Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Zheng-Rong Tiger Li Date

EVOLUTION OF CD8 T CELL EPITOPES OF INFLUENZA A VIRUS

By

Zheng-Rong Tiger Li
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Immunology and Molecular Pathogenesis

Rustom Antia, Ph.D.
Advisor

Jacob E. Kohlmeier, Ph.D.
Advisor

John D. Altman, Ph.D.
Committee Member

Anice C. Lowen, Ph.D.
Committee Member

Lance A. Waller, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

EVOLUTION OF CD8 T CELL EPITOPES OF INFLUENZA A VIRUS

By

ZHENG-RONG TIGER LI
Master of Science

Advisor: Rustom Antia, Ph.D.
Advisor: Jacob E. Kohlmeier, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Graduate Division of Biological and Biomedical Science
Immunology and Molecular Pathogenesis
2020

Abstract

EVOLUTION OF CD8 T CELL EPITOPES OF INFLUENZA A VIRUS

By Zheng-Rong Tiger Li

Influenza-specific CD8 T cells protect against severe pathology caused by heterosubtypic infection, mainly due to the conservation of influenza CD8 T cell epitopes. This conservation may be explained by (1) functional constraints, (2) small selection on the escaping mutant, and (3) human MHC polymorphism. We aim to model the interplay of these factors and quantify the parameters experimentally.

We constructed a population genetics model that incorporates fitness cost, selective advantage, and MHC allele frequency. Additionally, the development of compensatory immunity was considered by an ordinary differential equation system. Guided by the model, we designed a series of experiments to estimate the parameters. We introduced an escaping mutation to the nucleoprotein (NP) under PR8 (H1N1) back-ground. C57BL/6 mice of different immune status were challenged with a 1:1 mixture of wild-type and mutant PR8 and the lung viral load of each strain was measured by digital droplet PCR.

The population genetics model predicts that, within biologically reasonable range of parameters, it takes 3 to 10 years for an CD8 T cell-escaping mutant to reach 50% of prevalence in the population, if it does. The development of compensatory immunity further renders the invasion 'transiently' when the compensation level is beyond a threshold determined by above-mentioned parameters. Based on the area under in vivo viral growth curves, the selective advantage of the mutant PR8 is 0.27 [0.08, 0.2] in the intranasally primed C57BL/6 mice. Interestingly, the selective advantage significantly reduced in the intramuscularly primed C57BL/6 mice (0.04 [0.02, 0.06]).

We concluded that a CD8 T cell-escaping mutant either cannot invade or invades extremely slowly, if the invasion entirely depends on selection. The lower selective ad-vantage in intramuscularly primed mice suggests the lung-resident memory CD8 T cells may be the primary source of selection pressure on influenza virus.

EVOLUTION OF CD8 T CELL EPITOPES OF INFLUENZA A VIRUS

By

ZHENG-RONG TIGER LI
Master of Science

Advisor: Rustom Antia, Ph.D.
Advisor: Jacob E. Kohlmeier, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Graduate Division of Biological and Biomedical Science
Immunology and Molecular Pathogenesis
2020

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Introduction chapter is structured in four sections. In sections 1 and 2, I will briefly review the important features of influenza A virus (IAV) and the human immune system, respectively. Section 3 is dedicated to discussing the evolution of IAV and the role of adaptive immunity in shaping this evolution. In section 4, I will propose the scientific questions and specific aims of my dissertation research.

1.1 Influenza A Virus

Influenza viruses belong to the *Orthomyxoviridae* (*ortho-*: correct or normal; *myxa*: mucus) family and are categorized into four genera: A, B, C, and D [1]. All types of influenza viruses cause diseases in vertebrates, particularly in mammals and birds [2, 3]. This dissertation mainly focuses on influenza A virus (IAV), which has been a major public health concern for decades [4].

1.1.1 Virology

IAV is an enveloped virus with sphere, elliptical, filamentous, or sometimes irregular shape [5]. The genome of a complete influenza A virion consists of 8 single-stranded, negative-sense RNA gene segments [1]. Each of the gene segments encodes a full-length viral protein; RNA-dependent RNA polymerase (RdRP) subunit 1 (PB1), RdRP subunit 2 (PB2), RdRP acidic protein (PA), hemag-

glutinin (HA), nucleoprotein (NP), neuraminidase (NA), matrix (M) protein, and non-structural protein (NS) [1, 6, 7]. In addition, three of the gene segments (PB1, M, NS) contain another open reading frame and thus can be translated into shorter proteins: NS2 [8], which was later renamed as nuclear export protein (NEP) [9], M2 [10], and PB1-F2 [11]. Their specific functions will be discussed below.

The physical structure of the virion can be roughly dissected into three layers. The outer layer is a lipid envelope derived from the host cell membrane, from which the HA trimers and NA tetramers protrude; additionally, a small number of M2 tetramers transverse the envelope [12]. Beneath the envelope, the M proteins form a shell that supports the membrane and encapsulates the viral genome [12, 13]. Within the shell, the RNA strands wrap around the NP and are attached to the RdRP complex [6, 7, 12].

Infection of IAV starts from the binding of HA trimers to host cell receptors, mainly sialic acid moieties expressed on epithelial cells lining the respiratory tract [14, 15]. The virion enters host cells via receptor-mediated endocytosis [16]. As pH of the endosome drops, HA trimers undergo a conformational change, allowing the exposure of fusion peptide, which then inserts into the endosomal membrane to serve as an anchor, assisting the fusion of the viral envelope and endosomal membrane [17, 18, 19]. Meanwhile, the flux of protons into the endosome changes the conformation of M2 proteins, which function as proton channels that allow protons to enter the interior of virion [20], a process required for proper uncoating and release of the viral genome to the cytoplasm [21].

The negative-sense viral RNAs (vRNA) translocate into the nucleus [22], where the viral RdRP transcribes vRNAs to (1) complementary RNAs (cRNA) that serve as templates for the replication of viral genome and (2) viral messenger RNAs (mRNA) that are translated into viral proteins via cellular ribosomes [23]. HA, NA, and M2 proteins are transported through the *trans*-Golgi secretory pathway [24, 25, 26], while proteins required for replication and translocation of the viral

genome are sent back to the nucleus (reviewed in [27]). Newly synthesized vRNAs form a stable ribonucleoprotein complex by binding to the viral NP and RdRP [28] and are exported to the cytoplasm, facilitated by M1 [22] and NEP [9]. These mature gene segments are then packaged into the capsid formed by M1. The viral particles bud out and are released from the cell membrane as the NA cleaves the sialic acids on host membrane proteins (reviewed in [29, 30]).

The nature of the segmented RNA genome highlights two essential features of IAV. First, viruses with single-stranded RNA genomes tend to have higher mutation rates compared to retroviruses or viruses with double-stranded RNA or DNA genomes [31, 32]. A recent estimate for the mutation rate of IAV was within the magnitude of 10^{-4} substitutions per nucleotide per strand copied, or on average 2 to 3 mutations per replicated genome [33]. Second, the segmented genome allows exchange of gene segments between virions that co-infect the same cell, a process called *reassortment* [34]. Through gene reassortment, a highly pathogenic zoonotic IAV may acquire the ability to transmit between humans and trigger a deadly pandemic, such as the viruses responsible for the pandemics of H1N1 in 1918, H2N2 in 1957, H3N2 in 1968, and H1N1 in 2009 [35, 36] (see the next section for detailed discussions).

1.1.2 Ecology and Epidemiology

IAV is categorized into subtypes, termed HxNy, where x and y refer to the types of HA and NA, respectively. Today, 18 types of HA and 11 types of NA have been discovered in nature [1, 35]. Aquatic wild birds are the natural reservoirs of all IAV subtypes [2, 35] except H17 and H18, which have been exclusively found in bats [37, 38]. Some of the subtypes have established stable lineages in mammals. For example, the H1N1 and H3N2 subtypes have been co-circulating in human population for decades [2], and the H3N8 subtype is circulating in horse [39].

This host preference for birds or mammals is mainly determined by (1) the specificity of HA and

(2) the optimal temperature of viral replication [7, 15]. HAs of avian lineage IAVs have high affinity for α-2,3-linked sialic acid (α-2,3-SA) moieties, which are mainly expressed by the epithelial cells of the gastrointestinal (GI) tract of wild aquatic birds [40, 41, 42]. In addition, avian IAVs replicate best at 42°C, the core body temperature of birds [43]. In contrast, human lineage IAVs prefers to bind α-2,6-SA moieties, commonly found on epithelial cells in the human upper respiratory tract [44], and optimally replicate at 33°C [45]. Some mammals and domestic birds, such as swine, turkey, and quail, express both SA moieties in the respiratory epithelium, and thus may serve as a 'mixing vessel' for different lineages of IAVs [46, 47, 48]. In this section, I will focus on avian IAVs, human IAVs, and the concerns about spillover.

Avian IAVs

IAVs are transmitted among aquatic wild birds via the fecal-oral route and result in largely asymptomatic infection [49]. However, some lineages of H5 and H7 subtypes can cause severe pathology in domestic poultry, such as chickens and quail [35, 49]. These lineages are known as *highly pathogenic avian influenza* (HPAI), first detected in 1996 as an H5N1 subtype in geese in China [50]. In 2005, an unprecedented outbreak of HPAI H5N1 in Qinghai Lake in central China killed around 6,000 migratory birds, and this lineage subsequently spread to Europe and Africa through the Central Asian flyway [51, 52]. Since then, HPAI H5N1 infections have become frequent in domestic poultry in Asia, Europe, and Africa. In a report published by the United Nations Food and Agricultural Organizations in 2013, five asian countries (China, Viet Nam, Indonesia, Bangladesh, and eastern India), along with Egypt, were considered endemic of H5N1 [53]. In the U.S., an outbreak of HPAI in wild or domestic birds was detected from five northwestern states (California, Idaho, Oregon, Utah, and Washington) during December 2014 and January 2015. The viruses responsible were of H5N1, H5N2, and H5N8 subtypes, originated from Asia [54].

From 1997 to 2015, a total of 907 human HPAI H5N1 infections were documented in Asian and African countries, with around 70% hospitalization and 50% case-fatality rates [50]. The fact that most of the cases (∼80%) have had exposure to poultry, as well as the shutdown of poultry markets in Hong Kong in 1997 effectively stopping the outbreak of human HPAI H5N1 infection, implied this avian IAV lineage did not acquire the ability to sustain inter-human transmission [35, 50].

Another lineage of HPAI, the H7N9 subtype, was first detected as LPAI in chickens in China, 2013 [55]. In subsequent outbreaks in 2016 and afterwards, the increased pathogenicity of H7N9 subtype infections was determined to result from multiple mutations in HA and reassortments with IAVs of duck lineage [56]. From 2013 to 2017, 1,220 human HPAI H7N9 cases have been reported in China, with around 60% hospitalization and 40% case-fatality rates [57]. Similar to the HPAI H5N1, this lineage likely does not sustain inter-human transmission, as more than 70% of cases had poultry exposure [57, 58] and the closure of live poultry markets effectively controlled the outbreak [59].

Human IAVs

Up to the present, human lineages of IAVs are of H1, H2, and H3 subtypes, each of which has caused at least one pandemic since 1900 (Table 1.1) [4].

It is noteworthy that although the 'Russian flu' outbreak in 1977 was not considered a true pandemic, it led to the re-emergence of subtype H1N1, which co-circulated with the H3N2 subtype until 2009 [4]. Currently, the viruses responsible for seasonal influenza belong to IAV subtypes

Table 1.1: The four pandemics of IAV since 1900.

Year	Common name	Subtype	Mortality (global)	Mortality (U.S.)	End of circulation
1918	Spanish flu	H1N1	50-100 million	500,000	1957
1957	Asian flu	H2N2	2 million	70,000	1968
1968	Hong Kong flu	H3N2	1 million	30,000	currently circulating
2009	Swine flu	H1N1	280,000	12,000	currently circulating

H3N2 and H1N1, plus influenza type B [60, 61, 62]. The term *seasonal* refers to the regular pattern of outbreaks observed in the temperate zones, where the incidence usually peaks during cold and dry seasons, corresponding to November to next-year April in the Northern Hemisphere and May to September in the Southern Hemisphere [63]. Cases reported in warm seasons are mostly imported or sporadic. Seasonality is less apparent in the tropics; instead, transmission and infection of influenza viruses are considered year-round [64].

Typical symptoms of influenza infection include fever, cough, sore throat, runny or stuffy nose, muscle or body aches, headaches, and/or fatigue [62, 65]. A more severe case may develop pneumonia, either caused by the virus or a secondary bacterial infection, and, in some cases, organ failure, sepsis, and death [66]. A recent study integrated data from 47 countries and estimated that around 290,000 – 650,000 influenza-associated respiratory deaths occur annually worldwide, equal to 4.0 – 8.8 deaths per 100,000 individuals, with an upward trend in mortality rate with age ($<$ 65 years: 1.0 – 5.1; 65-74 years: 13.3 – 27.8; $\geqslant$ 75 years: 51.3 – 99.4) [67]. In the U.S., annual hospitalizations (a surrogate marker for disease severity) and deaths are estimated between 140,000 – 810,000 and 12,000 – 61,000, respectively, since 2010 [68].

The frequent contact between human and birds or mammals serving as 'mixing vessels' potentiates *spillover*[†] of IAV. In fact, IAV has long been considered a threat to public health because of several reasons:

1. The high mutation rate offers the virus opportunity to change its genetic compositions, which may increase pathogenicity and/or switch host specificity.

2. Reassortment may grant zoonotic IAVs the ability to sustain inter-human transmission as well as novel antigenicity.

[†] *Spillover* refers to the process through which a pathogen that originated from wild animals passes to humans, as exemplified by HIV, Ebola virus, SARS-CoV, MERS-CoV, and, the most recent SARS-CoV-2 that causes coronavirus disease 2019 (COVID-19) [69].

3. IAVs can be transmitted through air (including aerosol and droplet) and direct contact [70, 71, 72].

1.1.3 Treatment

Currently, there are two types of FDA-approved antiviral drugs for treating IAV infections: ion channel blockers (adamantanes), and neuraminidase inhibitors (NAIs, e.g., oseltamivir, zanamivir) [35]. As mentioned, decrease of pH inside the virion is an essential step for viral entry. Adamantanes function by blocking the ion channel (M2) and hence prevent proper release of viral genome from endosomes. This type of drug is not currently recommended by CDC, due to widespread resistance and its ineffectiveness in treating influenza B virus infection. Distinct from adamantanes, NAIs act by inhibiting the catalytic function of neuraminidase, which is required for release of virions from an infected cell. Treatment with NAIs has been proved to significantly reduce the symptoms and infection course when being given within 48 hours after symptom onset [73, 74, 75]. Although still recommended by CDC to treat severe IAV infection, NAI-resistant strains have been emerging since the 2007-2008 flu season in Europe [76, 77] and 2008-2009 in Japan [78].

The critical mutation conferring adamantane-resistance is the M2-S31N, which originated in 2004 and quickly replaced the wild-type M2 within two years. One of the mutations conferring oseltamivir resistance, NA-H274Y, emerged in 2007 and showed a similar pattern of rapid invasion. These rapid emergence of drug-resistant IAV strains was quite surprising, since neither drug was prescribed extensively in most countries. Several phylogenetic studies on the adamantane-resistance suggested that its fixation was unlikely due to the selection pressure of the drug; instead, the M2-S31N mutation may hitchhike along with the invasion of IAV strains with drifted HA antigenicity [79, 80, 81].

1.2 Immune Responses to IAV infection

All living organisms must defend against the invasion and colonization of foreign organisms [82]. Evolutionarily-advanced animals harbor a group of cells specialized for the defensive reaction [83]. These cells, collectively named leukocytes (*leuko-*: white; *-cyte*: cell), are further categorized into granulocytes, monocytes, and lymphocytes, based on their appearance after Wright's staining.

1.2.1 Overview of the vertebrate immune system

The vertebrate immune system has two arms: the innate immune response, and the adaptive immune response [84]. Broadly speaking, innate immunity responds to evolutionarily conserved molecules expressed by pathogens (also known as pathogen-associated molecular patterns, or PAMPs), and activation of the innate immune response begins shortly after pathogen invasion. These responses include but are not limited to:

- Detection of PAMPs by pattern recognition receptors (PRRs) and activation of downstream signaling pathways, for example, the toll-like receptors (TLRs) (reviewed in [85, 86]), the nucleotide-binding oligomerization domain-like receptors (NLRs) (reviewed in [87, 88]), and the product of retinoic acid-inducible gene I (RIG-I) [89, 90].
- Production of type I interferons (IFNs) that function to alert and transition the neighboring cells to virus-resistant state [91, 92, 93].
- Expression of pro-inflammatory and chemotactic cytokines, which induce local inflammatory reactions and recruit leukocytes from peripheral blood [94, 95, 96].
- Migration of leukocytes, mainly granulocytes and monocytes, into the infection site where they perform many different immune functions such as cytokine production, phagocytosis, antigen presentation, and tissue breakdown and repair [97].

In contrast, the T and B lymphocytes, the primary cellular components of the adaptive immune system, recognize unique molecular patterns specific to distinct pathogens, such as short peptides or parts of the three-dimensional conformation of a protein. These unique patterns are referred as *epitopes* and the molecules from which the epitopes originate as *antigens*. The adaptive immune response is featured by its effectiveness of detecting and clearing invasive pathogens; however, it generally takes 1 to 2 weeks to fully develop, much longer than the innate immune response (see next section).

1.2.2 Adaptive immunity

The two arms of adaptive immunity, cellular immunity mediated by T cells and humoral immunity mediated by B cells, play distinct but complementary roles in an immune response. Briefly, T cells are responsible for detecting and killing intracellular pathogens, such as viruses and certain species of bacteria (e.g., tuberculosis), as well as assisting the maturation of B cells and recruitment of innate immune cells. B cells, on the other hand, are specialized in producing antibody that binds to extracellular pathogens and foreign antigens (e.g., toxin) and promotes phagocytosis as well as formation of membrane-attack complex mediated by complement.

The kinetics of adaptive immune response was first described in the early 1970s, where the researchers monitored antibody and cellular cytotoxicity of rats following the inoculation of mouse- or rat-derived leukemia cells [98, 99, 100]. Out of 100 rats tested, 65 had tumor regression within 2 weeks. The cellular cytotoxicity arose from days 3-5, peaked by days 9-10, and declined to low level on day 14 post-inoculation [99, 100]. In contrast, the antibody titer peaked by day 10, declined slightly and peaked again on day 35, followed by a gradual decrease [98]. Additionally, a secondary inoculation of leukemia cells on either day 14 or 35 could raise the antibody titer but not the cellular cytotoxicity.

This boost of antibody titer by a secondary immunization is one of the most attractive, and potentially useful, features of adaptive immunity – the *immunological memory*, referring to the larger and quicker responses when the host encounters the same antigen again [101]. The next two sections will briefly summarize the basic biology of B cells and T cells.

B cells and antibody

In the 1960s, B cells were first described as the cells in birds' bursa of Fabricius that are associated with antibody production (hence the name 'B' cell) [102, 103]. In mammals, these cells are originated from and develop in the fetal liver [104] or bone marrow [105, 106]. Once fully developed, these naïve B cells – cells that have not yet encountered their cognate antigen – leave the bone marrow and mostly enter secondary lymphoid organs (SLOs), i.e., spleen and lymph nodes. In the SLOs, they may either become activated and differentiate into plasma or memory B cells, or they die due to not encountering their cognate antigen.

Antibody, also known as immunologlobulin (Ig), is a soluble protein produced by plasma cells, with the shape of a 'Y' and mass of 150 kDa as a monomer [107, 108]. It consists of two light chains and two heavy chains, covalently linked by disulfide bonds [108, 109, 110]. The two antigen biding sites are located at the tip of the branches, each of which is composed of the variable domains of one light chain and one heavy chain. In addition to the variable antigen-binding regions of the tips, there are constant domains, which make up the F_c region [108].

Based on the F_c region, antibodies are categorized into five classes: IgD, IgM, IgG, IgA, and IgE, each of which has its own unique features and immunological functions.

- IgD is mainly expressed on the surface of naïve B cells, serving as the B-cell receptor for activating signal transduction [111, 112, 113].

- IgM is expressed in two forms. The membrane-bound IgM monomer functions similarly to

IgD [112, 114], while the secretory IgM exists as a pentamer that bind to antigen and serves as a dock for the assembly of complement [115].

- IgG is the main type of Ig in the serum, secreted by plasma cells that have experienced affinity maturation and class switching [116]. IgG can bind to its cognate epitope with high affinity and specificity. If the epitope resides near the cell receptor biding site, IgG may prevent the antigen from interacting with host cells, a process called neutralization [117]. In addition, the Fc region of IgG can bind to Fc receptors expressed by various immune cells and trigger their effector functions, e.g., phagocytosis by macrophages, antibody-dependent cellular cytotoxicity (ADCC) by natural killer cells [118].

- IgA is also produced in two forms, serum IgA and secretory IgA [119]. The functions of serum IgA are similar to those of IgG [120]. Secretory IgA (sIgA), on the other hand, has multifaceted functions [121]. It is present at mucosal membranes as well as in external secretions (e.g., tears, mucus, milk) and exists as monomer, dimer, or tetramer, involving in both antimicrobial defense [122] and immune regulation [121].

- IgE is the least common Ig subtype in serum. It usually binds to the $F_c\epsilon R$ of mast cells. The presence of its cognate antigen aggregates neighboring IgE/$F_c\epsilon R$ complexes on the mast cell and triggers the secretion of histamine, a central mediator of allergic reactions [123, 124].

T cells

T cells were discovered in the late 1960s from a series of studies aiming to elucidate the role of thymus in mouse leukemia [125, 126]. Both B and T cells originate from the bone marrow as common lymphoid progenitor (CLP), which has the potential to become either lineage [127]. Migration of CLP to the thymus through circulation and the subsequent signals, such as Notch

[128], commit the progenitors to the T-cell lineage [129, 130]. Fully-developed naïve T cells migrate to the periphery or SLOs where they may encounter their cognate antigen presented by dendritic cells (DCs). A critical difference in antigen recognition between T and B cells is, unlike a B cell being able to recognize either DC-bound or free antigen, a T cell must recognize its cognate antigen bound by a self-originated major histocompatibility complex (MHC) molecule via its T-cell receptor (TCR). This phenomenon, known as *MHC restriction*, was first described by Zinkernagel and Doherty in 1974 [131], where they showed the presence of cytotoxic activity of mouse spleen cells required at least one homologous MHC allele. In the subsequent paper [132], they further suggested the recognition of self and foreign antigen are likely executed by one receptor, termed the *altered self* model [133]. MHCs, TCRs and T cell epitopes will be discussed in sections 1.2.3 and 1.2.4.

Upon recognizing the cognate antigen presented by DC, T cells proliferate and differentiate into several subsets, mainly the effector and memory cells. The selective proliferation of lymphocytes with certain antigen specificity was formally formulated by Sir Macfarlane Burnet in the context of antibody-secreting cells in 1959 [134]. One of the key experiments in support of this idea was from Nossal and Lederberg [135], who isolated cells from the lymph nodes of rats immunized with *Salmonellae adelaide* and *S. typhi* and demonstrated that the antibody produced by a single cell can immobilize at most one of the bacteria, suggesting each cell only forms one species of antibody. Evidences were accumulated in the next two decades and eventually led to the establishment of *clonal-selection* theory for the B cells [136, 137]. On the contrary, clonal selection among T cells *in vivo* was not fully accepted [138, 139] until 1986, when Denizot et al. [140] demonstrated a clear proliferation of cytotoxic Ly-2^+ (CD8α) cells in mice 10 days after injection of allogenic tumor cells, and argued that the non-specific Ly-2^- cells might explain the discrepancy in previous studies.[†]

[†]It is noteworthy that, from my perspective, this selective proliferation process in periphery is called *clonal expansion* in contemporary immunological language, while *clonal selection* is now referred to the positive and negative

According to which TCR genes are expressed, T cells are categorized into $\alpha\beta$T cells and $\gamma\delta$T cells [141]. The $\alpha\beta$T cells contribute to most of the diversity within the cellular immune response. Based on the co-receptor, they are further classified into CD8 T cells and CD4 T cells. Upon TCR activation, naïve CD8 T cells undergo clonal expansion and differentiate into cytotoxic T lymphocytes (CTLs), which are capable of killing cells displaying cognate epitopes, or memory CD8 T cells, which retain proliferative potential or effector functions [142, 143]. Similarly, naïve CD4 T cells differentiate into distinct types of helper T cells (Th) upon activation, guided by the cytokines produced by APCs [144]. In contrast, the TCRs of $\gamma\delta$T cells show less diversity and hence are considered to execute more innate-like immune responses [145, 146, 147].

Many studies have shown the importance of CD8 T cell immunity in controlling IAV infection; therefore, we will concentrate our discussion on the effector functions of CTLs, and the development, compartmentalization, and dynamics of memory CD8 T cells in section 1.2.5.

1.2.3 Major histocompatibility complex (MHC)

In mammalian immune system, MHC plays a critical role in cellular immunity by preventing unwanted reactions to normal self-originated antigens and inducing immune responses to intracellular pathogens. While developing in the thymus, T cell precursors that (i) do not bind to self MHC or (ii) bind to self-originated epitopes with too high affinity are eliminated from the T cell pool. These processes are called positive selection and negative selection, respectively [148, 149, 150].

There are two types of MHCs, the MHC class I (MHC-I) and class II (MHC-II). The endogenous antigens, which are mostly peptides derived from cellular proteins, are mostly presented to patrolling CTLs by MHC-I [151, 152, 153]. The recognition of cognate antigen presented by MHC-I triggers CTLs to (1) release perforin and granzyme B and (2) express Fas ligand (FasL), both of

selections in thymus and bone marrow that select for lymphocytes reactive to foreign antigens.

which initiate apoptosis pathways in the antigen-bearing cell [154, 155, 156]. On the contrary, expression of MHC-II is mostly limited to professional APCs (i.e., DCs, macrophages, and B cells), which take up and degrade pathogens or soluble antigens from the environment and present these exogenous antigens to CD4 T cells [151, 157].

Features of HLA genes

Human MHCs are encoded by the genes residing on chromosome 6, named human leukocyte antigen (HLA). Each type of MHC has three loci: HLA-A, -B, and -C for MHC-I, and HLA-DP, -DQ, and -DR for MHC-II, referring to the property of *polygeny* [158, 159]. Since human cells are diploid, each locus has two alleles and both of the alleles are equally expressed (*codominance*). As a result, a human cell may express at most 6 different MHC-I and 6 different MHC-II molecules, allowing for sufficient coverage of potential non-self epitopes.

Before we discuss the most profound HLA property, *polymorphism*, it is worth introducing the nomenclature of HLA alleles [160] (hla.alleles.org). Each HLA allele is named as follows: (1) a capital letter followed by an asterisk, (2) four sets of numbers separated by colons, and (3) a suffix capital letter if necessary.

(HLA-) Gene * Field 1 : Field 2 : Field 3 : Field 4 Suffix .

- Gene refers to the locus where this allele resides.
- Field 1 denotes the allele group to which this allele belongs.
- Field 2 names this specific allele.
- Field 3 indicates synonymous DNA changes within the coding region.
- Field 4 indicates any change outside the coding region.

- Suffix is used to denote change in expression, if any.

For example, A*02:01 indicates this allele resides in the HLA-A locus and belongs to the A*02 allele group. A*02:01:02 differs from A*02:01:01 by synonymous DNA mutation(s) within the coding region, and A*02:01:02:02 differs from A*02:01:02:01 by DNA change(s) outside the coding region. See Table 1.2 for the suffixes used to denote changes in expression and their meaning.

Table 1.2: Suffix of MHC nomenclature

Suffix	Origin	Meaning
N	Null	Allele is not expressed.
L	Low	Allele is expressed at lower level than normal.
S	Secreted	Allele is expressed in a soluble, secreted form.
C	Cytoplasmic	Allele is expressed in the cytoplasm instead of on the cell surface.
A	Aberrant	Allele whose expression is not yet determined.
Q	Questionable	The mutation in the allele has been shown to affect the expression levels in other alleles.

Polymorphism describes the existence of multiple alleles of a gene. HLA is the most polymorphic human gene. As of September 20, 2019, more than 3,000 productive alleles have been reported for each of the MHC-I loci, and 1,000-2,000 for each of the MHC-II loci (Table 1.3) [161].

Murine MHC genes reside on chromosome 17 and are named as histocompatibility 2 (H-2)

Table 1.3: Number of alleles of each MHC locus (updated to September 20, 2019)

	HLA class I			HLA class II		
Gene	A	B	C	DR*	DQ†	DP‡
Alleles	5,907	7,126	5,709	3,360	2,062	1,716
Proteins♯	3,720	4,604	3,470	2,359	1,303	1,071
Nulls▽	308	244	243	141	83	83

*DR includes DRA and DRB.
†DQ includes DQA1, DQA2, and DQB1.
‡DP includes DPA1, DPA2, DPB1, DPB2.
♯'Proteins' refers to the alleles that can be translated into functional MHCs.
▽'Nulls' refers to the alleles that are not translated into functional MHCs.

[162]. Depending on the mouse strain, there may be two to three loci for MHC-I (H-2D, H-2K, and some have H-2L) and one to two loci for MHC-II (H-2 I-A, and some have H-2 I-E). The MHC haplotypes, which are defined as the combinations of alleles on one chromosome, were arbitrarily defined in several common laboratory mouse strains; for example, the haplotypes of C57BL/6 and BALB/c strains are defined as H-2^b and H-2^d, respectively [163].

Structure of MHC

MHC-I is composed of an α chain and β_2 microglobulin (B2M). The α chain contains three domains, with the α_1 and α_2 domains forming the peptide-binding groove and the α_3 domain having a transmembrane region [164]. B2M non-covalently binds to the α_3 domain and helps stabilize the MHC-I structure [165]. A typical MHC-I molecule has 6 pockets, referred to as A through F, which non-covalently interact with the anchoring residues of the peptide [164, 166]. The positions and chemical properties of the anchors vary across MHC-I alleles, but the carboxy terminus is usually a hydrophobic amino acid, such as leucine, isoleucine, and valine [167, 168].

MHC-II is composed of an α chain and a β chain, both of which contain two domains [164]. The binding groove is formed by the α_1 and β_1 domains, while the α_2 and β_2 domains comprise the transmembrane region. The open-ended groove of MHC-II allows peptides to have a longer length, usually more than 13 amino acids [167]. Despite various peptide lengths, the section of the peptide embedded in the binding groove usually has amino acids of similar chemical properties at positions 1, 4, 6, and 9 [167].

The α_3 domain of MHC-I and α_2 domain of MHC-II also serve as docks for the binding of CD8 and CD4 co-receptors, respectively. Co-receptor binding prolongs the interaction between TCR and peptide-MHC, and is considered important for T cell activation [169].

1.2.4 T-cell receptor (TCR)

The TCR complex consists of: (1) two antigen-recognition proteins (the α and β chains of the TCR), and (2) a group of signaling proteins, collectively called CD3 [170, 171, 172]. The binding of TCR and peptide-MHC initiates conformational changes of the CD3 proteins, allowing signal transduction [173, 174, 175]. We will focus the discussion on TCR in this section.

As mentioned, the TCR is responsible for antigen recognition. To have sufficient breadth that covers as many epitopes as possible, the TCR repertoire is extremely diverse within an individual [176, 177, 178, 179]. There are three regions on TCR known to contact the epitope, named complementarity-determining regions (CDR) 1, 2, and 3, among which CDR3 is the most variable [180, 181]. The extremely high diversity is created by two mechanisms: (1) V(D)J recombination, and (2) P- and N-nucleotide addition.

V(D)J recombination was discovered in late 1970s and 1980s, first described by Tonegawa and colleagues in a series of studies aiming to resolve the gene sequence of murine antibody [182, 183, 184, 185]. In his landmark review published in 1983, Tonegawa summarized the organization of antibody gene sequences and proposed the potential mechanism for the V-J and V-D-J joining [186]. It is noteworthy that, he also mentioned the contribution of imprecise joining ends to antibody diversity, which was then named as P- and N-nucleotide addition [187, 188]. The fact that T cells employ the same mechanism to generate the TCR repertoire was proved one year later [189, 190, 191]. Let us first focus on TCR α chain. The human germline *TRA* gene has 70 to 80 variable (V_α) segments, 61 joining (J_α) segments, and one constant (C_α) segment [192]. During T-cell development, one V_α and one J_α are randomly picked and ligated [193, 194]. In addition, the enzyme complex adds nucleotides to the end of coding regions in palindromic (P-nucleotides) or non-template-encoded (N-nucleotides) manners, which further increases the junctional diversity. Production of the *TRB* gene follows the same rules as *TRA*. The *TRB* gene region has 52 V_β segments, two diverse (D)

segments, each of which is followed by an array of J_β (6 for $J_\beta 1$ and 7 for $J_\beta 2$) and C_β (1 for each) segments [195].

Based on the numbers of V, D, J segments, the heterodimer pairing of α:β chains, and P- and N-nucleotides, the total number of unique TCR sequences was estimated to be around 10^{18} [176]. However, most of these gene rearrangements may result in shift of reading frame or encode TCRs that cannot pass positive or negative selection; hence, the number of unique TCRs observed in the T-cell pool outside the thymus is much less. Qi et al. [178] estimated the diversity of the *TRB* gene to be 100 million (10^8), and this diversity decreases with age and T cell differentiation.

1.2.5 CD8 T cell responses to IAV infection

CD8 T cells detect IAV by recognizing viral epitopes presented by infected cells; thus, infection must occur in order to induce CD8 T cell responses. From this perspective, CD8 T cells, unlike antibodies, do not provide sterilizing immunity against IAV [196]. However, the high specificity of antibodies limits their potency for neutralizing a drifted or heterosubtypic virus, while many studies have shown influenza-specific memory CD8 T cells can broadly protect against pathology resulting from heterosubtypic IAV infection [196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210].

Viral kinetics during a primary IAV infection

Smith et al. [211] studied the kinetics of IAV replication in naïve mouse lungs (Figure 1.1). Using the 50% tissue culture infective dose ($TCID_{50}$) as the measure of viral titer, they described the kinetics of viral replication as having five phases:

1. Infection and undetectable virus (within 4 hours post-infection)
2. Exponential viral growth (4 hours to day 2 post-infection)
3. Peak of viral titer at around 10^6 $TCID_{50}$ ($\sim$day 2)

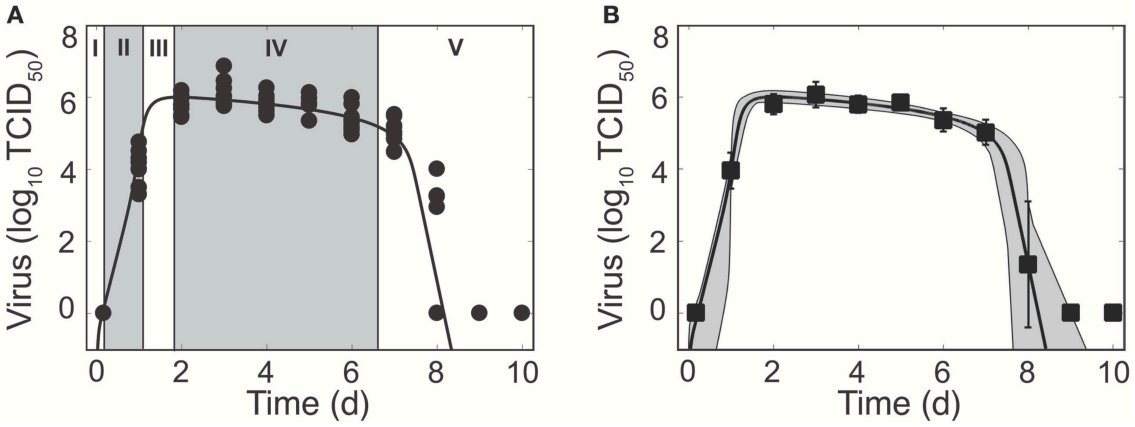

Figure 1.1: Kinetics of influenza A virus replication during primary infection of mouse (adopted from Smith et al. 2018). (A) Experimental data and curve of the best-fit model delineate the five phases of viral kinetics: I. Infection and undetectable; II. Exponential growth; III. Peak; IV. Slow decay; V. Rapid decay and clearance. (B) Curve of the best-fit model with 95% confidence interval.

4. Slow decay of virus (days 2 to 6)

5. Rapid decay and clearance of virus (days 7 to 9)

By fitting linear models to the slow and rapid decay phases, they estimated the viral titer decreased at rates of 0.2 and 3.8 $\log_{10}$-TCID_{50} per day. If the switch from slow to rapid decay reflected the induction of influenza-specific T cell responses, then these estimates suggested that cellular immunity is 25 times more effective than the innate response in promoting viral clearance. However, this model may not fully distinguish the contributions of resource limitations, innate immunity, and T cell responses, in viral clearance.

CD8 T cell responses in IAV infection

The CD8 T cell response against IAV infection begins as tissue-resident DCs migrate from the infected upper respiratory tract and lung, enter draining lymph nodes, and present viral epitopes to naïve CD8 T cells [212, 213]. Following activation and differentiation, virus-specific CTLs sense

the signals of C-X-C motif chemokine ligand (CXCL) 12 via C-X-C motif chemokine receptor (CXCR) 4 and migrate to the infection site [214], where they clear the virus-infected cells via extrinsic (Fas-FasL [215] or TRAIL-DR5 [216]) and/or intrinsic (release of granzymes and perforin [215, 217]) apoptotic pathways. It is believed that the presence of CTLs results in the rapid decay and clearance of virus after day 7 post-infection [213, 218, 219].

A fraction of activated CD8 T cells survive contraction of the effector T cell pool after clearance of IAV and differentiate into long-lived memory CD8 T cells [220, 221]. Based on their circulatory patterns, memory CD8 T cells are categorized into central (T_{CM}), effector (T_{EM}), and tissue-resident (T_{RM}) [222, 223, 224]. T_{CM} circulate between the SLOs and peripheral blood, featured by its $CCR7^{+}CD62^{+}$ phenotype and proliferative potential, whereas T_{EM} circulate between blood and non-lymphoid tissues (NLTs), express no CCR7 and heterogeneous amount of CD62L, and are epigenetically poised for immediate effector functions (i.e., cytotoxicity, cytokine production) instead of proliferative ability [225, 226, 227, 228, 229]. The T_{RM} compartment were first described in Masopust et al. [230] as long-lived memory cells in the NLTs of mice (e.g., lungs, liver, kidney) exhibiting effector levels of lytic activity *ex vivo*. These cells, as characterized by later studies in mice and humans [231, 232, 233, 234, 235, 236], do not enter circulation but dwell in and patrol the tissue and are specialized for the rapid release of cytolytic granules (e.g., perforin and granzymes) and secretion of cytokines that recruit and activate other immune cells from the circulation (reviewed in [222, 224, 237, 238]).

Studies regarding the development and dynamics of CD8 T_{RM} were mostly done in animal models, and it is now believed that CD8 T_{RM} residing in different NLTs need slightly different signals to develop [237], although the presence of antigen [239, 240], expression of surface markers associated with tissue retention [235] and downregulation of markers for tissue egression [241], as well as environments with interleukin (IL)-15 [242] and tissue growth factor (TGF)-β [243, 242]

are necessary in general. In the mouse skin, contact with conventional DCs is required [242], while in the lungs, monocyte-derived APCs are indispensable [244]. Furthermore, distinct from the skin, where CD8 T_{RM} maintain a stable population, CD8 T_{RM} in the lungs slowly decay after the resolution of infection [245, 246]. The mechanism behind this limited longevity in the lung remains unclear. Slütter et al. [245] had proposed that the decay of T_{RM} simply reflects the dynamics of CD8 T_{EM}; however, several lines of evidence, including parabiosis experiments, show that lung CD8 T_{RM} are not replenished by circulating memory CD8 T cells. Hayward et al. also demonstrated that the decay of IAV-specific CD8 T_{RM} is accelerated by an unrelated respiratory infection, e.g., Sendai virus infection (unpublished data). Lack of persistent antigen stimulation might be an additional mechanism leading to T_{RM} decay, since Uddbäck et al. [247] has shown the immunization with AdNP, which is a replication-deficient adenovirus 5 vector expressing the NP of A/Puerto Rico/8/34 (H1N1), results in persistent expression of NP and greatly improves the longevity of NP-specific lung CD8 T_{RM}.

CD8 T cell protection in heterosubtypic IAV infection

Although CD8 T cells do not provide sterilizing immunity, which completely prevents infection of the host, several studies have shown they protect the host from severe pathology and, to a lesser extent, limit virus transmission [218].

Mouse

An early study from 1994 compared the number of CD8 T cells in the bronchoalveolar lavage (BAL) to viral loads after a heterosubtypic IAV challenge[†] [197]. They showed the depletion of CD8 T cells in the upper or lower respiratory tracts partially reduced protection against heterosubtypic infection. Also, as the time from immunization increased, the efficacy of heterosubtypic protection

[†] *Heterosubtypic challenge* means re-infecting the subjects with an influenza virus whose subtype differs from the virus used for immunization; for example, immunizing with HKx31 (H3N2) and challenging with PR8 (H1N1). The advantage of this experimental system is to avoid the interference of antibody induced by the first infection.

declined. Two later studies recapitulated these findings and further established that the protection against pathology is mediated by CD8 T cells in the mucosa [198, 199]. Recent studies showed the CD8 T cell-mediated protection can be extended to highly pathogenic avian influenza, as immunization of mice with IAV of H3N2 or H1N1 subtypes can protect against lethal infection with H5N1 [204] or H7N9 [208].

Along these lines, Kreijtz et al. [204] found that memory CD8 T cells specific for the $NP_{366\text{-}374}$ (ASNENMDAM) and $PA_{224\text{-}233}$ (SCLENFRAYV) epitopes can respond to the slightly different epitopes harbored by the H5N1 strain (ASNENMEVM and SSLENFRAYV), a feature termed *cross-protection*. In addition, McMaster et al. [208] compared the weight loss and survival of mice primed with PR8 (H1N1) through intranasal (i.n.) or intramuscular (i.m.) routes. Given that the two routes induce similar numbers of systemic memory CD8 T cells but only the i.n. route induces lung T_{RM}, the finding that weight loss and mortality was similar between naïve and i.m.-primed mice highlighted the importance of lung T_{RM} in reducing severe pathology.

Interestingly, the epitope specificity of IAV-specific CD8 T cells may influence their protective potential, and may in some scenarios have a detrimental effect upon IAV challenge. Crowe et al. [248] showed that, even though PA_{244}-specific CD8 T cells are present at roughly the same number of NP_{366}-specific CD8 T cells following a primary IAV infection, PA_{244}-specific CD8 T cells may delay viral clearance upon heterosubtypic challenge, likely due to poor expression of the $PA_{244\text{-}233}$ epitope by airway epithelial cells.

Human

The importance of CD8 T cell immunity in IAV infection was first shown in a human challenge study [218]. This study showed that individuals who had CD8 T cells prior to IAV infection had lower viral loads in the nasal wash, a surrogate marker for transmission. However, this initial study did not find any correlation between the number of CD8 T cells and symptom scores. In contrast, a 2013

report from Sridhar et al. [207] analyzed data from a cohort study conducted during the outbreak of pandemic H1N1 and found the number of memory CD8 T cells in peripheral blood negatively correlated with the patients' symptom scores. The existence of cross-reactive IAV-specific CD8 T cells in humans are supported by two additional studies [208, 249], where a fraction of CD8 T cells isolated from subjects in North America can react to H7N9, which only sporadically spread in China and unlikely infected the subjects prior to the study. Recently, Koutsakos et al. showed the CD8 T cells can cross-react to type B and C influenza viruses, suggesting the protection might be even broader than what was expected previously [250].

CD8 T cell epitopes of IAV

Mouse

Most of the CD8 T cell epitopes of IAV are defined in the H-2^b background. The first epitope, D^b-restricted $NP_{366\text{-}374}$, was found by Townsend et al. in 1986 [251]. In late 1990s and early 2000s, D^b-$PA_{224\text{-}233}$ [252], K^b-$PB1_{703\text{-}711}$ [253], and a group of subdominant epitopes were discovered [254, 255]. Thomas et al. summarized the major and minor epitopes of PR8 discovered from mouse in a review in 2006 [256], which is recapitulated in Table 1.4. Recently, Wu et al. [257] used mass spectrum to discover additional epitopes in the H-2^b background.

Table 1.4: List of H-2^b-restricted epitopes of PR8

Epitope	Sequence	MHC-I restriction
$NP_{366\text{-}374}$	ASNENMETM	D^b
$PA_{224\text{-}233}$	SSLENFRAYV	D^b
$PB1_{703\text{-}711}$	SSYRRPVGI	K^b
$NS2_{114\text{-}121}$	RTFSFQLI	K^b
PB1-$F2_{72\text{-}80}$	LSLRNPILV	D^b
$M1_{128\text{-}135}$	MGLIYNRM	K^b

Human

As of Mar 27, 2020, 446 human CD8 T cell IAV epitopes have been experimentally verified and cata-

loged by the Immune Epitope Database (Table 1.5). A significant proportion of these epitopes were contributed by a few studies using broad, systematic approaches. Assarsson et al. [258] screened 4,080 peptides against peripheral blood mononuclear cells (PBMCs) from 44 human donors, across 6 MHC-I supertypes. They identified 38 MHC-I-restricted epitopes, most of which were derived from PB1 and M1. However, two later studies [259, 260], respectively focusing on the HLA-A2$^+$ and -A2$^-$ subjects, showed the majority of the IAV-specific CD8 T cell response was directed against NP and, to a lesser extent, M1.

Besides those immunogenic epitopes, Rimmelzwaan et al. [261] reviewed all the CD8 T cell-escaping mutations in IAV discovered before 2009 and separated them into two escaping mechanisms. The first type of mutation changes the MHC anchor residue thereby abrogating peptide-MHC binding; as a result, the epitope will not be presented at all. The second type of mutation changes the conformation of the TCR-contacting interface, and thus memory CD8 T cells cannot recognize the mutated epitope when presented by MHC-I.

Table 1.5: Numbers of experimentally verified human CD8 T cell epitopes in IAV (updated to March 27, 2020)

	PB1*	PB2	PA	HA†	NP	NA	M1‡	M2	NS1	NS2
Total	43	12	18	18	226	7	80	21	18	3

*Including epitopes derived from RdRP catalytic subunit
†Including epitopes derived from HA precursor
‡Including epitopes derived from matrix protein

1.3 Evolution of IAV

Influenza A virus is one of the fastest evolving pathogens. In this section, I will review studies on the genetic and antigenic evolution of IAVs.

1.3.1 Antigenic drift and shift

There are two mechanisms that result in antigenic change, antigenic drift and antigenic shift [1]. Most studies investigating these mechanisms focus on changes in hemagglutinin, which is the primary target of antibodies.

Antigenic drift

Antigenic drift results from one or a few nonsynonymous mutations, usually in the HA gene, that change the chemical properties and/or the local conformation. Although HA-specific antibodies have been postulated to be the driving force of antigenic drift, it was not initially clear how antibodies impose immune pressure on viruses. Hensley et al. [262] proposed that these substitutions mainly increase receptor binding affinity, enabling the HA-SA binding to outcompete the HA-Ab binding, and, coincidently, this often resulted in changes to antigenicity. This hypothesis is referred to as *generalized neutralization resistance*. However, Lee et al. [263] used deep sequencing to examine the quasispecies of IAV isolated from patients, and found most of the substitutions were located at the antigenic site, supporting that the mutations per se changed the antigenicity without necessarily changing the receptor binding affinity.

Smith et al. [264] used a multidimensional scaling technique to project the HI matrix, which consists of the hemagglutination-inhibition titers of a series of anti-sera against a panel of viruses, to two-dimensional plan. The distance between any two points on this map is the antigenic distance of two viruses. They found that antigenicity clustered into distinct patterns; moreover, after naming each cluster by its first isolate, it appears that 'cluster transition' happened every 2 to 8 years for H3N2 (Table 1.6). A profound implication of this mapping is that, rather than a continuous evolution as observed for the HA gene, antigenicity is evolving in a punctuated pattern [265].

This study created two important questions regarding the antigenic evolution of IAV. First,

Table 1.6: Antigenic clusters of human IAV H3N2 subtype

Antigenic clusters	Representative strain
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
FU05	Wisconsin/2005
BR07	Brisbane/2007
PE09	Perth/2009

certain mutations are associated with each antigenic cluster, but can the antigenic cluster be defined by these mutations? Second, how and why is the evolutionary pattern of antigenicity ('punctuated') distinct from that of nucleotides ('continuous'), given that changes in antigenicity resulted from mutations at nucleotide level?

The answer to the first question is 'yes'. Koel et al. [266] studied these mutations and found they essentially determined the antigenicity of HA. The second question is not yet fully answered. Koelle et al. [265] analyzed the antigenic clusters using phylogenetic information between neighboring clusters, and found some of the neighboring clusters had indeed originated from the same ancestral cluster. For example, the BE89 cluster gave rise to both BE92 and WU95 clusters, although the BE92 cluster died out after the WU95 cluster emerged. A more comprehensive study by Bedford et al. [267] integrated antigenic data into the phylogeny of IAV H3N2, H1N1, and two influenza B virus strains (Victoria and Yamagata), and found that H3N2 was evolving the fastest among the four.

Antigenic shift

Antigenic shift happens when an IAV acquires a novel gene segment encoding a surface protein with distinct antigenicity through reassortment. This primarily happens in IAV due to its segmented genome and the extensive animal reservoir [268]. Indeed, all documented IAV pandemics were considered to result from antigenic shift. For example, the pandemic of Asian flu (H2N2) in 1957 is thought to acquire HA, NA, and PB1 under the 1918 H1N1 background, and the pandemic of Hong Kong flu (H3N2) in 1968 may have originated from its predecessor, H2N2, by acquiring novel HA and PB1 gene segments from an avian IAV [269]. The well-studied 2009 pandemic H1N1 is also a reassortant, consisting of gene segments from four sources [270, 271]:

- PB2, PA: North American avian IAV
- PB1: Human H3N2 IAV
- HA, NP, NS: Classical American swine IAV
- NA, M: Eurasian 'avian-like' swine IAV

As mentioned in section 1.1.1, antigenic shift may give rise to a highly pathogenic virus of novel antigenicity, with almost no pre-existing herd immunity in the human population. This is an important public health concern.

Population dynamics of IAV in human

Phylogenetic information helped construct the population dynamics of IAV in human. As mentioned in section 1.1.2 Human IAVs, incidences of IAV infection show a clear seasonality [272, 273]. A critical observation is that IAV strains which caused epidemics in the northern temperate region usually were also observed during the next flu season in the southern temperate region [274]. This

led to the hypothesis that IAVs are transmitted from the northern to the southern hemispheres, while during the inter-epidemic seasons, the virus pool undergoes strong bottleneck, which shapes the genetic diversity and dynamics of IAV strains. However, two studies published in 2008 provided genetic evidence in support of an alternative explanation, the source-sink model [275, 276]. In this theory, the tropical regions, mainly east and Southeast Asia where no apparent seasonality of IAV infection is observed, serves as the source of genetically variable viruses due to continuous and intensive transmission. Some of the viruses spread to the subtropical and temperate regions, trigger epidemics under suitable climate conditions, and mostly die out there after the flu season (sink).

1.3.2 Evolution of CD8 T cell epitopes in IAV

Given the drawbacks of an antibody-based immunization strategy, novel vaccine formulations in development are attempting to incorporate antigens that may induce broadly-reactive and long-lasting protection [277, 278, 279, 280]. The conserved regions of HA and/or NA and T cell epitopes have been proposed as candidates for a 'universal influenza vaccine'. Indeed, it has been shown feasible to induce broadly neutralizing antibodies by engineered HA stem antigen [281]; however, these antibodies may select for escaping mutations on the HA stem *in vitro* [282]. By contrast, one study compared the substitution rates across the gene segments of IAVs currently circulating in humans (H3N2 and H1N1) and found the surface portions of HA and NA have much higher substitution rates than the other gene segments [283]. In addition, among all known CD8 T cell epitopes of NP in humans, only 6 escaping mutations have been found [261]. Provided CD8 T cells do protect against IAV infection, the overall conservation of CD8 T cell epitopes compared to antibody epitopes raises an important question: Why are CD8 T cell epitopes of IAV are so conserved?

Before discussing different hypotheses for the conservation of CD8 T cell epitopes, and the implications of these hypotheses, I would like to briefly review the three models of molecular evolution: Neutral evolution, negative selection, and positive selection.

Three models of molecular evolution

Neutral evolution describes the dynamics of mutations which do not bring any fitness change to the carrier. The dynamics of neutral evolution depend entirely on stochastic processes to determine whether these mutations are passed on to the next generation. When the mutations alter the fitness of the carrier, beneficial mutations are likely to be maintained through positive (Darwinian) selection, while detrimental mutations are likely to be removed via negative (purifying) selection [284].

Neutral evolution is best exemplified by synonymous mutations, which do not result in change of the amino acid sequence of the protein product. Indeed, a conventional statistic for testing which selection process a gene is undergoing is the ratio of the rate of substitutions at nonsynonymous sites (dN) to the rate of substitutions at synonymous sites (dS). A dN/dS ratio close to 1 indicates neutral evolution, while a ratio greater or less than 1 indicates positive or negative selection, respectively [285, 286].

Do CD8 T cell epitopes in IAV undergo selection?

Early studies [287, 288] showed that alanine replacements at anchor residues of immunodominant CD8 T cell epitopes (e.g., position 2 of $M1_{58\text{-}66}$, positions 2 and 9 of $NP_{418\text{-}426}$) had a strong detrimental effect on the virus, while the mutations at other residues, although also detrimental, were tolerated to various extents. From these data they concluded that functional constraints may limit variation in certain epitopes, especially at the anchor residues.

There are two implications from this study. First, it implied that negative selection, which purges mutations from the virus pool, maintains the overall conservation of CD8 T cell epitopes. Second, as summarized in a later review article [261], mutations that are not at anchor residues can result in escape, thereby possibly invading the virus pool if the selective advantage is large enough to overcome the fitness defect.

A limitation of the Berkhoff et al. studies was the lack of consideration given to compensating mutations that could rescue potential defects in viral fitness. Gong et al. [289] studied this question by sequentially introducing mutations into the NP of A/Aichi/2/1968 (H3N2) strain and measuring the ability of variants to replicate *in vitro*. Interestingly, although a single escaping mutation (L259S, R384G, or V280A) largely impaired the viral growth, viability was rescued to wild-type levels when co-mutations were introduced. In other words, escaping mutations are not necessarily accompanied with fitness defects in natural IAV evolution. Later, another study from the same group compared the phylogenies of NP and M genes between human- and swine-lineage IAVs [290]. The validity of this comparison relies on two presumptions:

1. The selection pressure from swine herd immunity is believed to be very low due to their short lifespan; therefore, the evolution of swine-lineage viruses is neutral.

2. The epitopes targeted by human cellular immunity differ from those targeted by swine; thus, the 'epitope region', which is defined as the set of nucleotide sites included in at least one of the known *human* CD8 T cell epitopes, is not under immune pressure mediated by swine CD8 T cells.

Indeed, Machkovech et al. observed that the epitope and non-epitope regions of swine-lineage IAVs had similar proportions of nucleotide sites with $dN/dS > 1$, supporting the neutrality assumption. It was not surprising that human-lineage IAVs generally had a higher proportion of sites with dN/dS

> 1 when compared to the swine-lineage viruses. The surprising finding was, among human-lineage IAVs, the epitope regions had *lower* proportions of nucleotide sites with dN/dS > 1 than the non-epitope regions. Although the implication was unclear since they did not report the average dN/dS ratios, it may suggest that the epitope regions are undergoing *less* positive selection than the non-epitope regions. To better separate the signal of positive selection from fitness defect, the authors designed a new statistic, F, denoting the average number of epitopes affected by one mutation. They found the average F of human-lineage IAVs was significantly higher than that of swine-lineage IAVs, suggesting more epitopes are changed by a single mutation in human IAVs. Accordingly, they concluded that CD8 T cell epitopes of human-lineage IAV are undergoing positive selection.

Changes in CD8 T cell antigenicity

To define the antigenic distance of CD8 T cell epitopes requires several careful calculations. Unlike antibodies, each person is likely to harbor a unique set of HLA genes. Thus, there is no a strong analogy of virus-antiserum pairs, nor a well-defined measure like HI titer, that can be universally applied across CD8 T cell epitopes. Woolthuis et al. [291] defined the antigenic distance of CD8 T cell epitopes by comparing the genome of each virus to the 'epitope compendium', which includes all experimentally verified CD8 T cell epitopes in human from the Immune Epitope Database. The antigenic distance between two viruses correlates inversely with the number of epitopes they have in common. Namely, the more epitopes two viruses have in common, the shorter the antigenic distance. If two viruses have exactly the same set of epitopes, the distance becomes 0.

Using this measure, they constructed an antigenic map for the subtypes of H1N1 (1932-1957, 1977-2015), H2N2 (1957-1968), H3N2 (1968-2015), and several zoonotic subtypes (H5N1, H7N9, H9N2). They found the antigenic evolution of CD8 T cell epitopes showed a continuous pattern, distinct from that of antibody epitopes. An interesting conclusion of their data from H3N2 viruses

was, even if IAVs may gain or lose epitopes in consecutive years, the average number of CD8 T cell epitopes per virus decreased over time, from ∼80 in 1968 to ∼65 in 2015. They postulated that the decrease of epitope number results from adaptation of IAVs to herd cellular immunity.

One caveat to this study is the compendium of *experimentally* verified CD8 T cell epitopes may misclassify a mutant epitope as 'lost' if the mutant has not yet been tested. Specifically, suppose E and E' are the wild-type and mutant forms of one epitope, and the mutation does not affect its antigenicity (i.e., E' remain as immunogenic as E), then ideally, both epitopes should belong to the compendium. However, if the antigenicity of E' has not been experimentally verified, it will not be included in the compendium based on the authors' inclusion criteria; as a result, all isolates harboring E' will be treated as 'loss of epitope'. This reporting bias may be more severe among the mutations recently acquired due to the lack of chance being tested.

Modeling the invasion of a CD8 T cell-escaping mutant

Despite the overall conservation of CD8 T cell epitopes, the six escaping mutants previously identified in human IAVs [261] replaced the wild type sequences strikingly rapidly – within 2 to 5 years. Gog et al. [292] proposed a model to explain the rapid invasion of one of these mutations, NP-R384G, which abrogates the binding of $NP_{380\text{-}388}$ to B*08:01 and $NP_{383\text{-}391}$ to B*27:05. The authors concluded that the stochastic persistence and bottleneck effect during the inter-epidemic trough (discussed above) offers the mutant epitope a chance to replace the wild-type epitope, given that escaping from CD8 T cells results in a longer period of infectiousness. Furthermore, a greater number of founders with the mutation (i.e., viruses that effectively initiate the next epidemic) will result in a higher probability and shorter duration of replacement. Nevertheless, as the source-sink model pointed out, the viruses that initiate epidemics in the temperate regions are likely imported from the tropics, where the inter-epidemic bottleneck is not as strong [276]. Thus, the rapid invasion

may stem from other evolutionary mechanisms.

1.4 Concluding Remarks

Here I want to summarize current knowledge regarding the evolution of CD8 T cell epitopes in IAV, and propose the research questions of this dissertation. It is clear that CD8 T cell epitopes in IAV are more conserved than antibody epitopes. However, escaping mutations in CD8 T cell epitopes still emerge, and they tend to replace the wild type epitopes rapidly. Given that people who have IAV-specific CD8 T cells have reduced transmission and/or are protected from severe disease during subsequent IAV infections, CD8 T cell immunity likely imposes selection pressure on the virus. This dissertation aims to address the reason why CD8 T cell epitopes of IAV are conserved, despite the selection pressure of CD8 T cell immunity.

Chapter 2

Modeling the invasion of a CD8 T cell-escaping IAV variant

2.1 Results

There are (at least) three factors that can affect the fitness of an escape-variant: (i) The fitness cost of the mutation; (ii) the extent of selection for the escape-variant in hosts carrying the *relevant* MHC-I allele(s) (i.e. one that presents the wild-type form of focal epitope); (iii) the frequency of hosts with the relevant MHC-I allele(s) in the population. We begin with a relatively simple population genetics model that allows us to examine how the interplay between these factors affects the rate of invasion of an escape-variant. This model assumes the selective advantage of the escape-variant in a host is fixed. We then consider why this assumption may not hold and examine the consequences of relaxing this assumption, using an epidemiological model for the spread of wild type and escape-variant.

2.1.1 Population genetics model

Let us consider the wild type (WT) and one escape-variant (MT) that carries a mutated focal epitope. Let h represents the set of MHC-I allele(s) that present the focal epitope of WT but not MT, and let f equal the frequency of h. The MHC-I alleles that present epitopes other than the

focal one, H, are at frequency $(1-f)$. With the usual assumptions for Hardy-Weinberg equilibrium, we have

	Host genotypes		
	HH	Hh	hh
Genotype frequency	$(1-f)^2$	$2f(1-f)$	f^2
Relative fitness of WT	1	1	1
Relative fitness of MT	$1-m$	$1-m+rs$	$1-m+s$

where m, s, r denote the fitness cost of MT in all hosts, the selective advantage of MT in hosts of hh genotype, and the dominance coefficient of the h allele, respectively. Here we have assumed that the WT is equally fit in all genotypes, i.e., while h alleles present the focal epitope, the H alleles present other epitopes carried by the virus, so that all alleles confer about the same level of resistance. This is not true for the MT: While the H alleles still successfully present its other epitopes, h does not present its focal epitope. The frequency of the MT in generation t, denoted by q_t, is given by

$$q_t = \frac{q_{t-1}K}{1-q_{t-1}+q_{t-1}K} \tag{2.1}$$

where $K = (1-f)^2(1-m)+2f(1-f)(1-m+rs)+f^2(1-m+s)$ is the mean fitness of MT in the host population. Equation 2.1 allows us to examine how the rate of invasion of the MT depends on: m, the fitness cost; f, the frequency of h; s, the selective advantage MT accrues in the hosts of genotype hh; and r, the dominance coefficient as the ratio of selective advantage of MT in Hh hosts to that in hh hosts. The dominance coefficient will depend on a number of factors, such as (1) the immunodominance of the focal epitope, and (2) the relationship between the magnitude of CD8 T cell responses and the rate of clearance of infected cells. In the absence of a quantitative understanding of these factors, we set r at an intermediate value ($r = 0.5$), and in section 2.3.1 we discuss the consequences of changing r to lower or higher values.

In Figure 2.1, we plot how the rate of invasion depends on the the selective advantage s of the MT (x-axis) and the frequency f of the MHC-I alleles in which MT has a selective advantage (y-

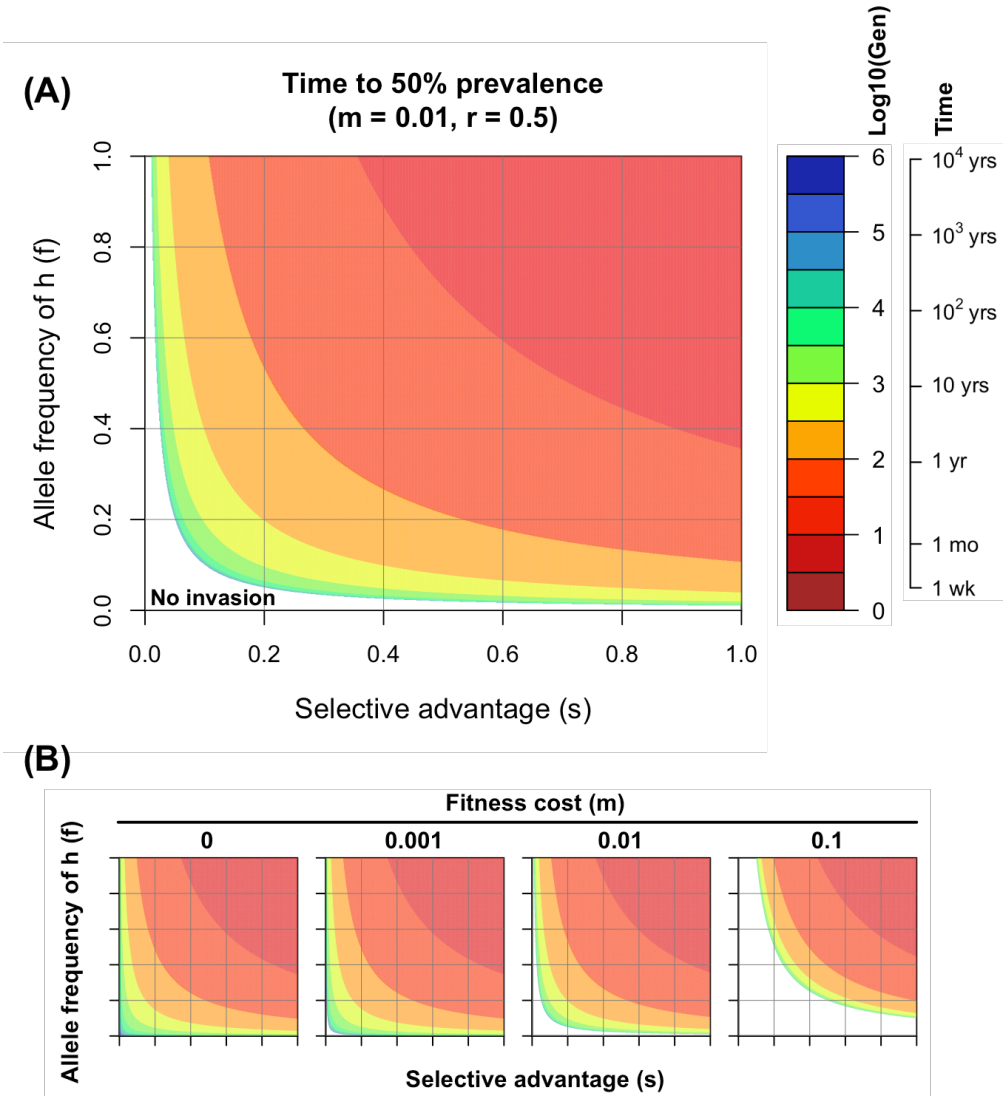

Figure 2.1: Rate of invasion of a CD8 T cell escape-variant. (A) Contour plots show the number of generations on a log scale required for the escape-variant to increase from 0.01 to 50% prevalence predicted by Equation 2.1. The approximate time for this to occur is calculated by assuming a serial interval of 3-4 days between infections (i.e. that there are about 100 generations per year). Fitness cost is set to 1% (i.e., $m = 0.01$). (B) We show the time for invasion under different fitness costs, which goes from 0 to 10%. We see that even when the mutant does not have a fitness cost ($m = 0$), it will invade relatively slowly if s and f are small due to the nature of CD8 T cell protection and extent of MHC polymorphism. Ticks on the axes in (B) indicate the same numbers as in (A).

axis). The rate of invasion is plotted on a log scale – it shows the $\log_{10}$ of the number of generations required for the MT to go from a prevalence of 0.01% to 50% in the host population. Given that the serial interval for influenza is about 3 to 4 days [293], 100-generation corresponds to about one year. In Panel A, we set the fitness cost to 1% (i.e. $m = 0.01$). There is a parameter regime (white region with low s and f) where the fitness cost is sufficient to prevent MT from invasion. When MT can invade, we see faster invasion when s or f increase.

Parameter m.

We show the effect of changing the fitness cost of MT (m) from 0 to 10% in Figure 2.1B. We see that even if the escape-variant does not have any fitness cost (i.e., $m = 0$), it invades relatively slowly when s and f are small (the region to the bottom left of the leftmost panel of Figure 2.1B). See section 2.3.1 for more details.

Parameter s.

Building on the earlier ideas proposed by Halloran et al. [294], immunity can provide protection by reducing susceptibility of immune individuals to infection (IE_S), as well as by reducing pathology (IE_P) and transmission (IE_I) in infected individuals. While the role of CD8 T cells in providing protection against influenza remains to be fully understood, a number of studies suggest that they play a significant role in reducing pathology (high IE_P). In humans, a greater number of IAV-specific CD8 T cells prior to heterosubtypic virus infection is associated with faster viral clearance [218] and fewer symptoms [207]. In support of human studies, mouse experiments have shown the cellular immunity induced by H1N1 and/or H3N2 is able to protect the hosts from lethal infection with avian H5N1 or H7N9 (see section 1.2.5). CD8 T cells are likely to be less effective in preventing infection (very low IE_S), although they may reduce the viral load during infection

(modest IE_I). Consequently, the selection pressure on a virus imposed by *all* CD8 T cell responses, $1-(1-\mathrm{IE}_S)(1-\mathrm{IE}_I)$, will be relatively low. Furthermore, since a virus has multiple CD8 T cell epitopes, the selective advantage of a variant that escapes CD8 T cell responses to a single epitope would be considerably smaller (see section 2.3.2 for details).

In conclusion, although CD8 T cells may provide some protection against severe pathology, escape-variants having a mutated CD8 T cell epitope are unlikely to have much selective advantage, even in the hosts with the relevant MHC-I that presents the wild-type epitope. In other words, we expect s to be small.

Parameter f.

Escape-variants will accrue advantages in hh hosts and to a lesser extent (proportional to the dominance coefficient r) in Hh hosts. For example, the R384G escaping mutation on the $\mathrm{NP}_{383-391}$ epitope is advantageous in the hosts who carry B*08:01 or B*27:05 but not in those who do not carry these alleles. In Figure 2.2, we show the distribution of experimentally-verified CD8 T cell epitopes derived from the nucleoprotein of human IAV retrieved from the Immune Epitope Database (iedb.org). It is clear that no single epitope is presented by all human leucocyte antigen (HLA) alleles, nor is there a single HLA allele presenting all epitopes. With this information and the frequencies of HLA alleles based on the Allele Frequency Net Database (allelefrequencies.net), we estimated the fraction of host population where the mutation at each amino acid residue would confer a selective advantage (top panel of Figure 2.2). We see that mutations typically confer selective advantage to the virus in only a small fraction of human population.

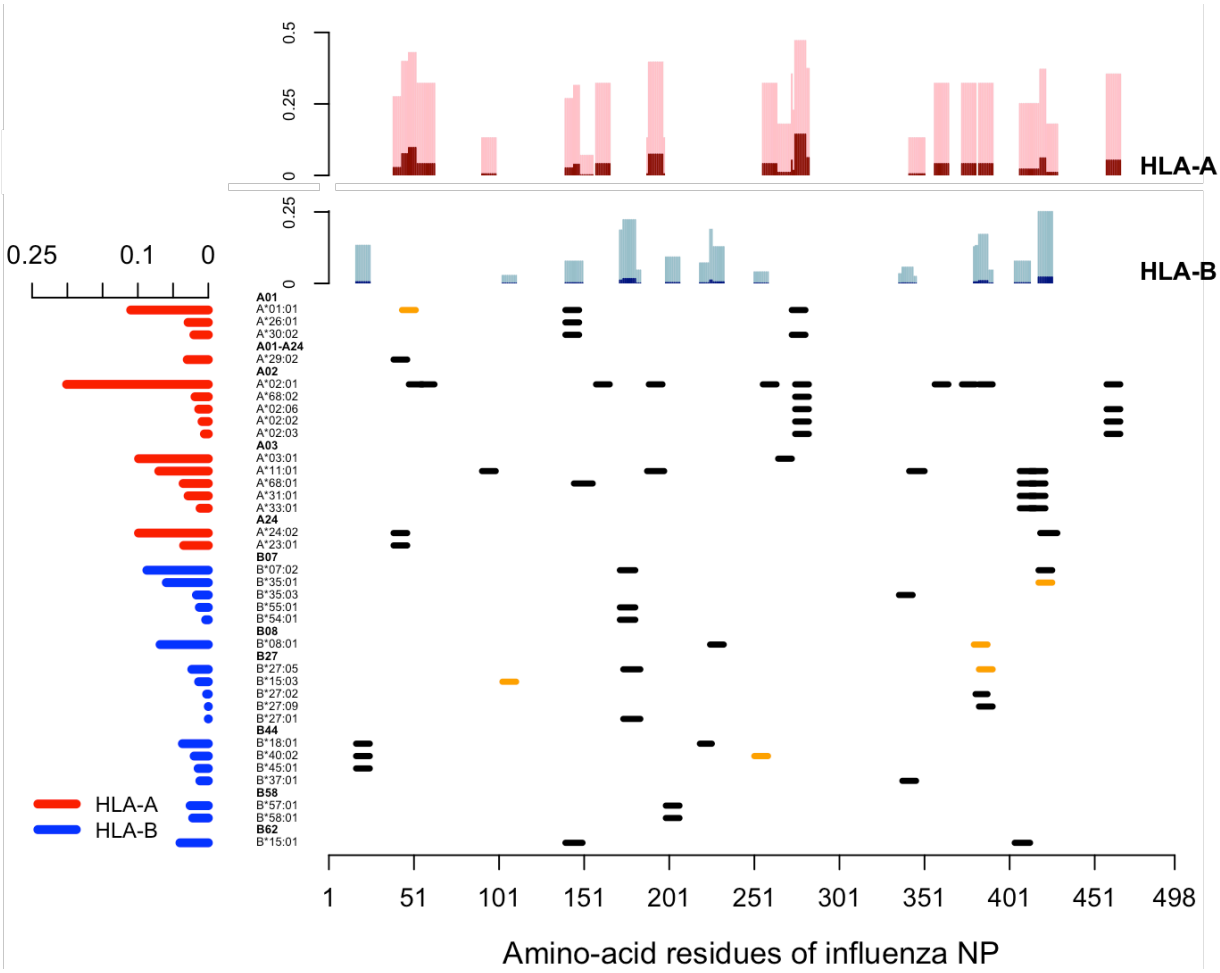

Figure 2.2: Distribution of CD8 T cell epitopes derived from the nucleoprotein of human IAV. In the main panel, each segment represents one unique CD8 T cell epitope derived from nucleoprotein of human IAV, aligned with its relevant HLA allele. Epitopes that have empirically verified escape-mutants reported are labeled in orange. The bar graph in left panel shows the weighted average frequency of the alleles, which are grouped into 11 supertypes (highlighted in boldface). The bar graph in the top panel shows the fractions of the population in which the virus with a mutation in the corresponding amino acid may have a selective advantage. The fractions of homozygote are shown in dark shades (dark red for HLA-A and dark blue for HLA-B) and the fractions of heterozygote are in light shades (pink for HLA-A and light blue for HLA-B).

Integrating parameters s, f, and m.

As mentioned in section 1.3.2, epistatic interactions allow the virus to potentially generate escape-variants to a given CD8 T cell epitope without engendering a substantial fitness cost. Our results show that, since s and f are small, these escape-variants will spread very slowly in the host population. For example, when $f = 0.1$ (i.e. the escape-variant has advantages in about 20% of the infections, who are of hh genotype ($f^2 = 0.01$) and of Hh genotype ($2f(1-f) = 0.18$)), an escape-variant would spend 10 years reaching 50% prevalence, even if its fitness is 10% more than the wild type (i.e. $s = 0.1$) in the hosts with relevant HLA and no fitness cost is accompanied (i.e., $m = 0$).

In conclusion, even if mutations that allow the virus to escape CD8 T cells specific for a given epitope have little or no fitness cost, escape-variants will only increase in frequency very slowly.

2.1.2 Epidemiological models.

Compensatory immunity reduces the selective advantage of mutants over time.

The population genetics framework described above assumes that the fitness of an escape-variant virus depends on host genotype and does not change over time. In particular, we assume that the fitness of escape-variant (MT) relative to the wild-type (WT) equals $(1-m+s)$, $(1-m+rs)$, and $(1-m)$ in the hosts of hh, Hh, and HH genotypes, respectively. However, in a host of hh or Hh genotype who has been infected by MT, recovered, and moved to the susceptible category (due to antigenic drift), we might expect the selective advantage (s) of MT to decrease. This decline in s could arise for at least two biological reasons. First, the mutated epitope is still presented by the MHC-I; the mutation simply changes the configuration of the epitope recognized by the CD8 T cell receptor [261]. In this case, the mutant epitope may induce a new set of CD8 T cells. Second, the absence of CD8 T cell response to one epitope could result in compensatory increases in responses to other epitopes.

We show the fitness of WT or MT infections in the hosts of different genotypes in Figure 2.3. In hosts of HH, the fitnesses of WT and MT are 1 and $(1-m)$, the same as in population genetics model. The fitnesses of MT in hosts of Hh and hh that are infected with MT for the first time are $(1-m+rs)$ and $(1-m+s)$ as described earlier. After the hosts of Hh and hh recover and regain susceptibility due to antigenic drift, the fitness of MT following reinfection with MT becomes $(1-m+rs(1-c))$ and $(1-m+s(1-c))$, where c denotes the extent of compensatory CD8 T cell responses. Parameter c ranges from 0 to 1, where 0 corresponds to no compensatory increase in responses to other epitopes and 1 corresponds to full compensation. The ranges for the fitness of MT in hosts of Hh and hh with prior WT and MT infections are shown by the dashed arrows in Figure 2.3A.

Epidemiology of infections with wild-type and escape-variant viruses.

We use a simple epidemiological model to describe the changes in frequencies of susceptible (S), infected (W and M for WT and MT infections) and immune (R) hosts. The subscript to S, W, M, and R populations indicates the viruses these hosts have been exposed to in the past. Individuals can be infected multiple times during their lifetime due to antigenic drift at antibody epitopes [295]. We incorporate this by letting individuals move from the immune (R) to susceptible (S) compartments at rate ω [296]21.

We consider the epidemiology of WT infections in individuals of different genotypes, all of which have the same structure as shown in Figure 2.3B. Prior to the introduction of MT, we assume that the WT is circulating and hosts have CD8 T cell immunity to the wild-type epitope. Due to antigenic drift, individuals typically get reinfected with a drifted strain every 5-10 years [295], and we choose the rate of loss of immunity corresponding to this duration ($\omega = 5 \times 10^{-4}$/day $\approx$ once every 5.5 years). We begin the simulations with WT infections at equilibrium.

Now, on the introduction of MT, the MT has fitness $(1-m)$, $(1-m+rs)$, or $(1-m+s)$ in the MT-infected hosts of the HH, Hh or hh genotypes, respectively. MT-infected individuals move to the immune category with a subscript of WM (i.e., R_{WM}). Individuals in R_{WM} become susceptible due to antigenic drift in the virus and move to S_{WM}. When individuals in S_{WM} are infected with the WT, they move to W_{WM} and the WT has fitness 1, while when individuals in S_{WM} are infected with the MT, they move to the M_{WM} and the MT has fitness $(1-m)$, $(1-m+rs(1-c))$, or $(1-m+s(1-c))$ according to the host's genotype. For simplicity, we incorporate the fitness in the transmissibility parameter (β).

In Figure 2.4, we explore how the escape-variant (MT) spreads through the host population following its introduction. In particular, we focus on how compensatory immunity changes the outcomes predicted by the population genetics models. In Panel 2.4A, we chose a simple scenario where the MT has a very small fitness cost ($m = 0.001$), a 5% and 2.5% selective advantages ($s = 0.05, rs = 0.025$) in the hosts of hh and Hh genotypes, which accounts for 1% and 18% of the population ($f = 0.1$), respectively, and compensatory immunity reduces the selective advantage by 90% ($c = 0.9$). In this scenario, we see that the MT now only transiently invades, and compensation in host immunity causes the frequency of MT to decline as the population-level immunity against the MT increases. In Panel 2.4B, we explore the consequences of changing the extent of compensation (c). We see that the initial rate of invasion is very similar to what is predicted by the population genetics model. However, once the MT has spread through the population, the outcome depends strongly on the extent of compensatory immunity described by the parameter c. If c is small, then the MT goes to fixation in a manner similar to that of the population genetics model. If c exceeds a threshold value c^* given by

$$c^* = 1 - \frac{m}{sf^2 + 2rsf(1-f)} \tag{2.2}$$

then the MT only transiently invades but declines to extinction in the long run. The competitive

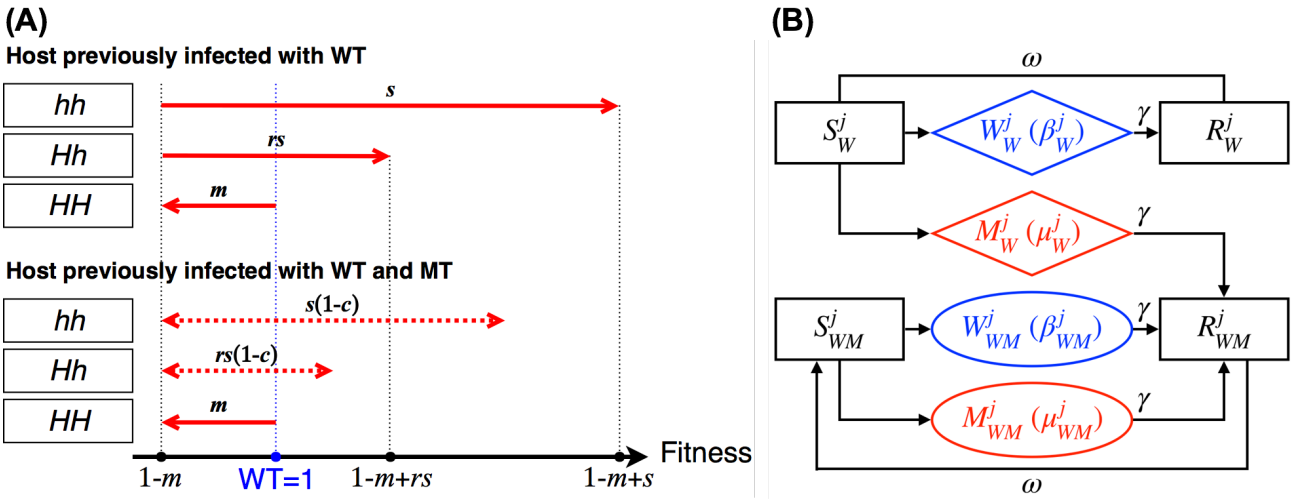

Figure 2.3: (A) Compensatory immunity alters the fitness of escape-variant infection in the Hh and hh individuals. We show the viral fitness in the hosts of different genotypes and infection histories (being previously infected by WT only or both WT and MT). The fitness of WT infection is 1 in all hosts (shown as the blue dot on the fitness scale). The fitness of MT in HH, Hh, and hh hosts are $(1-m)$, $(1-m+rs)$, and $(1-m+s)$, respectively (shown as the red arrows in the top panel and black dots on the fitness scale). Subsequent MT infections of Hh and hh hosts result in lower viral fitness, $(1-m+rs(1-c))$ and $(1-m+s(1-c))$, which is dependent on the extent of compensatory immunity, which reduces the selective advantage by c, and the dominance coefficient, r (shown as the red dashed arrows in the bottom panel). (B) Diagram illustrating the epidemiology of infections with WT (W shown in blue) and MT (M shown in red) in any of the three genotypes. Susceptible (S) and immune (R) hosts are indicated by S and R, respectively. The host genotype is indicated by the superscript j ($j=1,2,3$ for HH, Hh, and hh), and the prior infection status is indicated by the subscript (W for WT and M for MT).

exclusion between WT and MT infections is shown in Panel 2.4C, where we plot how the outcome depends on s, f, and c, given $r = 0.5$. Incorporating fitness parameters into the duration of infection give similar results (results not shown). Furthermore, we show these results are robust to the addition of seasonality (section 2.3.3).

In summary, the results of the epidemiological models show that the initial rate of invasion of the escape-variant is similar to that in the population genetics model described by Equation 2.1. However, at later time points, compensatory immunity reduces the rate of invasion. If the extent of CD8 T cell immunity to the escape-variant is sufficiently high, the outcome may even reversed if the overall selective advantage does not surmount fitness cost.

2.2 Materials and Methods

2.2.1 Population genetics model

The analytical solution is found by rearranging Equation 2.1 into

$$\frac{q_t}{1-q_t} = K\left(\frac{q_{t-1}}{1-q_{t-1}}\right) = K^t\left(\frac{q_0}{1-q_0}\right)$$

We then express t, the number of generation required for the MT to reach q_t, by

$$t = \frac{1}{\log K}\left[\log\left(\frac{q_t}{1-q_t}\right) - \log\left(\frac{q_0}{1-q_0}\right)\right]$$

The escape-variant can invade when $K > 1$, i.e.,

$$sf^2 + 2rsf(1-f) > m$$

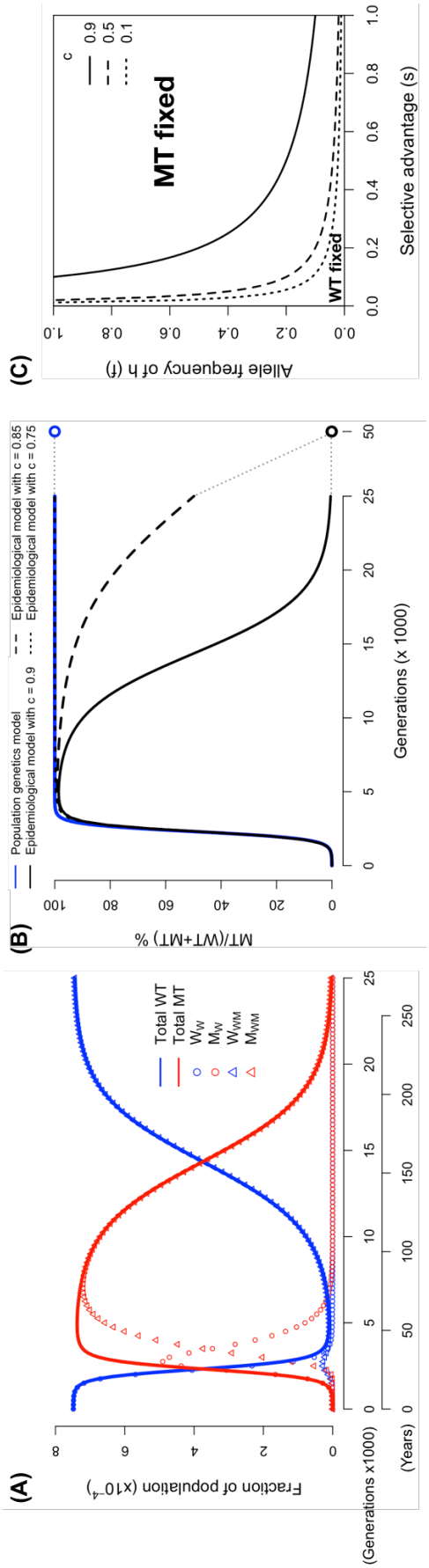

Figure 2.4: (A) The frequency of wild-type (WT) and escape-variant (MT) infections as a function of time. When CD8 T cell responses are compensated and thus decreases the selective advantage, the MT only invades transiently and goes extinct in the long run. ($m = 0.001$, $s = 0.05$, $f = 0.1$, $r = 0.5$, $c = 0.9$). (B) The MT prevalence predicted by epidemiological model with different extents of compensation (black lines), compared to the population genetics model (blue line). In all scenarios, the initial invasions of the MT (before it reaches 50% prevalence) are similar. After it reaches 50%, if the extent of compensation is smaller than the threshold (in our parameter setting, $c^* = 0.8$. See Equation 2.2), the MT goes to fixation only slightly slower than what predicted by the population genetics model. In contrast, if the extent of compensation is higher than the threshold, the MT invades transiently and becomes extinct in the long run. (C) The parameter regimes of s and f where either MT or WT becomes fixed (and the other becomes extinct). We see that as the extent of compensation rises, the parameter regime where MT becomes fixed shrinks.

2.2.2 Map of CD8 T cell epitope on influenza nucleoprotein

Epitope dataset was retrieved from Immune Epitope Database (iedb.org). We searched for MHC class I-restricted linear epitope of influenza A virus (ID: 11320, FLUAV) in humans with at least one positive T cell assay. We retrieved 1,220 records from IEDB, of which 514 were derived from NP. After excluding the records longer than 12 amino-acid residues or with no HLA allele information available, records with the same amino-acid sequence, with different sequences but at the same location of NP and presented by the same HLA allele, or nested under a longer record, were combined into one 'unique' epitope. In total, 64 unique epitopes were identified. Escaping mutations were identified from the literatures [249, 261].

HLA allele dataset reported by National Marrow Donor Program (NMDP) was retrieved from The Allele Frequency Net Database (iedb.org). We included all the alleles that have been reported to present at least one epitope in the epitope dataset, and calculated the average frequency weighted by sample sizes. In addition, since the alleles in one HLA supertype prefer amino acid with similar chemical property at certain residues of the epitopes, we grouped the HLA alleles based on the classification proposed by Sidney et al. [168].

2.2.3 Epidemiological model

The ordinary differential equations corresponding to Figure 2.3B are presented below.

$$
\begin{aligned}
\frac{\mathrm{d}S_W^j}{\mathrm{d}t} &= \omega R_W^j - S_W^j \sum_{j=1}^{3} \left(\beta_W^j W_W^j + \mu_W^j M_W^j + \beta_{WM}^j W_{WM}^j + \mu_{WM}^j M_{WM}^j \right) \\
\frac{\mathrm{d}S_{WM}^j}{\mathrm{d}t} &= \omega R_{WM}^j - S_{WM}^j \sum_{j=1}^{3} \left(\beta_W^j W_W^j + \mu_W^j M_W^j + \beta_{WM}^j W_{WM}^j + \mu_{WM}^j M_{WM}^j \right) \\
\frac{\mathrm{d}W_W^j}{\mathrm{d}t} &= S_W^j \sum_{j=1}^{3} \left(\beta_W^j W_W^j + \beta_{WM}^j W_{WM}^j \right) - \gamma W_W^j \\
\frac{\mathrm{d}M_W^j}{\mathrm{d}t} &= S_W^j \sum_{j=1}^{3} \left(\mu_W^j M_W^j + \mu_{WM}^j M_{WM}^j \right) - \gamma M_W^j \\
\frac{\mathrm{d}W_{WM}^j}{\mathrm{d}t} &= S_{WM}^j \sum_{j=1}^{3} \left(\beta_W^j W_W^j + \beta_{WM}^j W_{WM}^j \right) - \gamma W_{WM}^j \\
\frac{\mathrm{d}M_{WM}^j}{\mathrm{d}t} &= S_{WM}^j \sum_{j=1}^{3} \left(\mu_W^j M_W^j + \mu_{WM}^j M_{WM}^j \right) - \gamma M_{WM}^j \\
\frac{\mathrm{d}R_W^j}{\mathrm{d}t} &= \gamma W_W^j - \omega R_W^j \\
\frac{\mathrm{d}R_{WM}^j}{\mathrm{d}t} &= \gamma (M_W^j + W_{WM}^j + M_{WM}^j) - \omega R_{WM}^j
\end{aligned}
$$

where $j = 1, 2, 3$ denotes the genotype of HH, Hh, and hh, respectively. We started simulations from the equilibrium of WT infection, i.e.,

$$
{}^*S_W^j = \frac{\gamma}{\beta} \cdot \mathrm{Freq}(j), \quad {}^*I_W^j = \frac{\omega}{\omega + \gamma} \left(1 - \frac{\gamma}{\beta} \right) \cdot \mathrm{Freq}(j), \quad {}^*R_W^j = \frac{\gamma}{\omega + \gamma} \left(1 - \frac{\gamma}{\beta} \right) \cdot \mathrm{Freq}(j)
$$

where $\beta = \beta_W^1 = \beta_W^2 = \beta_W^3$ and $\mathrm{Freq}(j)$ is determined by f and r under Hardy-Weinberg Equilibrium. Values of parameters are listed in Table 2.1.

2.3 Appendix

2.3.1 Population genetics model

Figure 2.5 shows the number of generations for a CD8 T cell escape-variant (MT) reaches 50% from 0.01% of prevalence with different fitness costs ($m = 0, 0.001, 0.01, 0.1$) and dominance coefficients ($r = 0, 0.25, 0.75, 1$). As m increases, the MT requires higher s and/or f to overcome the deleterious effect. The increase in r changes the shape of the contours since it raises the selective advantage of MT in the heterozygotes. As a result, the threshold for MT invasion when r is large is lower than that when r is small.

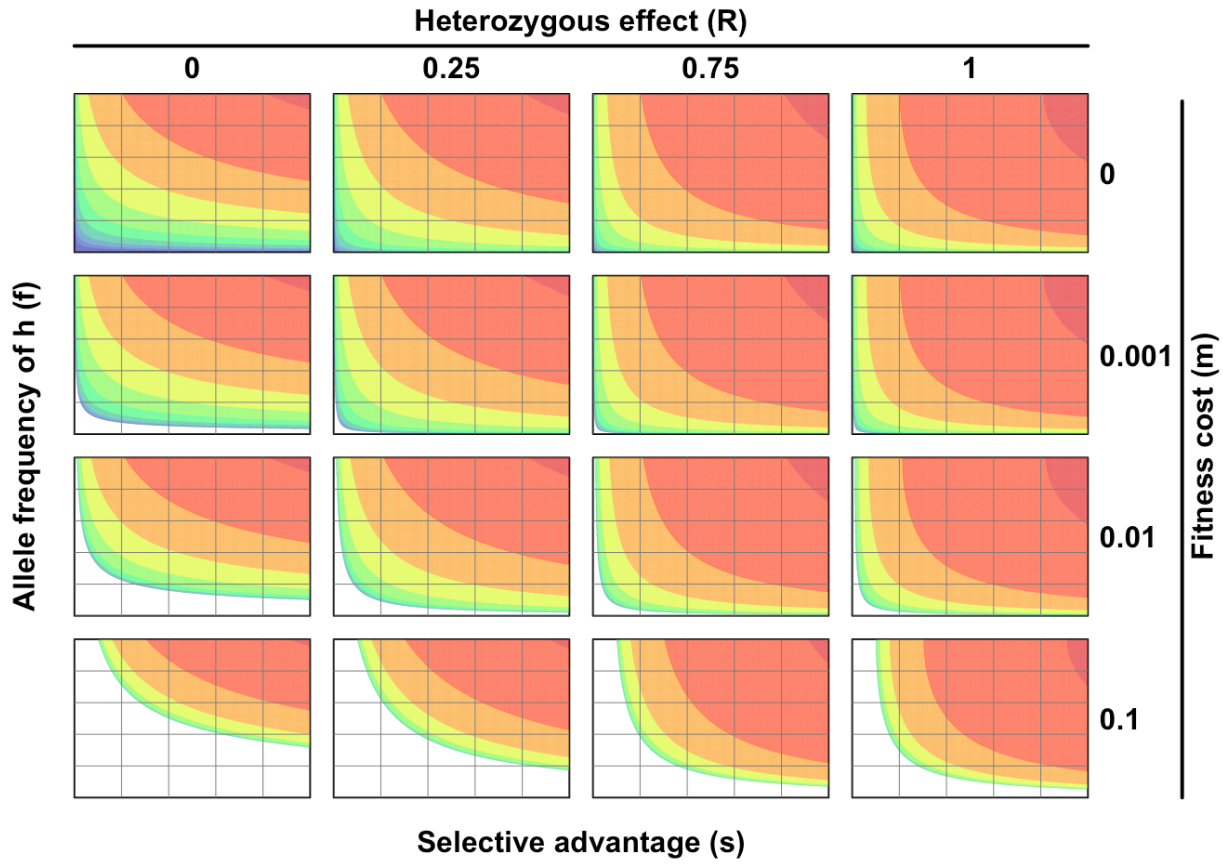

Figure 2.5: Number of generations for a CD8 T cell escape-variant reaches 50% from 0.01% of prevalence. The scale and color coding are the same as in Figure 2.1A.

2.3.2 Quantifying selection pressure on virus

We define three types of immunity effectiveness (IE) as follows:

- IE_S is the probability that an individual who have influenza-specific CD8 T cells does not get infected upon a contact with an infectious individual
- IE_P is the probability that an infected who have influenza-specific CD8 T cells does not develop symptoms.
- IE_I as the probability that an infected who have influenza-specific CD8 T cells does not spread the virus.

The selection pressure on virus ($\mathcal{S}$) is formulated by

$$1 - \mathcal{S} = (1 - \text{IE}_S)(1 - \text{IE}_I)$$

The viral fitness after selection, $1 - \mathcal{S}$, is the probability of transmission within the host population who have influenza-specific CD8 T cells, and can be expressed as the probability that an immune individual gets infected ($1 - \text{IE}_S$) and spreads the virus ($1 - \text{IE}_I$).

2.3.3 Effect of seasonality

Seasonality is incorporated into transmissibility by

$$\beta_i^j(t) = \beta_i^j \left[1 + A \sin\left(\frac{2\pi t}{T}\right)\right] \tag{2.3}$$

where i and j indicate the host's immune status and genotype, respectively; $0 \leqslant A \leqslant 1$ is the amplitude of oscillation, and T is the period [297]. Simulation of model with $A = 0.02$, $T = 365$ (annual outbreaks with β ranging from 0.392 to 0.408) is compared to the one with $A = 0$ (no

seasonality). Our results show that the mean of the oscillation is about the same as seen in the model with no seasonality (Figure 2.6).

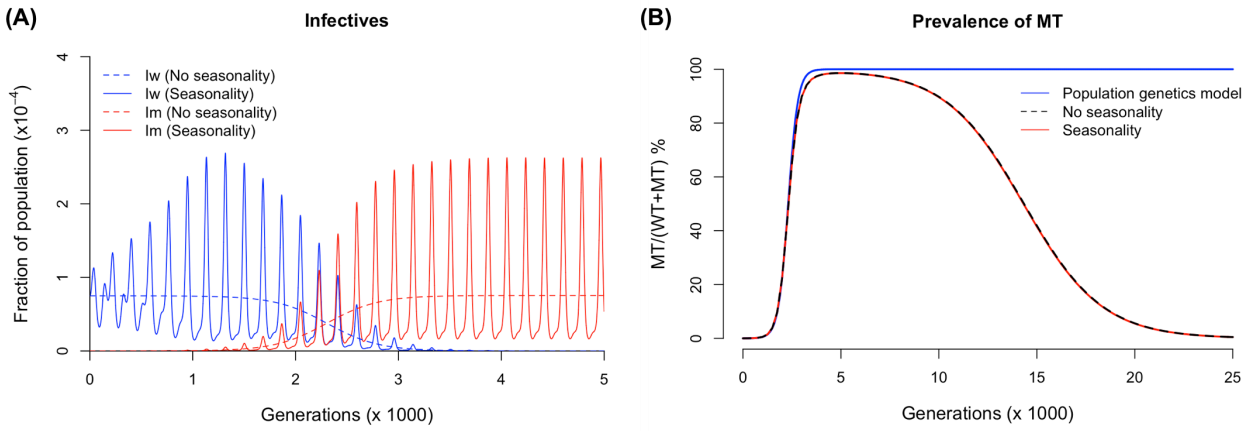

Figure 2.6: (A) Comparison between predictions of epidemiological model with ($A = 0.02$, $T = 365$ days in Equation 2.3) and without seasonality ($A = 0$). Other parameters are kept the same as in Figure 2.4A. (B) Comparison between predicted MT prevalence from population genetics model, epidemiological with and without seasonality.

Table 2.1: Model parameters

Parameter	Symbol	Value
Transmission rate of WT-infected (day^{-1})[†‡]	β	0.4
Recovery rate (day^{-1})	γ	0.25
Drifting rate (day^{-1})	ω	5×10^{-4}

[†] $\beta = \beta_W^j = \beta_{WM}^j$, $j = 1, 2, 3$.
[‡] See Table 2.2 for the setup of transmission parameters in MT-infected.

Table 2.2: Transmissibility according to genotypes and immune status

Compartment	Symbol	Value
M_W^1	μ_W^1	$\beta(1-m)$
M_W^2	μ_W^2	$\beta(1-m+rs)$
M_W^3	μ_W^3	$\beta(1-m+s)$
M_{WM}^1	μ_{WM}^1	$\beta(1-m)$
M_{WM}^2	μ_{WM}^2	$\beta(1-m+rs(1-c))$
M_{WM}^3	μ_{WM}^3	$\beta(1-m+s(1-c))$

Chapter 3

Quantifying the memory CD8 T cell-mediated immune selection pressure on influenza A virus *in vivo*

3.1 Results

3.1.1 *In vitro* viral growth

Prior studies have identified a mutation in the immunodominant PR8 influenza nucleoprotein $(NP)_{366\text{-}374}$ CD8 T cell epitope presented on H-2^b that prevents loading of the peptide onto MHC-I. We constructed wild-type (WT) and NP-N370Q mutant (MT) PR8 influenza viruses by reverse genetics and first compared viral growth in MDCK cell culture to determine if this mutation impacted viral fitness. The viral titers at five time points after inoculation are shown in Figure 3.1, where no significant difference was detected between virus strains (three-way ANOVA, p-value for virus strains = 0.797). A further logistic growth model fitting found no significant difference between the estimated model parameters associated with WT and MT PR8 viruses (Table 3.1). Therefore, we conclude that the NP-N370Q mutation does not result in a fitness defect *in vitro*.

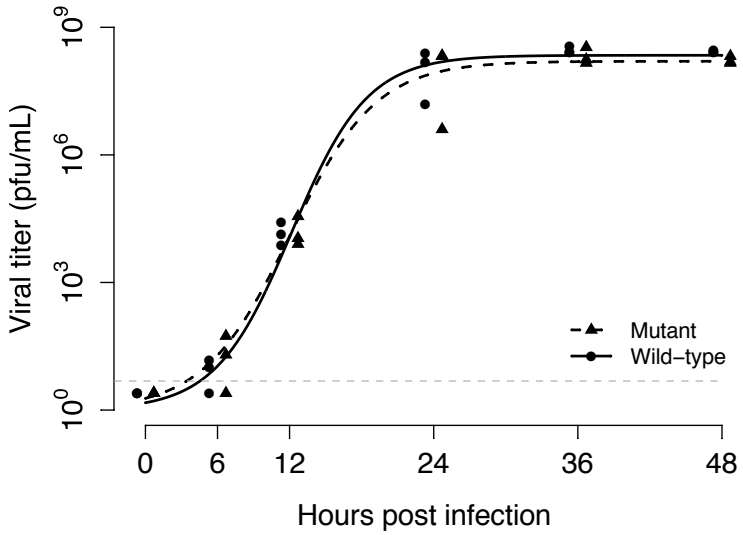

Figure 3.1: Viral growths of wild-type (WT) and NP-N370Q mutant (MT) PR8 influenza viruses in MDCK cell culture. MDCK cells were infected with WT or MT viruses at a multiplicity of infection (MOI) of 0.01. No significant difference was detected between virus strains (three-way ANOVA, p-value for virus strains = 0.797). The gray dotted line denotes the limit of detection (5 pfu/mL).

Table 3.1: The point estimates and 95% confidence intervals for the logistic growth model parameters

Virus strain	Growth rate	Capacity	Midpoint
Wild type	0.32 [0.24, 0.43]	8.3 [8.1, 8.6]	12.2 [11.5, 13.0]
Mutant	0.28 [0.21, 0.40]	8.2 [7.8, 8.6]	12.0 [11.0, 13.4]

Table 3.2: Comparison of viral load measured by droplet digital PCR (copy number/mL) and viral titer measured by plaque assay (pfu/mL)

Virus strain	RNA copy number (10^8 per mL)	Plaque forming unit (10^8 per mL)	Ratio (RNA copy / pfu)
Wild-type	85 (± 6.0)	10.5 (± 1.5)	8.1
Mutant	67 (± 8.3)	9.8 (± 1.5)	6.9

3.1.2 Validation of droplet digital PCR

In order to infer the impact of CD8 T cell-escaping mutations from in vivo experiments, we first had to design and validate a method for simultaneous measurement of WT and MT PR8 viruses within the same host. Thus, we employed a droplet digital PCR (ddPCR) system with probes specific for the WT or MT variants of the $NP_{366\text{-}374}$ epitope (See Supplemental Information). The assay is highly specific for the individual strains (Fig 3.2A), as the WT probe detected positive signals only from the WT viral RNA samples, and the MT probe was similarly specific for the MT viral RNA samples. This assay also robustly reflects the change in RNA concentration, as a 10-fold dilution of input viruses resulted in a 10-fold decrease in the final readouts indicating unbiased detection of WT and MT viruses (Fig 3.2B). Furthermore, the viral loads measured by ddPCR (in copies/mL) were consistent with viral titers measured by plaque assay (in pfu/mL), and there was no preferential detection of WT or MT viral RNA (Table 3.2). Thus, the ddPCR system allows us to accurately differentiate between WT and MT PR8 viruses in a mixed sample.

3.1.3 *In vivo* viral kinetics in primary infection

To determine if the MT PR8 virus has any fitness defect in vivo in the absence of influenza-specific memory CD8 T cell immunity, we infected naïve C57BL/6 (B6) mice with an equal mixture of WT and MT PR8 viruses (Fig 3.3A). In naïve B6 mice, the viral loads of WT and MT viruses in the lungs recapitulated the kinetics reported previously [211], where they (i) increased exponentially

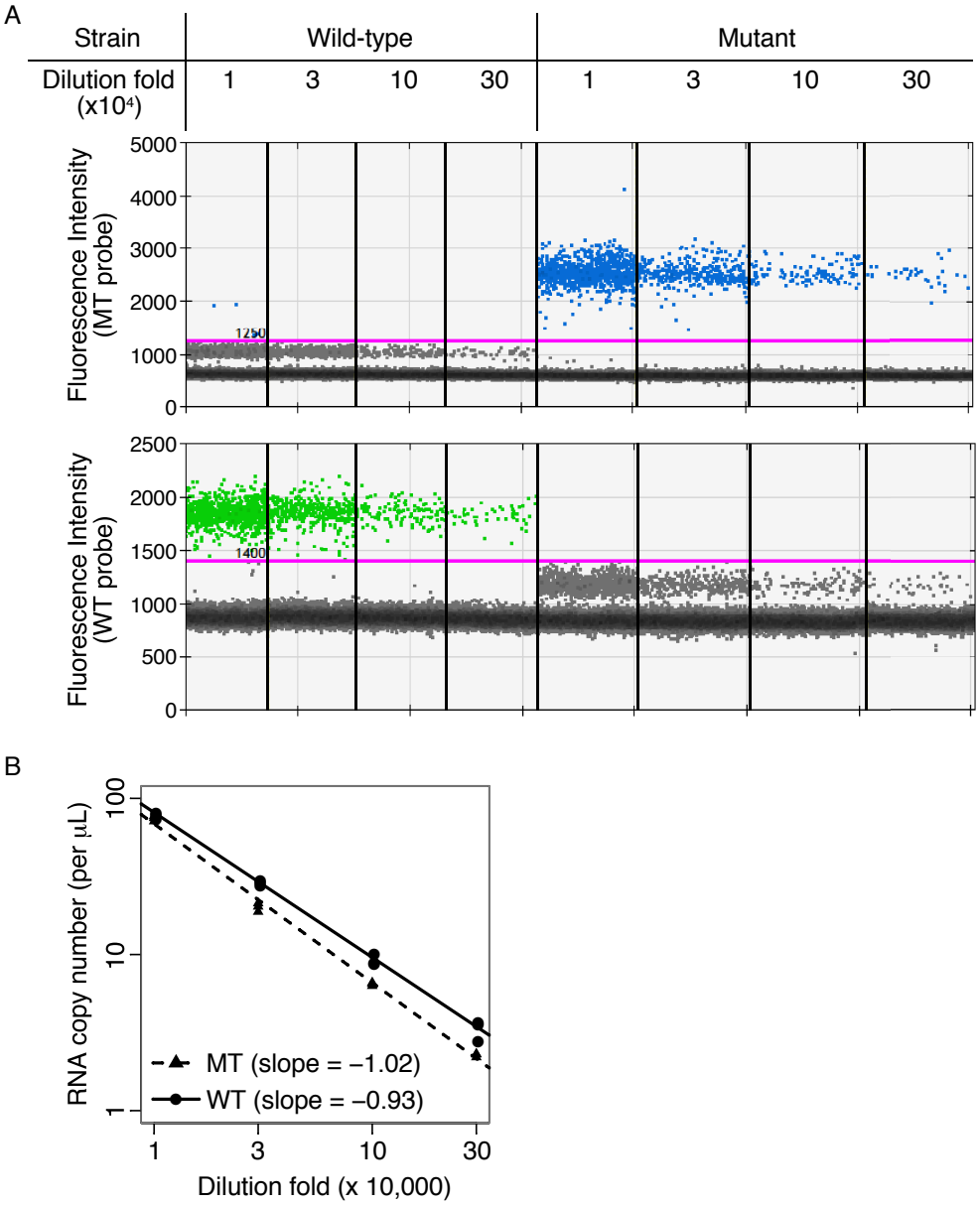

Figure 3.2: Testing the performance of droplet digital PCR (ddPCR) with serially diluted WT and MT PR8 virus stocks. (A) A representative figure of ddPCR analysis demonstrates a clear discrimination of positive signals detected by the MT probe (blue dots in the top panel) or the WT probe (green dots in the bot-tom panel) from negative signals (gray dots in both panels). (B) Regression lines quantified from the data in (A) over a range of virus dilutions.

and peaked around day 3 post-infection, (ii) slowly decayed from days 3 to 5, and (iii) quickly decayed and were cleared after day 5 (Fig 3.3B). Across the 5 time points being sampled, only on day 5 was the average log-transformed viral load of MT significantly higher than the WT ($p = 0.0007$).

We then defined the selection coefficient of MT as

$$\frac{\text{AUC}_t[\log(\text{MT})]}{\text{AUC}_t[\log(\text{WT})]} - 1$$

where AUC_t stands for the area under viral growth curve from days 0 to t. This statistic quantifies the fitness change associated with MT as a fraction of the fitness of WT. For instance, 0.1 indicates the MT has 10% higher fitness compared to WT, while -0.15 indicates the MT has 15% lower fitness compared to the WT. The sampling distribution of this statistic was assessed by the bootstrapping method (see section 3.2).

The selection coefficient of MT gradually increased from 0.001 on day 1 to 0.1 on day 9, while the 99% empirical bootstrap confidence interval (99% CI hereafter) did not contain zero only on day 9 (Fig 3.3C). These findings showed that the MT virus does not have a detectable fitness defect during the early stages of infection in vivo, but gains an advantage as the infection progresses, likely due to its ability to escape from the immunodominant $\text{NP}_{366\text{-}374}$-specific effector CD8 T cells.

3.1.4 *In vivo* viral kinetics in intranasally x31-immunized mice

To estimate the selective advantage of the MT virus under different settings of cellular immunity, we intranasally (i.n.) primed B6 mice with HKx31 (H3N2), challenged the mice 30 days later with an equal mixture of WT and MT PR8 (H1N1) viruses, and measured the viral kinetics (Fig 3.4A). Viral loads peaked on day 2, slowly decayed from days 2 to 4, and rapidly decayed and cleared from days 4 to 8. Despite a similar overall pattern to naïve mice, the peak viral loads were around 10-fold

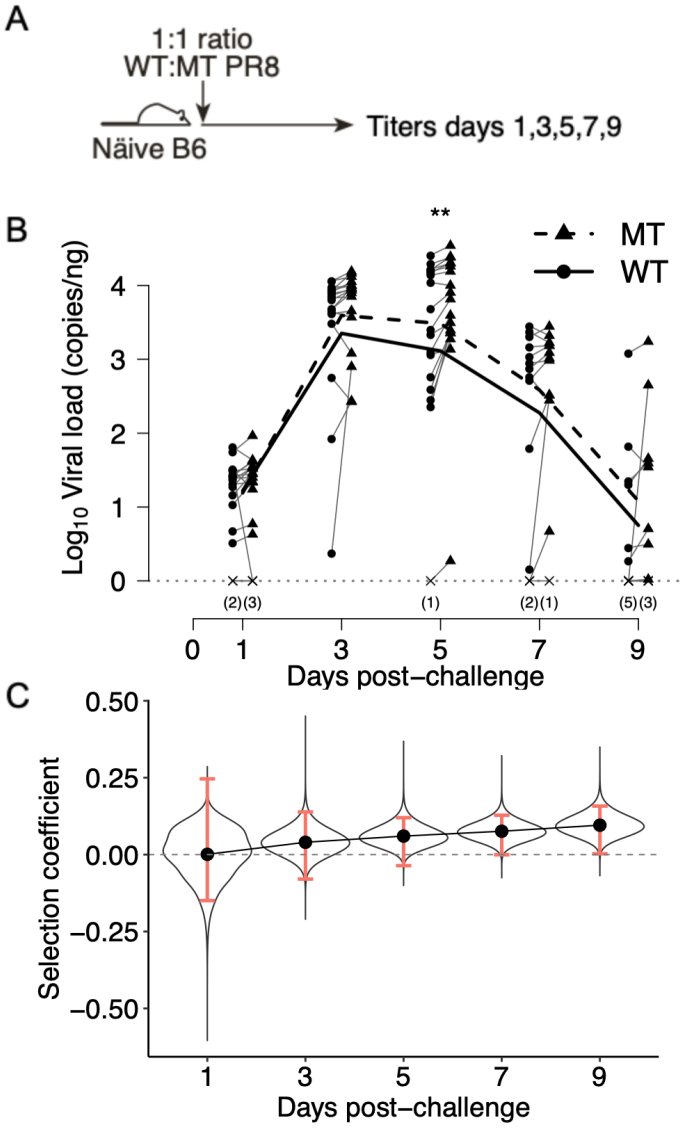

Figure 3.3: Comparison of the wild-type (WT) and mutant (MT) kinetics in primary infection. (A) Naïve C57BL/6 (B6) mice were inoculated with an equal mixture of WT and MT PR8 viruses, and the viral loads were measured on selected time points. (B) *In vivo* viral kinetics on days 1, 3, 5, 7, and 9 post-infection as measured by ddPCR. Detectable WT, detectable MT, and undetectable viral loads are marked by circles, triangles, and crosses, respectively. Paired data points are connected with a line. Numbers of undetectable data points are indicated in the parentheses. (C) The selection coefficient of MT and sampling distributions measured by AUC and bootstrapping. The point estimates and 99% CIs for the selection coefficient of MT are indicated by the dots and the red bars, respectively.

lower, and clearance was faster. The means of log-transformed MT viral loads were significantly higher than those of WT on days 1 and 4 ($p = 0.034$ and 0.0018, respectively) (Fig 3.4B). Likewise, the selection coefficient of MT continuously increased from 0.15 on day 1 to 0.27 on day 9, while the 99% CIs contained zero only on day 1 (Fig 3.4C). Thus, the MT acquired an advantage in the early stage of infection, and the advantage persisted and increased through the infection course.

To investigate whether an increased MHC diversity impacts the selective advantage of the MT virus, we applied the same prime-challenge procedure to CB6F1 (F1) mice, which are the offspring of C57BL/6 and BALB/c mouse strains, and harbor both H-2^d and H-2^b haplotypes (Fig 3.4D). Thus, F1 mice develop a broader influenza-specific CD8 T cell response that encompasses epitopes presented by both MHC alleles. We observed similar viral kinetics in F1 mice compared to B6 mice; however, the viruses were cleared even faster than B6 mice (no virus was detected on days 6 and 8), and the MT significantly outgrew WT only on day 4 ($p = 0.0054$) (Fig 3.4E).

When looking at the selection coefficient of the MT virus, we noticed two interesting differences between i.n.-primed B6 and F1 mice (Fig 3.4F). First, the point estimates were much lower on the first two days in F1 than B6 mice (day 1: 0.04 in F1 vs. 0.15 in B6; day 3: 0.1 in F1 vs. 0.17 in B6), and the 99% CIs contained zero on both days; however, this deviation became negligible on day 4. Second, after day 4, the selection coefficient of MT showed a pronounced plateau in F1 mice, corresponding to the fact that both WT and MT viruses were cleared on days 6 and 8. Together, these data suggest that escape from cellular immunity specific for the $NP_{366\text{-}374}$ epitope confers a fitness advantage to the MT virus early during heterosubtypic influenza infection. The impact from increased breadth of CD8 T cell response was not constant across the infection course; instead, it depends on the time of observation.

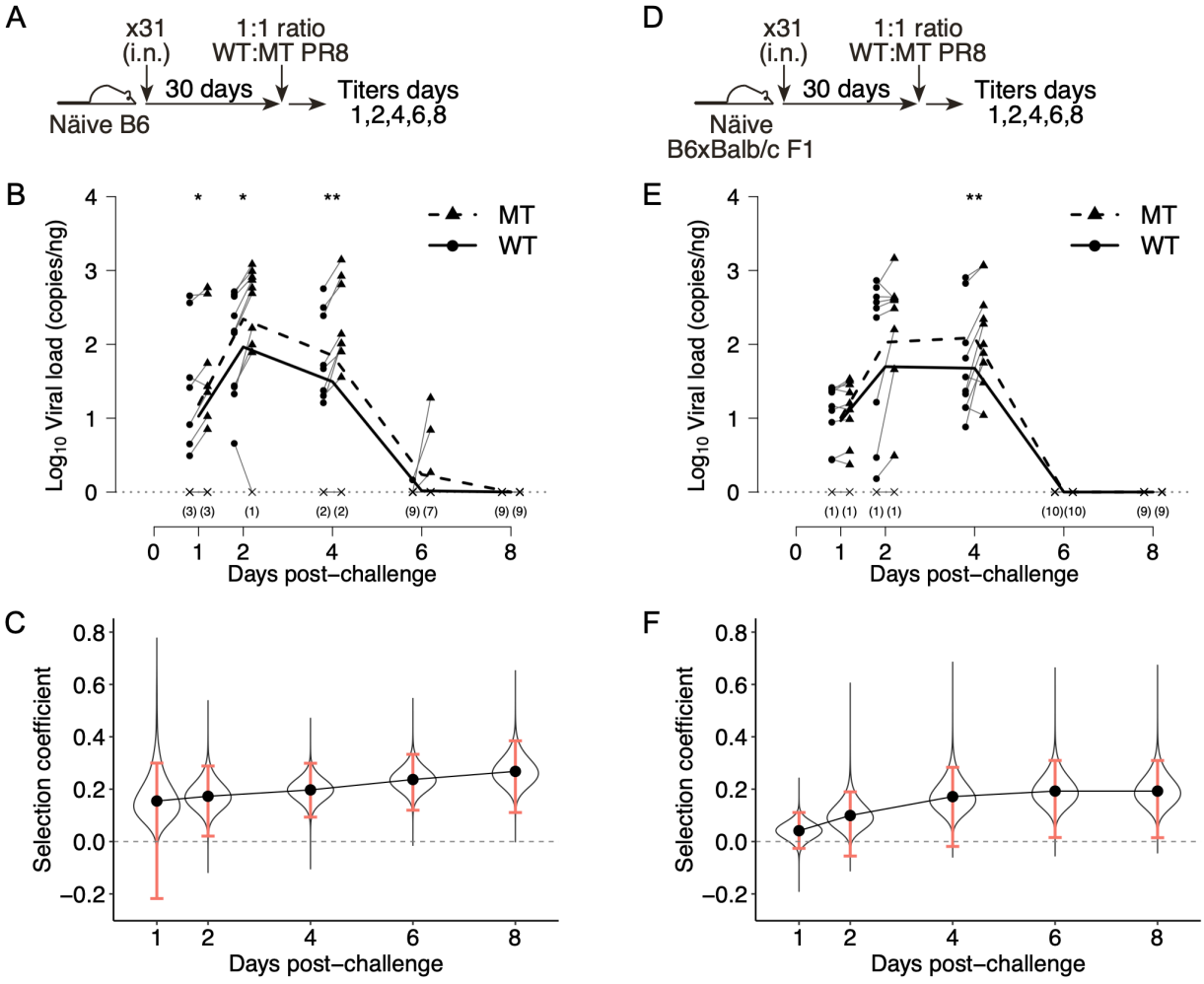

Figure 3.4: Viral kinetics in intranasally (i.n.) HKx31 (x31)-primed B6 (H-$2^{b/b}$) and CB6F1 (H-$2^{d/b}$) mice. (A) B6 mice were immunized i.n. with x31 and challenged with an equal mixture of WT and MT PR8 viruses 30 days later. (B) *In vivo* viral kinetics of i.n. x31-primed B6 mice on days 1, 2, 4, 6, and 8 post-challenge as measured by ddPCR. (C) The selection coefficient of MT in i.n. x31-primed B6 mice as measured AUC and bootstrapping. The 99% CIs did not contain zero except on day 1. (D) CB6F1 (F1) mice, which are the heterozygotes of H-2^b and H-2^d haplotypes, were immunized i.n. with x31 and challenged with an equal mixture of WT and MT PR8 viruses 30 days later. (E) *In vivo* viral kinetics of i.n. x31-primed F1 mice on days 1, 2, 4, 6, and 8 post-challenge as measured by ddPCR. (F) The selection coefficient of MT and sampling distribution in i.n. x31-primed F1 mice as measured by AUC and bootstrapping.

3.1.5 *In vivo* viral kinetics in AdNP-immunized mice

Intranasal priming of B6 mice with x31 generates memory CD8 T cells specific for multiple epitopes in addition to the immunodominant $NP_{366\text{-}374}$ epitope, including two additional immunodominant and at least eight subdominant epitopes [256, 257]. Therefore, even if the MT PR8 virus escapes detection from $NP_{366\text{-}374}$-specific CD8 T cells, it remains subject to recognition by CD8 T cells targeting other epitopes. However, some vaccine approaches are designed to focus the CD8 T cell response to a few immunodominant epitopes, raising the question of the selective advantage that can be gained by a mutant influenza virus that can escape a focused influenza-specific memory CD8 T cell repertoire. To address this question, we immunized B6 mice with a recombinant, replication-deficient adenovirus 5 expressing the PR8 influenza nucleoprotein (AdNP), which generates memory CD8 T cells specific for only NP-derived epitopes, mainly the immunodominant $NP_{366\text{-}374}$ epitope [247], and challenged them 30 days later with a mixture of WT and MT PR8 viruses (Fig 3.5A). Vastly different growth kinetics of the MT virus were observed compared to the WT virus (Fig 3.5B). The MT grew continuously and peaked around day 4, at a viral load 35-fold higher than WT. After day 4, both viruses decayed at the same rate exponentially, but neither were completely cleared on day 8. At each time point selected, the means of log-transformed MT viral loads were significantly higher than those of WT ($p = 0.031$, 0.0004, 0.0002, 0.0016, and 0.023 for days 1, 2, 4, 6, and 8, respectively). Consistent with this, the selection coefficient of MT continuously climbed throughout the infectious course from 0.24 on day 1 to 0.8 on day 8, and all the 99

3.1.6 *In vivo* viral kinetics in intramuscularly x31-immunized mice

Lung-resident memory CD8 T cells (lung T_{RM}) are important for mediating heterosubtypic immunity and controlling influenza virus replication due to their localization in the respiratory tract enabling rapid detection of infected cells. Thus, we hypothesize lung T_{RM} impose much of the early

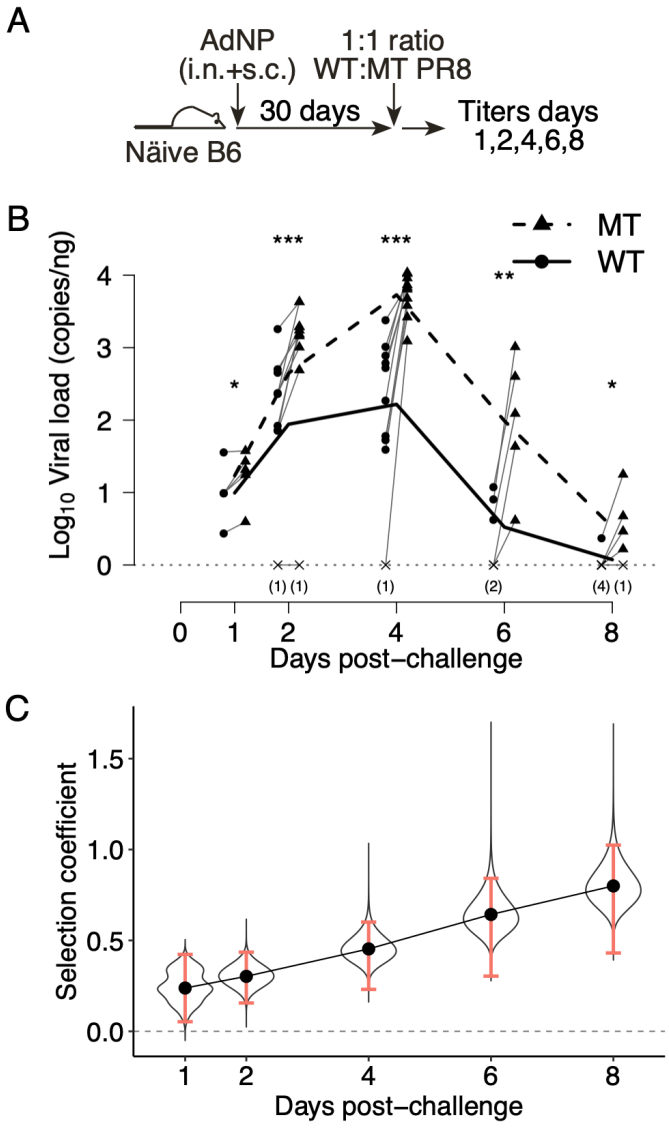

Figure 3.5: Viral kinetics in the AdNP-primed B6 mice, where only the $NP_{366\text{-}374}$-specific memory CD8 T cells were present. (A) B6 mice were i.n. and subcutaneously (s.c.) immunized with AdNP and challenged with an equal mixture of WT and MT PR8 viruses 30 days later. (B) *In vivo* viral kinetics of AdNP-primed B6 mice on days 1, 2, 4, 6, and 8 post-challenge as measured by ddPCR. (C) The selection coefficient of MT and sampling distributions in AdNP-primed B6 mice as measured by AUC and boot-strapping. All the 99% CIs did not contain zero.

cellular immune pressure on the virus. We tested this hypothesis by comparing the viral kinetics in 30-day i.n. x31-primed B6 mice with 30-day intramuscularly (i.m.) x31-primed B6 mice (Figure 3.6A), which we previously showed to have approximately the same number of influenza-specific systemic memory CD8 T cells as i.n.-primed mice but lack lung T_{RM} [208]. We observed that, in i.m. x31-primed mice, (i) the viral loads peaked at a later time point (day 4) and were about 6- and 12-fold higher than the peaks seen in i.n. x31-primed mice, (ii) the MT and WT viruses grew at the same rate between days 1-4 and peaked at similar levels, and (iii) after day 4, the WT virus was cleared faster than the MT virus ($p = 0.021$ on day 6, $p = 0.0095$ on day 8) (Figure 3.6B). Interestingly, the selection coefficient of MT showed a U-shape trend; it decreased from 0.12 on day 1 to 0.057 on day 4, and then increased to 0.22 on day 8 (Figure 3.6C). Nevertheless, these data showed the MT virus does not acquire the same increased advantage during the early stage of infection when the lung T_{RM} is absent. Thus, the selection pressure mediated by lung T_{RM} is the main driver of the outgrowth of MT virus observed in i.n. x31-primed B6 mice.

3.1.7 Estimation of selection coefficient

We summarized the selection coefficients of MT virus estimated from log-transformed viral kinetics in Table 3.3. Overall, the MT had higher fitness than the WT through the infection course, regardless of the context of pre-existing cellular immunity. The fitness gain ranged from 10% to 74%, depending on the immune settings. However, a stratifying analysis revealed that the fitness gain does not evenly distribute across the infection course. During the first 4 or 5 days, the MT had little advantage in naïve and i.m. x31-primed mice. This implied most of the advantage was acquired later during infection under these scenarios. In contrast, the MT acquired around 15% increase in fitness in both i.n. x31-primed B6 and F1 mice through days 0-4, but over the whole infection course this advantage increased to 24% in B6 mice while it was maintained in F1,

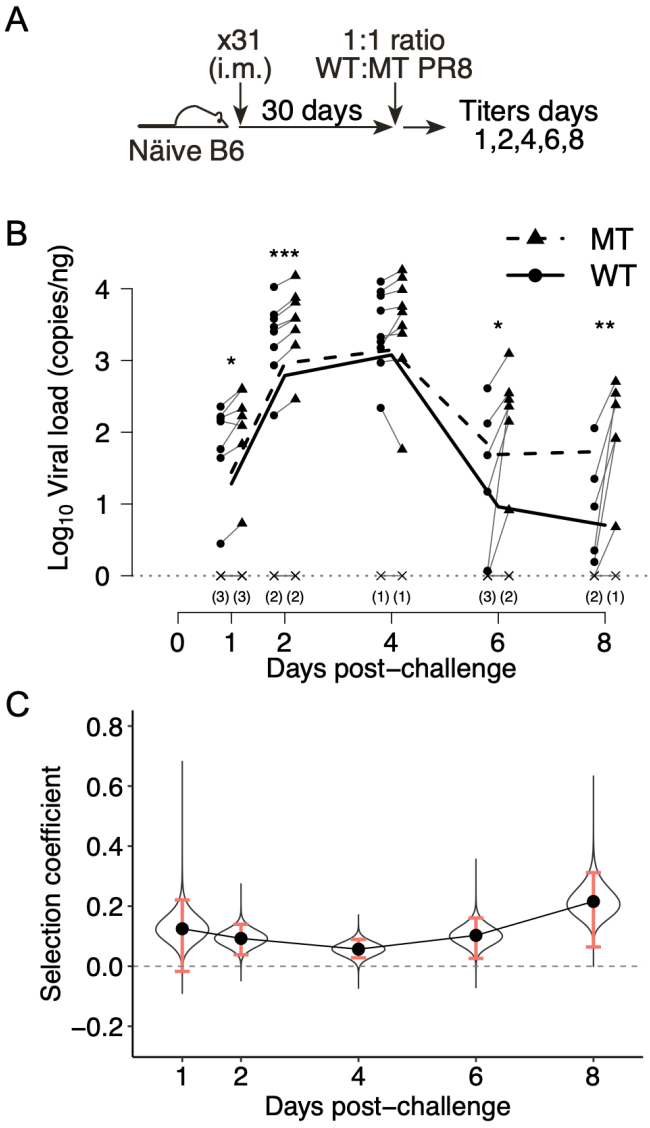

Figure 3.6: Viral kinetics in intramuscularly (i.m.) x31-primed B6 mice that lack lung TRM. (A) B6 mice were i.m. immunized with x31 and challenged with an equal mixture of WT and MT PR8 viruses 30 days later. (B) *In vivo* viral kinetics of i.m. x31-primed B6 mice on days 1, 2, 4, 6, and 8 post-challenge as measured by ddPCR. (C) The selection coefficient of MT and sampling distributions in i.m. x31-primed B6 mice as measured by AUC and bootstrapping. The 99% CIs did not contain zero except day 1.

corresponding to faster viral clearance in F1 mice. Lastly, the MT acquired a large and stable fitness gain in AdNP-primed B6 mice, at an average rate 8% per day. In summary, the NP-N370Q mutation confers no more than 25% increase in fitness when CD8 T cell immunity targets multiple epitopes, while it can confer up to 74% increase in fitness when only the $NP_{366\text{-}374}$ epitope is targeted.

Table 3.3: Point estimates and 99% empirical bootstrap confidence intervals for the selection coefficient of MT based on log-transformed viral kinetics

		Days 0-4		Days 0-8	
Immune status	Mouse	Estimate	99% CI	Estimate	99% CI
Naïve[†]	B6	0.06	[-0.036 , 0.12]	0.096	[0.003 , 0.16]
i.n. x31	B6	0.20	[0.093 , 0.30]	0.27	[0.11 , 0.38]
i.n. x31	F1	0.17	[-0.019 , 0.28]	0.19	[0.015 , 0.31]
AdNP	B6	0.45	[0.23 , 0.60]	0.80	[0.43 , 1.02]
i.m. x31	B6	0.057	[0.028 , 0.09]	0.22	[0.064 , 0.31]

[†]For naïve mice, point estimates and bootstrap were conducted over days 0-5 and days 0-9.

We explored the reason why escape from $NP_{366\text{-}374}$-specific CD8 T cells only confers modest selective advantage, focusing on CD8 T cell kinetics during secondary infection. We challenged i.n. x31-primed B6 mice with either WT or MT viruses alone and tracked the numbers and percentages of $CD44^{+}Tetramer^{+}$ CD8 T cells in the lung interstitium on days 0, 2, 4, and 7 (Figure 3.7A). The percentage of $D^{b}NP_{366}$-specific CD8 T cells increased from 5% to around 50% in the WT-challenged mice, but shrank to 1.5% in the MT-challenged mice. In contrast, the $D^{b}PA_{224}$- and $K^{b}PB1_{703}$-specific CD8 T cells increased from 14% to 40% in the MT-challenged mice, distinct from the modest increase of 10% to 13% observed in the WT-challenged mice. The kinetics in the airways had greater variation but followed the same pattern (Figure 3.7B). These data suggest the loss of $D^{b}NP_{366}$-specific CD8 T cell response might be compensated by the CD8 T cells against other epitopes.

Finally, we attempted to infer the transmission fitness of the viruses based on their replicative

fitness, which is estimated by the AUC of viral kinetics. Since the relationship between transmissibility and viral loads in the respiratory tract hasn't been well established, we conducted the AUC and bootstrap analysis assuming the relationship is linear or sigmoidal (Table 3.4. Also see Figure 3.8 for the sensitivity analysis on the choice of parameters of Hill function). Although different point estimates and 99% CIs were obtained, the choice of link function did not qualitatively change the inference about CD8 T cell-mediated selection pressure.

3.2 Materials and Methods

3.2.1 Viruses

The HKx31 (H3N2) virus was amplified in 10- to 11-day-old embryonated chicken eggs as previously described [208]. The wild-type (WT) PR8 (H1N1) virus and the NP-N370Q mutant (MT) virus were generated using the reverse genetics system [298].

3.2.2 Mice and infections

C57BL/6 (H-$2^{b/b}$) and CB6F1 (H-$2^{b/d}$) mice were purchased from Jackson Laboratory and inoculated at 8 weeks of age. For primary infection, mice were inoculated with 50 μL of 1:1 mixture of WT and MT PR8, 125 PFU for each (total dose is 250 PFU, corresponding to 1 LD50 of PR8). For secondary infection, mice were primed with (i) 3×10^4 EID_{50} of HKx31 in 30 μL HBSS intranasally (i.n.), (ii) 10^6 EID_{50} of HKx31 in 50 μL HBSS intramuscularly (i.m.), or (iii) AdNP virus in 30 μL of PBS both i.n. and subcutaneously (s.c.) [247]. Thirty days later, the immunized mice received a 50 uL of 1:1 mixture that consists of 1,250 PFU of each PR8 virus (total dose is 2,500 PFU, corresponding to 10 LD50). Challenged mice were euthanized with Avertin on the indicated days post-infection. All mouse studies were approved by the Institutional Animal Care and Use Committee of Emory University.

3.2.3 Measuring viral load using droplet digital PCR

RNA of virus stocks was isolated using QIAamp Viral RNA Mini Kit and stored at −80°C until use. For RNA isolation from infected mice, lungs were minced and preserved in RNALater at −80°C until all the samples of the same batch were collected. Total RNA was isolated from lung homogenates using Ambion RNA Isolation Kit and the concentration was measured by Nanodrop. For reverse transcription and ddPCR, cDNA from the total RNA of lungs or viral RNA was made using Maxima RT and universal influenza primers. The cDNA samples then underwent 10-fold serial dilution and were quantified by ddPCR with probes designed to discriminate wild-type and $NP_{366\text{-}374}$ mutant epitopes. The RNA copy number data were back-calculated to the RNA concentration (copy number per ng of total RNA, see section 3.3 Droplet digital PCR).

3.2.4 Quantitative analysis

Statistical tests

The *in vitro* viral titer data were tested by a three-way ANOVA to account for (i) the virus strain effect, (ii) temporal effect, and (iii) batch effect. The data were then used to estimate the parameters and the 95% confidence intervals of the logistic growth model. The log-transformed *in vivo* viral load data were tested by Student's t-test for paired data. All the analyses were done in R 3.6.1.

Area under the curve and bootstrapping

The area under the *in vivo* viral growth curve (AUC) represents the total amount of virus produced over the course of infection and serves as a marker for the replicative fitness. Based on our assumption that the transmission fitness is related to replicative fitness, we log-transformed the viral load data and calculated the AUCs of the mutant and the wild type viruses. Bootstrapping was performed one million times to assess the uncertainty of estimation (see section 3.3 Bootstrapping).

The 99% empirical bootstrap confidence intervals were calculated following basic bootstrap method [299].

3.3 Appendix

Maxima RT

Mix 4 μL of 5x RT buffer, 1 μL of universal F(A) primer (6 μM), 1 μL of dNTPs (10 mM), 1 μL of ribolock RI, 1 μL of Maxima RT with either 12 μL of viral RNA isolated from virus stock or 2 to 4 μg of total RNA isolated from the lungs on ice. Run for 30 minutes at 55°C, 10 minutes at 85°C, and hold at 10°C.

Droplet Digital PCR

A mixture containing 11 μL of 2x ddPCR supermix for probes, 1.1 μL of mixed forward and reverse primers (36 μM), 1.1 μL of wild-type probe, 1.1 μL of mutant probe, 3.3 μL of ultrapure water, and 4.4 μL of cDNA sample was used to amplify a fragment containing the WT or MT $NP_{366\text{-}374}$ epitope. 20 μL of this mixture was added to 70 μL of droplet generation oil, and after the droplet generation step, 40 μL of the suspension was used to perform ddPCR in a 96-well PCR plate. The PCR steps are listed below:

1. 95°C, 10 mins.
2. 94°C, 30 secs $\rightarrow$ 60°C, 1 min (35-40 cycles).
3. 98°C, 10 mins.
4. Hold on 4°C.

The fluorescent signal was read by a QX200 Droplet Reader (BioRad) and analyzed with QuantaSoft software. The gating for positive droplets was set according to the positive and negative controls on each plate. The viral loads (V, copy number/ng of RNA) were calculated by the formula

$$V = x \cdot \frac{100D}{LT}$$

where x is the readout of ddPCR (copy number/μL), D is the dilution factor, L is the volume of RNA used for RT (μL), T is the RNA concentration measured by Nanodrop (ng/μL), and 100 is a coefficient with unit μL.

1. Forward primer: `5'-GCATGCCATTCTGCCGCATT-3'`
2. Reverse primer: `5'-GCTGATTTGGCCCGCAGATG-3'`
3. Probe for the WT $NP_{366\text{-}374}$: `5'-/HEX/TA+GT+CT+CCATATTTTCATT+G+GAA+GC/BHQ-1/-3'`
4. Probe for the MT $NP_{366\text{-}374}$: `5'-/FAM/TA+GT+CT+CCATCTGTTCATT+G+GAA+GC/BHQ-1/-3'`

The nucleotides with a `+` prefix are locked nucleotides.

Surface staining and flow cytometry

Mice received 1.5 μg of CD3ε-PE/CF594 in 200 μL sterile PBS 5 minutes before euthanasia. The bronchoalveolar lavage (BAL), lungs, mediastinal lymph node (mLN), and spleen were harvested and processed as previously described [300, 246]. Cells were stained with D^bNP_{366}-BV421, D^bPA_{224}-PE, and K^bPB1_{703}-APC tetramers provided by NIH Tetramer Core and CD4-BV510, CD8α-BV785, CD44-A700, CD62L-BV605, CD69-A488, CD103-PerCP/Cy5.5, and Zombie-NIR. Samples were run on a Fortessa X-20 flow cytometer (BD Biosciences) and the percentages of tetramer-positive cells were calculated following analysis with FlowJo software.

Bootstrapping

For each immune setting, the bootstrapping on AUC was done by iterating the following procedure:

1. On each time point, of which n mice were measured, randomly sample n mice with replacement. Therefore, each sampled mouse will give one WT and one MT viral load.

2. Calculate the mean of log-transformed WT and MT viral loads, and then compute the AUCs of WT and MT with the means.

This procedure was repeated for a million times, and the bootstrapped AUCs were used to approximate the sampling distribution.

Sensitivity Analysis

To investigate the effects of link function on the estimates of selection coefficient in transmission, we also conducted the AUC and bootstrapping analysis on linear and Hill function, besides logarithm. Formally,

$$T_1(x) = ax + b$$

$$T_2(x) = a\log(x) + b$$

$$T_3(x) = \frac{a\log(x)^r}{x_0^r + \log(x)^r} + b$$

The constant b is assumed to be 0. In T_1, this assumption indicates transmission does not happen when there is no virus. In T_2 and T_3, it means transmission does not happen when the viral load is below 1 copy/ng, which is the detection limit of our ddPCR system. We notice that the choice of coefficient (a) does not affect the estimate of selection coefficient; however, the rate (r) and midpoint (x_0) of Hill function will. In Table 3.4 we adopted the estimates from Handel et al. [301], where $r = 4.8$ and $x_0 = 2.6$, and we ran sensitivity analyses on these two parameters.

Table 3.4: Point estimates and 99% empirical bootstrap confidence interval of the selection coefficient of MT under different immune settings and link functions

		Linear		Logarithmic		Hill	
Immune status	Mouse	Estimate	99% CI	Estimate	99% CI	Estimate	99% CI
Days 0-4							
Naïve†	B6	0.30	[0.16 , 0.42]	0.06	[-0.04 , 0.12]	0.08	[-0.06 , 0.15]
i.n. x31	B6	1.31	[0.69 , 1.71]	0.20	[0.09 , 0.30]	0.55	[0.06 , 0.77]
i.n. x31	F1	0.50	[-0.23 , 1.01]	0.17	[-0.02 , 0.28]	0.33	[-0.28 , 0.56]
AdNP	B6	7.04	[-2.06 , 10.35]	0.45	[0.23 , 0.60]	0.99	[0.12 , 1.43]
i.m. x31	B6	0.47	[0.28 , 0.60]	0.06	[0.03 , 0.09]	0.07	[0.01 , 0.13]
Days 0-8							
Naïve†	B6	0.31	[0.18 , 0.41]	0.10	[0.003 , 0.16]	0.11	[-0.01 , 0.19]
i.n. x31	B6	1.37	[0.79 , 1.76]	0.27	[0.11 , 0.38]	0.58	[0.09 , 0.80]
i.n. x31	F1	0.59	[-0.11 , 1.05]	0.19	[0.01 , 0.31]	0.42	[-0.27 , 0.69]
AdNP	B6	8.39	[-2.9 , 12.48]	0.80	[0.43 , 1.02]	1.48	[0.06 , 2.19]
i.m. x31	B6	0.49	[0.23 , 0.65]	0.26	[0.06 , 0.31]	0.25	[0.09 , 0.38]

†For naïve mice, point estimates and bootstrap were conducted over days 0-5 and days 0-9.

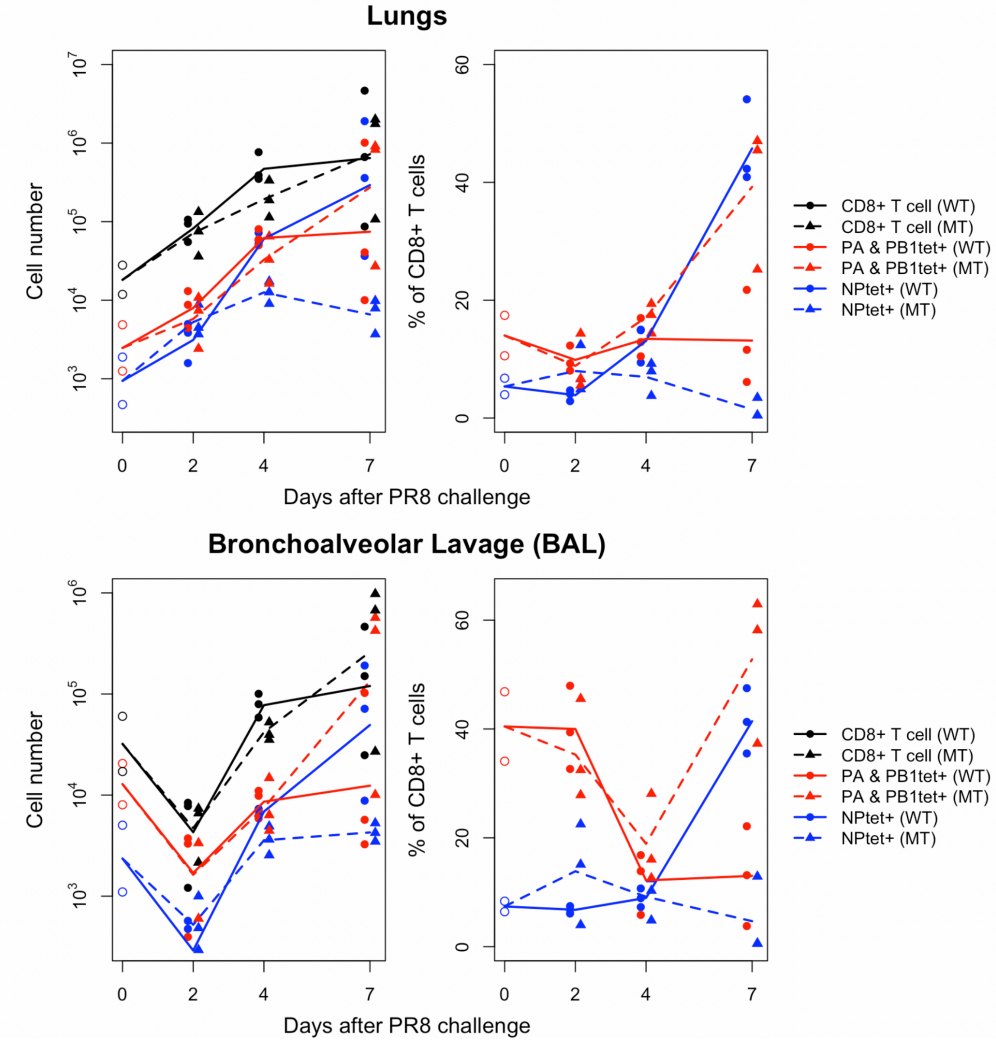

Figure 3.7: Number and percentage of antigen-specific CD8 T cells in the lung interstitium and airways following challenge with WT or MT viruses. (A, B) i.n. x31-primed B6 mice were rested for 30 days and infected with either WT or MT PR8 viruses. On days 0, 2, 4, and 7 post-challenge, the number and percentage of total CD8 T cells (black), FluPA- and FluPB1-specific CD8 T cells (red), and FluNP-specific CD8 T cells (blue) were assessed in mice challenged with WT (circle) or MT (triangle) viruses.

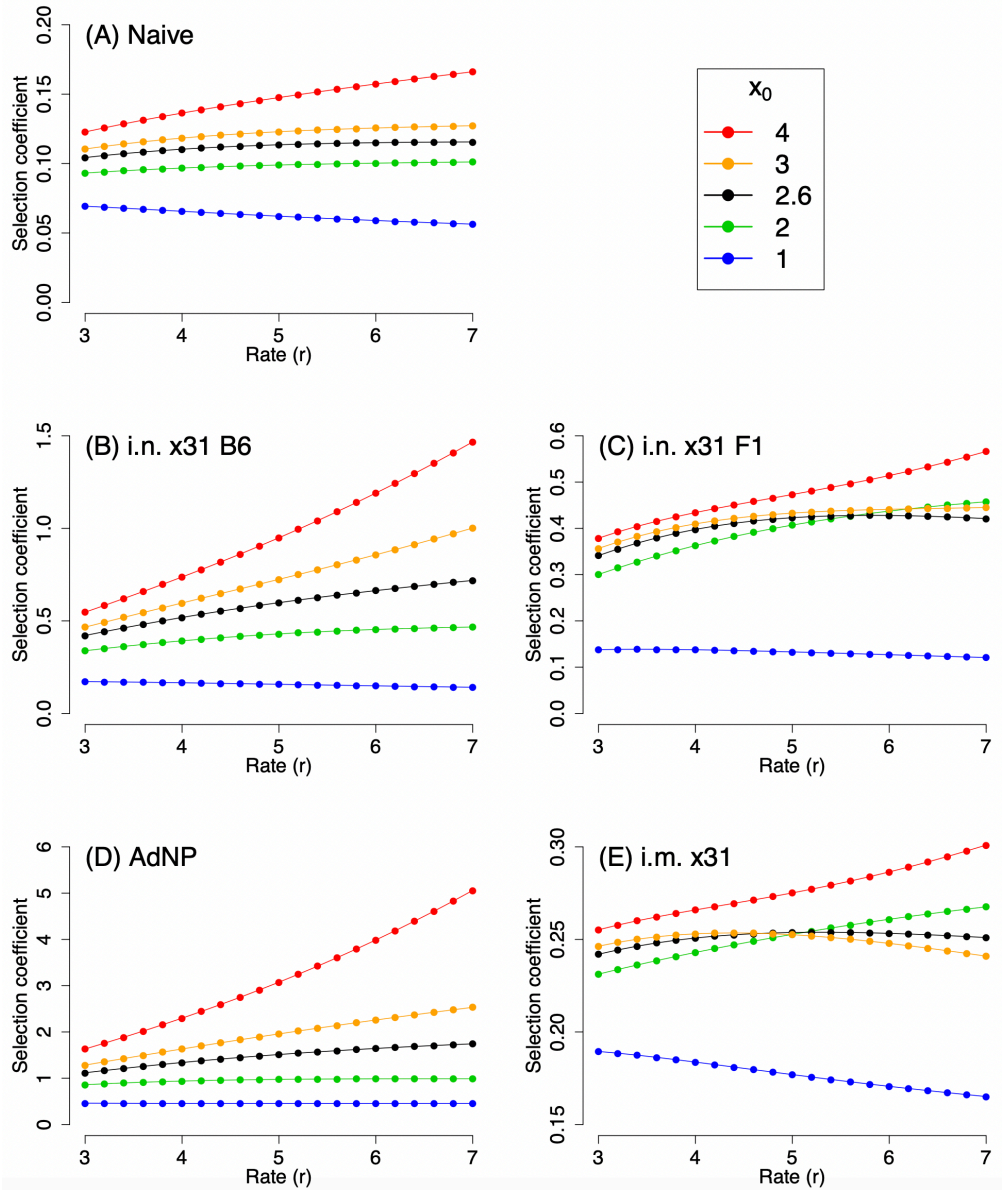

Figure 3.8: Sensitivity analysis on the rate (r) and midpoint (x_0) parameters of Hill function. The selection coefficients were calculated based on the whole infection course (naïve: days 0-9; others: days 0-8).

Chapter 4

Discussion

The Discussion chapter is structured in four sections. Section 1 recapitulates the scientific questions addressed in this dissertation, and the results will be discussed in Section 2. In Section 3, I will integrate the results using the framework of measuring viral fitness proposed by Wargo and Kurath (2012). Section 4 is devoted to discussing the implications of these results for the development of T cell-based vaccines.

4.1 Scientific questions

The continual change of hemagglutinin enables influenza A virus (IAV) to escape from antibody-mediated immunity and necessitates the annual update of IAV vaccines. Over the past decade, vaccines that induce cellular immunity against conserved T cell epitopes of IAV have been proposed a candidate approach for universal influenza vaccine.

Many studies have shown cellular immunity, CD8 T cells in particular, can cross-react to, and protect against heterosubtypic [197, 198, 199, 200, 201, 196, 203, 204, 205, 206, 207, 208, 209, 210] or even heterotypic [250] influenza viruses. Although CD8 T cells do not prevent viral entry as neutralizing antibodies do, they constrain viral replication by killing infected cells as well as releasing cytokines that induce resistance in neighboring cells [222, 224, 237, 238]; as a result, they reduce pathology and likely transmission, although direct evidence for diminished transmission is

currently limited. Yet, even if the viruses are under pressure from cellular immunity, the T cell epitopes do not mutate as quickly as the antibody epitopes at the population level. Therefore, the first question to be addressed is why CD8 T cell epitopes of IAV are conserved.

From the perspective of population genetics, the evolution of a gene is driven by two mechanisms: (1) Genetic drift that guides the dynamics of neutral alleles and (2) selection that keeps the beneficial alleles or purges the detrimental ones [284]. Indeed, Machkovech et al. [290] and Woolthuis et al. [291] have detected signs of positive selection among human IAVs, consistent with the idea that escaping cellular immunity will be adventageous to the virus. These findings give rise to a second question: Can we quantify the selection advantage a CD8 T cell-escaping variant would acquire in the presence of memory CD8 T cells?

4.2 Discussion of the results

4.2.1 Modeling project

In regards of the evolutionary mechanism underlying the conservation of T cell epitopes, one possibility that has been widely considered is that the nucleoprotein (NP) and matrix-1 (M1) protein, which harbor the bulk of CD8 T cell epitopes, are under strong constraints [287, 288]. In this view, mutations in CD8 T cell epitopes would have a high fitness cost or less likely to occur due to epistatic changes needed for escape-variants to have high fitness.

In our modeling study, we proposed two other mechanisms that could impact the fitness of viruses that escape CD8 T cell recognition. The first one is that escape from CD8 T cell responses against a single epitope provides only a relatively small selective advantage to the escape-variant, as other epitopes could compensate for the loss of a single epitope. The second one is that polymorphism in MHC-I genes at the population level restricts this small selective advantage to only a

fraction of individuals. We show that even if there is a minimal fitness cost to having a mutation in a CD8 T cell epitope, the latter two factors will result in a very low rate of invasion of the CD8 T cell escape-variant. We note that in the preceding models and discussion we consider a scenario where invasion of a CD8 T cell escape-variant is entirely driven by CD8 T cell-mediated selection. In reality, selection due to CD8 T cell immunity occurs in the context of a much stronger selection imposed by antibodies. The latter results in periodic replacement of virus strains over the course of a few seasons, leading to drifted virus strains with updated antigenicity. We consider the consequences of this in more detail later.

The conclusion of this study may seem to contradict the rapid invasion of a CD8 T cell-escaping mutation in the NP that was observed in H3N2 IAV between 1993-1994. This mutation alters the $NP_{383\text{-}391}$ epitope presented by HLA-B27:05 and $NP_{380\text{-}388}$ epitope presented by HLA-B08:01. The wild-type sequence has an arginine (R) at position 384, which forms an anchor residue at this site, and all 16 viruses isolated and sequenced during the 1992–1993 epidemic season had the wild-type sequence [302]. An arginine-to-glycine mutation at this site (R384G) abrogates the MHC-I binding and prevents antigen presentation, and this mutation rapidly swept through the population in the 1993–1994 epidemic season, where all 56 virus isolates had G at this residue. Gog et al. [292] suggested that the rapid fixation of the R384G mutation was due to a combination of a longer duration of infection that slows its decline compared to the wild type virus over the summer, and stochastic events. We propose an alternative hypothesis; that the selective advantage of the R384G mutation is irrelevant for invasion of the mutant, and hitchhiking of a randomly generated mutant would be sufficient to explain the data. The rapid invasion of R384G temporally matches the transition from BE92 to WU95 antigenic clusters (based on HA sequences) [264], suggesting that the NP-R384G mutation could have hitchhiked with the antigenically drifted WU95 strain. Additionally, during 2002–2005, the invasion of a valine-to-isoleucine (V425I) mutant of $NP_{418\text{-}426}$

epitope was temporally associated with the transition from FU02 to CA04 [267]. A similar rapid invasion pattern has also been observed on $NP_{251\text{-}259}$ (invasion of S259L) and $NP_{103\text{-}111}$ (alternating invasions of 103K and 103R), and all of these rapid invasions are consistent with the hitchhiking hypothesis.

We have intentionally used simple models because the empirical data does not include accurate measurements of many of the key parameters that govern the generation and spread of virus escape-variants. In these circumstances, the results of simpler models are typically more robust than those of complex models. Our models can be thought of as simplified representations of the dynamics in the tropical regions, where the virus pool is maintained by continuous transmission. Antigenic changes arising in the tropical regions have been suggested to drive the seasonal epidemics in the Northern and Southern Hemispheres [275, 276].

There are several differences in the ability of influenza viruses to evolve in response to antibody versus CD8 T cell immunity. First, antibodies can generate sterilizing immunity, preventing infection with a matched virus strain, and thus mediate strong selection for antibody escape-variants. In contrast, CD8 T cell immunity to influenza does not prevent infection and thus imposes less selection pressure to a CD8 T cell escape-variant. Second, an antibody escape-variant will gain a selective advantage in the majority of individuals with antibodies to the wild-type strain, while a CD8 T cell escape-variant will have a selective advantage only in individuals with particular MHC-I genotypes. Both these factors contribute to antibody, rather than CD8 T cell immunity, driving antigenic drift in influenza.

Aspects that might be included in more refined models include: the waning of CD8 T cell immunity over time, particularly due to the loss of resident memory cells from the respiratory tract, and the effect of stochasticity. More importantly, our modeling study pointed out the importance of measuring parameters such as the fitness of the wild-type and escape-variants simultaneously in

the same hosts, especially when the host has pre-existing CD8 T cell memory to the wild-type IAV. This is connected to the second part of the dissertation.

4.2.2 Experimental project

We measured the replicative fitness of wild-type PR8 (WT) and an immunodominant CD8 T cell-escaping mutant (MT), PR8 NP-N370Q, in mice under different settings of cellular immunity. The selection coefficient of the MT compared to the WT measures fitness change associated with the NP-N370Q mutation. Consistent with *in vitro* viral growth characteristics, the mutant virus does not show a significant fitness defect in the absence of pre-existing cellular immunity during the early stage of infection, as demonstrated by *in vivo* fitness measures in naïve B6 mice. During the late stage of infection, the MT acquired a slight advantage (around 10%), likely due to the increasing selection pressure from newly-induced effector $NP_{366\text{-}374}$-specific CD8 T cells that recognize only the wild-type virus (Smith et al. 2018). The selective advantage became significantly larger within intranasally x31-primed mice (16-24% in B6 and 15-17% in B6 x BALB/c F1), where the selection pressure of $NP_{366\text{-}374}$-specific memory CD8 T cells is present from the onset of infection. Summarizing the data from these scenarios showed the mutant virus has a 15-25% advantage in replication by escaping the immunodominant $NP_{366\text{-}374}$-specific memory CD8 T cell response, much smaller than the advantage it would acquire by escaping neutralizing antibodies, which block viral entry and thus prevent replication all together.

Why is the selective advantage conferred by escaping $NP_{366\text{-}374}$-specific memory CD8 T cells relatively small? A likely mechanism is that the ability to evade the $NP_{366\text{-}374}$-specific response is compensated by memory CD8 T cells specific for other immunodominant epitopes, namely, D^b-$PA_{224\text{-}233}$ [252] and K^b-$PB1_{703\text{-}711}$ [253]. This compensation was evident in B6 mice primed with x31 and challenged with either WT or MT alone, where expansion of CD8 T cells specific for

the additional immunodominant FluPA and FluPB1 epitopes was evident only in mice challenged with mutant virus. Overall, these data suggest that the breadth of epitope recognition by CD8 T cells enables the cellular immunity to at least partially compensate for the loss of a single immunodominant epitope. However, previous studies have shown the presentation of epitopes vary across different cell types [248], and that may also impact the advantage a virus may gain through an escaping mutation.

Interestingly, in MHC heterozygous (H-$2^{b/d}$) F1 mice, the advantage acquired by the MT started with a lower value, reached the same level as B6 mice on day 4, and leveled off thereafter. We conjectured that this difference stems from two mechanisms. First, $NP_{366\text{-}374}$-specific CD8 T cells contribute to a majority of the cellular immune response to H-2^b-restricted epitopes in the lungs during a secondary influenza infection [254, 303], and its loss may not be fully compensated by the additional H-2^d-restricted CD8 T cells. This would result in the similarity observed between B6 and F1 on day 4 after PR8 challenge. Second, the cellular immune response can be affected by the differential presentation of epitopes by distinct cell types. For instance, the $PA_{224\text{-}233}$ epitope is highly expressed by dendritic cells but weakly expressed by epithelial cells, resulting in delayed viral clearance after influenza infection [248]. Potentially, the H-2^d-restricted CD8 T cells may have different dynamics compared to the $NP_{366\text{-}374}$-specific CD8 T cells. This would result in the differences observed before day 2 and after day 6. Studies elucidating the dynamics of H-2^d-restricted CD8 T cells in BALB/c and CB6F1 mice during a recall response would be needed to test this argument. There is also a possibility that heterozygosity of immune-related genes in F1 mice has a broader impact on antiviral immunity beyond MHC restriction, although this would be anticipated to affect both the WT and MT viruses equally.

Given the strong interest in designing vaccines that can promote effective cellular immunity against influenza, it is important to understand how different scenarios of T cell memory can

impact the selective advantage of CD8 T cell-escaping mutations. Thus, we directly assessed the impact of two practical scenarios on viral fitness: a lack of lung-resident CD8 T cell memory, and a memory CD8 T cell response focused on a single epitope. Intramuscular (i.m.) injection is the conventional influenza vaccination route in humans; however, previous studies have shown the i.m. immunization route does not induce lung T_{RM}, which are essential for optimal cellular immunity against heterologous influenza challenge [204, 304, 208]. Our data reveal that lung CD8 T_{RM} are the primary source of selection pressure during the early stage of influenza virus infection, as the MT virus outgrew the WT virus in the i.n. x31-primed mice at all times over the course of infection, but failed to do so until day 6 post-challenge in the i.m. x31-primed mice. In contrast, when the influenza-specific memory CD8 T cell pool is directed against a single epitope as seen in AdNP-immunized mice, the mutant acquired a substantially large advantage during both the early stage (40%) or across the whole infection course (74%). These two cases demonstrate that, while inducing a population of lung CD8 TRM specific for one single epitope may be effective in controlling early viral replication [247], it would also impose a strong selection pressure on the virus.

4.3 Framework of measuring viral fitness

The conventional definition of fitness is the amount of progeny produced by an organism. Since the maintenance of a virus species requires within-host replication and between-host transmission, the term 'amount of progeny produced' could be defined under either contexts. Wargo and Kurath [305] proposed that, to measure viral fitness, one should consider three aspects: replicative fitness, transmission fitness, and epidemiologic fitness. Replicative fitness quantifies the competence of replication and can be readily tested *in vitro* or *in vivo*. Transmission fitness is a measure of the ability to spread to susceptible hosts within the population. A good index is the basic reproduction

number, R_0, which is the average number of secondary cases an index case gives rise to in a well-mixed, naïve population [306]. Epidemiologic fitness is the potential of the virus to become the dominant strain in the virus pool.

This dissertation aims to link the three measures of viral fitness in terms of escaping from cellular immunity. We first quantified the replicative fitness using the area under the curve (AUC) of *in vivo* viral growth and related it to transmission fitness, which then served as an essential parameter that dictates the dynamics of a CD8 T cell-escaping variant in the population genetics and the ordinary differential equation models. As estimated in section 3.1.7 (Table 3.3), the selective advantages of escape from $NP_{366\text{-}374}$-specific CD8 T cells are 16-24% and 15-17% within intranasally x31-primed B6 and F1 mice, respectively. Given the estimated heterozygous effect (0.7-0.9) of this epitope and the negligible fitness defect, it requires 3 to 10 years to reach 50% of prevalence in a host population with a cognate allele frequency of 0.2. Stated another way, the CD8 T cell-escaping mutant will require at least several years to sweep in and replace the wild-type virus, if the dynamics depends entirely on positive selection. It is noteworthy that the abovementioned scenario is likely to be the most favorable. In reality, human influenza-specific CD8 T cells may react to a broader set of epitopes; as such, the advantage of losing one epitope becomes less than what was measured in inbred mice.

A key presumption in these models is the link function between replicative and transmission fitness. As different link functions may end up with different estimates of transmission fitness, this limits the quantitative inference from integrating experimental data with the model. Although a sigmoidal function models the threshold and saturation commonly observed in biological systems, it includes more parameters that haven't been properly measured. In section 3.3, we used the estimates from [301], which related the amount of nasal discharge, a surrogate marker of transmissibility, to the viral titer in the discharge using the Hill function, with an important caveat

that nasal discharge may not be a sufficient measure for transmissibility. By running a sensitivity analysis based on Hill function, we show that the choice of parameters does change the estimates of selection coefficient and, thus, the quantitative inference of the time required for a CD8 T cell-escaping mutant virus to invade the circulating viral population.

Although there are robust animal transmission models of influenza available, such as ferrets and guinea pigs, their influenza-specific CD8 T cell profiles are not well characterized; in contrast, mice are an excellent model to investigate T cell immunity, but they do not transmit influenza virus. This dissertation wants to highlight the need for developing an influenza transmission animal model with well-characterized influenza-specific T cell profiles for linking the measurements of replicative, transmission, and epidemiologic viral fitness.

4.4 Implications on T cell-based vaccines

An ideal vaccine should be safe and should generate long-lasting protection. The current inactivated influenza vaccines (IIVs) are formulated with purified viral antigens [307], which induce IAV-specific antibodies but poorly stimulate IAV-specific CD8 T cells. When the antigenicity of the circulating and vaccine strains match, IIVs provide excellent protection against infection with minimal safety concern. The major drawback is that high selection pressure drives the virus to quickly evolve and hence limits the scope of protection.

This dissertation sheds light on the use of T cell-based vaccines and the consequences of vaccination. Given the assumption that transmission fitness is proportional to log-transformed replicative fitness, the selective advantage of one CD8 T cell-escaping mutation is limited to under 25% when the infection occurs through the natural route (i.e., airborne transmission), even if the mutant escapes from immunodominant CD8 T cells in MHC-homozygous individuals. This indicates that a vaccine inducing lung CD8 T_{RM} of sufficient breadth to recognize multiple epitopes could pro-

vide effective protection by controlling early viral replication and lessening immunopathology with limited selection pressure against a single epitope.

Many future studies are required for developing an effective T cell-based vaccine. As of now, a T cell-based vaccine may be formulated with attenuated viruses or a pool of selected epitopes administered with an adjuvant. Indeed, live-attenuated influenza vaccine (LAIV) has been approved in some countries for use in humans (e.g., FluMist) or poultry. Importantly, Zhou et al. [308] reported that a cold-adapted LAIV strain may reacquire pathogenicity when being serially passaged at increasing temperatures in vitro, raising the concern of reversion.

Another issue stems from the fact that protection of lung CD8 T_{RM} induced by IAV infection is transient. In mice, the number of lung CD8 T_{RM} decays within 6 to 7 months after initial infection, with a corresponding loss in protective cellular immunity following heterosubtypic influenza challenge [246]. In humans, a systematic review from Chung et al. [309] revealed the effectiveness[†] of quadrivalent LAIV against influenza was only 20% (95% C.I.: [-6, 39]), much less than that of the IIV (67%, 95% C.I.: [62, 72]). A potential dilemma may arise from the balance between the pursuit of long-lasting protection and the risk of increasing selection pressure on the virus, and modeling can serve as a useful tool to help researchers and policy makers to formulate the vaccine as well as the vaccination program that provide the optimal protection to the human population.

[†]The VE was measured by $100\% \times (1 - \text{odds ratio})$ in Chung et al. 2019.

Bibliography

[1] F. Krammer, G. J. D. Smith, R. A. M. Fouchier, M. Peiris, K. Kedzierska, P. C. Doherty, P. Palese, M. L. Shaw, J. Treanor, R. G. Webster, and A. Garcia-Sastre. Influenza. *Nat Rev Dis Primers*, 4(1):3, 2018.

[2] N. J. Cox and K. Subbarao. Global epidemiology of influenza: past and present. *Annu Rev Med*, 51:407–21, 2000.

[3] P. C. Doherty. *Pandemics*. What everyone needs to know. Oxford University Press, Oxford, 2013.

[4] N. J. Cox, S. C. Trock, and S. A. Burke. Pandemic preparedness and the influenza risk assessment tool (irat). *Curr Top Microbiol Immunol*, 385:119–36, 2014.

[5] T. Noda. Native morphology of influenza virions. *Front Microbiol*, 2:269, 2011.

[6] R. A. Lamb and P. W. Choppin. The gene structure and replication of influenza virus. *Annu Rev Biochem*, 52:467–506, 1983.

[7] R. A. Medina and A. Garcia-Sastre. Influenza a viruses: new research developments. *Nat Rev Microbiol*, 9(8):590–603, 2011.

[8] R. A. Lamb, P. R. Etkind, and P. W. Choppin. Evidence for a ninth influenza viral polypeptide. *Virology*, 91(1):60–78, 1978.

[9] R. E. O'Neill, J. Talon, and P. Palese. The influenza virus nep (ns2 protein) mediates the nuclear export of viral ribonucleoproteins. *EMBO J*, 17(1):288–96, 1998.

[10] R. A. Lamb and P. W. Choppin. Identification of a second protein (m2) encoded by rna segment 7 of influenza virus. *Virology*, 112(2):729–37, 1981.

[11] W. Chen, P. A. Calvo, D. Malide, J. Gibbs, U. Schubert, I. Bacik, S. Basta, R. O'Neill, J. Schickli, P. Palese, P. Henklein, J. R. Bennink, and J. W. Yewdell. A novel influenza a virus mitochondrial protein that induces cell death. *Nat Med*, 7(12):1306–12, 2001.

[12] N. M. Bouvier and P. Palese. The biology of influenza viruses. *Vaccine*, 26 Suppl 4:D49–53, 2008.

[13] I. T. Schulze. The structure of influenza virus. ii. a model based on the morphology and composition of subviral particles. *Virology*, 47(1):181–96, 1972.

[14] W. Weis, J. H. Brown, S. Cusack, J. C. Paulson, J. J. Skehel, and D. C. Wiley. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature*, 333(6172):426–31, 1988.

[15] Y. Suzuki, T. Ito, T. Suzuki, Jr. Holland, R. E., T. M. Chambers, M. Kiso, H. Ishida, and Y. Kawaoka. Sialic acid species as a determinant of the host range of influenza a viruses. *J Virol*, 74(24):11825–31, 2000.

[16] A. Yoshimura, K. Kuroda, K. Kawasaki, S. Yamashina, T. Maeda, and S. Ohnishi. Infectious cell entry mechanism of influenza virus. *J Virol*, 43(1):284–93, 1982.

[17] J. J. Skehel, P. M. Bayley, E. B. Brown, S. R. Martin, M. D. Waterfield, J. M. White, I. A. Wilson, and D. C. Wiley. Changes in the conformation of influenza virus hemagglutinin at the ph optimum of virus-mediated membrane fusion. *Proc Natl Acad Sci U S A*, 79(4):968–72, 1982.

[18] R. W. Doms, A. Helenius, and J. White. Membrane fusion activity of the influenza virus hemagglutinin. the low ph-induced conformational change. *J Biol Chem*, 260(5):2973–81, 1985.

[19] X. Han, J. H. Bushweller, D. S. Cafiso, and L. K. Tamm. Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. *Nat Struct Biol*, 8(8):715–20, 2001.

[20] L. H. Pinto, L. J. Holsinger, and R. A. Lamb. Influenza virus m2 protein has ion channel activity. *Cell*, 69(3):517–28, 1992.

[21] A. Yoshimura and S. Ohnishi. Uncoating of influenza virus in endosomes. *J Virol*, 51(2):497–504, 1984.

[22] K. Martin and A. Helenius. Transport of incoming influenza virus nucleocapsids into the nucleus. *J Virol*, 65(1):232–44, 1991.

[23] E. Fodor. The rna polymerase of influenza a virus: mechanisms of viral transcription and replication. *Acta Virol*, 57(2):113–22, 2013.

[24] C. S. Copeland, K. P. Zimmer, K. R. Wagner, G. A. Healey, I. Mellman, and A. Helenius. Folding, trimerization, and transport are sequential events in the biogenesis of influenza virus hemagglutinin. *Cell*, 53(2):197–209, 1988.

[25] T. Saito, G. Taylor, and R. G. Webster. Steps in maturation of influenza a virus neuraminidase. *J Virol*, 69(8):5011–7, 1995.

[26] J. D. Hull, R. Gilmore, and R. A. Lamb. Integration of a small integral membrane protein, m2, of influenza virus into the endoplasmic reticulum: analysis of the internal signal-anchor domain of a protein with an ectoplasmic nh2 terminus. *J Cell Biol*, 106(5):1489–98, 1988.

[27] S. Boulo, H. Akarsu, R. W. Ruigrok, and F. Baudin. Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. *Virus Res*, 124(1-2):12–21, 2007.

[28] K. Klumpp, R. W. Ruigrok, and F. Baudin. Roles of the influenza virus polymerase and nucleoprotein in forming a functional rnp structure. *EMBO J*, 16(6):1248–57, 1997.

[29] D. P. Nayak, E. K. Hui, and S. Barman. Assembly and budding of influenza virus. *Virus Res*, 106(2):147–65, 2004.

[30] J. S. Rossman and R. A. Lamb. Influenza virus assembly and budding. *Virology*, 411(2):229–36, 2011.

[31] S. Duffy, L. A. Shackelton, and E. C. Holmes. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet*, 9(4):267–76, 2008.

[32] R. Sanjuan, M. R. Nebot, N. Chirico, L. M. Mansky, and R. Belshaw. Viral mutation rates. *J Virol*, 84(19):9733–48, 2010.

[33] M. D. Pauly, M. C. Procario, and A. S. Lauring. A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza a viruses. *Elife*, 6, 2017.

[34] J. Steel and A. C. Lowen. Influenza a virus reassortment. *Curr Top Microbiol Immunol*, 385:377–401, 2014.

[35] R. G. Webster and E. A. Govorkova. Continuing challenges in influenza. *Ann N Y Acad Sci*, 1323:115–39, 2014.

[36] A. S. Monto and K. Fukuda. Lessons from influenza pandemics of the last 100 years. *Clin Infect Dis*, 70(5):951–957, 2020.

[37] S. Tong, X. Zhu, Y. Li, M. Shi, J. Zhang, M. Bourgeois, H. Yang, X. Chen, S. Recuenco, J. Gomez, L. M. Chen, A. Johnson, Y. Tao, C. Dreyfus, W. Yu, R. McBride, P. J. Carney, A. T. Gilbert, J. Chang, Z. Guo, C. T. Davis, J. C. Paulson, J. Stevens, C. E. Rupprecht, E. C. Holmes, I. A. Wilson, and R. O. Donis. New world bats harbor diverse influenza a viruses. *PLoS Pathog*, 9(10):e1003657, 2013.

[38] A. Mehle. Unusual influenza a viruses in bats. *Viruses*, 6(9):3438–49, 2014.

[39] A. Cullinane and J. R. Newton. Equine influenza–a global perspective. *Vet Microbiol*, 167(1-2):205–14, 2013.

[40] G. N. Rogers and J. C. Paulson. Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the h3 hemagglutinin based on species of origin. *Virology*, 127(2):361–73, 1983.

[41] G. N. Rogers and B. L. D'Souza. Receptor binding properties of human and animal h1 influenza virus isolates. *Virology*, 173(1):317–22, 1989.

[42] R. J. Connor, Y. Kawaoka, R. G. Webster, and J. C. Paulson. Receptor specificity in human, avian, and equine h2 and h3 influenza virus isolates. *Virology*, 205(1):17–23, 1994.

[43] B. R. Murphy, V. S. Hinshaw, D. L. Sly, W. T. London, N. T. Hosier, F. T. Wood, R. G. Webster, and R. M. Chanock. Virulence of avian influenza a viruses for squirrel monkeys. *Infect Immun*, 37(3):1119–26, 1982.

[44] J. M. Nicholls, A. J. Bourne, H. Chen, Y. Guan, and J. S. Peiris. Sialic acid receptor detection in the human respiratory tract: evidence for widespread distribution of potential binding sites for human and avian influenza viruses. *Respir Res*, 8:73, 2007.

[45] H. Stern and K. C. Tippett. Primary isolation of influenza viruses at 33 degrees c. *Lancet*, 1(7294):1301–2, 1963.

[46] T. Ito, J. N. Couceiro, S. Kelm, L. G. Baum, S. Krauss, M. R. Castrucci, I. Donatelli, H. Kida, J. C. Paulson, R. G. Webster, and Y. Kawaoka. Molecular basis for the generation in pigs of influenza a viruses with pandemic potential. *J Virol*, 72(9):7367–73, 1998.

[47] B. Kimble, G. R. Nieto, and D. R. Perez. Characterization of influenza virus sialic acid receptors in minor poultry species. *Virol J*, 7:365, 2010.

[48] R. K. Nelli, S. V. Kuchipudi, G. A. White, B. B. Perez, S. P. Dunham, and K. C. Chang. Comparative distribution of human and avian type sialic acid influenza receptors in the pig. *BMC Vet Res*, 6:4, 2010.

[49] R. G. Webster, W. J. Bean, O. T. Gorman, T. M. Chambers, and Y. Kawaoka. Evolution and ecology of influenza a viruses. *Microbiol Rev*, 56(1):152–79, 1992.

[50] S. Lai, Y. Qin, B. J. Cowling, X. Ren, N. A. Wardrop, M. Gilbert, T. K. Tsang, P. Wu, L. Feng, H. Jiang, Z. Peng, J. Zheng, Q. Liao, S. Li, P. W. Horby, J. J. Farrar, G. F. Gao, A. J. Tatem, and H. Yu. Global epidemiology of avian influenza a h5n1 virus infection in humans, 1997-2015: a systematic review of individual case data. *Lancet Infect Dis*, 16(7):e108–e118, 2016.

[51] H. Chen, G. J. Smith, S. Y. Zhang, K. Qin, J. Wang, K. S. Li, R. G. Webster, J. S. Peiris, and Y. Guan. Avian flu: H5n1 virus outbreak in migratory waterfowl. *Nature*, 436(7048):191–2, 2005.

[52] D. J. Prosser, P. Cui, J. Y. Takekawa, M. Tang, Y. Hou, B. M. Collins, B. Yan, N. J. Hill, T. Li, Y. Li, F. Lei, S. Guo, Z. Xing, Y. He, Y. Zhou, D. C. Douglas, W. M. Perry, and S. H. Newman. Wild bird migration across the qinghai-tibetan plateau: a transmission route for highly pathogenic h5n1. *PLoS One*, 6(3):e17622, 2011.

[53] FAO. *Lessons from HPAI : a technical stocktaking of outputs, outcomes, best practices and lessons learned from the fight against highly pathogenic avian influenza in Asia, 2005-2011.* FAO Animal Production and Health Paper. Rome, Italy, 2013.

[54] M. A. Jhung, D. I. Nelson, Control Centers for Disease, and Prevention. Outbreaks of avian influenza a (h5n2), (h5n8), and (h5n1) among birds–united states, december 2014-january 2015. *MMWR Morb Mortal Wkly Rep*, 64(4):111, 2015.

[55] R. Gao, B. Cao, Y. Hu, Z. Feng, D. Wang, W. Hu, J. Chen, Z. Jie, H. Qiu, K. Xu, X. Xu, H. Lu, W. Zhu, Z. Gao, N. Xiang, Y. Shen, Z. He, Y. Gu, Z. Zhang, Y. Yang, X. Zhao, L. Zhou, X. Li, S. Zou, Y. Zhang, X. Li, L. Yang, J. Guo, J. Dong, Q. Li, L. Dong, Y. Zhu, T. Bai, S. Wang, P. Hao, W. Yang, Y. Zhang, J. Han, H. Yu, D. Li, G. F. Gao, G. Wu, Y. Wang, Z. Yuan, and Y. Shu. Human infection with a novel avian-origin influenza a (h7n9) virus. *N Engl J Med*, 368(20):1888–97, 2013.

[56] L. Yang, W. Zhu, X. Li, M. Chen, J. Wu, P. Yu, S. Qi, Y. Huang, W. Shi, J. Dong, X. Zhao, W. Huang, Z. Li, X. Zeng, H. Bo, T. Chen, W. Chen, J. Liu, Y. Zhang, Z. Liang, W. Shi, Y. Shu, and D. Wang. Genesis and spread of newly emerged highly pathogenic h7n9 avian viruses in mainland china. *J Virol*, 91(23), 2017.

[57] X. Wang, H. Jiang, P. Wu, T. M. Uyeki, L. Feng, S. Lai, L. Wang, X. Huo, K. Xu, E. Chen, X. Wang, J. He, M. Kang, R. Zhang, J. Zhang, J. Wu, S. Hu, H. Zhang, X. Liu, W. Fu, J. Ou, S. Wu, Y. Qin, Z. Zhang, Y. Shi, J. Zhang, J. Artois, V. J. Fang, H. Zhu, Y. Guan, M. Gilbert, P. W. Horby, G. M. Leung, G. F. Gao, B. J. Cowling, and H. Yu. Epidemiology of avian influenza a h7n9 virus in human beings across five epidemics in mainland china, 2013-17: an epidemiological study of laboratory-confirmed case series. *Lancet Infect Dis*, 17(8):822–832, 2017.

[58] Q. Li, L. Zhou, M. Zhou, Z. Chen, F. Li, H. Wu, N. Xiang, E. Chen, F. Tang, D. Wang, L. Meng, Z. Hong, W. Tu, Y. Cao, L. Li, F. Ding, B. Liu, M. Wang, R. Xie, R. Gao, X. Li, T. Bai, S. Zou, J. He, J. Hu, Y. Xu, C. Chai, S. Wang, Y. Gao, L. Jin, Y. Zhang, H. Luo, H. Yu, J. He, Q. Li, X. Wang, L. Gao, X. Pang, G. Liu, Y. Yan, H. Yuan, Y. Shu, W. Yang, Y. Wang, F. Wu, T. M. Uyeki, and Z. Feng. Epidemiology of human infections with avian influenza a(h7n9) virus in china. *N Engl J Med*, 370(6):520–32, 2014.

[59] M. J. Ma, Y. Yang, and L. Q. Fang. Highly pathogenic avian h7n9 influenza viruses: Recent challenges. *Trends Microbiol*, 27(2):93–95, 2019.

[60] WHO: Influenza (Seasonal). Last updated: 2018/11/06.

[61] CDC: Types of Influenza Viruses. Last updated: 2019/11/18.

[62] C. Paules and K. Subbarao. Influenza. *Lancet*, 390(10095):697–708, 2017.

[63] L. Simonsen. The global impact of influenza on morbidity and mortality. *Vaccine*, 17 Suppl 1:S3–10, 1999.

[64] C. Viboud, W. J. Alonso, and L. Simonsen. Influenza in tropical regions. *PLoS Med*, 3(4):e89, 2006.

[65] CDC: Flu Symptoms & Complications. Last updated: 2019/09/18.

[66] N. Lee, K. W. Choi, P. K. Chan, D. S. Hui, G. C. Lui, B. C. Wong, R. Y. Wong, W. Y. Sin, W. M. Hui, K. L. Ngai, C. S. Cockram, R. W. Lai, and J. J. Sung. Outcomes of adults hospitalised with severe influenza. *Thorax*, 65(6):510–5, 2010.

[67] A. D. Iuliano, K. M. Roguski, H. H. Chang, D. J. Muscatello, R. Palekar, S. Tempia, C. Cohen, J. M. Gran, D. Schanzer, B. J. Cowling, P. Wu, J. Kyncl, L. W. Ang, M. Park, M. Redlberger-Fritz, H. Yu, L. Espenhain, A. Krishnan, G. Emukule, L. van Asten, S. Pereira da Silva, S. Aungkulanon, U. Buchholz, M. A. Widdowson, J. S. Bresee, and Network Global Seasonal Influenza-associated Mortality Collaborator. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet*, 391(10127):1285–1300, 2018.

[68] CDC: Disease Burden of Influenza. Last updated: 2020/04/17.

[69] David Quammen. *Spillover : animal infections and the next human pandemic*. W.W. Norton & Co., New York, 1st edition, 2012.

[70] A. Lowen and P. Palese. Transmission of influenza virus in temperate zones is predominantly by aerosol, in the tropics by contact: a hypothesis. *PLoS Curr*, 1:RRN1002, 2009.

[71] R. Tellier. Aerosol transmission of influenza a virus: a review of new studies. *J R Soc Interface*, 6 Suppl 6:S783–90, 2009.

[72] R. Tellier, Y. Li, B. J. Cowling, and J. W. Tang. Recognition of aerosol transmission of infectious agents: a commentary. *BMC Infect Dis*, 19(1):101, 2019.

[73] F. G. Hayden, A. D. Osterhaus, J. J. Treanor, D. M. Fleming, F. Y. Aoki, K. G. Nicholson, A. M. Bohnen, H. M. Hirst, O. Keene, and K. Wightman. Efficacy and safety of the neuraminidase inhibitor zanamivir in the treatment of influenzavirus infections. gg167 influenza study group. *N Engl J Med*, 337(13):874–80, 1997.

[74] K. G. Nicholson, F. Y. Aoki, A. D. Osterhaus, S. Trottier, O. Carewicz, C. H. Mercier, A. Rode, N. Kinnersley, and P. Ward. Efficacy and safety of oseltamivir in treatment of acute influenza: a randomised controlled trial. neuraminidase inhibitor flu treatment investigator group. *Lancet*, 355(9218):1845–50, 2000.

[75] J. J. Treanor, F. G. Hayden, P. S. Vrooman, R. Barbarash, R. Bettis, D. Riff, S. Singh, N. Kinnersley, P. Ward, and R. G. Mills. Efficacy and safety of the oral neuraminidase inhibitor oseltamivir in treating acute influenza: a randomized controlled trial. us oral neuraminidase study group. *JAMA*, 283(8):1016–24, 2000.

[76] S. H. Hauge, S. Dudman, K. Borgen, A. Lackenby, and O. Hungnes. Oseltamivir-resistant influenza viruses a (h1n1), norway, 2007-08. *Emerg Infect Dis*, 15(2):155–62, 2009.

[77] A. Meijer, A. Lackenby, O. Hungnes, B. Lina, S. van-der Werf, B. Schweiger, M. Opp, J. Paget, J. van-de Kassteele, A. Hay, M. Zambon, and Scheme European Influenza Surveillance. Oseltamivir-resistant influenza virus a (h1n1), europe, 2007-08 season. *Emerg Infect Dis*, 15(4):552–60, 2009.

[78] T. Baranovich, R. Saito, Y. Suzuki, H. Zaraket, C. Dapat, I. Caperig-Dapat, T. Oguma, II Shabana, T. Saito, H. Suzuki, and Group Japanese Influenza Collaborative Study. Emergence of h274y oseltamivir-resistant a(h1n1) influenza viruses in japan during the 2008-2009 season. *J Clin Virol*, 47(1):23–8, 2010.

[79] L. Simonsen, C. Viboud, B. T. Grenfell, J. Dushoff, L. Jennings, M. Smit, C. Macken, M. Hata, J. Gog, M. A. Miller, and E. C. Holmes. The genesis and spread of reassortment human influenza a/h3n2 viruses conferring adamantane resistance. *Mol Biol Evol*, 24(8):1811–20, 2007.

[80] M. I. Nelson, L. Simonsen, C. Viboud, M. A. Miller, and E. C. Holmes. The origin and global emergence of adamantane resistant a/h3n2 influenza viruses. *Virology*, 388(2):270–8, 2009.

[81] Y. Furuse, A. Suzuki, and H. Oshitani. Large-scale sequence analysis of m gene of influenza a viruses from different species: mechanisms for emergence and spread of amantadine resistance. *Antimicrob Agents Chemother*, 53(10):4457–63, 2009.

[82] J. A. Hoffmann, F. C. Kafatos, C. A. Janeway, and R. A. Ezekowitz. Phylogenetic perspectives in innate immunity. *Science*, 284(5418):1313–8, 1999.

[83] L. C. Smith and E. H. Davidson. The echinoderm immune system. characters shared with vertebrate immune systems and characters arising later in deuterostome phylogeny. *Ann N Y Acad Sci*, 712:213–26, 1994.

[84] A. L. Hughes and M. Yeager. Molecular evolution of the vertebrate immune system. *Bioessays*, 19(9):777–86, 1997.

[85] K. Takeda, T. Kaisho, and S. Akira. Toll-like receptors. *Annu Rev Immunol*, 21:335–76, 2003.

[86] T. Kawasaki and T. Kawai. Toll-like receptor signaling pathways. *Front Immunol*, 5:461, 2014.

[87] T. D. Kanneganti, M. Lamkanfi, and G. Nunez. Intracellular nod-like receptors in host defense and disease. *Immunity*, 27(4):549–59, 2007.

[88] J. M. Wilmanski, T. Petnicki-Ocwieja, and K. S. Kobayashi. Nlr proteins: integral members of innate immunity and mediators of inflammatory diseases. *J Leukoc Biol*, 83(1):13–30, 2008.

[89] A. Pichlmair, O. Schulz, C. P. Tan, T. I. Naslund, P. Liljestrom, F. Weber, and C. Reis e Sousa. Rig-i-mediated antiviral responses to single-stranded rna bearing 5'-phosphates. *Science*, 314(5801):997–1001, 2006.

[90] Y. M. Loo and Jr. Gale, M. Immune signaling by rig-i-like receptors. *Immunity*, 34(5):680–92, 2011.

[91] A. Isaacs, J. Lindenmann, and R. C. Valentine. Virus interference. ii. some properties of interferon. *Proc R Soc Lond B Biol Sci*, 147(927):268–73, 1957.

[92] C. Le Page, P. Genin, M. G. Baines, and J. Hiscott. Interferon activation and innate immunity. *Rev Immunogenet*, 2(3):374–86, 2000.

[93] A. K. Perry, G. Chen, D. Zheng, H. Tang, and G. Cheng. The host type i interferon response to viral and bacterial infections. *Cell Res*, 15(6):407–22, 2005.

[94] C. Gabay, C. Lamacchia, and G. Palmer. Il-1 pathways in inflammation and human diseases. *Nat Rev Rheumatol*, 6(4):232–41, 2010.

[95] T. Tanaka, M. Narazaki, and T. Kishimoto. Il-6 in inflammation, immunity, and disease. *Cold Spring Harb Perspect Biol*, 6(10):a016295, 2014.

[96] R. C. Russo, C. C. Garcia, M. M. Teixeira, and F. A. Amaral. The cxcl8/il-8 chemokine family and its receptors in inflammatory diseases. *Expert Rev Clin Immunol*, 10(5):593–619, 2014.

[97] S. Y. Tan and W. Weninger. Neutrophil migration in inflammation: intercellular signal relay and crosstalk. *Curr Opin Immunol*, 44:34–42, 2017.

[98] R. B. Herberman and M. E. Oren. Immune response to gross virus-induced lymphoma. i. kinetics of cytotoxic antibody response. *J Natl Cancer Inst*, 46(2):391–6, 1971.

[99] M. E. Oren, R. B. Herberman, and T. G. Canty. Immune response to gross virus-induced lymphoma. ii. kinetics of the cellular immune response. *J Natl Cancer Inst*, 46(3):621–8, 1971.

[100] M. O. De Landazuri and R. B. Herberman. Immune response to gross virus-induced lymphoma. 3. characteristics of the cellular immune response. *J Natl Cancer Inst*, 49(1):147–54, 1972.

[101] D. L. Farber, M. G. Netea, A. Radbruch, K. Rajewsky, and R. M. Zinkernagel. Immunological memory: lessons from the past and a look to the future. *Nat Rev Immunol*, 16(2):124–8, 2016.

[102] M. D. Cooper, R. D. Peterson, and R. A. Good. Delineation of the thymic and bursal lymphoid systems in the chicken. *Nature*, 205:143–6, 1965.

[103] M. D. Cooper, D. A. Raymond, R. D. Peterson, M. A. South, and R. A. Good. The functions of the thymus system and the bursa system in the chicken. *J Exp Med*, 123(1):75–102, 1966.

[104] J. J. Owen, M. D. Cooper, and M. C. Raff. In vitro generation of b lymphocytes in mouse foetal liver, a mammalian 'bursa equivalent'. *Nature*, 249(455):361–3, 1974.

[105] D. G. Osmond and G. J. Nossal. Differentiation of lymphocytes in mouse bone marrow. i. quantitative radioautographic studies of antiglobulin binding by lymphocytes in bone marrow and lymphoid tissues. *Cell Immunol*, 13(1):117–31, 1974.

[106] D. G. Osmond and G. J. Nossal. Differentiation of lymphocytes in mouse bone marrow. ii. kinetics of maturation and renewal of antiglobulin-binding cells studied by double labeling. *Cell Immunol*, 13(1):132–45, 1974.

[107] R. H. Pain. The molecular weights of the peptide chains of gamma-globulin. *Biochem J*, 88:234–9, 1963.

[108] S. Cohen and C. Milstein. Structure of antibody molecules. *Nature*, 214(5087):449–52 passim, 1967.

[109] G. M. Edelman and B. Benacerraf. On structural and functional relations between antibodies and proteins of the gamma-system. *Proc Natl Acad Sci U S A*, 48:1035–42, 1962.

[110] Jr. Small, P. A., J. E. Kehn, and M. E. Lamm. Polypeptide chains of rabbit gammaglobulin. *Science*, 142(3590):393–4, 1963.

[111] E. S. Vitetta and J. W. Uhr. Igd and b cell differentiation. *Immunol Rev*, 37:50–88, 1977.

[112] R. J. Benschop and J. C. Cambier. B cell development: signal transduction by antigen receptors and their surrogates. *Curr Opin Immunol*, 11(2):143–51, 1999.

[113] R. Ubelhart, E. Hug, M. P. Bach, T. Wossning, M. Duhren-von Minden, A. H. Horn, D. Tsiantoulas, K. Kometani, T. Kurosaki, C. J. Binder, H. Sticht, L. Nitschke, M. Reth, and H. Jumaa. Responsiveness of b cells is regulated by the hinge region of igd. *Nat Immunol*, 16(5):534–43, 2015.

[114] E. M. Fuentes-Panana, G. Bannish, and J. G. Monroe. Basal b-cell receptor signaling in b lymphocytes: mechanisms of regulation and role in positive selection, differentiation, and peripheral survival. *Immunol Rev*, 197:26–40, 2004.

[115] M. R. Ehrenstein and C. A. Notley. The importance of natural igm: scavenger, protector and regulator. *Nat Rev Immunol*, 10(11):778–86, 2010.

[116] G. Vidarsson, G. Dekkers, and T. Rispens. Igg subclasses and allotypes: from structure to effector functions. *Front Immunol*, 5:520, 2014.

[117] M. Knossow, M. Gaudier, A. Douglas, B. Barrere, T. Bizebard, C. Barbey, B. Gigant, and J. J. Skehel. Mechanism of neutralization of influenza virus infectivity by antibodies. *Virology*, 302(2):294–8, 2002.

[118] J. V. Ravetch and S. Bolland. Igg fc receptors. *Annu Rev Immunol*, 19:275–90, 2001.

[119] D. L. Delacroix, C. Dive, J. C. Rambaud, and J. P. Vaerman. Iga subclasses in various secretions and in serum. *Immunology*, 47(2):383–5, 1982.

[120] M. van Egmond, C. A. Damen, A. B. van Spriel, G. Vidarsson, E. van Garderen, and J. G. van de Winkel. Iga and the iga fc receptor. *Trends Immunol*, 22(4):205–11, 2001.

[121] B. Corthesy. Multi-faceted functions of secretory iga at mucosal surfaces. *Front Immunol*, 4:185, 2013.

[122] P. Brandtzaeg. Secretory iga: Designed for anti-microbial defense. *Front Immunol*, 4:222, 2013.

[123] H. J. Gould and B. J. Sutton. Ige in allergy and asthma today. *Nat Rev Immunol*, 8(3):205–17, 2008.

[124] S. J. Galli and M. Tsai. Ige and mast cells in allergic disease. *Nat Med*, 18(5):693–704, 2012.

[125] J. F. Miller and D. Osoba. Current concepts of the immunological function of the thymus. *Physiol Rev*, 47(3):437–520, 1967.

[126] G. F. Mitchell and J. F. Miller. Immunological activity of thymus and thoracic-duct lymphocytes. *Proc Natl Acad Sci U S A*, 59(1):296–303, 1968.

[127] M. Kondo, I. L. Weissman, and K. Akashi. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell*, 91(5):661–72, 1997.

[128] T. Ikawa, H. Kawamoto, S. Fujimoto, and Y. Katsura. Commitment of common t/natural killer (nk) progenitors to unipotent t and nk progenitors in the murine fetal thymus revealed by a single progenitor assay. *J Exp Med*, 190(11):1617–26, 1999.

[129] H. T. Petrie. Cell migration and the control of post-natal t-cell lymphopoiesis in the thymus. *Nat Rev Immunol*, 3(11):859–66, 2003.

[130] E. V. Rothenberg, J. E. Moore, and M. A. Yui. Launching the t-cell-lineage developmental programme. *Nat Rev Immunol*, 8(1):9–21, 2008.

[131] R. M. Zinkernagel and P. C. Doherty. Restriction of in vitro t cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature*, 248(5450):701–2, 1974.

[132] R. M. Zinkernagel and P. C. Doherty. Immunological surveillance against altered self components by sensitised t lymphocytes in lymphocytic choriomeningitis. *Nature*, 251(5475):547–8, 1974.

[133] P. C. Doherty and R. M. Zinkernagel. A biological role for the major histocompatibility antigens. *Lancet*, 1(7922):1406–9, 1975.

[134] F. M. Burnet. *The clonal selection theory of acquired immunity.* The Abraham Flexner lectures, 1958. Vanderbilt University Press, Nashville,, 1959.

[135] G. J. Nossal and J. Lederberg. Antibody production by single cells. *Nature*, 181(4620):1419–20, 1958.

[136] G. J. Nossal and B. L. Pike. Single cell studies on the antibody-forming potential of fractionated, hapten-specific b lymphocytes. *Immunology*, 30(2):189–202, 1976.

[137] G. L. Ada and G. Nossal. The clonal-selection theory. *Sci Am*, 257(2):62–9, 1987.

[138] N. A. Mitchison and S. Pettersson. Does clonal selection occur among t cells? *Ann Immunol (Paris)*, 134D(1):37–45, 1983.

[139] A. K. Kimura and H. Wigzell. Development and function of cytotoxic t lymphocytes (ctl). i. in vivo maturation of ctl precursors in the absence of detectable proliferation results as a normal consequence of alloimmunization. *J Immunol*, 130(5):2056–61, 1983.

[140] F. Denizot, A. Wilson, F. Battye, G. Berke, and K. Shortman. Clonal expansion of t cells: a cytotoxic t-cell response in vivo that involves precursor cell proliferation. *Proc Natl Acad Sci U S A*, 83(16):6089–92, 1986.

[141] P. C. Doherty, W. Allan, M. Eichelberger, and S. R. Carding. Roles of alpha beta and gamma delta t cell subsets in viral immunity. *Annu Rev Immunol*, 10:123–51, 1992.

[142] S. M. Kaech and R. Ahmed. Memory cd8+ t cell differentiation: initial antigen encounter triggers a developmental program in naive cells. *Nat Immunol*, 2(5):415–22, 2001.

[143] E. J. Wherry and R. Ahmed. Memory cd8 t-cell differentiation during viral infection. *J Virol*, 78(11):5535–45, 2004.

[144] J. Zhu, H. Yamane, and W. E. Paul. Differentiation of effector cd4 t cell populations (*). *Annu Rev Immunol*, 28:445–89, 2010.

[145] W. Haas, P. Pereira, and S. Tonegawa. Gamma/delta cells. *Annu Rev Immunol*, 11:637–85, 1993.

[146] Y. H. Chien, R. Jores, and M. P. Crowley. Recognition by gamma/delta t cells. *Annu Rev Immunol*, 14:511–32, 1996.

[147] A. C. Hayday. [gamma][delta] cells: a right time and a right place for a conserved third way of protection. *Annu Rev Immunol*, 18:975–1026, 2000.

[148] M. Blackman, J. Kappler, and P. Marrack. The role of the t cell receptor in positive and negative selection of developing t cells. *Science*, 248(4961):1335–41, 1990.

[149] T. K. Starr, S. C. Jameson, and K. A. Hogquist. Positive and negative selection of t cells. *Annu Rev Immunol*, 21:139–76, 2003.

[150] L. Klein, B. Kyewski, P. M. Allen, and K. A. Hogquist. Positive and negative selection of the t cell repertoire: what thymocytes see (and don't see). *Nat Rev Immunol*, 14(6):377–91, 2014.

[151] T. J. Braciale, L. A. Morrison, M. T. Sweetser, J. Sambrook, M. J. Gething, and V. L. Braciale. Antigen presentation pathways to class i and class ii mhc-restricted t lymphocytes. *Immunol Rev*, 98:95–114, 1987.

[152] E. W. Hewitt. The mhc class i antigen presentation pathway: strategies for viral immune evasion. *Immunology*, 110(2):163–9, 2003.

[153] T. H. Hansen and M. Bouvier. Mhc class i antigen presentation: learning from viral evasion strategies. *Nat Rev Immunol*, 9(7):503–13, 2009.

[154] P. A. Henkart. Lymphocyte-mediated cytotoxicity: two pathways and multiple effector molecules. *Immunity*, 1(5):343–6, 1994.

[155] M. Barry and R. C. Bleackley. Cytotoxic t lymphocytes: all roads lead to death. *Nat Rev Immunol*, 2(6):401–9, 2002.

[156] E. M. Janssen, E. E. Lemmens, N. Gour, R. A. Reboulet, D. R. Green, S. P. Schoenberger, and M. J. Pinkoski. Distinct roles of cytolytic effector molecules for antigen-restricted killing by ctl in vivo. *Immunol Cell Biol*, 88(7):761–5, 2010.

[157] J. Pieters. Mhc class ii restricted antigen presentation. *Curr Opin Immunol*, 9(1):89–96, 1997.

[158] R. D. Campbell and J. Trowsdale. Map of the human mhc. *Immunol Today*, 14(7):349–52, 1993.

[159] R. Horton, L. Wilming, V. Rand, R. C. Lovering, E. A. Bruford, V. K. Khodiyar, M. J. Lush, S. Povey, Jr. Talbot, C. C., M. W. Wright, H. M. Wain, J. Trowsdale, A. Ziegler, and S. Beck. Gene map of the extended human mhc. *Nat Rev Genet*, 5(12):889–99, 2004.

[160] S. G. Marsh, E. D. Albert, W. F. Bodmer, R. E. Bontrop, B. Dupont, H. A. Erlich, M. Fernandez-Vina, D. E. Geraghty, R. Holdsworth, C. K. Hurley, M. Lau, K. W. Lee, B. Mach, M. Maiers, W. R. Mayr, C. R. Muller, P. Parham, E. W. Petersdorf, T. Sasazuki, J. L. Strominger, A. Svejgaard, P. I. Terasaki, J. M. Tiercy, and J. Trowsdale. Nomenclature for factors of the hla system, 2010. *Tissue Antigens*, 75(4):291–455, 2010.

[161] J. Robinson, J. A. Halliwell, J. D. Hayhurst, P. Flicek, P. Parham, and S. G. Marsh. The ipd and imgt/hla database: allele variant databases. *Nucleic Acids Res*, 43(Database issue):D423–31, 2015.

[162] J. Klein. Seeds of time: fifty years ago peter a. gorer discovered the h-2 complex. *Immunogenetics*, 24(6):331–8, 1986.

[163] J. Klein, F. Figueroa, and C. S. David. H-2 haplotypes, genes and antigens: second listing. ii. the h-2 complex. *Immunogenetics*, 17(6):553–96, 1983.

[164] V. H. Engelhard. Structure of peptides associated with class i and class ii mhc molecules. *Annu Rev Immunol*, 12:181–207, 1994.

[165] G. R. Otten, E. Bikoff, R. K. Ribaudo, S. Kozlowski, D. H. Margulies, and R. N. Germain. Peptide and beta 2-microglobulin regulation of cell surface mhc class i conformation and expression. *J Immunol*, 148(12):3723–32, 1992.

[166] D. R. Madden, J. C. Gorga, J. L. Strominger, and D. C. Wiley. The three-dimensional structure of hla-b27 at 2.1 a resolution suggests a general mechanism for tight peptide binding to mhc. *Cell*, 70(6):1035–48, 1992.

[167] H. G. Rammensee. Chemistry of peptides associated with mhc class i and class ii molecules. *Curr Opin Immunol*, 7(1):85–96, 1995.

[168] J. Sidney, B. Peters, N. Frahm, C. Brander, and A. Sette. Hla class i supertypes: a revised and updated classification. *BMC Immunol*, 9:1, 2008.

[169] R. Konig. Interactions between mhc molecules and co-receptors of the tcr. *Curr Opin Immunol*, 14(1):75–83, 2002.

[170] A. de la Hera, U. Muller, C. Olsson, S. Isaaz, and A. Tunnacliffe. Structure of the t cell antigen receptor (tcr): two cd3 epsilon subunits in a functional tcr/cd3 complex. *J Exp Med*, 173(1):7–17, 1991.

[171] I. A. Wilson and K. C. Garcia. T-cell receptor structure and tcr complexes. *Curr Opin Struct Biol*, 7(6):839–48, 1997.

[172] M. S. Kuhns, M. M. Davis, and K. C. Garcia. Deconstructing the form and function of the tcr/cd3 complex. *Immunity*, 24(2):133–9, 2006.

[173] D. Gil, A. G. Schrum, B. Alarcon, and E. Palmer. T cell receptor engagement by peptide-mhc ligands induces a conformational change in the cd3 complex of thymocytes. *J Exp Med*, 201(4):517–22, 2005.

[174] R. M. Risueno, D. Gil, E. Fernandez, F. Sanchez-Madrid, and B. Alarcon. Ligand-induced conformational change in the t-cell receptor associated with productive immune synapses. *Blood*, 106(2):601–8, 2005.

[175] S. Minguet, M. Swamy, B. Alarcon, I. F. Luescher, and W. W. Schamel. Full activation of the t cell receptor requires both clustering and conformational changes at cd3. *Immunity*, 26(1):43–54, 2007.

[176] T. P. Arstila, A. Casrouge, V. Baron, J. Even, J. Kanellopoulos, and P. Kourilsky. A direct estimate of the human alphabeta t cell receptor diversity. *Science*, 286(5441):958–61, 1999.

[177] V. I. Zarnitsyna, B. D. Evavold, L. N. Schoettle, J. N. Blattman, and R. Antia. Estimating the diversity, completeness, and cross-reactivity of the t cell repertoire. *Front Immunol*, 4:485, 2013.

[178] Q. Qi, Y. Liu, Y. Cheng, J. Glanville, D. Zhang, J. Y. Lee, R. A. Olshen, C. M. Weyand, S. D. Boyd, and J. J. Goronzy. Diversity and clonal selection in the human t-cell repertoire. *Proc Natl Acad Sci U S A*, 111(36):13139–44, 2014.

[179] D. J. Laydon, C. R. Bangham, and B. Asquith. Estimating t-cell repertoire diversity: limitations of classical estimators and a new approach. *Philos Trans R Soc Lond B Biol Sci*, 370(1675), 2015.

[180] E. P. Rock, P. R. Sibbald, M. M. Davis, and Y. H. Chien. Cdr3 length in antigen-specific immune receptors. *J Exp Med*, 179(1):323–8, 1994.

[181] P. A. Moss and J. I. Bell. Sequence analysis of the human alpha beta t-cell receptor cdr3 region. *Immunogenetics*, 42(1):10–8, 1995.

[182] C. Brack, M. Hirama, R. Lenhard-Schuller, and S. Tonegawa. A complete immunoglobulin gene is created by somatic recombination. *Cell*, 15(1):1–14, 1978.

[183] O. Bernard, N. Hozumi, and S. Tonegawa. Sequences of mouse immunoglobulin light chain genes before and after somatic changes. *Cell*, 15(4):1133–44, 1978.

[184] S. Tonegawa, C. Brack, N. Hozumi, and V. Pirrotta. Organization of immunoglobulin genes. *Cold Spring Harb Symp Quant Biol*, 42 Pt 2:921–31, 1978.

[185] S. Tonegawa, A. M. Maxam, R. Tizard, O. Bernard, and W. Gilbert. Sequence of a mouse germ-line gene for a variable region of an immunoglobulin light chain. *Proc Natl Acad Sci U S A*, 75(3):1485–9, 1978.

[186] S. Tonegawa. Somatic generation of antibody diversity. *Nature*, 302(5909):575–81, 1983.

[187] F. W. Alt and D. Baltimore. Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three d-jh fusions. *Proc Natl Acad Sci U S A*, 79(13):4118–22, 1982.

[188] J. J. Lafaille, A. DeCloux, M. Bonneville, Y. Takagaki, and S. Tonegawa. Junctional sequences of t cell receptor gamma delta genes: implications for gamma delta t cell lineages and for a novel intermediate of v-(d)-j joining. *Cell*, 59(5):859–70, 1989.

[189] Y. H. Chien, N. R. Gascoigne, J. Kavaler, N. E. Lee, and M. M. Davis. Somatic recombination in a murine t-cell receptor gene. *Nature*, 309(5966):322–6, 1984.

[190] N. R. Gascoigne, Y. Chien, D. M. Becker, J. Kavaler, and M. M. Davis. Genomic organization and sequence of t-cell receptor beta-chain constant- and joining-region genes. *Nature*, 310(5976):387–91, 1984.

[191] M. Malissen, K. Minard, S. Mjolsness, M. Kronenberg, J. Goverman, T. Hunkapiller, M. B. Prystowsky, Y. Yoshikai, F. Fitch, T. W. Mak, and et al. Mouse t cell antigen receptor: structure and organization of constant and joining gene segments encoding the beta polypeptide. *Cell*, 37(3):1101–10, 1984.

[192] C. M. Croce, M. Isobe, A. Palumbo, J. Puck, J. Ming, D. Tweardy, J. Erikson, M. Davis, and G. Rovera. Gene for alpha-chain of human t-cell receptor: location on chromosome 14 region involved in t-cell neoplasms. *Science*, 227(4690):1044–7, 1985.

[193] M. Gellert. Molecular analysis of v(d)j recombination. *Annu Rev Genet*, 26:425–46, 1992.

[194] C. H. Bassing, W. Swat, and F. W. Alt. The mechanism and regulation of chromosomal v(d)j recombination. *Cell*, 109 Suppl:S45–55, 2002.

[195] N. Caccia, M. Kronenberg, D. Saxe, R. Haars, G. A. Bruns, J. Goverman, M. Malissen, H. Willard, Y. Yoshikai, M. Simon, and et al. The t cell receptor beta chain genes are located on chromosome 6 in mice and chromosome 7 in humans. *Cell*, 37(3):1091–9, 1984.

[196] T. J. Powell, T. Strutt, J. Reome, J. A. Hollenbaugh, A. D. Roberts, D. L. Woodland, S. L. Swain, and R. W. Dutton. Priming with cold-adapted influenza a does not prevent infection but elicits long-lived protection against supralethal challenge with heterosubtypic virus. *J Immunol*, 178(2):1030–8, 2007.

[197] S. Liang, K. Mozdzanowska, G. Palladino, and W. Gerhard. Heterosubtypic immunity to influenza type a virus in mice. effector mechanisms and their longevity. *J Immunol*, 152(4):1653–61, 1994.

[198] H. H. Nguyen, Z. Moldoveanu, M. J. Novak, F. W. van Ginkel, E. Ban, H. Kiyono, J. R. McGhee, and J. Mestecky. Heterosubtypic immunity to lethal influenza a virus infection is associated with virus-specific cd8(+) cytotoxic t lymphocyte responses induced in mucosa-associated tissues. *Virology*, 254(1):50–60, 1999.

[199] K. A. Benton, J. A. Misplon, C. Y. Lo, R. R. Brutkiewicz, S. A. Prasad, and S. L. Epstein. Heterosubtypic immunity to influenza a virus in mice lacking iga, all ig, nkt cells, or gamma delta t cells. *J Immunol*, 166(12):7437–45, 2001.

[200] A. C. Boon, G. de Mutsert, D. van Baarle, D. J. Smith, A. S. Lapedes, R. A. Fouchier, K. Sintnicolaas, A. D. Osterhaus, and G. F. Rimmelzwaan. Recognition of homo- and heterosubtypic variants of influenza a viruses by human cd8+ t lymphocytes. *J Immunol*, 172(4):2453–60, 2004.

[201] F. S. Quan, C. Huang, R. W. Compans, and S. M. Kang. Virus-like particle vaccine induces protective immunity against homologous and heterologous strains of influenza virus. *J Virol*, 81(7):3514–24, 2007.

[202] F. S. Quan, R. W. Compans, H. H. Nguyen, and S. M. Kang. Induction of heterosubtypic immunity to influenza virus by intranasal immunization. *J Virol*, 82(3):1350–9, 2008.

[203] L. Y. Lee, L. A. Ha do, C. Simmons, M. D. de Jong, N. V. Chau, R. Schumacher, Y. C. Peng, A. J. McMichael, J. J. Farrar, G. L. Smith, A. R. Townsend, B. A. Askonas, S. Rowland-Jones, and T. Dong. Memory t cells established by seasonal human influenza a infection cross-react with avian influenza a (h5n1) in healthy individuals. *J Clin Invest*, 118(10):3478–90, 2008.

[204] J. H. Kreijtz, R. Bodewes, J. M. van den Brand, G. de Mutsert, C. Baas, G. van Amerongen, R. A. Fouchier, A. D. Osterhaus, and G. F. Rimmelzwaan. Infection of mice with a human influenza a/h3n2 virus induces protective immunity against lethal infection with influenza a/h5n1 virus. *Vaccine*, 27(36):4983–9, 2009.

[205] Y. Furuya, J. Chan, M. Regner, M. Lobigs, A. Koskinen, T. Kok, J. Manavis, P. Li, A. Mullbacher, and M. Alsharifi. Cytotoxic t cells are the predominant players providing cross-protective immunity induced by gamma-irradiated influenza a viruses. *J Virol*, 84(9):4212–21, 2010.

[206] R. Bodewes, J. H. Kreijtz, M. L. Hillaire, M. M. Geelhoed-Mieras, R. A. Fouchier, A. D. Osterhaus, and G. F. Rimmelzwaan. Vaccination with whole inactivated virus vaccine affects the induction of heterosubtypic immunity against influenza virus a/h5n1 and immunodominance of virus-specific cd8+ t-cell responses in mice. *J Gen Virol*, 91(Pt 7):1743–53, 2010.

[207] S. Sridhar, S. Begom, A. Bermingham, K. Hoschler, W. Adamson, W. Carman, T. Bean, W. Barclay, J. J. Deeks, and A. Lalvani. Cellular immune correlates of protection against symptomatic pandemic influenza. *Nat Med*, 19(10):1305–12, 2013.

[208] S. R. McMaster, J. D. Gabbard, D. G. Koutsonanos, R. W. Compans, R. A. Tripp, S. M. Tompkins, and J. E. Kohlmeier. Memory t cells generated by prior exposure to influenza cross react with the novel h7n9 influenza virus and confer protective heterosubtypic immunity. *PLoS One*, 10(2):e0115725, 2015.

[209] K. D. Zens, J. K. Chen, and D. L. Farber. Vaccine-generated lung tissue-resident memory t cells provide heterosubtypic protection to influenza infection. *JCI Insight*, 1(10), 2016.

[210] C. S. Eickhoff, F. E. Terry, L. Peng, K. A. Meza, I. G. Sakala, D. Van Aartsen, L. Moise, W. D. Martin, J. Schriewer, R. M. Buller, A. S. De Groot, and D. F. Hoft. Highly conserved influenza t cell epitopes induce broadly protective immunity. *Vaccine*, 37(36):5371–5381, 2019.

[211] A. P. Smith, D. J. Moquin, V. Bernhauerova, and A. M. Smith. Influenza virus infection model with density dependence supports biphasic viral decay. *Front Microbiol*, 9:1554, 2018.

[212] J. E. Kohlmeier and D. L. Woodland. Immunity to respiratory viruses. *Annu Rev Immunol*, 27:61–82, 2009.

[213] T. S. Kim, J. Sun, and T. J. Braciale. T cell responses during influenza infection: getting and keeping control. *Trends Immunol*, 32(5):225–31, 2011.

[214] K. Lim, Y. M. Hyun, K. Lambert-Emo, T. Capece, S. Bae, R. Miller, D. J. Topham, and M. Kim. Neutrophil trails guide influenza-specific cd8(+) t cells in the airways. *Science*, 349(6252):aaa4352, 2015.

[215] D. J. Topham, R. A. Tripp, and P. C. Doherty. Cd8+ t cells clear influenza virus by perforin or fas-dependent processes. *J Immunol*, 159(11):5197–200, 1997.

[216] E. L. Brincks, A. Katewa, T. A. Kucaba, T. S. Griffith, and K. L. Legge. Cd8 t cells utilize trail to control influenza virus infection. *J Immunol*, 181(7):4918–25, 2008.

[217] S. P. Cullen and S. J. Martin. Mechanisms of granule-dependent killing. *Cell Death Differ*, 15(2):251–62, 2008.

[218] A. J. McMichael, F. M. Gotch, G. R. Noble, and P. A. Beare. Cytotoxic t-cell immunity to influenza. *N Engl J Med*, 309(1):13–7, 1983.

[219] E. J. Grant, S. M. Quinones-Parra, E. B. Clemens, and K. Kedzierska. Human influenza viruses and cd8(+) t cell responses. *Curr Opin Virol*, 16:132–142, 2016.

[220] J. T. Opferman, B. T. Ober, and P. G. Ashton-Rickardt. Linear differentiation of cytotoxic effectors into memory t lymphocytes. *Science*, 283(5408):1745–8, 1999.

[221] R. Antia, V. V. Ganusov, and R. Ahmed. The role of models in understanding cd8+ t-cell memory. *Nat Rev Immunol*, 5(2):101–11, 2005.

[222] S. N. Mueller, T. Gebhardt, F. R. Carbone, and W. R. Heath. Memory t cell subsets, migration patterns, and tissue residence. *Annu Rev Immunol*, 31:137–61, 2013.

[223] D. L. Farber, N. A. Yudanin, and N. P. Restifo. Human memory t cells: generation, compartmentalization and homeostasis. *Nat Rev Immunol*, 14(1):24–35, 2014.

[224] J. M. Schenkel and D. Masopust. Tissue-resident memory t cells. *Immunity*, 41(6):886–97, 2014.

[225] F. Sallusto, D. Lenig, R. Forster, M. Lipp, and A. Lanzavecchia. Two subsets of memory t lymphocytes with distinct homing potentials and effector functions. *Nature*, 401(6754):708–12, 1999.

[226] U. H. von Andrian and C. R. Mackay. T-cell function and migration. two sides of the same coin. *N Engl J Med*, 343(14):1020–34, 2000.

[227] A. Lanzavecchia and F. Sallusto. Dynamics of t lymphocyte responses: intermediates, effectors, and memory cells. *Science*, 290(5489):92–7, 2000.

[228] F. Sallusto, J. Geginat, and A. Lanzavecchia. Central memory and effector memory t cell subsets: function, generation, and maintenance. *Annu Rev Immunol*, 22:745–63, 2004.

[229] M. A. Williams and M. J. Bevan. Effector and memory ctl differentiation. *Annu Rev Immunol*, 25:171–92, 2007.

[230] D. Masopust, V. Vezys, A. L. Marzo, and L. Lefrancois. Preferential localization of effector memory cells in nonlymphoid tissue. *Science*, 291(5512):2413–7, 2001.

[231] T. Gebhardt, L. M. Wakim, L. Eidsmo, P. C. Reading, W. R. Heath, and F. R. Carbone. Memory t cells in nonlymphoid tissue that provide enhanced local immunity during infection with herpes simplex virus. *Nat Immunol*, 10(5):524–30, 2009.

[232] D. Masopust, D. Choo, V. Vezys, E. J. Wherry, J. Duraiswamy, R. Akondy, J. Wang, K. A. Casey, D. L. Barber, K. S. Kawamura, K. A. Fraser, R. J. Webby, V. Brinkmann, E. C. Butcher, K. A. Newell, and R. Ahmed. Dynamic t cell migration program provides resident memory within intestinal epithelium. *J Exp Med*, 207(3):553–64, 2010.

[233] X. Jiang, R. A. Clark, L. Liu, A. J. Wagers, R. C. Fuhlbrigge, and T. S. Kupper. Skin infection generates non-migratory memory cd8+ t(rm) cells providing global skin immunity. *Nature*, 483(7388):227–31, 2012.

[234] J. M. Schenkel, K. A. Fraser, V. Vezys, and D. Masopust. Sensing and alarm function of resident memory cd8(+) t cells. *Nat Immunol*, 14(5):509–13, 2013.

[235] T. Sathaliyawala, M. Kubota, N. Yudanin, D. Turner, P. Camp, J. J. Thome, K. L. Bickham, H. Lerner, M. Goldstein, M. Sykes, T. Kato, and D. L. Farber. Distribution and compartmentalization of human circulating and tissue-resident memory t cell subsets. *Immunity*, 38(1):187–97, 2013.

[236] J. J. Thome, N. Yudanin, Y. Ohmura, M. Kubota, B. Grinshpun, T. Sathaliyawala, T. Kato, H. Lerner, Y. Shen, and D. L. Farber. Spatial map of human t cell compartmentalization and maintenance over decades of life. *Cell*, 159(4):814–28, 2014.

[237] S. N. Mueller and L. K. Mackay. Tissue-resident memory t cells: local specialists in immune defence. *Nat Rev Immunol*, 16(2):79–89, 2016.

[238] P. C. Rosato, L. K. Beura, and D. Masopust. Tissue resident memory t cells and viral immunity. *Curr Opin Virol*, 22:44–50, 2017.

[239] T. N. Khan, J. L. Mooster, A. M. Kilgore, J. F. Osborn, and J. C. Nolz. Local antigen in nonlymphoid tissue promotes resident memory cd8+ t cell formation during viral infection. *J Exp Med*, 213(6):951–66, 2016.

[240] S. R. McMaster, A. N. Wein, P. R. Dunbar, S. L. Hayward, E. K. Cartwright, T. L. Denning, and J. E. Kohlmeier. Pulmonary antigen encounter regulates the establishment of tissue-resident cd8 memory t cells in the lung airways and parenchyma. *Mucosal Immunol*, 11(4):1071–1078, 2018.

[241] C. N. Skon, J. Y. Lee, K. G. Anderson, D. Masopust, K. A. Hogquist, and S. C. Jameson. Transcriptional downregulation of s1pr1 is required for the establishment of resident memory cd8+ t cells. *Nat Immunol*, 14(12):1285–93, 2013.

[242] L. K. Mackay, A. Rahimpour, J. Z. Ma, N. Collins, A. T. Stock, M. L. Hafon, J. Vega-Ramos, P. Lauzurica, S. N. Mueller, T. Stefanovic, D. C. Tscharke, W. R. Heath, M. Inouye, F. R. Carbone, and T. Gebhardt. The developmental pathway for cd103(+)cd8+ tissue-resident memory t cells of skin. *Nat Immunol*, 14(12):1294–301, 2013.

[243] N. Zhang and M. J. Bevan. Transforming growth factor-beta signaling controls the formation and maintenance of gut-resident memory t cells by regulating migration and retention. *Immunity*, 39(4):687–96, 2013.

[244] P. R. Dunbar, E. K. Cartwright, A. N. Wein, T. Tsukamoto, Z. R. Tiger Li, N. Kumar, I. E. Uddback, S. L. Hayward, S. Ueha, S. Takamura, and J. E. Kohlmeier. Pulmonary monocytes interact with effector t cells in the lung tissue to drive trm differentiation following viral infection. *Mucosal Immunol*, 13(1):161–171, 2020.

[245] B. Slutter, N. Van Braeckel-Budimir, G. Abboud, S. M. Varga, S. Salek-Ardakani, and J. T. Harty. Dynamics of influenza-induced lung-resident memory t cells underlie waning heterosubtypic immunity. *Sci Immunol*, 2(7), 2017.

[246] S. L. Hayward, C. D. Scharer, E. K. Cartwright, S. Takamura, Z. T. Li, J. M. Boss, and J. E. Kohlmeier. Environmental cues regulate epigenetic reprogramming of airway-resident memory cd8(+) t cells. *Nat Immunol*, 21(3):309–320, 2020.

[247] I. E. Uddback, L. M. Pedersen, S. R. Pedersen, M. A. Steffensen, P. J. Holst, A. R. Thomsen, and J. P. Christensen. Combined local and systemic immunization is essential for durable t-cell mediated heterosubtypic immunity against influenza a virus. *Sci Rep*, 6:20137, 2016.

[248] S. R. Crowe, S. C. Miller, R. M. Shenyo, and D. L. Woodland. Vaccination with an acidic polymerase epitope of influenza virus elicits a potent antiviral t cell response but delayed clearance of an influenza virus challenge. *J Immunol*, 174(2):696–701, 2005.

[249] S. Quinones-Parra, E. Grant, L. Loh, T. H. Nguyen, K. A. Campbell, S. Y. Tong, A. Miller, P. C. Doherty, D. Vijaykrishna, J. Rossjohn, S. Gras, and K. Kedzierska. Preexisting cd8+ t-cell immunity to the h7n9 influenza a virus varies across ethnicities. *Proc Natl Acad Sci U S A*, 111(3):1049–54, 2014.

[250] M. Koutsakos, P. T. Illing, T. H. O. Nguyen, N. A. Mifsud, J. C. Crawford, S. Rizzetto, A. A. Eltahla, E. B. Clemens, S. Sant, B. Y. Chua, C. Y. Wong, E. K. Allen, D. Teng, P. Dash, D. F. Boyd, L. Grzelak, W. Zeng, A. C. Hurt, I. Barr, S. Rockman, D. C. Jackson, T. C. Kotsimbos, A. C. Cheng, M. Richards, G. P. Westall, T. Loudovaris, S. I. Mannering, M. Elliott, S. G. Tangye, L. M. Wakim, J. Rossjohn, D. Vijaykrishna, F. Luciani, P. G. Thomas, S. Gras, A. W. Purcell, and K. Kedzierska. Human cd8(+) t cell cross-reactivity across influenza a, b and c viruses. *Nat Immunol*, 20(5):613–625, 2019.

[251] A. R. Townsend, J. Rothbard, F. M. Gotch, G. Bahadur, D. Wraith, and A. J. McMichael. The epitopes of influenza nucleoprotein recognized by cytotoxic t lymphocytes can be defined with short synthetic peptides. *Cell*, 44(6):959–68, 1986.

[252] G. T. Belz, W. Xie, J. D. Altman, and P. C. Doherty. A previously unrecognized h-2d(b)-restricted peptide prominent in the primary influenza a virus-specific cd8(+) t-cell response is much less apparent following secondary challenge. *J Virol*, 74(8):3486–93, 2000.

[253] G. T. Belz, W. Xie, and P. C. Doherty. Diversity of epitope and cytokine profiles for primary and secondary influenza a virus-specific cd8+ t cell responses. *J Immunol*, 166(7):4627–33, 2001.

[254] A. Vitiello, L. Yuan, R. W. Chesnut, J. Sidney, S. Southwood, P. Farness, M. R. Jackson, P. A. Peterson, and A. Sette. Immunodominance analysis of ctl responses to influenza pr8 virus reveals two new dominant and subdominant kb-restricted epitopes. *J Immunol*, 157(12):5555–62, 1996.

[255] S. S. Andreansky, J. Stambas, P. G. Thomas, W. Xie, R. J. Webby, and P. C. Doherty. Consequences of immunodominant epitope deletion for minor influenza virus-specific cd8+-t-cell responses. *J Virol*, 79(7):4329–39, 2005.

[256] P. G. Thomas, R. Keating, D. J. Hulse-Post, and P. C. Doherty. Cell-mediated protection in influenza infection. *Emerg Infect Dis*, 12(1):48–54, 2006.

[257] T. Wu, J. Guan, A. Handel, D. C. Tscharke, J. Sidney, A. Sette, L. M. Wakim, X. Y. X. Sng, P. G. Thomas, N. P. Croft, A. W. Purcell, and N. L. La Gruta. Quantification of epitope abundance reveals the effect of direct and cross-presentation on influenza ctl responses. *Nat Commun*, 10(1):2846, 2019.

[258] E. Assarsson, H. H. Bui, J. Sidney, Q. Zhang, J. Glenn, C. Oseroff, I. N. Mbawuike, J. Alexander, M. J. Newman, H. Grey, and A. Sette. Immunomic analysis of the repertoire of t-cell specificities for influenza a virus in humans. *J Virol*, 82(24):12241–51, 2008.

[259] C. Wu, D. Zanker, S. Valkenburg, B. Tan, K. Kedzierska, Q. M. Zou, P. C. Doherty, and W. Chen. Systematic identification of immunodominant cd8+ t-cell responses to influenza a virus in hla-a2 individuals. *Proc Natl Acad Sci U S A*, 108(22):9178–83, 2011.

[260] E. Grant, C. Wu, K. F. Chan, S. Eckle, M. Bharadwaj, Q. M. Zou, K. Kedzierska, and W. Chen. Nucleoprotein of influenza a virus is a major target of immunodominant cd8+ t-cell responses. *Immunol Cell Biol*, 91(2):184–94, 2013.

[261] G. F. Rimmelzwaan, J. H. Kreijtz, R. Bodewes, R. A. Fouchier, and A. D. Osterhaus. Influenza virus ctl epitopes, remarkably conserved and remarkably variable. *Vaccine*, 27(45):6363–5, 2009.

[262] S. E. Hensley, S. R. Das, A. L. Bailey, L. M. Schmidt, H. D. Hickman, A. Jayaraman, K. Viswanathan, R. Raman, R. Sasisekharan, J. R. Bennink, and J. W. Yewdell. Hemagglutinin receptor binding avidity drives influenza a virus antigenic drift. *Science*, 326(5953):734–6, 2009.

[263] J. M. Lee, R. Eguia, S. J. Zost, S. Choudhary, P. C. Wilson, T. Bedford, T. Stevens-Ayers, M. Boeckh, A. C. Hurt, S. S. Lakdawala, S. E. Hensley, and J. D. Bloom. Mapping person-to-person variation in viral mutations that escape polyclonal serum targeting influenza hemagglutinin. *Elife*, 8, 2019.

[264] D. J. Smith, A. S. Lapedes, J. C. de Jong, T. M. Bestebroer, G. F. Rimmelzwaan, A. D. Osterhaus, and R. A. Fouchier. Mapping the antigenic and genetic evolution of influenza virus. *Science*, 305(5682):371–6, 2004.

[265] K. Koelle, S. Cobey, B. Grenfell, and M. Pascual. Epochal evolution shapes the phylodynamics of interpandemic influenza a (h3n2) in humans. *Science*, 314(5807):1898–903, 2006.

[266] B. F. Koel, D. F. Burke, T. M. Bestebroer, S. van der Vliet, G. C. Zondag, G. Vervaet, E. Skepner, N. S. Lewis, M. I. Spronken, C. A. Russell, M. Y. Eropkin, A. C. Hurt, I. G. Barr, J. C. de Jong, G. F. Rimmelzwaan, A. D. Osterhaus, R. A. Fouchier, and D. J. Smith. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science*, 342(6161):976–9, 2013.

[267] T. Bedford, M. A. Suchard, P. Lemey, G. Dudas, V. Gregory, A. J. Hay, J. W. McCauley, C. A. Russell, D. J. Smith, and A. Rambaut. Integrating influenza antigenic dynamics with molecular evolution. *Elife*, 3:e01914, 2014.

[268] H. Kim, R. G. Webster, and R. J. Webby. Influenza virus: Dealing with a drifting and shifting pathogen. *Viral Immunol*, 31(2):174–183, 2018.

[269] E. D. Kilbourne. Influenza pandemics of the 20th century. *Emerg Infect Dis*, 12(1):9–14, 2006.

[270] R. J. Garten, C. T. Davis, C. A. Russell, B. Shu, S. Lindstrom, A. Balish, W. M. Sessions, X. Xu, E. Skepner, V. Deyde, M. Okomo-Adhiambo, L. Gubareva, J. Barnes, C. B. Smith, S. L. Emery, M. J. Hillman, P. Rivailler, J. Smagala, M. de Graaf, D. F. Burke, R. A. Fouchier, C. Pappas, C. M. Alpuche-Aranda, H. Lopez-Gatell, H. Olivera, I. Lopez, C. A. Myers, D. Faix, P. J. Blair, C. Yu, K. M. Keene, Jr. Dotson, P. D., D. Boxrud, A. R. Sambol, S. H. Abid, K. St George, T. Bannerman, A. L. Moore, D. J. Stringer, P. Blevins, G. J. Demmler-Harrison, M. Ginsberg, P. Kriner, S. Waterman, S. Smole, H. F. Guevara, E. A. Belongia, P. A. Clark, S. T. Beatrice, R. Donis, J. Katz, L. Finelli, C. B. Bridges, M. Shaw, D. B. Jernigan, T. M. Uyeki, D. J. Smith, A. I. Klimov, and N. J. Cox. Antigenic and genetic characteristics of swine-origin 2009 a(h1n1) influenza viruses circulating in humans. *Science*, 325(5937):197–201, 2009.

[271] G. J. Smith, D. Vijaykrishna, J. Bahl, S. J. Lycett, M. Worobey, O. G. Pybus, S. K. Ma, C. L. Cheung, J. Raghwani, S. Bhatt, J. S. Peiris, Y. Guan, and A. Rambaut. Origins and evolutionary genomics of the 2009 swine-origin h1n1 influenza a epidemic. *Nature*, 459(7250):1122–5, 2009.

[272] W. J. Alonso, C. Viboud, L. Simonsen, E. W. Hirano, L. Z. Daufenbach, and M. A. Miller. Seasonality of influenza in brazil: a traveling wave from the amazon to the subtropics. *Am J Epidemiol*, 165(12):1434–42, 2007.

[273] E. Lofgren, N. H. Fefferman, Y. N. Naumov, J. Gorski, and E. N. Naumova. Influenza seasonality: underlying causes and modeling theories. *J Virol*, 81(11):5429–36, 2007.

[274] M. I. Nelson, L. Simonsen, C. Viboud, M. A. Miller, and E. C. Holmes. Phylogenetic analysis reveals the global migration of seasonal influenza a viruses. *PLoS Pathog*, 3(9):1220–8, 2007.

[275] C. A. Russell, T. C. Jones, I. G. Barr, N. J. Cox, R. J. Garten, V. Gregory, I. D. Gust, A. W. Hampson, A. J. Hay, A. C. Hurt, J. C. de Jong, A. Kelso, A. I. Klimov, T. Kageyama, N. Komadina, A. S. Lapedes, Y. P. Lin, A. Mosterin, M. Obuchi, T. Odagiri, A. D. Osterhaus, G. F. Rimmelzwaan, M. W. Shaw, E. Skepner, K. Stohr, M. Tashiro, R. A. Fouchier, and D. J. Smith. The global circulation of seasonal influenza a (h3n2) viruses. *Science*, 320(5874):340–6, 2008.

[276] A. Rambaut, O. G. Pybus, M. I. Nelson, C. Viboud, J. K. Taubenberger, and E. C. Holmes. The genomic and epidemiological dynamics of human influenza a virus. *Nature*, 453(7195):615–9, 2008.

[277] S. Sridhar. Heterosubtypic t-cell immunity to influenza in humans: Challenges for universal t-cell influenza vaccines. *Front Immunol*, 7:195, 2016.

[278] F. Krammer. Novel universal influenza virus vaccine approaches. *Curr Opin Virol*, 17:95–103, 2016.

[279] C. I. Paules, H. D. Marston, R. W. Eisinger, D. Baltimore, and A. S. Fauci. The pathway to a universal influenza vaccine. *Immunity*, 47(4):599–603, 2017.

[280] L. D. Estrada and S. Schultz-Cherry. Development of a universal influenza vaccine. *J Immunol*, 202(2):392–398, 2019.

[281] A. Impagliazzo, F. Milder, H. Kuipers, M. V. Wagner, X. Zhu, R. M. Hoffman, R. van Meersbergen, J. Huizingh, P. Wanningen, J. Verspuij, M. de Man, Z. Ding, A. Apetri, B. Kukrer, E. Sneekes-Vriese, D. Tomkiewicz, N. S. Laursen, P. S. Lee, A. Zakrzewska, L. Dekking, J. Tolboom, L. Tettero, S. van Meerten, W. Yu, W. Koudstaal, J. Goudsmit, A. B. Ward, W. Meijberg, I. A. Wilson, and K. Radosevic. A stable trimeric influenza hemagglutinin stem as a broadly protective immunogen. *Science*, 349(6254):1301–6, 2015.

[282] N. Chai, L. R. Swem, M. Reichelt, H. Chen-Harris, E. Luis, S. Park, A. Fouts, P. Lupardus, T. D. Wu, O. Li, J. McBride, M. Lawrence, M. Xu, and M. W. Tan. Two escape mechanisms of influenza a virus to a broadly neutralizing stalk-binding antibody. *PLoS Pathog*, 12(6):e1005702, 2016.

[283] S. Bhatt, E. C. Holmes, and O. G. Pybus. The genomic rate of molecular adaptation of the human influenza a virus. *Mol Biol Evol*, 28(9):2443–51, 2011.

[284] John H. Gillespie. *Population genetics : a concise guide*. Johns Hopkins University Press, Baltimore, Md., 2nd edition, 2004.

[285] M. Kimura. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature*, 267(5608):275–6, 1977.

[286] R. Nielsen and Z. Yang. Estimating the distribution of selection coefficients from phylogenetic data with applications to mitochondrial and viral dna. *Mol Biol Evol*, 20(8):1231–9, 2003.

[287] E. G. Berkhoff, E. de Wit, M. M. Geelhoed-Mieras, A. C. Boon, J. Symons, R. A. Fouchier, A. D. Osterhaus, and G. F. Rimmelzwaan. Functional constraints of influenza a virus epitopes limit escape from cytotoxic t lymphocytes. *J Virol*, 79(17):11239–46, 2005.

[288] E. G. Berkhoff, E. de Wit, M. M. Geelhoed-Mieras, A. C. Boon, J. Symons, R. A. Fouchier, A. D. Osterhaus, and G. F. Rimmelzwaan. Fitness costs limit escape from cytotoxic t lymphocytes by influenza a viruses. *Vaccine*, 24(44-46):6594–6, 2006.

[289] L. I. Gong, M. A. Suchard, and J. D. Bloom. Stability-mediated epistasis constrains the evolution of an influenza protein. *Elife*, 2:e00631, 2013.

[290] H. M. Machkovech, T. Bedford, M. A. Suchard, and J. D. Bloom. Positive selection in cd8+ t-cell epitopes of influenza virus nucleoprotein revealed by a comparative analysis of human and swine viral lineages. *J Virol*, 89(22):11275–83, 2015.

[291] R. G. Woolthuis, C. H. van Dorp, C. Kesmir, R. J. de Boer, and M. van Boven. Long-term adaptation of the influenza a virus by escaping cytotoxic t-cell recognition. *Sci Rep*, 6:33334, 2016.

[292] J. R. Gog, G. F. Rimmelzwaan, A. D. Osterhaus, and B. T. Grenfell. Population dynamics of rapid fixation in cytotoxic t lymphocyte escape mutants of influenza a. *Proc Natl Acad Sci U S A*, 100(19):11143–7, 2003.

[293] B. J. Cowling, V. J. Fang, S. Riley, J. S. Malik Peiris, and G. M. Leung. Estimation of the serial interval of influenza. *Epidemiology*, 20(3):344–7, 2009.

[294] M. E. Halloran, C. J. Struchiner, and Jr. Longini, I. M. Study designs for evaluating different efficacy and effectiveness aspects of vaccines. *Am J Epidemiol*, 146(10):789–803, 1997.

[295] A. J. Kucharski, J. Lessler, J. M. Read, H. Zhu, C. Q. Jiang, Y. Guan, D. A. Cummings, and S. Riley. Estimating the life course of influenza a(h3n2) antibody responses from cross-sectional data. *PLoS Biol*, 13(3):e1002082, 2015.

[296] C. M. Pease. An evolutionary epidemiological mechanism, with applications to type a influenza. *Theor Popul Biol*, 31(3):422–52, 1987.

[297] K. Koelle, P. Khatri, M. Kamradt, and T. B. Kepler. A two-tiered model for simulating the ecological and evolutionary dynamics of rapidly evolving viruses, with an application to influenza. *J R Soc Interface*, 7(50):1257–74, 2010.

[298] B. M. Calderon, S. Danzy, G. K. Delima, N. T. Jacobs, K. Ganti, M. R. Hockman, G. L. Conn, A. C. Lowen, and J. Steel. Dysregulation of m segment gene expression contributes to influenza a virus host restriction. *PLoS Pathog*, 15(8):e1007892, 2019.

[299] A. C. Davison and D. V. Hinkley. *Bootstrap methods and their application*. Cambridge University Press, Cambridge ; New York, NY, USA, 1997.

[300] A. N. Wein, P. R. Dunbar, S. R. McMaster, Z. T. Li, T. L. Denning, and J. E. Kohlmeier. Il-36gamma protects against severe influenza infection by promoting lung alveolar macrophage survival and limiting viral replication. *J Immunol*, 201(2):573–582, 2018.

[301] A. Handel, Jr. Longini, I. M., and R. Antia. Neuraminidase inhibitor resistance in influenza: assessing the danger of its generation and spread. *PLoS Comput Biol*, 3(12):e240, 2007.

[302] J. T. Voeten, T. M. Bestebroer, N. J. Nieuwkoop, R. A. Fouchier, A. D. Osterhaus, and G. F. Rimmelzwaan. Antigenic drift in the influenza a virus (h3n2) nucleoprotein and escape from recognition by cytotoxic t lymphocytes. *J Virol*, 74(15):6800–7, 2000.

[303] K. J. Flynn, G. T. Belz, J. D. Altman, R. Ahmed, D. L. Woodland, and P. C. Doherty. Virus-specific cd8+ t cells in primary and secondary influenza pneumonia. *Immunity*, 8(6):683–91, 1998.

[304] T. Wu, Y. Hu, Y. T. Lee, K. R. Bouchard, A. Benechet, K. Khanna, and L. S. Cauley. Lung-resident memory cd8 t cells (trm) are indispensable for optimal cross-protection against pulmonary virus infection. *J Leukoc Biol*, 95(2):215–24, 2014.

[305] A. R. Wargo and G. Kurath. Viral fitness: definitions, measurement, and current insights. *Curr Opin Virol*, 2(5):538–45, 2012.

[306] Matthew James Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, Princeton, 2008.

[307] CDC: How Influenza (Flu) Vaccines Are Made. Last updated: 2019/12/09.

[308] B. Zhou, V. A. Meliopoulos, W. Wang, X. Lin, K. M. Stucker, R. A. Halpin, T. B. Stockwell, S. Schultz-Cherry, and D. E. Wentworth. Reversion of cold-adapted live attenuated influenza vaccine into a pathogenic virus. *J Virol*, 90(19):8454–63, 2016.

[309] J. R. Chung, B. Flannery, C. S. Ambrose, R. E. Begue, H. Caspard, L. DeMarcus, A. L. Fowlkes, G. Kersellius, A. Steffens, A. M. Fry, Team Influenza Clinical Investigation for Children Study, Project Influenza Incidence Surveillance, and U. S. Influenza Vaccine Effectiveness Network. Live attenuated and inactivated influenza vaccine effectiveness. *Pediatrics*, 143(2), 2019.