correspond to the panels below showing 2D-antibody landscape summaries constructed from HI measurements made against each serum sample. Also shown on the time line are coloured dots indicating dates of PCR-confirmed infections with different subtypes, as shown in the legend bottom left, red = H3N2, orange = H1N1, magenta = H1N1pdm09 (2009 pandemic) and green = influenza B. Red-shaded regions on the timeline mark the times at which the child was involved in a study cohort where they were actively monitored for influenza infection. Where a child was recorded as having received a vaccination, this is marked with a "V" on the timeline. Below the timeline, to give an idea of how many different influenza seasons a child has lived through, coloured circles mark the dates at which infections from the different subtypes were recorded in the wider cohort of all children studied in Managua. As described in chapter 2, the black line in each of the bottom panels represents the landscape height for each position on the antigenic summary path that was taken through the H3N2 antigenic clusters (and shown again on the antigenic map in figure 5.1). As in chapter 3, the antibody landscape measured in the first sample taken is shown shaded in grey (if indeed visible, which it is not in this plot), in subsequent samples red shading indicates increases, and beige decreases, compared with antibody reactivity measured in the previous sample. As in the previous chapters, dots along the x axis indicate the antigenic positions along the summary path of the subset of viruses used to generate each landscape, scaled negatively by distance from it. The sample date of each serum sample is also given above each of the lower panels.

was part of active surveillance for influenza infections are shaded in faint red along the timeline. Serum samples taken outside of periods of active influenza surveillance were generally collected because the individual was not yet enrolled in or had dropped out of an influenza study cohort but were part of a separate cohort monitoring for dengue infection, as is the case here.

The full set of timelines for each child are reproduced in appendix C.3. Interactive 3D versions of each timeline are also available online, with links provided in appendix C.5.

5.4.5 *First infection responses*

Of the 94 children in the antibody landscapes subset, 69 likely first infections were identified based on the criteria that a sample taken before the first PCR confirmed

H3N2 infection had detectable HI titres against fewer than 3 strains of the 32 strains tested (therefore accounting for small amounts of background HI measurement noise). The infections recorded from participant numbers 7555, 7543 and 6679 were also counted as likely first infections since the antibody reactivity present in the samples preceding the infections in these cases was characteristic of maternal immunity (see section 5.4.6), making a total of 72.

Responses by season

In general, first infection responses within a given season were remarkably stereotyped both in terms of the general pattern of reactivity and the magnitude of responses generated. This was despite the fact that the children were not related and came from different parts of Managua. Responses differed notably between seasons in keeping with the genetic differences found between each season.

Figure 5.7 shows the post-infection landscapes for the 68 cases of first infection where a sample was available at least 6 months post the infection date, split by season. Only samples taken at least 180 days post infection were chosen to allow for some short-term decay and compare instead the longer-term response to infection.

In season 1, antibody responses were directed primarily against California 2004, Wisconsin 2005 and Perth 2009-type strains, in keeping with the genetic analysis suggesting circulation of viruses with Wisconsin 2005-type characteristics in this season and allowing for some cross-reactivity against the previous and next antigenic cluster. Accordingly, children infected in seasons 3, 4 and 5 had responses that favoured Perth 2009-type viruses, again in keeping with the genetic analysis of viruses in these seasons.

In contrast to the other seasons, responses in season 2 were quite dichotomous with 7 of the first infection responses clearly highest against Perth 2009-type strains and 9 showing a mixed response against California 2004, Wisconsin 2005 and Perth 2009-type viruses. Although this dichotomy was almost confined mostly to season 2 responses, it should be noted that one child infected in season 4 and one child in season 5 also showed strong responses against California 2004 and Wisconsin 2005 in addition to Perth 2009-type viruses as can also be seen in the respective panels for season 4 and season 5 in figure 5.7.

Response similarity against genetic similarity

Of the 72 likely first infections identified, 19 were also part of the subset of samples sent for sequencing which allowed a small scale analysis of how antibody response similarity and genetic similarity of the causative virus compared. Figure 5.8 shows the result of a comparison of the sequence identity of the HA amino acid sequence of each infecting virus and the root-mean-squared difference between HI titres in the

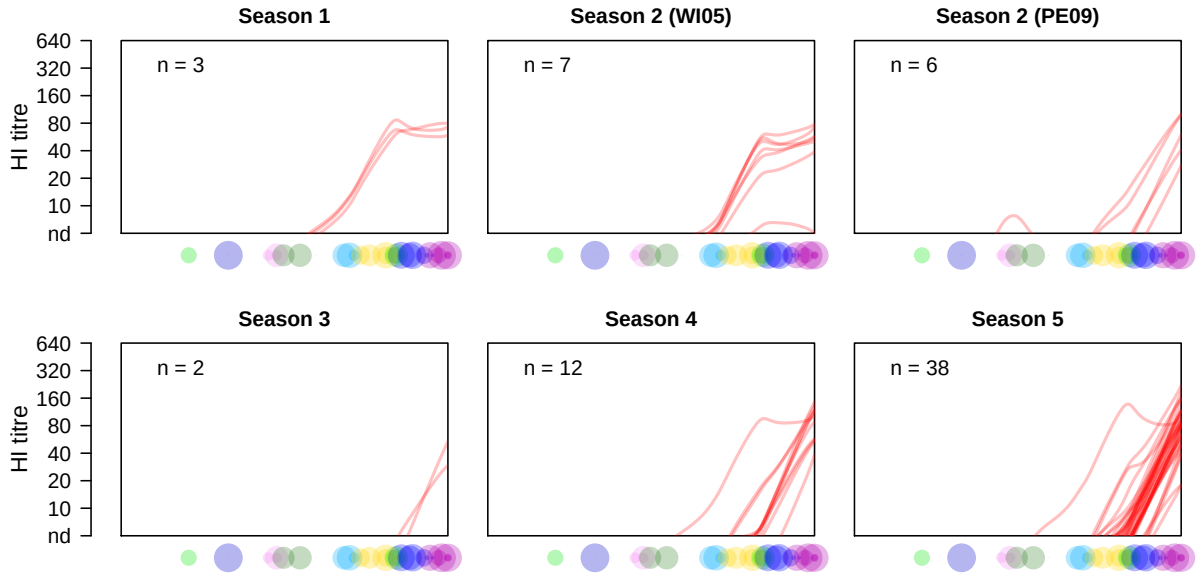


Figure 5.7: First infection responses shown by season. Post-infection landscapes for the 68 cases of possible first infection where a sample was available at least 6 months post the infection date, split by season. Each red line represents the landscape height for each individual for positions along the antigenic summary path defined in chapter 2 and shown again in figure 5.1. Responses in season 2 have been split by those showing joint responses against Wisconsin 2005 and Perth 2009 type strains - Season 2 (WI05) - and those showing responses predominantly against Perth 2009-type strains - Season 2 (PE09). As with previous plots, dots along the x axis indicate the antigenic positions along the summary path of the subset of viruses used to generate the landscapes, scaled negatively by distance from it.

associated first-infection responses. To avoid the distorting effect of post-infection samples taken at different phases of the initial dynamic antibody response to infection, samples were only included if taken more than 6 months after the first infection.

Responses cluster primarily along the lines already discussed, with either a very dominant response against Perth 2009-type viruses or a mixed response including the Wisconsin 2005-type strains. As noted before, genetically similar viruses typically seem associated with very similar antibody responses but there are also interesting exceptions. Of the 6 infections occurring in season 2, 4 children were from the group with a mixed response including both Wisconsin 2005 and Perth 2009 viruses while one child is from the group responding primarily against Perth 2009-type strains and these examples are also shown separately as part of the expanded section of figure 5.8. Despite the clear differences in response seen in sample 5906_01 compared to the others, the HA amino acid sequence of the isolated virus was either identical, or almost identical, to viruses isolated during the other infections (even the small differences to the virus from the infection associated with sample 5871_02 are not in parts of HA typically considered antigenic sites). A similar situation is seen with the last response in season 5 from subject number 7228. The response seen in this case is again focussed more around Wisconsin 2005-type strains, despite the fact that

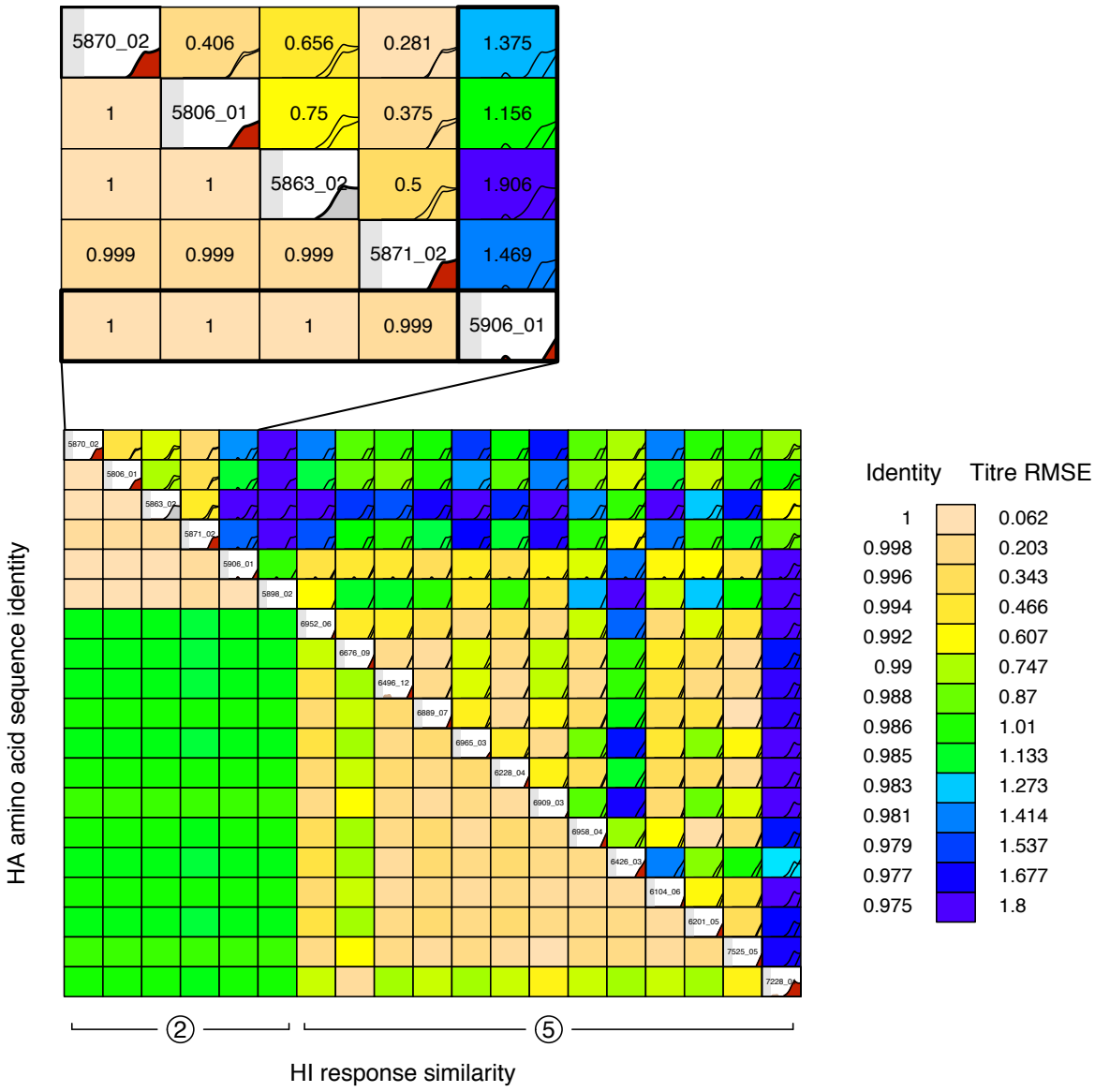


Figure 5.8: Genetic similarity of infecting virus compared against antibody response similarity for first infections. Comparison of the similarity of the RMSE of post-infection antibody titres against the HA amino-acid identity of the infecting virus for the 6 possible first infections in season 2 for which genetic information on the isolated strain was also available. Plots along the diagonal show the summary antibody landscape constructed from the next sample taken at least 6-months post infection (note that the third plot along is light grey rather than red to indicate that no pre-infection sample was available to compare post-infection responses to in this case). As for figure 5.3, the bottom left panels are coloured by the HA amino-acid sequence identity of the infecting viruses isolated. The top right panels now show pair-wise comparisons of the post-infection antibody landscapes, with each panel also coloured by the RMSE of differences between post-infection titres on the log scale, treating non-detectable titres as -1. The smaller plot above shows the results and data for samples taken post infection in season 2, excluding sample 5898_02 (which showed no detectable response to infection). Sample number 5906_01 is highlighted, a case which shows a very different antibody response to others in that season, even though the isolated virus has 100% HA amino acid sequence homology with 3 of the 4 other isolates.

subject 6676, infected with a strain with 99.9% HA amino acid sequence identity, shows a response focussed primarily on Perth 2009-type strains in-keeping with all other first infection responses from that season (again even the small differences in the HA amino acid sequence are outside of typical antigenic sites).

5.4.6 *Unusual patterns of antibody reactivity*

Some patterns of immunity were not readily attributable to the infection history of the children being monitored, in particular, detectable antibody titres against older antigenic variants that circulated long before a child was born.

Likely maternal antibodies

In some cases, despite samples taken no earlier than 6 months of age, these patterns of antibody reactivity were highly characteristic of what would be expected from maternal immunity. Two examples are shown in figure 5.9. In each of the suspected cases of maternal immunity, detectable titres are present against a range of the older strains in the very first sample, before any H3N2 infection had been recorded. Although 6 months is the point at which it has been estimated that HI titres due to maternal antibodies have mostly decayed below detectable levels [Bodewes et al., 2011], it is plausible that at the time these samples were taken (only 7-8 months after birth) detectable antibody may still remain, especially considering the greater sensitivity of the antibody landscapes approach compared to the analysis of single titrations. As would be expected of reactivity due to maternal antibodies, these titres against older strains decayed to non-detectable levels in the subsequent samples until the child went on to have a detectable H3N2 infection.

Titres against older strains

Other cases of reactivity against older viruses did not appear to be attributable to maternal antibodies and were harder to account for. In particular, while many samples returned non-detectable titres against all 32 strains tested, in 10 samples across 8 individuals, only the strain A/Netherlands/620/1989 returned a detectable titre. These titres were against samples taken before a first recorded influenza infection and, being a strain from the Sichuan 1987 antigenic cluster, not a strain that the children (all born after 2007) would be expected to have any reactivity against. In most cases the titre was non-detectable in other samples but in two cases it remained detectable in a following sample. Experimental error seems a likely explanation but in most cases the titre remained unchanged upon repeat testing of the sample so it remains unclear why this strain and no others should show such an unusual pattern of reactivity. The first two individuals in figure 5.10 show examples from the 8 individuals in which this pattern occurred.

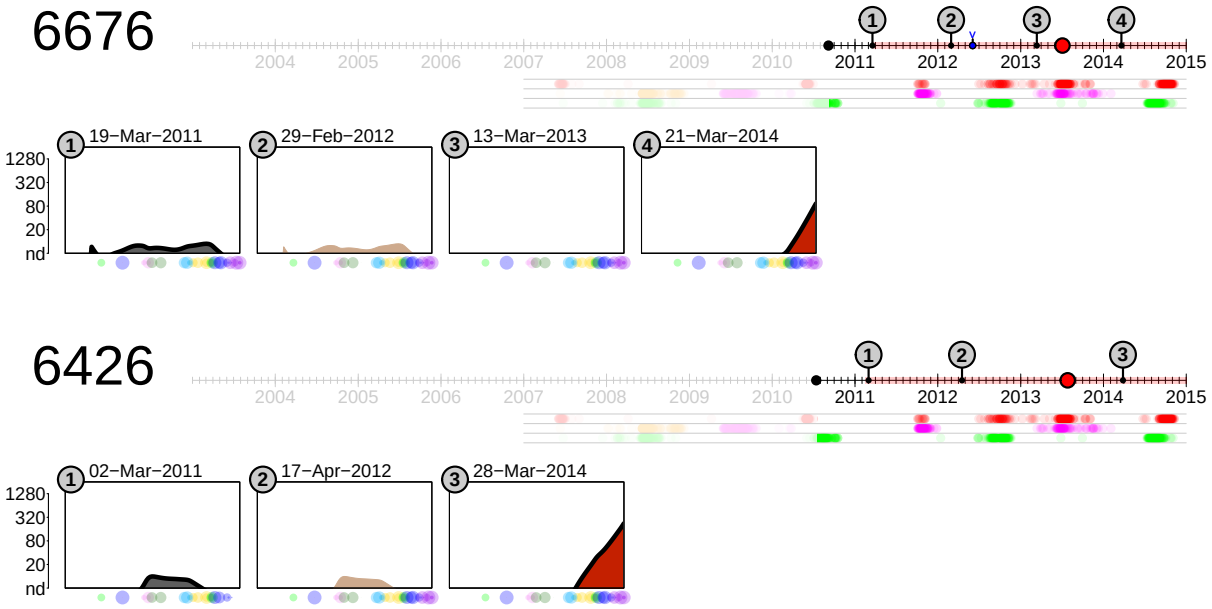


Figure 5.9: Timelines showing examples of likely maternal immunity. Two individuals where patterns of antibody reactivity in the first sample taken are characteristic of maternal immunity. Full caption as given for figure 5.6.

There was also an isolated example of a particularly unusual pattern of antibody response to a recorded H3N2 infection from subject 6301 and also shown in figure 5.10. Here it can be seen that following a PCR-confirmed H3N2 infection in 2011, titres increased not against contemporary strains but instead against older antigenic variants, decaying in the next sample before returning to roughly pre-infection levels in the final sample taken. Again, the same pattern of response was seen upon repeat titration of these samples. It is possible that the infecting virus was indeed an older variant perhaps surviving in an animal reservoir but this is hard to ascertain without genetic information on this particular strain.

Transient and non-detectable responses to infection

Further notable exceptions to the otherwise stereotyped nature of immune responses to first infections were children with patterns of antibody reactivity indicative of either a non-detectable or transient response to infection. These could be broken down as follows:

- 3 children with no detectable HI titres either before or after an H3N2 infection.
- 1 child with detectable HI titres before but not after an H3N2 infection.

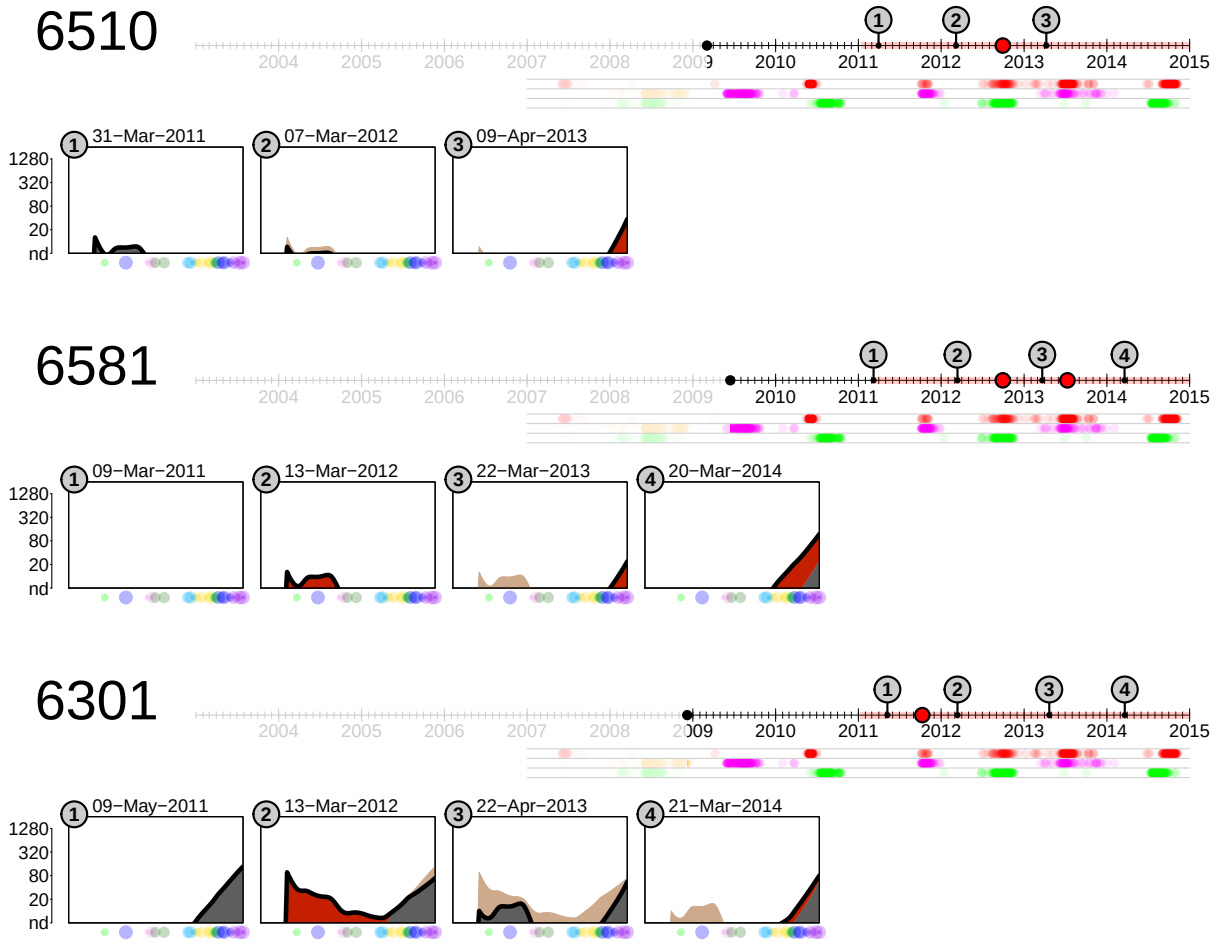


Figure 5.10: Timelines showing unusual patterns of antibody reactivity. Three individuals where patterns of antibody reactivity against the older strains titrated cannot be easily accounted for. Full caption as given for figure 5.6.

- 3 children with detectable HI titres that become completely undetectable in later samples.
- 2 children with detectable HI titres but no detectable increase after an H3N2 infection.
- 1 child with a detectable HI response in the sample following an H3N2 infection but no detectable titres in later samples.

Figure 5.11 shows examples of all of these categories of response and in most it is unclear whether the patterns represent methodological error or a genuine immunological phenomenon. In many of these cases for example, the unusual pattern could be explained if one of the samples were actually a switch with, or accidental repeat of, another sample in the series and an important test would be to re-aliquot samples

from the original collection tubes and repeat the measurements. In the case of participant 4524 it seems likely that the infection date of the H3N2 infection has been mis-recorded although upon further checking of participant records, no indication was found that this had occurred, but again such errors may be hard to detect depending upon the point at which they occurred.

The response pattern of participant 5898 is perhaps the most difficult to account for with methodological errors, since multiple different samples returned non-detectable titres against all 32 strains tested following an H3N2 infection indicated both by records of nasal and throat swabs PCR-positive for H3N2 and an associated transient antibody response in the subsequent sample. Interestingly, the sample after the infection in which HI titres are initially detectable was taken only 32 days after the onset of symptoms, within the period of the "transient response" identified following infections in chapter 3. In all other cases where no response at all was witnessed, for example participant 6428, the sample was taken relatively long-term post-infection, raising the possibility that these children may have all had a short-term detectable transient response to infection that was for some reason not maintained as longer-term detectable titres. Of note is that the final sample from participant 5898 once again shows detectable antibody titres despite the fact that it was only taken 248 days after a second recorded H3N2 infection, indicating that in this case, a long-term antibody response was sustained.

5.4.7 *Analysis of second infection responses*

22 participants in the Managua cohort had a PCR-confirmed infection that followed a sample in which detectable HI titres were already measurable (excluding individuals where pre-infection titres were thought to be characteristic of measurement error or maternal immunity - see section 5.4.5). For most individuals it was impossible to determine how many infections had occurred prior to the recorded infection but given the ages of the children and the infection rates witnessed in general in the cohort it seems reasonable to assume that in most, if not all, of these cases, only one infection had occurred. This would make most of the responses that occurred in the presence of prior immunity likely second infection responses. There were in addition, two individuals with patterns highly suggestive of either a sample swap (participant number 7308) or mis-recorded infection dates (participant number 4524) and these were discounted.

In general, the likely second infection responses in the Managua cohort followed the same patterns as the infection responses analysed in the Hà Nam cohort in chapter 2. Samples taken soon after infection showed strong boosting of titres which then decayed in later samples to leave a new stable landscape. The antigenic extent of the initial boosting response was more limited than that seen in the Hà Nam cohort,

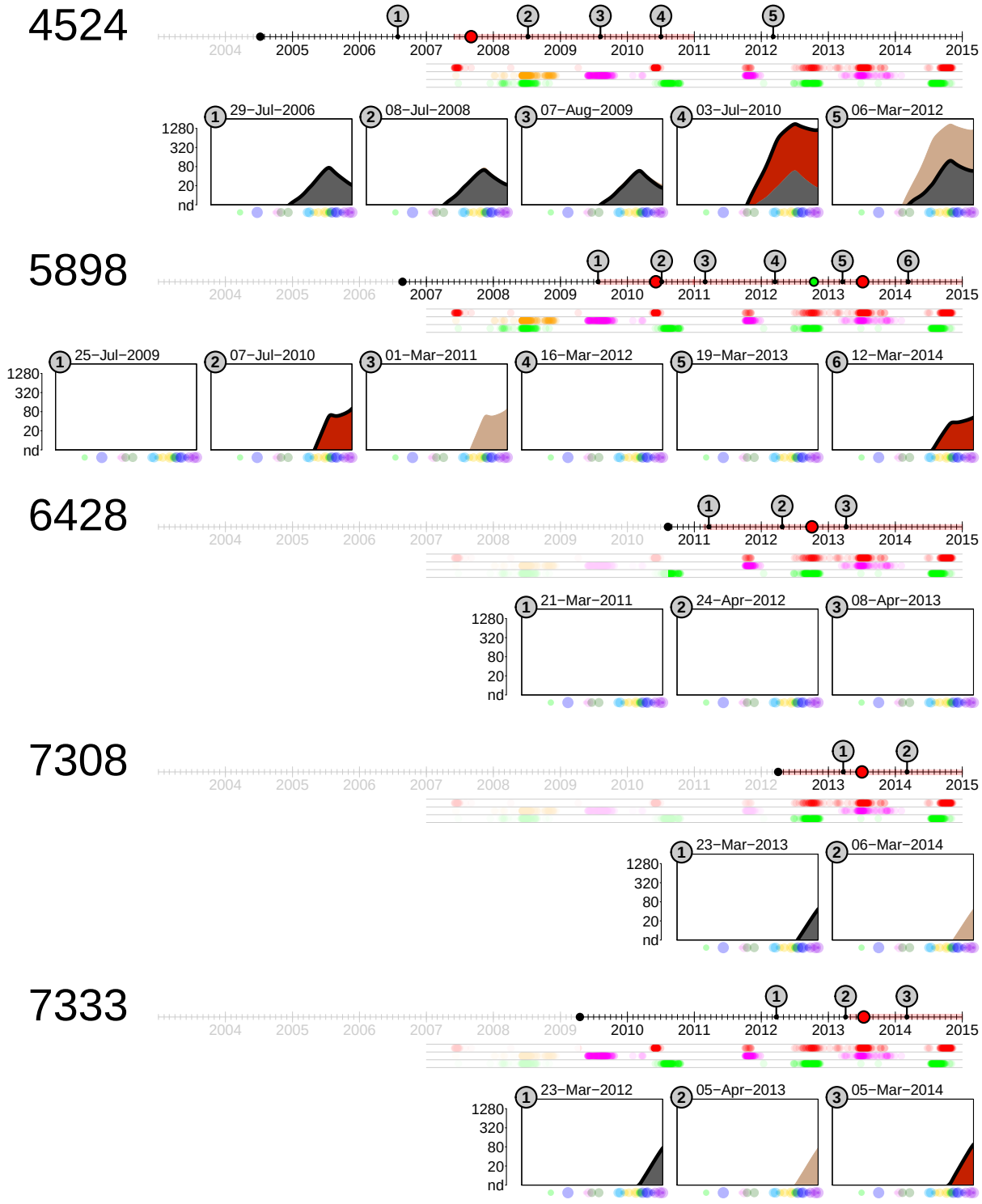


Figure 5.11: Timelines showing a lack of antibody response to infection. Five individuals where there was no significant change in antibody reactivity post-infection. Full caption as given for figure 5.6.

being again largely limited to antigenic regions against which there was already prior immunity (accounting for some cross-reactivity to preceding antigenic clusters).

First compared to second responses

Since such antigenically and genetically similar strains had circulated in each of the H3N2 seasons in Managua, there was the opportunity to examine the effect of prior immunity on antibody responses by comparing second infection responses to first infection responses that occurred in the same season.

Figure 5.12 compares the antibody titres from samples that were associated with first and second infection responses split by influenza season (due to their dichotomous nature, first infection responses in season 2 are split, as shown in figure 5.7, by those that include a substantial response against Wisconsin 2005-type viruses - labelled WI05 - and those that included a response predominantly against Perth 2009-type viruses - labelled PE09). Antibody reactivity is compared in terms of both individual and average antibody landscapes, and the average HI titre to strains from each antigenic cluster. Results are shown that correspond either to samples taken prior to a likely second infection (light blue), the next sample taken at least 6 months post the same likely second infections (dark blue), or samples taken at least 6 months post a likely first infection (red). Samples taken prior to a likely first infection are not shown since these had no measurable antibody reactivity against any of the strains tested (or reactivity attributable to maternal immunity).

As might be expected given previous findings in the literature, the results in figure 5.12 clearly show how prior immunity in the second infection responses appear to have skewed the antibody responses to include significant boosting of responses to previously encountered strains compared to the first infections. As found in previous studies [Davenport and Hennessy, 1956; Francis et al., 1947; St Groth and Webster, 1966b], this boosting of antibody titres against older strains seems to have come at no cost to the long-term antibody response to the infecting strain - post-infection antibody titres to viruses from the same antigenic cluster as the likely infecting virus (highlighted with a box on the rightmost plots) were similar for both first and second infections. In seasons 4 and 5, where titres post second infection do begin to be slightly larger against Perth 2009-type strains than those post-first infection, it seems that a likely explanation is that the selection of Perth 2009-type viruses used in the test set no longer provide such a good antigenic match to circulating strains and, had more recent viruses been included, similar results to the previous seasons would have been seen.

Strikingly, in season 1 - the one season where responses against future antigenic variants can be compared - post-infection titres against strains from the next antigenic cluster (Perth 2009) were actually lower in second infections compared to first. This

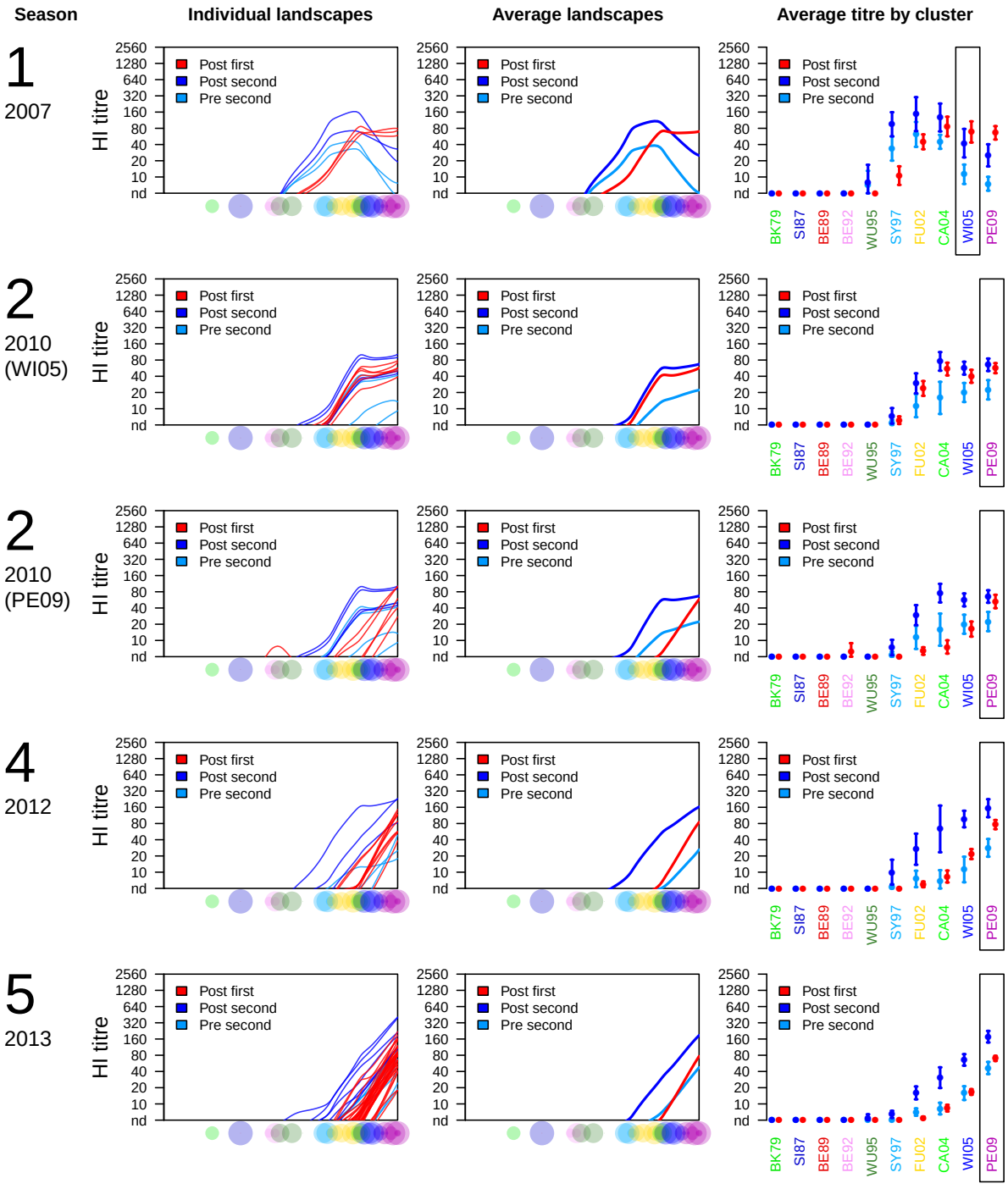


Figure 5.12: First compared to second infections split by season. Responses are shown by season (with year) as shown in the left column (with season 2 responses split as in figure 5.7). The second and third columns show individual and average antibody landscapes along the summary antigenic path described in chapter 2 and shown again in figure 5.1 are then shown for samples taken prior-to second infections and at least 6 months post first and second infections. The final column shows mean HI titre responses to the titrated H3N2 strains split by antigenic cluster. Error bars show the 95% confidence intervals for the mean. Boxes highlight the antigenic cluster of the likely-infecting strains.

is despite the fact that individuals having second infections had significantly higher pre-infection titres against Perth 2009-type viruses.

A possible explanation for diminished second infection responses to Perth 2009-type viruses is that individuals who had second infections were infected with strains that were more antigenically similar to the previous Wisconsin 2005 antigenic cluster. However, this seems unlikely given the genetic homogeneity of viruses circulating in season 1 shown in figure 5.3 and 5.5. A more plausible explanation seems instead that, as has been hypothesised previously [Cobey and Hensley, 2017; Linderman et al., 2014], the presence of prior immunity in some children has skewed their antibody response towards additional stimulation of memory B cells (with more cross-reactivity against older antigenic variants) at the cost of stimulation of naive B cells (with more cross-reactivity against future antigenic variants). If so, to my knowledge, this would be the first direct evidence of this effect in humans.

5.4.8 *The build up of antibody-mediated immunity*

A long-recognised pattern is that HI titres against viruses that circulated early in an individual's life tend to be higher than titres against strains that circulated later [Davenport and Hennessy, 1956]. Figure 5.13 shows the pre-vaccination antibody titres measured against all the strains in the 1997, 1998, 2009 and 2010 vaccine trials studied in chapter 4 plotted against the year age of each subject in the year of virus isolation, excluding viruses that were isolated after the year of sample collection or from samples taken from individuals born before 1968 - the first year of circulation of H3N2 viruses in the human population.

In keeping with previous findings, a clear peak can be seen in figure 5.13 against viruses that circulated earlier on in an individual's lifetime, with titres beginning to plateaux at a mean HI titre of around 20 to viruses circulating later. While some have speculated that this type of pattern arises from repeated boosting of first infection titres upon later influenza exposures maintaining titres against early encountered strains at a higher level [Davenport and Hennessy, 1956; Lessler et al., 2012], the responses measured in the Managua cohort give an opportunity to shed some light on this process.

The right panel of figure 5.14 shows the same plot as for figure 5.13 but applied to samples taken pre and at least 6 months post a likely second infection in the Managua cohort, while the left panel shows the plot for samples taken before and at least six months after an infection in the Hà Nam cohort studied in chapter 3 (which were generally in individuals of an age where it would be expected to be much later than a first or second infection). It is clear that in the samples taken before a second infection, titres have not yet reached the levels found in older individuals but, interestingly, the long-term boosting that occurs post-second infection appears already to be sufficient to establish the later pattern seen. In contrast, the infections

that occurred in the Hà Nam cohort did little to further build the pattern of high titres seen against viruses encountered early in life.

The inference from figure 5.14 is that titres against viruses encountered early in life are clearly boosted upon the second infection, but later infections beyond this do little to increase these titres further. This is in keeping with the findings in chapter 3 that, while short-term responses often boost titres against very antigenically distinct strains to which there was prior immunity, these responses did generally not translate to increases in HI titre to these strains longer-term.

Building on these inferences, figure 5.15 compares titres against the Hong Kong 1968 test strain that was used in the Hà Nam cohort study in chapter 3 and vaccine studies in chapter 4 (A/Bilthoven/16190/1968) against the year of birth of individuals titrated in these studies (namely the Hà Nam samples collected at the start of the study in 2007 or the pre-vaccination samples taken for the vaccine trials). If repeated boosting upon later infection were the key driver of higher titres against the first encountered strains, little difference would be expected for individuals born before 1968 (the first year of H3N2 circulation in the human population) since the amount of subsequent exposure to H3N2 viruses would have been broadly the same. Significantly however, it is clear that those born more than about 10 years before 1968 show titres against A/Bilthoven/16190/1968 that are no higher than the titres seen maintained against viruses encountered later in life in figures 5.13 and 5.14 (a titre of around 20). In contrast, for those individuals born shortly before 1968, titres do begin to reach higher than average levels as would perhaps be expected for a first H3N2 exposure.

Figure 5.15 seems to show that the higher titres seen against viruses encountered early in life are not due to boosting of these titres upon later infection, but rather a feature of first infection responses themselves. Where individuals were born long before H3N2 circulation and likely first infected with a virus from another subtype, the effect within the H3N2 subtype is lost. Adding to previous findings that first infections with a certain subtype can effect later protection against strains from a different subtype [Gostic et al., 2016], these results indicate that similarly, later antibody responses can also be detectably affected.

5.4.9 *Comparison of human and ferret immune responses*

It is difficult to do a direct comparison of human and ferret antibody responses using the data currently available since none of the available ferret antisera were raised against the same viruses as those infecting the humans, and HI titrations were made against a slightly different set of antigens. That being said, it is possible to try and approximate how the overall antigenic breadth of antibody response compares between ferrets and humans by comparing antibody responses by antigenic cluster.

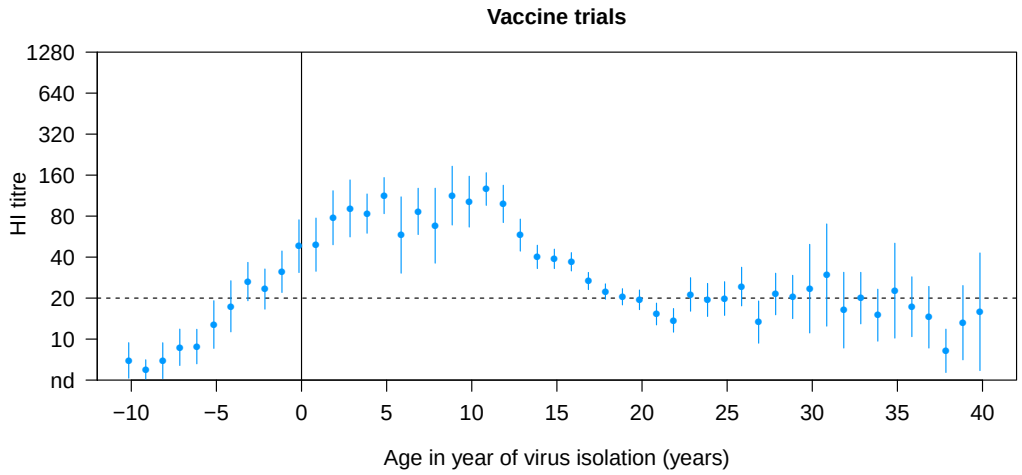


Figure 5.13: HI titre against age at time of virus circulation - vaccine trials. Average HI titres are shown grouped by age of individual in years at the time of virus circulation as measured in pre-vaccination samples from the 1997, 1998, 2009 and 2010 vaccination trials analysed in chapter 4. Circles show the mean titre and error bars show 95% confidence intervals. The horizontal dotted line marks an HI titre of 20.

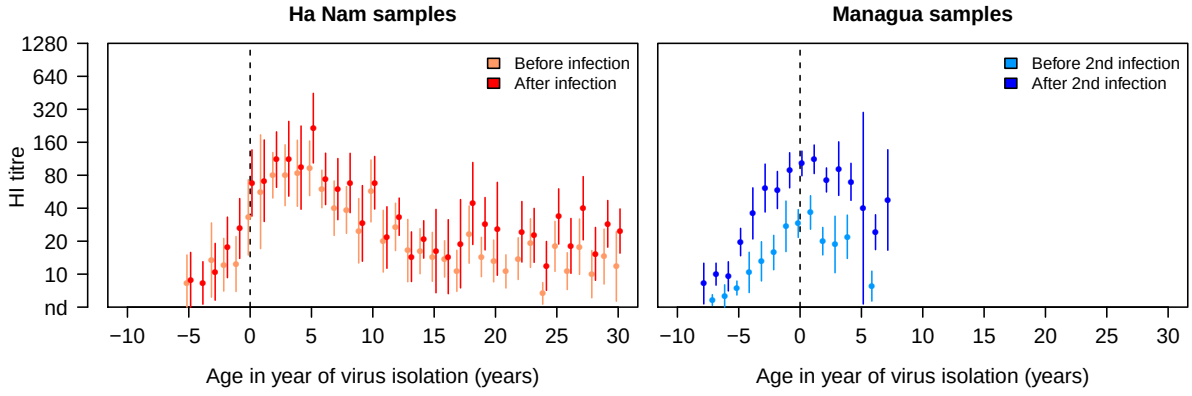


Figure 5.14: HI titre against age at time of virus circulation - Hà Nam & Managua cohorts. Average HI titres are shown grouped by age of individual in years at the time of virus circulation. Left panel: Averages across individuals in the Hà Nam cohort before and at least 6 months after seroconversion or PCR confirmed H3N2 infection. Right panel: Averages for individuals in the Managua cohort before and at least 6 months after a PCR-confirmed likely second infection. Circles show the mean titre and error bars show 95% confidence intervals.

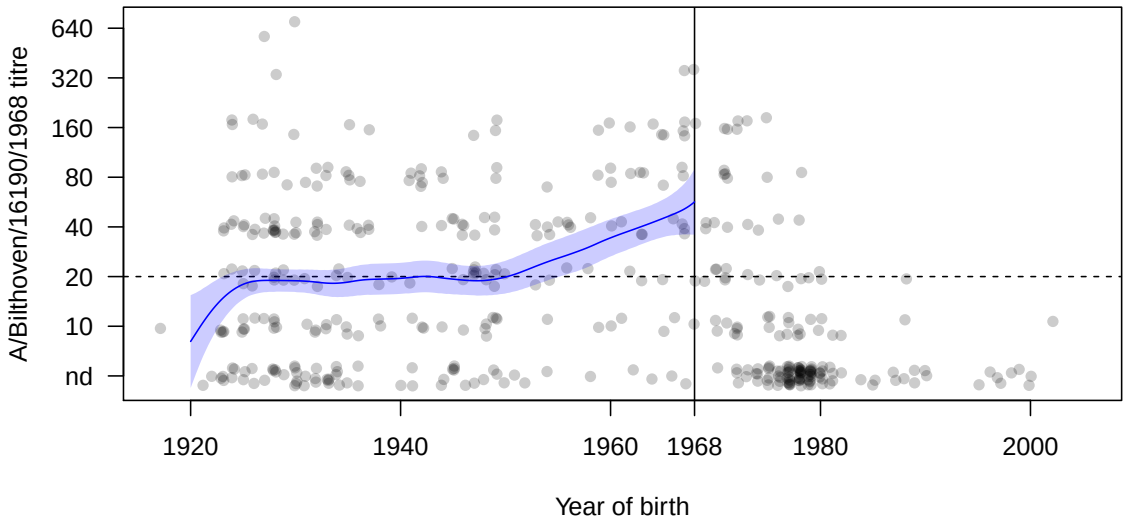


Figure 5.15: Year of birth against titre to A/Bilthoven/16190/1968. Individual titres are shown as transparent black circles with some jitter added in the x and y axes to help better distinguish the distribution of points. The solid black line marks 1968 - the first year of H3N2 circulation. The solid blue line shows a loess fit (span 0.5) to all the titres from individuals born in or before 1968. Light blue regions show 95% confidence intervals for the fit, based on 500 bootstrap repeats. The horizontal dotted line marks an HI titre of 20.

For each season, HI titres from ferret anti-sera raised against the antigen that best matched circulating strains (see section 5.4.3) were compared against first human infection responses. Because the ferret anti-sera were drawn 14 days post-infection and antibody titres later go on to decay, only human samples taken less than 40 days post infection were included in the analysis. Additionally, the responses from season 2 that measured a strong reaction against Wisconsin 2005-type strains indicative of a possible second infection were excluded (see section 5.4.5). Unfortunately no samples from seasons 3, 4 or 5 met these criteria but figure 5.16 shows the results for the 2 samples in season 1 and 3 samples in season 2 that did.

The average HI titre raised against viruses from each antigenic cluster is largely similar between ferret and human responses, as is the antigenic breadth in terms of number of clusters preceding the infecting against which there was a measurable response. On average the HI responses in the human sera were slightly lower than that seen in the ferret antisera, especially in season 2 but this may be expected since the average time post-infection that the samples were taken was 31 days in season 1 and 33 days in season 2, just over two weeks later than the ferrets and possibly accounting for some titre decay. On balance this analysis appears to offer little evidence for a substantial difference between humans and ferrets in terms of the strength or cross-reactivity of antibody responses following these infections.

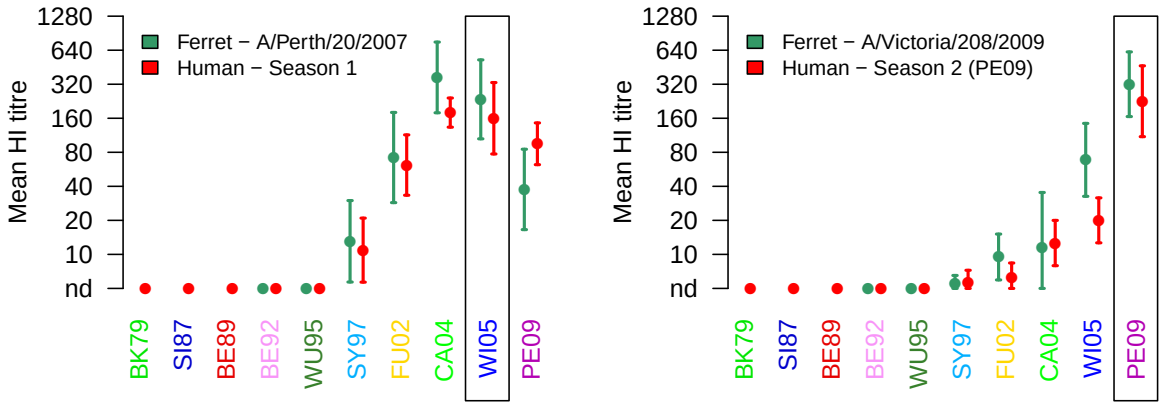


Figure 5.16: Comparison of human and ferret responses. Mean HI titre responses to titrated H3N2 strains split by antigenic cluster for likely first infections in the Managua cohort and for responses measured in a ferret antiserum. The left panel shows the mean HI titres post-infection for children who had a likely first infection in season 1 compared to a ferret antiserum raised against A/Perth/20/2007 (the closest genetic match of available ferret antisera to season 1 viruses). The right panel shows the mean HI titres post-infection for children who had a likely first infection in season 2 compared to a ferret antiserum raised against A/Victoria/208/2009 (the closest genetic match of available ferret antisera to season 2 viruses). Note that results from children identified in figure 5.7 as responding strongly to Wisconsin 2005-like viruses were excluded from the season 2 responses since it was unclear if these children had in fact had a prior infection. Error bars show the 95% confidence intervals for the mean. Boxes highlight the antigenic cluster of the infecting/likely-infecting strain.

An ideal test would be to raise ferret sera against the same viruses that had caused human infections. As it stands this dataset appears to offer no evidence of a substantial difference between ferret and human immune responses in the case of the particular infections analysed.

5.5 Discussion

5.5.1 *Variation in the first infection response*

The similarity of the antibody responses to first infection is a surprising finding given the heterogeneity of response to infection seen in the Hà Nam cohort in chapter 3 and to vaccination in chapter 4. It suggests that different exposure histories appear to have accounted for much of the variation seen in these previous cases rather than the other potential factors discussed such as genetic differences between individuals or variation in the severity of the infection. It also implies that in the absence of prior immunity, rather than a stochastic process, the overall antibody response to a given antigenic stimulus appears to be relatively stereotyped between different individuals. Also worth noting is that despite the large number of infection responses investigated here, we see no evidence for the seemingly universally broad responses produced by a minority of people in response to vaccination in chapter 4. The only individual showing something close to such a response is participant number 6301 (figure 5.10)

where the response pattern appears to be a rather more a haphazard one directed at older antigenic variants rather than a uniformly cross-reactive one.

Dichotomous responses in season 2

A clear exception to the otherwise relatively similar responses to first infection is the dichotomy of responses apparent in the first infections occurring in season 2, whereby approximately half the children responding to an apparent first infection showed a response clearly directed against viruses from the Perth 2009 antigenic cluster, while half showed a clearly mixed response, with additional reactivity against Wisconsin 2005-type strains.

The fact that circulating viruses sequenced from that season appeared to be exclusively of the Perth 2009 antigenic type, combined with the results showing very different responses following infection with the viruses that were almost genetically identical (figure 5.8), suggests that the simplest explanation of antigenic differences between infecting viruses is not the case. One possibility is that stochasticity of antibody responses to infection is indeed playing a role and that perhaps a proportion of children had generated responses to an epitope strongly shared between the season 2 viruses and viruses from the Wisconsin 2005 antigenic cluster, while others responded primarily against an epitope only found in Perth-2009-like strains. Given the lack of variation seen in the first infection responses of later seasons this seems like an improbable explanation, although possible if the shared epitope had mutated in subsequent seasons.

A more likely explanation seems to be that although none of the children had any detectable antibody titres in samples taken prior to the infection in season 2, some may nonetheless have had some antigenic experience of the previous season that shaped their subsequent response. Indeed, comparing the dates of birth of the subset of children showing high antibody responses to Wisconsin 2005-type strains to the epidemiological data on infection dates taken from the wider cohort, all of these children are old enough to have lived through and theoretically been exposed to viruses from the Wisconsin 2005 antigenic cluster.

Interestingly, had exposures occurred in these children it would have been at a very early stage of life when maternal antibodies acquired through passive placental transfer are still known to be circulating at significant levels in the blood [Bodewes et al., 2011; Leuridan and Van Damme, 2007; Leuridan et al., 2011; Watanaveeradej et al., 2003]. A resulting hypothesis is therefore that the antibody response to these exposures was dampened by the presence of cross-reactive maternal antibodies to the extent that no trace was detectable by serum HI titre. Upon secondary exposure however, memory B cells formed as a result of the first exposure again became stimulated, differentiating to form plasma cells secreting antibody more cross-reactive with

Wisconsin 2005-type strains. Evidence suggestive of exactly this process has already been found in studies of vaccination in young children [Glezen, 2003] and, if true, is an interesting example of how serum antibodies may not always be representative of the full B cell memory repertoire. Ultimately, follow-up experiments using additional assays to investigate whether evidence for prior immunity is found in certain samples will be required to answer these questions more conclusively.

Non-detectable and transient responses

The finding that certain children appear to have had no detectable long-term primary antibody response to a symptomatic and PCR-confirmed influenza infection is a surprising one. For later responses this would perhaps not be unexpected since prior immunity may deal with an infection without the necessity for significant production of additional antibodies. Following a first infection however, current understanding of B cell and antibody responses would predict that a durable antibody response should result.

It is important to understand if this is a real phenomenon and to rule out the possibility of confounding results due to sample switches or repeats (as seen in the Hà Nam cohort in chapter 3). There are a number of ways in which this could have happened, for example an accidental repeat of the sample taken from before the infection (a potential explanation for the responses seen in subjects 6428, 6613 and 7295), or after the infection (subjects 7308 and 7333), or even a switch of the before and after samples (subject 7308). Similarly, in the case of subject 4524 it seems possible that the date of infection may have been incorrectly recorded however records of both the original clinic visit and follow-up laboratory analysis corroborate the timing shown.

A separate explanation for the total lack of response seen in 6428, 6613 and 7295 is that they were infected with a later viral variant that did not share much cross-reactivity with the panel of strains tested. The latest variant tested was isolated in 2011 (A/Victoria/361/2011) but these infections occurred in 2012 and 2013. However, given no major antigenic change is thought to have occurred between these seasons and that other individuals infected in the same seasons show measurable titre responses against several of the strains tested, this seems unlikely.

Despite these possibilities, the pattern of responses seen in subject 5898 is harder to account for with titration errors or sample swaps since multiple samples taken after the first recorded H3N2 infection registered no detectable titres, while a sample taken 31 days after the first recorded infection showed evidence for a transient response that later disappears. Significantly, this first post-infection sample was taken within the window of time we would associate with the “dynamic” phase of the antibody response before the antibody landscape is again settled to a new level. This is so far the strongest piece of evidence that some children simply fail to maintain any

detectable antibody titres in the long-term, even after their very first infection. In the other cases of apparent non-response to infection the post-infection sample was not taken soon enough as to be in the time-window of the "dynamic" phase of the immune response, raising the possibility that these children responded in a similarly transient fashion yet failed to maintain their titres. Interestingly, 5898 goes on to have a response against a second infection that is still detectable 247 days afterwards. This would imply that the child is in principle still capable of generating a long-lived antibody response to infection although it will be informative to see if this response is also present in later follow-up samples taken in 2015 and 2016.

Repeat analysis of the potentially swapped samples may help to detect any errors that occurred upon aliquoting and shipment of serum samples while again, application of more sensitive assays to post infection samples would help to determine if there is any evidence for the development of lower levels of reactivity or reactivity against other antigenic components (such as stalk antibodies), that were not detectable via HI.

5.5.2 *Effect of prior immunity on secondary responses*

Although the number of post-second infection samples studied here were limited, they provide interesting insights into the effect of prior immunity on secondary responses. Previous studies have used animal models and artificial infection or vaccination to investigate antibody responses following primary and secondary exposures [Angelova and Shvartsman, 1982; Jensen et al., 1956; St Groth and Webster, 1966b; Webster, 1966]. This is the first example to my knowledge of a direct comparison in humans, and in response to natural infection. Many of the same phenomena described in seminal studies on the subject are also witnessed, with clear evidence that boosting of antibodies produced following a first infection is a dominant feature of secondary responses. Importantly, this is not just short term boosting, as described in previous studies, but we also witness long-term boosting of pre-infection antibody reactivity that is clearly distinct to first-infection responses.

Also notable is that average long-term post-infection antibody titres against strains from the infecting antigenic cluster were similar regardless of either presence or absence of prior immunity from a previous infection. This is also seen in the data of previous studies [Davenport and Hennessy, 1956; Francis et al., 1947; St Groth and Webster, 1966b], suggesting that prior-immunity does not ultimately reduce the antibody response against a secondary infection as some interpretations of the dogma of original antigenic sin have inferred [St Groth and Webster, 1966b]. Interestingly however, the results from the comparison of first and second infection responses in season 1 indicate that, while responses against the infecting cluster are not impaired, responses against future clusters can indeed be diminished. This is evidence for exactly the kind of "opportunity cost" of prior immunity hypothesised to have led to

susceptibility of certain portions of the population to emergent strains during the 2013-2014 influenza season [Linderman et al., 2014].

The evidence here supports an emerging picture of the antibody response to influenza whereby upon infection, antibody reactivity is generated partly through stimulation of memory B cells produced as a result of a prior infection but with cross-reactivity to the infecting strain (if such cells are present) and partly through direct stimulation of a naive B cell response. One might speculate that in terms of the ability to neutralise the infecting virus and maintain defence against re-exposure with the same strain, the relative contributions of naive and memory responses to this process have little impact since overall reactivity generated against the strain appears to remain the same (indeed it is interesting to note that the average post-infection titre maintained long-term against the infecting strain was almost exactly 40 - the classic cut-off for effective serological protection from infection). Due to the different patterns of cross-reactivity present in naive and memory responses however, contrasts become clear when examining antibody reactivity against antigenically variant strains. Ultimately it seems that by relying too heavily upon pre-existing immunological memory, the opportunity is missed to develop more comprehensive responses against the antigenic features present in the infecting strain with the result that protection against new variants is actually diminished.

5.5.3 *Implications for understanding the development of immunity*

The data presented help address questions regarding the dynamics of the development of immunity that were posed over 60 years ago by Davenport et. al who noted that titres against influenza strains tended to be highest against those that had circulated early in an individuals life - presumably the first influenza exposures [Davenport et al., 1953]. Since then, others have corroborated these findings [Francis, 1960; Lessler et al., 2012; St Groth and Webster, 1966a] and a popular interpretation has been that this pattern results from repeated boosting of antibodies cross-reactive against the first encountered strain [Cobey and Hensley, 2017; Francis, 1960; Lessler et al., 2012; Miller et al., 2013]. The findings presented in chapter 3, that although short-term responses tend to boost titres against the oldest strains encountered long-term responses do not, shed some doubt on this interpretation leading to the alternative explanation that first infections produce the largest antibody response and these high titres are then simply subsequently maintained. By comparing directly the patterns of immunity that follow first and second infections we can conclude that reality lies somewhere between these two extremes. Significant long-term boosting of titres to the first infecting strain does occur following a second infection, but the new evidence here shows that this is already enough to establish the long-term patterns observed and that later infections do little to further enhance these early titres, in contrast to other reports [Miller et al., 2013].

Interestingly, the definition of "first" infection appears to stretch across subtypes since even though individuals born long-prior to 1968 would have encountered such strains as their first H3N2 influenza subtype, this was still not sufficient to establish the pattern of distinctively high titres seen in those born at a time where they were likely to have encountered it as their first ever influenza infection. Despite the canonical dogma that different subtypes of haemagglutinin are not serologically cross-reactive, it appears that infection with H1N1 or H2N2 subtypes circulating prior to 1968 was able to diminish long-term titres formed against an early H3N2 strain. It has been recently reported how first infections can influence development of immunity between influenza groups [Gostic et al., 2016], but this is even despite the fact that both H1 and H2 are group 1 haemagglutinin types and H3 is group 2.

Why exactly first infection responses should generate higher titres is still unclear, one hypothesis is that without prior immunity the virus replicates on average to higher levels and thereby stimulates a larger antibody response compared to later infections. Cross-protection due to antibodies against viral components that are not neutralising but that are shared between subtypes would therefore explain why those individuals with prior infections with H1N1 or H2N2 viruses would not show this pattern of higher responses. It is striking however that titres against HK68 in those previously infected with H1N1 or H2N2 were not only lower than those having a first infection, but indistinguishable in magnitude from the average titre generated against later infections from the same subtype. This is even though it would surely be expected that first infections with a new subtype, even if not the very first influenza infection, must surely be more severe on average than those from later variants.

Ultimately, it will take more work to better understand exactly why the very first infection appears to hold such special status in determining patterns of later protection and antibody reactivity even across subtypes. Additional studies such as these but comparing antibody responses to first H3N2 infections in children previously infected with H1N1 or B may be one way to shed some more light on the mechanism, in particular whether the first H3N2 response itself would be diminished or instead the later boosting effect of second infections that drives it to higher than average levels.

Perhaps even more significant and relevant to our strategies for protection, it will also be important to understand the effect prior vaccination and simultaneous antigenic exposure to multiple influenza subtypes has on shaping later patterns of immunity. In particular, will this prime the immune response for development of high first infection titres against all seasonal influenza subtypes or somehow diminish this effect through substitution of the first infection with a less immunogenic vaccine antigen exposure?

6

Concluding Remarks

6.1 Chapter summary

The analyses throughout this thesis have covered an extensive range of serological data, including a staggering 59490 individual HI results from 1467 individual serum samples covering responses to a total of 168 PCR-confirmed influenza infections and 788 vaccinations.

In this chapter I would like to briefly bring together some of the key observations drawn from this extraordinary dataset, including insights into the immunological nature of the serological response, implications for global influenza vaccination and outlook for future work investigating these important subjects.

6.2 Outlook for antibody landscapes

The HI assay used for the majority of investigations in this thesis is typically considered to give a relatively coarse and variable measurement of individual antibody reactivity. The antibody landscapes methodology developed here has however demonstrated that, by combining results from many variant pathogen strains in a statistically appropriate way, it is possible to use data from such sources to achieve a much more detailed profile of individual immunity and immune change than would normally be expected.

Looking ahead, as a tool for the evaluation of different vaccination techniques antibody landscapes holds much promise, allowing a much more sensitive and comprehensive comparison of responses than traditional approaches. This may be especially beneficial when considering new approaches such as the effect of adjuvants [Gordon et al., 2016], vaccine composition [Moa et al., 2016] and dose sparing [Kenney et al., 2004], while in the case of universal vaccine candidates, the ability to accurately assess both magnitude and breadth would prove particularly useful.

In the field of seroepidemiology, the greater sensitivity to serological change and reduced dependence of results upon the particular selection of strains used could potentially make antibody landscapes a far more reliable indicator of infections in a population. While the workload associated with the extensive number of titrations required for antibody landscapes is for some applications prohibitively high, it is also likely that as newer technologies continue to be developed that allow high throughput analysis of serum reactivity to a greater number of strains, larger scale application of the antibody landscapes methodology will become feasible [Koopmans et al., 2012; Wang et al., 2015].

Perhaps most excitingly, while the investigations in this thesis have focussed on the antibody response to influenza I also hope that the principles uncovered and the techniques developed here will be more broadly applicable both to other aspects of immunity and to other pathogens where antigenic differences between different strains are being increasingly well quantified [Katzelnick et al., 2015].

6.3 Immunological observations

Individual variation was a common feature of the patterns of serological immunity observed in this thesis but several general observations can be made that hold surprisingly true for the extensive range of samples analysed. Taken together these general observations produce, loosely speaking, a set of “rules”

for the antibody response to the H3N2 influenza subtype and likely for other subtypes too:

1. Post infection or vaccination, a short-term response is stimulated that can be antigenically far broader than a typical novel antibody response, but that is generally limited by the extent of pre-existing antibody reactivity.
2. The larger short-term antibody response decays to a smaller steady-state long-term response over a course of time between 2 months and a year.
3. Long-term antibody responses are antigenically far narrower and more directed against strains that are antigenically similar to the causative virus.
4. Both short and long-term antibody titres against all strains remain at least as high after infection than before (although there may be little overall increase).
5. Once antibody titres have settled to form a long-term response, there is little further decay detectable.
6. There is no antibody response to H3N2 strains, detectable via HI, as a result of infection with other subtypes.

Although detailed investigation of any underlying cellular mechanisms was not a central focus, I feel that insights into the immunology of influenza in humans can be drawn from these features. Some of these inferences have been discussed in chapter 3, for example the fact that the antigenic breadth of the initial response is limited by the extent of prior immunity indicates, along with findings from the wider literature, that responses against older strains are characteristic of memory recall responses.

The dynamics of the short-term antibody response are in good keeping with the patterns of decay that would be expected from the estimated half-lives of short-lived plasma cells producing antibodies [Morell et al., 1970; Slifka et al., 1998], indicating a likely mechanism for this transient boost in titres against older viruses when memory B cells with some BCR-reactivity to an antigen become stimulated to divide and differentiate [Bernasconi et al., 2002]. The fact that in most cases increases in reactivity towards older viruses are short-lived indicates that this process of memory B cell stimulation and differentiation does not necessarily lead to additional numbers of long-lived plasma cells. In this regard, the observation that the antigenic breadth of long-term increases in antibody reactivity is more specific to strains similar to the infecting strain presumably reflects the greater degree of competition for limited survival opportunities amongst plasma cells with the potential to become long-lived, with

the consequence that only those with a higher threshold of BCR-affinity ultimately become sustained [Radbruch et al., 2006].

A common postulation in the immunological literature is that since total serum IgG levels do not increase over time, there must be a limited maximum number of survival niches and, as new antibody secreting long-lived plasma cells are produced, older ones must be displaced [Manz and Radbruch, 2002]. It could be imagined that in the case of influenza this would be reflected in a decrease in antibody titres to certain strains following a new infection. It is therefore also noteworthy that the long-term changes measured were almost always neutral or represented an increase - although it is of course possible simply that cells producing antibodies of a different specificity to those measured were displaced.

In keeping with other reports, after the initial dynamic response phase, there was generally little detectable further decay of antibody titres in the long term [Künzel et al., 1996; Yu et al., 2008], suggesting that antibody titres can be maintained for a very long time unchanged, perhaps indefinitely. However, this conclusion leads to an interesting question of how it came to be that some of the older individuals in the vaccine trials studied began the trial with such low titres against so many of the H3N2 strains tested. Whether this represents an unlikely lack of exposure to H3N2 strains, antibody responses directed against virus components not detectable with HI, or indeed a longer-term process of antibody decay that has occurred in these individuals is unclear.

Finally, the finding that infection with other influenza subtypes does not stimulate antibody responses to H3N2 measurable with HI is in keeping with previous findings from others [Grebe et al., 2008] but is here shown very clearly with regard to natural infection and measurement of responses against a wide range of H3N2 strains. However, the observation in chapter 5 that prior infections with other subtypes still appear to be able to affect later antibody responses to H3N2 blurs the boundaries between the antibody response to different influenza subtypes, and how they interact, in interesting new ways.

6.4 Cleaning up "original antigenic sin"

In addition to general immunological observations, the patterns identified in this thesis undoubtedly bear particular relation to the hypothesis of "original antigenic sin", a theme that has made repeated appearances [Francis, 1960]. In chapter 1 I argued that use of the term original antigenic sin to refer to many separate phenomena had caused confusion in the scientific literature, with the concept seemingly meaning slightly different things to different people. I feel it is worth dedicating some space in this final chapter to review the findings of this thesis in relation to evidence for and against specific facets of this long-held

dogma, and to hopefully cleanse original antigenic sin from some fallacies and ambiguity.

Later responses are biased towards stimulation of prior immunity but not necessarily antibodies produced against the first infecting strain.

In its loosest sense, original antigenic sin is often taken to mean that later antibody responses are somehow "skewed" or "biased" by infection history, in particular the first infecting strain [Cobey and Hensley, 2017; Lessler et al., 2012; St Groth and Webster, 1966b]. The results here show that undoubtedly, as well established by over 70 years of research, prior immune responses go on to strongly affect the response to later exposures [Davenport and Hennessy, 1956; Davenport and Hennessy, 1957; Jensen et al., 1956; Webster, 1966], however it is incorrect and overly simplistic to say that this is always manifested in a bias towards the immune response produced after the very first infection.

Responses in chapter 3 demonstrated the important distinction between fold-change responses and peak post-infection titres. As found in the data of previous studies, fold-change responses tended to favour the infecting strain but higher pre-infection titres to previous strains often meant that post-infection titres still peaked against earlier variants [Francis et al., 1947; Jensen et al., 1956; St Groth and Webster, 1966b]. Looking further ahead at the long-term response ultimately maintained post-infection, although initial antibody responses often included boosting of titres to all previously encountered strains, this effect was generally not maintained for very antigenically distinct early variants. Since antibody responses to the infecting strain also decayed over time, the ultimate result could still be that post-infection titres were highest against early strains. However, given that no long-term increase to these titres had occurred, it is misleading to suggest that this counts as a *bias* of responses to the first encountered viruses.

An illustration of these points is given in figure 6.1 where short and long-term post-infection titres and fold-change responses for one of the individuals infected in the Hà Nam cohort studied in chapter 3 are compared. In this case we can see that the long-term response is skewed towards some long-term boosting of *prior immunity*, but not towards long-term boosting of titres against the likely *first* infecting strain.

The fact that long-term increases in titres to the oldest encountered strains were often not maintained also has implications for other tenets of original antibody sin, for example that antibodies produced in response to a new exposure come primarily as a result of stimulation of memory B cells reactive against epitopes shared between the new and previously encountered strains, predominantly the first infection [Davenport et al., 1953; Francis, 1960; Jensen et al., 1956]. In

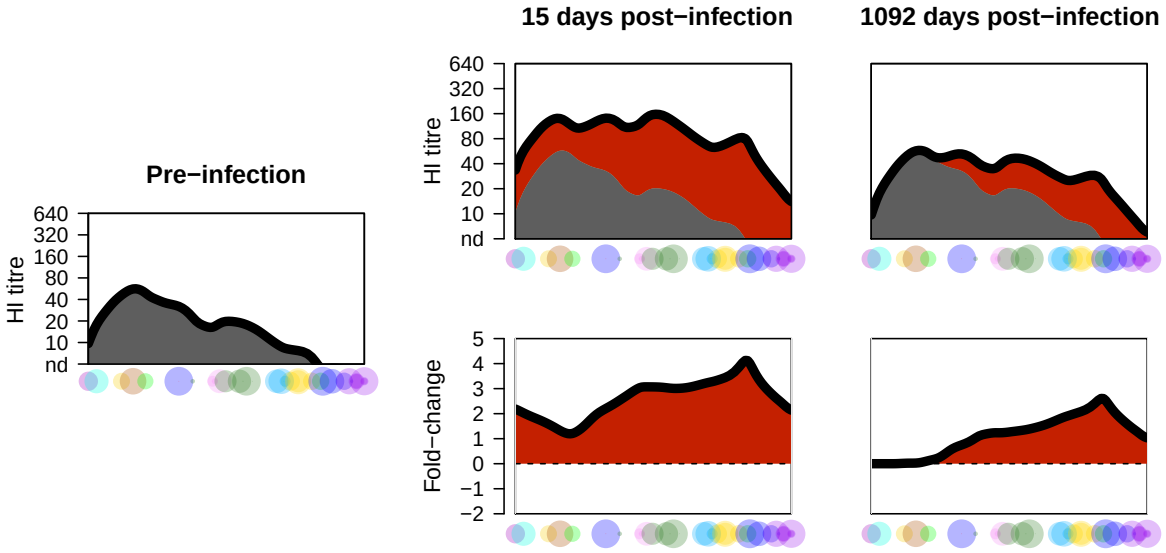


Figure 6.1: Example post-infection titre vs. fold-change response. The left-most panel shows the antibody landscape taken pre-infection for an individual in the hanam cohort. The panels to the right then go on to compare the short-term and long-term post-infection antibody landscape responses either in terms of the overall post-infection response (top panels) or the fold-change response (bottom panels). As described in chapter 2, the black lines in each plot represent the antibody landscape height for each position on the antigenic summary path taken through the H3N2 antigenic clusters. Dots along the x axis indicate the antigenic positions along the summary path of the subset of viruses used to generate each landscape, scaled negatively by distance from it. In the top panels, both the short-term and long-term responses are compared to the pre-infection sample with red-shading showing increases. Note that, although the likely infecting virus was from the Wisconsin 2005 or Perth 2009 antigenic cluster, post-infection responses peak against viruses from earlier antigenic clusters. Fold-change responses however, peak against viruses from the Wisconsin 2005-type strains.

its most extreme form, this aspect of original antigenic sin predicts that for an individual who was born during the Hong-Kong 1968 antigenic cluster, a large portion of the antibody response to a later infection with a Perth-2009-type virus should come from antibodies cross-reactive between both Hong-Kong 1968 and Perth 2009 strains.

Although no analysis of absorption studies was performed to investigate the binding specificity of newly produced antibodies, the fact that long-term responses in such cases generally included increases in titre to Perth-2009-type strains, with no joint rise in titres to a Hong-Kong-1968-type virus, suggests that the production of antibodies with such diverse cross-reactivity cannot be the source.

Mechanistically it can be imagined that the advantages memory B cells formed in response to the first infection have over naive B cell populations, both numerically and in terms of their ability to respond quickly to antigenic stimulation [Rajewsky and Schittek, 1990; Tangye et al., 2003], mean that they do indeed

dominate second and perhaps even third infection responses as has previously been reported. However, as the infecting virus becomes more and more antigenically distinct, BCR stimulation is reduced and, perhaps more significantly, these early memory B cell populations must compete with other memory cell populations formed as a result of later infections that would be expected to share more cross-reactivity with the new infecting strains. At some point, memory B cells produced in response to the first infection can no longer compete effectively with memory cells produced against strains with more cross-reactivity and these early populations cease being boosted long-term.

In general, as discussed in chapter 3, given the patterns of antibody response reported in this thesis, it is likely that by focussing on early responses to antigenically variant but still related strains, original studies may have overstated the dominance of the very first infection.

Prior immunity does not effect long-term antibody reactivity generated against an infecting strain, but may diminish reactivity to future antigenic variants.

Many authors have described the effect of antibody responses to influenza infection as a result of original antigenic sin (in this case taken to mean a strong boosting of responses to previously encountered strains) as suboptimal or in some way impaired [St Groth and Webster, 1966b]. Given the speed with which memory cell populations can respond to an invading pathogen and the fact that post-infection antibody reactivity to an infecting virus appears to be very similar regardless of the presence or absence of prior immunity, it is not immediately clear that such statements are justified. Where responses against later antigenic variants are compared however, direct evidence is found for the previously hypothesised effect that prior immunity can have in diminishing stimulation of novel antibody responses that would otherwise be more protective against future antigenic variants [Cobey and Hensley, 2017; Gostic et al., 2016; Linderman et al., 2014; Smith et al., 1999].

Evolutionarily speaking, it is interesting to consider why an immune system with these apparent short-comings may have persisted. In terms of protection, the stimulation of cross-reactive memory responses to combat an invading strain makes sense since responses can be far more rapid and likely to limit growth of the infecting virus and severity of disease. A possible explanation may simply be that by comparison, developing responses that are more likely to protect against future antigenic variants is very much a secondary concern. Given reports of the ability for antigen to persist for a long time after infection however [Zinkernagel and Hengartner, 2001], there are still unanswered questions as to why even though memory responses may dominate the initial response, such sequestered antigen could not be used to stimulate equally ro-

bust naive B cell responses in the long-term. To some extent this is exactly what is seen in the antibody responses described in chapter 3 - the initial response of the immune system to infection appears to be a massive stimulation of all antibody previously produced in response to infections with that influenza subtype, but a longer-term response that includes mainly only an increase in antibody reactivity to strains antigenically similar to the virus encountered. The findings in chapter 5 however show that although this process may help to maintain both effective initial protection and an antibody repertoire with relevant diversity, it does not seem to be enough to fully compensate for a loss of long-term overall reactivity compared to antibody responses produced to infections in the absence of prior immunity.

High titres to viruses circulating early in life are related to the distinct nature of the first infection response, not repeated boosting upon later exposures.

It is probably fair to say that this is an aspect of the concept of original antigenic sin where the waters get particularly muddy. Early researchers noted the pattern that individuals in cross-sectional serological studies tended to have the highest titres against strains circulating early in their lifetime and were understandably enthusiastic to link it to the other phenomena they observed, namely the strong boosting of titres to previously encountered antigens of the same subtype upon influenza exposure [Davenport et al., 1953]. One conclusion was that it was this repeated boosting that had built particularly high titres against the first strains.

Having examined in detail a range of short and long-term responses in this thesis I can now conclude that little evidence is found for the hypothesis that titres against early strains result from continual long-term boosting. Titres to first infection strains do indeed receive a long-term boost after a second infection but this is already sufficient to generate the high titre levels seen maintained in older individuals. Similarly, later infections in older individuals do boost titres to even very antigenically distinct previously encountered strains in the short-term, but this does not generally result in any further long-term increases. On the individual level there is considerable variation, but the average long-term build up of antibody titres to viruses therefore appears to be a process more akin to the schematic shown in figure 6.2 - titres are boosted to a given level after each infection and, aside from increases as a result of the second infection, the profile of antibody reactivity is mostly extended rather than continually boosted long-term.

As discussed in chapter 5, those likely to have had a first exposure to the H1N1 or H2N2 subtypes before the first circulation of H3N2 viruses in humans do not share the phenomenon of higher than average titres against early H3N2

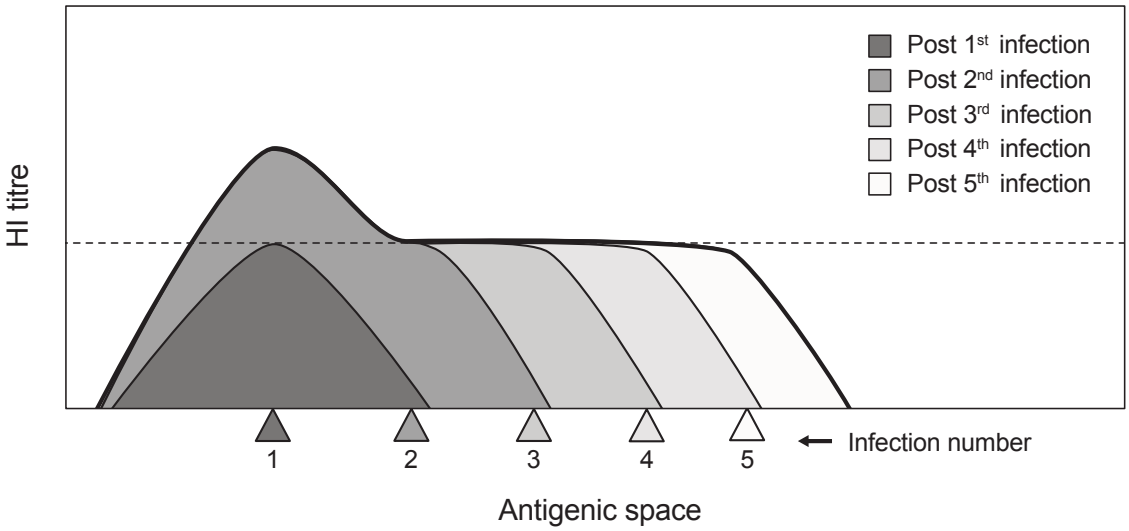


Figure 6.2: A schematic of the long-term build up of antibody reactivity over time. A series of 5 infections with antigenically evolving influenza viruses are imagined. Shaded regions indicate how antibody reactivity would be expected, on average, to build up over time. Post-infection titres tend to reach a consistent level against the infecting virus (marked by the horizontal dashed line). There is significant long-term boosting of first infection titres upon a second infection but little subsequent long-term change in these titres upon later infections, with antibody reactivity being extended rather than boosted over time.

viruses, implying that this phenomenon is instead a feature of first infections in particular rather than the responses of antibody titres to later infections. Presumably this finding would still hold in the case of first infections with another subtype not only because of birth before circulation of H3N2 but also for individuals born since 1968 who were by chance exposed to a seasonal or pandemic H1N1 virus before H3N2 (indeed whether this pattern is also present for those individuals first infected with an influenza B virus could tell us something about the physiological relevance of potential cross-reactivity between these different virus types). Such variation in first infecting virus would therefore also be a likely explanation as to why the pattern of high titres against the first encountered H3N2 strains was often absent at an individual level.

6.5 Implications for vaccination

Ultimately, a key motivation for generating a better understanding of the immune response and for the work in this thesis is how this knowledge can be used to generate improved immune responses and global protection through vacci-

nation. In this regard I hope some important insights have been generated and it is on this topic I would like to end.

In particular, I feel this work has highlighted that, while the current vaccine-strain strategy has undoubtedly provided considerable protection through attempts to always ensure the best antigenic match, by ignoring the effect of prior immunity on responses, this has led to a suboptimal approach. Chapter 4 highlighted a specific situation in which, due to pre-vaccination population immunity, the best antigenic match did not necessarily provide the best overall protection in terms of antibody responses, and use of a more antigenically advanced strain may have been beneficial. As perhaps the most significant point it also demonstrated that, in the face of low level circulation of novel antigenic variants at the time of strain selection, a decision to delay the vaccine update, which may at first glance appear a careful and conservative option, has the potential to be harmfully negligent when immunological memory and antibody recall responses are considered.

There is still much to be understood with regard to strategies such as pre-emptive vaccine strain updates. In particular, an understanding of the impact they may have on vulnerable individuals with immunodeficiency or absence of prior exposure is a key priority. However, we should at least aim to to move in the direction of a vaccine strain selection strategy that goes beyond simple antigenic matching. Progress along this path will likely require a lot of work to turn observational studies such as those performed here into a predictive understanding of how potential vaccine strain candidates would be likely to interact with prior immunity in the population - an understanding that would eventually allow optimisation of choices based on expected immune responses directly. It is an ambitious goal, but one that I feel is within sight. I hope the work presented in this thesis provides both a stepping stone towards it, and a demonstration of the potentially significant benefits it would hold for vaccine efficacy and individual protection should we reach it.

"Progress is impossible without change, and those who cannot change their minds cannot change anything."

- George Bernard Shaw [1944]

Appendix

A

List of H3N2 strains

Below is a list of all strains used in the construction of antibody landscapes. Datasets are as follows: 1 = Titrations for the hanam strains set 1 (more antigenically limited), 2 = Titrations for the hanam strains set 2 (more antigenically extensive), 3 = Strains used for the 1997 and 1998 vaccination studies, 4 = Strains used for the 2009 and 2010 vaccination studies, 5 = Strains used for the Nicaraguan children's cohort study.

Note that the "G145K" in A/Netherlands/178/1995 (G145K) refers to the fact that this virus was reverse engineered to change a Glycine (G) to a Lysine (K) in the HA gene at position 145. Similarly the "Egg" and "MDCK" occurring with respect to the strain A/Philippines/472/2002, indicates strains from the same original isolate but with different passaging histories, namely through egg or MDCK cell lines.

	Cluster	Strain name	Dataset					Accession no.
			1	2	3	4	5	
1	HK68	A/Bilthoven/16190/1968		×	×	×		KM821278
2	EN72	A/Bilthoven/21793/1972		×	×			KM821279
3	VI75	A/Bilthoven/1761/1976		×	×	×		KM821280
4	TX77	A/Bilthoven/2271/1976		×	×	×		KM821281
5	BK79	A/Netherlands/233/1982		×	×	×	×	KM821282
6	SI87	A/Netherlands/620/1989		×	×	×	×	KM821283
7	BE89	A/Netherlands/823/1992		×	×	×	×	KM821284
8	BE92	A/Beijing/32/1992	×	×	×		×	KM821292
9	BE92	A/Netherlands/179/1993			×			KM821285
10	BE92	A/Shangdong/9/1993			×	×		KM821294
11	BE92	A/Wellington/25/1993	×		×			KM821289
12	BE92	A/Wellington/96/1993			×			KM821290
13	BE92	A/Victoria/1/1993	×		×		×	KM821288
14	BE92	A/Wellington/1/1994		×	×	×	×	KM821286
15	BE92	A/Brisbane/22/1994			×			KM821287
16	BE92	A/Johannesburg/33/1994			×			KM821293
17	BE92	A/Victoria/9/1994			×			KM821291
18	WU95	A/Wuhan/359/1995			×			KM821302
19	WU95	A/Nanchang/933/1995		×	×	×		KM821307
20	WU95	A/Netherlands/178/1995 (G145K)		×	×			KM821295
21	WU95	A/Brisbane/22/1996	×		×			KM821296
22	WU95	A/Wellington/48/1996			×			KM821297
23	WU95	A/Auckland/5/1996		×	×			KM821303
24	WU95	A/Canberra/1/1996			×			KM821304
25	WU95	A/Christchurch/1/1996	×		×	×		KM821306
26	WU95	A/Victoria/47/1997	×	×	×	×	×	KM821298

	Cluster	Strain name	Dataset					Accession no.
			1	2	3	4	5	
27	WU95	A/Perth/5/1997	×	×	×		×	KM821299
28	WU95	A/Perth/9/1997	×	×	×		×	KM821300
29	WU95	A/Bangkok/1/1997			×			KM821305
30	WU95	A/Tasmania/1/1997	×		×			KM821301
31	SY97	A/Canberra/9/1997	×		×	×		KM821309
32	SY97	A/Sydney/5/1997		×	×		×	KM821316
33	SY97	A/Victoria/507/1998		×	×	×	×	KM821319
34	SY97	A/Netherlands/301/1999			×			KM821308
35	SY97	A/Panama/2007/1999			×			KM821317
36	SY97	A/Townsville/2/1999			×			KM821314
37	SY97	A/Christchurch/68/1999	×	×	×		×	KM821315
38	SY97	A/Victoria/577/1999			×			KM821320
39	SY97	A/Fujian/140/2000			×			KM821318
40	SY97	A/Sydney/228/2000			×			KM821313
41	SY97	A/South_Australia/53/2001	×	×	×		×	KM821312
42	SY97	A/South_Australia/84/2002	×	×	×		×	KM821311
43	SY97	A/Brisbane/5/2002			×			KM821310
44	FU02	A/Philippines/472/2002 (MDCK)		×	×		×	KM821322
45	FU02	A/Fujian/411/2002	×		×			KM821324
46	FU02	A/Philippines/472/2002 (Egg)			×			KM821330
47	FU02	A/Townsville/4/2002			×	×		KM821327
48	FU02	A/Fukuoka/55/2002			×			KM821333
49	FU02	A/Netherlands/213/2003			×			KM821321
50	FU02	A/Wyoming/3/2003			×			KM821328
51	FU02	A/Auckland/20/2003			×			KM821323
52	FU02	A/Townsville/36/2003			×			KM821325
53	FU02	A/Brisbane/342/2003			×			KM821326
54	FU02	A/Victoria/110/2004	×		×		×	KM821331
55	FU02	A/New_Caledonia/12/2004	×		×		×	KM821332
56	FU02	A/Wellington/1/2004	×	×	×	×	×	KM821329
57	CA04	A/California/7/2004	×		×		×	KM821334
58	CA04	A/Singapore/37/2004	×		×			KM821335
59	CA04	A/Victoria/523/2004		×	×		×	KM821336
60	CA04	A/New_York/55/2004			×	×	×	KM821338
61	CA04	A/Brisbane/3/2005	×		×		×	KM821337
62	WI05	A/Wisconsin/67/2005	×		×			KM821341
63	WI05	A/Brisbane/10/2007	×		×			KM821339
64	WI05	A/Perth/27/2007	×	×	×	×	×	KM821342
65	WI05	A/Uruguay/716/2007	×	×	×		×	KM821343

	Cluster	Strain name	Dataset					Accession no.
			1	2	3	4	5	
66	**WI05**	A/Hanoi/EL134/2008			×			KM821348
67	**WI05**	A/Hanoi/EL135/2008	×		×		×	KM821349
68	**WI05**	A/Hanoi/EL140/2008	×		×			KM821350
69	**WI05**	A/Victoria/500/2009		×	×	×		KM821340
70	**WI05**	A/Hanoi/EL196/2009	×	×	×		×	KM821351
71	**PE09**	A/Victoria/208/2009	×	×			×	KM821344
72	**PE09**	A/Perth/16/2009	×			×		KM821346
73	**PE09**	A/Hanoi/EL201/2009	×					KM821352
74	**PE09**	A/Hanoi/EL204/2009	×				×	KM821353
75	**PE09**	A/Hanoi/EL206/2009	×	×			×	KM821354
76	**PE09**	A/Victoria/8/2010	×					KM821345
77	**PE09**	A/Hanoi/EL442/2010	×				×	KM821355
78	**PE09**	A/Hanoi/EL443/2010	×					KM821356
79	**PE09**	A/Hanoi/EL444/2010	×					KM821357
80	**PE09**	A/Hanoi/VN053/2010	×	×			×	KM821358
81	**PE09**	A/Victoria/361/2011	×			×	×	KM821347

B

The HI assay

The HI assay protocol

All HI titrations analysed in this thesis were performed by collaborators at the Victorian Infectious Disease Reference Laboratory in Melbourne using the following protocol:

Ferret sera titrations

Sera from influenza-naive ferrets each infected with a single influenza H3N2 virus (viruses isolated between 1989 and 2010) were collected 13-20 (typically 14) days post-infection. These sera were tested with the haemagglutination inhibition (HI) assay against 74 H3N2 influenza viruses isolated globally between 1992 and 2011. The ferret antisera were tested against a fixed quantity of each influenza virus (standardised to 4 HA units/25 μL). Prior to testing, each ferret antiserum was treated to remove non-specific inhibitors, by combining 1 part antiserum with 4 parts receptor destroying enzyme (RDE) (100 units/25μL) (Seiken RDE II, Denka Seiken Co., Ltd.) and incubating at 37°C overnight. Five parts 1.6% (w/v) sodium citrate solution (Ajax Finechem Pty Ltd.) were then added to the RDE-treated serum and incubated for 30 minutes in a 56°C water bath to inactivate the RDE. The treatment of antiserum resulted in a ten-fold dilution. RDE-treated antiserum was then adsorbed with 5% (v/v) packed turkey RBC in a 10:1 ratio, mixed by inversion and incubated at room temperature for 30 minutes before centrifugation of the RBC pellet at 2000 rpm for 5 minutes followed by removal of the adsorbed serum supernatant. Serial two-fold dilutions of the treated antisera in CaMg-free Phosphate Buffered Saline (PBS), pH7.2 (Sigma-Aldrich), were used to determine the highest dilution of antiserum capable of blocking agglutination between test influenza viruses and 1% (v/v) turkey RBCs in a 96 well V-bottom Micro Test plate (Greiner Bio-One). This dilution is called the HI titre, or HI value. The lower detection limit of the titres was 10 (lower dilutions were not possible due to the dilution factor associated with RDE treatment).

Human sera titrations

To determine the antibody titre of a human serum, the HI assay was executed as above, but with 50 μL of the treated human serum of interest instead of the ferret antiserum. Following each HI assay, viruses were retitrated to verify that their HA titre was still at 4-8 HA units/25 μL. Ferret antisera from the Wuhan 1995 (WU95), Sydney 1997 (SY97), Fujian 2002 (FU02) and Wisconsin 2005 (WI05) antigenic clusters, and their homologous antigens with known end-point titres, were included in each assay as positive controls. Day 0 sera and PBS samples were tested in each assay and served as negative controls.

C

Additional figures

C.1 Antibody landscapes for the Hanam cohort study

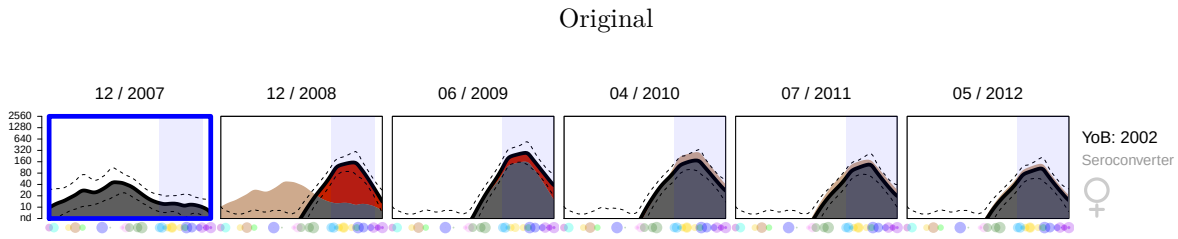
Shared caption for the figures in this section: Each row contains 2D summary antibody landscapes corresponding to samples taken from a single individual and the information to the right gives their date of birth, gender and reason for inclusion in the subset of test individuals. Each column contains landscapes corresponding to samples taken in a given year, with the year and month of the sample collections shown above. As described in chapter 2, the black lines represent the landscape height for each position on the antigenic summary path taken through the antigenic clusters. The antibody landscape measured in the first sample taken is shown shaded in grey, in subsequent samples, red shading indicates increases, and beige decreases, compared with the previously sampled year. Additionally shown are blue-shaded rectangles that indicate antigenic clusters that had circulated during an individual's life span until sample collection. Dots along the x axis indicate the antigenic positions along the summary path of the subset of viruses used to generate each landscape, scaled negatively by distance from it. The first sample after a PCR-confirmed influenza virus infection is marked with a coloured box and the number of days post collection of the PCR-positive swab shown. Different colours denote the different subtypes: red = H3N2, orange = H1N1, pink = H1N1p, green = influenza B. Where the set of titrated strains did not cover enough of H3N2 antigenic space to show a comparison of antibody landscapes to the previous sample across the full length of the summary path, the region is greyed out.

C.1.1 *Likely sample swaps*

Several of the antibody landscapes in the Hanam cohort appeared to fit better to the sequence of changes in antibody reactivity measured in other members of the same household. Four such examples were identified as shown in the following plots. For swaps 2, 3 and 4, where a likely pair for the swap could be identified, samples were switched before both analysis and other visualisations of these landscapes. For swap 1, where no likely pair could be identified (possibly the swap was with an individual in the cohort who was not part of the subset of individuals chosen for more detailed antibody-landscapes-based analysis), the sample was excluded. The likely sample swaps are highlighted in blue and in the individual plots, all these samples that were switched are labelled with an asterix.

Swap 1 - Household 047

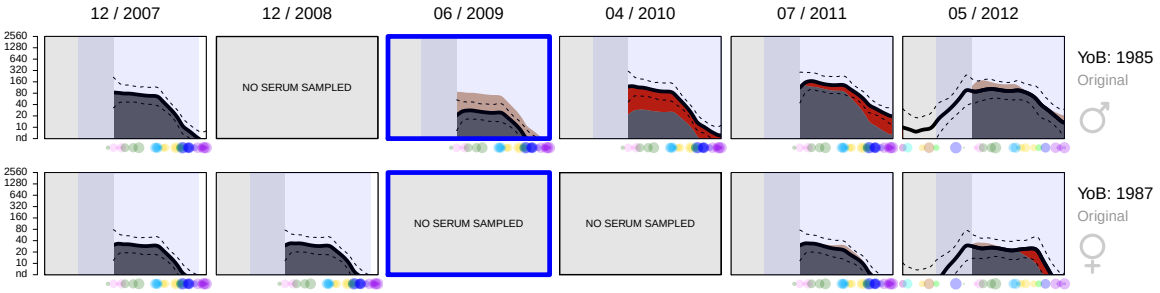
A potential sample swap in 2007 where the likely pair of the swap was not identified.



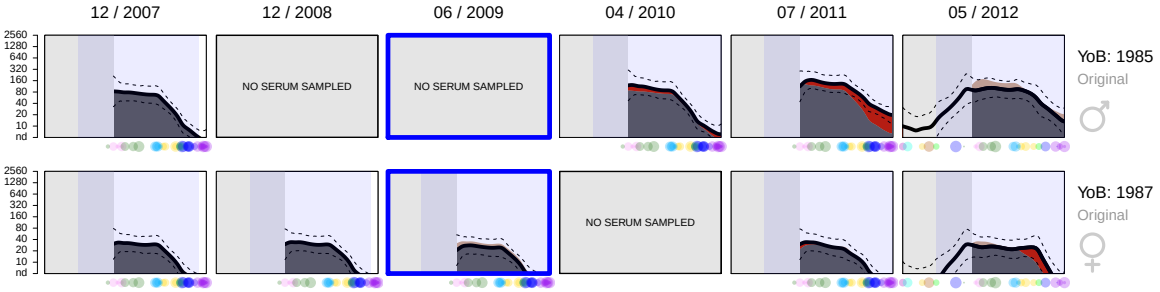
Swap 2 - Household 004

A potential sample swap in 2009 for two individuals within the same household.

Original



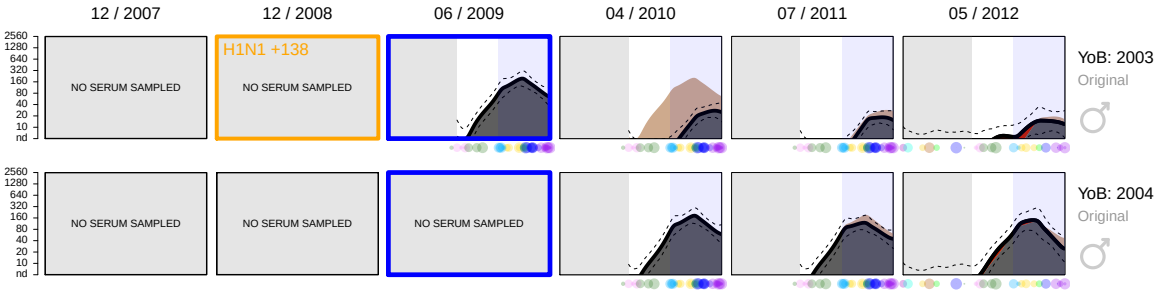
Switched



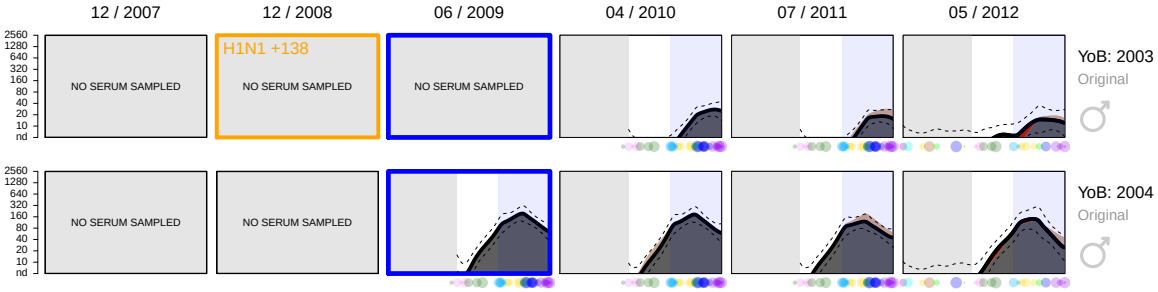
Swap 3 - Household 006

A potential sample swap in 2009 for two individuals within the same household.

Original



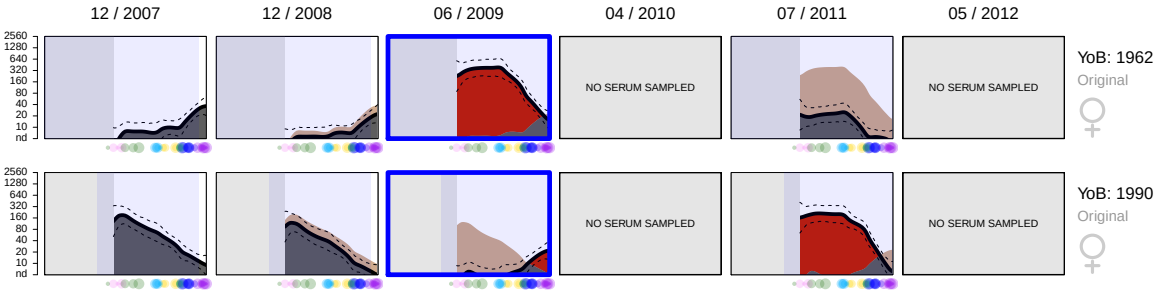
Switched



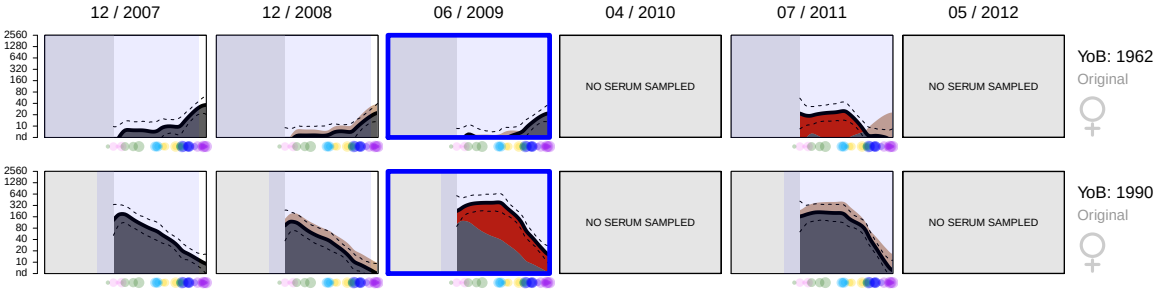
Swap 4 - Household 312

A potential sample swap in 2009 for two individuals within the same household.

Original



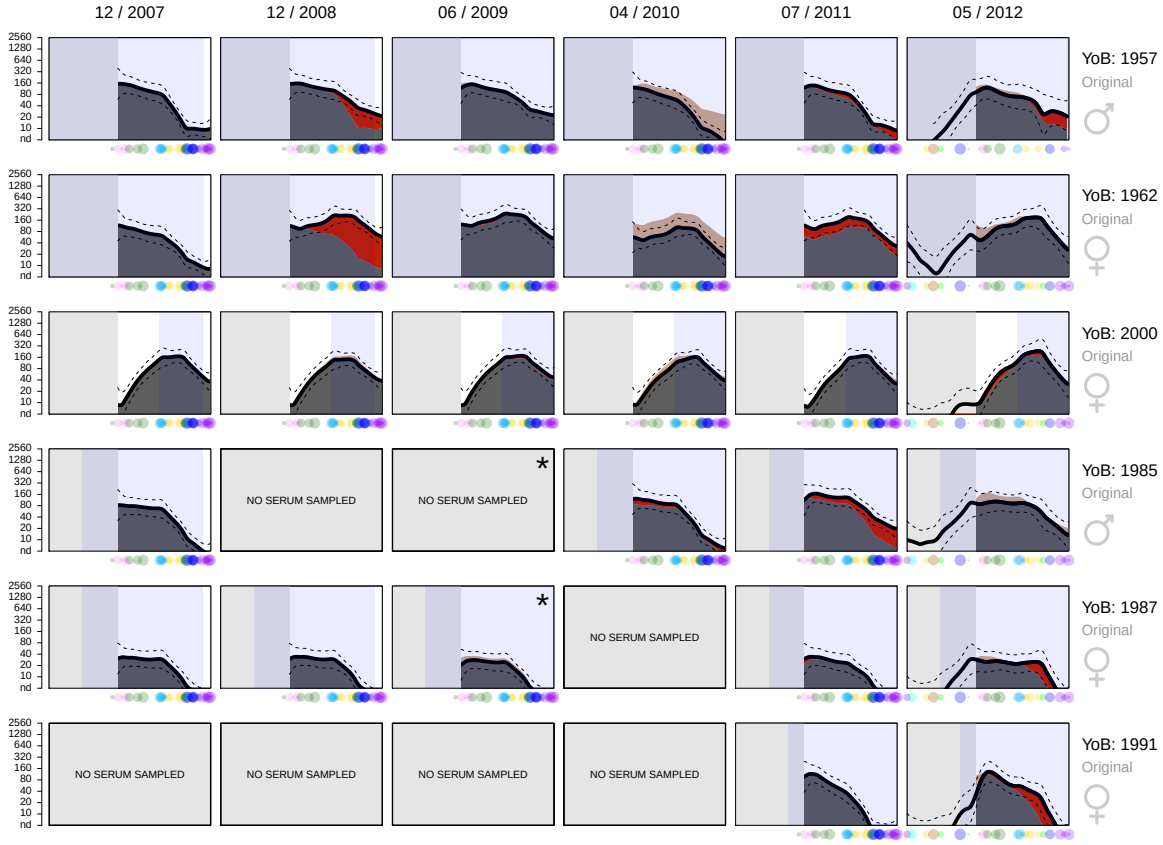
Switched



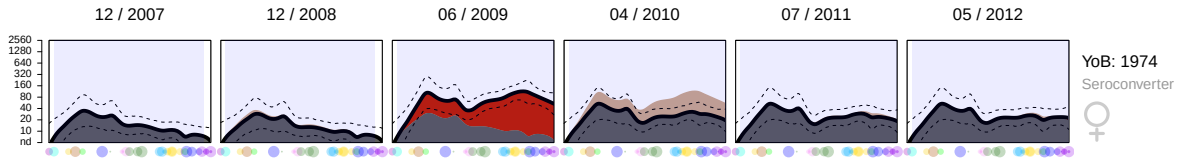
C.1.2 *Landscapes by household*

Antibody landscapes for serum samples analysed from individuals in the Hanam cohort are shown, split by household. Cases where samples were switched to correct for what was seen to be a likely sample swap are marked with an asterix.

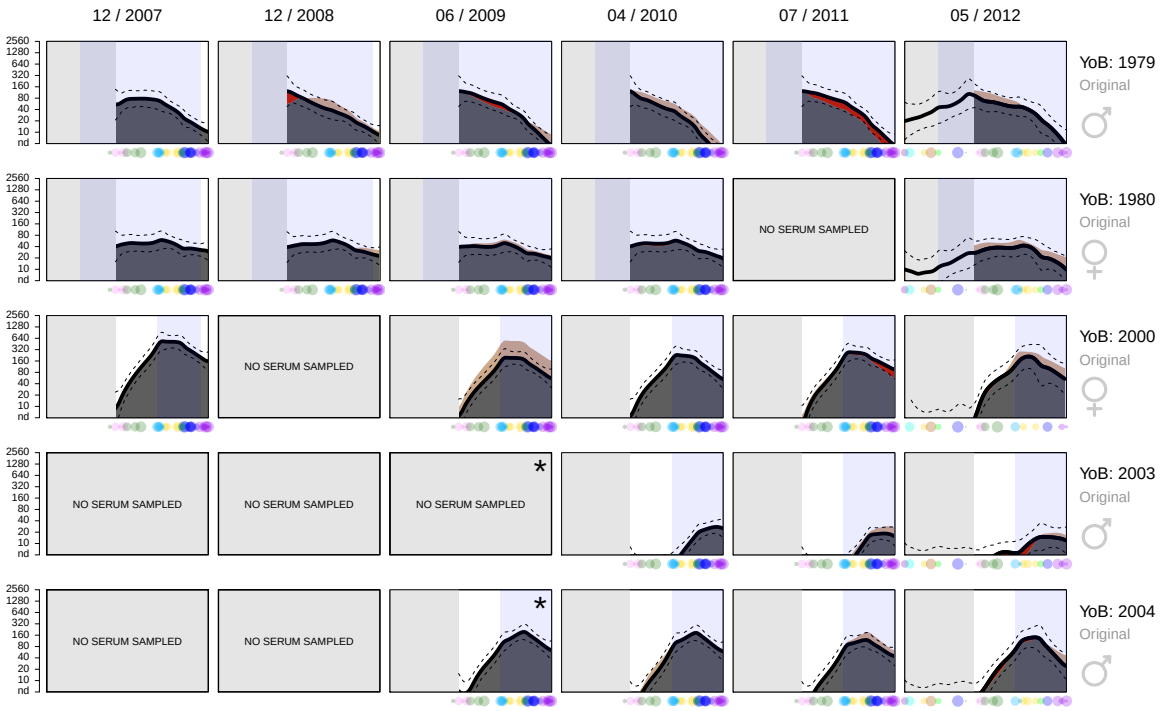
Household 004



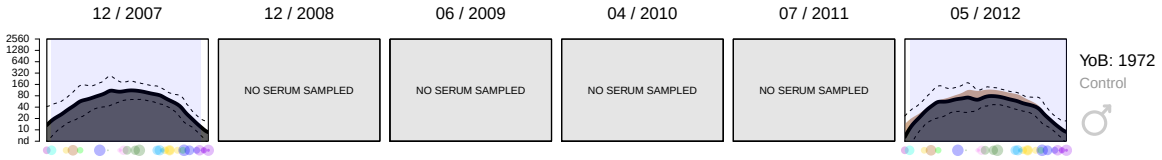
Household 010



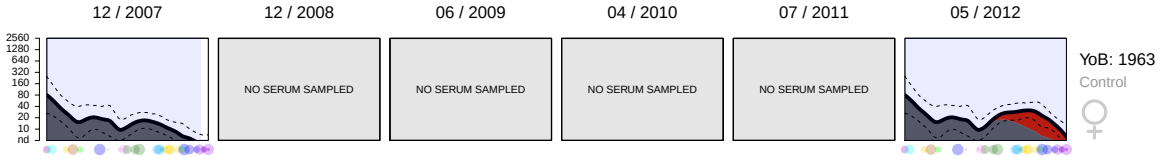
Household 006



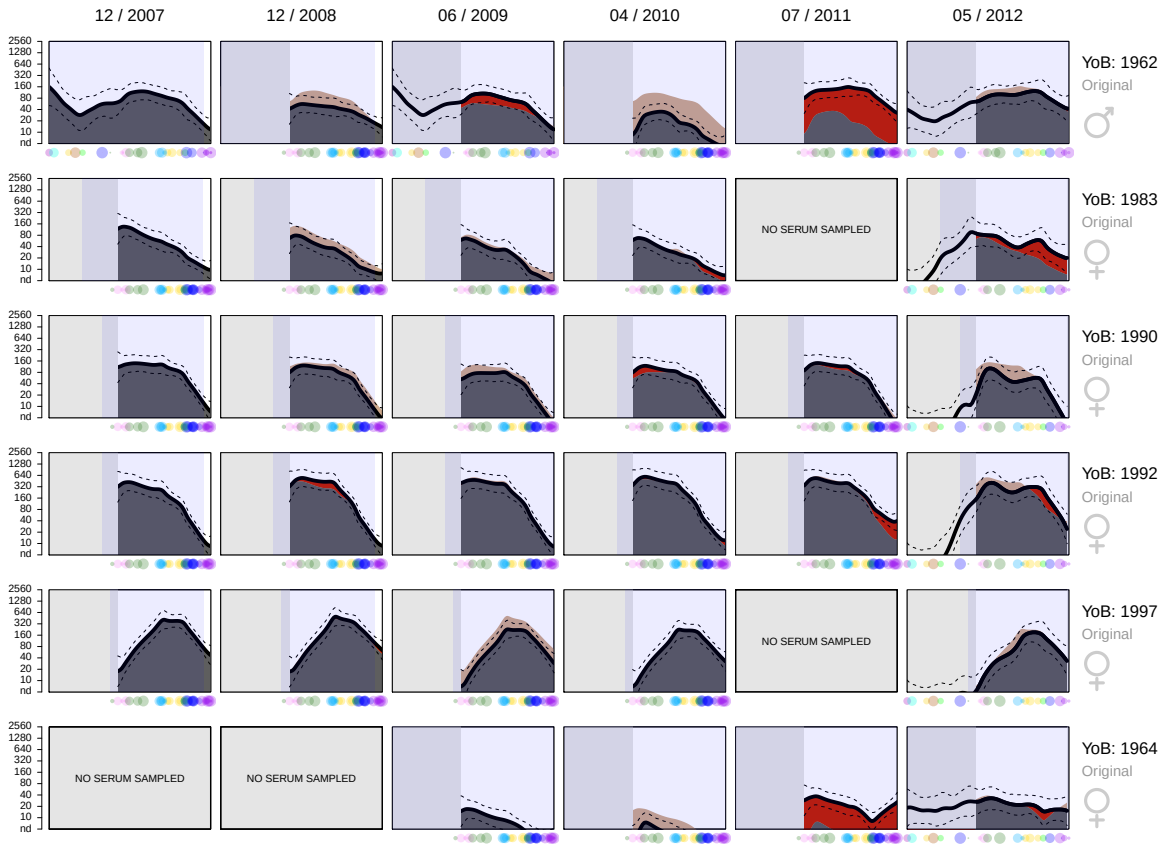
Household 011



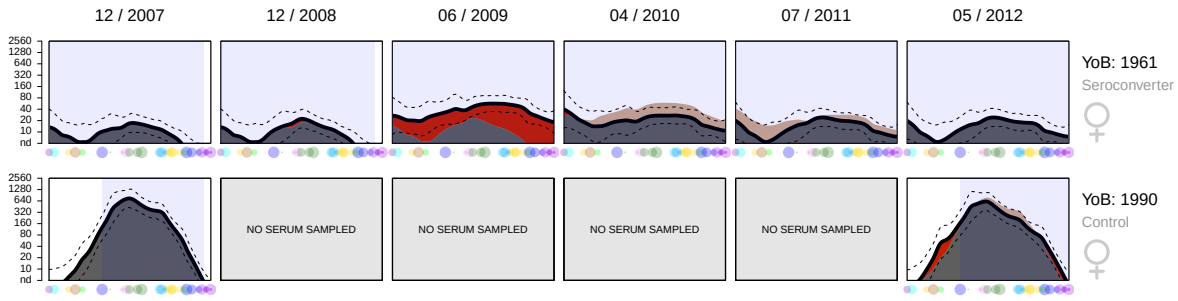
Household 028



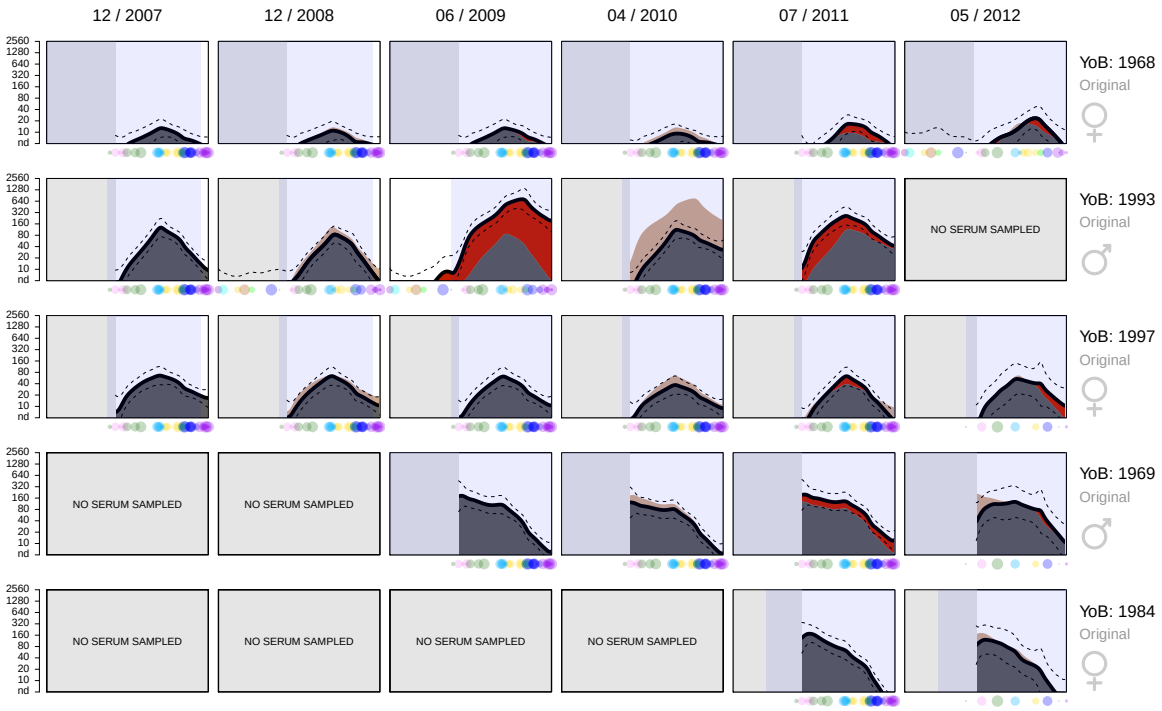
Household 055



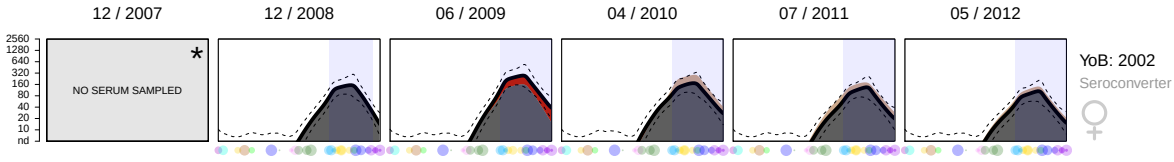
Household 097



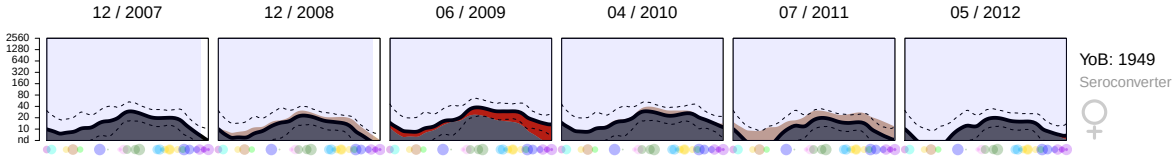
Household 284



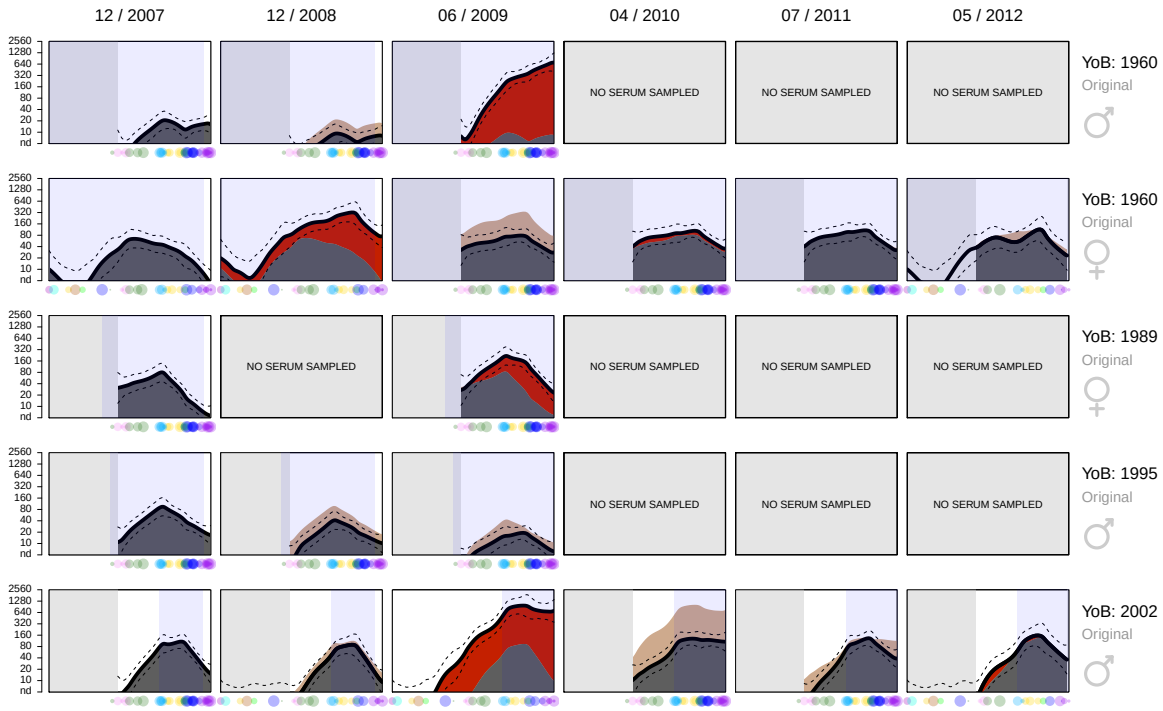
Household 047



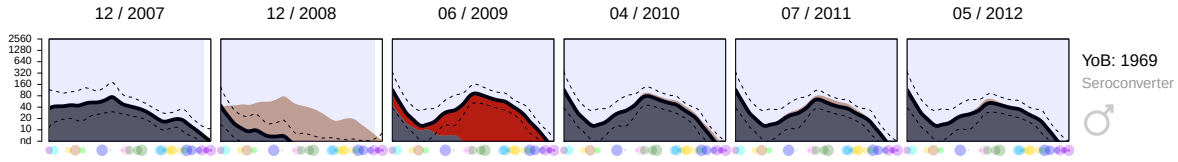
Household 060



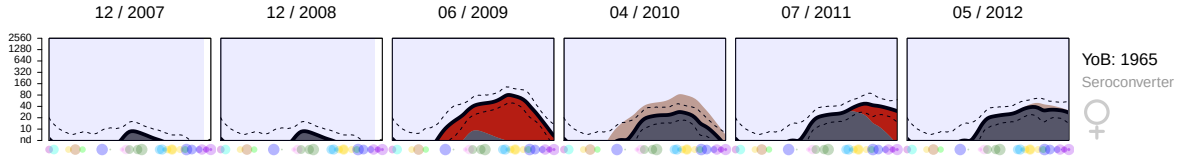
Household 273



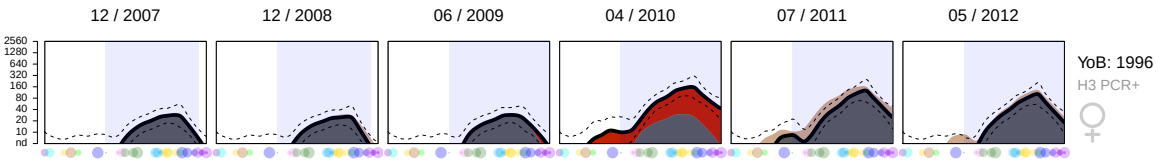
Household 054



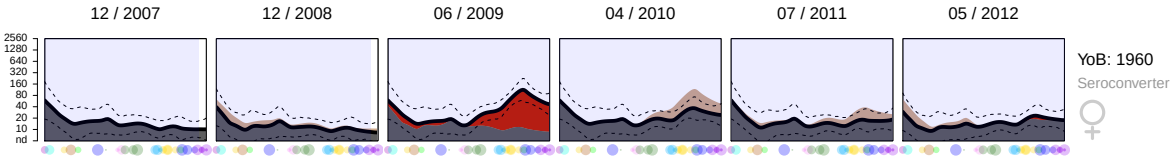
Household 082



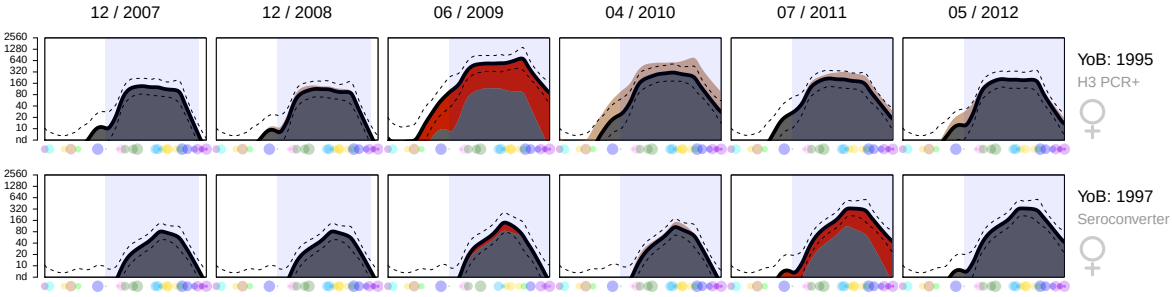
Household 107



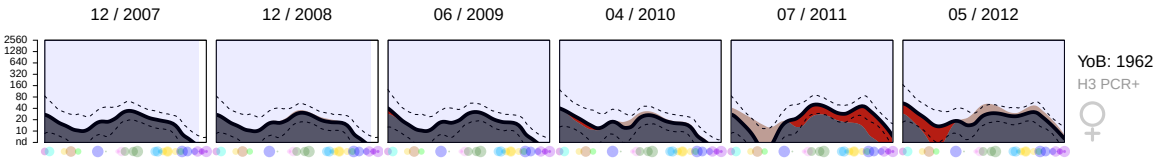
Household 117



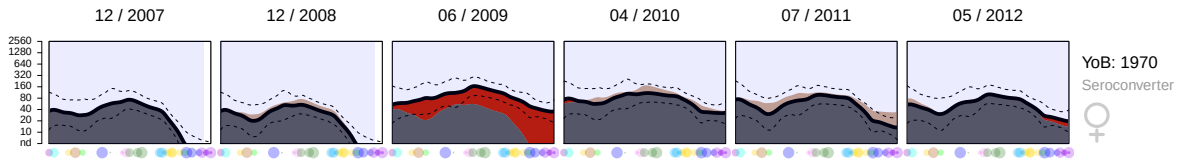
Household 134



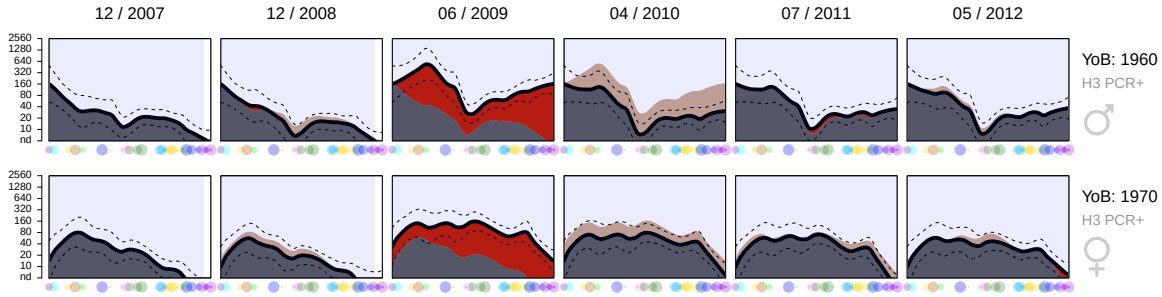
Household 138



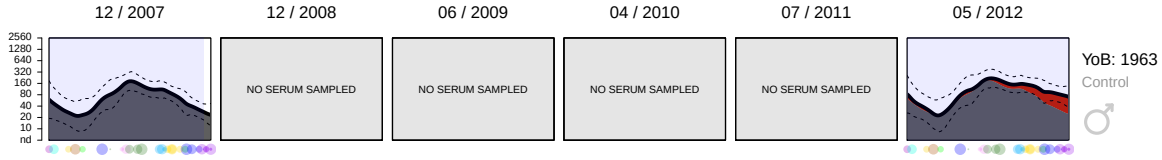
Household 140



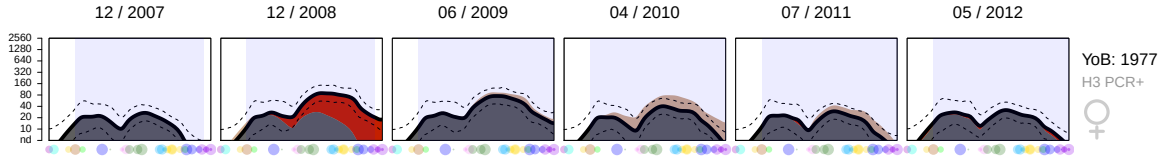
Household 145



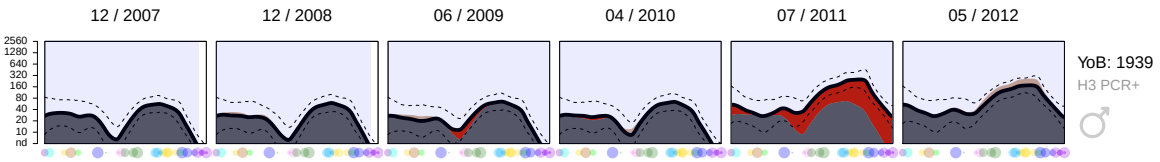
Household 147



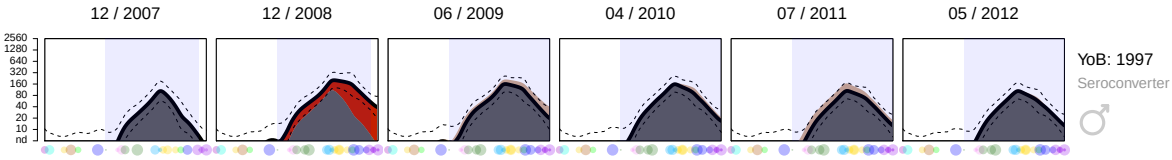
Household 148



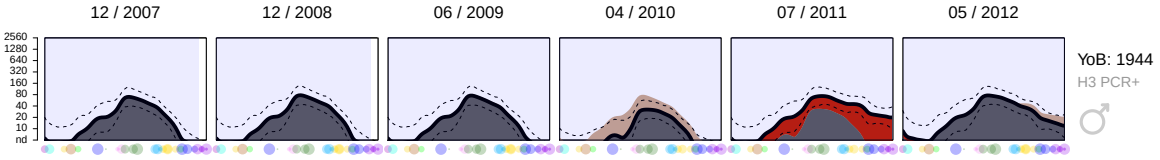
Household 150



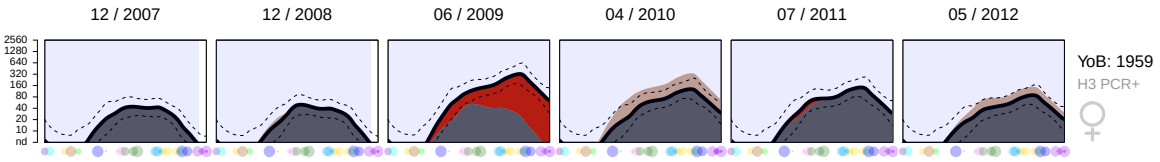
Household 173



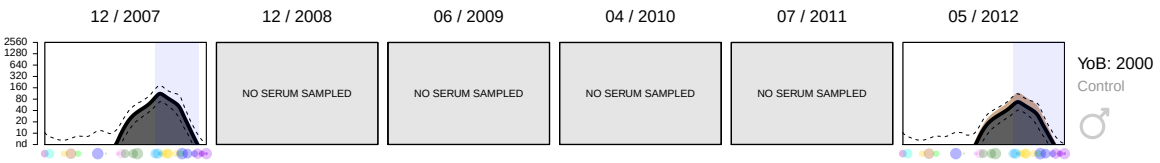
Household 181



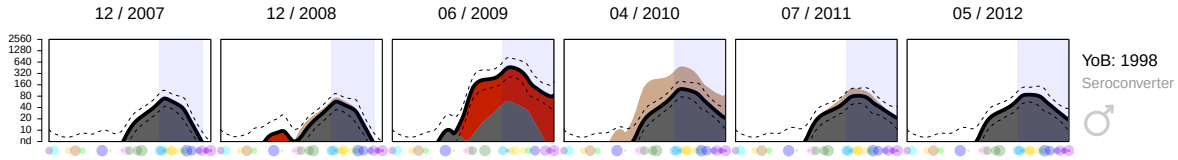
Household 188



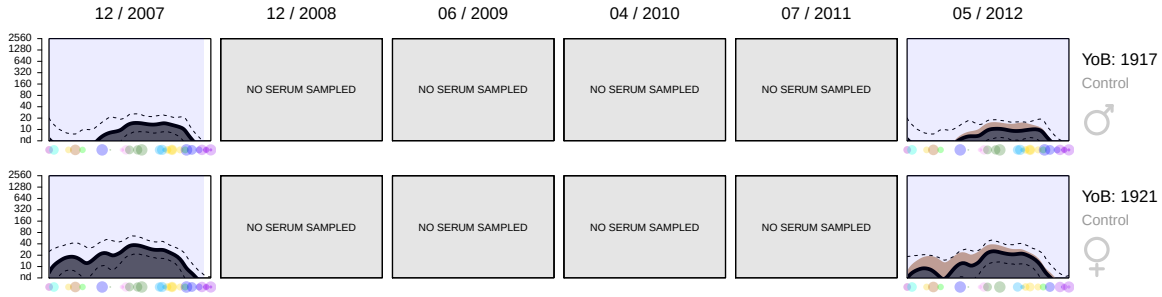
Household 193



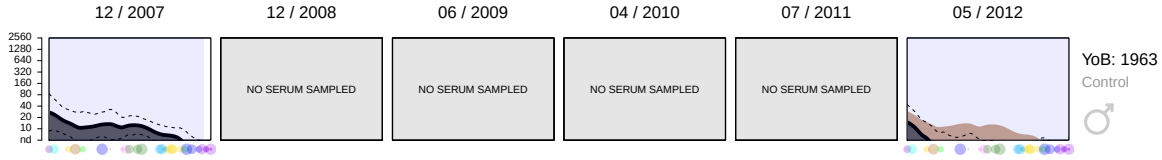
Household 212



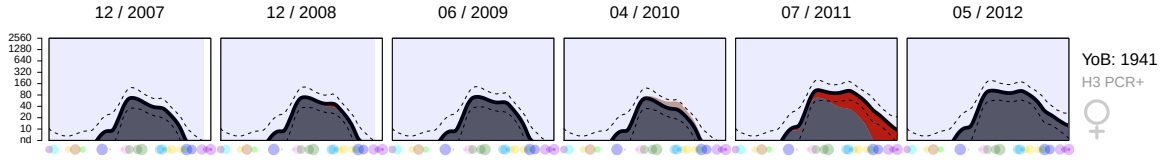
Household 231



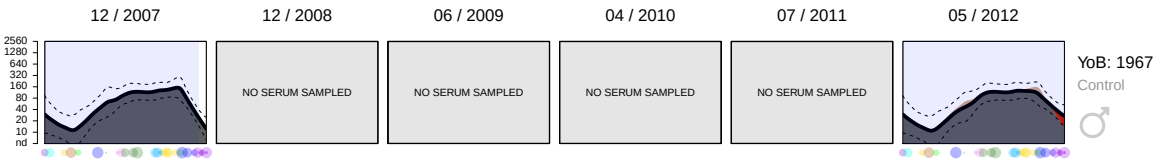
Household 240



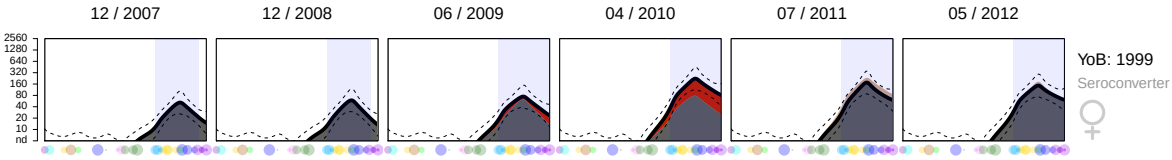
Household 246



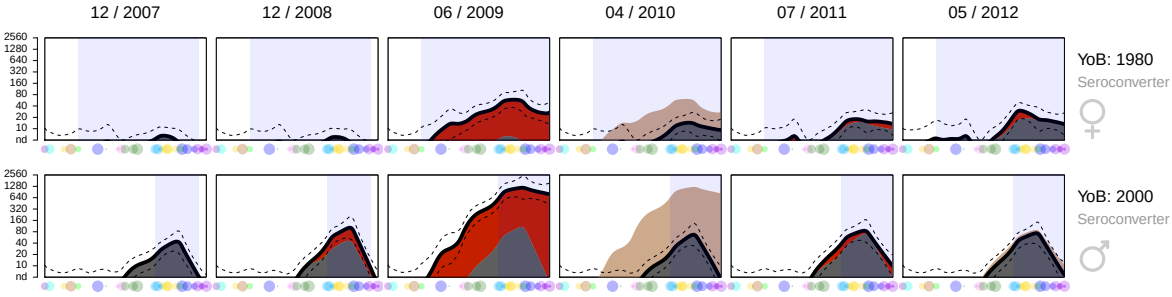
Household 252



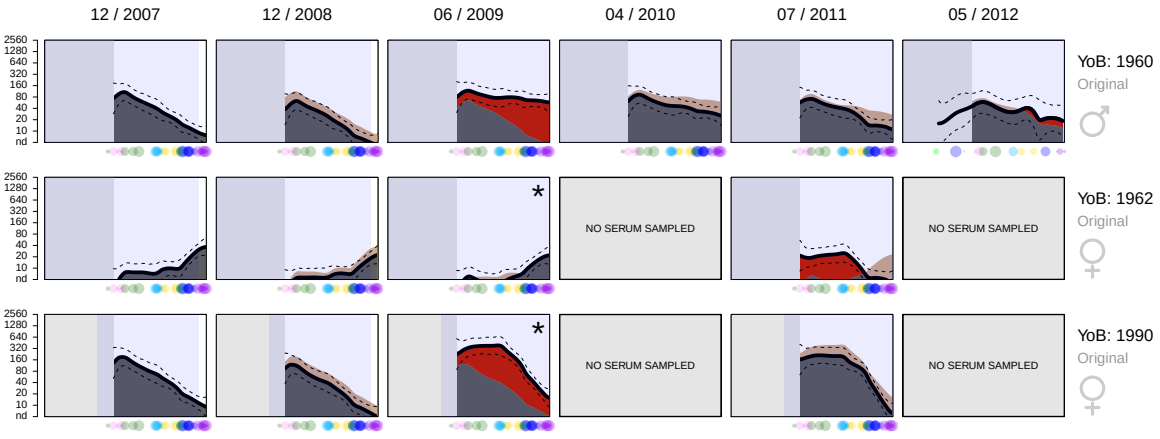
Household 256

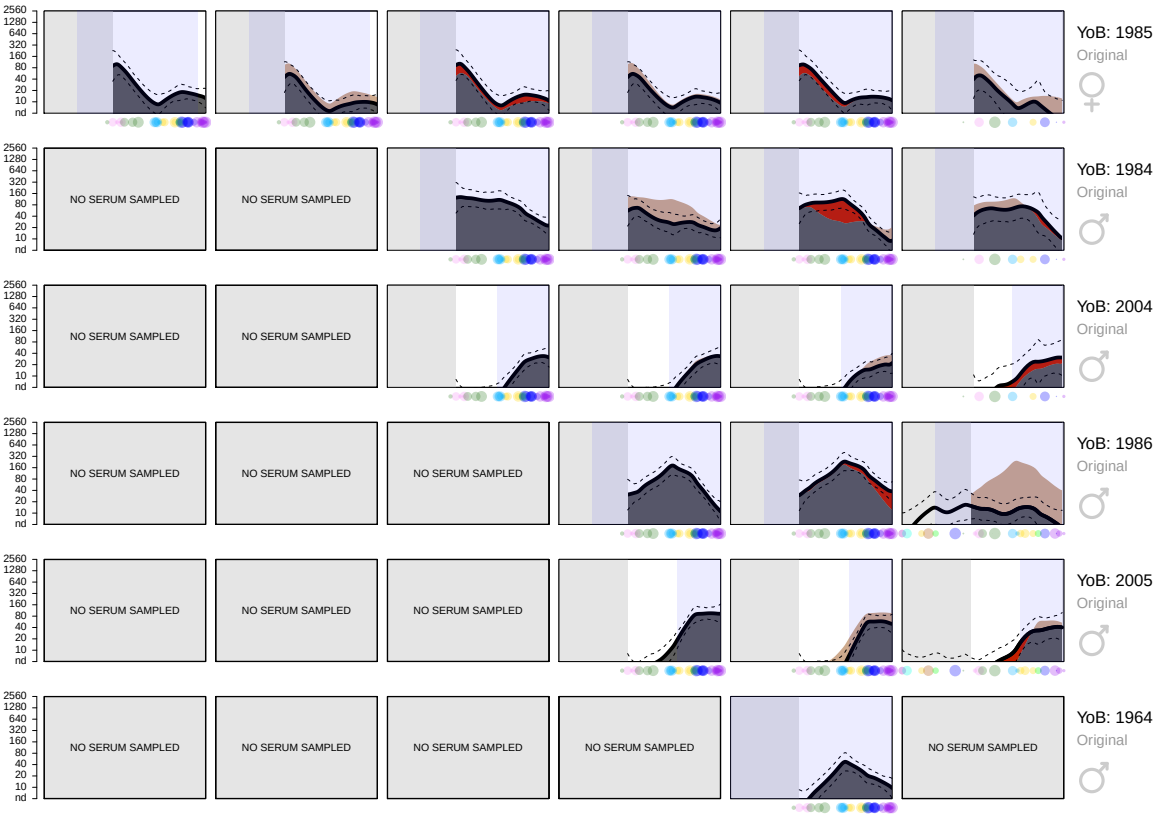


Household 283



Household 312

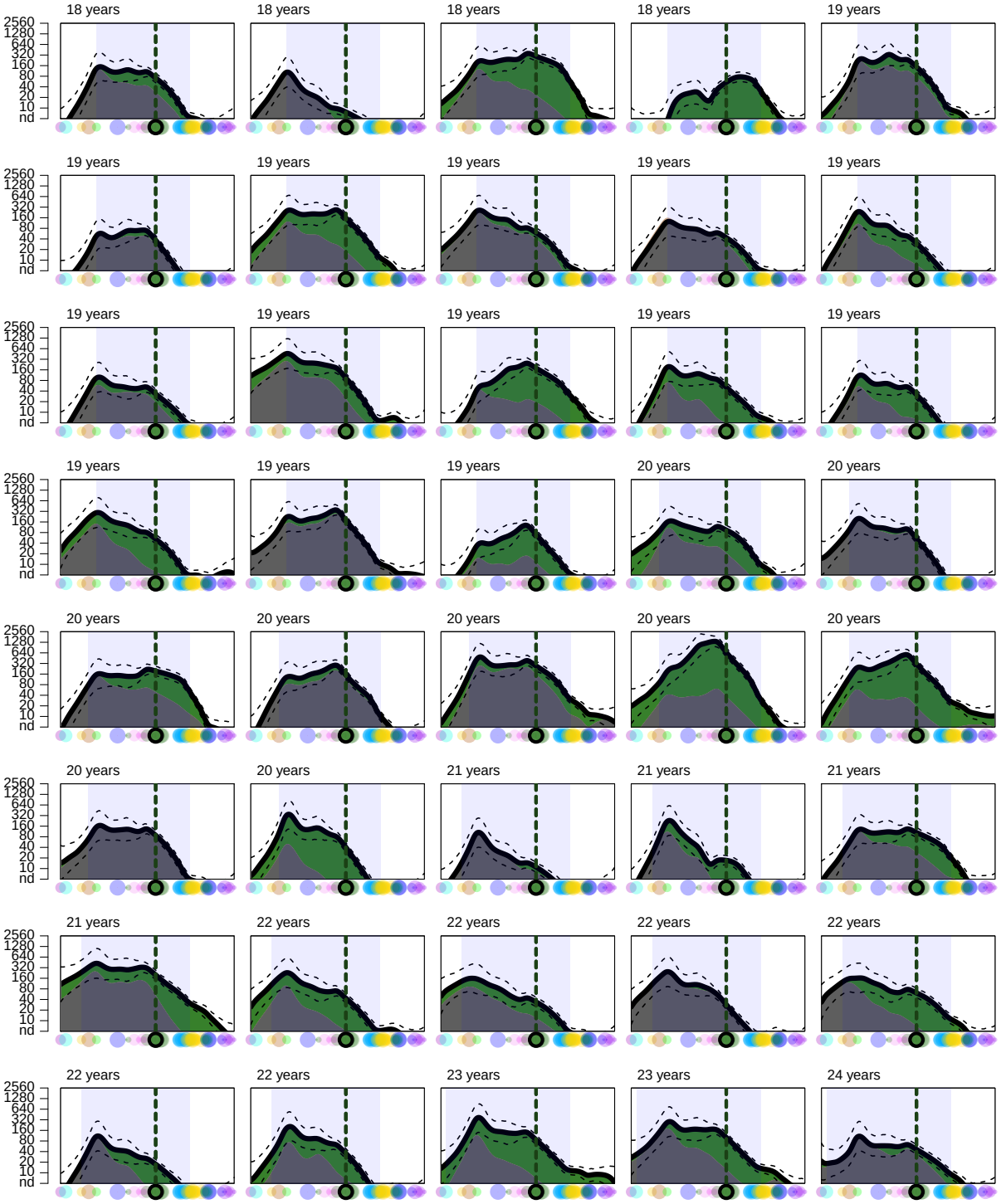


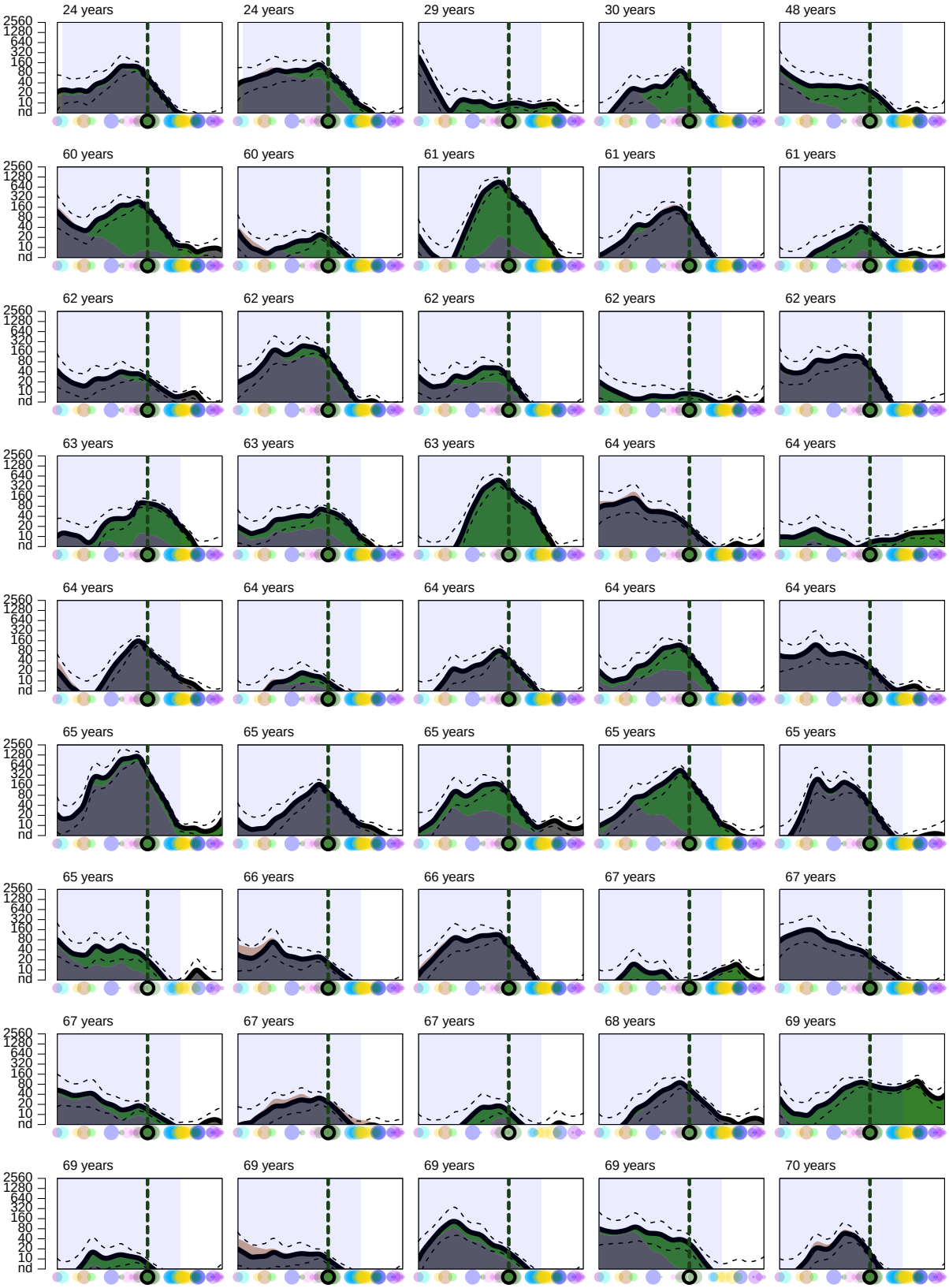

2560
1280
640
320
160
80
40
20
10
nd
NO SERUM SAMPLED
YoB: 1985
Original
YoB: 1984
Original
YoB: 2004
Original
YoB: 1986
Original
YoB: 2005
Original
YoB: 1964
Original

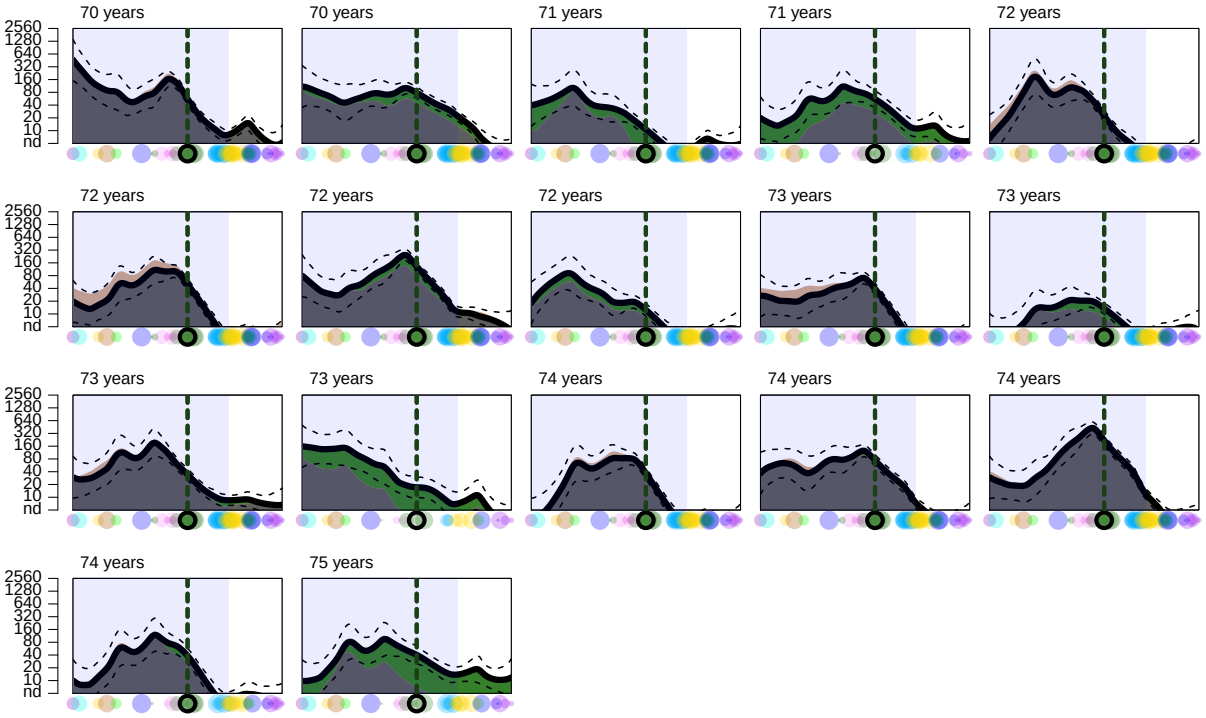
C.2 Antibody landscapes for the vaccination cohorts

Shared caption for the figures in this section: Antibody landscapes are shown for all individuals in the vaccine trial where enough serum was available to perform titrations against at least 20 different H3N2 strains, ordered by age of the individual. Unlike the plots of the Hanam and Mangua cohorts, each panel shows the antibody response for a different individual, comparing the post-vaccination antibody landscape to the one measured pre-vaccination. For each plot, the solid black line represents the landscape height along the antigenic summary path defined in chapter 2. Increases in antibody reactivity compared to the pre-vaccination sample are shown coloured either green (1997 vaccine trial), light blue (1998 vaccine trial), dark blue (2009 vaccine trial) or purple (2010 vaccine trial), while any beige regions represent a decrease. Dark grey regions represent areas where pre and post-vaccination antibody landscapes overlap. Additionally shown are blue-shaded rectangles that indicate antigenic clusters that had circulated during an individual's life span until sample collection, the age of the individual in years is given above the plots. Dots along the x axis indicate the antigenic positions along the summary path of the subset of viruses used to generate each landscape, scaled negatively by distance from it. Finally, in each panel the antigenic location of the H3N2 vaccine strain (1997 = A/Nanchang/933/1995, 1998 = A/Sydney/5/1997, 2009 = A/Brisbane/10/2007, 2010 = A/Perth/16/2009) that the individual received is highlighted on the x-axis and marked with a vertical dotted line.

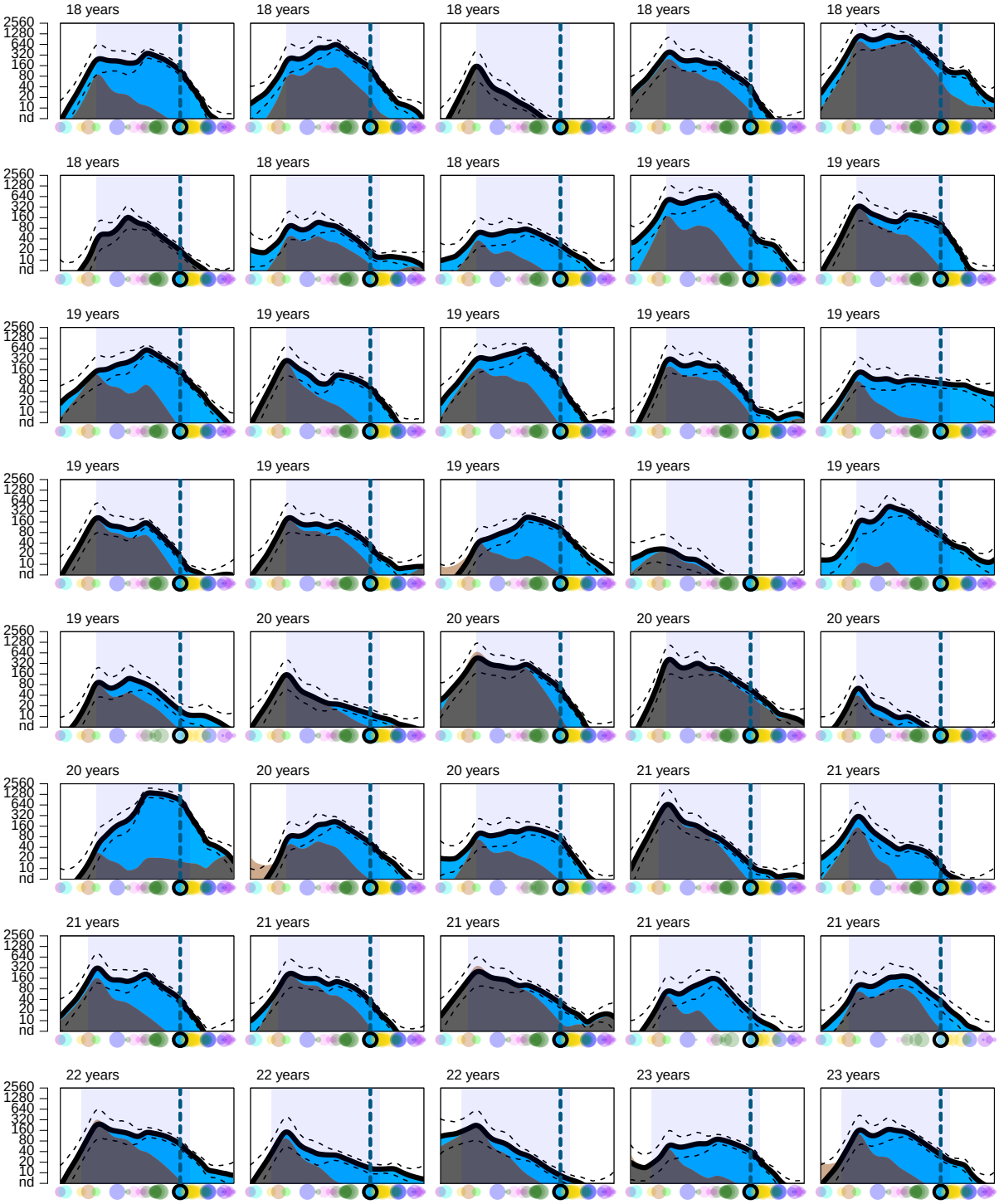
C.2.1 *1997 vaccination trial*

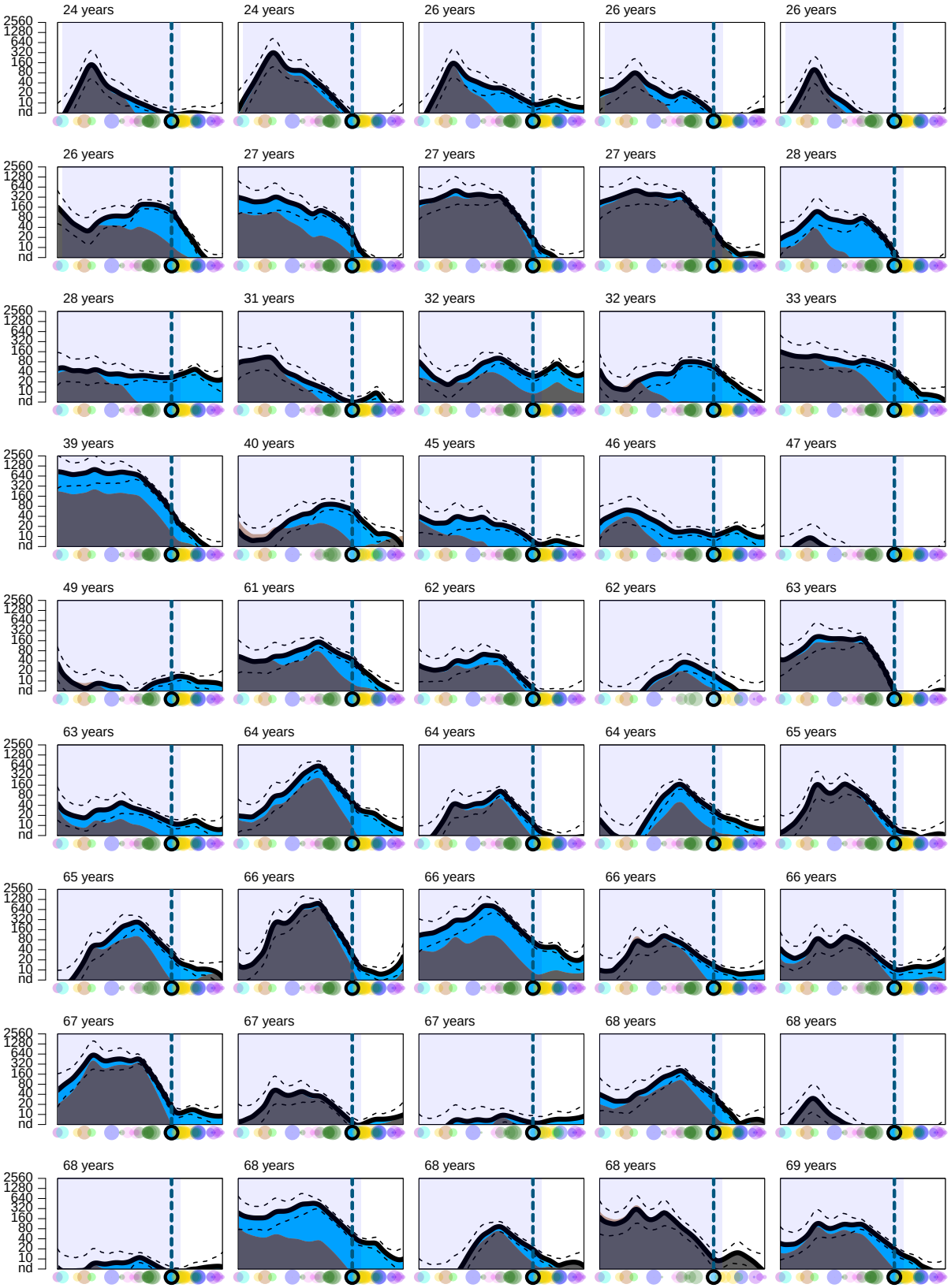


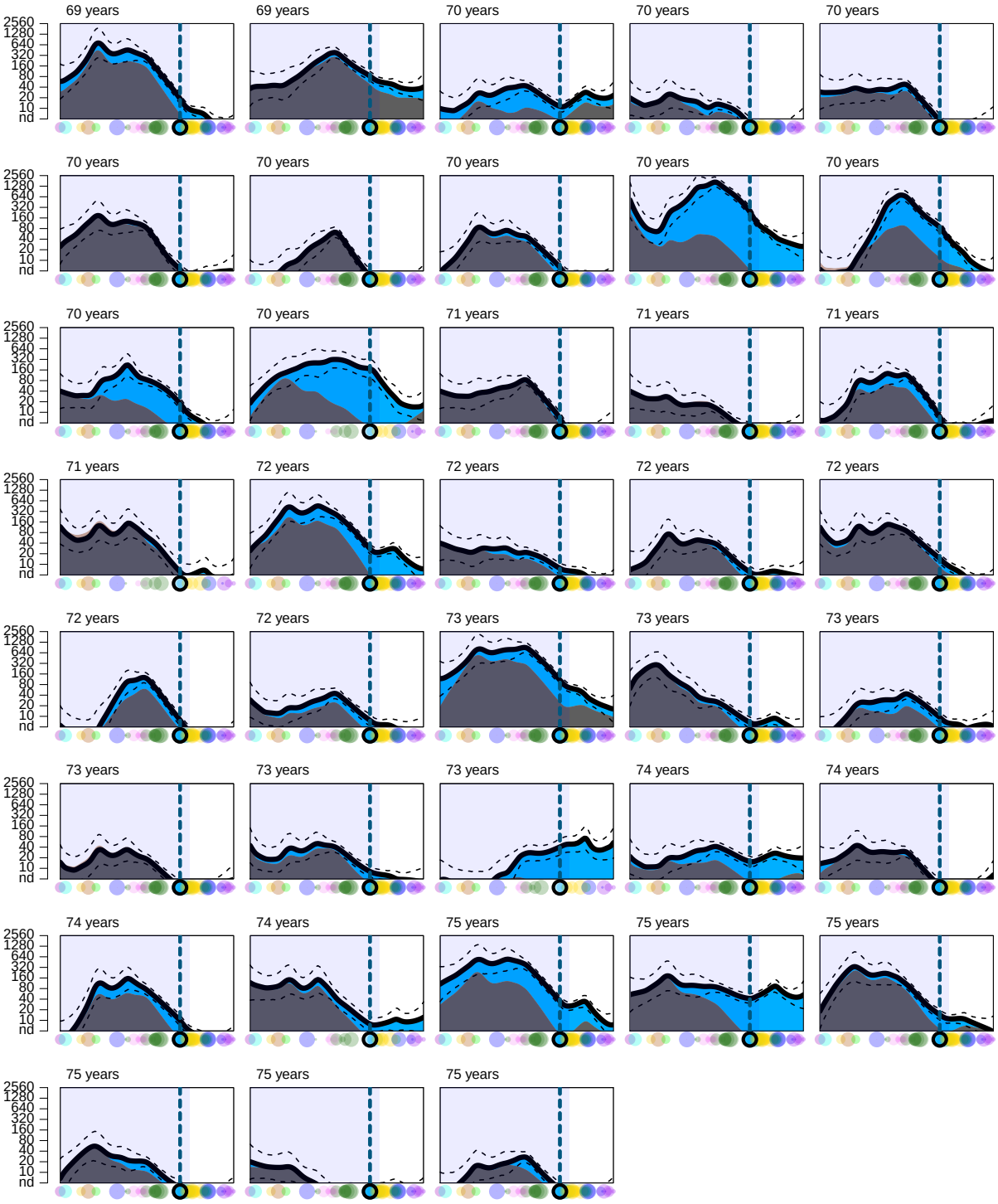

24 years
24 years
29 years
30 years
48 years
60 years
60 years
61 years
61 years
61 years
62 years
62 years
62 years
62 years
62 years
63 years
63 years
63 years
64 years
64 years
64 years
64 years
64 years
64 years
64 years
65 years
65 years
65 years
65 years
65 years
65 years
66 years
66 years
67 years
67 years
67 years
67 years
67 years
68 years
69 years
69 years
69 years
69 years
69 years
70 years
2560
1280
640
320
160
80
40
20
10
nd


70 years
70 years
71 years
71 years
72 years
72 years
72 years
72 years
73 years
73 years
73 years
73 years
74 years
74 years
74 years
74 years
75 years
2560
1280
640
320
160
80
40
20
10
nd

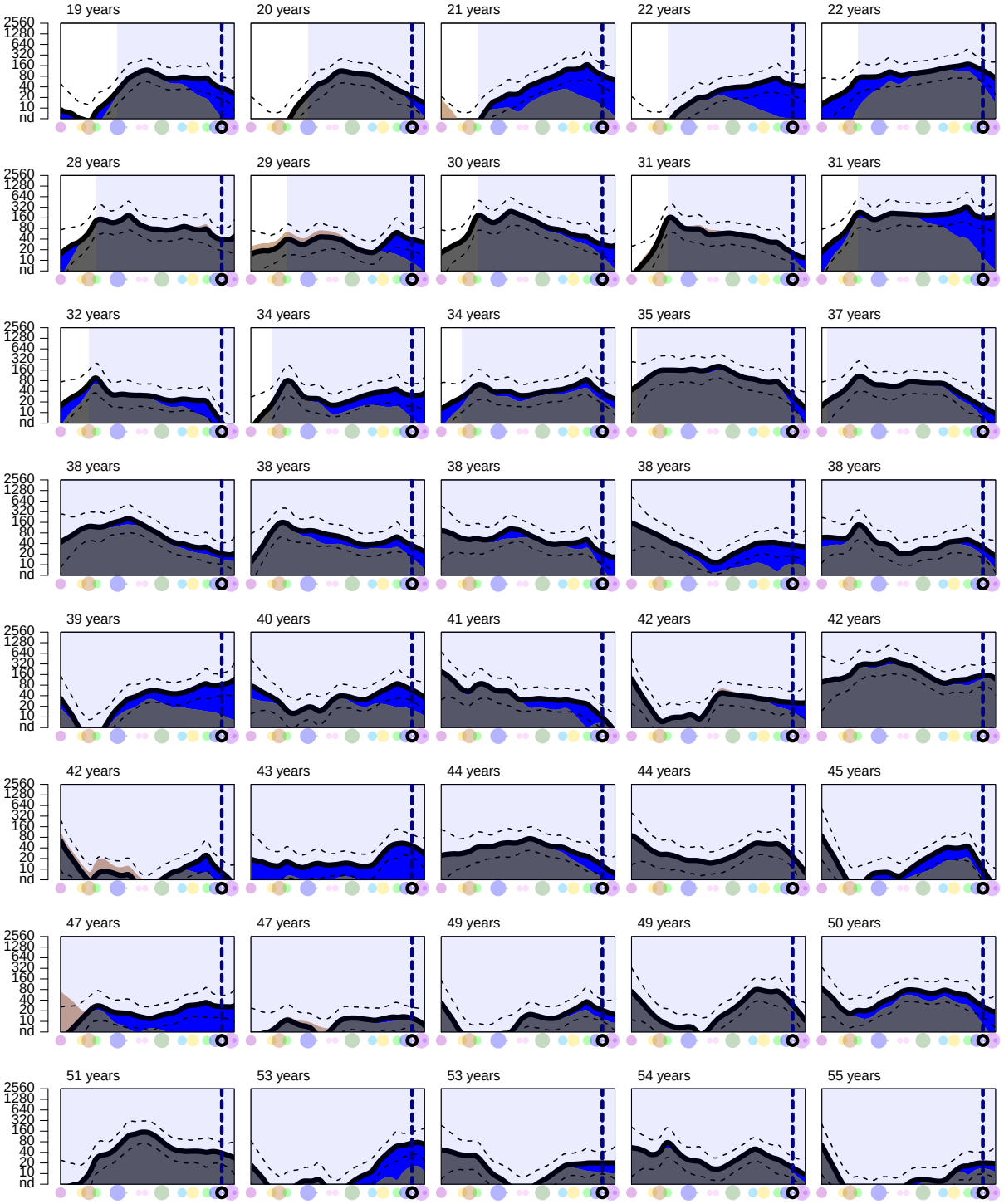
C.2.2 *1998 vaccination trial*

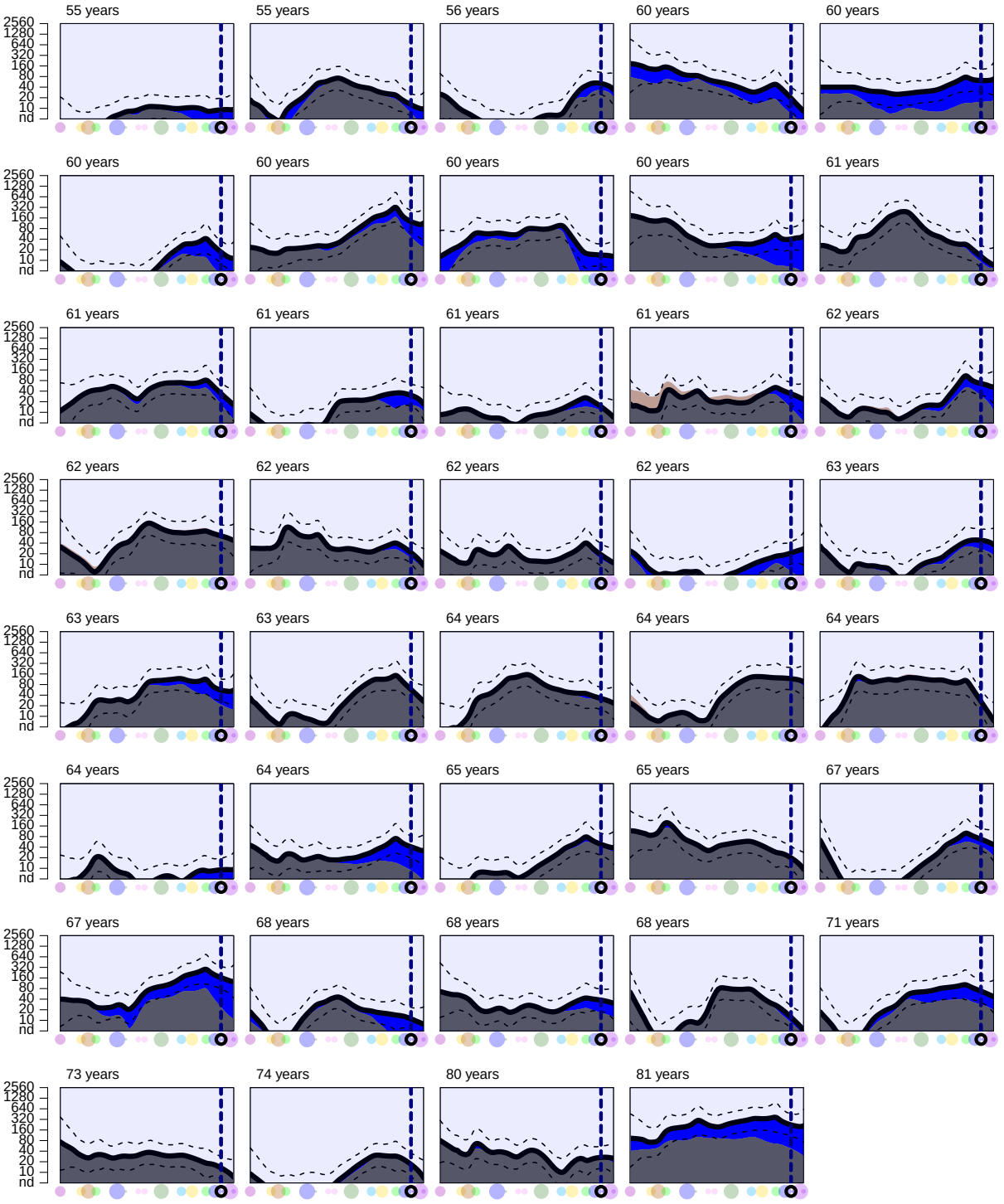



24 years
24 years
26 years
26 years
26 years
26 years
27 years
27 years
27 years
28 years
28 years
31 years
32 years
32 years
33 years
39 years
40 years
45 years
46 years
47 years
49 years
61 years
62 years
62 years
63 years
63 years
64 years
64 years
64 years
65 years
65 years
66 years
66 years
66 years
66 years
67 years
67 years
67 years
68 years
68 years
68 years
68 years
68 years
68 years
69 years
2560
1280
640
320
160
80
40
20
10
nd

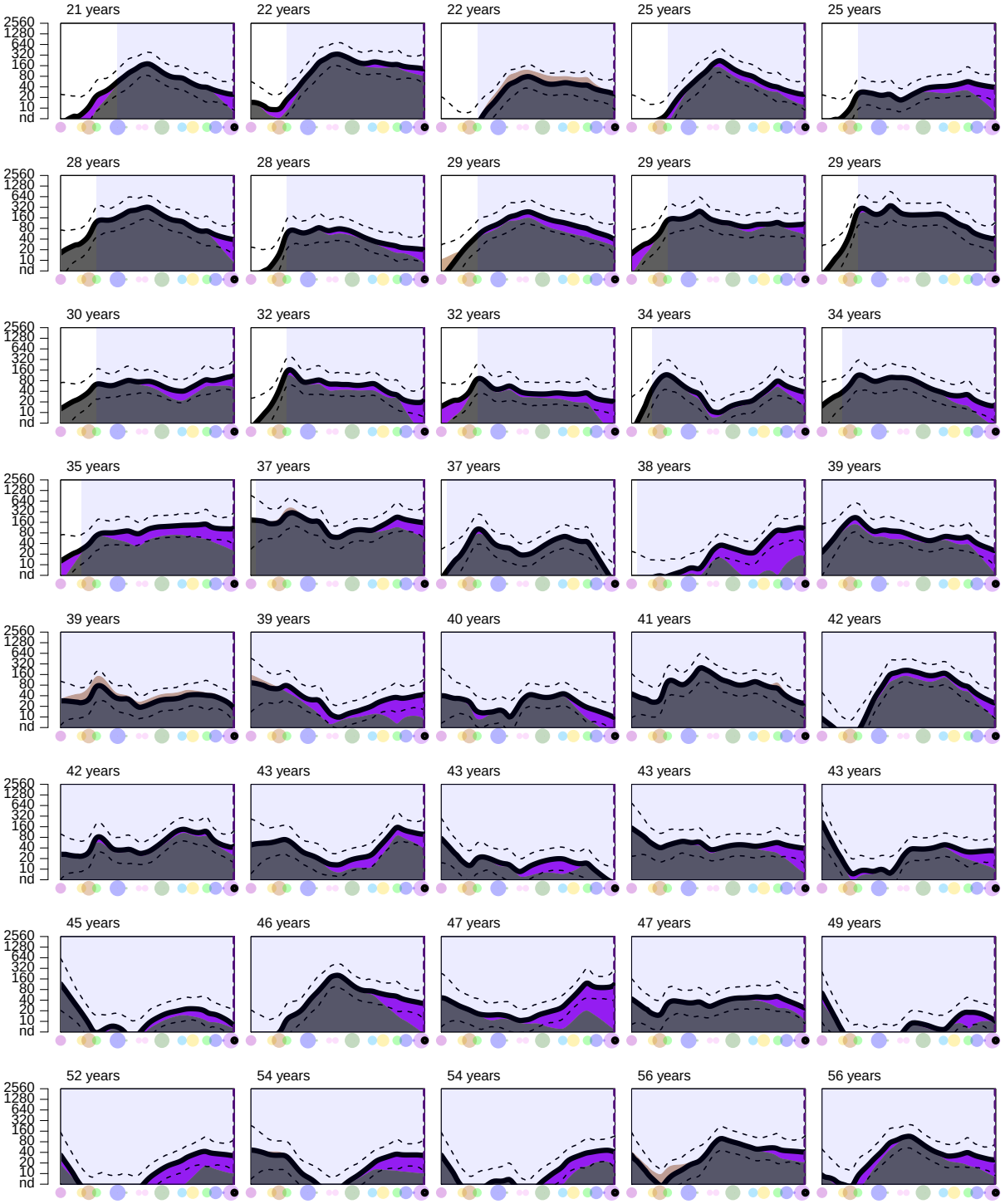

69 years
69 years
70 years
70 years
70 years
70 years
70 years
70 years
70 years
70 years
70 years
70 years
70 years
71 years
71 years
71 years
71 years
72 years
72 years
72 years
72 years
72 years
72 years
73 years
73 years
73 years
73 years
73 years
73 years
74 years
74 years
74 years
74 years
75 years
75 years
75 years
75 years
75 years
75 years
2560
1280
640
320
160
80
40
20
10
nd

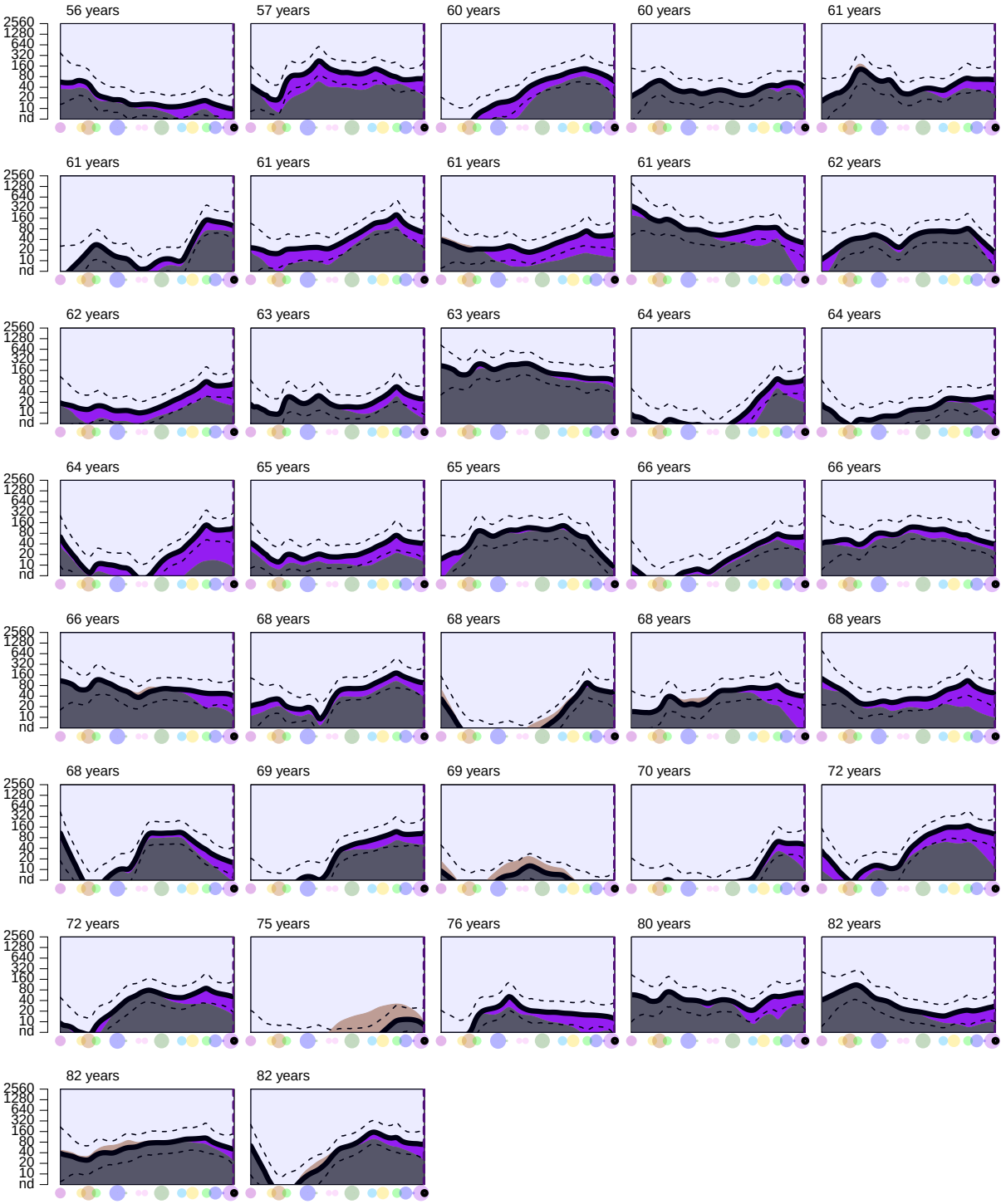
C.2.3 *2009 vaccination trial*




55 years
55 years
56 years
60 years
60 years
60 years
60 years
60 years
60 years
61 years
61 years
61 years
61 years
61 years
62 years
62 years
62 years
62 years
62 years
63 years
63 years
63 years
64 years
64 years
64 years
64 years
64 years
65 years
65 years
67 years
67 years
68 years
68 years
68 years
71 years
73 years
74 years
80 years
81 years
2560
1280
640
320
160
80
40
20
10
nd

C.2.4 *2010 vaccination trial*



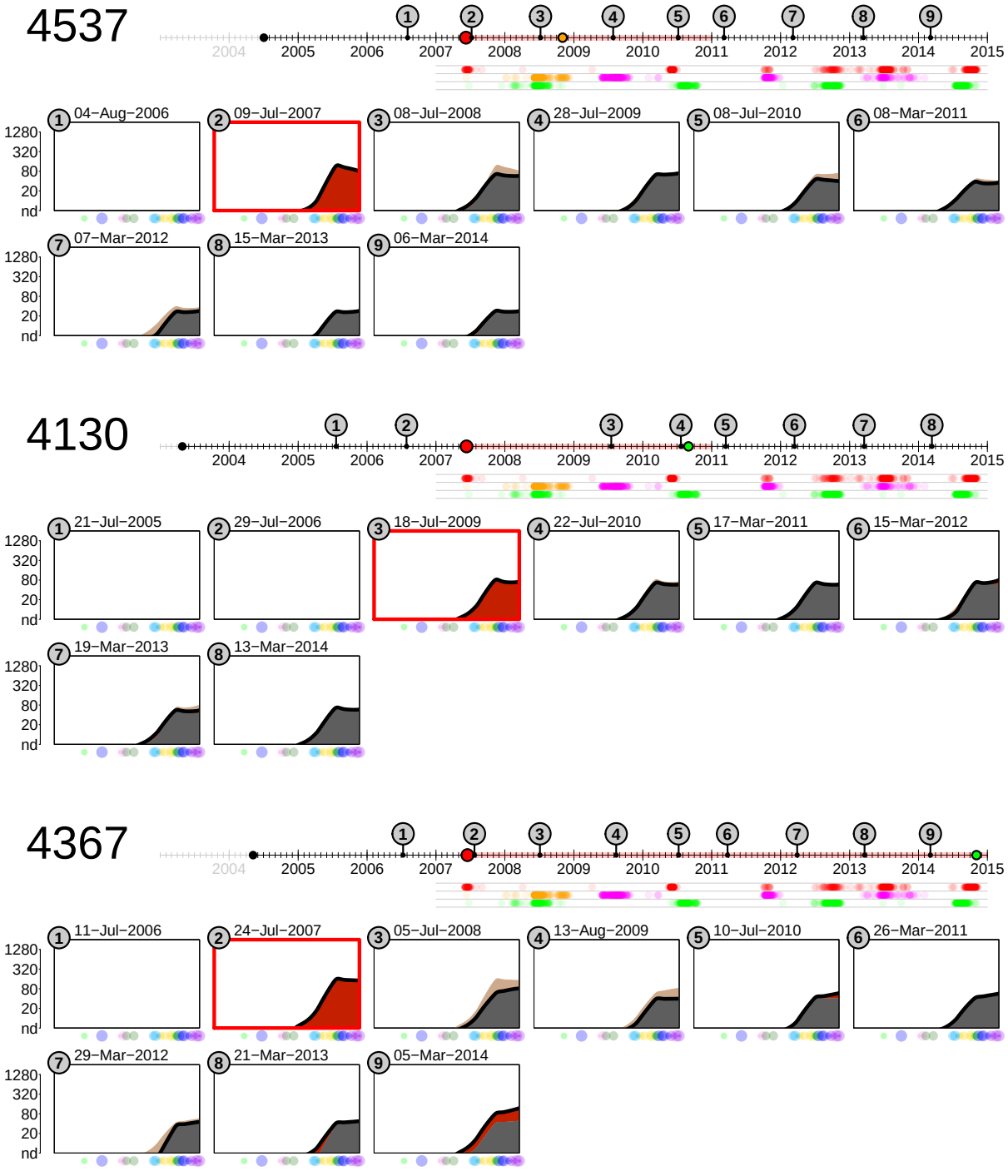

56 years
57 years
60 years
60 years
61 years
61 years
61 years
61 years
61 years
62 years
62 years
63 years
63 years
64 years
64 years
64 years
65 years
65 years
66 years
66 years
66 years
68 years
68 years
68 years
68 years
68 years
69 years
69 years
70 years
72 years
72 years
75 years
76 years
80 years
82 years
82 years
82 years
2560
1280
640
320
160
80
40
20
10
nd

C.3 Antibody landscapes for the Managua cohort study

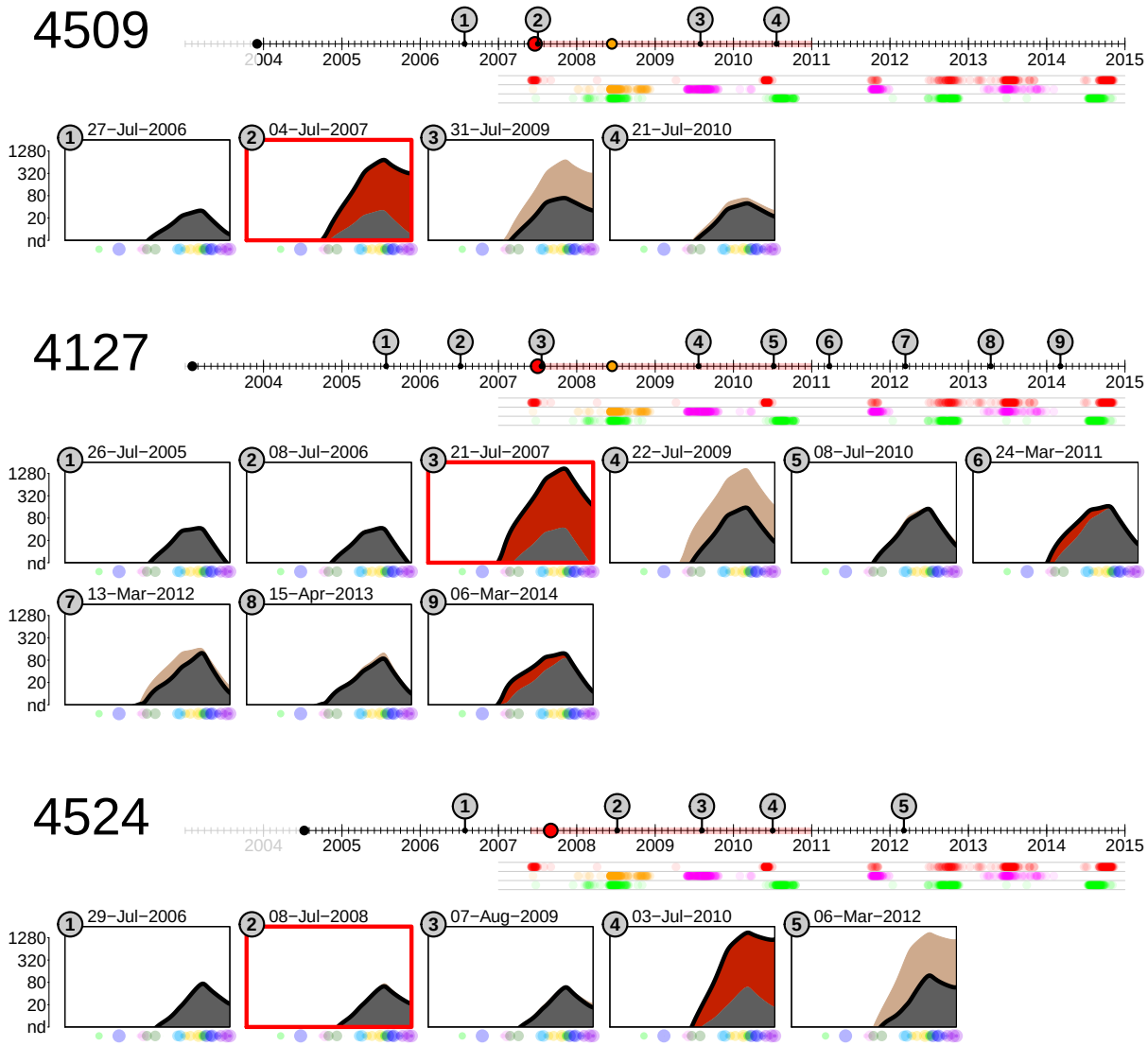
Shared caption for the figures in this section: At the top-left of the plot the subject ID is given for the child who's antibody responses are visualised. The top panel then shows a timeline of the child's life, with the date of birth marked as a large black circle to the left of the plot. Dates of serum sample collection are marked with a numbered circle and correspond to the panels below showing 2D-antibody landscape summaries constructed from HI measurements made against each serum sample. Also shown on the time line are coloured dots indicating dates of PCR-confirmed infections with different subtypes, red = H3N2, orange = H1N1, magenta = H1N1p (2009 pandemic) and green = influenza B. Red-shaded regions on the timeline mark the times at which the child was involved in a study cohort where they were actively monitored for influenza infection. Below the timeline, to give an idea of how many different influenza seasons a child has lived through, coloured circles mark the dates at which infections from the different subtypes were recorded in the wider cohort of all children studied in Managua. As described in chapter 2, the black line in each of the bottom panels represents the landscape height for each position on the antigenic summary path that was taken through the H3N2 antigenic clusters. The antibody landscape measured in the first sample taken is shown shaded in grey (if indeed visible, which it is not in many plots), in subsequent samples red shading indicates increases, and beige decreases, compared with antibody reactivity measured in the previous sample. Dots along the x axis indicate the antigenic positions along the summary path of the subset of viruses used to generate each landscape, scaled negatively by distance from it. The sample date of each serum sample is also given above each of the lower panels. Samples are split by season and infection type and ordered by date of the infection that occurred in that season.

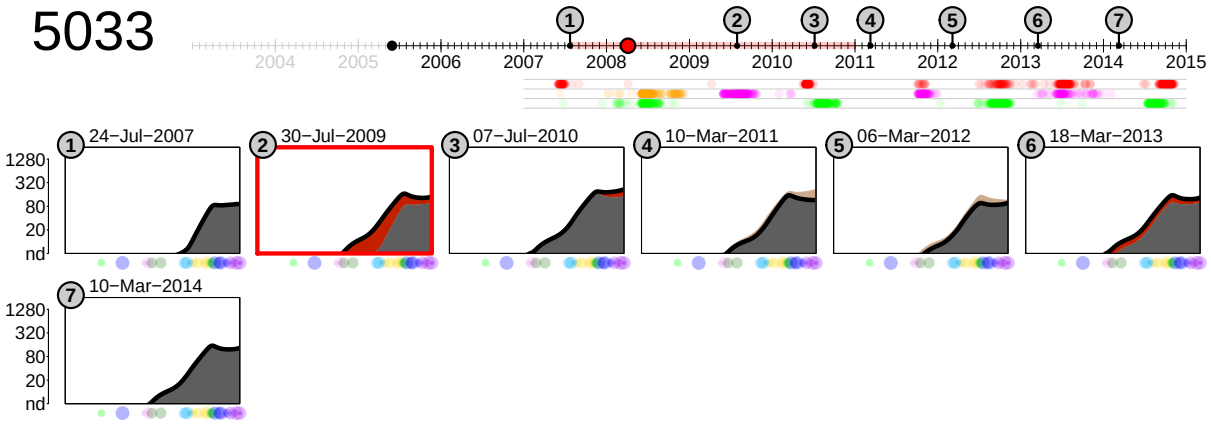
C.3.1 *Season 1*

Potential first infections



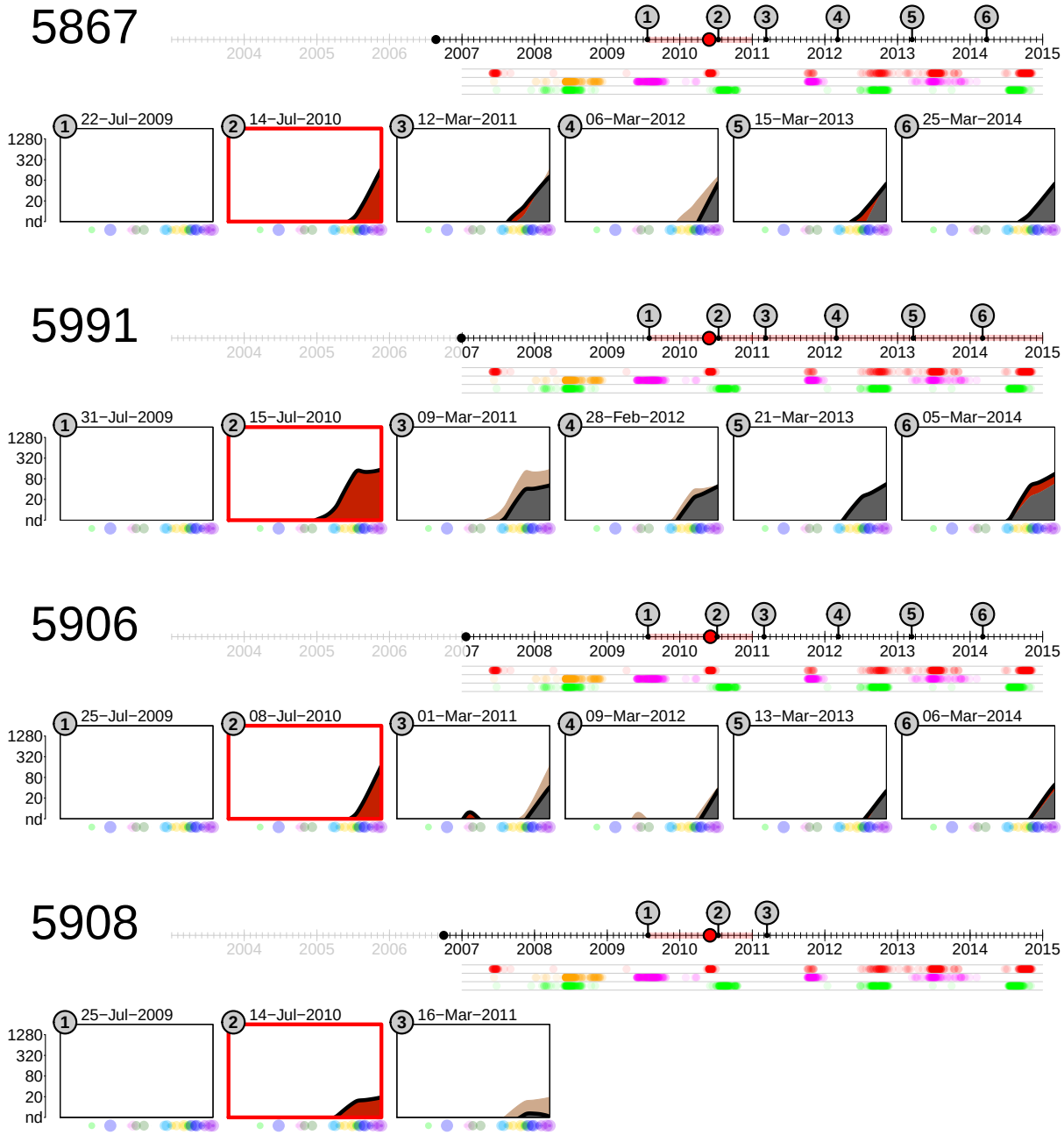
Potential second infections

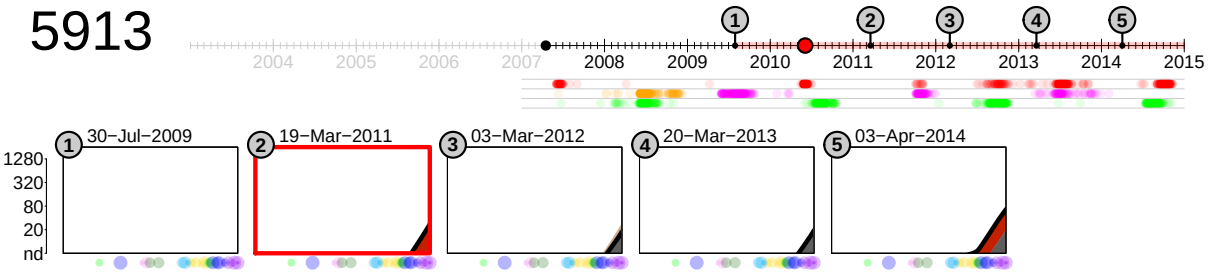


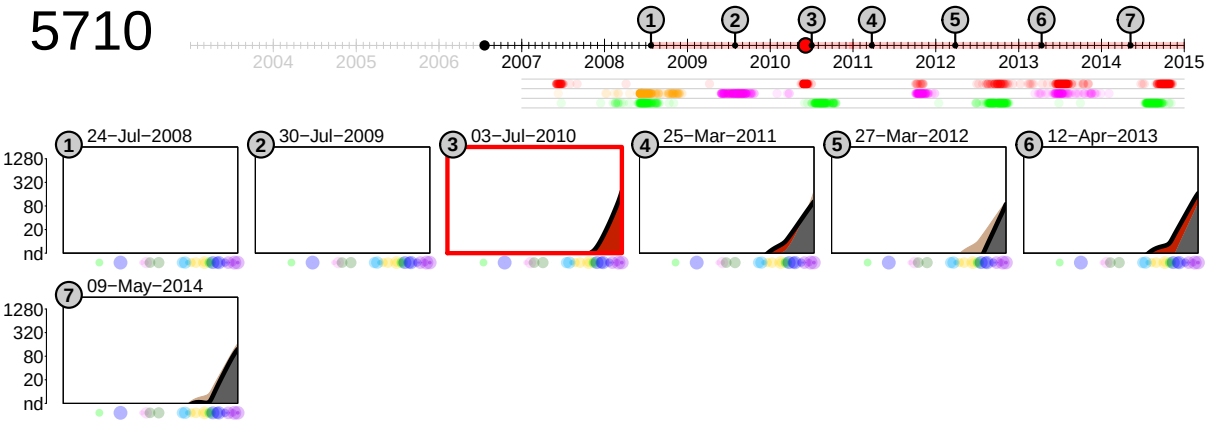

5033
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
5
6
7
1280
320
80
20
nd
24−Jul−2007
30−Jul−2009
07−Jul−2010
10−Mar−2011
06−Mar−2012
18−Mar−2013
10−Mar−2014

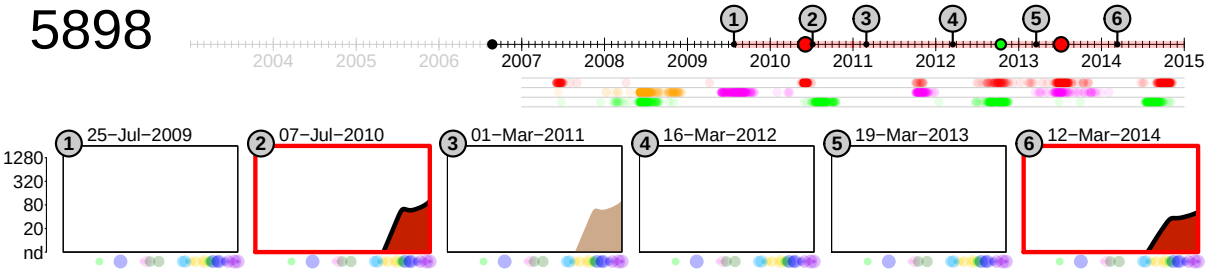
C.3.2 *Season 2*

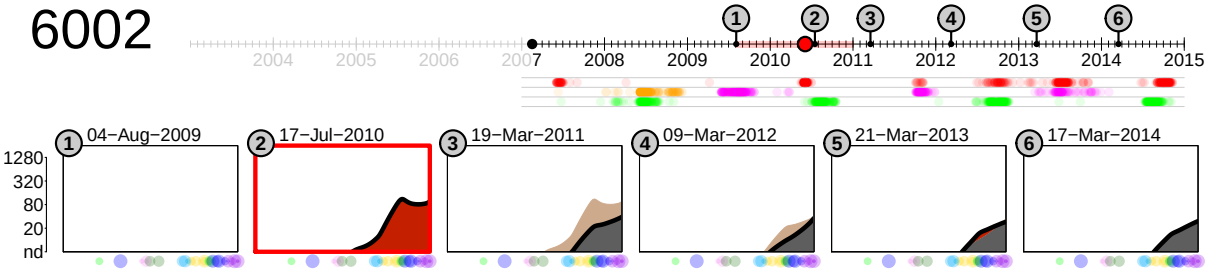
Potential first infections

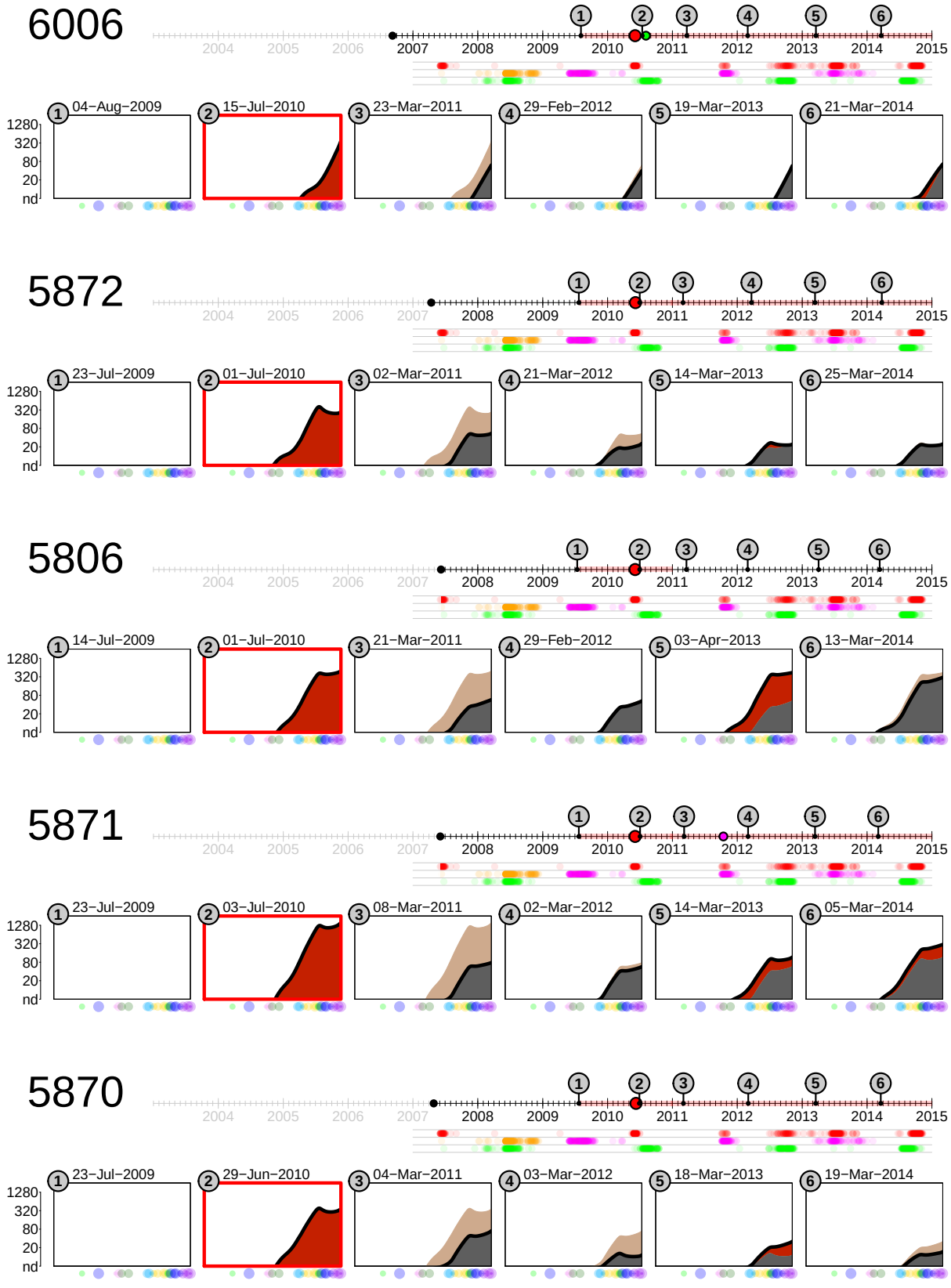


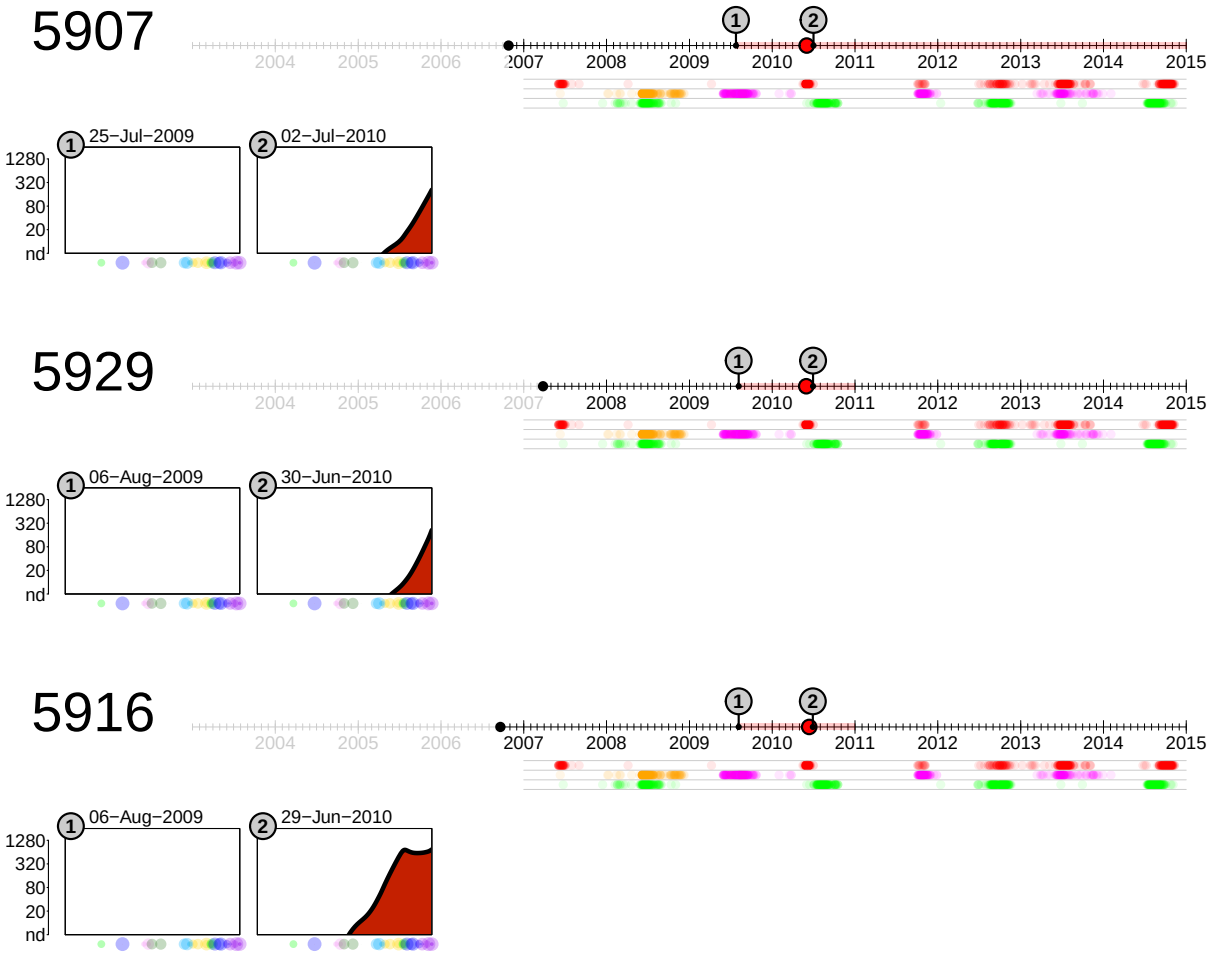

5913
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
5
1280
320
80
20
nd
1 30−Jul−2009
2 19−Mar−2011
3 03−Mar−2012
4 20−Mar−2013
5 03−Apr−2014


5710
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
5
6
7
1280
320
80
20
nd
1 24−Jul−2008
2 30−Jul−2009
3 03−Jul−2010
4 25−Mar−2011
5 27−Mar−2012
6 12−Apr−2013
7 09−May−2014

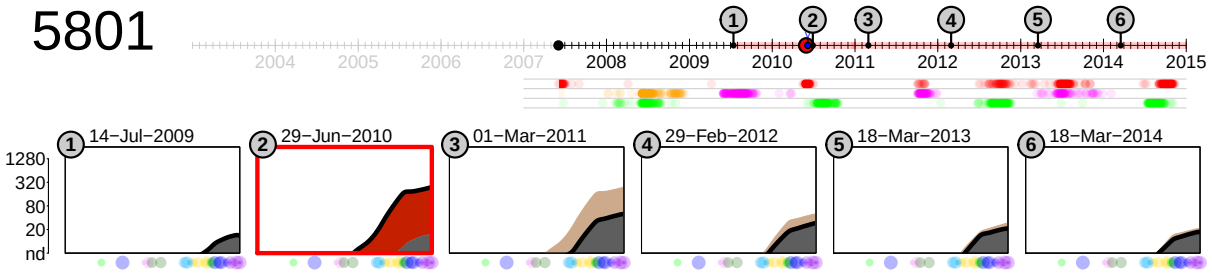

5898
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
5
6
1280
320
80
20
nd
1 25−Jul−2009
2 07−Jul−2010
3 01−Mar−2011
4 16−Mar−2012
5 19−Mar−2013
6 12−Mar−2014

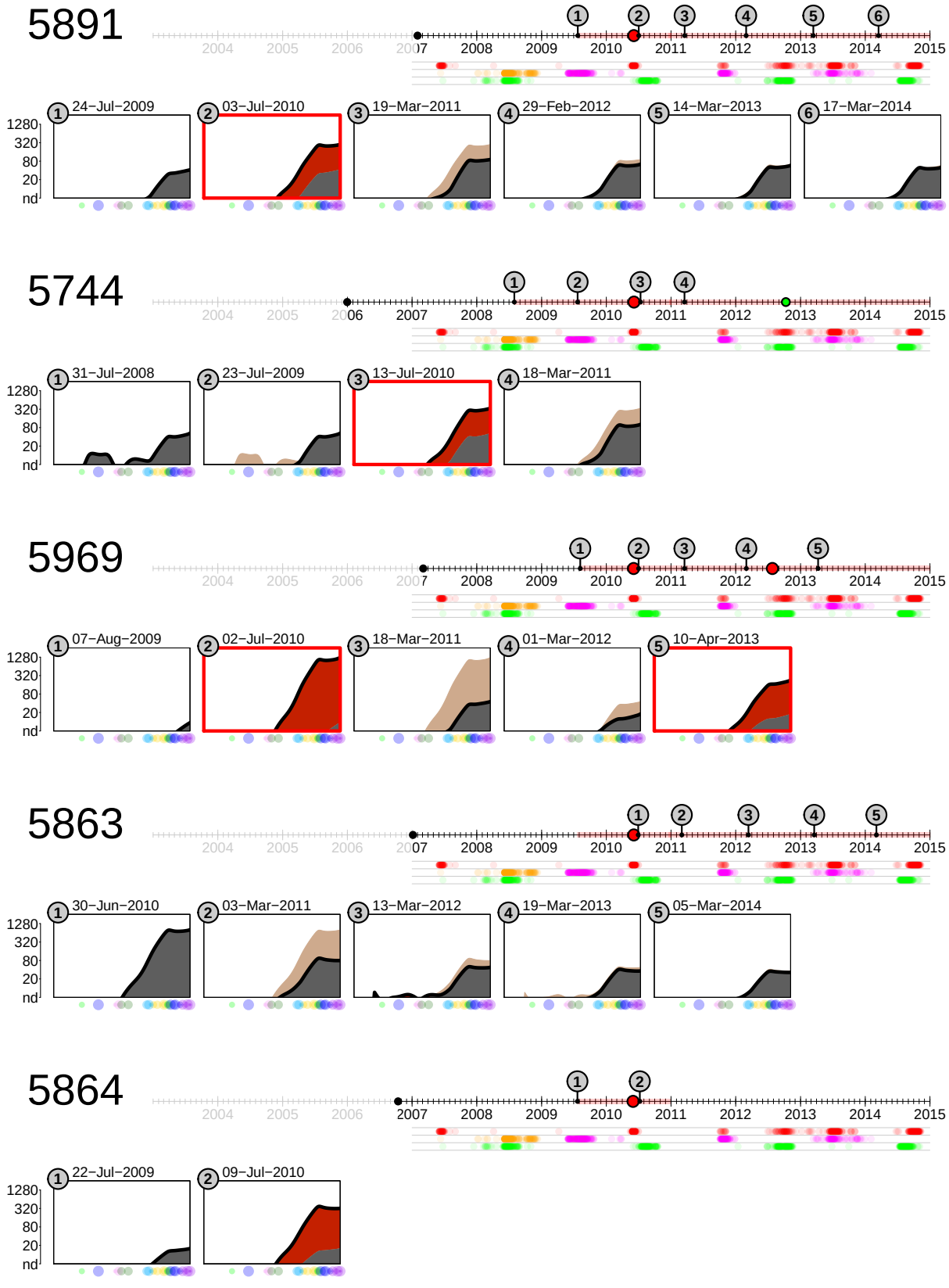

6002
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
5
6
1280
320
80
20
nd
1 04−Aug−2009
2 17−Jul−2010
3 19−Mar−2011
4 09−Mar−2012
5 21−Mar−2013
6 17−Mar−2014


6006
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 04–Aug–2009
2 15–Jul–2010
3 23–Mar–2011
4 29–Feb–2012
5 19–Mar–2013
6 21–Mar–2014
1280 320 80 20 nd
5872
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 23–Jul–2009
2 01–Jul–2010
3 02–Mar–2011
4 21–Mar–2012
5 14–Mar–2013
6 25–Mar–2014
1280 320 80 20 nd
5806
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 14–Jul–2009
2 01–Jul–2010
3 21–Mar–2011
4 29–Feb–2012
5 03–Apr–2013
6 13–Mar–2014
1280 320 80 20 nd
5871
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 23–Jul–2009
2 03–Jul–2010
3 08–Mar–2011
4 02–Mar–2012
5 14–Mar–2013
6 05–Mar–2014
1280 320 80 20 nd
5870
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 23–Jul–2009
2 29–Jun–2010
3 04–Mar–2011
4 03–Mar–2012
5 18–Mar–2013
6 19–Mar–2014
1280 320 80 20 nd



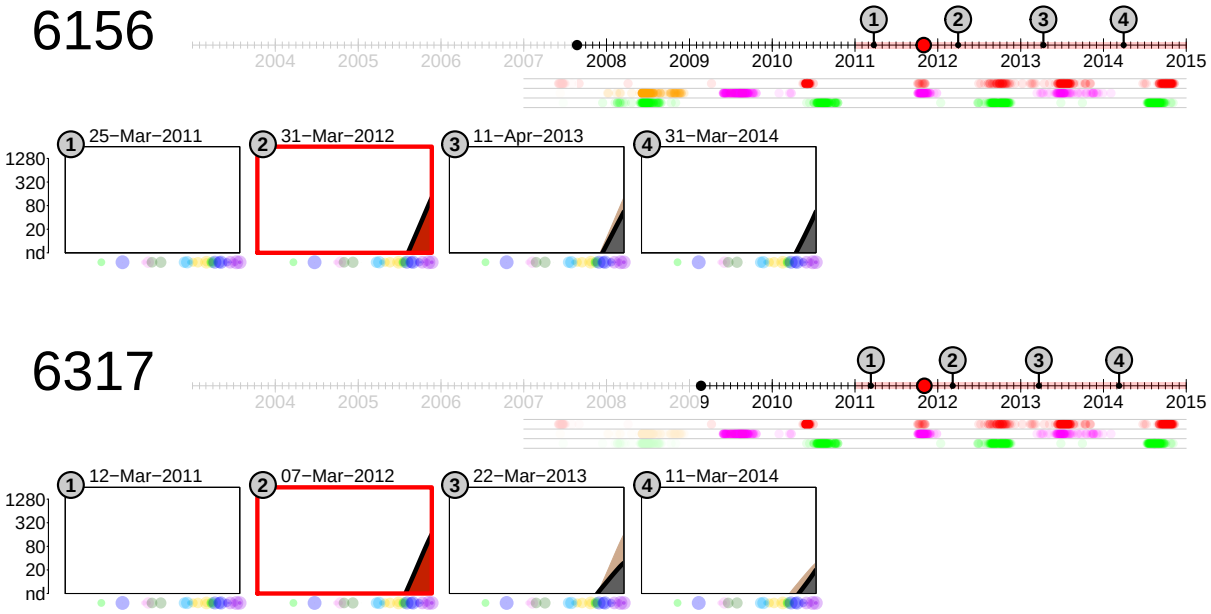
Potential second infections



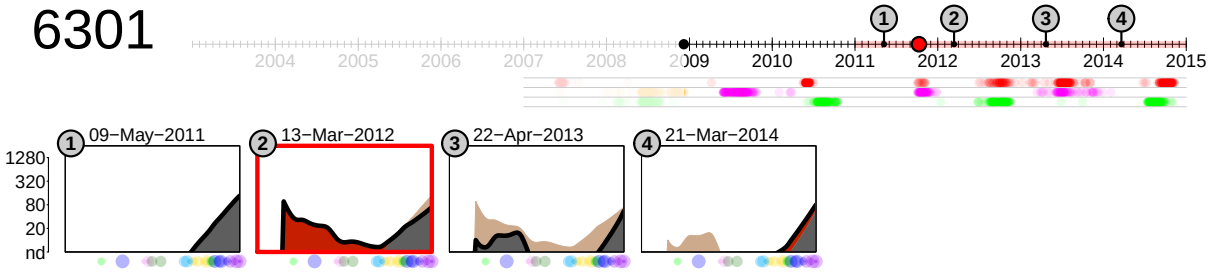

5891
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 24−Jul−2009
2 03−Jul−2010
3 19−Mar−2011
4 29−Feb−2012
5 14−Mar−2013
6 17−Mar−2014
1280
320
80
20
nd
5744
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 31−Jul−2008
2 23−Jul−2009
3 13−Jul−2010
4 18−Mar−2011
1280
320
80
20
nd
5969
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 07−Aug−2009
2 02−Jul−2010
3 18−Mar−2011
4 01−Mar−2012
5 10−Apr−2013
1280
320
80
20
nd
5863
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 30−Jun−2010
2 03−Mar−2011
3 13−Mar−2012
4 19−Mar−2013
5 05−Mar−2014
1280
320
80
20
nd
5864
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 22−Jul−2009
2 09−Jul−2010
1280
320
80
20
nd

C.3.3 *Season 3*

Potential first infections

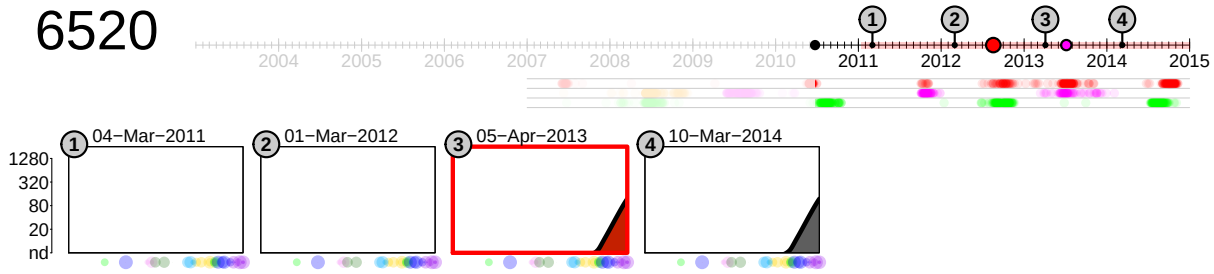


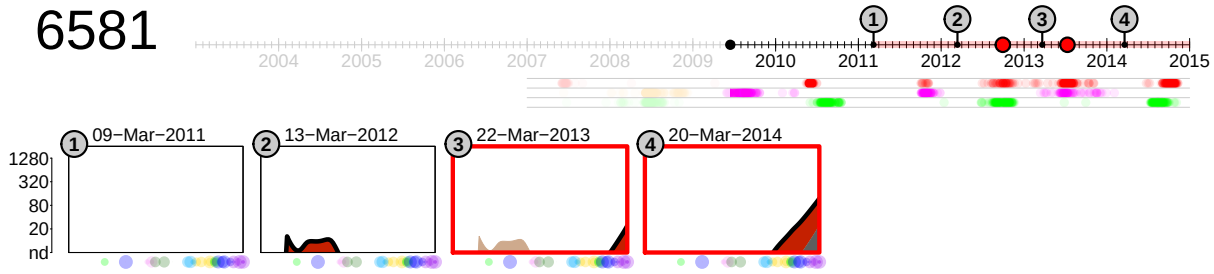
Potential second infections

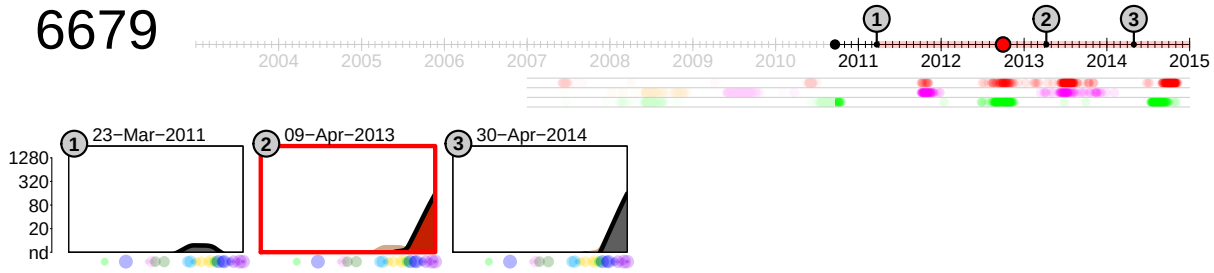


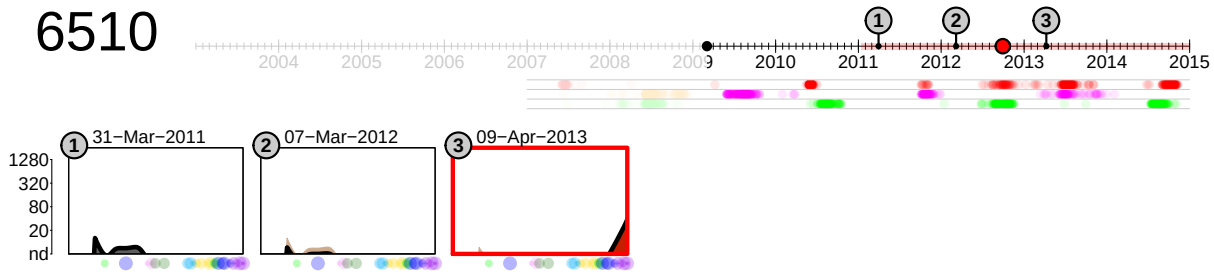
C.3.4 *Season 4*

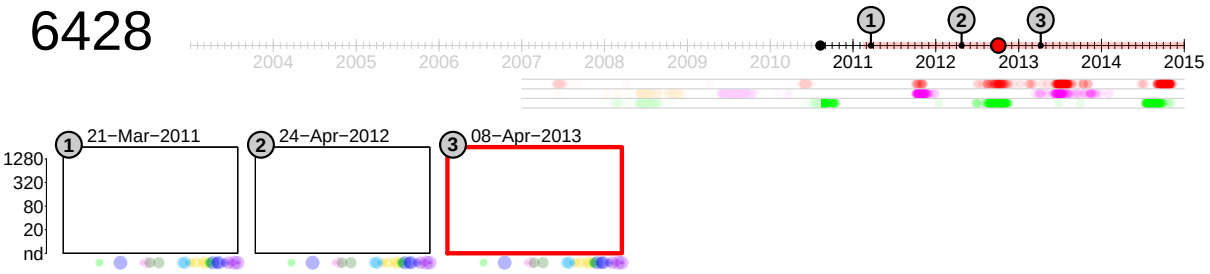
Potential first infections

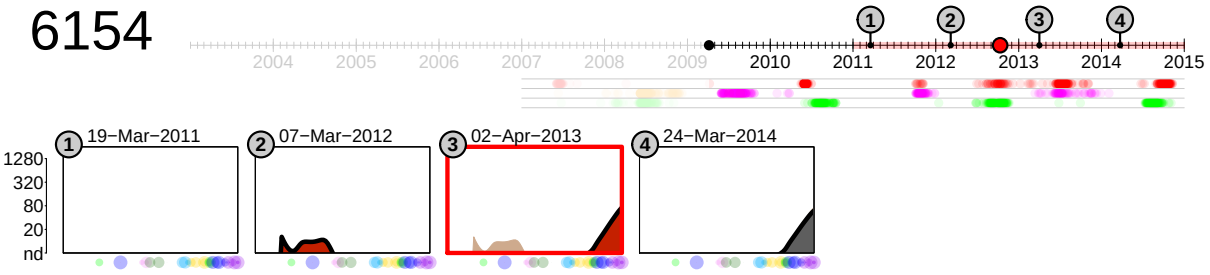


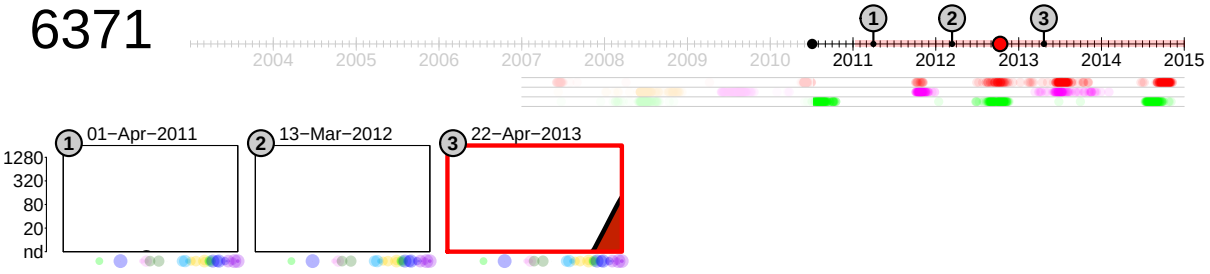


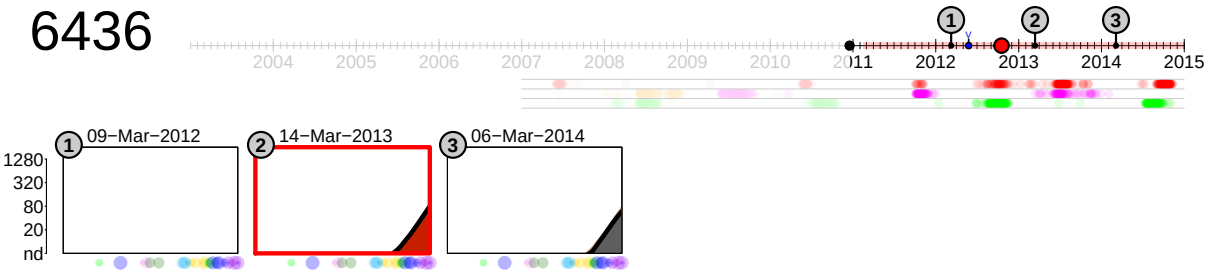


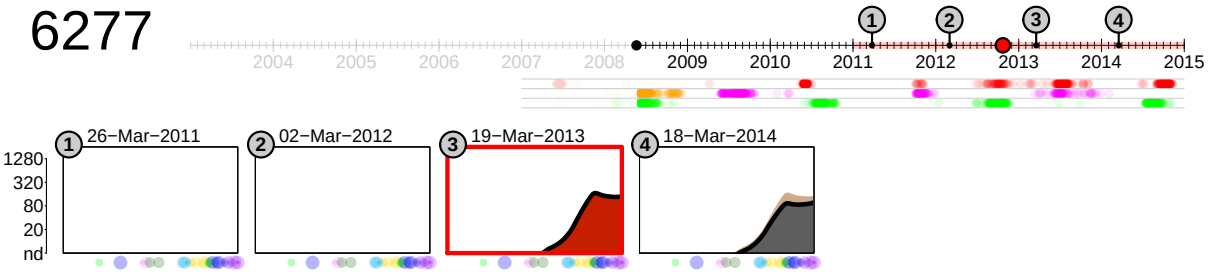


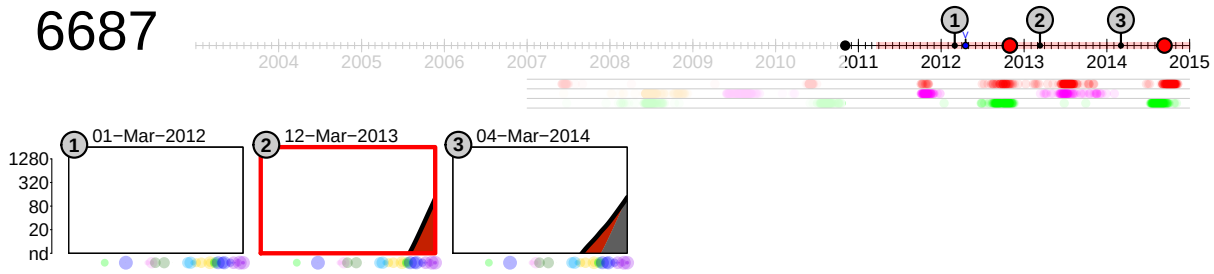

6428
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
1 21−Mar−2011
2 24−Apr−2012
3 08−Apr−2013
1280
320
80
20
nd

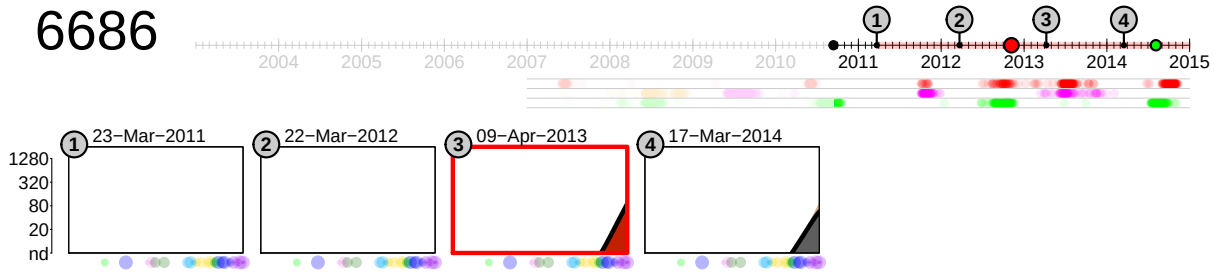

6154
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
1 19−Mar−2011
2 07−Mar−2012
3 02−Apr−2013
4 24−Mar−2014
1280
320
80
20
nd

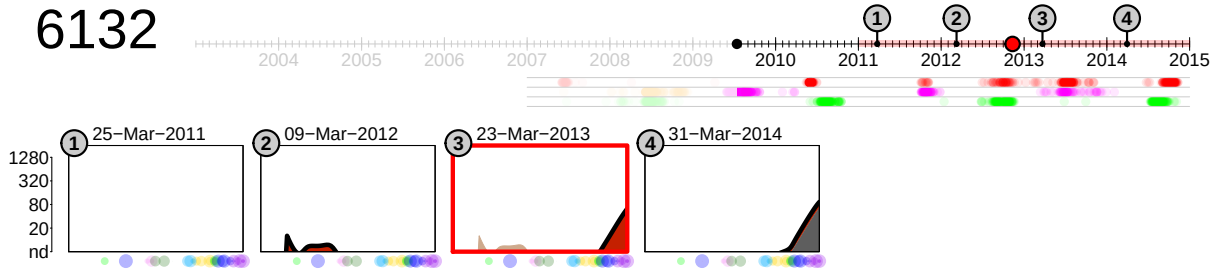

6371
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
1 01−Apr−2011
2 13−Mar−2012
3 22−Apr−2013
1280
320
80
20
nd

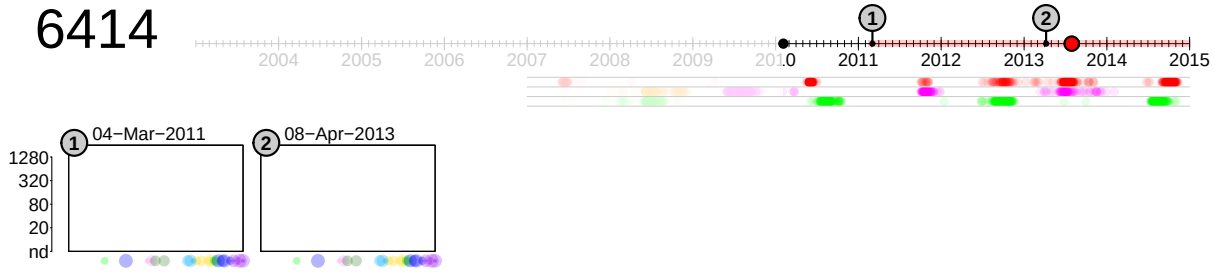

6436
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
1 09−Mar−2012
2 14−Mar−2013
3 06−Mar−2014
1280
320
80
20
nd

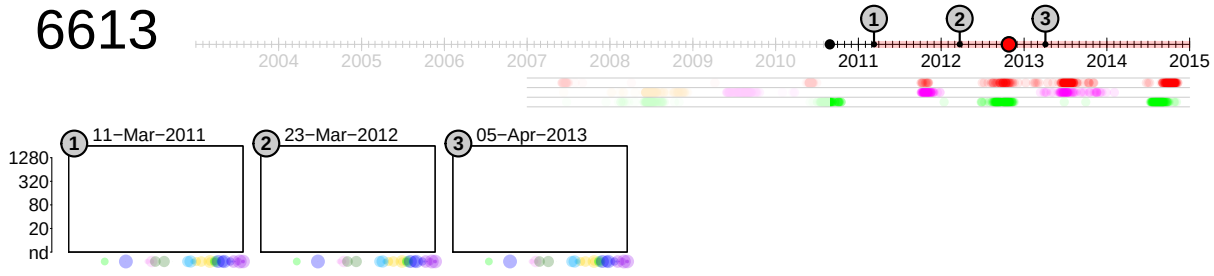

6277
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
1 26−Mar−2011
2 02−Mar−2012
3 19−Mar−2013
4 18−Mar−2014
1280
320
80
20
nd

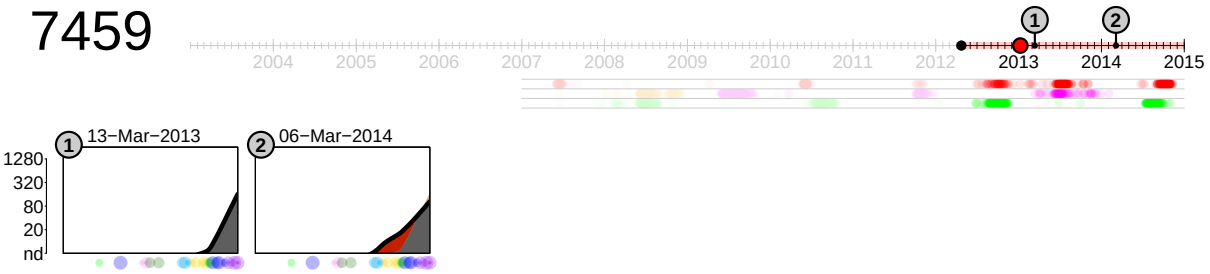

6687
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
01−Mar−2012
12−Mar−2013
04−Mar−2014
1280
320
80
20
nd


6686
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
23−Mar−2011
22−Mar−2012
09−Apr−2013
17−Mar−2014
1280
320
80
20
nd

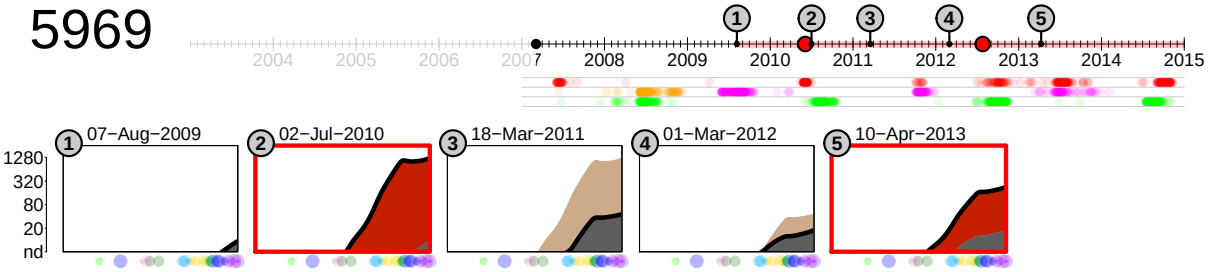

6132
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
25−Mar−2011
09−Mar−2012
23−Mar−2013
31−Mar−2014
1280
320
80
20
nd

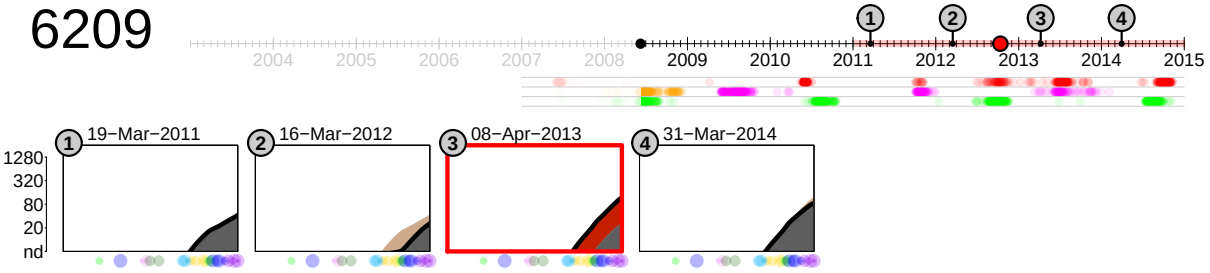

6414
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
04−Mar−2011
08−Apr−2013
1280
320
80
20
nd

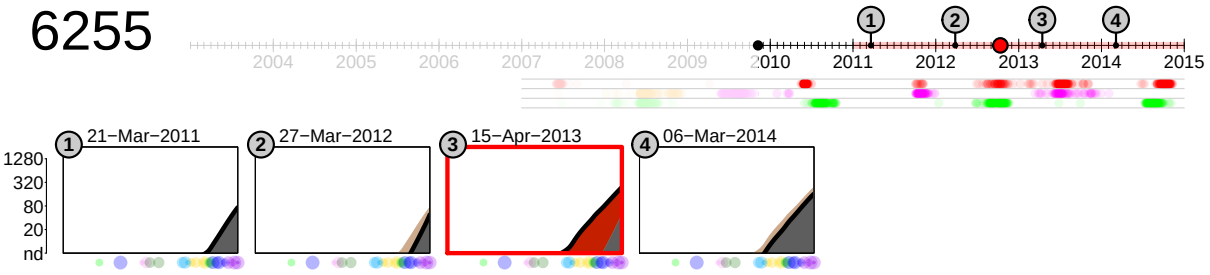

6613
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
11−Mar−2011
23−Mar−2012
05−Apr−2013
1280
320
80
20
nd



Potential second infections

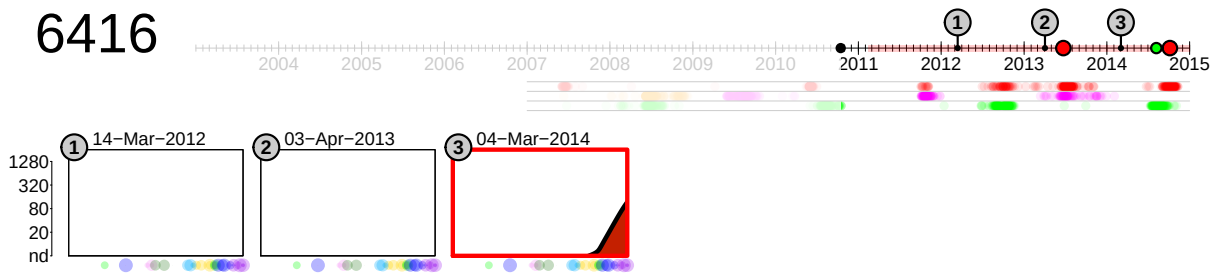


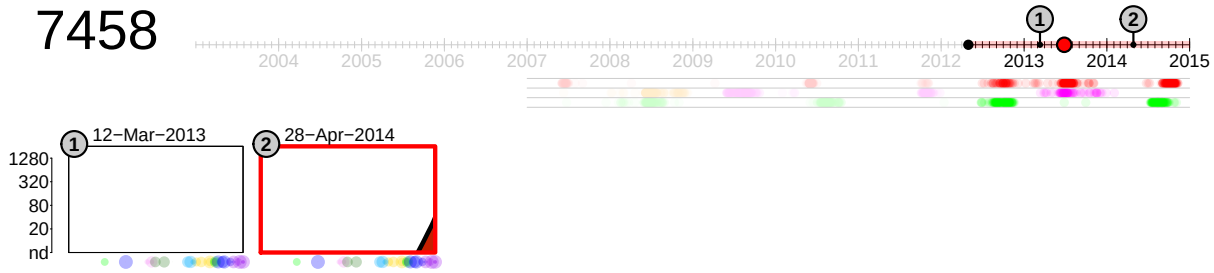


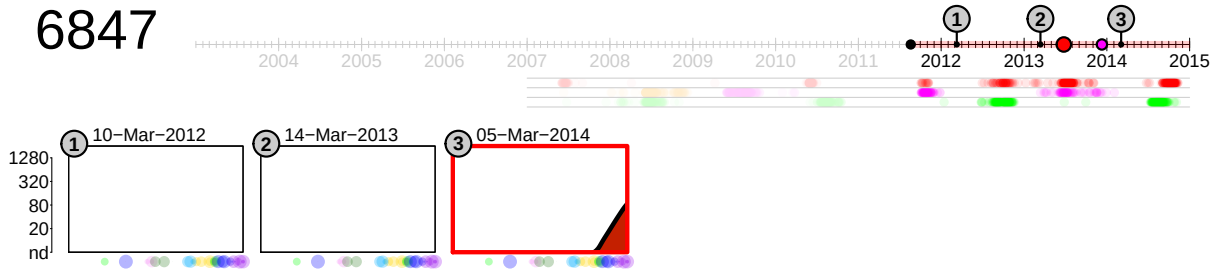


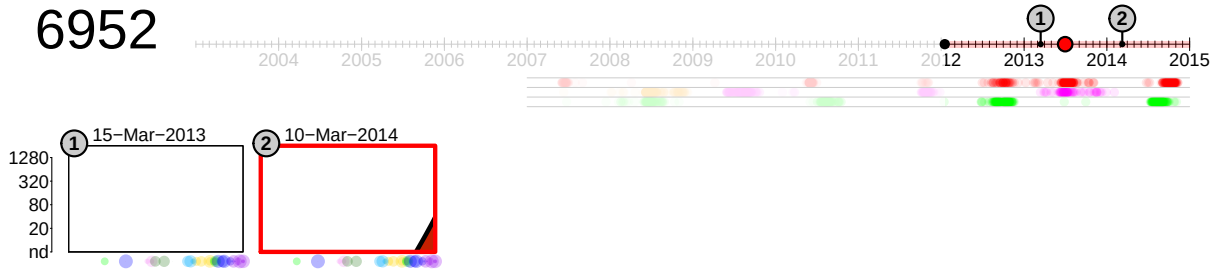
C.3.5 *Season 5*

Potential first infections

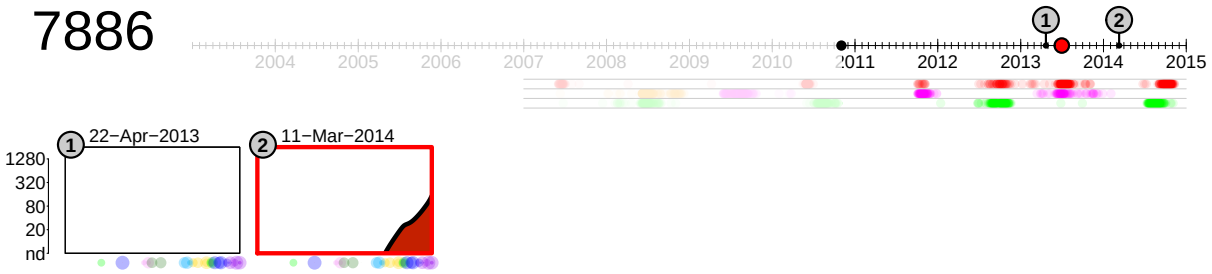




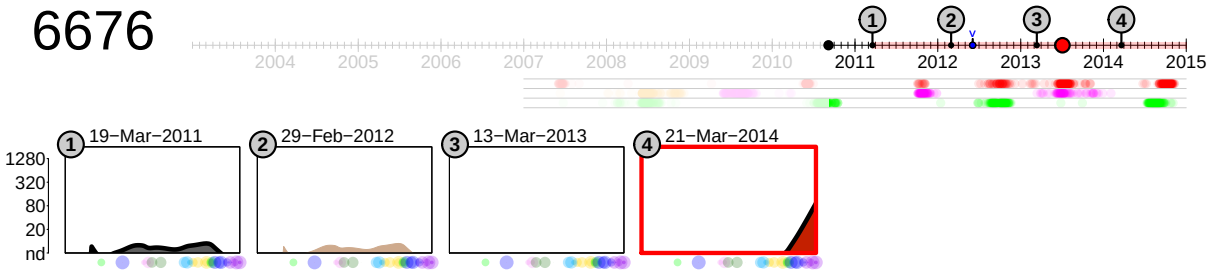




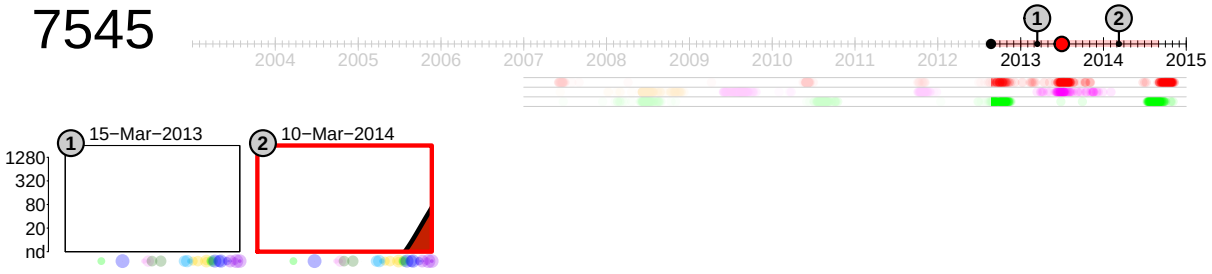
7886



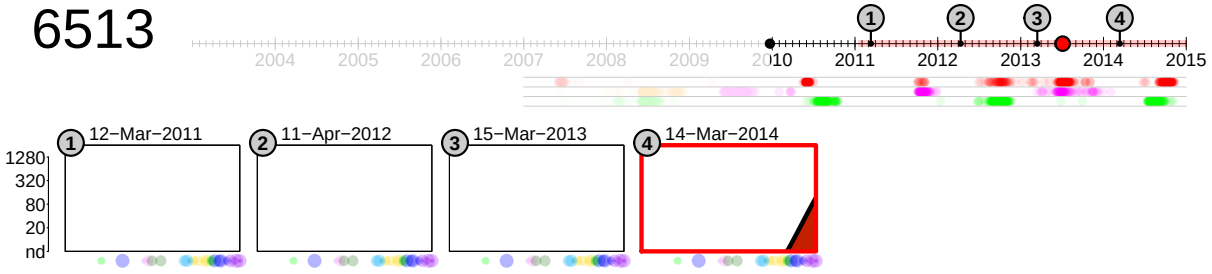
6676



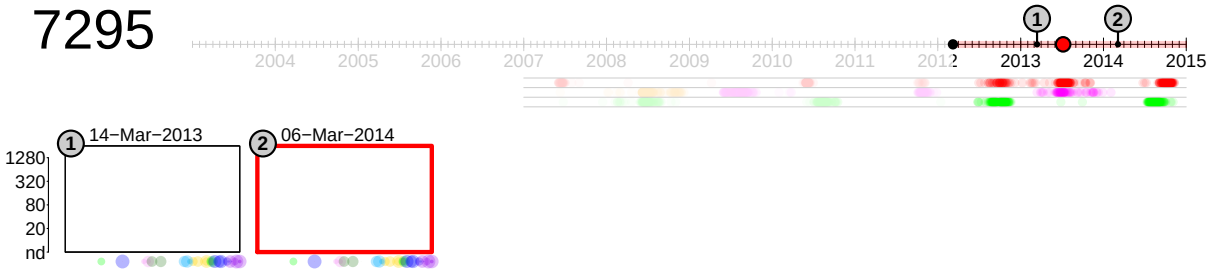
7545

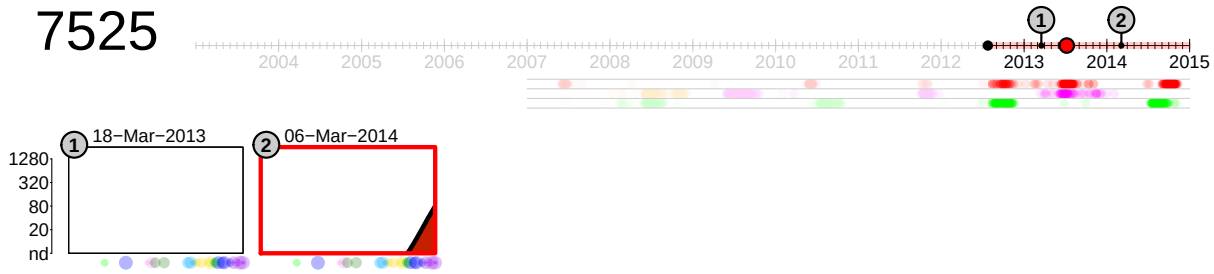


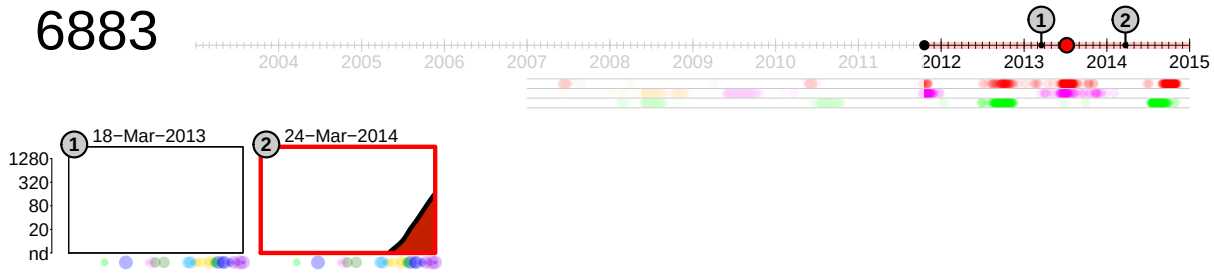
6513

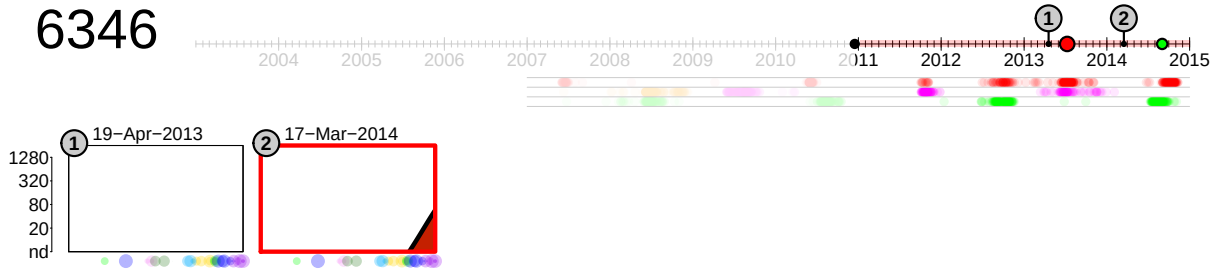


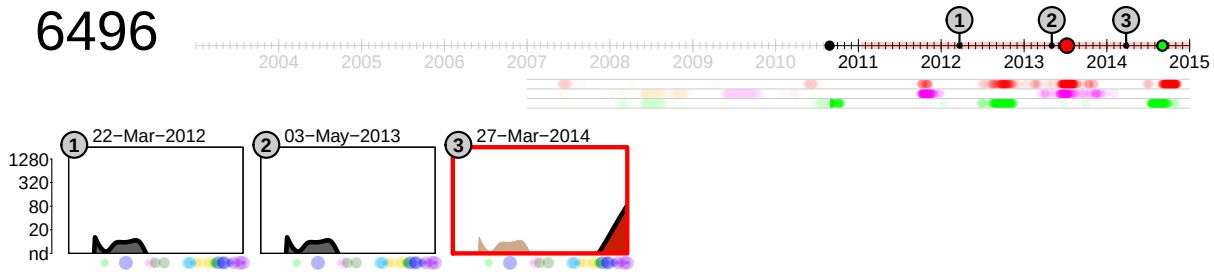
7295

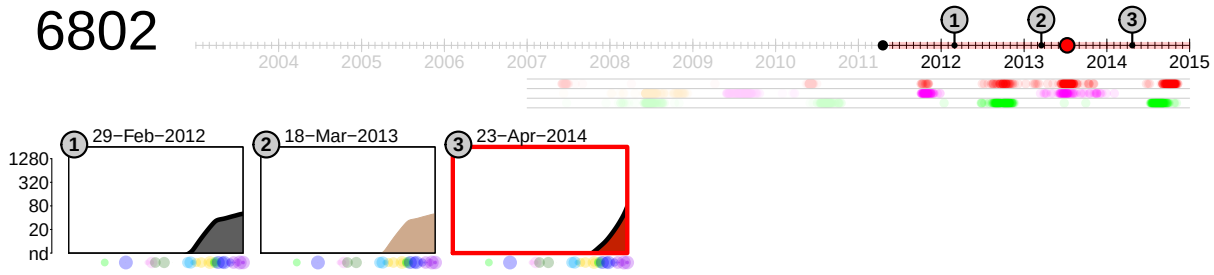



7525
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
18−Mar−2013
06−Mar−2014
1280
320
80
20
nd

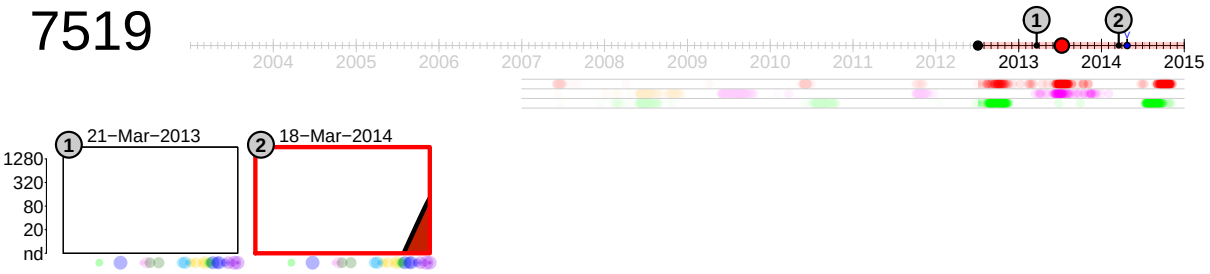

6883
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
18−Mar−2013
24−Mar−2014
1280
320
80
20
nd


6346
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
19−Apr−2013
17−Mar−2014
1280
320
80
20
nd

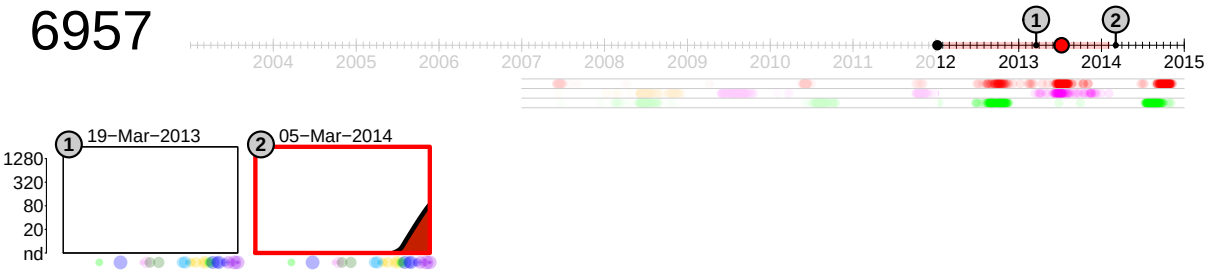

6496
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
22−Mar−2012
03−May−2013
27−Mar−2014
1280
320
80
20
nd


6802
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
29−Feb−2012
18−Mar−2013
23−Apr−2014
1280
320
80
20
nd

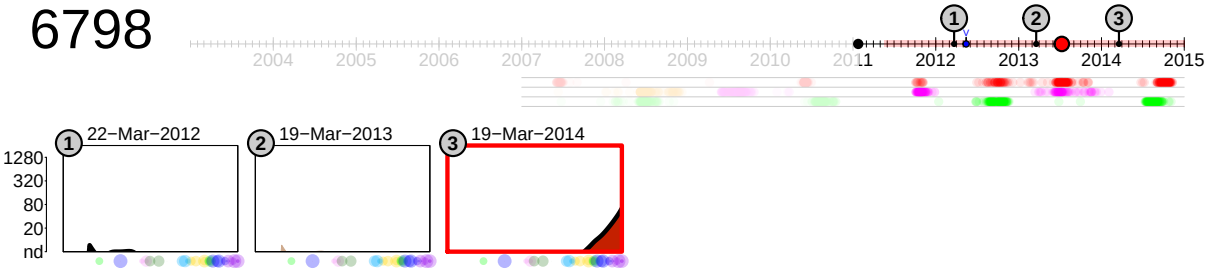
7519



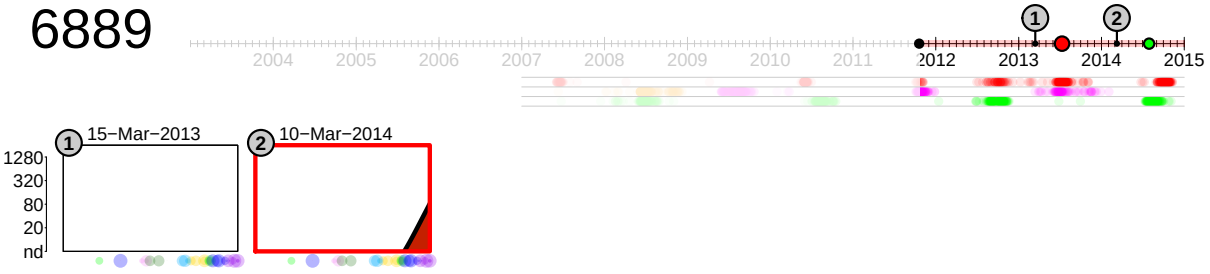
6957



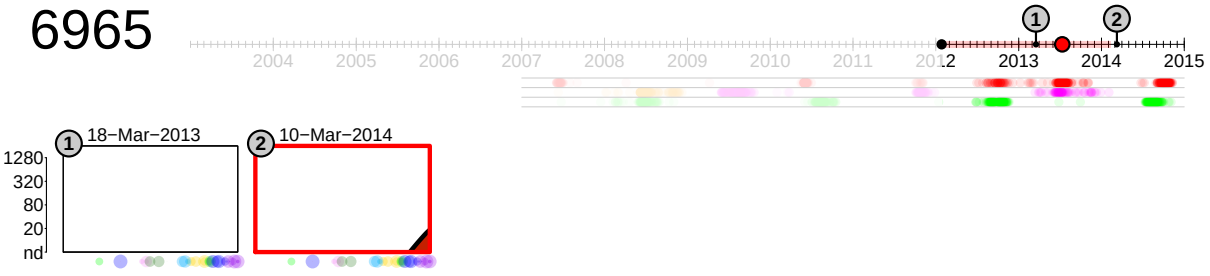
6798

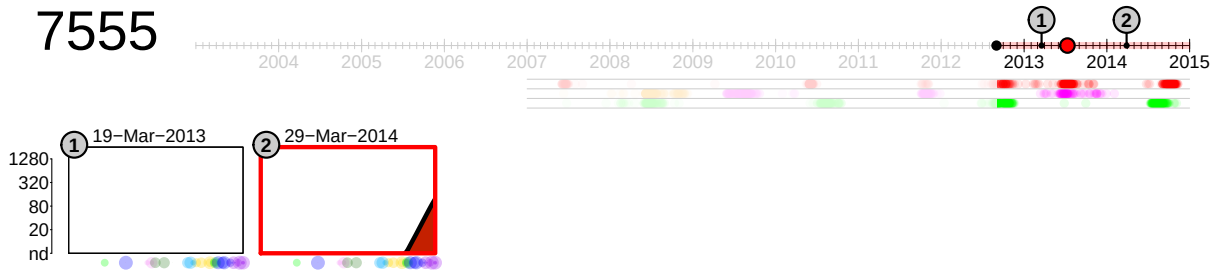


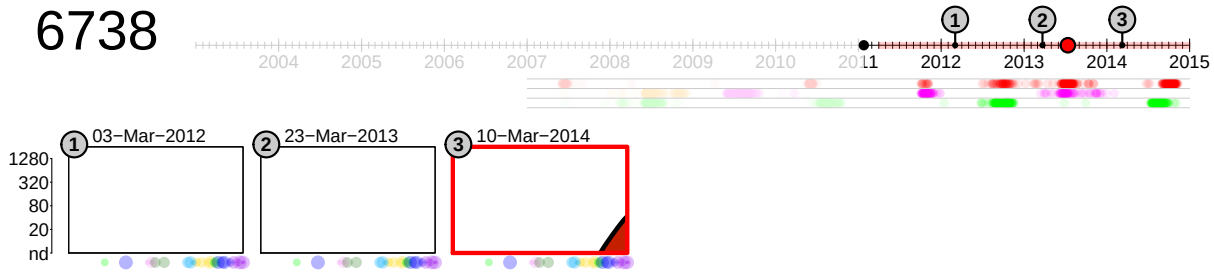
6889

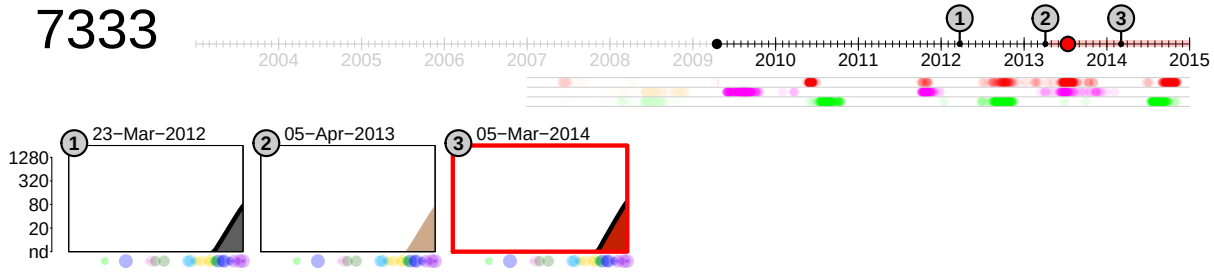


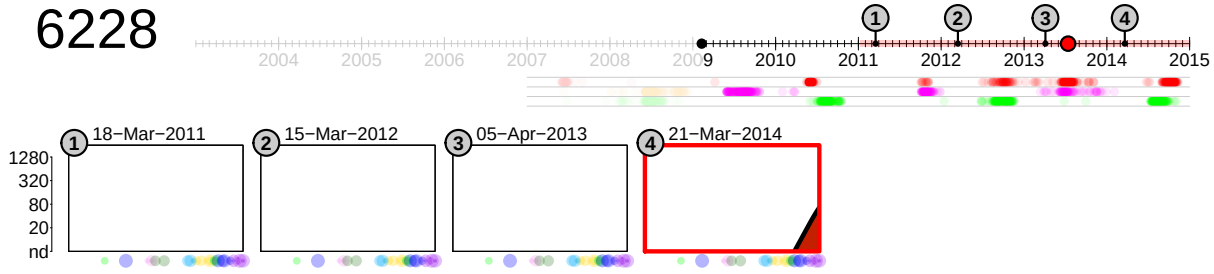
6965

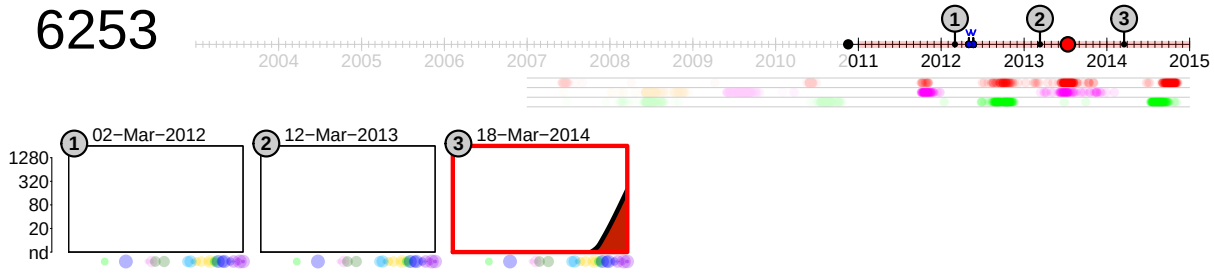


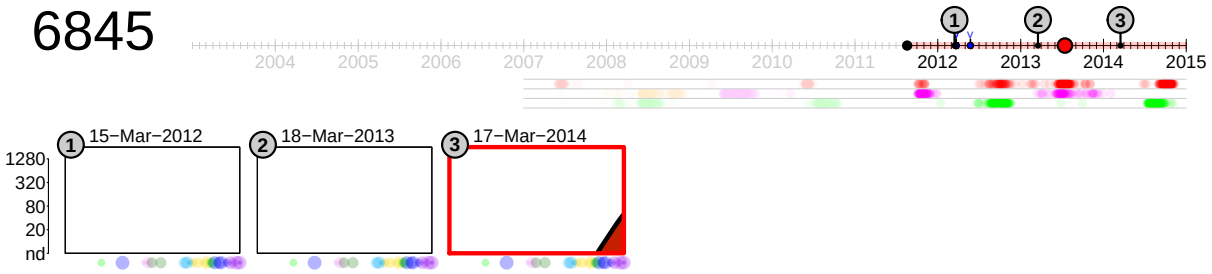

7555
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
1
19−Mar−2013
2
29−Mar−2014
1280
320
80
20
nd

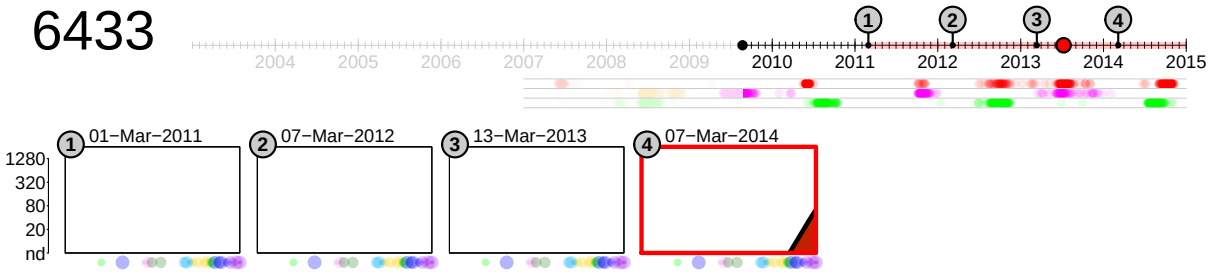

6738
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
1
03−Mar−2012
2
23−Mar−2013
3
10−Mar−2014
1280
320
80
20
nd

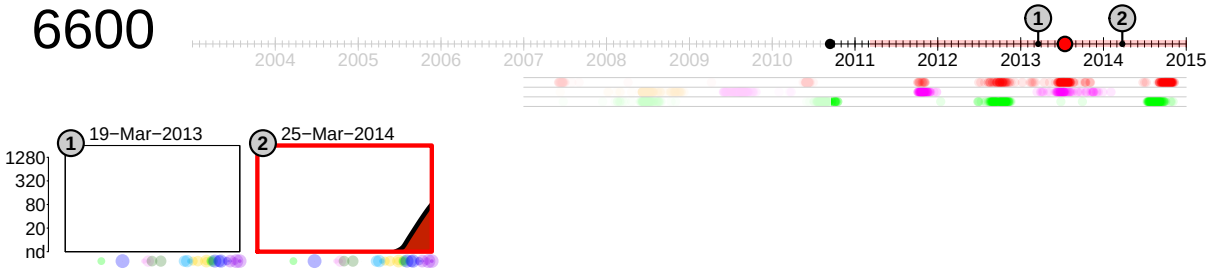

7333
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
1
23−Mar−2012
2
05−Apr−2013
3
05−Mar−2014
1280
320
80
20
nd

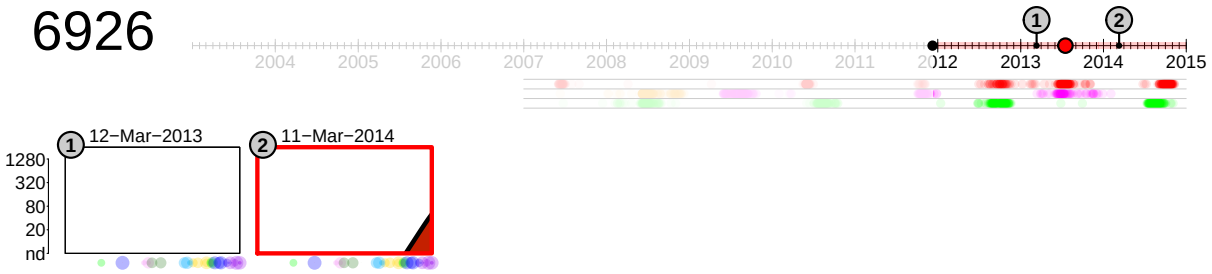

6228
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
1
18−Mar−2011
2
15−Mar−2012
3
05−Apr−2013
4
21−Mar−2014
1280
320
80
20
nd

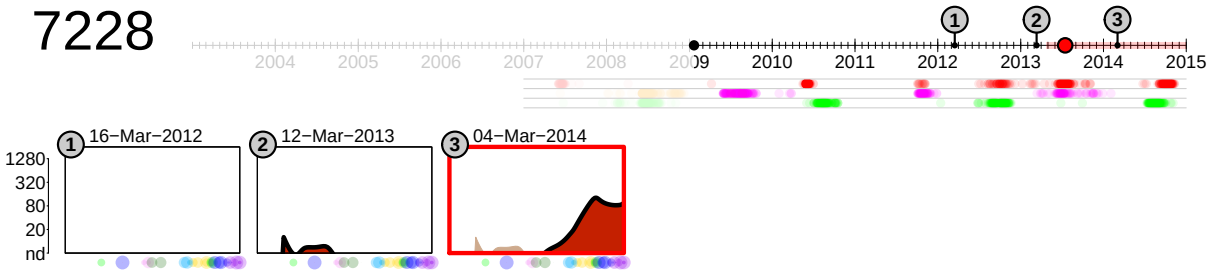

6253
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
1
02−Mar−2012
2
12−Mar−2013
3
18−Mar−2014
1280
320
80
20
nd

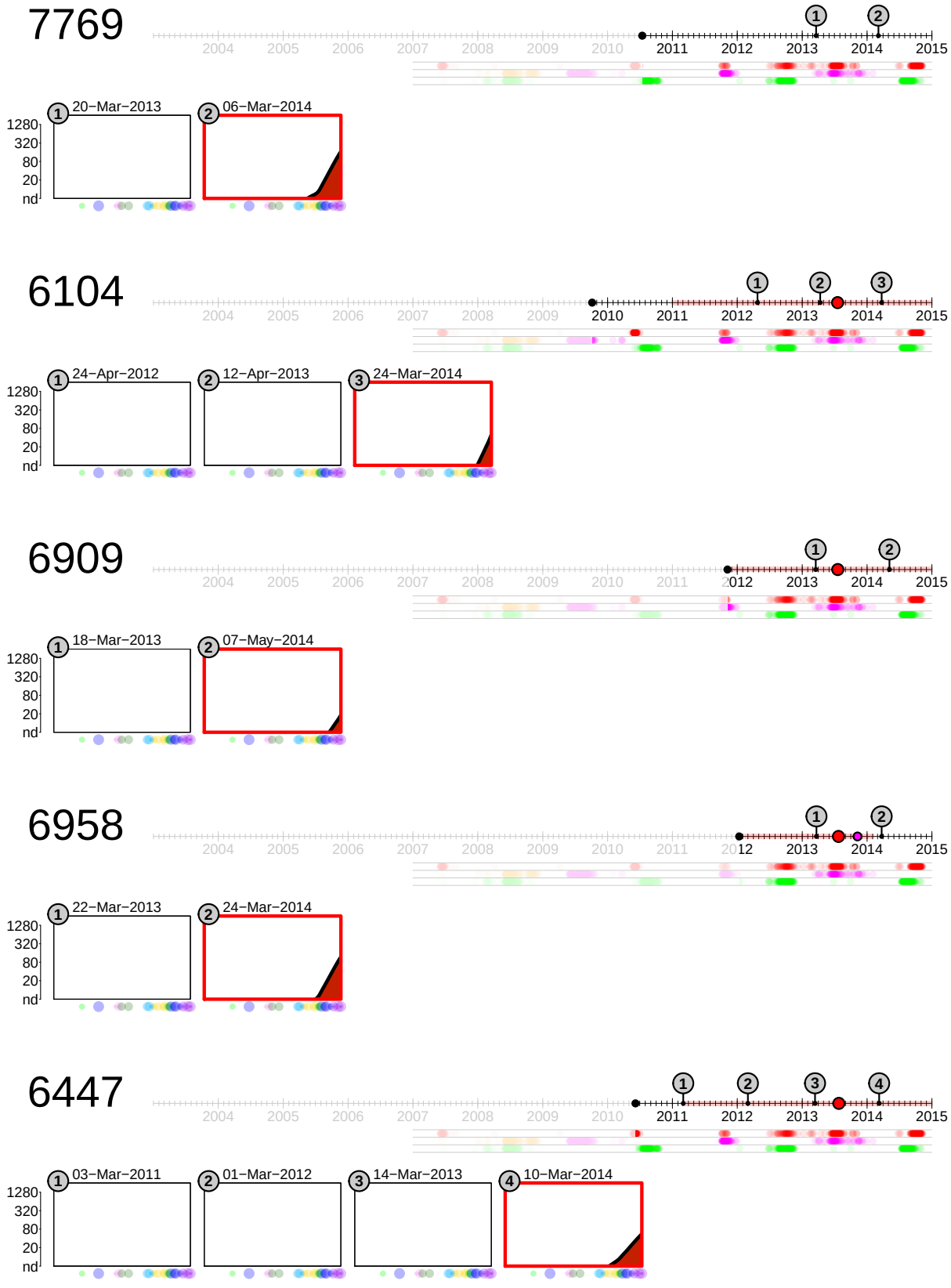

6845
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
15−Mar−2012
18−Mar−2013
17−Mar−2014
1280
320
80
20
nd

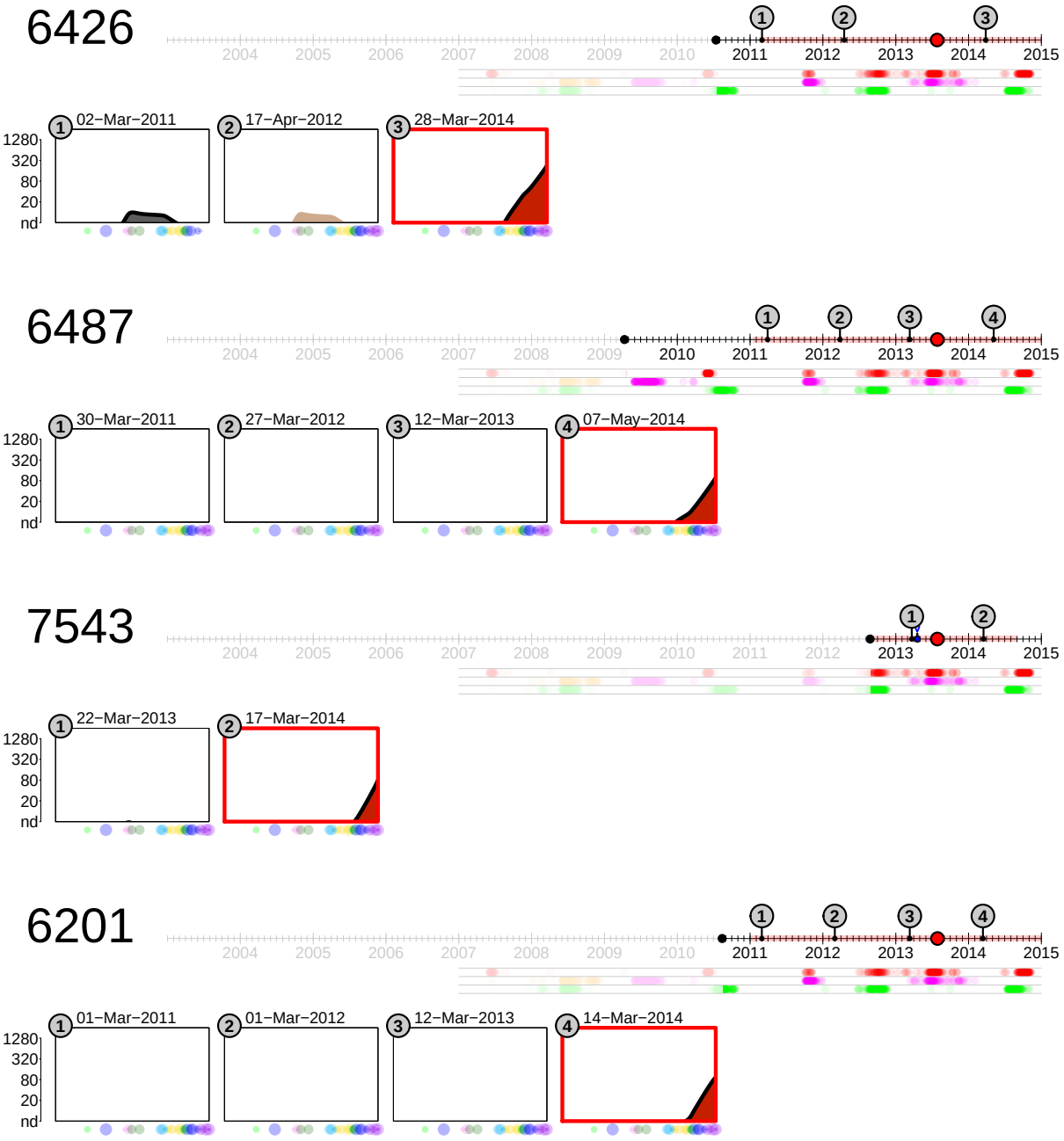

6433
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
4
01−Mar−2011
07−Mar−2012
13−Mar−2013
07−Mar−2014
1280
320
80
20
nd


6600
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
19−Mar−2013
25−Mar−2014
1280
320
80
20
nd

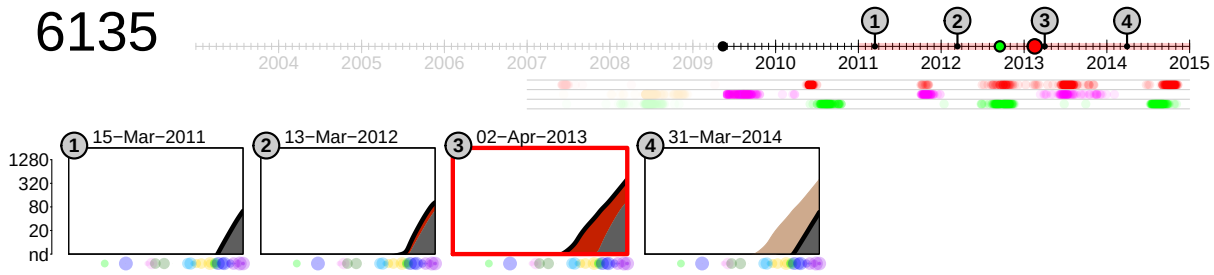

6926
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
12−Mar−2013
11−Mar−2014
1280
320
80
20
nd

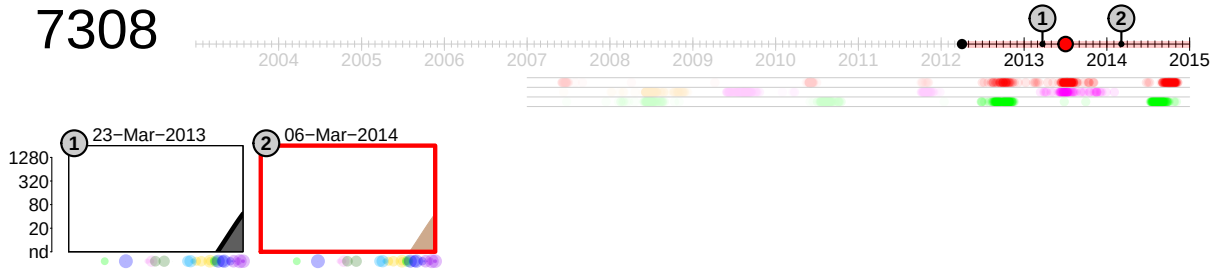

7228
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
1
2
3
16−Mar−2012
12−Mar−2013
04−Mar−2014
1280
320
80
20
nd

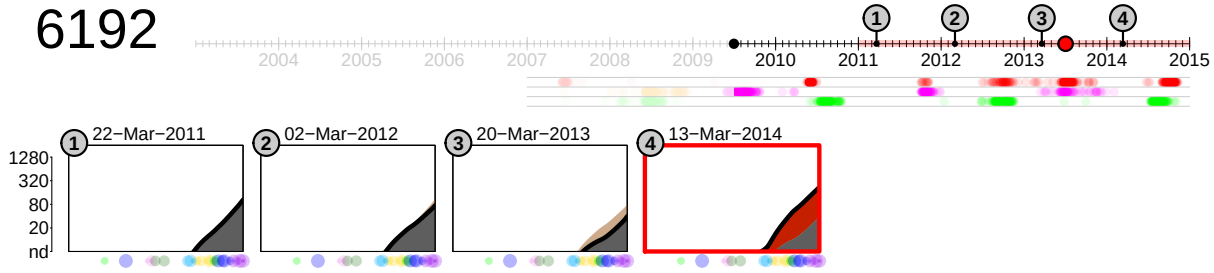

7769
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 20−Mar−2013
2 06−Mar−2014
1280 320 80 20 nd
6104
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 24−Apr−2012
2 12−Apr−2013
3 24−Mar−2014
1280 320 80 20 nd
6909
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 18−Mar−2013
2 07−May−2014
1280 320 80 20 nd
6958
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 22−Mar−2013
2 24−Mar−2014
1280 320 80 20 nd
6447
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 03−Mar−2011
2 01−Mar−2012
3 14−Mar−2013
4 10−Mar−2014
1280 320 80 20 nd

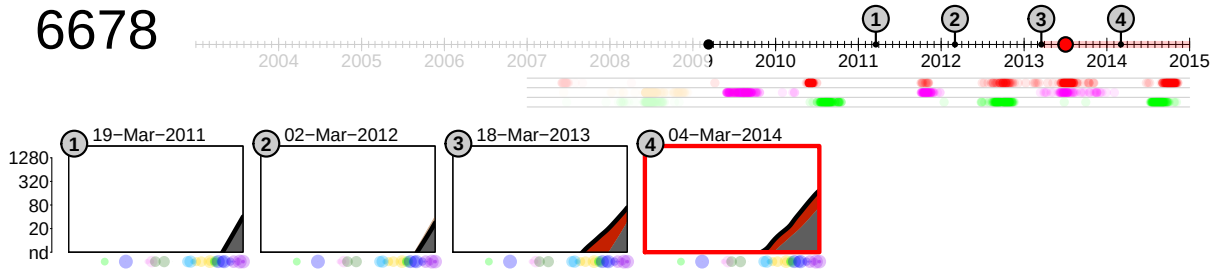

6426
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 2 3
1 02−Mar−2011
2 17−Apr−2012
3 28−Mar−2014
1280
320
80
20
nd
6487
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 2 3 4
1 30−Mar−2011
2 27−Mar−2012
3 12−Mar−2013
4 07−May−2014
1280
320
80
20
nd
7543
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 2
1 22−Mar−2013
2 17−Mar−2014
1280
320
80
20
nd
6201
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
1 2 3 4
1 01−Mar−2011
2 01−Mar−2012
3 12−Mar−2013
4 14−Mar−2014
1280
320
80
20
nd

Potential second infections

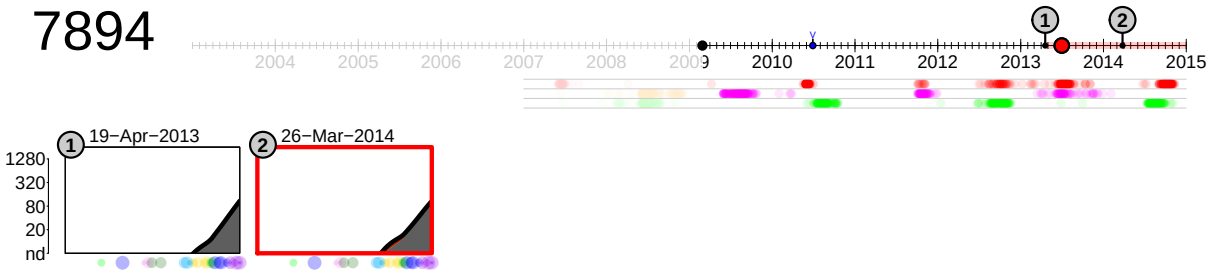




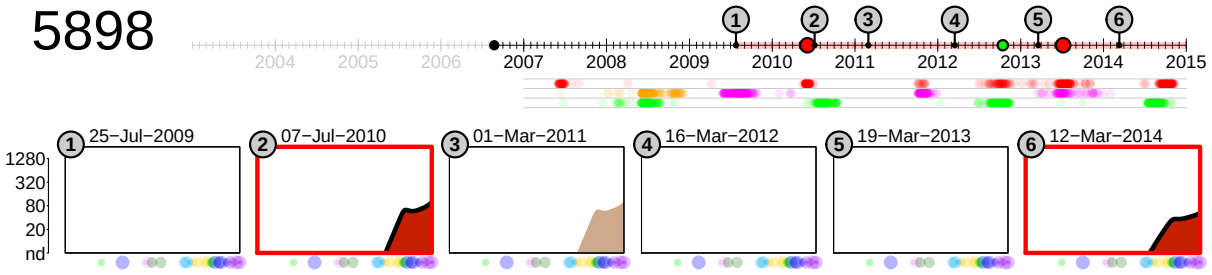




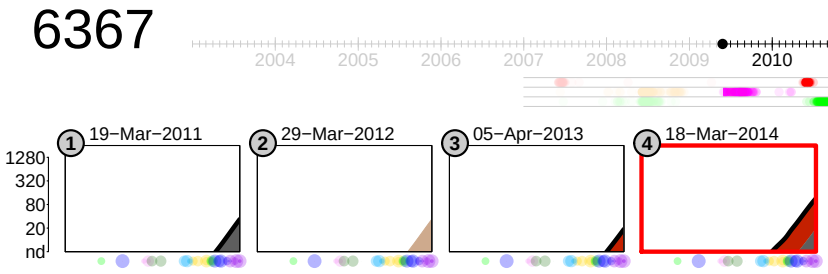
7894



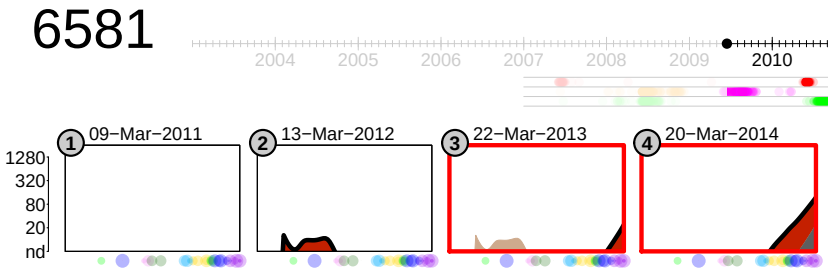
5898



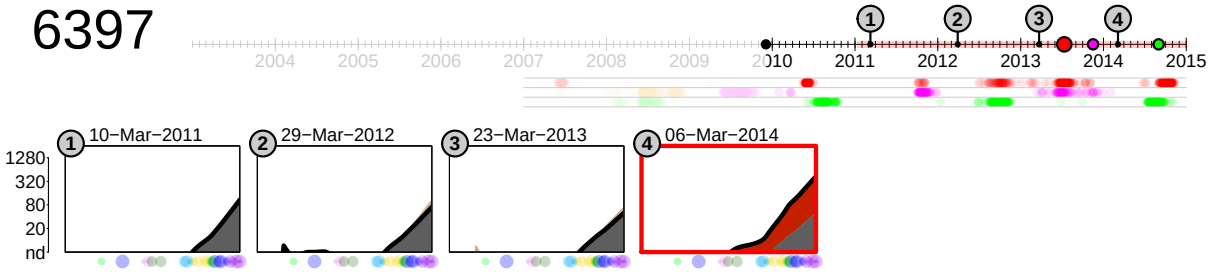
6367



6581



6397



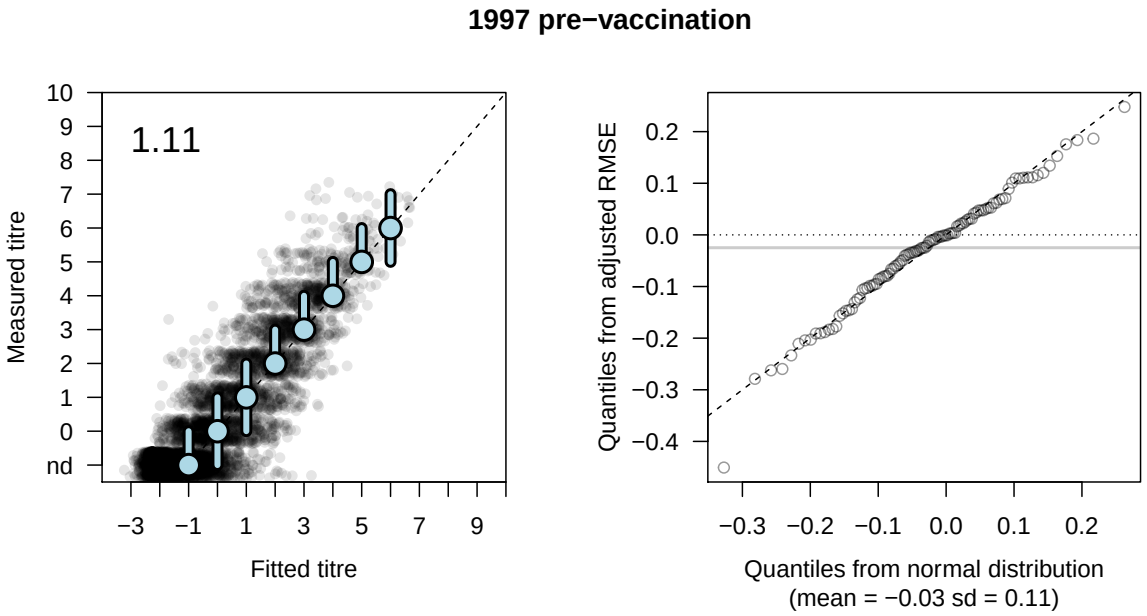
C.4 Diagnostic plots

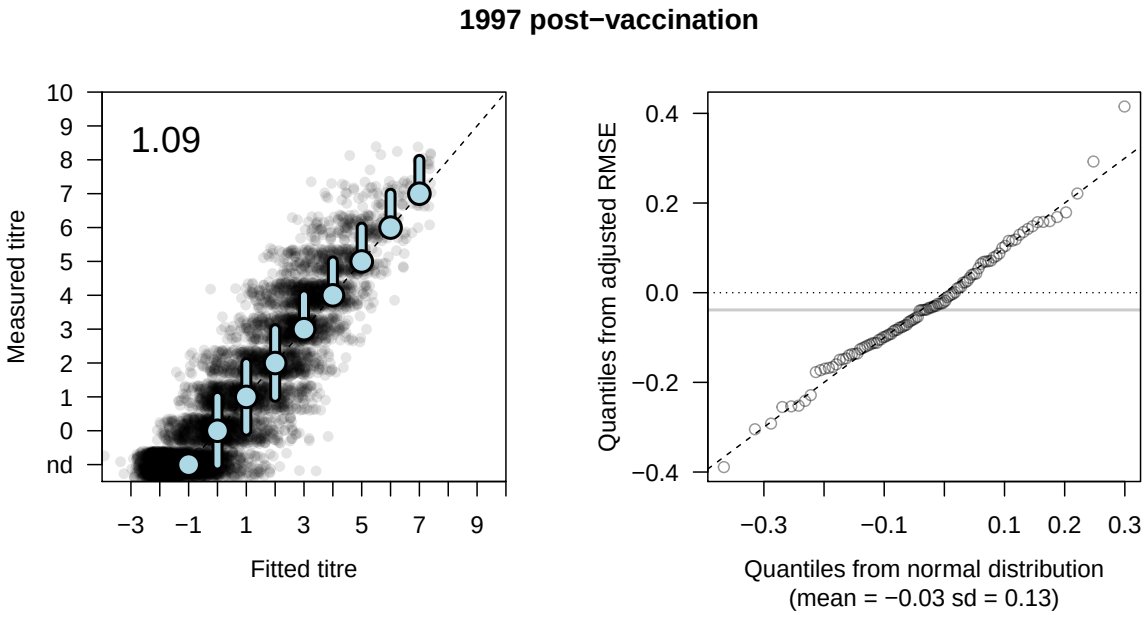
Shared caption for the figures in this section

Left panels: Diagnostic plots of predicted vs. measured titres. Predicted vs measured titres for antibody landscapes fit to serum samples. Larger light blue circles show the median measured titre for fitted titres rounded to the nearest whole titre and the light blue lines mark the interquartile ranges. Individual data points are shown in transparent grey and since the measured titres are all discrete values a small amount of jitter has been added in the y axis to better visualise the spread of data. The dashed black line shows the line of equality and the number in the top left of the plot gives the RMSE of fits to measurable titres.

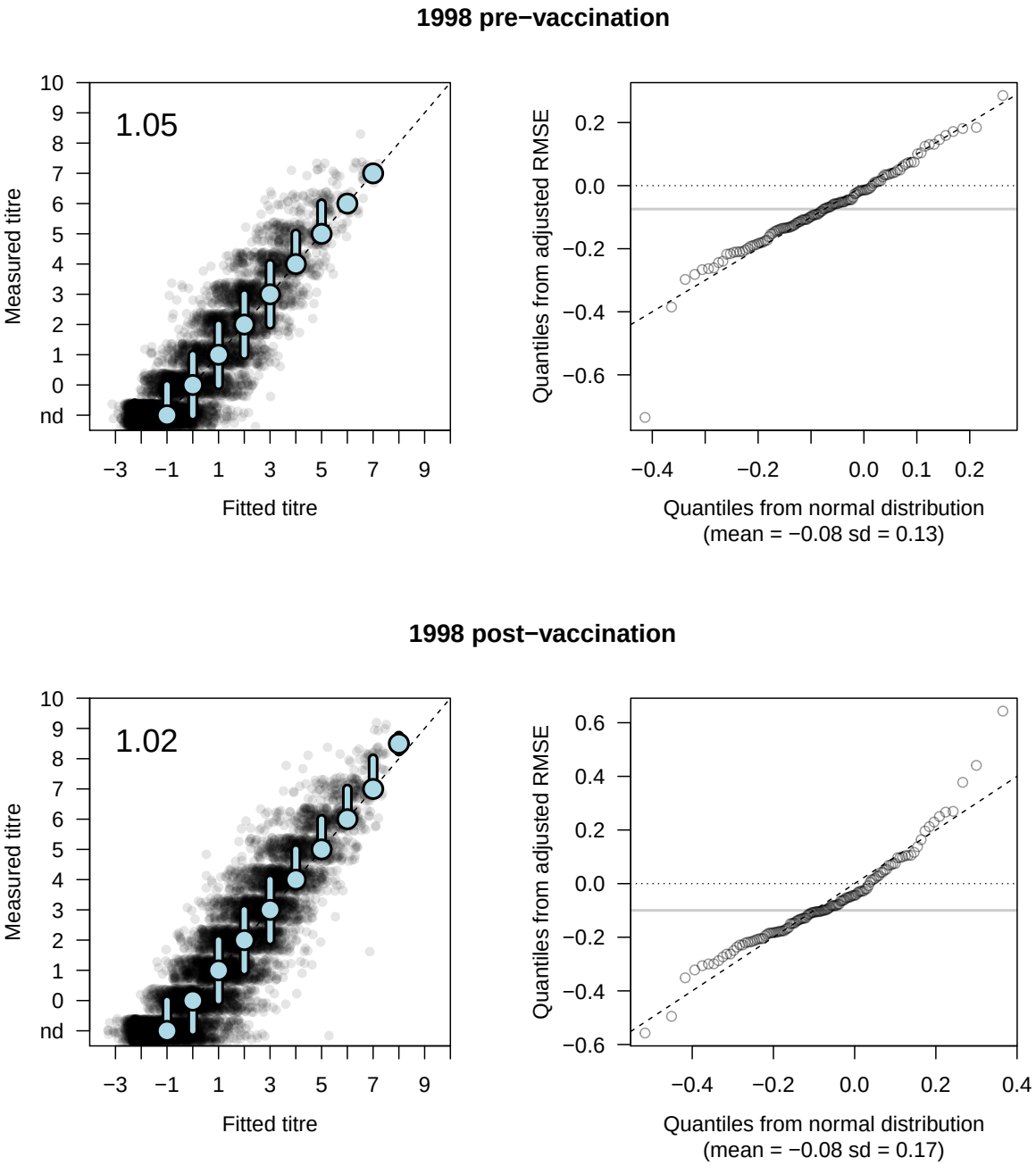
Right panels: Quantile-quantile plots to assess distribution of antibody landscape fit adjusted RMSEs. The distribution of the RMSEs compared to the expected average RMSE for each fit taking into account the number of non-detectable titres as described in chapter 2. In each case individual quantiles (each representing an individual antibody landscape fit) are shown as grey circles and the median adjusted RMSE is marked by the solid horizontal grey line. The dashed-black line shows the line of equality - if points are normally distributed you would expect them to lie on or close to this line.

1997 vaccination trial

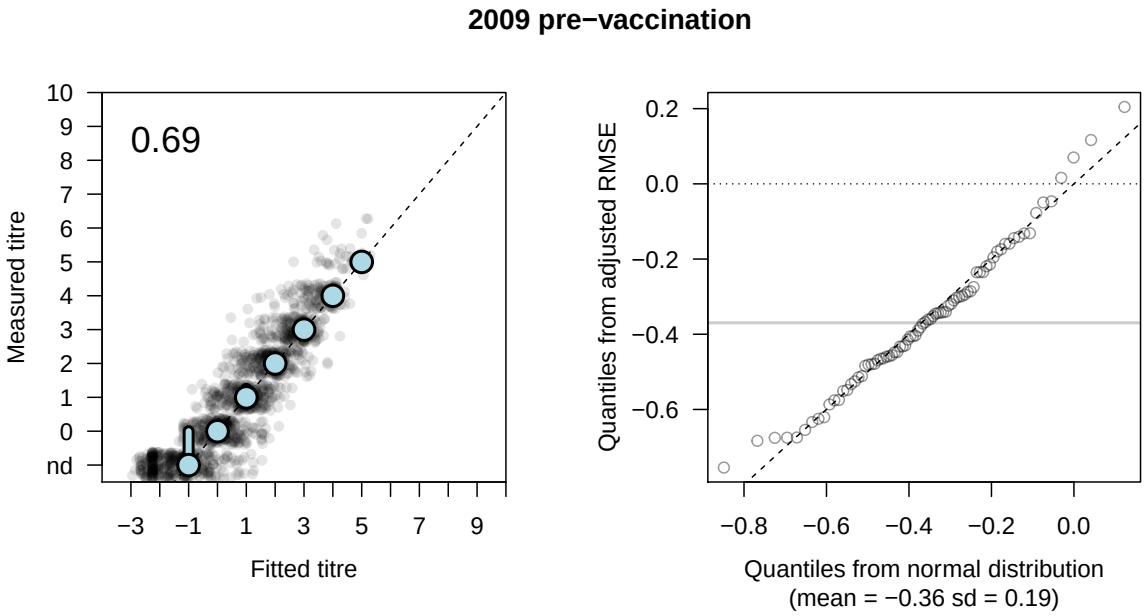


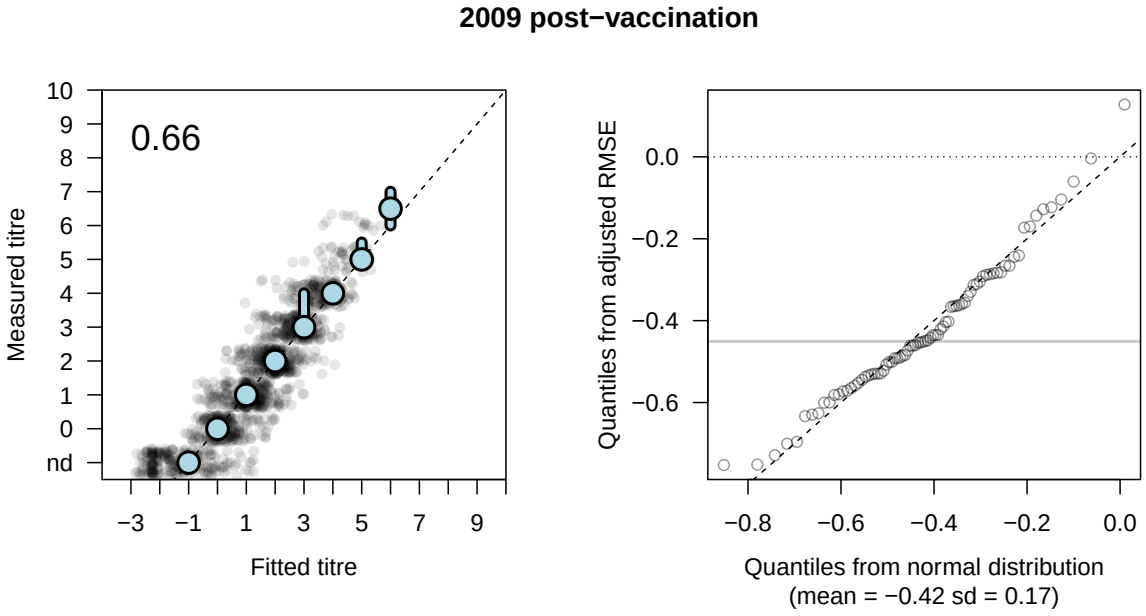


1998 vaccination trial

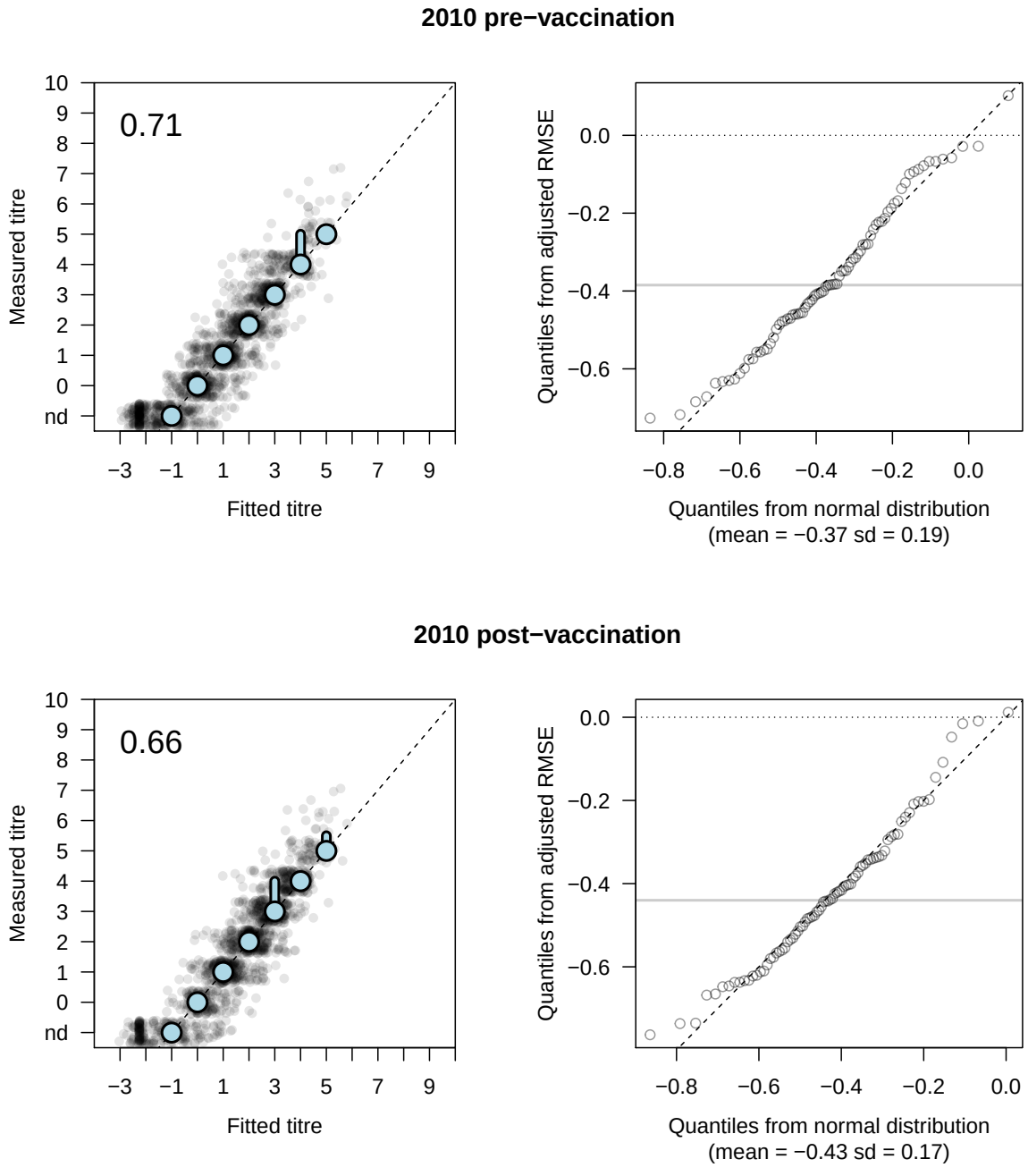


2009 vaccination trial

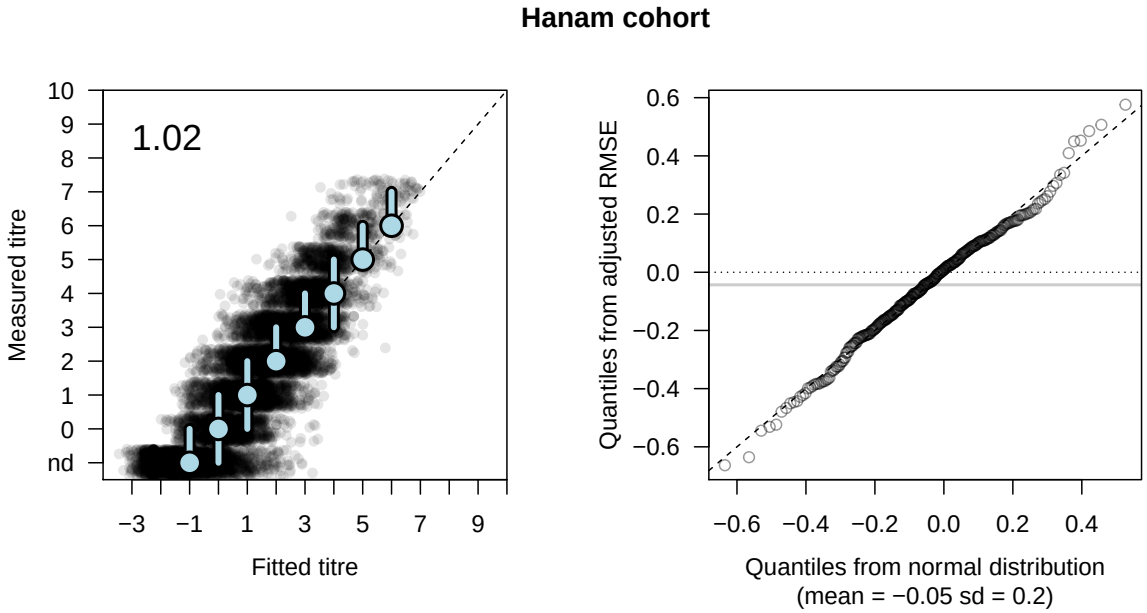




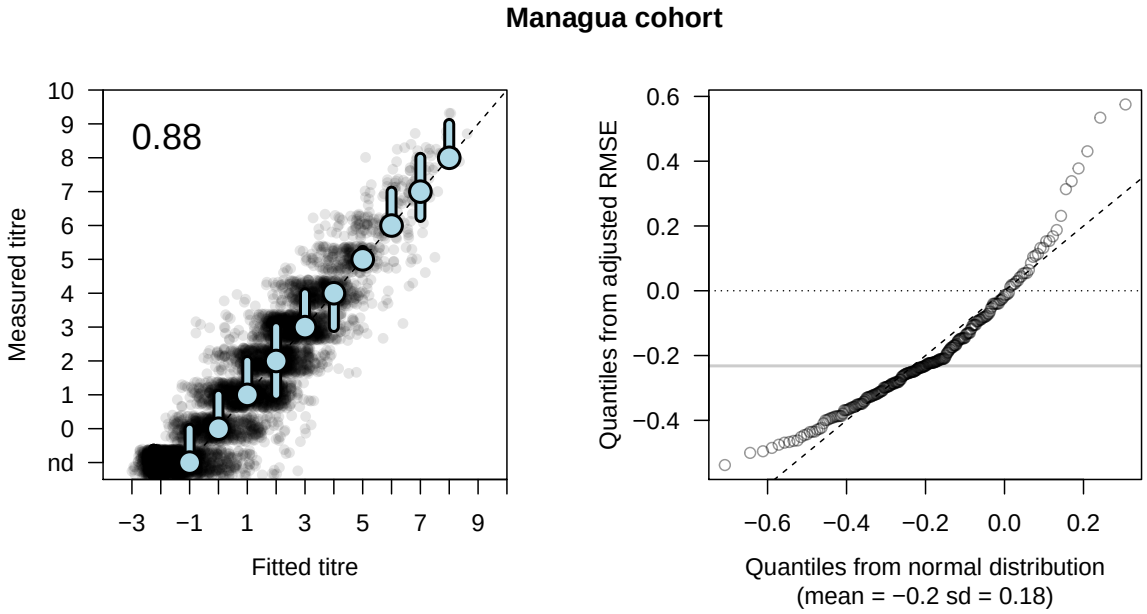
2010 vaccination trial



Hanam cohort



Managua cohort



C.5 Links to interactive plots

Interactive versions of many of the plots shown included in this thesis are available online at the website for the Center of Pathogen Evolution, Department of Zoology, University of Cambridge. Please visit:

http://antigenic-cartography.org/samwilks/phDthesis/

Bibliography

Adler, Daniel, Duncan Murdoch, and others (2017). *rgl: 3D Visualization Using OpenGL*. R package version 0.98.1. URL: `https://CRAN.R-project.org/package=rgl`.

Ahmed, Rafi and David Gray (1996). "Immunological Memory and Protective Immunity: Understanding Their Relation". In: *Science* 272.5258, pp. 54–60.

Air, G M, A J Gibbs, W G Laver, and R G Webster (1990). "Evolutionary changes in influenza B are not primarily governed by antibody selection." In: *Proceedings of the National Academy of Sciences* 87.10, pp. 3884–3888.

Allen, Christopher D C, Takaharu Okada, and Jason G Cyster (2007). "Germinal-Center Organization and Cellular Dynamics". In: *Immunity* 27.2, pp. 190–202.

Amanna, I J, N E Carlson, and M K Slifka (2007). "Duration of humoral immunity to common viral and vaccine antigens". In: *New England Journal of Medicine* 357.19, pp. 1903–1915.

Andrews, Sarah F, Yunping Huang, Kaval Kaur, Lyubov I Popova, Irvin Y Ho, Noel T Pauli, Carole J Henry Dunand, William M Taylor, Samuel Lim, Min Huang, Xinyan Qu, Jane-Hwei Lee, Marlene Salgado-Ferrer, Florian Krammer, Peter Palese, Jens Wrammert, Rafi Ahmed, and Patrick C Wilson (2015). "Immune history profoundly affects broadly protective B cell responses to influenza". In: *Science Translational Medicine* 7.316, –316ra192.

Angeletti, Davide, James S Gibbs, Matthew Angel, Ivan Kosik, Heather D Hickman, Gregory M Frank, Suman R Das, Adam K Wheatley, Madhu Prabhakaran, David J Leggat, Adrian B McDermott, and Jonathan W Yewdell (2017). "Defining B cell immunodominance to viruses". In: *Nature Immunology* 18.4, pp. 456–463.

Angelova, L A and Y S Shvartsman (1982). "Original Antigenic Sin to Influenza in Rats". In: *Immunology* 46.1, pp. 183–188.

Arriola, Carmen S, Lynnette Brammer, Scott Epperson, Lenee Blanton, Krista Kniss, Desiree Mustaquim, Craig Steffens, Rosaline Dhara, Michelle Leon, Alejandro Perez, Sandra S Chaves, Jackie Katz, Teresa Wallis, Julie Villanueva, Xiyan Xu, Anwar Isa Abd Elal, Larisa Gubareva, Nancy Cox, Lyn Finelli, Joseph Bresee, et al. (2014). "Update: influenza activity - United States, September 29, 2013-February 8, 2014." In: *MMWR. Morbidity and mortality weekly report* 63.7, pp. 148–154.

Aymard, M, M Valette, B Lina, and D Thouvenot (1999). "Surveillance and impact of influenza in Europe. Groupe Régional d'Observation de la Grippe and European Influenza Surveillance Scheme." In: *Vaccine* 17 Suppl 1, S30–41.

Bachmann, M F and G T Jennings (2010). "Vaccine delivery: a matter of size, geometry, kinetics and molecular patterns". In: *Nature Reviews Immunology*.

Belongia, Edward A, Burney A Kieke, James G Donahue, Robert T Greenlee, Amanda Balish, Angie Foust, Stephen Lindstrom, and David K Shay (2009). "Effectiveness of Inactivated Influenza Vaccines

Varied Substantially with Antigenic Match from the 2004–2005 Season to the 2006–2007 Season". In: *Journal of Infectious Diseases* 199.2, pp. 159–167.

Bernasconi, Nadia L, Elisabetta Traggiai, and Antonio Lanzavecchia (2002). "Maintenance of Serological Memory by Polyclonal Activation of Human Memory B Cells". In: *Science* 298.5601, pp. 2199–2202.

Beyer, Walter E P, A M Palache, G Lüchters, J Nauta, and A D M E Osterhaus (2004). "Seroprotection rate, mean fold increase, seroconversion rate: which parameter adequately expresses seroresponse to influenza vaccination?" In: *Virus Research* 103.1-2, pp. 125–132.

B.J., Grant, Rodrigues A.P.C., ElSawy K.M., McCammon J.A., and Caves L.S.D. (2006). "Bio3D: An R package for the comparative analysis of protein structures." In: *Bioinformatics* 22, pp. 2695–2696.

Bodewes, R, G De Mutsert, F R M van der Klis, M Ventresca, S Wilks, D J Smith, M Koopmans, R A M Fouchier, A D M E Osterhaus, and G F Rimmelzwaan (2011). "Prevalence of Antibodies against Seasonal Influenza A and B Viruses in Children in Netherlands". In: *Clinical and Vaccine Immunology* 18.3, pp. 469–476.

Bridges, Carolyn Buxton, William W Thompson, Martin I Meltzer, Gordon R Reeve, Walter J Talamonti, Nancy J Cox, Heather A Lilac, Henrietta Hall, Alexander Klimov, and Keiji Fukuda (2000). "Effectiveness and Cost-Benefit of Influenza Vaccination of Healthy Working Adults: A Randomized Controlled Trial". In: *Jama* 284.13, pp. 1655–1663.

Carrat, F and A Flahault (2007). "Influenza vaccine: the challenge of antigenic drift". In: *Vaccine*.

Carrat, F, E Vergu, N M Ferguson, M Lemaitre, S Cauchemez, S Leach, and A J Valleron (2008). "Time Lines of Infection and Disease in Human Influenza: A Review of Volunteer Challenge Studies". In: *American Journal of Epidemiology* 167.7, pp. 775–785.

Cauchemez, Simon, Peter Horby, Annette Fox, Le Quynh Mai, Le Thi Thanh, Pham Quang Thai, Le Nguyen Minh Hoa, Nguyen Tran Hien, and Neil M Ferguson (2012). "Influenza Infection Rates, Measurement Errors and the Interpretation of Paired Serology". In: *PLoS Pathogens* 8.12, e1003061.

CDC (2017). *Types of Influenza Viruses*. URL: `https://www.cdc.gov/flu/about/viruses/types.htm`.

Centers for Disease Control and Prevention (CDC) (1998). "Update: influenza activity–United States, 1997-98 season." In: *MMWR. Morbidity and mortality weekly report* 47.10, pp. 196–200.

Chen, J M, Y J Guo, K Y Wu, J F Guo, M Wang, J Dong, Y Zhang, Z Li, and Y L Shu (2007). "Exploration of the emergence of the Victoria lineage of influenza B virus". In: *Archives of Virology* 152.2, pp. 415–422.

Cleveland, W S and S J Devlin (1988). "Locally weighted regression: an approach to regression analysis by local fitting". In: *Journal of the American Statistical Association*.

Cobey, Sarah and Scott E Hensley (2017). "Immune history and influenza virus susceptibility". In: *Current Opinion in Virology* 22, pp. 105–111.

Connelly, Dawn (2015). "Effectiveness of the seasonal flu vaccine and new production methods". In: *The Pharmaceutical Journal* 295.7882.

Coudeville, Laurent, Fabrice Bailleux, Benjamin Riche, Françoise Megas, Philippe Andre, and René Ecochard (2010a). "Relationship between haemagglutination-inhibiting antibody titres and clinical protection against influenza: development and application of a bayesian random-effects model". In: *BMC Medical Research Methodology* 10.1, p. 18.

– (2010b). "Relationship between haemagglutination-inhibiting antibody titres and clinical protection against influenza: development and application of a bayesian random-effects model". In: *BMC Medical Research Methodology* 10.1, p. 18.

Davenport, F M and A V Hennessy (1956). "A serologic recapitulation of past experiences with influenza A; antibody response to monovalent vaccine". In: *Journal of Experimental Medicine* 104.1, pp. 85–97.

– (1957). "Predetermination by infection and by vaccination of antibody response to influenza virus vaccines". In: *Journal of Experimental Medicine* 106.6, pp. 835–850.

Davenport, F M, A V Hennessy, and T Francis (1953). "Epidemiologic and immunologic significance of age distribution of antibody to antigenic variants of influenza virus". In: *Journal of Experimental Medicine* 98.6, pp. 641–656.

Davenport, F M, A V Hennessy, and C H Stuart-Harris (1955). "Epidemiology of Influenza Comparative Serological Observations in England and the United States". In: *The Lancet*.

Dávila, Javier, Gerardo Chowell, Víctor H Borja-Aburto, Cécile Viboud, Concepciòn Grajales Muñiz, and Mark Miller (2014). "Substantial Morbidity and Mortality Associated with Pandemic A/H1N1 Influenza in Mexico, Winter 2013-2014: Gradual Age Shift and Severity". In: *PLoS Currents* 6.

Duffy, Ken R, Cameron J Wellard, John F Markham, Jie H S Zhou, Ross Holmberg, Edwin D Hawkins, Jhagvaral Hasbold, Mark R Dowling, and Philip D Hodgkin (2012). "Activation-Induced B Cell Fates Are Selected by Intracellular Stochastic Competition". In: *Science* 335.6066, pp. 338–341.

Fonville, J M, S L James, M Ventresca, L Xue, S Wilks, Y Wong, G van der Net, R Bodewes, C A Russell, G F Rimmelzwaan, A Mosterin, N Masurel, J C de Jong, W E P Beyer, F Pistoor, A Palache, A Hurt, I G Barr, A D M E Osterhaus, R A M Fouchier, et al. (2013). "Antibody Landscapes: Quantifying the Antibody Immune Response". In: *6th Orthomyxovirus Research Conference*.

Fonville, J M, S H Wilks, S L James, A Fox, M Ventresca, M Aban, L Xue, T C Jones, N M H Le, Q T Pham, N D Tran, Y Wong, A Mosterin, L C Katzelnick, D Labonte, T T Le, G van der Net, E Skepner, C A Russell, T D Kaplan, et al. (2014). "Antibody landscapes after influenza virus infection or vaccination". In: *Science* 346.6212, pp. 996–1000.

Fox, Annette, Le Quynh Mai, Le Thi Thanh, Marcel Wolbers, Nguyen Le Khanh Hang, Pham Quang Thai, Nguyen Thi Thu Yen, Le Nguyen Minh Hoa, Juliet E Bryant, Tran Nhu Duong, Dang Dinh Thoang, Ian G Barr, Heiman Wertheim, Jeremy Farrar, Nguyen Tran Hien, and Peter Horby (2015). "Hemagglutination inhibiting antibodies and protection against seasonal and pandemic influenza infection". In: *Journal of Infection* 70.2, pp. 187–196.

Francis, T (1960). "On the doctrine of original antigenic sin". In: *Proceedings of the American Philosophical Society*, pp. 572–578.

Francis, T, J E Salk, and J J Quilligan (1947). "Experience with Vaccination Against Influenza in the Spring of 1947: A Preliminary Report." In: *American journal of public health and the nation's health* 37.8, pp. 1013–1016.

George Bernard Shaw (1944). *Everybody's political what's what*, p. 330.

Gerdil, C (2003). "The annual production cycle for influenza vaccine". In: *Vaccine* 21.16, pp. 1776–1779.

Glezen, W (2003). "Effect of maternal antibodies on the infant immune response". In: *Vaccine* 21.24, pp. 3389–3392.

Goodeve, A, C W Potter, A Clark, R Jennings, G C Schild, and R Yetts (1983). "A graded-dose study of inactivated, surface antigen influenza B vaccine in volunteers: reactogenicity, antibody response and protection to challenge virus infection". In: *Journal of Hygiene* 90.1, pp. 107–115.

Goodwin, K, C Viboud, and L Simonsen (2006). "Antibody response to influenza vaccination in the elderly: A quantitative review". In: *Vaccine* 24.8, pp. 1159–1169.

Gordon, Aubree, Guillermina Kuan, William Aviles, Nery Sanchez, Sergio Ojeda, Brenda Lopez, Lionel Gresh, Angel Balmaseda, and Eva Harris (2015). "The Nicaraguan pediatric influenza cohort study: design, methods, use of technology, and compliance". In: *BMC Infectious Diseases* 15.1, p. 1917.

Gordon, David L, Dimitar Sajkov, Yoshikazu Honda-Okubo, Samuel H Wilks, Malet Aban, Ian G Barr, and Nikolai Petrovsky (2016). "Human Phase 1 trial of low-dose inactivated seasonal influenza vaccine formulated with Advax™ delta inulin adjuvant". In: *Vaccine* 34.33, pp. 3780–3786.

Gostic, Katelyn M, Monique Ambrose, Michael Worobey, and James O Lloyd-Smith (2016). "Potent protection against H5N1 and H7N9 influenza via childhood hemagglutinin imprinting". In: *Science* 354.6313, pp. 722–726.

Grebe, Kristie M, Jonathan W Yewdell, and Jack R Bennink (2008). "Heterosubtypic immunity to influenza A virus: where do we stand?" In: *Microbes and Infection* 10.9, pp. 1024–1029.

Hald, A (1999). "On the history of maximum likelihood in relation to inverse probability and least squares". In: *Statistical Science*.

Hannoun, C (2004). "Immunogenicity and protective efficacy of influenza vaccination". In: *Virus Research* 103.1-2, pp. 133–138.

Hannoun, Claude (2014). "The evolving history of influenza viruses and influenza vaccines". In: *Expert review of vaccines* 12.9, pp. 1085–1094.

Hay, A J, V Gregory, A R Douglas, and Y P Lin (2001). "The evolution of human influenza viruses." In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 356.1416, pp. 1861–1870.

Hennessy, A V, F M Davenport, and T Francis (1955). "Studies of antibodies to strains of influenza virus in persons of different ages in sera collected in a postepidemic period." In: *The Journal of Immunology* 75.5, pp. 401–409.

Hirst, G K (1943). "Studies of antigenic differences among strains of influenza A by means of red cell agglutination". In: *Journal of Experimental Medicine* 78.5, pp. 407–423.

Hobson, D, R L Curry, A S Beare, and A Ward-Gardner (1972). "The role of serum haemagglutination-inhibiting antibody in protection against challenge infection with influenza A2 and B viruses." In: *Journal of Hygiene* 70.4, pp. 767–777.

Holzmann, H, M Kundi, K Stiasny, J Clement, P McKenna, C Kunz, and F X Heinz (1996). "Correlation between ELISA, hemagglutination inhibition, and neutralization tests after vaccination against tick-borne encephalitis". In: *Journal of Medical Virology* 48.1, pp. 102–107.

Horby, P, L Q Mai, A Fox, P Q Thai, N Thi Thu Yen, L T Thanh, N Le Khanh Hang, T N Duong, D D Thoang, J Farrar, M Wolbers, and N T Hien (2012). "The Epidemiology of Interpandemic and Pandemic Influenza in Vietnam, 2007-2010: The Ha Nam Household Cohort Study I". In: *American Journal of Epidemiology* 175.10, pp. 1062–1074.

Horsfall, F L and E R Rickard (1941). "Neutralizing antibodies in human serum after influenza A - The lack of strain specificity in the immunological response". In: *Journal of Experimental Medicine* 74.5, pp. 433–439.

Hsieh, Ying-Hen, Chen-An Tsai, Chien-Yu Lin, Jin-Hua Chen, Chwan-Chuen King, Day-Yu Chao, and Kuang-Fu Cheng (2014). "Asymptomatic ratio for seasonal H1N1 influenza infection among schoolchildren in Taiwan". In: *BMC Infectious Diseases* 14.1, p. 80.

Huang, Yongsheng, Aimee K Zaas, Arvind Rao, Nicolas Dobigeon, Peter J Woolf, Timothy Veldman, N Christine Øien, Micah T McClain, Jay B Varkey, Bradley Nicholson, Lawrence Carin, Stephen Kingsmore, Christopher W Woods, Geoffrey S Ginsburg, and Alfred O Hero III (2011). "Temporal Dynamics of Host Molecular Responses Differentiate Symptomatic and Asymptomatic Influenza A Infection". In: *PLoS Genetics* 7.8, e1002234.

Jensen, K E, F M Davenport, A V Hennessy, and T Francis (1956). "Characterization of influenza antibodies by serum absorption." In: *Journal of Experimental Medicine* 104.2, pp. 199–209.

Jong, J C de, A M Palache, Walter E P Beyer, G F Rimmelzwaan, A C M Boon, and A D M E Osterhaus (2003). "Haemagglutination-inhibiting antibody to influenza virus." In: *Developments in biologicals* 115, pp. 63–73.

Jong, Jan de, Walter Beyer, Abraham Palache, Guus Rimmelzwaan, and Albert Osterhaus (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." In: *Journal of Medical Virology*, pp. 94–99.

Kassebaum, Nicholas J, Ryan M Barber, Zulfiqar A Bhutta, Lalit Dandona, Peter W Gething, Simon I Hay, Yohannes Kinfu, Heidi J Larson, Xiaofeng Liang, Stephen S Lim, Alan D Lopez, Rafael Lozano, George A Mensah, Ali H Mokdad, Mohsen Naghavi, Christine Pinho, Joshua A Salomon, Caitlyn Steiner, Theo Vos, Haidong Wang, et al. (2016). "Global, regional, and national levels of maternal mortality, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015". In: *The Lancet* 388.10053, pp. 1775–1812.

Katzelnick, Leah C, Judith M Fonville, Gregory D Gromowski, Jose Bustos Arriaga, Angela Green, Sarah L James, Louis Lau, Magelda Montoya, Chunling Wang, Laura A VanBlargan, Colin A Russell, Hlaing Myat Thu, Theodore C Pierson, Philippe Buchy, John G Aaskov, Jorge L Muñoz-Jordán, Nikos Vasilakis, Robert V Gibbons, Robert B Tesh, Albert D M E Osterhaus, et al. (2015). "Dengue viruses cluster antigenically but not as discrete serotypes." In: *Science* 349.6254, pp. 1338–1343.

Kenney, Richard T, Sarah A Frech, Larry R Muenz, Christina P Villar, and Gregory M Glenn (2004). "Dose sparing with intradermal injection of influenza vaccine." In: *New England Journal of Medicine* 351.22, pp. 2295–2301.

Kilbourne, E D, W G Laver, J L Schulman, and R G Webster (1968). "Antiviral activity of antiserum specific for an influenza virus neuraminidase." In: *Journal of Virology* 2.4, pp. 281–288.

Kilbourne, E D, B E Johansson, and B Grajower (1990). "Independent and disparate evolution in nature of influenza A virus hemagglutinin and neuraminidase glycoproteins." In: 87.2, pp. 786–790.

Kilbourne, Edwin D (2006). "Influenza Pandemics of the 20th Century". In: *Prevalence* 12.1, pp. 9–14.

Klimov, A, L Simonsen, K Fukuda, and N Cox (1999). *Surveillance and impact of influenza in the United States*. Vaccine.

Koel, Bjorn F, David F Burke, Theo M Bestebroer, Stefan van der Vliet, Gerben C M Zondag, Gaby Vervaet, Eugene Skepner, Nicola S Lewis, Monique I J Spronken, Colin A Russell, Mikhail Y Eropkin, Aeron C Hurt, Ian G Barr, Jan C de Jong, Guus F Rimmelzwaan, Albert D M E Osterhaus, Ron A M

Fouchier, and Derek J Smith (2013). “Substitutions Near the Receptor Binding Site Determine Major Antigenic Change During Influenza Virus Evolution”. In: *Science* 342.6161, pp. 976–979.

Koopmans, M, E de Bruin, G J Godeke, I Friesema, R van Gageldonk, M Schipper, A Meijer, R van Binnendijk, G F Rimmelzwaan, M D de Jong, A Buisman, J van Beek, D van de Vijver, and J Reimerink (2012). “Profiling of humoral immune responses to influenza viruses by using protein microarray”. In: *Clinical Microbiology and Infection* 18.8, pp. 797–807.

Krammer, Florian and Peter Palese (2013). “Influenza virus hemagglutinin stalk-based antibodies and vaccines”. In: *Current Opinion in Virology* 3.5, pp. 521–530.

Kucharski, Adam J, Justin Lessler, Jonathan M Read, Huachen Zhu, Chao Qiang Jiang, Yi Guan, Derek A T Cummings, and Steven Riley (2015). “Estimating the Life Course of Influenza A(H3N2) Antibody Responses from Cross-Sectional Data”. In: *PLOS Biol* 13.3, e1002082.

Künzel, W, H Glathe, H Engelmann, and Ch Van Hoecke (1996). “Kinetics of humoral antibody response to trivalent inactivated split influenza vaccine in subjects previously vaccinated or vaccinated for the first time”. In: *Vaccine* 14.12, pp. 1108–1110.

Kurosaki, Tomohiro, Kohei Kometani, and Wataru Ise (2015). “Memory B cells”. In: *Nature Reviews Immunology* 15.3, pp. 149–159.

Lambert, Linda C and Anthony S Fauci (2010). “Influenza vaccines for the future.” In: *New England Journal of Medicine* 363.21, pp. 2036–2044.

Lapedes, A and R Farber (2001). “The geometry of shape space: application to influenza”. In: *Journal of Theoretical Biology*.

Lessler, Justin, Steven Riley, Jonathan M Read, Shuying Wang, Huachen Zhu, Gavin J D Smith, Yi Guan, Chao Qiang Jiang, and Derek A T Cummings (2012). “Evidence for Antigenic Seniority in Influenza A (H3N2) Antibody Responses in Southern China”. In: *PLoS Pathogens* 8.7, e1002802.

Leuridan, E and P Van Damme (2007). “Passive transmission and persistence of naturally acquired or vaccine-induced maternal antibodies against measles in newborns”. In: *Vaccine* 25.34, pp. 6296–6304.

Leuridan, E, N Hens, V Hutse, M Aerts, and P Van Damme (2011). “Kinetics of maternal antibodies against rubella and varicella in infants”. In: *Vaccine* 29.11, pp. 2222–2226.

Li, Yang, Jaclyn L Myers, David L Bostick, Colleen B Sullivan, Jonathan Madara, Susanne L Linderman, Qin Liu, Donald M Carter, Jens Wrammert, Susanna Esposito, Nicola Principi, Joshua B Plotkin, Ted M Ross, Rafi Ahmed, Patrick C Wilson, and Scott E Hensley (2013). “Immune history shapes specificity of pandemic H1N1 influenza antibody responses”. In: *Journal of Experimental Medicine* 210.8, pp. 1493–1500.

Li, Yanming, Kerby Shedden, Brenda W. Gillespie, and John A. Gillespie (2016). *clikcorr: Censoring Data and Likelihood-Based Correlation Estimation*. R package version 1.0. URL: `https://CRAN.R-project.org/package=clikcorr`.

Linderman, Susanne L, Benjamin S Chambers, Seth J Zost, Kaela Parkhouse, Yang Li, Christin Herrmann, Ali H Ellebedy, Donald M Carter, Sarah F Andrews, Nai-Ying Zheng, Min Huang, Yunping Huang, Donna Strauss, Beth H Shaz, Richard L Hodinka, Gustavo Reyes-Terán, Ted M Ross, Patrick C Wilson, Rafi Ahmed, Jesse D Bloom, et al. (2014). "Potential antigenic explanation for atypical H1N1 infections among middle-aged adults during the 2013-2014 influenza season." In: *Proceedings of the National Academy of Sciences of the United States of America* 111.44, pp. 15798–15803.

Liu, Yong-Jun, Clarissa Barthélémy, Odette De Bouteiller, Christophe Arpin, Durand Isabelle, and Jacques Banchereau (1995). "Memory B cells from human tonsils colonize mucosal epithelium and directly present antigen to T cells by Rapid Up-Regulation of B7-1 and B7-2". In: *Immunity* 2.3, pp. 239–248.

Magill, T P and T Francis Jr (1936). "Antigenic differences in strains of human influenza virus". In: *Proceedings of the Society for Experimental Biology and Medicine*, pp. 463–466.

Maher, J A and J DeStefano (2004). "The ferret: an animal model to study influenza virus". In: *Lab animal* 33.9, p. 50.

Manz, Rudolf A and Andreas Radbruch (2002). "Plasma cells for a lifetime?" In: *European Journal of Immunology* 32.4, pp. 923–927.

Manz, Rudolf A, Andreas Thiel, and Andreas Radbruch (1997). "Lifetime of plasma cells in the bone marrow". In: *Nature* 388.6638, pp. 133–134.

Miller, Matthew S, Thomas J Gardner, Florian Krammer, Lauren C Aguado, Domenico Tortorella, Christopher F Basler, and Peter Palese (2013). "Neutralizing Antibodies Against Previously Encountered Influenza Virus Strains Increase over Time: A Longitudinal Analysis". In: *Science Translational Medicine* 5.198, 198ra107–198ra107.

Mitchcll, D M, A J McMichael, and J R Lamb (1985). "The immunology of influenza". In: *British Medical Bulletin* 41.1, pp. 80–85.

Moa, Aye M, Abrar A Chughtai, David J Muscatello, Robin M Turner, and C Raina MacIntyre (2016). "Immunogenicity and safety of inactivated quadrivalent influenza vaccine in adults: A systematic review and meta-analysis of randomised controlled trials". In: *Vaccine* 34.35, pp. 4092–4102.

Mögling, Ramona, Mathilde J Richard, Stefan van der Vliet, Ruud van Beek, Eefje J A Schrauwen, Monique I Spronken, Guus F Rimmelzwaan, and Ron A M Fouchier (2017). "Neuraminidase-mediated haemagglutination of recent human influenza A(H3N2) viruses is determined by arginine 150 flanking the neuraminidase catalytic site". In: *Journal of General Virology* 98.6, pp. 1274–1281.

Molinari, Noelle-Angelique M, Ismael R Ortega-Sanchez, Mark L Messonnier, William W Thompson, Pascale M Wortley, Eric Weintraub, and Carolyn B Bridges (2007). "The annual impact of seasonal influenza in the US: measuring disease burden and costs." In: *Vaccine* 25.27, pp. 5086–5096.

Morell, A, W D Terry, and T A Waldmann (1970). "Metabolic properties of IgG subclasses in man". In: *Journal of Clinical Investigation* 49.4, pp. 673–680.

Nelson, Martha I, Cécile Viboud, Lone Simonsen, Ryan T Bennett, Sara B Griesemer, Kirsten St George, Jill Taylor, David J Spiro, Naomi A Sengamalay, Elodie Ghedin, Jeffery K Taubenberger, and Edward C Holmes (2008). "Multiple Reassortment Events in the Evolutionary History of H1N1 Influenza A Virus Since 1918". In: *PLoS Pathogens* 4.2, e1000012.

Nobusawa, E and K Sato (2006). "Comparison of the mutation rates of human influenza A and B viruses". In: *Journal of Virology* 80.7, pp. 3675–3678.

Nordin, James, John Mullooly, Sung Poblete, Raymond Strikas, Richard Petrucci, Feifei Wei, Bill Rush, Benjamin Safirstein, Deborah Wheeler, and Kristin L Nichol (2001). "Influenza Vaccine Effectiveness in Preventing Hospitalizations and Deaths in Persons 65 Years or Older in Minnesota, New York, and Oregon: Data from 3 Health Plans". In: *Journal of Infectious Diseases* 184.6, pp. 665–670.

Nutt, Stephen L, Philip D Hodgkin, David M Tarlinton, and Lynn M Corcoran (2015). "The generation of antibody-secreting plasma cells". In: *Nature Reviews Immunology* 15.3, pp. 160–171.

Odendahl, Marcus, Henrik Mei, Bimba F Hoyer, Annett M Jacobi, Arne Hansen, Gwendolin Muehlinghaus, Claudia Berek, Falk Hiepe, Rudi Manz, Andreas Radbruch, and Thomas Dörner (2005). "Generation of migratory antigen-specific plasma blasts and mobilization of resident plasma cells in a secondary immune response". In: *Blood* 105.4, pp. 1614–1621.

Ohmit, Suzanne E, John C Victor, Judy R Rotthoff, Esther R Teich, Rachel K Truscon, Laura L Baum, Bhavya Rangarajan, Duane W Newton, Matthew L Boulton, and Arnold S Monto (2009). "Prevention of Antigenically Drifted Influenza by Inactivated and Live Attenuated Vaccines". In: *New England Journal of Medicine* 355.24, pp. 2513–2522.

Okuno, Y, Y Isegawa, F Sasao, and S Ueda (1993). "A common neutralizing epitope conserved between the hemagglutinins of influenza A virus H1 and H2 strains." In: *Journal of Virology* 67.5, pp. 2552–2558.

Osterholm, Michael T, Nicholas S Kelley, Alfred Sommer, and Edward A Belongia (2012). "Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis". In: *The Lancet Infectious Diseases* 12.1, pp. 36–44.

Pagas, H., P. Aboyoun, R. Gentleman, and S. DebRoy (2016). *Biostrings: String objects representing biological sequences, and matching algorithms*. R package version 2.42.1.

Palese, Peter and Taia T Wang (2011). "Why Do Influenza Virus Subtypes Die Out? A Hypothesis". In: *mBio* 2.5, e00150–11–e00150–11.

Pape, Kathryn A, Drew M Catron, Andrea A Itano, and Marc K Jenkins (2007). "The Humoral Immune Response Is Initiated in Lymph Nodes by B Cells that Acquire Soluble Antigen Directly in the Follicles". In: *Immunity* 26.4, pp. 491–502.

Paradis, E., J. Claude, and K. Strimmer (2004). "APE: analyses of phylogenetics and evolution in R language". In: *Bioinformatics* 20, pp. 289–290.

Parrish, Colin R, Pablo R Murcia, and Edward C Holmes (2015). "Influenza Virus Reservoirs and Intermediate Hosts: Dogs, Horses, and New Possibilities for Influenza Virus Exposure of Humans". In: *Journal of Virology* 89.6, pp. 2990–2994.

Parvin, J D, A Moscona, W T Pan, J M Leider, and P Palese (1986). "Measurement of the mutation rates of animal viruses: influenza A virus and poliovirus type 1." In: *Journal of Virology* 59.2, pp. 377–383.

Paul, W E (2013). *Fundamental immunology*. 7th ed. Lippincott Williams & Wilkins.

Pica, Natalie, Rong Hai, Florian Krammer, Taia T Wang, Jad Maamary, Dirk Eggink, Gene S Tan, Jens C Krause, Thomas Moran, Cheryl R Stein, David Banach, Jens Wrammert, Robert B Belshe, Adolfo García-Sastre, and Peter Palese (2012). "Hemagglutinin stalk antibodies elicited by the 2009 pandemic influenza virus as a mechanism for the extinction of seasonal H1N1 viruses." In: *Proceedings of the National Academy of Sciences of the United States of America* 109.7, pp. 2573–2578.

Plotkin, S A (2005). "Vaccines: past, present and future". In: *Nature Medicine* 10.4s, S5.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: `https://www.R-project.org/`.

Radbruch, Andreas, Gwendolin Muehlinghaus, Elke O Luger, Ayako Inamine, Kenneth G C Smith, Thomas Dörner, and Falk Hiepe (2006). "Competence and competition: the challenge of becoming a long-lived plasma cell". In: *Nature Reviews Immunology* 6.10, pp. 741–750.

Rajewsky, K and B Schittek (1990). "Maintenance of B-cell memory by long-lived cells generated from proliferating precursors". In: *Nature* 346.6286, p. 749.

Rambaut, Andrew, Oliver G Pybus, Martha I Nelson, Cécile Viboud, Jeffery K Taubenberger, and Edward C Holmes (2008). "The genomic and epidemiological dynamics of human influenza A virus". In: *Nature* 453.7195, pp. 615–619.

Raue, Andreas, Johan Karlsson, Maria Pia Saccomani, Mats Jirstrand, and Jens Timmer (2014). "Comparison of approaches for parameter identifiability analysis of biological systems." In: *Bioinformatics* 30.10, pp. 1440–1448.

Recker, Mario, Oliver G Pybus, Sean Nee, and Sunetra Gupta (2007). "The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types." In: *Proceedings of the National Academy of Sciences* 104.18, pp. 7711–7716.

Ripatti, Samuli and Juni Palmgren (2000). "Estimation of Multivariate Frailty Models Using Penalized Partial Likelihood". In: *Biometrics* 56.4, pp. 1016–1022.

Rota, P A, M L Hemphill, T Whistler, H L Regnery, and A P Kendal (1992). "Antigenic and Genetic Characterization of the Haemagglutinins of Recent Cocirculating Strains of Influenza B Virus". In: *Journal of General Virology* 73.10, pp. 2737–2742.

Rota, Paul A, Teresa R Wallis, Maurice W Harmon, Jennifer S Rota, Alan P Kendal, and Kuniaki Nerome (1990). "Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983". In: *Virology* 175.1, pp. 59–68.

Rowe, T, R A Abernathy, J Hu-Primmer, W W Thompson, X Lu, W Lim, K Fukuda, N J Cox, and J M Katz (1999a). "Detection of antibody to avian influenza A (H5N1) virus in human serum by using a combination of serologic assays." In: *Journal of Clinical Microbiology* 37.4, pp. 937–943.

– (1999b). "Detection of antibody to avian influenza A (H5N1) virus in human serum by using a combination of serologic assays." In: *Journal of Clinical Microbiology* 37.4, pp. 937–943.

Russell, Colin A, Terry C Jones, Ian G Barr, Nancy J Cox, Rebecca J Garten, Vicky Gregory, Ian D Gust, Alan W Hampson, Alan J Hay, Aeron C Hurt, Jan C de Jong, Anne Kelso, Alexander I Klimov, Tsutomu Kageyama, Naomi Komadina, Alan S Lapedes, Yi P Lin, Ana Mosterin, Masatsugu Obuchi, Takato Odagiri, et al. (2008). "The Global Circulation of Seasonal Influenza A (H3N2) Viruses". In: *Science* 320.5874, pp. 340–346.

Sandt, Carolien E van de, Rogier Bodewes, Guus F Rimmelzwaan, and Rory D de Vries (2015). "Influenza B viruses: not to be discounted". In: *Future microbiology* 10.9, pp. 1447–1465.

Schliep, Klaus, Potts, Alastair J., Morrison, David A., Grimm, and Guido W. (2017). "Intertwining phylogenetic trees and networks". In: *Methods in Ecology and Evolution* 8.10, pp. 1212–1220. ISSN: 2041-210X. DOI: 10.1111/2041-210X.12760. URL: http://dx.doi.org/10.1111/2041-210X.12760.

Schliep, K.P. (2011). "phangorn: phylogenetic analysis in R". In: *Bioinformatics* 27.4, pp. 592–593. URL: https://doi.org/10.1093/bioinformatics/btq706.

Skowronski, Danuta M, Naveed Z Janjua, Gaston De Serres, Suzana Sabaiduc, Alireza Eshaghi, James A Dickinson, Kevin Fonseca, Anne-Luise Winter, Jonathan B Gubbay, Mel Krajden, Martin Petric, Hugues Charest, Nathalie Bastien, Trijntje L Kwindt, Salaheddin M Mahmud, Paul Van Caeseele, and Yan Li (2014). "Low 2012–13 Influenza Vaccine Effectiveness Associated with Mutation in the Egg-Adapted H3N2 Vaccine Strain Not Antigenic Drift in Circulating Viruses". In: *PloS one* 9.3, e92153.

Slifka, M K and R Ahmed (1996). "Long-term humoral immunity against viruses: revisiting the issue of plasma cell longevity". In: *Trends in microbiology* 4.10, pp. 394–400.

Slifka, Mark K, Rustom Antia, Jason K Whitmire, and Rafi Ahmed (1998). "Humoral Immunity Due to Long-Lived Plasma Cells". In: *Immunity* 8.3, pp. 363–372.

Smith, D J, S Forrest, D H Ackley, and A S Perelson (1999). "Variable efficacy of repeated annual influenza vaccination." In: 96.24, pp. 14001–14006.

Smith, Derek J, Alan S Lapedes, Jan C de Jong, Theo M Bestebroer, Guus F Rimmelzwaan, Albert D M E Osterhaus, and Ron A M Fouchier (2004). "Mapping the antigenic and genetic evolution of influenza virus." In: *Science* 305.5682, pp. 371–376.

Smith, W, C H Andrewes, and P P Laidlaw (1933). "A virus obtained from influenza patients". In: *The Lancet* 222.5732, pp. 66–68.

Srivastava, Barkha, Paulina Błażejewska, Manuela Heßmann, Dunja Bruder, Robert Geffers, Susanne Mauel, Achim D Gruber, and Klaus Schughart (2009). "Host Genetic Background Strongly Influences the Response to Influenza A Virus Infections". In: *PloS one* 4.3, e4857.

St Groth, S Fazekas de and R G Webster (1966a). "Disquisitions on original antigenic sin: I. Evidence in man". In: *Journal of Experimental Medicine* 124.3, pp. 331–345.

– (1966b). *Disquisitions on original antigenic sin. II. Proof in lower creatures.* J Exp Med.

Stöhr, Klaus, Doris Bucher, Tony Colgate, and John Wood (2012). "Influenza Virus Surveillance, Vaccine Strain Selection, and Manufacture". In: *Influenza Virus*. Totowa, NJ: Humana Press, pp. 147–162.

Tangye, Stuart G, Danielle T Avery, Elissa K Deenick, and Philip D Hodgkin (2003). "Intrinsic Differences in the Proliferation of Naive and Memory Human B Cells as a Mechanism for Enhanced Secondary Immune Responses". In: *The Journal of Immunology* 170.2, pp. 686–694.

Terry M. Therneau and Patricia M. Grambsch (2000). *Modeling Survival Data: Extending the Cox Model*. New York: Springer. ISBN: 0-387-98784-3.

Therneau, Terry (2014). *deming: Deming, Thiel-Sen and Passing-Bablock Regression*. R package version 1.0-1. URL: `https://CRAN.R-project.org/package=deming`.

Therneau, Terry M (2015). *A Package for Survival Analysis in S*. version 2.38. URL: `https://CRAN.R-project.org/package=survival`.

Therneau, Terry M, Patricia M Grambsch, and V Shane Pankratz (2012). "Penalized Survival Models and Frailty". In: *Journal of Computational and Graphical Statistics* 12.1, pp. 156–175.

Tobin, James (1958). "Estimation of Relationships for Limited Dependent Variables". In: *Econometrica* 26.1, p. 24.

Tong, Suxiang, Xueyong Zhu, Yan Li, Mang Shi, Jing Zhang, Melissa Bourgeois, Hua Yang, Xianfeng Chen, Sergio Recuenco, Jorge Gomez, Li-Mei Chen, Adam Johnson, Ying Tao, Cyrille Dreyfus, Wenli Yu, Ryan McBride, Paul J Carney, Amy T Gilbert, Jessie Chang, Zhu Guo, et al. (2013). "New World Bats Harbor Diverse Influenza A Viruses". In: *PLoS Pathogens* 9.10, e1003657.

Tough, D F, P Borrow, and J Sprent (1996). “Induction of bystander T cell proliferation by viruses and type I interferon in vivo”. In: *Science* 272.5270, p. 1947.

Tricco, Andrea C, Ayman Chit, Charlene Soobiah, David Hallett, Genevieve Meier, Maggie H Chen, Mariam Tashkandi, Chris T Bauch, and Mark Loeb (2013). “Comparing influenza vaccine efficacy against mismatched and matched strains: a systematic review and meta-analysis”. In: *BMC Medicine* 11.1, p. 153.

Virelizier, Jean Louis (1975). “Host defenses against influenza virus: the role of anti-hemagglutinin antibody.” In: *The Journal of Immunology* 115.2, pp. 434–439.

Virelizier, Jean Louis, A C Allison, and G C Schild (1974). “Antibody responses to antigenic determinants of influenza virus hemagglutinin. II. Original antigenic sin: a bone marrow-derived lymphocyte memory phenomenon modulated by thymus-derived lymphocytes.” In: *Journal of Experimental Medicine* 140.6, pp. 1571–1578.

Wang, J, S P Hilchey, O Hyrien, N Huertas, and S Perry (2015). “Multi-Dimensional Measurement of Antibody-Mediated Heterosubtypic Immunity to Influenza”. In: *PloS one* 10.6, e0129858.

Watanaveeradej, V, T P Endy, and R Samakoses (2003). “Transplacentally transferred maternal-infant antibodies to dengue virus”. In: *The American journal of tropical medicine and hygiene* 69.2, pp. 123–128.

Webster, R G (1966). “Original antigenic sin in ferrets: the response to sequential infections with influenza viruses.” In: *The Journal of Immunology* 97.2, pp. 177–183.

Webster, R G, W G Laver, and E D Kilbourne (1968). “Reactions of Antibodies with Surface Antigens of Influenza Virus”. In: *Journal of General Virology* 3.3, pp. 315–326.

Webster, R G, W J Bean, O T Gorman, T M Chambers, and Y Kawaoka (1992). “Evolution and ecology of influenza A viruses.” In: *Microbiological Reviews* 56.1, pp. 152–179.

Weill, Jean-Claude, Simon Le Gallou, Yi Hao, and Claude-Agnès Reynaud (2013). “Multiple players in mouse B cell memory”. In: *Current Opinion in Immunology* 25.3, pp. 334–338.

Wikramaratna, P S and A Rambaut (2015). “Relationship between haemagglutination inhibition titre and immunity to influenza in ferrets”. In: *Vaccine*.

Wilks, Sam, Miranda de Graaf, Derek J Smith, and David F Burke (2012). “A review of influenza haemagglutinin receptor binding as it relates to pandemic properties”. In: *Vaccine* 30.29, pp. 4369–4376.

Wood, J (2003). “The influenza vaccine licensing process”. In: *Vaccine* 21.16, pp. 1786–1788.

World Health Organization (1997). “Influenza: antigenic analysis of recent influenza virus isolates and influenza activity in the southern hemisphere”. In: *Weekly Epidemiological Record* 72.39, pp. 293–293.

– (2009). “Recommended composition of influenza virus vaccines for use in the 2010 influenza season (southern hemisphere winter)”. In: *Weekly Epidemiological Record* 84.41, pp. 421–432.

Wu, Nicholas C and Ian A Wilson (2017). “A Perspective on the Structural and Functional Constraints for Immune Evasion: Insights from Influenza Virus”. In: *Journal of Molecular Biology* 429.17, pp. 2694–2709.

Yamashita, Makoto, Mark Krystal, Walter M Fitch, and Peter Palese (1988). “Influenza B virus evolution: Co-circulating lineages and comparison of evolutionary pattern with those of influenza A and C viruses”. In: *Virology* 163.1, pp. 112–122.

Yamayoshi, Seiya, Mariko Watanabe, Hideo Goto, and Yoshihiro Kawaoka (2016). “Identification of a Novel Viral Protein Expressed from the PB2 Segment of Influenza A Virus”. In: *Journal of Virology* 90.1, pp. 444–456.

Yu, Xiaocong, Tshidi Tsibane, Patricia A McGraw, Frances S House, Christopher J Keefer, Mark D Hicar, Terrence M Tumpey, Claudia Pappas, Lucy A Perrone, Osvaldo Martinez, James Stevens, Ian A Wilson, Patricia V Aguilar, Eric L Altschuler, Christopher F Basler, James E Crowe, and Jr (2008). “Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors”. In: *Nature* 455.7212, pp. 532–536.

Yuseff, Maria-Isabel, Paolo Pierobon, Anne Reversat, and Ana-Maria Lennon-Duménil (2013). “How B cells capture, process and present antigens: a crucial role for cell polarity”. In: *Nature Reviews Immunology* 13.7, pp. 475–486.

Zinkernagel, Rolf M and Hans Hengartner (2001). “Regulation of the Immune Response by Antigen”. In: *Science* 293.5528, pp. 251–253.